Gautam Ahuja

Assignment-1

Introduction to Machine Learning

Date

Page 1

I] Given $A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 1 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 2 & 4 & 1 & 10 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

(a) row reduced echloen matrix of $A = R$

we may write $IA = A$

now we create a augmented matrix of 5×10 and perform ~~redu~~ row reduction to get R.

$$\Rightarrow \left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 & 4 & 1 & 10 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \downarrow R_2 \to R_2 - R_1$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 & 4 & 1 & 10 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \downarrow R_4 \to R_4 - 2R_1$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ -2 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \downarrow R_2 \to R_2(-1)$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ -2 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} R_3 \to R_3 - R_2 \\ R_4 \to R_4 - R_2 \end{array}$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad R_3 \rightleftharpoons R_4$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ -3 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad R_5 \to R_5 - R_3$$

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ -3 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore$ matrix $P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 3 & -1 & 0 & -1 & 1 \end{bmatrix}$

(b) We know $R = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

It is in reduced echlon form.

The row space W of R is the set of all vectors that the rows of R span. which have pivots 1.

$$\therefore \text{Basis}(W) = \left\{ \begin{bmatrix}1\\2\\0\\3\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix} \right\}$$

(c) vectors b_1

To check if a vector $v \in$ span$\{W\}$ we check if $Ax = v$ has a solution in rref.

~~here~~ $A = \cancel{\begin{bmatrix}1&0&0\\2&0&0\\0&1&0\\3&4&0\\0&0&1\end{bmatrix}}$ ~~Converting to rref~~ $R_2 \to R_2 - 2R_1$, $R_4 \to R_4 - 3R_1$, $R_2 \rightleftharpoons R_3$, $R_4 - 4R_2$, $R_3 \rightleftharpoons R_5$ $\rightarrow \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\\0&0&0\end{bmatrix}$

vectors, b_1 b_2 b_3 b_4 b_5 after above operations:

$b_1' = \cancel{\begin{bmatrix}1\\0\\0\\-1\\-0\end{bmatrix}}$ $b_2' = \cancel{\begin{bmatrix}2\\0\\0\\-2\\-2\end{bmatrix}}$ $b_3' = \cancel{\begin{bmatrix}0\\-1\\0\\-3\\-1\end{bmatrix}}$ $b_4' = \cancel{\begin{bmatrix}3\\-7\\23\\10\\0\end{bmatrix}}$ $b_5' = \cancel{\begin{bmatrix}0\\0\\0\\1\\1\end{bmatrix}}$

$\Rightarrow Ax = b_1$; $Ax = b_2$; $Ax = b_3$; $Ax = b_4$; $Ax = b_5$

After rref for each we see,

$\{b_1, b_2, b_3\} \in$ span W

(d) $b_1 = \begin{bmatrix}1\\2\\0\\3\\4\end{bmatrix}$ $b_2 = \begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix}$ $b_3 = \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$

basis $B = \{v_1, v_2, v_3\}$

- $b_1 = \alpha v_1 + \beta v_2 + \gamma v_3$

$$\begin{bmatrix}1\\2\\0\\3\\4\end{bmatrix} = \alpha\begin{bmatrix}1\\2\\0\\3\\4\end{bmatrix} + \beta\begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix} + \gamma\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$$

$\Rightarrow$ On solving $\alpha = 1, \beta = 0, \gamma = 0$

$\therefore [b_1]_B = \begin{bmatrix}1\\0\\0\end{bmatrix}$

- $b_2 = \alpha v_1 + \beta v_2 + \gamma v_3$

$$\begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix} = \alpha\begin{bmatrix}1\\2\\0\\3\\4\end{bmatrix} + \beta\begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix} + \gamma\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$$

On solving

$\therefore [b_2]_B = \begin{bmatrix}0\\1\\0\end{bmatrix}$

- $b_3 = \alpha v_1 + \beta v_2 + \gamma v_3$

$$\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix} = \alpha\begin{bmatrix}1\\2\\0\\3\\4\end{bmatrix} + \beta\begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix} + \gamma\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$$

On solving

$\therefore [b_3]_B = \begin{bmatrix}0\\0\\1\end{bmatrix}$

(e) let $a_1 = (1,2,0,3,0)$
$a_2 = (1,2,-1,-1,0)$
$a_3 = (0,0,1,4,0)$
$a_4 = (2,4,1,10,1)$
$a_5 = (0,0,0,0,1)$

These are the row vectors of A

to write as a linear combination,

• $b_1 = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 + a_5 v_5$

$$\begin{bmatrix}1\\2\\0\\3\\0\end{bmatrix} = a_1\begin{bmatrix}1\\2\\0\\3\\0\end{bmatrix} + a_2\begin{bmatrix}1\\2\\-1\\-1\\0\end{bmatrix} + a_3\begin{bmatrix}0\\0\\1\\4\\0\end{bmatrix} + a_4\begin{bmatrix}2\\4\\1\\10\\1\end{bmatrix} + a_5\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$$

Solving the linear eq^ns we get : $\begin{bmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\\4\end{bmatrix}$

$\therefore\ b_1 = (1)v_1 + (0)v_2 + (0)v_3 + (0)v_4 + (4)v_5$

• $b_2 = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 + a_5 v_5$

Similarly on writing, we solve the linear eq^ns of $a_1\ a_2\ a_3\ a_4\ a_5$ we get :

$$\begin{bmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$$

$b_2 = (0)v_1 + (0)v_2 + (0)v_3 + (0)v_4 + (0)v_5$

$\Rightarrow$ b_2 is linearly independent of row space of A

• $b_3 = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 + a_5 v_5$

on solving. $\begin{bmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}$

$\Rightarrow\ b_3 = (0)v_1 + (0)v_2 + (0)v_3 + (0)v_4 + (1)v_4$

• vectors b_4 and b_5 are $\vec{0}$ vectors

$\therefore$ their combination $\begin{bmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{bmatrix} = \vec{0}$

(f) The null space of A is all solutions of $A\vec{x} = 0$
for this we convert A to rref as before
to get :

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

$\therefore$ the null space has ∞ many sol^n(s)

Here we have two $\vec{0}$ rows.
This implies there are 2 free variables

(g) One basis for null space of A is :

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad x_1 = -2x_2 - 3x_4$$
$$x_3 = 0x_2 - 4x_4$$
$$x_5 = 0x_2 + 0x_4$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\therefore \text{basis} = \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ -4 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

(h) In the matrix Y, the columns which can be written as a linear combination of column ~~space~~ of A vectors (or the vectors in $y \in$ Col Space(A)) are the one which has solution of $Ax = y$

(i) We check the rref of A = $\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

the range of A are the original vectors for which the pivot is 1

$$\therefore \text{range} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

(j) As we know, the row rank of a matrix is the number of non zero rows in the rref of matrix as done in part (b), the rank(A) = 3 as there are 3 non-zero rows in rref(A). Also we know col rank(A) = row rank(A), as done in class.

$\therefore$ Col rank(A) = 3.

the nullity is the # free variables (or dimension of ker(A)) as done in (g) nullity = 2

Date ___ / ___ / ___

$\therefore$ The relationship b/w rank & nullity (also know as rank - null theorem) is

~~statement~~ $\text{rank}(A) + \dim(\ker(A)) = n$ (# of cols)

$\Rightarrow 3 + 2 = 5$

$5 = 5$, the result holds hence the relation is verified.

(II) Given.

Space of square - integrable real functions

orthonormal basis : $\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}} \right\}$ $n = 1, 2, \ldots$

(a) to show :

$f(x) = f_s(x) + f_a(x)$

We may write $f(x)$ as : $f(x) = \frac{1}{2}(f(x) + f(x)) + \frac{1}{2}(f(x) - f(x))$

we will assume that

$\frac{1}{2}(f(x) + f(x)) \in S$ and $\frac{1}{2}(f(x) - f(x)) \in A$. and prove it true

for S : • if $f(x) \in S$

$\frac{1}{2}(f(x) + f(x)) \in S$, replace x by $-x$

$\Rightarrow \frac{1}{2}(f(-x) + f(-x))$

$= \frac{1}{2}(f(x) + f(x))$ ~~$\in S$~~ — (1)

• if $f(x) \in A$

$\Rightarrow \frac{1}{2}(f(x) + f(x))$ replace x by $-x$

$\Rightarrow -\frac{1}{2}(f(x) + f(x))$ — (2)

for A : • if $f(x) \in S$

$(\frac{1}{2} f(x) - f(x)) \in A$ replace x by $-x$

$= \frac{1}{2}(f(-x) - f(-x))$

$= \frac{1}{2}(f(x) - f(x))$ — (3)

• if ~~f~~$f(x) \in A$

$= \frac{1}{2}(f(-x) - f(-x))$

$= -\frac{1}{2}(f(x) - f(x))$ — (4)

Date __/__/__

- if $f(x) \in S$ then by eq^n ① ③

$$f(x) = \frac{1}{2}(f(x)+f(x)) + \frac{1}{2}(f(x) - f(x))$$
$$= f_s(x) + f_a(x)$$ Hence Proved.
$\because f_s(x) \in S$ & $f_a(x) \in A$

- if $f(x) \in A$ then by eq^n ② & ①

$$f(x) = -\frac{1}{2}(f(x)+f(x)) + \frac{1}{2}(f(x)-f(x))$$
$$\Rightarrow -f(x) = -\frac{1}{2}(f(x)+f(x)) - \frac{1}{2}(f(x)-f(x))$$
$$\Rightarrow f(x) = \frac{1}{2}(f(x)+f(x)) + \frac{1}{2}(f(x)+f(x))$$
$$\Rightarrow f(x) = f_s(x) + f_a(x)$$
$\because f_s(x) \in S$; $f_a(x) \in A$

Hence any $f(x)$ can be written as $f(x) = f_s(x) + f_a(x)$
$\because f_s(x) \in S$; $f_a(x) \in A$

(b) orthonormal basis for $L_2([-\pi, \pi]) = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}\right\}$

for S : $f(x) = f(-x)$

these are the even functions

$\therefore$ basis $= \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}\right\}$ for $n = 1, 2, \ldots$

for A : $f(x) = -f(-x)$ represents the odd functions

$\therefore$ basis $= \left\{\frac{\sin nx}{\sqrt{\pi}}\right\}$ for $n = 1, 2, \ldots$

This can also be confirmed as the set represents a fourier series for which $\int_{-\pi}^{\pi} f(x)^2 dx$ converges.

(c) from ① we know $f(x)$ can be written as sum of f_s & f_a

$\Rightarrow f(x) \subset f_s(x) + f_a(x)$ — ① $\because f_a(x) \in A$ & $f_s(x) \in S$

We may write $f_s(x) = \frac{f(x)+f(-x)}{2}$ & $f_a(x) = \frac{f(x)-f(-x)}{2}$

Also $f_s(x) + f_a(x) = \frac{f(x)+f(-x)+f(x)-f(-x)}{2} = f(x)$

$\Rightarrow f_s(x) + f_a(x) \subset f(x)$ — ②

from ① & ② $\because f(x) \in L_2([-\pi, \pi])$

if $A \oplus S \in L$ & $L \subset A \oplus S$ then we can write

$f(x) \in L_2([-\pi, \pi]) = S \oplus A$ Hence proved

Page No.

III Given.

B1	B2
2b	1b, 1w

$P(B_1) = \frac{1}{2} = P(B_2)$

since there is equal chance of picking either bag.

$P(b/B_1) = 1$ $\qquad$ $P(b/B_2) = \frac{1}{2}$

given drawn ball was black, to find Prob. it belongs to 1^{st} bag. This is a Baye's theorem problem

$$\Rightarrow P(B_1/b) = \frac{P(b/B_1)\,P(B_1)}{P(b/B_1)P(B_1) + P(b/B_2)P(B_2)}$$

$$= \frac{(1)(1/2)}{(1)(1/2) + (1/2)(1/2)}$$

$$= 1/2 \,/\, 3/4$$

$$P(B_1/b) = 2/3$$ Ans

IV Given $m \times n$ set of eqn's $Ax = b$ where $m > n$ rank $(A) = n$

The SVD of A can be expressed as $A = U \Sigma V^T$

$\because$ U = $m \times n$ orthonormal matrix of eigenvectors of AA^T

V^T = transpose $n \times n$ matrix containing orthonormal eigen vectors of A^TA

Σ = $n \times n$ diagonal matrix of singular values which are non-negative diagonal entries $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ which are sq. roots of eigenvalues of A^TA

$\rightarrow$ SVD $A = U \Sigma V^T$ (part (a))

To find the least sq. soln to system $Ax = b$, we follow the given steps. We want to minimize the error $\|Ax - b\|$ to find least square solutions $\Rightarrow \min \{\|Ax - b\|^2\}$

Proof of step 2: Set $b' = U^T b$

We can write $\|Ax - b\|^2$ as:

$\|U \Sigma V^T x - b\|^2$ $\quad$ $\{ \because A = U \Sigma V^T$

Date ___/___/______

$= \| U \Sigma V^T x + U^T b - U^T b - b \|^2$

$= \| U \Sigma V^T x - U^T b \|^2 + \| U^T b - b \|^2$

{ Using parallgram law
$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$
$u = U\Sigma V^T x - U^T b$; $v = U^T b - b$
on applying the law we get :
{ $\|U\Sigma V^T x - U^T b\|^2 + \|U^T b - b\|^2 =$
} $\|U\Sigma V^T x + U^T b - U^T b - b\|^2$

To minize the first term we equate to get x.

$\Rightarrow U \Sigma V^T x = U^T b$

$\Rightarrow x = V U^T b / \Sigma$

for part (d) ↳ (1) { as $UU^T = I$; $VV^T = I$ given orthonormal }

to minize $\| U^T b - b \|^2$, we differentiale wrt v

$\Rightarrow \nabla_b (U^T b - b) = 0$

$\rightarrow U^T U^T b - U^T b' = 0$

$\Rightarrow U^T b' = U^T U^T b$

U^T is $n \times n$, rank n invertable matrix

$\Rightarrow$ ~~$b' = (U^T U^T) U^T b$~~

~~$b = U U U^T b$~~

~~$\Rightarrow b = U^T b$~~

{ $(AB)^{-1} = B^{-1} A^{-1}$
{ $UU^T = I$

This is also the projection of b onto the colum space of U.

$\Rightarrow b' = (U^T)^{-1} U^T U^T b$

$\boxed{b' = U^T b}$ hence proved. part (b)

Then we compute $y_i = b'_i / \sigma_i$ (part (c))

$\forall i = 1, 2, \ldots, n$

We can clearly see that the resulting matrix Y is the minimum b_i.

Thus, putting in eq^n (1)

$x = V . \left(U^T b / \Sigma \right)$

$x = V . Y$

{ $\because y_i = U^T b_i / \sigma_i$
$Y = U^T b / \Sigma$

this follows from eq^n (1) that x is minimum solution to $\|Ax - b\|^2$

Hence all parts together contribute a least square solution

Page No.

Saathi

Date ___/___/____

[V] Given rank $(A) = r < n$, Similarly to Q.4

This is a general least-square problem., but since rank$(A) = r$, the Σ matrix will look like:

$$\begin{pmatrix} \sqrt{\lambda_1} & & & & & 0 \\ & \sqrt{\lambda_2} & & & & \\ & & \ddots & & & \\ & & & \sqrt{\lambda_r} & & \\ & & & & 0 & \\ 0 & & & & & \ddots 0 \end{pmatrix} = \begin{pmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & 0 \end{pmatrix}$$

$$\Rightarrow A_{m\times n} = U_{m\times n} \Sigma_{n\times n} V^T_{n\times n}$$

Again we follow the similar proof for minimizing $\|Ax-b\|^2$

The proof follows the same pattern

$\|U\Sigma V^T x - b\|^2$ is to be minimized.

$= \|U\Sigma V^T x - U^T b\|^2 + \|U^T b - b\|^2$

Minimizing both, $x = V U^T b / \Sigma$ — (1)

and $b' = U^T b$

then we compute $y_i = b'_i / \sigma_i$ $\forall$ i in $\{1, 2, \ldots r\}$

however since rank$(A) = r < n$; the rest of the value have to 0; as $\sigma_{r<i<n} = 0$ and cant be used for division

$$\therefore y_i = \begin{cases} b'_i/\sigma_i & \text{if } 1 < i \leq r \\ 0 & \text{otherwise} \end{cases}$$

The resulting matrix Y is the minimum $b_i \neq 0$

putting in eq^n (1) $x = VY$

gives the minima of $\|Ax-b\|^2$ and gives minimum $\|x\|$

Hence in the above procedure, all parts gives the solution x of minimum norm $\|x\| = Vy$.

N

Date ___/___/_____

[VI] for any X the PMF can be written as:

$$P_X(x) = \begin{cases} 1/52 & \text{if } x = 1 \\ \vdots & \\ 1/52 & \text{if } x = 52 \\ 0 & \text{otherwise} \end{cases}$$

$\because x_i$ is the card value X can take

$\therefore$

(a) $E[X_1] = \Sigma x \, P(X = x)$

$= 1\left(\frac{1}{52}\right) + 2\left(\frac{1}{52}\right) + \cdots + 52\left(\frac{1}{52}\right) + 0$

$= \frac{1}{52}(1 + 2 + \cdots + 52)$

$= \frac{(52)(53)}{52 \times 2} = 26.5$ which is same $\forall$ $E[X_i]$ $\because i \in [1, 50]$

(b) $Z = X_1 - 2X_2 + 3X_3$

since expectation can be linearly distributed

$\Rightarrow E[Z] = E[X_1] - 2E[X_2] + 3E[X_3]$

$= 26.5 - 2(26.5) + 3(26.5)$

$= 53$

(c) $Y = X_1 - X_2 + \cdots + X_{49} - X_{50}$

since +ve & -ve equally distributed & $E[X_i] = 26.5$

$\Rightarrow E[Y] = E[X_1] - E[X_2] + \cdots + E[X_{49}) - E[X_{50}]$

$= 0$

[VII] Given joint PMF

(a) marginal distribution: $P(X) = \begin{cases} 1/8 + 1/12 & , x = 1 \\ 1/8 + 1/12 + 1/30 & , x = 2 \\ 1/9 + 1/6 + 1/30 & , x = 3 \\ 1/9 + 0 + 4/15 & x = 4 \end{cases}$

$$P(x) = \begin{cases} 5/36 & , x = 1 \\ 31/180 & , x = 2 \\ 14/45 & , x = 3 \\ 17/45 & , x = 4 \\ 0 & , \text{otherwise} \end{cases} = \cancel{\begin{cases} 5/36 & \\ 31/180 & \\ 56/180 & x = 3 \\ 17/45 & x = 4 \\ 0 & \text{ow} \end{cases}}$$

Saathi

Date ___/___/_____

similarly.

$$P(y) = \begin{cases} 1/3 & , y=0 \\ 1/3 & , y=1 \\ 1/3 & , y=2 \\ 0 & , \text{otherwise} \end{cases}$$

(b) To X & Y to be independent, $P(X \cap Y) = P(X)\,P(Y)$ should hold.
$= P(X, Y)$

taking $X=1$ & $Y=0$,
$$P(X=1, Y=0) = \frac{1}{18}$$

$\cancel{P(X)} = P(X=1) = \frac{5}{36}$, $P(Y=0) = \frac{1}{3}$

$$\Rightarrow P(X)P(Y) = \frac{5}{108} \neq \frac{1}{18} = P(X,Y)$$

$\therefore$ X & Y are not independent

(c) to find $P(X, Y=2) = \begin{cases} 0 & x=1 \\ 1/30 & x=2 \\ 1/30 & x=3 \\ 4/15 & x=4 \end{cases}$

$E[X|Y=2]$ is given by conditional expectation of :

$$\sum_{k=1}^{4} k \cdot \frac{(P(X=k, Y=2))}{P(Y=2)}$$

$$= 1\left(\frac{0}{1/3}\right) + 2\left(\frac{1/30}{1/3}\right) + 3\left(\frac{1/30}{1/3}\right) + 4\left(\frac{4/15}{1/3}\right)$$

$$E[X|Y=2] = \frac{1}{5} + \frac{3}{10} + \frac{16}{5} = \frac{37}{10} = 3.7$$

(d) $E[X] = 1\left(\frac{5}{36}\right) + 2\left(\frac{31}{180}\right) + 3\left(\frac{14}{45}\right) + 4\left(\frac{17}{45}\right) = 2.928$
$= 527/180$

$$E[Y] = 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right) + 2\left(\frac{1}{3}\right) = 1$$

$$E[XY] = \sum_x \sum_y (xy)\, P(X=x, Y=y)$$

Saathi

Date ___/___/______

$= \sum\sum$ $(1)(0)P(X=1,Y=0) + (1)(1)P(X=1,Y=1) + (1)(2)P(X=1,Y=2)$
$+ (2)(0)P(X=2,Y=0) + (2)(1)P(X=2,Y=1) + (2)(2)P(X=2,Y=2)$
$+ (3)(0)P(X=3,Y=0) + (3)(1)P(X=3,Y=1) + (3)(2)P(X=3,Y=2)$
$+ (4)(0)P(X=4,Y=0) + (4)(1)P(X=4,Y=1) + (4)(2)P(X=4,Y=2)$
$= (1)(0)(1/18) + (1)(1)(1/12) + (1)(2)(0)$
$+ (2)(0)(1/18) + (2)(1)(1/12) + (2)(2)(1/30)$
$+ (3)(0)(1/9) + (3)(2)(1/6) + (3)(3)(1/30)$
$+ (4)(0)(1/9) + (4)(1)(0) + (4)(2)(4/15)$

$= \dfrac{229}{60}$

$= 3.816$

VIII Given $A \in \mathbb{R}^{m\times n}$ $\Rightarrow$ A is a $m\times n$ matrix $\Rightarrow A : \mathbb{R}^n \to \mathbb{R}^m$

~~to show~~ X st $A\vec{x} = 0$ is subspace of $\mathbb{R}^n$

For proving a subspace, the three conditions must be true :

(i) $\vec{0}_n \in X$ (ii) if $u, v \in X$ then $u+v \in X$

(iii) for some $c \in \mathbb{R}$, $\vec{u} \in \mathbb{R}^n$, $c\vec{u} \in X$

(i) $A\vec{0}_n = \vec{0}_m$, since the given sol^n $AX=0$, 0 is a sol^n

$\Rightarrow 0 \in X$; and $0 \in \mathbb{R}^n$

(ii) for some $\vec{u}, \vec{v} \in \mathbb{R}^n$, $\vec{u}+\vec{v} \in \mathbb{R}^n$

$\Rightarrow A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$

$= \vec{0}_m + \vec{0}_m$

$\left\{ \begin{array}{l} \circ \text{ distributive property} \\ \circ \ A\vec{x} = 0 \end{array} \right.$

here $\vec{u}+\vec{v} \in X$, and $\vec{u}+\vec{v} \in \mathbb{R}^n$

(iii) for some scalar $c \in \mathbb{R}$, $c\vec{u} \in \mathbb{R}^n$

$\Rightarrow$ ~~$A(c\vec{u})$~~ $A\vec{u} = \vec{0}_m$

$cA\vec{u} = \vec{0}_m$

$\Rightarrow A(c\vec{u}) = \vec{0}_m$

$\Rightarrow cu \in X$ and $cu \in \mathbb{R}^n$

$\therefore$ by above 3 X is a subspace of $\mathbb{R}^n$

this is the kernel of A.

Saathi

Date __/__/__

IX Given

$v_1 = [1, -1, 2, 0]$ $\qquad$ $v_3 = [1, -2, 3, -1]$

$v_2 = [1, 0, 1, 1]$ $\qquad$ $v_4 = [3, 1, 2, 4]$

(a) to check linear independence, the reduced echlon form of matrix formed by $v_i(s)$ should be of full rank.

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ -1 & 0 & -2 & 1 \\ 2 & 1 & 3 & 2 \\ 0 & 1 & -1 & 4 \end{bmatrix}^T \Rightarrow \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & -2 & 3 & -1 \\ 3 & 1 & 2 & 4 \end{bmatrix} \xrightarrow{R_3 - R_1} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 3 & 1 & 2 & 4 \end{bmatrix}$$

$$\xrightarrow{R_4 - 3R_1} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 0 & 4 & 4 & 4 \end{bmatrix} \xrightarrow{R_3 + R_2} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 4 & -4 & 4 \end{bmatrix} \xrightarrow{R_4 - 2R_2} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{aligned} x_1 - x_2 &= -2x_3 + x_4 \\ x_2 &= x_3 + x_4 \end{aligned}$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$\therefore$ basis $= \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$

Since the rank (A) = 2 not full, it has 2 basis, given vectors

They are NOT linearly Independent

(b) we know: row rank (A) = col rank (A) (as discussed in ~~call~~ class)

In above the col-matrix of vectors has rank 2, $\therefore$ row rank of ~~the~~ row matrix A = 2

this also matches with the rref of matrix formed by taking vectors as rows

(c) Basis of null space = $\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$

Basis of col space = $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -2 \\ 1 \end{pmatrix} \right\}$

Saathi

Date ___/___/_____

$\boxed{X}$ Given finite dimensional space $\mathbb{R}^n$

Orthonormal basis $\{x_1, x_2, \dots x_n\}$

$$y = \Sigma_i a_i x_i \qquad a_i = \langle x_i, y \rangle$$

minimize $|y - \hat{y}|$ $\because$ $\hat{y} = \sum_{i=1}^{k} \beta_i x_i$

Since $\{x_1, x_2, \dots x_n\}$ are orthogonal ; $x_i x_j = 0 \quad \forall i \neq j$

$= 1 \quad \forall i = j$

minimize $\left| \Sigma_i^k a_i x_i - \Sigma_i^k \beta_i x_i \right|^2$

Since it is a square & x_i is common, to minimize we should prove $a_i = \beta_i \quad \forall i = 1, \dots k$

k is sub space spanned on $\mathbb{R}^n$

$$y = a_1 x_1 + a_2 x_2 + \dots a_n x_n \quad \rightarrow n \text{ terms}$$
$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad \rightarrow k \text{ terms}$$

$\Rightarrow |y - \hat{y}| = |a_1 x_1 + a_2 x_2 + \dots + a_n x_n - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)|$

$= |(a_1 - \beta_1) x_1 + \dots + (a_k - \beta_k) x_k + a_{k+1} x_{k+1} + \dots + a_n x_n|$

for a vector x, $\|x\|^2 = \langle x, x \rangle = x^T x$

we minimize $|y - \hat{y}|^2$ (least square)

$\Rightarrow [(a_1 + \beta_1) x_1 + \dots + (a_k - \beta_k) x_k + \dots + a_n x_n]^T [(a_1 - \beta_1) x_1 + \dots + (a_k - \beta_k) x_k + \dots + \alpha_n x_n]$

$= [(a_1 - \beta_1) x_1^T + \dots + (a_k - \beta_k) x_k^T + \dots + a_n x_n^T] [(a_1 - \beta_1) x_1 + \dots + (a_k - \beta_k) x_k + \dots + a_n x_n]$

$= (a_1 - \beta_1) |x_1|^2 + (a_1 - \beta_1) x_1^T x_2 + \dots + (a_1 - \beta_1)(a_k - \beta_k) x_1^T x_k +$

$\dots (a_1 - \beta_1) x_1^T x_n + (a_2 - \beta_2)(a_1 - \beta_1) x_1^T x_2 + (a_2 - \beta_2)^2 |x_2|^2 +$

$\dots + (a_2 - \beta_2) x_n x^T + (a_3 - \beta_3)(a_1 - \beta_1) x_1^T x_3 + \dots$

$\dots + (a_k - \beta_k) x_k^2 + (a_k - \beta_k) x_k^T x_n + \dots$

We can use the identity $|x_i|^2 = 1 \quad \forall i = 1, \dots, n$

$x_i^T x_j = 0 \quad \forall i \neq j$ (orthonormal)

$\therefore$ the above becomes :

$= (a_1 - \beta_1)^2 + 0 + 0 + \dots + (a_2 - \beta_2)^2 + 0 + \dots + 0 + (a_3 - \beta_3)^2 +$

$\dots (a_k - \beta_k)^2$

$= (a_1 - \beta_1)^2 + (a_2 - \beta_2)^2 + \dots + (a_k - \beta_k)^2$

The above sum is always +ve to minimize $|y - \hat{y}|^2$, the expression should be $=$ zero which is only possible

$\{ \alpha_i = \beta_i \quad \forall i = 1, 2, \dots k. \}$