STAT 640: Homework 2

Due **Wednesday, February 2, 11:59pm MT** on the course Canvas webpage. Please follow the homework guidelines on the syllabus.

Name: Hannah Butler

Problem 1

Let $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ be an orthogonal matrix.

a. What values can $\det(\boldsymbol{Q})$ take?

Answer: Let's first establish a Lemma: *For a real $n \times n$ matrix $\boldsymbol{A}$,* $\det(\boldsymbol{A}) = \det)\boldsymbol{A}^T$.

Demi-Proof:

Consider an arbitrary, real, 2×2 matrix $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and its transpose $\boldsymbol{A}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. The determinant $|\boldsymbol{A}| = ad - bc$, which is equal to the determinant $|\boldsymbol{A}^T|$.

Now assume that this property, $|\boldsymbol{A}| = |\boldsymbol{A}^T|$, holds for any $n \times n$ matrix, $n \in \mathbb{N}$. It must be shown that this property also holds for an $(n+1) \times (n+1)$ matrix. To compute the determinant of an $(n+1) \times (n+1)$ matrix, we can sum the products of each element of the first row by the appropriate $n \times n$ matrices in the n remaining rows. On the other hand, considering the transpose, summing the products of the each element in the first column by the respective $n \times n$ matrices in the n remaining columns will result in the same thing, since these respective $n \times n$ matrices are transposes of the corresponding ones used in finding the determinant of the original matrix. Therefore, any square matrix has a determinant equal to it's transpose. $\square$

After establishing this property, it is easy to show what the values of $\det(\boldsymbol{Q})$ can be since $1 = \det(\boldsymbol{I}) = \det(\boldsymbol{Q}\boldsymbol{Q}^T) = \det(\boldsymbol{Q})^2$. Then the values that the determinant of an orthogonal matrix can be only ± 1.

---

b. What values can the eigenvalues of $\boldsymbol{Q}$ be?

Answer: Let λ be an eigenvalue of $\boldsymbol{Q}$. Then $\boldsymbol{Q}\boldsymbol{v} = \lambda\boldsymbol{v}$, and $\boldsymbol{v} \neq \boldsymbol{0}$. Then we have

$$\begin{aligned} \boldsymbol{v}^T\boldsymbol{Q}^T &= \lambda\boldsymbol{v}^T \\ \boldsymbol{v}^T\boldsymbol{Q}^T\boldsymbol{Q}^T\boldsymbol{v} &= \lambda^2\boldsymbol{v}^T\boldsymbol{v} \\ \boldsymbol{v}^T\boldsymbol{v} &= \boldsymbol{v}^T\boldsymbol{v}, \quad \boldsymbol{v}^T\boldsymbol{v} \text{ does not equal } \boldsymbol{0} \end{aligned}$$

That means that $\lambda^2 = 1$, so the eigenvalues of $\boldsymbol{Q}$ must be ± 1.

Problem 2

Prove Proposition 3.5: If $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is symmetric, then $\mathsf{tr}(\boldsymbol{A}^s) = \sum_{i=1}^{n} \lambda_i^s$, where λ_i are eigenvalues of $\boldsymbol{A}$. ($\boldsymbol{A}^s$ means $\boldsymbol{A}$ self-multiplied s times, for natural number s).

Answer:

$$\begin{aligned}
\mathsf{tr}(\boldsymbol{A}^s) &= \mathsf{tr}((\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T)^s) \\
&= \mathsf{tr}(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T \ldots \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T) \\
&= \mathsf{tr}(\boldsymbol{Q}\boldsymbol{\Lambda}^s\boldsymbol{Q}^T) \\
&= \mathsf{tr}(\boldsymbol{Q}\boldsymbol{Q}^T\boldsymbol{\Lambda}^s) \\
&= \mathsf{tr}(\boldsymbol{\Lambda}^s) \\
&= \sum_{i=1}^{n} \lambda_i^s
\end{aligned}$$

Problem 3

Consider the matrix $\boldsymbol{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}$. Assume $n > 3$. For **(a)** through **(c)**, describe the circumstances (i.e. a set of conditions on the values of $\{x_1, \ldots, x_n\}$) under which $\mathsf{rank}(\boldsymbol{X}) = r$ for the given values of r. Show why your answer is correct.

a. $r = 1$

Answer: $x_1 = x_2 = \cdots = x_n = a \in \mathbb{R}$. Then both the second and third columns will be scalar multiples of the first column. Namely $\boldsymbol{v}_2 = a\boldsymbol{v}_1$ and $\boldsymbol{v}_3 = a^2\boldsymbol{v}_1$. Hence, only one column will be linearly independent, and the rank of $\boldsymbol{X}$ is defined as the number of linearly independent columns.

---

b. $r = 2$

Answer: Consider the transpose of $\boldsymbol{X}$.

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \end{bmatrix}$$

It is known that $\mathsf{rank}(\boldsymbol{X}) = \mathsf{rank}(\boldsymbol{X}^T)$. We will have two linearly independent columns (and no more) if $x_i = a$ and $x_j = b, j \neq i$ for $a, b \in \mathbb{R}$. Indeed, $\mathcal{C}(\boldsymbol{X})$ will be spanned by the vectors $\{\boldsymbol{v}_1 = (1, \ldots, 1)^T, \boldsymbol{v}_2 = (a, a, \ldots, 0, \ldots, 0)^T, \boldsymbol{v}_3 = (0, 0, \ldots, b, \ldots, b)^T, \boldsymbol{v}_4 = (a^2, a^2, \ldots, 0, \ldots, 0)^T, \boldsymbol{v}_3 = (0, 0, \ldots, b^2, \ldots, b^2)^T\}$ (or some other permutation/combination of a's and 0's in $\boldsymbol{v}_2$ and b's in $\boldsymbol{v}_3$ in the positions corresponding with the positions of the 0's in $\boldsymbol{v}_2$). However, you can take the combination $\frac{1}{a}\boldsymbol{v}_2 + \frac{1}{b}\boldsymbol{v}_3$ to get $\boldsymbol{v}_1$, as well as express $v_4 = a\boldsymbol{v}_2$, and $v_5 = b\boldsymbol{v}_3$. So there are only two linearly independent vectors under these conditions.

---

c. $r = 3$

Answer: Again, considering the transpose of $\boldsymbol{X}$, we can achieve three linearly independent columns by having at least 3 distinct values among the x_i. Since the rank can not exceed 3, the values for the remaining x_i do not matter and will only add redundant information. It can be shown similarly to the case for $r = 2$, that there are only 3 vectors needed in order to express all the columns of $\boldsymbol{X}$, broken down by the value that x_i takes on.

Problem 4

This question implements the *power method* for finding eigenvectors. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$. Assume that the following are true:

- $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n > 0$.
- $\boldsymbol{x} \in \mathbb{R}^n$ is an arbitrary vector. This means that $\exists c_1, \ldots, c_n$ such that $\boldsymbol{x} = \sum_{i=1}^n c_i \boldsymbol{v}_i$.
- Assume $c_1 \neq 0$.

a. Derive an expression for $\boldsymbol{A}\boldsymbol{x}$ that involves only $c_1, \ldots, c_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, and $\lambda_1, \ldots, \lambda_n$.

Answer:

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{A} \sum_{i=1}^n c_i \boldsymbol{v}_i \\
&= \sum_{i=1}^n c_i \boldsymbol{A} \boldsymbol{v}_i \\
&= \sum_{i=1}^n c_i \lambda_i \boldsymbol{v}_i
\end{aligned}$$

---

b. For an arbitrary (whole) number $k \geq 1$, derive an expression for $\boldsymbol{A}^k \boldsymbol{x}$ that involves only $k, c_1, \ldots, c_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, and $\lambda_1, \ldots, \lambda_n$

Answer:

$$\begin{aligned}
\boldsymbol{A}^k \boldsymbol{x} &= \sum_{i=1}^n c_i \boldsymbol{A}^k \boldsymbol{v}_i \\
&= \boldsymbol{A}^{k-1} \sum_{i=1}^n c_i \boldsymbol{A} \boldsymbol{v}_i \\
&= \boldsymbol{A}^{k-1} \sum_{i=1}^n c_i \lambda_i \boldsymbol{v}_i \\
&= \quad \vdots \\
&= \sum_{i=1}^n c_i \lambda_i^k \boldsymbol{v}_i
\end{aligned}$$

---

c. Find $\lim_{k \to \infty} \left(\frac{\lambda_j}{\lambda_1} \right)^k$ for $j = 2, \ldots, n$.

Answer:

$$\lim_{k \to \infty} \left(\frac{\lambda_j}{\lambda_1} \right)^k = 0,$$

Since $\lambda_1 > \lambda_j$ for $j = 2, \ldots, n$.

---

d. Use your answers from (b) and (c) to show that $\lim_{k\to\infty} \frac{\boldsymbol{A}^k \boldsymbol{x}}{\lambda_1^k} = \alpha \boldsymbol{v}_1$ for some scalar α.

Answer:

$$\begin{aligned}
\lim_{k\to\infty} \frac{\boldsymbol{A}^k \boldsymbol{x}}{\lambda_1^k} &= \lim_{k\to\infty} \frac{c_1\lambda_1^k \boldsymbol{v}_1 + c_2\lambda_2^k \boldsymbol{v}_2 + \cdots + c_n\lambda_2^k \boldsymbol{v}_n}{\lambda_1^k} \\
&= \lim_{k\to\infty} \left(c_1 \left(\frac{\lambda_1}{\lambda_1}\right)^k + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \right) \\
&= 1
\end{aligned}$$

Problem 5

Consider the set of vectors (from Homework 1):

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \boldsymbol{v}_3 = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Let $\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{v}_4 \end{bmatrix}$. What is a basis for $\mathcal{N}(\boldsymbol{V})$ and what is the nullity of $\boldsymbol{V}$?

Answer: $\boldsymbol{v}_2 = \boldsymbol{v}_1 + \boldsymbol{v}_4$, so there are only 3 linearly independent vectors. Therefore, $\mathsf{rank}(\boldsymbol{V}) = 3$. Since the rank of $\boldsymbol{V}$ and the nullity of $\boldsymbol{V}$ add to the number of columns in $\boldsymbol{V}$, we have that the nullity must be 1. To find a basis for the null space, $\mathcal{N}(\boldsymbol{X})$, we can set up a system of equations such that $\langle \boldsymbol{v}_i, \boldsymbol{x} \rangle = 0$ for $i = 1, 2, 3, 4$:

$$\begin{aligned} 0 &= x_1 + x_2 + 2x_3 \\ 0 &= -x_1 + 2x_3 + x_4 \\ 0 &= x_1 + x_2 \\ 0 &= -x_1 + x_4 = 0 \end{aligned}$$

We have that $x_4 = x_1$ and $x_2 = -x_1$. Subbing these values into one of the first two equations we find that $x_3 = 0$.

Therefore, we can construct a basis for $\mathcal{N}(\boldsymbol{X})$ as $\{(a, -a, 0, a)^T : a \in \mathbb{R}\}$. A possible basis would be $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Problem 6

Let $\boldsymbol{X} \in \mathbb{R}^{n \times p}$ with $\mathsf{rank}(\boldsymbol{X}) > 0$, and $\mathsf{nullity}(\boldsymbol{X}) > 0$. Consider the follow spaces:

- $\mathcal{C}(\boldsymbol{X})$, $\mathcal{C}(\boldsymbol{X})^\perp$, $\mathcal{N}(\boldsymbol{X})$, $\mathcal{N}(\boldsymbol{X})^\perp$
- $\mathcal{C}(\boldsymbol{X}^\mathsf{T})$, $\mathcal{C}(\boldsymbol{X}^\mathsf{T})^\perp$, $\mathcal{N}(\boldsymbol{X}^\mathsf{T})$, $\mathcal{N}(\boldsymbol{X}^\mathsf{T})^\perp$

Not all of these spaces are distinct. Show which of these spaces are equivalent to one another.

Answer: Let $\boldsymbol{a} \in \mathcal{N}(\boldsymbol{X})$. Then $\boldsymbol{X}\boldsymbol{a} = \mathbf{0}$. Less succinctly,

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_p \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} \langle \boldsymbol{a}, \text{1st column of } \boldsymbol{X}^T \rangle \\ \vdots \\ \langle \boldsymbol{a}, n\text{th column of } \boldsymbol{X}^T \rangle \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This indicates that $\boldsymbol{a}$ is orthogonal to any vector in $\mathcal{C}(\boldsymbol{X}^T)$ for any $\boldsymbol{a} \in \mathcal{N}(\boldsymbol{X})$. Therefore, $\boldsymbol{a}$ must be in the orthogonal compliment $\mathcal{C}(\boldsymbol{X}^T)^\perp$ and $\mathcal{N}(\boldsymbol{X}) \subset \mathcal{C}(\boldsymbol{X}^T)^\perp$. These steps can be followed backwards to show that every element of $\mathcal{C}(\boldsymbol{X}^T)^\perp$ is also an element of $\mathcal{N}(\boldsymbol{X})$, hence they are equivalent.

Since $\boldsymbol{X}$ is an arbitrary matrix, let $\boldsymbol{Y} = \boldsymbol{X}^T$. Then by the established equivalence above, we have

$$\mathcal{N}(\boldsymbol{Y}) = \mathcal{C}(\boldsymbol{Y}^T)^\perp \implies \mathcal{N}(\boldsymbol{X}^T) = \mathcal{C}(\boldsymbol{X})^\perp$$

Since we have established equality between the spaces $\mathcal{N}(\boldsymbol{X})$ and $\mathcal{C}(\boldsymbol{X}^T)^\perp$, we know that any (non-zero) element of the orthogonal compliment of $\mathcal{C}(\boldsymbol{X}^T)^\perp$, $\mathcal{C}(\boldsymbol{X}^T)$, is also an element of the orthogonal compliment, $\mathcal{N}(\boldsymbol{X})^\perp$, of $\mathcal{N}(\boldsymbol{X})$ and vice-versa. So $\mathcal{C}(\boldsymbol{X}^T) = \mathcal{N}(\boldsymbol{X})^\perp$.

Using the second established equality, we can similarly establish that $\mathcal{C}(\boldsymbol{X}) = \mathcal{N}(\boldsymbol{X}^T)^\perp$.

Problem 7

Prove part 2 of Proposition 2.11. If $\boldsymbol{G}$ and $\boldsymbol{H}$ are generalized inverses of $\boldsymbol{X}^\top \boldsymbol{X}$, then $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^\top = \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^\top$.

Answer: In class we have already established the first part of Proposition 2.11: $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}$. We use this to prove the second part.

Let $\boldsymbol{y}$ be a non-zero vector in $\mathbb{R}^n$. Then $\boldsymbol{y} = \boldsymbol{v} + \boldsymbol{e}$ for $\boldsymbol{v} \in \mathcal{C}(\boldsymbol{X})$, and $\boldsymbol{e} \in \mathcal{N}(\boldsymbol{X})$. Then

$$\begin{aligned}
\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T(\boldsymbol{v}+\boldsymbol{e}) \\
&= \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{v} + \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{e} \\
&= \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{b} + \boldsymbol{0} && \boldsymbol{e} \textit{ is in } \mathcal{N}(\boldsymbol{X}), \boldsymbol{v} \text{ is a linear combination of columns in } \boldsymbol{X} \\
&= \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T\boldsymbol{e} && \textit{Part 1 of Proposition 2.11} \\
&= \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{b}+\boldsymbol{e}) \\
&= \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T(\boldsymbol{y}+\boldsymbol{e}) \\
&= \boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T\boldsymbol{y}
\end{aligned}$$

Therefore, $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T$ must be equal to $\boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T$
