CMSE 820 Homework 4

Wei-Chien Liao (liaowei2@msu.edu)

1. (a) ($\Rightarrow$) Assume A is positive semi-definite. Then $u^T Au \geq 0$ for any $u \in \mathbb{R}^n$. Let λ be an eigenvalue of A, so $Ax = \lambda x$. Then

$$x^T Ax = x^T(\lambda x) = \lambda x^T x = \lambda ||x||^2 \geq 0 \Rightarrow \lambda \geq 0. \tag{1}$$

Note that (1) holds for any eigenvalue of A. Hence, all the eigenvalues of A are non-negative.

($\Leftarrow$) Assume all eigenvalues of A are non-negative. Since A is symmetric, $A = Q^T \Lambda Q$ for some orthogonal matrix Q. Let $q_1, q_2, \cdots, q_n$ denote the columns of Q. Then $Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$ and $\{q_1, q_2, \cdots, q_n\}$ is a basis of $\mathbb{R}^n$. Let $x \in \mathbb{R}^n$. Then

$$x = \sum_{i=1}^{n} \alpha_i q_i, \alpha_i \in \mathbb{R}, \text{and } x^T = \sum_{i=1}^{n} \alpha_i q_i^T.$$

Compute $x^T Ax$ gives

$$\begin{aligned} x^T Ax &= \left(\sum_{i=1}^{n} \alpha_i q_i^T\right) A \left(\sum_{i=1}^{n} \alpha_i q_i\right) \\ &= \left(\sum_{i=1}^{n} \alpha_i q_i^T\right) \left(\sum_{i=1}^{n} \alpha_i A q_i\right) \\ &= \sum_{i=1}^{n} \alpha_i^2 q_i^T A q_i \\ &= \sum_{i=1}^{n} \alpha_i^2 q_i^T \lambda q_i = \sum_{i=1}^{n} \lambda_i \alpha_i^2 ||q_i||^2. \end{aligned}$$

Note that $\lambda_i \geq 0$, $\forall i = 1, \cdots, n$

$\Rightarrow \lambda_i \alpha_i^2 ||q_i||^2 \geq 0 \Rightarrow \sum_{i=1}^{n} \lambda_i \alpha_i^2 ||q_i||^2 \geq 0 \Rightarrow x^T Ax \geq 0.$

$\Rightarrow A$ is positive semi-definite.

$\square$

(b) ($\Rightarrow$) Assume A is positive semi-definite. Since A is symmetric, $A = Q^T \Lambda Q$ for some orthogonal matrix Q. Write

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Define

$$\sqrt{\Lambda} = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}. \end{bmatrix}$$

Note that $\lambda_i \geq 0$, $\forall i = 1, \cdots, n$ from (a), so $\sqrt{\Lambda} \in \mathbb{R}^{n \times n}$. Then

$$A = Q^T \Lambda Q = Q^T \sqrt{\Lambda}^T \sqrt{\Lambda} Q = \left(\sqrt{\Lambda} Q\right)^T \left(\sqrt{\Lambda} Q\right) = B^T B,$$

where $B := \sqrt{\Lambda} Q \in \mathbb{R}^{n \times n}$.

($\Leftarrow$) Assume $A = B^T B$ for some $B \in \mathbb{R}^{n \times n}$. Then

$$x^T A x = x^T B^T B x = (Bx)^T B x = ||Bx||^2 \geq 0,$$

so A is positive semi-definite.

$\square$

2. (a) Fix $x \in \mathbb{R}^n$. Compute $||Ax - \alpha x||^2$ gives

$$\begin{aligned}(Ax - \alpha x)^T (Ax - \alpha x) &= (x^T A - \alpha x^T)(Ax - \alpha x) \\ &= x^T A^2 x - 2\alpha x^T A x + \alpha^2 x^T x.\end{aligned}$$

Let $f(\alpha) = (x^T x)\alpha^2 - (2x^T A x)\alpha + x^T A^2 x$, then f is a quadratic function of α, and

$$\arg\min_{\alpha \in \mathbb{R}} ||Ax - \alpha x||^2 = \arg\min_{\alpha \in \mathbb{R}} f(\alpha).$$

First-order derivative of f is given by

$$f'(\alpha) = 2(x^T x)\alpha - 2x^T A x.$$

Let $f'(\alpha) = 0$, then

$$\alpha = \frac{x^T A x}{x^T x} = R_A(x).$$

$\square$

(b) Let $x \in \mathbb{R}^n$. Then

$$\begin{aligned}
\frac{\partial R_A(x)}{\partial x_j} &= \frac{\partial}{\partial x_j}\frac{x^T Ax}{x^T x} \\
&= \frac{\left(\frac{\partial}{\partial x_j} x^T Ax\right) x^T x - \left(\frac{\partial}{\partial x_j} x^T x\right) x^T Ax}{(x^T x)^2} \\
&= \frac{(2Ax)_j\, x^T x - (2x_j)\, x^T Ax}{(x^T x)^2} \\
&= \frac{2}{x^T x}\Big((Ax)_j - R_A(x)x_j\Big) = \frac{2}{x^T x}\left(Ax - R_A(x)x\right)_j .
\end{aligned}$$

Hence,

$$\nabla R_A(x) = \frac{2}{x^T x}\left(Ax - R_A(x)x\right).$$

Now let $0 \neq x_0 \in \mathbb{R}^n$, then

$$\begin{aligned}
&x_0 \text{ is a critical point of } R_A(x) \\
\Leftrightarrow &\nabla R_A(x_0) = 0 \\
\Leftrightarrow &\frac{2}{x_0^T x_0}\left(Ax_0 - R_A(x_0)x_0\right) = 0 \\
\Leftrightarrow &Ax_0 - R_A(x_0)x_0 = 0 \\
\Leftrightarrow &Ax_0 = R_A(x_0)x_0 \\
\Leftrightarrow &x_0 \text{ is an eigenvector of } A.
\end{aligned}$$

□

(c) Let $x_1, \cdots, x_{n-1}$ be the remaining $n-1$ eigenvectors of A with unit length, and
$Ax_i = \lambda_i x_i$, $i = 0, 1, \cdots, n-1$. Let $x \in \mathbb{R}^n$. Then $x = \sum_{i=0}^{n-1} a_i x_i$.

$$x - x_0 = (a_0 - 1)x_0 + \sum_{i=1}^{n-1} a_i x_i,$$

so,

$$||x - x_0||^2 = (a_0 - 1)^2 + \sum_{i=1}^{n-1} a_i^2.$$

It suffice to show that if $\left|\frac{a_i}{a_0}\right| < \varepsilon$ for all $i \neq 0$, $R_A(x) - \lambda_0 = O(\varepsilon^2)$.

$$\begin{aligned}
R_A(x) - \lambda_0 &= \frac{\sum_{i=0}^{n-1} \lambda_i a_i^2}{\sum_{i=0}^{n-1} a_i^2} - \lambda_0 \\
&= \frac{\sum_{i=0}^{n-1} (\lambda_i - \lambda_0) a_i^2}{\sum_{i=0}^{n-1} a_i^2} \\
&= \frac{\sum_{i=1}^{n-1} (\lambda_i - \lambda_0) a_i^2}{\sum_{i=0}^{n-1} a_i^2} \\
&\leq \frac{\sum_{i=1}^{n-1} \lambda_i a_i^2}{\sum_{i=0}^{n-1} a_i^2} \\
&\leq \frac{\sum_{i=1}^{n-1} \lambda_i a_i^2}{a_0^2} = \sum_{i=1}^{n-1} \lambda_i \left(\frac{a_i}{a_0}\right)^2 < \left(\sum_{i=1}^{n-1} \lambda_i\right) \varepsilon^2 = M\varepsilon^2,
\end{aligned}$$

where $M = \sum_{i=1}^{n-1} \lambda_i$ is a constant. Therefore, $R_A(x) - \lambda_0 = O(\varepsilon^2)$ and the proof is complete. □

3. (a) A is diagonal, so the diagonal entries are exactly the eigenvalues of A. May assume

$$A = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \Lambda,$$

since the diagonal entries of A can be reordered by apply some permutation matrix P and P is orthogonal. Then in this case, one can take $Q = \begin{bmatrix} e_1 & e_2 & \cdots e_n \end{bmatrix}$ so that $Q^T A Q = \Lambda$, where the column vector e_i has value 1 in the i-th entry and zero otherwise.

First consider the case

$$\max_{dim(V)=k} \min_{v \in V: ||v||=1} v^T A v.$$

Let $V_k = span\{e_1, \cdots, e_k\}$. Let $v \in V_k$ with $||v|| = 1$, then $v = \sum_{i=1}^k \alpha_i e_i$, $v^T = \sum_{i=1}^k \alpha_i e_i^T$, $\sum_{i=1}^k \alpha_i^2 = 1$, and

$$v^T A v = \left(\sum_{i=1}^k \alpha_i e_i^T\right)\left(\sum_{i=1}^k \alpha_i e_i\right) = \sum_{i=1}^k \alpha_i^2 e_i^T \lambda_i e_i = \sum_{i=1}^k \alpha_i^2 \lambda_i \geq \lambda_k.$$

Therefore,

$$\max_{dim(V)=k} \min_{v \in V: ||v||=1} v^T A v \geq \min_{v \in V_k: ||v||=1} v^T A v = \lambda_k.$$

On the other hand, let V be any k dimensional subspace.
Observe that $V \cap span\{e_k, \cdots, e_n\} \neq \phi$, since $\dim(V) = k$ and $\dim(span\{e_k, \cdots, e_n\}) = n - k + 1$.
If $v^* \in V \cap span\{e_k, \cdots, e_n\}$ and $||v^*|| = 1$, then $v^* = \sum_{i=k}^{n} \beta_i e_i$ and $\sum_{i=k}^{n} \beta_i^2 = 1$. Therefore,

$$\min_{v \in V:||v||=1} v^T A v \leq v^T A v = \sum_{i=k}^{n} \beta_i^2 e_i^T \lambda_i e_i = \sum_{i=k}^{n} \beta_i^2 \lambda_i \leq \lambda_k$$
$$\Rightarrow \max_{dim(V)=k} \min_{v \in V:||v||=1} v^T A v \leq \lambda_k.$$

Now consider the case

$$\min_{dim(V)=n-k+1} \max_{v \in V:||v||=1} v^T A v.$$

Let $V_{n-k+1} = span\{e_k, \cdots, e_k\}$. Then by similar arguments in the first case,

$$\min_{dim(V)=n-k+1} \max_{v \in V:||v||=1} v^T A v \leq \max_{v \in V_{n-k+1}:||v||=1} v^T A v = \lambda_k.$$

Now let V be any $n - k + 1$ dimensional subspace.
Observe that $V \cap span\{e_1, \cdots, e_k\} \neq \phi$.
If $v^* \in V \cap span\{e_1, \cdots, e_k\}$ and $||v^*|| = 1$, then $v^* = \sum_{i=1}^{k} \beta_i e_i$ and $\sum_{i=1}^{k} \beta_i^2 = 1$. Therefore,

$$\max_{v \in V:||v||=1} v^T A v \geq v^T A v = \sum_{i=1}^{k} \beta_i^2 \lambda_i \geq \lambda_k$$
$$\Rightarrow \min_{dim(V)=n-k+1} \max_{v \in V:||v||=1} v^T A v \geq \lambda_k.$$

□

(b) A is symmetric, so $A = Q^T \Lambda Q$ for some orthogonal matrix Q, where

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Then

$$\begin{aligned} \max_{dim(V)=k} \min_{v \in V:||v||=1} v^T A v &= \max_{dim(V)=k} \min_{v \in V:||v||=1} v^T Q^T \Lambda Q v \\ &= \max_{dim(V)=k} \min_{v \in V:||v||=1} (Qv)^T \Lambda Q v \text{ (Let } u = Qv) \\ &= \max_{dim(V)=k} \min_{u \in V:||u||=1} u^T \Lambda u. \end{aligned}$$

The second equality holds since Q is orthogonal, and this reduces to the first case verified in (a). Therefore,

$$\max_{dim(V)=k} \min_{v\in V:||v||=1} v^T A v = \max_{dim(V)=k} \min_{u\in V:||u||=1} u^T \Lambda u = \lambda_k.$$

On the other hand,

$$\begin{aligned}
\min_{dim(V)=n-k+1} \max_{v\in V:||v||=1} v^T A v &= \min_{dim(V)=n-k+1} \max_{v\in V:||v||=1} v^T Q^T \Lambda Q v \\
&= \min_{dim(V)=n-k+1} \max_{v\in V:||v||=1} (Qv)^T \Lambda Q v \text{ (Let } u = Qv\text{)} \\
&= \min_{dim(V)=n-k+1} \max_{u\in V:||u||=1} u^T \Lambda u.
\end{aligned}$$

This reduces to the second case verified in (a). Therefore,

$$\min_{dim(V)=n-k+1} \max_{v\in V:||v||=1} v^T A v = \min_{dim(V)=n-k+1} \max_{u\in V:||u||=1} u^T \Lambda u = \lambda_k.$$

$\square$

4. Observe that by the minimax theorem,

$$\begin{aligned}
\lambda_k(A+B) &= \max_{dim(V)=k} \min_{v\in V:||v||=1} v^T (A+B) v \\
&= \max_{dim(V)=k} \min_{v\in V:||v||=1} (v^T A v + v^T B v) \\
&\leq \max_{dim(V)=k} \min_{v\in V:||v||=1} (v^T A v + \lambda_1(B)) \\
&= \left(\max_{dim(V)=k} \min_{v\in V:||v||=1} v^T A v \right) + \lambda_1(B) = \lambda_k(A) + \lambda_1(B),
\end{aligned}$$

and

$$\begin{aligned}
\lambda_k(A+B) &= \max_{dim(V)=k} \min_{v\in V:||v||=1} v^T (A+B) v \\
&= \max_{dim(V)=k} \min_{v\in V:||v||=1} (v^T A v + v^T B v) \\
&\geq \max_{dim(V)=k} \min_{v\in V:||v||=1} (v^T A v + \lambda_n(B)) \\
&= \left(\max_{dim(V)=k} \min_{v\in V:||v||=1} v^T A v \right) + \lambda_n(B) = \lambda_k(A) + \lambda_n(B),
\end{aligned}$$

$\square$