1.

Sea

$$f(x) = \frac{1}{2}\sum_{j=1}^{m} r_j^2(x),$$

donde $r_j : \mathbb{R}^n \to \mathbb{R}$ para j = 1, ..., m. Si definimos la función **R** : $\mathbb{R}^n \to \mathbb{R}^m$ como **R = R(x)** = $(r_1(x), r_2(x), ..., r_m(x))^T$ y **J = J(x)** es la matriz jacobiana de R, muestre que

$$J^TJ = \sum_{i=1}^{m}(\nabla r_i)(\nabla r_i)^T,$$
$$J^TR = \sum_{i=1}^{m} r_i(\nabla r_i).$$

1.1.

$$J = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \dots & \frac{\partial r_1}{\partial x_n} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \dots & \frac{\partial r_2}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \frac{\partial r_m}{\partial x_2} & \dots & \frac{\partial r_m}{\partial x_n}, \end{bmatrix}$$

entonces,

$$J^TJ = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_2}{\partial x_1} & \dots & \frac{\partial r_m}{\partial x_1} \\ \frac{\partial r_1}{\partial x_2} & \frac{\partial r_2}{\partial x_2} & \dots & \frac{\partial r_m}{\partial x_2} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial r_1}{\partial x_n} & \frac{\partial r_2}{\partial x_n} & \dots & \frac{\partial r_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \dots & \frac{\partial r_1}{\partial x_n} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \dots & \frac{\partial r_2}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \frac{\partial r_m}{\partial x_2} & \dots & \frac{\partial r_m}{\partial x_n} \end{bmatrix},$$

finalmente,

$$J^TJ = \begin{bmatrix} \sum_{i=1}^{m} \frac{\partial r_i}{\partial x_1}\frac{\partial r_i}{\partial x_1} & \dots & \dots & \dots \\ \sum_{i=1}^{m} \frac{\partial r_i}{\partial x_2}\frac{\partial r_i}{\partial x_1} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} = \sum_{i=1}^{m}(\nabla r_i)(\nabla r_i)^T$$

1.2.

$$\begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_2}{\partial x_1} & \dots & \frac{\partial r_m}{\partial x_1} \\ \frac{\partial r_1}{\partial x_2} & \frac{\partial r_2}{\partial x_2} & \dots & \frac{\partial r_m}{\partial x_2} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial r_1}{\partial x_n} & \frac{\partial r_2}{\partial x_n} & \dots & \frac{\partial r_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} = \begin{bmatrix} r_1\frac{\partial r_1}{\partial x_1} + r_2\frac{\partial r_2}{\partial x_1} + ... + r_m\frac{\partial r_m}{\partial x_1} \\ r_1\frac{\partial r_1}{\partial x_2} + r_2\frac{\partial r_2}{\partial x_2} + ... + r_m\frac{\partial r_m}{\partial x_2} \\ \vdots \\ r_1\frac{\partial r_1}{\partial x_n} + r_2\frac{\partial r_2}{\partial x_n} + ... + r_m\frac{\partial r_m}{\partial x_n} \end{bmatrix} = \sum_{i=1}^{m} r_i \frac{\partial r_i}{\partial x_k}$$

2.

Sea R $\in \mathbb{R}^m, J \in \mathbb{R}^{m\times n}, I \in \mathbb{R}^{n\times n}$ la matriz identidad y $\mu > 0$. Muestre que $s = -(J^TJ + \mu I)^{-1}J^TR$ es la solución del problema de mínimos cuadrados

$$\min_{s\in\mathbb{R}^n} ||As + b||_2^2,$$

donde $A \in \mathbb{R}^{m+n} \times n$, $b \in \mathbb{R}^{m+n}$ están dadas por los bloques siguientes:

$$A = \begin{bmatrix} J \\ \mu I \end{bmatrix}, b = \begin{bmatrix} R \\ 0 \end{bmatrix}.$$

Reescribiendo f,

$$
\begin{aligned}
f(s) &= ||As+b||_2^2, \\
&= (s^T A^T + b^T)(As+b), \\
&= s^T(A^T A)s + b^T As + b^T b.
\end{aligned}
$$

Calculando el gradiente e igualando a cero para encontrar la solución: $\nabla f(s) = A^T As + A^T b = 0$. Se obtiene:

$$
\begin{aligned}
(A^T A)s &= -A^T b, \\
s &= -(A^T A)^{-1} A^T b.
\end{aligned}
$$

Si m = 2, n = 3:

$$
A = \begin{bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ \sqrt{\mu} & 0 & 0 \\ 0 & \sqrt{\mu} & 0 \\ 0 & 0 & \sqrt{\mu} \end{bmatrix}, b = \begin{bmatrix} r_1 \\ r_2 \\ 0 \\ 0 \\ 0 \end{bmatrix},
$$

entonces,

$$A^T b = J^T R.$$

De la misma forma, reemplazando $A^T A$ se obtiene $J^T J + \mu I$. Así, $s = -(J^T J + \mu I)^{-1} J^T R$.

3.

Algoritmo de Levenberg - Marquart

Este algoritmo se utiliza para resolver problemas de mínimos cuadrados no lineales, es una combinación entre el algoritmo de Gauss-Newton y el método de descenso de gradiente. En este algoritmo, la suma de los errores al cuadrado se reduce hacia la dirección de descenso más empinado.

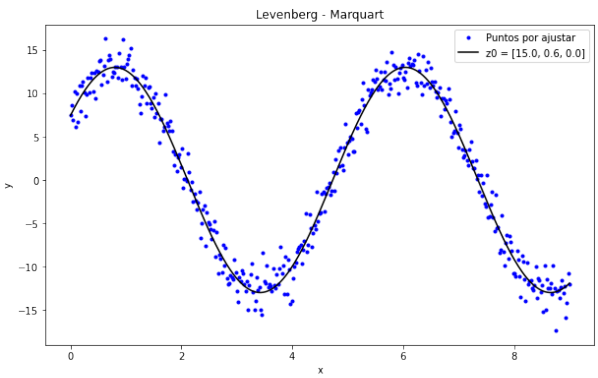

Figura 1: Ajuste utilizando el primer punto dado.

```
Valor inicial:  45454.05280978729
El algoritmo convergió en la iteración:  7
z_{k}:  [12.99599948  1.19938759 -5.67328839]
f(z_{k}):  45454.05280978729
||p_{k}||:  0.0001393870447214434
Número de iteraciones:  7
```

Figura 2: Descripción de los resultados.

Utilizando el punto inicial [15.0, 0.6, 0.0], el algoritmo converge rápidamente y en muy pocas iteraciones.

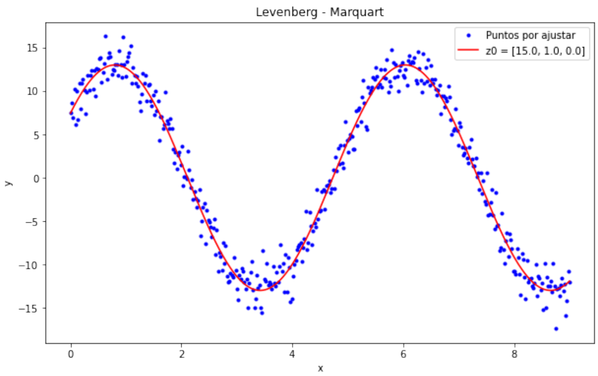

Figura 3: Ajuste con otro punto inicial.

```
Valor inicial:  40807.16289819636
El algoritmo convergió en la iteración:  34
z_{k}:  [-12.99605762  -1.19936531  -6.89318307]
f(z_{k}):  40807.16289819636
||p_{k}||:  2.006993321640592e-05
Número de iteraciones:  34
```

Figura 4: Descripción de los resultados.

Utilizando este punto se muestra que el número de iteraciones necesarias para converger aumentó casi cinco veces más, sin embargo, la norma del vector p se redujó significativamente.

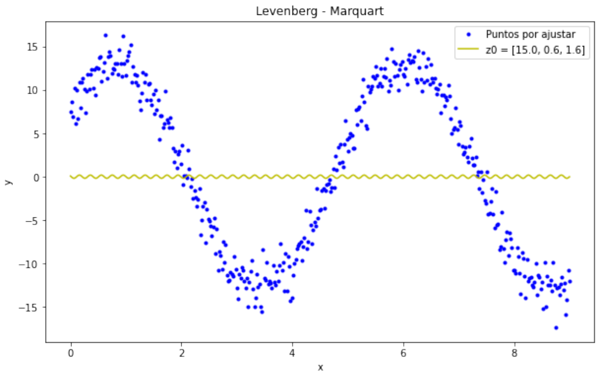

Figura 5: Último punto de ajuste.

```
Valor inicial:  37048.62007346928
z_{k}:  [ -0.19071224 -31.807208    72.68240361]
f(z_{k}):  37048.62007346928
||p_{k}||:  0.5978166364652602
Número de iteraciones:  5000
```

Figura 6: Descripción de los resultados.

En este caso, el algoritmo no convergió e hizo todas las iteraciones, y como se muestra en la gráfica, no llegó a buenos resultados después de 5000 iteraciones.