Exercises: Linear Algebra Recall (for Optimization)

1. Diagonalization and spectrum

Exercise 1 (Spectrum and invertibility). *Let $A \in \mathbb{R}^{n\times n}$.*

(a) Show that A is invertible if and only if $0 \notin \sigma(A)$.

(b) Show that if A is invertible then every eigenvalue of A^{-1} is of the form $1/\lambda$ where $\lambda \in \sigma(A)$.

Answers:

(a) A invertible $\Leftrightarrow \det(A) \neq 0$. But $0 \in \sigma(A)$ $\Leftrightarrow \det(A - 0 \cdot I) = \det(A) = 0$.

(b) If $Ax = \lambda x$ with $x \neq 0$ and $\lambda \neq 0$, apply A^{-1}: $x = \lambda A^{-1}x$, so $A^{-1}x = (1/\lambda)x$.

Exercise 2 (Distinct eigenvalues $\Rightarrow$ diagonalizable). *Assume $A \in \mathbb{R}^{n\times n}$ has n* distinct *real eigenvalues. Show that A is diagonalizable over $\mathbb{R}$.*

Answers: Eigenvectors associated with distinct eigenvalues are linearly independent. With n distinct eigenvalues we get n independent eigenvectors, hence a basis, hence $A = PDP^{-1}$.

Exercise 3 (Orthogonal diagonalization is special). *Give an example of a diagonalizable (real) matrix A that is* not *orthogonally diagonalizable (i.e., there is no orthogonal Q with $A = Q\Lambda Q^\top$).*

Answers: Take any non-symmetric diagonalizable matrix, e.g. $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ (distinct eigenvalues $1, 2$ so diagonalizable). If $A = Q\Lambda Q^\top$, then A would be symmetric (since $Q\Lambda Q^\top$ is always symmetric), contradiction.

2. PSD order and spectral theorem

Exercise 4 (Loewner order is a partial order). *On S_n (real symmetric matrices), define $A \preceq B$ iff $B - A \succeq 0$. Show that $\preceq$ is a partial order on S_n (reflexive, antisymmetric, transitive).*

Answers: Reflexive: $A - A = 0 \succeq 0$. Transitive: if $B - A \succeq 0$ and $C - B \succeq 0$ then $C - A = (C - B) + (B - A) \succeq 0$. Antisymmetric: if $A \preceq B$ and $B \preceq A$, then $B - A \succeq 0$ and $A - B \succeq 0$ so for all x, $x^\top(B - A)x \geq 0$ and ≤ 0, hence $x^\top(B - A)x = 0$ for all x which implies $B - A = 0$.

Exercise 5 (PSD/PD and eigenvalues). *Let $A \in S_n$ with spectral decomposition $A = Q\Lambda Q^\top$.*

(a) Prove $A \succeq 0$ if and only if all eigenvalues satisfy $\lambda_i \geq 0$.

(b) Prove $A \succ 0$ if and only if all eigenvalues satisfy $\lambda_i > 0$.

(c) Assume $A \succeq 0$. Define $A^{1/2} = Q\Lambda^{1/2}Q^\top$ with $\Lambda^{1/2} = \mathrm{diag}(\sqrt{\lambda_i})$. Show that $A^{1/2}A^{1/2} = A$.

Answers: Write $x = Qz$ so $x^\top Ax = z^\top \Lambda z = \sum_i \lambda_i z_i^2$. Then:

(a) $x^\top Ax \geq 0$ for all x iff $\sum_i \lambda_i z_i^2 \geq 0$ for all z, which holds iff all $\lambda_i \geq 0$.

(b) Similar: strict inequality for all $x \neq 0$ iff all $\lambda_i > 0$. (c) $A^{1/2}A^{1/2} = Q\Lambda^{1/2}(Q^\top Q)\Lambda^{1/2}Q^\top = Q\Lambda Q^\top = A$.

Exercise 6 (Rayleigh quotient extrema). *Let $A \in S_n$ and define $R_A(x) = \dfrac{x^\top Ax}{x^\top x}$ for $x \neq 0$. Show that*

$$\lambda_{\min}(A) = \min_{\|x\|_2=1} x^\top Ax, \qquad \lambda_{\max}(A) = \max_{\|x\|_2=1} x^\top Ax.$$

Answers: With $A = Q\Lambda Q^\top$ and $x = Qz$ with $\|z\|_2 = 1$, we get $x^\top Ax = z^\top \Lambda z = \sum_i \lambda_i z_i^2$, a convex combination of eigenvalues, hence lies in $[\lambda_{\min}, \lambda_{\max}]$ and the endpoints are reached by $z = e_{i^*}$.

Exercise 7 (Eigenvalue bounds from PSD order)**.** *Let $A, B \in S_n$ and assume $A \preceq B$. Prove that $\lambda_{\min}(A) \le \lambda_{\min}(B)$ and $\lambda_{\max}(A) \le \lambda_{\max}(B)$.*

Answers: Use the Rayleigh characterization: $\lambda_{\max}(A) = \max_{\|x\|=1} x^\top Ax \le \max_{\|x\|=1} x^\top Bx = \lambda_{\max}(B)$ since $x^\top Ax \le x^\top Bx$ for all x. Similarly for $\lambda_{\min}$ using min.

3. Orthogonal projectors

Exercise 8 (Projector onto a subspace given an orthonormal basis)**.** *Let $Q \in \mathbb{R}^{n\times k}$ have orthonormal columns ($Q^\top Q = I_k$). Define $P = QQ^\top$.*

(a) Prove that P is an orthogonal projector.

(b) Prove that $\mathrm{Im}(P) = \mathrm{Im}(Q)$.

(c) Show that for all x, Px is the unique minimizer of $\min_{y\in\mathrm{Im}(Q)} \|x - y\|_2$.

Answers:

(a) $P^\top = (QQ^\top)^\top = QQ^\top$. Also $P^2 = QQ^\top QQ^\top = Q(Q^\top Q)Q^\top = QQ^\top = P$.

(b) $\mathrm{Im}(P) \subseteq \mathrm{Im}(Q)$ since $Px = Q(Q^\top x)$; and $Qw = P(Qw)$ so $\mathrm{Im}(Q) \subseteq \mathrm{Im}(P)$.

(c) Standard projection theorem: $x - Px \perp \mathrm{Im}(Q)$ because for any w, Qw satisfies $(x - Px)^\top Qw = x^\top Qw - x^\top QQ^\top Qw = 0$.

Exercise 9 (Projection onto a hyperplane)**.** *Let $a \in \mathbb{R}^n$ with $a \ne 0$ and consider the hyperplane $H = \{x \in \mathbb{R}^n : a^\top x = b\}$.*

(a) Show that the Euclidean projection of x_0 onto H is

$$\Pi_H(x_0) = x_0 - \frac{a^\top x_0 - b}{\|a\|_2^2}\, a.$$

(b) Deduce the projector onto the subspace $\{x : a^\top x = 0\}$.

Answers: (a) Solve $\min_x \frac{1}{2}\|x - x_0\|_2^2$ s.t. $a^\top x = b$ by KKT: $x - x_0 + \lambda a = 0$ so $x = x_0 - \lambda a$. Enforce constraint: $a^\top x_0 - \lambda\|a\|^2 = b$, hence $\lambda = (a^\top x_0 - b)/\|a\|^2$. (b) Take $b = 0$.

4. Norms and inner products

Exercise 10 (Norm inequalities)**.** *Let $x \in \mathbb{R}^n$. Prove*

$$\|x\|_\infty \le \|x\|_2 \le \|x\|_1 \le \sqrt{n}\,\|x\|_2 \le n\,\|x\|_\infty.$$

Answers: $\|x\|_\infty \le \|x\|_2$ since $\|x\|_2^2 = \sum_i x_i^2 \ge \max_i x_i^2$. Also $\|x\|_2 \le \|x\|_1$ since $\sum_i x_i^2 \le (\sum_i |x_i|)^2$. Next, Cauchy–Schwarz gives $\|x\|_1 = \langle |x|, \mathbf{1}\rangle \le \||x|\|_2 \|\mathbf{1}\|_2 = \sqrt{n}\,\|x\|_2$. Finally $\|x\|_2^2 = \sum_i x_i^2 \le n\,\|x\|_\infty^2$.

Exercise 11 (Cauchy–Schwarz equality case)**.** *Let $\langle\cdot,\cdot\rangle$ be an inner product and $\|x\| = \sqrt{\langle x, x\rangle}$. Prove that for $x \ne 0$ and $y \ne 0$, equality in $|\langle x, y\rangle| \le \|x\|\,\|y\|$ holds if and only if $y = \alpha x$ for some $\alpha \in \mathbb{R}$.*

Answers: Consider $\|y - \alpha x\|^2 \ge 0$ and choose $\alpha = \langle x, y\rangle/\|x\|^2$ to minimize. Then $0 \le \|y\|^2 - \langle x, y\rangle^2/\|x\|^2$, i.e. $\langle x, y\rangle^2 \le \|x\|^2\|y\|^2$. Equality iff the minimizer achieves $\|y - \alpha x\| = 0$.

Exercise 12 ($\|\cdot\|_Q$ norm and eigenvalue bounds)**.** *Let $Q \in S_n^+$ and define $\|x\|_Q := \sqrt{x^\top Qx}$.*

(a) Show that $\|\cdot\|_Q$ is a norm if and only if $Q \succ 0$.

(b) Assume $Q \succ 0$. Show that

$$\|x\|_Q \le \sqrt{\lambda_{\max}(Q)}\,\|x\|_2, \qquad \|x\|_2 \le \frac{1}{\sqrt{\lambda_{\min}(Q)}}\,\|x\|_Q.$$

Answers:

(a) If $Q \succ 0$, then $x^\top Qx > 0$ for $x \ne 0$ so definiteness holds; homogeneity and triangle inequality follow since $\|x\|_Q = \|Q^{1/2}x\|_2$. If Q is singular, any nonzero $x \in \ker(Q)$ has $\|x\|_Q = 0$ so it is only a seminorm.

(b) Use $x^\top Qx \le \lambda_{\max}(Q)\|x\|_2^2$ and $x^\top Qx \ge \lambda_{\min}(Q)\|x\|_2^2$ (Rayleigh quotient bounds).

5. Dual norms and operator norms

Exercise 13 (Dual norm basics)**.** *Let $\|\cdot\|$ be a norm and define its dual norm by $\|y\|_\star := \sup_{\|x\|\le 1} y^\top x$.*

(a) Prove the generalized Cauchy–Schwarz inequality: $|y^\top x| \le \|y\|_\star\,\|x\|$.

(b) Compute the duals of $\|\cdot\|_2$, $\|\cdot\|_1$, and $\|\cdot\|_\infty$.

Answers:

(a) If $x = 0$ it is trivial. Otherwise set $u = x/\|x\|$ so $\|u\| \le 1$ and $y^\top x = \|x\| \, y^\top u \le \|x\| \sup_{\|v\|\le 1} y^\top v = \|x\| \, \|y\|_\star$.

(b) 2-norm is self-dual. For $\|\cdot\|_1$, the maximizer over $\|x\|_1 \le 1$ concentrates mass on the largest coordinate of y, giving $\|y\|_\star = \|y\|_\infty$. For $\|\cdot\|_\infty$, the maximizer over $\|x\|_\infty \le 1$ takes $x_i = \operatorname{sign}(y_i)$, giving $\|y\|_\star = \|y\|_1$.

Exercise 14 (Induced operator norm)**.** *Given a vector norm $\|\cdot\|$, define*

$$\|A\|_{op} := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

(a) Show that $\|A\|_{op} = \sup_{\|x\|\le 1} \|Ax\|$.

(b) Prove submultiplicativity: $\|AB\|_{op} \le \|A\|_{op} \, \|B\|_{op}$.

(c) For the Euclidean norm, show that $\|A\|_2^2 = \lambda_{\max}(A^\top A)$.

Answers: (a) By homogeneity, scale any nonzero x to unit norm. (b) For any $x \neq 0$, $\|ABx\| \le \|A\| \, \|Bx\| \le \|A\| \, \|B\| \, \|x\|$, then take supremum. (c) Use $\|Ax\|_2^2 = x^\top A^\top Ax$ and Rayleigh quotient: $\sup_{\|x\|_2=1} x^\top A^\top Ax = \lambda_{\max}(A^\top A)$.

6. Linear systems and factorizations

Exercise 15 (LU and triangular solves)**.** *Assume $A = LU$ with L lower triangular and U upper triangular, both invertible. Show that solving $Ax = b$ reduces to two triangular solves.*

Answers: Solve $Ly = b$ by forward substitution, then solve $Ux = y$ by backward substitution.

Exercise 16 (Cholesky and SPD)**.** *(a) Let $A \in S_n$ and assume $A = LL^\top$ for some invertible lower triangular L. Prove that $A \succ 0$.*

(b) Compute the Cholesky factorization of $A = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$.

(c) Reuse your factorization to solve $Ax = b$ for $b = (2, 1)^\top$ and for $b = (0, 1)^\top$.

Answers: (a) For $x \neq 0$, $x^\top Ax = x^\top LL^\top x = \|L^\top x\|_2^2 > 0$ since $L^\top$ invertible.
(b) Write $L = \begin{pmatrix} \ell_{11} & 0 \\ \ell_{21} & \ell_{22} \end{pmatrix}$. Matching $LL^\top$ gives $\ell_{11} = 2$, $\ell_{21} = 1$, $\ell_{22} = \sqrt{2}$.
(c) For $b = (2, 1)$: solve $Ly = b$ gives $y = (1, 0)$, then $L^\top x = y$ gives $x = (1/2, 0)$. For $b = (0, 1)$: $Ly = b$ gives $y = (0, 1/\sqrt{2})$, then $L^\top x = y$ gives $x = (-1/4, \, 1/2)$.

7. Least squares problems

Exercise 17 (Solve least squares via thin QR)**.** *Let $A \in \mathbb{R}^{m\times n}$ with $m \ge n$ and assume A has full column rank. Let $A = QR$ be a thin QR factorization with $Q \in \mathbb{R}^{m\times n}$, $Q^\top Q = I_n$, and $R \in \mathbb{R}^{n\times n}$ upper triangular (hence invertible).*

(a) Let $P := QQ^\top$. Prove the identity

$$\|Ax-b\|_2^2 = \|Rx-Q^\top b\|_2^2 + \|(I-P)b\|_2^2 \qquad \forall x \in \mathbb{R}^n.$$

(b) Deduce that the unique minimizer of $\min_x \|Ax - b\|_2^2$ satisfies

$$Rx^\star = Q^\top b,$$

and explain why this can be solved by backward substitution.

Answers:

(a) Write $Ax = QRx$ and decompose b into its orthogonal components: $b = Pb + (I - P)b$ with $Pb \in \operatorname{Im}(Q)$ and $(I - P)b \perp \operatorname{Im}(Q)$. Since $QRx \in \operatorname{Im}(Q)$, we have

$$QRx - b = (QRx - Pb) - (I - P)b,$$

and the two terms are orthogonal, so by Pythagoras:

$$\|QRx - b\|_2^2 = \|QRx - Pb\|_2^2 + \|(I - P)b\|_2^2.$$

Now use $Pb = QQ^\top b$ and the isometry on $\operatorname{Im}(Q)$: for any $u \in \mathbb{R}^n$, $\|Qu\|_2^2 = u^\top Q^\top Qu = \|u\|_2^2$. Thus

$$\|QRx-QQ^\top b\|_2^2 = \|Q(Rx-Q^\top b)\|_2^2 = \|Rx-Q^\top b\|_2^2.$$

(b) The term $\|(I-P)b\|_2^2$ does not depend on x, so minimizing $\|Ax-b\|_2^2$ is equivalent to minimizing $\|Rx-Q^\top b\|_2^2$. Since R is invertible and upper triangular, the minimizer is the unique solution of $Rx = Q^\top b$, obtained by backward substitution.

Exercise 18 (Normal equations square the conditioning). *Assume $A \in \mathbb{R}^{m\times n}$ has full column rank. Recall $\kappa_2(A) = \|A\|_2\|A^\dagger\|_2$ (where $A^\dagger = (A^\top A)^{-1}A^\top$) and $\kappa_2(A^\top A) = \|A^\top A\|_2\|(A^\top A)^{-1}\|_2$.*

(a) Show that $\|A^\top A\|_2 = \|A\|_2^2$.

(b) Show that $\|(A^\top A)^{-1}\|_2 = \|A^\dagger\|_2^2$ and deduce

$$\kappa_2(A^\top A) = \kappa_2(A)^2.$$

(c) Interpret why this suggests avoiding normal equations in finite precision.

Answers:

(a) $\|A\|_2^2 = \sup_{\|x\|_2=1}\|Ax\|_2^2 = \sup_{\|x\|_2=1} x^\top A^\top Ax = \lambda_{\max}(A^\top A) = \|A^\top A\|_2$ since $A^\top A \succeq 0$.

(b) With SVD $A = U\Sigma V^\top$ (full column rank), we have $A^\top A = V\Sigma^\top\Sigma V^\top$ with eigenvalues σ_i^2. Thus $(A^\top A)^{-1}$ has eigenvalues $1/\sigma_i^2$ so $\|(A^\top A)^{-1}\|_2 = 1/\sigma_{\min}(A)^2$. Also $A^\dagger = V\Sigma^{-1}U^\top$, hence $\|A^\dagger\|_2 = 1/\sigma_{\min}(A)$, giving $\|(A^\top A)^{-1}\|_2 = \|A^\dagger\|_2^2$. Therefore $\kappa_2(A^\top A) = \|A\|_2^2\,\|A^\dagger\|_2^2 = \kappa_2(A)^2$.

(c) Forming $A^\top A$ can drastically worsen conditioning, so rounding errors get amplified much more; QR (or SVD) typically yields a more stable route to the LS solution.

8. SVD, conditioning, and numerical stability

Exercise 19 (SVD on a simple matrix). *Let $A = \begin{pmatrix}1 & 0\\ 0 & \varepsilon\end{pmatrix}$ with $\varepsilon > 0$.*

(a) Compute the singular values of A.

(b) Compute $\kappa_2(A)$.

Answers:

(a) $A^\top A = \operatorname{diag}(1,\varepsilon^2)$ so singular values are $\sigma_1 = 1$, $\sigma_2 = \varepsilon$.

(b) $\kappa_2(A) = \sigma_1/\sigma_2 = 1/\varepsilon$.