Recitation 2

DS-GA 1013 Mathematical Tools for Data Science

1. Let X be a random vector taking values in $\mathbb{R}^n$ with mean $\vec{\mu} \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. If $A \in \mathbb{R}^{m \times n}$ and $\vec{b} \in \mathbb{R}^m$ what are the mean and covariance matrix of $AX + \vec{b}$?

> **Solution:** The mean is $A\vec{\mu} + \vec{b}$ and the covariance is given by $A\Sigma A^T$.
>
> To see the mean, note that
>
> $$E[A_{i,:}X] = E[\sum_{k=1}^{n} A_{ik}X[k]] = \sum_{k=1}^{n} A_{ik}E[X[k]] = A_{i,:}E[X],$$
>
> by the linearity of expectation. Applying this to every row shows $E[AX] = AE[X]$.
>
> To see the covariance, recall that $\mathrm{Cov}(X) = E[(X - \vec{\mu})(X - \vec{\mu})^T]$. Thus we have
>
> $$\begin{aligned} \mathrm{Cov}(AX + \vec{b}) &= E[(AX + \vec{b} - (A\vec{\mu} + \vec{b}))(AX + \vec{b} - (A\vec{\mu} + \vec{b}))^T] \\ &= E[(A(X - \vec{\mu}))(A(X - \vec{\mu}))^T] \\ &= E[A(X - \vec{\mu})(X - \vec{\mu})^T A^T] \\ &= AE[(X - \vec{\mu})(X - \vec{\mu})^T]A^T \\ &= A\Sigma A^T, \end{aligned}$$
>
> by linearity of expectation twice (for A on the left and A^T on the right).

2. Prove or disprove the following statements:

   1. There exists a matrix $A \in \mathbb{R}^{n \times n}$ with no eigenvalues.
   2. The sum of two eigenvalues of a matrix A is also an eigenvalue of A.
   3. Two square matrices A and B are said to be similar, if there is some invertible matrix P such that $B = P^{-1}AP$. A and B have share eigenvalues.
   4. Covariance matrices are positive semi-definite. i.e, $x^T\Sigma x \geq 0$ for all x.
   5. Covariance matrices have non-negative eigenvalues.

> **Solution:**
>
> 1. Yes. Consider a rotation matrix R that rotates by $\pi/2$ for example. Rv can never align with v. Note that the matrix will have eigenvalues if we change to a complex field.
> 2. False.
> 3. Let v be an eigenvector of A with eigenvalue λ
>
> $$\begin{aligned} B(P^{-1}v) &= P^{-1}AP(P^{-1}v) \\ &= P^{-1}\lambda v \\ &= \lambda(P^{-1}v) \end{aligned}$$

4. $\Sigma = E(zz^T)$.
$$x^T\Sigma x = E[x^T zz^T x] = E[(x^T z)^2] \geq 0$$
5. Let u be an eigenvector.
$$0 \leq u^T\Sigma u = \lambda\|u\|^2$$
Therefore, $\lambda \geq 0$

3. Suppose $D \in \mathbb{R}^{n\times n}$ is diagonal. Give a vector v with $\|v\|_2 = 1$ such that $\|Dv\|_2$ is maximized.

Solution:
$$\|Dv\|_2^2 = \Sigma_i D_i i^2 v_i^2 \leq (\max_i D_{ii})^2 \Sigma v_i^2 = (\max_i D_{ii})^2$$
v should be e_i with i along the max entry of $|D|$

4. Suppose $A \in \mathbb{R}^{n\times n}$ be symmetric. Give a vector v with $\|v\|_2 = 1$ such that $\|Av\|_2$ is maximized.

Solution: Since A is symmetric, we can apply spectral theorem.
$$v = \Sigma\alpha_i u_i$$
$$\begin{aligned}\|Av\|_2^2 &= \|\Sigma\alpha_i\lambda_i u_i\|_2^2 \\ &= \Sigma(\lambda_i\alpha_i)^2 \\ &\leq (\max_i \lambda_i^2)\Sigma\alpha_i^2 \\ &= \max_i \lambda_i^2\end{aligned}$$

5. Let $A \in \mathbb{R}^{n\times n}$ have an eigenvalue λ Prove that
$$E_\lambda = \{v \in \mathbb{R}^n : Av = \lambda v\}$$
is a subspace of $\mathbb{R}^n$. E_λ is called an eigenspace of A corresponding to λ.

Solution:

- $A0 = 0 = \lambda 0$ So $0 \in E_\lambda$
- If $v, w \in E_\lambda$ then
$$A(v+w) = Av + Aw = \lambda v + \lambda w = \lambda(v+w)$$
showing that $(v+w) \in E_\lambda$
- For $v \in E_\lambda$ and $c \in \mathbb{R}$
$$A(cv) = c(Av) = c(\lambda v) = \lambda(cv)$$
showing that $cv \in E_\lambda$

Below, for a matrix $A \in \mathbb{R}^{n\times n}$ known to have real eigenvalues, let $\lambda_1(A) \geq \lambda_2(A) \geq \cdots$ denote the eigenvalues of A.

6. (Courant-Fischer Min-Max Theorem) Let $A \in \mathbb{R}^{n\times n}$ be symmetric.

    1. Let S be a subspace of $\mathbb{R}^n$ with $\dim(S) = k > 0$. Prove that

    $$\min_{\substack{\vec{x}\in S \\ \|\vec{x}\|=1}} \vec{x}^T A\vec{x} \leq \lambda_k(A).$$

    2. Prove that the bound $\lambda_k(A)$ is achievable if we choose the correct subspace S of dimension $k \geq 1$. That is, prove that

    $$\max_{S:\dim(S)=k} \min_{\substack{\vec{x}\in S \\ \|\vec{x}\|=1}} \vec{x}^T A\vec{x} = \lambda_k(A).$$

Solution:

Solution.

1. Let T denote the subspace of $\mathbb{R}^n$ spanned by $\vec{v}_k, \dots, \vec{v}_n$, the eigenvectors of A corresponding to $\lambda_k, \dots, \lambda_n$. Then $\dim(T) + \dim(S) = n + 1 > n$ so there is unit length vector $\vec{w} \in T \cap S$. Writing

$$\vec{w} = \alpha_k \vec{w}_k + \cdots + \alpha_n \vec{w}_n$$

for some $\alpha_k, \dots, \alpha_n \in \mathbb{R}$ we have

$$\vec{w}^T A\vec{w} = \sum_{i=k}^{n} \lambda_k \alpha_k^2 \leq \lambda_k \sum_{i=k}^{n} \alpha_k^2 = \lambda_k.$$

2. Choose S to be the span of $\vec{v}_1, \dots, \vec{v}_k$, the eigenvectors corresponding to $\lambda_1, \dots, \lambda_k$. Let unit length vector $\vec{w} \in S$ be given as

$$\vec{w} = \alpha_1 \vec{v}_1 + \cdots + \alpha_k \vec{v}_k$$

for some $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Then

$$\vec{w}^T A\vec{w} = \sum_{i=1}^{k} \lambda_i \alpha_i^2 \geq \lambda_k \sum_{i=1}^{k} \alpha_i^2 = \lambda_k.$$

7. (Weyl's Inequalities) Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric.

    1. For any j, prove that $\lambda_j(A + B) \leq \lambda_j(A) + \lambda_1(B)$.
    2. For any j, prove that $\lambda_j(A + B) \geq \lambda_j(A) + \lambda_n(B)$.
    3. For any j, prove that $|\lambda_j(A + B) - \lambda_j(A)| \leq \|B\|$.

Solution:

Solution. Below let $_1, \dots, _n$, $\vec{v}_1, \dots, \vec{v}_n$, and $\vec{w}_1, \dots, \vec{w}_n$ denote the ordered eigenvectors for A, B, and $A + B$ respectively.

1. Choose a unit vector $\vec{x}$ in

$$\mathrm{Span}(_j, \ldots, _n) \cap \mathrm{Span}(\vec{w}_1, \ldots, \vec{w}_j)$$

(possible since the dimensions add up to $n+1$). Then we have

$$\lambda_j(A+B) \leq \vec{x}^T(A+B)\vec{x} = \vec{x}^T A\vec{x} + \vec{x}^T B\vec{x} \leq \lambda_j(A) + \lambda_1(B).$$

2. First note that

$$\lambda_j(-A) = -\lambda_{n-j+1}(A)$$

since the order of the eigenvalues is reversed. By the previous part (applied to $-A$ and $-B$) we have

$$\lambda_{n-j+1}(-A-B) \leq \lambda_{n-j+1}(-A) + \lambda_n(-B).$$

Applying the note above we obtain

$$\lambda_j(A+B) \geq \lambda_j(A) + \lambda_1(B).$$

3. This follows from the previous two parts noting that $|\lambda_j(B)| \leq \|B\|$ for all j.

8. (Simple Variant of Davis-Kahan) Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric. Let λ be an eigenvalue of A with unit length eigenvector $\vec{v}$, and let μ be an eigenvalue of B with unit length eigenvector $\vec{w}$. Prove that

$$|\vec{v}^T\vec{w}||\lambda - \mu| \leq \|A - B\|.$$

Solution:

Solution.

$$\begin{aligned}
(\vec{v}^T\vec{w})(\lambda-\mu) &= (\lambda\vec{v})^T\vec{w} - \vec{v}^T(\mu\vec{w}) \\
&= (A\vec{v})^T\vec{w} - \vec{v}^T(B\vec{w}) \\
&= \vec{v}^T A\vec{w} - \vec{v}^T B\vec{w} \\
&= \vec{v}^T(A-B)\vec{w},
\end{aligned}.$$

Taking absolute values and applying Cauchy-Schwarz we see this is bounded above by $\|A - B\|$.