Advanced Methods for Information Representation
Exercise Session 3 – Solutions

Linear operators, adjoints and projections

University of Brescia

This document contains the complete solutions to the exercises discussed in Exercise Session 3. The notation follows *Foundations of Signal Processing*.

Useful facts and notation

We recall the definitions used repeatedly in this session.

A linear operator $A : H_0 \to H_1$ is bounded if

$$\|A\| := \sup_{\|x\|=1} \|Ax\| < \infty.$$

The adjoint $A^* : H_1 \to H_0$ is characterized by

$$\langle Ax, y \rangle_{H_1} = \langle x, A^* y \rangle_{H_0}.$$

A projection operator satisfies

$$P^2 = P,$$

and it is an *orthogonal* projection when, in addition,

$$P^* = P.$$

For real matrices, $P^* = P^T$.

Conceptual message

Every orthogonal projection is a projection, but not every projection is orthogonal. The difference is geometric: an oblique projection reaches the target subspace along a non-orthogonal direction.

1. Operator norm of a matrix: direct optimization

Exercise 1. *Consider*

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2,$$

with the Euclidean norm.

(a) Compute the induced operator norm $\|A\|_2$ by writing a generic unit vector as

$$x(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and optimizing over θ.

(b) Find all unit vectors at which the maximum is attained.

(c) Compute the smallest possible value of $\|Ax\|_2$ *for* $\|x\|_2 = 1$.

Solution

A generic unit vector of $\mathbb{R}^2$ is

$$x(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

Therefore

$$Ax(\theta) = \begin{pmatrix} 3\cos\theta + \sin\theta \\ \cos\theta + 3\sin\theta \end{pmatrix}.$$

Instead of differentiating the square root, optimize its square:

$$\begin{aligned} \|Ax(\theta)\|_2^2 &= (3\cos\theta + \sin\theta)^2 + (\cos\theta + 3\sin\theta)^2 \\ &= 10(\cos^2\theta + \sin^2\theta) + 12\sin\theta\cos\theta \\ &= 10 + 6\sin(2\theta). \end{aligned}$$

Since $-1 \le \sin(2\theta) \le 1$,

$$4 \le \|Ax(\theta)\|_2^2 \le 16.$$

Hence

$$\boxed{\|A\|_2 = 4}.$$

The maximum occurs when

$$\sin(2\theta) = 1, \qquad \theta = \frac{\pi}{4} + k\pi,$$

so the maximizing unit vectors are

$$\boxed{x = \pm\frac{1}{\sqrt{2}}(1,1)^T}.$$

Likewise the minimum occurs when $\sin(2\theta) = -1$, i.e.

$$\theta = -\frac{\pi}{4} + k\pi,$$

and therefore

$$\boxed{\min_{\|x\|_2=1} \|Ax\|_2 = 2}, \qquad x = \pm\frac{1}{\sqrt{2}}(1,-1)^T.$$

Conceptual message

This is the definition of the induced norm used literally: parametrize the unit sphere, compute $\|Ax\|$, and optimize. The spectral method introduced next packages this optimization into an eigenvalue problem.

2. Spectral formula for the Euclidean operator norm

Exercise 2 (FSP, Exercise 2.26)**.** *Let* $A \in \mathbb{C}^{M\times N}$*, with the Euclidean norm on domain and codomain. Prove that*

$$\boxed{\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)}}.$$

Also show that any unit eigenvector of A^*A associated with $\lambda_{\max}$ is a maximizing direction for the operator norm.

Solution

The matrix A^*A is Hermitian and positive semidefinite because

$$(A^*A)^* = A^*A, \qquad x^*A^*Ax = \|Ax\|_2^2 \geq 0.$$

Hence it admits an orthonormal eigenbasis. Write

$$A^*A = U\Lambda U^*,$$

where

$$\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_N), \qquad \lambda_i \geq 0.$$

For $\|x\|_2 = 1$, set $y = U^*x$. Since U is unitary, $\|y\|_2 = 1$, and

$$\begin{aligned} \|Ax\|_2^2 &= x^*A^*Ax \\ &= y^*\Lambda y \\ &= \sum_{i=1}^{N} \lambda_i |y_i|^2 \\ &\leq \lambda_{\max} \sum_{i=1}^{N} |y_i|^2 = \lambda_{\max}. \end{aligned}$$

Thus

$$\|A\|_2^2 \leq \lambda_{\max}(A^*A).$$

Choose now $x = v_{\max}$, a unit eigenvector satisfying

$$A^*Av_{\max} = \lambda_{\max} v_{\max}.$$

Then

$$\|Av_{\max}\|_2^2 = v_{\max}^* A^*Av_{\max} = \lambda_{\max}.$$

Therefore the upper bound is attained and

$$\boxed{\|A\|_2^2 = \lambda_{\max}(A^*A)}, \qquad \boxed{\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)}}.$$

Conceptual message

The matrix A^*A converts the geometric optimization problem into a Rayleigh-quotient problem. Its eigenvalues are the squared singular values of A.

3. A complex matrix: compute the operator norm using A^*A

Exercise 3. *Let*

$$A = \begin{pmatrix} 1 & \mathrm{j} & 0 \\ 1 & 0 & \mathrm{j} \end{pmatrix} : \mathbb{C}^3 \to \mathbb{C}^2, \qquad \mathrm{j}^2 = -1.$$

(a) Compute A^* and A^*A.

(b) Find the eigenvalues of A^*A.

(c) Compute $\|A\|_2$ using the result of Exercise 2.

(d) Find one unit vector at which the maximum is attained.

Solution

Taking the Hermitian transpose,

$$A^* = \begin{pmatrix} 1 & 1 \\ -\mathrm{j} & 0 \\ 0 & -\mathrm{j} \end{pmatrix}.$$

Therefore

$$A^*A = \begin{pmatrix} 2 & \mathrm{j} & \mathrm{j} \\ -\mathrm{j} & 1 & 0 \\ -\mathrm{j} & 0 & 1 \end{pmatrix}.$$

Its characteristic polynomial is

$$\det(\lambda I - A^*A) = \lambda(\lambda - 1)(\lambda - 3),$$

so

$$\sigma(A^*A) = \{0, 1, 3\}.$$

By Exercise 2,

$$\boxed{\|A\|_2 = \sqrt{3}}.$$

For $\lambda_{\max} = 3$, the equations

$$(A^*A - 3I)x = 0$$

give

$$x_1 = x_2 = -\frac{\mathrm{j}}{2}x_0.$$

Choosing the normalization $\|x\|_2 = 1$ yields, for instance,

$$\boxed{x_{\max} = \begin{pmatrix} \sqrt{2/3} \\ -\mathrm{j}/\sqrt{6} \\ -\mathrm{j}/\sqrt{6} \end{pmatrix}}.$$

Indeed

$$\|Ax_{\max}\|_2^2 = 3.$$

Conceptual message

This is the typical workflow for a general complex matrix: form the Hermitian matrix A^*A, compute its largest eigenvalue, and take the square root.

4. Properties of the adjoint operator

Exercise 4 (adjoint algebra)**.** *Let $A : H_0 \to H_1$ and $B : H_1 \to H_2$ be bounded linear operators between Hilbert spaces, and assume the relevant adjoints exist. Prove the following properties.*

(a) The adjoint is unique and $(A^*)^* = A$.

(b) If $C : H_0 \to H_1$ is bounded, then

$$(A + C)^* = A^* + C^*.$$

(c)

$$(BA)^* = A^* B^*.$$

(d) A^*A and AA^* are self-adjoint and positive semidefinite.

(e) If A is invertible, then

$$(A^{-1})^* = (A^*)^{-1}.$$

(f)

$$\|A^*\| = \|A\|.$$

Solution

Recall the defining identity

$$\langle Ax, y\rangle_{H_1} = \langle x, A^* y\rangle_{H_0}.$$

(a) Uniqueness and involution. If $T_1, T_2 : H_1 \to H_0$ both satisfy the adjoint identity, then

$$\langle x, (T_1 - T_2)y\rangle = 0 \qquad \forall x, y.$$

For fixed y, choose $x = (T_1 - T_2)y$. Then

$$\|(T_1 - T_2)y\|^2 = 0,$$

so $T_1 = T_2$.
Moreover, by conjugate symmetry of the inner product,

$$\langle A^* y, x\rangle_{H_0} = \langle y, Ax\rangle_{H_1},$$

which says precisely that A is the adjoint of A^*. Hence

$$\boxed{(A^*)^* = A}.$$

(b) Sum. For every x, y,

$$\begin{aligned}\langle (A + C)x, y\rangle &= \langle Ax, y\rangle + \langle Cx, y\rangle \\ &= \langle x, A^* y\rangle + \langle x, C^* y\rangle \\ &= \langle x, (A^* + C^*)y\rangle.\end{aligned}$$

By uniqueness,

$$\boxed{(A + C)^* = A^* + C^*}.$$

(c) Product. For $x \in H_0$ and $z \in H_2$,

$$\begin{aligned}\langle BAx, z\rangle_{H_2} &= \langle Ax, B^* z\rangle_{H_1} \\ &= \langle x, A^* B^* z\rangle_{H_0}.\end{aligned}$$

Therefore

$$\boxed{(BA)^* = A^* B^*}.$$

(d) Self-adjointness and positivity. Using the product rule,

$$(A^*A)^* = A^*(A^*)^* = A^*A,$$

and similarly $(AA^*)^* = AA^*$. Furthermore,

$$\langle A^*Ax, x\rangle = \|Ax\|^2 \geq 0,$$

and

$$\langle AA^*y, y\rangle = \|A^*y\|^2 \geq 0.$$

Thus both operators are positive semidefinite.

(e) Inverse. Since $AA^{-1} = I$ and $A^{-1}A = I$, taking adjoints gives

$$(A^{-1})^*A^* = I, \qquad A^*(A^{-1})^* = I.$$

Hence

$$\boxed{(A^{-1})^* = (A^*)^{-1}}.$$

(f) Equality of the norms. Use the Hilbert-space identity

$$\|z\| = \sup_{\|x\|=1} |\langle z, x\rangle|.$$

For $\|y\| = 1$,

$$\begin{aligned}\|A^*y\| &= \sup_{\|x\|=1} |\langle A^*y, x\rangle| \\ &= \sup_{\|x\|=1} |\langle y, Ax\rangle| \\ &\leq \|A\|.\end{aligned}$$

Thus $\|A^*\| \leq \|A\|$. Apply the same inequality to A^* and use $(A^*)^* = A$:

$$\|A\| \leq \|A^*\|.$$

Therefore

$$\boxed{\|A^*\| = \|A\|}.$$

Conceptual message

For matrices these identities become the familiar rules for the Hermitian transpose:

$$(A + C)^* = A^* + C^*, \qquad (BA)^* = A^*B^*, \qquad (A^{-1})^* = (A^*)^{-1}.$$

The abstract adjoint is exactly the Hilbert-space generalization of conjugate transpose.

5. Adjoint of a local-averaging operator

Exercise 5 (local averaging)**.** *For*

$$I_k = \left[k - \frac{1}{2}, k + \frac{1}{2}\right), \qquad k \in \mathbb{Z},$$

define

$$A : L^2(\mathbb{R}) \to \ell^2(\mathbb{Z}), \qquad (Ax)_k = \int_{I_k} x(t)\, dt.$$

(a) Show that A is bounded and compute $\|A\|$.

(b) Compute the adjoint A^* explicitly.

(c) Show that $AA^* = I$ on $\ell^2(\mathbb{Z})$.

(d) Show that $P = A^*A$ is the orthogonal projection of $L^2(\mathbb{R})$ onto the subspace of functions that are constant on every interval I_k.

Solution

(a) Boundedness. By Cauchy–Schwarz on each interval I_k, whose length is 1,

$$|(Ax)_k|^2 = \left| \int_{I_k} x(t)\, dt \right|^2 \leq \int_{I_k} |x(t)|^2\, dt.$$

Summing over k gives

$$\|Ax\|_{\ell^2}^2 \leq \sum_k \int_{I_k} |x(t)|^2 dt = \|x\|_{L^2}^2 .$$

Thus $\|A\| \leq 1$. Equality is obtained, for example, with $x = \mathbf{1}_{I_0}$, so

$$\boxed{\|A\| = 1}.$$

(b) Adjoint. For $y = (y_k) \in \ell^2(\mathbb{Z})$,

$$\begin{aligned} \langle Ax, y \rangle_{\ell^2} &= \sum_{k \in \mathbb{Z}} \left(\int_{I_k} x(t)\, dt \right) y_k^* \\ &= \int_{\mathbb{R}} x(t) \left(\sum_{k \in \mathbb{Z}} y_k \mathbf{1}_{I_k}(t) \right)^* dt. \end{aligned}$$

Therefore

$$\boxed{(A^*y)(t) = \sum_{k \in \mathbb{Z}} y_k \mathbf{1}_{I_k}(t).}$$

The adjoint converts a sequence into a staircase function.

(c) $AA^* = I$. On I_k, A^*y is equal to the constant y_k, hence

$$(AA^*y)_k = \int_{I_k} y_k\, dt = y_k.$$

So

$$\boxed{AA^* = I_{\ell^2}}.$$

Notice that this does *not* imply $A^*A = I_{L^2}$.

(d) The operator $P = A^*A$. Explicitly,

$$(Px)(t) = \sum_{k \in \mathbb{Z}} \left(\int_{I_k} x(s)\, ds \right) \mathbf{1}_{I_k}(t).$$

Moreover,

$$P^2 = A^*(AA^*)A = A^*A = P,$$

and

$$P^* = (A^*A)^* = A^*A = P.$$

Thus P is an orthogonal projection. Its range consists exactly of the staircase functions that are constant on each I_k.

Common pitfall

The adjoint is not the same thing as the inverse. Here A^* is a right inverse of A because $AA^* = I$, but it is not a two-sided inverse since $A^*A \neq I$ on all of $L^2(\mathbb{R})$.

6. Oblique versus orthogonal projection

Exercise 6. *Consider*

$$P = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

(a) Verify that $P^2 = P$ but $P^T \neq P$.

(b) Determine $\mathrm{Ran}(P)$ *and* $\mathrm{Ker}(P)$.

(c) Explain geometrically why P is an oblique projection.

(d) Find the orthogonal projector Q onto the same range.

(e) For $x = (6, 6, 8)^T$, compute Px and Qx.

Solution

Direct multiplication gives

$$P^2 = P,$$

while, for instance, $P_{12} = 1/2$ and $P_{21} = 0$, so $P^T \neq P$. Therefore P is a projection, but not an orthogonal projection.
For $y = Px$,

$$y_0 = \frac{x_0 + x_1 - x_2}{2}, \qquad y_1 = x_1, \qquad y_2 = \frac{-x_0 + x_1 + x_2}{2}.$$

Hence

$$y_0 + y_2 = y_1.$$

Thus

$$\mathrm{Ran}(P) = S := \{y \in \mathbb{R}^3 : y_0 + y_2 = y_1\}.$$

Solving $Px = 0$ gives

$$x_1 = 0, \qquad x_0 = x_2,$$

so

$$\mathrm{Ker}(P) = \mathrm{span}\{(1, 0, 1)^T\}.$$

The plane S can be written as

$$y_0 - y_1 + y_2 = 0,$$

so a normal vector is

$$n = (1, -1, 1)^T.$$

Since

$$\mathrm{Ker}(P) \neq S^\perp = \mathrm{span}\{n\},$$

the projection direction is not perpendicular to the target plane: P is oblique.
The orthogonal projector onto $S = n^\perp$ is

$$Q = I - \frac{nn^T}{\|n\|_2^2} = I - \frac{1}{3}nn^T,$$

therefore

$$\boxed{Q = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}}.$$

For $x = (6, 6, 8)^T$,

$$\boxed{Px = (2, 6, 4)^T},$$

whereas

$$\boxed{Qx = \left(\frac{10}{3}, \frac{26}{3}, \frac{16}{3}\right)^T}.$$

Conceptual message

The equation $P^2 = P$ identifies a projection algebraically. The extra condition $P^* = P$ identifies the orthogonal projection geometrically.

7. All projections onto the range of a matrix

Exercise 7 (adapted from FSP, Exercise 2.32)**.** *Let*

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^3.$$

Find all projection operators onto $\mathrm{Ran}(B)$ *and identify the orthogonal one.*

Solution

Start with the structural observation.
If $A : \mathbb{R}^3 \to \mathbb{R}^2$ is a left inverse of B, so

$$AB = I_2,$$

then

$$P := BA$$

is a projection, because

$$P^2 = BABA = B(AB)A = BA = P.$$

Moreover,

$$\mathrm{Ran}(P) = \mathrm{Ran}(B).$$

Indeed $Px \in \mathrm{Ran}(B)$ for every x, while for any $Bu \in \mathrm{Ran}(B)$,

$$P(Bu) = B(AB)u = Bu.$$

Now write

$$A = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \end{pmatrix}.$$

The condition $AB = I_2$ gives

$$a_{00} + a_{01} = 1, \qquad a_{01} + a_{02} = 0,$$

$$a_{10} + a_{11} = 0, \qquad a_{11} + a_{12} = 1.$$

Set

$$\alpha = a_{01}, \qquad \beta = a_{11}.$$

Then every left inverse is

$$A_{\alpha,\beta} = \begin{pmatrix} 1-\alpha & \alpha & -\alpha \\ -\beta & \beta & 1-\beta \end{pmatrix}.$$

Hence all the projections onto $\mathrm{Ran}(B)$ are

$$P_{\alpha,\beta} = BA_{\alpha,\beta} = \begin{pmatrix} 1-\alpha & \alpha & -\alpha \\ 1-\alpha-\beta & \alpha+\beta & 1-\alpha-\beta \\ -\beta & \beta & 1-\beta \end{pmatrix}.$$

For orthogonality we need

$$P_{\alpha,\beta}^T = P_{\alpha,\beta}.$$

From the off-diagonal entries,

$$\alpha = \beta, \qquad \alpha = 1 - \alpha - \beta.$$

Thus

$$\boxed{\alpha = \beta = \frac{1}{3}}.$$

Therefore the unique orthogonal projector is

$$\boxed{P_\perp = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}}.$$

This is exactly the matrix Q found in Exercise 6.

Conceptual message

There are infinitely many projections onto the same subspace, because one may project along many different complementary directions. There is only one orthogonal projection onto a fixed closed subspace.

8. Why the orthogonal projection is the best approximation

Exercise 8. *Let S be a closed subspace of a Hilbert space H, and let Q be the orthogonal projector onto S. Prove that for every $x \in H$ and every $s \in S$,*

$$\|x - Qx\| \le \|x - s\|,$$

with equality only when $s = Qx$.

Then compare the two projections from Exercise 6 for $x = (6, 6, 8)^T$.

Solution

Since $Qx \in S$ and $x - Qx \in S^\perp$, for every $s \in S$ we have

$$Qx - s \in S$$

and therefore

$$(x - Qx) \perp (Qx - s).$$

Now decompose

$$x - s = (x - Qx) + (Qx - s).$$

Pythagoras gives

$$\|x - s\|^2 = \|x - Qx\|^2 + \|Qx - s\|^2 \geq \|x - Qx\|^2 .$$

Thus

$$\boxed{\|x - Qx\| \leq \|x - s\|}.$$

Equality holds iff $\|Qx - s\| = 0$, i.e. iff $s = Qx$.
For the concrete vector in Exercise 6,

$$Px = (2, 6, 4)^T, \qquad Qx = \left(\frac{10}{3}, \frac{26}{3}, \frac{16}{3}\right)^T .$$

The squared errors are

$$\|x - Px\|_2^2 = \|(4, 0, 4)\|_2^2 = 32,$$

while

$$\|x - Qx\|_2^2 = \|(8/3, -8/3, 8/3)\|_2^2 = \frac{64}{3} < 32.$$

So the orthogonal projection is genuinely the closest point in the plane.

Backup exercise A. Projection by restriction of the domain

Additional exercise

For a measurable set $E \subset \mathbb{R}$, define

$$(P_E f)(t) = \mathbf{1}_E(t) f(t), \qquad f \in L^2(\mathbb{R}).$$

(a) Prove that P_E is an orthogonal projection.

(b) Identify its range and kernel.

(c) If E_1 and E_2 are disjoint up to sets of measure zero, prove that

$$\mathrm{Ran}(P_{E_1}) \perp \mathrm{Ran}(P_{E_2}).$$

(d) State the condition under which

$$L^2(\mathbb{R}) = \mathrm{Ran}(P_{E_1}) \oplus \mathrm{Ran}(P_{E_2}).$$

Solution

We have

$$P_E^2 f = \mathbf{1}_E^2 f = \mathbf{1}_E f = P_E f,$$

so P_E is idempotent. Also

$$\langle P_E f, g \rangle = \int_E f(t) g(t)^* dt = \langle f, P_E g \rangle ,$$

so $P_E^* = P_E$. Hence it is an orthogonal projection.
Its range is

$$\mathrm{Ran}(P_E) = \{f \in L^2(\mathbb{R}) : f = 0 \text{ a.e. on } E^c\},$$

and its kernel is

$$\mathrm{Ker}(P_E) = \{f \in L^2(\mathbb{R}) : f = 0 \text{ a.e. on } E\}.$$

If $E_1 \cap E_2$ has measure zero, then for f supported on E_1 and g supported on E_2,

$$\langle f, g\rangle = 0.$$

Finally,

$$L^2(\mathbb{R}) = \mathrm{Ran}(P_{E_1}) \oplus \mathrm{Ran}(P_{E_2})$$

precisely when E_1, E_2 form a measurable partition of $\mathbb{R}$ up to null sets:

$$E_1 \cap E_2 = \varnothing \text{ a.e.}, \qquad E_1 \cup E_2 = \mathbb{R} \text{ a.e.}$$

Backup exercise B. Unitary operators

Additional exercise

Let $U : H \to H$ be invertible. Show that the following are equivalent:

(i) $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all $x, y \in H$;

(ii) $U^*U = I$;

(iii) $U^{-1} = U^*$.

Verify these properties for a rotation matrix in $\mathbb{R}^2$.

Solution

By the definition of adjoint,

$$\langle Ux, Uy\rangle = \langle x, U^*Uy\rangle .$$

Thus preservation of inner products for every x, y is equivalent to

$$U^*U = I.$$

If U is invertible, this is equivalent to

$$U^* = U^{-1}.$$

For

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

one checks directly that

$$R_\theta^T R_\theta = I, \qquad R_\theta^{-1} = R_\theta^T = R_{-\theta}.$$