ECE C247 Homework 01

Chih-En Lin

1. (a)

(i) Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\Rightarrow AA^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{bmatrix}$$

$$= I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{cases} a^2+b^2 = 1 \\ c^2+d^2 = 1 \\ ac+bd = 0 \end{cases}$$

Assume $a = \frac{1}{2}$, $d = \frac{-1}{2}$

$$\Rightarrow \begin{cases} b = \frac{\sqrt{3}}{2} \\ c = \frac{\sqrt{3}}{2} \end{cases} \Rightarrow A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} \end{bmatrix}$$

$$Ax = \lambda x$$

$$\Rightarrow (A - \lambda I)x = 0$$

$$\det(A - \lambda I) = 0$$

$$\Rightarrow \det\left(\begin{bmatrix} \frac{1}{2} - \lambda & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} - \lambda \end{bmatrix} \right) = 0$$

$$\Rightarrow \left(\frac{1}{2} - \lambda\right)\left(\frac{-1}{2} - \lambda\right) - \frac{3}{4} = 0$$

$$\Rightarrow \lambda^2 - 1 = 0$$

$$\Rightarrow \lambda = \pm 1$$

eigenvalues: $1, -1$

① eigenvector with $\lambda_1 = 1$:

$(A - I)x = 0$

$$\Rightarrow \begin{bmatrix} \frac{1}{2} - 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$\Rightarrow x_1 = \sqrt{3}\, x_2$

$\Rightarrow$ Let $x_2 = 1 \Rightarrow x_1 = \sqrt{3}$

normalize

$$\Rightarrow x = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \lambda_1 = 1$$

② eigenvector with $\lambda_2 = -1$

$(A + I)x = 0$

$$\Rightarrow \begin{bmatrix} \frac{1}{2} + 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} + 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} \frac{3}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$\Rightarrow x_2 = -\sqrt{3}\, x_1$

$\Rightarrow$ Let $x_1 = 1 \Rightarrow x_2 = -\sqrt{3}$

normalize

$$\Rightarrow x = \begin{bmatrix} \frac{1}{2} \\ \frac{-\sqrt{3}}{2} \end{bmatrix}, \quad \lambda_2 = -1$$

We can notice that eigenvalues $|\lambda| = 1$,
eigenvectors are orthogonal. #

(ii) $AA^T = I$

$\Rightarrow A^T A = I$

$Ax = \lambda x$

$\Rightarrow (Ax)^T (Ax) = (\lambda x)^T (\lambda x)$

$\Rightarrow x^T A^T A x = |\lambda|^2 x^T x$

$\Rightarrow x^T x = |\lambda|^2 x^T x$ $\because$ eigenvectors are normalized

$\Rightarrow x^T x = 1$

$\Rightarrow |\lambda|^2 = 1$ #

(iii) Assume λ_1, λ_2 are the eigenvalues of A

x_1, x_2 are the eigenvectors of A

$\lambda_1 \neq \lambda_2$

$Ax_1 = \lambda_1 x_1 , \quad Ax_2 = \lambda_2 x_2$

$(Ax_1)^H (Ax_2) = (\lambda_1 x_1)^H (\lambda_2 x_2)$

$\Rightarrow x_1^H A^H A x_2 = \bar{\lambda}_1 \lambda_2 x_1^H x_2$

$(\because A^H A = I)$

$\Rightarrow x_1^H x_2 = \bar{\lambda}_1 \lambda_2 x_1^H x_2$

$\Rightarrow (\bar{\lambda}_1 \lambda_2 - 1) x_1^H x_2 = 0$

From (ii), $\bar{\lambda}_1 \lambda_2 = 1$ only if $\lambda_1 = \lambda_2$

$\Rightarrow x_1^H x_2 = 0$

it means the eigenvectors are orthogonal. #

(iv)

Vector x will be rotated or reflected,

but the length will not be changed. #

(b)

(i) The left singular vectors of A are the eigenvectors of AA^T.

The right singular vectors of A are the eigenvectors of A^TA.

Singular Value Decomposition:

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

$$\begin{aligned} AA^T &= (U\Sigma V^T)(U\Sigma V^T)^T \\ &= U\Sigma V^T V \Sigma^T U^T \quad \left(\because \begin{array}{l} V^TV = I \\ \Sigma^T = \Sigma \end{array}\right) \\ &= U\Sigma^2 U^T \end{aligned}$$

$$\begin{aligned} A^TA &= (U\Sigma V^T)^T(U\Sigma V^T) \\ &= V\Sigma^T U^T U \Sigma V^T \quad \left(\because \begin{array}{l} U^TU = I \\ \Sigma^T = \Sigma \end{array}\right) \\ &= V\Sigma^2 V^T \end{aligned}$$

#

(ii) The singular values of A are the square root of the eigenvalues of AA^T and A^TA.

(c)

(i) False. ex: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ eigenvalues are all 1.

(ii) False. ex: the eigenvectors of $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, but their sum $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not an eigenvector.

(iii) True. $Ax = \lambda x$

$\Rightarrow x^T A x = x^T(\lambda x) = x^T x \lambda \geq 0$

$\because$ $x^T x$ is definitely non-negative

$\Rightarrow$ λ is also non-negative.

(iv) True. ex: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $\rightarrow$ rank : 2

$\rightarrow$ has only one distinct eigenvalue 1

(v) True. ex: the eigenvectors of $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ corresponding to the same eigenvalue 1.

Their sum $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is also an eigenvector.

2. (a) H denotes Head, T denotes Tail

(i)

$$P(H50 \mid T) = \frac{P(T \mid H50) \cdot P(H50)}{P(H50, T) + P(H60, T)}$$

$$= \frac{P(T \mid H50) \cdot P(H50)}{P(T \mid H50) \cdot P(H50) + P(T \mid H60) \cdot P(H60)}$$

$$= \frac{0.5 \cdot 0.5}{0.5 \cdot 0.5 + 0.4 \cdot 0.5}$$

$$= \frac{25}{45}$$

$$= \frac{5}{9} \approx 0.5556$$ #

(ii)

$$P(H50 \mid THHH) = \frac{P(THHH \mid H50) \cdot P(H50)}{P(H50, THHH) + P(H60, THHH)}$$

$$= \frac{P(THHH \mid H50) \cdot P(H50)}{P(THHH \mid H50) \cdot P(H50) + P(THHH \mid H60) \cdot P(H60)}$$

$$= \frac{(0.5)^4 \cdot 0.5}{(0.5)^4 \cdot 0.5 + (0.4 \cdot (0.6)^3) \cdot 0.5}$$

$$= \frac{3125}{3125 + 4320}$$

$$= \frac{625}{1489} \approx 0.4197$$ #

(iii) Let A denotes 9 Heads and 1 Tail in 10 flips, regardless the order.

①

$$P(H50|A) = \frac{P(A|H50)\cdot P(H50)}{P(H50, A) + P(H55, A) + P(H60, A)}$$

$$= \frac{P(A|H50)\cdot P(H50)}{P(A|H50)\cdot P(H50) + P(A|H55)\cdot P(H55) + P(A|H60)\cdot P(H60)}$$

$$= \frac{((0.5)^9\cdot 0.5)\cdot \frac{1}{3}}{((0.5)^9\cdot 0.5)\cdot \frac{1}{3} + ((0.55)^9\cdot 0.45)\cdot \frac{1}{3} + ((0.6)^9\cdot 0.4)\cdot \frac{1}{3}}$$

≈ 0.1379 #

②

$$P(H55|A) = \frac{P(A|H55)\cdot P(H55)}{P(H50, A) + P(H55, A) + P(H60, A)}$$

$$= \frac{P(A|H55)\cdot P(H55)}{P(A|H50)\cdot P(H50) + P(A|H55)\cdot P(H55) + P(A|H60)\cdot P(H60)}$$

$$= \frac{((0.55)^9\cdot 0.45)\cdot \frac{1}{3}}{((0.5)^9\cdot 0.5)\cdot \frac{1}{3} + ((0.55)^9\cdot 0.45)\cdot \frac{1}{3} + ((0.6)^9\cdot 0.4)\cdot \frac{1}{3}}$$

≈ 0.2927 #

③

$$P(H60 \mid A) = \frac{P(A \mid H60) \cdot P(H60)}{P(H50, A) + P(H55, A) + P(H60, A)}$$

$$= \frac{P(A \mid H60) \cdot P(H60)}{P(A \mid H50) \cdot P(H50) + P(A \mid H55) \cdot P(H55) + P(A \mid H60) \cdot P(H60)}$$

$$= \frac{((0.6)^9 \cdot 0.4) \cdot \frac{1}{3}}{((0.5)^9 \cdot 0.5) \cdot \frac{1}{3} + ((0.55)^9 \cdot 0.45) \cdot \frac{1}{3} + ((0.6)^9 \cdot 0.4) \cdot \frac{1}{3}}$$

$$\approx 0.5694$$ #

(b)

$$P(\text{Preg} \mid \text{Pos}) = \frac{P(\text{Pos} \mid \text{Preg}) \cdot P(\text{Preg})}{P(\text{Preg}, \text{Pos}) + P(\text{NotPreg}, \text{Pos})}$$

$$= \frac{P(\text{Pos} \mid \text{Preg}) \cdot P(\text{Preg})}{P(\text{Pos} \mid \text{Preg}) \cdot P(\text{Preg}) + P(\text{Pos} \mid \text{NotPreg}) \cdot P(\text{NotPreg})}$$

$$= \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.1 \cdot 0.99}$$

$$= \frac{1}{11} \approx 0.0909$$ #

If a woman received a positive test, only 9 % probability that she is pregnant. It is a bad test. Because a large proportion (99 %) of the female population is not pregnant, but the test returns "positive" 10 %, it will cause many incorrect tests. #

(c)

$$E(x_i + b) = E(x_i) + b$$

$$E(Ax_i) = E\left(\sum_{j=1}^{n} A_{i,j} x_j\right)$$

$$= \sum_{j=1}^{n} A_{i,j} E(x_j)$$

$$= \sum_{j=1}^{n} A_{i,j} E(x)_j$$

$$= [A \cdot E(x)]_i$$

$$\Rightarrow E(Ax) = A \cdot E(x)$$

$$\therefore E(Ax + b) = E(Ax) + b$$

$$= AE(x) + b \quad \#$$

(d)

$$\text{cov}(x) = E((x - E(x))(x - E(x))^T)$$

$$\text{cov}(Ax+b) = E((Ax+b-E(Ax+b))(Ax+b-E(Ax+b))^T)$$

From (c) $= E((Ax+b - AE(x)-b)(Ax+b-AE(x)-b)^T)$

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

$$= E(A(x - E(x))(x - E(x))^T A^T)$$

$$= A\,E((x - E(x))(x - E(x))^T)\,A^T$$

$$= A\,\text{cov}(x)\,A^T \quad \#$$

3. (a)

$$\nabla_x x^T A y = A y \quad \#$$

(b)

$$\nabla_y x^T A y = A^T x \quad \#$$

(c)

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

$$= x y^T \quad \#$$

(d)

$$\nabla_x f = \nabla_x (x^T A x + b^T x)$$

$$= \nabla_x (x^T A x) + \nabla_x (b^T x)$$

$$= A x + A^T x + b \quad \#$$

(e)

Assume $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$

$$tr(AB) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} b_{j,i}$$

$$\frac{\partial tr(AB)}{\partial a_{i,j}} = b_{j,i}$$

$$\nabla_A f = \nabla_A tr(AB) = B^T \quad \#$$

4.

$$\frac{1}{2}\sum_{i=1}^{n}\|y^{(i)}-Wx^{(i)}\|^2$$

$$= \frac{1}{2}\sum_{i=1}^{n}(y^{(i)}-Wx^{(i)})^T(y^{(i)}-Wx^{(i)})$$

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

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

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

Only take out the terms with W

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

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

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

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

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

$$= -tr(WXY^T) + \frac{1}{2}tr(WXX^TW^T)$$

From Hint, $\frac{\partial f(W)}{\partial W}$

$$= -(XY^T)^T + \frac{1}{2}\left(W(XX^T)^T + W(XX^T)\right)$$

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

$$= -YX^T + WXX^T = 0$$

$$\Rightarrow WXX^T = YX^T$$

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

$$= YX^{-1}$$

#