MATH 6610 Companion Notes

Numerical Analysis

Companion text for MATH 6610

Contents

Part I

Algebraic and Numerical Foundations

Chapter 1

Linear Algebraic Foundations

This chapter explores linear algebra from a numerical analysis point of view. In particular, vectors and matrices can be interpreted both as algebraic and geometric objects. This chapter fixes notation and develops the geometric tools that will be used throughout the course.

1.1 Vectors, matrices, and the ambient geometry

We work primarily over the fields $\mathbb{R}$ and $\mathbb{C}$. If $z = x + iy \in \mathbb{C}$, then its complex conjugate is $\overline{z} = x - iy$. For a scalar, the notation z^* has the same meaning as $\overline{z}$. For a vector or matrix, $*$ denotes the conjugate transpose; over $\mathbb{R}$ it reduces to the ordinary transpose.

A vector $\mathbf{x} \in \mathbb{C}^n$ has entries $x_1, \dots, x_n$. A matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ has entries a_{jk}, with $1 \le j \le m$ and $1 \le k \le n$. If $\mathbf{B} \in \mathbb{C}^{n\times k}$, their product is the $m \times k$ matrix determined by

$$(\mathbf{AB})_{j\ell} = \sum_{s=1}^{n} a_{js} b_{s\ell}. \tag{1.1}$$

When $\mathbf{A}$ and $\mathbf{B}$ can be multiplied in this way, we say they are of *conforming sizes*. Write the jth row of a matrix as $\mathbf{r}_j^*$ and its ℓth column as $\mathbf{c}_\ell$. For column vectors $\mathbf{c}$ and $\mathbf{r}$, the matrix $\mathbf{cr}^*$ is an *outer product*; if both vectors are nonzero, it has rank one.

Problem 1.1 (Matrix multiplication as inner and outer products). Let $\mathbf{A} \in \mathbb{C}^{K\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times L}$. Using only the entrywise definition (1.1), prove the following statements.

1. If the rows of $\mathbf{A}$ are $\mathbf{r}_k^*$ and the columns of $\mathbf{B}$ are $\mathbf{c}_\ell$, then

$$(\mathbf{AB})_{k\ell} = \mathbf{r}_k^* \mathbf{c}_\ell.$$

2. If $\mathbf{C} \in \mathbb{C}^{K\times n}$ has columns $\mathbf{c}_j$ and $\mathbf{R} \in \mathbb{C}^{n\times L}$ has rows $\mathbf{r}_j^*$, then

$$\mathbf{CR} = \sum_{j=1}^{n} \mathbf{c}_j \mathbf{r}_j^*.$$

3. Determine and justify the maximum possible rank of $\mathbf{CR}$ in terms of K, n, L.

Thus the ℓth column of $\mathbf{AB}$ is obtained by applying $\mathbf{A}$ to the ℓth column of $\mathbf{B}$. This operator viewpoint is especially useful in numerical linear algebra: an $m \times n$ matrix is a linear map from $\mathbb{C}^n$ to $\mathbb{C}^m$.

The standard inner product on $\mathbb{C}^n$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{j=1}^{n} x_j \overline{y_j}. \tag{1.2}$$

Our convention makes the inner product linear in its first argument and conjugate-linear in its second. It induces the Euclidean norm

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Vectors are orthogonal, written $\mathbf{x} \perp \mathbf{y}$, when $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. A collection is orthonormal when its members are pairwise orthogonal and each has norm one.

For a nonzero vector $\mathbf{v}$, the orthogonal projection of $\mathbf{x}$ onto its span is

$$\operatorname{Proj}_{\mathbf{v}} \mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}. \tag{1.3}$$

In a complex vector space an angle must be defined with some care because an inner product need not be real. One useful principal angle is specified by

$$\cos\theta = \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\|_2 \, \|\mathbf{y}\|_2}, \qquad 0 \le \theta \le \frac{\pi}{2}.$$

For real vectors one can omit the absolute value and recover the usual signed angle in $[0, \pi]$.

1.1.1 The four fundamental subspaces

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$. Its four fundamental subspaces are

$$\begin{aligned} \operatorname{range}(\mathbf{A}) &\subseteq \mathbb{C}^m, & \operatorname{corange}(\mathbf{A}) &:= \operatorname{range}(\mathbf{A}^*) \subseteq \mathbb{C}^n, \\ \mathcal{K}(\mathbf{A}) &:= \{\mathbf{x} \in \mathbb{C}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}, & \operatorname{coker}(\mathbf{A}) &:= \mathcal{K}(\mathbf{A}^*) \subseteq \mathbb{C}^m. \end{aligned}$$

The *rank* of $\mathbf{A}$ is the dimension of its range:

$$\operatorname{rank}(\mathbf{A}) := \dim \operatorname{range}(\mathbf{A}).$$

The row-rank/column-rank identity gives

$$\dim \operatorname{range}(\mathbf{A}) = \dim \operatorname{range}(\mathbf{A}^*) = \operatorname{rank}(\mathbf{A}).$$

Here *corange* is the row space, represented as a subspace of $\mathbb{C}^n$, and *cokernel* is used in the numerical-linear-algebra sense of left nullspace. The latter convention differs from the quotient-space definition sometimes used in abstract algebra.

To see the basic geometry, write the conjugate-transposed rows of $\mathbf{A}$ as $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathbb{C}^n$. The equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ says precisely that $\langle \mathbf{x}, \mathbf{w}_j \rangle = 0$ for every j. Consequently,

$$\mathcal{K}(\mathbf{A}) = \operatorname{range}(\mathbf{A}^*)^\perp.$$

Applying the same observation to $\mathbf{A}^*$ gives the companion identity $\mathcal{K}(\mathbf{A}^*) = \operatorname{range}(\mathbf{A})^\perp$.

Problem 1.2 (Fundamental-subspace inclusions for a product). Let $\mathbf{A}$ and $\mathbf{B}$ have conforming sizes. Prove that

1. $\operatorname{range}(\mathbf{AB}) \subseteq \operatorname{range}(\mathbf{A})$;
2. $\operatorname{corange}(\mathbf{AB}) \subseteq \operatorname{corange}(\mathbf{B})$;
3. $\mathcal{K}(\mathbf{AB}) \supseteq \mathcal{K}(\mathbf{B})$;
4. $\operatorname{coker}(\mathbf{AB}) \supseteq \operatorname{coker}(\mathbf{A})$.

Theorem 1.1 (Fundamental theorem of linear algebra)**.** *For every* $\mathbf{A} \in \mathbb{C}^{m\times n}$*,*

$$\begin{aligned} \mathbb{C}^n &= \operatorname{corange}(\mathbf{A}) \oplus \mathcal{K}(\mathbf{A}), & n &= \dim \operatorname{corange}(\mathbf{A}) + \dim \mathcal{K}(\mathbf{A}), \\ \mathbb{C}^m &= \operatorname{range}(\mathbf{A}) \oplus \operatorname{coker}(\mathbf{A}), & m &= \dim \operatorname{range}(\mathbf{A}) + \dim \operatorname{coker}(\mathbf{A}), \end{aligned}$$

and both direct sums are orthogonal.

The theorem says that every input splits uniquely into a part visible to $\mathbf{A}$ and a part annihilated by $\mathbf{A}$. Likewise, every possible output splits into a component that $\mathbf{A}$ can produce and a component orthogonal to all such outputs. Later, the singular value decomposition will provide orthonormal bases adapted simultaneously to all four spaces.

A rank-one perturbation changes a matrix by an outer product $\mathbf{uw}^*$. Although such a perturbation can change invertibility, its special form makes the change in the inverse equally structured when the updated matrix remains nonsingular.

Problem 1.3 (Sherman–Morrison rank-one update). Let $\mathbf{u}, \mathbf{w} \in \mathbb{C}^n$ be nonzero.

1. Prove that $\mathbf{I} + \mathbf{uw}^*$ is singular if and only if $\mathbf{w}^*\mathbf{u} = -1$.
2. If $\mathbf{w}^*\mathbf{u} \neq -1$, prove that $(\mathbf{I} + \mathbf{uw}^*)^{-1} - \mathbf{I}$ has rank one and find an explicit formula for the inverse.
3. Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be invertible. Give a necessary and sufficient condition for $\mathbf{A} + \mathbf{uw}^*$ to be invertible. When it is invertible, derive a formula expressing its inverse as a rank-one perturbation of $\mathbf{A}^{-1}$.

1.2 Norms and the measurement of size

Definition 1.2. Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A map $\|\cdot\| : V \to \mathbb{R}$ is a *norm* when, for all $\mathbf{x}, \mathbf{y} \in V$ and all scalars c,

1. $\|\mathbf{x}\| \geq 0$;
2. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$;
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;
4. $\|c\mathbf{x}\| = |c|\ \|\mathbf{x}\|$.

Problem 1.4 (When $\|\mathbf{Ax}\|$ is a norm). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and let $\|\cdot\|$ be any norm on $\mathbb{C}^n$. Determine necessary and sufficient conditions on $\mathbf{A}$ (possibly including conditions on m and/or n) under

Problem 1.4 (continued).
which

$$\mathbf{x} \longmapsto \|\mathbf{A}\mathbf{x}\|$$

is a norm on $\mathbb{C}^n$, and prove your claim.

The standard vector norms are the ℓ^p norms:

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}, \qquad 1 \le p < \infty, \tag{1.4}$$

$$\|\mathbf{x}\|_\infty = \max_{1\le j\le n} |x_j|. \tag{1.5}$$

The norm $\|\cdot\|_2$ is special because it comes from the inner product in (1.2).

Problem 1.5 (Sharp ℓ^∞–ℓ^2 equivalence). Find the sharp constants c_n and C_n such that

$$c_n \|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2 \le C_n \|\mathbf{x}\|_\infty \qquad (\mathbf{x} \in \mathbb{C}^n).$$

Identify nonzero vectors attaining equality in each bound.

Theorem 1.3 (Cauchy–Schwarz). *For* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 .$$

Proof. If $\mathbf{y} = \mathbf{0}$ there is nothing to prove. Otherwise subtract the projection of $\mathbf{x}$ onto $\mathbf{y}$ and use nonnegativity:

$$0 \le \left\| \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{y}\|_2^2}\mathbf{y} \right\|_2^2 = \|\mathbf{x}\|_2^2 - \frac{|\langle \mathbf{x}, \mathbf{y}\rangle|^2}{\|\mathbf{y}\|_2^2}.$$

Rearranging gives the claim. □

Cauchy–Schwarz supplies the triangle inequality for $\|\cdot\|_2$: expanding $\|\mathbf{x}+\mathbf{y}\|_2^2$ and bounding the two cross terms gives $(\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)^2$. The remaining norm axioms are immediate, so this also completes the standard proof that $\|\cdot\|_2$ is a norm.

1.2.1 Pythagoras and orthogonal decompositions

If $\mathbf{x}_1 \perp \mathbf{x}_2$, then

$$\begin{aligned} \|\mathbf{x}_1 + \mathbf{x}_2\|_2^2 &= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2\rangle \\ &= \|\mathbf{x}_1\|_2^2 + \|\mathbf{x}_2\|_2^2 , \end{aligned}$$

because the cross terms vanish. More generally, for mutually orthogonal $\mathbf{x}_1, \dots, \mathbf{x}_k$,

$$\left\| \sum_{j=1}^{k} \mathbf{x}_j \right\|_2^2 = \sum_{j=1}^{k} \left\|\mathbf{x}_j\right\|_2^2 . \tag{1.6}$$

This identity explains why orthogonal coordinates are so valuable computationally: the size of a sum is recovered without cancellation-prone cross terms.

1.2.2 Matrix norms

One can regard a matrix simply as a vector with two indices. Vectorizing its entries therefore gives the entrywise norms

$$\|\mathbf{A}\|_{p,p} \coloneqq \|\mathrm{vec}(\mathbf{A})\|_p .$$

A mixed entrywise norm may first measure each column in ℓ^p and then measure the resulting vector in ℓ^q:

$$\|\mathbf{A}\|_{p,q} = \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}|^p \right)^{q/p} \right)^{1/q}, \tag{1.7}$$

with the usual maximum replacing a sum when an exponent is infinity. The particularly important case $p = q = 2$ is the Frobenius norm,

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 = \mathrm{tr}(\mathbf{A}^*\mathbf{A}).$$

It comes from the Frobenius inner product

$$\langle \mathbf{A}, \mathbf{B} \rangle_F = \mathrm{tr}(\mathbf{B}^*\mathbf{A}).$$

The operator viewpoint leads to a different and often more useful construction. For $1 \le p \le \infty$, define the induced matrix norm by

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}. \tag{1.8}$$

Equivalently, this is the smallest constant C for which $\|\mathbf{A}\mathbf{x}\|_p \le C \|\mathbf{x}\|_p$ for every $\mathbf{x}$.

Problem 1.6 (Induced 1- and ∞-norm formulas). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ have columns $\mathbf{c}_1, \dots, \mathbf{c}_n$ and rows $\mathbf{r}_1^*, \dots, \mathbf{r}_m^*$. Prove that $\mathbf{A} : \mathbb{C}^n \to \mathbb{C}^m$ satisfies,

$$\|\mathbf{A}\|_1 = \max_{1\le j\le n} \left\|\mathbf{c}_j\right\|_1, \qquad \|\mathbf{A}\|_\infty = \max_{1\le i\le m} \|\mathbf{r}_i\|_1 .$$

Problem 1.7 (Matrix-norm submultiplicativity). A matrix norm $\|\cdot\|$ is *submultiplicative* if $\|\mathbf{A}\mathbf{B}\| \le \|\mathbf{A}\| \, \|\mathbf{B}\|$ for every conforming pair, including rectangular matrices. Prove that

1. every induced p-norm, $1 \le p \le \infty$, is submultiplicative;
2. the Frobenius norm is submultiplicative.

For induced norms, submultiplicativity follows immediately from the defining operator inequality:

$$\|\mathbf{A}\mathbf{B}\|_p \le \|\mathbf{A}\|_p \, \|\mathbf{B}\|_p .$$

Example 1.4 (Computing an induced 2-norm)**.** Let

$$\mathbf{A} = \begin{pmatrix} -2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Then

$$\mathbf{A}^*\mathbf{A} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

For $\mathbf{x} = (x_1, x_2)^T$,

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_2^2 &= 5\big(|x_1|^2 + |x_2|^2\big) + 8\,\mathrm{Re}(\overline{x_1}x_2) \\ &\le 9\big(|x_1|^2 + |x_2|^2\big) = 9\,\|\mathbf{x}\|_2^2 . \end{aligned}$$

Equality is obtained, for example, when $x_1 = x_2 \neq 0$. Hence $\|\mathbf{A}\|_2 = 3$. Chapter 2 will explain this calculation through singular values.

1.2.3 Equivalence of norms

Theorem 1.5 (Equivalence of finite-dimensional norms)**.** *If V is finite-dimensional and $\|\cdot\|_*$, $\|\cdot\|_+$ are two norms on V, then there are constants $c, k > 0$, independent of $\mathbf{x}$, such that*

$$c\,\|\mathbf{x}\|_* \le \|\mathbf{x}\|_+ \le k\,\|\mathbf{x}\|_* \qquad \textit{for every } \mathbf{x} \in V.$$

The finite-dimensional hypothesis is essential. The theorem says that no finite-dimensional norm has a fundamentally different notion of smallness, although the constants can depend strongly on the dimension. That dependence is why the choice of norm still matters in estimates.

For example,

$$\frac{1}{\sqrt{n}}\,\|\mathbf{x}\|_1 \le \|\mathbf{x}\|_2 \le \|\mathbf{x}\|_1 . \tag{1.9}$$

The right inequality follows by expanding $\|\mathbf{x}\|_1^2$; the left is Cauchy–Schwarz applied to the vectors $(|x_1|, \dots, |x_n|)^T$ and $(1, \dots, 1)^T$. Both constants are sharp: a one-sparse vector gives equality on the right, while a vector whose entries have equal modulus gives equality on the left.

Guided exercise 1.6. For $\mathbf{A} \in \mathbb{C}^{m\times n}$, prove the sharp induced-norm comparison

$$\frac{1}{\sqrt{m}}\,\|\mathbf{A}\|_1 \le \|\mathbf{A}\|_2 \le \sqrt{n}\,\|\mathbf{A}\|_1 .$$

For sharpness, test a matrix with one nonzero column consisting entirely of ones, and a matrix with one nonzero row consisting entirely of ones.

1.3 Unitary and orthogonal matrices

Definition 1.7. A matrix $\mathbf{U} \in \mathbb{C}^{n\times n}$ is *unitary* when $\mathbf{U}^*\mathbf{U} = \mathbf{I}$. A real matrix satisfying $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ is called *orthogonal*.

The columns of a unitary matrix form an orthonormal basis of $\mathbb{C}^n$. Because a square matrix with a left inverse is invertible,

$$\mathbf{U}^{-1} = \mathbf{U}^*, \qquad \mathbf{U}\mathbf{U}^* = \mathbf{I}. \tag{1.10}$$

Problem 1.8 (Triangular unitary matrices are diagonal). Prove that a square matrix that is both triangular and unitary must be diagonal. Show also that every diagonal entry has modulus

Problem 1.8 (continued).
one.

Problem 1.9 (Unitary invariance of vector and matrix norms).

1. If $\mathbf{U} \in \mathbb{C}^{n\times n}$ is unitary, prove that $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ for every $\mathbf{x}$.
2. Prove that the induced 2-norm and the Frobenius norm satisfy
$$\|\mathbf{UAV}\| = \|\mathbf{A}\|$$
for all matrices $\mathbf{A}$ and all conforming unitary $\mathbf{U}, \mathbf{V}$.
3. If $\mathbf{W} \in \mathbb{C}^{m\times n}$ satisfies $\mathbf{W}^*\mathbf{W} = \mathbf{I}_n$, prove that $\|\mathbf{Wx}\|_2 = \|\mathbf{x}\|_2$ even when $m \neq n$.

In particular,
$$\|\mathbf{Ux}\|_2^2 = \mathbf{x}^*\mathbf{U}^*\mathbf{Ux} = \|\mathbf{x}\|_2^2.$$
Thus unitary matrices are isometries. Real orthogonal matrices describe rigid rotations and reflections fixing the origin; unitary matrices are the complex analogue, also allowing changes of complex phase.

1.4 Projection matrices

Suppose that $V, W \subseteq \mathbb{C}^N$ are complementary subspaces:
$$\mathbb{C}^N = V \oplus W.$$
Every $\mathbf{x}$ has a unique decomposition $\mathbf{x} = \mathbf{v} + \mathbf{w}$ with $\mathbf{v} \in V$ and $\mathbf{w} \in W$. The map
$$\mathbf{P}(\mathbf{v} + \mathbf{w}) = \mathbf{v}$$
is the projection *onto* V *along* W. Its range is V, its kernel is W, and the complement condition is what makes it defined uniquely on the whole ambient space.

Theorem 1.8 (Existence and uniqueness of a projection)**.** *If* $V \cap W = \{\mathbf{0}\}$ *and* $\dim V + \dim W = N$, *then there is a unique* $\mathbf{P} \in \mathbb{C}^{N\times N}$ *with* $\operatorname{range}(\mathbf{P}) = V$ *and* $\mathcal{K}(\mathbf{P}) = W$.

Problem 1.10 (A projection is determined by range and kernel). Let $\mathbf{P}$ be a projection with range V and kernel W. Prove directly that any other projection with the same range and kernel equals $\mathbf{P}$.

To make the construction explicit, let the columns of $\mathbf{V} \in \mathbb{C}^{N\times r}$ be a basis of V and the columns of $\mathbf{W} \in \mathbb{C}^{N\times(N-r)}$ a basis of W. Then $[\mathbf{V}\ \mathbf{W}]$ is invertible, and
$$\mathbf{P} = \begin{bmatrix}\mathbf{V} & \mathbf{0}_{N\times(N-r)}\end{bmatrix}\begin{bmatrix}\mathbf{V} & \mathbf{W}\end{bmatrix}^{-1}. \tag{1.11}$$
The formula first finds the coordinates of $\mathbf{x}$ in the combined basis, then discards its W-coordinates.

Theorem 1.9 (Algebraic characterization)**.** *A square matrix* $\mathbf{P}$ *is a projection if and only if*
$$\mathbf{P}^2 = \mathbf{P}.$$

Proof. A projection acts as the identity on its range, so applying it twice has the same effect as applying it once. Conversely, if $\mathbf{P}^2 = \mathbf{P}$, then

$$\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I} - \mathbf{P})\mathbf{x}.$$

The first term lies in range($\mathbf{P}$), the second lies in $\mathcal{K}(\mathbf{P})$, and their intersection is $\{\mathbf{0}\}$. Thus these spaces are complementary and $\mathbf{P}$ is their associated projection. □

A general projection can increase Euclidean length substantially when its range and kernel are nearly aligned. The distinguished case occurs when they are orthogonal.

Problem 1.11 (Constructing an oblique projector of prescribed norm). Let $M \in [1, \infty)$ and $n \geq 2$. Construct explicitly a projection $\mathbf{P} \in \mathbb{C}^{n\times n}$ such that $\|\mathbf{P}\|_2 = M$, and verify both its idempotence and its norm.

Definition 1.10. An *orthogonal projector* is a projection satisfying

$$\mathcal{K}(\mathbf{P}) = \text{range}(\mathbf{P})^{\perp}.$$

Problem 1.12 (Hermitian idempotents are orthogonal projectors). Starting from the geometric definition above, prove that a projection $\mathbf{P}$ is an orthogonal projector if and only if $\mathbf{P}$ is Hermitian.

Theorem 1.11. *A matrix* $\mathbf{P}$ *is an orthogonal projector if and only if*

$$\mathbf{P}^2 = \mathbf{P} \qquad \textit{and} \qquad \mathbf{P}^* = \mathbf{P}.$$

For every $\mathbf{x}$ *it also satisfies* $\|\mathbf{P}\mathbf{x}\|_2 \leq \|\mathbf{x}\|_2$.

Proof. For an orthogonal projector, $\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I} - \mathbf{P})\mathbf{x}$ is an orthogonal sum. Pythagoras gives

$$\|\mathbf{x}\|_2^2 = \|\mathbf{P}\mathbf{x}\|_2^2 + \|(\mathbf{I} - \mathbf{P})\mathbf{x}\|_2^2,$$

which proves nonexpansiveness. In orthonormal bases adapted to the range and kernel, $\mathbf{P}$ has diagonal form diag$(1, \dots, 1, 0, \dots, 0)$, hence is Hermitian. Conversely, if $\mathbf{P}$ is Hermitian and idempotent, then for $\mathbf{v} = \mathbf{P}\mathbf{x}$ and $\mathbf{w} \in \mathcal{K}(\mathbf{P})$,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{P}\mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{x}, \mathbf{P}^*\mathbf{w} \rangle = 0.$$

Thus its range is orthogonal to its kernel. □

Problem 1.13 (Orthogonal projection minimizes distance). Let $\mathbf{P}$ be the orthogonal projector onto a subspace $V \subseteq \mathbb{C}^n$. For every $\mathbf{x} \in \mathbb{C}^n$, prove that

$$\|\mathbf{x} - \mathbf{P}\mathbf{x}\|_2 = \min_{\mathbf{v}\in V} \|\mathbf{x} - \mathbf{v}\|_2.$$

Identify which $\mathbf{v} \in V$ achieves equality.

Guided exercise 1.12. Let $\mathbf{q}_1, \dots, \mathbf{q}_r$ be orthonormal and set $\mathbf{Q} = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_r]$. Show that $\mathbf{P} = \mathbf{Q}\mathbf{Q}^*$ is the orthogonal projector onto range($\mathbf{Q}$). Identify $\mathbf{I} - \mathbf{P}$ geometrically.

1.5 Permutation matrices

A permutation is a bijection $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$. When viewed as a map on the indices of a vector, this induces a linear operator that permutes coordinates of a vector, represented by the matrix $\mathbf{P}$ determined by

$$\mathbf{P}\mathbf{e}_j = \mathbf{e}_{\pi(j)}.$$

Its columns are a rearrangement of the standard orthonormal basis, so every permutation matrix is unitary (and, being real, orthogonal):

$$\mathbf{P}^* = \mathbf{P}^{-1}.$$

Products of permutation matrices are again permutation matrices, reflecting composition of the underlying bijections. These matrices will later encode row and column exchanges.

Guided exercise 1.13. Write the 3×3 matrix associated with the cycle $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$. Compute its inverse in two ways: from the inverse permutation and from its transpose.

Problems for Chapter 1

The following statements repeat the problems from their inline locations. Each return link leads back to the first occurrence.

Problem 1.1 (Matrix multiplication as inner and outer products). Let $\mathbf{A} \in \mathbb{C}^{K\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times L}$. Using only the entrywise definition (1.1), prove the following statements.

1. If the rows of $\mathbf{A}$ are $\mathbf{r}_k^*$ and the columns of $\mathbf{B}$ are $\mathbf{c}_\ell$, then

$$(\mathbf{AB})_{k\ell} = \mathbf{r}_k^*\mathbf{c}_\ell.$$

2. If $\mathbf{C} \in \mathbb{C}^{K\times n}$ has columns $\mathbf{c}_j$ and $\mathbf{R} \in \mathbb{C}^{n\times L}$ has rows $\mathbf{r}_j^*$, then

$$\mathbf{CR} = \sum_{j=1}^{n} \mathbf{c}_j\mathbf{r}_j^*.$$

3. Determine and justify the maximum possible rank of $\mathbf{CR}$ in terms of K, n, L.

Return to the inline occurrence of Problem 1.1.

Problem 1.2 (Fundamental-subspace inclusions for a product). Let $\mathbf{A}$ and $\mathbf{B}$ have conforming sizes. Prove that

1. $\mathrm{range}(\mathbf{AB}) \subseteq \mathrm{range}(\mathbf{A})$;
2. $\mathrm{corange}(\mathbf{AB}) \subseteq \mathrm{corange}(\mathbf{B})$;
3. $\mathcal{K}(\mathbf{AB}) \supseteq \mathcal{K}(\mathbf{B})$;
4. $\mathrm{coker}(\mathbf{AB}) \supseteq \mathrm{coker}(\mathbf{A})$.

Return to the inline occurrence of Problem 1.2.

Problem 1.3 (Sherman–Morrison rank-one update). Let $\mathbf{u}, \mathbf{w} \in \mathbb{C}^n$ be nonzero.

1. Prove that $\mathbf{I} + \mathbf{u}\mathbf{w}^*$ is singular if and only if $\mathbf{w}^*\mathbf{u} = -1$.
2. If $\mathbf{w}^*\mathbf{u} \neq -1$, prove that $(\mathbf{I} + \mathbf{u}\mathbf{w}^*)^{-1} - \mathbf{I}$ has rank one and find an explicit formula for the inverse.
3. Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be invertible. Give a necessary and sufficient condition for $\mathbf{A} + \mathbf{u}\mathbf{w}^*$ to be invertible. When it is invertible, derive a formula expressing its inverse as a rank-one perturbation of $\mathbf{A}^{-1}$.

Return to the inline occurrence of Problem 1.3.

Problem 1.4 (When $\|\mathbf{Ax}\|$ is a norm). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and let $\|\cdot\|$ be any norm on $\mathbb{C}^n$. Determine necessary and sufficient conditions on $\mathbf{A}$ (possibly including conditions on m and/or n) under which

$$\mathbf{x} \longmapsto \|\mathbf{Ax}\|$$

is a norm on $\mathbb{C}^n$, and prove your claim.

Return to the inline occurrence of Problem 1.4.

Problem 1.5 (Sharp ℓ^∞–ℓ^2 equivalence). Find the sharp constants c_n and C_n such that

$$c_n \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq C_n \|\mathbf{x}\|_\infty \qquad (\mathbf{x} \in \mathbb{C}^n).$$

Identify nonzero vectors attaining equality in each bound.

Return to the inline occurrence of Problem 1.5.

Problem 1.6 (Induced 1- and ∞-norm formulas). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ have columns $\mathbf{c}_1, \dots, \mathbf{c}_n$ and rows $\mathbf{r}_1^*, \dots, \mathbf{r}_m^*$. Prove that $\mathbf{A} : \mathbb{C}^n \to \mathbb{C}^m$ satisfies,

$$\|\mathbf{A}\|_1 = \max_{1\leq j\leq n} \left\|\mathbf{c}_j\right\|_1, \qquad \|\mathbf{A}\|_\infty = \max_{1\leq i\leq m} \|\mathbf{r}_i\|_1.$$

Return to the inline occurrence of Problem 1.6.

Problem 1.7 (Matrix-norm submultiplicativity). A matrix norm $\|\cdot\|$ is *submultiplicative* if $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$ for every conforming pair, including rectangular matrices. Prove that

1. every induced p-norm, $1 \leq p \leq \infty$, is submultiplicative;
2. the Frobenius norm is submultiplicative.

Return to the inline occurrence of Problem 1.7.

Problem 1.8 (Triangular unitary matrices are diagonal). Prove that a square matrix that is both triangular and unitary must be diagonal. Show also that every diagonal entry has modulus one.

Return to the inline occurrence of Problem 1.8.

Problem 1.9 (Unitary invariance of vector and matrix norms).

1. If $\mathbf{U} \in \mathbb{C}^{n\times n}$ is unitary, prove that $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ for every $\mathbf{x}$.

Problem 1.9 (continued).

2. Prove that the induced 2-norm and the Frobenius norm satisfy

$$\|\mathbf{UAV}\| = \|\mathbf{A}\|$$

for all matrices $\mathbf{A}$ and all conforming unitary $\mathbf{U}, \mathbf{V}$.

3. If $\mathbf{W} \in \mathbb{C}^{m\times n}$ satisfies $\mathbf{W}^*\mathbf{W} = \mathbf{I}_n$, prove that $\|\mathbf{Wx}\|_2 = \|\mathbf{x}\|_2$ even when $m \neq n$.

Return to the inline occurrence of Problem 1.9.

Problem 1.10 (A projection is determined by range and kernel). Let $\mathbf{P}$ be a projection with range V and kernel W. Prove directly that any other projection with the same range and kernel equals $\mathbf{P}$.

Return to the inline occurrence of Problem 1.10.

Problem 1.11 (Constructing an oblique projector of prescribed norm). Let $M \in [1, \infty)$ and $n \geq 2$. Construct explicitly a projection $\mathbf{P} \in \mathbb{C}^{n\times n}$ such that $\|\mathbf{P}\|_2 = M$, and verify both its idempotence and its norm.

Return to the inline occurrence of Problem 1.11.

Problem 1.12 (Hermitian idempotents are orthogonal projectors). Starting from the geometric definition above, prove that a projection $\mathbf{P}$ is an orthogonal projector if and only if $\mathbf{P}$ is Hermitian.

Return to the inline occurrence of Problem 1.12.

Problem 1.13 (Orthogonal projection minimizes distance). Let $\mathbf{P}$ be the orthogonal projector onto a subspace $V \subseteq \mathbb{C}^n$. For every $\mathbf{x} \in \mathbb{C}^n$, prove that

$$\|\mathbf{x} - \mathbf{Px}\|_2 = \min_{\mathbf{v}\in V} \|\mathbf{x} - \mathbf{v}\|_2 .$$

Identify which $\mathbf{v} \in V$ achieves equality.

Return to the inline occurrence of Problem 1.13.

Chapter 2

Spectral Structure and the Singular Value Decomposition

Before addressing algorithms for computing eigenvalues, we discuss what eigenvalues reveal, what they fail to reveal, and which matrix structures make them especially useful. This chapter develops similarity, Hermitian and normal spectral theory, variational characterizations, and the singular value decomposition (SVD).

2.1 Eigenvalues, eigenvectors, and eigenspaces

Definition 2.1. Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. A scalar-vector pair $(\lambda, \mathbf{v}) \in \mathbb{C} \times (\mathbb{C}^n \setminus \{\mathbf{0}\})$ is an *eigenpair* when

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \tag{2.1}$$

The set of eigenvalues is the *spectrum*, denoted $\operatorname{spec}(\mathbf{A})$. For an eigenvalue λ, its eigenspace is

$$E_\lambda = \mathcal{K}(\mathbf{A} - \lambda\mathbf{I}).$$

The zero vector belongs to every eigenspace but is not itself an eigenvector. Even when $\mathbf{A}$ is real, its eigenvalues and eigenvectors may be complex. Over $\mathbb{C}$, the characteristic polynomial

$$p_{\mathbf{A}}(z) = \det(z\mathbf{I} - \mathbf{A})$$

has degree n, so $\mathbf{A}$ has n eigenvalues when algebraic multiplicity is counted. Finding the roots of this polynomial is valuable on paper but is a poor blueprint for a numerical eigensolver.

The *algebraic multiplicity* a_λ is the multiplicity of λ as a root of $p_{\mathbf{A}}$. The *geometric multiplicity* is

$$g_\lambda = \dim E_\lambda.$$

We always have $1 \le g_\lambda \le a_\lambda$. An eigenvalue is *simple* when $a_\lambda = g_\lambda = 1$ and *defective* when $g_\lambda < a_\lambda$. A matrix with a defective eigenvalue is called defective.

Proposition 2.2. *Eigenvectors belonging to distinct eigenvalues are linearly independent. Every eigenspace is invariant under* $\mathbf{A}$*.*

Proof. For invariance, if $\mathbf{x} \in E_\lambda$, then $\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \in E_\lambda$. For independence, suppose $\mathbf{v}_1, \dots, \mathbf{v}_k$ correspond to distinct eigenvalues and take a nontrivial relation $\sum_{j=1}^{k} c_j \mathbf{v}_j = \mathbf{0}$. If the eigenvectors were linearly depedent, then not all the c_j can be zero; without loss assume $c_1 \neq 0$. Applying $\mathbf{A} - \lambda_k \mathbf{I}$ removes the last term and gives a shorter relation

$$\sum_{j=1}^{k-1} c_j(\lambda_j - \lambda_k)\mathbf{v}_j = \mathbf{0}.$$

Repeatedly applying this argument implies that $c_1 = 0$, achieving a contradiction. □

Problem 2.1 (Spectrum and diagonalization of a projection). Let $\mathbf{P} \in \mathbb{C}^{n\times n}$ be a projection of rank r, where $0 \leq r \leq n$. Determine its eigenvalues and their algebraic multiplicities. Prove that every projection is diagonalizable and construct an eigenbasis from vectors in its range and kernel.

Problem 2.2 (Gershgorin discs). Let $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{n\times n}$ and define

$$D_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq \sum_{j\neq i} |a_{ij}| \right\}, \qquad 1 \leq i \leq n.$$

Prove that

$$\operatorname{spec}(\mathbf{A}) \subseteq \bigcup_{i=1}^{n} D_i.$$

2.2 Similarity and diagonalization

Definition 2.3. Matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ are *similar* if there is an invertible $\mathbf{S}$ such that

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$$

Similar matrices represent the same linear operator in different bases. If $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent eigenvectors, set

$$\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n], \qquad \mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Writing all eigenvector equations at once gives

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}. \tag{2.2}$$

Since $\mathbf{V}$ is invertible,

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \qquad \mathbf{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}. \tag{2.3}$$

Theorem 2.4. *A square matrix is diagonalizable if and only if it is not defective, or equivalently, if and only if it has a basis of eigenvectors.*

Problem 2.3 (Determinant and trace from eigenvalues). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \dots, \lambda_n$, counted with algebraic multiplicity. Prove that

$$\det(\mathbf{A}) = \prod_{j=1}^{n} \lambda_j, \qquad \operatorname{tr}(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j.$$

Your argument must also cover defective matrices. Deduce that, when $\mathbf{A}$ is invertible, $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.

Similarity preserves the characteristic polynomial because

$$\begin{aligned}\det(z\mathbf{I} - \mathbf{S}^{-1}\mathbf{A}\mathbf{S}) &= \det\bigl(\mathbf{S}^{-1}(z\mathbf{I} - \mathbf{A})\mathbf{S}\bigr) \\ &= \det(z\mathbf{I} - \mathbf{A}).\end{aligned}$$

Thus it preserves eigenvalues, including algebraic multiplicity, whether or not the matrix is diagonalizable. In particular,

$$\det(\mathbf{A}) = \prod_{j=1}^{n} \lambda_j, \qquad \operatorname{tr}(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j. \tag{2.4}$$

Defective matrices do not admit (2.3), but they possess a more delicate canonical form. For each distinct eigenvalue λ_j, write $a_j = a_{\lambda_j}$ and $g_j = g_{\lambda_j}$. There exist positive Jordan-block sizes $s_{j,1}, \dots, s_{j,g_j}$, determined by $\mathbf{A}$ up to their order, satisfying

$$\sum_{\ell=1}^{g_j} s_{j,\ell} = a_j.$$

Then $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$ for an invertible $\mathbf{V}$, where

$$\mathbf{J} = \bigoplus_{j} \bigoplus_{\ell=1}^{g_j} \left(\lambda_j \mathbf{I}_{s_{j,\ell}} + \mathbf{N}_{s_{j,\ell}}\right), \tag{2.5}$$

and $\mathbf{N}_s$ has ones on its first superdiagonal and zeros elsewhere. The familiar matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is the simplest defective Jordan block. Jordan form is structurally revealing, but a *unitary* change of coordinates is preferable for numerical geometry.

For vectors $\mathbf{x} \in \mathbb{C}^m$ and $\mathbf{y} \in \mathbb{C}^n$, their tensor product $\mathbf{x} \otimes \mathbf{y} \in \mathbb{C}^{mn}$ is obtained by stacking the vectors $x_j\mathbf{y}$. For matrices $\mathbf{A} = (a_{jk}) \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$, the *Kronecker product* $\mathbf{A} \otimes \mathbf{B}$ is the $m \times m$ block matrix whose (j,k) block is $a_{jk}\mathbf{B}$.

Problem 2.4 (Kronecker-product eigenpairs). Suppose $\mathbf{A} \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$ are diagonalizable. Express all eigenvalues and an eigenbasis of $\mathbf{A} \otimes \mathbf{B}$ in terms of eigenpairs of $\mathbf{A}$ and $\mathbf{B}$.

2.3 Hermitian matrices and the spectral theorem

Definition 2.5. A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is *Hermitian* when $\mathbf{A} = \mathbf{A}^*$. A real Hermitian matrix is symmetric.

Theorem 2.6 (Spectral theorem for Hermitian matrices)**.** *A matrix* $\mathbf{A}$ *is Hermitian if and only if its eigenvalues are real and it has an orthonormal basis of eigenvectors. Equivalently, there are a unitary* $\mathbf{U}$ *and a real diagonal* $\mathbf{\Lambda}$ *such that*

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*. \tag{2.6}$$

A matrix $\mathbf{A}$ is *skew-Hermitian* when $\mathbf{A}^* = -\mathbf{A}$.

Problem 2.5 (Spectral theorem for skew-Hermitian matrices). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Prove that $\mathbf{A}$ is skew-Hermitian if and only if it is unitarily diagonalizable and every eigenvalue is purely imaginary (with zero allowed).

If $\mathbf{u}_j$ is the jth column of $\mathbf{U}$, then

$$\mathbf{A} = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \mathbf{u}_j^*. \tag{2.7}$$

This expansion gives a geometric interpretation: resolve a vector in an orthonormal frame, scale its jth coordinate by λ_j, and rotate back.

The *spectral radius* is

$$\rho(\mathbf{A}) = \max_{1\le j\le n} |\lambda_j|.$$

For Hermitian $\mathbf{A}$, orthonormal coordinates and Pythagoras show that

$$\|\mathbf{A}\mathbf{x}\|_2^2 = \sum_{j=1}^{n} |\lambda_j|^2 \left|\langle \mathbf{x}, \mathbf{u}_j\rangle\right|^2 \le \rho(\mathbf{A})^2 \|\mathbf{x}\|_2^2.$$

Equality at an eigenvector belonging to an eigenvalue of largest modulus gives

$$\|\mathbf{A}\|_2 = \rho(\mathbf{A}). \tag{2.8}$$

2.3.1 Positive definiteness, energy norms, and square roots

A Hermitian matrix is *positive definite* when every eigenvalue is strictly positive, and *positive semidefinite* when every eigenvalue is nonnegative. For a positive-definite matrix, the formula

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\mathbf{x}^*\mathbf{A}\mathbf{x}} \tag{2.9}$$

defines a norm. Indeed, if $\mathbf{x} = \sum_j c_j \mathbf{u}_j$, then

$$\mathbf{x}^*\mathbf{A}\mathbf{x} = \sum_{j=1}^{n} \lambda_j |c_j|^2 > 0$$

for nonzero $\mathbf{x}$, and (2.9) is the Euclidean norm after the coordinate scaling $c_j \mapsto \sqrt{\lambda_j} c_j$.

The same scaling constructs the unique positive-definite square root:

$$\mathbf{A}^{1/2} = \mathbf{U}\,\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})\mathbf{U}^*. \tag{2.10}$$

It satisfies $(\mathbf{A}^{1/2})^2 = \mathbf{A}$. Positivity fixes the choice of scalar square root and hence gives uniqueness among positive-definite square roots.

2.3.2 Quadratic forms and spectral subspaces

For Hermitian $\mathbf{A}$, the real-valued function

$$Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^*\mathbf{A}\mathbf{x}$$

is its Hermitian quadratic form. Partition an orthonormal eigenbasis according to positive, negative, and zero eigenvalues, and let V^+, V^-, and V^0 be the spans of the respective groups. Then

$$\mathbb{C}^n = V^+ \oplus V^- \oplus V^0,$$

and

$$Q_{\mathbf{A}}(\mathbf{x}) \begin{cases} > 0, & \mathbf{x} \in V^+ \setminus \{\mathbf{0}\}, \\ < 0, & \mathbf{x} \in V^- \setminus \{\mathbf{0}\}, \\ = 0, & \mathbf{x} \in V^0. \end{cases}$$

More generally, if $\mathbf{x} = \sum_j c_j \mathbf{u}_j$, then

$$Q_{\mathbf{A}}(\mathbf{x}) = \sum_j \lambda_j \left|c_j\right|^2.$$

Thus the spectrum records the sign geometry of a Hermitian quadratic form. Notice, however, that $V^0 = \mathcal{K}(\mathbf{A})$ need not be the complete zero set of $Q_{\mathbf{A}}$: when $\mathbf{A}$ is indefinite, positive and negative spectral components can cancel and produce nonzero isotropic vectors.

2.4 Rayleigh quotients and variational eigenvalue formulas

For any $\mathbf{A} \in \mathbb{C}^{n\times n}$ and nonzero $\mathbf{x}$, define the *Rayleigh quotient*

$$R_{\mathbf{A}}(\mathbf{x}) = \frac{\mathbf{x}^*\mathbf{A}\mathbf{x}}{\mathbf{x}^*\mathbf{x}}. \tag{2.11}$$

It is unchanged by nonzero scaling of $\mathbf{x}$, and an eigenpair satisfies $R_{\mathbf{A}}(\mathbf{v}) = \lambda$. The *numerical range* is

$$W(\mathbf{A}) = \{R_{\mathbf{A}}(\mathbf{x}) \mid \mathbf{x} \neq \mathbf{0}\}.$$

Equivalently, it is enough to use unit vectors. Hence $\operatorname{spec}(\mathbf{A}) \subseteq W(\mathbf{A})$.

Theorem 2.7 (Toeplitz–Hausdorff). *The numerical range $W(\mathbf{A})$ is a compact, convex subset of $\mathbb{C}$.*

Compactness follows because the unit sphere is compact and the Rayleigh quotient is continuous there. Convexity is more subtle. For Hermitian matrices, however, the whole set is immediately visible from the spectral expansion. Writing a unit vector as $\mathbf{x} = \sum_j c_j \mathbf{u}_j$ gives

$$R_{\mathbf{A}}(\mathbf{x}) = \sum_{j=1}^{n} \lambda_j \left|c_j\right|^2, \qquad \sum_{j=1}^{n} \left|c_j\right|^2 = 1. \tag{2.12}$$

Thus the quotient is a convex combination of the eigenvalues. If they are ordered $\lambda_1 \leq \cdots \leq \lambda_n$, then

$$\lambda_1 \leq R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_n, \qquad W(\mathbf{A}) = [\lambda_1, \lambda_n]. \tag{2.13}$$

For a general square matrix, define its numerical radius and Hermitian part by

$$w(\mathbf{A}) = \sup_{z\in W(\mathbf{A})} |z|, \qquad H(\mathbf{A}) = \frac{\mathbf{A} + \mathbf{A}^*}{2}.$$

Problem 2.6 (Numerical-range identities). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$.

1. Prove that $\rho(\mathbf{A}) \le w(\mathbf{A}) \le \|\mathbf{A}\|_2$.
2. For $\theta \in \mathbb{R}$, prove that $W(e^{i\theta}\mathbf{A}) = e^{i\theta}W(\mathbf{A})$.
3. Prove that $W(\mathbf{A}^*) = \{\overline{z} \mid z \in W(\mathbf{A})\}$.
4. Show that $W(H(\mathbf{A}))$ is the interval obtained by projecting $W(\mathbf{A})$ onto the real axis, and express its endpoints through the extreme eigenvalues of $H(\mathbf{A})$.

For a direction $e^{i\theta}$, set

$$H_\theta = H(e^{-i\theta}\mathbf{A}).$$

Then $\lambda_{\max}(H_\theta)$ is the support value $\max_{z\in W(\mathbf{A})} \operatorname{Re}(e^{-i\theta}z)$. If $\mathbf{x}_\theta$ is a corresponding unit eigenvector, the Rayleigh quotient $\mathbf{x}_\theta^*\mathbf{A}\mathbf{x}_\theta$ is a boundary point whose supporting line has outward normal $e^{i\theta}$.

Problem 2.7 (Computing a numerical range by support lines). Design and implement an algorithm that samples angles θ and uses the largest eigenpair of H_θ to approximate the boundary of $W(\mathbf{A})$. Give pseudocode, explain how angular resolution affects the output, and test at least three matrices, including nilpotent Jordan blocks of at least three sizes.
For the size-s nilpotent Jordan block $J_s(0)$, investigate and justify the disk suggested by the computation. More generally, if

$$\mathbf{J} = \bigoplus_{\ell} J_{s_\ell}(\lambda_\ell),$$

describe $W(\mathbf{J})$ Explain why this description of $W(\mathbf{J})$ does not in general determine $W(\mathbf{A})$ for a matrix merely similar to $\mathbf{J}$.

The interior eigenvalues also have variational descriptions.

Theorem 2.8 (Courant–Fischer min–max theorem). *Let* $\mathbf{A}$ *be Hermitian with* $\lambda_1 \le \cdots \le \lambda_n$. *For* $1 \le k \le n$,

$$\lambda_k = \min_{\substack{V\subseteq\mathbb{C}^n\\ \dim V = k}} \max_{\mathbf{x}\in V\setminus\{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x}), \tag{2.14}$$

$$\lambda_k = \max_{\substack{V\subseteq\mathbb{C}^n\\ \dim V = n-k+1}} \min_{\mathbf{x}\in V\setminus\{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x}). \tag{2.15}$$

The first outer extremum is attained by

$$V = \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\},$$

and the second is attained by

$$V = \operatorname{span}\{\mathbf{u}_k, \dots, \mathbf{u}_n\}.$$

For the first formula, every k-dimensional V contains a nonzero vector orthogonal to $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$. Its Rayleigh quotient is at least λ_k. On the proposed extremizing subspace, every quotient is at most λ_k. These two inequalities prove (2.14); the second formula follows by the symmetric argument.

If $\mathbf{Q} \in \mathbb{C}^{n\times r}$ has orthonormal columns, then $\mathbf{B} = \mathbf{Q}^*\mathbf{A}\mathbf{Q}$ is a *compression* of $\mathbf{A}$. Applying the min–max theorem to the subspace range($\mathbf{Q}$) yields the following consequence.

Problem 2.8 (Rayleigh–Ritz exactness). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ and let $\mathbf{Q} \in \mathbb{C}^{n\times p}$ have orthonormal columns. Suppose $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ and $\mathbf{v} \in \operatorname{range}(\mathbf{Q})$. Prove that $\mathbf{c} = \mathbf{Q}^*\mathbf{v}$ is an eigenvector of $\mathbf{Q}^*\mathbf{A}\mathbf{Q}$ with eigenvalue λ, and that $\mathbf{v} = \mathbf{Q}\mathbf{c}$. Explain why the converse requires the Ritz residual $\mathbf{A}\mathbf{Q}\mathbf{c} - \lambda\mathbf{Q}\mathbf{c}$ to vanish.

Theorem 2.9 (Cauchy interlacing)**.** *Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be Hermitian with eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. If an $(n-1)\times(n-1)$ compression $\mathbf{B}$ has eigenvalues $\mu_1 \le \cdots \le \mu_{n-1}$, then*

$$\lambda_j \le \mu_j \le \lambda_{j+1}, \qquad 1 \le j \le n-1.$$

Guided exercise 2.10. Use (2.12) to prove that $R_{\mathbf{A}}(\mathbf{x}) = \lambda_1$ exactly when $\mathbf{x}$ lies in the λ_1 eigenspace (allowing multiplicity). State and prove the analogous claim for λ_n.

2.5 Normal matrices and Schur form

Definition 2.11. A square matrix is *normal* when it commutes with its adjoint:

$$\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}.$$

Hermitian and unitary matrices are important examples. Normality is precisely the algebraic condition that guarantees a unitary eigenvector basis.

Every complex square matrix has an eigenvector. After normalizing one such vector, it can be completed to an orthonormal basis; induction on the remaining compression suggests the following triangularization.

Problem 2.9 (Proving Schur decomposition). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. Prove that there are a unitary $\mathbf{U}$ and an upper-triangular $\mathbf{T}$ such that

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*.$$

Theorem 2.12 (Schur decomposition)**.** *For every $\mathbf{A} \in \mathbb{C}^{n\times n}$, there are a unitary $\mathbf{U}$ and an upper triangular $\mathbf{T}$ such that*

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*. \tag{2.16}$$

Unitary similarity preserves normality. A direct entrywise argument shows that a triangular normal matrix must be diagonal.

Problem 2.10 (Spectral theorem for normal matrices). Prove that a complex square matrix is unitarily diagonalizable if and only if it is normal.

Theorem 2.13 (Spectral theorem for normal matrices)**.** *A complex square matrix is unitarily diagonalizable if and only if it is normal.*

Indeed, applying the triangular-normal observation to the Schur factor $\mathbf{T}$ proves one implication. Conversely, $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^*$ plainly commutes with its adjoint.

For a normal matrix, the induced 2-norm, spectral radius, and numerical radius

$$w(\mathbf{A}) = \max_{z\in W(\mathbf{A})} |z|$$

all coincide. Normality is also equivalent to

$$\|\mathbf{A}\mathbf{x}\|_2 = \|\mathbf{A}^*\mathbf{x}\|_2 \quad \text{for every } \mathbf{x}.$$

It is further equivalent to the existence of a unitary $\mathbf{Z}$ such that

$$\mathbf{A} = \mathbf{A}^*\mathbf{Z}. \tag{2.17}$$

Indeed, a unitary diagonalization permits the phases of the eigenvalues to be placed in $\mathbf{Z}$. Conversely, (2.17) gives

$$\mathbf{A}\mathbf{A}^* = (\mathbf{A}^*\mathbf{Z})(\mathbf{Z}^*\mathbf{A}) = \mathbf{A}^*\mathbf{A}.$$

These statements explain the special role of normal matrices: up to a rigid unitary change of coordinates, their action is completely described by their eigenvalues.

2.6 The singular value decomposition

Diagonalization is restricted to square matrices and can require a badly distorted eigenvector basis. The SVD instead uses independent unitary changes of coordinates in the input and output spaces, and therefore applies to every matrix.

Theorem 2.14 (Singular value decomposition)**.** *For every $\mathbf{A} \in \mathbb{C}^{m\times n}$, there are unitary matrices $\mathbf{U} \in \mathbb{C}^{m\times m}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ such that*

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*, \tag{2.18}$$

where $\boldsymbol{\Sigma} \in \mathbb{C}^{m\times n}$ is diagonal with entries

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_p \ge 0, \qquad p = \min\{m, n\}.$$

The σ_j are the singular values, the columns $\mathbf{u}_j$ of $\mathbf{U}$ are left singular vectors, and the columns $\mathbf{v}_j$ of $\mathbf{V}$ are right singular vectors.

Construction when $m \ge n$. The Gram matrix $\mathbf{A}^*\mathbf{A}$ is Hermitian positive semidefinite, because

$$\mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|_2^2 \ge 0.$$

Choose an orthonormal eigenbasis $\mathbf{v}_1, \dots, \mathbf{v}_n$ with eigenvalues $\sigma_1^2 \ge \cdots \ge \sigma_n^2 \ge 0$. If $r = \operatorname{rank}(\mathbf{A})$, define, only for $1 \le j \le r$,

$$\mathbf{u}_j = \frac{\mathbf{A}\mathbf{v}_j}{\sigma_j}.$$

For $i, j \le r$,

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \frac{\mathbf{v}_j^*\mathbf{A}^*\mathbf{A}\mathbf{v}_i}{\sigma_i\sigma_j} = \delta_{ij},$$

so these left singular vectors are orthonormal. Complete them to an orthonormal basis of $\mathbb{C}^m$. For $j > r$, $\mathbf{A}\mathbf{v}_j = \mathbf{0}$, and hence the column relations combine into $\mathbf{A}\mathbf{V} = \mathbf{U}\boldsymbol{\Sigma}$. Right multiplication by $\mathbf{V}^*$ gives (2.18). The case $m < n$ follows by applying the construction to $\mathbf{A}^*$. □

The singular values are unique. Singular vectors are not: paired complex phases may be changed, and within a repeated positive singular-value subspace the corresponding left and right bases may be changed by the same unitary transformation. The left and right nullspace bases may be chosen independently.

The outer-product form is

$$\mathbf{A} = \sum_{j=1}^{p} \sigma_j \mathbf{u}_j \mathbf{v}_j^*. \tag{2.19}$$

If $r = \operatorname{rank}(\mathbf{A})$, exactly r singular values are positive, and deleting the zero terms gives the reduced SVD

$$\mathbf{A} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^* = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*. \tag{2.20}$$

Problem 2.11 (Eigenvalues versus singular values). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$.

1. If $\mathbf{A}$ is Hermitian positive semidefinite, prove that its eigenvalues and singular values agree when both are ordered nonincreasingly.
2. If $\mathbf{A}$ is normal, prove that its singular values are the moduli of its eigenvalues, counted with multiplicity and arranged nonincreasingly.

For the second statement, write $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$. Then $\mathbf{A}^*\mathbf{A} = \mathbf{U}\,|\mathbf{\Lambda}|^2\,\mathbf{U}^*$. For a nonnormal matrix there is no such direct correspondence between eigenvalues and singular values.

Unitary invariance of the Euclidean norm immediately gives

$$\|\mathbf{A}\|_2 = \sigma_1. \tag{2.21}$$

Problem 2.12 (Frobenius norm from singular values). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and let $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_p)^T$, where $p = \min\{m, n\}$. Prove that

$$\|\mathbf{A}\|_F^2 = \sum_{j=1}^{p} \sigma_j^2 = \|\boldsymbol{\sigma}\|_2^2.$$

The *minimum gain* of a square matrix is

$$\min_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_{\min}(\mathbf{A}).$$

Problem 2.13 (An inf–sup characterization of $\sigma_{\min}$). For $\mathbf{A} \in \mathbb{C}^{n\times n}$ define

$$\beta = inf_{\mathbf{x}\neq\mathbf{0}}\, sup_{\mathbf{y}\neq\mathbf{0}} \frac{|\mathbf{x}^*\mathbf{A}\mathbf{y}|}{\|\mathbf{x}\|_2\,\|\mathbf{y}\|_2}.$$

Show that $\beta = \sigma_{\min}(\mathbf{A})$ and hence that $\beta > 0$ if and only if $\mathbf{A}$ is invertible. Treat separately the case in which the vector whose direction would maximize the inner product is zero.

Problem 2.14 (Singular-value interlacing under compression). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{Q} \in \mathbb{C}^{(m-1)\times m}$ satisfy $\mathbf{Q}\mathbf{Q}^* = \mathbf{I}_{m-1}$. Write the singular values of $\mathbf{A}$ as $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$, where $p = \min\{m, n\}$, and those of $\mathbf{QA}$ as $\mu_1 \geq \cdots \geq \mu_q \geq 0$, where $q = \min\{m-1, n\}$. Prove the interlacing bounds

$$\sigma_j \geq \mu_j \geq \sigma_{j+1}, \qquad 1 \leq j \leq q,$$

using the convention $\sigma_{p+1} = 0$.

Problem 2.15 (Full-rank matrices are dense). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$, let $\|\cdot\|$ be any matrix norm, and let $\epsilon > 0$. Prove that there is a matrix $\mathbf{B}$ of rank $\min\{m, n\}$ such that

$$\|\mathbf{B} - \mathbf{A}\| < \epsilon.$$

A factorization $\mathbf{A} = \mathbf{US}$ with $\mathbf{U}$ unitary and $\mathbf{S}$ Hermitian positive semidefinite is a *right polar decomposition*; a factorization $\mathbf{A} = \mathbf{RU}$ is a *left polar decomposition*.

Problem 2.16 (Unitary logarithms and polar decomposition).

1. Prove that every unitary $\mathbf{U} \in \mathbb{C}^{n\times n}$ can be written as $\mathbf{U} = e^{i\boldsymbol{\Theta}}$ for a Hermitian $\boldsymbol{\Theta}$. (Since $\boldsymbol{\Theta}$ is unitarily diagonalizable, we interpret $e^{i\boldsymbol{\Theta}}$ through the usual functional calculus where $\boldsymbol{\Theta} \mapsto e^{i\boldsymbol{\Theta}}$ is computed by shifting the complex exponential to operate on the individual eigenvalues.)
2. Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Prove that

$$\mathbf{A} = \mathbf{R}e^{i\boldsymbol{\Theta}} = e^{i\boldsymbol{\Theta}}\mathbf{S},$$

where $\mathbf{R}$ and $\mathbf{S}$ are Hermitian positive semidefinite and unitarily similar. Identify $\mathbf{R}$ and $\mathbf{S}$ in terms of $\mathbf{A}$.

If $\mathbf{A}$ is square and invertible, every singular value is positive and

$$\mathbf{A}^{-1} = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^*.$$

2.6.1 The four fundamental subspaces revisited

The SVD displays orthonormal bases for all four fundamental subspaces:

$$\begin{aligned}
\operatorname{range}(\mathbf{A}) &= \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_r\}, \\
\operatorname{coker}(\mathbf{A}) = \mathcal{K}(\mathbf{A}^*) &= \operatorname{span}\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}, \\
\operatorname{corange}(\mathbf{A}) = \operatorname{range}(\mathbf{A}^*) &= \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}, \\
\mathcal{K}(\mathbf{A}) &= \operatorname{span}\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}.
\end{aligned}$$

Thus a right singular vector describes an input direction, its singular value describes the stretching in that direction, and the corresponding left singular vector describes the output direction.

For the reduced SVD define the *Moore–Penrose pseudoinverse* by

$$\mathbf{A}^+ = \widetilde{\mathbf{V}}\widetilde{\boldsymbol{\Sigma}}^{-1}\widetilde{\mathbf{U}}^*.$$

Problem 2.17 (Moore–Penrose pseudoinverse). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ have rank r and pseudoinverse as above.

1. Prove that $\mathbf{A}^+ = \mathbf{A}^{-1}$ when $\mathbf{A}$ is invertible.
2. Prove that $\mathbf{A}\mathbf{A}^+$, $\mathbf{A}^+\mathbf{A}$, $\mathbf{I}_m - \mathbf{A}\mathbf{A}^+$, and $\mathbf{I}_n - \mathbf{A}^+\mathbf{A}$ are orthogonal projectors. Identify their ranges and kernels using the four fundamental subspaces.
3. Prove the identity $\mathbf{A} = \mathbf{A}\mathbf{A}^+\mathbf{A}$.
4. Determine exactly when $\mathbf{A}^+\mathbf{A} = \mathbf{I}_n$ and when $\mathbf{A}\mathbf{A}^+ = \mathbf{I}_m$.
5. For nonzero $\mathbf{A}$, decide whether a uniform perturbation bound can control the relative change
$$\frac{\|(\mathbf{A}+\mathbf{E})^+ - \mathbf{A}^+\|}{\|\mathbf{A}^+\|}$$
by $\|\mathbf{E}\| / \|\mathbf{A}\|$ for arbitrary small perturbations. Explain the role of rank changes.

2.7 Low-rank approximation

Low-rank matrices store less information and often isolate dominant structure. This makes them useful for data compression, denoising, simplification, and interpretable modeling. Given $k < \min\{m, n\}$, consider
$$\min_{\substack{\mathbf{B}\in\mathbb{C}^{m\times n}\\ \operatorname{rank}(\mathbf{B})\le k}} \|\mathbf{A} - \mathbf{B}\|. \tag{2.22}$$

Definition 2.15. A matrix norm is *unitarily invariant* when
$$\|\mathbf{A}\| = \|\mathbf{Q}\mathbf{A}\mathbf{Z}\|$$
for all unitary $\mathbf{Q}, \mathbf{Z}$ of appropriate sizes. Such a norm depends only on the singular values. Both the induced 2-norm and Frobenius norm are unitarily invariant.

For $0 \le k < \operatorname{rank}(\mathbf{A})$, define the truncated SVD
$$\mathbf{A}_k = \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^*, \tag{2.23}$$
with $\mathbf{A}_0 = \mathbf{0}$.

Problem 2.18 (Eckart–Young in the induced 2-norm). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $0 \le k < \operatorname{rank}(\mathbf{A})$. Prove that the truncated SVD (2.23) is a minimizer of
$$\min_{\operatorname{rank}(\mathbf{B})\le k} \|\mathbf{A} - \mathbf{B}\|_2$$
and that the minimum value is σ_{k+1}. Do not assume the minimizer is unique. For the lower bound, use $\dim \mathcal{K}(\mathbf{B}) \ge n - k$ to find a unit vector in both $\mathcal{K}(\mathbf{B})$ and $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$.

Theorem 2.16 (Eckart–Young–Mirsky)**.** *Let $k < \operatorname{rank}(\mathbf{A})$. For every unitarily invariant norm, a solution of* (2.22) *is the truncated SVD* (2.23)*.*

The theorem says that the best approximation is obtained by retaining the k strongest input-output directions and discarding the rest. In particular,

$$\|\mathbf{A} - \mathbf{A}_k\|_2 = \sigma_{k+1},$$
$$\|\mathbf{A} - \mathbf{A}_k\|_F = \left(\sum_{j=k+1}^{p} \sigma_j^2 \right)^{1/2}.$$

For the Frobenius norm the minimizer is unique exactly when $\sigma_k > \sigma_{k+1}$; uniqueness for a general unitarily invariant norm requires additional care.

Guided exercise 2.17. Use the orthogonality of the rank-one matrices $\mathbf{u}_j\mathbf{v}_j^*$ in the Frobenius inner product to derive the displayed Frobenius error formula. Explain geometrically why a repeated value $\sigma_k = \sigma_{k+1}$ permits more than one optimal rank-k subspace.

2.7.1 Principal component analysis

Let $\mathbf{B} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n] \in \mathbb{C}^{m\times n}$ be a data matrix whose columns are observations. Define the sample mean and centered matrix by

$$\mathbf{b}_0 = \frac{1}{n}\sum_{j=1}^{n} \mathbf{b}_j, \qquad \mathbf{A} = \mathbf{B} - \mathbf{b}_0\mathbf{1}^*,$$

where $\mathbf{1}^*$ is the $1 \times n$ row of ones. If $\mathbf{U}_k$ contains the first k left singular vectors of $\mathbf{A}$, then PCA uses the "scores" $\mathbf{c}_j = \mathbf{U}_k^*(\mathbf{b}_j - \mathbf{b}_0)$ and reconstructions/approximations $\tilde{\mathbf{b}}_j = \mathbf{U}_k\mathbf{c}_j + \mathbf{b}_0$.

Problem 2.19 (Principal component analysis). With the centered data matrix and notation above, let $P_{\mathbf{U}} = \mathbf{U}\mathbf{U}^*$ for any matrix $\mathbf{U}$ with orthonormal columns.

1. Prove that
$$P_{\mathbf{U}_k}\mathbf{A} = \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^*.$$

2. Prove that $P_{\mathbf{U}_k}\mathbf{A}$ is a solution of
$$\min_{\substack{\mathbf{U}\in\mathbb{C}^{m\times k}\\ \mathbf{U}^*\mathbf{U}=\mathbf{I}_k}} \|\mathbf{A} - P_{\mathbf{U}}\mathbf{A}\|_F^2.$$

3. For a k-dimensional subspace V, define
$$\mathrm{var}_V(\mathbf{B}) = \sum_{j=1}^{n} \left\| P_V(\mathbf{b}_j - \mathbf{b}_0) \right\|_2^2.$$

   Prove that $V = \mathrm{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ maximizes this variance over all k-dimensional subspaces.

Problem 2.20 (Eigenfaces and low-rank data experiments). Use the Yale Face Database dataset on Kaggle, or an equivalent labeled face dataset, for the following reproducible experiments. State the dataset, preprocessing, selected observations, values of k, and error measure.

Problem 2.20 (continued).

1. Treat each of several grayscale images separately as a matrix. Plot rank-k truncated-SVD approximations for several k and report the reconstruction error. This is image compression by low-rank SVD, not PCA.
2. Vectorize a collection of images as columns of a data matrix, subtract the sample mean, and compute centered rank-k PCA reconstructions. Compare reconstruction accuracy across several k with the single-image experiment.
3. Plot the centered PCA scores in two and three dimensions. Describe any visible relationship between the embeddings and available face labels.

Explain that a rank-k factored approximation to an $m \times n$ matrix uses $\mathcal{O}(k(m+n))$ storage rather than $\mathcal{O}(mn)$.

Problems for Chapter 2

The following statements repeat the problems from their inline locations. Each return link leads back to the first occurrence.

Problem 2.1 (Spectrum and diagonalization of a projection). Let $\mathbf{P} \in \mathbb{C}^{n\times n}$ be a projection of rank r, where $0 \le r \le n$. Determine its eigenvalues and their algebraic multiplicities. Prove that every projection is diagonalizable and construct an eigenbasis from vectors in its range and kernel.

Return to the inline occurrence of Problem 2.1.

Problem 2.2 (Gershgorin discs). Let $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{n\times n}$ and define

$$D_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \le \sum_{j\ne i} |a_{ij}| \right\}, \qquad 1 \le i \le n.$$

Prove that

$$\operatorname{spec}(\mathbf{A}) \subseteq \bigcup_{i=1}^{n} D_i.$$

Return to the inline occurrence of Problem 2.2.

Problem 2.3 (Determinant and trace from eigenvalues). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \dots, \lambda_n$, counted with algebraic multiplicity. Prove that

$$\det(\mathbf{A}) = \prod_{j=1}^{n} \lambda_j, \qquad \operatorname{tr}(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j.$$

Your argument must also cover defective matrices. Deduce that, when $\mathbf{A}$ is invertible, $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.

Return to the inline occurrence of Problem 2.3.

Problem 2.4 (Kronecker-product eigenpairs). Suppose $\mathbf{A} \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$ are diagonalizable. Express all eigenvalues and an eigenbasis of $\mathbf{A} \otimes \mathbf{B}$ in terms of eigenpairs of $\mathbf{A}$ and

Problem 2.4 (continued).

B.

Return to the inline occurrence of Problem 2.4.

Problem 2.5 (Spectral theorem for skew-Hermitian matrices). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Prove that $\mathbf{A}$ is skew-Hermitian if and only if it is unitarily diagonalizable and every eigenvalue is purely imaginary (with zero allowed).

Return to the inline occurrence of Problem 2.5.

Problem 2.6 (Numerical-range identities). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$.

1. Prove that $\rho(\mathbf{A}) \le w(\mathbf{A}) \le \|\mathbf{A}\|_2$.
2. For $\theta \in \mathbb{R}$, prove that $W(e^{i\theta}\mathbf{A}) = e^{i\theta} W(\mathbf{A})$.
3. Prove that $W(\mathbf{A}^*) = \{\overline{z} \mid z \in W(\mathbf{A})\}$.
4. Show that $W(H(\mathbf{A}))$ is the interval obtained by projecting $W(\mathbf{A})$ onto the real axis, and express its endpoints through the extreme eigenvalues of $H(\mathbf{A})$.

Return to the inline occurrence of Problem 2.6.

Problem 2.7 (Computing a numerical range by support lines). Design and implement an algorithm that samples angles θ and uses the largest eigenpair of H_θ to approximate the boundary of $W(\mathbf{A})$. Give pseudocode, explain how angular resolution affects the output, and test at least three matrices, including nilpotent Jordan blocks of at least three sizes.

For the size-s nilpotent Jordan block $J_s(0)$, investigate and justify the disk suggested by the computation. More generally, if

$$\mathbf{J} = \bigoplus_{\ell} J_{s_\ell}(\lambda_\ell),$$

describe $W(\mathbf{J})$ Explain why this description of $W(\mathbf{J})$ does not in general determine $W(\mathbf{A})$ for a matrix merely similar to $\mathbf{J}$.

Return to the inline occurrence of Problem 2.7.

Problem 2.8 (Rayleigh–Ritz exactness). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ and let $\mathbf{Q} \in \mathbb{C}^{n\times p}$ have orthonormal columns. Suppose $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ and $\mathbf{v} \in \operatorname{range}(\mathbf{Q})$. Prove that $\mathbf{c} = \mathbf{Q}^*\mathbf{v}$ is an eigenvector of $\mathbf{Q}^*\mathbf{A}\mathbf{Q}$ with eigenvalue λ, and that $\mathbf{v} = \mathbf{Q}\mathbf{c}$. Explain why the converse requires the Ritz residual $\mathbf{A}\mathbf{Q}\mathbf{c} - \lambda\mathbf{Q}\mathbf{c}$ to vanish.

Return to the inline occurrence of Problem 2.8.

Problem 2.9 (Proving Schur decomposition). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. Prove that there are a unitary $\mathbf{U}$ and an upper-triangular $\mathbf{T}$ such that

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*.$$

Return to the inline occurrence of Problem 2.9.

Problem 2.10 (Spectral theorem for normal matrices). Prove that a complex square matrix is unitarily diagonalizable if and only if it is normal.

Return to the inline occurrence of Problem 2.10.

Problem 2.11 (Eigenvalues versus singular values). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$.

1. If $\mathbf{A}$ is Hermitian positive semidefinite, prove that its eigenvalues and singular values agree when both are ordered nonincreasingly.
2. If $\mathbf{A}$ is normal, prove that its singular values are the moduli of its eigenvalues, counted with multiplicity and arranged nonincreasingly.

Return to the inline occurrence of Problem 2.11.

Problem 2.12 (Frobenius norm from singular values). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and let $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_p)^T$, where $p = \min\{m, n\}$. Prove that

$$\|\mathbf{A}\|_F^2 = \sum_{j=1}^{p} \sigma_j^2 = \|\boldsymbol{\sigma}\|_2^2.$$

Return to the inline occurrence of Problem 2.12.

Problem 2.13 (An inf–sup characterization of $\sigma_{\min}$). For $\mathbf{A} \in \mathbb{C}^{n\times n}$ define

$$\beta = inf_{\mathbf{x}\neq\mathbf{0}}\, sup_{\mathbf{y}\neq\mathbf{0}} \frac{|\mathbf{x}^*\mathbf{A}\mathbf{y}|}{\|\mathbf{x}\|_2\,\|\mathbf{y}\|_2}.$$

Show that $\beta = \sigma_{\min}(\mathbf{A})$ and hence that $\beta > 0$ if and only if $\mathbf{A}$ is invertible. Treat separately the case in which the vector whose direction would maximize the inner product is zero.

Return to the inline occurrence of Problem 2.13.

Problem 2.14 (Singular-value interlacing under compression). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{Q} \in \mathbb{C}^{(m-1)\times m}$ satisfy $\mathbf{Q}\mathbf{Q}^* = \mathbf{I}_{m-1}$. Write the singular values of $\mathbf{A}$ as $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$, where $p = \min\{m, n\}$, and those of $\mathbf{Q}\mathbf{A}$ as $\mu_1 \geq \cdots \geq \mu_q \geq 0$, where $q = \min\{m-1, n\}$. Prove the interlacing bounds

$$\sigma_j \geq \mu_j \geq \sigma_{j+1}, \qquad 1 \leq j \leq q,$$

using the convention $\sigma_{p+1} = 0$.

Return to the inline occurrence of Problem 2.14.

Problem 2.15 (Full-rank matrices are dense). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$, let $\|\cdot\|$ be any matrix norm, and let $\epsilon > 0$. Prove that there is a matrix $\mathbf{B}$ of rank $\min\{m, n\}$ such that

$$\|\mathbf{B} - \mathbf{A}\| < \epsilon.$$

Return to the inline occurrence of Problem 2.15.

Problem 2.16 (Unitary logarithms and polar decomposition).

1. Prove that every unitary $\mathbf{U} \in \mathbb{C}^{n\times n}$ can be written as $\mathbf{U} = e^{i\boldsymbol{\Theta}}$ for a Hermitian $\boldsymbol{\Theta}$. (Since $\boldsymbol{\Theta}$ is unitarily diagonalizable, we interpret $e^{i\boldsymbol{\Theta}}$ through the usual functional calculus where $\boldsymbol{\Theta} \mapsto e^{i\boldsymbol{\Theta}}$ is computed by shifting the complex exponential to operate on the individual eigenvalues.)
2. Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Prove that

$$\mathbf{A} = \mathbf{R}e^{i\boldsymbol{\Theta}} = e^{i\boldsymbol{\Theta}}\mathbf{S},$$

Problem 2.16 (continued).

where $\mathbf{R}$ and $\mathbf{S}$ are Hermitian positive semidefinite and unitarily similar. Identify $\mathbf{R}$ and $\mathbf{S}$ in terms of $\mathbf{A}$.

Return to the inline occurrence of Problem 2.16.

Problem 2.17 (Moore–Penrose pseudoinverse). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ have rank r and pseudoinverse as above.

1. Prove that $\mathbf{A}^+ = \mathbf{A}^{-1}$ when $\mathbf{A}$ is invertible.
2. Prove that $\mathbf{A}\mathbf{A}^+$, $\mathbf{A}^+\mathbf{A}$, $\mathbf{I}_m - \mathbf{A}\mathbf{A}^+$, and $\mathbf{I}_n - \mathbf{A}^+\mathbf{A}$ are orthogonal projectors. Identify their ranges and kernels using the four fundamental subspaces.
3. Prove the identity $\mathbf{A} = \mathbf{A}\mathbf{A}^+\mathbf{A}$.
4. Determine exactly when $\mathbf{A}^+\mathbf{A} = \mathbf{I}_n$ and when $\mathbf{A}\mathbf{A}^+ = \mathbf{I}_m$.
5. For nonzero $\mathbf{A}$, decide whether a uniform perturbation bound can control the relative change

$$\frac{\|(\mathbf{A}+\mathbf{E})^+ - \mathbf{A}^+\|}{\|\mathbf{A}^+\|}$$

by $\|\mathbf{E}\| / \|\mathbf{A}\|$ for arbitrary small perturbations. Explain the role of rank changes.

Return to the inline occurrence of Problem 2.17.

Problem 2.18 (Eckart–Young in the induced 2-norm). Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $0 \le k < \operatorname{rank}(\mathbf{A})$. Prove that the truncated SVD (2.23) is a minimizer of

$$\min_{\operatorname{rank}(\mathbf{B})\le k} \|\mathbf{A} - \mathbf{B}\|_2$$

and that the minimum value is σ_{k+1}. Do not assume the minimizer is unique. For the lower bound, use $\dim \mathcal{K}(\mathbf{B}) \ge n - k$ to find a unit vector in both $\mathcal{K}(\mathbf{B})$ and $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$.

Return to the inline occurrence of Problem 2.18.

Problem 2.19 (Principal component analysis). With the centered data matrix and notation above, let $P_{\mathbf{U}} = \mathbf{U}\mathbf{U}^*$ for any matrix $\mathbf{U}$ with orthonormal columns.

1. Prove that

$$P_{\mathbf{U}_k}\mathbf{A} = \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^*.$$

2. Prove that $P_{\mathbf{U}_k}\mathbf{A}$ is a solution of

$$\min_{\substack{\mathbf{U}\in\mathbb{C}^{m\times k}\\ \mathbf{U}^*\mathbf{U}=\mathbf{I}_k}} \|\mathbf{A} - P_{\mathbf{U}}\mathbf{A}\|_F^2 .$$

3. For a k-dimensional subspace V, define

$$\operatorname{var}_V(\mathbf{B}) = \sum_{j=1}^{n} \left\|P_V(\mathbf{b}_j - \mathbf{b}_0)\right\|_2^2 .$$

Problem 2.19 (continued).

Prove that $V = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ maximizes this variance over all k-dimensional subspaces.

Return to the inline occurrence of Problem 2.19.

Problem 2.20 (Eigenfaces and low-rank data experiments). Use the Yale Face Database dataset on Kaggle, or an equivalent labeled face dataset, for the following reproducible experiments. State the dataset, preprocessing, selected observations, values of k, and error measure.

1. Treat each of several grayscale images separately as a matrix. Plot rank-k truncated-SVD approximations for several k and report the reconstruction error. This is image compression by low-rank SVD, not PCA.
2. Vectorize a collection of images as columns of a data matrix, subtract the sample mean, and compute centered rank-k PCA reconstructions. Compare reconstruction accuracy across several k with the single-image experiment.
3. Plot the centered PCA scores in two and three dimensions. Describe any visible relationship between the embeddings and available face labels.

Explain that a rank-k factored approximation to an $m \times n$ matrix uses $\mathcal{O}(k(m+n))$ storage rather than $\mathcal{O}(mn)$.

Return to the inline occurrence of Problem 2.20.