KRISH PATEL 605 796 227

ECE 147 Homework 1

Q1) a) i) Q = real orthogonal matrix

Property $\rightarrow$ $Q = Q^T$ [If orthogonal].

Thus, Q^T is also an orthogonal matrix. Also, as $Q^T = Q^{-1}$, thus Q^{-1} is also orthogonal

ii) $Qv = \lambda v$ where λ = eigenvalue

$$\|Qv\| = \|\lambda v\|$$

$$\|Q\| \cdot \|v\| = \|\lambda\| \cdot \|v\|$$

Norm $\|\lambda\| = \|Q\| = 1$

(As columns are perpendicular to each other, each column has a norm of 1. $\|Q\| = 1$

iii) $\text{Det}(A) \cdot \text{Det}(B) = \text{Det}(AB)$ $\qquad (Q^TQ) = I$

$$\det(Q^T) \cdot \text{Det}(Q) = \det(Q^TQ) = \det(I) = 1$$

$Q^T = Q$,

therefore $(\det(Q))^2 = 1$

$$\det(Q) = \pm 1$$

iv) Need to show $\|Qx\| = \|x\|$ where x is a vector.

$$\|Qx\| = \sqrt{(Qx)_1^2 + (Qx)_2^2 + (Qx)_3^2 \cdots}$$

$$= \sqrt{x^TQ^TQx} = \sqrt{x^TIx} = \sqrt{x^Tx}$$

$$\sqrt{x^Tx} = \|x\|$$

Thus, $\|Qx\| = \sqrt{x^Tx} = \|x\|$

And therefore, is a length preserving transformation.

Scanned with CamScanner

b) i) A is a matrix (square)

$A = U\Sigma V^T$

U & V are orthogonal, Σ is diagonal matrix

[Left singular ~~values~~/vectors U are eigenvectors of AA^T.]

Eigenvalues of AA^T are the squares of the singular values in Σ

$$AA^T = U\Sigma^T V^T V\Sigma U^T = U\Sigma^T\Sigma U^T$$
$$A^TA = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

[Right singular vectors of A are the eigenvectors of A^TA.]

ii) $A = U\Sigma V^T$ Singular values $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ ....

Eigenvalues of AA^T

$$\lambda_1 = \sigma_1^2, \lambda_2 = \sigma_2^2, \lambda_3 = \sigma_3^2 \ldots$$

Eigenvalues of A^TA

$$\lambda_1 = \sigma_1^2 \quad \lambda_2 = \sigma_2^2 \quad \lambda_3 = \sigma_3^2$$

c) i) False. The identity matrix for any given n have the same eigenvalues 1.

ii) Consider the matrix $A = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$

Eigenvectors: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, Eigenvector = $[0 \quad 1]$

" " value = 2 " " value = 4

however $[1, 0] + [0, 1] = [1, 1]$ is not an eigenvector for A. Hence **FALSE**

iii) Eigenvalues of a positive semidefinite matrix are non negative
Thus TRUE

iv) TRUE → The rank of a matrix can exceed the number of distinct eigenvalues, which may be lesser than or equal to the number of distinct eigenvalues.

v) TRUE → Non zero sum of ~~as~~ two eigenvectors corresponding to the same eigenvalue λ is always an eigenvector.

(Q2) PROBABILITY REFRESHER :

1 Never hit = $\sum^{\infty} (1-p)^{n-1}(1-p)$

Q2) a) i) $P(A \text{ not hit}) = \sum_{n=1}^{\infty} P(A \text{ not hit}, n)$

If duel ends in n rounds and A is not hit, then

$\underset{1}{\underline{BM}}$ $\underset{2}{\underline{BM}}$ $\underset{3}{\underline{BM}}$ $\underset{4}{\underline{BM}}$ .... $\underset{n-1}{\underline{BM}}$. $\underset{n}{\underline{B\text{ Hits}}}$

$$P(\overset{A}{\cancel{\text{no}}} \text{ not hit}, n) = (1-P_A)^{(n-1)} (1-P_B)^{n-1} \cdot P_A (1-P_B)$$

$$= P_A (1-P_A)^{n-1} (1-P_B)^{n}$$

So Total Prob $= P_A \sum_{n=1}^{\infty} (1-P_A)^{n-1} (1-P_B)^{n}$

which is a geometric series.

Thus, $P(\overset{A}{\text{not}} \text{ hit}) = \dfrac{P_A (1-P_B)}{1-(P_A-1)(P_B-1)}$

ii) P(Both duellists hit) → In the similar manner to the above part

$$\sum_{n=1}^{\infty} (1-P_A)^{n-1} (1-P_B)^{n-1} P_A P_B$$

$$= P_A P_B \sum_{n=1}^{\infty} (1-P_A)^{n-1} (1-P_B)^{n-1}$$

$$= \frac{P_A P_B}{1-(1-P_A)(1-P_B)}$$

iii) Duel can end at n^{th} round in 3 possible ways, either A or B is hit, or both are hit in the same round.

P (duel ends n^{th} round)

$$= (1-P_A)^{n-1} (1-P_B)^{n-1} (1-P_A) P_B + (1-P_A)^{n-1} (1-P_B)^{n-1} (1-P_B) P_A$$

$$+ (1-P_A)^{n-1} (1-P_B)^{n-1} P_A P_B$$

$$= \left[(1-P_A)^{n-1} (1-P_B)^{n-1}\right] \left[1-(1-P_A)(1-P_B)\right]$$

iv) duel ends after n^{th} round, from (i)

$$\frac{P(n, A \text{ not hit})}{P(A \text{ not hit})} = \frac{(1-P_A)^{n-1}(1-P_B)^{n} \cdot P_A}{\left[\frac{P_A(1-P_B)}{1-(1-P_A)(1-P_B)}\right]}$$

$$= \Big(1-(P_A-1)(P_B-1)\Big)\Big((1-P_A)(1-P_B)\Big)^{n-1}$$

v) duel ends after n^{th} round, given both are hit

$$\frac{P(n, A \,\&\, B \text{ are hit})}{P(A \,\&\, B \text{ are hit})} = \frac{(1-P_A)^{n-1}(1-P_B)^{n-1} P_A P_B}{\left[\frac{(P_A P_B)}{1-(1-P_A)(1-P_B)}\right]}$$

$$= \Big(1-(1-P_A)(1-P_B)\Big)\Big[(1-P_A)(1-P_B)\Big]^{n-1}$$

b) ~~(b)~~ i) lets define X_i as random Bernouilli variable:

E : i^{th} member is ECE
C : i^{th} " " CSE
M : i^{th} " " Math

$$\cancel{X_i} = \begin{cases} 1 & i^{th} \text{ member is isolated} \\ 0 & i^{th} \text{ member isn't isolated} \end{cases}$$

$$X = \sum_{i=1}^{18} X_i \qquad E[X_i] = P(X_i = 1)$$

$$= P(i|E)P(E) + P(i|C)P(C) + P(i|M)P(M)$$

$$= \frac{12}{17} \times \frac{11}{16} \times \frac{1}{3} \times 3 \text{ (For all departments)}$$

$$= 45/91 \qquad E[X] = 18 \times \cancel{\tfrac{45}{91}} = \frac{33}{68} = 8.735$$

from linearity of expectation.

ii) $$y_i = \begin{cases} 1, & i^{th} \text{ faculty is semi happy} \\ 0, & i^{th} \text{ faculty isn't semi happy.} \end{cases}$$

(Q2)c) Defining events

D: man has dangerous type of disease.

T: man has a positive LSA test.

$P(T|D) = 0.9$ $\quad P(T|D^c) = 0.01$ $\quad P(D = 0.0005$

i) Using Bayes theorem:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)}$$

$$= \frac{0.9 \times 0.0005}{0.9 \times 0.0005 + 0.01 \times 0.9995} \approx 0.043$$

ii) $$P(D|T^c) = \frac{P(T^c|D)P(D)}{P(T^c)}$$

$$\frac{0.1 \times 0.0005}{0.1 \times 0.0005 + 0.99 \times 0.9995} = 0.00050528$$

d) $$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = E(Ax + b) = E(Ax) + E(b)$$

$$= AE(x) + b$$

e) $\operatorname{cov}(x) = E((x - Ex)(x - Ex)^T)$

$\operatorname{cov}(Ax + B)$

$$= E((Ax + B - E(Ax + b))(Ax + B - E(Ax + b))^T)$$

$$= E((Ax - AE(x))(Ax - AE(x))^T)$$

$$= A\,E(\{x - E[x]\}\{x - E[x]\}^T)A^T$$

$$= A \operatorname{cov}(x) A^T$$

Scanned with CamScanner

3) Multivariate derivatives:

a) $\nabla_x x^T A y = Ay$ → Using matrix cookbook property

b) $\nabla_y x^T A y = (x^T A)^T = \underline{A^T x}$

c) $\nabla_A x^T A y \rightarrow \begin{bmatrix} \frac{\partial x^T A y}{\partial a_{1,1}} & \frac{\partial x^T A y}{\partial a_{1,2}} & \frac{\partial x^T A y}{\partial a_{1,3}} & \cdots & \frac{\partial x^T A y}{\partial a_{1,m}} \\ \vdots & \vdots & & \ddots & \\ \frac{\partial x^T A y}{\partial a_{n,1}} & \frac{\partial x^T A y}{\partial a_{n,2}} & \cdots & & \frac{\partial x^T A y}{\partial a_{n,m}} \end{bmatrix} = xy^T$

d) $x \in \mathbb{R}^n$ $\quad f = x^T A x + b^T x$

$\nabla_x f = \underline{(Ax + A^T x) + b}$

e) $f = tr(AB)$ $\quad \nabla_A f = ?$

$tr(AB) =$ sum of element wise products of A and B^T.

thus, $\nabla_A f = \underline{B^T}$

f) $f = tr(BA + A^T B + A^2 B)$

$= tr(BA) + tr(A^T B) + tr(A^2 B)$

$\nabla_A f = B^T + B + 2AB$

g) $f = \|A + \lambda B\|_F^2$ Frobenius. $\quad \nabla_A f = ?$

$f = tr((A + \lambda B)^T (A + \lambda B))$

$f = tr(A^T A + \lambda A^T B + \lambda B^T A + \lambda^2 B^T B)$

$\Rightarrow \nabla_A f = 2A + \lambda B + \lambda B + 0$

$= \underline{2A + 2\lambda B}$

Scanned with CamScanner

Q4) $\min_{w} \frac{1}{2}\sum_{i=1}^{n} \left\| y^{(i)} - Wx^{(i)} \right\|^2$

Obj function $= \cancel{\min} \frac{1}{2}\sum_{k=1}^{n} (y^{(i)} - Wx^{(i)})^T (y^{(i)} - Wx^{(i)})$

Terms dependent on w

$= \frac{1}{2}\sum_{k=1}^{n} \left(-2y^{(i)T} W x^{(i)} + x^{(i)T} W^T W x^{(i)}\right)$

Taking the trace

$= \sum_{k=1}^{n} -tr(Wx^{(i)} y^{(i)T}) + \frac{1}{2} tr(Wx^{(i)} x^{(i)T} W^T)$

Vectorization→

$g = -tr(WXY^T) - \frac{1}{2} tr(WXX^T W^T)$ — ①

$\nabla_W g = -YX^T + \frac{1}{2}(WXX^T + WXX^T)$

equating to 0.

$YX^T = WXX^T$

$W = YX^T (XX^T)^{-1}$

Q5) $\frac{1}{2}\sum_{i=1}^{N} (y^{(i)} - \theta^T \hat{x}^{(i)})^2 + \frac{\lambda}{2}\|\theta\|^2$

Taking ① from the above equation

$J(\theta) = -tr(\theta XY^T) - \frac{1}{2} tr(\theta XX^T \theta^T) + \frac{\lambda}{2} tr(\theta^T \theta)$

$\nabla_\theta J(\theta) = -YX^T + \theta XX^T + \frac{2\lambda}{2}\theta = 0$

$\theta(XX^T + \lambda I) = YX^T \qquad \theta^* = YX^T (XX^T + \lambda I)^{-1}$

Scanned with CamScanner