1.

a)

i.

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

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

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

$$\det\begin{pmatrix} \frac{1}{\sqrt{2}} - \lambda & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} - \lambda \end{pmatrix} = 0$$

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

$$-\frac{1}{2} - \cancel{\frac{1}{\sqrt{2}}\lambda} + \cancel{\frac{1}{\sqrt{2}}\lambda} + \lambda^2 - \frac{1}{2} = 0$$

$$\lambda^2 = 1$$

$$\lambda = \pm 1$$

$\lambda_1 = 1$ $\quad Ax = \lambda x$

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

$$\frac{1}{\sqrt{2}}(x_1 + x_2) = x_1 \Rightarrow x_2 = \sqrt{2}\,x_1 - x_1 \Rightarrow x_2 - (\sqrt{2} - 1)x_1 = 0$$

$$\frac{1}{\sqrt{2}}(x_1 - x_2) = x_2 \Rightarrow x_1 = \sqrt{2}x_2 + x_2 \Rightarrow x_1 - (\sqrt{2}+1)x_2 = 0$$

$x_2 = 1$

$x_1 = \sqrt{2} + 1$ $\quad a = \begin{bmatrix} \sqrt{2}+1 \\ 1 \end{bmatrix}$

$\lambda_2 = -1$

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

$$\frac{1}{\sqrt{2}}(x_1 + x_2) = -x_1 \Rightarrow x_2 = -\sqrt{2}x_1 - x_1 \Rightarrow x_2 + (1+\sqrt{2})x_1 = 0$$

$$\frac{1}{\sqrt{2}}(x_1 - x_2) = -x_2 \Rightarrow x_1 = -\sqrt{2}x_2 + x_2 \Rightarrow x_1 + (-1+\sqrt{2})x_2 = 0$$

$x_2 = 1$

$x_1 = 1 - \sqrt{2}$ $\quad b = \begin{bmatrix} 1-\sqrt{2} \\ 1 \end{bmatrix}$

ii. $\quad Ax = \lambda x$

$$\|Ax\|^2 = \|\lambda x\|^2$$

$$(Ax)^T (Ax) = \|\lambda x\|^2$$

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

$$x^T I x = |\lambda|^2 \|x\|^2$$

$$\cancel{\|x\|^2} = |\lambda|^2 \cancel{\|x\|^2}$$

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

$$|\lambda| = 1$$

iii. $Ax_1 = \lambda_1 x_1 \quad Ax_2 = \lambda_2 x_2$

$x_1^T x_2 = 0$

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

$$x_1^T x_2 = \lambda_1 \lambda_2 x_1^T x_2$$

$$\lambda_1 \lambda_2 x_1^T x_2 - x_1^T x_2 = 0$$

$$x_1^T x_2 (\lambda_1 \lambda_2 - 1) = 0$$

$\Rightarrow x_1^T x_2 = 0$ since $|\lambda_1| = 1$ and $|\lambda_2| = 1$ and they are distinct so $\lambda_1 = 1$ $\lambda_2 = -1$

iv. A either rotates x or reflects x, or does some combination of both.

b.

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

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

ii. Singular values of A are the square roots of the eigenvalues of A^TA and AA^T

c.

i. False, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ has eigenvalues 1, 1 ii $Ax_1 = \lambda_1 x_1$

$Ax_2 = \lambda_2 x_2$

$A(x_1 + x_2) = Ax_1 + Ax_2 = \lambda_1 x_1 + \lambda_2 x_2$

$\boxed{\text{False}}$, impossible to rewrite in the form $Ax = \lambda x$ unless $\lambda_1 = \lambda_2$

iii. $x^T A x \geq 0$

since we know $Ax = \lambda x$

$x^T \lambda x \geq 0$

$\lambda \|x\|^2 \geq 0$

since $\|x\|^2 > 0 \Rightarrow \lambda > 0$

$\boxed{\text{True}}$

iv. I_2 has 1 distinct eigenvalue: 1, but its rank is 2.

$\boxed{\text{True}}$

v. $Ax_1 = \lambda_1 x_1$

$Ax_2 = \lambda_2 x_2$

$A(x_1 + x_2) = \lambda_1 x_1 + \lambda_2 x_2 \qquad \lambda_1 = \lambda_2 = \lambda$

$A(x_1 + x_2) = \lambda x_1 + \lambda x_2$

$= \lambda(x_1 + x_2)$

$\Rightarrow x_1 + x_2$ is an eigenvector

IF $\lambda_1 = \lambda_2$

True

2.

a). Let $x = H50 \quad y = H60$

i. h = heads, t = tails

$$p(t|x) = 0.5 \quad p(t|y) = 0.6$$

$$p(x) = 0.5 \quad p(y) = 0.5$$

$$p(x|t) = \frac{p(t|x) \cdot p(x)}{p(t)}$$

$$= \frac{p(t|x) \cdot p(x)}{p(t|x)p(x) + p(t|y)p(y)}$$

$$= \frac{(0.5)(0.5)}{(0.5)(0.5) + (0.4)(0.5)}$$

$$= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{2}{10}} = \frac{\frac{1}{4}}{\frac{5}{20} + \frac{4}{20}} = \frac{\frac{1}{\cancel{4}}}{\frac{9}{\cancel{20}\,5}}$$

$$= \frac{5}{9}$$

ii. Let $a = THHH$

$$p(a|x) = (0.5)^4 = 0.0625$$

$$p(\ |y) = (0.4)^1(0.6)^3 = 0.0864$$

$$p(x|a) = \frac{p(a|x)\,p(x)}{}$$

$$p(x|a) = \frac{\quad}{p(a)}$$

$$= \frac{p(a|x)\,p(x)}{p(a|x)p(x) + p(a|y)p(y)}$$

$$= \frac{(0.0625)(0.5)}{(0.0625)(0.5) + (0.0864)(0.5)}$$

$$= \frac{0.0625}{0.0625 + 0.0864} = \boxed{0.42}$$

iii Let b = Heads 9 times

z = H55

$$p(b|x) = (0.5)^{10}$$
$$p(b|y) = (0.4)^1 (0.6)^9$$
$$p(b|z) = (0.45)^1 (0.55)^9$$

$$p(x|b) = \frac{(0.5)^{10}}{(0.4)^1(0.6)^9 + (0.45)^1(0.55)^9 + (0.5)^{10}}$$

$$p(y|b) = \frac{(0.4)^1(0.6)^9}{(0.4)^1(0.6)^9 + (0.45)^1(0.55)^9 + (0.5)^{10}}$$

$$p(z|b) = \frac{(0.45)^1(0.55)^9}{(0.4)^1(0.6)^9 + (0.45)^1(0.55)^9 + (0.5)^{10}}$$

b. X = pregnant Y = positive test

$$p(x|y) = \frac{p(y|x)\,p(x)}{p(y)}$$

$$= \frac{p(y|x)\,p(x)}{p(y|x)p(x) + p(y|\neg x)p(\neg x)}$$

$$= \frac{(0.99)(0.01)}{(0.99)(0.01) + (0.1)(0.99)}$$

$$= \boxed{0.091}$$

There is a relatively high number of nonpregnant women and false positive rate is relatively high.

c. $E(Ax+b) = A\,E(x) + b$

d. $$\mathrm{Cov}(x) = E\left((x - Ex)(x - Ex)^T\right)$$

$$\mathrm{cov}(Ax+b) = E\left((Ax+b - E(Ax+b))(Ax+b - E(Ax+b))^T\right)$$

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

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

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

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

3.

a) $f = x^T A y = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & & a_{1m} \\ & & \\ a_{n1} & & a_{nm} \end{bmatrix} \begin{bmatrix} y_1 \\ \\ y_m \end{bmatrix}$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} x_i a_{ij} y_j$$

$$\nabla_x f = \begin{bmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} a_{1j} y_j \\ \vdots \\ \sum_{j=1}^{m} a_{nj} y_j \end{bmatrix} = \boxed{Ay}$$

b) $f = x^T A y = \sum_{j=1}^{m} \sum_{i=1}^{n} x_i a_{ij} y_j$

$$\nabla_y f = \begin{bmatrix} \partial f / \partial y_1 \\ \vdots \\ \partial f / \partial y_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i a_{i1} \\ \vdots \\ \sum_{i=1}^{n} x_i a_{im} \end{bmatrix} = \boxed{A^T x}$$

c) $f = x^T A y = \sum_{j=1}^{m} \sum_{i=1}^{n} x_i a_{ij} y_j$

$$\nabla_A f = \begin{bmatrix} \partial f / \partial a_{11} & \cdots & \partial f / \partial a_{1m} \\ \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_m \\ \vdots & & \vdots \end{bmatrix}$$

$$\begin{bmatrix} \cdots & & \cdots \\ \partial f/\partial a_{n1} & & \partial f/\partial a_{nm} \end{bmatrix} \begin{bmatrix} & \vdots & \\ x_n y_1 & & x_n y_m \end{bmatrix}$$

$$= \boxed{x y^T}$$

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

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

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

e) $f = tr(AB)$ $\quad A \in \mathbb{R}^{n \times m} \quad B \in \mathbb{R}^{m \times n}$

$$= tr\left(\begin{bmatrix} \sum_{i=1}^{m} a_{1i} b_{i1} & \cdots & \sum_{i=1}^{m} a_{1i} b_{in} \\ \vdots & & \vdots \\ \sum_{i=1}^{m} a_{ni} b_{i1} & \cdots & \sum_{i=1}^{m} a_{ni} b_{in} \end{bmatrix}\right)$$

$$f = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ji} b_{ij}$$

$$\nabla_A f = \begin{bmatrix} \partial f/\partial a_{11} & \cdots & \partial f/\partial a_{1m} \\ \vdots & & \vdots \\ \partial f/\partial a_{n1} & \cdots & \partial f/\partial a_{nm} \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & & \vdots \\ b_{1n} & & b_{mn} \end{bmatrix} = \boxed{B^T}$$

4.

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

$$= \frac{1}{2}(Y - XW)^T (Y - XW)$$

$$= \frac{1}{2}(Y^T Y - 2Y^T XW + W^T X^T X W)$$

$$\frac{\partial L}{\partial W} = \frac{1}{2}(0 - 2X^T Y + 2X^T X W)$$

$$= -X^T Y + X^T X W$$

$$-X^T Y + X^T X W = 0$$

$$X^T X W = X^T Y$$

$$\boxed{W = (X^T X)^{-1} X^T Y}$$