1.

(a) i. $A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$

$AA^T = I$

$A - \lambda I$

$= \begin{pmatrix} \frac{1}{\sqrt{2}} - \lambda & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{pmatrix}$

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

$\Rightarrow \lambda_1 = \frac{1+j}{\sqrt{2}}$

$\lambda_2 = \frac{1-j}{\sqrt{2}}$

$A - \lambda_1 I$

$= \begin{pmatrix} -\frac{j}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{j}{\sqrt{2}} \end{pmatrix}$

$\Rightarrow$ eigenvector $v_1 = \begin{pmatrix} -\frac{j}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

$A - \lambda_2 I$

$= \begin{pmatrix} \frac{j}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{j}{\sqrt{2}} \end{pmatrix}$

$\Rightarrow$ eigenvector $v_2 = \begin{pmatrix} \frac{j}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

ii. $Av = \lambda v$

$\Rightarrow \|Av\|^2 = \|\lambda v\|^2 = |\lambda|^2 \|v\|^2$

$\|Av\|^2$

$= (Av)^T (Av)$

$= (v^T A^T)(Av)$

$= v^T A^T A v$

$AA^T = I$

$\Rightarrow AA^T A = A$

$\left.\begin{array}{l} \Rightarrow A(A^T A - I) = 0 \\ A \neq 0 \end{array}\right\}$

$\Rightarrow A^T A = I$

$\Rightarrow \|Av\|^2$

$= v^T v = \|v\|^2$

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

iii. Let λ_1, λ_2 be distinct eigenvalues of A corresponding to e-vects v_1, v_2

$\left.\begin{array}{l} Av_1 = \lambda_1 v_1 \\ Av_2 = \lambda v_2 \end{array}\right\} \Rightarrow (Av_1)^T Av_2 = (\lambda v_1)^T (\lambda v_2)$

$\Rightarrow v_1^T A^T A v_2 = \lambda_1 \lambda_2 v_1^T v_2$

$\Rightarrow v_1^T v_2 = \lambda_1 \lambda_2 v_1^T v_2$

$\Rightarrow (\lambda_1 \lambda_2 - 1) v_1^T v_2 = 0$

Since ~~$v_1 \neq v_2$~~ $\lambda_1 \neq \lambda_2$ and $|\lambda| = 1$,

$\lambda_1 \lambda_2 \neq 1$

So $v_1^T v_2 = 0$, namely orthogonal

iv. Its norm would not change. However, it will be rotated or flipped.

(b) i. Denote SVD of A is $A = U \Sigma V^T$
$= U_1 S V_1^T$

AA^T

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

$= U_1 S S^T U_1^T$

$A^T A$

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

$= V_1 S^T S V_1^T$

The left singular vectors of A are e-vects of AA^T. The right singular vectors of A are e-vects of $A^T A$.

ii. The singular value of A is the square root of e-vals of AA^T and $A^T A$.

(c) i. False. Zero linear operator 0 has ~~only~~ one eigenvalue 0. Identity linear operator has only one distinct eigenvalue, 1.

ii. False. Sum would change direction of e-vect.

iii. True.

iiii. False. The argument is false when counting repeated e-vals.

v. ~~True~~. False.

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

with two e-vals 1, 1 corresponding to e-vect

$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$

However, $\begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix}$ is not an e-vect.

2.

(a) i. $p(H50 \mid T)$

$= \frac{p(H50, T)}{p(T)}$

~~$= \frac{p(T \mid H50)\, p(H50)}{\frac{1}{2}(p(H50, T) + p(H60, T))}$~~

~~$= \frac{0.5 \times 0.5}{\frac{1}{2}(0.5 + 0.6)}$~~

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

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

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

$= \frac{5}{9}$

ii. $p(H50 \mid THHH)$

$= \frac{p(H50, THHH)}{p(H50, THHH) + p(H60, THHH)}$

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

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

$= \frac{625}{1489}$

$= 0.4197$

iii. $p(H50 \mid \geq 9H)$

$= \frac{p(\geq 9H \mid H50)\, p(H50)}{p(\geq 9H \mid H50) p(H50) + p(\geq 9H \mid H55) p(H55) + p(\geq 9H \mid H60) p(H60)}$

$= \frac{(0.5^{10} \times 11) \times \frac{1}{3}}{(0.5^{10} \times 11) \times \frac{1}{3} + (0.55^{10} + 10 \times 0.45 \times 0.55^9) \times \frac{1}{3} + (0.6^{10} + 10 \times 0.4 \times 0.6^9) \times \frac{1}{3}}$

$= 0.1337$

Similarly,

$p(H55 \mid \geq 9H)$

$= \frac{(0.55^{10} + 10 \times 0.45 \times 0.55^9) \times \frac{1}{3}}{(0.5^{10} \times 11) \times \frac{1}{3} + (0.55^{10} + 10 \times 0.45 \times 0.55^9) \times \frac{1}{3} + (0.6^{10} + 10 \times 0.4 \times 0.6^9) \times \frac{1}{3}}$

$= 0.2894$

$p(H60 \mid \geq 9H)$

$= \frac{(0.6^{10} + 10 \times 0.4 \times 0.6^9) \times \frac{1}{3}}{(0.5^{10} \times 11) \times \frac{1}{3} + (0.55^{10} + 10 \times 0.45 \times 0.55^9) \times \frac{1}{3} + (0.6^{10} + 10 \times 0.4 \times 0.6^9) \times \frac{1}{3}}$

$= 0.5769$

(b) $p(preg \mid pos)$ ~~$p(pos \mid preg)$~~

$$= \frac{p(pos, preg)}{\cancel{p(preg)}\, p(pos)}$$

$$= \frac{p(preg)\, p(pos \mid preg)}{p(pos \mid preg)\, p(preg) + p(pos \mid not\ preg)\, p(not\ preg)}$$

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

$$= 0.0909$$

That makes sense because fall-out ratio is too high (10%).

(c) $E(Ax_i)$

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

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

$$= \left(\sum_{j=1}^{n} A_{ij} E(x)_j\right)$$

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

$$\Rightarrow E(Ax) = AE(x)$$

$$\Rightarrow E(Ax+b) = E(Ax) + b = AE(x) + b$$

(d) $\mathrm{cov}(Ax+b)$

$$= E\left((Ax+b-AE(x)-b)(Ax+b-(AE(x)+b))^T\right)$$

$$= E\left((Ax-AE(x))(Ax-(AE(x)))^T\right)$$

$$= E\left(A(x-E(x))((x-E(x)))^T A^T\right)$$

$$= AE\left((x-E(x))((x-E(x)))^T\right)A^T$$

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

3.

(a) $\nabla_x x^T A y = Ay$

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

(c) $\nabla_A x^T A y$

$$= \begin{bmatrix} \frac{\partial x^T A y}{\partial a_{11}} & \cdots & \frac{\partial x^T A y}{\partial a_{1m}} \\ \vdots & & \vdots \\ \frac{\partial x^T A y}{\partial a_{n1}} & \cdots & \frac{\partial x^T A y}{\partial a_{nm}} \end{bmatrix}$$

$$= xy^T$$

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

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

$$= Ax + A^T x + b$$

(e) Suppose $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times q}$

$\mathrm{tr}(AB)$

$$= \mathrm{tr}\left(\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} [b_1 \ \cdots \ b_q]\right)$$

$$= \mathrm{tr}\left(\begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_q \\ \vdots & & & \vdots \\ a_m b_1 & \cdots & & a_m b_q \end{bmatrix}\right)$$

$$= \sum_{i=1}^{n} a_{1i} b_{i1} + \sum_{i=1}^{n} a_{2i} b_{i2} + \cdots$$

$$\Rightarrow \frac{\partial\, \mathrm{tr}(AB)}{\partial A}$$

$$= \begin{bmatrix} \frac{\partial\, \mathrm{tr}(AB)}{\partial a_{11}} & \cdots & \frac{\partial\, \mathrm{tr}(AB)}{\partial a_{1n}} \\ \vdots & & \vdots \\ \frac{\partial\, \mathrm{tr}(AB)}{\partial a_{m1}} & \cdots & \frac{\partial\, \mathrm{tr}(AB)}{\partial a_{mn}} \end{bmatrix}$$

$$\underline{\underline{m=q}} \begin{cases} \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{n1} \\ \vdots & & & \vdots \\ b_{1n} & \cdots & & b_{nn} \end{bmatrix} & \text{if } n = m \\ \cancel{\begin{bmatrix} b_{11} & \cdots & b_{n1} & 0 \\ \vdots & & \vdots & \\ b_{1n} & \cdots & b_{nn} & \\ 0 & \cdots & 0 & 0 \end{bmatrix}} & \text{if } n < m \ \ \cancel{\text{otherwise}} \\ \begin{bmatrix} b_{11} & \cdots & b_{m1} & 0 \\ \vdots & & \vdots & \vdots \\ b_{1m} & \cdots & b_{mm} & 0 \end{bmatrix} & \text{otherwise} \end{cases}$$

4. $f(W) = \frac{1}{2} \sum_{i=1}^{n} \| y^{(i)} - W x^{(i)} \|^2$

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

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

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

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

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

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

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

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

$$\frac{\partial f}{\partial W} = 0 \Rightarrow W = YX^T (XX^T)^{-1}$$