239 AS HW #1

1.)

a)

i) $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

$$AA^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det(A - \lambda I) = 0 \Rightarrow \det \begin{vmatrix} \frac{1}{\sqrt{2}} - \lambda & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{vmatrix} = 0$$

$$\left(\frac{1}{\sqrt{2}} - \lambda\right)\left(\frac{1}{\sqrt{2}} - \lambda\right) + \frac{1}{2} = 0 = \lambda^2 - \frac{2}{\sqrt{2}}\lambda + 1 \Rightarrow \lambda = \frac{1 \pm i}{\sqrt{2}}$$

$$Av_1 = \lambda_1 v_1, \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1-i}{\sqrt{2}} \begin{bmatrix} a \\ b \end{bmatrix}, \cancel{\frac{1}{\sqrt{2}}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \cancel{\frac{1}{\sqrt{2}}}(1-i)\begin{bmatrix} a \\ b \end{bmatrix}$$

$a - b = (1-i)a \quad a - b = a - ai, \; b = ai$

$a + b = (1-i)b \quad a + b = b - bi, \; a + ai = ai + a, \; 0 = 0$

inf solution, to keep orthonormal, $\boxed{a = \frac{1}{\sqrt{2}}, \; b = \frac{i}{\sqrt{2}}}$

$$Av_2 = \lambda_2 v_2, \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = 1+i \begin{bmatrix} c \\ d \end{bmatrix}$$

$c - d = (1+i)c, \; c - d = c + ci, \; d = -ci$

$c + d = (1+i)d, \; c + d = d + di, \; c - ci = -ci + c, \; 0 = 0$

$\Rightarrow \boxed{c = \frac{1}{\sqrt{2}}, \; d = -\frac{1}{\sqrt{2}} i}$

$$\boxed{\begin{aligned} \lambda_1 &= \frac{1-i}{\sqrt{2}}, & v_1 &= \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} i\right]^T \\ \lambda_2 &= \frac{1+i}{\sqrt{2}}, & v_2 &= \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} i\right]^T \end{aligned}}$$

ii) $Av = \lambda v$

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

$||Av||^2 = (Av)^T Av = \bar{v}^T A^T A v = v^T v = ||v||^2$

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

iii) For general eigenvector v of normal Matrix A

$Av = \lambda v$

$Av - \lambda v = 0$

$||(A - \lambda I)v|| = 0$ For normal A, $A - \lambda I$ is normal

$||(A - \lambda I)^* v|| = 0$

$A^* v - \bar{\lambda} v = 0$

$A^* v = \bar{\lambda} v$

$\lambda_1 x^T y = (\lambda_1 x)^T y = (Ax)^T y = x^T A^* y$

$= x^T \bar{\lambda}_2 y = (\lambda_2 x)^T y = \lambda_2 x^T y$

$\lambda_1 x^T y - \lambda_2 x^T y = 0 \quad (\lambda_1 - \lambda_2) x^T y = 0$

$\lambda_1 \neq \lambda_2 \therefore \lambda_1 - \lambda_2 \neq 0 \therefore x^T y = 0$

iv) Multiplication of a vector by an orthogonal matrix is a simple transformation in Euclidean space, such as a rotation or reflection

b i) - Left-singular vectors of A are the eigenvectors of AA^T

- Right-singular vectors of A are the eigenvectors of $A^T A$

ii) - The non zero singular values of A are the square roots of the eigen values of $A^T A$ & AA^T

c.)

i) false - Only orthogonal matrices do.
Counter example => identity matrix => two eigvals of 1

ii) false - it is a linear combination of the existing eigen vectors and therefore not orthogonal, giving no span information.

iii) true

iv) false - rank is the dimension of the vector space spanned by a matrix, which is equal to or less than number of non zero eigenvalues

v) true - the result is still scaled by a linear factor when A is applied

2.

a.

i) $P(H50 \mid T) = \dfrac{P(T \mid H50)\,P(H50)}{P(T)}$

$P(T \mid H50) = .5$, $P(H50) = P(H60) = .5$

$P(T) = P(T \mid H50)\,P(H50) + P(T \mid H60)\,P(H60)$

$= .5 \cdot .5 + .4 \cdot .5$

$= .45$

$P(H50 \mid T) = \dfrac{.5 \cdot .5}{.45} = .56$

ii)

$P(H50 \mid THHH) = \dfrac{P(THHH \mid H50) \cdot P(H50)}{P(THHH)}$

$P(THHH) = P(THHH \mid H50)\,P(H50) + P(THHH \mid H60)\,P(H60)$

$= (.5)^4 \cdot .5 + (.4)(.6)^3 \cdot .5$

$= .0744$

$P(THHH \mid H50) = .5^4 = .0625$

$P(H50 \mid THHH) = \dfrac{.0625 \cdot .5}{.0744} = .42$

iii) $P(H_9T) = P(H_9T \mid H50)\,P(H50) + P(H_9T \mid H55)\,P(H55)$
$+ P(H_9T \mid H60)\,P(H60)$

$= (.5)^{10} \cdot \frac{1}{3} + (.55)^9 \cdot .45 \cdot \frac{1}{3} + (.6)^9 \cdot .4 \cdot \frac{1}{3}$

$= .0624$

$P(H50 \mid H_9T) = \dfrac{P(H_9T \mid H50)\,P(H50)}{P(H_9T)} = .5^{10} \cdot \frac{1}{3} / .0624 = .138$

$P(H55 \mid H_9T) = \dfrac{P(H_9T \mid H55)\,P(H55)}{P(H_9T)} = .55^9 \cdot .45 \cdot \frac{1}{3} / .0624 = .293$

$P(H60 \mid H_9T) = \dfrac{P(H_9T \mid H60)\,P(H60)}{P(H_9T)} = .6^9 \cdot .4 \cdot \frac{1}{3} / .0624 = .569$

b.) $P(1|P) = .99$

$P(1|NP) = .1$

$P(P) = .01$

$P(NP) = .99$

$P(1) = P(1|P)P(P) + P(1|NP)P(NP)$

$= .99 \cdot .01 + .1 \cdot .99$

$= .1089$

$$P(P|1) = \frac{P(1|P)P(P)}{P(1)} = \frac{.99 \cdot .01}{.1089} = .091$$

This makes sense because given the amount of population that is not pregnant, many more positive results will be a result of a false positive on a non-pregnant person.

c.) $\mathbb{E}(Ax+b) = \mathbb{E}(Ax) + \mathbb{E}(b)$

$= A\mathbb{E}(x) + b$

d.) $\operatorname{cov}(Ax+b) = \mathbb{E}[(Ax+b-\mathbb{E}[Ax+b])(Ax+b-\mathbb{E}[Ax+b])^T]$

$= \mathbb{E}[(Ax+b-(A\mathbb{E}(x)+b))(Ax+b-(A\mathbb{E}(x)+b))^T]$

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

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

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

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

3. $x \in \mathbb{R}^n \quad y \in \mathbb{R}^m \quad A \in \mathbb{R}^{n \times m}$

a.)

$$\nabla_x \, x^T A y = \nabla_x \left[\sum_i^n \sum_j^m A_{ij} x_i y_j \right]$$

$$\frac{\partial(x^T A y)}{\partial x_1} = \sum_j^m A_{1j} y_j = (Ay)_1 \Rightarrow \frac{\partial(x^T A y)}{\partial x_n} = (Ay)_n$$

$$\boxed{\frac{\partial(x^T A y)}{\partial x} = \begin{bmatrix} (Ay)_1 \\ (Ay)_2 \\ \vdots \\ (Ay)_n \end{bmatrix} = Ay}$$

b.) $\nabla_y \, x^T A y = \nabla_y \left[\sum_i^n \sum_j^m A_{ij} x_i y_j \right]$

$$\frac{\partial(x^T A y)}{\partial y_1} = \sum_i A_{i1} x_i = (x^T A)_1$$

$$\boxed{\frac{\partial(x^T A y)}{\partial y} = \begin{bmatrix} (x^T A)_1 \\ \vdots \\ (x^T A)_m \end{bmatrix} = x^T A}$$

c.) $\nabla_A \, x^T A y = \nabla_A \left[\sum_i^n \sum_j^m A_{ij} x_i y_j \right]$

$$\frac{\partial x^T A y}{\partial A_{11}} = x_1 y_1$$

$$\boxed{\frac{\partial x^T A y}{\partial A} = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_m \\ \vdots & \ddots & \\ x_n y_1 & \cdots & x_n y_m \end{bmatrix} = x y^T}$$

d.) $f = x^T A x + b^T x =$

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

$$x^T A x = \sum_i \sum_j A_{ij} x_i x_j$$

$$\frac{\partial x^T A x}{\partial x_1} = 2A_{11} x_1 + \sum_{j \neq 1}^{n} A_{1j} x_j + \sum_{i \neq 1}^{m} A_{i1} x_i = (A + A^T) x$$

$$b^T x = \sum_i b_i x_i \quad \frac{\partial (b^T x)}{\partial x_1} = b_1 \; \ldots \; \frac{\partial (b^T x)}{\partial x} = b$$

$$\Rightarrow \boxed{\nabla_x f = (A + A^T) x + b}$$

e.) $f = tr(AB)$

$$AB = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} \sum a_{1i} b_{i1} & & \\ & \ddots & \\ & & \sum a_{ni} b_{in} \end{bmatrix}$$

$$tr(AB) = \sum_j \sum_i a_{ji} b_{ij}$$

$$\frac{\partial (tr(AB))}{\partial a_{11}} = b_{11} \quad \frac{\partial (tr(AB))}{\partial a_{12}} = b_{21} \quad \frac{\partial (tr(AB))}{\partial a_{nm}} = b_{mn}$$

$$\boxed{\nabla_A tr(AB) = B^T}$$

4. $\hat{y} = Wx$ $\quad LS \Rightarrow \min_{W} \frac{1}{2} \sum_{i=1}^{n} \|y^i - Wx^i\|^2$

Argument is vector $\therefore$ Euclidean norm = Frobenius Norm

$$\Rightarrow LS \Rightarrow \min_{W} \frac{1}{2} \sum_{i=1}^{n} \left(\|y^i - Wx^i\|_F\right)^2 = \min_{W} \frac{1}{2} \sum_{i=1}^{n} tr\left[(y^i - Wx^i)(y^i - Wx^i)^T\right]$$

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

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

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

deriving wrt W

$$= -\frac{\partial}{\partial W} \sum tr(Wx^i y^{iT}) + \frac{1}{2} \frac{\partial}{\partial W} \sum tr(W x^i x^{iT} W^T)$$

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

$(x_i x^T)^T = x_i x^T$

$$= -\sum y^i x^{iT} + \sum_{i=1}^{n} Wx^i x^{iT}$$

Set eq to zero

$$0 = -\sum y^i x^{iT} + W \sum x^i x^{iT} \Rightarrow W \sum x^i x^{iT} = \sum y^i x^{iT}$$

$$\boxed{W = \left(\sum_{i=1}^{n} x^i x^{iT}\right)^{-1} \sum_{i=1}^{n} y^i x^{iT}}$$