1. Use least-squares regression to fit a straight line $y = mx + c$ by computing the slope m and intercept c as follows:

Given $n = 5$

Step 1: For each (x, y) point calculate x^2 and xy

Step 2: Sum all x, y, x^2 and xy, which gives us $\sum x$, $\sum y$, $\sum x^2$ and $\sum xy$ ($\sum$ means "sum up")

$$\sum x_i = 26 \tag{1}$$

$$\sum y_i = 41 \tag{2}$$

$$\sum x_i^2 = 168 \tag{3}$$

$$\sum x_i y_i = 263 \tag{4}$$

Step 3: Calculate Slope m:

$$m = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{5 \times 263 - 26 \times 41}{5 \times 168 - 26 \times 26} = 1.1583 \tag{5}$$

Step 4: Calculate Intercept c:

$$c = \frac{\sum y_i - m \sum x_i}{n} = \frac{41 - 1.5183 \times 26}{5} = 0.305 \tag{6}$$

Step 5: Assemble the equation of a line:

$$y = mx + c = y = 1.5183x + 0.305 \tag{7}$$

2. Exact Line Search using Steepest Descent Method:

Given the intuition that the negative gradient $-\nabla f_k$ can be an effective search direction, steepest descent follows the idea and establishes a systematic method for minimizing the objective function. Setting $-\nabla f_k$ as the direction, steepest descent computes the step-length α_k by minimizing a single-variable objective function. More specifically, the steps of Steepest Descent Method are as follows:

Steepest Descent Algorithm

Set a starting point x_0

Set a convergence criterium $\epsilon > 0$

Set $k = 0$

Set the maximum iteration N

While $k \leq N$:

$\nabla f(x_k) = \frac{\partial f(x)}{\partial x}\Big|_{x=x_k}$

If $\nabla f(x_k) \leq \epsilon$:

Break

End if

$\alpha_k = \arg\min_{\alpha} f(x_k - \alpha \nabla f(x_k))$

$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$

$k = k + 1$

End while

Return $x_k, f(x_k)$

3. Let us proceed with the exact line search using steepest descent method with 5 iterations.

First iteration:

We have $\nabla f(x) = \begin{bmatrix} 1 + 2x_2 + 4x_1 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$.
Starting from $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ gives $-\nabla f(x_0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
We then have $f(x_0 - \alpha \nabla f(x_0)) = f(-\alpha, \alpha) = \alpha^2 - 2\alpha$.

Taking partial derivative of the above equation with respect to α and set it to zero to find the minimizer, $\alpha_0 = 1$.

Therefore, $x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \alpha_0 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Second iteration:

Given $x_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, we have $-\nabla f(x_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
Then from $\min_{\alpha} f(x_1 - \alpha \nabla f(x_1)) = 5\alpha^2 - 2\alpha - 1$, finding the minimizer, $\alpha_1 = 0.2$.

Hence, $x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix}$.

Third iteration:

Given $x_2 = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix}$, we have have $-\nabla f(x_2) = \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix}$.

Then from $\min_{\alpha} f(x_2 - \alpha \nabla f(x_2)) = 0.04\alpha^2 - 0.08\alpha - 1.2$, finding the minimizer, $\alpha_2 = 1$.

Hence, $x_3 = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} + \alpha_2 \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix} = \begin{bmatrix} -1.0 \\ 1.4 \end{bmatrix}$.

Fourth iteration:

Given $x_3 = \begin{bmatrix} -1.0 \\ 1.4 \end{bmatrix}$, we have have $-\nabla f(x_3) = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$.

Then from $\min_{\alpha} f(x_3 - \alpha \nabla f(x_3)) = 0.2\alpha^2 - 0.08\alpha - 1.24$, finding the minimizer, $\alpha_3 = 0.2$.

Hence, $x_4 = \begin{bmatrix} -1.0 \\ 1.4 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix} = \begin{bmatrix} -0.96 \\ 1.44 \end{bmatrix}$.

Fifth iteration:

Given $x_4 = \begin{bmatrix} -0.96 \\ 1.44 \end{bmatrix}$, we have have $-\nabla f(x_4) = \begin{bmatrix} -0.04 \\ 0.04 \end{bmatrix}$.

Then from $\min_{\alpha} f(x_4 - \alpha \nabla f(x_4)) = 0.0016\alpha^2 - 0.0032\alpha - 1.248$, finding the minimizer, $\alpha_4 = 1$.

Hence, $x_5 = \begin{bmatrix} -0.96 \\ 1.44 \end{bmatrix} + \alpha_4 \begin{bmatrix} -0.04 \\ 0.04 \end{bmatrix} = \begin{bmatrix} -1.00 \\ 1.48 \end{bmatrix}$.

Termination:

At this point, we find $\nabla f(x_5) = \begin{bmatrix} -0.04 \\ -0.04 \end{bmatrix}$.

Check to see if the convergence is sufficient by evaluating $||\nabla f(x_5)||$:

$||\nabla f(x_5)|| = \sqrt{(-0.04)^2 + (-0.04)^2} = 0.0565$.

Since 0.0565 is relatively small and is close enough to zero, the line search is complete.

The derived optimal solution is $x^* = \begin{bmatrix} -1.00 \\ 1.48 \end{bmatrix}$, and the optimal objective value is found to be -1.25.

4. Solve the First Order Necessary Condition to find stationary point *.

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - x_2 - 15 \\ -x_1 + 6x_2 - 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 15 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\implies \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

Now Check Second Order Sufficiency Condition as follows:

The Hessian at x^* is: $\nabla^2 f(x_1^*, x_2^*) = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}$

The eigenvalues of the Hessian come out to be 4.236 and 1.764, and as both eigenvalues are positive, the Hessian is positive definite which means that $f(x_1, x_2)$ has a local minima at $x^* = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$.