Homework 2

1. *Max–min and min–max characterization of eigenvalues.* Let A be a symmetric $n \times n$ matrix, with eigendecomposition

$$A = Q\,\mathbf{diag}(\lambda)Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T.$$

The matrix Q is orthogonal ($Q^TQ = QQ^T = I$) with columns $q_1, \ldots, q_n$, and $\mathbf{diag}(\lambda)$ is the diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ on its diagonal. We assume the eigenvalues are sorted as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

We denote by $\mu_1(X) \geq \mu_2(X) \geq \cdots \geq \mu_m(X)$ the eigenvalues of X^TAX, where X is an $n \times m$ matrix. In this problem we show that

$$\begin{bmatrix} \mu_1(X) \\ \mu_2(X) \\ \vdots \\ \mu_m(X) \end{bmatrix} \preceq \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} \qquad (1)$$

for all matrices $X \in \mathbf{R}^{n\times m}$ with orthonormal columns ($X^TX = I$). The inequality is a component-wise vector inequality, *i.e.*, equivalent to the m scalar inequalities

$$\mu_1(X) \leq \lambda_1, \qquad \mu_2(X) \leq \lambda_2, \qquad \ldots, \qquad \mu_m(X) \leq \lambda_m.$$

(a) Suppose $X \in \mathbf{R}^{n\times m}$ is given and satisfies $X^TX = I$. We drop the argument X in $\mu_i(X)$, and write the eigendecomposition of the $m \times m$ matrix X^TAX as

$$X^TAX = \sum_{i=1}^{m} \mu_i v_i v_i^T.$$

The vectors $v_1, \ldots, v_m$ are orthonormal eigenvectors, and $\mu_1, \ldots, \mu_m$ are the eigenvalues, sorted as $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_m$.

Suppose $1 \leq k \leq m$. Denote by $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix}$ the matrix with the first k eigenvectors $v_1, \ldots, v_k$ as its columns. Verify the following expressions for μ_k:

$$\begin{aligned} \mu_k &= \inf_{y_1^2+\cdots+y_k^2=1} (\mu_1 y_1^2 + \cdots + \mu_k y_k^2) \\ &= \inf_{y_1^2+\cdots+y_k^2=1} y^TV^T(X^TAX)Vy \\ &= \inf_{y_1^2+\cdots+y_k^2=1} y^TV^TX^TQ\,\mathbf{diag}(\lambda)Q^TXVy \\ &= \inf_{y_1^2+\cdots+y_k^2=1} \sum_{i=1}^{n} \lambda_i (q_i^TXVy)^2, \end{aligned} \qquad (2)$$

where $y = (y_1, \ldots, y_k) \in \mathbf{R}^k$.

(b) From the last expression (2), show that $\mu_k \leq \lambda_k$.
Hint. Consider a vector $\tilde{y} \in \mathbf{R}^k$ that that satisfies $\|\tilde{y}\|_2 = 1$ and

$$q_1^T X V \tilde{y} = 0, \qquad q_2^T X V \tilde{y} = 0, \qquad \ldots, \qquad q_{k-1}^T X V \tilde{y} = 0.$$

Show that

$$\mu_k \leq \sum_{i=1}^{n} \lambda_i (q_i^T X V \tilde{y})^2 \leq \lambda_k.$$

Since the inequality (1) holds with equality for the matrix

$$X = [\ q_1 \quad q_2 \quad \cdots \quad q_m\ ], \tag{3}$$

we can conclude that the set

$$S = \{(\mu_1(X), \ldots, \mu_m(X)) \mid X \in \mathbf{R}^{n \times m},\ X^T X = I\}$$

has a *maximum element*, given by $(\lambda_1, \ldots, \lambda_m)$. Applying this result to $-A$, we see that the set S also has a *minimum element*, given by $(\lambda_{n-m+1}, \ldots, \lambda_n)$. This result is known as the Courant–Fischer min–max theorem.

As an application, it follows that the matrix X given in (3) is a solution of the (nonconvex) optimization problem

$$\begin{array}{ll} \text{maximize} & f(X) \\ \text{subject to} & X^T X = I, \end{array}$$

with variable $X \in \mathbf{R}^{n \times m}$, for the following functions:

$$\begin{aligned} f(X) &= \lambda_{\max}(X^T A X) \\ &= \mu_1(X), \\ f(X) &= \lambda_{\min}(X^T A X) \\ &= \mu_m(X), \\ f(X) &= \mathbf{tr}(X^T A X) \\ &= \mu_1(X) + \cdots + \mu_m(X), \end{aligned}$$

and, more generally, any function $f(X) = \sum_{k=1}^{m} h_k(\mu_k(X))$ where h_k is a nondecreasing function.

2. Exercise T3.1.

3. Exercise T3.18 (b).

4. Exercise A3.10.

5. Exercise T3.19 (a).

6. Exercise A3.20 (a).

7. Exercise A6.8.