# Matemáticas III

## Mínimos cuadrados
## Factorización QR
## Semana 09


**Hermes Pantoja Carhuavilca**
(hpantoja@utec.edu.pe)
**Brigida Molina Carabaño**
(bmolina@utec.edu.pe)
**Rósulo Pérez Cupe**
(rperezc@utec.edu.pe)
**Asistente: Victor Anhuaman**
(vanhuaman@utec.edu.pe)


UTEC UNIVERSIDAD DE INGENIERÍA & TECNOLOGÍA

# Temas

# Objetivo

Aproximar funciones utilizando el método de mínimos cuadrados, así como también la descomposición QR.

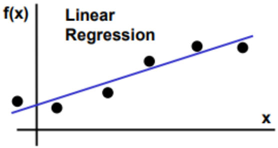

# 1 AJUSTE POR MÍNIMOS CUADRADOS

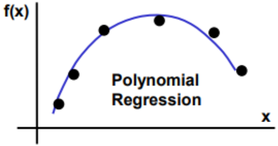
# Logros

1. Aproxima las funciones utilizando el método de los mínimos cuadrados.
2. Resuelve el problema de mínimos cuadrados mediante ecuaciones normales.
3. Realiza la descomposición QR de una matriz.

# Formalización de Contenidos

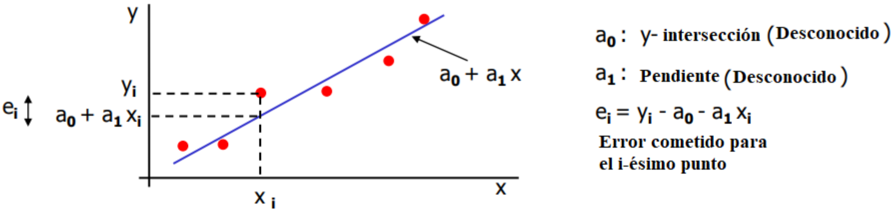
El ajuste de curva permite expresar un conjunto de puntos de datos discretos como una función continua. Generalmente los datos son experimentales, que pueden tener una cantidad significativa de error y no es necesario encontrar una función que pase por todos los puntos discretos.



# Ajuste lineal por Mínimos Cuadrados

1. Ajustar a una recta un conjunto de datos (pares ordenados) $(x_1; y_1)$, $(x_2; y_2)$ $(x_3; y_3), \dots, (x_n; y_n)$




2. Minimizar el error para obtener un mejor ajuste lineal (encontrar $a_0$ y $a_1$). Para ello minimizamos la suma de los cuadrados de los errores cometidos.

# Continuación...

Consideremos las expresiones

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - a_1 x_i - a_0)^2$$

Determinar $a_0$ y $a_1$ minimizando $S_r$.
Derivamos $S_r$ con respecto a $a_0$ y $a_1$ e igualando a cero.

$$\frac{\partial S_r}{\partial a_0} = -2 \sum_{i=1}^{n} (y_i - a_1 x_i - a_0) = 0$$

$$\frac{\partial S_r}{\partial a_1} = -2 \sum_{i=1}^{n} (y_i - a_1 x_i - a_0) x_i = 0$$

# Continuación...

Despejando convenientemente y considerando las variables $a_0$ y $a_1$ siendo el resto de elementos conocido, tenemos planteado las ecuaciones normales:

$$\begin{cases} na_0 + \left(\sum_{i=1}^{n} x_i\right) a_1 = \sum_{i=1}^{n} y_i \\ \left(\sum_{i=1}^{n} x_i\right) a_0 + \left(\sum_{i=1}^{n} x_i^2\right) a_1 = \left(\sum_{i=1}^{n} x_i y_i\right) \end{cases}$$

Al resolver el sistema anterior obtenemos los coeficientes $a_0$ y $a_1$

# Bondad del ajuste lineal

Consideremos las expresiones

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \qquad S_t = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Donde $\bar{x}$ e $\bar{y}$ son las medias de $x$ e $y$ respectivamente. Entonces

$$r^2 = \frac{S_t - S_r}{S_t}$$

o en términos de la nube de puntos

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

# Bondad del ajuste lineal

- A la cantidad $r^2$ se le conoce como **coeficiente de Determinación**.
- Siendo $r$ el **coeficiente de Correlación**.
- Un ajuste perfecto ocurre cuando se cumple $r^2 = r = 1$.
- El coeficiente $r^2$ determina la calidad del ajuste, algunos autores sugieren que debe ser mayor a 0.7.

# Actividad 1

Halle la recta que mejor se ajusta a los siguientes datos (mostrados en la tabla), determinando asímismo la calidad del ajuste

| Tiempo (s) | 0 | 3 | 4 | 7 | 9 |
|---|---|---|---|---|---|
| Distancia (m) | 0 | 30 | 60 | 90 | 120 |

Consideremos $x_i = Tiempo$ e $y_i = Distancia$.

# Desarrollo

Según las necesidades del cálculo elaboramos la siguiente tabla

| $x_i$ | $y_i$ | $x_i^2$ | $y_i^2$ | $x_i y_i$ | $(y_i - \bar{y})^2$ | $y$ | $e_i^2$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 3600 | -1.707 | 2.9149 |
| 3 | 30 | 9 | 900 | 90 | 900 | 38.537 | 72.873 |
| 4 | 60 | 16 | 3600 | 240 | 0 | 51.951 | 64.783 |
| 7 | 90 | 49 | 8100 | 630 | 900 | 92.195 | 4.8186 |
| 9 | 120 | 81 | 14400 | 1080 | 3600 | 119.02 | 0.9518 |
| 23 | 300 | 155 | 27000 | 2040 | 9000 | | 146.34 |

En la última fila de la tabla se tiene la suma de las respectivas columnas, las cuales utilizamos para determinar $n$, $\bar{y}$, $a$, $b$, $S_t$, $S_r$ y $r^2$ [ejercicio].

De la tabla tenemos:

$$n = 5, \quad \sum_{i=1}^{5} x_i = 23, \quad \sum_{i=1}^{5} y_i = 300, \quad \sum_{i=1}^{5} x_i^2 = 155, \quad \sum_{i=1}^{5} x_i y_i = 2040$$

De donde formamos la ecuación normal:

$$\begin{cases} n a_0 + \left(\sum_{i=1}^{5} x_i\right) a_1 & = \sum_{i=1}^{5} y_i \\ \left(\sum_{i=1}^{5} x_i\right) a_0 + \left(\sum_{i=1}^{5} x_i^2\right) a_1 & = \sum_{i=1}^{5} x_i y_i \end{cases}$$

Es decir:

$$\begin{cases} 5a_0 + 23a_1 = 300 \\ 23a_0 + 155a_1 = 2040 \end{cases}$$

Resolviendo obtenemos:

$$a_1 = 13.41463415 \qquad ; a_0 = -1.707317073$$

# Cálculo de los parámetros

$$n = 5 \qquad \bar{y} = 60 \qquad a_1 = 13.41463415 \qquad a_0 = -1.707317073$$

$$S_t = 9000 \qquad S_r = 146.34$$

en consecuencia, la recta ajustada por mínimos cuadrados sería

$$y = 13.41463415x - 1.707317073$$

y el coeficiente de determinación

$$r^2 = 0.983739837.$$


# 2 FACTORIZACIÓN QR

# Formalización de Contenidos

## Factorización QR por Gram-Schmidt

La descomposición $QR$ (Llamado también factorización $QR$ ) de una matriz es una descomposición de una matriz en una matriz ortogonal y una matriz triangular superior. Una descomposición QR de una matriz real $A$ (no necesariamente cuadrada) es una descomposición de $A$ como

$$A = QR$$

donde $Q$ es una matriz ortogonal ($Q^T Q = I$) del mismo tamaño que $A$, las columnas de $Q$ son ortonormales, y $R$ es una matriz triangular superior. Si $A$ es una matriz cuadrada y no singular, entonces ésta factorización es única.

# Procedimiento

$$A = [a_1 | a_2 | \dots | a_n]$$

Entonces,

$$u_1 = a_1, \qquad q_1 = \frac{u_1}{|u_1|}$$

$$u_2 = a_2 - (a_2.q_1).q_1, \qquad q_2 = \frac{u_2}{|u_2|}$$

$$u_{k+1} = a_{k+1} - (a_{k+1}.q_1)q_1 - \dots - (a_{k+1}.q_k)q_k, \qquad q_{k+1} = \frac{u_{k+1}}{|u_{k+1}|}$$

# Continuación...

El resultado de la factorización es:

$$A = [a_1|a_2|\dots|a_n] = [q_1|q_2|\dots|q_n]\begin{bmatrix} a_1.q_1 & a_2.q_1 & \dots & a_n.q_1 \\ 0 & a_2.q_2 & \dots & a_n.q_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n.q_n \end{bmatrix} = QR$$

# Actividad 2

## P1

Hallar la factorización QR de la matriz

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

# Solución

Con los vectores $a_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $a_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Luego:

# Continuación...

$$u_1 = a_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad q_1 = \frac{u_1}{|u_1|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$u_2 = a_2 - (a_2.q_1)q_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{-1}{2} \\ 1 \end{bmatrix} \qquad q_2 = \frac{u_2}{|u_2|} = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

# Continuación...

Así

$$Q = [q_1 \mid q_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

$$R = \begin{bmatrix} a_1.q_1 & a_2.q_1 \\ 0 & a_2.q_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} \end{bmatrix}$$

Por lo tanto:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} \end{bmatrix}$$

# Actividad 3

El nivel del agua en el mar de Lima esta principalmente determinado por la marea. Se han tomado las siguientes medidas

| **t** | 0 | 2 | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|
| **h(t)** | 1 | 1.6 | 1.4 | 0.6 | 0.2 | 0.8 |

Con t medido en horas,

# Continuación...

1. Halle la recta $h_1(t) = a_1 + a_0 t$ que mejor ajuste a los datos por mínimos cuadrados
2. Determinar el **Coeficiente de Determinación** e indicar que tan buen ajuste es.

# Continuación...

Utilice la descomposicion QR(usar Matlab), para resolverlo.
Ajuste los datos anteriores a la siguiente forma:
**Caso 1:**

$$h_1(t) = a_1 + a_0 t$$

**Caso 2:**

$$h_2(t) = a_0 + a_1 sen(\pi t/6) + a_2 cos(\pi t/6)$$

Del sistema matricial $Xa = y$, descomponer X en QR

# Continuación...

1. Hallar la suma de elementos de la primera fila de R, para los casos 1 y 2.
2. Hallar: $a_1 \times a_0$, para el caso 1.
3. Hallar X del sistema matricial $Xa = y$, para los casos 1 y 2
4. Hallar la suma de elementos de la primera fila de Q, para los casos 1 y 2
5. Hallar $a_0 + a_1 + a_2$, para el caso 2.

# Gracias por su atención