Exercises for Section 2

Any references made are to the lecture notes, except when the numbers are preceded with an 'S', such as in "See Figure S1", which then refers to the present document. Comments and corrections to `tom.berrett@warwick.ac.uk`. Exercises marked with a † will be discussed in the seminars.

Exercise 2.1. Prove parts 1, 2, 7, 8 and 9 from Theorem 2.1.1.

Exercise 2.2. For $i = 1, 2$, let A_i be a matrix given in block form as

$$A_i = \begin{pmatrix} B_i & C_i \\ D_i & E_i \end{pmatrix},$$

where the B_i are $p \times p$, the C_i are $p \times q$, the D_i are $q \times p$ and the E_i are $q \times q$. Show that

$$A_1 A_2 = \begin{pmatrix} B_1 B_2 + C_1 D_2 & B_1 C_2 + C_1 E_2 \\ D_1 B_2 + E_1 D_2 & D_1 C_2 + E_1 E_2 \end{pmatrix}$$

Show also that

$$A_i^T = \begin{pmatrix} B_i^T & D_i^T \\ C_i^T & E_i^T \end{pmatrix}.$$

Exercise 2.3. Show that the eigenvectors and eigenvalues of the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

are $\lambda_1 = 2$, $\lambda_2 = 1$, $\mathbf{v}_1 = (1, 0)^\intercal, \mathbf{v}_2 = (0, 1)^\intercal$.

Exercise 2.4 (†). Find the eigenvectors and eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Exercise 2.5. Find the eigenvectors and eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

Exercise 2.6. Let A be a symmetric matrix. Define A^k using the eigendecomposition of A and show that:

- $A^{-1}A = I$
- $A^{1/2}A^{1/2} = A$
- $A^{-1/2}A^{-1/2} = A^{-1}$

Exercise 2.7 (†)**.** Let A be a $p \times q$ matrix with SVD $A = ULV^\intercal$. Let k be the number of non-zero singular values of A, $\mathbf{u}_i$, respectively $\mathbf{v}_j$, be the i-th column of U, respectively the j-th column of V. Show that

$$A = \sum_{i=1}^{k} l_i(A)\mathbf{u}_i\mathbf{v}_i^\intercal.$$

Exercise 2.8. Let $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in \mathbb{R}^3$ be orthonormal and non-zero. Give the SVD of the matrix $A = \mathbf{u}_1\mathbf{u}_1^\intercal - 2\mathbf{u}_2\mathbf{u}_2^\intercal - \pi\mathbf{u}_3\mathbf{u}_3^\intercal$. What is $l_1(A)$?

Exercise 2.9 (†)**.** Let X be a $n \times p$ real matrix. Show that

$$\|X\|_F^2 = \sum_{i=1}^{\min(n,p)} l_i^2(X),$$

and

$$\|X - X_k\|_F^2 = \sum_{i=k+1}^{\min(n,p)} l_i^2(X),$$

where $l_i^2(X) = (l_i(X))^2$. Both $\|\cdot\|_F$ and $l_i(\cdot)$ are defined in Section 2.2. X_k is defined in Theorem 2.2.3.

Exercise 2.10 (non examinable)**.** Let X be an $n \times p$ matrix with singular values $l_k(X)$, $k = 1, \ldots, \min(n,p)$. Show that

$$l_k(X) = \inf_{\substack{S \leq \mathbb{R}^p \\ \dim(S) = n-k+1}} \sup_{\substack{v \in S \\ \|v\|=1}} \|Xv\|_2,$$

where the infimum is over all linear subspaces S of $\mathbb{R}^p$ with dimension $n - k + 1$.

Exercise 2.11 (non examinable)**.** Use the previous exercise to prove *Weyl's inequalities*: for real matrices A and B of the same size,

$$l_{i+j-1}(A + B) \leq l_i(A) + l_j(B), \quad i, j \geq 1; i + j - 1 \leq p.$$

Exercise 2.12 (not examinable)**.** Use Weyl's inequalities to prove Theorem 2.2.3.

Exercise 2.13. Prove Lemma 2.4.3.

Exercise 2.14. Let Σ be a (non-singular) symmetric positive-definite matrix. The Mahalanobis distance between two points $\mathbf{x}$ and $\mathbf{y}$ is defined as $d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^\intercal\Sigma^{-1}(\mathbf{x} - \mathbf{y})}$. Show that $d(\mathbf{x}, \mathbf{y})$ is indeed a distance, that is that it satisfies the three defining properties of a distance. You may assume that the Euclidean distance satisfies the triangular inequality.

Exercise 2.15. Show that both $A^\mathsf{T}A$ and AA^T are symmetric positive semi-definite.

Exercise 2.16. Show that for any $p \times q$ matrix B with $B^\mathsf{T}B = I_q$, we have $\|B\|_F^2 = q$ and

$$0 \leq \sum_{j=1}^{q} (B)_{ij}^2 \leq 1, \quad \forall i = 1, \ldots, p,$$

that is, the rows of B have norm at most 1.

Exercise 2.17. Assume B is symmetric positive semi-definite. Show that $B = R^\mathsf{T}R$ for some square matrix R that is **not** necessarily symmetric, and that if B is symmetric positive definite, then R is invertible.

Solutions

Any references made are to the lecture notes, except when the numbers are preceded with an 'S', such as in "See Figure S1", which then refers to the present document.

Solution to Exercise 2.3 Since we are dealing with 2×2 matrices, the 'characteristic polynomial' $A - \lambda I = 0$ is just a quadratic in λ, so it's easy to find eigenvalues and eigenvectors. That is, we first solve $|A - \lambda I| = \mathbf{0}$, which gives us the solutions λ_1, λ_2. The eigenvector $\mathbf{v}_i$ corresponding to λ_i is found by solving the system of linear equations $(A - \lambda_i I)\mathbf{v}_i = \mathbf{0}$.

Alternatively, to **show** that $\mathbf{v}$ is an eigenvector of A, just evaluate $A\mathbf{v}$ and show that it equals $\lambda\mathbf{v}$ for the appropriate λ.

Solution to Exercise 2.4 The solution is similar to the previous exercise (but you might need to solve the system of equations to find the eigenvectors).

Solution to Exercise 2.5 The solution is similar to the previous exercise (but you might need to solve the system of equations to find the eigenvectors).

Solution to Exercise 2.6 Let $A = V\Lambda V^\intercal$ be the eigendecomposition of A, where Λ is a diagonal matrix with eigenvalues as diagonal entries and V is a $p \times p$ matrix with eigenvectors as its columns. We define $A^k = V\Lambda^k V^\intercal$.

From the definition it follows that $A^{1/2}A^{1/2} = V\Lambda^{1/2}V^\intercal V\Lambda^{1/2}V^\intercal = V\Lambda^{1/2}\Lambda^{1/2}V^\intercal = V\Lambda V^\intercal = A$, where the second equality follows from the eigenvectors being orthogonal and having unit norm (so that $V^\intercal V = I$, the identity matrix). The arguments for $A^{-1/2}A^{-1/2}$ and $A^{-1}A$ are analogous.

Solution to Exercise 2.7 Discussed in seminar.

Solution to Exercise 2.8 The SVD is given by $A = \mathbf{u}_1\mathbf{u}_1^\intercal + 2(-\mathbf{u}_2)\mathbf{u}_2^\intercal + \pi(-\mathbf{u}_3)\mathbf{u}_3^\intercal$, thus $l_1(A) = \pi$ and $U = (-\mathbf{u}_3, -\mathbf{u}_2, \mathbf{u}_1)$, $V = (\mathbf{u}_3, \mathbf{u}_2, \mathbf{u}_1)$ and $L = \operatorname{diag}(\pi, 2, 1)$, and $A = ULV^\intercal$.

Solution to Exercise 2.9 Discussed in seminar.

Solution to Exercise 2.10 Please come to office hours with attempted proof for the solution. Hint: for any two subspaces S, S' of $\mathbb{R}^p$ we have $\dim(S \cap S') = \dim(S) + \dim(S') - \dim(S \cup S') \geq \dim(S) + \dim(S') - p$.

Solution to Exercise 2.11 Please come to office hours with attempted proof for the solution.

Solution to Exercise 2.12 Without loss of generality, suppose $n \geq p$. Now assume A has rank at most k. Then $l_{k+1}(A) = 0$ and

$$l_{i+k}(X) \leq l_i(X - A) + l_{k+1}(A) = l_i(X - A).$$

Since the squared Frobenius norm of a matrix is the sum of its singular values squared (see exercises),

$$\sum_{i=k+1}^{n} l_i^2(X) \leq \sum_{i=1}^{n-k} l_i^2(X - A) \leq \|X - A\|_F^2.$$

The left-hand side is equal to $\|X - X_k\|_F^2$ (see exercises), and this finishes the proof.

Solution to Exercise 2.13 Please come to office hours with attempted proof for the solution.

Solution to Exercise 2.14 Symmetry and positivity follow immediately from Σ being SPD. For the triangle inequality, notice that $d(x, y) = |\Sigma^{-1/2}(x - y)|$, and therefore

$$\begin{aligned} d(x, y) = |\Sigma^{-1/2}(x - z + z - y)| &= |\Sigma^{-1/2}(x - z) + \Sigma^{-1/2}(z - y)| \\ &\leq |\Sigma^{-1/2}(x - z)| + |\Sigma^{-1/2}(z - y)| \\ &= d(x, z) + d(z, y), \end{aligned}$$

for any $z \in \mathbb{R}^p$.

Solution to Exercise 2.15 This is basic matrix algebra — please go back to your lecture notes from previous years.

Solution to Exercise 2.16 Please come to office hours with attempted proof for the solution.

Solution to Exercise 2.17 Use $B = E\Lambda E^{\mathsf{T}}$ by spectral decomposition, and define $R = \Lambda^{1/2}E^{\mathsf{T}}$.