Solutions for Test 3: Computational Linear Algebra

Problem 1 (points: 5)

Case 1: $\mathbf{A}$ is not invertible

$\mathbf{A}$ is not invertible $\Rightarrow$ det($\mathbf{A}$) = 0 $\Rightarrow$ rank($\mathbf{A}$) < n. Also we have, rank($\mathbf{AB}$) $\leq$ rank($\mathbf{A}$) < n $\Rightarrow$ $\mathbf{AB}$ is not invertible $\Rightarrow$ det($\mathbf{AB}$) = 0 = det($\mathbf{A}$)det($\mathbf{B}$) ($\because$ det($\mathbf{A}$) = 0).

Case 2: $\mathbf{A}$ is invertible

Define $g : (\mathbb{R}^n)^n \to \mathbb{R}$, $g(\mathbf{B}) = g(\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n) = \frac{\text{det}(\mathbf{AB})}{\text{det}(\mathbf{A})}$, where $\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n$ are the columns of $\mathbf{B}$. Note that $\mathbf{AB} = [\mathbf{A}\boldsymbol{b}_1\ \mathbf{A}\boldsymbol{b}_2\ \ ...\ \ \mathbf{A}\boldsymbol{b}_n]$. We verify that g satisfies the three properties.

(1) $g(\mathbf{I}) = g(\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n) = \frac{\text{det}(\mathbf{AI})}{\text{det}(\mathbf{A})} = 1.$

(2) For any $\boldsymbol{b}_1, ..., \boldsymbol{b}_n, \boldsymbol{c}_1 \in \mathbb{R}^n$ we have,

$$\begin{aligned} g(\boldsymbol{b}_1 + \boldsymbol{c}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n) &= \frac{\text{det}(\mathbf{A}(\boldsymbol{b}_1 + \boldsymbol{c}_1), \mathbf{A}\boldsymbol{b}_2, ..., \mathbf{A}\boldsymbol{b}_n)}{\text{det}(\mathbf{A})} \\ &= \frac{1}{\text{det}(\mathbf{A})}[\text{det}(\mathbf{A}\boldsymbol{b}_1, \mathbf{A}\boldsymbol{b}_2, ..., \mathbf{A}\boldsymbol{b}_n) + \text{det}(\mathbf{A}\boldsymbol{c}_1, \mathbf{A}\boldsymbol{b}_2, ..., \mathbf{A}\boldsymbol{b}_n)](\because \text{ det is linear}) \\ &= g(\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n) + g(\boldsymbol{c}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n). \end{aligned}$$

Also for any $\alpha \in \mathbb{R}$,

$$\begin{aligned} g(\alpha\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n) &= \frac{1}{\text{det}(\mathbf{A})}\text{det}(\alpha\mathbf{A}\boldsymbol{b}_1, \mathbf{A}\boldsymbol{b}_2, ..., \mathbf{A}\boldsymbol{b}_n) \\ &= \alpha\frac{1}{\text{det}(\mathbf{A})}\text{det}(\mathbf{A}\boldsymbol{b}_1, \mathbf{A}\boldsymbol{b}_2, ..., \mathbf{A}\boldsymbol{b}_n)(\because \text{ det is linear in the first argument}) \\ &= \alpha\ g(\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n). \end{aligned}$$

$\Rightarrow$ g is linear in its first argument. Similarly g is linear in each of its argument.

(3) Suppose $\widehat{\mathbf{B}} = [...\boldsymbol{b}_j, .., \boldsymbol{b}_i, ..]$ is obtained by switching the ith and jth columns of $\mathbf{B}$. Then,

$$\begin{aligned} g(\widehat{\mathbf{B}}) = g(...\boldsymbol{b}_j, .., \boldsymbol{b}_i, ..) &= \frac{1}{\text{det}(\mathbf{A})}\text{det}(..., \mathbf{A}\boldsymbol{b}_j, .., \mathbf{A}\boldsymbol{b}_i, ...) \\ &= -\frac{1}{\text{det}(\mathbf{A})}\text{det}(..., \mathbf{A}\boldsymbol{b}_i, .., \mathbf{A}\boldsymbol{b}_j, ...)\ (\because \text{ det is alternating}) \\ &= -\frac{1}{\text{det}(\mathbf{A})}\text{det}(\mathbf{AB}) = -g(\mathbf{B}) \Rightarrow g \text{ is alternating.} \end{aligned}$$

From (1), (2) and (3), $g(\mathbf{B}) = \text{det}(\mathbf{B}), \forall \mathbf{B} \in \mathcal{M}_n \Rightarrow \text{det}(\mathbf{B}) = \text{det}(\mathbf{AB})/\text{det}(\mathbf{A})$.

Problem 1 (Alternative solution)

Case 1: $\mathbf{A}$ is not invertible

$\mathbf{A}$ is not invertible $\Rightarrow$ det($\mathbf{A}$) = 0 $\Rightarrow$ rank($\mathbf{A}$) < n. Also we have, rank($\mathbf{AB}$) $\leq$ rank($\mathbf{A}$) < n $\Rightarrow$ $\mathbf{AB}$ is not invertible $\Rightarrow$ det($\mathbf{AB}$) = 0 = det($\mathbf{A}$)det($\mathbf{B}$) ($\because$ det($\mathbf{A}$) = 0).

Case 2: $\mathbf{A}$ is invertible

Invertible matrices can be written as the product of Elementary matrices. Suppose $\mathbf{A}$ is an elementary operation of interchanging two rows. Then $\mathbf{AB}$ interchanges two rows of $\mathbf{B}$. Since the determinant function is alternating (property (3)), det($\mathbf{AB}$) = - det($\mathbf{B}$). But also det($\mathbf{A}$) = -1 (by properties (1) and (3)). So we can write, det($\mathbf{AB}$) = det($\mathbf{A}$)det($\mathbf{B}$). Now on extending this result to a general case where $\mathbf{A} = \mathbf{E}_n\mathbf{E}_{n-1}...\mathbf{E}_2\mathbf{E}_1$, where $\mathbf{E}_i$'s are elementary matrices.

$$\text{det}(\mathbf{AB}) = \text{det}(\mathbf{E}_n\mathbf{E}_{n-1}.....\mathbf{E}_2\mathbf{E}_1\mathbf{B}) = \text{det}(\mathbf{E}_n(\mathbf{E}_{n-1}.....\mathbf{E}_2\mathbf{E}_1\mathbf{B})). \tag{1}$$

Now by using the case when $\mathbf{A}$ was just a single elementary matrix, we can write (1) as,

$$\begin{aligned}\text{det}(\mathbf{AB}) &= \text{det}(\mathbf{E}_n)\text{det}(\mathbf{E}_{n-1}.....\mathbf{E}_2\mathbf{E}_1\mathbf{B})\\ &= \text{det}(\mathbf{E}_n)\text{det}(\mathbf{E}_{n-1}).....\text{det}(\mathbf{E}_2)\text{det}(\mathbf{E}_1)\text{det}(\mathbf{B})\\ &= \text{det}(\mathbf{E}_n)\text{det}(\mathbf{E}_{n-1}).....\text{det}(\mathbf{E}_3)\text{det}(\mathbf{E}_2\mathbf{E}_1)\text{det}(\mathbf{B})\\ &= \text{det}(\mathbf{E}_n\mathbf{E}_{n-1}.....\mathbf{E}_2\mathbf{E}_1)\text{det}(\mathbf{B})\\ &= \text{det}(\mathbf{A})\text{det}(\mathbf{B}).\end{aligned}$$

Problem 2 (points: 5)

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_k\}$ and $\{\boldsymbol{w}_1, \boldsymbol{w}_2, ..., \boldsymbol{w}_k\}$ be orthogonal bases of $\mathbb{V}$.

$$\begin{aligned}\varphi(\boldsymbol{v}_1, .., \boldsymbol{v}_n) &= \sum_{i=1}^{k} \boldsymbol{v}_i^\top \mathbf{A}\boldsymbol{v}_i = \sum_{i=1}^{k}\left(\sum_{j=1}^{k}\langle \boldsymbol{w}_j, \boldsymbol{v}_i\rangle \boldsymbol{w}_j\right)^\top \mathbf{A}\left(\sum_{l=1}^{k}\langle \boldsymbol{w}_l, \boldsymbol{v}_i\rangle \boldsymbol{w}_l\right)\\ &= \sum_{j=1}^{k}\sum_{l=1}^{k}\left(\sum_{i=1}^{k}\langle \boldsymbol{w}_j, \boldsymbol{v}_i\rangle\langle \boldsymbol{v}_i, \boldsymbol{w}_l\rangle\right)\boldsymbol{w}_j^\top \mathbf{A}\boldsymbol{w}_l.\end{aligned}$$

For $j = l$,
$\sum_{i=1}^{k}\langle \boldsymbol{w}_j, \boldsymbol{v}_i\rangle\langle \boldsymbol{v}_i, \boldsymbol{w}_l\rangle = \sum_{i=1}^{k}\langle \boldsymbol{w}_j, \boldsymbol{v}_i\rangle^2 = \|\boldsymbol{w}_j\|^2 = 1.$
For $j \neq l$,
$\sum_{i=1}^{k}\langle \boldsymbol{w}_j, \boldsymbol{v}_i\rangle\langle \boldsymbol{v}_i, \boldsymbol{w}_l\rangle = \boldsymbol{w}_j^\top(\sum_{i=1}^{k}\boldsymbol{v}_i\boldsymbol{v}_i^\top)\boldsymbol{w}_l = \boldsymbol{w}_j^\top \boldsymbol{w}_l = 0.$
Hence, $\varphi(\boldsymbol{v}_1, .., \boldsymbol{v}_n) = \sum_{j=1}^{k}\boldsymbol{w}_j^\top \mathbf{A}\boldsymbol{w}_j = \varphi(\boldsymbol{w}_1, .., \boldsymbol{w}_n)$.

Problem 3 (points: 10)

Define $f : \mathcal{M}_n \times \mathcal{M}_n \to \mathbb{R}$ by $f(\mathbf{X}, \mathbf{Y}) = \text{trace}(\mathbf{X}^\top\mathbf{Y})$. We will show that f is an inner product on $\mathcal{M}_n$.

(1) $f(\mathbf{X}, \mathbf{X})$ = trace($\mathbf{X}^\top\mathbf{X}$) = $\sum_{i=1}^{i=n}\boldsymbol{x}_i^\top\boldsymbol{x}_i$ = $\sum_{i=1}^{i=n}\|\boldsymbol{x}_i\|_2^2$, where $\boldsymbol{x}_1, .., \boldsymbol{x}_n$ are the columns of $\mathbf{X}$. Thus $f(\mathbf{X}, \mathbf{X}) \geq 0, \forall\mathbf{X}$ and $f(\mathbf{X}, \mathbf{X}) = 0$ if and only if $\|\boldsymbol{x}_i\|_2^2 = 0\ \ \forall i, i.e., \mathbf{X} = 0$.

(2) For any $\mathbf{X}, \mathbf{Y} \in \mathcal{M}_n$, $f(\mathbf{X}, \mathbf{Y})$ = trace($\mathbf{X}^\top\mathbf{Y}$) = $\sum_{i=1}^{i=n}\boldsymbol{x}_i^\top\boldsymbol{y}_i$ = $\sum_{i=1}^{i=n}\boldsymbol{y}_i^\top\boldsymbol{x}_i$ = trace($\mathbf{Y}^\top\mathbf{X}$) $\Rightarrow$ $f(\mathbf{X}, \mathbf{Y}) = f(\mathbf{Y}, \mathbf{X})$.

(3) For any fixed $\mathbf{A} \in \mathcal{M}_n$, $f(\mathbf{A}, \mathbf{X})$ is linear in $\mathbf{X}$, since, $f(\mathbf{A}^\top(\alpha\mathbf{X}_1 + \beta\mathbf{X}_2))$ = trace($\alpha\mathbf{A}^\top\mathbf{X}_1 + \beta\mathbf{A}^\top\mathbf{X}_2$) = αtrace($\mathbf{A}^\top\mathbf{X}_1$) + βtrace($\mathbf{A}^\top\mathbf{X}_2$) = $\alpha f(\mathbf{A}, \mathbf{X}_1) + \beta f(\mathbf{A}, \mathbf{X}_2)$.

From (1), (2) and (3) $f(\mathbf{X}, \mathbf{Y})$ = trace($\mathbf{X}^\top\mathbf{Y}$) is an inner product on $\mathcal{M}_n$. Hence $g(\mathbf{X}) := \sqrt{f(\boldsymbol{x}, \boldsymbol{x})}$ = (trace($\mathbf{X}^\top\mathbf{X}$))$^{1/2}$ is a valid norm on $\mathcal{M}_n$.
Now to prove the second part, let $\mathbf{A} \in \mathcal{M}_n$ be fixed. For all $\mathbf{X} \in \mathcal{M}_n$, we get, |trace($\mathbf{AX}$)| = |trace(($\mathbf{A}^\top)^\top\mathbf{X}$))| = $|\langle\mathbf{A}^\top, \mathbf{X}\rangle| \leq \|\mathbf{A}^\top\|\|\mathbf{X}\|$ (Cauchy Schwarz). Take $c = \|\mathbf{A}^\top\| = \sqrt{\text{trace}(\mathbf{A}^\top\mathbf{A})}$ to get the result.

Problem 4 (points: 10)

Proving the first part,

(a) Suppose there exists $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n \in \mathbb{R}^n$ such that $\mathbf{A}_{ij} = \boldsymbol{v}_i^\top \boldsymbol{v}_j$, we have to show that eigen values of the symmetric matrix $\mathbf{A}$ are non negative.

$\boldsymbol{x}^\top \mathbf{A}\boldsymbol{x} = \sum_{i,j} \mathbf{A}_{ij}\boldsymbol{x}_i\boldsymbol{x}_j = \sum_{i,j} \boldsymbol{x}_i\boldsymbol{x}_j\boldsymbol{v}_i^\top \boldsymbol{v}_j$, where $\boldsymbol{x}_i$ is the ith element of $\boldsymbol{x}$. Define $\boldsymbol{y} = \sum_i \boldsymbol{x}_i\boldsymbol{v}_i$, then, $0 \leq \boldsymbol{y}^\top \boldsymbol{y} = \sum_{ij} \boldsymbol{x}_i\boldsymbol{x}_j\boldsymbol{v}_i^\top \boldsymbol{v}_j = \boldsymbol{x}^\top \mathbf{A}\boldsymbol{x}$. Hence $\boldsymbol{x}^\top \mathbf{A}\boldsymbol{x} \geq 0 \; \forall \boldsymbol{x}$.

(b) Suppose the eigen values of the symmetric matrix $\mathbf{A}$ are non-negative, we have to show that there exists $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n \in \mathbb{R}^n$ such that $\mathbf{A}_{ij} = \boldsymbol{v}_i^\top \boldsymbol{v}_j$.

Since the matrix $\mathbf{A}$ is symmetric, by spectral decomposition, we get, $\mathbf{A} = \sum_i \lambda_i \boldsymbol{x}_i\boldsymbol{x}_i^\top$. Define, $\boldsymbol{y}_i = \sqrt{\lambda_i}\boldsymbol{x}_i$ ($\lambda_i' s$ are non negative). Then, $\mathbf{A} = \sum_i \boldsymbol{y}_i\boldsymbol{y}_i^\top$. Define $\mathbf{B}$ to be the matrix whose columns are $\boldsymbol{y}_i$, then $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$ be the rows of $\mathbf{B}$, then by definition of matrix multiplication, $\mathbf{A}_{ij} = \boldsymbol{v}_i^\top \boldsymbol{v}_j$.

Proving the second part,

Since $\mathbf{A}$ and $\mathbf{B}$ have non-negative eigenvalues, by the above result in (b), there exist $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathbb{R}^n$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$ such that $\mathbf{A}_{ij} = \boldsymbol{u}_i^\top \boldsymbol{u}_j$ and $\mathbf{B}_{ij} = \boldsymbol{v}_i^\top \boldsymbol{v}_j$.

Let $\mathbf{C}$ be defined as $\mathbf{C}_{ij} = \mathbf{A}_{ij}\mathbf{B}_{ij}$.

We claim that for all $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\boldsymbol{x}^\top \mathbf{C}\boldsymbol{x} \geqslant 0$. From this, we can conclude that the eigenvalues of $\mathbf{C}$ are non-negative. Indeed,

$$\boldsymbol{x}^\top \mathbf{C}\boldsymbol{x} = \sum_{ij} x_i x_j \mathbf{C}_{ij} = \sum_{ij} x_i x_j (\boldsymbol{u}_i^\top \boldsymbol{u}_j)(\boldsymbol{v}_i^\top \boldsymbol{v}_j) = \langle \mathbf{X}, \mathbf{X} \rangle \geqslant 0,$$

where $\mathbf{X} = \sum_i x_i \boldsymbol{u}_i \boldsymbol{v}_i^\top$ and $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{trace}(\mathbf{X}^\top \mathbf{Y})$ is the inner-product in problem 3.

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆