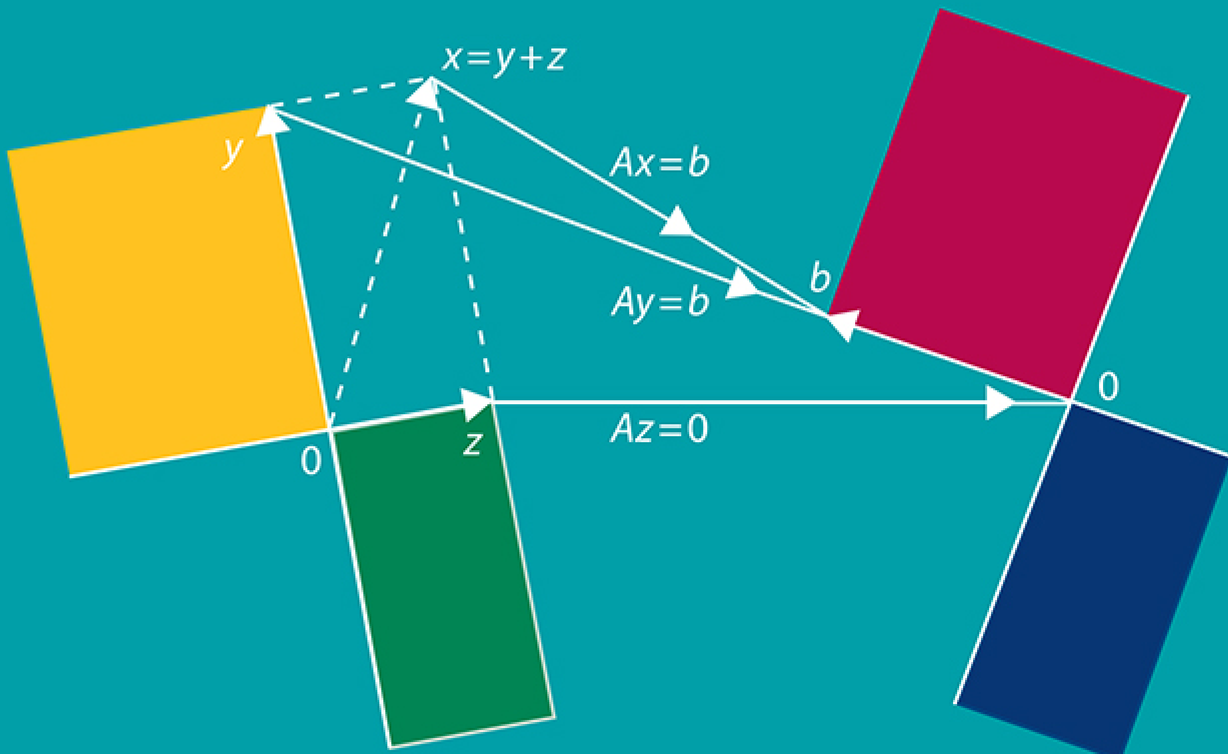
Introduction to
LINEAR ALGEBRA
FIFTH EDITION



GILBERT STRANG

INTRODUCTION

TO

LINEAR

ALGEBRA

Fifth Edition

MANUAL FOR INSTRUCTORS

Gilbert Strang

Massachusetts Institute of Technology

math.mit.edu/linearalgebra

web.mit.edu/18.06

video lectures: ocw.mit.edu

math.mit.edu/∼gs

www.wellesleycambridge.com
email: linearalgebrabook@gmail.com

Wellesley - Cambridge Press

Box 812060

Wellesley, Massachusetts 02482

Problem Set 1.1, page 8

1 The combinations give (a) a line in $\mathbf{R}^3$ (b) a plane in $\mathbf{R}^3$ (c) all of $\mathbf{R}^3$.

2 $\boldsymbol{v}+\boldsymbol{w}=(2,3)$ and $\boldsymbol{v}-\boldsymbol{w}=(6,-1)$ will be the diagonals of the parallelogram with $\boldsymbol{v}$ and $\boldsymbol{w}$ as two sides going out from $(0,0)$.

3 This problem gives the diagonals $\boldsymbol{v}+\boldsymbol{w}$ and $\boldsymbol{v}-\boldsymbol{w}$ of the parallelogram and asks for the sides: The opposite of Problem 2. In this example $\boldsymbol{v}=(3,3)$ and $\boldsymbol{w}=(2,-2)$.

4 $3\boldsymbol{v}+\boldsymbol{w}=(7,5)$ and $c\boldsymbol{v}+d\boldsymbol{w}=(2c+d,c+2d)$.

5 $\boldsymbol{u}+\boldsymbol{v}=(-2,3,1)$ and $\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}=(0,0,0)$ and $2\boldsymbol{u}+2\boldsymbol{v}+\boldsymbol{w}=($ add first answers$)=(-2,3,1)$. The vectors $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ are in the same plane because a combination gives $(0,0,0)$. Stated another way: $\boldsymbol{u}=-\boldsymbol{v}-\boldsymbol{w}$ is in the plane of $\boldsymbol{v}$ and $\boldsymbol{w}$.

6 The components of every $c\boldsymbol{v}+d\boldsymbol{w}$ add to zero because the components of $\boldsymbol{v}$ and of $\boldsymbol{w}$ add to zero. $c=3$ and $d=9$ give $(3,3,-6)$. There is no solution to $c\boldsymbol{v}+d\boldsymbol{w}=(3,3,6)$ because $3+3+6$ is not zero.

7 The nine combinations $c(2,1)+d(0,1)$ with $c=0,1,2$ and $d=(0,1,2)$ will lie on a lattice. If we took all whole numbers c and d, the lattice would lie over the whole plane.

8 The other diagonal is $\boldsymbol{v}-\boldsymbol{w}$ (or else $\boldsymbol{w}-\boldsymbol{v}$). Adding diagonals gives $2\boldsymbol{v}$ (or $2\boldsymbol{w}$).

9 The fourth corner can be $(4,4)$ or $(4,0)$ or $(-2,2)$. Three possible parallelograms!

10 $\boldsymbol{i}-\boldsymbol{j}=(1,1,0)$ is in the base (x-y plane). $\boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}=(1,1,1)$ is the opposite corner from $(0,0,0)$. Points in the cube have $0\le x\le 1$, $0\le y\le 1$, $0\le z\le 1$.

11 Four more corners $(1,1,0),(1,0,1),(0,1,1),(1,1,1)$. The center point is $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$. Centers of faces are $(\frac{1}{2},\frac{1}{2},0),(\frac{1}{2},\frac{1}{2},1)$ and $(0,\frac{1}{2},\frac{1}{2}),(1,\frac{1}{2},\frac{1}{2})$ and $(\frac{1}{2},0,\frac{1}{2}),(\frac{1}{2},1,\frac{1}{2})$.

12 The combinations of $\boldsymbol{i}=(1,0,0)$ and $\boldsymbol{i}+\boldsymbol{j}=(1,1,0)$ fill the xy plane in xyz space.

13 Sum = zero vector. Sum $=-2{:}00$ vector $=8{:}00$ vector. 2:00 is 30° from horizontal $=(\cos\frac{\pi}{6},\sin\frac{\pi}{6})=(\sqrt{3}/2,1/2)$.

14 Moving the origin to 6:00 adds $\boldsymbol{j}=(0,1)$ to every vector. So the sum of twelve vectors changes from $\mathbf{0}$ to $12\boldsymbol{j}=(0,12)$.

15 The point $\frac{3}{4}\boldsymbol{v} + \frac{1}{4}\boldsymbol{w}$ is three-fourths of the way to $\boldsymbol{v}$ starting from $\boldsymbol{w}$. The vector $\frac{1}{4}\boldsymbol{v} + \frac{1}{4}\boldsymbol{w}$ is halfway to $\boldsymbol{u} = \frac{1}{2}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}$. The vector $\boldsymbol{v} + \boldsymbol{w}$ is $2\boldsymbol{u}$ (the far corner of the parallelogram).

16 All combinations with $c + d = 1$ are on the line that passes through $\boldsymbol{v}$ and $\boldsymbol{w}$. The point $\boldsymbol{V} = -\boldsymbol{v} + 2\boldsymbol{w}$ is on that line but it is beyond $\boldsymbol{w}$.

17 All vectors $c\boldsymbol{v} + c\boldsymbol{w}$ are on the line passing through $(0, 0)$ and $\boldsymbol{u} = \frac{1}{2}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}$. That line continues out beyond $\boldsymbol{v} + \boldsymbol{w}$ and back beyond $(0, 0)$. With $c \geq 0$, half of this line is removed, leaving a *ray* that starts at $(0, 0)$.

18 The combinations $c\boldsymbol{v} + d\boldsymbol{w}$ with $0 \leq c \leq 1$ and $0 \leq d \leq 1$ *fill the parallelogram* with sides $\boldsymbol{v}$ and $\boldsymbol{w}$. For example, if $\boldsymbol{v} = (1, 0)$ and $\boldsymbol{w} = (0, 1)$ then $c\boldsymbol{v} + d\boldsymbol{w}$ fills the unit square. But when $\boldsymbol{v} = (a, 0)$ and $\boldsymbol{w} = (b, 0)$ these combinations only fill a segment of a line.

19 With $c \geq 0$ and $d \geq 0$ we get the infinite "cone" or "wedge" between $\boldsymbol{v}$ and $\boldsymbol{w}$. For example, if $\boldsymbol{v} = (1, 0)$ and $\boldsymbol{w} = (0, 1)$, then the cone is the whole quadrant $x \geq 0$, $y \geq 0$. *Question*: What if $\boldsymbol{w} = -\boldsymbol{v}$? The cone opens to a half-space. But the combinations of $\boldsymbol{v} = (1, 0)$ and $\boldsymbol{w} = (-1, 0)$ only fill a line.

20 (a) $\frac{1}{3}\boldsymbol{u} + \frac{1}{3}\boldsymbol{v} + \frac{1}{3}\boldsymbol{w}$ is the center of the triangle between $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$; $\frac{1}{2}\boldsymbol{u} + \frac{1}{2}\boldsymbol{w}$ lies between $\boldsymbol{u}$ and $\boldsymbol{w}$ (b) To fill the triangle keep $c \geq 0$, $d \geq 0$, $e \geq 0$, and $c + d + e = \mathbf{1}$.

21 The sum is $(\boldsymbol{v} - \boldsymbol{u}) + (\boldsymbol{w} - \boldsymbol{v}) + (\boldsymbol{u} - \boldsymbol{w}) =$ **zero vector**. Those three sides of a triangle are in the same plane!

22 The vector $\frac{1}{2}(\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w})$ is *outside* the pyramid because $c + d + e = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} > 1$.

23 *All* vectors are combinations of $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ as drawn (not in the same plane). Start by seeing that $c\boldsymbol{u} + d\boldsymbol{v}$ fills a plane, then adding $e\boldsymbol{w}$ fills all of $\mathbf{R}^3$.

24 The combinations of $\boldsymbol{u}$ and $\boldsymbol{v}$ fill one plane. The combinations of $\boldsymbol{v}$ and $\boldsymbol{w}$ fill another plane. Those planes meet in a *line*: *only the vectors* $c\boldsymbol{v}$ are in both planes.

25 (a) For a line, choose $\boldsymbol{u} = \boldsymbol{v} = \boldsymbol{w} =$ any nonzero vector (b) For a plane, choose $\boldsymbol{u}$ and $\boldsymbol{v}$ in different directions. A combination like $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$ is in the same plane.

26 Two equations come from the two components: $c + 3d = 14$ and $2c + d = 8$. *The solution is* $c = 2$ *and* $d = 4$. Then $2(1, 2) + 4(3, 1) = (14, 8)$.

27 A four-dimensional cube has $2^4 = 16$ corners and $2 \cdot 4 = 8$ three-dimensional faces and 24 two-dimensional faces and 32 edges in Worked Example **2.4 A**.

28 There are **6** unknown numbers $v_1, v_2, v_3, w_1, w_2, w_3$. The six equations come from the components of $\boldsymbol{v} + \boldsymbol{w} = (4, 5, 6)$ and $\boldsymbol{v} - \boldsymbol{w} = (2, 5, 8)$. Add to find $2\boldsymbol{v} = (6, 10, 14)$ so $\boldsymbol{v} = (3, 5, 7)$ and $\boldsymbol{w} = (1, 0, -1)$.

29 Two combinations out of infinitely many that produce $\boldsymbol{b} = (0, 1)$ are $-2\boldsymbol{u} + \boldsymbol{v}$ and $\frac{1}{2}\boldsymbol{w} - \frac{1}{2}\boldsymbol{v}$. **No**, three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in the x-y plane could fail to produce $\boldsymbol{b}$ if all three lie on a line that does not contain $\boldsymbol{b}$. *Yes*, if one combination produces $\boldsymbol{b}$ then two (and infinitely many) combinations will produce $\boldsymbol{b}$. This is true even if $\boldsymbol{u} = \boldsymbol{0}$; the combinations can have different $c\boldsymbol{u}$.

30 The combinations of $\boldsymbol{v}$ and $\boldsymbol{w}$ fill the plane *unless* $\boldsymbol{v}$ *and* $\boldsymbol{w}$ *lie on the same line through* $(0, 0)$. Four vectors whose combinations fill 4-dimensional space: one example is the "standard basis" $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$, and $(0, 0, 0, 1)$.

31 The equations $c\boldsymbol{u} + d\boldsymbol{v} + e\boldsymbol{w} = \boldsymbol{b}$ are

$$\begin{array}{rrrl} 2c & -d & & = 1 \\ -c & +2d & -e & = 0 \\ & -d & +2e & = 0 \end{array} \qquad \begin{array}{l} \text{So } d = 2e \\ \text{then } c = 3e \\ \text{then } 4e = 1 \end{array} \qquad \begin{array}{l} c = 3/4 \\ d = 2/4 \\ e = 1/4 \end{array}$$

Problem Set 1.2, page 18

1 $\boldsymbol{u} \cdot \boldsymbol{v} = -2.4 + 2.4 = 0$, $\boldsymbol{u} \cdot \boldsymbol{w} = -.6 + 1.6 = 1$, $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w} = 0 + 1$, $\boldsymbol{w} \cdot \boldsymbol{v} = 4 - 6 = -2 = \boldsymbol{v} \cdot \boldsymbol{w}$.

2 $\|\boldsymbol{u}\| = 1$ and $\|\boldsymbol{v}\| = 5$ and $\|\boldsymbol{w}\| = \sqrt{5}$. Then $|\boldsymbol{u} \cdot \boldsymbol{v}| = 0 < (1)(5)$ and $|\boldsymbol{v} \cdot \boldsymbol{w}| = 10 < 5\sqrt{5}$, confirming the Schwarz inequality.

3 Unit vectors $\boldsymbol{v}/\|\boldsymbol{v}\| = (\frac{4}{5}, \frac{3}{5}) = (0.8, 0.6)$. The vectors $\boldsymbol{w}, (2,-1)$, and $-\boldsymbol{w}$ make $0^\circ, 90^\circ, 180^\circ$ angles with $\boldsymbol{w}$ and $\boldsymbol{w}/\|\boldsymbol{w}\| = (1/\sqrt{5}, 2/\sqrt{5})$. The cosine of θ is $\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \cdot \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} = 10/5\sqrt{5}$.

4 (a) $\boldsymbol{v} \cdot (-\boldsymbol{v}) = \mathbf{-1}$ (b) $(\boldsymbol{v}+\boldsymbol{w}) \cdot (\boldsymbol{v}-\boldsymbol{w}) = \boldsymbol{v}\cdot\boldsymbol{v} + \boldsymbol{w}\cdot\boldsymbol{v} - \boldsymbol{v}\cdot\boldsymbol{w} - \boldsymbol{w}\cdot\boldsymbol{w} = 1 + (\quad) - (\quad) - 1 = \mathbf{0}$ so $\theta = 90^\circ$ (notice $\boldsymbol{v}\cdot\boldsymbol{w} = \boldsymbol{w}\cdot\boldsymbol{v}$) (c) $(\boldsymbol{v}-2\boldsymbol{w})\cdot(\boldsymbol{v}+2\boldsymbol{w}) = \boldsymbol{v}\cdot\boldsymbol{v} - 4\boldsymbol{w}\cdot\boldsymbol{w} = 1 - 4 = \mathbf{-3}$.

5 $\boldsymbol{u}_1 = \boldsymbol{v}/\|\boldsymbol{v}\| = (1,3)/\sqrt{10}$ and $\boldsymbol{u}_2 = \boldsymbol{w}/\|\boldsymbol{w}\| = (2,1,2)/3$. $\boldsymbol{U}_1 = (3,-1)/\sqrt{10}$ is perpendicular to $\boldsymbol{u}_1$ (and so is $(-3,1)/\sqrt{10}$). $\boldsymbol{U}_2$ could be $(1,-2,0)/\sqrt{5}$: There is a whole plane of vectors perpendicular to $\boldsymbol{u}_2$, and a whole circle of unit vectors in that plane.

6 All vectors $\boldsymbol{w} = (c, 2c)$ are perpendicular to $\boldsymbol{v}$. They lie on a line. All vectors (x,y,z) with $x+y+z=0$ lie on a *plane*. All vectors perpendicular to $(1,1,1)$ and $(1,2,3)$ lie on a *line* in 3-dimensional space.

7 (a) $\cos\theta = \boldsymbol{v}\cdot\boldsymbol{w}/\|\boldsymbol{v}\|\|\boldsymbol{w}\| = 1/(2)(1)$ so $\theta = 60^\circ$ or $\pi/3$ radians (b) $\cos\theta = 0$ so $\theta = 90^\circ$ or $\pi/2$ radians (c) $\cos\theta = 2/(2)(2) = 1/2$ so $\theta = 60^\circ$ or $\pi/3$ (d) $\cos\theta = -1/\sqrt{2}$ so $\theta = 135^\circ$ or $3\pi/4$.

8 (a) False: $\boldsymbol{v}$ and $\boldsymbol{w}$ are any vectors in the plane perpendicular to $\boldsymbol{u}$ (b) True: $\boldsymbol{u}\cdot(\boldsymbol{v}+2\boldsymbol{w}) = \boldsymbol{u}\cdot\boldsymbol{v} + 2\boldsymbol{u}\cdot\boldsymbol{w} = 0$ (c) True, $\|\boldsymbol{u}-\boldsymbol{v}\|^2 = (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v})$ splits into $\boldsymbol{u}\cdot\boldsymbol{u} + \boldsymbol{v}\cdot\boldsymbol{v} = \mathbf{2}$ when $\boldsymbol{u}\cdot\boldsymbol{v} = \boldsymbol{v}\cdot\boldsymbol{u} = 0$.

9 If $v_2w_2/v_1w_1 = -1$ then $v_2w_2 = -v_1w_1$ or $v_1w_1 + v_2w_2 = \boldsymbol{v}\cdot\boldsymbol{w} = 0$: perpendicular! The vectors $(1,4)$ and $(1,-\frac{1}{4})$ are perpendicular.

10 Slopes $2/1$ and $-1/2$ multiply to give -1: then $\boldsymbol{v}\cdot\boldsymbol{w} = 0$ and the vectors (the directions) are perpendicular.

11 $\boldsymbol{v}\cdot\boldsymbol{w} < 0$ means angle $> 90^\circ$; these $\boldsymbol{w}$'s fill half of 3-dimensional space.

12 $(1,1)$ perpendicular to $(1,5) - c(1,1)$ if $(1,1)\cdot(1,5) - c(1,1)\cdot(1,1) = 6 - 2c = 0$ or $\boldsymbol{c = 3}$; $\boldsymbol{v}\cdot(\boldsymbol{w} - c\boldsymbol{v}) = 0$ if $c = \boldsymbol{v}\cdot\boldsymbol{w}/\boldsymbol{v}\cdot\boldsymbol{v}$. Subtracting $c\boldsymbol{v}$ is the key to constructing a perpendicular vector.

13 The plane perpendicular to $(1,0,1)$ contains all vectors $(c,d,-c)$. In that plane, $\boldsymbol{v} = (1,0,-1)$ and $\boldsymbol{w} = (0,1,0)$ are perpendicular.

14 One possibility among many: $\boldsymbol{u} = (1,-1,0,0), \boldsymbol{v} = (0,0,1,-1), \boldsymbol{w} = (1,1,-1,-1)$ and $(1,1,1,1)$ are perpendicular to each other. "We can rotate those $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in their 3D hyperplane and they will stay perpendicular."

15 $\frac{1}{2}(x+y) = (2+8)/2 = 5$ and $5 > 4$; $\cos\theta = 2\sqrt{16}/\sqrt{10}\sqrt{10} = \mathbf{8/10}$.

16 $\|\boldsymbol{v}\|^2 = 1+1+\cdots+1 = 9$ so $\|\boldsymbol{v}\| = \mathbf{3}$; $\boldsymbol{u} = \boldsymbol{v}/3 = (\frac{1}{3},\ldots,\frac{1}{3})$ is a unit vector in 9D; $\boldsymbol{w} = (1,-1,0,\ldots,0)/\sqrt{2}$ is a unit vector in the 8D hyperplane perpendicular to $\boldsymbol{v}$.

17 $\cos\alpha = 1/\sqrt{2}$, $\cos\beta = 0$, $\cos\gamma = -1/\sqrt{2}$. For any vector $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ the cosines with $(1,0,0)$ and $(0,0,1)$ are $\cos^2\alpha+\cos^2\beta+\cos^2\gamma = (v_1^2+v_2^2+v_3^2)/\|\boldsymbol{v}\|^2 = 1$.

18 $\|\boldsymbol{v}\|^2 = 4^2 + 2^2 = 20$ and $\|\boldsymbol{w}\|^2 = (-1)^2 + 2^2 = 5$. Pythagoras is $\|(3,4)\|^2 = 25 = 20 + 5$ for the length of the hypotenuse $\boldsymbol{v} + \boldsymbol{w} = (3,4)$.

19 Start from the rules $(1), (2), (3)$ for $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$ and $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w})$ and $(c\boldsymbol{v}) \cdot \boldsymbol{w}$. Use rule (2) for $(\boldsymbol{v}+\boldsymbol{w}) \cdot (\boldsymbol{v}+\boldsymbol{w}) = (\boldsymbol{v}+\boldsymbol{w}) \cdot \boldsymbol{v} + (\boldsymbol{v}+\boldsymbol{w}) \cdot \boldsymbol{w}$. By rule (1) this is $\boldsymbol{v} \cdot (\boldsymbol{v}+\boldsymbol{w}) + \boldsymbol{w} \cdot (\boldsymbol{v}+\boldsymbol{w})$. Rule (2) again gives $\boldsymbol{v}\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{w} + \boldsymbol{w}\cdot\boldsymbol{v} + \boldsymbol{w}\cdot\boldsymbol{w} = \boldsymbol{v}\cdot\boldsymbol{v} + 2\boldsymbol{v}\cdot\boldsymbol{w} + \boldsymbol{w}\cdot\boldsymbol{w}$. Notice $\boldsymbol{v}\cdot\boldsymbol{w} = \boldsymbol{w}\cdot\boldsymbol{v}$! The main point is to feel free to open up parentheses.

20 We know that $(\boldsymbol{v}-\boldsymbol{w}) \cdot (\boldsymbol{v}-\boldsymbol{w}) = \boldsymbol{v}\cdot\boldsymbol{v} - 2\boldsymbol{v}\cdot\boldsymbol{w} + \boldsymbol{w}\cdot\boldsymbol{w}$. The Law of Cosines writes $\|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos\theta$ for $\boldsymbol{v}\cdot\boldsymbol{w}$. Here θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$. When $\theta < 90^\circ$ this $\boldsymbol{v}\cdot\boldsymbol{w}$ is positive, so in this case $\boldsymbol{v}\cdot\boldsymbol{v} + \boldsymbol{w}\cdot\boldsymbol{w}$ is larger than $\|\boldsymbol{v}-\boldsymbol{w}\|^2$.

Pythagoras changes from equality $a^2+b^2 = c^2$ to *inequality* when $\theta < 90^\circ$ or $\theta > 90^\circ$.

21 $2\boldsymbol{v}\cdot\boldsymbol{w} \le 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|$ leads to $\|\boldsymbol{v}+\boldsymbol{w}\|^2 = \boldsymbol{v}\cdot\boldsymbol{v} + 2\boldsymbol{v}\cdot\boldsymbol{w} + \boldsymbol{w}\cdot\boldsymbol{w} \le \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| + \|\boldsymbol{w}\|^2$. This is $(\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2$. Taking square roots gives $\|\boldsymbol{v}+\boldsymbol{w}\| \le \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$.

22 $v_1^2w_1^2 + 2v_1w_1v_2w_2 + v_2^2w_2^2 \le v_1^2w_1^2 + v_1^2w_2^2 + v_2^2w_1^2 + v_2^2w_2^2$ is true (cancel 4 terms) because the difference is $v_1^2w_2^2 + v_2^2w_1^2 - 2v_1w_1v_2w_2$ which is $(v_1w_2 - v_2w_1)^2 \ge 0$.

23 $\cos\beta = w_1/\|\boldsymbol{w}\|$ and $\sin\beta = w_2/\|\boldsymbol{w}\|$. Then $\cos(\beta - a) = \cos\beta\cos\alpha + \sin\beta\sin\alpha = v_1w_1/\|\boldsymbol{v}\|\|\boldsymbol{w}\| + v_2w_2/\|\boldsymbol{v}\|\|\boldsymbol{w}\| = \boldsymbol{v}\cdot\boldsymbol{w}/\|\boldsymbol{v}\|\|\boldsymbol{w}\|$. This is $\cos\theta$ because $\beta - \alpha = \theta$.

24 Example 6 gives $|u_1||U_1| \le \frac{1}{2}(u_1^2 + U_1^2)$ and $|u_2||U_2| \le \frac{1}{2}(u_2^2 + U_2^2)$. The whole line becomes $.96 \le (.6)(.8) + (.8)(.6) \le \frac{1}{2}(.6^2 + .8^2) + \frac{1}{2}(.8^2 + .6^2) = 1$. True: $.96 < 1$.

25 The cosine of θ is $x/\sqrt{x^2+y^2}$, near side over hypotenuse. Then $|\cos\theta|^2$ is not greater than 1: $x^2/(x^2+y^2) \le 1$.

26 The vectors $\boldsymbol{w} = (x, y)$ with $(1, 2) \cdot \boldsymbol{w} = x + 2y = 5$ lie on a line in the xy plane. The shortest $\boldsymbol{w}$ on that line is $(1, 2)$. (The Schwarz inequality $\|\boldsymbol{w}\| \ge \boldsymbol{v} \cdot \boldsymbol{w}/\|\boldsymbol{v}\| = \sqrt{5}$ is an equality when $\cos\theta = 0$ and $\boldsymbol{w} = (1, 2)$ and $\|\boldsymbol{w}\| = \sqrt{5}$.)

27 The length $\|\boldsymbol{v} - \boldsymbol{w}\|$ is between 2 and 8 (triangle inequality when $\|\boldsymbol{v}\| = 5$ and $\|\boldsymbol{w}\| = 3$). The dot product $\boldsymbol{v} \cdot \boldsymbol{w}$ is between -15 and 15 by the Schwarz inequality.

28 Three vectors in the plane could make angles greater than $90°$ with each other: for example $(1, 0), (-1, 4), (-1, -4)$. Four vectors could *not* do this ($360°$ total angle). How many can do this in $\mathbf{R}^3$ or $\mathbf{R}^n$? Ben Harris and Greg Marks showed me that the answer is $n + 1$. The vectors from the center of a regular simplex in $\mathbf{R}^n$ to its $n + 1$ vertices all have negative dot products. If $n+2$ vectors in $\mathbf{R}^n$ had negative dot products, project them onto the plane orthogonal to the last one. Now you have $n + 1$ vectors in $\mathbf{R}^{n-1}$ with negative dot products. Keep going to 4 vectors in $\mathbf{R}^2$: no way!

29 For a specific example, pick $\boldsymbol{v} = (1, 2, -3)$ and then $\boldsymbol{w} = (-3, 1, 2)$. In this example $\cos\theta = \boldsymbol{v} \cdot \boldsymbol{w}/\|\boldsymbol{v}\|\|\boldsymbol{w}\| = -7/\sqrt{14}\sqrt{14} = -1/2$ and $\theta = 120°$. This always happens when $x + y + z = 0$:

$$\boldsymbol{v} \cdot \boldsymbol{w} = xz + xy + yz = \frac{1}{2}(x + y + z)^2 - \frac{1}{2}(x^2 + y^2 + z^2)$$

This is the same as $\boldsymbol{v} \cdot \boldsymbol{w} = 0 - \dfrac{1}{2}\,\|\boldsymbol{v}\|\|\boldsymbol{w}\|$. Then $\cos\theta = \dfrac{1}{2}$.

30 Wikipedia gives this proof of geometric mean $G = \sqrt[3]{xyz} \le$ arithmetic mean $A = (x + y + z)/3$. First there is equality in case $x = y = z$. Otherwise A is somewhere between the three positive numbers, say for example $z < A < y$.

Use the known inequality $g \le a$ for the *two* positive numbers x and $y + z - A$. Their mean $a = \frac{1}{2}(x + y + z - A)$ is $\frac{1}{2}(3A - A)$ = same as A! So $a \ge g$ says that

$A^3 \geq g^2A = x(y+z-A)A$. But $(y+z-A)A = (y-A)(A-z) + yz > yz$. Substitute to find $A^3 > xyz = G^3$ as we wanted to prove. Not easy!

There are many proofs of $G = (x_1x_2\cdots x_n)^{1/n} \leq A = (x_1 + x_2 + \cdots + x_n)/n$. In calculus you are maximizing G on the plane $x_1 + x_2 + \cdots + x_n = n$. The maximum occurs when all x's are equal.

31 The columns of the 4 by 4 "Hadamard matrix" (times $\frac{1}{2}$) are perpendicular unit vectors:

$$\frac{1}{2}H = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

32 The commands $V =$ **randn** $(3, 30)$; $D =$ **sqrt** (**diag** $(V' * V)$); $U = V \backslash D$; will give 30 random unit vectors in the columns of U. Then $u' * U$ is a row matrix of 30 dot products whose average absolute value may be close to $2/\pi$.

Problem Set 1.3, page 29

1 $2\boldsymbol{s}_1 + 3\boldsymbol{s}_2 + 4\boldsymbol{s}_3 = (2, 5, 9)$. The same vector $\boldsymbol{b}$ comes from S times $\boldsymbol{x} = (2, 3, 4)$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} (\text{row } 1) \cdot \boldsymbol{x} \\ (\text{row } 2) \cdot \boldsymbol{x} \\ (\text{row } 2) \cdot \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix}.$$

2 The solutions are $y_1 = 1$, $y_2 = 0$, $y_3 = 0$ (right side = column 1) and $y_1 = 1$, $y_2 = 3$, $y_3 = 5$. That second example illustrates that the first n odd numbers add to n^2.

3
$$\begin{aligned} y_1 &= B_1 \\ y_1 + y_2 &= B_2 \\ y_1 + y_2 + y_3 &= B_3 \end{aligned} \quad \text{gives} \quad \begin{aligned} y_1 &= B_1 \\ y_2 &= -B_1 + B_2 \\ y_3 &= -B_2 + B_3 \end{aligned} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}$$

The inverse of $S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ is $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$: **independent** columns in A and S!

4 The combination $0\boldsymbol{w}_1 + 0\boldsymbol{w}_2 + 0\boldsymbol{w}_3$ always gives the zero vector, but this problem looks for other *zero* combinations (then the vectors are *dependent*, they lie in a plane): $\boldsymbol{w}_2 = (\boldsymbol{w}_1 + \boldsymbol{w}_3)/2$ so one combination that gives zero is $\frac{1}{2}\boldsymbol{w}_1 - \boldsymbol{w}_2 + \frac{1}{2}\boldsymbol{w}_3 = \mathbf{0}$.

5 The rows of the 3 by 3 matrix in Problem 4 must also be *dependent*: $\boldsymbol{r}_2 = \frac{1}{2}(\boldsymbol{r}_1 + \boldsymbol{r}_3)$. The column and row combinations that produce $\mathbf{0}$ are the same: this is unusual. Two solutions to $y_1\boldsymbol{r}_1 + y_2\boldsymbol{r}_2 + y_3\boldsymbol{r}_3 = \mathbf{0}$ are $(Y_1, Y_2, Y_3) = (1, -2, 1)$ and $(2, -4, 2)$.

6 $c = \mathbf{3}$ $\quad \begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 7 & 4 & \mathbf{3} \end{bmatrix}$ has column $3 =$ column $1 -$ column 2

$c = \mathbf{-1}$ $\quad \begin{bmatrix} 1 & 0 & \mathbf{-1} \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ has column $3 = -$ column $1 +$ column 2

$c = \mathbf{0}$ $\quad \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 2 & 1 & 5 \\ 3 & 3 & 6 \end{bmatrix}$ has column $3 = 3$ (column 1) $-$ column 2

7 All three rows are perpendicular to the solution $\boldsymbol{x}$ (the three equations $\boldsymbol{r}_1 \cdot \boldsymbol{x} = 0$ and $\boldsymbol{r}_2 \cdot \boldsymbol{x} = 0$ and $\boldsymbol{r}_3 \cdot \boldsymbol{x} = 0$ tell us this). Then the whole plane of the rows is perpendicular to $\boldsymbol{x}$ (the plane is also perpendicular to all multiples $c\boldsymbol{x}$).

8
$$\begin{aligned} x_1 - 0 &= b_1 \\ x_2 - x_1 &= b_2 \\ x_3 - x_2 &= b_3 \\ x_4 - x_3 &= b_4 \end{aligned} \qquad \begin{aligned} x_1 &= b_1 \\ x_2 &= b_1 + b_2 \\ x_3 &= b_1 + b_2 + b_3 \\ x_4 &= b_1 + b_2 + b_3 + b_4 \end{aligned} \quad = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = A^{-1}\boldsymbol{b}$$

9 The cyclic difference matrix C has a line of solutions (in 4 dimensions) to $C\boldsymbol{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ when } \boldsymbol{x} = \begin{bmatrix} c \\ c \\ c \\ c \end{bmatrix} = \text{ any constant vector.}$$

10
$$\begin{aligned} z_2 - z_1 &= b_1 \\ z_3 - z_2 &= b_2 \\ 0 - z_3 &= b_3 \end{aligned} \qquad \begin{aligned} z_1 &= -b_1 - b_2 - b_3 \\ z_2 &= \quad\;\; - b_2 - b_3 \\ z_3 &= \qquad\quad\;\; - b_3 \end{aligned} = \begin{bmatrix} -1 & -1 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \Delta^{-1}\boldsymbol{b}$$

11 The forward differences of the squares are $(t+1)^2 - t^2 = t^2 + 2t + 1 - t^2 = 2t + 1$. Differences of the nth power are $(t+1)^n - t^n = t^n - t^n + nt^{n-1} + \cdots$. The leading term is the derivative nt^{n-1}. The binomial theorem gives all the terms of $(t+1)^n$.

12 Centered difference matrices of *even size* seem to be invertible. Look at eqns. 1 and 4:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \quad \begin{matrix} \text{First} \\ \text{solve} \\ x_2 = b_1 \\ -x_3 = b_4 \end{matrix} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -b_2 - b_4 \\ b_1 \\ -b_4 \\ b_1 + b_3 \end{bmatrix}$$

13 *Odd size*: The five centered difference equations lead to $b_1 + b_3 + b_5 = 0$.

$$\begin{aligned} x_2 \qquad &= b_1 \\ x_3 - x_1 &= b_2 \\ x_4 - x_2 &= b_3 \\ x_5 - x_3 &= b_4 \\ - x_4 &= b_5 \end{aligned}$$

Add equations $1, 3, 5$

The left side of the sum is zero

The right side is $b_1 + b_3 + b_5$

There cannot be a solution unless $b_1 + b_3 + b_5 = 0$.

14 An example is $(a, b) = (3, 6)$ and $(c, d) = (1, 2)$. We are given that the ratios a/c and b/d are equal. Then $ad = bc$. Then (when you divide by bd) the ratios a/b and c/d must also be equal!

Problem Set 2.1, page 41

1 The row picture for $A = I$ has 3 perpendicular planes $x = 2$ and $y = 3$ and $z = 4$. Those are perpendicular to the x and y and z axes: $z = 4$ is a horizontal plane at height 4.

The column vectors are $\boldsymbol{i} = (1, 0, 0)$ and $\boldsymbol{j} = (0, 1, 0)$ and $\boldsymbol{k} = (0, 0, 1)$. Then $\boldsymbol{b} = (2, 3, 4)$ is the linear combination $2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}$.

2 The planes in a row picture are the same: $2x = 4$ is $x = 2$, $3y = 9$ is $y = 3$, and $4z = 16$ is $z = 4$. The solution is the same point $\boldsymbol{X} = \boldsymbol{x}$. The three column vectors are changed; but the same combination (coefficients z, produces $\boldsymbol{b} = 34$), $(4, 9, 16)$.

3 The solution is not changed! The second plane and row 2 of the matrix and all columns of the matrix (vectors in the column picture) are changed.

4 If $z = 2$ then $x + y = 0$ and $x - y = 2$ give the point $(x, y, z) = (1, -1, 2)$. If $z = 0$ then $x + y = 6$ and $x - y = 4$ produce $(5, 1, 0)$. Halfway between those is $(3, 0, 1)$.

5 If x, y, z satisfy the first two equations they also satisfy the third equation $=$ sum of the first two. The line **L** of solutions contains $\boldsymbol{v} = (1, 1, 0)$ and $\boldsymbol{w} = (\frac{1}{2}, 1, \frac{1}{2})$ and $\boldsymbol{u} = \frac{1}{2}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}$ and all combinations $c\boldsymbol{v} + d\boldsymbol{w}$ with $c + d = 1$. (Notice that requirement $c + d = 1$. If you allow all c and d, you get a plane.)

6 Equation $1 +$ equation $2 -$ equation 3 is now $0 = -4$. The intersection line L of planes 1 and 2 misses plane 3: *no solution*.

7 Column $3 =$ Column 1 makes the matrix singular. For $\boldsymbol{b} = (2, 3, 5)$ the solutions are $(x, y, z) = (1, 1, 0)$ or $(0, 1, 1)$ and you can add any multiple of $(-1, 0, 1)$. $\boldsymbol{b} = (4, 6, c)$ needs $c = 10$ for solvability (then $\boldsymbol{b}$ lies in the plane of the columns and the three equations add to $0 = 0$).

8 Four planes in 4-dimensional space normally meet at a *point*. The solution to $A\boldsymbol{x} = (3, 3, 3, 2)$ is $\boldsymbol{x} = (0, 0, 1, 2)$ if A has columns $(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)$. The equations are $x + y + z + t = 3, y + z + t = 3, z + t = 3, t = 2$. Solve them in reverse order!

9 (a) $A\boldsymbol{x} = (18, 5, 0)$ and (b) $A\boldsymbol{x} = (3, 4, 5, 5)$.

10 Multiplying as linear combinations of the columns gives the same $A\boldsymbol{x} = (18, 5, 0)$ and $(3, 4, 5, 5)$. By rows or by columns: **9** separate multiplications when A is 3 by 3.

11 $A\boldsymbol{x}$ equals $(14, 22)$ and $(0, 0)$ and $(9, 7)$.

12 $A\boldsymbol{x}$ equals (z, y, x) and $(0, 0, 0)$ and $(3, 3, 6)$.

13 (a) $\boldsymbol{x}$ has n components and $A\boldsymbol{x}$ has m components (b) Planes from each equation in $A\boldsymbol{x} = \boldsymbol{b}$ are in n-dimensional space. The columns of A are in m-dimensional space.

14 $2x+3y+z+5t = 8$ is $A\boldsymbol{x} = \boldsymbol{b}$ with the 1 by 4 matrix $A = [\,2 \quad 3 \quad 1 \quad 5\,]$: *one row*. The solutions (x, y, z, t) fill a 3D "plane" in 4 dimensions. It could be called a *hyperplane*.

15 (a) $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ = "identity" (b) $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ = "permutation"

16 90° rotation from $R = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, 180° rotation from $R^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I$.

17 $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ produces $\begin{bmatrix} y \\ z \\ x \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ recovers $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Q is the *inverse* of P. Later we write $QP = I$ and $Q = P^{-1}$.

18 $E = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ and $E = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ subtract the first component from the second.

19 $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ and $E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$, $E\boldsymbol{v} = \begin{bmatrix} 3 \\ 4 \\ 8 \end{bmatrix}$ and $E^{-1}E\boldsymbol{v}$ recovers $\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$.

20 $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ projects onto the x-axis and $P_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ projects onto the y-axis. The vector $\boldsymbol{v} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ projects to $P_1\boldsymbol{v} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ and $P_2P_1\boldsymbol{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

21 $R = \frac{1}{2}\begin{bmatrix}\sqrt{2} & -\sqrt{2}\\ \sqrt{2} & \sqrt{2}\end{bmatrix}$ rotates all vectors by 45° . The columns of R are the results from rotating $(1,0)$ and $(0,1)$!

22 The dot product $A\boldsymbol{x} = [\,1 \;\; 4 \;\; 5\,]\begin{bmatrix}x\\ y\\ z\end{bmatrix}$ = (1 by 3)(3 by 1) is zero for points (x, y, z) on a plane in three dimensions. The 3 columns of A are one-dimensional vectors.

23 $A = [\,1 \;\; 2 \;\; ; \;\; 3 \;\; 4\,]$ and $\boldsymbol{x} = [\,5 \;\; -2\,]'$ or $[\,5 \;\; ; \;\; -2\,]$ and $\boldsymbol{b} = [\,1 \;\; 7\,]'$ or $[\,1 \;\; ; \;\; 7\,]$. $\boldsymbol{r} = \boldsymbol{b} - A * \boldsymbol{x}$ prints as two zeros.

24 $A * \boldsymbol{v} = [\,3 \;\; 4 \;\; 5\,]'$ and $\boldsymbol{v}' * \boldsymbol{v} = 50$. But $\boldsymbol{v} * A$ gives an error message from 3 by 1 times 3 by 3.

25 **ones**$(4,4)$ $*$ **ones**$(4,1)$ = column vector $[\,4 \;\; 4 \;\; 4 \;\; 4\,]'$; $B * \boldsymbol{w} = [\,10 \;\; 10 \;\; 10 \;\; 10\,]'$.

26 The row picture has two lines meeting at the solution $(4,2)$. The column picture will have $4(1,1) + 2(-2,1) =$ **4**(column 1) + **2**(column 2) = right side $(0,6)$.

27 The row picture shows **2 planes** in **3-dimensional space**. The column picture is in **2-dimensional space**. The solutions normally fill a ***line in*** 3-***dimensional space***.

28 The row picture shows four *lines* in the 2D plane. The column picture is in *four*-dimensional space. No solution unless the right side is a combination of *the two columns*.

29 $\boldsymbol{u}_2 = \begin{bmatrix}.7\\ .3\end{bmatrix}$ and $\boldsymbol{u}_3 = \begin{bmatrix}.65\\ .35\end{bmatrix}$. The components add to 1. They are always positive. Their components still add to 1.

30 $\boldsymbol{u}_7$ and $\boldsymbol{v}_7$ have components adding to 1; they are close to $\boldsymbol{s} = (.6, .4)$. $\begin{bmatrix}.8 & .3\\ .2 & .7\end{bmatrix}\begin{bmatrix}.6\\ .4\end{bmatrix} = \begin{bmatrix}.6\\ .4\end{bmatrix}$ = *steady state* $\boldsymbol{s}$. No change when multiplied by $\begin{bmatrix}.8 & .3\\ .2 & .7\end{bmatrix}$.

31 $M = \begin{bmatrix}8 & 3 & 4\\ 1 & 5 & 9\\ 6 & 7 & 2\end{bmatrix} = \begin{bmatrix}5+u & 5-u+v & 5-v\\ 5-u-v & 5 & 5+u+v\\ 5+v & 5+u-v & 5-u\end{bmatrix}$; $M_3(1,1,1) = (15,15,15)$; $M_4(1,1,1,1) = (34,34,34,34)$ because $1 + 2 + \cdots + 16 = 136$ which is $4(34)$.

32 A is singular when its third column $\boldsymbol{w}$ is a combination $c\boldsymbol{u} + d\boldsymbol{v}$ of the first columns. A typical column picture has $\boldsymbol{b}$ outside the plane of $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$. A typical row picture has the intersection line of two planes parallel to the third plane. *Then no solution.*

33 $\boldsymbol{w} = (5,7)$ is $5\boldsymbol{u} + 7\boldsymbol{v}$. Then $A\boldsymbol{w}$ equals 5 times $A\boldsymbol{u}$ plus 7 times $A\boldsymbol{v}$. **Linearity** means: When w is a combination of $\boldsymbol{u}$ and $\boldsymbol{v}$, then $A\boldsymbol{w}$ is the same combination of $A\boldsymbol{u}$ and $A\boldsymbol{v}$.

34 $\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ has the solution $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \\ 8 \\ 6 \end{bmatrix}$.

35 $\boldsymbol{x} = (1, \ldots, 1)$ gives $S\boldsymbol{x} =$ sum of each row $= 1 + \cdots + 9 = 45$ for Sudoku matrices. 6 row orders $(1,2,3)$, $(1,3,2)$, $(2,1,3)$, $(2,3,1)$, $(3,1,2)$, $(3,2,1)$ are in Section 2.7. The same 6 permutations of *blocks* of rows produce Sudoku matrices, so $6^4 = 1296$ orders of the 9 rows all stay Sudoku. (And also 1296 permutations of the 9 columns.)

Problem Set 2.2, page 53

1 Multiply equation 1 by $\ell_{21} = \frac{10}{2} = \mathbf{5}$ and subtract from equation 2 to find $2x + 3y = 1$ (unchanged) and $-6y = 6$. The pivots to circle are 2 and -6.

2 $-6y = 6$ gives $y = -1$. Then $2x + 3y = 1$ gives $x = 2$. Multiplying the right side $(1, 11)$ by 4 will multiply the solution by 4 to give the new solution $(x, y) = (8, -4)$.

3 Subtract $-\frac{1}{2}$ (or add $\frac{1}{2}$) times equation 1. The new second equation is $\mathbf{3}\boldsymbol{y} = \mathbf{3}$. Then $y = 1$ and $x = 5$. If the right side changes sign, so does the solution: $(x, y) = (-5, -1)$.

4 Subtract $\ell = \frac{c}{a}$ times equation 1 from equation 2. The new second pivot multiplying y is $d - (cb/a)$ or $(ad - bc)/a$. Then $y = (ag - cf)/(ad - bc)$. Notice the "determinant of A" $= ad - bc$. It must be nonzero for this division.

5 $6x + 4y$ is 2 times $3x + 2y$. There is no solution unless the right side is $2 \cdot 10 = 20$. Then all the points on the line $3x+2y = 10$ are solutions, including $(0, 5)$ and $(4, -1)$. The two lines in the row picture are the same line, containing all solutions.

6 Singular system if $b = 4$, because $4x + 8y$ is 2 times $2x + 4y$. Then $g = 32$ makes the lines $2x + 4y = 16$ and $4x + 8y = 32$ become the *same*: infinitely many solutions like $(8, 0)$ and $(0, 4)$.

7 If $a = 2$ elimination must fail (two parallel lines in the row picture). The equations have no solution. With $a = 0$, elimination will stop for a row exchange. Then $3y = -3$ gives $y = -1$ and $4x + 6y = 6$ gives $x = 3$.

8 If $k = 3$ elimination must fail: no solution. If $k = -3$, elimination gives $0 = 0$ in equation 2: infinitely many solutions. If $k = 0$ a row exchange is needed: one solution.

9 On the left side, $6x - 4y$ is 2 times $(3x - 2y)$. Therefore we need $b_2 = 2b_1$ on the right side. Then there will be infinitely many solutions (two parallel lines become one single line in the row picture). The column picture has both columns along the same line.

10 The equation $y = 1$ comes from elimination (subtract $x + y = 5$ from $x + 2y = 6$). Then $x = 4$ and $5x - 4y = 20 - 4 = c = 16$.

11 (a) Another solution is $\frac{1}{2}(x+X, y+Y, z+Z)$. (b) If 25 planes meet at two points, they meet along the whole line through those two points.

12 Elimination leads to this upper triangular system; then comes back substitution.

$$\begin{aligned} 2x + 3y + z &= 8 \\ y + 3z &= 4 \\ 8z &= 8 \end{aligned} \quad \text{gives} \quad \begin{aligned} x &= 2 \\ y &= 1 \\ z &= 1 \end{aligned}$$

If a zero is at the start of row 2 or row 3, that avoids a row operation.

13

$$\begin{aligned} 2x - 3y \phantom{{}+z} &= 3 \\ 4x - 5y + z &= 7 \\ 2x - y - 3z &= 5 \end{aligned} \quad \text{gives} \quad \begin{aligned} 2x - 3y &= 3 \\ y + z &= 1 \\ 2y + 3z &= 2 \end{aligned} \quad \text{and} \quad \begin{aligned} 2x - 3y &= 3 \\ y + z &= 1 \\ -5z &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} x &= 3 \\ y &= 1 \\ z &= 0 \end{aligned}$$

Here are steps $1, 2, 3$: Subtract 2 $\times$ row 1 from row 2, subtract 1 $\times$ row 1 from row 3, subtract 2 $\times$ row 2 from row 3

14 Subtract 2 times row 1 from row 2 to reach $(d-10)y-z=2$. Equation (3) is $y-z=3$. If $d=10$ exchange rows 2 and 3. If $d=11$ the system becomes singular.

15 The second pivot position will contain $-2-b$. If $b=-2$ we exchange with row 3. If $b=-1$ (singular case) the second equation is $-y-z=0$. But equation (3) is the same so there is a *line of solutions* $(x,y,z)=(1,1,-1)$.

16 (a) **Example of 2 exchanges**

$$\begin{aligned} 0x+0y+2z&=4\\ x+2y+2z&=5\\ 0x+3y+4z&=6 \end{aligned}$$

(exchange 1 and 2, then 2 and 3)

(b) **Exchange but then breakdown**

$$\begin{aligned} 0x+3y+4z&=4\\ x+2y+2z&=5\\ 0x+3y+4z&=6 \end{aligned}$$

(rows 1 and 3 are not consistent)

17 If row $1=$ row 2, then row 2 is zero after the first step; exchange the zero row with row 3 and row 3 has no pivot. If column $2=$ column 1, then column 2 has no pivot.

18 *Example* $x+2y+3z=0$, $4x+8y+12z=0$, $5x+10y+15z=0$ has 9 different coefficients but rows 2 and 3 become $0=0$: infinitely many solutions to $A\boldsymbol{x}=\mathbf{0}$ but almost surely no solution to $A\boldsymbol{x}=\boldsymbol{b}$ for a random $\boldsymbol{b}$.

19 Row 2 becomes $3y-4z=5$, then row 3 becomes $(q+4)z=t-5$. If $q=-4$ the system is singular—no third pivot. Then if $t=5$ the third equation is $0=0$ which allows infinitely many solutions. Choosing $z=1$ the equation $3y-4z=5$ gives $y=3$ and equation 1 gives $x=-9$.

20 Singular if row 3 is a combination of rows 1 and 2. From the end view, the three planes form a triangle. This happens if rows $1{+}2{=}$row 3 on the left side but not the right side: $x+y+z=0$, $x-2y-z=1$, $2x-y=4$. No parallel planes but still no solution. The three planes in the row picture form a triangular tunnel.

21 (a) Pivots $2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}$ in the equations $2x+y=0, \frac{3}{2}y+z=0, \frac{4}{3}z+t=0, \frac{5}{4}t=5$ after elimination. Back substitution gives $t=4, z=-3, y=2, x=-1$. (b) If the off-diagonal entries change from $+1$ to -1, the pivots are the same. The solution is $(1,2,3,4)$ instead of $(-1,2,-3,4)$.

22 The fifth pivot is $\frac{6}{5}$ for both matrices (1's or -1's off the diagonal). The nth pivot is $\frac{n+1}{n}$.

23 If ordinary elimination leads to $x + y = 1$ and $2y = 3$, the original second equation could be $2y + \ell(x + y) = 3 + \ell$ for any ℓ. Then ℓ will be the multiplier to reach $2y = 3$, by subtracting ℓ times equation 1 from equation 2.

24 Elimination fails on $\begin{bmatrix} a & 2 \\ a & a \end{bmatrix}$ if $a = 2$ or $a = 0$. (You could notice that the determinant $a^2 - 2a$ is zero for $a = 2$ and $a = 0$.)

25 $a = 2$ (equal columns), $a = 4$ (equal rows), $a = 0$ (zero column).

26 Solvable for $s = 10$ (add the two pairs of equations to get $a+b+c+d$ on the left sides, 12 and $2 + s$ on the right sides). So 12 must agree with $2 + s$, which makes $s = 10$. The four equations for a, b, c, d are **singular**! Two solutions are $\begin{bmatrix} 1 & 3 \\ 1 & 7 \end{bmatrix}$ and $\begin{bmatrix} 0 & 4 \\ 2 & 6 \end{bmatrix}$,

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

27 Elimination leaves the diagonal matrix diag$(3, 2, 1)$ in $3x = 3, 2y = 2, z = 2$. Then $x = 1, y = 1, z = 2$.

28 $A(2, :) = A(2, :) - 3 * A(1, :)$ subtracts 3 times row 1 from row 2.

29 The average pivots for rand(3) *without* row exchanges were $\frac{1}{2}, 5, 10$ in one experiment—but pivots 2 and 3 can be arbitrarily large. Their averages are actually infinite ! *With row exchanges* in MATLAB's **lu** code, the averages $.75$ and $.50$ and $.365$ are much more stable (and should be predictable, also for **randn** with normal instead of uniform probability distribution for the numbers in A).

30 If $A(5, 5)$ is 7 not 11, then the last pivot will be 0 not 4.

31 Row j of U is a combination of rows $1, \ldots, j$ of A (when there are no row exchanges). If $A\boldsymbol{x} = \mathbf{0}$ then $U\boldsymbol{x} = \mathbf{0}$ (not true if $\boldsymbol{b}$ replaces $\mathbf{0}$). U just keeps the diagonal of A when A is *lower triangular.*

32 The question deals with 100 equations $A\boldsymbol{x} = \mathbf{0}$ when A is singular.

(a) Some linear combination of the 100 rows is **the row of 100 zeros**.

(b) Some linear combination of the 100 **columns** is **the column of zeros**.

(c) A very singular matrix has all ones: $A =$ **ones**(100). A better example has 99 random rows (or the numbers $1^i, \ldots, 100^i$ in those rows). The 100th row could be the sum of the first 99 rows (or any other combination of those rows with no zeros).

(d) The row picture has 100 planes **meeting along a common line through 0**. The column picture has 100 vectors all in the same 99-dimensional hyperplane.

Problem Set 2.3, page 66

1 $E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 7 & 1 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

2 $E_{32}E_{21}\boldsymbol{b} = (1, -5, -35)$ but $E_{21}E_{32}\boldsymbol{b} = (1, -5, 0)$. When E_{32} comes first, row 3 feels no effect from row 1.

3 $\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ $M = E_{32}E_{31}E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 10 & -2 & 1 \end{bmatrix}$.

Those E's are in the right order to give $MA = U$.

4 Elimination on column 4: $\boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \overset{E_{21}}{\rightarrow} \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix} \overset{E_{31}}{\rightarrow} \begin{bmatrix} 1 \\ -4 \\ 2 \end{bmatrix} \overset{E_{32}}{\rightarrow} \begin{bmatrix} 1 \\ -4 \\ 10 \end{bmatrix}$. The original $A\boldsymbol{x} = \boldsymbol{b}$ has become $U\boldsymbol{x} = \boldsymbol{c} = (1, -4, 10)$. Then back substitution gives $\boldsymbol{z} = -5, y = \frac{1}{2}, x = \frac{1}{2}$. This solves $A\boldsymbol{x} = (1, 0, 0)$.

5 Changing a_{33} from 7 to 11 will change the third pivot from 5 to 9. Changing a_{33} from 7 to 2 will change the pivot from 5 to *no pivot*.

6 Example: $\begin{bmatrix} 2 & 3 & 7 \\ 2 & 3 & 7 \\ 2 & 3 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}$. If all columns are multiples of column 1, there is no second pivot.

7 To reverse E_{31}, **add** 7 times row **1** to row **3**. The inverse of the elimination matrix $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\mathbf{7} & 0 & 1 \end{bmatrix}$ is $E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{7} & 0 & 1 \end{bmatrix}$. Multiplication confirms $EE^{-1} = I$.

8 $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $M* = \begin{bmatrix} a & b \\ c - \ell a & d - \ell b \end{bmatrix}$. $\det M* = a(d - \ell b) - b(c - \ell a)$ reduces to $ad - bc$! Subtracting row 1 from row 2 doesn't change $\det M$.

9 $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$. After the exchange, we need E_{31} (not E_{21}) to act on the new row 3.

10 $E_{13} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $E_{31}E_{13} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Test on the identity matrix!

11 An example with two negative pivots is $A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$. The diagonal entries can change sign during elimination.

12 The first product is $\begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$ rows and also columns reversed. The second product is $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & 2 & -3 \end{bmatrix}$.

13 (a) E times the third column of B is the third column of EB. A column that starts at zero will stay at zero. (b) E could add row 2 to row 3 to change a zero row to a nonzero row.

14 E_{21} has $-\ell_{21} = \frac{1}{2}$, E_{32} has $-\ell_{32} = \frac{2}{3}$, E_{43} has $-\ell_{43} = \frac{3}{4}$. Otherwise the E's match I.

15 $a_{ij} = 2i - 3j$: $A = \begin{bmatrix} -1 & -4 & -7 \\ 1 & -2 & -5 \\ 3 & \mathbf{0} & -3 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & -4 & -7 \\ 0 & -6 & -12 \\ 0 & \mathbf{-12} & -24 \end{bmatrix}$. The zero became -12, an example of *fill-in*. To remove that -12, choose $E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$.

Every 3 by 3 matrix with entries $a_{ij} = ci + dj$ is **singular**!

16 (a) The ages of X and Y are x and y: $x - 2y = 0$ and $x + y = 33$; $x = 22$ and $y = 11$
(b) The line $y = mx + c$ contains $x = 2$, $y = 5$ and $x = 3$, $y = 7$ when $2m + c = 5$ and $3m + c = 7$. Then $m = 2$ is the slope.

17 The parabola $y = a + bx + cx^2$ goes through the 3 given points when $\begin{aligned} a + \ b + \ c &= 4 \\ a + 2b + 4c &= 8 \\ a + 3b + 9c &= 14 \end{aligned}$.

Then $a = 2$, $b = 1$, and $c = 1$. This matrix with columns $(1,1,1)$, $(1,2,3)$, $(1,4,9)$ is a "Vandermonde matrix."

18 $EF = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix}$, $FE = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b+ac & c & 1 \end{bmatrix}$, $E^2 = \begin{bmatrix} 1 & 0 & 0 \\ 2a & 1 & 0 \\ 2b & 0 & 1 \end{bmatrix}$, $F^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3c & 1 \end{bmatrix}$.

19 $PQ = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. In the opposite order, two row exchanges give $QP = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $P^2 = I$. If M exchanges rows 2 and 3 then $M^2 = I$ (also $(-M)^2 = I$). There are many square roots of I: Any matrix $M = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ has $M^2 = I$ if $a^2 + bc = 1$.

20 (a) Each column of EB is E times a column of B (b) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{4} \\ \mathbf{1} & \mathbf{2} & \mathbf{4} \end{bmatrix} =$ $\begin{bmatrix} 1 & 2 & 4 \\ \mathbf{2} & \mathbf{4} & \mathbf{8} \end{bmatrix}$. All rows of EB are *multiples* of $\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$.

21 **No.** $E = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ give $EF = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ but $FE = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

22 (a) $\sum a_{3j} x_j$ (b) $a_{21} - a_{11}$ (c) $a_{21} - 2a_{11}$ (d) $(EA\boldsymbol{x})_1 = (A\boldsymbol{x})_1 = \sum a_{1j} x_j$.

23 $E(EA)$ subtracts 4 times row 1 from row 2 (EEA does the row operation twice). AE subtracts 2 times column 2 of A from column 1 (multiplication by E on the right side acts on **columns** instead of rows).

24 $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} 2 & 3 & \mathbf{1} \\ 4 & 1 & \mathbf{17} \end{bmatrix} \to \begin{bmatrix} 2 & 3 & \mathbf{1} \\ 0 & -5 & \mathbf{15} \end{bmatrix}$. The triangular system is $\begin{aligned} 2x_1 + 3x_2 &= 1 \\ -5x_2 &= 15 \end{aligned}$

Back substitution gives $x_1 = 5$ and $x_2 = -3$.

25 The last equation becomes $0 = 3$. If the original 6 is 3, then row 1 + row 2 = row 3. Then the last equation is $0 = 0$ and the system has infinitely many solutions.

26 (a) Add two columns $\boldsymbol{b}$ and $\boldsymbol{b}^*$ to get $[A \;\; \boldsymbol{b} \;\; \boldsymbol{b}^*]$. The example has

$$\begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{bmatrix} \to \boldsymbol{x} = \begin{bmatrix} -7 \\ 2 \end{bmatrix} \text{ and } \boldsymbol{x}^* = \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$$

27 (a) No solution if $d = 0$ and $c \neq 0$ (b) Many solutions if $d = 0 = c$. No effect from a, b.

28 $A = AI = A(BC) = (AB)C = IC = C$. That middle equation is crucial.

29 $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$ subtracts each row from the next row. The result $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}$ still has multipliers $= 1$ in a 3 by 3 Pascal matrix. The product M of all elimination matrices is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix}$. This "alternating sign Pascal matrix" is on page 91.

30 (a) $E = A^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ will reduce row 2 of EM to $[2 \;\; 3]$.

(b) Then $F = B^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ will reduce row 1 of FEM to $[1 \;\; 1]$.

(c) Then $E = A^{-1}$ twice will reduce row 2 of $EEFEM$ to $[0 \;\; 1]$

(d) Now $EEFEM = B$. Move E's and F's to get $\boldsymbol{M} = \boldsymbol{ABAAB}$. This question focuses on positive integer matrices M with $ad - bc = 1$. The same steps make the entries smaller and smaller until M is a product of A's and B's.

31 $E_{21} = \begin{bmatrix} 1 & & & \\ a & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $E_{32} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & b & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $E_{43} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & c & 1 \end{bmatrix}$,

$$E_{43}\, E_{32}\, E_{21} = \begin{bmatrix} 1 & & & \\ a & 1 & & \\ ab & b & 1 & \\ abc & bc & c & 1 \end{bmatrix}$$

Problem Set 2.4, page 77

1 If all entries of A, B, C, D are 1, then $BA = 3$ **ones**(5) is 5 by 5; $AB = 5$ **ones**(3) is 3 by 3; $ABD = 15$ **ones**$(3, 1)$ is 3 by 1. DC and $A(B + C)$ are not defined.

2 (a) A (column 2 of B) (b) (Row 1 of A) B (c) (Row 3 of A)(column 5 of B) (d) (Row 1 of C)D(column 1 of E).

3 $AB + AC$ is the same as $A(B + C) = \begin{bmatrix} 3 & 8 \\ 6 & 9 \end{bmatrix}$. (*Distributive law*).

4 $A(BC) = (AB)C$ by the *associative law*. In this example both answers are $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Column 1 of AB and row 2 of C are zero (then multiply columns times rows).

5 (a) $A^2 = \begin{bmatrix} 1 & 2b \\ 0 & 1 \end{bmatrix}$ and $A^n = \begin{bmatrix} 1 & nb \\ 0 & 1 \end{bmatrix}$. (b) $A^2 = \begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix}$ and $A^n = \begin{bmatrix} 2^n & 2^n \\ 0 & 0 \end{bmatrix}$.

6 $(A + B)^2 = \begin{bmatrix} 10 & 4 \\ 6 & 6 \end{bmatrix} = A^2 + AB + BA + B^2$. But $A^2 + 2AB + B^2 = \begin{bmatrix} 16 & 2 \\ 3 & 0 \end{bmatrix}$.

7 (a) True (b) False (c) True (d) False: usually $(AB)^2 = ABAB \neq A^2B^2$.

8 The rows of DA are 3 (row 1 of A) and 5 (row 2 of A). Both rows of EA are row 2 of A. The columns of AD are 3 (column 1 of A) and 5 (column 2 of A). The first column of AE is zero, the second is column 1 of A + column 2 of A.

9 $AF = \begin{bmatrix} a & a+b \\ c & c+d \end{bmatrix}$ and $E(AF)$ equals $(EA)F$ because matrix multiplication is *associative*.

10 $FA = \begin{bmatrix} a+c & b+d \\ c & d \end{bmatrix}$ and then $E(FA) = \begin{bmatrix} a+c & b+d \\ a+2c & b+2d \end{bmatrix}$. $E(FA)$ is not the same as $F(EA)$ because multiplication is not commutative: $EF \neq FE$.

11 Suppose EA does the row operation and then $(EA)F$ does the column operation (because F is multiplying from the right). The associative law says that $(EA)F = E(AF)$ so the column operation can be done first!

12 (a) $B = 4I$ (b) $B = 0$ (c) $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (d) Every row of B is $1, 0, 0$.

13 $AB = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = BA = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ gives $\boldsymbol{b} = \boldsymbol{c} = \mathbf{0}$. Then $AC = CA$ gives $\boldsymbol{a} = \boldsymbol{d}$. The only matrices that commute with B and C (and all other matrices) are multiples of I: $A = aI$.

14 $(A-B)^2 = (B-A)^2 = A(A-B) - B(A-B) = A^2 - AB - BA + B^2$. In a typical case (when $AB \neq BA$) the matrix $A^2 - 2AB + B^2$ is different from $(A-B)^2$.

15 (a) True (A^2 is only defined when A is square).

(b) False (if A is m by n and B is n by m, then AB is m by m and BA is n by n).

(c) True by part (b).

(d) False (take $B = 0$).

16 (a) mn (use every entry of A) (b) $mnp = p\times$part (a) (c) n^3 (n^2 dot products).

17 (a) Use only column 2 of B (b) Use only row 2 of A (c)–(d) Use row 2 of first A.

Column 2 of $AB = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ Row 2 of $AB = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ Row 2 of $A^2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$

Row 2 of $A^3 = \begin{bmatrix} 3 & -2 \end{bmatrix}$

18 $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ has $a_{ij} = \min(i,j)$. $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ has $a_{ij} = (-1)^{i+j} =$ "alternating sign matrix". $A = \begin{bmatrix} 1/1 & 1/2 & 1/3 \\ 2/1 & 2/2 & 2/3 \\ 3/1 & 3/2 & 3/3 \end{bmatrix}$ has $a_{ij} = i/j$. This will be an example of a *rank one matrix*: 1 column $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^{\mathrm{T}}$ multiplies 1 row $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$.

19 Diagonal matrix, lower triangular, symmetric, all rows equal. Zero matrix fits all four.

20 (a) a_{11} (b) $\ell_{31} = a_{31}/a_{11}$ (c) $a_{32} - \left(\dfrac{a_{31}}{a_{11}}\right) a_{12}$ (d) $a_{22} - \left(\dfrac{a_{21}}{a_{11}}\right) a_{12}$.

21 $A^2 = \begin{bmatrix} 0&0&4&0 \\ 0&0&0&4 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}$, $A^3 = \begin{bmatrix} 0&0&0&8 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}$, $A^4 =$ zero matrix for *strictly triangular* A.

Then $A\boldsymbol{v} = A\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 2y \\ 2z \\ 2t \\ 0 \end{bmatrix}$, $A^2\boldsymbol{v} = \begin{bmatrix} 4z \\ 4t \\ 0 \\ 0 \end{bmatrix}$, $A^3\boldsymbol{v} = \begin{bmatrix} 8t \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $A^4\boldsymbol{v} = \mathbf{0}$.

22 $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has $A^2 = -I$; $BC = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;

$DE = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = -ED$. You can find more examples.

23 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has $A^2 = 0$. Note: Any matrix $A =$ column times row $= \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ will have $A^2 = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} = 0$ if $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} = 0$. $A = \begin{bmatrix} 0&1&0 \\ 0&0&1 \\ 0&0&0 \end{bmatrix}$ has $A^2 = \begin{bmatrix} 0&0&1 \\ 0&0&0 \\ 0&0&0 \end{bmatrix}$ but $A^3 = 0$; strictly triangular as in Problem 21.

24 $(A_1)^n = \begin{bmatrix} 2^n & 2^n - 1 \\ 0 & 1 \end{bmatrix}$, $(A_2)^n = 2^{n-1}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $(A_3)^n = \begin{bmatrix} a^n & a^{n-1}b \\ 0 & 0 \end{bmatrix}$.

25 $\begin{bmatrix} a&b&c \\ d&e&f \\ g&h&i \end{bmatrix}\begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix} = \begin{bmatrix} a \\ d \\ g \end{bmatrix}\begin{bmatrix} 1&0&0 \end{bmatrix} + \begin{bmatrix} d \\ e \\ h \end{bmatrix}\begin{bmatrix} 0&1&0 \end{bmatrix} + \begin{bmatrix} c \\ f \\ i \end{bmatrix}\begin{bmatrix} 0&0&1 \end{bmatrix}$.

26 **Columns of A times rows of B** $\begin{bmatrix}1\\2\\2\end{bmatrix}\begin{bmatrix}3 & 3 & 0\end{bmatrix}+\begin{bmatrix}0\\4\\1\end{bmatrix}\begin{bmatrix}1 & 2 & 1\end{bmatrix}=\begin{bmatrix}3&3&0\\6&6&0\\6&6&0\end{bmatrix}+\begin{bmatrix}0&0&0\\4&8&4\\1&2&1\end{bmatrix}=$

$$\begin{bmatrix}3&3&0\\10&14&4\\7&8&1\end{bmatrix}=AB.$$

27 (a) (row 3 of A) $\cdot$ (column 1 or 2 of B) and (row 3 of A) $\cdot$ (column 2 of B) are all zero.

(b) $\begin{bmatrix}x\\x\\0\end{bmatrix}\begin{bmatrix}0 & x & x\end{bmatrix}=\begin{bmatrix}0&x&x\\0&x&x\\0&0&0\end{bmatrix}$ and $\begin{bmatrix}x\\x\\x\end{bmatrix}\begin{bmatrix}0 & 0 & x\end{bmatrix}=\begin{bmatrix}0&0&x\\0&0&x\\0&0&x\end{bmatrix}$: **both upper**.

28 A times B with cuts $\quad A\begin{bmatrix}\,|\,|\,|\,|\,\end{bmatrix},\quad \begin{bmatrix}\\ \text{———}\\ \\ \end{bmatrix}B,\quad \begin{bmatrix}\\ \text{———}\\ \\ \end{bmatrix}\begin{bmatrix}\,|\,|\,|\,|\,\end{bmatrix},\quad \begin{bmatrix}\,|\,|\,|\,\end{bmatrix}\begin{bmatrix}\text{———}\\ \\ \text{———}\end{bmatrix}$

4 cols — 2 rows — 2 rows – 4 cols — 3 cols – 3 rows

29 $E_{21}=\begin{bmatrix}1&0&0\\1&1&0\\0&0&1\end{bmatrix}$ and $E_{31}=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}$ produce zeros in the $2,1$ and $3,1$ entries.

Multiply E's to get $E=E_{31}E_{21}=\begin{bmatrix}1&0&0\\1&1&0\\-4&0&1\end{bmatrix}$. Then $EA=\begin{bmatrix}2&1&0\\0&1&1\\0&1&3\end{bmatrix}$ is the result of both E's since $(E_{31}E_{21})A=E_{31}(E_{21}A)$.

30 In **29**, $\boldsymbol{c}=\begin{bmatrix}-2\\8\end{bmatrix}$, $D=\begin{bmatrix}0&1\\5&3\end{bmatrix}$, $D-\boldsymbol{c}\boldsymbol{b}/a=\begin{bmatrix}1&1\\1&3\end{bmatrix}$ in the lower corner of EA.

31 $\begin{bmatrix}A&-B\\B&A\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\\boldsymbol{y}\end{bmatrix}=\begin{bmatrix}A\boldsymbol{x}-B\boldsymbol{y}\\B\boldsymbol{x}+A\boldsymbol{y}\end{bmatrix}$ real part / imaginary part. Complex matrix times complex vector needs **4** real times real multiplications.

32 A times $X=[\,\boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \boldsymbol{x}_3\,]$ will be the identity matrix $I=[\,A\boldsymbol{x}_1\ \ A\boldsymbol{x}_2\ \ A\boldsymbol{x}_3\,]$.

33 $\boldsymbol{b} = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}$ gives $\boldsymbol{x} = 3\boldsymbol{x}_1 + 5\boldsymbol{x}_2 + 8\boldsymbol{x}_3 = \begin{bmatrix} 3 \\ 8 \\ 16 \end{bmatrix}$; $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ will have those $\boldsymbol{x}_1 = (1,1,1), \boldsymbol{x}_2 = (0,1,1), \boldsymbol{x}_3 = (0,0,1)$ as columns of its "inverse" A^{-1}.

34 $A*$ **ones** $= \begin{bmatrix} a+b & a+b \\ c+d & c+d \end{bmatrix}$ agrees with **ones** $*A = \begin{bmatrix} a+c & b+b \\ a+c & b+d \end{bmatrix}$ when $b = c$ and $a = d$

Then $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$. These are the matrices that commute with $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

35 $S = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$, $S^2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix}$,

aba, ada	**cba, cda**	These show
bab, bcb	**dab, dcb**	16 2-step
abc, adc	**cbc, cdc**	paths in
bad, bcd	**dad, dcd**	the graph

36 Multiplying AB =(m by n)(n by p) needs mnp multiplications. Then $(AB)C$ needs mpq more. Multiply $BC =$ (n by p)(p by q) needs npq and then $A(BC)$ needs mnq.

(a) If m, n, p, q are $2, 4, 7, 10$ we compare $(2)(4)(7) + (2)(7)(10) = \mathbf{196}$ with the larger number $(2)(4)(10) + (4)(7)(10) = \mathbf{360}$. So AB first is better, we want to multiply that 7 by 10 matrix by as few rows as possible.

(b) If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are N by 1, then $(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v})\boldsymbol{w}^{\mathrm{T}}$ needs $2N$ multiplications but $\boldsymbol{u}^{\mathrm{T}}(\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}})$ needs N^2 to find $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ and N^2 more to multiply by the row vector $\boldsymbol{u}^{\mathrm{T}}$. Apologies to use the transpose symbol so early.

(c) We are comparing $mnp + mpq$ with $mnq + npq$. Divide all terms by $mnpq$: Now we are comparing $q^{-1} + n^{-1}$ with $p^{-1} + m^{-1}$. This yields a simple important rule. If matrices A and B are multiplying $\boldsymbol{v}$ for $AB\boldsymbol{v}$, **don't multiply the matrices first**. Better to multiply $B\boldsymbol{v}$ and then $A(B\boldsymbol{v})$.

37 The proof of $(AB)\boldsymbol{c} = A(B\boldsymbol{c})$ used the column rule for matrix multiplication—this rule is clearly linear, column by column.

Even for nonlinear transformations, $A(B(\boldsymbol{c}))$ would be the *"composition"* of A with B (applying B then A). This composition $A \circ B$ is just AB for matrices.

One of many uses for the associative law: The left-inverse B = right-inverse C from $B = B(AC) = (BA)C = C$.

38 (a) Multiply the columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ by the rows $\boldsymbol{a}_1^{\mathrm{T}}, \ldots, \boldsymbol{a}_m^{\mathrm{T}}$ and add the resulting matrices.

(b) $A^{\mathrm{T}}CA = c_1\boldsymbol{a}_1\boldsymbol{a}_1^{\mathrm{T}} + \cdots + c_m\boldsymbol{a}_m\boldsymbol{a}_m^{\mathrm{T}}$. Diagonal C makes it neat.

Problem Set 2.5, page 92

1 $A^{-1} = \begin{bmatrix} 0 & \frac{1}{4} \\ \frac{1}{3} & 0 \end{bmatrix}$ and $B^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{bmatrix}$ and $C^{-1} = \begin{bmatrix} 7 & -4 \\ -5 & 3 \end{bmatrix}$.

2 For the first, a simple row exchange has $P^2 = I$ so $P^{-1} = P$. For the second, $P^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Always P^{-1} = "transpose" of P, coming in Section 2.7.

3 $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} .5 \\ -.2 \end{bmatrix}$ and $\begin{bmatrix} t \\ z \end{bmatrix} = \begin{bmatrix} -.2 \\ .1 \end{bmatrix}$ so $A^{-1} = \frac{1}{10}\begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$. This question solved $AA^{-1} = I$ column by column, the main idea of Gauss-Jordan elimination. For a different matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, you could find a first column for A^{-1} but not a second column—so A would be singular (*no inverse*).

4 The equations are $x + 2y = 1$ and $3x + 6y = 0$. No solution because 3 times equation 1 gives $3x + 6y = 3$.

5 An upper triangular U with $U^2 = I$ is $U = \begin{bmatrix} 1 & a \\ 0 & -1 \end{bmatrix}$ for any a. And also $-U$.

6 (a) Multiply $AB = AC$ by A^{-1} to find $B = C$ (since A is invertible) (b) As long as $B - C$ has the form $\begin{bmatrix} x & y \\ -x & -y \end{bmatrix}$, we have $AB = AC$ for $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

7 (a) In $A\boldsymbol{x} = (1, 0, 0)$, equation 1 + equation 2 − equation 3 is $0 = 1$ (b) Right sides must satisfy $b_1 + b_2 = b_3$ (c) Row 3 becomes a row of zeros—no third pivot.

8 (a) The vector $\boldsymbol{x} = (1, 1, -1)$ solves $A\boldsymbol{x} = \mathbf{0}$ (b) After elimination, columns 1 and 2 end in zeros. Then so does column $3 =$ column $1 + 2$: no third pivot.

9 Yes, B is invertible (A was just multiplied by a permutation matrix P). If you exchange rows 1 and 2 of A to reach B, you exchange **columns** 1 and 2 of A^{-1} to reach B^{-1}. In matrix notation, $B = PA$ has $B^{-1} = A^{-1}P^{-1} = A^{-1}P$ for this P.

10 $A^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1/5 \\ 0 & 0 & 1/4 & 0 \\ 0 & 1/3 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \end{bmatrix}$ and $B^{-1} = \begin{bmatrix} 3 & -2 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & -7 & 6 \end{bmatrix}$ (invert each block of B)

11 (a) If $B = -A$ then certainly $A + B$ = zero matrix is not invertible.

(b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ are both singular but $A + B = I$ is invertible.

12 Multiply $C = AB$ on the left by A^{-1} and on the right by C^{-1}. Then $A^{-1} = BC^{-1}$.

13 $M^{-1} = C^{-1}B^{-1}A^{-1}$ so multiply on the left by C and the right by A : $B^{-1} = CM^{-1}A$.

14 $B^{-1} = A^{-1}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = A^{-1}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$: subtract *column* 2 of A^{-1} from *column* 1.

15 If A has a column of zeros, so does BA. Then $BA = I$ is impossible. There is no A^{-1}.

16 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix}$. The inverse of each matrix is the other divided by $ad-bc$

17 $E_{32}E_{31}E_{21} = \begin{bmatrix} 1 & & \\ & 1 & \\ & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -1 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -1 & 1 & \\ 0 & -1 & 1 \end{bmatrix} = E.$

Reverse the order and change -1 to $+1$ to get inverses $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & 1 & 1 \end{bmatrix} =$ $L = E^{-1}$. Notice the 1's unchanged by multiplying inverses in this order.

18 $A^2B = I$ can also be written as $A(AB) = I$. Therefore A^{-1} is AB.

19 The $(1,1)$ entry requires $4a - 3b = 1$; the $(1,2)$ entry requires $2b - a = 0$. Then $b = \frac{1}{5}$ and $a = \frac{2}{5}$. For the 5 by 5 case $5a - 4b = 1$ and $2b = a$ give $b = \frac{1}{6}$ and $a = \frac{2}{6}$.

20 $A * \text{ones}(4,1) = A$ (column of 1's) is the zero vector so A cannot be invertible.

21 Six of the sixteen $0-1$ matrices are invertible: I and P and all four with three 1's.

22 $\begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix};$

$$\begin{bmatrix} 1 & 4 & 1 & 0 \\ 3 & 9 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 1 & 0 \\ 0 & -3 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -3 & 4/3 \\ 0 & 1 & 1 & -1/3 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

23 $[\boldsymbol{A}\ \ \boldsymbol{I}] = \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & -1/2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \to$

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & -1/2 & 1 & 0 \\ 0 & 0 & 4/3 & 1/3 & -2/3 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3/2 & 0 & -3/4 & 3/2 & -3/4 \\ 0 & 0 & 4/3 & 1/3 & -2/3 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & -1 & 1/2 \\ 0 & 3/2 & 0 & -3/4 & 3/2 & -3/4 \\ 0 & 0 & 4/3 & 1/3 & -2/3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & -1/2 & 1/4 \\ 0 & 1 & 0 & -1/2 & 1 & -1/2 \\ 0 & 0 & 1 & 1/4 & -1/2 & 3/4 \end{array}\right] =$$

$[\boldsymbol{I}\ \boldsymbol{A}^{-1}]$.

24 $$\begin{bmatrix} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & a & 0 & 1 & 0 & -b \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & -a & ac-b \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

25 $$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}^{-\mathbf{1}} = \frac{1}{4}\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix};\quad B\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so B^{-1} does not exist.

26 $E_{21}A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. $E_{12}E_{21}A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.

Multiply by $D = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$ to reach $DE_{12}E_{21}A = I$. Then $A^{-1} = DE_{12}E_{21} = \frac{1}{2}\begin{bmatrix} 6 & -2 \\ -2 & 1 \end{bmatrix}$.

27 $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$ (notice the sign changes); $A^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

28 $$\begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 2 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & -1 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1/2 & 1/2 \\ 0 & 1 & 1/2 & 0 \end{bmatrix}.$$

This is $\begin{bmatrix} I & A^{-1} \end{bmatrix}$: row exchanges are certainly allowed in Gauss-Jordan.

29 (a) True (If A has a row of zeros, then every AB has too, and $AB = I$ is impossible).

(b) False (the matrix of all ones is singular even with diagonal 1's.

(c) True (the inverse of A^{-1} is A and the inverse of A^2 is $(A^{-1})^2$).

30 Elimination produces the pivots a and $a-b$ and $a-b$. $A^{-1} = \dfrac{1}{a(a-b)}\begin{bmatrix} a & 0 & -b \\ -a & a & 0 \\ 0 & -a & a \end{bmatrix}$.

The matrix C is not invertible if $c = 0$ or $c = 7$ or $c = 2$.

31 $A^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. When the triangular A has $1, -1, 1, -1, \ldots$ on successive diagonals, A^{-1} is *bidiagonal* with 1's on the diagonal and first superdiagonal.

32 $\boldsymbol{x} = (1, 1, \ldots, 1)$ has $\boldsymbol{x} = P\boldsymbol{x} = Q\boldsymbol{x}$ so $(P - Q)\boldsymbol{x} = \mathbf{0}$. Permutations do not change this all-ones vector.

33 $\begin{bmatrix} I & 0 \\ -C & I \end{bmatrix}$ and $\begin{bmatrix} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix}$ and $\begin{bmatrix} -D & I \\ I & 0 \end{bmatrix}$.

34 A can be invertible with diagonal zeros (example to find). B is singular because each row adds to zero. The all-ones vector $\boldsymbol{x}$ has $B\boldsymbol{x} = \mathbf{0}$.

35 The equation $LDLD = I$ says that $LD =$ pascal $(4, 1)$ is its own inverse.

36 hilb(6) is not the exact Hilbert matrix because fractions are rounded off. So inv(hilb(6)) is not the exact inverse either.

37 The three Pascal matrices have $P = LU = LL^{\mathrm{T}}$ and then inv$(P) =$ inv$(L^{\mathrm{T}})*$inv(L).

38 $A\boldsymbol{x} = \boldsymbol{b}$ has many solutions when $A =$ ones $(4, 4) =$ *singular* and $\boldsymbol{b} =$ ones $(4, 1)$. $A\backslash\boldsymbol{b}$ in MATLAB will pick the shortest solution $\boldsymbol{x} = (1, 1, 1, 1)/4$. This is the only solution that is a combination of the rows of A (later it comes from the "pseudoinverse" $A^+ =$ pinv(A) which replaces A^{-1} when A is singular). Any vector that solves $A\boldsymbol{x} = \mathbf{0}$ could be added to this particular solution $\boldsymbol{x}$.

39 The inverse of $A = \begin{bmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -b & 0 \\ 0 & 0 & 1 & -c \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is $A^{-1} = \begin{bmatrix} 1 & a & ab & abc \\ 0 & 1 & b & bc \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix}$. (This would be a good example for the cofactor formula $A^{-1} = C^{\mathrm{T}} / \det A$ in Section 5.3)

40 $$\begin{bmatrix} 1 & & & \\ a & 1 & & \\ b & 0 & 1 & \\ c & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & d & 1 & \\ 0 & e & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & f & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ a & 1 & & \\ b & d & 1 & \\ c & e & f & 1 \end{bmatrix}$$

In this order the multipliers a, b, c, d, e, f are unchanged in the product (**important for** $A = LU$ **in Section 2.6**).

41 4 by 4 still with $T_{11} = 1$ has pivots $1, 1, 1, 1$; reversing to $T^* = UL$ makes $T^*_{44} = 1$.

$$T = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \quad \text{and} \quad T^{-1} = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

42 Add the equations $C\boldsymbol{x} = \boldsymbol{b}$ to find $0 = b_1 + b_2 + b_3 + b_4$. So C is singular. Same for $F\boldsymbol{x} = \boldsymbol{b}$.

43 The block pivots are A and $S = D - CA^{-1}B$ (and $d - cb/a$ is the correct second pivot of an ordinary 2 by 2 matrix). The example problem has Schur complement $S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 3 & 3 \end{bmatrix} = \begin{bmatrix} -5 & -6 \\ -6 & -5 \end{bmatrix}$.

44 Inverting the identity $A(I + BA) = (I + AB)A$ gives $(I + BA)^{-1}A^{-1} = A^{-1}(I + AB)^{-1}$. So $I{+}BA$ and $I{+}AB$ are both invertible or both singular when A is invertible. (This remains true also when A is singular : Chapter 6 will show that AB and BA have the same nonzero eigenvalues, and we are looking here at the eigenvalue -1.)

Problem Set 2.6, page 104

1 $\ell_{21} = 1$ multiplied row 1; $L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ times $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \boldsymbol{c}$ is $A\boldsymbol{x} = \boldsymbol{b}$: $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. L multiplies $U\boldsymbol{x} = \boldsymbol{c}$ to give $A\boldsymbol{x} = \boldsymbol{b}$.

2 $L\boldsymbol{c} = \boldsymbol{b}$ is $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$, solved by $\boldsymbol{c} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ as elimination goes forward.

$U\boldsymbol{x} = \boldsymbol{c}$ is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, solved by $\boldsymbol{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ in back substitution.

3 $\ell_{31} = 1$ and $\ell_{32} = 2$ (and $\ell_{33} = 1$): reverse steps to get $A\boldsymbol{u} = \boldsymbol{b}$ from $U\boldsymbol{x} = \boldsymbol{c}$: 1 times $(x+y+z = 5)+2$ times $(y+2z = 2)+1$ times $(z = 2)$ gives $x+3y+6z = 11$.

4 $L\boldsymbol{c} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix}$; $\quad U\boldsymbol{x} = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 2 \\ & & 1 \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{x} \\ \\ \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$; $\boldsymbol{x} = \begin{bmatrix} 5 \\ -2 \\ 2 \end{bmatrix}$.

5 $EA = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 6 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix} = U.$

With E^{-1} as L, $A = LU = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 6 & 3 & 5 \end{bmatrix}$.

6 $\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -6 \end{bmatrix} = U$. Then $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} U$ is the same as $E_{21}^{-1} E_{32}^{-1} U = LU$. The multipliers $\ell_{21} = \ell_{32} = 2$ fall into place in L.

7 $E_{32}E_{31}E_{21}\ A = \begin{bmatrix} 1 & & \\ & 1 & \\ & -\mathbf{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ -\mathbf{3} & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -\mathbf{2} & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{bmatrix}$. This is $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = U$. Put those multipliers $2, 3, 2$ into L. Then $A = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{2} & 1 & 0 \\ \mathbf{3} & \mathbf{2} & 1 \end{bmatrix} U = LU$.

8 $E = E_{32}E_{31}E_{21} = \begin{bmatrix} 1 & & \\ -a & 1 & \\ ac-b & -c & 1 \end{bmatrix}$ is mixed but L is $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & & \\ a & 1 & \\ b & c & 1 \end{bmatrix}$.

9 2 by 2: $d = 0$ not allowed; $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ \ell & 1 & \\ m & n & 1 \end{bmatrix} \begin{bmatrix} d & e & g \\ & f & h \\ & & i \end{bmatrix}$ $d = 1, e = 1$, then $\ell = 1$; $f = 0$ is not allowed; **no pivot in row 2**

10 $c = 2$ leads to zero in the second pivot position: exchange rows and not singular. $c = 1$ leads to zero in the third pivot position. In this case the matrix is *singular*.

11 $A = \begin{bmatrix} 2 & 4 & 8 \\ 0 & 3 & 9 \\ 0 & 0 & 7 \end{bmatrix}$ has $L = I$ (A is already upper triangular) and $D = \begin{bmatrix} 2 & & \\ & 3 & \\ & & 7 \end{bmatrix}$;

$A = LU$ has $U = A$; $A = LDU$ has $U = D^{-1}A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$ with 1's on the diagonal.

12 $A = \begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = LDU$; $\boldsymbol{U}$ **is** $\boldsymbol{L}^{\mathrm{T}}$

$$\begin{bmatrix} 1 & & \\ 4 & 1 & \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 0 \\ 0 & -4 & 4 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 4 & 1 & \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & -4 & \\ & & 4 \end{bmatrix} \begin{bmatrix} 1 & 4 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = LDL^{\mathrm{T}}.$$

13 $\begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & a & a & a \\ & b-a & b-a & b-a \\ & & c-b & c-b \\ & & & d-c \end{bmatrix}$. Need $a \neq 0$ All of the $b \neq a$ multipliers $c \neq b$ are $\ell_{ij} = 1$ $d \neq c$ for this A

14 $\begin{bmatrix} a & r & r & r \\ a & b & s & s \\ a & b & c & t \\ a & b & c & d \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & r & r & r \\ & b-r & s-r & s-r \\ & & c-s & t-s \\ & & & d-t \end{bmatrix}$. Need $a \neq 0$, $b \neq r$, $c \neq s$, $d \neq t$

15 $\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \boldsymbol{c} = \begin{bmatrix} 2 \\ 11 \end{bmatrix}$ gives $\boldsymbol{c} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Then $\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ gives $\boldsymbol{x} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$.

$A\boldsymbol{x} = \boldsymbol{b}$ is $LU\boldsymbol{x} = \begin{bmatrix} 2 & 4 \\ 8 & 17 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 11 \end{bmatrix}$. Eliminate to $\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \boldsymbol{c}$.

16 $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \boldsymbol{c} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ gives $\boldsymbol{c} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$. Then $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$ gives $\boldsymbol{x} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$.

Those are forward elimination and back substitution for $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

17 (a) L goes to I (b) I goes to L^{-1} (c) LU goes to U. Elimination multiplies by L^{-1}!

18 (a) Multiply $LDU = L_1D_1U_1$ by inverses to get $L_1^{-1}LD = D_1U_1U^{-1}$. The left side is lower triangular, the right side is upper triangular $\Rightarrow$ both sides are diagonal.

(b) L, U, L_1, U_1 have diagonal 1's so $D = D_1$. Then $L_1^{-1}L$ and U_1U^{-1} are both I.

19 $\begin{bmatrix} 1 & & \\ 1 & 1 & \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ & 1 & 1 \\ & & 1 \end{bmatrix} = LIU; \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix} = L \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix} U.$

A tridiagonal matrix A has **bidiagonal factors** L and U.

20 A tridiagonal T has 2 nonzeros in the pivot row and only one nonzero below the pivot (one operation to find ℓ and then one for the new pivot!). Only $2n$ operations for elimination on a tridiagonal matrix. T=bidiagonal L times bidiagonal U.

21 For the first matrix A, L keeps the 3 zeros at the start of rows. But U may not have the upper zero where $A_{24} = 0$. For the second matrix B, L keeps the bottom left zero at the start of row 4. U keeps the upper right zero at the start of column 4. One zero in A and two zeros in B are filled in.

22 Eliminating *upwards*, $\begin{bmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 4 & 2 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix} = L$. We reach a *lower* triangular L, and the multipliers are in an *upper* triangular U. $A = \boldsymbol{UL}$ with $U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

23 The 2 by 2 upper submatrix A_2 has the first two pivots $5, 9$. Reason: Elimination on A starts in the upper left corner with elimination on A_2.

24 The upper left blocks all factor at the same time as A: A_k is $L_k U_k$. So $A = LU$ is possible only if all those blocks A_k are invertible.

25 The i, j entry of L^{-1} is j/i for $i \geq j$. And $L_{i\,i-1}$ is $(1-i)/i$ below the diagonal

26 $(K^{-1})_{ij} = j(n-i+1)/(n+1)$ for $i \geq j$ (and symmetric): Multiply K^{-1} by $n+1$ (the determinant of K) to see all whole numbers.

Problem Set 2.7, page 117

1 $A = \begin{bmatrix} 1 & 0 \\ 9 & 3 \end{bmatrix}$ has $A^{\mathrm{T}} = \begin{bmatrix} 1 & 9 \\ 0 & 3 \end{bmatrix}$, $A^{-1} = \begin{bmatrix} 1 & 0 \\ -3 & 1/3 \end{bmatrix}$, $(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1} = \begin{bmatrix} 1 & -3 \\ 0 & 1/3 \end{bmatrix}$;

$A = \begin{bmatrix} 1 & c \\ c & 0 \end{bmatrix}$ has $A^{\mathrm{T}} = A$ and $A^{-1} = \frac{1}{c^2}\begin{bmatrix} 0 & c \\ c & -1 \end{bmatrix} = (A^{-1})^{\mathrm{T}}$.

2 $(AB)^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = B^{\mathrm{T}}A^{\mathrm{T}}$. This answer is different from $A^{\mathrm{T}}B^{\mathrm{T}}$ (except when $AB = BA$ and transposing gives $B^{\mathrm{T}}A^{\mathrm{T}} = A^{\mathrm{T}}B^{\mathrm{T}}$).

3 (a) $((AB)^{-1})^{\mathrm{T}} = (B^{-1}A^{-1})^{\mathrm{T}} = (A^{-1})^{\mathrm{T}}(B^{-1})^{\mathrm{T}}$. This is also $(A^{\mathrm{T}})^{-1}(B^{\mathrm{T}})^{-1}$.

(b) If U is upper triangular, so is U^{-1}: then $(U^{-1})^{\mathrm{T}}$ is *lower* triangular.

4 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has $A^2 = 0$. But the diagonal of $A^{\mathrm{T}}A$ has dot products of columns of A with *themselves*. If $A^{\mathrm{T}}A = 0$, zero dot products $\Rightarrow$ zero columns $\Rightarrow$ $A =$ zero matrix.

5 (a) $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 5$

(b) This is the row $\boldsymbol{x}^{\mathrm{T}}A = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$ times $\boldsymbol{y}$.

(c) This is also the row $\boldsymbol{x}^{\mathrm{T}}$ times $A\boldsymbol{y} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$.

6 $M^{\mathrm{T}} = \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}$; $M^{\mathrm{T}} = M$ needs $A^{\mathrm{T}} = A$ *and* $B^{\mathrm{T}} = C$ *and* $D^{\mathrm{T}} = D$.

7 (a) False: $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ is symmetric only if $A = A^{\mathrm{T}}$.

(b) False: The transpose of AB is $B^{\mathrm{T}}A^{\mathrm{T}} = BA$. So $(AB)^{\mathrm{T}} = AB$ needs $BA = AB$.

(c) True: Invertible symmetric matrices have symmetric inverses! Easiest proof is to transpose $AA^{-1} = I$.

(d) True: $(ABC)^{\mathrm{T}}$ is $C^{\mathrm{T}}B^{\mathrm{T}}A^{\mathrm{T}}(= CBA$ for symmetric matrices A, B, and C).

8 The 1 in row 1 has n choices; then the 1 in row 2 has $n-1$ choices . . . ($n!$ overall).

9 $$P_1P_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ but } P_2P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
If P_3 and P_4 exchange *different* pairs of rows, $P_3P_4 = P_4P_3 =$ both exchanges.

10 $(3,1,2,4)$ and $(2,3,1,4)$ keep 4 in place; 6 more even P's keep 1 or 2 or 3 in place; $(2,1,4,3)$ and $(3,4,1,2)$ and $(4,3,2,1)$ exchange 2 pairs. $(1,2,3,4)$ makes 12.

11 $$PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 0 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$ is upper triangular. Multiplying A *on the right* by a permutation matrix P_2 exchanges the *columns* of A. To make this A lower triangular, we also need P_1 to exchange rows 2 and 3:
$$P_1AP_2 = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} A \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 5 & 4 & 0 \\ 3 & 2 & 1 \end{bmatrix}.$$

12 $(P\boldsymbol{x})^{\mathrm{T}}(P\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}P^{\mathrm{T}}P\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ since $P^{\mathrm{T}}P = I$. In general $P\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{x}\cdot P^{\mathrm{T}}\boldsymbol{y} \neq \boldsymbol{x}\cdot P\boldsymbol{y}$:
Non-equality where $P \neq P^{\mathrm{T}}$: $$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

13 A cyclic $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ or its transpose will have $P^3 = I : (1,2,3) \to (2,3,1) \to (3,1,2) \to (1,2,3)$. The permutation $\widehat{P} = \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix}$ for the same P has $\widehat{P}^4 = \widehat{P} \neq I$.

14 The "reverse identity" P takes $(1,\ldots,n)$ into $(n,\ldots,1)$. When rows and also columns are reversed, the $1,1$ and n,n entries of A change places in PAP. So do the $1,n$ and $n,1$ entries. In general $(PAP)_{ij}$ is $(A)_{n-i+1,n-j+1}$.

15 (a) If P sends row 1 to row 4, then P^{T} sends row 4 to row 1 (b) $P = \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix} = P^{\mathrm{T}}$ with $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ moves all rows: 1 and 2 are exchanged, 3 and 4 are exchanged.

16 $A^2 - B^2$ (but not $(A+B)(A-B)$, this is different) and also ABA are symmetric if A and B are symmetric.

17 (a) $S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = S^{\mathrm{T}}$ is not invertible (b) $S = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ needs row exchange (c) $S = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ has pivots $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$: no real square root.

18 (a) $5+4+3+2+1 = 15$ independent entries if $S = S^{\mathrm{T}}$ (b) L has 10 and D has 5; total 15 in LDL^{T} (c) Zero diagonal if $A^{\mathrm{T}} = -A$, leaving $4+3+2+1 = 10$ choices.

19 (a) The transpose of $A^{\mathrm{T}}SA$ is $A^{\mathrm{T}}S^{\mathrm{T}}A^{\mathrm{T}\,\mathrm{T}} = A^{\mathrm{T}}SA = n$ by n when $S^{\mathrm{T}} = S$ (any m by n matrix A) (b) $(A^{\mathrm{T}}A)_{jj} =$ (column j of A)$\cdot$ (column j of A) $=$ (length squared of column j) ≥ 0.

20 $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$; $\begin{bmatrix} 1 & b \\ b & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & c-b^2 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ & 1 & -\frac{2}{3} \\ & & 1 \end{bmatrix} = \boldsymbol{LDL}^{\mathrm{T}}.$$

21 Elimination on a symmetric 3 by 3 matrix leaves a symmetric lower right 2 by 2 matrix. $\begin{bmatrix} 2 & 4 & 8 \\ 4 & 3 & 9 \\ 8 & 9 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$ lead to $\begin{bmatrix} -5 & -7 \\ -7 & -32 \end{bmatrix}$ and $\begin{bmatrix} d-b^2 & e-bc \\ e-bc & f-c^2 \end{bmatrix}$: symmetric!

22 $\begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} A = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ & 1 & 1 \\ & & -1 \end{bmatrix}$; $\begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} A = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ & -1 & 1 \\ & & 1 \end{bmatrix}$

23 $A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = P$ and $L = U = I$. This cyclic P exchanges rows 1-2 then rows 2-3 then rows 3-4.

24 $PA = LU$ is $\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 8 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & 1/3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ & 3 & 8 \\ & & -2/3 \end{bmatrix}$. If we wait to exchange and a_{12} is the pivot, $A = L_1 P_1 U_1 = \begin{bmatrix} 1 & & \\ 3 & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$.

25 One way to decide even vs. odd is to count all pairs that P has in the wrong order. Then P is even or odd when that count is even or odd. Hard step: Show that an exchange always switches that count! Then 3 or 5 exchanges will leave that count odd.

26 (a) $E_{21} = \begin{bmatrix} 1 & & \\ -3 & 1 & \\ & & 1 \end{bmatrix}$ puts 0 in the $2, 1$ entry of $E_{21}A$. Then $E_{21}AE_{21}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 9 \end{bmatrix}$ is still symmetric, with zero also in its 1, 2 entry. (b) Now use $E_{32} = \begin{bmatrix} 1 & & \\ & 1 & \\ & -4 & 1 \end{bmatrix}$ to make the 3, 2 entry zero and $E_{32}E_{21}AE_{21}^{\mathrm{T}}E_{32}^{\mathrm{T}} = D$ also has zero in its 2, 3 entry. Key point: Elimination from both sides gives the symmetric LDL^{T} directly.

27 $A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \end{bmatrix} = A^{\mathrm{T}}$ has $0, 1, 2, 3$ in every row. I don't know any rules for a symmetric construction like this "Hankel matrix" with constant antidiagonals.

28 Reordering the rows and/or the columns of $\begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}$ will move the entry **a**. So the result cannot be the transpose (which doesn't move **a**).

29 (a) Total currents are $A^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_{BC} \\ y_{CS} \\ y_{BS} \end{bmatrix} = \begin{bmatrix} y_{BC} + y_{BS} \\ -y_{BC} + y_{CS} \\ -y_{CS} - y_{BS} \end{bmatrix}$.

(b) Either way $(A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}\boldsymbol{y}) = x_B y_{BC} + x_B y_{BS} - x_C y_{BC} + x_C y_{CS} - x_S y_{CS} - x_S y_{BS}$. Six terms.

30 $\begin{bmatrix} 1 & 50 \\ 40 & 1000 \\ 2 & 50 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\boldsymbol{x}$; $A^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 1 & 40 & 2 \\ 50 & 1000 & 50 \end{bmatrix} \begin{bmatrix} 700 \\ 3 \\ 3000 \end{bmatrix} = \begin{bmatrix} 6820 \\ 188000 \end{bmatrix}$ $\begin{matrix} \text{1 truck} \\ \text{1 plane} \end{matrix}$

31 $A\boldsymbol{x} \cdot \boldsymbol{y}$ is the *cost of inputs* while $\boldsymbol{x} \cdot A^{\mathrm{T}}\boldsymbol{y}$ is the *value of outputs*.

32 $P^3 = I$ so three rotations for $360°$; P rotates every $\boldsymbol{v}$ around the $(1, 1, 1)$ line by $120°$.

33 $\begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = \boldsymbol{EH} =$ (elementary matrix) times (symmetric matrix).

34 $L(U^{\mathrm{T}})^{-1}$ is lower triangular times lower triangular, so *lower triangular*. The transpose of $U^{\mathrm{T}}DU$ is $U^{\mathrm{T}}D^{\mathrm{T}}U^{\mathrm{T}\,\mathrm{T}} = U^{\mathrm{T}}DU$ again, so $U^{\mathrm{T}}DU$ is *symmetric*. The factorization multiplies lower triangular by symmetric to get LDU which is A.

35 These are groups: Lower triangular with diagonal 1's, diagonal invertible D, permutations P, orthogonal matrices with $Q^{\mathrm{T}} = Q^{-1}$.

36 Certainly B^{T} is northwest. B^2 is a full matrix! B^{-1} is southeast: $\begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{-1} \end{bmatrix}$. The rows of B are in reverse order from a lower triangular L, so $B = PL$. Then $B^{-1} = L^{-1}P^{-1}$ has the *columns* in reverse order from L^{-1}. So B^{-1} is *southeast*. Northwest $B = PL$ times southeast PU is $(PLP)U =$ upper triangular.

37 There are $n!$ permutation matrices of order n. Eventually *two powers of P must be the same permutation*. And if $P^r = P^s$ then $P^{r-s} = I$. Certainly $r - s \leq n!$

$P = \begin{bmatrix} P_2 & \\ & P_3 \end{bmatrix}$ is 5 by 5 with $P_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and $P^6 = I$.

38 To split the matrix M into (symmetric S) + (anti-symmetric A), the only choice is $S = \frac{1}{2}(M + M^{\mathrm{T}})$ and $A = \frac{1}{2}(M - M^{\mathrm{T}})$.

39 Start from $Q^{\mathrm{T}}Q = I$, as in $\begin{bmatrix} \boldsymbol{q}_1^{\mathrm{T}} \\ \boldsymbol{q}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(a) The diagonal entries give $\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{q}_1 = 1$ and $\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{q}_2 = 1$: *unit vectors*

(b) The off-diagonal entry is $\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{q}_2 = 0$ (and in general $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j = 0$)

(c) The leading example for Q is the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Problem Set 3.1, page 131

Note An interesting "max-plus" vector space comes from the real numbers **R** combined with $-\infty$. Change addition to give $x + y = \mathbf{max}(\boldsymbol{x}, \boldsymbol{y})$ and change multiplication to $xy = \textbf{usual } \boldsymbol{x} + \boldsymbol{y}$. Which y is the zero vector that gives $x + \mathbf{0} = \max(x, \mathbf{0}) = x$ for every x?

1 $\boldsymbol{x} + \boldsymbol{y} \neq \boldsymbol{y} + \boldsymbol{x}$ and $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) \neq (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$ and $(c_1 + c_2)\boldsymbol{x} \neq c_1\boldsymbol{x} + c_2\boldsymbol{x}$.

2 When $c(x_1, x_2) = (cx_1, 0)$, the only broken rule is 1 times $\boldsymbol{x}$ equals $\boldsymbol{x}$. Rules (1)-(4) for addition $\boldsymbol{x} + \boldsymbol{y}$ still hold since addition is not changed.

3 (a) $c\boldsymbol{x}$ may not be in our set: not closed under multiplication. Also no $\mathbf{0}$ and no $-\boldsymbol{x}$
(b) $c(\boldsymbol{x} + \boldsymbol{y})$ is the usual $(xy)^c$, while $c\boldsymbol{x} + c\boldsymbol{y}$ is the usual $(x^c)(y^c)$. Those are equal. With $c = 3$, $x = 2$, $y = 1$ this is $3(\mathbf{2} + \mathbf{1}) = 8$. The zero vector is the number 1.

4 The zero vector in matrix space $\mathbf{M}$ is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; $\frac{1}{2}A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ and $-A = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix}$. The smallest subspace of $\mathbf{M}$ containing the matrix A consists of all matrices cA.

5 (a) One possibility: The matrices cA form a subspace not containing B (b) Yes: the subspace must contain $A - B = I$ (c) Matrices whose main diagonal is all zero.

6 When $\boldsymbol{f}(x) = x^2$ and $\boldsymbol{g}(x) = 5x$, the combination $3\boldsymbol{f} - 4\boldsymbol{g}$ in function space is $\boldsymbol{h}(x) = 3\boldsymbol{f}(x) - 4\boldsymbol{g}(x) = 3x^2 - 20x$.

7 Rule 8 is broken: If $c\boldsymbol{f}(x)$ is defined to be the usual $\boldsymbol{f}(cx)$ then $(c_1 + c_2)\boldsymbol{f} = \boldsymbol{f}((c_1 + c_2)x)$ is not generally the same as $c_1\boldsymbol{f} + c_2\boldsymbol{f} = \boldsymbol{f}(c_1x) + \boldsymbol{f}(c_2x)$.

8 If $(\boldsymbol{f} + \boldsymbol{g})(x)$ is the usual $\boldsymbol{f}(\boldsymbol{g}(x))$ then $(\boldsymbol{g} + \boldsymbol{f})x$ is $\boldsymbol{g}(\boldsymbol{f}(x))$ which is different. In Rule 2 both sides are $\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{h}(x)))$. Rule 4 is broken because there might be no inverse function $\boldsymbol{f}^{-1}(x)$ such that $\boldsymbol{f}(\boldsymbol{f}^{-1}(x)) = x$. If the inverse function exists it will be the vector $-\boldsymbol{f}$.

9 (a) The vectors with integer components allow addition, but not multiplication by $\frac{1}{2}$
(b) Remove the x axis from the xy plane (but leave the origin). Multiplication by any c is allowed but not all vector additions: $(1, 1) + (-1, 1) = (0, 2)$ is removed.

10 The only subspaces are (a) the plane with $b_1 = b_2$ (d) the linear combinations of $\boldsymbol{v}$ and $\boldsymbol{w}$ (e) the plane with $b_1 + b_2 + b_3 = 0$.

11 (a) All matrices $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ (b) All matrices $\begin{bmatrix} a & a \\ 0 & 0 \end{bmatrix}$ (c) All diagonal matrices.

12 For the plane $x + y - 2z = 4$, the sum of $(4, 0, 0)$ and $(0, 4, 0)$ is not on the plane. (The key is that this plane does not go through $(0, 0, 0)$.)

13 The parallel plane $\mathbf{P}_0$ has the equation $x + y - 2z = 0$. Pick two points, for example $(2, 0, 1)$ and $(0, 2, 1)$, and their sum $(2, 2, 2)$ is in $\mathbf{P}_0$.

14 (a) The subspaces of $\mathbf{R}^2$ are $\mathbf{R}^2$ itself, lines through $(0, 0)$, and $(0, 0)$ by itself (b) The subspaces of $\mathbf{R}^4$ are $\mathbf{R}^4$ itself, three-dimensional planes $\boldsymbol{n} \cdot \boldsymbol{v} = 0$, two-dimensional subspaces ($\boldsymbol{n}_1 \cdot \boldsymbol{v} = 0$ and $\boldsymbol{n}_2 \cdot \boldsymbol{v} = 0$), one-dimensional lines through $(0, 0, 0, 0)$, and finally $(0, 0, 0, 0)$ by itself, which is the zero subspace **Z**.

15 (a) Two planes through $(0, 0, 0)$ probably intersect in a line through $(0, 0, 0)$
(b) The plane and line probably intersect in the point $(0, 0, 0)$
(c) If $\boldsymbol{x}$ and $\boldsymbol{y}$ are in both $\boldsymbol{S}$ and $\boldsymbol{T}$, $\boldsymbol{x} + \boldsymbol{y}$ and $c\boldsymbol{x}$ are in both subspaces.

16 The smallest subspace containing a plane $\mathbf{P}$ and a line $\mathbf{L}$ is *either* $\mathbf{P}$ (when the line $\mathbf{L}$ is in the plane $\mathbf{P}$) *or* $\mathbf{R}^3$ (when $\mathbf{L}$ is not in $\mathbf{P}$).

17 (a) The invertible matrices do not include the zero matrix, so they are not a subspace
(b) The sum of singular matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is not singular: not a subspace.

18 (a) *True*: The symmetric matrices do form a subspace (b) *True*: The matrices with $A^{\mathrm{T}} = -A$ do form a subspace (c) *False*: The sum of two unsymmetric matrices could be symmetric.

19 The column space of A is the x-axis = all vectors $(x, 0, 0)$: a *line*. The column space of B is the xy plane = all vectors $(x, y, 0)$. The column space of C is the line of vectors $(x, 2x, 0)$.

20 (a) Elimination leads to $0 = b_2 - 2b_1$ and $0 = b_1 + b_3$ in equations 2 and 3: Solution only if $b_2 = 2b_1$ and $b_3 = -b_1$ (b) Elimination leads to $0 = b_1 + b_3$ in equation 3: Solution only if $b_3 = -b_1$.

21 A combination of the columns of C is also a combination of the columns of A. Then $C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ have the same column space. $B = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ has a different column space. The key word is "space".

22 (a) Solution for every $\boldsymbol{b}$ (b) Solvable only if $b_3 = 0$ (c) Solvable only if $b_3 = b_2$.

23 The extra column $\boldsymbol{b}$ enlarges the column space unless $\boldsymbol{b}$ is *already in* the column space. $[\,A \;\; \boldsymbol{b}\,] = \begin{bmatrix} 1 & 0 & \mathbf{1} \\ 0 & 0 & \mathbf{1} \end{bmatrix}$ (larger column space) (no solution to $A\boldsymbol{x} = \boldsymbol{b}$) $\begin{bmatrix} 1 & 0 & \mathbf{1} \\ 0 & 1 & \mathbf{1} \end{bmatrix}$ ($\boldsymbol{b}$ is in column space) ($A\boldsymbol{x} = \boldsymbol{b}$ has a solution)

24 The column space of AB is *contained in* (possibly equal to) the column space of A. The example $B =$ zero matrix and $A \neq 0$ is a case when $AB =$ zero matrix has a smaller column space (it is just the zero space $\mathbf{Z}$) than A.

25 The solution to $A\boldsymbol{z} = \boldsymbol{b} + \boldsymbol{b}^*$ is $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$. If $\boldsymbol{b}$ and $\boldsymbol{b}^*$ are in $\boldsymbol{C}(A)$ so is $\boldsymbol{b} + \boldsymbol{b}^*$.

26 The column space of any invertible 5 by 5 matrix is $\mathbf{R}^5$. The equation $A\boldsymbol{x} = \boldsymbol{b}$ is always solvable (by $\boldsymbol{x} = A^{-1}\boldsymbol{b}$) so every $\boldsymbol{b}$ is in the column space of that invertible matrix.

27 (a) *False*: Vectors that are *not* in a column space don't form a subspace. (b) *True*: Only the zero matrix has $\boldsymbol{C}(A) = \{\mathbf{0}\}$. (c) *True*: $\boldsymbol{C}(A) = \boldsymbol{C}(2A)$. (d) *False*: $\boldsymbol{C}(A - I) \neq \boldsymbol{C}(A)$ when $A = I$ or $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (or other examples).

28 $A = \begin{bmatrix} 1 & 1 & \mathbf{0} \\ 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & \mathbf{2} \\ 1 & 0 & \mathbf{1} \\ 0 & 1 & \mathbf{1} \end{bmatrix}$ do not have $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in $\boldsymbol{C}(A)$. $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \end{bmatrix}$ has $\boldsymbol{C}(A) =$ line in $\mathbf{R}^3$.

29 When $A\boldsymbol{x} = \boldsymbol{b}$ is solvable for all $\boldsymbol{b}$, every $\boldsymbol{b}$ is in the column space of A. So that space is $\boldsymbol{C}(A) = \mathbf{R}^9$.

30 (a) If $\boldsymbol{u}$ and $\boldsymbol{v}$ are both in $\boldsymbol{S}+\boldsymbol{T}$, then $\boldsymbol{u}=\boldsymbol{s}_1+\boldsymbol{t}_1$ and $\boldsymbol{v}=\boldsymbol{s}_2+\boldsymbol{t}_2$. So $\boldsymbol{u}+\boldsymbol{v}=(\boldsymbol{s}_1+\boldsymbol{s}_2)+(\boldsymbol{t}_1+\boldsymbol{t}_2)$ is also in $\boldsymbol{S}+\boldsymbol{T}$. And so is $c\boldsymbol{u}=c\boldsymbol{s}_1+c\boldsymbol{t}_1$: $\boldsymbol{S}+\boldsymbol{T}=$ *subspace*.

(b) If $\boldsymbol{S}$ and $\boldsymbol{T}$ are different lines, then $\boldsymbol{S}\cup\boldsymbol{T}$ is just the two lines (*not a subspace*) but $\boldsymbol{S}+\boldsymbol{T}$ is the whole plane that they span.

31 If $\boldsymbol{S}=\boldsymbol{C}(A)$ and $\boldsymbol{T}=\boldsymbol{C}(B)$ then $\boldsymbol{S}+\boldsymbol{T}$ is the column space of $M=[\,A\quad B\,]$.

32 The columns of AB are combinations of the columns of A. So all columns of $[\,A\quad AB\,]$ are already in $\boldsymbol{C}(A)$. But $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ has a larger column space than $A^2=\begin{bmatrix}0&0\\0&0\end{bmatrix}$. For square matrices, the column space is $\mathbf{R}^n$ exactly when A is *invertible*.

Problem Set 3.2, page 142

1 (a) $U=\begin{bmatrix}1&2&2&4&6\\0&0&1&2&3\\0&0&0&0&0\end{bmatrix}$ Free variables x_2,x_4,x_5 Pivot variables x_1,x_3 (b) $U=\begin{bmatrix}2&4&2\\0&4&4\\0&0&0\end{bmatrix}$ Free x_3 Pivot x_1,x_2

2 (a) Free variables x_2,x_4,x_5 and solutions $(-2,1,0,0,0)$, $(0,0,-2,1,0)$, $(0,0,-3,0,1)$

(b) Free variable x_3: solution $(1,-1,1)$. Special solution for each free variable.

3 $R=\begin{bmatrix}1&2&0&0&0\\0&0&1&2&3\\0&0&0&0&0\end{bmatrix}$, $R=\begin{bmatrix}1&0&-1\\0&1&1\\0&0&0\end{bmatrix}$, R has the same nullspace as U and A.

4 (a) Special solutions $(3,1,0)$ and $(5,0,1)$ (b) $(3,1,0)$. Total of pivot and free is n.

5 (a) *False*: Any singular square matrix would have free variables (b) *True*: An invertible square matrix has *no* free variables. (c) *True* (only n columns to hold pivots)

(d) *True* (only m rows to hold pivots)

6 $\begin{bmatrix}0&\mathbf{1}&1&1&1&1&1\\0&0&0&\mathbf{1}&1&1&1\\0&0&0&0&\mathbf{1}&1&1\\0&0&0&0&0&0&0\end{bmatrix}$ $\begin{bmatrix}\mathbf{1}&1&1&1&1&1&1\\0&0&\mathbf{1}&1&1&1&1\\0&0&0&0&0&\mathbf{1}&1\\0&0&0&0&0&0&\mathbf{1}\end{bmatrix}$ $\begin{bmatrix}0&0&0&\mathbf{1}&1&1&1\\0&0&0&0&0&\mathbf{1}&1\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\end{bmatrix}$

7 $\begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}, \begin{bmatrix} 0 & \mathbf{1} & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Notice the identity matrix in the pivot columns of these *reduced* row echelon forms R.

8 If column 4 of a 3 by 5 matrix is all zero then x_4 is a *free* variable. Its special solution is $\boldsymbol{x} = (0, 0, 0, 1, 0)$, because 1 will multiply that zero column to give $A\boldsymbol{x} = \mathbf{0}$.

9 If column $1 =$ column 5 then x_5 is a free variable. Its special solution is $(-1, 0, 0, 0, 1)$.

10 If a matrix has n columns and r pivots, there are $\boldsymbol{n} - \boldsymbol{r}$ special solutions. The nullspace contains only $\boldsymbol{x} = \mathbf{0}$ when $\boldsymbol{r} = \boldsymbol{n}$. The column space is all of $\mathbf{R}^m$ when $r = m$. All those statements are important!

11 The nullspace contains only $\boldsymbol{x} = \mathbf{0}$ when A has 5 pivots. Also the column space is $\mathbf{R}^5$, because we can solve $A\boldsymbol{x} = \boldsymbol{b}$ and every $\boldsymbol{b}$ is in the column space.

12 $A = [1 \ \ -3 \ \ -1]$ gives the plane $x - 3y - z = 0$; y and z are free variables. The special solutions are $(3, 1, 0)$ and $(1, 0, 1)$.

13 Fill in $\mathbf{12}$ then $\mathbf{4}$ then $\mathbf{1}$ to get the complete solution in $\mathbf{R}^3$ to $x - 3y - z = 12$:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \mathbf{12} \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} \mathbf{4} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} \mathbf{1} \\ 0 \\ 1 \end{bmatrix} = \boldsymbol{x}_{\text{particular}} + \boldsymbol{x}_{\text{nullspace}}.$$

14 Column 5 is sure to have no pivot since it is a combination of earlier columns. With 4 pivots in the other columns, the special solution is $\boldsymbol{s} = (1, 0, 1, 0, 1)$. The nullspace contains all multiples of this vector $\boldsymbol{s}$ (this nullspace is a line in $\mathbf{R}^5$).

15 To produce special solutions $(2, 2, 1, 0)$ and $(3, 1, 0, 1)$ with free variables x_3, x_4: $R = \begin{bmatrix} 1 & 0 & -2 & -3 \\ 0 & 1 & -2 & -1 \end{bmatrix}$ and A can be any invertible 2 by 2 matrix times this R.

16 The nullspace of $A = \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$ is the line through the special solution $\begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$.

17 $A = \begin{bmatrix} 1 & 0 & -1/2 \\ 1 & 3 & -2 \\ 5 & 1 & -3 \end{bmatrix}$ has $\begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$ in $\boldsymbol{C}(A)$ and $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ in $\boldsymbol{N}(A)$. *Which other A's?*

18 This construction is impossible for 3 by 3 ! 2 pivot columns and 2 free variables.

19 $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$ has $(1, 1, 1)$ in $\boldsymbol{C}(A)$ and only the line (c, c, c, c) in $\boldsymbol{N}(A)$.

20 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has $\boldsymbol{N}(A) = \boldsymbol{C}(A)$. Notice that $\text{rref}(A^{\mathrm{T}}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not A^{T}.

21 If nullspace $=$ column space (with r pivots) then $n - r = r$. If $n = 3$ then $3 = 2r$ is impossible.

22 If A times every column of B is zero, the column space of B is contained in the *nullspace* of A. An example is $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$. Here $\boldsymbol{C}(B)$ *equals* $\boldsymbol{N}(A)$. For $B = 0, \boldsymbol{C}(B)$ is smaller than $\boldsymbol{N}(A)$.

23 For $A =$ random 3 by 3 matrix, R is almost sure to be I. For 4 by 3, R is most likely to be I with a fourth row of zeros. What is R for a random 3 by 4 matrix?

24 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ shows that (a)(b)(c) are all false. Notice $\text{rref}(A^{\mathrm{T}}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

25 If $\boldsymbol{N}(A) =$ line through $\boldsymbol{x} = (2, 1, 0, 1)$, A has *three pivots* (4 columns and 1 special solution). Its reduced echelon form can be $R = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ (add any zero rows).

26 $R = [1 \ \ -2 \ \ -3]$, $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $R = I$. Any zero rows come after those rows.

27 (a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) All 8 matrices are R's !

28 One reason that R is the same for A and $-A$: They have the same nullspace. (They also have the same row space. They also have the same column space, but that is not required for two matrices to share the same R. R tells us the nullspace and row space.)

29 The nullspace of $B = [\,A \quad A\,]$ contains all vectors $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{y} \\ -\boldsymbol{y} \end{bmatrix}$ for $\boldsymbol{y}$ in $\mathbf{R}^4$.

30 If $C\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x} = \mathbf{0}$ and $B\boldsymbol{x} = \mathbf{0}$. So $\boldsymbol{N}(C) = \boldsymbol{N}(A) \cap \boldsymbol{N}(B) =$ *intersection*.

31 (a) rank 1 (b) rank 2 because every row is a combination of $(1, 2, 3, 4)$ and $(1, 1, 1, 1)$

(c) rank 1 because all columns are multiples of $(1, 1, 1)$

32 $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$: $y_1 - y_3 + y_4 = -y_1 + y_2 + +y_5 = -y_2 + y_4 + y_6 = -y_4 - y_5 - y_6 = 0$. These equations add to $0 = 0$. Free variables y_3, y_5, y_6: watch for flows around loops. The solutions to $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$ are combinations of $(-1, 0, 0, 1, -1, 0)$ and $(0, 0, -1, -1, 0, 1)$ and $(0, -1, 0, 0, 1, -1)$. Those are flows around the 3 small loops.

33 (a) and (c) are correct; (b) is completely false; (d) is false because R might have 1's in nonpivot columns.

34 $R_A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ $R_B = \begin{bmatrix} R_A & R_A \end{bmatrix}$ $R_C \longrightarrow \begin{bmatrix} R_A & 0 \\ 0 & R_A \end{bmatrix} \longrightarrow$ Zero rows go to the bottom

35 If all pivot variables come last then $R = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}$. The nullspace matrix is $N = \begin{bmatrix} I \\ 0 \end{bmatrix}$.

36 I think $R_1 = A_1, R_2 = A_2$ is true. But $R_1 - R_2$ may have -1's in some pivots.

37 A and A^{T} have the same rank r = number of pivots. But *pivcol* (the column number) is 2 for this matrix A and 1 for A^{T}: $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

38 Special solutions in $N = [\,-2 \quad -4 \quad 1 \quad 0;\ -3 \quad -5 \quad 0 \quad 1\,]$ and $[\,1 \quad 0 \quad 0;\ 0 \quad -2 \quad 1\,]$.

39 The new entries keep rank 1: $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & \mathbf{4} & \mathbf{8} \\ 4 & \mathbf{8} & \mathbf{16} \end{bmatrix}$, $B = \begin{bmatrix} \mathbf{2} & 6 & \mathbf{-3} \\ 1 & \mathbf{3} & \mathbf{-3/2} \\ 2 & 6 & -3 \end{bmatrix}$, $M = \begin{bmatrix} a & b \\ c & \boldsymbol{bc/a} \end{bmatrix}$.

40 If A has rank 1, the column space is a *line* in $\mathbf{R}^m$. The nullspace is a *plane* in $\mathbf{R}^n$ (given by one equation). The nullspace matrix N is n by $n-1$ (with $n-1$ special solutions in its columns). The column space of A^{T} is a *line* in $\mathbf{R}^n$.

41 $\begin{bmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 4 & 8 & 8 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 2 & 6 & 4 \\ -1 & -1 & -3 & -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 & 2 \end{bmatrix}$

42 With rank 1, the second row of R is a zero row.

43 Invertible r by r submatrices
Use pivot rows and columns $S = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$ and $S = [\,1\,]$ and $S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

44 P has $\operatorname{rank} r$ (the same as A) because elimination produces the same pivot columns.

45 The rank of R^{T} is also r. The example matrix A has $\operatorname{rank} 2$ with invertible S:

$$P = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 2 & 7 \end{bmatrix} \qquad P^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 2 \\ 3 & 6 & 7 \end{bmatrix} \qquad S^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \qquad S = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}.$$

46 The product of rank one matrices has rank one or zero. These particular matrices have rank$(AB) = 1$; rank$(AM) = 1$ except $AM = 0$ if $c = -1/2$.

47 $(\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})(\boldsymbol{w}\boldsymbol{z}^{\mathrm{T}}) = \boldsymbol{u}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w})\boldsymbol{z}^{\mathrm{T}}$ has rank one unless the inner product is $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = 0$.

48 (a) By matrix multiplication, each column of AB is A times the corresponding column of B. So if column j of B is a combination of earlier columns, then column j of AB is the same combination of earlier columns of AB. Then rank $(AB) \le$ rank (B). No new pivot columns! (b) The rank of B is $r = 1$. Multiplying by A cannot increase this rank. The rank of AB stays the same for $A_1 = I$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. It drops to zero for $A_2 = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$.

49 If we know that $\text{rank}(B^{\text{T}} A^{\text{T}}) \leq \text{rank}(A^{\text{T}})$, then since rank stays the same for transposes, (apologies that this fact is not yet proved), we have $\text{rank}(AB) \leq \text{rank}(A)$.

50 We are given $AB = I$ which has rank n. Then rank$(AB) \leq$ rank(A) forces rank$(A) = n$. This means that A is invertible. The right-inverse B is also a left-inverse: $BA = I$ and $B = A^{-1}$.

51 Certainly A and B have at most $\text{rank}\,2$. Then their product AB has at most $\text{rank}\,2$. Since BA is 3 by 3, it cannot be I even if $AB = I$.

52 (a) A and B will both have the same nullspace and row space as the R they share.

(b) A equals an *invertible* matrix times B, when they share the same R. A key fact!

53 $A = (\text{pivot columns})(\text{nonzero rows of } R) = \begin{bmatrix} 1 & 0 \\ 1 & 4 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 8 \end{bmatrix}$. $B = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{matrix} \text{columns} \\ \text{times rows} \end{matrix} = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 3 \end{bmatrix}$

54 If $c = 1, R = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ has x_2, x_3, x_4 free. If $c \neq 1, R = \begin{bmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ has x_3, x_4 free. Special solutions in $N = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (for $c = 1$) and $N = \begin{bmatrix} -2 & -2 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ (for $c \neq 1$). If $c = 1, R = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and x_1 free; if $c = 2, R = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$

and x_2 free; $R = I$ if $c \neq 1, 2$. Special solutions in $N = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $(c = 1)$ or $N = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $(c = 2)$ or $N = 2$ by 0 empty matrix.

55 $A = \begin{bmatrix} I & I \end{bmatrix}$ has $N = \begin{bmatrix} I \\ -I \end{bmatrix}$; $B = \begin{bmatrix} I & I \\ 0 & 0 \end{bmatrix}$ has the same N; $C = \begin{bmatrix} I & I & I \end{bmatrix}$ has $N = \begin{bmatrix} -I & -I \\ I & 0 \\ 0 & I \end{bmatrix}$.

56 $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} =$ (pivot column) (first row) $= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$

57 The m by n matrix Z has r ones to start its main diagonal. Otherwise Z is all zeros.

58 $R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} r \text{ by } r & r \text{ by } n-r \\ m-r \text{ by } r & m-r \text{ by } n-r \end{bmatrix}$; **rref**$(R^{\mathrm{T}}) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$; **rref**$(R^{\mathrm{T}}R) =$ same R

59 $R = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ has $R^{\mathrm{T}}R = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and this matrix row reduces to $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} R \\ \text{zero row} \end{bmatrix}$. Always $R^{\mathrm{T}}R$ has the same nullspace as R, so its row reduced form must be R with $n - m$ extra zero rows. R is determined by its nullspace and shape !

60 The *row-column reduced echelon form* is always $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$; I is r by r.

Problem Set 3.3, page 158

1 $\begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 2 & 5 & 7 & 6 & \mathbf{b}_2 \\ 2 & 3 & 5 & 2 & \mathbf{b}_3 \end{bmatrix} \to \begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 0 & 1 & 1 & 2 & \mathbf{b}_2 - \mathbf{b}_1 \\ 0 & -1 & -1 & -2 & \mathbf{b}_3 - \mathbf{b}_1 \end{bmatrix} \to \begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 0 & 1 & 1 & 2 & \mathbf{b}_2 - \mathbf{b}_1 \\ 0 & 0 & 0 & 0 & \mathbf{b}_3 + \mathbf{b}_2 - \mathbf{2b}_1 \end{bmatrix}$

$A\boldsymbol{x} = \boldsymbol{b}$ has a solution when $b_3 + b_2 - 2b_1 = 0$; the column space contains all combinations of $(2, 2, 2)$ and $(4, 5, 3)$. **This is the plane** $b_3 + b_2 - 2b_1 = 0$ (!). The nullspace contains all combinations of $\boldsymbol{s}_1 = (-1, -1, 1, 0)$ and $\boldsymbol{s}_2 = (2, -2, 0, 1)$; $\boldsymbol{x}_{complete} = \boldsymbol{x}_p + c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2$;

$\begin{bmatrix} R & \boldsymbol{d} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & -2 & 4 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ gives the particular solution $x_p = (4, -1, 0, 0)$.

2 $\begin{bmatrix} 2 & 1 & 3 & \mathbf{b}_1 \\ 6 & 3 & 9 & \mathbf{b}_2 \\ 4 & 2 & 6 & \mathbf{b}_3 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 3 & \mathbf{b}_1 \\ 0 & 0 & 0 & \mathbf{b}_2 - \mathbf{3b}_1 \\ 0 & 0 & 0 & \mathbf{b}_3 - \mathbf{2b}_1 \end{bmatrix}$ Then $[\,R \quad \boldsymbol{d}\,] = \begin{bmatrix} 1 & 1/2 & 3/2 & \mathbf{5} \\ 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$

$A\boldsymbol{x} = \boldsymbol{b}$ has a solution when $b_2 - 3b_1 = 0$ and $b_3 - 2b_1 = 0$; $\boldsymbol{C}(A) =$ line through $(2, 6, 4)$ which is the intersection of the planes $b_2 - 3b_1 = 0$ and $b_3 - 2b_1 = 0$; the nullspace contains all combinations of $\boldsymbol{s}_1 = (-1/2, 1, 0)$ and $\boldsymbol{s}_2 = (-3/2, 0, 1)$; particular solution $\boldsymbol{x}_p = \boldsymbol{d} = (5, 0, 0)$ and complete solution $\boldsymbol{x}_p + c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2$.

3 $\boldsymbol{x}_{\text{complete}} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$. The matrix is singular but the equations are still solvable; $\boldsymbol{b}$ is in the column space. Our particular solution has free variable $y = 0$.

4 $\boldsymbol{x}_{\text{complete}} = \boldsymbol{x}_p + \boldsymbol{x}_n = (\frac{1}{2}, 0, \frac{1}{2}, 0) + x_2(-3, 1, 0, 0) + x_4(0, 0, -2, 1)$.

5 $\begin{bmatrix} 1 & 2 & -2 & b_1 \\ 2 & 5 & -4 & b_2 \\ 4 & 9 & -8 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -2 & b_1 \\ 0 & 1 & 0 & b_2 - 2b_1 \\ 0 & 0 & 0 & b_3 - 2b_1 - b_2 \end{bmatrix}$ solvable if $b_3 - 2b_1 - b_2 = 0$.

Back-substitution gives the particular solution to $A\boldsymbol{x} = \boldsymbol{b}$ and the special solution to $A\boldsymbol{x} = \boldsymbol{0}$: $\boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_2 \\ b_2 - 2b_1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

6 (a) Solvable if $b_2 = 2b_1$ and $3b_1 - 3b_3 + b_4 = 0$. Then $\boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_3 \\ b_3 - 2b_1 \end{bmatrix} = \boldsymbol{x}_p$

(b) Solvable if $b_2 = 2b_1$ and $3b_1 - 3b_3 + b_4 = 0$. $\boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_3 \\ b_3 - 2b_1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

7 $\begin{bmatrix} 1 & 3 & 1 & b_1 \\ 3 & 8 & 2 & b_2 \\ 2 & 4 & 0 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 & b_2 \\ 0 & -1 & -1 & b_2 - 3b_1 \\ 0 & -2 & -2 & b_3 - 2b_1 \end{bmatrix}$ One more step gives $[\,0\ \ 0\ \ 0\ \ 0\,] =$ row 3 − 2 (row 2) + 4(row 1) **provided** $\boldsymbol{b_3 - 2b_2 + 4b_1 = 0}$.

8 (a) Every $\boldsymbol{b}$ is in $\boldsymbol{C}(A)$: *independent rows*, only the zero combination gives $\boldsymbol{0}$.

(b) We need $b_3 = 2b_2$, because $(\text{row } 3) - 2(\text{row } 2) = \boldsymbol{0}$.

9 $L\begin{bmatrix} U & \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{bmatrix}$

$= \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$; particular $\boldsymbol{x}_p = (-9, 0, 3, 0)$ means $-9(1, 2, 3) + 3(3, 8, 7) = (0, 6, -6)$. This is $A\boldsymbol{x}_p = \boldsymbol{b}$.

10 $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ has $\boldsymbol{x}_p = (2, 4, 0)$ and $\boldsymbol{x}_{\text{null}} = (c, c, c)$. Many possible A !

11 A 1 by 3 system has at least **two** free variables. But $\boldsymbol{x}_{\text{null}}$ in Problem 10 only has **one**.

12 (a) If $A\boldsymbol{x}_1 = \boldsymbol{b}$ and $A\boldsymbol{x}_2 = \boldsymbol{b}$ then $\boldsymbol{x}_1 - \boldsymbol{x}_2$ and also $\boldsymbol{x} = \boldsymbol{0}$ solve $A\boldsymbol{x} = \boldsymbol{0}$

(b) $A(2\boldsymbol{x}_1 - 2\boldsymbol{x}_2) = \boldsymbol{0}, A(2\boldsymbol{x}_1 - \boldsymbol{x}_2) = \boldsymbol{b}$

13 (a) The particular solution $\boldsymbol{x}_p$ is always multiplied by 1 (b) Any solution can be $\boldsymbol{x}_p$

(c) $\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$. Then $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is shorter (length $\sqrt{2}$) than $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ (length 2)

(d) The only "homogeneous" solution in the nullspace is $\boldsymbol{x}_n = \boldsymbol{0}$ when A is invertible.

14 If column 5 has no pivot, x_5 is a *free* variable. The zero vector *is not* the only solution to $A\boldsymbol{x} = \mathbf{0}$. If this system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution, it has *infinitely many* solutions.

15 If row 3 of U has no pivot, that is a *zero row*. $U\boldsymbol{x} = \boldsymbol{c}$ is only solvable provided $\boldsymbol{c}_3 = 0$. $A\boldsymbol{x} = \boldsymbol{b}$ *might not be solvable*, because U may have other zero rows needing more $c_i = 0$.

16 The largest rank is 3. Then there is a pivot in every *row*. The solution *always exists*. The column space is $\mathbf{R}^3$. An example is $A = [\,I \quad F\,]$ for any 3 by 2 matrix F.

17 The largest rank of a 6 by 4 matrix is 4. Then there is a pivot in every *column*. The solution is *unique* (if there is a solution). The nullspace contains only the *zero vector*. An example is $A = R = [\,I \quad F\,]$ for any 4 by 2 matrix F.

18 Rank $= 2$; rank $= 3$ unless $q = 2$ (then rank $= 2$). Transpose has the same rank!

19 Both matrices A have rank 2. Always $A^{\mathrm{T}}A$ and AA^{T} have **the same rank** as A.

20 $A = LU = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 1 & 0 \\ 0 & -3 & 0 & 1 \end{bmatrix}$; $A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \mathbf{2} & -2 & 3 \\ 0 & 0 & 11 & -5 \end{bmatrix}$.

21 (a) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. The second equation in part (b) removed one special solution from the nullspace.

22 If $A\boldsymbol{x}_1 = \boldsymbol{b}$ and also $A\boldsymbol{x}_2 = \boldsymbol{b}$ then $A(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \mathbf{0}$ and we can add $\boldsymbol{x}_1 - \boldsymbol{x}_2$ to any solution of $A\boldsymbol{x} = \boldsymbol{B}$: the solution $\boldsymbol{x}$ is not unique. But there will be **no solution** to $A\boldsymbol{x} = \boldsymbol{B}$ if $\boldsymbol{B}$ is not in the column space.

23 For $A, q = 3$ gives rank 1, every other q gives rank 2. For $B, q = 6$ gives rank 1, every other q gives rank 2. These matrices cannot have rank 3.

24 (a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix} [\,x\,] = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ has 0 or 1 solutions, depending on $\boldsymbol{b}$ (b) $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,b\,]$ has infinitely many solutions for every b (c) There are 0 or ∞ solutions when A

has rank $r < m$ and $r < n$: the simplest example is a zero matrix. (d) *one* solution for all $\boldsymbol{b}$ when A is square and invertible (like $A = I$).

25 (a) $r < m$, always $r \le n$ (b) $r = m, r < n$ (c) $r < m, r = n$ (d) $r = m = n$.

26 $\begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{bmatrix} \to R = \begin{bmatrix} \mathbf{1} & 0 & -2 \\ 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix} \to R = I$.

27 $R = I$ when A is square and invertible—so for a triangular matrix, all diagonal entries must be nonzero.

28 $\begin{bmatrix} 1 & 2 & 3 & \mathbf{0} \\ 0 & 0 & 4 & \mathbf{0} \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \end{bmatrix}$; $\boldsymbol{x}_n = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$; $\begin{bmatrix} 1 & 2 & 3 & \mathbf{5} \\ 0 & 0 & 4 & \mathbf{8} \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & \mathbf{-1} \\ 0 & 0 & 1 & \mathbf{2} \end{bmatrix}$.

Free $x_2 = 0$ gives $\boldsymbol{x}_p = (-1, 0, 2)$ because the pivot columns contain I.

29 $[R \ \boldsymbol{d}] = \begin{bmatrix} 1 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$ leads to $\boldsymbol{x}_n = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $[R \ \boldsymbol{d}] = \begin{bmatrix} 1 & 0 & 0 & \mathbf{-1} \\ 0 & 0 & 1 & \mathbf{2} \\ 0 & 0 & 0 & \mathbf{5} \end{bmatrix}$:

this has no solution because of the 3rd equation

30 $\begin{bmatrix} 1 & 0 & 2 & 3 & \mathbf{2} \\ 1 & 3 & 2 & 0 & \mathbf{5} \\ 2 & 0 & 4 & 9 & \mathbf{10} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 3 & \mathbf{2} \\ 0 & 3 & 0 & -3 & \mathbf{3} \\ 0 & 0 & 0 & 3 & \mathbf{6} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 0 & \mathbf{-4} \\ 0 & 1 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 0 & 1 & \mathbf{2} \end{bmatrix}$; $\begin{bmatrix} -4 \\ 3 \\ 0 \\ 2 \end{bmatrix}$; $\boldsymbol{x}_n = x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$.

31 For $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$, the only solution to $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is $\boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. B cannot exist since 2 equations in 3 unknowns cannot have a unique solution.

32 $A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 5 \end{bmatrix}$ factors into $LU = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 2 & 2 & 1 & \\ 1 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and the rank is $r = 2$. The special solution to $A\boldsymbol{x} = \mathbf{0}$ and $U\boldsymbol{x} = \mathbf{0}$ is $\boldsymbol{s} = (-7, 2, 1)$. Since

$\boldsymbol{b} = (1,3,6,5)$ is also the last column of A, a particular solution to $A\boldsymbol{x} = \boldsymbol{b}$ is $(0,0,1)$ and the complete solution is $\boldsymbol{x} = (0,0,1) + c\boldsymbol{s}$. (Or use the particular solution $\boldsymbol{x}_p = (7,-2,0)$ with free variable $x_3 = 0$.)

For $\boldsymbol{b} = (1,0,0,0)$ elimination leads to $U\boldsymbol{x} = (1,-1,0,1)$ and the fourth equation is $0 = 1$. No solution for this $\boldsymbol{b}$.

33 If the complete solution to $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \end{bmatrix}$ then $A = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$.

34 (a) If $\boldsymbol{s} = (2,3,1,0)$ is the only special solution to $A\boldsymbol{x} = \boldsymbol{0}$, the complete solution is $\boldsymbol{x} = c\boldsymbol{s}$ (a line of solutions). The rank of A must be $4 - 1 = 3$.

(b) The fourth variable x_4 is *not free* in $\boldsymbol{s}$, and R must be $\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(c) $A\boldsymbol{x} = \boldsymbol{b}$ can be solved for all $\boldsymbol{b}$, because A and R have *full row rank* $r = 3$.

35 For the $-1, 2, -1$ matrix K(9 by 9) and constant right side $\boldsymbol{b} = (10, \cdots, 10)$, the solution $\boldsymbol{x} = K^{-1}\boldsymbol{b} = (45, 80, 105, 120, 125, 120, 105, 80, 45)$ rises and falls along the parabola $x_i = 50i - 5i^2$. (A formula for K^{-1} is later in the text.)

36 If $A\boldsymbol{x} = \boldsymbol{b}$ and $C\boldsymbol{x} = \boldsymbol{b}$ have the same solutions, A and C have the same shape and the same nullspace (take $\boldsymbol{b} = \boldsymbol{0}$). If $\boldsymbol{b} =$ column 1 of A, $\boldsymbol{x} = (1, 0, \ldots, 0)$ solves $A\boldsymbol{x} = \boldsymbol{b}$ so it solves $C\boldsymbol{x} = \boldsymbol{b}$. Then A and C share column 1. Other columns too: $A = C$!

37 The column space of R (m by n with rank r) spanned by its r pivot columns (the first r columns of an m by m identity matrix).

Problem Set 3.4, page 175

1 $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0$ gives $c_3 = c_2 = c_1 = 0$. So those 3 column vectors are independent. But $\begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} \\ c \\ \\ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is solved by $\boldsymbol{c} = (1, 1, -4, 1)$. Then $\boldsymbol{v}_1 + \boldsymbol{v}_2 - 4\boldsymbol{v}_3 + \boldsymbol{v}_4 = \mathbf{0}$ (dependent).

2 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are independent (the -1's are in different positions). All six vectors in $\mathbf{R}^4$ are on the plane $(1, 1, 1, 1) \cdot \boldsymbol{v} = 0$ so no four of these six vectors can be independent.

3 If $a = 0$ then column $1 = \mathbf{0}$; if $d = 0$ then $b(\text{column } 1) - a(\text{column } 2) = \mathbf{0}$; if $f = 0$ then all columns end in zero (they are all in the xy plane, they must be dependent).

4 $U\boldsymbol{x} = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ gives $z = 0$ then $y = 0$ then $x = 0$ (by back substitution). A square triangular matrix has independent columns (invertible matrix) *when its diagonal has no zeros*.

5 (a) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -18/5 \end{bmatrix}$: invertible $\Rightarrow$ independent columns.

(b) $\begin{bmatrix} 1 & 2 & -3 \\ -3 & 1 & 2 \\ 2 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -3 \\ 0 & 7 & -7 \\ 0 & -7 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -3 \\ 0 & 7 & -7 \\ 0 & 0 & 0 \end{bmatrix}$; $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ columns add to $\mathbf{0}$.

6 Columns 1, 2, 4 are independent. Also 1, 3, 4 and 2, 3, 4 and others (but not 1, 2, 3). Same column numbers (not same columns!) for A. This is because $EA = U$ for the matrix E that subtracts 2 times row 1 from row 4. So A and U have the same nullspace (same dependencies of columns).

7 The sum $\boldsymbol{v}_1 - \boldsymbol{v}_2 + \boldsymbol{v}_3 = \mathbf{0}$ because $(\boldsymbol{w}_2 - \boldsymbol{w}_3) - (\boldsymbol{w}_1 - \boldsymbol{w}_3) + (\boldsymbol{w}_1 - \boldsymbol{w}_2) = \mathbf{0}$. So the difference are *dependent* and the difference matrix is singular: $A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$.

8 If $c_1(\boldsymbol{w}_2 + \boldsymbol{w}_3) + c_2(\boldsymbol{w}_1 + \boldsymbol{w}_3) + c_3(\boldsymbol{w}_1 + \boldsymbol{w}_2) = \mathbf{0}$ then $(c_2 + c_3)\boldsymbol{w}_1 + (c_1 + c_3)\boldsymbol{w}_2 + (c_1 + c_2)\boldsymbol{w}_3 = \mathbf{0}$. Since the $\boldsymbol{w}$'s are independent, $c_2 + c_3 = c_1 + c_3 = c_1 + c_2 = 0$. The only solution is $c_1 = c_2 = c_3 = 0$. Only this combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ gives $\mathbf{0}$.

(changing -1's to 1's for the matrix A in solution **7** above makes A invertible.)

9 (a) The four vectors in $\mathbf{R}^3$ are the columns of a 3 by 4 matrix A. There is a nonzero solution to $A\boldsymbol{x} = \mathbf{0}$ because there is at least one free variable (b) Two vectors are dependent if $[\, \boldsymbol{v}_1 \;\; \boldsymbol{v}_2 \,]$ has rank 0 or 1. (OK to say "they are on the same line" or "one is a multiple of the other" but *not* "$\boldsymbol{v}_2$ is a multiple of $\boldsymbol{v}_1$" —since $\boldsymbol{v}_1$ might be $\mathbf{0}$.) (c) A nontrivial combination of $\boldsymbol{v}_1$ and $\mathbf{0}$ gives $\mathbf{0}$: $0\boldsymbol{v}_1 + 3(0, 0, 0) = \mathbf{0}$.

10 The plane is the nullspace of $A = [\, 1 \;\; 2 \;\; -3 \;\; -1 \,]$. Three free variables give three independent solutions $(x, y, z, t) = (2, -1, 0, 0)$ and $(3, 0, 1, 0)$ and $(1, 0, 0, 1)$. Combinations of those special solutions give more solutions (all solutions).

11 (a) Line in $\mathbf{R}^3$ (b) Plane in $\mathbf{R}^3$ (c) All of $\mathbf{R}^3$ (d) All of $\mathbf{R}^3$.

12 $\boldsymbol{b}$ is in the column space when $A\boldsymbol{x} = \boldsymbol{b}$ has a solution; $\boldsymbol{c}$ is in the row space when $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}$ has a solution. *False*. The zero vector is always in the row space.

13 The column space and row space of A and U all have the same dimension = 2. *The row spaces of A and U are the same*, because the rows of U are combinations of the rows of A (and vice versa!).

14 $\boldsymbol{v} = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) + \frac{1}{2}(\boldsymbol{v} - \boldsymbol{w})$ and $\boldsymbol{w} = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) - \frac{1}{2}(\boldsymbol{v} - \boldsymbol{w})$. The two pairs *span* the same space. They are a basis when $\boldsymbol{v}$ and $\boldsymbol{w}$ are *independent*.

15 The n independent vectors span a space of dimension n. They are a *basis* for that space. If they are the columns of A then m is *not less* than n $(m \geq n)$. *Invertible* if $m = n$.

16 These bases are not unique! (a) $(1,1,1,1)$ for the space of all constant vectors (c,c,c,c) (b) $(1,-1,0,0),(1,0,-1,0),(1,0,0,-1)$ for the space of vectors with sum of components = 0 (c) $(1,-1,-1,0),(1,-1,0,-1)$ for the space perpendicular to $(1,1,0,0)$ and $(1,0,1,1)$ (d) The columns of I are a basis for its column space, the empty set is a basis (by convention) for $\boldsymbol{N}(I)=\boldsymbol{Z}=$ {zero vector}.

17 The column space of $U=\begin{bmatrix}1&0&1&0&1\\0&1&0&1&0\end{bmatrix}$ is $\mathbf{R}^2$ so take any bases for $\mathbf{R}^2$; (row 1 and row 2) or (row 1 and row $1+$ row 2) or (row 1 and $-$ row 2) are bases for the row space of U.

18 (a) The 6 vectors *might not* span $\mathbf{R}^4$ (b) The 6 vectors *are not* independent (c) Any four *might be* a basis.

19 n independent columns $\Rightarrow$ rank n. Columns span $\mathbf{R}^m\Rightarrow$ rank m. Columns are basis for $\mathbf{R}^m\Rightarrow$ *rank* $=m=n$. The rank counts the number of *independent* columns.

20 One basis is $(2,1,0)$, $(-3,0,1)$. A basis for the intersection with the xy plane is $(2,1,0)$. The normal vector $(1,-2,3)$ is a basis for the line perpendicular to the plane.

21 (a) The only solution to $A\boldsymbol{x}=\mathbf{0}$ is $\boldsymbol{x}=\mathbf{0}$ because *the columns are independent* (b) $A\boldsymbol{x}=\boldsymbol{b}$ is solvable because *the columns span* $\mathbf{R}^5$. Key point: A basis gives exactly one solution for every $\boldsymbol{b}$.

22 (a) True (b) False because the basis vectors for $\mathbf{R}^6$ might not be in **S**.

23 Columns 1 and 2 are bases for the (**different**) column spaces of A and U; rows 1 and 2 are bases for the (**equal**) row spaces of A and U; $(1,-1,1)$ is a basis for the (**equal**) nullspaces.

24 (a) *False* $A=[\,1\ \ 1\,]$ has dependent columns, independent row (b) *False* Column space $\neq$ row space for $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ (c) *True*: Both dimensions $=2$ if A is invertible, dimensions $=0$ if $A=0$, otherwise dimensions $=1$ (d) *False*, columns may be dependent, in that case not a basis for $\boldsymbol{C}(A)$.

25 A has rank 2 if $c = 0$ and $d = 2$; $B = \begin{bmatrix} c & d \\ d & c \end{bmatrix}$ has rank 2 except when $c = d$ or $c = -d$.

26 (a) Basis for all diagonal matrices: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) Add $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ = basis for symmetric matrices.

(c) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.

These are simple bases (among many others) for (a) diagonal matrices (b) symmetric matrices (c) skew-symmetric matrices. The dimensions are $3, 6, 3$.

27 I, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$;

echelon matrices do *not* form a subspace; they ***span*** the upper triangular matrices (not every U is an echelon matrix).

28 $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$; $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}$.

29 (a) The invertible matrices span the space of all 3 by 3 matrices (b) The rank one matrices also span the space of all 3 by 3 matrices (c) I by itself spans the space of all multiples cI.

30 $\begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}$. **Dimension** $= 4$.

31 (a) $y(x) =$ constant C (b) $y(x) = 3x$. (c) $y(x) = 3x + C = y_p + y_n$ solves $y' = 3$.

32 $y(0) = 0$ requires $A + B + C = 0$. One basis is $\cos x - \cos 2x$ and $\cos x - \cos 3x$.

33 (a) $y(x) = e^{2x}$ is a basis for all solutions to $y' = 2y$ (b) $y = x$ is a basis for all solutions to $dy/dx = y/x$ (First-order linear equation $\Rightarrow$ 1 basis function in solution space).

34 $y_1(x), y_2(x), y_3(x)$ can be $x, 2x, 3x$ $(\dim 1)$ or $x, 2x, x^2$ $(\dim 2)$ or x, x^2, x^3 $(\dim 3)$.

35 Basis 1, x, x^2, x^3, for cubic polynomials; basis $x-1$, x^2-1, x^3-1 for the subspace with $p(1) = 0$.

36 Basis for $\mathbf{S}$: $(1, 0, -1, 0)$, $(0, 1, 0, 0)$, $(1, 0, 0, -1)$; basis for $\mathbf{T}$: $(1, -1, 0, 0)$ and $(0, 0, 2, 1)$; $\mathbf{S} \cap \mathbf{T} =$ multiples of $(3, -3, 2, 1) =$ nullspace for 3 equations in $\mathbf{R}^4$ has dimension 1.

37 The subspace of matrices that have $AS = SA$ has dimension *three*. The 3 numbers a, b, c can be chosen independently in A.

38 (a) No, 2 vectors don't span $\mathbf{R}^3$ (b) No, 4 vectors in $\mathbf{R}^3$ are dependent (c) Yes, a basis (d) No, these three vectors are dependent

39 If the 5 by 5 matrix $[\,A \;\; \boldsymbol{b}\,]$ is invertible, $\boldsymbol{b}$ is not a combination of the columns of A: no solution to $A\boldsymbol{x} = \boldsymbol{b}$. If $[\,A \;\; \boldsymbol{b}\,]$ is singular, and the 4 columns of A are independent (rank 4), $\boldsymbol{b}$ *is* a combination of those columns. In this case $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.

40 (a) The functions $y = \sin x$, $y = \cos x$, $y = e^x$, $y = e^{-x}$ are a basis for solutions to $d^4y/dx^4 = y(x)$.

(b) A particular solution to $d^4y/dx^4 = y(x) + 1$ is $y(x) = -\mathbf{1}$. The complete solution is $y(x) = -1 + c_1 \sin x + c_2 \cos x + c_3 e^x + c_4 e^{-x}$ (or use another basis for the nullspace of the 4th derivative).

41 $$\boldsymbol{I} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} - \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} + \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} + \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} - \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}. \quad \begin{matrix} \text{The six } P\text{'s} \\ \text{are dependent} \end{matrix}.$$

Those five are independent: The 4th has $P_{11} = 1$ and cannot be a combination of the others. Then the 2nd cannot be (from $P_{32} = 1$) and also 5th ($P_{32} = 1$). Continuing, a nonzero combination of all five could not be zero. Further challenge: How many independent 4 by 4 permutation matrices?

42 The dimension of $\boldsymbol{S}$ spanned by all rearrangements of $\boldsymbol{x}$ is (a) zero when $\boldsymbol{x} = \mathbf{0}$ (b) one when $\boldsymbol{x} = (1, 1, 1, 1)$ (c) three when $\boldsymbol{x} = (1, 1, -1, -1)$ because all rearrangements of this $\boldsymbol{x}$ are perpendicular to $(1, 1, 1, 1)$ (d) four when the $\boldsymbol{x}$'s are not equal and don't add to zero. **No $\boldsymbol{x}$ gives** dim $\boldsymbol{S} = 2$. I owe this nice problem to Mike Artin—the answers are the same in higher dimensions: $0, 1, n-1, n$.

43 The problem is to show that the $\boldsymbol{u}$'s, $\boldsymbol{v}$'s, $\boldsymbol{w}$'s together are independent. We know the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s together are a basis for $\boldsymbol{V}$, and the $\boldsymbol{u}$'s and $\boldsymbol{w}$'s together are a basis for $\boldsymbol{W}$. Suppose a combination of $\boldsymbol{u}$'s, $\boldsymbol{v}$'s, $\boldsymbol{w}$'s gives $\mathbf{0}$. **To be proved**: All coefficients $=$ zero.

Key idea: In that combination giving $\mathbf{0}$, the part $\boldsymbol{x}$ from the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s is in $\boldsymbol{V}$. So the part from the $\boldsymbol{w}$'s is $-\boldsymbol{x}$. This part is now in $\boldsymbol{V}$ and also in $\boldsymbol{W}$. But if $-\boldsymbol{x}$ is in $\boldsymbol{V} \cap \boldsymbol{W}$ it is a combination of $\boldsymbol{u}$'s only. Now the combination giving $\mathbf{0}$ uses only $\boldsymbol{u}$'s and $\boldsymbol{v}$'s (independent in $\boldsymbol{V}$!) so all coefficients of $\boldsymbol{u}$'s and $\boldsymbol{v}$'s must be zero. Then $\boldsymbol{x} = \mathbf{0}$ and the coefficients of the $\boldsymbol{w}$'s are also zero.

44 The inputs to multiplication by an m by n matrix fill $\mathbf{R}^n$: dimension n. The outputs (column space!) have dimension r. The nullspace has $n - r$ special solutions. The formula becomes $r + (n - r) = n$.

45 If the left side of $\dim(\mathbf{V}) + \dim(\mathbf{W}) = \dim(\mathbf{V} \cap \mathbf{W}) + \dim(\mathbf{V} + \mathbf{W})$ is greater than n, then $\dim(\mathbf{V} \cap \mathbf{W})$ must be greater than zero. So $\mathbf{V} \cap \mathbf{W}$ contains nonzero vectors.

Oh here is a more basic approach: Put a basis for **V** and then a basis for **W** in the columns of a matrix A. Then A has more columns than rows and there is a nonzero solution to $A\boldsymbol{x} = \mathbf{0}$. That $\boldsymbol{x}$ gives a combination of the **V** columns $=$ a combination of the **W** columns.

46 If $A^2 =$ zero matrix, this says that each column of A is in the nullspace of A. If the column space has dimension r, the nullspace has dimension $10 - r$. So we must have $r \leq 10 - r$ and this leads to $r \leq 5$.

Problem Set 3.5, page 190

1 (a) Row and column space dimensions $= 5$, nullspace dimension $= 4$, $\dim(\boldsymbol{N}(A^{\mathrm{T}}))$ $= 2$ sum $5 + 5 + 4 + 2 = 16 = m + n$

(b) Column space is $\mathbf{R}^3$; left nullspace contains only $\mathbf{0}$.

2 A: Row space basis $=$ row $1 = (1, 2, 4)$; nullspace $(-2, 1, 0)$ and $(-4, 0, 1)$; column space basis $=$ column $1 = (1, 2)$; left nullspace $(-2, 1)$. B: Row space basis $=$ both rows $= (1, 2, 4)$ and $(2, 5, 8)$; column space basis $=$ two columns $= (1, 2)$ and $(2, 5)$; nullspace $(-4, 0, 1)$; left nullspace basis is empty because the space contains only $\boldsymbol{y} = \mathbf{0}$: the rows of B are independent.

3 Row space basis $=$ first two rows of U; column space basis $=$ pivot columns (of A not U) $= (1, 1, 0)$ and $(3, 4, 1)$; nullspace basis $(1, 0, 0, 0, 0)$, $(0, 2, -1, 0, 0)$, $(0, 2, 0, -2, 1)$; left nullspace $(1, -1, 1) =$ last row of $E^{-1} = L$.

4 (a) $\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) Impossible: $r+(n-r)$ must be 3 (c) $[1 \quad 1]$ (d) $\begin{bmatrix} 9 & -3 \\ 3 & -1 \end{bmatrix}$

(e) *Impossible* Row space $=$ column space requires $m = n$. Then $m - r = n - r$; nullspaces have the same dimension. Section 4.1 will prove $\boldsymbol{N}(A)$ and $\boldsymbol{N}(A^{\mathrm{T}})$ orthogonal to the row and column spaces respectively—here those are the same space.

5 $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$ has those rows spanning its row space. $B = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$ has the same rows spanning its nullspace and $AB^{\mathrm{T}} = 0$.

6 A: dim $\mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{1}$: Rows $(0, 3, 3, 3)$ and $(0, 1, 0, 1)$; columns $(3, 0, 1)$ and $(3, 0, 0)$; nullspace $(1, 0, 0, 0)$ and $(0, -1, 0, 1)$; $\boldsymbol{N}(A^{\mathrm{T}})$ $(0, 1, 0)$. B: dim $\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{2}$ Row space (1), column space $(1, 4, 5)$, nullspace: empty basis, $\boldsymbol{N}(A^{\mathrm{T}})$ $(-4, 1, 0)$ and $(-5, 0, 1)$.

7 Invertible 3 by 3 matrix A: row space basis $=$ column space basis $= (1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$; nullspace basis and left nullspace basis are *empty*. Matrix $B = \begin{bmatrix} A & A \end{bmatrix}$: row space basis $(1, 0, 0, 1, 0, 0)$, $(0, 1, 0, 0, 1, 0)$ and $(0, 0, 1, 0, 0, 1)$; column space basis

$(1,0,0)$, $(0,1,0)$, $(0,0,1)$; nullspace basis $(-1,0,0,1,0,0)$ and $(0,-1,0,0,1,0)$ and $(0,0,-1,0,0,1)$; left nullspace basis is empty.

8 $\begin{bmatrix} I & 0 \end{bmatrix}$ and $\begin{bmatrix} I & I; & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \end{bmatrix} = 3$ by 2 have *row space dimensions* $= 3, 3, 0 =$ column space dimensions; *nullspace dimensions* $2, 3, 2$; *left nullspace dimensions* $0, 2, 3$.

9 (a) Same row space and nullspace. So rank (dimension of row space) is the same

(b) Same column space and left nullspace. Same rank (dimension of column space).

10 For **rand** (3), almost surely rank$= 3$, nullspace and left nullspace contain only $(0,0,0)$. For **rand** $(3,5)$ the rank is almost surely 3 and the dimension of the nullspace is 2.

11 (a) No solution means that $r < m$. Always $r \le n$. Can't compare m and n here.

(b) Since $m - r > 0$, the left nullspace must contain a nonzero vector.

12 A neat choice is $\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $r + (n - r) = n = 3$ does not match $2 + 2 = 4$. Only $\boldsymbol{v} = \mathbf{0}$ is in both $\boldsymbol{N}(A)$ and $\boldsymbol{C}(A^{\mathrm{T}})$.

13 (a) *False*: Usually row space $\neq$ column space (they do not have the same dimension!)

(b) *True*: A and $-A$ have the same four subspaces

(c) *False* (choose A and B same size and invertible: then they have the same four subspaces)

14 Row space basis can be the nonzero rows of U: $(1,2,3,4)$, $(0,1,2,3)$, $(0,0,1,2)$; nullspace basis $(0,1,-2,1)$ as for U; column space basis $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ (happen to have $\boldsymbol{C}(A) = \boldsymbol{C}(U) = \mathbf{R}^3$); left nullspace has empty basis.

15 After a row exchange, the row space and nullspace stay the same; $(2,1,3,4)$ is in the new left nullspace after the row exchange.

16 If $A\boldsymbol{v} = \mathbf{0}$ and $\boldsymbol{v}$ is a row of A then $\boldsymbol{v} \cdot \boldsymbol{v} = 0$. So $\boldsymbol{v} = \mathbf{0}$.

17 Row space $= yz$ plane; column space $= xy$ plane; nullspace $= x$ axis; left nullspace $= z$ axis. For $I + A$: Row space $=$ column space $= \mathbf{R}^3$, both nullspaces contain only the zero vector.

18 Row $3-2$ row $2+$ row $1=$ zero row so the vectors $c(1,-2,1)$ are in the left nullspace. The same vectors happen to be in the nullspace (an accident for this matrix).

19 (a) Elimination on $A\boldsymbol{x}=\mathbf{0}$ leads to $0=b_3-b_2-b_1$ so $(-1,-1,1)$ is in the left nullspace. (b) 4 by 3: Elimination leads to $b_3-2b_1=0$ and $b_4+b_2-4b_1=0$, so $(-2,0,1,0)$ and $(-4,1,0,1)$ are in the left nullspace. *Why*? Those vectors multiply the matrix to give *zero rows* in $\boldsymbol{v}A$. Section 4.1 will show another approach: $A\boldsymbol{x}=\boldsymbol{b}$ is solvable ($\boldsymbol{b}$ is in $\boldsymbol{C}(A)$) exactly when $\boldsymbol{b}$ is orthogonal to the left nullspace.

20 (a) Special solutions $(-1,2,0,0)$ and $(-\frac{1}{4},0,-3,1)$ are perpendicular to the rows of R (and rows of ER). (b) $A^{\mathrm{T}}\boldsymbol{y}=\mathbf{0}$ has 1 independent solution = last row of E^{-1}. ($E^{-1}A=R$ has a zero row, which is just the transpose of $A^{\mathrm{T}}\boldsymbol{y}=\mathbf{0}$).

21 (a) $\boldsymbol{u}$ and $\boldsymbol{w}$ (b) $\boldsymbol{v}$ and $\boldsymbol{z}$ (c) rank <2 if $\boldsymbol{u}$ and $\boldsymbol{w}$ are dependent or if $\boldsymbol{v}$ and $\boldsymbol{z}$ are dependent (d) The rank of $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}+\boldsymbol{w}\boldsymbol{z}^{\mathrm{T}}$ is 2.

22 $$A=\begin{bmatrix} & \\ \boldsymbol{u} & \boldsymbol{w} \\ & \end{bmatrix}\begin{bmatrix} \boldsymbol{v}^{\mathrm{T}} \\ \boldsymbol{z}^{\mathrm{T}} \end{bmatrix}=\begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}=\begin{bmatrix} 3 & 2 \\ 4 & 2 \\ 5 & 1 \end{bmatrix}$$ $\boldsymbol{u},\boldsymbol{w}$ span column space; $\boldsymbol{v},\boldsymbol{z}$ span row space

23 As in Problem 22: Row space basis $(3,0,3),(1,1,2)$; column space basis $(1,4,2)$, $(2,5,7)$; the rank of (3 by 2) times (2 by 3) cannot be larger than the rank of either factor, so rank ≤ 2 and the 3 by 3 product is not invertible.

24 $A^{\mathrm{T}}\boldsymbol{y}=\boldsymbol{d}$ puts $\boldsymbol{d}$ in the *row space* of A; unique solution if the *left nullspace* (nullspace of A^{T}) contains only $\boldsymbol{y}=\mathbf{0}$.

25 (a) *True* (A and A^{T} have the same rank) (b) *False* $A=[\,1\quad 0\,]$ and A^{T} have very different left nullspaces (c) *False* (A can be invertible and unsymmetric even if $\boldsymbol{C}(A)=\boldsymbol{C}(A^{\mathrm{T}})$) (d) *True* (The subspaces for A and $-A$ are always the same. If $A^{\mathrm{T}}=A$ or $A^{\mathrm{T}}=-A$ they are also the same for A^{T})

26 Choose $d=bc/a$ to make $\begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}$ a rank-1 matrix. Then the row space has basis (a,b) and the nullspace has basis $(-b,a)$. Those two vectors are perpendicular !

27 B and C (checkers and chess) both have rank 2 if $p\neq 0$. Row 1 and 2 are a basis for the row space of C, $B^{\mathrm{T}}\boldsymbol{y}=\mathbf{0}$ has 6 special solutions with -1 and 1 separated by a zero;

$\boldsymbol{N}(C^{\mathrm{T}})$ has $(-1,0,0,0,0,0,0,1)$ and $(0,-1,0,0,0,0,1,0)$ and columns $3,4,5,6$ of I; $\boldsymbol{N}(C)$ is a challenge: one vector in $\boldsymbol{N}(C)$ is $(1,0,\ldots,0,-1)$.

28 $a_{11}=1, a_{12}=0, a_{13}=1, a_{22}=0, a_{32}=1, a_{31}=0, a_{23}=1, a_{33}=0, a_{21}=1$. (Need to specify the five moves).

29 The subspaces for $A=\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ are pairs of orthogonal lines ($\boldsymbol{v}$ and $\boldsymbol{v}^{\perp}$, $\boldsymbol{u}$ and $\boldsymbol{u}^{\perp}$). If B has those same four subspaces then $B=cA$ with $c\neq 0$.

30 (a) $AX=0$ if each column of X is a multiple of $(1,1,1)$; $\dim(\text{nullspace})=3$. (b) If $AX=B$ then all columns of B add to zero; dimension of the B's $=6$. (c) $3+6=\dim(M^{3\times 3})=9$ entries in a 3 by 3 matrix.

31 The key is equal row spaces. First row of A = combination of the rows of B: only possible combination (notice I) is 1 (row 1 of B). Same for each row so $F=G$.

Problem Set 4.1, page 202

1 Both nullspace vectors will be orthogonal to the row space vector in $\mathbf{R}^3$. The column space of A and the nullspace of A^{T} are perpendicular lines in $\mathbf{R}^2$ because rank $= 1$.

2 The nullspace of a 3 by 2 matrix with rank 2 is $\mathbf{Z}$ (only the zero vector because the 2 columns are independent). So $\boldsymbol{x}_n = \mathbf{0}$, and row space $= \mathbf{R}^2$. Column space $=$ plane perpendicular to left nullspace $=$ line in $\mathbf{R}^3$ (because the rank is 2).

3 (a) One way is to use these two columns directly: $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & -3 & 1 \\ -3 & 5 & -2 \end{bmatrix}$

(b) Impossible because $\boldsymbol{N}(A)$ and $\boldsymbol{C}(A^{\mathrm{T}})$ are orthogonal subspaces: $\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ is not orthogonal to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ in $\boldsymbol{C}(A)$ and $\boldsymbol{N}(A^{\mathrm{T}})$ is impossible: not perpendicular

(d) Rows orthogonal to columns makes A times $A =$ zero matrix ρ. An example is $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

(e) $(1, 1, 1)$ in the nullspace (columns add to the zero vector) and also $(1, 1, 1)$ is in the row space: no such matrix.

4 If $AB = 0$, the columns of B are in the *nullspace* of A and the rows of A are in the *left nullspace* of B. If rank $= 2$, all those four subspaces have dimension at least 2 which is impossible for 3 by 3.

5 (a) If $A\boldsymbol{x} = \boldsymbol{b}$ has a solution and $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$, then $\boldsymbol{y}$ is perpendicular to $\boldsymbol{b}$. $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}\boldsymbol{y}) = 0$. This says again that $\boldsymbol{C}(A)$ is orthogonal to $\boldsymbol{N}(A^{\mathrm{T}})$. (b) If $A^{\mathrm{T}}\boldsymbol{y} = (1, 1, 1)$ has a solution, $(1, 1, 1)$ is a combination of the rows of A. It is in the **row space** and is orthogonal to every $\boldsymbol{x}$ in the **nullspace**.

6 Multiply the equations by $y_1, y_2, y_3 = 1, 1, -1$. Now the equations add to $0 = 1$ so there is no solution. In subspace language, $\boldsymbol{y} = (1, 1, -1)$ is in the left nullspace. $A\boldsymbol{x} = \boldsymbol{b}$ would need $0 = (\boldsymbol{y}^{\mathrm{T}} A)\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 1$ but here $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 1$.

7 Multiply the 3 equations by $\boldsymbol{y} = (1, 1, -1)$. Then $x_1 - x_2 = 1$ plus $x_2 - x_3 = 1$ *minus* $x_1 - x_3 = 1$ is $0 = 1$. Key point: This $\boldsymbol{y}$ in $\boldsymbol{N}(A^{\mathrm{T}})$ is not orthogonal to $\boldsymbol{b} = (1, 1, 1)$ so $\boldsymbol{b}$ is not in the column space and $A\boldsymbol{x} = \boldsymbol{b}$ has *no solution*.

8 Figure 4.3 has $\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n$, where $\boldsymbol{x}_r$ is in the row space and $\boldsymbol{x}_n$ is in the nullspace. Then $A\boldsymbol{x}_n = \mathbf{0}$ and $A\boldsymbol{x} = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r$. The example has $\boldsymbol{x} = (1, 0)$ and row space = line through $(1, 1)$ so the splitting is $\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n = \left(\frac{1}{2}, \frac{1}{2}\right) + \left(\frac{1}{2}, -\frac{1}{2}\right)$. All $A\boldsymbol{x}$ are in $\boldsymbol{C}(A)$.

9 $A\boldsymbol{x}$ is always in the *column space* of A. If $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x}$ is also in the *nullspace* of A^{T}. Those subspaces are perpendicular. So $A\boldsymbol{x}$ is perpendicular to itself. Conclusion: $A\boldsymbol{x} = \mathbf{0}$ if $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$.

10 (a) With $A^{\mathrm{T}} = A$, the column and row spaces are the *same*. The nullspace is always perpendicular to the row space. (b) $\boldsymbol{x}$ is in the nullspace and $\boldsymbol{z}$ is in the column space = row space: so these "eigenvectors" $\boldsymbol{x}$ and $\boldsymbol{z}$ have $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{z} = 0$.

11 ***For A***: The nullspace is spanned by $(-2, 1)$, the row space is spanned by $(1, 2)$. The column space is the line through $(1, 3)$ and $\boldsymbol{N}(A^{\mathrm{T}})$ is the perpendicular line through $(3, -1)$. ***For B***: The nullspace of B is spanned by $(0, 1)$, the row space is spanned by $(1, 0)$. The column space and left nullspace are the same as for A.

12 $\boldsymbol{x} = (2, 0)$ splits into $\boldsymbol{x}_r + \boldsymbol{x}_n = (1, -1) + (1, 1)$. Notice $\boldsymbol{N}(A^{\mathrm{T}})$ is the $y - z$ plane.

13 $V^{\mathrm{T}}W =$ zero matrix makes each column of V orthogonal to each column of W. This means: each basis vector for $\boldsymbol{V}$ is orthogonal to each basis vector for $\boldsymbol{W}$. Then *every* $\boldsymbol{v}$ in $\boldsymbol{V}$ (combinations of the basis vectors) is orthogonal to *every* $\boldsymbol{w}$ in $\boldsymbol{W}$.

14 $A\boldsymbol{x} = B\widehat{\boldsymbol{x}}$ means that $[\, A \quad B\,] \begin{bmatrix} \boldsymbol{x} \\ -\widehat{\boldsymbol{x}} \end{bmatrix} = \mathbf{0}$. Three homogeneous equations (zero right hand sides) in four unknowns always have a nonzero solution. Here $\boldsymbol{x} = (3, 1)$ and

$\widehat{\boldsymbol{x}} = (1, 0)$ and $A\boldsymbol{x} = B\widehat{\boldsymbol{x}} = (5, 6, 5)$ is in both column spaces. Two planes in $\mathbf{R}^3$ must share a line.

15 A p-dimensional and a q-dimensional subspace of $\mathbf{R}^n$ share at least a line if $\boldsymbol{p} + \boldsymbol{q} > \boldsymbol{n}$. (The $p + q$ basis vectors of $\boldsymbol{V}$ and $\boldsymbol{W}$ cannot be independent, so same combination of the basis vectors of $\boldsymbol{V}$ is also a combination of the basis vectors of $\boldsymbol{W}$.)

16 $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$ leads to $(A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{y} = 0$. Then $\boldsymbol{y} \perp A\boldsymbol{x}$ and $\boldsymbol{N}(A^{\mathrm{T}}) \perp \boldsymbol{C}(A)$.

17 If $\boldsymbol{S}$ is the subspace of $\mathbf{R}^3$ containing only the zero vector, then $\boldsymbol{S}^{\perp}$ is all of $\mathbf{R}^3$. If $\boldsymbol{S}$ is spanned by $(1, 1, 1)$, then $\boldsymbol{S}^{\perp}$ is the plane spanned by $(1, -1, 0)$ and $(1, 0, -1)$. If $\boldsymbol{S}$ is spanned by $(2, 0, 0)$ and $(0, 0, 3)$, then $\boldsymbol{S}^{\perp}$ is the line spanned by $(0, 1, 0)$.

18 $\boldsymbol{S}^{\perp}$ contains all vectors perpendicular to those two given vectors. So $\boldsymbol{S}^{\perp}$ is the nullspace of $A = \begin{bmatrix} 1 & 5 & 1 \\ 2 & 2 & 2 \end{bmatrix}$. Therefore $\boldsymbol{S}^{\perp}$ is a *subspace* even if $\boldsymbol{S}$ is not.

19 $\boldsymbol{L}^{\perp}$ is the 2-*dimensional subspace* (*a plane*) in $\mathbf{R}^3$ perpendicular to $\boldsymbol{L}$. Then $(\boldsymbol{L}^{\perp})^{\perp}$ is a 1-*dimensional subspace* (*a line*) perpendicular to $\boldsymbol{L}^{\perp}$. In fact $(\boldsymbol{L}^{\perp})^{\perp}$ is $\boldsymbol{L}$.

20 If $\boldsymbol{V}$ is the whole space $\mathbf{R}^4$, then $\boldsymbol{V}^{\perp}$ contains only the *zero vector*. Then $(\boldsymbol{V}^{\perp})^{\perp} =$ all vectors perpendicular to the zero vector $= \mathbf{R}^4 = \boldsymbol{V}$.

21 For example $(-5, 0, 1, 1)$ and $(0, 1, -1, 0)$ span $\boldsymbol{S}^{\perp}$ =nullspace of $A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix}$.

22 $(1, 1, 1, 1)$ is a basis for the line $\boldsymbol{P}^{\perp}$ orthogonal to P. $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ has $\boldsymbol{P}$ as its nullspace and $\boldsymbol{P}^{\perp}$ as its row space.

23 $\boldsymbol{x}$ in $\boldsymbol{V}^{\perp}$ is perpendicular to every vector in $\boldsymbol{V}$. Since $\boldsymbol{V}$ contains all the vectors in $\boldsymbol{S}$, $\boldsymbol{x}$ is perpendicular to every vector in $\boldsymbol{S}$. So every $\boldsymbol{x}$ in $\boldsymbol{V}^{\perp}$ is also in $\boldsymbol{S}^{\perp}$.

24 $AA^{-1} = I$: Column 1 of A^{-1} is orthogonal to rows $2, 3, \ldots, n$ and therefore to the space spanned by those rows.

25 If the columns of A are unit vectors, all mutually perpendicular, then $A^{\mathrm{T}}A = I$. Simple but important! We write Q for such a matrix.

26 $A = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$, This example shows a matrix with perpendicular columns. $A^{\mathrm{T}}A = 9I$ is *diagonal*: $(A^{\mathrm{T}}A)_{ij} =$ (column i of A) $\cdot$ (column j of A). When the columns are *unit vectors*, then $A^{\mathrm{T}}A = I$.

27 The lines $3x + y = b_1$ and $6x + 2y = b_2$ are **parallel**. They are the same line if $b_2 = 2b_1$. In that case (b_1, b_2) is perpendicular to $(\mathbf{-2, 1})$. The nullspace of the 2 by 2 matrix is the line $3x + y = \mathbf{0}$. One particular vector in the nullspace is $(\mathbf{-1, 3})$.

28 (a) $(1, -1, 0)$ is in both planes. Normal vectors are perpendicular, but planes still intersect! Two planes in $\mathbf{R}^3$ can't be orthogonal. (b) Need *three* orthogonal vectors to span the whole orthogonal complement in $\mathbf{R}^5$. (c) Lines in $\mathbf{R}^3$ can meet at the zero vector without being orthogonal.

29 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{bmatrix}$; A has $\boldsymbol{v} = (1, 2, 3)$ in row and column spaces B has $\boldsymbol{v}$ in its column space and nullspace. $\boldsymbol{v}$ **can not** be in the nullspace and row space, or in the left nullspace and column space. These spaces are orthogonal and $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} \neq 0$.

30 When $AB = 0$, every column of B is multiplied by A to give zero. So the column space of B is contained in the nullspace of A. Therefore the dimension of $\boldsymbol{C}(B) \leq$ dimension of $\boldsymbol{N}(A)$. This means $\text{rank}(B) \leq 4 - \text{rank}(A)$.

31 null(N') produces a basis for the *row space* of A (perpendicular to $\boldsymbol{N}(A)$).

32 We need $\boldsymbol{r}^{\mathrm{T}}\boldsymbol{n} = 0$ and $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\ell} = 0$. All possible examples have the form $a\boldsymbol{c}\boldsymbol{r}^{\mathrm{T}}$ with $a \neq 0$.

33 Both $\boldsymbol{r}$'s must be orthogonal to both $\boldsymbol{n}$'s, both $\boldsymbol{c}$'s must be orthogonal to both $\boldsymbol{\ell}$'s, each pair ($\boldsymbol{r}$'s, $\boldsymbol{n}$'s, $\boldsymbol{c}$'s, and $\boldsymbol{\ell}$'s) must be independent. Fact: All A's with these subspaces have the form $[\boldsymbol{c}_1\ \boldsymbol{c}_2]M[\boldsymbol{r}_1\ \boldsymbol{r}_2]^{\mathrm{T}}$ for a 2 by 2 invertible M.

You must take $[\boldsymbol{c}_1, \boldsymbol{c}_2]$ times $[\boldsymbol{r}_1, \boldsymbol{r}_2]^{\mathrm{T}}$.

Problem Set 4.2, page 214

1 (a) $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 5/3$; $\boldsymbol{p} = 5\boldsymbol{a}/3 = (5/3, 5/3, 5/3)$; $\boldsymbol{e} = (-2, 1, 1)/3$

(b) $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = -1$; $\boldsymbol{p} = \boldsymbol{a}$; $\boldsymbol{e} = \mathbf{0}$.

2 (a) The projection of $\boldsymbol{b} = (\cos\theta, \sin\theta)$ onto $\boldsymbol{a} = (1, 0)$ is $\boldsymbol{p} = (\cos\theta, 0)$
(b) The projection of $\boldsymbol{b} = (1,1)$ onto $\boldsymbol{a} = (1,-1)$ is $\boldsymbol{p} = (0,0)$ since $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 0$.

The picture for part (a) has the vector $\boldsymbol{b}$ at an angle θ with the horizontal $\boldsymbol{a}$. The picture for part (b) has vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ at a 90° angle.

3 $P_1 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ and $P_1\boldsymbol{b} = \frac{1}{3}\begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$. $P_2 = \frac{1}{11}\begin{bmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{bmatrix}$ and $P_2\boldsymbol{b} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$.

4 $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_2 = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. P_1 projects onto $(1,0)$, P_2 projects onto $(1,-1)$. $P_1P_2 \neq 0$ and $P_1 + P_2$ is not a projection matrix. $(P_1 + P_2)^2$ is different from $P_1 + P_2$.

5 $P_1 = \frac{1}{9}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix}$ and $P_2 = \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$.

P_1 and P_2 are the projection matrices onto the lines through $\boldsymbol{a}_1 = (-1, 2, 2)$ and $\boldsymbol{a}_2 = (2, 2, -1)$. $P_1P_2 =$ *zero matrix because* $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$.

6 $\boldsymbol{p}_1 = (\frac{1}{9}, -\frac{2}{9}, -\frac{2}{9})$ and $\boldsymbol{p}_2 = (\frac{4}{9}, \frac{4}{9}, -\frac{2}{9})$ and $\boldsymbol{p}_3 = (\frac{4}{9}, -\frac{2}{9}, \frac{4}{9})$. So $\boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 = \boldsymbol{b}$.

7 $P_1 + P_2 + P_3 = \frac{1}{9}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix} + \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix} + \frac{1}{9}\begin{bmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{bmatrix} = I.$

We *can* add projections onto *orthogonal vectors* to get the projection matrix onto the larger space. This is important.

8 The projections of $(1,1)$ onto $(1,0)$ and $(1,2)$ are $\boldsymbol{p}_1 = (1,0)$ and $\boldsymbol{p}_2 = \frac{3}{5}(1,2)$. Then $\boldsymbol{p}_1 + \boldsymbol{p}_2 \neq \boldsymbol{b}$. The sum of projections is not a projection because $(1,0)$ and $(2,1)$ are *not orthogonal*.

9 Since A is invertible, $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ separates into $AA^{-1}(A^{\mathrm{T}})^{-1}A^{\mathrm{T}} = I$. And I is the projection matrix onto all of $\mathbf{R}^2$.

10 $P_2 = \dfrac{\boldsymbol{a}_2\boldsymbol{a}_2^{\mathrm{T}}}{\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{a}_2} = \begin{bmatrix} 0.2 & 0.4 \\ 0.4 & 0.8 \end{bmatrix}$, $P_2\boldsymbol{a}_1 = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}$, $P_1 = \dfrac{\boldsymbol{a}_1\boldsymbol{a}_1^{\mathrm{T}}}{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_1P_2\boldsymbol{a}_1 = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix}$. This is not $\boldsymbol{a}_1 = (1, 0)$

No, $\boldsymbol{P_1P_2} \neq (P_1P_2)^2$.

11 (a) $\boldsymbol{p} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = (2, 3, 0)$, $\boldsymbol{e} = (0, 0, 4)$, $A^{\mathrm{T}}\boldsymbol{e} = \boldsymbol{0}$

(b) $\boldsymbol{p} = (4, 4, 6)$ and $\boldsymbol{e} = \boldsymbol{0}$ because $\boldsymbol{b}$ is in the column space of A.

12 $P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ = projection matrix onto the column space of A (the xy plane)

$P_2 = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ = Projection matrix $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ onto the second column space. Certainly $(P_2)^2 = P_2$. A true projection matrix.

13 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, P = square matrix = $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\boldsymbol{p} = P\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$.

14 The projection of this $\boldsymbol{b}$ = column 1 of A onto the column space of A is $\boldsymbol{b}$ itself because $\boldsymbol{b}$ is in that column space. But P is not necessarily I.

$A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 0 \end{bmatrix}$ gives $P = \dfrac{1}{21}\begin{bmatrix} 5 & 8 & -4 \\ 8 & 17 & 2 \\ -4 & 2 & 20 \end{bmatrix}$ and $\boldsymbol{b} = P\boldsymbol{b} = \boldsymbol{p} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$ = 2 (column 1 of A).

15 $2A$ has the same column space as A. Then P is the same for A and $2A$m but $\widehat{\boldsymbol{x}}$ for $2A$ is *half* of $\widehat{\boldsymbol{x}}$ for A.

16 $\frac{1}{2}(1, 2, -1) + \frac{3}{2}(1, 0, 1) = (2, 1, 1)$. So $\boldsymbol{b}$ is in the plane. Projection shows $P\boldsymbol{b} = \boldsymbol{b}$.

17 If $P^2 = P$ then $(\boldsymbol{I} - \boldsymbol{P})^{\mathbf{2}} = (I - P)(I - P) = I - PI - IP + P^2 = \boldsymbol{I} - \boldsymbol{P}$. When P projects onto the column space, $I - P$ projects onto the *left nullspace*.

18 (a) $I - P$ is the projection matrix onto $(1, -1)$ in the perpendicular direction to $(1, 1)$

(b) $I - P$ projects onto the plane $x + y + z = 0$ perpendicular to $(1, 1, 1)$.

19 For any basis vectors in the plane $x - y - 2z = 0$, say $(1, 1, 0)$ and $(2, 0, 1)$, the matrix $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ is $\begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

20 $\boldsymbol{e} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$, $Q = \dfrac{\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}} = \begin{bmatrix} 1/6 & -1/6 & -1/3 \\ -1/6 & 1/6 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}$, $I - Q = \begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

21 $\big(A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\big)^2 = A(A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}A)(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. So $P^2 = P$. $P\boldsymbol{b}$ is in the column space (where P projects). Then its projection $P(P\boldsymbol{b})$ is also $P\boldsymbol{b}$.

22 $P^{\mathrm{T}} = (A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}})^{\mathrm{T}} = A((A^{\mathrm{T}}A)^{-1})^{\mathrm{T}}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = P$. ($A^{\mathrm{T}}A$ is symmetric!)

23 If A is invertible then its column space is all of $\mathbf{R}^n$. So $P = I$ and $\boldsymbol{e} = \mathbf{0}$.

24 The nullspace of A^{T} is *orthogonal* to the column space $\boldsymbol{C}(A)$. So if $A^{\mathrm{T}}\boldsymbol{b} = \mathbf{0}$, the projection of $\boldsymbol{b}$ onto $\boldsymbol{C}(A)$ should be $\boldsymbol{p} = \mathbf{0}$. Check $P\boldsymbol{b} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = A(A^{\mathrm{T}}A)^{-1}\mathbf{0}$.

25 **The column space of P is the space that P projects onto**. The column space of A always contains all outputs $A\boldsymbol{x}$ and here the outputs $P\boldsymbol{x}$ fill the subspace S. Then rank of P = dimension of $S = n$.

26 A^{-1} exists since the rank is $r = m$. Multiply $A^2 = A$ by A^{-1} to get $A = I$.

27 If $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x}$ is in the **nullspace of A^{T}**. But $A\boldsymbol{x}$ is always in the **column space of A**. To be in both of those perpendicular spaces, $A\boldsymbol{x}$ must be zero. So A and $A^{\mathrm{T}}A$ have the *same nullspace*: $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ exactly when $A\boldsymbol{x} = \mathbf{0}$.

28 $P^2 = P = P^{\mathrm{T}}$ give $P^{\mathrm{T}}P = P$. Then the $(2, 2)$ entry of P equals the $(2, 2)$ entry of $P^{\mathrm{T}}P$. But the $(2, 2)$ entry of $P^{\mathrm{T}}P$ is the length squared of column 2.

29 $A = B^{\mathrm{T}}$ has independent columns, so $A^{\mathrm{T}}A$ (which is BB^{T}) must be invertible.

30 (a) The column space is the line through $\boldsymbol{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ so $P_C = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \dfrac{1}{25}\begin{bmatrix} 9 & 12 \\ 12 & 25 \end{bmatrix}$.

The formula $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ needs independent columns—this A has dependent columns. The update formula is correct.

(b) The row space is the line through $\boldsymbol{v} = (1,2,2)$ and $P_R = \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}/\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}$. Always $P_C A = A$ (columns of A project to themselves) and $AP_R = A$. Then $\boldsymbol{P_C A P_R = A}$.

31 *Test*: The error $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$ must be perpendicular to all the $\boldsymbol{a}$'s.

32 Since $P_1\boldsymbol{b}$ is in $\boldsymbol{C}(A)$ and P_2 projects onto that column space, $P_2(P_1\boldsymbol{b})$ equals $P_1\boldsymbol{b}$. So $P_2P_1 = P_1 = \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$ where $\boldsymbol{a} = (1,2,0)$.

33 Each $\boldsymbol{b}_1$ to $\boldsymbol{b}_{99}$ is multiplied by $\frac{1}{999} - \frac{1}{1000}\left(\frac{1}{999}\right) = \frac{999}{1000}\frac{1}{999} = \frac{1}{1000}$. The last pages of the book discuss least squares and the Kalman filter.

Problem Set 4.3, page 229

1 $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$ give $A^{\mathrm{T}}A = \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix}$ and $A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$.

$A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ gives $\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\boldsymbol{p} = A\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$ and $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix}$
$E = \|\boldsymbol{e}\|^2 = \mathbf{44}$

2 $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. This $A\boldsymbol{x} = \boldsymbol{b}$ is unsolvable. Project $\boldsymbol{b}$ to $\boldsymbol{p} = P\boldsymbol{b} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$; When $\boldsymbol{p}$ replaces $\boldsymbol{b}$,

$\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ exactly solves $A\widehat{\boldsymbol{x}} = \boldsymbol{p}$.

3 In Problem 2, $\boldsymbol{p} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = (1,5,13,17)$ and $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = (-1,3,-5,3)$. This $\boldsymbol{e}$ is perpendicular to both columns of A. This shortest distance $\|\boldsymbol{e}\|$ is $\sqrt{44}$.

4 $E = (C + \mathbf{0}D)^2 + (C + \mathbf{1}D - 8)^2 + (C + \mathbf{3}D - 8)^2 + (C + \mathbf{4}D - 20)^2$. Then $\partial E/\partial C = 2C + 2(C + D - 8) + 2(C + 3D - 8) + 2(C + 4D - 20) = 0$ and $\partial E/\partial D = 1 \cdot 2(C + D - 8) + 3 \cdot 2(C + 3D - 8) + 4 \cdot 2(C + 4D - 20) = 0$. These two normal equations are again $\begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$.

5 $E = (C-0)^2 + (C-8)^2 + (C-8)^2 + (C-20)^2$. $A^{\mathrm{T}} = [\,1 \;\; 1 \;\; 1 \;\; 1\,]$ and $A^{\mathrm{T}}A = [\,4\,]$. $A^{\mathrm{T}}\boldsymbol{b} = [\,36\,]$ and $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = \mathbf{9} =$ best height C for the horizontal line. Errors $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = (-9, -1, -1, 11)$ still add to zero.

6 $\boldsymbol{a} = (1, 1, 1, 1)$ and $\boldsymbol{b} = (0, 8, 8, 20)$ give $\widehat{x} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 9$ and the projection is $\widehat{x}\boldsymbol{a} = \boldsymbol{p} = (9, 9, 9, 9)$. Then $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{a} = (-9, -1, -1, 11)^{\mathrm{T}}(1, 1, 1, 1) = 0$ and the shortest distance from $\boldsymbol{b}$ to the line through $\boldsymbol{a}$ is $\|\boldsymbol{e}\| = \sqrt{204}$.

7 Now the 4 by 1 matrix in $A\boldsymbol{x} = \boldsymbol{b}$ is $A = [\,0 \;\; 1 \;\; 3 \;\; 4\,]^{\mathrm{T}}$. Then $A^{\mathrm{T}}A = [\,26\,]$ and $A^{\mathrm{T}}\boldsymbol{b} = [\,112\,]$. Best $D = 112/26 = 56/13$.

8 $\widehat{\boldsymbol{x}} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 56/13$ and $\boldsymbol{p} = (56/13)(0, 1, 3, 4)$. $(C, D) = (9, 56/13)$ don't match $(C, D) = (1, 4)$ from Problems 1-4. Columns of A were not perpendicular so we can't project separately to find C and D.

9 Parabola Project $\boldsymbol{b}$ 4D to 3D $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = \begin{bmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \\ 400 \end{bmatrix}$.

Figure 4.9 (a) is fitting 4 points and 4.9 (b) is a projection in $\mathbf{R}^4$: same problem!

10 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. Then $\begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ 47 \\ -28 \\ 5 \end{bmatrix}$. **Exact cubic so $\boldsymbol{p} = \boldsymbol{b}$, $\boldsymbol{e} = \mathbf{0}$.** This Vandermonde matrix gives exact interpolation by a cubic at $0, 1, 3, 4$

11 (a) The best line $x = 1 + 4t$ gives the center point $\widehat{\boldsymbol{b}} = 9$ at center time, $\widehat{t} = 2$.

(b) The first equation $Cm + D\sum t_i = \sum b_i$ divided by m gives $C + D\widehat{t} = \widehat{\boldsymbol{b}}$. This shows: The best line goes through $\widehat{\boldsymbol{b}}$ at time $\widehat{t}$.

12 (a) $\boldsymbol{a} = (1, \ldots, 1)$ has $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = m$, $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = b_1 + \cdots + b_m$. Therefore $\widehat{x} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/m$ is the **mean** of the b's (their average value)

(b) $\boldsymbol{e} = \boldsymbol{b} - \widehat{x}\boldsymbol{a}$ and $\|\boldsymbol{e}\|^2 = (b_1 - \text{ mean })^2 + \cdots + (b_m - \text{ mean })^2 =$ **variance** (denoted by $\boldsymbol{\sigma^2}$).

(c) $\boldsymbol{p} = (3, 3, 3)$ and $\boldsymbol{e} = (-2, -1, 3)$ $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{e} = 0$. Projection matrix $P = \dfrac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

13 $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(\boldsymbol{b} - A\boldsymbol{x}) = \widehat{\boldsymbol{x}} - \boldsymbol{x}$. This tells us: When the components of $A\boldsymbol{x} - \boldsymbol{b}$ add to zero, so do the components of $\widehat{\boldsymbol{x}} - \boldsymbol{x}$: Unbiased.

14 The matrix $(\widehat{\boldsymbol{x}} - \boldsymbol{x})(\widehat{\boldsymbol{x}} - \boldsymbol{x})^{\mathrm{T}}$ is $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(\boldsymbol{b} - A\boldsymbol{x})(\boldsymbol{b} - A\boldsymbol{x})^{\mathrm{T}}A(A^{\mathrm{T}}A)^{-1}$. When the average of $(\boldsymbol{b} - A\boldsymbol{x})(\boldsymbol{b} - A\boldsymbol{x})^{\mathrm{T}}$ is $\sigma^2 I$, the average of $(\widehat{\boldsymbol{x}} - \boldsymbol{x})(\widehat{\boldsymbol{x}} - \boldsymbol{x})^{\mathrm{T}}$ will be the *output covariance matrix* $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\sigma^2 A(A^{\mathrm{T}}A)^{-1}$ which simplifies to $\sigma^2(A^{\mathrm{T}}A)^{-1}$. That gives the average of the squared output errors $\widehat{\boldsymbol{x}} - \boldsymbol{x}$.

15 When A has 1 column of 4 ones, Problem 14 gives the expected error $(\widehat{x} - x)^2$ as $\sigma^2(A^{\mathrm{T}}A)^{-1} = \sigma^2/4$. By taking m measurements, the variance drops from σ^2 to σ^2/m. This leads to the **Monte Carlo method** in Section 12.1.

16 $\dfrac{1}{10}b_{10} + \dfrac{\mathbf{9}}{\mathbf{10}}\widehat{x}_9 = \dfrac{1}{10}(b_1 + \cdots + b_{10})$. Knowing $\widehat{x}_9$ avoids adding all ten b's.

17 $\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 21 \end{bmatrix}$. The solution $\widehat{\boldsymbol{x}} = \begin{bmatrix} \mathbf{9} \\ \mathbf{4} \end{bmatrix}$ comes from $\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix}$.

18 $\boldsymbol{p} = A\widehat{\boldsymbol{x}} = (5, 13, 17)$ gives the heights of the closest line. The vertical errors are $\boldsymbol{b} - \boldsymbol{p} = (2, -6, 4)$. This error $\boldsymbol{e}$ has $P\boldsymbol{e} = P\boldsymbol{b} - P\boldsymbol{p} = \boldsymbol{p} - \boldsymbol{p} = \mathbf{0}$.

19 If $\boldsymbol{b} =$ error $\boldsymbol{e}$ then $\boldsymbol{b}$ is perpendicular to the column space of A. Projection $\boldsymbol{p} = \mathbf{0}$.

20 The matrix A has columns $1, 1, 1$ and $-1, 1, 2$. If $\boldsymbol{b} = A\widehat{\boldsymbol{x}} = (5, 13, 17)$ then $\widehat{\boldsymbol{x}} = (9, 4)$ and $\boldsymbol{e} = \mathbf{0}$ since $\boldsymbol{b} = 9\,(\text{column } 1) + 4\,(\text{column } 2)$ is *in the column space of* A.

21 $\boldsymbol{e}$ is in $\boldsymbol{N}(A^{\mathrm{T}})$; $\boldsymbol{p}$ is in $\boldsymbol{C}(A)$; $\widehat{\boldsymbol{x}}$ is in $\boldsymbol{C}(A^{\mathrm{T}})$; $\boldsymbol{N}(A) = \{\mathbf{0}\}$ = zero vector only.

22 The least squares equation is $\begin{bmatrix} 5 & \mathbf{0} \\ \mathbf{0} & 10 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$. Solution: $C = 1$, $D = -1$. The best line is $b = 1 - t$. Symmetric t's $\Rightarrow$ diagonal $A^{\mathrm{T}}A \Rightarrow$ easy solution.

23 $\boldsymbol{e}$ is orthogonal to $\boldsymbol{p}$ in $\mathbf{R}^m$; then $\|\boldsymbol{e}\|^2 = e^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{p}) = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{p}$.

24 The derivatives of $\|A\boldsymbol{x} - \boldsymbol{b}\|^2 = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} - 2\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}$ (this last term is constant) are zero when $2A^{\mathrm{T}}A\boldsymbol{x} = 2A^{\mathrm{T}}\boldsymbol{b}$, or $\boldsymbol{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.

25 3 points on a linewill give **equal slopes** $(b_2 - b_1)/(t_2 - t_1) = (b_3 - b_2)/(t_3 - t_2)$. Linear algebra: Orthogonal to the columns $(1, 1, 1)$ and (t_1, t_2, t_3) is $\boldsymbol{y} = (t_2 - t_3, t_3 - t_1, t_1 - t_2)$ in the left nullspace of A. $\boldsymbol{b}$ is in the column space! Then $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ is the same equal slopes condition written as $(b_2 - b_1)(t_3 - t_2) = (b_3 - b_2)(t_2 - t_1)$.

26 The unsolvable equations for $C + Dx + Ey = (0, 1, 3, 4)$ at the 4 corners are $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}$. Then $A^{\mathrm{T}}A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

and $A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 8 \\ -2 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -3/2 \end{bmatrix}$. At $x, y = 0, 0$ the best plane $2 - x - \frac{3}{2}y$ has height $C = \mathbf{2} =$ average of $0, 1, 3, 4$.

27 The shortest link connecting two lines in space is *perpendicular to those lines*.

28 If A has dependent columns, then $A^{\mathrm{T}}A$ is not invertable and the usual formula $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ will fail. Replace A in that formula by the matrix B that keeps *only the pivot columns of* A.

29 Only 1 plane contains $\mathbf{0}, \boldsymbol{a}_1, \boldsymbol{a}_2$ unless $\boldsymbol{a}_1, \boldsymbol{a}_2$ are *dependent*. Same test for $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1}$. If they are dependent, there is a vector $\boldsymbol{v}$ perpendicular to all the $\boldsymbol{a}$'s. Then they all lie on the plane $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = 0$ going through $\boldsymbol{x} = (0, 0, \ldots, 0)$.

30 When A has orthogonal columns $(1,\ldots,1)$ and $(T_1,\ldots,T_m)$, the matrix $A^{\mathrm{T}}A$ is **diagonal** with entries m and $T_1^2+\cdots+T_m^2$. Also $A^{\mathrm{T}}b$ has entries $b_1+\cdots+b_m$ and $T_1b_1+\cdots+T_mb_m$. The solution with that diagonal $A^{\mathrm{T}}A$ is just the given $\widehat{\boldsymbol{x}}=(C,D)$.

Problem Set 4.4, page 242

1 (a) *Independent* (b) *Independent* and *orthogonal* (c) *Independent* and *orthonormal*. For orthonormal vectors, (a) becomes $(1,0)$, $(0,1)$ and (b) is $(.6,.8)$, $(.8,-.6)$.

2 Divide by length 3 to get $\boldsymbol{q}_1=(\frac{2}{3},\frac{2}{3},-\frac{1}{3})$. $\boldsymbol{q}_2=(-\frac{1}{3},\frac{2}{3},\frac{2}{3})$. $Q^{\mathrm{T}}Q=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ but $QQ^{\mathrm{T}}=\begin{bmatrix}5/9&2/9&-4/9\\2/9&8/9&2/9\\-4/9&2/9&5/9\end{bmatrix}$.

3 (a) $A^{\mathrm{T}}A$ will be $16I$ (b) $A^{\mathrm{T}}A$ will be diagonal with entries $1^2,2^2,3^2=1,4,9$.

4 (a) $Q=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$, $QQ^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\neq I$. Any Q with $n<m$ has $QQ^{\mathrm{T}}\neq I$.

(b) $(1,0)$ and $(0,0)$ are *orthogonal*, not *independent*. Nonzero orthogonal vectors *are* independent. (c) From $\boldsymbol{q}_1=(1,1,1)/\sqrt{3}$ my favorite is $\boldsymbol{q}_2=(1,-1,0)/\sqrt{2}$ and $\boldsymbol{q}_3=(1,1,-2)/\sqrt{6}$.

5 *Orthogonal* vectors are $(1,-1,0)$ and $(1,1,-1)$. *Orthonormal* after dividing by their lengths: $\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0\right)$ and $\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}}\right)$.

6 Q_1Q_2 is orthogonal because $(Q_1Q_2)^{\mathrm{T}}Q_1Q_2=Q_2^{\mathrm{T}}Q_1^{\mathrm{T}}Q_1Q_2=Q_2^{\mathrm{T}}Q_2=I$.

7 When Gram-Schmidt gives Q with orthonormal columns, $Q^{\mathrm{T}}Q\widehat{\boldsymbol{x}}=Q^{\mathrm{T}}\boldsymbol{b}$ becomes $\widehat{\boldsymbol{x}}=Q^{\mathrm{T}}\boldsymbol{b}$. No cost to solve the normal equations!

8 If $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are *orthonormal* vectors in $\mathbf{R}^5$ then $\boldsymbol{p}=(\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1+(\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_2$ is closest to $\boldsymbol{b}$. The error $\boldsymbol{e}=\boldsymbol{b}-\boldsymbol{p}$ is orthogonal to $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$.

9 (a) $Q=\begin{bmatrix}.8&-.6\\.6&.8\\0&0\end{bmatrix}$ has $P=QQ^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$ = projection on the xy plane.

(b) $(QQ^{\mathrm{T}})(QQ^{\mathrm{T}}) = Q(Q^{\mathrm{T}}Q)Q^{\mathrm{T}} = QQ^{\mathrm{T}}$.

10 (a) If $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ are *orthonormal* then the dot product of $\boldsymbol{q}_1$ with $c_1\boldsymbol{q}_1 + c_2\boldsymbol{q}_2 + c_3\boldsymbol{q}_3 = \mathbf{0}$ gives $c_1 = 0$. Similarly $c_2 = c_3 = 0$. This proves : *Independent* q's

(b) $Q\boldsymbol{x} = \mathbf{0}$ leads to $Q^{\mathrm{T}}Q\boldsymbol{x} = \mathbf{0}$ which says $\boldsymbol{x} = \mathbf{0}$.

11 (a) Two *orthonormal* vectors are $\boldsymbol{q}_1 = \frac{1}{10}(1, 3, 4, 5, 7)$ and $\boldsymbol{q}_2 = \frac{1}{10}(-7, 3, 4, -5, 1)$

(b) Closest projection in the plane = *projection* $QQ^{\mathrm{T}}(1, 0, 0, 0, 0) = (0.5, -0.18, -0.24, 0.4, 0)$.

12 (a) Orthonormal $\boldsymbol{a}$'s: $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{a}_1^{\mathrm{T}}(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3) = x_1(\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1) = x_1$

(b) Orthogonal $\boldsymbol{a}$'s: $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{a}_1^{\mathrm{T}}(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3) = x_1(\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1)$. Therefore $x_1 = \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1$

(c) x_1 is the first component of A^{-1} times $\boldsymbol{b}$ (A is 3 by 3 and invertible).

13 The multiple to subtract is $\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$. Then $\boldsymbol{B} = \boldsymbol{b} - \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$.

14 $\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 \end{bmatrix} \begin{bmatrix} \|\boldsymbol{a}\| & \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b} \\ 0 & \|\boldsymbol{B}\| \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 0 & 2\sqrt{2} \end{bmatrix} = QR.$

15 (a) Gram-Schmidt chooses $\boldsymbol{q}_1 = \boldsymbol{a}/||\boldsymbol{a}|| = \frac{\mathbf{1}}{\mathbf{3}}(1, 2, -2)$ and $\boldsymbol{q}_2 = \frac{1}{3}(2, 1, 2)$. Then $\boldsymbol{q}_3 = \frac{1}{3}(2, -2, -1)$.

(b) The nullspace of A^{T} contains $\boldsymbol{q}_3$

(c) $\widehat{\boldsymbol{x}} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(1, 2, 7) = (1, 2)$.

16 $\boldsymbol{p} = (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})\boldsymbol{a} = 14\boldsymbol{a}/49 = 2\boldsymbol{a}/7$ is the projection of $\boldsymbol{b}$ onto $\boldsymbol{a}$. $\boldsymbol{q}_1 = \boldsymbol{a}/\|\boldsymbol{a}\| = \boldsymbol{a}/7$ is $(4, 5, 2, 2)/7$. $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{p} = (-1, 4, -4, -4)/7$ has $\|\boldsymbol{B}\| = 1$ so $\boldsymbol{q}_2 = \boldsymbol{B}$.

17 $\boldsymbol{p} = (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})\boldsymbol{a} = (3, 3, 3)$ and $\boldsymbol{e} = (-2, 0, 2)$. Then Gram-Schmidt will choose $\boldsymbol{q}_1 = (1, 1, 1)/\sqrt{3}$ and $\boldsymbol{q}_2 = (-1, 0, 1)/\sqrt{2}$.

18 $\boldsymbol{A} = \boldsymbol{a} = (1, -1, 0, 0)$; $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{p} = (\frac{1}{2}, \frac{1}{2}, -1, 0)$; $\boldsymbol{C} = \boldsymbol{c} - \boldsymbol{p}_A - \boldsymbol{p}_B = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1)$. Notice the pattern in those orthogonal $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$. In $\mathbf{R}^5$, $\boldsymbol{D}$ would be $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, -1)$.

Gram-Schmidt would go on to normalize $\boldsymbol{q}_1 = \boldsymbol{A}/||\boldsymbol{A}||, \boldsymbol{q}_2 = \boldsymbol{B}/||\boldsymbol{B}||, \boldsymbol{q}_3 = \boldsymbol{C}/||\boldsymbol{C}||$.

19 If $A = QR$ then $A^{\mathrm{T}}A = R^{\mathrm{T}}Q^{\mathrm{T}}QR = R^{\mathrm{T}}R =$ *lower* triangular times *upper* triangular (this Cholesky factorization of $A^{\mathrm{T}}A$ uses the same R as Gram-Schmidt!). The example has $A = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 2 & 4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix} = QR$ and the same R appears in $A^{\mathrm{T}}A = \begin{bmatrix} 9 & 9 \\ 9 & 18 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix} = R^{\mathrm{T}}R.$

20 (a) *True* because $Q^{\mathrm{T}}Q = I$ leads to $(Q^{-1})(Q^{-1}) = I$.

(b) *True*. $Q\boldsymbol{x} = x_1\boldsymbol{q}_1 + x_2\boldsymbol{q}_2$. $||Q\boldsymbol{x}||^2 = x_1^2 + x_2^2$ because $\boldsymbol{q}_1 \cdot \boldsymbol{q}_2 = 0$. Also $||Q\boldsymbol{x}||^2 = \boldsymbol{x}^{\mathrm{T}}Q^{\mathrm{T}}Q\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

21 The orthonormal vectors are $\boldsymbol{q}_1 = (1,1,1,1)/2$ and $\boldsymbol{q}_2 = (-5,-1,1,5)/\sqrt{52}$. Then $\boldsymbol{b} = (-4,-3,3,0)$ projects to $\boldsymbol{p} = (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1 + (\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_2 = (-7,-3,-1,3)/2$. And $\boldsymbol{b} - \boldsymbol{p} = (-1,-3,7,-3)/2$ is orthogonal to both $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$.

22 $A = (1,1,2),\ B = (1,-1,0),\ C = (-1,-1,1)$. These are not yet unit vectors. As in Problem 18, Gram-Schmidt will divide by $||\boldsymbol{A}||$ and $||\boldsymbol{B}||$ and $||\boldsymbol{C}||$.

23 You can see why $\boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix} = QR$. This Q is just a permutation matrix—certainly orthogonal.

24 (a) One basis for the subspace $\boldsymbol{S}$ of solutions to $x_1 + x_2 + x_3 - x_4 = 0$ is the 3 special solutions $\boldsymbol{v}_1 = (-1,1,0,0)$, $\boldsymbol{v}_2 = (-1,0,1,0)$, $\boldsymbol{v}_3 = (1,0,0,1)$

(b) Since $\boldsymbol{S}$ contains solutions to $(1,1,1,-1)^{\mathrm{T}}\boldsymbol{x} = 0$, a basis for $\boldsymbol{S}^{\perp}$ is $(1,1,1,-1)$

(c) Split $(1,1,1,1)$ into $\boldsymbol{b}_1 + \boldsymbol{b}_2$ by projection on $\boldsymbol{S}^{\perp}$ and $\boldsymbol{S}$: $\boldsymbol{b}_2 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$ and $\boldsymbol{b}_1 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2})$.

25 This question shows 2 by 2 formulas for QR; breakdown $R_{22} = 0$ for singular A. Nonsingular example $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \cdot \frac{1}{\sqrt{5}}\begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}$.

Singular example $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}}\begin{bmatrix} 2 & 2 \\ 0 & \mathbf{0} \end{bmatrix}$.
The Gram-Schmidt process breaks down when $ad - bc = 0$.

26 $(\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{C}^*)\boldsymbol{q}_2 = \dfrac{\boldsymbol{B}^{\mathrm{T}}\boldsymbol{c}}{\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}}\boldsymbol{B}$ because $\boldsymbol{q}_2 = \dfrac{\boldsymbol{B}}{\|\boldsymbol{B}\|}$ and the extra $\boldsymbol{q}_1$ in $\boldsymbol{C}^*$ is orthogonal to $\boldsymbol{q}_2$.

27 When $\boldsymbol{a}$ and $\boldsymbol{b}$ are not orthogonal, the projections onto these lines *do not add* to the projection onto the plane of $\boldsymbol{a}$ and $\boldsymbol{b}$. We must use the orthogonal $\boldsymbol{A}$ and $\boldsymbol{B}$ (or orthonormal $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$) to be allowed to add projections on those lines.

28 There are $\frac{1}{2}m^2n$ multiplications to find the numbers r_{kj} and the same for v_{ij}.

29 $\boldsymbol{q}_1 = \frac{1}{3}(2, 2, -1)$, $\boldsymbol{q}_2 = \frac{1}{3}(2, -1, 2)$, $\boldsymbol{q}_3 = \frac{1}{3}(1, -2, -2)$.

30 The columns of the wavelet matrix W are *orthonormal*. Then $W^{-1} = W^{\mathrm{T}}$. This is a useful orthonormal basis with many zeros.

31 (a) $c = \frac{1}{2}$ normalizes all the orthogonal columns to have unit length (b) The projection $(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})\boldsymbol{a}$ of $\boldsymbol{b} = (1, 1, 1, 1)$ onto the first column is $\boldsymbol{p}_1 = \frac{1}{2}(-1, 1, 1, 1)$. (Check $\boldsymbol{e} = \mathbf{0}$.) To project onto the plane, add $\boldsymbol{p}_2 = \frac{1}{2}(1, -1, 1, 1)$ to get $(0, 0, 1, 1)$.

32 $Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ reflects across x axis, $Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ across plane $y + z = 0$.

33 Orthogonal and lower triangular $\Rightarrow \pm 1$ on the main diagonal and zeros elsewhere.

34 (a) $Q\boldsymbol{u} = (I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{u} = \boldsymbol{u} - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}$. This is $-\boldsymbol{u}$, provided that $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}$ equals 1 (b) $Q\boldsymbol{v} = (I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{v} = \boldsymbol{u} - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u}$, provided that $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$.

35 Starting from $\boldsymbol{A} = (1, -1, 0, 0)$, the orthogonal (not orthonormal) vectors $\boldsymbol{B} = (1, 1, -2, 0)$ and $\boldsymbol{C} = (1, 1, 1, -3)$ and $\boldsymbol{D} = (1, 1, 1, 1)$ are in the directions of $\boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4$. The 4 by 4 and 5 by 5 matrices with *integer orthogonal columns* (not orthogonal rows, since not orthonormal Q!) are

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & -3 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 & 1 \\ 0 & 0 & -3 & 1 & 1 \\ 0 & 0 & 0 & -4 & 1 \end{bmatrix}$$

36 $[Q, R] = \boldsymbol{qr}(A)$ produces from A (m by n of rank n) a "*full-size*" square $Q = [\, Q_1 \;\; Q_2 \,]$ and $\begin{bmatrix} R \\ 0 \end{bmatrix}$. The columns of Q_1 are the orthonormal basis from Gram-Schmidt of the *column space* of A. The $m - n$ columns of Q_2 are an orthonormal basis for the *left nullspace* of A. Together the columns of $Q = [\, Q_1 \;\; Q_2 \,]$ are an orthonormal basis for $\mathbf{R}^m$.

37 This question describes the next $\boldsymbol{q}_{n+1}$ in Gram-Schmidt using the matrix Q with the columns $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ (instead of using those $\boldsymbol{q}$'s separately). Start from $\boldsymbol{a}$, *subtract its projection* $\boldsymbol{p} = Q^{\mathrm{T}}\boldsymbol{a}$ onto the earlier $\boldsymbol{q}$'s, *divide by the length* of $\boldsymbol{e} = \boldsymbol{a} - Q^{\mathrm{T}}\boldsymbol{a}$ to get $\boldsymbol{q}_{n+1} = \boldsymbol{e}/\|\boldsymbol{e}\|$.

Problem Set 5.1, page 254

1 $\det(2A) = 2^4 \det A = 8$; $\det(-A) = (-1)^4 \det A = \frac{1}{2}$; $\det(A^2) = \frac{1}{4}$; $\det(A^{-1}) = 2$.

2 $\det(\frac{1}{2}A) = (\frac{1}{2})^3 \det A = -\frac{1}{8}$ and $\det(-A) = (-1)^3 \det A = 1$; $\det(A^2) = 1$; $\det(A^{-1}) = -1$.

3 (a) *False*: $\det(I + I)$ is not $1 + 1$ (except when $n = 1$) (b) *True*: The product rule extends to ABC (use it twice) (c) *False*: $\det(4A)$ is $4^n \det A$
(d) *False*: $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $AB - BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is invertible.

4 Exchange rows 1 and 3 to show $|J_3| = -1$. Exchange rows 1 and 4, then rows 2 and 3 to show $|J_4| = 1$.

5 $|J_5| = 1$ by exchanging row 1 with 5 and row 2 with 4. $|J_6| = -1$, $|J_7| = -1$. Determinants $1, 1, -1, -1$ repeat in cycles of length 4 so the determinant of J_{101} is $+1$.

6 To prove Rule 6, multiply the zero row by $t = 2$. The determinant is multiplied by 2 (Rule 3) but the matrix is the same. So $2\det(A) = \det(A)$ and $\det(A) = 0$.

7 $\det(Q) = 1$ for rotation and $\det(Q) = 1 - 2\sin^2\theta - 2\cos^2\theta = -1$ for reflection.

8 $Q^{\mathrm{T}}Q = I \Rightarrow |Q^{\mathrm{T}}|\,|Q| = |Q|^2 = 1 \Rightarrow |Q| = \pm 1$; Q^n stays orthogonal so its determinant can't blow up as $n \to \infty$.

9 $\det A = 1$ from two row exchanges . $\det B = 2$ (subtract rows 1 and 2 from row 3, then columns 1 and 2 from column 3). $\det C = 0$ (equal rows) even though $C = A + B$!

10 If the entries in every row add to zero, then $(1, 1, \ldots, 1)$ is in the nullspace: singular A has $\det = 0$. (The columns add to the zero column so they are linearly dependent.) If every row adds to one, then rows of $A - I$ add to zero (not necessarily $\det A = 1$).

11 $CD = -DC \Rightarrow \det CD = (-1)^n \det DC$ and *not* just $-\det DC$. If n is even then $\det CD = \det DC$ and we can have an invertible CD.

12 $\det(A^{-1})$ divides twice by $ad - bc$ (once for each row). This gives $\det A^{-1} = \dfrac{ad - bc}{(ad - bc)^2} = \dfrac{1}{ad - bc}$.

13 Pivots $1, 1, 1$ give determinant $= \mathbf{1}$; pivots $1, -2, -3/2$ give determinant $= \mathbf{3}$.

14 $\det(A) = \mathbf{36}$ and the 4 by 4 second difference matrix has det $= \mathbf{5}$.

15 The first determinant is $\mathbf{0}$, the second is $1 - 2t^2 + t^4 = (1 - t^2)^2$.

16 A singular rank one matrix has determinant $= 0$. The skew-symmetric K also has $\det K = 0$ (see $\#\mathbf{17}$): a skew-symmetric matrix K of odd order 3.

17 Any 3 by 3 skew-symmetric K has $\det(K^{\mathrm{T}}) = \det(-K) = (-1)^3\det(K)$. This is $-\det(K)$. But always $\det(K^{\mathrm{T}}) = \det(K)$. So we must have $\det(K) = 0$ for 3 by 3.

18 $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix} = \begin{vmatrix} b-a & b^2-a^2 \\ c-a & c^2-a^2 \end{vmatrix}$ (to reach 2 by 2, eliminate a and a^2 in row 1 by column operations)—subtract a and a^2 times column 1 from columns 2 and 3. Factor out $b - a$ and $c - a$ from the 2 by 2: $(b-a)(c-a)\begin{vmatrix} 1 & b+a \\ 1 & c+a \end{vmatrix} = (b-a)(c-a)(c-b)$.

19 For triangular matrices, just multiply the diagonal entries: $\det(U) = 6, \det(U^{-1}) = \frac{1}{6}$, and $\det(U^2) = 36$. 2 by 2 matrix: $\det(U) = ad, \det(U^2) = a^2d^2$. If $ad \neq 0$ then $\det(U^{-1}) = 1/ad$.

20 $\det \begin{bmatrix} a - Lc & b - Ld \\ c - \ell a & d - \ell b \end{bmatrix}$ reduces to $(ad - bc)(1 - L\ell)$. The determinant changes if you do two row operations at once.

21 We can exchange rows using the three elimination steps in the problem, followed by multiplying row 1 by -1. So Rules 5 and 3 give Rule 2. (Since Rules 4 and 3 give 5, they also give Rule 2.)

22 $\det(A) = 3, \det(A^{-1}) = \frac{1}{3}, \det(A - \lambda I) = \lambda^2 - 4\lambda + 3$. The numbers $\lambda = 1$ and $\lambda = 3$ give $\det(A - \lambda I) = 0$. The (singular !) matrices are

$$A - I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } A - 3I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Note to instructor: You could explain that this is the reason determinants come before eigenvalues. Identify $\lambda = 1$ and $\lambda = 3$ as the eigenvalues of A.

23 $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$ has det$(A) = 10$, $A^2 = \begin{bmatrix} 18 & 7 \\ 14 & 11 \end{bmatrix}$, det$(A^2) = 100$, $A^{-1} = \frac{1}{10}\begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}$ has det $\frac{1}{10}$. det$(A - \lambda I) = \lambda^2 - 7\lambda + 10 = 0$ when $\lambda = \mathbf{2}$ or $\mathbf{5}$; those are eigenvalues.

24 Here $A = LU$ with $\det(L) = 1$ and $\det(U) = -6 =$ product of pivots, so also $\det(A) = -6$. $\det(U^{-1}L^{-1}) = -\frac{1}{6} = 1/\det(A)$ and $\det(U^{-1}L^{-1}A)$ is $\det I = 1$.

25 When the i, j entry is ij, row $2 = 2$ times row 1 so $\det A = 0$.

26 When the ij entry is $i + j$, row $3 -$ row $2 =$ row$2 -$ row 1 so A is singular: $\det A = 0$.

27 $\det A = abc$, $\det B = -abcd$, $\det C = a(b-a)(c-b)$ by doing elimination.

28 (a) *True*: $\det(AB) = \det(A)\det(B) = 0$ (b) *False*: A row exchange gives $-\det =$ product of pivots. (c) *False*: $A = 2I$ and $B = I$ have $A - B = I$ but the determinants have $2^n - 1 \neq 1$ (d) *True*: $\det(AB) = \det(A)\det(B) = \det(BA)$.

29 A is rectangular so $\det(A^{\mathrm{T}}A) \neq (\det A^{\mathrm{T}})(\det A)$: these determinants are not defined. In fact if A is tall and thin ($m > n$), then $\det(A^{\mathrm{T}}A)$ adds up $|\det B|^2$ where the B's are all the n by n submatrices of A.

30 Derivatives of $f = \ln(ad - bc)$:

$$\begin{bmatrix} \partial f/\partial a & \partial f/\partial c \\ \partial f/\partial b & \partial f/\partial d \end{bmatrix} = \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = A^{-1}.$$

31 The Hilbert determinants are 1, 8×10^{-2}, 4.6×10^{-4}, 1.6×10^{-7}, 3.7×10^{-12}, 5.4×10^{-18}, 4.8×10^{-25}, 2.7×10^{-33}, 9.7×10^{-43}, 2.2×10^{-53}. Pivots are ratios of determinants so the 10th pivot is near 10^{-10}. The Hilbert matrix is numerically difficult (*ill-conditioned*). Please see the Figure 7.1 and Section 8.3.

32 Typical determinants of **rand**(n) are $10^6, 10^{25}, 10^{79}, 10^{218}$ for $n = 50, 100, 200, 400$. **randn**(n) with normal distribution gives $10^{31}, 10^{78}, 10^{186}$, **Inf** which means $\geq 2^{1024}$. MATLAB allows $1.999999999999999 \times 2^{1023} \approx 1.8 \times 10^{308}$ but one more 9 gives **Inf**!

33 I now know that maximizing the determinant for $1, -1$ matrices is **Hadamard's problem** (1893): see Brenner in American Math. Monthly volume 79 (1972) 626-630. Neil Sloane's wonderful On-Line Encyclopedia of Integer Sequences (**research.att.com/∼ njas**) includes the solution for small n (and more references) when the problem is changed to $0, 1$ matrices. That sequence A003432 starts from $n = 0$ with 1, 1, 1, 2, 3, 5, 9. Then the $1, -1$ maximum for size n is 2^{n-1} times the $0, 1$ maximum for size $n-1$ (so $(32)(5) = \mathbf{160}$ for $n = 6$ in sequence **A003433**).

To reduce the $1, -1$ problem from 6 by 6 to the $0, 1$ problem for 5 by 5, multiply the six rows by ± 1 to put $+1$ in column 1. Then subtract row 1 from rows 2 to 6 to get a 5 by 5 submatrix S with entries -2 and 0. Then divide S by -2.

Here is an advanced MATLAB code that finds a $1, -1$ matrix with largest $\det A = 48$ for $n = 5$:

```
n = 5; p = (n - 1)^2; A0 = ones(n); maxdet = 0;
for k = 0 : 2^p - 1
Asub = rem(floor(k.*2.^(-p+1:0)),2); A = A0; A(2:n,2:n) = 1 - 2*
reshape(Asub, n-1, n-1);
if abs(det(A)) > maxdet, maxdet = abs(det(A)); maxA = A;
end
end
```

Output: maxA = $\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 \end{matrix}$ maxdet = 48.

34 Reduce B by row operations to [row 3; row 2; row 1]. Then $\det B = -6$ (odd permutation from the order of the rows in A).

Problem Set 5.2, page 266

1 $\det A = 1 + 18 + 12 - 9 - 4 - 6 = 12$, the rows of A are independent; $\det B = 0$, row 1 + row 2 = row 3 so the rows of B are linearly dependent; $\det C = -1$, so C has independent rows ($\det C$ has one term, an odd permutation).

2 $\det A = -2$, independent; $\det B = 0$, dependent; $\det C = -1$, independent but $\det D = 0$ because its submatrix B has dependent rows.

3 The problem suggests 3 ways to see that $\det A = 0$: All cofactors of row 1 are zero. A has rank ≤ 2. Each of the 6 terms in $\det A$ is zero. Notice also that column 2 has no pivot.

4 $a_{11}a_{23}a_{32}a_{44}$ gives -1, because the terms $a_{23}a_{32}$ have columns 2 and 3 in reverse order. $a_{14}a_{23}a_{32}a_{41}$ which has 2 row exchanges gives $+1$, $\det A = 1 - 1 = 0$. Using the same entries but now taken from B, $\det B = 2{\cdot}4{\cdot}4{\cdot}2 - 1{\cdot}4{\cdot}4{\cdot}1 = 64 - 16 = 48$.

5 Four zeros in the same row guarantee $\det = 0$ (and also four zeros in the same column). $A = I$ has 12 zeros (this is the maximum with $\det \neq 0$).

6 (a) If $a_{11} = a_{22} = a_{33} = 0$ then 4 terms will be zeros (b) 15 terms must be zero. Effectively we are counting the permutations that make everyone move; $2, 3, 1$ and $3, 1, 2$ for $n = 3$ mean that the other 4 permutations take a term from the diagonal of A; so those terms are 0 when the diagonal is all zeros.

7 $5!/2 = 60$ permutation matrices (half of $5! = 120$ permutations) have $\det = +1$. Move row 5 of I to the top; then starting from $(5, 1, 2, 3, 4)$ elimination will do four row exchanges on P.

8 If $\det A \neq 0$, then certainly some term $a_{1\alpha}a_{2\beta}\cdots a_{n\omega}$ in the big formula is not zero! Move rows 1, 2, . . ., n into rows α, β, . . ., ω. Then all these nonzero a's will be on the main diagonal.

9 The big formula has six terms all ± 1: say q are -1 and $6-q$ are 1. Then $\det A = -q+6-q =$ even (so $\det A = 5$ is impossible). Also $\det A = 6$ is impossible. All 3 even permutations like $a_{11}a_{22}a_{33}$ would have to give $+1$ (so an even number of -1's in the matrix). But all 3 odd permutations like $a_{11}a_{23}a_{32}$ would have to give -1 (so an odd number of -1's in the matrix). We can't have it both ways, so $\det A = 4$ is best possible and not hard to arrange: put -1's on the main diagonal.

10 The $4!/2 = 12$ even permutations are $(1,2,3,4), (2,1,4,3), (3,1,4,2), (4,3,2,1)$, and 8 P's with one number in place and even permutation of the other three numbers: examples are $1,3,4,2$ and $1,4,2,3$.

$\det(I + P_{\text{even}})$ is always 16 or 4 or 0 (16 comes from $I + I$).

11 $C = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. $D = \begin{bmatrix} 0 & 42 & -35 \\ 0 & -21 & 14 \\ -3 & 6 & -3 \end{bmatrix}$. $\det B = 1(0)+2(42)+3(-35) = -21$. Puzzle: $\det D = 441 = (-21)^2$. *Why* is $\det(\text{cofactor matrix}) = (\det\ \text{matrix})^{n-1}$?

12 $C = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ and $AC^{\mathrm{T}} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. Therefore $A^{-1} = \frac{1}{4}C^{\mathrm{T}} = C^{\mathrm{T}}/\det A$.

13 (a) $C_1 = 0,\ C_2 = -1,\ C_3 = 0,\ C_4 = 1$ (b) $C_n = -C_{n-2}$ by cofactors of row 1 then cofactors of column 1. Therefore $C_{10} = -C_8 = C_6 = -C_4 = C_2 = -1$.

14 For the matrices in Problem 13 to produce nonzeros in the big formula, we must choose 1's from column 2 then column 1, column 4 then column 3,and so on. Therefore n must be even to have $\det \neq 0$. The number of row exchanges is $n/2$ so the overall determinant is $C_n = (-1)^{n/2}$.

15 The $1,1$ cofactor of the n by n matrix is E_{n-1}. The $1,2$ cofactor has a single 1 in its first column, with cofactor E_{n-2}: sign gives $-E_{n-2}$. So $E_n = E_{n-1} - E_{n-2}$. Then E_1 to E_6 is $1,\ 0,\ -1,\ -1,\ 0,\ 1$ and this cycle of six will repeat: $E_{100} = E_4 = -1$.

16 The $1,1$ cofactor of the n by n matrix is F_{n-1}. The $1,2$ cofactor has a 1 in column 1, with cofactor F_{n-2}. Multiply by $(-1)^{1+2}$ and also (-1) from the $1,2$ entry to find $\boldsymbol{F_n = F_{n-1} + F_{n-2}}$. So these determinants are Fibonacci numbers.

17 Use cofactors along row 4 instead of row 1 (last row instead of first).
$$|B_4| = 2\det\begin{bmatrix} 1 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix} + \det\begin{bmatrix} 1 & -1 & \\ -1 & 2 & \\ & -1 & -1 \end{bmatrix} = 2|B_3| - \det\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$
So $|B_4| = 2|B_3| - |B_2|$.

18 Rule 3 (linearity in row 1) gives $|B_n| = |A_n| - |A_{n-1}| = (n+1) - n = 1$.

19 Since x, x^2, x^3 are all in the same row, they never multiply each other in $\det V_4$. The determinant is zero at $x = a$ or b or c because of equal rows! So $\det V$ has factors $(x-a)(x-b)(x-c)$. Multiply by the cofactor V_3. The Vandermonde matrix $V_{ij} = (x_i)^{j-1}$ is for fitting a polynomial $p(\boldsymbol{x}) = \boldsymbol{b}$ at the points x_i. It has $\det V =$ product of all $x_k - x_m$ for $k > m$.

20 $G_2 = -1$, $G_3 = 2$, $G_4 = -3$, and $G_n = (-1)^{n-1}(n-1)$. One way to reach that G_n is to multiply the n eigenvalues $-1, -1, \ldots, -1, n-1$ of the matrix. Is there a good choice of row operations to produce this determinant G_n ?

21 $S_1 = 3, S_2 = 8, S_3 = 21$. The rule looks like every second number in Fibonacci's sequence $\ldots 3, 5, 8, 13, 21, 34, 55, \ldots$ so the guess is $S_4 = 55$. Following the solution to Problem 30 with 3's instead of 2's on the diagonal confirms $S_4 = 81+1-9-9-9 = 55$. Problem 32 directly proves $S_n = F_{2n+2}$.

22 Changing 3 to 2 in the corner reduces the determinant F_{2n+2} by 1 times the cofactor of that corner entry. This cofactor is the determinant of S_{n-1} (one size smaller) which is F_{2n}. Therefore changing 3 to 2 changes the determinant to $F_{2n+2} - F_{2n}$ which is Fibonacci's F_{2n+1}.

23 (a) If we choose an entry from B we must choose an entry from the zero block; result zero. This leaves entries from A times entries from D leading to $(\det A)(\det D)$
(b) and (c) Take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. See $\#25$.

24 (a) All the lower triangular blocks L_k have 1's on the diagonal and $\det = 1$. Then use $A_k = L_k U_k$ to find $\det U_k = \det A_k = 2, 6, -6$ for $k = 1, 2, 3$

(b) Equation (3) in this section gives the kth pivot as $\det A_k/\det A_{k-1}$. So $\det A_k = 5, 6, 7$ gives pivot $d_k = 5/1, 6/5, 7/6$.

25 Problem 23 gives $\det \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} = 1$ and $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = |A|$ times $|D - CA^{-1}B|$. By the product rule this is $|AD - ACA^{-1}B|$. **If $AC = CA$** this is $|AD - CAA^{-1}B| = \mathbf{det}(\boldsymbol{AD} - \boldsymbol{CB})$.

26 If A is a row and B is a column then $\det M = \det AB =$ dot product of A and B. If A is a column and B is a row then AB has rank 1 and $\det M = \det AB = 0$ (unless $m = n = 1$). This block matrix M is invertible when AB is invertible which certainly requires $m \le n$.

27 (a) $\det A = a_{11}C_{11} + \cdots + a_{1n}C_{1n}$. Derivative with respect to $a_{11} =$ cofactor C_{11}.

28 Row $1 - 2$ row $2 +$ row $3 = 0$ so this matrix is singular and $\det A$ is zero.

29 There are five nonzero products, all 1's with a plus or minus sign. Here are the (row, column) numbers and the signs: $+\,(1,1)(2,2)(3,3)(4,4)\ +\ (1,2)(2,1)(3,4)(4,3)\ -\ (1,2)(2,1)(3,3)(4,4)\ -\ (1,1)(2,2)(3,4)(4,3)\ -\ (1,1)(2,3)(3,2)(4,4)$. Total -1.

30 The 5 products in solution 29 change to $16 + 1 - 4 - 4 - 4$ since A has 2's and -1's:

$$(2)(2)(2)(2) + (-1)(-1)(-1)(-1) - (-1)(-1)(2)(2) - (2)(2)(-1)(-1) - (2)(-1)(-1)(2) = \mathbf{5} = \boldsymbol{n} + \mathbf{1}.$$

31 $\det P = -1$ because the cofactor of $P_{14} = 1$ in row one has sign $(-1)^{1+4}$. The big formula for $\det P$ has only one term $(1 \cdot 1 \cdot 1 \cdot 1)$ with minus sign because three exchanges take $4, 1, 2, 3$ into $1, 2, 3, 4$; $\det(P^2) = (\det P)(\det P) = +1$ so

$$\det \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ is } \textit{not right}.$$

32 The problem is to show that $F_{2n+2} = 3F_{2n} - F_{2n-2}$. Keep using Fibonacci's rule:

$$F_{2n+2} = F_{2n+1} + F_{2n} = F_{2n} + F_{2n-1} + F_{2n} = 2F_{2n} + (F_{2n} - F_{2n-2}) = 3F_{2n} - F_{2n-2}.$$

33 The difference from 20 to 19 multiplies its 3 by 3 cofactor $= 1$: then $\det$ drops by 1.

34 (a) The last three rows must be dependent because only 2 columns are nonzero

(b) In each of the 120 terms: Choices from the last 3 rows must use 3 different columns; at least one of those choices will be zero.

35 Subtracting 1 from the n, n entry subtracts its cofactor C_{nn} from the determinant. That cofactor is $C_{nn} = 1$ (smaller Pascal matrix). Subtracting 1 from 1 leaves 0.

Problem Set 5.3, page 283

1 (a) $|A| = \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} = 3, \quad |B_1| = \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} = 6, \quad |B_2| = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$ so $x_1 = -6/3 = -2$ and $x_2 = 3/3 = 1$ (b) $|A| = 4, |B_1| = 3, |B_2| = 2, |B_3| = 1$. Therefore $x_1 = 3/4$ and $x_2 = -1/2$ and $x_3 = 1/4$.

2 (a) $y = \begin{vmatrix} a & 1 \\ c & 0 \end{vmatrix} / \begin{vmatrix} a & b \\ c & d \end{vmatrix} = c/(ad - bc)$ (b) $y = \det B_2 / \det A = (fg - id)/D$. That is because B_2 with $(1, 0, 0)$ in column 2 has $\det B_2 = fg - id$.

3 (a) $x_1 = 3/0$ and $x_2 = -2/0$: *no solution* (b) $x_1 = x_2 = \mathbf{0}/\mathbf{0}$: *undetermined*.

4 (a) $x_1 = \det\big([\, \boldsymbol{b} \;\; \boldsymbol{a}_2 \;\; \boldsymbol{a}_3 \,]\big) / \det A$, if $\det A \neq 0$. This is $|B_1|/|A|$.

(b) The determinant is linear in its first column so $|x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3\boldsymbol{a}_2\boldsymbol{a}_3|$ splits into $x_1|\boldsymbol{a}_1 \; \boldsymbol{a}_2 \; \boldsymbol{a}_3| + x_2|\boldsymbol{a}_2 \; \boldsymbol{a}_2 \; \boldsymbol{a}_3| + x_3|\boldsymbol{a}_3 \; \boldsymbol{a}_2 \; \boldsymbol{a}_3|$. The last two determinants are zero because of repeated columns, leaving $x_1|\boldsymbol{a}_1 \; \boldsymbol{a}_2 \; \boldsymbol{a}_3|$ which is $x_1 \det A$.

5 If the first column in A is also the right side b then $\det A = \det B_1$. Both B_2 and B_3 are singular since a column is repeated. Therefore $x_1 = |B_1|/|A| = 1$ and $x_2 = x_3 = 0$.

6 (a) $\begin{bmatrix} 1 & -\frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & -\frac{7}{3} & 1 \end{bmatrix}$ (b) $\frac{1}{4}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. An invertible symmetric matrix has a symmetric inverse.

7 If all cofactors $= 0$ then A^{-1} would be the zero matrix if it existed; cannot exist. (And also, the cofactor formula gives $\det A = 0$.) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has no zero cofactors but it is not invertible.

8 $C = \begin{bmatrix} 6 & -3 & 0 \\ 3 & 1 & -1 \\ -6 & 2 & 1 \end{bmatrix}$ and $AC^{\mathrm{T}} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. This is $(\det A)I$ and $\det A = 3$. The $1, 3$ cofactor of A is 0. Then $C_{31} = 4$ or 100: no change.

9 If we know the cofactors and $\det A = 1$, then $C^{\mathrm{T}} = A^{-1}$ and also $\det A^{-1} = 1$. Now A is the inverse of C^{T}, so A can be found from the cofactor matrix for C.

10 Take the determinant of $AC^{\mathrm{T}} = (\det A)I$. The left side gives $\det AC^{\mathrm{T}} = (\det A)(\det C)$ while the right side gives $(\det A)^n$. Divide by $\det A$ to reach $\det C = (\det A)^{n-1}$.

11 The cofactors of A are integers. Division by $\det A = \pm 1$ gives integer entries in A^{-1}.

12 Both $\det A$ and $\det A^{-1}$ are integers since the matrices contain only integers. But $\det A^{-1} = 1/\det A$ so $\det A$ must be 1 or -1.

13 $A = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$ has cofactor matrix $C = \begin{bmatrix} -1 & 2 & 1 \\ 3 & -6 & 2 \\ 1 & 3 & -1 \end{bmatrix}$ and $A^{-1} = \frac{1}{5}C^{\mathrm{T}}$.

14 (a) Lower triangular L has cofactors $C_{21} = C_{31} = C_{32} = 0$ (b) $C_{12} = C_{21}$, $C_{31} = C_{13}, C_{32} = C_{23}$ make S^{-1} symmetric. (c) Orthogonal Q has cofactor matrix $C = (\det Q)(Q^{-1})^{\mathrm{T}} = \pm Q$ *also orthogonal*. Note $\det Q = 1$ or -1.

15 For $n = 5$, C contains *25* cofactors and each 4 by 4 cofactor has *24* terms. Each term needs *3* multiplications: total 1800 multiplications vs.125 for Gauss-Jordan.

16 (a) Area $\begin{vmatrix} \mathbf{3} & \mathbf{2} \\ \mathbf{1} & \mathbf{4} \end{vmatrix} = 10$ (b) and (c) Area $10/2 = 5$, these triangles are half of the parallelogram in (a).

17 Volume$=\begin{vmatrix} \mathbf{3} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{3} \end{vmatrix}=\mathbf{20}$. Area of faces $=$ length of cross product $= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \mathbf{3} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{1} \end{vmatrix} =$ $-2\boldsymbol{i} - 2\boldsymbol{j} + 8\boldsymbol{k}$ length$=\mathbf{6\sqrt{2}}$

18 (a) Area $\frac{1}{2}\begin{vmatrix} \mathbf{2} & \mathbf{1} & \mathbf{1} \\ \mathbf{3} & \mathbf{4} & \mathbf{1} \\ \mathbf{0} & \mathbf{5} & \mathbf{1} \end{vmatrix} = 5$ (b) $5 +$ new triangle area $\frac{1}{2}\begin{vmatrix} \mathbf{2} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{5} & \mathbf{1} \\ \mathbf{-1} & \mathbf{0} & \mathbf{1} \end{vmatrix} = 5 + 7 = 12$.

19 $\begin{vmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{2} & \mathbf{3} \end{vmatrix} = 4 = \begin{vmatrix} \mathbf{2} & \mathbf{2} \\ \mathbf{1} & \mathbf{3} \end{vmatrix}$ because the transpose has the same determinant. See $\#22$.

20 The edges of the hypercube have length $\sqrt{1+1+1+1} = 2$. The volume $\det H$ is $2^4 = \mathbf{16}$. ($H/2$ has orthonormal columns. Then $\det(H/2) = 1$ leads again to $\det H = 16$ in 4 dimensions.)

21 The maximum volume $L_1 L_2 L_3 L_4$ is reached when the edges are orthogonal in $\mathbf{R}^4$. With entries 1 and -1 all lengths are $\sqrt{4} = 2$. The maximum determinant is $2^4 = 16$, achieved in Problem 20. For a 3 by 3 matrix, $\det A = (\sqrt{3})^3$ can't be achieved by ± 1. $\rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$.

22 This question is still waiting for a solution! An 18.06 student showed me how to transform the parallelogram for A to the parallelogram for A^{T}, without changing its area. (Edges slide along themselves, so no change in baselength or height or area.)

23 $A^{\mathrm{T}}A = \begin{bmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} & 0 & 0 \\ 0 & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} & 0 \\ 0 & 0 & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{c} \end{bmatrix}$ has $\begin{array}{lcl} \det A^{\mathrm{T}}A & = & (\|a\|\|b\|\|c\|)^2 \\ \det A & = & \pm\|a\|\|b\|\|c\| \end{array}$

24 The box has height 4 and volume $= \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 4 \end{bmatrix} = 4$. $\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}$ and $(\boldsymbol{k} \cdot \boldsymbol{w}) = 4$.

25 The n-dimensional cube has 2^n corners, $n2^{n-1}$ edges and $2n$ $(n-1)$-dimensional faces. Coefficients from $(2+x)^n$ in Worked Example **2.4A**. Cube from $2I$ has volume 2^n.

26 The pyramid has volume $\frac{1}{6}$. The 4-dimensional pyramid has volume $\frac{1}{24}$ (and $\frac{1}{n!}$ in $\mathbf{R}^n$)

27 $x = r\cos\theta, y = r\sin\theta$ give $J = r$. This is the r in polar area $r\,dr\,d\theta$. The columns are orthogonal and their lengths are 1 and r.

28 $J = \begin{vmatrix} \sin\varphi\cos\theta & \rho\cos\varphi\sin\theta & -\rho\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \\ \cos\varphi & -\rho\sin\varphi & \theta \end{vmatrix} = \rho^2\sin\varphi$. This Jacobian is needed for triple integrals inside spheres. Those integrals have $\rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$.

29 From x,y to r,θ: $\begin{vmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial\theta/\partial x & \partial\theta/\partial y \end{vmatrix} = \begin{vmatrix} x/r & y/r \\ -y/r^2 & x/r^2 \end{vmatrix} = \begin{vmatrix} \cos\theta & \sin\theta \\ (-\sin\theta)/r & (\cos\theta)/r \end{vmatrix}$ $= \dfrac{1}{r} = \dfrac{1}{\text{Jacobian in } \mathbf{27}}$. The surprise was that $\frac{dr}{dx}$ and $\frac{dx}{dr}$ are both $\frac{x}{r}$.

30 The triangle with corners $(0,0),(6,0),(1,4)$ has area $(6)(4)/2 = 12$. Rotated by $\theta = 60^\circ$ the area is *unchanged*. The determinant of the rotation matrix is

$$J = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \begin{vmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{vmatrix} = 1.$$

31 Base area $||\boldsymbol{u} \times \boldsymbol{v}|| = 10$, height $||\boldsymbol{w}|| \cos\theta = 2$, volume $(10)(2) = 20$.

32 The volume of the box is $\det \begin{bmatrix} 2 & 4 & 0 \\ -1 & 3 & 0 \\ 1 & 2 & 2 \end{bmatrix} = 20$, agreeing with Problem 31.

33 $\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = u_1 \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}$. This is $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$.

34 $(\boldsymbol{w} \times \boldsymbol{u}) \cdot \boldsymbol{v} = (\boldsymbol{v} \times \boldsymbol{w}) \cdot \boldsymbol{u} = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$: *Even permutation* of $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ keeps the same determinant. Odd permutations like $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{v}$ will reverse the sign.

35 $S = (2,1,-1)$, area $\|PQ \times PS\| = \|(-2,-2,-1)\| = \sqrt{2^2 + 2^2 + 1^2} = 3$. The other four corners of the box can be $(0,0,0)$, $(0,0,2)$, $(1,2,2)$, $(1,1,0)$. The volume of the tilted box with edges along P, Q, and R is $|\det| = 1$.

36 If $(1,1,0)$, $(1,2,1)$, (x,y,z) are in a plane the volume is $\det \begin{bmatrix} x & y & z \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = x - y + z = 0$. The "box" with those edges is flattened to zero height.

37 $\det \begin{bmatrix} x & y & z \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix} = 7x - 5y + z$ will be *zero* when (x,y,z) is a combination of $(2,3,1)$ and $(1,2,3)$. The plane containing those two vectors has equation $7x - 5y + z = 0$. Volume = zero because the 3 box edges out from $(0,0,0)$ lie in a plane.

38 Doubling each row multiplies the volume by 2^n. Then $2\det A = \det(2A)$ only if $n=1$.

39 $AC^{\mathrm{T}} = (\det A)I$ gives $(\det A)(\det C) = (\det A)^n$. Then $\det A = (\det C)^{1/3}$ with $n = 4$. With $\det A^{-1} = 1/\det A$, construct A^{-1} using the cofactors. *Invert to find* A.

40 The cofactor formula adds 1 by 1 determinants (which are just entries) *times* their cofactors of size $n-1$. Jacobi discovered that this formula can be generalized. For $n = 5$, Jacobi multiplied each 2 by 2 determinant from rows 1-2 (with columns $a < b$) times a 3 by 3 determinant from rows 3-5 (using the remaining columns $c < d < e$).

The key question is $+$ or $-$ sign (as for cofactors). The product is given a $+$ sign when a, b, c, d, e is an even permutation of 1, 2, 3, 4, 5. This gives the correct determinant $+1$ for that permutation matrix. More than that, all other P that permute a, b and separately c, d, e will come out with the correct sign when the 2 by 2 determinant for columns a, b multiplies the 3 by 3 determinant for columns c, d, e.

41 The Cauchy-Binet formula gives the determinant of a square matrix AB (and AA^{T} in particular) when the factors A, B are rectangular. For (2 by 3) times (3 by 2) there are 3 products of 2 by 2 determinants from A and B (printed in boldface):

$$\begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} & c \\ \boldsymbol{d} & \boldsymbol{e} & f \end{bmatrix} \begin{bmatrix} \boldsymbol{g} & \boldsymbol{j} \\ \boldsymbol{h} & \boldsymbol{k} \\ i & \ell \end{bmatrix} \qquad \begin{bmatrix} \boldsymbol{a} & b & \boldsymbol{c} \\ \boldsymbol{d} & e & \boldsymbol{f} \end{bmatrix} \begin{bmatrix} \boldsymbol{g} & \boldsymbol{j} \\ h & k \\ \boldsymbol{i} & \boldsymbol{\ell} \end{bmatrix} \qquad \begin{bmatrix} a & \boldsymbol{b} & \boldsymbol{c} \\ d & \boldsymbol{e} & \boldsymbol{f} \end{bmatrix} \begin{bmatrix} g & j \\ \boldsymbol{h} & \boldsymbol{k} \\ \boldsymbol{i} & \boldsymbol{\ell} \end{bmatrix}$$

$$\textit{Check} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 7 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \qquad AB = \begin{bmatrix} 14 & 30 \\ 30 & 66 \end{bmatrix}$$

Cauchy-Binet: $(4-2)(4-2) + (7-3)(7-3) + (14-12)(14-12) = \mathbf{24}$

det of AB : $(14)(66) - (30)(30) = \mathbf{24}$

42 A 5 by 5 tridiagonal matrix has cofactor $C_{11} = 4$ by 4 tridiagonal matrix. Cofactor C_{12} has only one nonzero at the top of column 1. That nonzero multiplies its 3 by 3 cofactor which is tridiagonal. So $\det A = a_{11}C_{11} + a_{12}C_{12} =$ tridiagonal determinants of sizes 4 and 3. The number F_n of nonzero terms in $\det A$ follows Fibonacci's rule $F_n = F_{n-1} + F_{n-2}$.

Problem Set 6.1, page 298

1 The eigenvalues are 1 and 0.5 for A, 1 and 0.25 for A^2, 1 and 0 for A^∞. Exchanging the rows of A changes the eigenvalues to 1 and -0.5 (the trace is now $0.2 + 0.3$). Singular matrices stay singular during elimination, so $\lambda = 0$ does not change.

2 A has $\lambda_1 = -1$ and $\lambda_2 = 5$ with eigenvectors $x_1 = (-2, 1)$ and $x_2 = (1, 1)$. The matrix $A + I$ has the same eigenvectors, with eigenvalues increased by 1 to $\mathbf{0}$ and $\mathbf{6}$. That zero eigenvalue correctly indicates that $A + I$ is singular.

3 A has $\lambda_1 = 2$ and $\lambda_2 = -1$ (check trace and determinant) with $\boldsymbol{x}_1 = (1, 1)$ and $\boldsymbol{x}_2 = (2, -1)$. A^{-1} has the same eigenvectors, with eigenvalues $1/\lambda = \frac{1}{2}$ and -1.

4 $\det(A - \lambda I) = \lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2)$. Then A has $\lambda_1 = -3$ and $\lambda_2 = 2$ (check trace $= -1$ and determinant $= -6$) with $\boldsymbol{x}_1 = (3, -2)$ and $\boldsymbol{x}_2 = (1, 1)$. A^2 has the *same eigenvectors* as A, with eigenvalues $\lambda_1^2 = 9$ and $\lambda_2^2 = 4$.

5 A and B have eigenvalues 1 and 3 (their diagonal entries : triangular matrices). $A + B$ has $\lambda^2 + 8\lambda + 15 = 0$ and $\lambda_1 = 3$, $\lambda_2 = 5$. Eigenvalues of $A + B$ *are not equal* to eigenvalues of A plus eigenvalues of B.

6 A and B have $\lambda_1 = 1$ and $\lambda_2 = 1$. AB and BA have $\lambda^2 - 4\lambda + 1$ and the quadratic formula gives $\lambda = 2 \pm \sqrt{3}$. Eigenvalues of AB *are not equal* to eigenvalues of A times eigenvalues of B. Eigenvalues of AB and BA ***are equal*** (this is proved at the end of Section 6.2).

7 The eigenvalues of U (on its diagonal) are the *pivots* of A. The eigenvalues of L (on its diagonal) are all 1's. The eigenvalues of A *are not* the same as the pivots.

8 (a) Multiply $A\boldsymbol{x}$ to see $\lambda\boldsymbol{x}$ which reveals λ (b) Solve $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ to find $\boldsymbol{x}$.

9 (a) Multiply by A: $A(A\boldsymbol{x}) = A(\lambda\boldsymbol{x}) = \lambda A\boldsymbol{x}$ gives $A^2\boldsymbol{x} = \boldsymbol{\lambda}^2\boldsymbol{x}$

(b) Multiply by A^{-1}: $\boldsymbol{x} = A^{-1}A\boldsymbol{x} = A^{-1}\lambda\boldsymbol{x} = \lambda A^{-1}\boldsymbol{x}$ gives $A^{-1}\boldsymbol{x} = \frac{1}{\boldsymbol{\lambda}}\boldsymbol{x}$

(c) Add $I\boldsymbol{x} = \boldsymbol{x}$: $(A + I)\boldsymbol{x} = (\boldsymbol{\lambda} + \mathbf{1})\boldsymbol{x}$.

10 $\det(A - \lambda I) = d^2 - 1.4\lambda + 0.4$ so A has $\lambda_1 = 1$ and $\lambda_2 = 0.4$ with $\boldsymbol{x}_1 = (1, 2)$ and $\boldsymbol{x}_2 = (1, -1)$. A^∞ has $\lambda_1 = 1$ and $\lambda_2 = 0$ (same eigenvectors). A^{100} has $\lambda_1 = 1$ and $\lambda_2 = (0.4)^{100}$ which is near zero. So A^{100} is very near A^∞: same eigenvectors and close eigenvalues.

11 Columns of $A - \lambda_1 I$ are in the nullspace of $A - \lambda_2 I$ because $M = (A - \lambda_2 I)(A - \lambda_1 I)$ is the zero matrix **[**this is the *Cayley-Hamilton Theorem* in Problem 6.2.30**]**. Notice that M has *zero eigenvalues* $(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_1) = 0$ and $(\lambda_2 - \lambda_2)(\lambda_2 - \lambda_1) = 0$. So those columns solve $(A - \lambda_2 I)\,\boldsymbol{x} = \mathbf{0}$, they are eigenvectors.

12 The projection matrix P has $\lambda = 1, 0, 1$ with eigenvectors $(1, 2, 0)$, $(2, -1, 0)$, $(0, 0, 1)$. Add the first and last vectors: $(1, 2, 1)$ also has $\lambda = 1$. The whole column space of P contains eigenvectors with $\lambda = 1$! *Note* $P^2 = P$ leads to $\lambda^2 = \lambda$ so $\lambda = 0$ or 1.

13 (a) $P\boldsymbol{u} = (\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}) = \boldsymbol{u}$ so $\lambda = 1$ (b) $P\boldsymbol{v} = (\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{v} = \boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}) = \mathbf{0}$
(c) $\boldsymbol{x}_1 = (-1, 1, 0, 0)$, $\boldsymbol{x}_2 = (-3, 0, 1, 0)$, $\boldsymbol{x}_3 = (-5, 0, 0, 1)$ all have $P\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}$.

14 $\det(Q - \lambda I) = \lambda^2 - 2\lambda\cos\theta + 1 = 0$ when $\lambda = \cos\theta \pm i\sin\theta = e^{i\theta}$ and $e^{-i\theta}$. Check that $\lambda_1\lambda_2 = 1$ and $\lambda_1 + \lambda_2 = 2\cos\theta$. Two eigenvectors of this rotation matrix are $\boldsymbol{x}_1 = (1, i)$ and $\boldsymbol{x}_2 = (1, -i)$ (more generally $c\boldsymbol{x}_1$ and $d\boldsymbol{x}_2$ with $cd \neq 0$).

15 The other two eigenvalues are $\lambda = \frac{1}{2}(-1 \pm i\sqrt{3})$. The three eigenvalues are $1, 1, -1$.

16 Set $\lambda = 0$ in $\det(A - \lambda I) = (\lambda_1 - \lambda)\ldots(\lambda_n - \lambda)$ to find $\det A = (\lambda_1)(\lambda_2)\cdots(\lambda_n)$.

17 $\lambda_1 = \frac{1}{2}(a + d + \sqrt{(a-d)^2 + 4bc})$ and $\lambda_2 = \frac{1}{2}(a + d - \sqrt{})$ add to $a + d$. If A has $\lambda_1 = 3$ and $\lambda_2 = 4$ then $\det(A - \lambda I) = (\lambda - 3)(\lambda - 4) = \lambda^2 - 7\lambda + 12$.

18 These 3 matrices have $\lambda = 4$ and 5, trace 9, $\det 20$: $\begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix}, \begin{bmatrix} 3 & 2 \\ -1 & 6 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ -3 & 7 \end{bmatrix}$.

19 (a) rank $= 2$ (b) $\det(B^{\mathrm{T}}B) = 0$ (d) eigenvalues of $(B^2 + I)^{-1}$ are $1, \frac{1}{2}, \frac{1}{5}$.

20 $A = \begin{bmatrix} 0 & 1 \\ \mathbf{-28} & \mathbf{11} \end{bmatrix}$ has trace 11 and determinant 28, so $\lambda = 4$ and 7. Moving to a 3 by 3 companion matrix, for eigenvalues $1, 2, 3$ we want $\det(C - \lambda I) = (1 - \lambda)(2 - \lambda)(3 - \lambda)$. Multiply out to get $-\lambda^3 + 6\lambda^2 - 11\lambda + 6$. To get those numbers $6, -11, 6$ from a companion matrix you just put them into the last row:

$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \mathbf{6} & \mathbf{-11} & \mathbf{6} \end{bmatrix}$ Notice the trace $6 = 1 + 2 + 3$ and determinant $6 = (1)(2)(3)$.

21 $(A - \lambda I)$ has the same determinant as $(A - \lambda I)^{\mathrm{T}}$ because every square matrix has $\det M = \det M^{\mathrm{T}}$. Pick $M = A - \lambda I$.

$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ have *different eigenvectors*.

22 The eigenvalues must be $\lambda = \mathbf{1}$ (because the matrix is Markov), $\mathbf{0}$ (for singular), $-\frac{\mathbf{1}}{\mathbf{2}}$ (so sum of eigenvalues = trace $= \frac{1}{2}$).

23 $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}.$ Always A^2 is the zero matrix if $\lambda = 0$ and 0, by the Cayley-Hamilton Theorem in Problem 6.2.30.

24 $\lambda = \mathbf{0, 0, 6}$ (*notice rank* 1 *and trace* 6). Two eigenvectors of $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ are perpendicular to $\boldsymbol{v}$ and the third eigenvector is $\boldsymbol{u}$: $\boldsymbol{x}_1 = (0, -2, 1)$, $\boldsymbol{x}_2 = (1, -2, 0)$, $\boldsymbol{x}_3 = (1, 2, 1)$.

25 When A and B have the same n λ's and $\boldsymbol{x}$'s, look at any combination $\boldsymbol{v} = c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$. Multiply by A and B: $A\boldsymbol{v} = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n$ **equals** $B\boldsymbol{v} = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n$ **for all vectors** $\boldsymbol{v}$. So $A = B$.

26 The block matrix has $\lambda = 1, 2$ from B and $\lambda = 5, 7$ from D. All entries of C are multiplied by zeros in $\det(A - \lambda I)$, so C has no effect on the eigenvalues of the block matrix.

27 A has rank 1 with eigenvalues $0, 0, 0, 4$ (the 4 comes from the trace of A). C has rank 2 (ensuring two zero eigenvalues) and $(1, 1, 1, 1)$ is an eigenvector with $\lambda = 2$. With trace 4, the other eigenvalue is also $\lambda = 2$, and its eigenvector is $(1, -1, 1, -1)$.

28 Subtract from $0, 0, 0, 4$ in Problem 27. $B = A - I$ has $\lambda = -1, -1, -1, 3$ and $C = I - A$ has $\lambda = 1, 1, 1, -3$. Both have $\det = -3$.

29 A is triangular: $\lambda(A) = 1, 4, 6$; $\lambda(B) = 2, \sqrt{3}, -\sqrt{3}$; C has rank one: $\lambda(C) = 0, 0, 6$.

30 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ c+d \end{bmatrix} = (\boldsymbol{a}+\boldsymbol{b}) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $\lambda_2 = \boldsymbol{d} - \boldsymbol{b}$ to produce the correct trace $(a+b)+(d-b) = a+d$.

31 Eigenvector $(1,3,4)$ for A with $\lambda = 11$ and eigenvector $(3,1,4)$ for PAP^{T} with $\lambda = 11$. Eigenvectors with $\lambda \neq 0$ must be in the column space since $A\boldsymbol{x}$ is always in the column space, and $\boldsymbol{x} = A\boldsymbol{x}/\lambda$.

32 (a) $\boldsymbol{u}$ is a basis for the nullspace (we know $A\boldsymbol{u} = 0\boldsymbol{u}$); $\boldsymbol{v}$ and $\boldsymbol{w}$ give a basis for the column space (we know $A\boldsymbol{v}$ and $A\boldsymbol{w}$ are in the column space).

(b) $A(\boldsymbol{v}/3 + \boldsymbol{w}/5) = 3\boldsymbol{v}/3 + 5\boldsymbol{w}/5 = \boldsymbol{v} + \boldsymbol{w}$. So $\boldsymbol{x} = \boldsymbol{v}/3 + \boldsymbol{w}/5$ is a particular solution to $A\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w}$. Add any $c\boldsymbol{u}$ from the nullspace

(c) If $A\boldsymbol{x} = \boldsymbol{u}$ had a solution, $\boldsymbol{u}$ would be in the column space: wrong dimension 3.

33 Always $(\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})\,\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u})$ so $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ with $\lambda = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}$. (watch numbers $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}$, vectors $\boldsymbol{u}$, matrices $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$!!) If $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} = 0$ then $A^2 = \boldsymbol{u}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{v}^{\mathrm{T}}$ is the zero matrix and $\lambda^2 = 0, 0$ and $\lambda = 0, 0$ and trace $(A) = 0$. This zero trace also comes from adding the diagonal entries of $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$:

$$A = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 \\ u_2v_1 & u_2v_2 \end{bmatrix} \quad \text{has trace} \quad u_1v_1 + u_2v_2 = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} = 0$$

34 $\det(P - \lambda I) = 0$ gives the equation $\lambda^4 = 1$. This reflects the fact that $P^4 = I$. The solutions of $\lambda^4 = 1$ are $\lambda = 1, i, -1, -i$. The real eigenvector $\boldsymbol{x}_1 = (1,1,1,1)$ is not changed by the permutation P. Three more eigenvectors are $(1, i, i^2, i^3)$ and $(1, -1, 1, -1)$ and $(1, -i, (-i)^2, (-i)^3)$.

35 The six 3 by 3 permutation matrices include $P = I$ and three single row exchange matrices P_{12}, P_{13}, P_{23} and two double exchange matrices like $P_{12}P_{13}$. Since $P^{\mathrm{T}}P = I$ gives $(\det P)^2 = 1$, the determinant of P is 1 or -1. The pivots are always 1 (but there may be row exchanges). The trace of P can be 3 (for $P = I$) or 1 (for row exchange) or 0 (for double exchange). The possible eigenvalues are 1 and -1 and $e^{2\pi i/3}$ and $e^{-2\pi i/3}$.

36 $AB - BA = I$ can happen only for infinite matrices. If $A^{\mathrm{T}} = A$ and $B^{\mathrm{T}} = -B$ then

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\,(AB - BA)\,\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\,(A^{\mathrm{T}}B + B^{\mathrm{T}}A)\,\boldsymbol{x} \le ||A\boldsymbol{x}||\,||B\boldsymbol{x}|| + ||B\boldsymbol{x}||\,||A\boldsymbol{x}||.$$

Therefore $||A\boldsymbol{x}||\,||B\boldsymbol{x}|| \ge \frac{1}{2}||\boldsymbol{x}||^2$ and $(||A\boldsymbol{x}||/||\boldsymbol{x}||)\,(||B\boldsymbol{x}||/||\boldsymbol{x}||) \ge \frac{1}{2}$.

37 $\lambda_1 = e^{2\pi i/3}$ and $\lambda_2 = e^{-2\pi i/3}$ give $\det \lambda_1\lambda_2 = 1$ and trace $\lambda_1 + \lambda_2 = -1$. $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ with $\theta = \dfrac{2\pi}{3}$ has this trace and det. So does every $M^{-1}AM$!

38 (a) Since the columns of A add to 1, one eigenvalue is $\lambda = 1$ and the other is $c - 0.6$ (to give the correct trace $c + 0.4$).

(b) If $c = 1.6$ then both eigenvalues are 1, and all solutions to $(A - I)\,\boldsymbol{x} = \boldsymbol{0}$ are multiples of $\boldsymbol{x} = (1, -1)$. In this case A has rank 1.

(c) If $c = 0.8$, the eigenvectors for $\lambda = 1$ are multiples of (1, 3). Since all powers A^n also have column sums $= 1$, A^n will approach $\dfrac{1}{4}\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ = rank-1 matrix A^{∞} with eigenvalues $1, 0$ and correct eigenvectors. $(1, 3)$ and $(1, -1)$.

Problem Set 6.2, page 314

1 Eigenvectors in $\boldsymbol{X}$ and eigenvalues in Λ. Then $A = X\Lambda X^{-1}$ is $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

The second matrix has $\lambda = 0$ (rank 1) and $\lambda = 4$ (trace $= 4$). Then $A = X\Lambda X^{-1}$ is $\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$.

2 Put the eigenvectors in X and eigenvalues $2, 5$ in Λ. $A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$.

3 If $A = X\Lambda X^{-1}$ then the eigenvalue matrix for $A + 2I$ is $\Lambda + 2I$ and the eigenvector matrix is still X. So $A + 2I = S(\Lambda + 2I)X^{-1} = X\Lambda X^{-1} + X(2I)X^{-1} = A + 2I$.

4 (a) False: We are not given the λ's (b) True (c) True (d) False: For this we would need the eigenvectors of X

5 With $X = I$, $A = X\Lambda X^{-1} = \Lambda$ is a diagonal matrix. If X is triangular, then X^{-1} is triangular, so $X\Lambda X^{-1}$ is also triangular.

6 The columns of S are nonzero multiples of (2,1) and (0,1): either order. The same eigenvector matrices diagonalize A and A^{-1}.

7 $A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} /2 = \begin{bmatrix} \lambda_1 + \lambda_2 & \lambda_1 - \lambda_2 \\ \lambda_1 - \lambda_2 & \lambda_1 + \lambda_2 \end{bmatrix} /2.$

These are the matrices $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$, their eigenvectors are $(1,1)$ and $(1,-1)$.

8 $A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \dfrac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}.$

$X\Lambda^k X^{-1} = \dfrac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

The second component is $F_k = (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)$.

9 (a) The equations are $\begin{bmatrix} G_{k+2} \\ G_{k+1} \end{bmatrix} = A \begin{bmatrix} G_{k+1} \\ G_k \end{bmatrix}$ with $A = \begin{bmatrix} .5 & .5 \\ 1 & 0 \end{bmatrix}$. This matrix has $\lambda_1 = 1,\ \lambda_2 = -\frac{1}{2}$ with $\boldsymbol{x}_1 = (1,1),\ \boldsymbol{x}_2 = (1,-2)$

(b) $A^n = X\Lambda^n X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & (-.5)^n \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \to A^\infty = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

10 The rule $F_{k+2} = F_{k+1} + F_k$ produces the pattern: even, odd, odd, even, odd, odd, . . .

11 (a) *True* (no zero eigenvalues) (b) *False* (repeated $\lambda = 2$ may have only one line of eigenvectors) (c) *False* (repeated λ may have a full set of eigenvectors)

12 (a) False: don't know if $\lambda = 0$ or not.

(b) True: an eigenvector is missing, which can only happen for a repeated eigenvalue.

(c) True: We know there is only one line of eigenvectors.

13 $A = \begin{bmatrix} 8 & 3 \\ -3 & 2 \end{bmatrix}$ (or other), $A = \begin{bmatrix} 9 & 4 \\ -4 & 1 \end{bmatrix}$, $A = \begin{bmatrix} 10 & 5 \\ -5 & 0 \end{bmatrix}$; only eigenvectors are $\boldsymbol{x} = (c, -c)$.

14 The rank of $A - 3I$ is $r = 1$. Changing any entry except $a_{12} = 1$ makes A diagonalizable (the new A will have two different eigenvalues)

15 $A^k = X\Lambda^k X^{-1}$ approaches zero **if and only if every** $|\boldsymbol{\lambda}| < \mathbf{1}$; A_1 is a Markov matrix so $\lambda_{\text{max}} = 1$ and $A_1^k \to A_1^\infty$, A_2 has $\lambda = .6 \pm .3$ so $A_2^k \to 0$.

16 A_1 is $X\Lambda X^{-1}$ with $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & .2 \end{bmatrix}$ and $X = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; $\Lambda^k \to \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Then $A_1 k = X\Lambda^k X^{-1} \to \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$: *steady state*.

17 A_2 is $X\Lambda X^{-1}$ with $\Lambda = \begin{bmatrix} .9 & 0 \\ 0 & .3 \end{bmatrix}$ and $X = \begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$; $A_2^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = (.9)^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

$A_2^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = (.3)^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Then $A_2^{10} \begin{bmatrix} 6 \\ 0 \end{bmatrix} = (.9)^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} + (.3)^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ because $\begin{bmatrix} 6 \\ 0 \end{bmatrix}$ is the sum of $\begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

18 $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = X\Lambda X^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ and

$$A^k = X\Lambda^k X^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Multiply those last three matrices to get $A^k = \frac{1}{2} \begin{bmatrix} 1+3^k & 1-3^k \\ 1-3^k & 1+3^k \end{bmatrix}$.

19 $B^k = X\Lambda^k X^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}^k \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 5^k & 5^k - 4^k \\ 0 & 4^k \end{bmatrix}$.

20 $\det A = (\det X)(\det \Lambda)(\det X^{-1}) = \det \Lambda = \lambda_1 \cdots \lambda_n$. This proof ($\det$ = product of λ's) works when A is *diagonalizable*. The formula is always true.

21 trace $XY = (aq + bs) + (cr + dt)$ is equal to $(qa + rc) + (sb + td)$ = trace YX. Diagonalizable case: the trace of $X\Lambda X^{-1}$ = trace of $(\Lambda X^{-1})X = \Lambda$: *sum of the* λ's.

22 $AB - BA = I$ is impossible since trace AB − trace BA = *zero* $\neq$ trace I. $AB - BA = C$ is possible when trace $(C) = 0$. For example $E = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ has $EE^{\mathrm{T}} - E^{\mathrm{T}}E = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = C$ with trace zero.

23 If $A = X\Lambda X^{-1}$ then $B = \begin{bmatrix} A & 0 \\ 0 & 2A \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 2\Lambda \end{bmatrix} \begin{bmatrix} X^{-1} & 0 \\ 0 & X^{-1} \end{bmatrix}$. So B has the original λ's from A and the additional eigenvalues $2\lambda_1, \ldots, 2\lambda_n$ from $2A$.

24 The A's form a subspace since cA and $A_1 + A_2$ all have the same X. When $X = I$ the A's with those eigenvectors give the subspace of **diagonal matrices**. The dimension of that matrix space is 4 since the matrices are 4 by 4.

25 If A has columns $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ then column by column, $A^2 = A$ means every $A\boldsymbol{x}_i = \boldsymbol{x}_i$. All vectors in the column space (combinations of those columns $\boldsymbol{x}_i$) are eigenvectors with $\lambda = 1$. Always the nullspace has $\lambda = 0$ (A might have dependent columns, so there could be less than n eigenvectors with $\lambda = 1$). Dimensions of those spaces $\boldsymbol{C}(A)$ and $\boldsymbol{N}(A)$ add to n by the Fundamental Theorem, so A *is diagonalizable* (n independent eigenvectors altogether).

26 Two problems: The nullspace and column space can overlap, so $\boldsymbol{x}$ could be in both. There may not be r independent eigenvectors in the column space.

27 $R = X\sqrt{\Lambda}X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} /2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has $R^2 = A$.
$\sqrt{B}$ needs $\lambda = \sqrt{9}$ and $\sqrt{-1}$, trace (their sum) is not real so $\sqrt{B}$ cannot be real. Note that $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ has *two* imaginary eigenvalues $\sqrt{-1} = i$ and $-i$, real trace 0, real square root $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

28 The factorizations of A and B into $X\Lambda X^{-1}$ are the same. So $A = B$. (This is the same as Problem 6.1.25, expressed in matrix form.)

29 $A = X\Lambda_1 X^{-1}$ and $B = X\Lambda_2 X^{-1}$. Diagonal matrices always give $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$. Then $AB = BA$ from

$$X\Lambda_1 X^{-1} X\Lambda_2 X^{-1} = X\boldsymbol{\Lambda_1\Lambda_2} X^{-1} = X\boldsymbol{\Lambda_2\Lambda_1} X^{-1} = X\Lambda_2 X^{-1} X\Lambda_1 X^{-1} = BA.$$

30 (a) $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ has $\lambda = a$ and $\lambda = d$: $(A - aI)(A - dI) = \begin{bmatrix} 0 & b \\ 0 & d-a \end{bmatrix} \begin{bmatrix} a-d & b \\ 0 & 0 \end{bmatrix}$ $= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. (b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ has $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $A^2 - A - I = 0$ is true, matching $\lambda^2 - \lambda - 1 = 0$ as the Cayley-Hamilton Theorem predicts.

31 When $A = X\Lambda X^{-1}$ is diagonalizable, the matrix $A - \lambda_j I = X(\Lambda - \lambda_j I)X^{-1}$ will have 0 in the j, j diagonal entry of $\Lambda - \lambda_j I$. In the product $p(A) = (A - \lambda_1 I)\cdots(A - \lambda_n I)$, each inside X^{-1} cancels X. This leaves X times (*product of diagonal matrices* $\Lambda - \lambda_j I$) times X^{-1}. That product is the zero matrix because the factors produce a zero in each diagonal position. Then $p(A) =$ zero matrix, which is the Cayley-Hamilton Theorem. (If A is not diagonalizable, one proof is to take a sequence of diagonalizable matrices approaching A.)

Comment I have also seen this Caley-Hamilton proof but I am not convinced:

Apply the formula $AC^{\mathrm{T}} = (\det A)I$ from Section 5.3 to $A - \lambda I$ with variable λ. Its cofactor matrix C will be a polynomial in λ, since cofactors are determinants:

$$(A - \lambda I)\ \text{cof}\ (A - \lambda I)^{\mathrm{T}} = \det(A - \lambda I)I = p(\lambda)I.$$

"For fixed A, this is an identity between two matrix polynomials." Set $\lambda = A$ to find the zero matrix on the left, so $p(A) =$ zero matrix on the right—which is the Cayley-Hamilton Theorem.

I am not certain about the key step of substituting a matrix for λ. If other matrices B are substituted for λ, does the identity remain true? If $AB \neq BA$, even the order of multiplication seems unclear . . .

32 If $AB = BA$, then B has the same eigenvectors $(1,0)$ and $(0,1)$ as A. So B is also diagonal $b = c = 0$. The nullspace for the following equation is 2-dimensional:
$$AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -\boldsymbol{b} \\ \boldsymbol{c} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$
Those 4 equations $0 = 0, -b = 0, c = 0, 0 = 0$ have a 4 by 4 coefficient matrix with rank $4 - 2 = 2$.

33 B has $\lambda = i$ and $-i$, so B^4 has $\lambda^4 = 1$ and 1 and $B^{1024} = I$.

C has $\lambda = (1 \pm \sqrt{3}i)/2$. This λ is $\exp(\pm\pi i/3)$ so $\lambda^3 = -1$ and -1. Then $C^3 = -I$ which leads to $C^{1024} = (-I)^{341}C = -C$.

34 The eigenvalues of $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ are $\lambda = e^{i\theta}$ and $e^{-i\theta}$ (trace $2\cos\theta$ and determinant $= 1$). Their eigenvectors are $(1, -i)$ and $(1, i)$:

$$A^n = X\Lambda^n X^{-1} = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \begin{bmatrix} e^{in\theta} & \\ & e^{-in\theta} \end{bmatrix} \begin{bmatrix} i & -1 \\ i & 1 \end{bmatrix} /2i$$
$$= \begin{bmatrix} (e^{in\theta} + e^{-in\theta})/2 & \cdots \\ (e^{in\theta} - e^{-in\theta})/2i & \cdots \end{bmatrix} = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

Geometrically, n rotations by θ give one rotation by $n\theta$.

35 Columns of X times rows of ΛX^{-1} gives a sum of r rank-1 matrices ($r =$ rank of A).

36 Multiply $\text{ones}(n) * \text{ones}(n) = n * \text{ones}(n)$. This leads to $C = \boldsymbol{-1/(n+1)}$.

$$AA^{-1} = (\text{eye}(n) + \text{ones}(n)) * (\text{eye}(n) + C * \text{ones}(n))$$
$$= \text{eye}(n) + (1 + C + Cn) * \text{ones}(n) = \text{eye}(n).$$

Problem Set 6.3, page 332

1 Eigenvalues 4 and 1 with eigenvectors $(1,0)$ and $(1,-1)$ give solutions $\boldsymbol{u}_1 = e^{4t}\begin{bmatrix}1\\0\end{bmatrix}$ and $\boldsymbol{u}_2 = e^{t}\begin{bmatrix}1\\-1\end{bmatrix}$. If $\boldsymbol{u}(0) = \begin{bmatrix}5\\-2\end{bmatrix} = 3\begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}1\\-1\end{bmatrix}$, then $\boldsymbol{u}(t) = 3e^{4t}\begin{bmatrix}1\\0\end{bmatrix} + 2e^{t}\begin{bmatrix}1\\-1\end{bmatrix}$.

2 $z(t) = 2e^t$ solves $dx/dt = z$ with $z(0) = 2$. Then $dy/dt = 4y - 6e^t$ with $y(0) = 5$ gives $y(t) = 3e^{4t} + 2e^t$ as in Problem 1.

3 (a) If every column of A adds to zero, this means that the rows add to the zero row. So the rows are dependent, and A is singular, and $\lambda = 0$ is an eigenvalue.

(b) The eigenvalues of $A = \begin{bmatrix}-2 & 3\\ 2 & -3\end{bmatrix}$ are $\lambda_1 = 0$ with eigenvector $\boldsymbol{x}_1 = (3,2)$ and $\lambda_2 = -5$ (to give trace $= -5$) with $\boldsymbol{x}_2 = (1,-1)$. Then the usual 3 steps:

1. Write $u(0) = \begin{bmatrix}4\\1\end{bmatrix}$ as $\begin{bmatrix}3\\2\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix} = \boldsymbol{x}_1 + \boldsymbol{x}_2 =$ combination of eigenvectors
2. The solutions follow those eigenvectors: $e^{0t}\boldsymbol{x}_1$ and $e^{-5t}\boldsymbol{x}_2$
3. The solution $\boldsymbol{u}(t) = \boldsymbol{x}_1 + e^{-5t}\boldsymbol{x}_2$ has steady state $\boldsymbol{x}_1 = (3,2)$ since $e^{-5t} \to 0$.

4 $d(v+w)/dt = (w-v) + (v-w) = 0$, so the total $v + w$ is constant. $A = \begin{bmatrix}-1 & 1\\ 1 & -1\end{bmatrix}$ has $\begin{matrix}\lambda_1 = 0\\ \lambda_2 = -2\end{matrix}$ with $\boldsymbol{x}_1 = \begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix}1\\-1\end{bmatrix}$.

$$\begin{bmatrix}v(0)\\w(0)\end{bmatrix} = \begin{bmatrix}30\\10\end{bmatrix} = 20\begin{bmatrix}1\\1\end{bmatrix} + 10\begin{bmatrix}1\\-1\end{bmatrix} \text{ leads to } \begin{matrix}v(1) = 20 + 10e^{-2} & v(\infty) = 20\\ w(1) = 20 - 10e^{-2} & w(\infty) = 20\end{matrix}$$

5 $\dfrac{d}{dt}\begin{bmatrix}v\\w\end{bmatrix} = \begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}$ has $\lambda = 0$ and $\lambda = +2$: $v(t) = \mathbf{20 + 10e^{2t}} \to -\infty$ as $t \to \infty$.

6 $A = \begin{bmatrix}a & 1\\ 1 & a\end{bmatrix}$ has real eigenvalues $a+1$ and $a-1$. These are both negative if $\boldsymbol{a < -1}$. In this case the solutions of $\boldsymbol{u}' = A\boldsymbol{u}$ approach zero.

$B = \begin{bmatrix} b & -1 \\ 1 & b \end{bmatrix}$ has complex eigenvalues $b+i$ and $b-i$. These have negative real parts if $\boldsymbol{b} < \mathbf{0}$. In this case and all solutions of $\boldsymbol{v}' = B\boldsymbol{v}$ approach zero.

7 A projection matrix has eigenvalues $\lambda = 1$ and $\lambda = 0$. Eigenvectors $P\boldsymbol{x} = \boldsymbol{x}$ fill the subspace that P projects onto: here $\boldsymbol{x} = (1, 1)$. Eigenvectors with $P\boldsymbol{x} = \mathbf{0}$ fill the perpendicular subspace: here $\boldsymbol{x} = (1, -1)$. For the solution to $\boldsymbol{u}' = -P\boldsymbol{u}$,

$$\boldsymbol{u}(0) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \boldsymbol{u}(t) = e^{-t} \begin{bmatrix} 2 \\ 2 \end{bmatrix} + e^{0t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ approaches } \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

8 $\begin{bmatrix} 6 & -2 \\ 2 & 1 \end{bmatrix}$ has $\lambda_1 = 5,\ \boldsymbol{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix},\ \lambda_2 = 2,\ \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$; rabbits $r(t) = 20e^{5t} + 10e^{2t}$, $w(t) = 10e^{5t} + 20e^{2t}$. The ratio of rabbits to wolves approaches $20/10$; e^{5t} dominates.

9 (a) $\begin{bmatrix} 4 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ i \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -i \end{bmatrix}$. (b) Then $u(t) = 2e^{it} \begin{bmatrix} 1 \\ i \end{bmatrix} + 2e^{-it} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 4\cos t \\ 4\sin t \end{bmatrix}$.

10 $\dfrac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix}$. This correctly gives $y' = y'$ and $y'' = 4y + 5y'$. $A = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix}$ has $\det(A - \lambda I) = \lambda^2 - 5\lambda - 4 = 0$. Directly substituting $y = e^{\lambda t}$ into $y'' = 5y' + 4y$ also gives $\lambda^2 = 5\lambda + 4$ and the same two values of λ. Those values are $\frac{1}{2}(5 \pm \sqrt{41})$ by the quadratic formula.

11 The series for e^{At} is $e^{At} = I + t \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \text{ zeros } = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$.

Then $\begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix} \begin{bmatrix} y(0) + y'(0)t \\ y'(0) \end{bmatrix}$. This $y(t) = y(0) + y'(0)t$ solves the equation—the factor t tells us that A had only one eigenvector: not diagonalizable.

12 $A = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix}$ has trace 6, det 9, $\lambda = 3$ and 3 with *one* independent eigenvector $(1, 3)$. Substitute $y = te^{3t}$ to show that this gives the needed second solution ($y = e^{3t}$ is the first solution).

13 (a) $y(t) = \cos 3t$ and $\sin 3t$ solve $y'' = -9y$. It is $\mathbf{3}\cos 3t$ that starts with $y(0) = 3$ and $y'(0) = 0$. (b) $A = \begin{bmatrix} 0 & 1 \\ -9 & 0 \end{bmatrix}$ has det $= 9$: $\lambda = 3i$ and $-3i$ with eigenvectors $x = \begin{bmatrix} 1 \\ 3i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -3i \end{bmatrix}$. Then $\boldsymbol{u}(t) = \frac{3}{2}e^{3it}\begin{bmatrix} 1 \\ 3i \end{bmatrix} + \frac{3}{2}e^{-3it}\begin{bmatrix} 1 \\ -3i \end{bmatrix} = \begin{bmatrix} 3\cos 3t \\ \mathbf{-9\sin 3t} \end{bmatrix}$.

14 When A is skew-symmetric, the derivative of $||u(t)||^2$ is *zero*. Then $\|\boldsymbol{u}(t)\| = \|e^{At}\boldsymbol{u}(0)\|$ stays at $\|\boldsymbol{u}(0)\|$. So e^{At} is matrix *orthogonal*.

15 $\boldsymbol{u}_p = 4$ and $\boldsymbol{u}(t) = ce^t + 4$. For the matrix equation, the particular solution $\boldsymbol{u}_p = A^{-1}\boldsymbol{b}$ is $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and $\boldsymbol{u}(t) = c_1 e^t \begin{bmatrix} 1 \\ t \end{bmatrix} + c_2 e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

16 Substituting $\boldsymbol{u} = e^{ct}\boldsymbol{v}$ gives $ce^{ct}\boldsymbol{v} = Ae^{ct}\boldsymbol{v} - e^{ct}\boldsymbol{b}$ or $(A - cI)\boldsymbol{v} = \boldsymbol{b}$ or $\boldsymbol{v} = (A - cI)^{-1}\boldsymbol{b}$ = particular solution. If c is an eigenvalue then $A - cI$ is not invertible.

17 (a) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. These show the unstable cases (a) $\lambda_1 < 0$ and $\lambda_2 > 0$ (b) $\lambda_1 > 0$ and $\lambda_2 > 0$ (c) $\lambda = a \pm ib$ with $a > 0$

18 $d/dt(e^{At}) = A + A^2t + \frac{1}{2}A^3t^2 + \frac{1}{6}A^4t^3 + \cdots = A(I + At + \frac{1}{2}A^2t^2 + \frac{1}{6}A^3t^3 + \cdots)$. This is exactly Ae^{At}, the derivative we expect.

19 $e^{Bt} = I + Bt$ (short series with $B^2 = 0$) $= \begin{bmatrix} \mathbf{1} & \boldsymbol{-4t} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$. Derivative $= \begin{bmatrix} 0 & -4 \\ 0 & 0 \end{bmatrix} = B$.

20 The solution at time $t + T$ is $e^{A(t+T)}\boldsymbol{u}(0)$. Thus e^{At} times e^{AT} equals $e^{A(t+T)}$.

21 $\begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{0} \end{bmatrix}$ diagonalizes $A = X\Lambda X^{-1}$.

Then $e^{At} = Xe^{\Lambda t}X^{-1} = \begin{bmatrix} 1 & 4 \\ 0 & -1 \end{bmatrix}$; $\begin{bmatrix} 1 & 4 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \boldsymbol{e^t} & 0 \\ 0 & \mathbf{1} \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} e^t & 4e^t - 4 \\ 0 & 1 \end{bmatrix}$.

22 $A^2 = A$ gives $e^{At} = I + At + \frac{1}{2}\boldsymbol{At^2} + \frac{1}{6}At^3 + \cdots = I + (e^t - 1)A = \begin{bmatrix} e^t & 4e^t - 4 \\ 0 & 1 \end{bmatrix}$.

23 $e^A = \begin{bmatrix} e & 4(e-1) \\ 0 & 1 \end{bmatrix}$ from **21** and $e^B = \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}$ from **19**. By direct multiplication $e^A e^B \neq e^B e^A \neq e^{A+B} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$.

24 $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$. Then $e^{At} = \begin{bmatrix} \boldsymbol{e^t} & \frac{1}{2}(e^{3t} - e^t) \\ 0 & \boldsymbol{e^{3t}} \end{bmatrix}$.

At $t = 0, e^{At} = I$ and $\Lambda e^{At} = A$.

25 The matrix has $A^2 = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} = A$. *Then all* $A^n = A$. So $e^{At} = I + (t + t^2/2! + \cdots)A = I + (e^t - 1)A = \begin{bmatrix} e^t & 3(e^t - 1) \\ 0 & 0 \end{bmatrix}$ as in Problem 22.

26 (a) The inverse of e^{At} is e^{-At} (b) If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then $e^{At}\boldsymbol{x} = e^{\lambda t}\boldsymbol{x}$ and $e^{\lambda t} \neq 0$.
To see $e^{At}\boldsymbol{x}$, write $(I + At + \frac{1}{2}A^2t^2 + \cdots)\boldsymbol{x} = (1 + \lambda t + \frac{1}{2}\lambda^2 t^2 + \cdots)\boldsymbol{x} = e^{\lambda t}\boldsymbol{x}$.

27 $(x, y) = (e^{4t}, e^{-4t})$ is a growing solution. The correct matrix for the exchanged $\boldsymbol{u} = \begin{bmatrix} y \\ x \end{bmatrix}$ is $\begin{bmatrix} 2 & -2 \\ -4 & 0 \end{bmatrix}$. It *does* have the same eigenvalues as the original matrix.

28 Invert $\begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix}$ to produce $\boldsymbol{U}_{n+1} = \begin{bmatrix} 1 & 0 \\ -\Delta t & 1 \end{bmatrix} \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \boldsymbol{U}_n = \begin{bmatrix} 1 & \Delta t \\ -\Delta t & 1 - (\Delta t)^2 \end{bmatrix} \boldsymbol{U}_n$.

At $\Delta t = 1$, $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ has $\lambda = e^{i\pi/3}$ and $e^{-i\pi/3}$. Both eigenvalues have $\lambda^6 = 1$ so $\boldsymbol{A^6 = I}$. Therefore $\boldsymbol{U}_6 = A^6\boldsymbol{U}_0$ comes exactly back to $\boldsymbol{U}_0$.

29 First A has $\lambda = \pm i$ and $A^4 = I$.
Second A has $\lambda = -1, -1$ and $A^n = (-1)^n \begin{bmatrix} 1 - 2n & -2n \\ 2n & 2n + 1 \end{bmatrix}$ Linear growth.

30 With $a = \Delta t/2$ the trapezoidal step is $\boldsymbol{U}_{n+1} = \dfrac{1}{1 + a^2} \begin{bmatrix} 1 - a^2 & 2a \\ -2a & 1 - a^2 \end{bmatrix} \boldsymbol{U}_n$.

That matrix has orthonormal columns $\Rightarrow$ orthogonal matrix $\Rightarrow \|\boldsymbol{U}_{n+1}\| = \|\boldsymbol{U}_n\|$

31 (a) If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then the infinite cosine series gives $(\cos A)\boldsymbol{x} = (\cos\lambda)\boldsymbol{x}$

(b) $\lambda(A) = 2\pi$ and 0 so $\cos\lambda = 1$ and 1 which means that $\cos A = I$

(c) $\boldsymbol{u}(t) = 3(\cos 2\pi t)(1,1) + 1(\cos 0t)(1,-1)$ [$\boldsymbol{u}' = A\boldsymbol{u}$ has **exp**, $\boldsymbol{u}'' = A\boldsymbol{u}$ has **cos**]

32 For proof 2, square the start of the series to see $(I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3)^2 = I + 2A + \frac{1}{2}(2A)^2 + \frac{1}{6}(2A)^3 + \cdots$. The diagonalizing proof is easiest when it works (needing diagonalizable A).

Problem Set 6.4, page 345

Note A way to complete the proof at the end of page 334, (perturbing the matrix to produce distinct eigenvalues) is now on the course website: "*Proofs of the Spectral Theorem*." **math.mit.edu/linearalgebra**.

1 The first is ASA^{T}: symmetric but eigenvalues are different from 1 and -1 for S.

The second is ASA^{-1}: same eigenvalues as S but not symmetric.

The third is $ASA^{\mathrm{T}} = ASA^{-1}$: **symmetric with the same eigenvalues as S**.

This needed $B = A^{\mathrm{T}} = A^{-1}$ to be an **orthogonal matrix**.

2 (a) ASB stays symmetric like S when $B = A^{\mathrm{T}}$

(b) ASB is similar to S when $B = A^{-1}$

To have both (a) and (b) we need $B = A^{\mathrm{T}} = A^{-1}$ to be an **orthogonal matrix**

3 $A = \begin{bmatrix} 1 & 3 & 6 \\ 3 & 3 & 3 \\ 6 & 3 & 5 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix}$ $= \frac{1}{2}(A + A^{\mathrm{T}}) + \frac{1}{2}(A - A^{\mathrm{T}})$ $=$ **symmetric** $+$ **skew-symmetric**.

4 $(A^{\mathrm{T}}CA)^{\mathrm{T}} = A^{\mathrm{T}}C^{\mathrm{T}}(A^{\mathrm{T}})^{\mathrm{T}} = A^{\mathrm{T}}CA$. When A is 6 by 3, C will be 6 by 6 and the triple product $A^{\mathrm{T}}CA$ is 3 by 3.

5 $\lambda = 0, 4, -2$; unit vectors $\pm(0,1,-1)/\sqrt{2}$ and $\pm(2,1,1)/\sqrt{6}$ and $\pm(1,-1,-1)/\sqrt{3}$.

6 $\lambda = 10$ and -5 in $\Lambda = \begin{bmatrix} 10 & 0 \\ 0 & -5 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ have to be normalized to unit vectors in $Q = \dfrac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.

7 $Q = \dfrac{1}{3} \begin{bmatrix} 2 & 1 & 2 \\ 2 & -2 & -1 \\ -1 & -2 & 2 \end{bmatrix}$. The columns of Q are unit eigenvectors of S. Each unit eigenvector could be multiplied by -1

8 $S = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$ has $\lambda = 0$ and 25 so the columns of Q are the two eigenvectors: $Q = \begin{bmatrix} .8 & .6 \\ -.6 & .8 \end{bmatrix}$ or we can exchange columns or reverse the signs of any column.

9 (a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ has $\lambda = -1$ and 3 (b) The pivots $1, 1 - b^2$ have the same signs as the λ's (c) The trace is $\lambda_1 + \lambda_2 = 2$, so S can't have two negative eigenvalues.

10 If $A^3 = 0$ then all $\lambda^3 = 0$ so all $\lambda = 0$ as in $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. If A is *symmetric* then $A^3 = Q\Lambda^3 Q^{\mathrm{T}} = 0$ requires $\Lambda = 0$. The only symmetric A is $Q\,0\,Q^{\mathrm{T}} =$ zero matrix.

11 If λ is complex then $\overline{\lambda}$ is also an eigenvalue $(A\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}})$. Always $\lambda + \overline{\lambda}$ is real. The trace is real so the third eigenvalue of a 3 by 3 real matrix must be real.

12 If $\boldsymbol{x}$ is not real then $\lambda = \boldsymbol{x}^{\mathrm{T}} A\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ is *not* always real. Can't assume real eigenvectors!

13 $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = 2 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + 4 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$; $\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} = 0 \begin{bmatrix} .64 & -.48 \\ -.48 & .36 \end{bmatrix} + 25 \begin{bmatrix} .36 & .48 \\ .48 & .64 \end{bmatrix}$

14 $\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix}$ is an Q matrix so $P_1 + P_2 = \boldsymbol{x}_1\boldsymbol{x}_1^{\mathrm{T}} + \boldsymbol{x}_2\boldsymbol{x}_2^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \boldsymbol{x}_2^{\mathrm{T}} \end{bmatrix} = I$; also $P_1 P_2 = \boldsymbol{x}_1(\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{x}_2)\boldsymbol{x}_2^{\mathrm{T}} =$ zero matrix.

Second proof: $P_1 P_2 = P_1(I - P_1) = P_1 - P_1 = 0$ since $P_1^2 = P_1$.

15 $A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ has $\lambda = ib$ and $-ib$. The block matrices $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$ and $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ are also skew-symmetric with $\lambda = ib$ (twice) and $\lambda = -ib$ (twice).

16 M is skew-symmetric and **orthogonal**; λ's must be i, i, $-i$, $-i$ to have trace zero.

17 $A = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}$ has $\lambda = 0, 0$ and only one independent eigenvector $\boldsymbol{x} = (i, 1)$. The good property for complex matrices is not $A^{\mathrm{T}} = A$ (symmetric) but $\overline{A}^{\mathrm{T}} = A$ (Hermitian with real eigenvalues and orthogonal eigenvectors: see Problem 22 and Section 9.2).

18 (a) If $A\boldsymbol{z} = \lambda\boldsymbol{y}$ and $A^{\mathrm{T}}\boldsymbol{y} = \lambda\boldsymbol{z}$ then $B[\,\boldsymbol{y};\;\; -\boldsymbol{z}\,] = [\,-A\boldsymbol{z};\;\; A^{\mathrm{T}}\boldsymbol{y}\,] = -\lambda[\,\boldsymbol{y};\;\; -\boldsymbol{z}\,]$. So $-\lambda$ is also an eigenvalue of B. (b) $A^{\mathrm{T}}A\boldsymbol{z} = A^{\mathrm{T}}(\lambda\boldsymbol{y}) = \lambda^2\boldsymbol{z}$. (c) $\lambda = -1, -1, 1, 1$; $\boldsymbol{x}_1 = (1, 0, -1, 0)$, $\boldsymbol{x}_2 = (0, 1, 0, -1)$, $\boldsymbol{x}_3 = (1, 0, 1, 0)$, $\boldsymbol{x}_4 = (0, 1, 0, 1)$.

19 The eigenvalues of $S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ are $0, \sqrt{2}, -\sqrt{2}$ by Problem 16 with

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \end{bmatrix}, \boldsymbol{x}_3 = \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \end{bmatrix}.$$

20 **1.** $\boldsymbol{y}$ is in the nullspace of S and $\boldsymbol{x}$ is in the column space (that is also row space because $S = S^{\mathrm{T}}$). The nullspace and row space are perpendicular so $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} = 0$.

2. If $S\boldsymbol{x} = \lambda\boldsymbol{x}$ and $S\boldsymbol{y} = \beta\boldsymbol{y}$ then shift S by βI to have a zero eigenvalue that matches Step **1**.$(S - \beta I)\boldsymbol{x} = (\lambda - \beta)\boldsymbol{x}$ and $(S - \beta I)\boldsymbol{y} = \mathbf{0}$ and again $\boldsymbol{x}$ is perpendicular to y.

21 S has $X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; B has $X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2d \end{bmatrix}$. Perpendicular for A. Not perpendicular for S since $B^{\mathrm{T}} \neq B$

22 $S = \begin{bmatrix} 1 & 3+4i \\ 3-4i & 1 \end{bmatrix}$ is a *Hermitian matrix* ($\overline{S}^{\mathrm{T}} = S$). Its eigenvalues 6 and -4 are *real*. Adjust equations (1)–(2) in the text to prove that λ is always real when $\overline{S}^{\mathrm{T}} = S$:

$S\boldsymbol{x} = \lambda\boldsymbol{x}$ leads to $\overline{S}\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$. Transpose to $\overline{x}^{\mathrm{T}}S = \overline{x}^{\mathrm{T}}\overline{\lambda}$ using $\overline{S}^{\mathrm{T}} = S$.

Then $\overline{\boldsymbol{x}}^{\mathrm{T}}S\boldsymbol{x} = \overline{\boldsymbol{x}}^{\mathrm{T}}\lambda\boldsymbol{x}$ and also $\overline{\boldsymbol{x}}^{\mathrm{T}}S\boldsymbol{x} = \overline{\boldsymbol{x}}^{\mathrm{T}}\overline{\lambda}\boldsymbol{x}$. So $\lambda = \overline{\lambda}$ is real.

23 (a) False. $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (b) True from $A^{\mathrm{T}} = Q\Lambda Q^{\mathrm{T}} = A$ (c) True from $S^{-1} = Q\Lambda^{-1}Q^{\mathrm{T}}$ (d) False!

24 A and A^{T} have the same λ's but the *order* of the $\boldsymbol{x}$'s can change. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has $\lambda_1 = i$ and $\lambda_2 = -i$ with $\boldsymbol{x}_1 = (1, i)$ first for A but $\boldsymbol{x}_1 = (1, -i)$ is first for A^{T}.

25 A is invertible, orthogonal, permutation, diagonalizable, Markov; B is projection, diagonalizable, Markov. A allows $QR, X\Lambda X^{-1}, Q\Lambda Q^{\mathrm{T}}$; B allows $X\Lambda X^{-1}$ and $Q\Lambda Q^{\mathrm{T}}$.

26 Symmetry gives $Q\Lambda Q^{\mathrm{T}}$ if $b = 1$; repeated λ and no X if $b = -1$; singular if $b = 0$.

27 Orthogonal and symmetric requires $|\lambda| = 1$ and λ real, so $\lambda = \pm 1$. Then $S = \pm I$ or
$S = Q\Lambda Q^{\mathrm{T}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$.

28 Eigenvectors $(1, 0)$ and $(\mathbf{1}, \mathbf{1})$ give a 45° angle even with A^{T} very close to A.

29 The roots of $\lambda^2 + b\lambda + c = 0$ are $\frac{1}{2}(-b \pm \sqrt{b^2 - 4ac})$. Then $\lambda_1 - \lambda_2$ is $\sqrt{b^2 - 4c}$. For $\det(A + tB - \lambda I)$ we have $b = -3 - 8t$ and $c = 2 + 16t - t^2$. The minimum of $b^2 - 4c$ is $1/17$ at $t = 2/17$. Then $\lambda_2 - \lambda_1 = 1/\sqrt{17}$: close but not equal !

30 $S = \begin{bmatrix} 4 & 2+i \\ 2-i & 0 \end{bmatrix} = \overline{S}^{\mathrm{T}}$ has real eigenvalues $\lambda = 5$ and -1 with trace $= 4$ and $\det = -5$. The solution to **20** proves that λ is real when $\overline{S}^{\mathrm{T}} = S$ is Hermitian.

31 (a) $A = Q\Lambda\overline{Q}^{\mathrm{T}}$ times $\overline{A}^{\mathrm{T}} = Q\overline{\Lambda}^{\mathrm{T}}\overline{Q}^{\mathrm{T}}$ equals $\overline{A}^{\mathrm{T}}$ times A because $Q = \overline{Q}^{\mathrm{T}}$ and $\Lambda\overline{\Lambda}^{\mathrm{T}} = \overline{\Lambda}^{\mathrm{T}}\Lambda$ (diagonal!) (b) Step 2: The $1, 1$ entries of $\overline{T}^{\mathrm{T}}\, T$ and $T\overline{T}^{\mathrm{T}}$ are $|a|^2$ and $|a|^2 + |b|^2$. Equally makes $b = 0$ and $T = \Lambda$.

32 a_{11} is $\left[q_{11} \dots q_{1n}\right] \left[\lambda_1\overline{q}_{11} \dots \lambda_n\overline{q}_{1n}\right]^{\mathrm{T}} \leq \lambda_{\max}\left(|q_{11}|^2 + \cdots + |q_{1n}|^2\right) = \lambda_{\max}$.

33 (a) $\boldsymbol{x}^{\mathrm{T}}(A\boldsymbol{x}) = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x} = -\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$. (b) $\overline{\boldsymbol{z}}^{\mathrm{T}}A\boldsymbol{z}$ is pure imaginary, its real part is $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} + \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{y} = 0 + 0$ (c) $\det A = \lambda_1 \dots \lambda_n \geq 0$: pairs of λ's $= ib, -ib$.

34 Since S is diagonalizable with eigenvalue matrix $\Lambda = 2I$, the matrix S itself has to be $X\Lambda X^{-1} = X(2I)X^{-1} = 2I$. (The unsymmetric matrix $[2\ 1\ ;\ 0\ 2]$ also has $\lambda = 2, 2$.)

35 (a) $S^{\mathrm{T}} = S$ and $S^{\mathrm{T}}S = I$ lead to $S^2 = I$.

(b) The only possible eigenvalues of S are 1 and -1.

(c) $\Lambda = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$ so $S = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \Lambda \begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{bmatrix} = Q_1Q_1^{\mathrm{T}} - Q_2Q_2^{\mathrm{T}}$ with $Q_1^{\mathrm{T}}Q_2 = 0$.

36 $(A^{\mathrm{T}}SA)^{\mathrm{T}} = A^{\mathrm{T}}S^{\mathrm{T}}A^{\mathrm{TT}} = A^{\mathrm{T}}SA$. This matrix $A^{\mathrm{T}}SA$ may have different eigenvalues from S, but the "inertia theorem" says that the two sets of eigenvalues have the same signs. The inertia = number of (positive, zero, negative) eigenvalues is the same for S and $A^{\mathrm{T}}SA$.

37 Substitute $\lambda = a$ to find $\det(S - aI) = a^2 - a^2 - ca + ac - b^2 = -b^2$ (negative). The parabola crosses at the eigenvalues λ because they have $\det(S - \lambda I) = 0$.

Problem Set 6.5, page 358

1 Suppose $a > 0$ and $ac > b^2$ so that also $c > b^2/a > 0$.

(i) The eigenvalues have the *same sign* because $\lambda_1\lambda_2 = \det = ac - b^2 > 0$.

(ii) That sign is *positive* because $\lambda_1 + \lambda_2 > 0$ (it equals the trace $a + c > 0$).

2 Only $S_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}$ has two positive eigenvalues since $101 > 10^2$.

$\boldsymbol{x}^{\mathrm{T}}S_1\boldsymbol{x} = 5x_1^2 + 12x_1x_2 + 7x_2^2$ is negative for example when $x_1 = 4$ and $x_2 = -3$: A_1 is not positive definite as its determinant confirms; S_2 has trace c_0; S_3 has $\det = 0$.

3 Positive definite for $-3 < b < 3$ $\begin{bmatrix}1 & 0\\ b & 1\end{bmatrix}\begin{bmatrix}1 & b\\ 0 & 9-b^2\end{bmatrix} = \begin{bmatrix}1 & 0\\ b & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 9-b^2\end{bmatrix}\begin{bmatrix}1 & b\\ 0 & 1\end{bmatrix} = LDL^{\mathrm{T}}$

Positive definite for $c > 8$ $\begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}\begin{bmatrix}2 & 4\\ 0 & c-8\end{bmatrix} = \begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & c-8\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix} = LDL^{\mathrm{T}}.$

Positive definite for $c > b$ $L = \begin{bmatrix}1 & 1\\ -b/c & 0\end{bmatrix}$ $D = \begin{bmatrix}c & 0\\ 0 & c-b/c\end{bmatrix}$ $S = LDL^{\mathrm{T}}$.

4 $f(x,y) = x^2 + 4xy + 9y^2 = (x+2y)^2 + 5y^2;\ x^2 + 6xy + 9y^2 = (x+3y)^2.$

5 $x^2 + 4xy + 3y^2 = (x+2y)^2 - y^2 =$ *difference of squares* is negative at $x = 2$, $y = -1$, where the first square is zero.

6 $A = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ produces $f(x,y) = \begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \boldsymbol{2xy}$. A has $\lambda = 1$ and -1. Then A is an *indefinite matrix* and $f(x,y) = 2xy$ has a *saddle point*.

7 $A^{\mathrm{T}}A = \begin{bmatrix}1 & 2\\ 2 & 13\end{bmatrix}$ and $A^{\mathrm{T}}A = \begin{bmatrix}6 & 5\\ 5 & 6\end{bmatrix}$ are positive definite; $A^{\mathrm{T}}A = \begin{bmatrix}2 & 3 & 3\\ 3 & 5 & 4\\ 3 & 4 & 5\end{bmatrix}$ is singular (and positive semidefinite). The first two A's have independent columns. The 2 by 3 A cannot have full column rank 3, with only 2 rows; $A^{\mathrm{T}}A$ is singular.

8 $S = \begin{bmatrix}3 & 6\\ 6 & 16\end{bmatrix} = \begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}$. Pivots $3, 4$ outside squares, ℓ_{ij} inside. $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} = 3(x+2y)^2 + 4y^2$

9 $S = \begin{bmatrix}4 & -4 & 8\\ -4 & 4 & -8\\ 8 & -8 & 16\end{bmatrix}$ has only one pivot $= 4$, rank $S = 1$, eigenvalues are $24, 0, 0, \det S = 0$.

10 $S = \begin{bmatrix}2 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}$ has pivots $2, \frac{3}{2}, \frac{4}{3}$; $T = \begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{bmatrix}$ is singular; $T\begin{bmatrix}1\\ 1\\ 1\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$.

11 Corner determinants $|S_1| = 2$, $|S_2| = 6$, $|S_3| = 30$. The pivots are $2/1, 6/2, 30/6$.

12 S is positive definite for $c > 1$; determinants $c, c^2 - 1$, and $(c-1)^2(c+2) > 0$. T is *never* positive definite (determinants $d - 4$ and $-4d + 12$ are never both positive).

13 $S = \begin{bmatrix} 1 & 5 \\ 5 & 10 \end{bmatrix}$ is an example with $a + c > 2b$ but $ac < b^2$, so not positive definite.

14 The eigenvalues of S^{-1} are positive because they are $1/\lambda(S)$. Also the entries of S^{-1} pass the determinant tests. And $\boldsymbol{x}^{\mathrm{T}} S^{-1} \boldsymbol{x} = (S^{-1}\boldsymbol{x})^{\mathrm{T}} S (S^{-1}\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$.

15 Since $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} > 0$ and $\boldsymbol{x}^{\mathrm{T}} T \boldsymbol{x} > 0$ we have $\boldsymbol{x}^{\mathrm{T}}(S+T)\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}} T \boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$. Then $S + T$ is a positive definite matrix. The second proof uses the test $S = A^{\mathrm{T}} A$ (independent columns in A): If $S = A^{\mathrm{T}} A$ and $T = B^{\mathrm{T}} B$ pass this test, then $S + T = \begin{bmatrix} A & B \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} A \\ B \end{bmatrix}$ also passes, and must be positive definite.

16 $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ is zero when $(x_1, x_2, x_3) = (0, 1, 0)$ because of the zero on the diagonal. Actually $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ goes *negative* for $\boldsymbol{x} = (1, -10, 0)$ because the second pivot is *negative*.

17 If a_{jj} were smaller than all λ's, $S - a_{jj} I$ would have all eigenvalues > 0 (positive definite). But $S - a_{jj} I$ has a *zero* in the (j, j) position; impossible by Problem 16.

18 If $S\boldsymbol{x} = \lambda \boldsymbol{x}$ then $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = \lambda \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}$. If S is positive definite this leads to $\lambda = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} / \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} > 0$ (ratio of positive numbers). So positive energy $\Rightarrow$ positive eigenvalues.

19 All cross terms are $\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j = 0$ because symmetric matrices have orthogonal eigenvectors. So positive eigenvalues $\Rightarrow$ positive energy.

20 (a) The determinant is positive; all $\lambda > 0$ (b) All projection matrices except I are singular (c) The diagonal entries of D are its eigenvalues (d) $S = -I$ has $\det = +1$ when n is even.

21 S is positive definite when $s > 8$; T is positive definite when $t > 5$ by determinants.

22 $A = \dfrac{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}{\sqrt{2}} \begin{bmatrix} \sqrt{9} & \\ & \sqrt{1} \end{bmatrix} \dfrac{\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}{\sqrt{2}} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$; $A = Q \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} Q^{\mathrm{T}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

23 $x^2/a^2 + y^2/b^2$ is $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ when $S = \mathrm{diag}(1/a^2, 1/b^2)$. Then $\lambda_1 = 1/a^2$ and $\lambda_2 = 1/b^2$ so $a = 1/\sqrt{\lambda_1}$ and $b = 1/\sqrt{\lambda_2}$. The ellipse $9x^2 + 16y^2 = 1$ has axes with half-lengths $a = \frac{1}{3}$ and $b = \frac{1}{4}$. The points $(\frac{1}{3}, 0)$ and $(0, \frac{1}{4})$ are at the ends of the axes.

24 The ellipse $x^2 + xy + y^2 = 1$ has axes with half-lengths $1/\sqrt{\lambda} = \sqrt{2}$ and $\sqrt{2/3}$.

25 $S = C^{\mathrm{T}}C = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix}$; $\begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$

26 The Cholesky factors $C = \left(L\sqrt{D}\right)^{\mathrm{T}} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{5} \end{bmatrix}$ have *square roots* of the pivots from D. Note again $C^{\mathrm{T}}C = LDL^{\mathrm{T}} = S$.

27 Writing out $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}LDL^{\mathrm{T}}\boldsymbol{x}$ gives $ax^2+2bxy+cy^2 = a(x+\frac{b}{a}y)^2+\frac{ac-b^2}{a}y^2$. So the LDL^{T} from elimination is exactly the same as *completing the square*. The example $2x^2+8xy+10y^2 = 2(x+2y)^2+2y^2$ with pivots $2, 2$ outside the squares and multiplier 2 inside.

28 $\det S = (1)(10)(1) = 10$; $\lambda = 2$ and 5; $\boldsymbol{x}_1 = (\cos\theta, \sin\theta)$, $\boldsymbol{x}_2 = (-\sin\theta, \cos\theta)$; the λ's are positive. So S is positive definite.

29 $S_1 = \begin{bmatrix} 6x^2 & 2x \\ 2x & 2 \end{bmatrix}$ is semidefinite; $f_1 = (\frac{1}{2}x^2 + y)^2 = 0$ on the curve $\frac{1}{2}x^2 + y = 0$; $S_2 = \begin{bmatrix} 6x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is indefinite at $(0, 1)$ where first derivatives $= 0$. Then $x = 0, y = 1$ is a saddle point of the function $f_2(x, y)$.

30 $ax^2 + 2bxy + cy^2$ has a saddle point if $ac < b^2$. The matrix is *indefinite* ($\lambda < 0$ and $\lambda > 0$) because the determinant $ac - b^2$ is *negative*.

31 If $c > 9$ the graph of z is a bowl, if $c < 9$ the graph has a saddle point. When $c = 9$ the graph of $z = (2x + 3y)^2$ is a "trough" staying at zero along the line $2x + 3y = 0$.

32 Orthogonal matrices, exponentials e^{At}, matrices with $\det = 1$ are groups. Examples of subgroups are orthogonal matrices with $\det = 1$, exponentials e^{An} for integer n. Another subgroup: lower triangular elimination matrices E with diagonal 1's.

33 A product ST of symmetric positive definite matrices comes into many applications. The "*generalized*" eigenvalue problem $K\boldsymbol{x} = \lambda M\boldsymbol{x}$ has $ST = M^{-1}K$. (Often we use

eig(K, M) without actually inverting M.) All eigenvalues λ are positive:

$$ST\boldsymbol{x} = \lambda\boldsymbol{x} \text{ gives } (T\boldsymbol{x})^{\mathrm{T}}ST\boldsymbol{x} = (T\boldsymbol{x})^{\mathrm{T}}\lambda x. \text{ Then } \lambda = \boldsymbol{x}^{\mathrm{T}}T^{\mathrm{T}}ST\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}T\boldsymbol{x} > 0.$$

34 The five eigenvalues of K are $2 - 2\cos\frac{k\pi}{6} = 2 - \sqrt{3},\ 2 - 1,\ 2,\ 2 + 1,\ 2 + \sqrt{3}$. The product of those eigenvalues is $6 = \det K$.

35 Put parentheses in $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}CA\boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}}C(A\boldsymbol{x})$. Since C is assumed positive definite, this energy can drop to zero only when $A\boldsymbol{x} = \mathbf{0}$. Sine A is assumed to have independent columns, $A\boldsymbol{x} = \mathbf{0}$ only happens when $\boldsymbol{x} = \mathbf{0}$. Thus $A^{\mathrm{T}}CA$ has positive energy and is positive definite.

My textbooks *Computational Science and Engineering* and *Introduction to Applied Mathematics* start with many examples of $A^{\mathrm{T}}CA$ in a wide range of applications. I believe this is a unifying concept from linear algebra.

36 (a) The eigenvectors of $\lambda_1 I - S$ are $\lambda_1 - \lambda_1, \lambda_1 - \lambda_2, \ldots, \lambda_1 - \lambda_n$. Those are ≥ 0; $\lambda_1 I - S$ is semidefinite.

(b) Semidefinite matrices have energy $\boldsymbol{x}^{\mathrm{T}}(\lambda_1 I - S)\,\boldsymbol{x}_2 \geq 0$. Then $\lambda_1\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \geq \boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$.

(c) Part (b) says $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \leq \lambda_1$ for all $\boldsymbol{x}$. Equality at the eigenvector with $S\boldsymbol{x} = \lambda_1\boldsymbol{x}$.

37 Energy $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} = a\,(x_1+x_2+x_3)^2 + c\,(x_2-x_3)^2 \geq 0$ if $a \geq 0$ and $c \geq 0$: semidefinite. The matrix has rank ≤ 2 and determinant $= 0$; cannot be positive definite for any a and c.

Problem Set 6.6, page 360

1 $B = GCG^{-1} = GF^{-1}AFG^{-1}$ so $M = FG^{-1}$. C similar to A and $B \Rightarrow A$ similar to B.

2 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ is similar to $B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = M^{-1}AM$ with $M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

3 $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = M^{-1}AM$;

$B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$;

$B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

4 A has no repeated λ so it can be diagonalized: $S^{-1}AS = \Lambda$ makes A similar to Λ.

5 $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ are similar (they all have eigenvalues 1 and 0).
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is by itself and also $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is by itself with eigenvalues 1 and -1.

6 *Eight families* of similar matrices: six matrices have $\lambda = 0, 1$ (one family); three matrices have $\lambda = 1, 1$ and three have $\lambda = 0, 0$ (two families each!); one has $\lambda = 1, -1$; one has $\lambda = 2, 0$; two matrices have $\lambda = \frac{1}{2}(1 \pm \sqrt{5})$ (they are in one family).

7 (a) $(M^{-1}AM)(M^{-1}\boldsymbol{x}) = M^{-1}(A\boldsymbol{x}) = M^{-1}\mathbf{0} = \mathbf{0}$ (b) The nullspaces of A and of $M^{-1}AM$ have the same *dimension*. Different vectors and different bases.

8 Same Λ Same S But $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ have the same line of eigenvectors and the same eigenvalues $\lambda = 0, 0$.

9 $A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $A^3 = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$, every $A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$. $A^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

10 $J^2 = \begin{bmatrix} c^2 & 2c \\ 0 & c^2 \end{bmatrix}$ and $J^k = \begin{bmatrix} c^k & kc^{k-1} \\ 0 & c^k \end{bmatrix}$; $J^0 = I$ and $J^{-1} = \begin{bmatrix} c^{-1} & -c^{-2} \\ 0 & c^{-1} \end{bmatrix}$.

11 $\boldsymbol{u}(0) = \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} v(0) \\ w(0) \end{bmatrix}$. The equation $\dfrac{d\boldsymbol{u}}{dt} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \boldsymbol{u}$ has $\dfrac{dv}{dt} = \lambda v + w$ and $\dfrac{dw}{dt} = \lambda w$. Then $w(t) = 2e^{\lambda t}$ and $v(t)$ must include $2te^{\lambda t}$ (this comes from the repeated λ). To match $v(0) = 5$, the solution is $v(t) = 2te^{\lambda t} + 5e^{\lambda t}$.

12 If $M^{-1}JM=K$ then $JM=\begin{bmatrix} m_{21} & m_{22} & m_{23} & m_{24} \\ 0 & \mathbf{0} & \mathbf{0} & 0 \\ m_{41} & m_{42} & m_{43} & m_{44} \\ 0 & 0 & 0 & 0 \end{bmatrix} = MK=\begin{bmatrix} \mathbf{0} & m_{12} & m_{13} & \mathbf{0} \\ 0 & m_{22} & m_{23} & 0 \\ 0 & m_{32} & m_{33} & 0 \\ 0 & m_{42} & m_{43} & 0 \end{bmatrix}$.

That means $m_{21} = m_{22} = m_{23} = m_{24} = 0$. M is not invertible, J not similar to K.

13 The five 4 by 4 Jordan forms with $\lambda = 0, 0, 0, 0$ are $J_1 =$ zero matrix and

$$J_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad J_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$J_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad J_5 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 12 showed that J_3 and J_4 are *not similar*, even with the same rank. Every matrix with all $\lambda = 0$ is "*nilpotent*" (its nth power is $A^n =$ zero matrix). You see $J^4 = 0$ for these matrices. How many possible Jordan forms for $n = 5$ and all $\lambda = 0$?

14 (1) Choose $M_i =$ reverse diagonal matrix to get $M_i^{-1} J_i M_i = M_i^{\mathrm{T}}$ in each block (2) M_0 has those diagonal blocks M_i to get $M_0^{-1} J M_0 = J^{\mathrm{T}}$. (3) $A^{\mathrm{T}} = (M^{-1})^{\mathrm{T}} J^{\mathrm{T}} M^{\mathrm{T}}$ equals $(M^{-1})^{\mathrm{T}} M_0^{-1} J M_0 M^{\mathrm{T}} = (M M_0 M^{\mathrm{T}})^{-1} A (M M_0 M^{\mathrm{T}})$, and A^{T} *is similar to* A.

15 $\det(M^{-1}AM - \lambda I) = \det(M^{-1}AM - M^{-1}\lambda IM)$. This is $\det(M^{-1}(A - \lambda I)M)$. By the product rule, the determinants of M and M^{-1} cancel to leave $\det(A - \lambda I)$.

16 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is similar to $\begin{bmatrix} d & c \\ b & a \end{bmatrix}$; $\begin{bmatrix} b & a \\ d & c \end{bmatrix}$ is similar to $\begin{bmatrix} c & d \\ a & b \end{bmatrix}$. So two pairs of similar matrices but $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is not similar to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$: different eigenvalues!

17 (a) *False*: Diagonalize a nonsymmetric $A = S\Lambda S^{-1}$. Then Λ is symmetric and similar

(b) *True*: A singular matrix has $\lambda = 0$. (c) *False*: $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ are similar

(they have $\lambda = \pm 1$) (d) *True*: Adding I increases all eigenvalues by 1

18 $AB = B^{-1}(BA)B$ so AB is similar to BA. *If* $AB\boldsymbol{x} = \lambda\boldsymbol{x}$ *then* $BA(B\boldsymbol{x}) = \lambda(B\boldsymbol{x})$.

19 Diagonal blocks 6 by 6, 4 by 4; AB has the same eigenvalues as BA plus $6 - 4$ zeros.

20 (a) $A = M^{-1}BM \Rightarrow A^2 = (M^{-1}BM)(M^{-1}BM) = M^{-1}B^2M$. So A^2 is similar to B^2. (b) A^2 equals $(-A)^2$ but A may not be similar to $B = -A$ (it could be!). (c) $\begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}$ is diagonalizable to $\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$ because $\lambda_1 \neq \lambda_2$, so these matrices are similar. (d) $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ has only one eigenvector, so not diagonalizable (e) PAP^{T} is similar to A.

21 J^2 has three 1's down the *second* superdiagonal, and *two* independent eigenvectors for $\lambda = 0$. Its 5 by 5 Jordan form is $\begin{bmatrix} J_3 & \\ & J_2 \end{bmatrix}$ with $J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $J_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Note to professors: An interesting question: *Which matrices A have (complex) square roots $R^2 = A$?* If A is invertible, no problem. But any Jordan blocks for $\lambda = 0$ must have sizes $n_1 \geq n_2 \geq \ldots \geq n_k \geq n_{k+1} = 0$ that come in pairs like 3 and 2 in this example: $n_1 = (n_2$ or $n_2+1)$ and $n_3 = (n_4$ or $n_4+1)$ and so on.

A list of all 3 by 3 and 4 by 4 Jordan forms could be $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, $\begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$,

$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}$ (for any numbers a, b, c) with $3, 2, 1$ eigenvectors; $\operatorname{diag}(a, b, c, d)$ and $\begin{bmatrix} a & 1 & & \\ & a & & \\ & & b & \\ & & & c \end{bmatrix}$,

$\begin{bmatrix} a & 1 & & \\ & a & & \\ & & b & 1 \\ & & & b \end{bmatrix}$, $\begin{bmatrix} a & 1 & & \\ & a & 1 & \\ & & a & \\ & & & b \end{bmatrix}$, $\begin{bmatrix} a & 1 & & \\ & a & 1 & \\ & & a & 1 \\ & & & a \end{bmatrix}$ with $4, 3, 2, 1$ eigenvectors.

22 If all roots are $\lambda = 0$, this means that $\det(A - \lambda I)$ must be just λ^n. The Cayley-Hamilton Theorem in Problem 6.2.32 immediately says that A^n = zero matrix. The key example is a single n by n Jordan block (with $n-1$ ones above the diagonal): Check directly that J^n = zero matrix.

23 Certainly Q_1R_1 is similar to $R_1Q_1 = Q_1^{-1}(Q_1R_1)Q_1$. Then $A_1 = Q_1R_1 - cs^2I$ is similar to $A_2 = R_1Q_1 - cs^2I$.

24 A could have eigenvalues $\lambda = 2$ and $\lambda = \frac{1}{2}$ (A could be diagonal). Then A^{-1} has the same two eigenvalues (and is similar to A).

Problem Set 6.7, page 371

1 $A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}} = \dfrac{\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \dfrac{\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}{\sqrt{5}}$

2 This $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is a 2 by 2 matrix of rank 1. Its row space has basis $\boldsymbol{v}_1$, its nullspace has basis $\boldsymbol{v}_2$, its column space has basis $\boldsymbol{u}_1$, its left nullspace has basis $\boldsymbol{u}_2$:

$$\text{Row space} \quad \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ Nullspace } \quad \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{Column space} \quad \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \ \boldsymbol{N}(A^{\mathrm{T}}) \quad \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

3 If A has rank 1 then so does $A^{\mathrm{T}}A$. The only nonzero eigenvalue of $A^{\mathrm{T}}A$ is its trace, which is the sum of all a_{ij}^2. (Each diagonal entry of $A^{\mathrm{T}}A$ is the sum of a_{ij}^2 down one column, so the trace is the sum down all columns.) Then $\sigma_1 =$ square root of this sum, and $\sigma_1^2 =$ this sum of all a_{ij}^2.

4 $A^{\mathrm{T}}A = AA^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ has eigenvalues $\sigma_1^2 = \dfrac{3+\sqrt{5}}{2}$, $\sigma_2^2 = \dfrac{3-\sqrt{5}}{2}$. But A is indefinite

$\sigma_1 = (1+\sqrt{5})/2 = \lambda_1(A)$, $\sigma_2 = (\sqrt{5}-1)/2 = -\lambda_2(A)$; $\boldsymbol{u}_1 = \boldsymbol{v}_1$ but $\boldsymbol{u}_2 = -\boldsymbol{v}_2$.

5 A proof that ***eigshow*** finds the SVD. When $\boldsymbol{V}_1 = (1,0), \boldsymbol{V}_2 = (0,1)$ the demo finds $A\boldsymbol{V}_1$ and $A\boldsymbol{V}_2$ at some angle θ. A 90° turn by the mouse to $\boldsymbol{V}_2, -\boldsymbol{V}_1$ finds $A\boldsymbol{V}_2$ and $-A\boldsymbol{V}_1$ at the angle $\pi - \theta$. Somewhere between, the constantly orthogonal $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ must produce $A\boldsymbol{v}_1$ and $A\boldsymbol{v}_2$ *at angle* $\pi/2$. Those orthogonal directions give $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.

6 $AA^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has $\sigma_1^2 = 3$ with $\boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\sigma_2^2 = 1$ with $\boldsymbol{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ has $\sigma_1^2 = 3$ with $\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $\sigma_2^2 = 1$ with $\boldsymbol{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$;

and $\boldsymbol{v}_3 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. Then $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = [\,\boldsymbol{u}_1 \ \ \boldsymbol{u}_2\,] \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} [\,\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{v}_3\,]^{\mathrm{T}}$.

7 The matrix A in Problem 6 had $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$ in Σ. The smallest change to rank 1 is **to make** $\boldsymbol{\sigma}_2 = \mathbf{0}$. In the factorization

$$A = U\Sigma V^{\mathrm{T}} = \boldsymbol{u}_1\sigma_1\boldsymbol{v}_1^{\mathrm{T}} + \boldsymbol{u}_2\sigma_2\boldsymbol{v}_2^{\mathrm{T}}$$

this change $\sigma_2 \to 0$ will leave the closest rank–1 matrix as $\boldsymbol{u}_1\sigma_1\boldsymbol{v}_1^{\mathrm{T}}$. See Problem 14 for the general case of this problem.

8 The number $\sigma_{\max}(A^{-1})\sigma_{\max}(A)$ is the same as $\sigma_{\max}(A)/\sigma_{\min}(A)$. This is certainly ≥ 1. It equals 1 if all σ's are equal, and $A = U\Sigma V^{\mathrm{T}}$ is a multiple of an orthogonal matrix. The ratio $\sigma_{\max}/\sigma_{\min}$ is the important **condition number** of A studied in Section 9.2.

9 $A = UV^{\mathrm{T}}$ since all $\sigma_j = 1$, which means that $\Sigma = I$.

10 A rank–1 matrix with $A\boldsymbol{v} = 12\boldsymbol{u}$ would have $\boldsymbol{u}$ in its column space, so $A = \boldsymbol{u}\boldsymbol{w}^{\mathrm{T}}$ for some vector $\boldsymbol{w}$. I intended (but didn't say) that $\boldsymbol{w}$ is a multiple of the unit vector $\boldsymbol{v} = \frac{1}{2}(1,1,1,1)$ in the problem. Then $A = 12\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ to get $A\boldsymbol{v} = 12\boldsymbol{u}$ when $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = 1$.

11 If A has orthogonal columns $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ of lengths $\sigma_1, \ldots, \sigma_n$, then $A^{\mathrm{T}}A$ will be diagonal with entries $\sigma_1^2, \ldots, \sigma_n^2$. So the σ's are definitely the singular values of A (as expected). The eigenvalues of that diagonal matrix $A^{\mathrm{T}}A$ are the columns of I, so $V = I$ in the SVD. Then the $\boldsymbol{u}_i$ are $A\boldsymbol{v}_i/\sigma_i$ which is the unit vector $\boldsymbol{w}_i/\sigma_i$.

The SVD of this A with orthogonal columns is $A = U\Sigma V^{\mathrm{T}} = (A\Sigma^{-1})(\Sigma)(I)$.

12 Since $A^{\mathrm{T}} = A$ we have $\sigma_1^2 = \lambda_1^2$ and $\sigma_2^2 = \lambda_2^2$. But λ_2 is negative, so $\sigma_1 = 3$ and $\sigma_2 = 2$. The unit eigenvectors of A are the same $\boldsymbol{u}_1 = \boldsymbol{v}_1$ as for $A^{\mathrm{T}}A = AA^{\mathrm{T}}$ and $\boldsymbol{u}_2 = -\boldsymbol{v}_2$ (notice the sign change because $\sigma_2 = -\lambda_2$, as in Problem 4).

13 Suppose the SVD of R is $R = U\Sigma V^{\mathrm{T}}$. Then multiply by Q to get $A = QR$. So the SVD of this A is $(QU)\Sigma V^{\mathrm{T}}$. (Orthogonal Q times orthogonal U = orthogonal QU.)

14 The smallest change in A is to set its smallest singular value σ_2 to zero. See # **7**.

15 The singular values of $A + I$ are not $\sigma_j + 1$. They come from eigenvalues of $(A+I)^{\mathrm{T}}(A+I)$.

16 This simulates the random walk used by *Google* on billions of sites to solve $A\boldsymbol{p} = \boldsymbol{p}$. It is like the power method of Section 9.3 except that it follows the links in one "walk" where the vector $p_k = A^k p_0$ averages over all walks.

17 $A = U\Sigma V^{\mathrm{T}} =$ [cosines including $\boldsymbol{u}_4$] **diag**(**sqrt**$(2-\sqrt{2}, 2, 2+\sqrt{2})$) [sine matrix]$^{\mathrm{T}}$. $AV = U\Sigma$ says that differences of sines in V are cosines in U times σ's.

The SVD of the *derivative* on $[0,\pi]$ with $f(0)=0$ has $\boldsymbol{u} = \sin nx$, $\sigma = n$, $\boldsymbol{v} = \cos nx$!

Problem Set 7.1, page 370

1 $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ has rank **1** with $\boldsymbol{u}^{\mathrm{T}} = \boldsymbol{v}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$. Those vectors have $||\boldsymbol{u}||^2 = ||\boldsymbol{v}||^2 = 30$ so the SVD has a division by $\sqrt{30}$ to reach $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$. Multiply by $\sigma_1 = 30$ to recover A.

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} = 30 \frac{\boldsymbol{u}}{\sqrt{30}} \frac{\boldsymbol{v}^{\mathrm{T}}}{\sqrt{30}} = U\Sigma V^{\mathrm{T}} \quad (1 \text{ column in } U \text{ and } V).$$

B has rank $r = 2$. The first two columns of B are independent (the pivot columns). Column 3 is a combination 2 (col 2) – (col 1). Column 4 is 3 (col 2) – 2 (col 1) :

$$B = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{bmatrix} \quad \begin{matrix} (\text{col } 1)(\text{row } 1)^{\mathrm{T}} \\ + \\ (\text{col } 2)(\text{row } 2)^{\mathrm{T}} \end{matrix}$$

Those pivot columns come from the first half of the book : *not orthogonal* ! They don't give the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s of the SVD. For that we need eigenvalues and eigenvectors of $B^{\mathrm{T}}B$ and BB^{T}.

2 All the singular values of I are $\boldsymbol{\sigma} = \mathbf{1}$. We cannot leave out any of the terms $\boldsymbol{u}_i \cdot \boldsymbol{v}_i^{\mathrm{T}}$ without making an error of size 1. And the matrix $A = I$ starts with size 1 ! None of the SVD pieces can be left out.

Notice that the SVD is $I = (U)(I)(U^{\mathrm{T}})$ so that $U = V$. The natural choice for the SVD is just $U\Sigma V^{\mathrm{T}} = III$. But we could actually choose any orthogonal matrix U. (The eigenvectors of I are very far from unique—many choices ! Any orthogonal matrix U holds orthonormal eigenvectors of I.)

One possible rank 5 flag with a 3 by 3 cross of zeros is $A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & 1 & 1 & 1 \\ 1 & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

3 $$\begin{bmatrix}1&2&1&1\\2&2&2&2\\1&2&1&1\end{bmatrix} = \begin{bmatrix}1&0\\0&1\\1&0\end{bmatrix} = \begin{bmatrix}1&2&1&1\\2&2&2&2\end{bmatrix}$$

$$\begin{bmatrix}1&2&2\\1&3&3\end{bmatrix} = \begin{bmatrix}1&2\\1&3\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&1\end{bmatrix} = \begin{bmatrix}\text{pivot}\\\text{columns}\end{bmatrix}\begin{bmatrix}\text{rows}\\\text{of } R\end{bmatrix}$$

4 $BB^{\mathrm{T}} = \begin{bmatrix}1&2&2\\1&3&3\end{bmatrix}\begin{bmatrix}1&1\\2&3\\2&3\end{bmatrix} = \begin{bmatrix}9&13\\13&19\end{bmatrix}$. Trace $\mathbf{28}$, Determinant $\mathbf{2}$.

$B^{\mathrm{T}}B = \begin{bmatrix}1&1\\2&3\\2&3\end{bmatrix}\begin{bmatrix}1&2&2\\1&3&3\end{bmatrix} = \begin{bmatrix}2&5&5\\5&13&13\\5&13&13\end{bmatrix}$. Trace $\mathbf{28}$, Determinant $\mathbf{0}$.

With a small singular value $\sigma_2 \approx \frac{1}{\sqrt{14}}$, B is compressible. But we don't just keep the first row and column of B. The good row $\boldsymbol{v}_1$ and column $\boldsymbol{u}_1$ are eigenvectors of $B^{\mathrm{T}}B$ and BB^{T}.

5 My hand calculation produced $A^{\mathrm{T}}A = \begin{bmatrix}7&10&7\\10&16&10\\7&10&7\end{bmatrix}$ and $\det(A^{\mathrm{T}}A - \lambda I) = -\lambda^3 + 30\lambda^2 - 24\lambda$.

This gives $\lambda = 0$ as one eigenvalue of $A^{\mathrm{T}}A$ (correct). The others are:

$$\lambda^2 - 30\lambda + 24 = 0 \quad \text{gives} \quad \lambda = 15 \pm \sqrt{15^2 - 24} \approx 15 \pm 14 = 29 \text{ and } 1.$$

So $\sigma_1 \approx \sqrt{29}$ and $\sigma_2 = 1$. The **svd** (A) command in MATLAB will give accurate σ's and U and V.

6 The matrix A has trace 4 and determinant 0. So its eigenvalues are 4 and 0—*not used in the SVD*! The matrix $A^{\mathrm{T}}A$ has trace 25 and determinant 0, so $\lambda_1 = 25 = \sigma_1^2$ gives $\sigma_1 = 5$.

The eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ of $A^{\mathrm{T}}A$ (a symmetric matrix!) are orthogonal:

$$\begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{25} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \mathbf{0} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Similarly AA^{T} has orthogonal eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2$:

$$\begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{25} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \mathbf{0} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

7 Multiply both sides of $A = U\Sigma V^{\mathrm{T}}$ by the matrix V to get $\boldsymbol{AV} = \boldsymbol{U\Sigma}$. Column by column this says that $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$. Notice that Σ goes on the **right side** of U when we want to multiply every *column* of U by its singular value σ_i.

8 The text found $\lambda_1 = \sigma_1^2 = \frac{1}{2}\left(3+\sqrt{5}\right)$ and then $\sigma_1 = \frac{1}{2}\left(1+\sqrt{5}\right)$. Then $\sigma_1 + 1$ equals σ_1^2. Also $\lambda_2 = \sigma_2^2 = \frac{1}{2}\left(3-\sqrt{5}\right)$ and $\sigma_2 = \frac{1}{2}\left(\sqrt{5}-1\right)$ and $\sigma_1 - \sigma_2 = \frac{1}{2} + \frac{1}{2} = 1$. (**Why don't we choose** $\sigma_2 = \frac{1}{2}\left(1-\sqrt{5}\right)$?).

9 The 20 by 40 random matrices are $A =$ **rand** $(20, 40)$ and $B =$ **randn** $(20, 40)$. With those random choices the 20 rows are independent with probability 1. Notice for these *continuous probabilities, this does not mean that the rows are* ***always*** *independent* ! A random determinant might be 0 even when the probability of nonzero is 1.

MATLAB again gives the singular values of a random A and B.

By averaging 100 samples you would begin to see the expected distribution of σ's, which is highly imortant in "random matrix theory".

Problem Set 7.2, page 379

1 $A = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix}$ has eigenvalues 0 and 0; $A^{\mathrm{T}}A = \begin{bmatrix} 0 & 0 \\ 0 & 16 \end{bmatrix}$ has eigenvalues $\lambda = 16$ and 0. Then $\sigma_1(A) = \sqrt{16} = 4$. The eigenvectors of $A^{\mathrm{T}}A$ and AA^{T} are the columns of $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Then $U\Sigma V^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix} = A.$

$A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$ gives $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix}$ with $\lambda_1 = 16$ and $\lambda_2 = 1$. Same U and V.

Then $U\Sigma V^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix} = A.$

2 $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$ leads to $A^{\mathrm{T}}A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ with eigenvectors in $V = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

$\sigma_1^2 = 8 \quad \boldsymbol{u}_1 = \dfrac{A\boldsymbol{v}_1}{\sigma_1} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 4 \\ 0 \end{bmatrix} \dfrac{1}{\sigma_1}$ has unit vector $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\sigma_1 = 2\sqrt{2}$

$\sigma_2^2 = 2 \quad \boldsymbol{u}_2 = \dfrac{A\boldsymbol{v}_2}{\sigma_2} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 2 \end{bmatrix} \dfrac{1}{\sigma_2}$ has unit vector $\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\sigma_2 = \sqrt{2}$

The full SVD is $A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & \\ & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} / \sqrt{2}.$

3 Problem 7.2.2 happens to have $AA^{\mathrm{T}} =$ diagonal matrix $\begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$. So its eigenvectors $(1, 0)$ and $(0, 1)$ go in $U = I$. Its eigenvalues are $\sigma_1^2 = 8$ and $\sigma_2^2 = 2$. The rows of A are orthogonal but not *orthonormal*. So $A^{\mathrm{T}}A$ is not diagonal and V is not I.

4 $AA^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has $\sigma_1^2 = 3$ with $\boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\sigma_2^2 = 1$ with $\boldsymbol{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ has $\sigma_1^2 = 3$ with $\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $\sigma_2^2 = 1$ with $\boldsymbol{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$

and $\boldsymbol{v}_3 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. Then $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = U\Sigma.$

5 (a) $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ has $\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in its row space and $\boldsymbol{u}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ in its column space. Those are unit vectors.

Since $A^{\mathrm{T}}A = \begin{bmatrix} 10 & 10 \\ 10 & 10 \end{bmatrix}$ has $\lambda_1 = 20$ and $\lambda_2 = 0$, A itself has $\sigma_1 = \sqrt{20}$ and has no σ_2. (Remember that the r singular values have to be strictly positive!)

(b) If we want square matrices U and V, choose $\boldsymbol{u}_2$ and $\boldsymbol{v}_2$ orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$:

$$U = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} \quad \text{and} \quad V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

6 If $A = U\Sigma V^{\mathrm{T}}$ then $A^{\mathrm{T}} = V\Sigma^{\mathrm{T}}U^{\mathrm{T}}$ and $A^{\mathrm{T}}A = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}$. This is a diagonalization $V\Lambda V^{\mathrm{T}}$ with $\Lambda = \Sigma^{\mathrm{T}}\Sigma$ (so each $\sigma_i^2 = \lambda_i$). Similarly $AA^{\mathrm{T}} = U\Sigma\Sigma^{\mathrm{T}}U^{\mathrm{T}}$ is a diagonalization of AA^{T}. We see that the eigenvalues in $\Sigma\Sigma^{\mathrm{T}}$ are the same $\sigma_i^2 = \lambda_i$.

7 This small question is a key to everything. It is based on the associative law $(AA^{\mathrm{T}})A = A(A^{\mathrm{T}}A)$. Here we are applying both sides to an eigenvector $\boldsymbol{v}$ of $A^{\mathrm{T}}A$:

$$(AA^{\mathrm{T}})A\boldsymbol{v} = A(A^{\mathrm{T}}A)\boldsymbol{v} = A\lambda\boldsymbol{v} = \lambda A\boldsymbol{v}.$$

So **$A\boldsymbol{v}$ is an eigenvector of AA^{T}** with the same eigenvalue λ.

8 $A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}} = \dfrac{\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \dfrac{\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}{\sqrt{5}}$

9 This $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is a 2 by 2 matrix of rank 1. Its row space has basis $\boldsymbol{v}_1$, its nullspace has basis $\boldsymbol{v}_2$, its column space has basis $\boldsymbol{u}_1$, its left nullspace has basis $\boldsymbol{u}_2$:

$$\text{Row space} \quad \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{Nullspace} \quad \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{Column space} \quad \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \boldsymbol{N}(A^{\mathrm{T}}) \quad \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

10 If A has rank 1 then so does $A^{\mathrm{T}}A$. The only nonzero eigenvalue of $A^{\mathrm{T}}A$ is its trace, which is the sum of all a_{ij}^2. (Each diagonal entry of $A^{\mathrm{T}}A$ is the sum of a_{ij}^2 down one column, so the trace is the sum down all columns.) Then $\sigma_1 =$ square root of this sum, and $\sigma_1^2 =$ this sum of all a_{ij}^2.

11 $A^{\mathrm{T}}A = AA^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ has eigenvalues $\sigma_1^2 = \dfrac{3+\sqrt{5}}{2}$, $\sigma_2^2 = \dfrac{3-\sqrt{5}}{2}$. But A is indefinite $\sigma_1 = (1+\sqrt{5})/2 = \lambda_1(A)$, $\boldsymbol{\sigma_2 = (\sqrt{5}-1)/2 = -\lambda_2(A)}$; $\boldsymbol{u}_1 = \boldsymbol{v}_1$ but $\boldsymbol{u}_2 = -\boldsymbol{v}_2$.

12 A proof that ***eigshow*** finds the SVD. When $\boldsymbol{V}_1 = (1,0), \boldsymbol{V}_2 = (0,1)$ the demo finds $A\boldsymbol{V}_1$ and $A\boldsymbol{V}_2$ at some angle θ. A $90°$ turn by the mouse to $\boldsymbol{V}_2, -\boldsymbol{V}_1$ finds $A\boldsymbol{V}_2$ and $-A\boldsymbol{V}_1$ at the angle $\pi - \theta$. Somewhere between, the constantly orthogonal $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ must produce $A\boldsymbol{v}_1$ and $A\boldsymbol{v}_2$ *at angle* $\pi/2$. Those orthogonal directions give $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.

13 The number $\sigma_{\max}(A^{-1})\sigma_{\max}(A)$ is the same as $\sigma_{\max}(A)/\sigma_{\min}(A)$. This is certainly ≥ 1. It equals 1 if all σ's are equal, and $A = U\Sigma V^{\mathrm{T}}$ is a multiple of an orthogonal matrix. The ratio $\sigma_{\max}/\sigma_{\min}$ is the important **condition number** of A studied in Section 9.2.

14 $A = UV^{\mathrm{T}}$ since all $\sigma_j = 1$, which means that $\Sigma = I$.

15 A rank–1 matrix with $A\boldsymbol{v} = 12\boldsymbol{u}$ would have $\boldsymbol{u}$ in its column space, so $A = \boldsymbol{u}\boldsymbol{w}^{\mathrm{T}}$ for some vector $\boldsymbol{w}$. I intended (but didn't say) that $\boldsymbol{w}$ is a multiple of the unit vector $\boldsymbol{v} = \frac{1}{2}(1,1,1,1)$ in the problem. Then $A = 12\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ to get $A\boldsymbol{v} = 12\boldsymbol{u}$ when $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = 1$.

16 If A has orthogonal columns $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ of lengths $\sigma_1, \ldots, \sigma_n$, then $A^{\mathrm{T}}A$ will be diagonal with entries $\sigma_1^2, \ldots, \sigma_n^2$. So the σ's are definitely the singular values of A (as expected). The eigenvalues of that diagonal matrix $A^{\mathrm{T}}A$ are the columns of I, so $V = I$ in the SVD. Then the $\boldsymbol{u}_i$ are $A\boldsymbol{v}_i/\sigma_i$ which is the unit vector $\boldsymbol{w}_i/\sigma_i$.

The SVD of this A with orthogonal columns is $A = U\Sigma V^{\mathrm{T}} = (A\Sigma^{-1})(\Sigma)(I)$.

17 Since $A^{\mathrm{T}} = A$ we have $\sigma_1^2 = \lambda_1^2$ and $\sigma_2^2 = \lambda_2^2$. But λ_2 is negative, so $\sigma_1 = 3$ and $\sigma_2 = 2$. The unit eigenvectors of A are the same $\boldsymbol{u}_1 = \boldsymbol{v}_1$ as for $A^{\mathrm{T}}A = AA^{\mathrm{T}}$ and $\boldsymbol{u}_2 = -\boldsymbol{v}_2$ (notice the sign change because $\sigma_2 = -\lambda_2$, as in Problem 11).

18 Suppose the SVD of R is $R = U\Sigma V^{\mathrm{T}}$. Then multiply by Q to get $A = QR$. So the SVD of this A is $(QU)\Sigma V^{\mathrm{T}}$. (Orthogonal Q times orthogonal $U =$ orthogonal QU.)

19 The smallest change in A is to set its smallest singular value σ_2 to zero.

20 $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 100 & 1 \end{bmatrix} = \begin{bmatrix} 10001 & 100 \\ 100 & 1 \end{bmatrix}$ has eigenvalues $\lambda(A^{\mathrm{T}}A) = \sigma^2(A)$.

$\lambda^2 - 10002\lambda + 1 = 0$ gives $\lambda = 5001 \pm \sqrt{(5001)^2 - 1} \approx 5001 \pm \left(5001 - \dfrac{1}{10002}\right)$.

So $\lambda \approx 10002$ and $1/10002$ and $\sigma \approx 100.01$ and $1/100.01$. Check $\sigma_1\sigma_2 \approx 1 = \det A$.

21 The singular values of $A + I$ are not $\sigma_j + 1$. They come from eigenvalues of $(A+I)^{\mathrm{T}}(A+I)$. Test the diagonal matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$.

22 Since Q_1 and U are orthogonal, so is Q_1U. (check: $(Q_1U)^{\mathrm{T}}(Q_1U) = U^{\mathrm{T}}Q_1^{\mathrm{T}}Q_1U = U^{\mathrm{T}}U = I$.) So the SVD of the matrix $Q_1AQ_2^{\mathrm{T}}$ is just $Q_1U\Sigma V^{\mathrm{T}}Q_2^{\mathrm{T}} = (Q_1U)\Sigma(Q_2V)^{\mathrm{T}}$ and Σ is the same as for A. The matrices A and $Q_1AQ_2^{\mathrm{T}}$ and Σ are all "*isometric*" = sharing the same Σ.

23 The singular values of Q are the eigenvalues of $Q^{\mathrm{T}}Q = I$ (therefore all 1's).

24 (a) From $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} = 3x_1^2 + 2x_1x_2 + 3x_2^2$ you can see that $S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Its eigenvalues are 4 and 2. The maximum of $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ is $\mathbf{4}$.

(b) The 1 by 2 matrix $A = \begin{bmatrix} 1 & 4 \end{bmatrix}$ leads to $\dfrac{||A\boldsymbol{x}||^2}{||\boldsymbol{x}||^2} = \dfrac{(x_1 + 4x_2)^2}{x_1^2 + x_2^2}$. The maximum value is $\sigma_1^2(A)$. For this matrix $A = \begin{bmatrix} 1 & 4 \end{bmatrix}$ that singular value squared is $\sigma_1^2 = 17$. This is because $AA^{\mathrm{T}} = \begin{bmatrix} 17 \end{bmatrix}$ and also $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 4 \\ 4 & 16 \end{bmatrix}$ has $\lambda = 17$ and 0.

25 The minimum value of $\dfrac{\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ is the *smallest* eigenvalue of S. The eigenvector is the minimizing $\boldsymbol{x}$. That eigenvector gives $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\lambda_{\text{min}}\boldsymbol{x}$.

Since $\dfrac{||A\boldsymbol{x}||^2}{||\boldsymbol{x}||^2} = \dfrac{\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ we see that the minimizing $\boldsymbol{x}$ is an **eigenvector of $\boldsymbol{A^{\mathrm{T}}A}$** (and not usually an eigenvector of A).

26 From $AV = U\Sigma$ we know that $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ = first column of V goes to $2 \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ = first column of $U\Sigma$. Similarly the second column $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ goes to $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$. The two outputs are orthogonal and they are the axes of an ellipse. With $\theta = 30°$ those axes are $\begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$ going out from $(0, 0)$ at $30°$ and $\frac{1}{2}\begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix}$ going out at $120°$. Comparing to the picture in Section 7.4, the first step would be a reflection (not a rotation), then a stretch by factors 2 and 1, then a $30°$ rotation.

27 Start from $A = U\Sigma V^{\mathrm{T}}$. The columns of U are a basis for the column space of A, and so are the columns of C, so $U = CF$ for some invertible r by r matrix F.

Similarly the columns of V are a basis for the row space of A and so are the columns of B, so $V = BG$ for some invertible r by r matrix G.

Then $A = U\Sigma V^{\mathrm{T}} = C(F\Sigma G^{\mathrm{T}})B^{\mathrm{T}} = CMB^{\mathrm{T}}$ and $M = F\Sigma G^{\mathrm{T}}$ is r by r and invertible.

Problem Set 7.3, page 391

1 The row averages of A_0 are 3 and 0. Therefore

$$A = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\mathrm{T}}}{4} = \frac{1}{4}\begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues of S are $\lambda_1 = \frac{10}{4}$ and $\lambda_2 = \frac{4}{4} = 1$. The top eigenvector of S is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

I think this means that a **vertical line** is closer to the five points $(2, -1), \ldots, (-2, -1)$ in the columns of A than any other line through the origin $(0, 0)$.

2 Now the row averages of A_0 are $\frac{1}{2}$ and 2. Therefore

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & 1 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\mathrm{T}}}{5} = \frac{1}{5}\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 4 \end{bmatrix}.$$

Again the rows of A are accidentally orthogonal (because of the special patterns of those rows). This time the top eigenvector of S is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. So a **horizontal line** is closer to the six points $\left(\frac{1}{2}, -1\right), \ldots, \left(-\frac{1}{2}, -1\right)$ from the columns of A than any other line through the center point $(0, 0)$.

3 $A_0 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 2 \end{bmatrix}$ has row averages 2 and so $A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}$. Then $S = \dfrac{1}{2}AA^{\mathrm{T}} = \dfrac{1}{2}\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$.

Then trace $(S) = \frac{1}{2}\,(8)$ and $\det(S) = \left(\frac{1}{2}\right)^2(3)$. The eigenvalues $\lambda(S)$ are $\frac{1}{2}$ times the roots of $\lambda^2 - 8\lambda + 3 = 0$. Those roots are $4 \pm \sqrt{16-3}$. Then the σ's are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$.

4 This matrix A with orthogonal rows has $S = \dfrac{AA^{\mathrm{T}}}{n-1} = \dfrac{1}{3}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

With λ's in descending order $\lambda_1 > \lambda_2 > \lambda_3$, the eigenvectors are $(0, 1, 0)$ and $(0, 0, 1)$ and $(1, 0, 0)$. The first eigenvector shows the $\boldsymbol{u}_1$ direction. Combined with the second eigenvector $\boldsymbol{u}_2$, the best plane is the yz plane.

These problems are examples where the sample **correlation matrix** (rescaling S so all its diagonal entries are 1) would be the identity matrix. If we think the original scaling is not meaningful and the rows should have the same length, then there is no reason to choose $\boldsymbol{u}_1 = (0, 1, 0)$ from the 8 in row 2.

5 The correlation matrix DSD which has 1's on the diagonal is

$$DSD = \begin{bmatrix} \frac{1}{2} & & \\ & \frac{1}{2} & \\ & & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & & \\ & \frac{1}{2} & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}.$$

6 Working with letters instead of numbers, the correlation matrix $C = DSD$ is

$$\begin{bmatrix} 1 & c_{12} & c_{13} \\ c_{12} & 1 & c_{23} \\ c_{13} & c_{23} & 1 \end{bmatrix} \text{ with } c_{12} = \frac{S_{12}}{\sigma_1\sigma_2} \text{ and } c_{13} = \frac{S_{13}}{\sigma_1\sigma_3} \text{ and } c_{23} = \frac{S_{23}}{\sigma_2\sigma_3}.$$

$$\text{Then } D = \begin{bmatrix} 1/\sigma_1 & & \\ & 1/\sigma_2 & \\ & & 1/\sigma_3 \end{bmatrix} \text{ gives } DSD = C.$$

7 From each row of A_0, subtract the average of that row (the average grade for that course) from the 10 grades in that row. This produces the centered matrix A. Then the sample covariance matrix is $S = \frac{1}{9}AA^{\mathrm{T}}$. The leading eigenvector of the 5 by 5 matrix S tells the weights on the 5 courses to produce the "*eigencourse*". This is the course whose grades have the most information (the greatest variance).

If a course gives everyone an A, the variance is zero and that course is not included in the eigencourse. We are looking for most information not best grade.

Problem Set 7.4, page 398

1 $A^{\mathrm{T}}A = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$ has $\lambda = 50$ and 0, $\boldsymbol{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\boldsymbol{v}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$; $\sigma_1 = \sqrt{50}$.

2 Orthonormal bases: $\boldsymbol{v}_1$ for row space, $\boldsymbol{v}_2$ for nullspace, $\boldsymbol{u}_1$ for column space, $\boldsymbol{u}_2$ for $N(A^{\mathrm{T}})$. All matrices with those four subspaces are multiples cA, since the subspaces are just lines. Normally many more matrices share the same 4 subspaces. (For example, all n by n invertible matrices share $\mathbf{R}^n$ as their column space.)

3 $A = QS = \frac{1}{\sqrt{50}}\begin{bmatrix} 7 & -1 \\ 1 & 7 \end{bmatrix} \frac{1}{\sqrt{50}}\begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$. S is semidefinite because A is singular.

4 $A^+ = V\begin{bmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}} = \frac{1}{50}\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$; $A^+A = \begin{bmatrix} .2 & .4 \\ .4 & .8 \end{bmatrix}$, $AA^+ = \begin{bmatrix} .1 & .3 \\ .3 & .9 \end{bmatrix}$.

5 $A^{\mathrm{T}}A = \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$ has $\lambda = 18$ and 2, $\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\sigma_1 = \sqrt{18}$ and $\sigma_2 = \sqrt{2}$.

6 $AA^{\mathrm{T}} = \begin{bmatrix} 18 & 0 \\ 0 & 2 \end{bmatrix}$ has $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The same $\sqrt{18}$ and $\sqrt{2}$ go into Σ.

7 $\begin{bmatrix} \sigma_1\boldsymbol{u}_1 & \sigma_2\boldsymbol{u}_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \boldsymbol{v}_2^{\mathrm{T}} \end{bmatrix} = \sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}} + \sigma_2\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$. In general this is $\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^{\mathrm{T}}$.

8 $A = U\Sigma V^{\mathrm{T}}$ splits into QK (polar): $Q = UV^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $K = V\Sigma V^{\mathrm{T}} = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$.

9 A^+ is A^{-1} because A is invertible. Pseudoinverse equals inverse when A^{-1} exists!

10 $A^{\mathrm{T}}A = \begin{bmatrix} 9 & 12 & 0 \\ 12 & 16 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ has $\lambda = 25, 0, 0$ and $\boldsymbol{v}_1 = \begin{bmatrix} .6 \\ .8 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} .8 \\ -.6 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Here $A = [\,3\;\;4\;\;0\,]$ has rank 1 and $AA^{\mathrm{T}} = [\,25\,]$ and $\sigma_1 = 5$ is the only singular value in $\Sigma = [\,5\;\;0\;\;0\,]$.

11 $A = [\,1\,]\,[\,5\;\;0\;\;0\,]V^{\mathrm{T}}$ and $A^+ = V\begin{bmatrix} .2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .12 \\ .16 \\ 0 \end{bmatrix}$; $A^+A = \begin{bmatrix} .36 & .48 & 0 \\ .48 & .64 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $AA^+ = [\,1\,]$

12 The zero matrix has no pivots or singular values. Then $\Sigma =$ same 2 by 3 zero matrix and the pseudoinverse is the 3 by 2 zero matrix.

13 If $\det A = 0$ then rank$(A) < n$; thus rank$(A^+) < n$ and $\det A^+ = 0$.

14 This problem explains why the matrix A transforms the circle of unit vectors $||\boldsymbol{x}|| = 1$ into an **ellipse** of vectors $\boldsymbol{y} = A\boldsymbol{x}$. The reason is that $\boldsymbol{x} = A^{-1}\boldsymbol{y}$ and the vectors with $||A^{-1}\boldsymbol{y}|| = 1$ do lie on an ellipse:

$$||A^{-1}\,\boldsymbol{y}||^2 = 1 \quad \text{is} \quad \boldsymbol{y}^{\mathrm{T}}\,(A^{-1})^{\mathrm{T}}\,A^{-1}\,\boldsymbol{y} = 1 \quad \text{or} \quad \boldsymbol{y}^{\mathrm{T}}\,(AA^{\mathrm{T}})^{-1}\,\boldsymbol{y} = 1.$$

That matrix $(AA^{\mathrm{T}})^{-1}$ is *symmetric positive definite* (A is assumed invertible).

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{gives} \quad AA^{\mathrm{T}} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad (AA^{\mathrm{T}})^{-1} = \frac{1}{9}\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}.$$

So the ellipse $||A^{-1}\,\boldsymbol{y}||^2 = 1$ of outputs $\boldsymbol{y} = A\boldsymbol{x}$ has equation $5y_1^2 - 8y_1y_2 + 5y_2^2 = 9$.

The singular values of this positive definite A are its eigenvalues 3 and 1.

The ellipse $||A^{-1}\,\boldsymbol{y}|| = 1$ has semi-axes of lengths $1/3$ and $1/1$.

15 (a) $A^{\mathrm{T}}A$ is singular (b) This $\boldsymbol{x}^+$ in the row space does give $A^{\mathrm{T}}A\boldsymbol{x}^+ = A^{\mathrm{T}}\boldsymbol{b}$ (c) If $(1,-1)$ in the nullspace of A is added to $\boldsymbol{x}^+$, we get another solution to $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$. But this $\widehat{\boldsymbol{x}}$ is longer than $\boldsymbol{x}^+$ because the added part is orthogonal to $\boldsymbol{x}^+$ in the row space and $||\widehat{\boldsymbol{x}}||^2 = ||\boldsymbol{x}^+||^2 + ||\text{added part from nullspace}||^2$.

16 $\boldsymbol{x}^+$ in the row space of A is perpendicular to $\widehat{\boldsymbol{x}} - \boldsymbol{x}^+$ in the nullspace of $A^{\mathrm{T}}A =$ nullspace of A. The right triangle has $c^2 = a^2 + b^2$.

17 $AA^+\boldsymbol{p} = \boldsymbol{p},\ AA^+\boldsymbol{e} = \boldsymbol{0},\ A^+A\boldsymbol{x}_r = \boldsymbol{x}_r,\ A^+A\boldsymbol{x}_n = \boldsymbol{0}$.

18 $A^+ = V\Sigma^+U^{\mathrm{T}}$ is $\frac{1}{5}[\,.6\ \ .8\,] = [\,.12\ \ .16\,]$ and $A^+A = [\,1\,]$ and $AA^+ = \begin{bmatrix} .36 & .48 \\ .48 & .64 \end{bmatrix} =$ projection.

19 L is determined by ℓ_{21}. Each eigenvector in X is determined by one number. The counts are $1+3$ for LU, $1+2+1$ for LDU, $1+3$ for QR (notice 1 rotation angle), $1+2+1$ for $U\Sigma V^{\mathrm{T}}$, $2+2+0$ for $X\Lambda X^{-1}$.

20 LDL^{T} and $Q\Lambda Q^{\mathrm{T}}$ are determined by $\mathbf{1+2+0}$ numbers because A is *symmetric*. **Note** Problem 20 should have referred to Problem 19 not 18.

21 Check the formula for A^+A using A^+ and A:

$$A^+A = \left(\sum_1^r \frac{\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{T}}}{\sigma_i}\right)\left(\sum_1^r \sigma_j\boldsymbol{u}_j\boldsymbol{v}_j^{\mathrm{T}}\right) = \sum_1^r \boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{T}} \text{ because } \boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_j = 0 \text{ when } i \neq j$$

Then every $\boldsymbol{u}_i^{\mathrm{T}}\,\boldsymbol{u}_i = 1$ (unit vector) so $A^+A = \sum_1^r \boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}$ is correct.

$$\text{Similarly } AA^+ = \left(\sum_1^r \sigma_j\,\boldsymbol{u}_j\,\boldsymbol{v}_j^{\mathrm{T}}\right)\left(\sum_1^r \frac{\boldsymbol{v}_i\,\boldsymbol{u}_i^{\mathrm{T}}}{\sigma_i}\right) = \sum_1^r \boldsymbol{u}_i\,\boldsymbol{u}_i^{\mathrm{T}}.$$

22 $M = \begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} A\boldsymbol{v} \\ A^{\mathrm{T}}\boldsymbol{u} \end{bmatrix} = \sigma \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$. Thus $\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$ is an eigenvector.

The singular values of A are *eigenvalues* of this block matrix.

Problem Set 8.1, page 407

1 With $\boldsymbol{w} = \mathbf{0}$ linearity gives $T(\boldsymbol{v} + \mathbf{0}) = T(\boldsymbol{v}) + T(\mathbf{0})$. Thus $T(\mathbf{0}) = \mathbf{0}$. With $c = -1$ linearity gives $T(-\mathbf{0}) = -T(\mathbf{0})$. This is a second proof that $T(\mathbf{0}) = \mathbf{0}$.

2 Combining $T(c\boldsymbol{v}) = cT(\boldsymbol{v})$ and $T(d\boldsymbol{w}) = dT(\boldsymbol{w})$ with addition gives $T(c\boldsymbol{v} + d\boldsymbol{w}) = cT(\boldsymbol{v}) + dT(\boldsymbol{w})$. Then one more addition gives $cT(\boldsymbol{v}) + dT(\boldsymbol{w}) + eT(\boldsymbol{u})$.

3 (d) is not linear.

4 (a) $S(T(\boldsymbol{v})) = \boldsymbol{v}$ (b) $S(T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2)) = S(T(\boldsymbol{v}_1)) + S(T(\boldsymbol{v}_2))$.

5 Choose $\boldsymbol{v} = (1, 1)$ and $\boldsymbol{w} = (-1, 0)$. Then $T(\boldsymbol{v}) + T(\boldsymbol{w}) = (\boldsymbol{v} + \boldsymbol{w})$ but $T(\boldsymbol{v} + \boldsymbol{w}) = (0, 0)$.

6 (a) $T(\boldsymbol{v}) = \boldsymbol{v}/\|\boldsymbol{v}\|$ does not satisfy $T(\boldsymbol{v} + \boldsymbol{w}) = T(\boldsymbol{v}) + T(\boldsymbol{w})$ or $T(c\boldsymbol{v}) = cT(\boldsymbol{v})$ (b) and (c) are linear (d) satisfies $T(c\boldsymbol{v}) = cT(\boldsymbol{v})$.

7 (a) $T(T(\boldsymbol{v})) = \boldsymbol{v}$ (b) $T(T(\boldsymbol{v})) = \boldsymbol{v} + (2, 2)$ (c) $T(T(\boldsymbol{v})) = -\boldsymbol{v}$ (d) $T(T(\boldsymbol{v})) = T(\boldsymbol{v})$.

8 (a) The range of $T(v_1, v_2) = (v_1 - v_2, 0)$ is the line of vectors $(c, 0)$. The nullspace is the line of vectors (c, c). (b) $T(v_1, v_2, v_3) = (v_1, v_2)$ has Range $\mathbf{R}^2$, kernel $\{(0, 0, \boldsymbol{v}_3)\}$ (c) $T(\boldsymbol{v}) = \mathbf{0}$ has Range $\{\mathbf{0}\}$, kernel $\mathbf{R}^2$ (d) $T(v_1, v_2) = (v_1, v_1)$ has Range = multiples of $(1, 1)$, kernel = multiples of $(1, -1)$.

9 If $T(v_1, v_2, v_3) = (v_2, v_3, v_1)$ then $T(T(\boldsymbol{v})) = (\boldsymbol{v}_3, \boldsymbol{v}_1, \boldsymbol{v}_2)$; $T^3(\boldsymbol{v}) = \boldsymbol{v}$; $T^{100}(\boldsymbol{v}) = T(\boldsymbol{v})$.

10 (a) $T(1, 0) = \mathbf{0}$ (b) $(0, 0, 1)$ is not in the range (c) $T(0, 1) = \mathbf{0}$.

11 For multiplication $T(\boldsymbol{v}) = A\boldsymbol{v}$: $\boldsymbol{V} = \mathbf{R}^n$, $\boldsymbol{W} = \mathbf{R}^m$; the outputs fill the column space; $\boldsymbol{v}$ is in the kernel if $A\boldsymbol{v} = \mathbf{0}$.

12 $T(\boldsymbol{v}) = (4, 4)$; $(2, 2)$; $(2, 2)$; if $\boldsymbol{v} = (a, b) = b(1, 1) + \frac{a-b}{2}(2, 0)$ then $T(\boldsymbol{v}) = b(2, 2) + (0, 0)$.

13 The *distributive law* (page 69) gives $A(M_1 + M_2) = AM_1 + AM_2$. The *distributive law* over c's gives $A(cM) = c(AM)$.

14 This A is invertible. Multiply $AM = 0$ and $AM = B$ by A^{-1} to get $M = 0$ and $M = A^{-1}B$. The kernel contains only the zero matrix $M = 0$.

15 This A is *not* invertible. $AM = I$ is impossible. $A\begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. The range contains only matrices AM whose columns are multiples of $(1, 3)$.

16 No matrix A gives $A\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. To professors: Linear transformations on matrix space come from **4** by **4** matrices. Those in Problems 13–15 were special.

17 For $T(M) = MT$ (a) $T^2 = I$ is True (b) True (c) True (d) False.

18 $T(I) = 0$ but $M = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = T(M)$; these M's fill the range. Every $M = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$ is in the kernel. Notice that $\dim$ (range) $+ \dim$ (kernel) $= 3 + 1 = \dim$ (input space of 2 by 2 M's).

19 $T(T^{-1}(M)) = M$ so $T^{-1}(M) = A^{-1}MB^{-1}$.

20 (a) Horizontal lines stay horizontal, vertical lines stay vertical (b) House squashes onto a line (c) Vertical lines stay vertical because $T(1, 0) = (a_{11}, 0)$.

21 $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ doubles the width of the house. $A = \begin{bmatrix} .7 & .7 \\ .3 & .3 \end{bmatrix}$ *projects* the house (since $A^2 = A$ from trace $= 1$ and $\lambda = 0, 1$). The projection is onto the column space of $A =$ line through $(.7, .3)$. $U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ will *shear* the house horizontally: The point at (x, y) moves over to $(x + y, y)$.

22 (a) $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ with $d > 0$ leaves the house AH sitting straight up (b) $A = 3I$ expands the house by 3 (c) $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ rotates the house.

23 $T(\boldsymbol{v}) = -\boldsymbol{v}$ rotates the house by 180° around the origin. Then the affine transformation $T(\boldsymbol{v}) = -\boldsymbol{v} + (1, 0)$ shifts the rotated house one unit to the right.

24 A code to add a chimney will be gratefully received!

25 This code needs a correction: add spaces between $-10\ 10\ -10\ 10$

26 $\begin{bmatrix}1 & 0\\0 & .1\end{bmatrix}$ compresses vertical distances by 10 to 1. $\begin{bmatrix}.5 & .5\\.5 & .5\end{bmatrix}$ projects onto the 45° line. $\begin{bmatrix}.5 & .5\\-.5 & .5\end{bmatrix}$ rotates by 45° clockwise and contracts by a factor of $\sqrt{2}$ (the columns have length $1/\sqrt{2}$). $\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}$ has determinant -1 so the house is "flipped and sheared." One way to see this is to factor the matrix as LDL^{T}:

$$\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}1 & \\ & -1\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = \text{(shear) (flip left-right) (shear).}$$

27 Also **30** emphasizes that circles are transformed to ellipses (see figure in Section 6.7).

28 A code that adds two eyes and a smile will be included here with public credit given!

29 (a) $ad - bc = 0$ (b) $ad - bc > 0$ (c) $|ad - bc| = 1$. If vectors to two corners transform to themselves then by linearity $T = I$. (Fails if one corner is $(0, 0)$.)

30 Linear transformations keep straight lines straight! And two parallel edges of a square (edges differing by a fixed $\boldsymbol{v}$) go to two parallel edges (edges differing by $T(\boldsymbol{v})$). So the output is a parallelogram.

Problem Set 8.2, page 418

1 For $S\boldsymbol{v} = d^2\boldsymbol{v}/dx^2$
$\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4 = 1, x, x^2, x^3$
$S\boldsymbol{v}_1 = S\boldsymbol{v}_2 = \boldsymbol{0}, S\boldsymbol{v}_3 = 2\boldsymbol{v}_1, S\boldsymbol{v}_4 = 6\boldsymbol{v}_2;$
The matrix for S is $B = \begin{bmatrix}0 & 0 & 2 & 0\\0 & 0 & 0 & 6\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$.

2 $Sv = d^2v/dx^2 = 0$ for linear functions $\boldsymbol{v}(\boldsymbol{x}) = a + bx$. All $(a, b, 0, 0)$ are in the nullspace of the second derivative matrix B.

3 (Matrix A)$^2 = B$ when (transformation T)$^2 = S$ and output basis = input basis.

4 The third derivative matrix has **6** in the $(1, 4)$ position; since the third derivative of x^3 is 6. This matrix also comes from AB. The fourth derivative of a cubic is zero, and B^2 is the zero matrix.

5 $T(\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3) = 2\boldsymbol{w}_1 + \boldsymbol{w}_2 + 2\boldsymbol{w}_3$; A times $(1, 1, 1)$ gives $(2, 1, 2)$.

6 $\boldsymbol{v} = c(\boldsymbol{v}_2 - \boldsymbol{v}_3)$ gives $T(\boldsymbol{v}) = \boldsymbol{0}$; nullspace is $(0, c, -c)$; solutions $(1, 0, 0) + (0, c, -c)$.

7 $(1, 0, 0)$ is not in the column space of the matrix A, and $\boldsymbol{w}_1$ is not in the range of the linear transformation T. Key point: *Column space* of matrix matches *range* of transformation.

8 We don't know $T(\boldsymbol{w})$ unless the $\boldsymbol{w}$'s are the same as the $\boldsymbol{v}$'s. In that case the matrix is A^2.

9 Rank of $A = 2 =$ dimension of the *range* of T. The outputs $A\boldsymbol{v}$ (column space) match the outputs $T(\boldsymbol{v})$ (the range of T). The "output space" $\boldsymbol{W}$ is like $\mathbf{R}^m$: it contains all outputs but may not be filled up.

10 The matrix for T is $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. For the output $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ choose input $\boldsymbol{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = A^{-1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. This means: For the output $\boldsymbol{w}_1$ choose the input $\boldsymbol{v}_1 - \boldsymbol{v}_2$.

11 $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ so $T^{-1}(\boldsymbol{w}_1) = \boldsymbol{v}_1 - \boldsymbol{v}_2, T^{-1}(\boldsymbol{w}_2) = \boldsymbol{v}_2 - \boldsymbol{v}_3, T^{-1}(\boldsymbol{w}_3) = \boldsymbol{v}_3$. The columns of A^{-1} describe T^{-1} from $\boldsymbol{W}$ back to $\boldsymbol{V}$. The only solution to $T(\boldsymbol{v}) = 0$ is $\boldsymbol{v} = 0$.

12 (c) $T^{-1}(T(\boldsymbol{w}_1)) = \boldsymbol{w}_1$ is wrong because $\boldsymbol{w}_1$ is not generally in the input space.

13 (a) $T(\boldsymbol{v}_1) = \boldsymbol{v}_2, T(\boldsymbol{v}_2) = \boldsymbol{v}_1$ is its own inverse (b) $T(\boldsymbol{v}_1) = \boldsymbol{v}_1, T(\boldsymbol{v}_2) = 0$ has $T^2 = T$ (c) If $T^2 = I$ for part (a) and $T^2 = T$ for part (b), then T must be I.

14 (a) $\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$ = inverse of (a) (c) $A\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ must be $2A\begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

15 (a) $M = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$ transforms $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} r \\ t \end{bmatrix}$ and $\begin{bmatrix} s \\ u \end{bmatrix}$; this is the "easy" direction. (b) $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$ transforms in the inverse direction, back to the standard basis vectors. (c) $ad = bc$ will make the forward matrix singular and the inverse impossible.

16 $MW = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -1 \\ -7 & 3 \end{bmatrix}$.

17 Recording basis vectors is done by a *Permutation matrix*. Changing lengths is done by a *positive diagonal matrix*.

18 $(a, b) = (\cos\theta, -\sin\theta)$. Minus sign from $Q^{-1} = Q^{\mathrm{T}}$.

19 $M = \begin{bmatrix} 1 & 1 \\ 4 & 5 \end{bmatrix}$; $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$ = first column of M^{-1} = coordinates of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in basis $\begin{bmatrix} 1 \\ 4 \end{bmatrix}\begin{bmatrix} 1 \\ 5 \end{bmatrix}$.

20 $\boldsymbol{w}_2(x) = 1 - x^2$; $\boldsymbol{w}_3(x) = \frac{1}{2}(x^2 - x)$; $\boldsymbol{y} = 4\boldsymbol{w}_1 + 5\boldsymbol{w}_2 + 6\boldsymbol{w}_3$.

21 $\boldsymbol{w}$'s to $\boldsymbol{v}$'s: $\begin{bmatrix} 0 & 1 & 0 \\ .5 & 0 & -.5 \\ .5 & -1 & .5 \end{bmatrix}$. $\boldsymbol{v}$'s to $\boldsymbol{w}$'s: inverse matrix $= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$. *The key idea*: The matrix multiplies the coordinates in the $\boldsymbol{v}$ basis to give the coordinates in the $\boldsymbol{w}$ basis.

22 The 3 equations to match $4, 5, 6$ at $x = a, b, c$ are $\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$. This Vandermonde determinant equals $(b - a)(c - a)(c - b)$. So a, b, c must be distinct to have $\det \neq 0$ and one solution A, B, C.

23 The matrix M with these nine entries must be invertible.

24 Start from $A = QR$. Column 2 is $\boldsymbol{a}_2 = r_{12}\boldsymbol{q}_1 + r_{22}\boldsymbol{q}_2$. This gives $\boldsymbol{a}_2$ as a combination of the $\boldsymbol{q}$'s. So the change of basis matrix is $\boldsymbol{R}$.

25 Start from $A = LU$. Row 2 of A is ℓ_{21}(row 1 of U) $+ \ell_{22}$ (row 2 of U). The change of basis matrix is always *invertible*, because basis goes to basis.

26 The matrix for $T(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i$ is $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$.

27 If T is not invertible, $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)$ is not a basis. We couldn't choose $\boldsymbol{w}_i = T(\boldsymbol{v}_i)$.

28 (a) $\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$ gives $T(\boldsymbol{v}_1) = \boldsymbol{0}$ and $T(\boldsymbol{v}_2) = 3\boldsymbol{v}_1$. (b) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ gives $T(\boldsymbol{v}_1) = \boldsymbol{v}_1$ and $T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{v}_1$ (which combine into $T(\boldsymbol{v}_2) = \boldsymbol{0}$ by *linearity*).

29 $T(x, y) = (x, -y)$ is reflection across the x-axis. Then reflect across the y-axis to get $S(x, -y) = (-x, -y)$. Thus $ST = -I$.

30 S takes (x, y) to $(-x, y)$. $S(T(\boldsymbol{v})) = (\boldsymbol{-1}, \boldsymbol{2})$. $S(\boldsymbol{v}) = (-2, 1)$ and $T(S(\boldsymbol{v})) = (\boldsymbol{1}, \boldsymbol{-2})$.

31 Multiply the two reflections to get $\begin{bmatrix} \cos 2(\theta - \alpha) & -\sin 2(\theta - \alpha) \\ \sin 2(\theta - \alpha) & \cos 2(\theta - \alpha) \end{bmatrix}$ which is *rotation by* $2(\theta - \alpha)$. In words: $(1, 0)$ is reflected to have angle 2α, and that is reflected again to angle $2\theta - 2\alpha$.

32 The matrix for T in this basis is $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

33 Multiplying by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gives $T(\boldsymbol{v}_1) = A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = a\boldsymbol{v}_1 + c\boldsymbol{v}_3$. Similarly $T(\boldsymbol{v}_2) = a\boldsymbol{v}_2 + c\boldsymbol{v}_4$ and $T(\boldsymbol{v}_3) = b\boldsymbol{v}_1 + d\boldsymbol{v}_3$ and $T(\boldsymbol{v}_4) = b\boldsymbol{v}_2 + d\boldsymbol{v}_4$. The matrix for T in this basis is $\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$

34 False: We will not know $T(\boldsymbol{v})$ for *energy* $\boldsymbol{v}$ unless the n $\boldsymbol{v}$'s are linearly independent.

Problem Set 8.3, page 429

1 For this matrix J, the rank of $J - 3I$ is **3** so the dimension of the nullspace is only 1. There is only 1 independent eigenvector even though $\lambda = 3$ is a *double root* of $\det(J - \lambda I) = 0$: a repeated eigenvalue.

$$J = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}.$$

2 $J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is similar to all other 2 by 2 matrices A that have **2** zero eigenvalues but only **1** independent eigenvector. Then $J = B_1^{-1} A_1 B_1$ is the same as $B_1 J = A_1 B_1$:

$$B_1 J = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = A_1 B_1$$

$$B_2 J = \begin{bmatrix} 4 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -8 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 2 & 0 \end{bmatrix} = A_2 B_2$$

3 Every matrix is similar to its transpose (same eigenvalues, same multiplicity, more than that the same Jordan form). In this example

$$BJ = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = J^{\mathrm{T}} B.$$

4 Here J and K are *different* Jordan forms (block sizes $2, 2$ versus block sizes $3, 1$). Even though J and K have the same λ's (all zero) and same rank, J and K are *not similar*. If $BK = JB$ then B is *not invertible*:

$$BK = B \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & b_{11} & b_{12} & 0 \\ 0 & b_{21} & b_{22} & 0 \\ 0 & b_{31} & b_{32} & 0 \\ 0 & b_{41} & b_{42} & 0 \end{bmatrix}$$

$$JB = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} B = \begin{bmatrix} b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & 0 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Those right hand sides agree only if $b_{21} = 0, b_{41} = 0, b_{24} = 0, b_{44} = 0, b_{22} = 0$, $b_{42} = 0$. But then also $b_{11} = b_{22} = 0$ and $b_{31} = b_{42} = 0$. So the first column has $b_{11} = b_{21} = b_{31} = b_{41} = 0$ and B is not invertible.

5 If A^3 is the zero matrix then every eigenvalue of A is $\lambda = 0$ (because $A\boldsymbol{x} = \lambda\boldsymbol{x}$ leads to $\boldsymbol{0} = A^3\boldsymbol{x} = \lambda^3\boldsymbol{x}$). The Jordan form J will also have $J^3 = 0$ because $J = B^{-1}AB$ has $J^3 = B^{-1}A^3B = 0$. The blocks of J must become zero blocks in J^3. So those blocks of J can be

$$\begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{but not} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{pmatrix} \text{third power} \\ \text{is not zero} \end{pmatrix}$$

The rank of J (and A) is largest if every block is 3 by 3 of rank 2. Then rank $\leq \dfrac{2}{3}n$.

If $A^n =$ zero matrix then A is *not invertible* and rank $(A) < n$.

6 This question substitutes $u_1 = te^{\lambda t}$ and $u_2 = e^{\lambda t}$ to show that u_1, u_2 solve the system $\boldsymbol{u}' = J\boldsymbol{u}$:

$$\begin{aligned} u_1' &= \lambda u_1 + u_2 & \qquad e^{\lambda t} + t\lambda e^{\lambda t} &= \lambda(te^{\lambda t}) + (e^{\lambda t}) \\ u_2' &= \qquad \lambda u_2 & \lambda e^{\lambda t} &= \qquad \lambda(e^{\lambda t}) \,. \end{aligned}$$

Certainly $u_1 = 0$ and $u_2 = 1$ at $t = 0$, so we have the solution and it involves $te^{\lambda t}$ (the factor t appears because λ is a double eigenvalue of J).

7 The equation $u_{k+2} - 2\lambda u_{k+1} + \lambda^2 u_k$ is certainly solved by $u_k = \lambda^k$. But this is a second order equation and there must be another solution. In analogy with $te^{\lambda t}$ for the differential equation in 8.3.6, that second solution is $u_k = k\lambda^k$. Check:

$$(k+2)\lambda^{k+2} - 2\lambda(k+1)\lambda^{k+1} + \lambda^2(k)\lambda^k = \big[k+2-2(k+1)+k\big]\lambda^{k+2} = 0.$$

8 $\lambda^3 = 1$ has 3 roots $\lambda = 1$ and $e^{2\pi i/3}$ and $e^{4\pi i/3}$. Those are $\mathbf{1}, \boldsymbol{\lambda}, \boldsymbol{\lambda^2}$ if we take $\lambda = e^{2\pi i/3}$. The Fourier matrix is

$$F_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \lambda & \lambda^2 \\ 1 & \lambda^2 & \lambda^4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{8\pi i/3} \end{bmatrix}.$$

9 A 3 by 3 circulant matrix has the form on page 425:

$$C = \begin{bmatrix} c_0 & c_1 & c_2 \\ c_2 & c_0 & c_1 \\ c_1 & c_2 & c_0 \end{bmatrix} \text{ with } C \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (c_0 + c_1 + c_2) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \end{bmatrix} = (c_0 + c_1\lambda + c_2\lambda^2) \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \end{bmatrix} \qquad C \begin{bmatrix} 1 \\ \lambda^2 \\ \lambda^4 \end{bmatrix} = (c_0 + c_1\lambda^2 + c_2\lambda^4) \begin{bmatrix} 1 \\ \lambda^2 \\ \lambda^4 \end{bmatrix}.$$

Those 3 eigenvalues of C are exactly the 3 components of $F\boldsymbol{c} = F \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$,

10 The Fourier cosine coefficient c_3 is in formula (7) with integrals from $-\pi$ to π. Because f drops to zero at $x = L$, the integral stops at L:

$$a_3 = \frac{\int f(x)\cos 3x\,dx}{\int (\cos 3x)^2\,dx} = \frac{1}{\pi}\int_{-L}^{L} (1)(\cos 3x)\,dx = \frac{1}{3\pi}\Big[\sin 3x\Big]_{x=-L}^{x=L} = \frac{2\sin 3L}{3\pi}.$$

Note that we should have defined $f(x) = 0$ for $L < |x| < \pi$ (not 2π !).

Problem Set 9.1, page 436

1 (a)(b)(c) have sums 4, $-2+2i$, $2\cos\theta$ and products 5, $-2i$, 1. Note $(e^{i\theta})(e^{-i\theta}) = 1$.

2 In polar form these are $\sqrt{5}e^{i\theta}$, $5e^{2i\theta}$, $\frac{1}{\sqrt{5}}e^{-i\theta}$, $\sqrt{5}$.

3 The absolute values are $r = 10$, 100, $\frac{1}{10}$, and 100. The angles are θ, 2θ, $-\theta$ and -2θ.

4 $|z\times w| = 6$, $|z+w| \le 5$, $|z/w| = \frac{2}{3}$, $|z-w| \le 5$.

5 $a+ib = \frac{\sqrt{3}}{2}+\frac{1}{2}i$, $\frac{1}{2}+\frac{\sqrt{3}}{2}i$, i, $-\frac{1}{2}+\frac{\sqrt{3}}{2}i$; $w^{12} = 1$.

6 $1/z$ has absolute value $1/r$ and angle $-\theta$; $(1/r)e^{-i\theta}$ times $re^{i\theta}$ equals 1.

7 $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix}$ $\begin{bmatrix} ac-bd \\ bc+ad \end{bmatrix}$ **real part** **imaginary part** $\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$ is the matrix form of $(1+3i)(1-3i) = 10$.

8 $\begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$ gives complex matrix $=$ vector multiplication $(A_1 + iA_2)(\boldsymbol{x}_1 + i\boldsymbol{x}_2) = \boldsymbol{b}_1 + i\boldsymbol{b}_2$.

9 $2+i$; $(2+i)(1+i) = 1+3i$; $e^{-i\pi/2} = -i$; $e^{-i\pi} = -1$; $\frac{1-i}{1+i} = -i$; $(-i)^{103} = i$.

10 $z+\overline{z}$ is real; $z-\overline{z}$ is pure imaginary; $z\overline{z}$ is positive; $z/\overline{z}$ has absolute value 1.

11 $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ includes aI (which just adds a to the eigenvalues and $b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. So the eigenvectors are $\boldsymbol{x}_1 = (1, i)$ and $\boldsymbol{x}_2 = (1, -i)$. The eigenvalues are $\lambda_1 = a+bi$ and $\lambda_2 = a-bi$. We see $\overline{\boldsymbol{x}}_1 = \boldsymbol{x}_2$ and $\overline{\lambda}_1 = \lambda_2$ as expected for real matrices with complex eigenvalues.

12 (a) When $a = b = d = 1$ the square root becomes $\sqrt{4c}$; λ is complex if $c < 0$ (b) $\lambda = 0$ and $\lambda = a+d$ when $ad = bc$ (c) the λ's can be real and different.

13 Complex λ's when $(a+d)^2 < 4(ad-bc)$; write $(a+d)^2 - 4(ad-bc)$ as $(a-d)^2 + 4bc$ which is positive when $bc > 0$.

14 The symmetric block matrix has real eigenvalues; so $i\lambda$ is real and λ is pure imaginary.

15 (a) $2e^{i\pi/3}$, $4e^{2i\pi/3}$ (b) $e^{2i\theta}$, $e^{4i\theta}$ (c) $7e^{3\pi i/2}$, $49e^{3\pi i}$ $(=-49)$ (d) $\sqrt{50}e^{-\pi i/4}$, $50e^{-\pi i/2}$.

16 $r = 1$, angle $\frac{\pi}{2} - \theta$; multiply by $e^{i\theta}$ to get $e^{i\pi/2} = i$.

17 $a + ib = 1, i, -1, -i, \pm\frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}$. The root $\overline{w} = w^{-1} = e^{-2\pi i/8}$ is $1/\sqrt{2} - i/\sqrt{2}$.

18 1, $e^{2\pi i/3}$, $e^{4\pi i/3}$ are cube roots of 1. The cube roots of -1 are -1, $e^{\pi i/3}$, $e^{-\pi i/3}$. Altogether six roots of $z^6 = 1$.

19 $\cos 3\theta = \text{Re}[(\cos\theta + i\sin\theta)^3] = \cos^3\theta - 3\cos\theta\sin^2\theta$; $\sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta$.

20 If the conjugate $\overline{z} = 1/z$ then $|z|^2 = 1$ and z is any point $e^{i\theta}$ on the unit circle.

21 e^i is at angle $\theta = 1$ on the unit circle; $|i^e| = 1^e$; Infinitely many $i^e = e^{i(\pi/2 + 2\pi n)e}$.

22 (a) Unit circle (b) Spiral in to $e^{-2\pi}$ (c) Circle continuing around to angle $\theta = 2\pi^2$.

Problem Set 9.2, page 443

1 $\|\boldsymbol{u}\| = \sqrt{9} = 3$, $\|\boldsymbol{v}\| = \sqrt{3}$, $\boldsymbol{u}^{\mathrm{H}}\boldsymbol{v} = 3i + 2$, $\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u} = -3i + 2$ (this is the conjugate of $\boldsymbol{u}^{\mathrm{H}}\boldsymbol{v}$).

2 $A^{\mathrm{H}}A = \begin{bmatrix} 2 & 0 & 1+i \\ 0 & 2 & 1+i \\ 1-i & 1-i & 2 \end{bmatrix}$ and $AA^{\mathrm{H}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ are Hermitian matrices. They share the eigenvalues 4 and 2.

3 $\boldsymbol{z}$ = multiple of $(1+i, 1+i, -2)$; $A\boldsymbol{z} = \boldsymbol{0}$ gives $\boldsymbol{z}^{\mathrm{H}}A^{\mathrm{H}} = \boldsymbol{0}^{\mathrm{H}}$ so $\boldsymbol{z}$ (not $\overline{\boldsymbol{z}}$!) is orthogonal to all columns of A^{H} (using complex inner product $\boldsymbol{z}^{\mathrm{H}}$ times columns of A^{H}).

4 The four fundamental subspaces are now $\boldsymbol{C}(A)$, $\boldsymbol{N}(A)$, $\boldsymbol{C}(A^{\mathrm{H}})$, $\boldsymbol{N}(A^{\mathrm{H}})$. A^{H} ***and not*** A^{T}.

5 (a) $(A^{\mathrm{H}}A)^{\mathrm{H}} = A^{\mathrm{H}}A^{\mathrm{HH}} = A^{\mathrm{H}}A$ again (b) If $A^{\mathrm{H}}A\boldsymbol{z} = \boldsymbol{0}$ then $(\boldsymbol{z}^{\mathrm{H}}A^{\mathrm{H}})(A\boldsymbol{z}) = 0$. This is $\|A\boldsymbol{z}\|^2 = 0$ so $A\boldsymbol{z} = \boldsymbol{0}$. The nullspaces of A and $A^{\mathrm{H}}A$ are always the ***same***.

6 (a) False (c) False $A = Q = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (b) True: $-i$ is not an eigenvalue when $S = S^{\mathrm{H}}$.

7 cS is still Hermitian *for real* c; $(iS)^{\mathrm{H}} = -iS^{\mathrm{H}} = -iS$ is skew-Hermitian.

8 This P is invertible and unitary. $P^2 = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$, $P^3 = \begin{bmatrix} -i & & \\ & -i & \\ & & -i \end{bmatrix} = -iI$. Then $P^{100} = (-i)^{33}P = -iP$. The eigenvalues of P are the roots of $\lambda^3 = -i$, which are i and $ie^{2\pi i/3}$ and $ie^{4\pi i/3}$.

9 One unit eigenvector is certainly $\boldsymbol{x}_1 = (1,1,1)$ with $\lambda_1 = i$. The other eigenvectors are $\boldsymbol{x}_2 = (1, w, w^2)$ and $\boldsymbol{x}_3 = (1, w^2, w^4)$ with $w = e^{2\pi i/3}$. *The eigenvector matrix is the Fourier matrix* F_3. The eigenvectors of any unitary matrix like P are orthogonal (using the correct complex form $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$ of the inner product).

10 $(1,1,1)$, $(1, e^{2\pi i/3}, e^{4\pi i/3})$, $(1, e^{4\pi i/3}, e^{2\pi i/3})$ are orthogonal (complex inner product!) because P is an orthogonal matrix—and therefore its eigenvector matrix is unitary.

11 If $Q^{\mathrm{H}}Q = I$ then $Q^{-1}(Q^{\mathrm{H}})^{-1} = Q^{-1}(Q^{-1})^{\mathrm{H}} = I$ so Q^{-1} is also unitary. Also $(QU)^{\mathrm{H}}(QU) = U^{\mathrm{H}}Q^{\mathrm{H}}QU = U^{\mathrm{H}}U = I$ so QU is unitary.

12 Determinant = product of the eigenvalues *(all real)*. And $A = A^{\mathrm{H}}$ gives $\det A = \overline{\det A}$.

13 $(\boldsymbol{z}^{\mathrm{H}}A^{\mathrm{H}})(A\boldsymbol{z}) = \|A\boldsymbol{z}\|^2$ is positive unless $A\boldsymbol{z} = \boldsymbol{0}$. When A has independent columns this means $\boldsymbol{z} = \boldsymbol{0}$; so $A^{\mathrm{H}}A$ is positive definite.

14 $S = \dfrac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1+i \\ 1+i & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \dfrac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ -1-i & 1 \end{bmatrix}$.

15 $K =$(iA^{T} in Problem 14) $= \dfrac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1-i \\ 1-i & 1 \end{bmatrix} \begin{bmatrix} 2i & 0 \\ 0 & -i \end{bmatrix} \dfrac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1+i \\ -1+i & 1 \end{bmatrix}$; λ's are imaginary.

16 $U = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta - i\sin\theta \end{bmatrix} \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$ has $|\lambda| = 1$.

17 $U = \dfrac{1}{L} \begin{bmatrix} 1+\sqrt{3} & -1+i \\ 1+i & 1+\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \dfrac{1}{L} \begin{bmatrix} 1+\sqrt{3} & 1-i \\ -1-i & 1+\sqrt{3} \end{bmatrix}$ with $L^2 = 6+2\sqrt{3}$. Unitary means $|\lambda| = 1$. $U = U^{\mathrm{H}}$ gives real λ. Then trace zero gives $\lambda = 1$ and -1.

18 The $\boldsymbol{v}$'s are columns of a unitary matrix U, so U^{H} is U^{-1}. Then $\boldsymbol{z} = UU^{\mathrm{H}}\boldsymbol{z} =$ (multiply by columns) $= \boldsymbol{v}_1(\boldsymbol{v}_1^{\mathrm{H}}\boldsymbol{z}) + \cdots + \boldsymbol{v}_n(\boldsymbol{v}_n^{\mathrm{H}}\boldsymbol{z})$: a typical orthonormal expansion.

19 $\boldsymbol{z} = (1, i, -2)$ completes an orthogonal basis for $\mathbf{C}^3$. So does any $e^{i\theta}\boldsymbol{z}$.

20 $S = A + iB = (A + iB)^{\mathrm{H}} = A^{\mathrm{T}} - iB^{\mathrm{T}};$ A is symmetric but B is skew-symmetric.

21 $\mathbf{C}^n$ has dimension n; the columns of any unitary matrix are a basis. For example use the columns of iI: $(i, 0, \ldots, 0), \ldots, (0, \ldots, 0, i)$

22 $[\,1\,]$ and $[\,-1\,]$; any $[\,e^{i\theta}\,]$; $\begin{bmatrix} a & b + ic \\ b - ic & d \end{bmatrix}$; $\begin{bmatrix} w & e^{i\phi}\overline{z} \\ -z & e^{i\phi}\overline{w} \end{bmatrix}$ with $|w|^2 + |z|^2 = 1$ and any angle ϕ

23 The eigenvalues of A^{H} are *complex conjugates* of the eigenvalues of A: $\det(A - \lambda I) = 0$ gives $\det(A^{\mathrm{H}} - \overline{\lambda}I) = 0$.

24 $(I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})^{\mathrm{H}} = I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}$ and also $(I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}})^2 = I - 4\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}} + 4\boldsymbol{u}(\boldsymbol{u}^{\mathrm{H}}\boldsymbol{u})\boldsymbol{u}^{\mathrm{H}} = I$. The rank-1 matrix $\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}$ projects onto the line through $\boldsymbol{u}$.

25 Unitary $U^{\mathrm{H}}U = I$ means $(A^{\mathrm{T}} - iB^{\mathrm{T}})(A + iB) = (A^{\mathrm{T}}A + B^{\mathrm{T}}B) + i(A^{\mathrm{T}}B - B^{\mathrm{T}}A) = I$.

$A^{\mathrm{T}}A + B^{\mathrm{T}}B = I$ and $A^{\mathrm{T}}B - B^{\mathrm{T}}A = 0$ which makes the block matrix orthogonal.

26 We are given $A + iB = (A + iB)^{\mathrm{H}} = A^{\mathrm{T}} - iB^{\mathrm{T}}$. Then $A = A^{\mathrm{T}}$ and $B = -B^{\mathrm{T}}$. So that $\begin{bmatrix} \boldsymbol{A} & -\boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix}$ is symmetric.

27 $SS^{-1} = I$ gives $(S^{-1})^{\mathrm{H}}S^{\mathrm{H}} = I$. Therefore $(S^{-1})^{\mathrm{H}}$ is $(S^{\mathrm{H}})^{-1} = S^{-1}$ and S^{-1} is Hermitian.

28 If U has (complex) orthonormal columns, then $U^{\mathrm{H}}U = I$ and U is *unitary*. If those columns are eigenvectors of A, then $A = U\Lambda U^{-1} = U\Lambda U^{\mathrm{H}}$ is *normal*. The direct test for a normal matrix (which is $AA^{\mathrm{H}} = A^{\mathrm{H}}A$ because diagonals could be real!) and Λ^{H} surely commute:

$$AA^{\mathrm{H}} = (U\Lambda U^{\mathrm{H}})(U\Lambda^{\mathrm{H}}U^{\mathrm{H}}) = U(\Lambda\Lambda^{\mathrm{H}})U^{\mathrm{H}} = U(\Lambda^{\mathrm{H}}\Lambda)U^{\mathrm{H}} = (U\Lambda^{\mathrm{H}}U^{\mathrm{H}})(U\Lambda U^{\mathrm{H}}) = A^{\mathrm{H}}A.$$

An easy way to construct a normal matrix is $1 + i$ times a symmetric matrix. Or take $A = S + iT$ where the real symmetric S and T commute (Then $A^{\mathrm{H}} = S - iT$ and $AA^{\mathrm{H}} = A^{\mathrm{H}}A$).

Problem Set 9.3, page 450

1 Equation (3) (the FFT) is correct using $i^2 = -1$ in the last two rows and three columns.

2 $$F^{-1} = \begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 1 & 1 & & \\ 1 & i^2 & & \\ & & 1 & 1 \\ & & 1 & i^2 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 1 & & 1 & \\ & 1 & & 1 \\ 1 & & -1 & \\ & -i & & i \end{bmatrix} = \frac{1}{4}F^{\mathrm{H}}.$$

3 $$F = \begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ 1 & i^2 & & \\ & & 1 & 1 \\ & & 1 & i^2 \end{bmatrix} \begin{bmatrix} 1 & & 1 & \\ & 1 & & 1 \\ 1 & & -1 & \\ & -i & & i \end{bmatrix}$$ permutation last.

4 $$D = \begin{bmatrix} 1 & & \\ & e^{2\pi i/6} & \\ & & e^{4\pi i/6} \end{bmatrix}$$ (note 6 not 3) and $$F_3 \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{2\pi i/3} \end{bmatrix}.$$

5 $F^{-1}\boldsymbol{w} = \boldsymbol{v}$ and $F^{-1}\boldsymbol{v} = \boldsymbol{w}/4$. Delta vector $\leftrightarrow$ all-ones vector.

6 $$(F_4)^2 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \end{bmatrix}$$ and $(F_4)^4 = 16I$. Four transforms recover the signal!

7 $$\boldsymbol{c} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix} = F\boldsymbol{c}. \text{ Also } C = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \end{bmatrix} = FC.$$
Adding $c + C$ gives $(1,1,1,1)$ to $(4,0,0,0) = 4$ (delta vector).

8 $\boldsymbol{c} \to (1,1,1,1,0,0,0,0) \to (4,0,0,0,0,0,0,0) \to (4,0,0,0,4,0,0,0) = F_8\boldsymbol{c}$.
$\boldsymbol{C} \to (0,0,0,0,1,1,1,1) \to (0,0,0,0,4,0,0,0) \to (4,0,0,0,-4,0,0,0) = F_8\boldsymbol{C}$.

9 If $w^{64} = 1$ then w^2 is a 32nd root of 1 and $\sqrt{w}$ is a 128th root of 1: Key to FFT.

10 For every integer n, the nth roots of 1 add to zero. For even n, they cancel in pairs. For any n, use the geometric series formula $1 + w + \cdots + w^{n-1} = (w^n - 1)/(w - 1) = 0$. In particular for $n = 3$, $1 + (-1 + i\sqrt{3})/2 + (-1 - i\sqrt{3})/2 = 0$.

11 The eigenvalues of P are $1, i, i^2 = -1$, and $i^3 = -i$. Problem 11 displays the eigenvectors. And also $\det(P - \lambda I) = \lambda^4 - 1$.

12 $\Lambda = \text{diag}(1, i, i^2, i^3)$; $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and P^{T} lead to $\lambda^3 - 1 = 0$.

13 $e_1 = c_0 + c_1 + c_2 + c_3$ and $e_2 = c_0 + c_1 i + c_2 i^2 + c_3 i^3$; E contains the four eigenvalues of $C = FEF^{-1}$ because F contains the eigenvectors.

14 Eigenvalues $e_1 = 2 - 1 - 1 = 0$, $e_2 = 2 - i - i^3 = 2$, $e_3 = 2 - (-1) - (-1) = 4$, $e_4 = 2 - i^3 - i^9 = 2$. Just transform column 0 of C. Check trace $0 + 2 + 4 + 2 = 8$.

15 Diagonal E needs n multiplications, Fourier matrix F and F^{-1} need $\frac{1}{2}n \log_2 n$ multiplications each by the **FFT**. The total is much less than the ordinary n^2 for C times $\boldsymbol{x}$.

16 The row $1, \overline{w}^k, \overline{w}^{2k}, \ldots$ in $\overline{F}$ is the same as the row $1, w^{N-k}, w^{N-2k}, \ldots$ in F because $w^{N-k} = e^{(2\pi i/N)(N-k)}$ is $e^{2\pi i} e^{-(2\pi i/N)k} = 1$ times $\overline{w}^k$. So F and $\overline{F}$ have the **same rows in reversed order** (except for row 0 which is all ones).

17

0	000 reverses to $000 = \mathbf{0}$		
1	001 reverses to $100 = \mathbf{4}$		
2	010 reverses to $010 = \mathbf{2}$	**Now evens come before odds**!
3	011 reverses to $110 = \mathbf{6}$		
4	100 reverses to $001 = \mathbf{1}$		
5	101 reverses to $101 = \mathbf{5}$		
6	110 reverses to $011 = \mathbf{3}$		
7	111 reverses to $111 = \mathbf{7}$		

Problem Set 10.1, page 459

1 $A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$; nullspace contains $\begin{bmatrix} c \\ c \\ c \end{bmatrix}$; $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is not orthogonal to that nullspace.

2 $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ for $\boldsymbol{y} = (1, -1, 1)$; current along edge 1, edge 3, back on edge 2 (full loop).

3 Elimination on $b_1[\,A \quad \boldsymbol{b}\,] = \begin{bmatrix} -1 & 1 & 0 & b_1 \\ -1 & 0 & 1 & b_2 \\ 0 & -1 & 1 & b_3 \end{bmatrix}$ leads to $[\,U \quad \boldsymbol{c}\,] =$ $\begin{bmatrix} -1 & 1 & 0 & b_1 \\ 0 & -1 & 1 & b_2 - b_1 \\ 0 & 0 & 0 & b_3 - b_2 + b_1 \end{bmatrix}$. The nonzero rows of U come from edges 1 and 3 in a tree. The zero row comes from the loop (all 3 edges).

4 For the matrix in Problem 3, $A\boldsymbol{x} = \boldsymbol{b}$ is solvable for $\boldsymbol{b} = (1, 1, 0)$ and not solvable for $\boldsymbol{b} = (1, 0, 0)$. For solvable $\boldsymbol{b}$ (in the column space), $\boldsymbol{b}$ must be orthogonal to $\boldsymbol{y} = (1, -1, 1)$; that combination of rows is the zero row, and $b_1 - b_2 + b_3 = 0$ is the third equation after elimination.

5 Kirchhoff's Current Law $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{f}$ is solvable for $\boldsymbol{f} = (1, -1, 0)$ and not solvable for $\boldsymbol{f} = (1, 0, 0)$; $\boldsymbol{f}$ must be orthogonal to $(1, 1, 1)$ in the nullspace: $f_1 + f_2 + f_3 = 0$.

6 $A^{\mathrm{T}}A\boldsymbol{x} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix} = \boldsymbol{f}$ produces $\boldsymbol{x} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} c \\ c \\ c \end{bmatrix}$; potentials $\boldsymbol{x} = 1, -1, 0$ and currents $-A\boldsymbol{x} = 2,\ 1,\ -1$; $\boldsymbol{f}$ sends 3 units from node 2 into node 1.

7 $A^{\mathrm{T}} \begin{bmatrix} 1 & & \\ & 2 & \\ & & 2 \end{bmatrix} A = \begin{bmatrix} 3 & -1 & -2 \\ -1 & 3 & -2 \\ -2 & -2 & 4 \end{bmatrix}$; $\boldsymbol{f} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ yields $\boldsymbol{x} = \begin{bmatrix} 5/4 \\ 1 \\ 7/8 \end{bmatrix} +$ any $\begin{bmatrix} c \\ c \\ c \end{bmatrix}$; potentials $\boldsymbol{x} = \frac{5}{4}, 1, \frac{7}{8}$ and currents $-CA\boldsymbol{x} = \frac{1}{4}, \frac{3}{4}, \frac{1}{4}$.

8 $A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$ leads to $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$ solving $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$.

9 Elimination on $A\boldsymbol{x} = \boldsymbol{b}$ always leads to $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ in the zero rows of U and R: $-b_1 + b_2 - b_3 = 0$ and $b_3 - b_4 + b_5 = 0$ (those $\boldsymbol{y}$'s are from Problem 8 in the left nullspace). This is Kirchhoff's *Voltage* Law around the two *loops*.

10 The echelon form of A is $U = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ The nonzero rows of U keep edges 1, 2, 4. Other spanning trees from edges, 1, 2, 5; 1, 3, 4; 1, 3, 5; 1, 4, 5; 2, 3, 4; 2, 3, 5; 2, 4, 5.

11 $A^{\mathrm{T}}A = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$ diagonal entry = number of edges into the node
the trace is 2 times the number of nodes
off-diagonal entry $= -1$ if nodes are connected
$A^{\mathrm{T}}A$ is the **graph Laplacian**, $A^{\mathrm{T}}CA$ is **weighted** by C

12 (a) The nullspace and rank of $A^{\mathrm{T}}A$ and A are always the same (b) $A^{\mathrm{T}}A$ is always positive semidefinite because $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = \|A\boldsymbol{x}\|^2 \geq 0$. Not positive definite because rank is only 3 and $(1, 1, 1, 1)$ is in the nullspace (c) Real eigenvalues all ≥ 0 because positive semidefinite.

13 $A^{\mathrm{T}}CA\boldsymbol{x} = \begin{bmatrix} 4 & -2 & -2 & 0 \\ -2 & 8 & -3 & -3 \\ -2 & -3 & 8 & -3 \\ 0 & -3 & -3 & 6 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$ gives four potentials $\boldsymbol{x} = (\frac{5}{12}, \frac{1}{6}, \frac{1}{6}, 0)$
I grounded $x_4 = 0$ and solved for $\boldsymbol{x}$
currents $\boldsymbol{y} = -CA\boldsymbol{x} = (\frac{2}{3}, \frac{2}{3}, 0, \frac{1}{2}, \frac{1}{2})$

14 $A^{\mathrm{T}}CA\boldsymbol{x} = \mathbf{0}$ for $\boldsymbol{x} = c(1, 1, 1, 1) = (c, c, c, c)$. If $A^{\mathrm{T}}CA\boldsymbol{x} = \boldsymbol{f}$ is solvable, then $\boldsymbol{f}$ in the column space (= row space by symmetry) must be orthogonal to $\boldsymbol{x}$ in the nullspace: $f_1 + f_2 + f_3 + f_4 = 0$.

15 The number of loops in this connected graph is $n - m + 1 = \mathbf{7 - 7 + 1 = 1}$. What answer if the graph has two separate components (no edges between)?

16 Start from (4 nodes) $-$ (6 edges) $+$ (3 loops) $= 1$. If a new node connects to 1 old node, $5 - 7 + 3 = 1$. If the new node connects to 2 old nodes, a new loop is formed: $5 - 8 + 4 = 1$.

17 (a) 8 independent columns (b) $\boldsymbol{f}$ must be orthogonal to the nullspace so f's add to zero (c) Each edge goes into 2 nodes, 12 edges make diagonal entries sum to 24.

18 A *complete graph* has $5 + 4 + 3 + 2 + 1 = 15$ edges. With n nodes that count is $1 + \cdots + (n - 1) = n(n - 1)/2$. Tree has 5 edges.

Problem Set 10.2, page 472

1 Det $A_0^{\mathrm{T}} C_0 A_0 = \begin{bmatrix} c_1 + c_2 & -c_2 & 0 \\ -c_2 & c_2 + c_3 & -c_3 \\ 0 & -c_3 & c_3 + c_4 \end{bmatrix}$ is by direct calculation. Set $c_4 = 0$ to find det $A_1^{\mathrm{T}} C_1 A_1 = c_1 c_2 c_3$.

2 $(A_1^{\mathrm{T}} C_1 A_1)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1^{-1} & & \\ & c_2^{-1} & \\ & & c_3^{-1} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} =$

$$\begin{bmatrix} c_1^{-1} & c_1^{-1} & c_1^{-1} \\ c_1^{-1} & c_1^{-1} + c_2^{-1} & c_1^{-1} + c_2^{-1} \\ c_1^{-1} & c_1^{-1} + c_2^{-1} & c_1^{-1} + c_2^{-1} + c_3^{-1} \end{bmatrix}.$$

3 The rows of the free-free matrix in equation (9) add to $[\,0 \quad 0 \quad 0\,]$ so the right side needs $f_1 + f_2 + f_3 = 0$. $\boldsymbol{f} = (-1, 0, 1)$ gives $c_2 u_1 - c_2 u_2 = -1, c_3 u_2 - c_3 u_3 = -1, 0 = 0$. Then $\boldsymbol{u}_{\text{particular}} = (-c_2^{-1} - c_3^{-1}, -c_3^{-1}, 0)$. Add any multiple of $\boldsymbol{u}_{\text{nullspace}} = (1, 1, 1)$.

4 $\displaystyle\int -\frac{d}{dx}\left(c(x)\frac{du}{dx}\right) dx = -\left[c(x)\frac{du}{dx}\right]_0^1 = 0$ (bdry cond) so we need $\displaystyle\int f(x)\, dx = 0$.

5 $-\dfrac{dy}{dx} = f(x)$ gives $y(x) = C - \displaystyle\int_0^x f(t)dt$. Then $y(1) = 0$ gives $C = \displaystyle\int_0^1 f(t)dt$ and $y(x) = \displaystyle\int_x^1 f(t)dt$. If the load is $f(x) = 1$ then the displacement is $y(x) = 1 - x$.

6 Multiply $A_1^{\mathrm{T}} C_1 A_1$ as columns of A_1^{T} times c's times rows of A_1. The first 3 by 3 "*element matrix*" $c_1 E_1 = [\,1 \;\; 0 \;\; 0\,]^{\mathrm{T}} c_1 [\,1 \;\; 0 \;\; 0\,]$ has c_1 in the top left corner.

7 For 5 springs and 4 masses, the 5 by 4 A has two nonzero diagonals: all $a_{ii} = 1$ and $a_{i+1,i} = -1$. With $C = \mathsf{diag}(c_1, c_2, c_3, c_4, c_5)$ we get $K = A^{\mathrm{T}} C A$, symmetric tridiagonal with diagonal entries $K_{ii} = c_i + c_{i+1}$ and off-diagonals $K_{i+1,i} = -c_{i+1}$. With $C = I$ this K is the $-1, 2, -1$ matrix and $K(2, 3, 3, 2) = (1, 1, 1, 1)$ solves $K\boldsymbol{u} = \mathsf{ones}(4, 1)$. ($K^{-1}$ will solve $K\boldsymbol{u} = \mathsf{ones}(4)$.)

8 The solution to $-u'' = 1$ with $u(0) = u(1) = 0$ is $u(x) = \frac{1}{2}(x - x^2)$. At $x = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$ this gives $\boldsymbol{u} = 2, 3, 3, 2$ (discrete solution in Problem 7) times $(\Delta x)^2 = 1/25$.

9 $-u'' = mg$ has complete solution $u(x) = A + Bx - \frac{1}{2} mgx^2$. From $u(0) = 0$ we get $A = 0$. From $u'(1) = 0$ we get $B = mg$. Then $u(x) = \frac{1}{2} mg(2x - x^2)$ at $x = \frac{1}{3}, \frac{2}{3}, \frac{3}{3}$ equals $mg/6, 4mg/9, mg/2$. This $u(x)$ is *not* proportional to the discrete $u = (3mg, 5mg, 6mg)$ at the meshpoints. This imperfection is because the discrete problem uses a 1-sided difference, less accurate at the free end. Perfect accuracy is recovered by a centered difference (discussed on page 21 of my CSE textbook).

10 (added in later printing, changing **10-11** below into **11-12**). The solution in this fixed-fixed case is $(2.25, 2.50, 1.75)$ so the second mass moves furthest.

11 The two graphs of 100 points are "discrete parabolas" starting at $(0, 0)$: symmetric around 50 in the fixed-fixed case, ending with slope zero in the fixed-free case.

12 Forward/backward/centered for du/dx has a big effect because that term has the large coefficient. MATLAB: $E = \mathsf{diag}(\mathsf{ones}(6, 1), 1)$; $K = 64 * (2 * \mathsf{eye}(7) - E - E')$; $D = 80 * (E - \mathsf{eye}(7))$; $(K + D)\backslash\mathsf{ones}(7, 1)$; % forward; $(K - D')\backslash\mathsf{ones}(7, 1)$; % backward; $(K + D/2 - D'/2)\backslash\mathsf{ones}(7, 1)$; % centered is usually the best: more accurate

Problem Set 10.3, page 480

1 Eigenvalues $\lambda = 1$ and $.75$; $(A - I)\boldsymbol{x} = 0$ gives the steady state $\boldsymbol{x} = (.6, .4)$ with $A\boldsymbol{x} = \boldsymbol{x}$.

2 $A = \begin{bmatrix} .6 & -1 \\ .4 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & .75 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -.4 & .6 \end{bmatrix}$; $A^\infty = \begin{bmatrix} .6 & -1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -.4 & .6 \end{bmatrix} = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}$.

3 $\lambda = 1$ and $.8$, $\boldsymbol{x} = (1, 0)$; 1 and $-.8$, $\boldsymbol{x} = (\frac{5}{9}, \frac{4}{9})$; $1, \frac{1}{4}$, and $\frac{1}{4}$, $\boldsymbol{x} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

4 A^{T} always has the eigenvector $(1, 1, \ldots, 1)$ for $\lambda = 1$, because each row of A^{T} adds to 1. (Note again that many authors use row vectors multiplying Markov matrices. So they transpose our form of A.)

5 The steady state eigenvector for $\lambda = 1$ is $(0, 0, 1)$ = everyone is dead.

6 Add the components of $A\boldsymbol{x} = \lambda\boldsymbol{x}$ to find sum $s = \lambda s$. If $\lambda \neq 1$ the sum must be $s = 0$.

7 $(.5)^k \to 0$ gives $A^k \to A^\infty$; any $A = \begin{bmatrix} .6 + .4a & .6 - .6a \\ .4 - .4a & .4 + .6a \end{bmatrix}$ with $\begin{matrix} a \leq 1 \\ .4 + .6a \geq 0 \end{matrix}$

8 If P = cyclic permutation and $\boldsymbol{u}_0 = (1, 0, 0, 0)$ then $\boldsymbol{u}_1 = (0, 0, 1, 0)$; $\boldsymbol{u}_2 = (0, 1, 0, 0)$; $\boldsymbol{u}_3 = (1, 0, 0, 0)$; $\boldsymbol{u}_4 = \boldsymbol{u}_0$. The eigenvalues $1, i, -1, -i$ are all *on the unit circle*. This Markov matrix contains zeros; a *positive* matrix has *one* largest eigenvalue $\lambda = 1$.

9 M^2 is still nonnegative; $[\,1 \ \cdots \ 1\,]M = [\,1 \ \cdots \ 1\,]$ so multiply on the right by M to find $[\,1 \ \cdots \ 1\,]M^2 = [\,1 \ \cdots \ 1\,] \Rightarrow$ columns of M^2 add to 1.

10 $\lambda = 1$ and $a + d - 1$ from the trace; steady state is a multiple of $\boldsymbol{x}_1 = (b, 1 - a)$.

11 Last row $.2, .3, .5$ makes $A = A^{\mathrm{T}}$; rows also add to 1 so $(1, \ldots, 1)$ is also an eigenvector of A.

12 B has $\lambda = 0$ and $-.5$ with $\boldsymbol{x}_1 = (.3,\ .2)$ and $\boldsymbol{x}_2 = (-1, 1)$; A has $\lambda = 1$ so $A - I$ has $\lambda = 0$. $e^{-.5t}$ approaches zero and the solution approaches $c_1 e^{0t}\boldsymbol{x}_1 = c_1\boldsymbol{x}_1$.

13 $\boldsymbol{x} = (1, 1, 1)$ is an eigenvector when the row sums are equal; $A\boldsymbol{x} = (.9, .9, .9)$

14 $(I-A)(I+A+A^2+\cdots) = (I+A+A^2+\cdots)-(A+A^2+A^3+\cdots) = \boldsymbol{I}$. This says that $I + A + A^2 + \cdots$ is $(I-A)^{-1}$. When $A = \begin{bmatrix} 0 & .5 \\ 1 & 0 \end{bmatrix}$, $A^2 = \frac{1}{2}I$, $A^3 = \frac{1}{2}A$, $A^4 = \frac{1}{4}I$ and the series adds to $\begin{bmatrix} 1+\frac{1}{2}+\cdots & \frac{1}{2}+\frac{1}{4}+\cdots \\ 1+\frac{1}{2}+\cdots & 1+\frac{1}{2}+\cdots \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} = (I-A)^{-1}$.

15 The first two A's have $\lambda_{\max} < 1$; $\boldsymbol{p} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} 130 \\ 32 \end{bmatrix}$; $I - \begin{bmatrix} .5 & 1 \\ .5 & 0 \end{bmatrix}$ has no inverse.

16 $\lambda = 1$ (Markov), 0 (singular), .2 (from trace). Steady state $(.3, .3, .4)$ and $(30, 30, 40)$.

17 *No*, A has an eigenvalue $\lambda = 1$ and $(I-A)^{-1}$ does not exist.

18 The Leslie matrix on page 435 has $\det(A-\lambda I) = \det \begin{bmatrix} F_1-\lambda & F_2 & F_3 \\ P_1 & -\lambda & 0 \\ 0 & P_2 & -\lambda \end{bmatrix} = -\lambda^3 + F_1\lambda^2 + F_2P_1\lambda + F_3P_1P_2$. This is negative for large λ. It is positive at $\lambda = 1$ provided that $F_1 + F_2P_1 + F_3P_1P_2 > 1$. Under this key condition, $\det(A-\lambda I)$ must be zero at some λ between 1 and ∞. That eigenvalue means that the population grows (under this condition connecting F's and P's reproduction and survival rates).

19 Λ times $X^{-1}\Delta X$ has the same diagonal as $X^{-1}\Delta X$ times Λ because Λ is diagonal.

20 If $B > A > 0$ and $A\boldsymbol{x} = \lambda_{\max}(A)\boldsymbol{x} > 0$ then $B\boldsymbol{x} > \lambda_{\max}(A)\boldsymbol{x}$ and $\lambda_{\max}(B) > \lambda_{\max}(A)$.

of C = four components of $F\boldsymbol{c}$. Circulants are special!

Problem Set 10.4, page 489

1 Feasible set = line segment $(6,0)$ to $(0,3)$; minimum cost at $(6,0)$, maximum at $(0,3)$.

2 Feasible set has corners $(0,0), (6,0), (2,2), (0,6)$. Minimum cost $2x - y$ at $(6,0)$.

3 Only two corners $(4,0,0)$ and $(0,2,0)$; let $x_i \to -\infty$, $x_2 = 0$, and $x_3 = x_1 - 4$.

4 From $(0,0,2)$ move to $\boldsymbol{x} = (0,1,1.5)$ with the constraint $x_1 + x_2 + 2x_3 = 4$. The new cost is $3(1) + 8(1.5) = \$15$ so $r = -1$ is the reduced cost. The simplex method also checks $\boldsymbol{x} = (1,0,1.5)$ with cost $5(1) + 8(1.5) = \$17$; $r = 1$ means more expensive.

5 Cost $= 20$ at start $(4, 0, 0)$; keeping $x_1+x_2+2x_3 = 4$ move to $(3, 1, 0)$ with cost 18 and $r = -2$; or move to $(2, 0, 1)$ with cost 17 and $r = -3$. Choose x_3 as entering variable and move to $(0, 0, 2)$ with cost 14. Another step will reach $(0, 4, 0)$ with minimum cost 12.

6 If we reduce the Ph.D. cost to \$1 or \$2 (below the student cost of \$3), the job will go to the Ph.D. with cost vector $\boldsymbol{c} = (2, 3, 8)$ the Ph.D. takes 4 hours $(x_1 + x_2 + 2x_3 = 4)$ and charges \$8.

The teacher in the dual problem now has $y \le 2, y \le 3, 2y \le 8$ as constraints $A^{\mathrm{T}}\boldsymbol{y} \le \boldsymbol{c}$ on the charge of y per problem. So the dual has maximum at $y = 2$. The dual cost is also \$8 for 4 problems and maximum = minimum.

7 $\boldsymbol{x} = (2, 2, 0)$ is a corner of the feasible set with $x_1+x_2+2x_3 = 4$ and the new constraint $2x_1 + x_2 + x_3 = 6$. The cost of this corner is $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = (5, 3, 8) \cdot (2, 2, 0) = 16$. Is this the minimum cost?

Compute the reduced cost r *if* $x_3 = 1$ *enters* (x_3 was previously zero). The two constraint equations now require $x_1 = 3$ and $x_2 = -1$. With $\boldsymbol{x} = (3, -1, 1)$ the new cost is $3.5 - 1.3 + 1.8 = 20$. This is higher than 16, so the original $\boldsymbol{x} = (2, 2, 0)$ was optimal.

Note that $x_3 = 1$ led to $x_2 = -1$ and a negative x_2 is not allowed. If x_3 reduced the cost (it didn't) we would not have used as much as $x_3 = 1$.

8 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} \le \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{x} = (A^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{x} \le \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$. The first inequality needed $\boldsymbol{y} \ge 0$ and $A\boldsymbol{x}-\boldsymbol{b} \ge 0$.

Problem Set 10.5, page 494

1 $\int_0^{2\pi} \cos((j + k)x)\, dx = \left[\frac{\sin((j+k)x)}{j+k}\right]_0^{2\pi} = 0$ and similarly $\int_0^{2\pi} \cos((j - k)x)\, dx = 0$ Notice $j - k \ne 0$ in the denominator. If $j = k$ then $\int_0^{2\pi} \cos^2 jx\, dx = \pi$.

2 Three integral tests show that $1, x, x^2 - \frac{1}{3}$ are orthogonal on the interval $[-1,\ 1]$: $\int_{-1}^{1}(1)(x)\ \ dx = 0, \int_{-1}^{1}(1)(x^2 - \frac{1}{3})\ \ dx = 0, \int_{-1}^{1}(x)(x^2 - \frac{1}{3})\ \ dx = 0$. Then

$2x^2 = 2(x^2 - \frac{1}{3}) + 0(x) + \frac{2}{3}(1)$. Those coefficients $2, 0, \frac{2}{3}$ can come from integrating $f(x) = 2x^2$ times the 3 basis functions and dividing by their lengths squared—in other words using $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/a^{\mathrm{T}}a$ for functions (where $\boldsymbol{b}$ is $f(x)$ and $\boldsymbol{a}$ is 1 or x or $x^2 - \frac{1}{3}$) exactly as for vectors.

3 One example orthogonal to $\boldsymbol{v} = (1, \frac{1}{2}, \ldots)$ is $\boldsymbol{w} = (2, -1, 0, 0, \ldots)$ with $\|\boldsymbol{w}\| = \sqrt{5}$.

4 $\int_{-1}^{1}(1)(x^3 - cx)\,dx = 0$ and $\int_{-1}^{1}(x^2 - \frac{1}{3})(x^3 - cx)\,dx = 0$ for all c (odd functions). Choose c so that $\int_{-1}^{1} x(x^3 - cx)\,dx = [\frac{1}{5}x^5 - \frac{c}{3}x^3]_{-1}^{1} = \frac{2}{5} - c\frac{2}{3} = 0$. Then $c = \frac{3}{5}$.

5 The integrals lead to the Fourier coefficients $a_1 = 0,\ b_1 = 4/\pi,\ b_2 = 0$.

6 From eqn. (3) $a_k = 0$ and $b_k = 4/\pi k$ (*odd* k). The square wave has $\|f\|^2 = 2\pi$. Then eqn. (6) is $2\pi = \pi(16/\pi^2)(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots)$. That infinite series equals $\pi^2/8$.

7 The $-1, 1$ odd square wave is $f(x) = x/|x|$ for $0 < |x| < \pi$. Its Fourier series in equation (8) is $4/\pi$ times $[\sin x + (\sin 3x)/3 + (\sin 5x/5) + \cdots]$. The sum of the first N terms has an interesting shape, close to the square wave except where the wave jumps between -1 and 1. At those jumps, the Fourier sum spikes the wrong way to ± 1.09 (the *Gibbs phenomenon*) before it takes the jump with the true $f(x)$.

This happens for the Fourier sums of all functions with jumps. It makes shock waves hard to compute. You can see it clearly in a graph of the sum of 10 terms.

8 $\|\boldsymbol{v}\|^2 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2$ so $\|\boldsymbol{v}\| = \sqrt{2}$; $\|\boldsymbol{v}\|^2 = 1 + a^2 + a^4 + \cdots = 1/(1 - a^2)$ so $\|\boldsymbol{v}\| = 1/\sqrt{1 - a^2}$; $\int_0^{2\pi}(1 + 2\sin x + \sin^2 x)\,dx = 2\pi + 0 + \pi$ so $\|f\| = \sqrt{3\pi}$.

9 (a) $f(x) = (1 + \text{square wave})/2$ so the a's are $\frac{1}{2}$, 0, $0, \ldots$ and the b's are $2/\pi$, 0, $-2/3\pi, 0, 2/5\pi, \ldots$ (b) $a_0 = \int_0^{2\pi} x\,dx/2\pi = \pi$, all other $a_k = 0$, $b_k = -2/k$.

10 The integral from $-\pi$ to π or from 0 to 2π (or from any a to $a + 2\pi$) is over one complete period of the function. If $f(x)$ is periodic this changes $\int_0^{2\pi} f(x)\,dx$ to $\int_0^{\pi} f(x)\,dx + \int_{-\pi}^{0} f(x)\,dx$. If $f(x)$ is **odd**, those integrals cancel to give $\int f(x)\,dx = 0$ over one period.

11 $\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos 2x$; $\cos(x + \frac{\pi}{3}) = \cos x \cos\frac{\pi}{3} - \sin x \sin\frac{\pi}{3} = \frac{1}{2}\cos x - \frac{\sqrt{3}}{2}\sin x$.

12 $$\frac{d}{dx}\begin{bmatrix}1\\ \cos x\\ \sin x\\ \cos 2x\\ \sin 2x\end{bmatrix}=\begin{bmatrix}0\\ -\sin x\\ \cos x\\ -2\sin 2x\\ 2\cos 2x\end{bmatrix}=\begin{bmatrix}0&0&0&0&0\\0&0&-1&0&0\\0&1&0&0&0\\0&0&0&0&-2\\0&0&0&2&0\end{bmatrix}\begin{bmatrix}1\\ \cos x\\ \sin x\\ \cos 2x\\ \sin 2x\end{bmatrix}.$$ This shows the differentiation matrix.

13 The square pulse with $F(x)=1/h$ for $-x \le h/2 \le x$ is an even function, so all sine coefficients b_k are zero. The average a_0 and the cosine coefficients a_k are

$$a_0=\frac{1}{2\pi}\int_{-h/2}^{h/2}(1/h)dx=\frac{1}{2\pi}$$

$$a_k=\frac{1}{\pi}\int_{-h/2}^{h/2}(1/h)\cos kx dx=\frac{2}{\pi kh}\left(\sin\frac{kh}{2}\right) \text{ which is } \frac{1}{\pi}\text{ sinc}\left(\frac{kh}{2}\right)$$

(introducing the sinc function $(\sin x)/x$). As h approaches zero, the number $x=kh/2$ approaches zero, and $(\sin x)/x$ approaches 1. So all those a_k approach $1/\pi$.

The limiting "delta function" contains an equal amount of all cosines: a very irregular function.

Problem Set 10.6, page 500

1 (x,y,z) has homogeneous coordinates (cx,cy,cz,c) for $c=1$ and all $c\neq 0$.

2 For an affine transformation we also need T (origin), because $T(\mathbf{0})$ need not be $\mathbf{0}$ for affine T. Including this translation by $T(\mathbf{0})$, $(x,y,z,1)$ is transformed to $xT(\boldsymbol{i})+yT(\boldsymbol{j})+zT(\boldsymbol{k})+T(\mathbf{0})$.

3 $$TT_1=\begin{bmatrix}1&&&\\&1&&\\&&1&\\1&4&3&1\end{bmatrix}\begin{bmatrix}1&&&\\&1&&\\&&1&\\0&2&5&1\end{bmatrix}=\begin{bmatrix}1&&&\\&1&&\\&&1&\\1&6&8&1\end{bmatrix}$$ is translation along $(1,6,8)$.

4 $S=\text{diag}\,(c,c,c,1)$; row 4 of ST and TS is $1,4,3,1$ and $c,4c,3c,1$; use $\boldsymbol{v}TS$!

5 $S = \begin{bmatrix} 1/8.5 & & \\ & 1/11 & \\ & & 1 \end{bmatrix}$ for a 1 by 1 square, starting from an 8.5 by 11 page.

6 $$[\,x\;y\;z\;1\,]\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ -1 & -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 1 \end{bmatrix} = [\,x\;y\;z\;1\,]\begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ -2 & -2 & -4 & 1 \end{bmatrix}.$$
The first matrix translates by $(-1,-1,-2)$. The second matrix rescales by 2.

7 The three parts of Q in equation (1) are $(\cos\theta)I$ and $(1-\cos\theta)\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}$ and $-\sin\theta(\boldsymbol{a}\times)$. Then $Q\boldsymbol{a} = \boldsymbol{a}$ because $\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = \boldsymbol{a}$(unit vector) and $\boldsymbol{a}\times\,\boldsymbol{a} = \boldsymbol{0}$.

8 If $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 0$ and those three parts of Q (Problem 7) multiply $\boldsymbol{b}$, the results in $Q\boldsymbol{b}$ are $(\cos\theta)\boldsymbol{b}$ and $\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{0}$ and $(-\sin\theta)\boldsymbol{a}\times\,\boldsymbol{b}$. The component along $\boldsymbol{b}$ is $(\cos\theta)\boldsymbol{b}$.

9 $\boldsymbol{n} = \left(\frac{2}{3},\frac{2}{3},\frac{1}{3}\right)$ has $P = I - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} = \frac{1}{9}\begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix}$. Notice $\|\boldsymbol{n}\| = 1$.

10 We can choose $(0,0,3)$ on the plane and multiply $T_-PT_+ = \frac{1}{9}\begin{bmatrix} 5 & -4 & -2 & 0 \\ -4 & 5 & -2 & 0 \\ -2 & -2 & 8 & 0 \\ 6 & 6 & 3 & 9 \end{bmatrix}$.

11 $(3,3,3)$ projects to $\frac{1}{3}\big(-1,-1,4\big)$ and $(3,3,3,1)$ projects to $\big(\frac{1}{3},\frac{1}{3},\frac{5}{3},1\big)$. Row vectors!

12 The projection of a square onto a plane is a parallelogram (or a line segment). The sides of the square are perpendicular, but their projections may not be ($\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = 0$ but $(P\boldsymbol{x})^{\mathrm{T}}(P\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}P^{\mathrm{T}}P\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}P\boldsymbol{y}$ may be nonzero).

13 That projection of a cube onto a plane produces a hexagon.

14 $$(3,3,3)(I - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}) = \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\begin{bmatrix} 1 & -8 & -4 \\ -8 & 1 & -4 \\ -4 & -4 & 7 \end{bmatrix} = \left(-\frac{11}{3},-\frac{11}{3},-\frac{1}{3}\right).$$

15 $(3,3,3,1) \to (3,3,0,1) \to \left(-\frac{7}{3},-\frac{7}{3},-\frac{8}{3},1\right) \to \left(-\frac{7}{3},-\frac{7}{3},\frac{1}{3},1\right)$.

16 Just subtracting vectors would give $\boldsymbol{v} = (x,y,z,0)$ ending in 0 (not 1). In homogeneous coordinates, add a **vector** to a point.

17 Space is rescaled by $1/c$ because (x,y,z,c) is the same point as $(x/c, y/c, z/c, 1)$.

Problem Set 10.7, page 507

1 **Multiplying** n whole numbers gives an odd number only when *all* n *numbers are odd*. This translates to multiplication (*mod* 2). Multiplying n 1's or 0's gives 1 only when all n numbers are 1.

2 **Adding** n whole numbers gives an odd number only when the n numbers include *an odd number of odd numbers*. For addition of 1's and 0's (*mod* 2), the answer is odd when the number of 1's is odd.

3 (a) We are given that $y_1 - x_1$ and $y_2 - x_2$ are both divisible by p. Then their sum $y_1 + y_2 - x_1 - x_2$ is divisible by p.

(b) $5 \equiv 2$ (*mod* 3) and $8 \equiv 2$ (*mod* 3) add to $13 \equiv 4$ (*mod* 3). The number 1 is smaller than 4 and $13 \equiv 1$ (*mod* 3).

5 If $y - x$ is divisible by p then $x - y$ is also divisible by p. In other words, if $y - x = mp$ then $x - y = (-m)p$.

6 $A = \begin{bmatrix} 5 & 5 \\ 5 & 10 \end{bmatrix}$ is an invertible matrix but (*mod* 5) A becomes the zero matrix.

7 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ are invertible : 6 out of 16 possible 0-1 matrices.

8 Yes, $A\boldsymbol{x} = \mathbf{0}$ (*mod* 11) says that every row of A is orthogonal to every $\boldsymbol{x}$ in the nullspace (*mod* 11). But a basis for the usual $\mathbf{N}(A)$ could include vectors that are zero (*mod* 11).

9 For simplicity, number the letters as they appear in the message :

THISWHOLEBOOKISINCODE $= \quad 123/452/678/966/(10)34/3(11)(12)/6(13)8$.

Multiply each block by this L to obtain Hill's cipher.

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{Cipher} = 1\ 3\ 6/4\ 9\ 11/6\ 13\ 21/9\ 15\ 21/10\ 13\ 17/3\ 14\ 26/6\ 19\ 27.$$

If the cipher is *mod* p then replace each number by the correct number from 0 to $p-1$.

To decode, first multiply by L^{-1}. Then what to do??

10 First you have to discover the block size ($=$ matrix size) and also the matrix L itself. Start with a guess for the block size. Then the plaintext and the coded cipher tell you a series of matrix-vector products $L\boldsymbol{x} \equiv \boldsymbol{b}$. If the text is long enough (and the blocks are not too long) this is enough information to find L—or to show that the block size must be wrong, when there is no L that gets all correct blocks $L\boldsymbol{x} \equiv \boldsymbol{b}$.

The extra difficulty is to find the value of p.

Problem Set 11.1, page 516

1 Without exchange, pivots .001 and 1000; with exchange, 1 and -1. When the pivot is larger than the entries below it, all $|\ell_{ij}| = \dfrac{|\text{entry}|}{|\text{pivot}|} \leq 1$. $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$.

2 The exact inverse of hilb(3) is $A^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$.

3 $A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 11/6 \\ 13/12 \\ 47/60 \end{bmatrix} = \begin{bmatrix} 1.833 \\ 1.083 \\ 0.783 \end{bmatrix}$ compares with $A\begin{bmatrix} 0 \\ 6 \\ -3.6 \end{bmatrix} = \begin{bmatrix} 1.80 \\ 1.10 \\ 0.78 \end{bmatrix}$. $\|\Delta \boldsymbol{b}\| < .04$ but $\|\Delta \boldsymbol{x}\| > 6$.
The difference $(1,1,1) - (0,6,-3.6)$ is in a direction $\Delta\boldsymbol{x}$ that has $A\Delta\boldsymbol{x}$ near zero.

4 The largest $\|\boldsymbol{x}\| = \|A^{-1}\boldsymbol{b}\|$ is $\|A^{-1}\| = 1/\lambda_{\text{min}}$ since $A^{\mathrm{T}} = A$; largest error $10^{-16}/\lambda_{\text{min}}$.

5 Each row of U has at most w entries. Use w multiplications to substitute components of $\boldsymbol{x}$ (already known from below) and divide by the pivot. Total for n rows $< wn$.

6 The triangular L^{-1}, U^{-1}, R^{-1} need $\frac{1}{2}n^2$ multiplications. Q needs n^2 to multiply the right side by $Q^{-1} = Q^{\mathrm{T}}$. So $QR\boldsymbol{x} = \boldsymbol{b}$ takes 1.5 times longer than $LU\boldsymbol{x} = \boldsymbol{b}$.

7 $UU^{-1} = I$: Back substitution needs $\frac{1}{2}j^2$ multiplications on column j, using the j by j upper left block. Then $\frac{1}{2}(1^2 + 2^2 + \cdots + n^2) \approx \frac{1}{2}(\frac{1}{3}n^3) =$ total to find U^{-1}.

8 $\begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 2 \\ 0 & -1 \end{bmatrix} = U$ with $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $L = \begin{bmatrix} 1 & 0 \\ .5 & 1 \end{bmatrix}$;

$A \to \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = U$ with

$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and $L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ .5 & -.5 & 1 \end{bmatrix}$.

9 $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ has cofactors $C_{13} = C_{31} = C_{24} = C_{42} = 1$ and $C_{14} = C_{41} = -1$. A^{-1} is a full matrix!

10 With 16-digit floating point arithmetic the errors $\|\boldsymbol{x} - \boldsymbol{x}_{\text{computed}}\|$ for $\varepsilon = 10^{-3}$, 10^{-6}, 10^{-9}, 10^{-12}, 10^{-15} are of order 10^{-16}, 10^{-11}, 10^{-7}, 10^{-4}, 10^{-3}.

11 (a) $\cos\theta = 1/\sqrt{10}$, $\sin\theta = -3/\sqrt{10}$, $R = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 5 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 10 & 14 \\ 0 & 8 \end{bmatrix}$.

(b) A has eigenvalues 4 and 2. Put one of the unit eigenvectors in row 1 of Q: either

$Q = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $QAQ^{-1} = \begin{bmatrix} 2 & -4 \\ 0 & 4 \end{bmatrix}$ **or**

$Q = \dfrac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$ and $QAQ^{-1} = \begin{bmatrix} 4 & -4 \\ 0 & 2 \end{bmatrix}$.

12 When A is multiplied by a plane rotation Q_{ij}, this changes the $2n$ (not n^2) entries in rows i and j. Then multiplying on the right by $(Q_{ij})^{-1} = (Q_{ij})^{\mathrm{T}}$ changes the $2n$ entries in columns i and j.

13 $Q_{ij}A$ uses $4n$ multiplications (2 for each entry in rows i and j). By factoring out $\cos\theta$, the entries 1 and $\pm\tan\theta$ need only $2n$ multiplications, which leads to $\frac{2}{3}n^3$ for QR.

14 The $(2,1)$ entry of $Q_{21}A$ is $\frac{1}{3}(-\sin\theta + 2\cos\theta)$. This is zero if $\sin\theta = 2\cos\theta$ or $\tan\theta = 2$. Then the $2, 1, \sqrt{5}$ right triangle has $\sin\theta = 2/\sqrt{5}$ and $\cos\theta = 1/\sqrt{5}$.

Every 3 by 3 rotation with $\det Q = +1$ is the product of 3 plane rotations.

15 This problem shows how elimination is more expensive (the nonzero multipliers in L and LL are counted by **nnz**(L) and **nnz**(LL)) when we spoil the tridiagonal K by a random permutation.

If on the other hand we start with a poorly ordered matrix K, an improved ordering is found by the code **symamd** discussed in this section.

16 The "red-black ordering" puts rows and columns 1 to 10 in the odd-even order $1, 3, 5, 7, 9, 2, 4, 6, 8, 10$. When K is the $-1, 2, -1$ tridiagonal matrix, odd points are connected only to even points (and 2 stays on the diagonal, connecting every point to itself):

$$K = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \cdot & \cdot & \cdot \\ & & -1 & 2 \end{bmatrix} \text{ and } PKP^{\mathrm{T}} = \begin{bmatrix} 2I & D \\ D^{\mathrm{T}} & 2I \end{bmatrix} \text{ with}$$

$$D = \begin{bmatrix} -1 & & & & \\ -1 & -1 & & & \\ 0 & -1 & -1 & & \\ & & -1 & -1 & \\ & & & -1 & -1 \end{bmatrix} \begin{matrix} 1 \text{ to } 2 \\ 3 \text{ to } 2,4 \\ 5 \text{ to } 4,6 \\ 7 \text{ to } 6,8 \\ 9 \text{ to } 8,10 \end{matrix}$$

17 Jeff Stuart's **Shake a Stick** activity has long sticks representing the graphs of two linear equations in the x-y plane. The matrix is nearly singular and Section 9.2 shows how to compute its condition number $c = \|A\|\|A^{-1}\| = \sigma_{\max}/\sigma_{\min} \approx 80,000$:

$$A = \begin{bmatrix} 1 & 1.0001 \\ 1 & 1.0000 \end{bmatrix} \quad \|A\| \approx 2 \quad A^{-1} = 10000 \begin{bmatrix} -1 & 1.0001 \\ 1 & -1 \end{bmatrix} \quad \begin{matrix} \|A^{-1}\| \approx 20000 \\ c \approx 40000. \end{matrix}$$

Problem Set 11.2, page 522

1 $\|A\| = 2$, $\|A^{-1}\| = 2$, $c = 4$; $\|A\| = 3$, $\|A^{-1}\| = 1$, $c = 3$; $\|A\| = 2 + \sqrt{2} = \lambda_{\max}$ for positive definite A, $\|A^{-1}\| = 1/\lambda_{\min}$, comd $= (2+\sqrt{2})/(2-\sqrt{2}) = 5.83$.

2 $\|A\| = 2$, $c = 1$; $\|A\| = \sqrt{2}$, $c = \infty$ (singular matrix); $A^{\mathrm{T}}A = 2I$, $\|A\| = \sqrt{2}$, $c = 1$.

3 For the first inequality replace $\boldsymbol{x}$ by $B\boldsymbol{x}$ in $\|A\boldsymbol{x}\| \le \|A\|\|\boldsymbol{x}\|$; the second inequality is just $\|B\boldsymbol{x}\| \le \|B\|\|\boldsymbol{x}\|$. Then $\|AB\| = \max(\|AB\boldsymbol{x}\|/\|\boldsymbol{x}\|) \le \|A\|\|B\|$.

4 $1 = \|I\| = \|AA^{-1}\| \le \|A\|\|A^{-1}\| = c(A)$.

5 If $\Lambda_{\max} = \Lambda_{\min} = 1$ then all $\Lambda_i = 1$ and $A = SIS^{-1} = I$. The only matrices with $\|A\| = \|A^{-1}\| = 1$ are *orthogonal matrices*.

6 All orthogonal matrices have norm 1, so $\|A\| \leq \|Q\|\|R\| = \|R\|$ and in reverse $\|R\| \leq \|Q^{-1}\|\|A\| = \|A\|$. Then $\|A\| = \|R\|$. Inequality is usual in $\|A\| < \|L\|\|U\|$ when $A^{\mathrm{T}}A \neq AA^{\mathrm{T}}$. Use **norm** on a random A.

7 The triangle inequality gives $\|A\boldsymbol{x} + B\boldsymbol{x}\| \leq \|A\boldsymbol{x}\| + \|B\boldsymbol{x}\|$. Divide by $\|\boldsymbol{x}\|$ and take the maximum over all nonzero vectors to find $\|A + B\| \leq \|A\| + \|B\|$.

8 If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\| = |\lambda|$ for that particular vector $\boldsymbol{x}$. When we maximize the ratio $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\|$ over all vectors we get $\|A\| \geq |\lambda|$.

9 $A+B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has $\rho(A) = 0$ and $\rho(B) = 0$ but $\rho(A+B) = 1$. The triangle inequality $\|A + B\| \leq \|A\| + \|B\|$ fails for $\rho(A)$. $AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ has $\rho(AB) > \rho(A)\,\rho(B)$; thus $\rho(A) = \max|\lambda(A)|$ = spectral radius is not a norm.

10 (a) The condition number of A^{-1} is $\|A^{-1}\|\|(A^{-1})^{-1}\|$ which is $\|A^{-1}\|\|A\| = c(A)$. (b) Since $A^{\mathrm{T}}A$ and AA^{T} have the same nonzero eigenvalues, A^{T} has the same norm as A.

11 Use the quadratic formula for $\lambda_{\max}/\lambda_{\min}$, which is $c = \sigma_{\max}/\sigma_{\min}$ since this $A = A^{\mathrm{T}}$ is positive definite:

$$c(A) = \left(1.00005 + \sqrt{(1.00005)^2 - .0001}\right) \Big/ \left(1.00005 - \sqrt{}\right) \approx 40{,}000.$$

12 $\det(2A)$ is not $2\det A$; $\det(A + B)$ is not always less than $\det A + \det B$; taking $|\det A|$ does not help. The only reasonable property is $\det AB = (\det A)(\det B)$. The condition number should not change when A is multiplied by 10.

13 The residual $\boldsymbol{b} - A\boldsymbol{y} = (10^{-7}, 0)$ is much smaller than $\boldsymbol{b} - A\boldsymbol{z} = (.0013, .0016)$. But $\boldsymbol{z}$ is much closer to the solution than $\boldsymbol{y}$.

14 $\det A = 10^{-6}$ so $A^{-1} = 10^3 \begin{bmatrix} 659 & -563 \\ -913 & 780 \end{bmatrix}$: $\|A\| > 1,\ \|A^{-1}\| > 10^6$, then $c > 10^6$.

15 $\boldsymbol{x} = (1,1,1,1,1)$ has $\|\boldsymbol{x}\| = \sqrt{5}, \|\boldsymbol{x}\|_1 = 5, \|\boldsymbol{x}\|_\infty = 1$. $\boldsymbol{x} = (.1,.7,.3,.4,.5)$ has $\|\boldsymbol{x}\| = 1, \|\boldsymbol{x}\|_1 = 2$ (sum), $\|\boldsymbol{x}\|_\infty = .7$ (largest).

16 $x_1^2+\cdots+x_n^2$ is not smaller than $\max(x_i^2)$ and not larger than $(|x_1|+\cdots+|x_n|)^2 = \|\boldsymbol{x}\|_1^2$. $x_1^2+\cdots+x_n^2 \le n\,\max(x_i^2)$ so $\|\boldsymbol{x}\| \le \sqrt{n}\|\boldsymbol{x}\|_\infty$. Choose $y_i = \operatorname{sign} x_i = \pm 1$ to get $\|\boldsymbol{x}\|_1 = \boldsymbol{x} \cdot \boldsymbol{y} \le \|\boldsymbol{x}\|\|\boldsymbol{y}\| = \sqrt{n}\|\boldsymbol{x}\|$. The vector $\boldsymbol{x} = (1,\ldots,1)$ has $\|\boldsymbol{x}\|_1 = \sqrt{n}\,\|\boldsymbol{x}\|$.

17 For the ℓ^∞ norm, the largest component of $\boldsymbol{x}$ plus the largest component of $\boldsymbol{y}$ is not less than $\|\boldsymbol{x}+\boldsymbol{y}\|_\infty =$ largest component of $\boldsymbol{x}+\boldsymbol{y}$.

For the ℓ^1 norm, each component has $|x_i + y_i| \le |x_i| + |y_i|$. Sum on $i = 1$ to n: $\|\boldsymbol{x}+\boldsymbol{y}\|_1 \le \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1$.

18 $|x_1| + 2|x_2|$ is a norm but $\min(|x_1|,|x_2|)$ is not a norm. $\|\boldsymbol{x}\| + \|\boldsymbol{x}\|_\infty$ is a norm; $\|\boldsymbol{Ax}\|$ is a norm provided A is invertible (otherwise a nonzero vector has norm zero; for rectangular A we require independent columns to avoid $\|A\boldsymbol{x}\| = 0$).

19 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = x_1y_1 + x_2y_2 + \cdots \le (\max|y_i|)(|x_1| + |x_2| + \cdots) = ||\boldsymbol{x}||_1\,||\boldsymbol{y}||_\infty$.

20 With $\lambda_j = 2 - 2\cos(j\pi/n{+}1)$, the largest eigenvalue is $\lambda_n \approx 2+2 = 4$. The smallest is $\lambda_1 = 2 - 2\cos(\pi/n{+}1) \approx \left(\frac{\pi}{n+1}\right)^2$, using $2\cos\theta \approx 2 - \theta^2$. So the condition number is $c = \lambda_{\max}/\lambda_{\min} \approx (4/\pi^2)\,n^2$, growing with n.

Problem Set 11.3, page 531

1 The iteration $\boldsymbol{x}_{k+1} = (I - A)\boldsymbol{x}_k + \boldsymbol{b}$ has $S = I$ and $T = I - A$ and $S^{-1}T = I - A$.

2 If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then $(I - A)\boldsymbol{x} = (1-\lambda)\boldsymbol{x}$. Real eigenvalues of $B = I - A$ have $|1-\lambda| < 1$ provided λ is between 0 and 2.

3 This matrix A has $I - A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ which has $|\lambda| = 2$. The iteration diverges.

4 Always $\|AB\| \le \|A\|\|B\|$. Choose $A = B$ to find $\|B^2\| \le \|B\|^2$. Then choose $A = B^2$ to find $\|B^3\| \le \|B^2\|\|B\| \le \|B\|^3$. Continue (or use induction) to find $\|B^k\| \le \|B\|^k$. Since $\|B\| \ge \max|\lambda(B)|$ it is no surprise that $\|B\| < 1$ gives convergence.

5 $A\boldsymbol{x} = \mathbf{0}$ gives $(S - T)\boldsymbol{x} = \mathbf{0}$. Then $S\boldsymbol{x} = T\boldsymbol{x}$ and $S^{-1}T\boldsymbol{x} = \boldsymbol{x}$. Then $\lambda = 1$ means that the errors do not approach zero. We can't expect convergence when A is singular and $A\boldsymbol{x} = \boldsymbol{b}$ is unsolvable!

6 Jacobi has $S^{-1}T = \frac{1}{3}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ with $|\lambda|_{\max} = \frac{1}{3}$. Small problem, fast convergence.

7 Gauss-Seidel has $S^{-1}T = \begin{bmatrix} 0 & \frac{1}{3} \\ 0 & \frac{1}{9} \end{bmatrix}$ with $|\lambda|_{\max} = \frac{1}{9}$ which is ($|\lambda|_{\max}$ for Jacobi)$^{\mathbf{2}}$.

8 Jacobi has $S^{-1}T = \begin{bmatrix} a & \\ & d \end{bmatrix}^{-1} \begin{bmatrix} 0 & -b \\ -c & 0 \end{bmatrix} = \begin{bmatrix} 0 & -b/a \\ -c/d & 0 \end{bmatrix}$ with $|\lambda| = |bc/ad|^{1/2}$.

Gauss-Seidel has $S^{-1}T = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} 0 & -b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -b/a \\ 0 & -bc/ad \end{bmatrix}$ with $|\lambda| = |bc/ad|$.

So Gauss-Seidel is twice as fast to converge if $|\lambda| < 1$ (or to explode if $|bc| > |ad|$).

9 Gauss-Seidel will converge for the $-1, 2, -1$ matrix. $|\lambda|_{\max} = \cos^2\left(\frac{\pi}{n+1}\right)$ is given on page 527, together with the improvement from successive overrelaxation.

10 If the iteration gives all $x_i^{\text{new}} = x_i^{\text{old}}$ then the quantity in parentheses is zero, which means $A\boldsymbol{x} = \boldsymbol{b}$. For Jacobi change $\boldsymbol{x}^{\text{new}}$ on the right side to x^{old}.

11 $\boldsymbol{u}_k/\lambda_1^k = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2(\lambda_2/\lambda_1)^k + \cdots + c_n\boldsymbol{x}_n(\lambda_n/\lambda_1)^k \to c_1\boldsymbol{x}_1$ if all ratios $|\lambda_i/\lambda_1| < 1$. The largest ratio controls the rate of convergence (when k is large). $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has $|\lambda_2| = |\lambda_1|$ and no convergence.

12 The eigenvectors of A and also A^{-1} are $\boldsymbol{x}_1 = (.75, .25)$ and $\boldsymbol{x}_2 = (1, -1)$. The inverse power method converges to a multiple of $\boldsymbol{x}_2$, since $|1/\lambda_2| > |1/\lambda_1|$.

13 In the jth component of $A\boldsymbol{x}_1$, $\lambda_1 \sin\frac{j\pi}{n+1} = 2\sin\frac{j\pi}{n+1} - \sin\frac{(j-1)\pi}{n+1} - \sin\frac{(j+1)\pi}{n+1}$. The last two terms combine into $-2\sin\frac{j\pi}{n+1}\cos\frac{\pi}{n+1}$. Then $\lambda_1 = 2 - 2\cos\frac{\pi}{n+1}$.

14 $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ produces $\boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{u}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$, $\boldsymbol{u}_3 = \begin{bmatrix} 14 \\ -13 \end{bmatrix}$.

This is converging to the eigenvector direction $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with largest eigenvalue $\lambda = 3$. Divide $\boldsymbol{u}_k$ by $\|\boldsymbol{u}_k\|$ to keep unit vectors.

15 $A^{-1} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ gives $\boldsymbol{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\boldsymbol{u}_2 = \frac{1}{9}\begin{bmatrix} 5 \\ 4 \end{bmatrix}$, $\boldsymbol{u}_3 = \frac{1}{27}\begin{bmatrix} 14 \\ 13 \end{bmatrix} \rightarrow \boldsymbol{u}_\infty = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$.

16 $R = Q^{\mathrm{T}}A = \begin{bmatrix} 1 & \cos\theta\sin\theta \\ 0 & -\sin^2\theta \end{bmatrix}$ and $A_1 = RQ = \begin{bmatrix} \cos\theta(1+\sin^2\theta) & -\sin^3\theta \\ \boldsymbol{-\sin^3\theta} & -\cos\theta\sin^2\theta \end{bmatrix}$.

17 If A is orthogonal then $Q = A$ and $R = I$. Therefore $A_1 = RQ = A$ again, and the "QR method" doesn't move from A. But shift A slightly and the method goes quickly to Λ.

18 If $A - cI = QR$ then $A_1 = RQ + cI = Q^{-1}(QR + cI)Q = Q^{-1}AQ$. No change in eigenvalues from the shift and shift back, because A_1 is similar to A.

19 Multiply $A\boldsymbol{q}_j = b_{j-1}\boldsymbol{q}_{j-1} + a_j\boldsymbol{q}_j + b_j\boldsymbol{q}_{j+1}$ by $\boldsymbol{q}_j^{\mathrm{T}}$ to find $\boldsymbol{q}_j^{\mathrm{T}}A\boldsymbol{q}_j = a_j$ (because the $\boldsymbol{q}$'s are orthonormal). The matrix form (multiplying by columns) is $AQ = QT$ where T is *tridiagonal*. The entries down the diagonals of T are the a's and b's.

20 Theoretically the $\boldsymbol{q}$'s are orthonormal. In reality this important algorithm is not very stable. We must stop every few steps to reorthogonalize—or find another more stable way to orthogonalize the sequence $\boldsymbol{q}, A\boldsymbol{q}, A^2\boldsymbol{q}, \ldots$

21 If A is symmetric then $A_1 = Q^{-1}AQ = Q^{\mathrm{T}}AQ$ is also symmetric. $A_1 = RQ = R(QR)R^{-1} = RAR^{-1}$ has R and R^{-1} upper triangular, so A_1 cannot have nonzeros on a lower diagonal than A. If A is tridiagonal and symmetric then (by using symmetry for the upper part of A_1) the matrix $A_1 = RAR^{-1}$ is also tridiagonal.

22 From the last line of code, $\boldsymbol{q}_2$ is in the direction of $\boldsymbol{v} = A\boldsymbol{q}_1 - h_{11}\boldsymbol{q}_1 = A\boldsymbol{q}_1 - (\boldsymbol{q}_1^{\mathrm{T}}A\boldsymbol{q}_1)\boldsymbol{q}_1$. The dot product with $\boldsymbol{q}_1$ is zero. This is Gram-Schmidt with $A\boldsymbol{q}_1$ as the second input vector; we subtract from $A\boldsymbol{q}_1$ its projection onto the first vector $\boldsymbol{q}_1$.

Note The three lines after the short "pseudocodes" describe two key properties of conjugate gradients—the residuals $\boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k$ are orthogonal and the search directions are A-orthogonal ($\boldsymbol{d}_i^{\mathrm{T}} A \boldsymbol{d}_k = 0$). Then each new approximation $\boldsymbol{x}_{k+1}$ is the **closest vector to $\boldsymbol{x}$** among all combinations of $\boldsymbol{b}, A\boldsymbol{b} \ldots, A^k \boldsymbol{b}$. Ordinary iteration $S\boldsymbol{x}_{k+1} = T\boldsymbol{x}_k + \boldsymbol{b}$ does *not* find this best possible combination $\boldsymbol{x}_{k+1}$.

23 The solution is straightforward and important. Since $H = Q^{-1}AQ = Q^{\mathrm{T}}AQ$ is *symmetric* if $A = A^{\mathrm{T}}$, and since H has only one lower diagonal by construction, then H has only *one upper diagonal*: H is tridiagonal and all the recursions in Arnoldi's method have only 3 terms.

24 $H = Q^{-1}AQ$ is *similar* to A, so H has the same eigenvalues as A (at the end of Arnoldi). When Arnoldi is stopped sooner because the matrix size is large, the eigenvalues of H_k (called *Ritz values*) are close to eigenvalues of A. This is an important way to compute approximations to λ for large matrices.

25 In principle the conjugate gradient method converges in 100 (or 99) steps to the exact solution $\boldsymbol{x}$. But it is slower than elimination and its all-important property is to give good approximations to $\boldsymbol{x}$ much sooner. (Stopping elimination part way leaves you nothing.) The problem asks how close $\boldsymbol{x}_{10}$ and $\boldsymbol{x}_{20}$ are to $\boldsymbol{x}_{100}$, which equals $\boldsymbol{x}$ except for roundoff errors.

26 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1.1 \end{bmatrix}$ has $A^n = \begin{bmatrix} 1 & q \\ 0 & (1.1)^n \end{bmatrix}$ with $q = 1 + 1.1 + \cdots + (1.1)^{n-1} = (1.1^n - 1)/(1.1 - 1) \approx \mathbf{10\,(1.1)^n}$. So the growing part of A^n is $(1.1)^n \begin{bmatrix} 0 & 10 \\ 0 & 1 \end{bmatrix}$ with $||A^n|| \approx \sqrt{101}$ times 1.1^n for larger n.

Problem Set 12.1, page 544

1 When 7 is added to every output, the mean increases by 7 and the variance does not change (because new variance comes from (distance)^2 to the new mean).

New sample mean and new expected mean : Add 7. New variance : No change.

2 If we add $\frac{1}{3}$ to $\frac{1}{7}$ (fraction of integers divisible by 3 *plus* fraction divisible by 7) we have **double counted** the integers divisible by both 3 and 7. This is a fraction $\frac{1}{21}$ of all integers (because these double counted numbers are multiples of 21). So the fraction divisible by 3 or 7 or both is

$$\frac{1}{3}+\frac{1}{7}-\frac{1}{21}=\frac{7}{21}+\frac{3}{21}-\frac{1}{21}=\frac{9}{21}=\mathbf{\frac{3}{7}}.$$

3 In the numbers from 1 to 1000, each group of ten numbers will contain each possible ending $x = 1, 2, 3, \ldots, 0$. So those endings all have the same probability $p_i = \frac{1}{10}$. Expected mean of that last digit x :

$$m = \mathrm{E}\,[x] = \sum p_i\, x_i = \frac{1}{10}\sum_{i=0}^{9} i = \frac{45}{10} = \mathbf{4.5}$$

The best way to find the variance $\sigma^2 = 8.25$ is **in the last line below and in problem 12.1.7**. The slower way to find σ^2 is

$$\sigma^2 = \mathrm{E}\,[(x-4.5)^2] = \sum_{i=0}^{9} p_i(x_i-4.5)^2 = \frac{1}{10}\sum_{i=0}^{9}(i-4.5)^2$$

We can separate $\left(i-4.5\right)^2$ into $\left(i^2-9i+(4.5)^2\right)$ and add from $i=0$ to $i=9$:

$$\frac{1}{10}\left(\sum_{0}^{9} i^2 - 9\sum_{0}^{9} i + \sum_{0}^{9}(4.5)^2\right) = \frac{1}{10}\left(285-9(45)+10(4.5)^2\right)$$
$$= \frac{1}{10}(285-405+202.5) = \frac{82.5}{10} = 8.25 = \frac{33}{4}.$$

Notice that 202.5 is half of 405—like Nm^2 and $2Nm^2$ in equation (4), page 536. **I should have extended equation** (4) **to its best form :**

$$\boldsymbol{\sigma^2 = \mathrm{E}\,[(x-m)^2] = \mathrm{E}\,[x^2] - m^2}$$

That quickly gives $\frac{285}{10} - (4.5)^2 = 8.25 =$ same answer.

4 For numbers ending in $0, 1, 2, \ldots, 9$ the squares end in $x = 0, 1, 4, 9, 6, 5, 6, 9, 4, 1$. So the probabilities of $x = 0$ and 5 are $p = \frac{1}{10}$ and the probabilities of $x = 1, 4, 6, 9$ are $p = \frac{1}{5}$. The mean is

$$m = \sum p_i\, x_i = \frac{0}{0} + \frac{5}{10} + \frac{1}{5}\,(1 + 4 + 6 + 9) = 4.5 = \text{ same as before.}$$

The variance using the improvement of equation (4) is

$$\begin{aligned}\sigma^2 \quad &= \mathrm{E}\,[x^2] - m^2 = \frac{1}{10}\,0^2 + \frac{1}{10}\,5^2 + \frac{1}{5}\,(1^2 + 4^2 + 6^2 + 9^2) - m^2 \\ &= \frac{25}{10} + \frac{134}{5} - 20.25 = \mathbf{9.05}\end{aligned}$$

5 For numbers from 1 to 1000, the first digit is $x = 1$ for 1000 and 100-199 and 10-19 and 1 (112 times). The first digit is $x = 2$ for 200-299 and 20-29 and 2 (111 times). The other first digits $x = 3$ to 9 also happen (111 times). Total (1000 times) is correct.

The average first digit is the mean, close to $\frac{1}{9}\big(1 + 2 + \cdots + 9\big) = 5$:

$$m = \sum p_i\, x_i = \frac{112}{1000}\,(1) + \frac{111}{1000}\,(2+3+\cdots+9) = \frac{112 + 111(44)}{1000} = \frac{4996}{1000} = 4.996 \approx 5.$$

The variance is

$$\sigma^2 = \mathrm{E}\,[(x - m)^2] = \mathrm{E}\,[x^2] - m^2 = \frac{112}{1000}\,(1^2) + \frac{111}{1000}(2^2 + \cdots + 9^2) - m^2$$

$$= \frac{112 + 111(284)}{1000} - m^2 \approx \frac{31635}{1000} - 5^2 = \mathbf{6.635}.$$

6 The first digits of $157^2, 312^2, 696^2$, and 602^2 are $\mathbf{2, 9, 4, 3}$, The sample mean is $\frac{1}{4}(2 + 9 + 4 + 3) = \frac{18}{4} = \mathbf{4.5}$. The sample variance with $N - 1 = 3$ is

$$S^2 = \frac{1}{3}\Big[(-2.5)^2 + (4.5)^2 + (-.5)^2 + (-1.5)^2\Big] = \frac{1}{3}\Big[29\Big].$$

7 This question is about the fast way to compute σ^2 using m^2. The mean m is probably already computed:

$$\begin{aligned}\sigma^2 &= \sum p_i\,(x_i - m)^2 = \sum p_i\,(x_i^2 - 2mx_i + m^2) \\ &= \sum p_i x_i^2 - 2m \sum p_i x_i + m^2 \sum p_i \\ &= \sum p_i x_i^2 - 2m^2 + m^2 = \sum \boldsymbol{p_i x_i^2 - m^2} = \mathrm{E}\,[\boldsymbol{x^2}] - \boldsymbol{m^2}.\end{aligned}$$

8 For $N = 24$ samples, all equal to $x = 20$,

$$\boldsymbol{\mu} = \frac{1}{N}\sum x_i = \frac{24}{24}(20) = \mathbf{20} \quad \text{and} \quad \boldsymbol{S^2} = \frac{1}{N-1}\sum (x_i - \mu)^2 = \mathbf{0}$$

For 12 samples of $x = 20$ and 12 samples of $x = 21$,

$$\boldsymbol{\mu} = \frac{12(20) + 12(21)}{24} = \mathbf{20.5} \quad \text{and} \quad \boldsymbol{S^2} = \frac{1}{N-1}\sum (x_i - \mu)^2 = \frac{1}{23}\,24\left(\frac{1}{2}\right)^2 = \frac{\mathbf{6}}{\mathbf{23}}.$$

9 This question asks you to set up a random 0-1 generator and run it a million times to find the average $A_{1000000}$.

One way is to use MATLAB's **rand** command with a uniform distribution between 0 and 1. Add $\frac{1}{2}$ to go between 0.5 and 1.5, then find the integer part (0 or 1). Using your computed average A_N (its mean is $m = \frac{1}{2}$ since 0 and 1 are equally likely for every sample) find the normalized variable X :

$$X = \frac{A_N - \frac{1}{2}}{2\sqrt{N}} = \frac{A_N - \frac{1}{2}}{2000} \quad \text{for } N = \text{ one million.}$$

10 The average number of heads in N fair coin flips is $m = N/2$. This is obvious—but how to derive it from probabilities p_0 to p_N of 0 to N heads? We have to compute

$$m = 0p_0 + 1p_1 + \cdots + Np_N \quad \text{with} \quad p_i = \frac{b_i}{2^N} = \frac{1}{2^N}\,\frac{N!}{i!\,(N-i)!}$$

A useful fact is $\boldsymbol{p_i = p_{N-i}}$. The probability of $\boldsymbol{i}$ **heads** equals the probability of $\boldsymbol{i}$ **tails**. If we take just those two terms in m, they give

$$ip_i + (N-i)p_{N-i} = ip_i + (N-i)p_i = Np_i$$

So we can compute m two ways and add :

$$\begin{aligned} m &= 0p_0 + 1p_1 + \cdots + (N-1)p_{N-1} + Np_N \\ m &= Np_0 + (N-1)p_1 + \cdots + 1p_{N-1} + 0p_0 \\ 2m &= Np_0 + Np_1 + \cdots + Np_{N-1} + Np_N \\ &= N(p_0 + p_1 + \cdots + p_{N-1} + p_N) = \boldsymbol{N}. \end{aligned}$$

Then $\boldsymbol{m = N/2}$. The average number of heads is $N/2$.

11 $\mathbf{E}\,[\boldsymbol{x^2}]$ $= \mathrm{E}\,[(x-m)^2 + 2xm - m^2]$

$$= \mathrm{E}\,[(x-m)^2] + 2m\,\mathrm{E}\,[x] - m^2\,\mathrm{E}\,[1]$$

$$= \sigma^2 + 2m^2 - m^2 = \boldsymbol{\sigma^2 + m^2}$$

12 The first step multiplies two independent 1-dimensional integrals (each one from $-\infty$ to ∞) to produce a 2-dimensional integral over the whole plane :

$$2\pi \int_{-\infty}^{\infty} p(x)\,dx \int_{-\infty}^{\infty} p(y)\,dy = 2\pi \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x)\,p(y)\,dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2}e^{-y^2/2}\,dxdy.$$

The second step changes to polar coordinates ($x = r\cos\theta, y = r\sin\theta, dxdy = r\,dr\,d\theta$, $x^2+y^2 = r^2$ with $0 \le \theta \le 2\pi$ and $0 \le r \le \infty$). Notice $-x^2/2 - y^2/2 = -r^2/2$:

$$\iint_{\text{plane}} e^{-r^2/2}\,r\,dr\,d\theta = \int_{\theta=0}^{2\pi}\int_{r=0}^{\infty} e^{-r^2/2}\,r\,dr\,d\theta$$

The r and θ integrals give the answers 1 and 2π :

$$\int_{r=0}^{\infty} e^{-r^2/2}\,r\,dr = \left[-e^{-r^2/2}\right]_{r=0}^{\infty} = 1 \qquad \int_{\theta=0}^{2\pi} 1\,d\theta = 2\pi.$$

The trick was to get $e^{-r^2/2}\,r\,dr$ (which is a perfect derivative of $-e^{-r^2/2}$) by combining $e^{-x^2/2}\,dx$ and $e^{-y^2/2}\,dy$ (which can *not* be separately integrated from a to b).

Problem Set 12.2, page 554

1 (a) Mean $m = \mathrm{E}\,[x] = (0)(1-p) + (1)(p) = p$ when the probability of heads is p. Here $x = 0$ for tails and $x = 1$ for heads. Notice that $0^2 = 0$ and $1^2 = 1$ so $\mathrm{E}\,[x^2] = \mathrm{E}\,[x] = p$.

$$\text{Variance}\ \ \boldsymbol{\sigma^2} = \mathrm{E}\,[x^2] - m^2 = \boldsymbol{p - p^2}$$

(b) These are independent flips ! So the N by N covariance matrix V is diagonal. The diagonal entries are the variances $\sigma^2 = p - p^2$ for each flip. Then the rule $(16-17-18)$ gives the overall variance of the sum from N flips :

$$\text{overall variance } = [1\ \ 1 \ldots 1]\, V \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = N\sigma^2 = N(p - p^2)$$

This is just saying : Add the variances for the N independent experiments. Here those N experiments just repeat one experiment.

2 I am just imitating equation (2) in the text. Now the experiments are numbered 3 and 5. They have means m_3 and m_5. The covariance σ_{35} adds up **joint probabilities** p_{ij} times (distance $x_i - m_3$) times (distance $y_j - m_5$). Here x_i and y_j are outputs from experiments 3 and 5 :

$$\boldsymbol{\sigma_{35}} = \sum_{\text{all}} \sum_{i,j} p_{ij}\,(x_i - m_3)\,(y_j - m_5).$$

3 The 3 by 3 covariance matrix V will be a sum of rank one matrices $V_{ijk} = UU^{\mathrm{T}}$ multiplied by the joint probability p_{ijk} of outputs x_i, y_j, z_k. I am copying equation (12) :

$$V = \sum_{\text{all}} \sum_{i,} \sum_{j,\ k} p_{ijk}\, UU^{\mathrm{T}} \qquad U = \begin{bmatrix} \text{output } x_i - \text{mean } \overline{x} \\ \text{output } y_j - \text{mean } \overline{y} \\ \text{output } z_k - \text{mean } \overline{z} \end{bmatrix}$$

These matrices UU^{T} = column times row are positive semidefinite with rank 1 (unless $U = \mathbf{0}$). The sum V is positive *definite* unless the 3 experiments are dependent.

Notice that the means $\overline{x}, \overline{y}, \overline{z} = m_1, m_2, m_3$ have to be computed before the variances.

4 We are told that the 3 experiments are *independent*. Then the *covariances are zero* off the main diagonal of V. This covariance matrix only shows "covariances with itself" = "variances" $\sigma_1^2, \sigma_2^2, \sigma_3^2$ on the main diagonal.

$$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}.$$

5 The point is that some output $X = x_i$ must occur. So the possibilities are $Y = y_j$ and $X = x_1$, or $Y = y_j$ and $X = x_2$, or $Y = y_j$ and $X = x_3$ et cetera. The total probability of $Y = y_j$ is the sum of the conditional probabilities that $Y = y_j$ when $X = x_i$.

Here is another way to say this **law of total probability**. When $B_1, B_2, \ldots$ are separate disjoint outcomes that together account for all possible outcomes, then for any A

$$\text{Prob}\,(A) = \sum_i \text{Prob}\,(A \cap B_i) = \sum_i \text{Prob}\,(A|B_i)\,\text{Prob}\,(B_i).$$

6 Prob $(A|B)$ = **conditional probability** of A given B satisfies this axiom:

$$\text{Prob}\,(A \textit{ and } B) = \text{Prob}\,(A|B)\,\text{Prob}\,(B).$$

Reason: If both A and B occur, then B must occur—and knowing that B occurs, Prob $(A|B)$ gives the probability that A also occurs.

This axiom is saying that $\boldsymbol{p_{ij} = \textbf{Prob}\,(A|B)\,p_i}$

where B is the event $x = x_i$ which has Prob $(B) = p_i$.

7 The joint probabilities p_{ij} = Prob $(x = x_i$ and $y = y_j)$ are in the matrix P:

$$P = \begin{bmatrix} 0.1 & 0.3 \\ 0.2 & 0.4 \end{bmatrix} \text{ with entries adding to } 1.$$

Problem **6** says that Prob $(Y = y_2|X = x_1) = \dfrac{p_{12}}{p_{11} + p_{12}} = \dfrac{0.3}{0.1 + 0.3} = \dfrac{\mathbf{3}}{\mathbf{4}}$.

Problem **5** says that Prob $(X = x_1) = p_{11} + p_{12} = 0.1 + 0.3 = \mathbf{0.4}$.

8 This product rule of conditional probability is the axiom in Solution 12.2.6 above:

$$\text{Prob }(A \textbf{ and } B) = \text{Prob }(A \textbf{ given } B) \text{ times Prob }(B).$$

9 This discussion of Bayes' Theorem is much too compressed! Let me reproduce three equations from Wolfram MathWorld. Here A and B are possible "sets" = "outcomes from an experiment" and the simple-looking identity $(*)$ connects conditional and unconditional probabilities.

We know from **8** that $\quad$ Prob $(A$ **and** $B) =$ Prob $(A$ given $B)$ times Prob (B)

Reversing A and B gives $\quad$ Prob $(A$ **and** $B) =$ Prob $(B$ given $A)$ times Prob (A)

$$(*) \text{ Therefore Prob}\,(B \text{ given } A) = \frac{\text{Prob}\,(A \text{ given } B)\,\text{Prob}\,(B)}{\text{Prob}\,(A)}$$

MathWorld gives this extension to non-overlapping sets $A_1, \ldots, A_n$ whose union is A:

$$\text{Prob}\,(A_i \text{ given } A) = \frac{\text{Prob}\,(A_i)\,\text{Prob}\,(A \text{ given } A_i)}{\sum_j \text{Prob}\,(A_j)\,\text{Prob}\,(A \text{ given } A_j)}$$

Problem Set 12.3, page 560

1 The two equations from two measurements are

$$\begin{matrix} x = b_1 \\ x = b_2 \end{matrix} \quad \text{or} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \text{or} \quad A\boldsymbol{x} = \boldsymbol{b}.$$

The covariance matrix V is diagonal since the measurements are independent:

$$V = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}.$$

The weighted least squares equation is $A^{\mathrm{T}} V^{-1} A\widehat{\boldsymbol{x}} = A^{\mathrm{T}} V^{-1}\boldsymbol{b}$.

$$A^{\mathrm{T}} V^{-1} A = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

$$A^{\mathrm{T}} V^{-1} \boldsymbol{b} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{b_1}{\sigma_1^2} + \frac{b_2}{\sigma_2^2}$$

Then $\widehat{\boldsymbol{x}}$ is the ratio of those numbers:

$$\widehat{\boldsymbol{x}} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

The variance of that estimate $\widehat{x}$ should be written as in (13) :

$$\mathrm{E}\left[(\widehat{\boldsymbol{x}}-\boldsymbol{x})(\widehat{\boldsymbol{x}}-\boldsymbol{x})^{\mathrm{T}}\right]=(A^{\mathrm{T}}V^{-1}A)^{-1}=\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)^{-1}.$$

2 (a) In the limit $\sigma_2 \to 0$ the ratio $\widehat{\boldsymbol{x}}$ approaches b_2 because :

$$(\text{Multiply } \widehat{\boldsymbol{x}} \text{ above and below by } \sigma_1^2\,\sigma_2^2) \quad \widehat{\boldsymbol{x}}=\frac{b_1\sigma_2^2+b_2\sigma_1^2}{\sigma_2^2\;+\;\sigma_1^2}\to\frac{b_2\sigma_1^2}{\sigma_1^2}=\boldsymbol{b_2}.$$

The second equation $x=b_2$ is 100% accurate if its variance is $\sigma_2=0$.

(b) If $\sigma_2\to\infty$ then $1/\sigma_2^2\to 0$ and $\widehat{\boldsymbol{x}}\to\dfrac{b_1/\sigma_1^2}{1/\sigma_1^2}=\boldsymbol{b_1}$. We are getting *no information* from the totally unreliable measurement $x=b_2$.

3 The key fact of **independence** is in the equation $p(x,y)=p(x)\,p(y)$. Then

$$\iint p(x,y)\,dx\,dy=\iint p(x)\,p(y)\,dx\,dy=\int p(x)\,dx\int p(y)\,dy=(1)\,(1)=\mathbf{1}.$$

$$\begin{aligned}\iint(x+y)\,p(x,y)\,dx\,dy &= \iint x\,p(x)\,p(y)\,dx\,dy+\iint y\,p(x)\,p(y)\,dx\,dy\\ &= \int x\,p(x)\,dx\int p(y)\,dy+\int p(x)\,dx\int y\,p(y)\,dy\\ &= (m_x)\,(1)+(1)\,(m_y)=m_x+m_y.\end{aligned}$$

4 Continue Problem 3 to find variances σ_x^2 and σ_y^2 and to see covariance $\sigma_{xy}=0$:

$$\iint(x-m_x)^2\,p(x,y)\,dx\,dy=\int(x-m_x)^2\,p(x)\,dx\int p(y)\,dy=\boldsymbol{\sigma_x^2}$$

$$\iint(x-m_x)\,(y-m_y)\,p(x,y)\,dx\,dy=\int(x-m_x)\,p(x)\,dx\int(y-m_y)\,p(y)\,dy=(0)\,(0).$$

5 We are inverting a 2 by 2 matrix using $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}=\dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a\end{bmatrix}$:

$$V^{-1}=\begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}^{-1}=\frac{1}{\sigma_1^2\,\sigma_2^2-\sigma_{12}^2}\begin{bmatrix}\sigma_2^2 & -\sigma_{12}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}= \qquad \rho=\frac{\sigma_{12}}{\sigma_1\,\sigma_2}$$

$$\frac{1}{\sigma_1^2\,\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_2^2 & -\sigma_{12}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}=\frac{1}{1-\rho^2}\begin{bmatrix}1/\sigma_1^2 & -\rho/\sigma_1\,\sigma_2\\ -\rho/\sigma_1\,\sigma_2 & 1/\sigma_2^2\end{bmatrix}$$

6 The right hand side of $\widehat{x}_{k+1}$ shows the **gain factor** $\boldsymbol{1/(k+1)}$:

$$\widehat{x}_k+\frac{1}{k+1}(b_{k+1}-\widehat{x}_k)=\frac{b_1+\cdots+b_k}{k}+\frac{\boldsymbol{1}}{\boldsymbol{k+1}}\left(b_{k+1}-\frac{b_1+\cdots+b_k}{k}\right)=\frac{b_1+\cdots+b_{k+1}}{k+1}$$

Check that each number $b_1, b_2, \ldots, b_k, b_{k+1}$ is correctly divided by $k+1$:

$$\frac{1}{k}-\frac{1}{k+1}\,\frac{1}{k}=\frac{1}{k}\left(\frac{k+1\;\;-1}{k+1}\right)=\frac{1}{k+1}.$$

7 We are considering the case when all the measurements $b_1, b_2, \ldots, b_{k+1}$ have the same variance σ^2. We know that the correct variance of their average is $W_{k+1}=\sigma^2/(k+1)$. We want to see how this answer comes from equation (18) when we have the correct $W_k=\sigma^2/k$ from the previous step, and we update to W_{k+1} :

(18) says that $W_{k+1}^{-1}=W_k^{-1}+A_{k+1}^{\mathrm{T}}\,V_{k+1}^{-1}\,A_{k+1}=\dfrac{k}{\sigma^2}+[1]\,[1/\sigma^2]\,[1]=\dfrac{k}{\sigma^2}+\dfrac{1}{\sigma^2}=\dfrac{k+1}{\sigma^2}.$

So $W_{k+1}=\sigma^2/(k+1)$ is correct at the new step (and forever by induction).

8 The three equations have variances σ^2, s^2, σ^2 and they have *zero covariances*. (This makes the step-by-step Kalman filter possible.) We can divide the equations by σ, s, σ to produce *unit variances* (which lead to ordinary unweighted least squares). We are given $F=1$:

$$\begin{bmatrix} 1/\sigma & 0 \\ -1/s & 1/s \\ 0 & 1/\sigma \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \end{bmatrix}=\begin{bmatrix} b_0/\sigma \\ 0 \\ b_1/\sigma \end{bmatrix} \text{ is our } A\boldsymbol{x}=\boldsymbol{b}.$$

The normal equation (now unweighted) is $A^{\mathrm{T}}A\,\widehat{\boldsymbol{x}}=A^{\mathrm{T}}\boldsymbol{b}$:

$$\begin{bmatrix} \dfrac{1}{\sigma^2}+\dfrac{1}{s^2} & -\dfrac{1}{s^2} \\ -\dfrac{1}{s^2} & \dfrac{1}{\sigma^2}+\dfrac{1}{s^2} \end{bmatrix}\begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{bmatrix}=\begin{bmatrix} \dfrac{b_0}{\sigma^2} \\ \dfrac{b_1}{\sigma^2} \end{bmatrix}.$$

The determinant of this $A^{\mathrm{T}}A$ is $\det=\dfrac{1}{\sigma^4}+\dfrac{2}{\sigma^2 s^2}$. The solution is

$$\widehat{x}_1=\frac{1}{\det}\left(\frac{b_0}{\sigma^4}+\frac{b_0}{\sigma^2 s^2}+\frac{b_1}{\sigma^2 s^2}\right)\qquad \widehat{x}_2=\frac{1}{\det}\left(\frac{b_0}{\sigma^2 s^2}+\frac{b_1}{\sigma^2 s^2}+\frac{b_1}{\sigma^4}\right).$$

9 With $A = I$ and $\boldsymbol{u}^{\mathrm{T}} = \boldsymbol{v}^{\mathrm{T}} = [1 \;\; 2 \;\; 3]$ we can use the direct formula for M^{-1}:

$$(I - \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = I + \frac{\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}}{1 - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}} = I + \frac{1}{1-14}\begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix}$$

$$= \begin{bmatrix} 1-\frac{1}{13} & \frac{2}{13} & \frac{3}{13} \\ \frac{2}{13} & 1-\frac{4}{13} & \frac{6}{13} \\ \frac{3}{13} & \frac{6}{13} & 1-\frac{9}{13} \end{bmatrix}. \text{ Multiply } \boldsymbol{b} = \begin{bmatrix}2\\1\\4\end{bmatrix} \text{ to get } \boldsymbol{y} = \begin{bmatrix}2\\1\\4\end{bmatrix} - \frac{16}{13}\begin{bmatrix}1\\2\\3\end{bmatrix} = \frac{1}{13}\begin{bmatrix}10\\-19\\4\end{bmatrix}.$$

Instead of this formula for $(I = \boldsymbol{u}\,\boldsymbol{v}^{\mathrm{T}})^{-1}$, try steps 1 and 2:

Step 1 with $A = I$ gives $\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{z} = \boldsymbol{u}$.

Step 2 gives $\boldsymbol{y} = \boldsymbol{b} - \dfrac{\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{u}}{13}\,\boldsymbol{u} = \begin{bmatrix}2\\1\\4\end{bmatrix} - \dfrac{16}{13}\begin{bmatrix}1\\2\\3\end{bmatrix}$ as before.

10 We are asked to check that $M\boldsymbol{y} = \boldsymbol{b}$ using the update formula. Start with

$$M\boldsymbol{y} = (A - \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})\left(\boldsymbol{x} + \frac{\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x}}{c}\,\boldsymbol{z}\right)$$
$$= A\boldsymbol{x} - \boldsymbol{u}\,(\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x}) + \frac{\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x}\,A\boldsymbol{z}}{c} - \frac{\boldsymbol{u}\,(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{z})\,(\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x})}{c}$$

Since $A\boldsymbol{x} = \boldsymbol{b}$ we hope the other 3 terms combine to give zero when $A\boldsymbol{z} = \boldsymbol{u}$

$$\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x}\left[-1 + \frac{1}{c} - \frac{\boldsymbol{v}^{\mathrm{T}}\boldsymbol{z}}{c}\right] = \frac{\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{x}}{c}\left[-c + 1 - \boldsymbol{v}^{\mathrm{T}}\,\boldsymbol{z}\right] = \boldsymbol{0} \text{ from the formula for } c$$

11 Multiply **columns times rows** to see that the new $\boldsymbol{v}$ changes $A^{\mathrm{T}}A$ to

$$\begin{bmatrix} A^{\mathrm{T}} & \boldsymbol{v} \end{bmatrix}\begin{bmatrix} A \\ \boldsymbol{v}^{\mathrm{T}} \end{bmatrix} = A^{\mathrm{T}}A + \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$$

So adding the new row to A (and of course the new column to A^{T}) has increased $A^{\mathrm{T}}\,A$ by the rank one matrix $\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$.

The book is ending with matrix multiplication! We could allow changes of rank r :

When A changes to $M = A - UW^{-1}V$, its inverse changes to

$$\boldsymbol{M^{-1}} = A^{-1} + A^{-1}\,U(W - VA^{-1}\,U)^{-1}\,VA^{-1}.$$

This change has rank r when $W_{r\times r}$ and $V_{r\times n}$ and $U_{n\times r}$ all have rank r.