

Data 5100 Assignment 6

Hrishabh
Kulkarni

classmate

Date _____
Page _____

Ques. 1:

(a) To show: $I - P \rightarrow$ Projection Matrix, if P is a proj. Matrix.

Matrix $P \in \mathbb{R}^{n \times n} \rightarrow$ Proj. Matrix if: $P = P^T$ and $P = P^2$

$$Q = I - P$$

Check if Q has both same properties as proj. matrix.

check if, $Q = Q^T$ and $Q \cdot Q = Q$

$$\begin{aligned} \text{① } Q^T &= (I - P)^T \\ &= I^T - P^T \end{aligned}$$

$$= I - P \quad (\text{since } I^T = I \text{ and } P^T = P)$$

$$\therefore Q^T = Q \quad (\text{unitary})$$

\therefore Confirming that $Q = Q^T$

So, Q is Symmetric.

$$TA^T \cdot (A^T A) A = Q$$

$$\begin{aligned} \text{② } Q^2 &= (I - P)(I - P)(A^T A) A \\ &= I - 2P + P^2 \end{aligned}$$

As $P \cdot P = P$, substituting in P^2

$$\begin{aligned} \therefore Q^2 &= I - 2P + P \\ &= I - P \end{aligned}$$

$$= Q$$

$$9 = 9$$

$$9 = 9$$

\therefore Confirming that $Q^2 = Q \cdot (A^T A) A = Q$

$$TA^T \cdot (A^T A) \cdot (A^T A) A = Q$$

$Q = I - P$ is Symmetric and satisfies $Q^2 = Q$ as well.

So, Q is a Projection Matrix.

(b) To Show: If $U \in \mathbb{R}^{n \times n}$ is Orthogonal mat., then UU^T is a projection matrix.

Mat. $U \rightarrow$ Orthogonal mat., so, $U^T U = I \rightarrow$ each col. of U is unit length.

\rightarrow To check $UU^T = (U^T)^T$

$$(UU^T)^T = U(U^T)^T$$

$$= UU^T$$

Confirming $UU^T = (U^T)^T$, so, it's Symmetric.

→ check $UUT \cdot UUT^T = UU^T$:

$$(U^T)(UUT) = U(U^T U)^T$$

$$= U \cdot I \cdot U^T$$

$$= UU^T$$

Confirming the above.

Since, UUT is symmetric and satisfies $(U^T)(UUT) = UU^T$,

It is a Projection Matrix.

(c) To show: $A \in \mathbb{R}^{n \times k}$ is full rank with $k \leq n$,
then $A(A^T A)^{-1} A^T$ → is a projection Matrix.

$$P = A(A^T A)^{-1} A^T$$

→ check if $P = P^T$ and $P^2 = P$

$$\begin{aligned} P^T &= ((A(A^T A)^{-1} A^T)^T)^T \\ &= A((A^T A)^{-1})^T A^T \\ &= A(A^T A)^{-1} A^T \\ &= P \end{aligned}$$

$$\text{So, } P^2 = P$$

$$\rightarrow P \cdot P = P$$

$$P^2 = (A(A^T A)^{-1} A^T)(A(A^T A)^{-1} A^T)$$

$$P^2 = A(A^T A)^{-1} (A^T A)(A^T A)^{-1} A^T$$

$$\text{Now as } Q = A(A^T A)^{-1} A^T$$

$$= P$$

we can say that

So, finally $A(A^T A)^{-1} A^T$ is a projection matrix.

(A) To show: When are two different incoming light spectra, \vec{p} and \vec{p}_0 , visually distinguishable from each other?

Two light spectra \vec{p} and \vec{p}_0 are visually indistinguishable if they produce the same response in all three types of cones (L, M, S).

Each spectrum produces a response vector (Lcone, Mcone, Scone) where:

$$L_{\text{cone}} = \sum_{i=1}^{20} l_i p_i, \quad M_{\text{cone}} = \sum_{i=1}^{20} m_i p_i, \quad S_{\text{cone}} = \sum_{i=1}^{20} s_i p_i$$

p_i : Power of light in the i^{th} wavelength.

l_i, m_i, s_i : special spectral response constants for L, M, S cones.

Matrix Equation: $\vec{C} \cdot \vec{p} = \vec{C} \cdot \vec{p}_0$ where \vec{C} is a matrix with rows $(l_i), (m_i), (s_i)$ and \vec{p}, \vec{p}_0 are vectors representing the spectra.

\vec{p} and \vec{p}_0 are indistinguishable if $\vec{C} \cdot (\vec{p} - \vec{p}_0) = 0$.

Example: If two spectra \vec{p} and \vec{p}_0 have the same cone response values for L, M, and S, they will appear as the same color to the human eye, even if they are different in terms of the actual wavelengths.

Example: $n=3$, $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$

Constants for cones are given by:

$$\vec{C} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ s_1 & s_2 & s_3 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 & 0.3 \\ 0.2 & 0.9 & 0.4 \\ 0.3 & 0.2 & 0.7 \end{bmatrix}$$

The cone responses for spectrum \vec{P} are now:

$$\begin{bmatrix} \text{Lcone} \\ M_{\text{cone}} \\ S_{\text{cone}} \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ s_1 & s_2 & s_3 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}$$

Assuming, $\vec{P} = \begin{bmatrix} 10 \\ 5 \\ 2 \end{bmatrix}$ and $\vec{P}_0 = \begin{bmatrix} 8 \\ 6 \\ 4 \end{bmatrix}$

$$\begin{bmatrix} \text{Lcone} \\ M_{\text{cone}} \\ S_{\text{cone}} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 & 0.3 \\ 0.2 & 0.9 & 0.4 \\ 0.3 & 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 10 \\ 5 \\ 2 \end{bmatrix}$$

$$\begin{aligned} \text{Lcone} &= 0.8 \times 10 + 0.1 \times 5 + 0.3 \times 2 \\ &= 0.2 \times 10 + 0.9 \times 5 + 0.4 \times 2 \\ &= 0.3 \times 10 + 0.2 \times 5 + 0.7 \times 2 \end{aligned}$$

$$= \begin{bmatrix} 8.1 \\ 7.3 \\ 5.4 \end{bmatrix}$$

$$\begin{bmatrix} \text{Lcone} \\ M_{\text{cone}} \\ S_{\text{cone}} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 & 0.3 \\ 0.2 & 0.9 & 0.4 \\ 0.3 & 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 8 \\ 6 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 \times 8 + 0.1 \times 6 + 0.3 \times 4 \\ 0.2 \times 8 + 0.9 \times 6 + 0.4 \times 4 \\ 0.3 \times 8 + 0.2 \times 6 + 0.7 \times 4 \end{bmatrix}$$

$$\begin{aligned} &= 6.4 + 0.6 + 1.2 \\ &= 1.6 + 5.4 + 1.6 \\ &= 2.4 + 1.2 + 2.8 \end{aligned}$$

Conclusion:
They are different, however!
Some appear to be same.

$$= \begin{bmatrix} 8.2 \\ 8.6 \\ 6.4 \end{bmatrix}$$

(B) To show; In color matching problem, can an observer always adjust the intensities of 3 primary lights to match a test light.

→ Observer must find a combination of intensities

$a_1, a_2, a_3 \rightarrow$ for 3 primary lights
with spectra $\vec{u}, \vec{v}, \vec{w}$

to match a test light spectrum \vec{p}_{test} :

$$\rightarrow \vec{p}_{\text{match}} = a_1 \vec{u} + a_2 \vec{v} + a_3 \vec{w}$$

→ This is a linear combination of primary light spectra to produce a new spectrum \vec{p}_{match} that matches the test light.

→ A match can always be found if cone response vector $\rightarrow (L, M, S)$

for primary lights $\rightarrow \vec{u}, \vec{v}, \vec{w}$
span the same space,

as the test light's cone response.

→ Primary light spectra must be linearly independent.

for them to span all possible cone responses.

If they are linearly independent, the observer can adjust a_1, a_2, a_3 to match any test light.

Example:

If primary light $\rightarrow \vec{u}, \vec{v}, \vec{w}$ do not span all cone responses, it might be impossible to match certain test lights, as there would be cone responses that cannot be reached through any combination of $\vec{u}, \vec{v}, \vec{w}$.

Example:

$$\vec{U} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \vec{V} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{W} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{P}_{\text{test}} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$$

The goal is to find scalars a_1, a_2, a_3 :

$$\vec{P}_{\text{match}} = a_1 \vec{U} + a_2 \vec{V} + a_3 \vec{W} = \vec{P}_{\text{test}}$$

Part C → python code done based on above both emphonation and understandings.

Soh. 3:

(a) Solving using Least-Square:

$$y_i = \vec{a}_i^T \vec{x} + \alpha + \beta iT + w_i \quad \text{[Formulate Model with Bias and Drift]}$$

where,

\vec{a}_i → observation vector at i^{th} measurement.

\vec{x} → vector of unknown temperature values.

α → constant bias (offset).

$\beta \cdot T$ → drift term, which is dependent on time $t = iT$.

w_i → random noise

Reformulate as a linear system: Rewrite the model in matrix form to isolate \vec{x}, α, β as:

$$y = A \vec{x} + \alpha \cdot 1 + \beta \cdot T \cdot [1, 2, \dots, m-1]^T + w$$

Now, we'll define an augmented matrix that includes \vec{x} , α , and β as unknown. Matrix equation can be written as:

$$y = [A \ 1^T \cdot [0, 1, 2, \dots, m-1]^T] \begin{bmatrix} \vec{x} \\ \alpha \\ \beta \end{bmatrix} + w$$

All augmented matrix B and vector of unknowns \vec{z} :

$$y = B\vec{z} + w$$

The Least Square Solution :

We want to find $\vec{z} = [\vec{x} \ \alpha \ \beta]^T$ that minimize the error. we use least square approach: $\vec{z} = (B^T B)^{-1} B^T y$.

Applying Conditions : For least-squares solution to work, $B^T B$ must be invertible.

This requires that:

$$\rightarrow m \geq n+2$$

$\rightarrow B$ must be full rank (rank of B should be $n+2$).

Example of Conditions met :

If $m=5$, $n=3$ to consider.

$$m \geq n+2$$

and if A is designed with linearly independent rows and columns, then B will be full rank.

- (b) Covered in the py code from my understanding.

Soln. 4:(a) To show: The weighted cost function $\Rightarrow \|D(A\vec{x} - \vec{b})\|^2$ \rightarrow Weighted least squares cost function is :

$$\sum_{i=1}^m w_i (\vec{a}_i^T \vec{x} - b_i)^2$$

Expressing the above in matrix form where,

 $\|D(A\vec{x} - \vec{b})\|^2$, where D is a diagonal matrix containing square roots of weights.

1. Cost function can be written as :

$$\sum_{i=1}^m w_i (\tau_i)^2 \quad \text{where } \tau_i = \vec{a}_i^T \vec{x} - b_i$$

2. This sum can be expressed as a norm squared :

$$= \sum_{i=1}^m (\sqrt{w_i} \tau_i)^2 = \|\sqrt{W}(A\vec{x} - \vec{b})\|^2$$

3. Let $D = \sqrt{W} = \text{diag}(\sqrt{w_1}, \sqrt{w_2}, \dots, \sqrt{w_m})$,

So, the weighted cost function becomes :

$$\|D(A\vec{x} - \vec{b})\|^2$$

where D is the diagonal matrix formed from the square roots of the weights w_i .

Example : If $A \rightarrow 3 \times 2$ matrix
 and $\vec{b} \rightarrow$ vector
 size 3

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and $w_1=2, w_2=3, w_3=1$,

Then the diagonal matrix D would be ↴

$$D = \text{diag}(\sqrt{2}, \sqrt{3}, \sqrt{1})$$

(b) To Show : if A has linearly independent columns, then $B=DA$ also has linearly independent columns.

Explanations : If A has linearly independent columns, it means that there is no non-trivial linear combination of columns of A which results in zero vector.

Now, consider matrix $B=DA$, where $D \rightarrow \text{Diagonal Matrix}$ with non-zero elements (since all $w_i > 0$).

Proof :

1. Suppose columns of B are linearly dependent. Then there exists a non-zero vector \vec{c} such that :

$$\vec{B}\vec{c} = 0 \text{ or equivalently } D\vec{A}\vec{c} = 0$$

2. Since D is diagonal with non-zero entries, multiplying by $D^{-1}0$ (it's inverse)

$$\vec{A}\vec{c} = 0$$

3. But, since A has linearly independent columns, this implies that $\vec{c} = 0$, which is a contradiction.

Thus, the columns of B are also linearly independent.

(C) Find formula for solution to weighted least square problem.

Standard Least Squares Solutions :

The soln. to least square problem $\rightarrow \vec{x} = (A^T A)^{-1} A^T \vec{b}$

Weighted Least Squares Soln :

For weighted least squares problem, the cost function to minimize is:

$$\|D(A\vec{x} - \vec{b})\|^2$$

The corresponding normal equation becomes:

$$(A^T W A) \vec{x} = A^T W \vec{b}$$

where,

$$W = \text{diag}(w_1, w_2, \dots, w_m)$$

Thus, the solution is:

$$\vec{x} = (A^T W A)^{-1} A^T W \vec{b}$$

Example:

Using the previous matrix A and vector \vec{b} , and weight matrix W , the weighted least squares solution can be computed by substituting into the formula above (I have read the example for reference) (but not written here).

Soln. 5: Problem : You are fitting a Least Squares regression model.

So, Least Squares Regression minimizes the sum of squared errors:

$$\min_{\vec{x}} \| A \vec{x} - \vec{b} \|^2$$

where, $A \rightarrow$ Feature Matrix

$\vec{x} \rightarrow$ Vector of fit coefficient

$\vec{b} \rightarrow$ Target Vector.

New Feature

$$\vec{a}_{n+1} = \frac{1}{n} \sum_{i=1}^n \vec{a}_i$$

This

Means that new feature is just average of existing features $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$.

Issues with adding \vec{a}_{n+1} :

rr Collinearity :

\vec{a}_{n+1} is a linear combination of the other features,

so that we can say it is perfectly collinear.

rr Singular Matrix :

Collinearity makes $A^T A$ singular or poor with conditions, which can lead to numerical instability.

rr Redundancy / Repeatedness :

The new feature adds no new information only redundancy, which does not improve the model.

Conclusion :

If this the answer should be No.

It is not a good idea because adding \vec{a}_{n+1} introduces collinearity, making the matrix $A^T A \rightarrow$ singular.

This leads \rightarrow numerical issues

This can reduce the performance of the model.

The new feature does not include provide any new information \rightarrow which may lead to \rightarrow redundancy in the model.

Example :

Two feature vectors in A: $\vec{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

The new feature would be:

$$\vec{a}_{n+1} = \frac{1}{2} (\vec{a}_1 + \vec{a}_2) = \frac{1}{2} \begin{bmatrix} 1+4 \\ 2+5 \\ 3+6 \end{bmatrix} = \begin{bmatrix} 2 \cdot S \\ 3 \cdot S \\ 4 \cdot S \end{bmatrix}$$

Adding this feature to Matrix A: will look like

$$A = \begin{bmatrix} 1 & 4 & 2 \cdot S \\ 2 & 5 & 3 \cdot S \\ 3 & 6 & 4 \cdot S \end{bmatrix}$$

So, we can't solve the least square problem

A has cols that are linearly dependent, as 3rd col. is aug.. So, $A^T A$ is singular.