7.2 Positive Definite Matrices and the SVD

This chapter about applications of $A^{\mathrm{T}}A$ depends on two important ideas in linear algebra. These ideas have big parts to play, we focus on them now.

1. Positive definite symmetric matrices (both $A^{\mathrm{T}}A$ and $A^{\mathrm{T}}CA$ are positive definite)

2. Singular Value Decomposition ($A = U\Sigma V^{\mathrm{T}}$ gives perfect bases for the 4 subspaces)

Those are orthogonal matrices U and V in the SVD. Their columns are orthonormal eigenvectors of AA^{T} and $A^{\mathrm{T}}A$. The entries in the diagonal matrix Σ are the *square roots* of the eigenvalues. The matrices AA^{T} and $A^{\mathrm{T}}A$ have the same nonzero eigenvalues.

Section 6.5 showed that the eigenvectors of these symmetric matrices are orthogonal. I will show now that *the eigenvalues of $A^{\mathrm{T}}A$ are positive*, if A has independent columns.

Start with $A^{\mathrm{T}}A\boldsymbol{x} = \lambda\boldsymbol{x}$. Then $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = \lambda\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$. Therefore $\lambda = ||A\boldsymbol{x}||^2/||\boldsymbol{x}||^2 > 0$

I separated $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}$ into $(A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x}) = ||A\boldsymbol{x}||^2$. We don't have $\lambda = 0$ because $A^{\mathrm{T}}A$ is invertible (since A has independent columns). The eigenvalues must be positive.

Those are the key steps to understanding positive definite matrices. They give us three tests on S—three ways to recognize when a symmetric matrix S is positive definite :

Positive	**1.**	All the eigenvalues of S are positive.
definite	**2.**	The "energy" $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$ is positive for all nonzero vectors $\boldsymbol{x}$.
symmetric	**3.**	S has the form $S = A^{\mathrm{T}}A$ with independent columns in A.

There is also a test on the pivots (all > 0) and a test on n determinants (all > 0).

Example 1 Are these matrices positive definite ? When their eigenvalues are positive, construct matrices A with $S = A^{\mathrm{T}}A$ and find the positive energy $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$.

$$\text{(a)}\quad S = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{(b)}\quad S = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \qquad \text{(c)}\quad S = \begin{bmatrix} 4 & 5 \\ 5 & 4 \end{bmatrix}$$

Solution The answers are *yes*, *yes*, and *no*. The eigenvalues of those matrices S are

(a) 4 and 1 : *positive* (b) 9 and 1 : *positive* (c) 9 and -1 : *not positive*.

A quicker test than eigenvalues uses **two determinants** : the 1 by 1 determinant S_{11} and the 2 by 2 determinant of S. Example (b) has $S_{11} = \mathbf{5}$ and det $S = 25 - 16 = \mathbf{9}$ (*pass*). Example (c) has $S_{11} = \mathbf{4}$ but det $S = 16 - 25 = \mathbf{-9}$ (*fail the test*).

Positive energy is equivalent to positive eigenvalues, when S is symmetric. Let me test the energy $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ in all three examples. Two examples pass and the third fails :

$$[x_1 \quad x_2]\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4x_1^2 + x_2^2 > 0 \qquad \text{Positive energy when } \boldsymbol{x} \neq \boldsymbol{0}$$

$$[x_1 \quad x_2]\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5x_1^2 + 8x_1x_2 + 5x_2^2 \qquad \text{Positive energy when } \boldsymbol{x} \neq \boldsymbol{0}$$

$$[x_1 \quad x_2]\begin{bmatrix} 4 & 5 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4x_1^2 + 10x_1x_2 + 4x_2^2 \qquad \text{Energy } \boldsymbol{-2} \text{ when } \boldsymbol{x} = (1, -1)$$

Positive energy is a fundamental property. This is the best definition of *positive definiteness*.

When the eigenvalues are positive, there will be many matrices A that give $A^{\mathrm{T}}A = S$. One choice of A is symmetric and positive definite! Then $A^{\mathrm{T}}A$ is A^2, and this choice $A = \sqrt{S}$ is a true square root of S. The successful examples (a) and (b) have $\boldsymbol{S = A^2}$:

$$\boxed{\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}$$

We know that all symmetric matrices have the form $S = V\Lambda V^{\mathrm{T}}$ with orthonormal eigenvectors in V. The diagonal matrix Λ has a square root $\sqrt{\Lambda}$, when all eigenvalues are positive. In this case $A = \sqrt{S} = V\sqrt{\Lambda}V^{\mathrm{T}}$ is the symmetric positive definite square root:

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \sqrt{S}\sqrt{S} = (V\sqrt{\Lambda}V^{\mathrm{T}})(V\sqrt{\Lambda}V^{\mathrm{T}}) = V\sqrt{\Lambda}\sqrt{\Lambda}V^{\mathrm{T}} = \boldsymbol{S} \quad \text{because } V^{\mathrm{T}}V = I.$$

Starting from this unique square root $\sqrt{S}$, other choices of A come easily. Multiply $\sqrt{S}$ by any matrix Q that has orthonormal columns (so that $Q^{\mathrm{T}}Q = I$). Then $Q\sqrt{S}$ is another choice for A (not a symmetric choice). In fact all choices come this way :

$$A^{\mathrm{T}}A = (Q\sqrt{S})^{\mathrm{T}}\,(Q\sqrt{S}) = \sqrt{S}Q^{\mathrm{T}}Q\sqrt{S} = S. \tag{1}$$

I will choose a particular Q in Example 1, to get particular choices of A.

Example 1 (*continued*) Choose $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ to multiply $\sqrt{S}$. Then $A = Q\sqrt{S}$.

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix} \qquad \text{has} \quad S = A^{\mathrm{T}}A = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 2 & 1 \end{bmatrix} \qquad \text{has} \quad S = A^{\mathrm{T}}A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

Positive Semidefinite Matrices

Positive *semidefinite* matrices include positive definite matrices, and more. Eigenvalues of S can be zero. Columns of A can be dependent. The energy $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$ can be zero—*but not negative*. This gives new equivalent conditions on a (possibly singular) matrix $S = S^{\mathrm{T}}$.

1′ All eigenvalues of S satisfy $\lambda \geq 0$ (semidefinite allows zero eigenvalues).

2′ The energy is nonnegative for every $\boldsymbol{x} : \boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} \geq 0$ (zero energy is allowed).

3′ S has the form $A^{\mathrm{T}}A$ (every A is allowed; its columns can be dependent).

Example 2 The first two matrices are singular and positive semidefinite—but not the third :

$$\text{(d)} \quad S = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{(e)} \quad S = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \qquad \text{(f)} \quad S = \begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix}.$$

The eigenvalues are $1, 0$ and $8, 0$ and $-8, 0$. The energies $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$ are x_2^2 and $4(x_1 + x_2)^2$ and $-4(x_1 - x_2)^2$. So the third matrix is actually *negative* semidefinite.

Singular Value Decomposition

Now we start with A, square or rectangular. Applications also start this way—the matrix comes from the model. The SVD splits any matrix into *orthogonal* U times *diagonal* Σ times *orthogonal* V^{T}. Those orthogonal factors will give orthogonal bases for the four fundamental subspaces associated with A.

Let me describe the goal for any m by n matrix, and then how to achieve that goal.

Find orthonormal bases $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ **for** $\mathbf{R}^n$ **and** $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ **for** $\mathbf{R}^m$ **so that**

$$\boldsymbol{A}\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1 \quad \ldots \quad \boldsymbol{A}\boldsymbol{v}_r = \sigma_r \boldsymbol{u}_r \qquad \boldsymbol{A}\boldsymbol{v}_{r+1} = \boldsymbol{0} \quad \ldots \quad \boldsymbol{A}\boldsymbol{v}_n = \boldsymbol{0} \tag{2}$$

The rank of A is r. Those requirements in (4) are expressed by a multiplication $\boldsymbol{AV} = \boldsymbol{U\Sigma}$. The r nonzero singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ are on the diagonal of Σ :

$$\boldsymbol{AV} = \boldsymbol{U\Sigma} \qquad A\begin{bmatrix} \\ \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_r & \ldots & \boldsymbol{v}_n \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_r & \ldots & \boldsymbol{u}_m \\ \\ \end{bmatrix} \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{bmatrix} \tag{3}$$

The last $n - r$ vectors in V are a basis for the nullspace of A. The last $m - r$ vectors in U are a basis for the nullspace of A^{T}. The diagonal matrix Σ is m by n, with r nonzeros.

Remember that $V^{-1} = V^{\mathrm{T}}$, because the columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are orthonormal in $\mathbf{R}^n$:

$$\textbf{Singular Value Decomposition} \qquad \boldsymbol{AV} = \boldsymbol{U\Sigma} \quad \textbf{becomes} \quad \boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathbf{T}}. \tag{4}$$

The SVD has orthogonal matrices U and V, containing eigenvectors of AA^{T} and $A^{\mathrm{T}}A$.

Comment. A square matrix is diagonalized by its eigenvectors : $A\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$ is like $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$. But even if A has n eigenvectors, they may not be orthogonal. We need *two bases*—an input basis of $\boldsymbol{v}$'s in $\mathbf{R}^n$ and an output basis of $\boldsymbol{u}$'s in $\mathbf{R}^m$. With two bases, any m by n matrix can be diagonalized. The beauty of those bases is that they can be chosen orthonormal. Then $U^{\mathrm{T}}U = I$ and $V^{\mathrm{T}}V = I$.

The $\boldsymbol{v}$'s are eigenvectors of the symmetric matrix $S = A^{\mathrm{T}}A$. We can guarantee their orthogonality, so that $\boldsymbol{v}_j^{\mathrm{T}}\boldsymbol{v}_i = 0$ for $j \neq i$. That matrix S is positive semidefinite, so its eigenvalues are $\sigma_i^2 \geq 0$. **The key to the SVD is that $A\boldsymbol{v}_j$ is orthogonal to $A\boldsymbol{v}_i$** :

$$\textbf{Orthogonal } \boldsymbol{u}\textbf{'s} \quad (A\boldsymbol{v}_j)^{\mathrm{T}}(A\boldsymbol{v}_i) = \boldsymbol{v}_j^{\mathrm{T}}(A^{\mathrm{T}}A\boldsymbol{v}_i) = \boldsymbol{v}_j^{\mathrm{T}}(\sigma_i^2\boldsymbol{v}_i) = \begin{cases} \sigma_i^2 & \text{if } j = i \\ \mathbf{0} & \text{if } j \neq i \end{cases} \tag{5}$$

This says that the vectors $\boldsymbol{u}_i = A\boldsymbol{v}_i/\sigma_i$ are orthonormal for $i = 1, \ldots, r$. They are a basis for the column space of A. And the $\boldsymbol{u}$'s are eigenvectors of the symmetric matrix AA^{T}, which is usually different from $S = A^{\mathrm{T}}A$ (but the eigenvalues $\sigma_1^2, \ldots, \sigma_r^2$ are the same).

Example 3 Find the input and output eigenvectors $\boldsymbol{v}$ and $\boldsymbol{u}$ for the rectangular matrix A :

$$A = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{bmatrix} = U\Sigma V^{\mathrm{T}}.$$

Solution Compute $S = A^{\mathrm{T}}A$ and its unit eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$. The eigenvalues σ^2 are $8, 2, 0$ so the positive singular values are $\sigma_1 = \sqrt{8}$ and $\sigma_2 = \sqrt{2}$:

$$A^{\mathrm{T}}A = \begin{bmatrix} 5 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{has} \quad \boldsymbol{v}_1 = \frac{1}{2}\begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \frac{1}{2}\begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The outputs $\boldsymbol{u}_1 = A\boldsymbol{v}_1/\sigma_1$ and $\boldsymbol{u}_2 = A\boldsymbol{v}_2/\sigma_2$ are also orthonormal, with $\sigma_1 = \sqrt{8}$ and $\sigma_2 = \sqrt{2}$. Those vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are in the column space of A :

$$\boldsymbol{u}_1 = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{bmatrix}\frac{\boldsymbol{v}_1}{\sqrt{8}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{u}_2 = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{bmatrix}\frac{\boldsymbol{v}_2}{\sqrt{2}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then $U = I$ and the Singular Value Decomposition for this 2 by 3 matrix is $U\Sigma V^{\mathrm{T}}$:

$$A = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{\sqrt{8}} & 0 & 0 \\ 0 & \boldsymbol{\sqrt{2}} & 0 \end{bmatrix}\frac{1}{2}\begin{bmatrix} \sqrt{2} & \sqrt{2} & 0 \\ \sqrt{2} & -\sqrt{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}^{\mathrm{T}}.$$

The Fundamental Theorem of Linear Algebra

I think of the SVD as the final step in the Fundamental Theorem. First come the *dimensions* of the four subspaces in Figure 7.3. Then come the *orthogonality* of those pairs of subspaces. Now come the *orthonormal bases of $\boldsymbol{v}$'s and $\boldsymbol{u}$'s that diagonalize A*:

SVD $\quad \boxed{\begin{aligned} A\boldsymbol{v}_j &= \sigma_j\boldsymbol{u}_j \quad \text{for } j \le r \\ A\boldsymbol{v}_j &= \boldsymbol{0} \quad\quad\ \text{for } j > r \end{aligned}} \qquad \boxed{\begin{aligned} A^{\mathrm{T}}\boldsymbol{u}_j &= \sigma_j\boldsymbol{v}_j \quad \text{for } j \le r \\ A^{\mathrm{T}}\boldsymbol{u}_j &= \boldsymbol{0} \quad\quad\ \text{for } j > r \end{aligned}}$

Multiplying $A\boldsymbol{v}_j = \sigma_j\boldsymbol{u}_j$ by A^{T} and dividing by σ_j gives that equation $\boldsymbol{A^{\mathrm{T}}u_j = \sigma_j v_j}$.

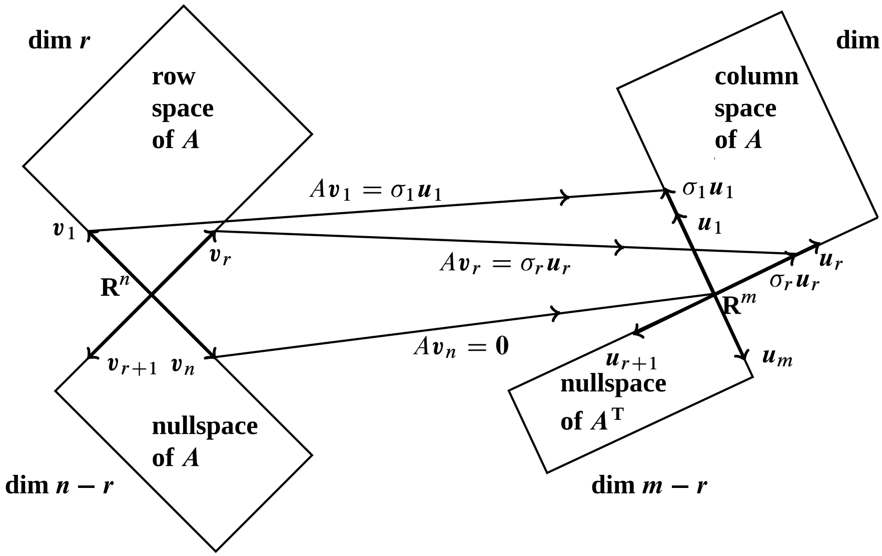

Figure 7.3: Orthonormal bases of $\boldsymbol{v}$'s and $\boldsymbol{u}$'s that diagonalize A : m by n with rank r.

The "**norm**" of A is its largest singular value : $\boldsymbol{||A|| = \sigma_1}$. This measures the largest possible ratio of $||A\boldsymbol{v}||$ to $||\boldsymbol{v}||$. That ratio of lengths is a maximum when $\boldsymbol{v} = \boldsymbol{v}_1$ and $A\boldsymbol{v} = \sigma_1\boldsymbol{u}_1$. This singular value σ_1 is a much better measure for the size of a matrix than the largest eigenvalue. An extreme case can have zero eigenvalues and just one eigenvector $(1, 1)$ for A. But $A^{\mathrm{T}}A$ can still be large : if $\boldsymbol{v} = (1, -1)$ then $A\boldsymbol{v}$ is 200 times larger.

$$A = \begin{bmatrix} 100 & -100 \\ 100 & -100 \end{bmatrix} \text{ has } \lambda_{\max} = \boldsymbol{0}. \quad \text{But } \boldsymbol{\sigma_{\max}} = \textbf{norm of } \boldsymbol{A} = \boldsymbol{200}. \tag{6}$$

The Condition Number

A valuable property of $A = U\Sigma V^{\mathrm{T}}$ is that it puts the pieces of A *in order of importance*. Multiplying a column $\boldsymbol{u}_i$ times a row $\sigma_i \boldsymbol{v}_i^{\mathrm{T}}$ produces one piece of the matrix. There will be r nonzero pieces from r nonzero σ's, when A has rank r. The pieces add up to A, when we multiply columns of U times rows of ΣV^{T} :

The pieces have rank 1
$$A = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 \boldsymbol{v}_1^{\mathrm{T}} \\ \dots\dots \\ \sigma_r \boldsymbol{v}_r^{\mathrm{T}} \end{bmatrix} = \boldsymbol{u}_1(\sigma_1 \boldsymbol{v}_1^{\mathrm{T}}) + \cdots + \boldsymbol{u}_r(\sigma_r \boldsymbol{v}_r^{\mathrm{T}}). \qquad (7)$$

The first piece gives the norm of A which is σ_1. The last piece gives the norm of A^{-1}, which is $1/\sigma_n$ when A is invertible. The **condition number** is σ_1 times $1/\sigma_n$:

$$\boxed{\textbf{Condition number of } \boldsymbol{A} \qquad \boldsymbol{c(A)} = ||A||\ ||A^{-1}|| = \frac{\boldsymbol{\sigma_1}}{\boldsymbol{\sigma_n}}.} \qquad (8)$$

This number $c(A)$ is the key to numerical stability in solving $A\boldsymbol{v} = \boldsymbol{b}$. When A is an orthogonal matrix, the symmetric $S = A^{\mathrm{T}}A$ is the identity matrix. So all singular values of an orthogonal matrix are $\sigma = 1$. At the other extreme, a singular matrix has $\sigma_n = 0$. In that case $c = \infty$. Orthogonal matrices have the best condition number $c = 1$.

Data Matrices : Application of the SVD

"*Big data*" is the linear algebra problem of this century (and we won't solve it here). Sensors and scanners and imaging devices produce enormous volumes of information. Making decisive sense of that data is *the* problem for a world of analysts (mathematicians and statisticians of a new type). Most often the data comes in the form of a matrix.

The usual approach is by PCA—*Principal Component Analysis*. That is essentially the SVD. The first piece $\sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}}$ holds the most information (in statistics this piece has the greatest variance). It tells us the most. The Chapter 7 Notes include references.

■ REVIEW OF THE KEY IDEAS ■

1. Positive definite symmetric matrices have positive eigenvalues and pivots and energy.
2. $S = A^{\mathrm{T}}A$ is positive definite if and only if A has independent columns.
3. $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x})$ is zero when $A\boldsymbol{x} = \boldsymbol{0}$. $A^{\mathrm{T}}A$ can be positive *semidefinite*.
4. The SVD is a factorization $A = U\Sigma V^{\mathrm{T}} =$ (*orthogonal*) (*diagonal*) (*orthogonal*).
5. The columns of V and U are eigenvectors of $A^{\mathrm{T}}A$ and AA^{T} (singular vectors of A).
6. Those orthonormal bases achieve $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ and A is diagonalized.
7. The largest piece of $A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{T}}$ gives the norm $||A|| = \sigma_1$.

Problem Set 7.2

1 For a 2 by 2 matrix, suppose the 1 by 1 and 2 by 2 determinants a and $ac - b^2$ are positive. Then $c > b^2/a$ is also positive.

(i) λ_1 and λ_2 have the *same sign* because their product $\lambda_1\lambda_2$ equals ______.

(i) That sign is positive because $\lambda_1 + \lambda_2$ equals ______.

Conclusion: The tests $a > 0, ac - b^2 > 0$ guarantee positive eigenvalues λ_1, λ_2.

2 Which of S_1, S_2, S_3, S_4 has two positive eigenvalues? Use a and $ac - b^2$, don't compute the λ's. Find an $\boldsymbol{x}$ with $\boldsymbol{x}^{\mathrm{T}} S_1 \boldsymbol{x} < 0$, confirming that A_1 fails the test.

$$S_1 = \begin{bmatrix} 5 & 6 \\ 6 & 7 \end{bmatrix} \quad S_2 = \begin{bmatrix} -1 & -2 \\ -2 & -5 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 10 \\ 10 & 100 \end{bmatrix} \quad S_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}.$$

3 For which numbers b and c are these matrices positive definite?

$$S = \begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix} \qquad S = \begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix} \qquad S = \begin{bmatrix} c & b \\ b & c \end{bmatrix}.$$

4 What is the energy $q = ax^2 + 2bxy + cy^2 = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ for each of these matrices? Complete the square to write q as a sum of squares $d_1(\quad)^2 + d_2(\quad)^2$.

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}.$$

5 $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = 2x_1x_2$ certainly has a saddle point and not a minimum at $(0,0)$. What symmetric matrix S produces this energy? What are its eigenvalues?

6 Test to see if $A^{\mathrm{T}}A$ is positive definite in each case:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

7 Which 3 by 3 symmetric matrices S and T produce these quadratic energies?

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = 2\left(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3\right). \quad \text{Why is } S \text{ positive definite?}$$
$$\boldsymbol{x}^{\mathrm{T}} T \boldsymbol{x} = 2\left(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3\right). \quad \text{Why is } T \text{ semidefinite?}$$

8 Compute the three upper left determinants of S to establish positive definiteness. (The first is 2.) Verify that their ratios give the second and third pivots.

$$\textbf{Pivots = ratios of determinants} \qquad S = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{bmatrix}.$$

9 For what numbers c and d are S and T positive definite? Test the 3 determinants:

$$S = \begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & d & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

10 *If S is positive definite then S^{-1} is positive definite.* Best proof: The eigenvalues of S^{-1} are positive because ______. *Second proof* (only for 2 by 2):

$$\text{The entries of } S^{-1} = \frac{1}{ac-b^2}\begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \quad \text{pass the determinant tests} \quad \_\_\_\_\_\_.$$

11 *If S and T are positive definite, their sum $S+T$ is positive definite.* Pivots and eigenvalues are not convenient for $S+T$. Better to prove $\boldsymbol{x}^{\mathrm{T}}(S+T)\boldsymbol{x} > 0$.

12 A positive definite matrix *cannot have a zero* (or even worse, a negative number) *on its diagonal*. Show that this matrix fails to have $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} > 0$:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & \mathbf{0} & 2 \\ 1 & 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{is not positive when } (x_1, x_2, x_3) = (\quad,\quad,\quad).$$

13 A diagonal entry a_{jj} of a symmetric matrix cannot be smaller than all the λ's. If it were, then $A - a_{jj}I$ would have ______ eigenvalues and would be positive definite. But $A - a_{jj}I$ has a ______ on the main diagonal.

14 Show that *if all $\lambda > 0$ then $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x} > 0$.* We must do this for *every* nonzero $\boldsymbol{x}$, not just the eigenvectors. So write $\boldsymbol{x}$ as a combination of the eigenvectors and *explain why all "cross terms" are $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$.* Then $\boldsymbol{x}^{\mathrm{T}}S\boldsymbol{x}$ is

$$(c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n)^{\mathrm{T}}(c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n x_n) = c_1^2\lambda_1\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{x}_1 + \cdots + c_n^2\lambda_n\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{x}_n > 0.$$

15 Give a quick reason why each of these statements is true:

(a) Every positive definite matrix is invertible.

(b) The only positive definite projection matrix is $P = I$.

(c) A diagonal matrix with positive diagonal entries is positive definite.

(d) A symmetric matrix with a positive determinant might not be positive definite!

16 With positive pivots in D, the factorization $S = LDL^{\mathrm{T}}$ becomes $L\sqrt{D}\sqrt{D}L^{\mathrm{T}}$. (Square roots of the pivots give $D = \sqrt{D}\sqrt{D}$.) Then $A = \sqrt{D}L^{\mathrm{T}}$ yields the ***Cholesky factorization*** $S = A^{\mathrm{T}}A$ which is "symmetrized LU":

$$\text{From} \quad A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{find } S. \qquad \text{From} \quad S = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} \quad \text{find } A = \mathbf{chol}(S).$$

17 Without multiplying $S = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, find

(a) the determinant of S (b) the eigenvalues of S
(c) the eigenvectors of S (d) a reason why S is symmetric positive definite.

18 For $F_1(x,y) = \frac{1}{4}x^4 + x^2y + y^2$ and $F_2(x,y) = x^3 + xy - x$ find the second derivative matrices H_1 and H_2:

$$\textbf{Test for minimum} \qquad H = \begin{bmatrix} \partial^2F/\partial x^2 & \partial^2F/\partial x\partial y \\ \partial^2F/\partial y\partial x & \partial^2F/\partial y^2 \end{bmatrix} \text{ is positive definite}$$

H_1 is positive definite so F_1 is concave up ($=$ convex). Find the minimum point of F_1 and the saddle point of F_2 (look only where first derivatives are zero).

19 The graph of $z = x^2 + y^2$ is a bowl opening upward. The graph of $z = x^2 - y^2$ is *a saddle*. The graph of $z = -x^2 - y^2$ is a bowl opening downward. What is a test on a, b, c for $z = ax^2 + 2bxy + cy^2$ to have a saddle point at $(0,0)$?

20 Which values of c give a bowl and which c give a saddle point for the graph of $z = 4x^2 + 12xy + cy^2$? Describe this graph at the borderline value of c.

21 When S and T are symmetric positive definite, ST might not even be symmetric. But its eigenvalues are still positive. Start from $ST\boldsymbol{x} = \lambda\boldsymbol{x}$ and take dot products with $T\boldsymbol{x}$. Then prove $\lambda > 0$.

22 Suppose C is positive definite (so $\boldsymbol{y}^{\mathrm{T}}C\boldsymbol{y} > 0$ whenever $\boldsymbol{y} \neq \boldsymbol{0}$) and A has independent columns (so $A\boldsymbol{x} \neq \boldsymbol{0}$ whenever $\boldsymbol{x} \neq \boldsymbol{0}$). Apply the energy test to $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}CA\boldsymbol{x}$ to show that $A^{\mathrm{T}}CA$ is positive definite: *the crucial matrix in engineering.*

23 Find the eigenvalues and unit eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ of $A^{\mathrm{T}}A$. Then find $\boldsymbol{u}_1 = A\boldsymbol{v}_1/\sigma_1$:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \text{ and } A^{\mathrm{T}}A = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix} \text{ and } AA^{\mathrm{T}} = \begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix}.$$

Verify that $\boldsymbol{u}_1$ is a unit eigenvector of AA^{T}. Complete the matrices U, Σ, V.

$$\textbf{SVD} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}}.$$

24 Write down orthonormal bases for the four fundamental subspaces of this A.

25 (a) Why is the trace of $A^{\mathrm{T}}A$ equal to the sum of all a_{ij}^2?

(b) For every rank-one matrix, why is $\sigma_1^2 =$ sum of all a_{ij}^2?

26 Find the eigenvalues and unit eigenvectors of $A^{\mathrm{T}}A$ and AA^{T}. Keep each $A\boldsymbol{v} = \sigma\boldsymbol{u}$:

$$\textbf{Fibonacci matrix} \qquad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Construct the singular value decomposition and verify that A equals $U\Sigma V^{\mathrm{T}}$.

27 Compute $A^{\mathrm{T}}A$ and AA^{T} and their eigenvalues and unit eigenvectors for V and U.

$$\textbf{Rectangular matrix} \qquad A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Check $AV = U\Sigma$ (this will decide $\pm$ signs in U). Σ *has the same shape as* A.

28 Construct the matrix with rank one that has $A\boldsymbol{v} = 12\boldsymbol{u}$ for $\boldsymbol{v} = \frac{1}{2}(1,1,1,1)$ and $\boldsymbol{u} = \frac{1}{3}(2,2,1)$. Its only singular value is $\sigma_1 =$ ______.

29 Suppose A is invertible (with $\sigma_1 > \sigma_2 > 0$). Change A by ***as small a matrix as possible*** to produce a singular matrix A_0. Hint: U and V do not change.

$$\text{From} \qquad A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{\mathrm{T}} \qquad \text{find the nearest } A_0.$$

30 The SVD for $A + I$ doesn't use $\Sigma + I$. *Why is* $\sigma(A+I)$ *not just* $\sigma(A) + I$?

31 Multiply $A^{\mathrm{T}}A\boldsymbol{v} = \sigma^2\boldsymbol{v}$ by A. Put in parentheses to show that $A\boldsymbol{v}$ is an eigenvector of AA^{T}. We divide by its length $||A\boldsymbol{v}|| = \sigma$ to get the unit eigenvector $\boldsymbol{u}$.

32 My favorite example of the SVD is when $Av(x) = dv/dx$, with the endpoint conditions $v(0) = 0$ and $v(1) = 0$. We are looking for orthogonal functions $v(x)$ so that their derivatives $Av = dv/dx$ are also orthogonal. The perfect choice is $v_1 = \sin \pi x$ and $v_2 = \sin 2\pi x$ and $v_k = \sin k\pi x$. Then each u_k is a cosine.

The derivative of v_1 is $Av_1 = \pi \cos \pi x = \pi u_1$. The singular values are $\sigma_1 = \pi$ and $\sigma_k = k\pi$. Orthogonality of the sines (and orthogonality of the cosines) is the foundation for Fourier series.

You may object to $AV = U\Sigma$. The derivative $A = d/dx$ is not a matrix! The orthogonal factor V has functions $\sin k\pi x$ in its columns, not vectors. The matrix U has cosine functions $\cos k\pi x$. Since when is this allowed? One answer is to refer you to the **chebfun** package on the web. This extends linear algebra to matrices whose columns are functions—not vectors.

Another answer is to replace d/dx by a first difference matrix A. Its shape will be $N+1$ by N. A has 1's down the diagonal and -1's on the diagonal below. Then $AV = U\Sigma$ has discrete sines in V and discrete cosines in U. For $N = 2$ those will be sines and cosines of 30° and 60° in $\boldsymbol{v}_1$ and $\boldsymbol{u}_1$.

****** Can you construct the difference matrix A (3 by 2) and $A^{\mathrm{T}}A$ (2 by 2)? The discrete sines are $\boldsymbol{v}_1 = (\sqrt{3}/2, \sqrt{3}/2)$ and $\boldsymbol{v}_2 = (\sqrt{3}/2, -\sqrt{3}/2)$. Test that $A\boldsymbol{v}_1$ is orthogonal to $A\boldsymbol{v}_2$. What are the singular values σ_1 and σ_2 in Σ ?