Problem Set 6 due October 21, at 10 AM, on Gradescope (via Stellar)

Please list all of your sources: collaborators, written materials (other than our textbook and lecture notes) and online materials (other than Gilbert Strang's videos on OCW).

Give complete solutions, providing justifications for every step of the argument. Points will be deducted for insufficient explanation or answers that come out of the blue

Problem 1: Consider the matrix $A = \begin{bmatrix} 2 & 4 & 0 \\ 3 & 5 & 1 \\ 0 & -2 & 2 \end{bmatrix}$.

(1) Compute the projection matrices P_C and P_R onto the column and row spaces of A, respectively. *(10 points)*

(2) Compute $P_C A$ and $A P_R$ and give a geometric explanation of your answer. *(10 points)*

Solution:(1) The matrix A has rank 2, because the columns are linearly dependent:

$$\begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} = 2 \cdot \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

The first and last columns are linearly independent, so they form a basis of the column space. So let

$$A' = \begin{bmatrix} 2 & 0 \\ 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Using these we can compute $P_C = A'(A'^T A')^{-1} A'^T$.

$$(A'^T A')^{-1} = \begin{bmatrix} 13 & 3 \\ 3 & 5 \end{bmatrix}^{-1} = \frac{1}{56} \begin{bmatrix} 5 & -3 \\ -3 & 13 \end{bmatrix}$$

$$P_C = A'(A'^T A')^{-1} A'^T = \begin{bmatrix} 2 & 0 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \frac{1}{56} \begin{bmatrix} 5 & -3 \\ -3 & 13 \end{bmatrix} \begin{bmatrix} 2 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 5 & 6 & -3 \\ 6 & 10 & 2 \\ -3 & 2 & 13 \end{bmatrix}$$

Similarly, a basis of the row space is given by the first and third rows, so we should consider the matrix:

$$A'' = \begin{bmatrix} 2 & 0 \\ 4 & -2 \\ 0 & 2 \end{bmatrix}$$

and use it to compute $P_C = A''(A''^T A'')^{-1} A''^T$:

$$(A''^T A'')^{-1} = \begin{bmatrix} 20 & -8 \\ -8 & 8 \end{bmatrix}^{-1} = \frac{1}{24} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$$

$$P_R = \begin{bmatrix} 2 & 0 \\ 4 & -2 \\ 0 & 2 \end{bmatrix} \frac{1}{24} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 \\ 0 & -2 & 2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & -1 \\ 2 & -1 & 5 \end{bmatrix}$$

(2) We have:

$$P_C A = \frac{1}{14} \begin{bmatrix} 5 & 6 & -3 \\ 6 & 10 & 2 \\ -3 & 2 & 13 \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 & 0 \\ 3 & 5 & 1 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 0 \\ 3 & 5 & 1 \\ 0 & -2 & 2 \end{bmatrix}$$

$$AP_R = \begin{bmatrix} 2 & 4 & 0 \\ 3 & 5 & 1 \\ 0 & -2 & 2 \end{bmatrix} \frac{1}{6} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & -1 \\ 2 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 0 \\ 3 & 5 & 1 \\ 0 & -2 & 2 \end{bmatrix}$$

so we conclude that $P_C A = AP_R = A$. Geometrically, the formula $P_C A = A$ simply means that projecting any vector $A\boldsymbol{v}$ onto the column space of A does not change this vector (which makes sense, because any vector $A\boldsymbol{v}$ is already in the column space). The formula $AP_R = A$ is proved analogously, by transposition (the row space of A is the same as the column space of A^T).

Grading Rubric:

- Correct basis for column space *(2 points)*
- Correct choice for A' for the formula of P_C *(1 point)*
- Correct P_C *(2 points)*
- Correct argument that $P_C = P_R$ or in same distribution of points as for P_C *(5 points)*
- Correct computation $P_C A = A$ *(2 points)*
- Correct geometric argument for the above *(3 points)*
- Correct argument reducing $AP_R = A$ to previous result or same distribution of points for the direct computation *(5 points)*

Problem 2: (1) Orthogonal matrices have the property that $Q^T Q = 1$. Prove that the product of two general orthogonal matrices Q_1 and Q_2 is an orthogonal matrix. *(10 points)*

(2) Suppose you have non-zero mutually orthogonal vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$. Prove that they must be linearly independent. *(10 points).*

Solution:(1) Let Q_1, Q_2 be orthogonal matrices. Then $(Q_1 Q_2)^T Q_1 Q_2 = Q_2^T Q_1^T Q_1 Q_2 = Q_2^T Q_2 = I$, so $Q_1 Q_2$ is an orthogonal matrix.
(2) Suppose there is a 0 linear combination, ie $\sum \lambda_i q_i = 0$. Then if we take the dot product with q_i, we get $0 = \lambda_i q_i \cdot q_i$. Then as $q_i \neq 0$ $q_i \cdot q_i \neq 0$, so $\lambda_i = 0$. Thus any 0 linear combination is trivial, thus q_i are linearly independent as required.

Grading Rubric:

- Correct proof of Part (1) *(10 points)*

- Partial credit if there is an attempt of a proof in the correct direction *(∼5 points)*
- Correct proof of Part (2) *(10 points)*
- Partial credit if there is an attempt of a proof in the correct direction, eg setting up the 0 linear combination *(∼5 points)*

Problem 3: (1) Use Gram-Schmidt to compute an orthonormal basis of $\mathbb{R}^3$ that includes the vector $\boldsymbol{q}_1 = \frac{1}{3}\begin{bmatrix}1\\2\\2\end{bmatrix}$. *(10 points)*

(2) Compute the $A = QR$ factorization of the matrix:

$$A = \begin{bmatrix}1 & 1 & 0\\1 & 2 & 1\\1 & 1 & 3\\1 & 2 & 4\end{bmatrix}$$

(where Q is orthogonal and R is square upper triangular). *(10 points)*

Solution: First we complete q_1 to a basis of $\mathbb{R}^3$, by adding $v_2 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ and $v_3 = \begin{bmatrix}0\\1\\0\end{bmatrix}$. Note that this is clearly a basis as these span the whole $\mathbb{R}^3$ which is a 3 dimensional space.

$$q_2' = v_2 - (q_1 \cdot v_2)q_1 = \frac{1}{9}\begin{bmatrix}8\\-2\\-2\end{bmatrix}$$

So $q_2 = \frac{1}{6\sqrt{2}}\begin{bmatrix}8\\-2\\-2\end{bmatrix}$ Further

$$q_3' = v_3 - (q_1 \cdot v_3)q_1 - (q_2 \cdot v_3)q_2 = \frac{1}{2}\begin{bmatrix}0\\1\\-1\end{bmatrix}$$

So $q_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\end{bmatrix}$ and q_1, q_2 and q_3 form an orthonormal basis including q_1.

(2) We apply Gram-Schmidt to the columns v_1, v_2 and v_3 of this matrix A.

$q_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$.

$$q_2' = v_2 - (q_1 \cdot v_2)q_1 = \frac{1}{2}\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix} = q_2$$

Continuing we compute

$$q_3' = v_3 - (q_1 \cdot v_3 q_1) - (q_2 \cdot v_3) q_2 = \frac{3}{2} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

Thus $q_3 = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

These linear combinations give the QR factorization

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

Grading Rubric:

- Correct set up of Gram-Schmidt *(2 points)*
- Correct computation of second basis vector *(4 points)*
- Minor mistakes in computation of second basis vector *(2 points)*
- Correct computation of third basis vector *(4 points)*
- Minor mistakes in computation of third basis vector *(2 points)*
- Correct computation of orthonormal set *(5 points)*
- Minor mistakes in computation of orthonormal set *(3 points)*
- Correct computation of factorization $A = QR$ *(5 points)*
- Minor mistakes in computation of factorization $A = QR$ *(3 points)*

Problem 4: Consider a length 1 vector $\boldsymbol{a} \in \mathbb{R}^n$ (so $||\boldsymbol{a}|| = 1$), and look at the linear transformation:

$$\phi : \mathbb{R}^n \to \mathbb{R}^n \qquad \text{corresponding to the matrix} \qquad A = I - 2\boldsymbol{a}\boldsymbol{a}^T$$

(1) Compute $\boldsymbol{a}^T\boldsymbol{a}$ and show that the matrix A is orthogonal. *(10 points)*

(2) What is the subspace of $\mathbb{R}^n$ fixed by ϕ, i.e. the subspace:

$$\{\boldsymbol{v} \in \mathbb{R}^n \text{ such that } \phi(\boldsymbol{v}) = \boldsymbol{v}\}$$

(5 points)

(3) Compute $\phi(\boldsymbol{a})$ and describe the linear transformation ϕ geometrically (i.e. say what it is called in plain English, and draw a picture in the $n = 3$ case). *(10 points)*

Solution: Note that for a vectors v and w, $v^T w = v \cdot w$. Thus $a^T a = a \cdot a = \|a\|^2 = 1$.
Thus to prove A is orthogonal

$$A^T A = (I - 2aa^T)^T (I - 2aa^T) = (I - 2aa^T)^2 = I - 4aa^T + 4aa^T aa^T = I - 4aa^T + 4aa^T = I$$

Thus this is orthogonal as required.

(2) Note that the space we want to compute is $N(A - I)$. If v is orthogonal to a then $a^T v = a \cdot v = 0$ so $Av = v$. Thus this gives us an $n - 1$ dimensional vector subspace of $N(A - I)$.
Further $A - I \neq 0$ so $N(A - I) \neq \mathbb{R}^n$, thus $N(A - I) =< a >^{\perp}$

(3) $Aa = a - 2aa^T a = a - 2a = -a$. So we can see this is the orthogonal reflection on the hyperspace perpendicular to a. For the case $n = 3$ draw a plane in 3 space and the transformation is given by a reflection on this plane

Grading Rubric:

- Correct computation of $a^T a = 1$ *(3 points)*
- Correct set up of $A^T A$ *(3 points)*
- Correct proof that $A^T A = I$ *(4 points)*
- Description that the orthogonal space to a is fixed by ϕ *(3 points)*
- Correct argument that these are all the fixed vectors *(2 points)*
- Correct computation that $\phi(a) = -a$ *(3 points)*
- Description of this as a reflection on the orthogonal hyperspace to a *(4 points)*
- Drawing in 3 dimensions desscribing a reflection on a plane *(3 points)*

Problem 5: Consider the function:

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \qquad f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - 2y + 2 \\ 3x + y - 2 \end{bmatrix}$$

(1) Explain why f is <u>not</u> a linear transformation. *(5 points)*

(2) Find a linear transformation $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ and translations $\sigma, \tau : \mathbb{R}^2 \to \mathbb{R}^2$ such that:

$$f = \sigma \circ \phi \qquad \text{and} \qquad f = \phi \circ \tau$$

(a translation is a function $g : \mathbb{R}^2 \to \mathbb{R}^2$ of the form $g(\boldsymbol{v}) = \boldsymbol{v} + \boldsymbol{a}$ for a fixed vector $\boldsymbol{a}$). *(10 points)*

Solution: (1) Note that $f(0) \neq 0$ so it can not be linear. Another argument is that $f(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x' \\ y' \end{bmatrix}) \neq f(\begin{bmatrix} x \\ y \end{bmatrix}) + f(\begin{bmatrix} x' \\ y' \end{bmatrix})$. Thirdly it can be checked that $f(\begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix}) \neq \lambda f(\begin{bmatrix} x' \\ y' \end{bmatrix})$

(2) Note

$$f(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x - 2y \\ 3x + y \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

So if we write $A = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}$ and $a = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$. We can define $\phi(v) = Av$, $\sigma(v) = v + a$ and $\tau(v) = v + A^{-1}a$. Clearly σ and τ are translations and ϕ is a linear transformation. Further $\sigma(\phi(v)) = \sigma(Av) = Av + a = f(v)$ and $\phi(\tau(v)) = \phi(v + A^{-1}a) = A(v + A^{-1}a) = Av + a$. So the only thing missing to compute is

$$A^{-1}a = \frac{1}{7} \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$

Grading Rubric:

- Correct argument that f is not linear *(5 points)*
- Find correct linear transformation *(3 points)*
- Correct translation σ *(3 points)*
- Correct translation τ *(3 points)*
- Correct checking that f is given by the composition as given *(1 point)*