Henry Trinh

1. a. Show that $E[y] = E[Ax+b] = AE[x] + b$. In other words, let $y = Ax+b$ be a random vector, show that the expectation is linear.

From lecture 1, expected values differ for discrete and continuous probability distributions (slide 53). Hence, we consider both cases below:

Discrete: $E(f(x)) = \sum_x f(x_i)p(x_i)$. Let $f(x) = y$.

$E[y] = \sum_y (Ax+b)P(y)$, where $P(y)$ is the probability mass function

$= \sum_y AxP(y) + \sum_y bP(y) = A\sum_y xP(y) + b\sum_y P(y)$

Since $\sum_y P(y) = 1$ (sum of all probabilities),

$E[y] = A\sum_y x \cdot P(y) + b$. And, $\sum xP(x) = E(x)$, so

$\therefore \boxed{E[y] = AE[x] + b, \text{ as desired}}$

Continuous: $E(f(x)) = \int_x f(x_i)p(x_i)$

Then, $E[y] = \int_{-\infty}^{\infty} y \cdot f(y)\,dy$, where $f(y)$ is the prob. density function

Since $y = Ax+b$, by substitution:

$E[y] = \int_{-\infty}^{\infty} (Ax+b) \cdot f(y)\,dy$

$= \int_{-\infty}^{\infty} Axf(y)\,dy + \int_{-\infty}^{\infty} b \cdot f(y)\,dy$

$= A\int_{-\infty}^{\infty} xf(y)\,dy + b\int_{-\infty}^{\infty} f(y)\,dy.$

Since $\int_{-\infty}^{\infty} f(y)\,dy = 1$ (integral of entire density function),

$E[y] = A\int_{-\infty}^{\infty} x f(y)\,dy + b$. And $E[x] = \int_{-\infty}^{\infty} xf(x)\,dx$, so

$\therefore \boxed{E[y] = AE[x] + b, \text{ as desired}}$

Hence, for all cases, the expectation is linear.

b. Show that $cov[y] = cov[Ax+b] = A\,cov[x]A^T = A\Sigma A^T$.

Slide 55 depicts the covariance matrix, but no corresponding formula. From Kent State University, we are given that:

$$cov(x) = E[(x-E[x])(x-E[x])^T].$$

So, we can now determine the cov. matrix of $y = Ax+b$:

$$cov(y) = E[(y-E[y])(y-E[y])]^T$$

$$= cov(Ax+b) = E[(Ax+b-E[Ax+b])(Ax+b-E[Ax+b])^T]$$

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

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

By the definition of the covariance matrix, this becomes

$$= A \cdot cov(x) \cdot A^T.$$

Since $cov(x) = \Sigma$ (slide 59), we thus see that

$\therefore$ $\boxed{cov(y) = A\Sigma A^T, \text{ as desired}}$

2. We are given dataset $D = \{(x,y)\} = \{(0,1), (2,3), (3,6), (4,8)\}$

a. Determine the least squares estimate $y = \theta^T x$ using Cramer's rule.

From slide 29, we are given the following: ⊙ Find $y = mx + b$

	x_i	y_i	x_i^2	$x_i y_i$
	0	1	0	0
	2	3	4	6
	3	6	9	18
	4	8	16	32
Σ:	9	18	29	56

⊛ $\theta = (X^T X)^{-1} X^T \vec{y}$

Let $x = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$

↑ intercepts

Since $X^T X \theta = X^T y$ (slide 36), we say that $A\theta = B$, where $A = X^T X$ and $B = X^T y$. {θ is the vector of coefficients}.

Then, $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1\cdot1 + 1\cdot1 + 1\cdot1 + 1\cdot1 & 1\cdot0 + 1\cdot2 + 1\cdot3 + 1\cdot4 \\ 0\cdot1 + 2\cdot1 + 3\cdot1 + 4\cdot1 & 0\cdot0 + 2\cdot2 + 3\cdot3 + 4\cdot4 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$

$\uparrow a \quad \uparrow b$

and $x^T y = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix}}_{2\times4} \underbrace{\begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}}_{4\times1} = \underbrace{\begin{bmatrix} 1\cdot1 + 1\cdot3 + 1\cdot6 + 1\cdot8 \\ 0\cdot1 + 2\cdot3 + 3\cdot6 + 4\cdot8 \end{bmatrix}}_{2\times1} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$

$\uparrow c$

Now, we assume that the coefficient matrix is invertible; hence $\det = a_1 b_2 - b_1 a_2$ is non-zero (slide 24). Then, we see that:

$$\theta_0 = \frac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18\cdot29 - 56\cdot9}{4\cdot29 - 9\cdot9} = \frac{18}{35} \quad \Big| \quad \theta_1 = \frac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{4\cdot56 - 18\cdot9}{4\cdot29 - 9\cdot9} = \frac{62}{35}$$

$\therefore \boxed{y = \frac{62}{35}x + \frac{18}{35}}$

b. Use the normal equations, $\theta = (X^T X)^{-1} X^T \vec{y}$ {Slide 28}.

$$\theta = \underbrace{\underbrace{\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1}}_{2\times2} \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix}}_{2\times4} \underbrace{\begin{bmatrix} 1 \\ 3 \\ 4 \\ 8 \end{bmatrix}}_{4\times1}}_{2\times1} = \frac{1}{4\cdot29 - 9\cdot9} \begin{bmatrix} 29 & -9 \\ -9 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \\ 8 \end{bmatrix}$$

$$= \frac{1}{35} \begin{bmatrix} 29\cdot1 + (-9)\cdot0 & 29\cdot1 + (-9)\cdot2 & 29\cdot1 + (-9)\cdot3 & 29\cdot1 + (-9)4 \\ (-9)1 + 0(4) & -9\cdot1 + 4(2) & -9\cdot1 + 4(3) & -9\cdot1 + 4(4) \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \\ 8 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 29 & 11 & 2 & -7 \\ -9 & -1 & 3 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

$$= \frac{1}{35} \begin{bmatrix} 29\cdot1 + 11\cdot3 + 2\cdot6 \cdot (-7)\cdot8 \\ -9\cdot1 + (-1)\cdot3 + 3\cdot6 + 7\cdot8 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 18 \\ 62 \end{bmatrix} \Rightarrow \boxed{\begin{bmatrix} 18/35 \\ 62/35 \end{bmatrix}, \text{as desired}}$$