WORKSHEET
DISCUSSION SECTION 10
NOT TO BE SUBMITTED

(1) This problem explains how to find a line of best fit. Given n data points (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), the *linear least squares fit* is the function $f(x) = mx + b$ that minimizes the sum of squares (error term)

$$E(m, b) = \sum_{j=1}^{n} (y_j - f(x_j))^2.$$

The idea is to solve for m and b that make $E(m, b)$ as small as possible; in this problem $x_1, \ldots, x_n$, and $y_1, \ldots, y_n$ are constants.

(a) Rewrite the formula for $E(m, b)$ in terms of m and b.
(b) Compute the partial derivatives $\partial E/\partial m$ and $\partial E/\partial b$.
(c) Under mild hypotheses, the function $E(m, b)$ attains a minimum. Show that it occurs at m and b satisfying the equations below.

$$m \sum_{j=1}^{n} x_j + bn = \sum_{j=1}^{n} y_j$$

$$m \sum_{j=1}^{n} x_j^2 + b \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} x_j y_j$$

Hint: find the critical point of $E(m, b)$.

Now suppose we are given the two data points $(1, 1)$ and $(2, 3)$ (so $n = 2$).

(d) What is the line of best fit?
(e) Use the equations in (c) to re-derive your answer.

Now suppose we are given four data points $(0, 0)$, $(1, 2)$, $(2, 1)$, and $(3, 3)$.

(f) Plot these points in the xy plane and guess the line of best fit.
(g) Compute the line of best fit using your equations in (c).

(2) Let $f(x, y) = xy$. This problem explores what the mixed partial f_{xy} measures. It may help clarify the meaning of the discriminant.

(a) Compute $f_x(0, n)$ for $n = -2, -1, 0, 1, 2$.

Now consider the vertical traces of $z = xy$ obtained by intersecting $z = xy$ with the planes $y = -2, -1, 0, 1, 2$.

(b) Sketch the tangent lines to these traces at $(0, -2)$, $(0, -1)$, $(0, 0)$, $(0, 1)$, $(0, 2)$. Hint: their slopes are x-partial derivatives of f.
(c) What happens to the slopes of these tangent lines as y increases?
(d) Compute f_{xy} and explain why it confirms your answer in (b).

Conclusion: f_{xy} measures how the tangent lines of traces twist. If they twist enough, then we get a saddle point.

(e) Think about how this problem works for $f(x, y) = (x^3 + x)(y^3 + y)$.

1a. $E(m, b) = \sum_{j=1}^{n} (y_j - (mx_j + b))^2$

b. We use linearity of the derivative (i.e., $(f + cg)' = f' + cg'$) to obtain:

$$\frac{\partial E}{\partial m} = -\sum_{j=1}^{n} 2(y_j - mx_j - b)x_j, \quad \frac{\partial E}{\partial b} = -\sum_{j=1}^{n} 2(y_j - mx_j - b)$$

c. $\frac{\partial E}{\partial m} = 0 = \frac{\partial E}{\partial b}$

$$\Rightarrow \sum_i x_i(y_i - mx_i - b) = 0 = \sum (y_i - mx_i - b)$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\sum_i x_i y_i = b\sum_i x_i + m\sum_i x_i^2 \qquad \sum_i y_i = nb + m\sum_i x_i$$

d. $(1, 1), (2, 3)$

$$m = \frac{3-1}{2-1} = 2 \Rightarrow y = 2x - 1$$ which has zero error (goes through both points)

e. $1 \cdot 1 + 2 \cdot 3 = b(1 + 2) + m(1^2 + 2^2)$

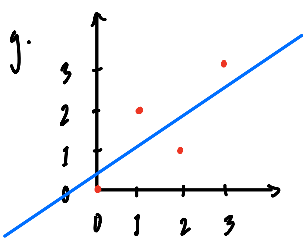
$\Rightarrow 7 = 3b + 5m$

$1 + 3 = 2b + m(1 + 2)$

$\Rightarrow 4 = 2b + 3m$

$m = 2, b = -1$

f, g.



something like this

$x_i = 0, 1, 2, 3$

$y_i = 0, 2, 1, 3$

$0 \cdot 0 + 1 \cdot 2 + 2 \cdot 1 + 3 \cdot 3 = b(0 + 1 + 2 + 3) + m(0^2 + 1^2 + 2^2 + 3^2)$

$0 + 2 + 1 + 3 = 4b + m(0 + 1 + 2 + 3)$

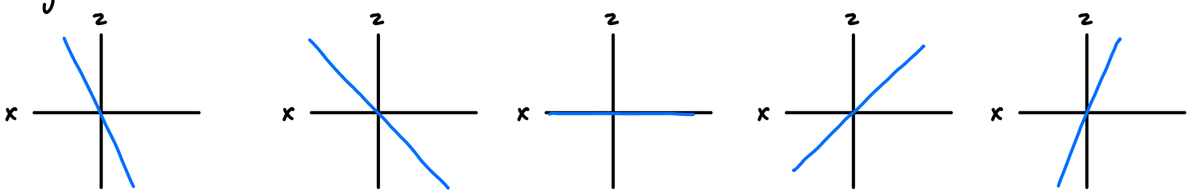
$$\Rightarrow \begin{cases} 13 = 6b + 14m \\ 6 = 4b + 6m \end{cases} \Rightarrow \begin{cases} 13 = 6b + 14m \\ 3 = 2b + 3m \end{cases}$$

$\Rightarrow 13 = 3(3 - 3m) + 14m \Rightarrow m = 4/5, \ b = 3/10$

$$\Rightarrow \boxed{y = \frac{4}{5}x + \frac{3}{10}}$$

2a. $f_x(0, n) = n \quad (n = -2, -1, 0, 1, 2)$

b. $y = -2$: $\quad -1$: $\quad 0$: $\quad 1$: $\quad 2$:



c. slope increases as y increases

d. $f_{xy}(0, 0) = 1 > 0$

e. $f_x(0, y) = (3x^2 + 1)(y^3 + y)\big|_{x=0} = y^3 + y$: near 0, slope increases as y increases

$f_{xy}(0, 0) = (3x^2 + 1)(3y^2 + 1)\big|_{x=y=0} = 1 > 0.$