CSE 446: HW 0
Eric Boris: 1976637
Collaborators: Brittan Robinett, Micah Witthaus

1.

We are given the probability of having a certain disease $P(x) = 0.0001$. From this we know the probability of not having the disease $P(\overline{x}) = 0.9999$. We are given the probability of testing positive given that we have the disease $P(y|x) = 0.99$ and we are given the probability of testing negative given that we don't have the disease $P(\overline{y}|\overline{x}) = 0.99$. From these we know the probability of testing positive given that we don't have the disease $P(y|\overline{x}) = 0.01$ and the probability of testing negative given that have the disease $P(\overline{y}|x) = 0.01$. We use Bayes' Rule to determine our probability of having the disease given that we test positive $P(x|y)$.

$$\begin{aligned}
P(x|y) &= \frac{P(x,y)}{P(y)} && \text{Bayes' Rule} \\
&= \frac{P(x)P(y|x)}{P(\overline{x})P(y|\overline{x}) + P(x)P(y|x)} \\
&= \frac{(0.0001)(0.99)}{(0.9999)(0.01) + (0.0001)(0.99)} && \text{Substitution} \\
&= \frac{1}{102} \\
&\approx 0.0098
\end{aligned}$$

2.

a.

Show that if $E[Y|X = x] = x$ that $Cov(X, Y) = E[(X - E[X])^2]$.

$$\begin{aligned}
Cov(X, Y) &= E[X, Y] - E[X]E[Y] \\
&= \sum_x P(X = x)E[XY|X] \\
&= \sum_x P(X = x)xE[Y|X] \\
&= \sum_x P(X = x)x^2 && \text{Law of total expectation} \\
&= E[X^2] - E[X]E[Y] \\
&= E[X^2] - E[X]\sum_x P(X = x)E[Y|X = x] \\
&= E[X^2] - E[X]\sum_x P(X = x)x \\
&= E[X^2] - E[X]E[X] && \text{Law of total expectation} \\
&= E[X^2] - E[X]^2 \\
&= E[(X - E[X])^2]
\end{aligned}$$

b.

Show that if X and Y are independent that $Cov(X,Y) = 0$.

$$\begin{aligned}
Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\
&= E[XY] - E[X]E[Y] \\
&= \sum_{x,y} P(X = x, Y = y)xy - E[X]E[Y] \\
&= \sum_{x,y} P(X = x)P(Y = y)xy - E[X]E[Y] \\
&= (\sum_{x} P(X = x)x)(\sum_{y} P(Y = y)y) - E[X]E[Y] \\
&= E[X]E[Y] - E[X]E[Y] \\
&= 0
\end{aligned}$$

3.

a.

Show that $h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)\,dx$.

$$\begin{aligned}
H(z) &= P(Z \leq z) && \text{Let H() be the CDF of Z} \\
H(z) &= P(X + Y \leq z) \\
H(z) &= \int\int_{x+y\leq z} F_{XY}(x,y)\,dx\,dy \\
H(z) &= \int_{-\infty}^{\infty} (\int_{-\infty}^{z-x} f(x)f(y)\,dy)\,dx \\
H(z) &= \int_{-\infty}^{\infty} f(x)(\int_{-\infty}^{z-x} g(y)\,dy)\,dx \\
H(z) &= \int_{-\infty}^{\infty} f(x)G(z-x)\,dx && \text{Let G() be the CDF of Y} \\
\frac{d}{dz}H(z) &= \frac{d}{dz}(\int_{-\infty}^{\infty} f(x)G(z-x)\,dx) \\
h(z) &= \int_{-\infty}^{\infty} f(x)\frac{d}{dz}(G(z-x))\,dx \\
h(z) &= \int_{-\infty}^{\infty} f(x)g(z-x)\,dx
\end{aligned}$$

b.

Simplify h(z).

$$\begin{aligned}
h(z) &= \int_{-\infty}^{\infty} f(x)g(z-x)\,dx && \text{From a.} \\
&= \int_{0}^{1} g(z-x)\,dx && \text{Since f(x)=0 for x<0 and x>1 and f(x)=1 for x} \in \text{[0,1]}
\end{aligned}$$

There are four cases to consider:

1. $z < 0$: When $z < 0$ then $g(z - x) = 0$ and thus $h(z) = 0$.
2. $0 \leq z \leq 1$: When $0 \leq z \leq 1$ then $h(z) = \int_0^z g(z - x)\,dx$ and thus $h(z) = z$.
3. $1 < z \leq 2$: When $1 < z \leq 2$ then $h(z) = \int_{z-1}^1 g(z - x)\,dx$ and thus $h(z) = 2 - z$.
4. $z > 2$: When $z > 2$ then $g(z - x) = 0$ and thus $h(z) = 0$.

Thus:

$$h(z) = \begin{cases} z & \text{if } 0 \leq z \leq 1 \\ 2 - z & \text{if } 1 < z \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

4.

Solve for a.

$$\begin{aligned} Y &= aX + b \\ Var(Y) &= Var(aX + b) \\ Var(Y) &= a^2 Var(X) \\ 1 &= a^2 Var(X) \\ 1 &= a^2 \sigma^2 \\ 1 &= \pm a\sigma \\ a &= \pm \frac{1}{\sigma} \end{aligned}$$

Solve for b.

$$\begin{aligned} Y &= aX + b \\ E[Y] &= E[aX + b] \\ E[Y] &= aE[X] + b \\ 0 &= aE[X] + b \\ 0 &= a\mu + b \\ b &= -a\mu \\ b &= -(\pm \frac{1}{\sigma})\mu \\ b &= \pm \frac{\mu}{\sigma} \end{aligned}$$

5.

Find $E[\sqrt{n}(\mu_n - \mu)]$.

$$\begin{aligned}
E[\sqrt{n}(\mu_n - \mu)] &= E[\sqrt{n}(\mu_n - \mu)] \\
&= \sqrt{n}E[(\mu_n - \mu)] && \text{By Linearity of Expectation} \\
&= \sqrt{n}(E[\mu_n] - E[\mu]) \\
&= \sqrt{n}(E[\mu_n] - \mu) \\
&= \sqrt{n}(E[(\frac{1}{n}\sum_{i=1}^{n} x_i)] - \mu) \\
&= \sqrt{n}((\frac{1}{n}\sum_{i=1}^{n} E[x_i]) - \mu) \\
&= \sqrt{n}((\frac{1}{n}\sum_{i=1}^{n} \mu) - \mu) \\
&= \sqrt{n}((\frac{1}{n}n\mu) - \mu) \\
&= \sqrt{n}(\mu - \mu) \\
&= \sqrt{n}(0) \\
&= 0
\end{aligned}$$

Find $Var(\sqrt{n}(\mu_n - \mu))$.

$$\begin{aligned}
Var(\sqrt{n}(\mu_n - \mu)) &= Var(\sqrt{n}(\mu_n - \mu)) \\
&= (\sqrt{n})^2 Var(\mu_n - \mu) \\
&= n(Var(\mu_n) - Var(\mu)) \\
&= n(Var(\mu_n) - 0 \\
&= nVar(\frac{1}{n}\sum_{i=1}^{n} x_i) \\
&= n\frac{1}{n^2}\sum_{i=1}^{n} Var(x_i) \\
&= n\frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 \\
&= n\frac{1}{n^2}n\sigma^2 \\
&= \sigma^2
\end{aligned}$$

6.

a.

Find $E[\hat{F}_n(x)]$.

$$
\begin{aligned}
E[\hat{F}_n(x)] &= E[\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \leq x)] \\
&= \frac{1}{n}\sum_{i=1}^{n}E[\mathbf{1}(X_i \leq x)] \\
&= \frac{1}{n}\sum_{i=1}^{n}F(x) && X_i \text{ is i.i.d for } i \in [1,n] \\
&= \frac{1}{n}nF(x) \\
&= F(x)
\end{aligned}
$$

b.

Show $Var(\hat{F}_n(x)) = \frac{F(x)(1-F(x))}{n}$.

$$
\begin{aligned}
Var(\hat{F}_n(x)) &= Var(\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \leq x)) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}Var(\mathbf{1}(X_i \leq x)) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}(E[(\mathbf{1}(X_i \leq x))^2] - (E[\mathbf{1}(X_i \leq x)])^2) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}(E[\mathbf{1}(X_i \leq x)] - (E[\mathbf{1}(X_i \leq x)])^2) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}(F(x) - (F(x))^2) && X_i \text{ is i.i.d for } i \in [1,n] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}F(x)(1-F(x)) \\
&= \frac{1}{n^2}nF(x)(1-F(x)) \\
&= \frac{F(x)(1-F(x))}{n}
\end{aligned}
$$

c.

Show that for all $x \in \mathbf{R}$ that $E[(\hat{F}_n(x) - F(x))^2] \leq \frac{1}{4n}$. Suppose $x \in \mathbf{R}$. We know from b.:

$$
\begin{aligned}
E[(\hat{F}_n(x) - F(x))^2] &= Var(\hat{F}_n(x)) \\
&= \frac{F(x)(1-F(x)}{n} \\
&= \frac{F(x) - F(x)^2}{n}
\end{aligned}
$$

We know $max\{\frac{F(x)-F(x)^2}{n}\}$ occurs at $max\{F(x)-F(x)^2\}$. To find $max\{F(x)-F(x)^2\}$ we define $g(F(x)) = F(x) - F(x)^2$, and take $g'(F(x))$ and $g''(F(x))$.

$$\begin{aligned}
g(F(x)) &= F(x) - F(x)^2 \\
g'(F(x)) &= 1 - 2F(x) \\
g''(F(x)) &= -2
\end{aligned}$$

From $g'(F(x))$ we see a convexity at $F(x) = \frac{1}{2}$ and from $g''(F(x))$ we see that that convexity is a maximum. Therefore $max\{F(x)-F(x)^2\} = (\frac{1}{2}) - (\frac{1}{2})^4 = \frac{1}{4}$ and $max\{\frac{F(x)-F(x)^2}{n}\} = \frac{(\frac{1}{4})}{n} = \frac{1}{4n}$. Thus, $E[(\hat{F}_n(x) - F(x))^2] \leq \frac{1}{4n}$ since $max\{\frac{F(x)-F(x)^2}{n}\} = E[(\hat{F}_n(x) - F(x))^2]$.

7.

Let H(x) be the be the CDF of Y. Find H(x).

$$\begin{aligned}
H(x) &= P(maxX_i \leq x) \\
&= \prod_{i=1}^{n} P(X_i \leq x) \\
&= x^n
\end{aligned}$$

Let h(x) be the PDF of Y. Find h(x).

$$\begin{aligned}
h(x) &= \frac{d}{dx}H(x) \\
&= \frac{d}{dx}(x^n) \\
&= nx^{n-1}
\end{aligned}$$

Find E[Y].

$$\begin{aligned}
E[Y] &= E[max\{X_1, ..., X_n\}] \\
&= \int_0^1 xh(x)\,dx \\
&= \int_0^1 x(nx^{n-1})\,dx \\
&= \int_0^1 nx^n\,dx \\
&= n\int_0^1 x^n\,dx \\
&= n[\frac{x^{n+1}}{n+1}|_0^1] \\
&= n[\frac{1}{n+1}] \\
&= \frac{n}{n+1}
\end{aligned}$$

8.

Given a random variable X with $E[X] = \mu$ and $E[(X-\mu)^2] = \sigma^2$. For any $x \geq 0$ show that $P(X \geq \mu + \sigma x) \leq \frac{1}{x^2}$.

$$\begin{aligned}
P(X \geq x) &\leq \frac{E[X]}{x} \\
P(X^2 \geq x^2) &\leq \frac{E[X^2]}{x^2} \\
P((x-\mu)^2 \geq x^2) &\leq \frac{E[(x-\mu)^2}{x^2} \\
P(|x-\mu| \geq x) &\leq \frac{E[(x-\mu)^2}{x^2} \\
P(|x-\mu| \geq x) &\leq \frac{\sigma^2}{x^2} \\
P(|x-\mu| \geq \sigma x) &\leq \frac{1}{x^2} \\
P(|x-\mu| \geq \sigma x) &= P(x \geq \mu + \sigma x) + P(x \geq \mu - \sigma x) \\
P(|x-\mu| \geq \sigma x) &\geq P(x \geq \mu + \sigma x) \\
P(|x-\mu| \leq \sigma x) &\leq \frac{1}{x^2}
\end{aligned}$$

$$P(x \leq \mu + \sigma x) \leq \frac{1}{x^2}$$

9.

a.

Write A in row echelon form.

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 2 \\ 1 & 1 & 2 \end{bmatrix} && \text{R2 - R1} \Rightarrow \text{R2} \\
&= \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{bmatrix} && \text{R3 - R1} \Rightarrow \text{R3} \\
&= \begin{bmatrix} 1 & 0 & 3 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{bmatrix} && \text{R1 + R2} \Rightarrow \text{R1} \\
&= \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix} && \tfrac{R2}{2} \Rightarrow \text{R2} \\
&= \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} && \text{R2 + R3} \Rightarrow \text{R3}
\end{aligned}$$

Since there are two nonzero rows in ref(A), rank(A)=2.

Write B in row echelon from.

$$\begin{aligned}
B &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -2 \\ 1 & 1 & 2 \end{bmatrix} && \text{R2 - R1} \Rightarrow \text{R2} \\
&= \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -2 \\ 0 & -1 & -1 \end{bmatrix} && \text{R3 - R1} \Rightarrow \text{R3} \\
&= \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} && \tfrac{R2}{2} \Rightarrow \text{R2} \\
&= \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} && \text{R1 - R2} \Rightarrow \text{R3}
\end{aligned}$$

Since there are two nonzero rows in ref(B), rank(B)=2.

b.

Since there are two pivot columns in ref(A), basis(A)=2.
Since there are two pivot columns in ref(B), basis(B)=2.

10.

a.

$$\begin{aligned}
Ac &= \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1*0+1*2+1*4 \\ 1*2+1*4+1*2 \\ 1*3+1*3+1*1 \end{bmatrix} \\
&= \begin{bmatrix} 6 \\ 8 \\ 7 \end{bmatrix}
\end{aligned}$$

b.

Find A^{-1}.

$$A^{-1} = \left[\begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 1 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} 2 & 4 & 2 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \qquad \text{R1} \Leftrightarrow \text{R2}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \qquad \frac{R1}{2} \Rightarrow \text{R1}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 0 & -3 & -2 & 0 & \frac{-3}{2} & 1 \end{array}\right] \qquad \text{R3 - 3R1} \Rightarrow \text{R3}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 0 & -3 & -2 & 0 & \frac{-3}{2} & 1 \end{array}\right] \qquad \text{R1 - R2} \Rightarrow \text{R1}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & -3 & -2 & 0 & \frac{-3}{2} & 1 \end{array}\right] \qquad \frac{R2}{2} \Rightarrow \text{R1}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 4 & \frac{3}{2} & \frac{-3}{2} & 1 \end{array}\right] \qquad \text{3R2 + R3} \Rightarrow \text{R3}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{8} & \frac{-3}{8} & \frac{1}{4} \end{array}\right] \qquad \frac{R3}{4} \Rightarrow \text{R3}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{8} & \frac{-5}{8} & \frac{3}{4} \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{8} & \frac{-3}{8} & \frac{1}{4} \end{array}\right] \qquad \text{3R3 + R1} \Rightarrow \text{R1}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{8} & \frac{-5}{8} & \frac{3}{4} \\ 0 & 1 & 0 & \frac{-1}{4} & \frac{3}{4} & \frac{-1}{2} \\ 0 & 0 & 1 & \frac{3}{8} & \frac{-3}{8} & \frac{1}{4} \end{array}\right] \qquad \text{-2R3 + R2} \Rightarrow \text{R2}$$

$$= \begin{bmatrix} \frac{1}{8} & \frac{-5}{8} & \frac{3}{4} \\ \frac{-1}{4} & \frac{3}{4} & \frac{-1}{2} \\ \frac{3}{8} & \frac{-3}{8} & \frac{1}{4} \end{bmatrix}$$

Find $\mathrm{x} = \mathrm{A}^{-1}\mathrm{B}$.

$$\begin{aligned} x &= A^{-1}B \\ &= \begin{bmatrix} \frac{1}{8} & \frac{-5}{8} & \frac{3}{4} \\ \frac{-1}{4} & \frac{3}{4} & \frac{-1}{2} \\ \frac{3}{8} & \frac{-3}{8} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} -2 \\ -2 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} -2 * \frac{1}{8} - 2 * \frac{-5}{8} - 4 * \frac{3}{4} \\ -2 * \frac{-1}{4} - 2 * \frac{3}{4} - 4 * \frac{-1}{2} \\ -2 * \frac{3}{8} - 2 * \frac{-3}{8} - 4 * \frac{1}{4} \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

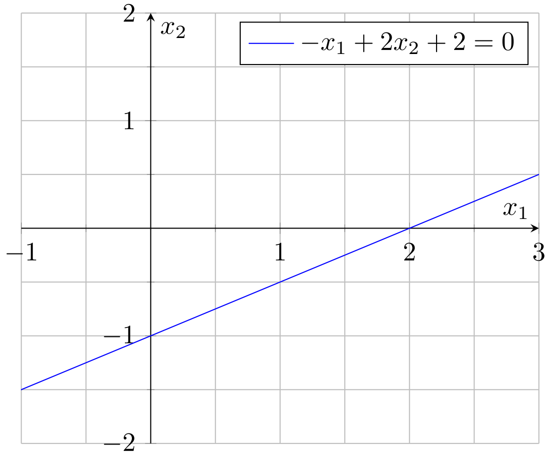
11.

a.

Solve for x_2.

$$
\begin{aligned}
Wx + b &= 0 \\
\begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 2 &= 0 \\
-x_1 + 2x_2 + 2 &= 0 \\
-x_1 + 2x_2 &= -2 \\
2x_2 &= x_1 - 2 \\
x_2 &= \frac{1}{2}x_1 - 1
\end{aligned}
$$

Plot the hyperplane.

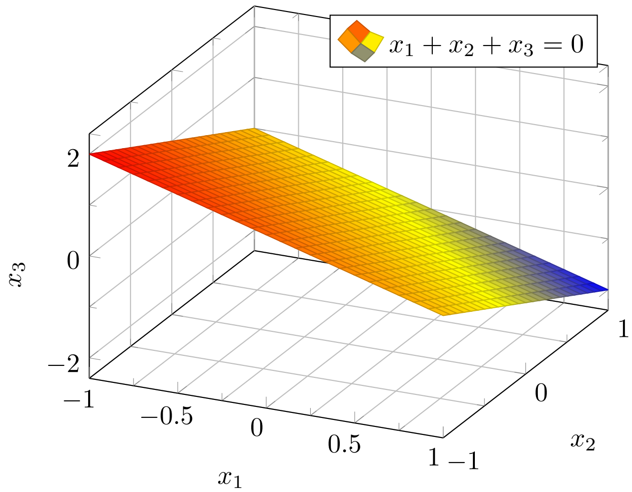


b.

Solve for x_3.

$$
\begin{aligned}
Wx + b &= 0 \\
\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 0 &= 0 \\
x_1 + x_2 + x_3 &= 0 \\
x_3 = -x_1 - x_2
\end{aligned}
$$

Plot the hyperplane.



c.

Find $\min_x ||x_0 - x||^2$ such that $w^T x + b = 0$. Let $\tilde{x}_0$ be the minimizer to the problem.

$$
\begin{aligned}
||x_0 - \tilde{x}_0|| &= |\frac{w^T(x_0 - \tilde{x}_0)}{||w||}| \\
||x_0 - \tilde{x}_0||^2 &= |\frac{w^T(x_0 - \tilde{x}_0)}{||w||}|^2 \\
&= |\frac{w^T x_0 + b}{||w||}|^2
\end{aligned}
$$

12.

a.

$$
\begin{aligned}
f(x,y) &= x^T A x + y^T B x + c \\
&= [x_1, ..., x_n] \begin{bmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} + [y_1, ..., y_n] \begin{bmatrix} b_{11} & ... & b_{1n} \\ ... & ... & ... \\ b_{n1} & ... & b_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} + c \\
&= [\sum_{i=1}^n x_i A_{i1}, ..., \sum_{i=1}^n x_i A_{in}] \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} + [\sum_{i=1}^n y_i B_{i1}, ..., \sum_{i=1}^n y_i B_{in}] \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} + c \\
&= \sum_{j=1}^n x_j [\sum_{i=1}^n x_i A_{ij}] + \sum_{j=1}^n x_j [\sum_{i=1}^n y_i B_{ij}] + c \\
&= \sum_{j=1}^n x_j [\sum_{i=1}^n x_i A_{ij} + \sum_{i=1}^n y_i B_{ij}] + c \\
&= \sum_{j=1}^n x_j [\sum_{i=1}^n [x_i A_{ij} + y_i B_{ij}]] + c
\end{aligned}
$$

b.

Find $\frac{\partial f(x,y)}{\partial x_k}$ for $1 \leq k \leq n$.

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x_k} &= \frac{\partial}{\partial x_k}(\sum_{j=1}^{n} x_j[\sum_{i=1}^{n}[x_i A_{ij} + y_i B_{ij}]] + c) \\
&= \frac{\partial}{\partial x_k}\sum_{j=1}^{n} x_j[\sum_{i=1}^{n}[x_i A_{ij} + y_i B_{ij}]] \\
&= \sum_{j=1}^{n}(\frac{\partial}{\partial x_k}(x_j)(\sum_{i=1}^{n}(x_i A_{ij} + y_i B_{ij}))) + (\sum_{j=i}^{n} x_j(\frac{\partial}{\partial x_k}(\sum_{i=1}^{n}(x_i A_{ij} + y_i B_{ij})))) \\
&= (0 + ... + 0 + \frac{\partial}{\partial x_k}(x_k)(\sum_{i=1}^{n}(x_i A_{ik} + y_i B_{ik})) + 0 + ... + 0) + (\sum_{j=i}^{n} x_j(\sum_{i=1}^{n}\frac{\partial}{\partial x_x}(x_i A_{ij} + y_i B_{ij}))) \\
&= \frac{\partial}{\partial x_k}(x_k)(\sum_{i=1}^{n}(x_i A_{ik} + y_i B_{ik})) + (\sum_{j=1}^{n} x_j(\sum_{i=1}^{n}\frac{\partial}{\partial x_k}(x_i A_{ij}) + 0)) \\
&= \frac{\partial}{\partial x_k}(x_k)(\sum_{i=1}^{n}(x_i A_{ik} + y_i B_{ik})) + (\sum_{j=1}^{n} x_j(0 + ... + 0 + \frac{\partial}{\partial x_k}(x_k A_{kj}) + 0 + ... + 0)) \\
&= \sum_{i=1}^{n}(x_i A_{ik} + y_i B_{ik}) + \sum_{j=1}^{n} x_j A_{kj}
\end{aligned}$$

Express $\nabla_x f(x,y)$ in index notation.

$$\begin{aligned}
\nabla_x f(x,y) &= [\frac{\partial f(x,y)}{\partial x_1}, ..., \frac{\partial f(x,y)}{\partial x_n}] \\
&= [(\sum_{i=1}^{n}(x_i A_{i1} + y_i B_{i1}) + \sum_{j=1}^{n} x_j A_{1j}), ..., (\sum_{i=1}^{n}(x_i A_{in} + y_i B_{in}) + \sum_{j=1}^{n} x_j A_{nj})]
\end{aligned}$$

Express $\nabla_x f(x,y)$ in vector notation.

$$\nabla_x f(x,y) = (A + A^T)x + B^T y$$

c.

Find $\frac{\partial f(x,y)}{\partial y_k}$ for $1 \leq k \leq n$.

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y_k} &= \frac{\partial}{\partial y_k}(\sum_{j=1}^{n} x_j[\sum_{i=1}^{n}[x_i A_{ij} + y_i B_{ij}]] + c) \\
&= \frac{\partial}{\partial y_k}\sum_{j=1}^{n} x_j[\sum_{i=1}^{n}[x_i A_{ij} + y_i B_{ij}]] + 0 \\
&= (\sum_{j=1}^{n} \frac{\partial}{\partial y_k}(x_j)(\sum_{i=1}^{n}(x_i A_{ij} + y_i B_{ij}))) + (\sum_{j=1}^{n} x_j(\frac{\partial}{\partial y_k}\sum_{i=1}^{n}(x_i A_{ij} + y_i B_{ij}))) \\
&= 0 + (\sum_{j=1}^{n} x_j(\sum_{i=1}^{n} \frac{\partial}{\partial y_k}(x_i A_{ij} + y_i B_{ij})) \\
&= \sum_{j=1}^{n} x_j(\sum_{i=1}^{n}(0 + \frac{\partial}{\partial y_k}(y_i B_{ij}))) \\
&= \sum_{j=1}^{n} x_j(\sum_{i=1}^{n}(0 + ... + 0 + \frac{\partial}{\partial y_k}(y_k B_{kj}))) \\
&= \sum_{j=1}^{n} x_j B_{kj}
\end{aligned}$$

Express $\nabla_y f(x,y)$ in index notation.

$$\begin{aligned}
\nabla_y f(x,y) &= [\frac{\partial f(x,y)}{\partial y_1}, \frac{\partial f(x,y)}{\partial y_2}, ..., \frac{\partial f(x,y)}{\partial y_n}] \\
&= [(\sum_{j=1}^{n} x_j B_{1j}), (\sum_{j=1}^{n} x_j B_{2j}), ..., (\sum_{j=1}^{n} x_j B_{nj})]
\end{aligned}$$

Express $\nabla_y f(x,y)$ in index notation.

$$\nabla_y f(x,y) = Bx$$

13.

Show Tr(AB) = T(BA).

$$\begin{aligned}
Tr(AB) &= Tr(AB) \\
&= \sum_{i=1}^{n}(AB)_{ii} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} b_{ji} \\
&= \sum_{j=1}^{m}\sum_{i=1}^{n} b_{ji} a_{ij} \\
&= \sum_{j=1}^{m}(BA)_{jj} \\
&= Tr(BA)
\end{aligned}$$

14.

a.

Find the minimum and maximum rank of $\sum_{i=1}^{n} v_i v_i^T$. Since v_i is non-zero for all $1 \leq i \leq n$ the minimum rank of $\sum_{i=1}^{n} v_i v_i^T$ is 1. The maximum rank is dependent on the smallest dimesion after summation, or $min\{n, d\}$.

b.

Find the minimum and maximum rank of V. The answer to this is clearly the same as the above and for the same reasons. Thus, $rank_{min}(V) = 1$. and $rank_{max}(V) = min\{n, d\}$.

c.

Find the minimum and maximum rank of $\sum_{i=1}^{n}(Av_i)(Av_i)^T$. The minimum rank is 0 since A could be a zero matrix. The maximum rank is $min\{n, d\}$.

d.

Find the minimum and maximum rank of AV. The answer to this is the same as above. Thus, $rank_{min}(AV) = 0$ and $rank_{max}(AV) = min\{n, d\}$. And if V is rank d? Then the minimum rank remains the same, $rank_{min}(AV) = 0$, and since $d \leq n$, $rank_{max}(AV) = d$.

15.

a. and b.

Code:

```
import numpy as np

A = np.matrix([[0, 2, 4], [2, 4, 2], [3, 3, 1]])
b = np.array([[-2], [-2], [-4]])
c = np.array([[1], [1], [1]])

print('a. A^{-1} = \n', A.getI(), '\n')
print('b. A^{-1} * b = \n', A.getI() * b, '\n')
print('b. A * c = \n', A * c)
```

Output:

```
a. A^{-1} =
 [[ 0.125 -0.625  0.75 ]
 [-0.25   0.75  -0.5  ]
 [ 0.375 -0.375  0.25 ]]

b. A^{-1} * b =
 [[-2.]
 [ 1.]
 [-1.]]

b. A * c =
 [[6]
 [8]
 [7]]
```

16.

Solve for n.

$$\sqrt{E[(\hat{F}_n(x) - F(x))^2]} \leq \frac{1}{400}$$

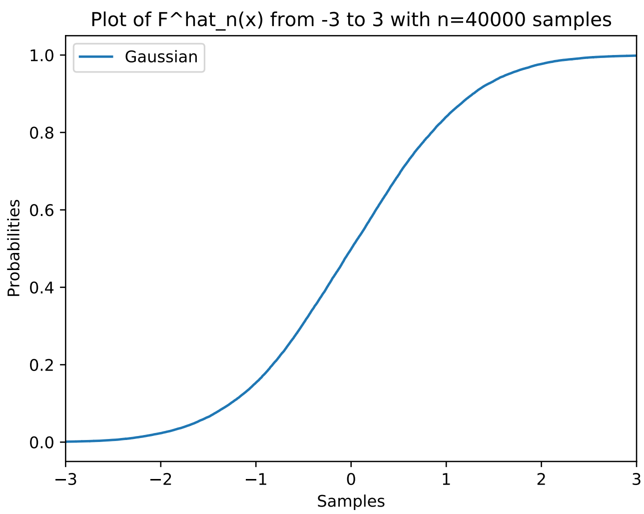
$$\sqrt{\frac{1}{4n}} = \frac{1}{400} \qquad \text{Since } max\{E(\hat{F}_n(x) - F(x))^2\} = \frac{1}{4n}$$

$$\frac{1}{4n} = \frac{1}{160000}$$

$$\frac{1}{n} = \frac{1}{40000}$$

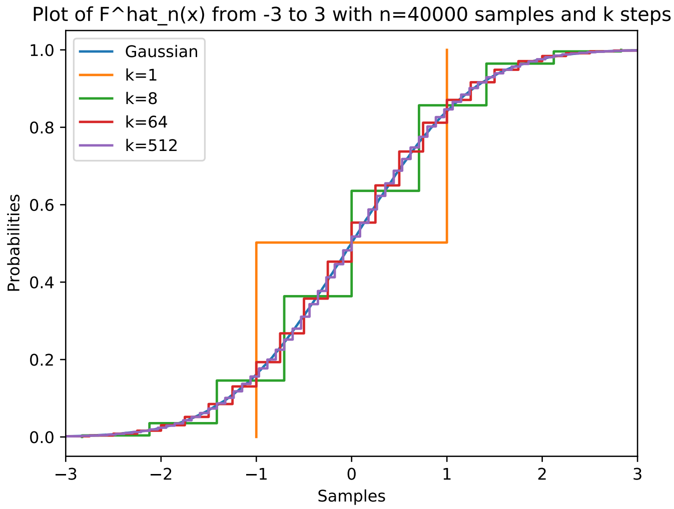
$$n = 40000$$

a.




```
import numpy as np
import matplotlib.pyplot as plt

n = 40000
Z = np.random.randn(n)

# Draw the graph.
plt.step(sorted(Z), np.arange(1, n+1) / float(n), label='Gaussian')

axes = plt.gca()
axes.set_xlim((-3, 3))

plt.title('Plot of F^hat_n(x) from -3 to 3 with n=40000 samples')
plt.legend()
plt.xlabel('Samples')
plt.ylabel('Probabilities')
plt.savefig('q16a.pdf')

```

b.




```python
import numpy as np
import matplotlib.pyplot as plt

n = 40000
Z = np.random.randn(n)

# Draw the same graph from a.
plt.step(sorted(Z), np.arange(1, n+1) / float(n), label='Gaussian')

# Draw the step graphs for b.
for k in [1, 8, 64, 512]:
    Z_k = np.sum(np.sign(np.random.randn(n, k)) * np.sqrt(1.0 / k), axis=1)
    plt.step(sorted(Z_k), np.arange(1, n+1) / float(n), label=f'k={k}')

axes = plt.gca()
axes.set_xlim((-3, 3))

plt.title('Plot of F^hat_n(x) from -3 to 3 with n=40000 samples and k steps')
plt.legend()
plt.xlabel('Samples')
plt.ylabel('Probabilities')
plt.savefig('q16b.pdf')
```