ECE 133B HW3

Lawrence Liu

May 3, 2023

Problem 1

(a)

We can write this as:

$$\text{trace}(X^T AX) = \sum_{i=1}^{k} x_i^T A x_i$$

Where x_i is the ith column of X. We also have that $x_i^T x_j = \delta_{ij}$, so we have that $x_i, ..., x_k$ and the eigenvector corresponding to the kth largest eigenvalues of A. Thus, we have that the maximum of

$$\text{trace}(X^T AX) \geq \sum_{i=1}^{k} \lambda_i$$

Where λ_i is the ith largest eigenvalue of A.

(b)

From the same logic we have that $x_i, ..., x_k$ are the eigenvectors corresponding to the k smallest eigenvalues of A. Thus, we have that the minimum of

$$\text{trace}(X^T AX) \leq \sum_{i=1}^{k} \lambda_{n+k+1-i}$$

Where λ_i is the ith largest eigenvalue of A.

(c)

Let us start from the case where $k = 1$, then we have that we want to maximize $x^T Ax$ subject to $x^T x = 1$. This is maximized when x is the eigenvalue corresponding to the largest eigenvalue of A, and the resulting $x^T Ax = \lambda_1$. Now let us extend this to the case of $k = 2$, then we have that

$$\det\big(X^T AX\big) = \lambda_1(x_2^T Ax_2) - (x_2^T Ax_1)$$

Therefore we can see that we can simultaneously maximize $x_2^T Ax_2$ and minimize $x_2^T Ax_1$ with x_2 being the eigenvector corresponding to the second largest eigenvalue of A. Thus, we have that the maximum of

$$\det\big(X^T AX\big) = \lambda_1\lambda_2$$

Where λ_1 and λ_2 are the two largest eigenvalues of A. Let us extend this to the case of k dimensions, let us assume that we have that

$$\det\big(X_{k-1}^T AX_{k-1}\big) = \lambda_1 ... \lambda_{k-1}$$

And that

$$X_{k-1} = \begin{bmatrix} v_{k-1} & v_{k-2} & ... & v_1 \end{bmatrix}$$

$$X_{k-1}^T AX_{k-1} = \begin{bmatrix} \lambda_{k-1} & 0 & ... & 0 \\ 0 & \lambda_{k-2} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_1 \end{bmatrix}$$

Where v_i is the eigenvector corresponding to the ith largest eigenvalue of A, λ_i. Now let us add the kth dimension, then we have that

$$X_k = \begin{bmatrix} x_k & v_{k-1} & v_{k-2} & ... & v_1 \end{bmatrix}$$

And then we will have that

$$X_k^T AX_k = \begin{bmatrix} x_k^T Ax_k & x_k^T Av_{k-1} & ... & x_k^T Av_1 \\ x_k^T Av_{k-1} & \lambda_{k-1} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_k^T Av_1 & 0 & ... & \lambda_1 \end{bmatrix}$$

Therefore from the Lebiniz formula we have that

$$\det\left(X_k^T A X_k\right) = \left(x_k^T A x_k - \sum_{i=1}^{k-1} x_k^T A v_i\right) \prod_{i=1}^{k-1} \lambda_i$$

Therefore we can see that we need to maximize $x_k^T A x_k - \sum_{i=1}^{k-1} x_k^T A v_i$, which we can do by setting x_k to be the eigenvector corresponding to the kth largest eigenvalue of A. Thus, we have that

$$\det\left(X_k^T A X_k\right) = \lambda_1 ... \lambda_k$$

And thus we get that the maximum of $\det\left(X^T A X\right)$ is $\prod_{i=1}^{k} \lambda_i$ with the maximum being achieved when the columns of X are the eigenvectors corresponding to the k largest eigenvalues of A.

(d)

We have that

$$\begin{aligned}
||X^T A X||_F &= \sqrt{\text{trace}\left((X^T A X)(X^T A X)^T\right)} \\
&= \sqrt{\text{trace}\left((X^T A X)^2\right)} \\
&= \sqrt{\text{trace}(X^T A X)^2 - 2\sum_{i<j} \lambda_i' \lambda_j'} \\
&= \sqrt{\sum_{i=1}^{k} \lambda_i^2}
\end{aligned}$$

Where λ_i' is the ith eigenvalue of $X^T A X$, since the eigenvalues of $X^T A X$ are the same as the eigenvalues of A, we have that this is maximized when the columns of X are the eigenvectors corresponding to the k largest eigenvalues of A. And the resulting value of $||X^T A X||_F$ is $\sqrt{\sum_{i=1}^{k} \lambda_i^2}$.

Problem 2

(a)

We have that

$$\begin{aligned}||A-tI||_F &= \sqrt{\text{trace}\left((A-tI)(A-tI)^T\right)} \\ &= \sqrt{\text{trace}(A^2-2tA+t^2I)} \\ &= \sqrt{\text{trace}(A^2)-2t\text{trace}(A)+t^2\text{trace}(I)} \\ &= \sqrt{\sum_{i=1}^{n}\lambda_i^2-2t\sum_{i=1}^{n}\lambda_i+nt^2}\end{aligned}$$

To maximize we take the derivative with respect to t and set it to zero

$$\begin{aligned}\frac{\partial}{\partial t}||A-tI||_F &= \frac{1}{2||A-tI||_F}\left(-2\sum_{i=1}^{n}\lambda_i+2nt\right) \\ 0 &= \frac{1}{2||A-tI||_F}\left(-2\sum_{i=1}^{n}\lambda_i+2nt\right) \\ t &= \boxed{\frac{1}{n}\sum_{i=1}^{n}\lambda_i}\end{aligned}$$

In doing so we ignored the trivial case where $||A-tI||_F = 0$, since this would imply that $A = tI$ and thus A would be $A = aI$ for some a and thus $t = a$.

(b)

We have that the eigenvalues of $A - tI$ are $\lambda_i - t$ for $i = 1, ..., n$. Therefore we have that

$$||A-tI||_2 = \max_{i=1,\ldots,n}|\lambda_i - t|$$

Therefore we have to minimize this, therefore we have that $t = \boxed{\frac{\lambda_i+\lambda_n}{2}}$ where λ_n is the smallest eigenvalue of A and λ_i is the largest eigenvalue of A.

Problem 3

We have that from Courant-Fischer

$$\lambda_{max}(A+B) = \max\left\{\frac{x^T(A+B)x}{x^Tx} : x \neq 0\right\}$$

We have that

$$\max\left\{\frac{x^T(A+B)x}{x^Tx} : x \neq 0\right\} \leq \max\left\{\frac{x^TAx}{x^Tx} : x \neq 0\right\} + \max\left\{\frac{x^TBx}{x^Tx} : x \neq 0\right\}$$

$$\lambda_{max}(A+B) \leq \lambda_{max}(A) + \lambda_{max}(B)$$

And

$$\max\left\{\frac{x^T(A+B)x}{x^Tx} : x \neq 0\right\} \geq \max\left\{\frac{x^TAx}{x^Tx} : x \neq 0\right\} + \min\left\{\frac{x^TBx}{x^Tx} : x \neq 0\right\}$$

$$\lambda_{max}(A+B) \geq \lambda_{max}(A) + \lambda_{min}(B)$$

Therefore we have

$$\lambda_{max}(A) + \lambda_{min}(B) \leq \lambda_{max}(A+B) \leq \lambda_{max}(A) + \lambda_{max}(B)$$

Problem 4

(a)

We have that:

$$\begin{aligned} B &= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} 0 & \Sigma \\ \Sigma^T & 0 \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix} \\ &= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} 0 & \Sigma V^T \\ \Sigma^T U^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & U\Sigma V^T \\ V\Sigma^T U^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \end{aligned}$$

(b)

We have that

$$\begin{aligned} B^TB &= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} \Sigma\Sigma^T & 0 \\ 0 & \Sigma^T\Sigma \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix} \\ &= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \mathrm{diag}(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2, 0, ..., 0, \sigma_1^2, \sigma_2^2, ..., \sigma_n^2) \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix} \end{aligned}$$

Therefore we have that the eigendecomposition of B is

$$B = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \mathrm{diag}(\sigma_1, \sigma_2, ..., \sigma_n, 0, ..., 0, -\sigma_1, -\sigma_2, ..., -\sigma_n) \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix}$$

(c)

The eigenvalues are $\sigma_1, \sigma_2, ..., \sigma_n$ and $-\sigma_1, -\sigma_2, ..., -\sigma_n$.