UNIVERSIDAD DE GUANAJUATO

Reconocimiento de patrones

Dr. Mario Alberto Ibarra Manzano
ibarram@ugto.mx
Oficina: 217

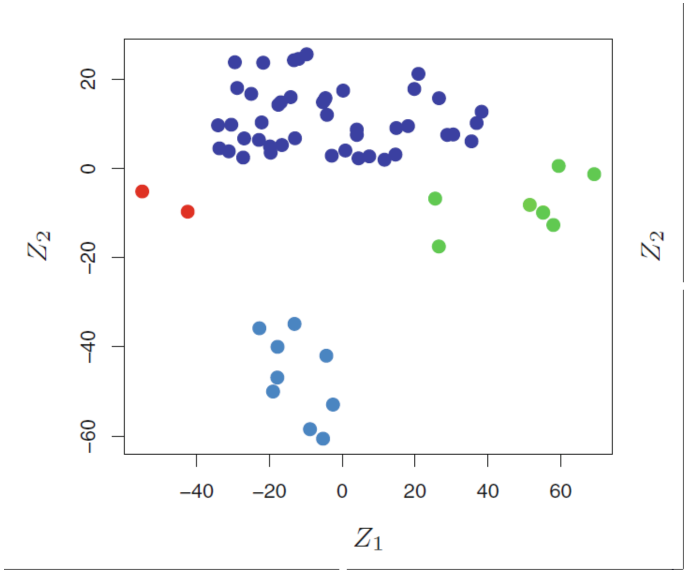
Recocimiento de patrones es el conjunto de herramientas que permiten seleccionar, ajustar, aplicar, implementar y comparar un estimador para analizar o predecir el comportamiento de un sistema o proceso.

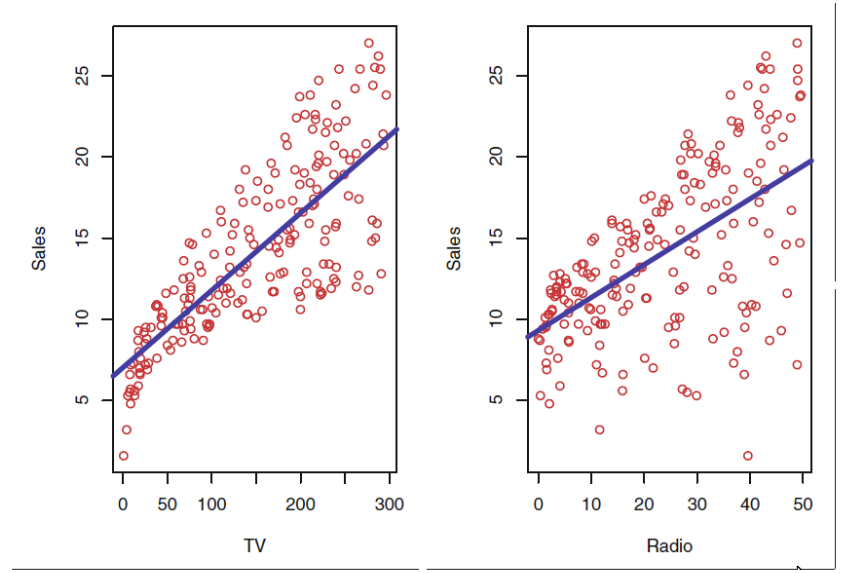
Las etapas generales en la formulación de un estimador son:

1. Selección apropiada de la familia del modelo y el modelo a utilizar en función de la naturaleza del problema.
2. Estimación de los parámetros desconocidos del modelo seleccionado en función de las variables conocidas y el comportamiento del proceso.
3. Evaluación de la eficiencia (desempeño) de las hipótesis sobre los parámetros desconocidos.

¿Qué deseo modelar?

La naturaleza del problema (selección de la rama del problema)





(a) Clasificación

1. Detección de fallas en sistemas electromecánicos
2. Diagnóstico médico
3. Detección de objetos

(b) Regresión

1. Modelado de sensores
2. Predicción del clima
3. Modelado de sistemas de transporte

¿Cuáles son las características que definen el proceso?

Principales métodos de recocimiento de patrones

- Modelos lineales
- Modelos no lineales
- Métodos basados en arboles
- Métodos basados en kernel
- Modelos dinámicos
- Métodos de agrupamiento

Orden del método de recocimiento de patrones

- Primer orden (datos brutos)
- Segundo orden (operador de transformación fundamental)
- Alto orden (modelo descriptivo de los datos)

¿Cuáles son las variables que determinan el proceso?

Definición de las variables que influye en el proceso

- Número de variable (dimensionalidad del problema)
- Naturaleza de las variables
- Número de muestras disponibles
- Proceso paramétrico o no paramétrico

Determinación de parámetros del modelo

$$\{X_1, \dots, X_n, \theta\}$$

$$\hat{\theta}_n = g(X_1, \dots X_n)$$

$$bias(\hat{\theta}_n) = \mathbb{E}_\theta(\hat{\theta}_n) - \theta$$

$$\hat{\theta}_n \xrightarrow{P} \theta \quad consistente$$

$$se = se(\hat{\theta}_n) = \sqrt{\mathbb{V}(\hat{\theta}_n)}$$

$$MSE = bias^2(\hat{\theta}_n) + \mathbb{V}(\hat{\theta}_n)$$

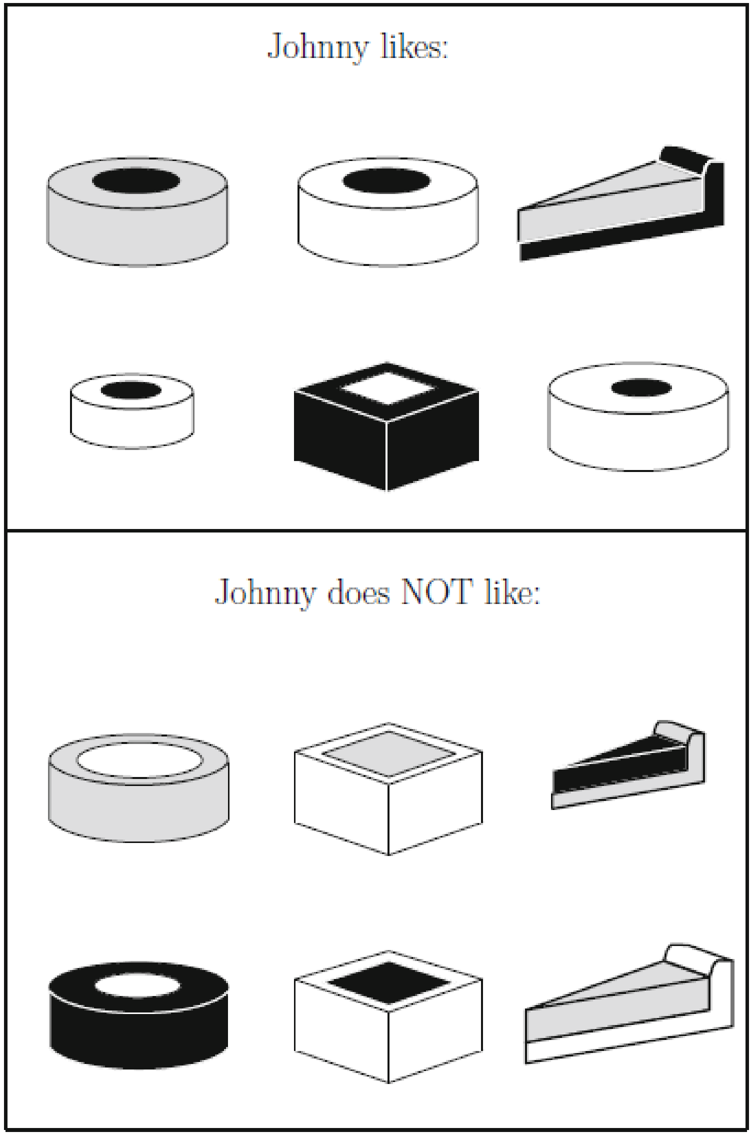
$$Si\ bias \to 0\ \&\ se \to 0\ con\ n \to \infty\ entonces\ \hat{\theta}_n\ es\ consistente\ \hat{\theta}_n \xrightarrow{P} \theta$$

$$MSE \to 0$$

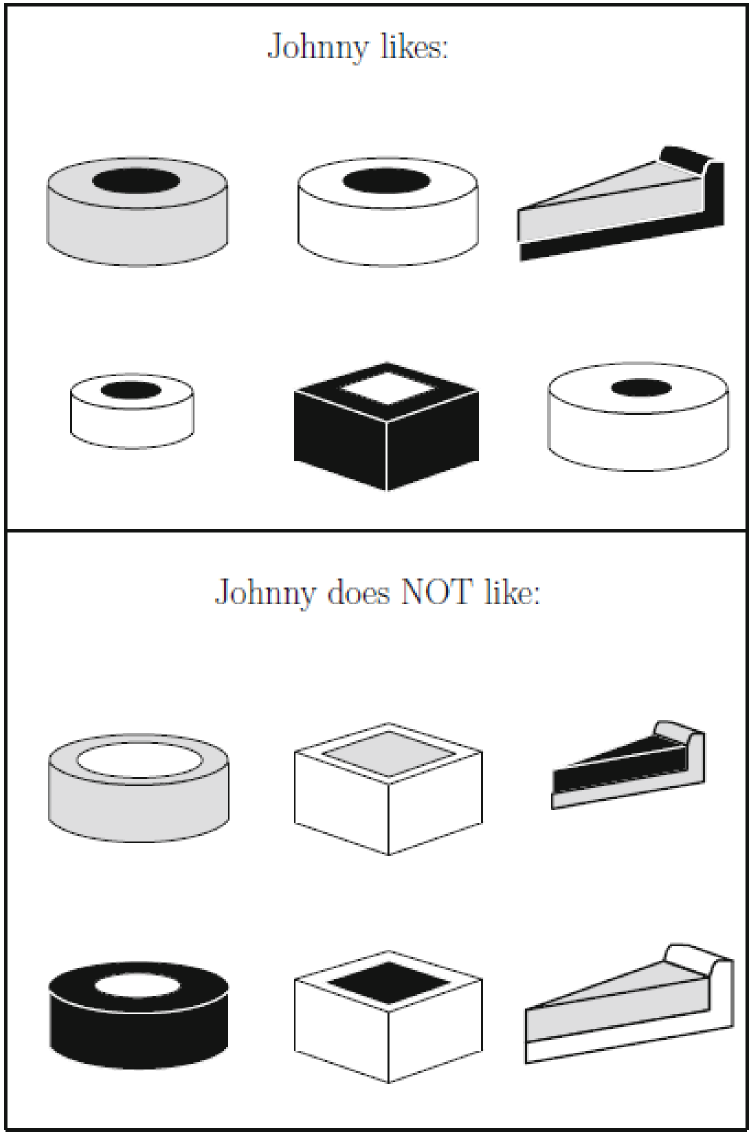
$$\frac{\hat{\theta}_n - \theta}{se} \to N(0,1)$$

Distribution	Mean	Variance
Point mass at a	a	0
Bernoulli(p)	p	$p(1-p)$
Binomial(n,p)	np	$np(1-p)$
Geometric(p)	$1/p$	$(1-p)/p^2$
Poisson(λ)	λ	λ
Uniform(a,b)	$(a+b)/2$	$(b-a)^2/12$
Normal(μ,σ^2)	μ	σ^2
Exponential(β)	β	β^2
Gamma(α,β)	$\alpha\beta$	$\alpha\beta^2$
Beta(α,β)	$\alpha/(\alpha+\beta)$	$\alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$
t_ν	0 (if $\nu>1$)	$\nu/(\nu-2)$ (if $\nu>2$)
χ^2_p	p	$2p$
Multinomial(n,p)	np	see below
Multivariate Normal(μ,Σ)	μ	Σ

Ejemplo



Ejemplo



Atributos (Características)

1. Forma (Shape)
   1. Circulo (Circle)
   2. Cuadrado (Square)
   3. Triángulo (Triangle)
2. Color (Colour)
   1. Blanco (White)
   2. Negro (Dark)
   3. Gris (Gray)
3. Textura (Texture)
   1. ?
4. Tamaño (Size)
   1. Chico (Thin)
   2. Grande (Thick)
5. Posición (Position)
   1. Interno (Crust)
   2. Externo (Filling)

Ejemplo

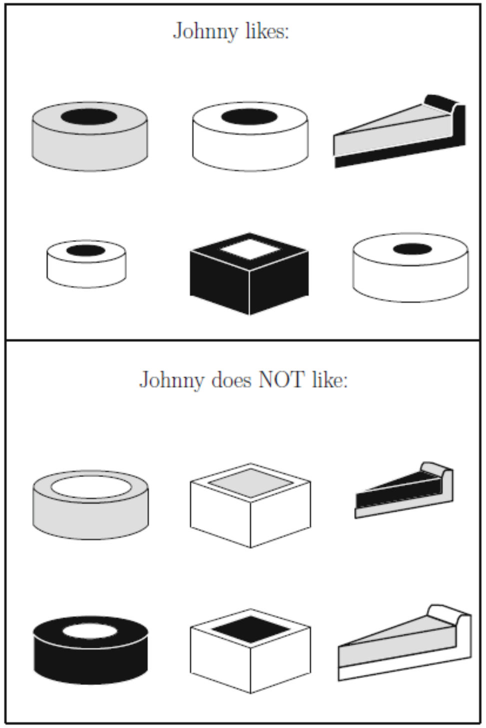


Table 1.1 The twelve training examples expressed in a matrix form

Example	Shape	Crust		Filling		Class
		Size	Shade	Size	Shade	
ex1	Circle	Thick	Gray	Thick	Dark	pos
ex2	Circle	Thick	White	Thick	Dark	pos
ex3	Triangle	Thick	Dark	Thick	Gray	pos
ex4	Circle	Thin	White	Thin	Dark	pos
ex5	Square	Thick	Dark	Thin	White	pos
ex6	Circle	Thick	White	Thin	Dark	pos
ex7	Circle	Thick	Gray	Thick	White	neg
ex8	Square	Thick	White	Thick	Gray	neg
ex9	Triangle	Thin	Gray	Thin	Dark	neg
ex10	Circle	Thick	Dark	Thick	White	neg
ex11	Square	Thick	White	Thick	Dark	neg
ex12	Triangle	Thick	White	Thick	Gray	neg

Vector de Atributos: Forma, Tamaño externo, Color externo, Tamaño interno, Color interno

Attribute vectors: Shape, Crust size, Crust shade, Filling size, Filling shade

Vector de Atributos:

1. Forma (3 valores)
   1. Circulo
   2. Cuadrado
   3. Triángulo
2. Tamaño externo (2 valores)
   1. Pequeño
   2. Grande
3. Color externo (3 valores)
   1. Blanco
   2. Gris
   3. Negro
4. Tamaño interno (2 valores)
   1. Pequeño
   2. Grande
5. Color interno (3 valores)
   1. Blanco
   2. Gris
   3. Negro

Forma y Tamaño externo

$$3 \times 2 = 6$$

Forma, Tamaño externo y Color externo

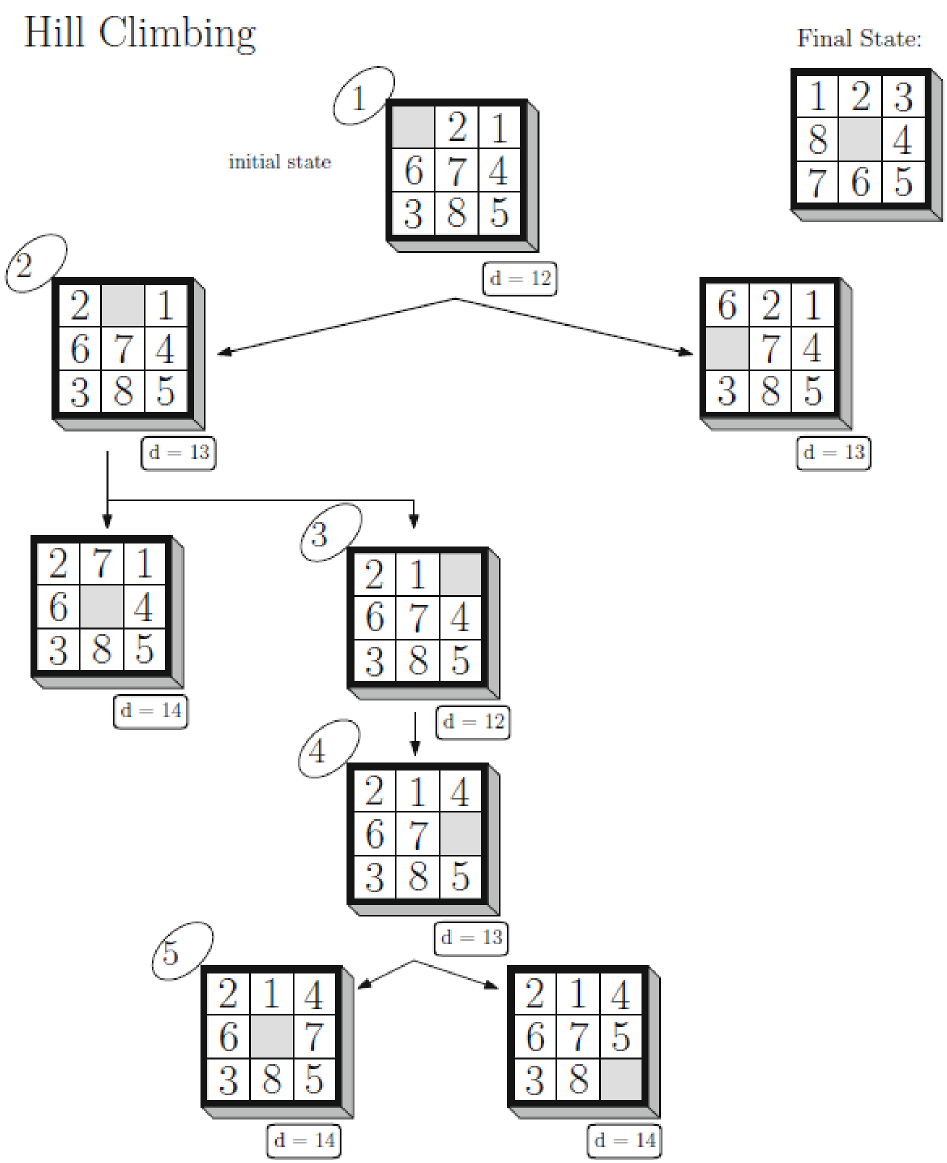
$$3 \times 2 \times 3 = 18$$

Todos los atributos

$$3 \times 2 \times 3 \times 2 \times 3 = 108$$

Para 108 diferentes ejemplos

$$2^{108} \; posibles \; subconjuntos$$



Algoritmo de búsqueda Hill-climbing:

1. Crea dos listas, *L* y *Ls*. Al *principio*, *L* contiene solo el estado inicial, y *Ls* está vacía.
2. Sea *n* el primer elemento de *L*. Compare este estado con el estado final. Si son idénticos, detenerse y terminar con éxito.
3. Aplique a *n* todos los operadores de búsqueda disponibles, obteniendo así un conjunto de nuevos estados. Descartar aquellos estados que ya existen en *Ls*. En cuanto al resto, ordénelos por la función de evaluación y colóquelos al frente de *L*.
4. Transfiera *n* de *L* a la lista, *Ls*, de los estados que se han investigado.
5. Si *L=Ø*, detener y terminar con error. De lo contrario, regresar a 2.

Ejemplo

Table 1.1 The twelve training examples expressed in a matrix form

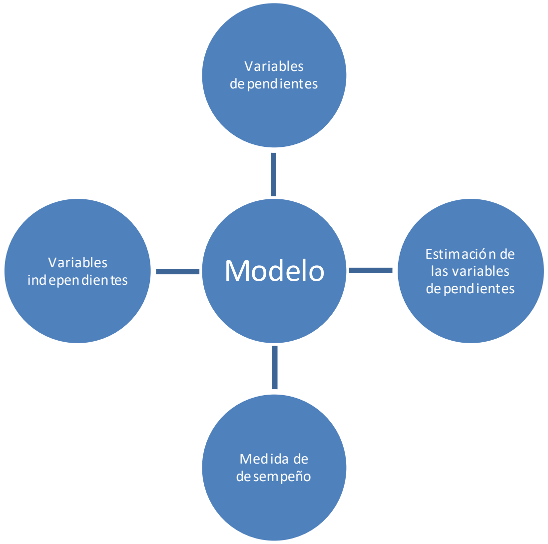
Example	Shape	Crust		Filling		Class
		Size	Shade	Size	Shade	
ex1	Circle	Thick	Gray	Thick	Dark	pos
ex2	Circle	Thick	White	Thick	Dark	pos
ex3	Triangle	Thick	Dark	Thick	Gray	pos
ex4	Circle	Thin	White	Thin	Dark	pos
ex5	Square	Thick	Dark	Thin	White	pos
ex6	Circle	Thick	White	Thin	Dark	pos
ex7	Circle	Thick	Gray	Thick	White	neg
ex8	Square	Thick	White	Thick	Gray	neg
ex9	Triangle	Thin	Gray	Thin	Dark	neg
ex10	Circle	Thick	Dark	Thick	White	neg
ex11	Square	Thick	White	Thick	Dark	neg
ex12	Triangle	Thick	White	Thick	Gray	neg

Algoritmo de búsqueda Hill-climbing:

1. Crea dos listas, *L* y *Ls*. Al *principio*, *L* contiene solo el estado inicial, y *Ls* está vacía.
2. Sea *n* el primer elemento de *L*. Compare este estado con el estado final. Si son idénticos, detenerse y terminar con éxito.
3. Aplique a *n* todos los operadores de búsqueda disponibles, obteniendo así un conjunto de nuevos estados. Descartar aquellos estados que ya existen en *Ls*. En cuanto al resto, ordénelos por la función de evaluación y colóquelos al frente de *L*.
4. Transfiera *n* de *L* a la lista, *Ls*, de los estados que se han investigado.
5. Si *L=Ø*, detener y terminar con error. De lo contrario, regresar a 2.

El objetivo principal de un modelo de regresión consiste en encontrar la función que satisfaga la relación entre un conjunto de datos. El conjunto de datos esta integrado por dos tipos de variables, las variables independientes o de entrada y las variables dependientes o de salida. Las variables dependientes e independientes puede ser continuas, discretas, binarias o conjuntos, en algunos casos esta variables pueden describir de forma temporal, espacial o alguna característica del modelo de interés.

En los modelos de regresión es importante conocer la naturaleza del problema, además de contar con una medida para evaluar el desempeño que permita discriminar las mejoras del mismo.



Dado un conjunto de puntos $\{x_i, y_i\}$ es necesario obtener un modelo $f(x_i)$ tal que satisfaga la medida de desempeño.

$$\begin{Bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_i & y_i \\ \vdots & \vdots \\ x_n & y_n \end{Bmatrix}$$

$$\hat{y} = f(x)$$

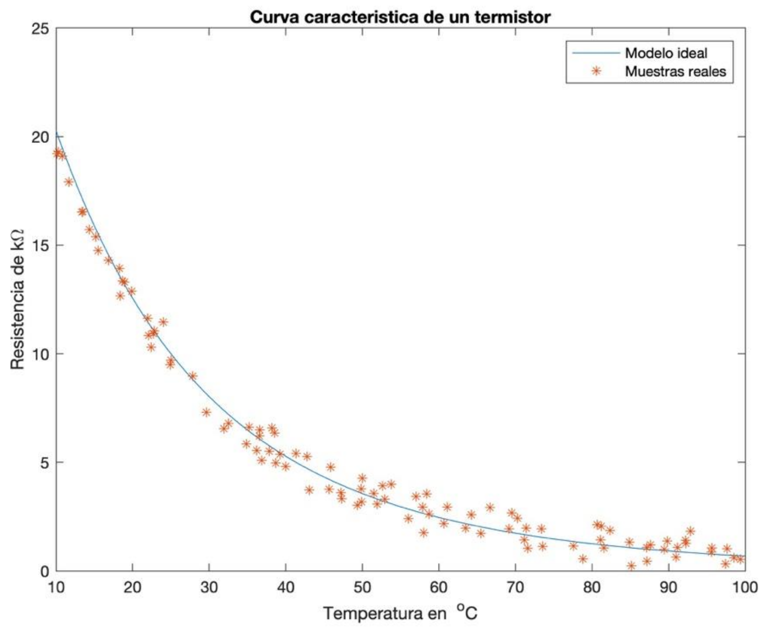
$$\hat{y_i} = f(x_i)$$

$$\varepsilon_i = y_i - \hat{y_i} = y_i - f(x_i)$$

$$SSE = \sum \varepsilon^2 = \sum \left(y_i - f(x_i)\right)^2$$

$$MSE = \frac{SSE}{n}$$

$$R^2 = 1 - \frac{MSE}{\sigma_y^2} \qquad \sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n} \qquad \bar{y} = \frac{\sum y_i}{n}$$



$$f(x) = mx + b$$

$$\mathrm{SSE} = \sum \left(y_i - f(x_i)\right)^2$$

$$SSE = \sum (y_i - mx_i - b)^2$$

$$SSE = \sum \left(y_i^2 - 2mx_iy_i - 2by_i + m^2x_i^2 + 2mbx_i + b^2\right)$$

$$\mathrm{SSE} = \sum y_i^2 - 2m\sum x_iy_i - 2b\sum y_i + m^2\sum x_i^2 + 2mb\sum x_i + nb^2$$

$$\frac{\partial SSE}{\partial b} = 0 \qquad \frac{\partial SSE}{\partial m} = 0$$

$$\frac{\partial SSE}{\partial b} = -2\sum y_i + 2m\sum x_i + 2nb \qquad \frac{\partial SSE}{\partial m} = -2\sum x_iy_i + 2m\sum x_i^2 + 2b\sum x_i$$

$$\frac{\partial SSE}{\partial b} = -2\sum y_i + 2m\sum x_i + 2nb$$

$$\frac{\partial SSE}{\partial m} = -2\sum x_i y_i + 2m\sum x_i^2 + 2b\sum x_i$$

$$m\sum x_i + nb = \sum y_i$$

$$m\sum x_i^2 + b\sum x_i = \sum x_i y_i$$

$$\begin{bmatrix} \sum x_i & n \\ \sum x_i^2 & \sum x_i \end{bmatrix} \times \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

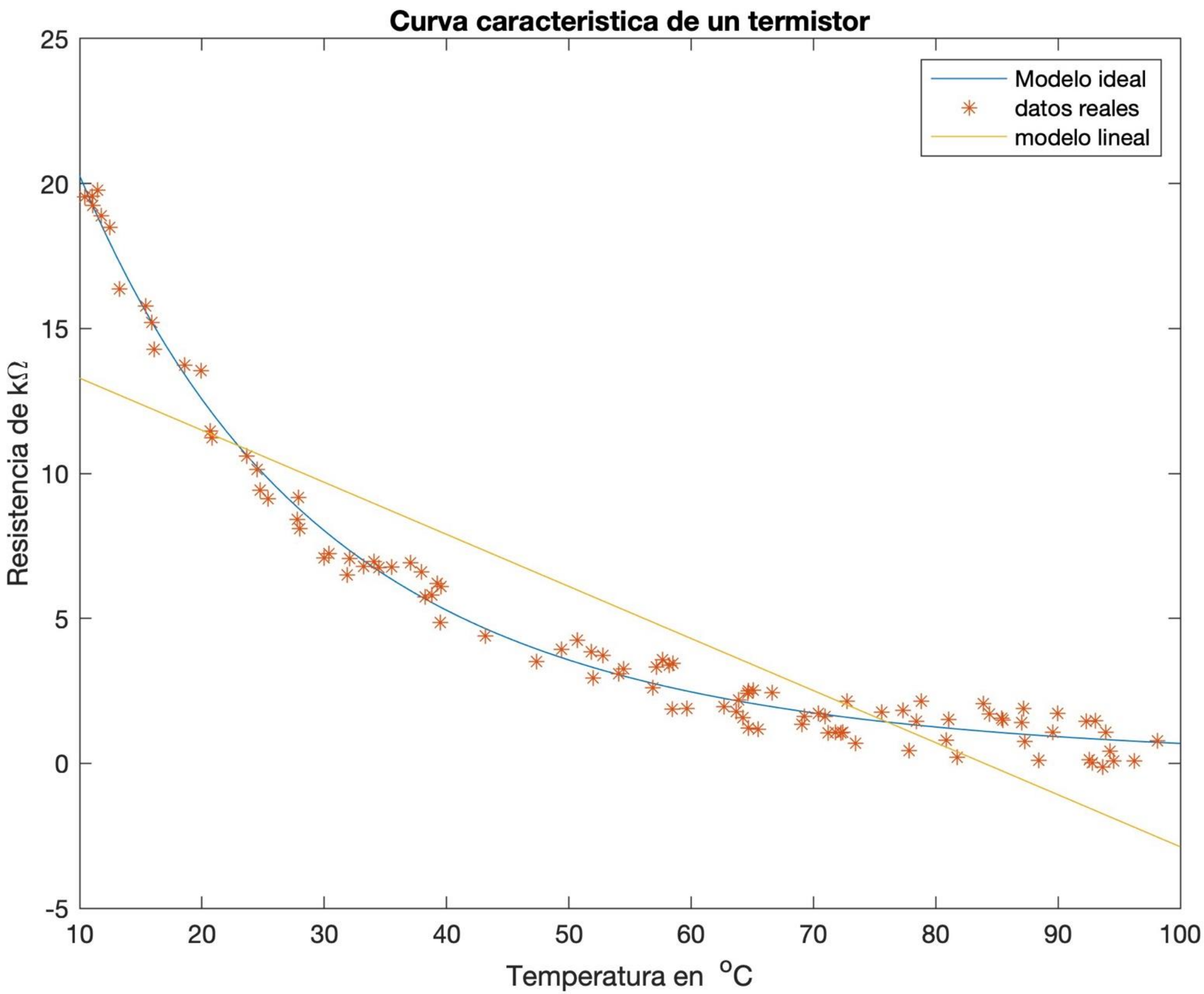
$$b = \frac{\sum y_i - m\sum x_i}{n} = \bar{y} - m\bar{x}$$

$$m = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$SSE = \sum \varepsilon^2 = \sum \big(y_i - f(x_i)\big)^2$$

$$R^2 = 1 - \frac{MSE}{\sigma_y^2}$$

$$\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n}$$


Curva caracteristica de un termistor
Modelo ideal
datos reales
modelo lineal
Resistencia de kΩ
Temperatura en °C
25
20
15
10
5
0
-5
10
20
30
40
50
60
70
80
90
100

$$\begin{bmatrix} \sum x_i & n \\ \sum x_i^2 & \sum x_i \end{bmatrix} \times \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$\begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_n) \end{bmatrix} = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & n \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}$$

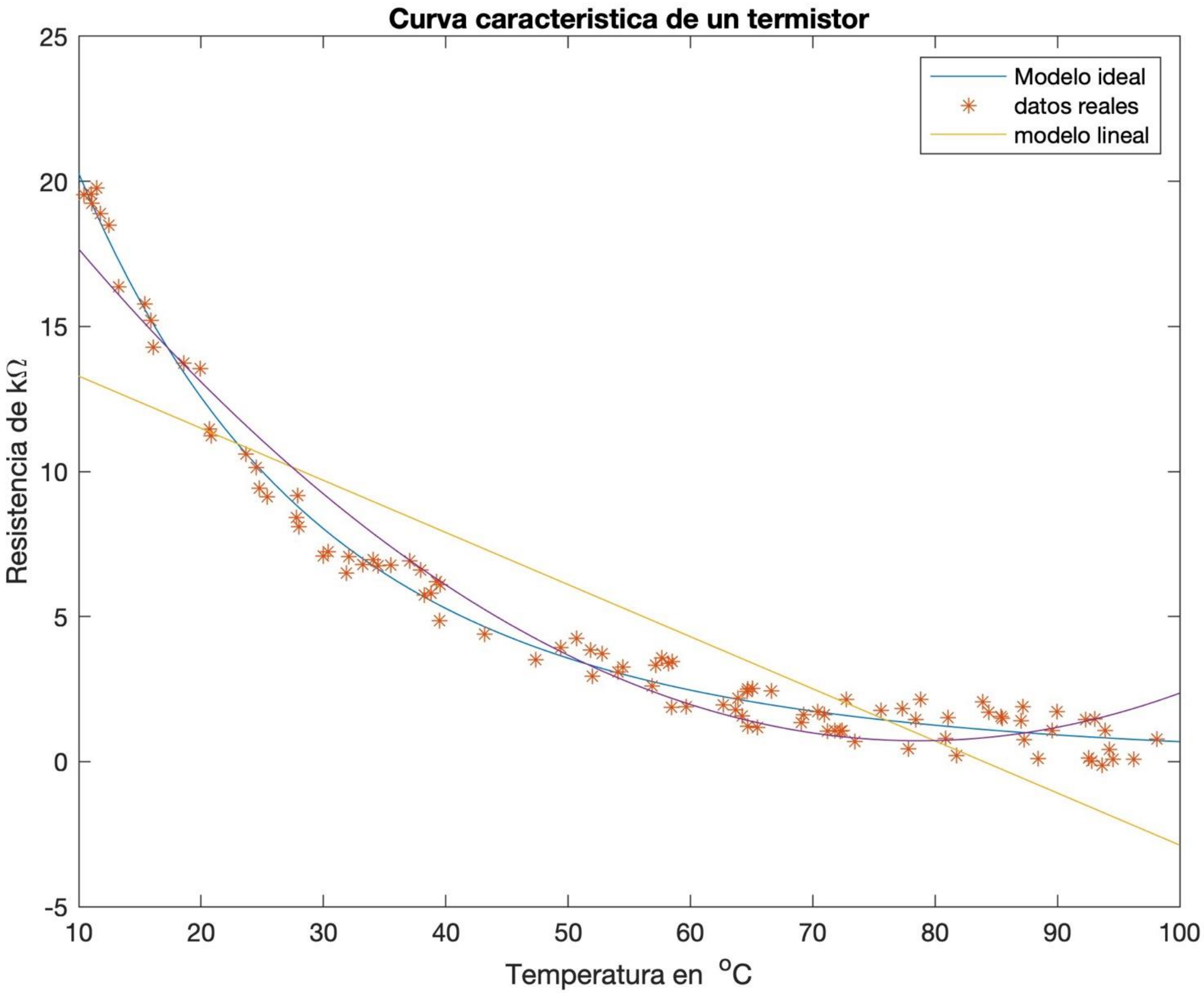
$$f_2(x) = a_2x^2 + a_1x + a_0 = \begin{bmatrix} x^2 & x & 1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}$$

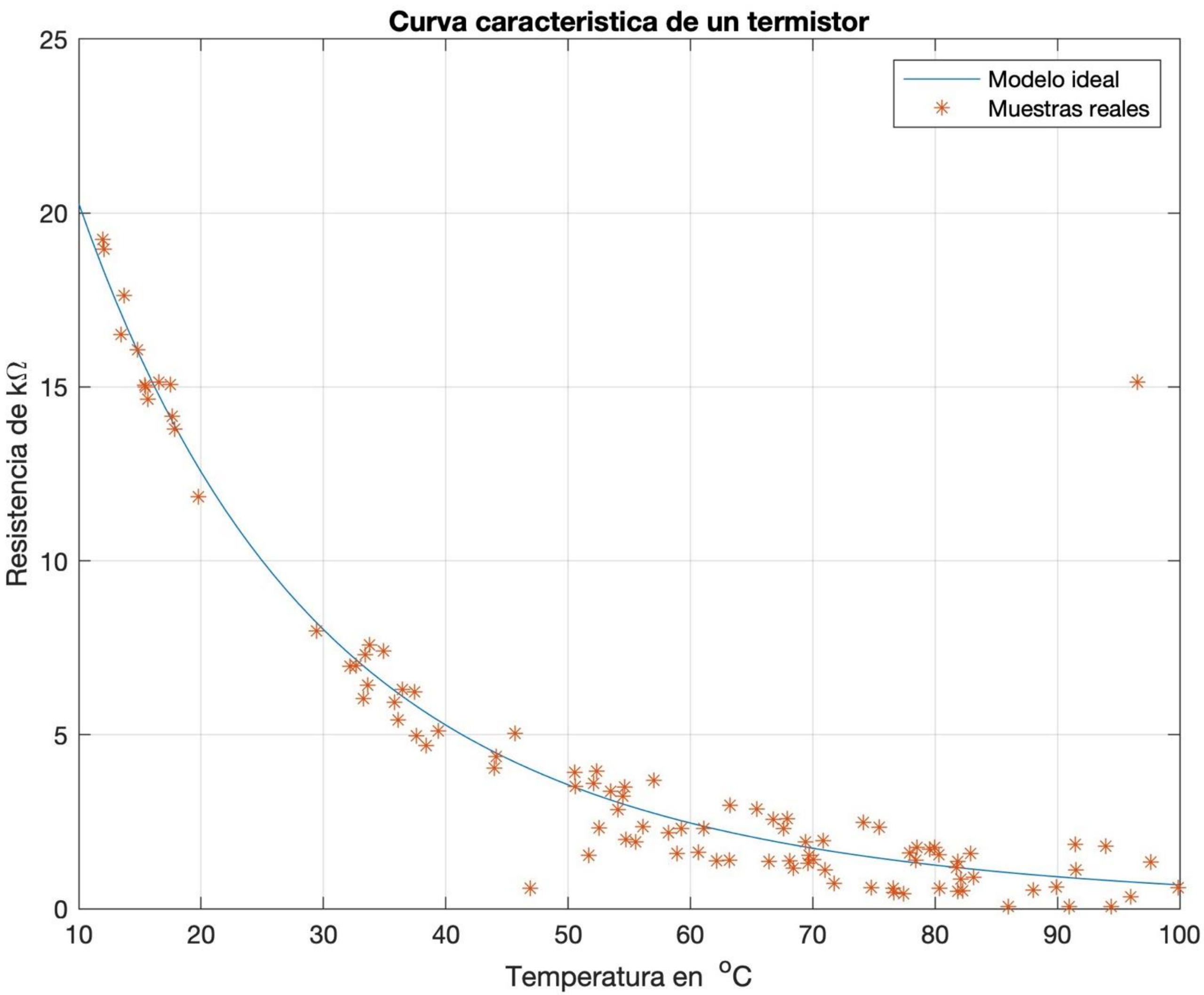
$$\begin{bmatrix} \sum x_i^2 & \sum x_i & n \\ \sum x_i^3 & \sum x_i^2 & \sum x_i \\ \sum x_i^4 & \sum x_i^3 & \sum x_i^2 \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

$$SSE = \sum \varepsilon^2 = \sum \left(y_i - f(x_i)\right)^2$$

$$R^2 = 1 - \frac{MSE}{\sigma_y^2}$$

$$\sigma_y^2 = \sum (y_i - \bar{y})^2$$


Curva caracteristica de un termistor
Modelo ideal
datos reales
modelo lineal
Resistencia de kΩ
Temperatura en °C
25
20
15
10
5
0
-5
10
20
30
40
50
60
70
80
90
100



Curva caracteristica de un termistor
Modelo ideal
Muestras reales
Resistencia de kΩ
Temperatura en °C
0
5
10
15
20
25
10
20
30
40
50
60
70
80
90
100

$$\widehat{y_{i,j}} = f(x_i \in \{x_i, y_i\} | i \neq j)$$

$$f_2(x) = a_2x^2 + a_1x + a_0 = \begin{bmatrix} x^2 & x & 1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}$$

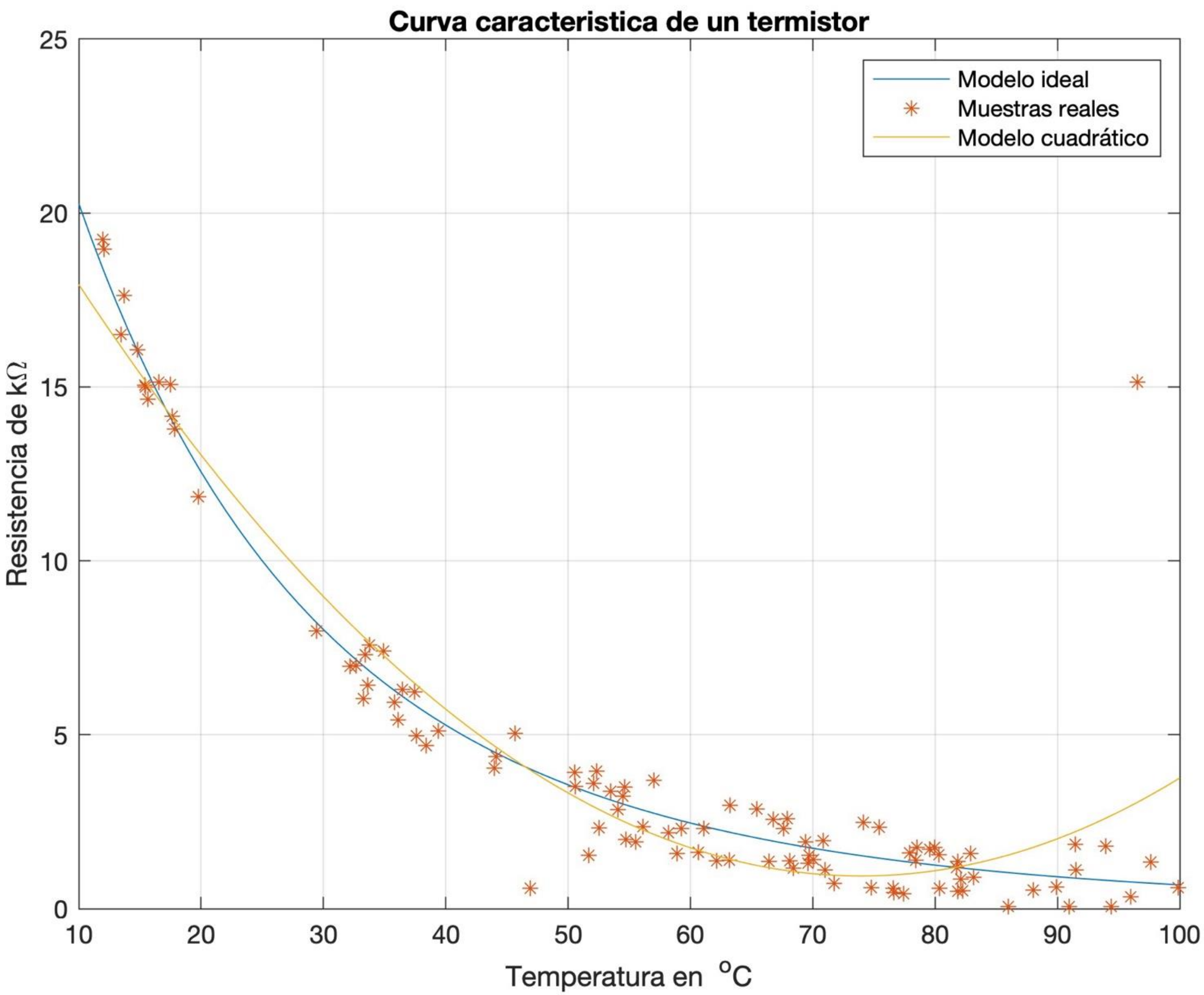
$$\begin{bmatrix} \sum x_i^2 & \sum x_i & n \\ \sum x_i^3 & \sum x_i^2 & \sum x_i \\ \sum x_i^4 & \sum x_i^3 & \sum x_i^2 \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

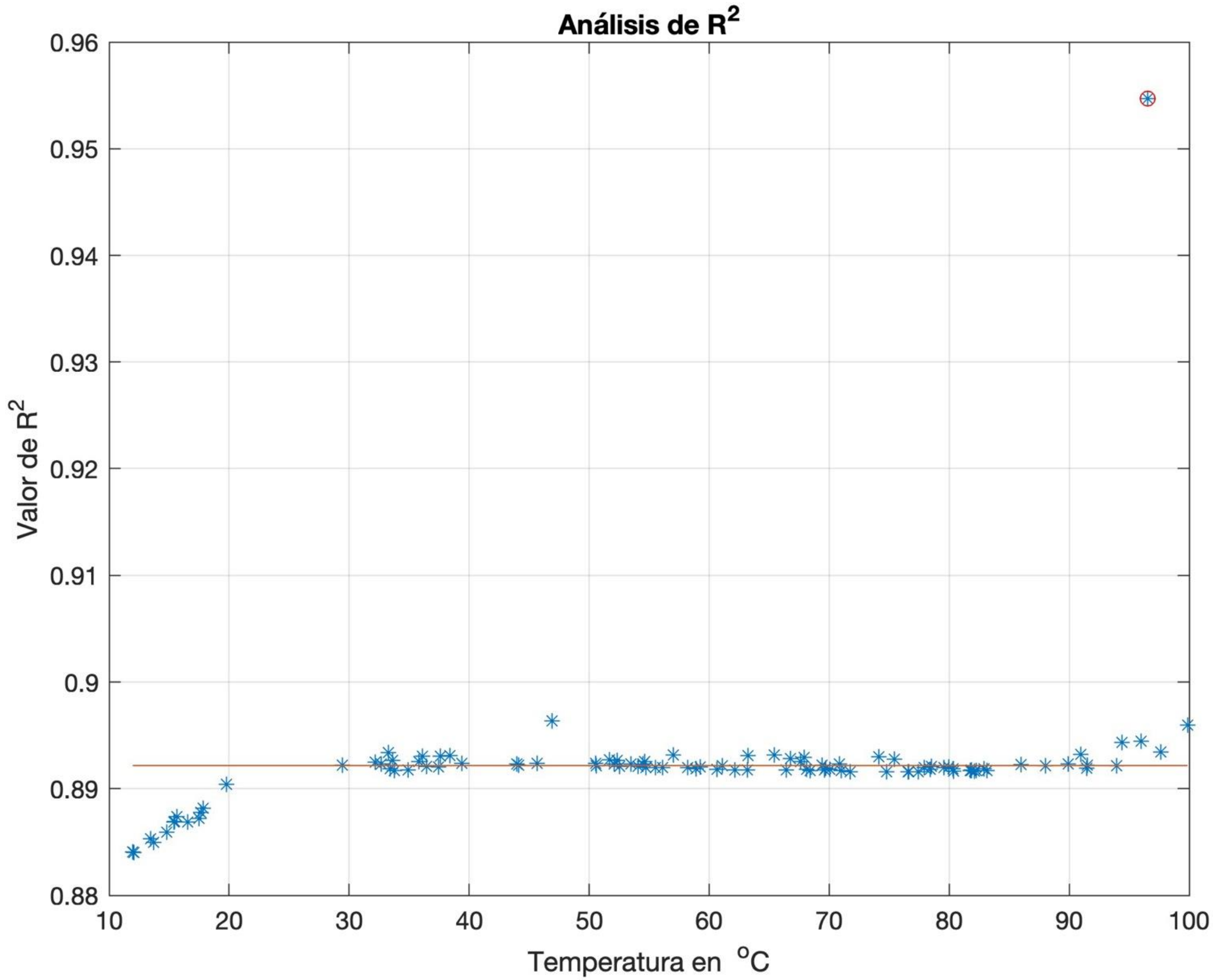
$$SSE = \sum \varepsilon^2 = \sum \left(y_i - f(x_i)\right)^2$$

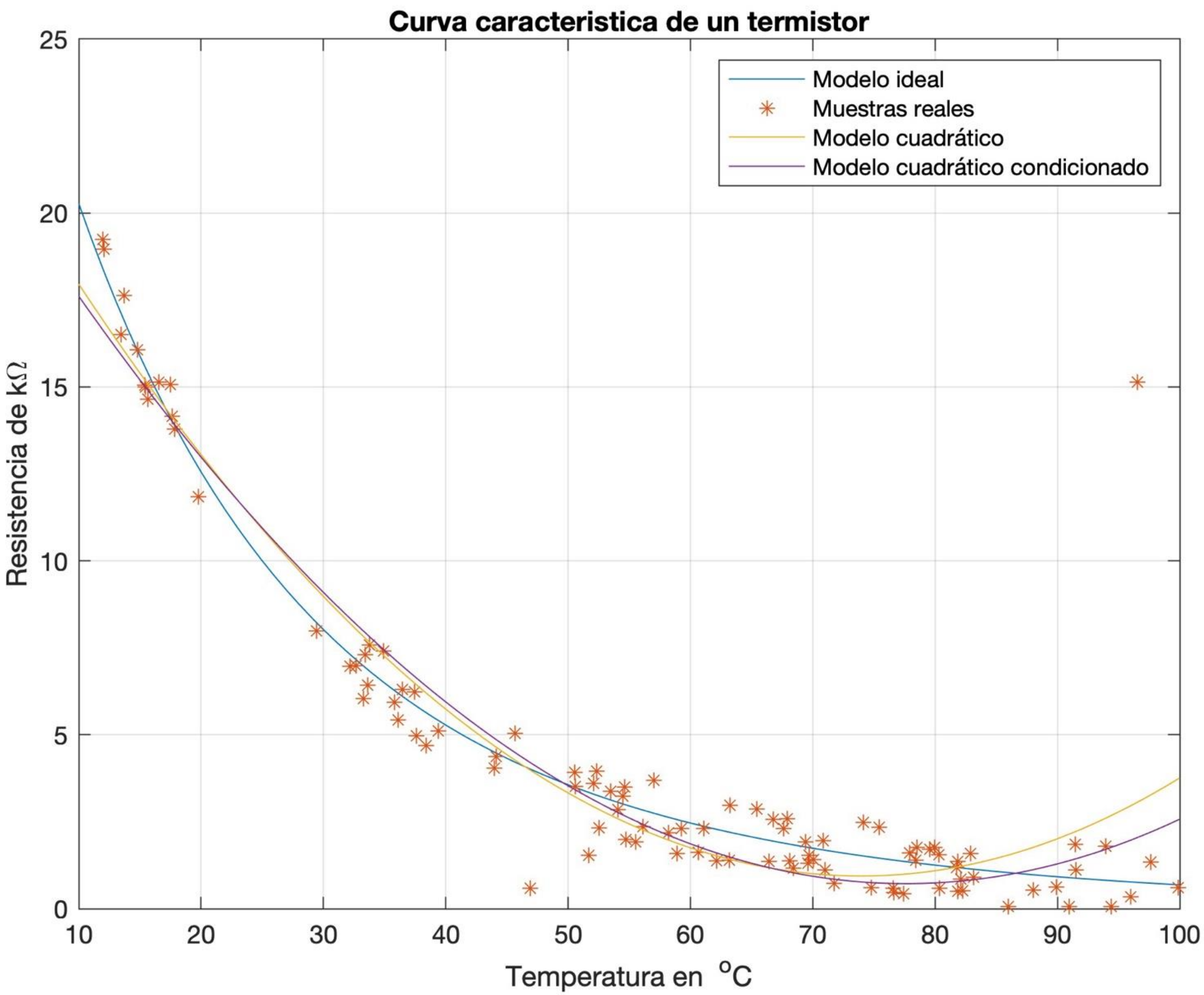
$$\sigma_y^2 = \sum (y_i - \bar{y})^2$$

$$R^2 = 1 - \frac{MSE}{\sigma_y^2}$$

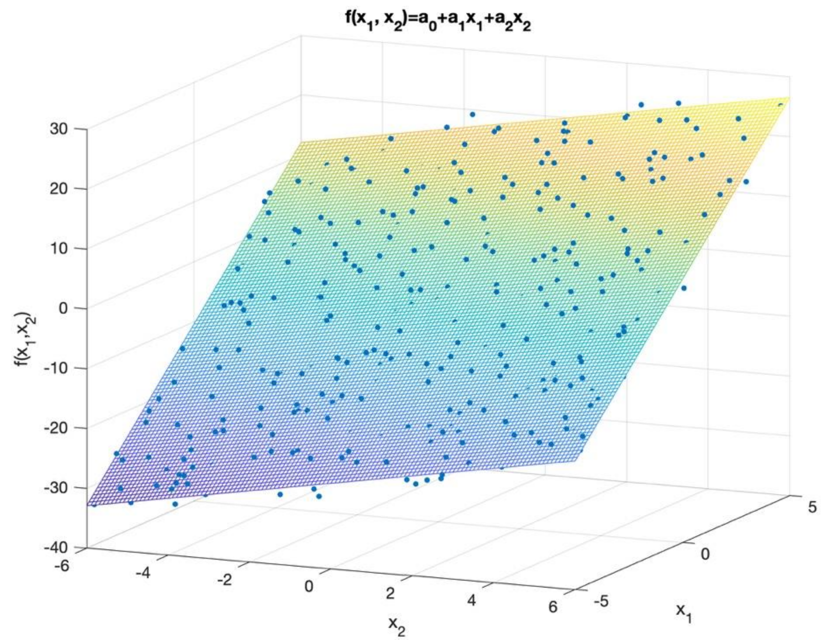
Ajuste de polinomios y funcional / Regresión




Análisis de R2
Valor de R2
Temperatura en °C
0.96
0.95
0.94
0.93
0.92
0.91
0.9
0.89
0.88
10
20
30
40
50
60
70
80
90
100


Curva caracteristica de un termistor
Modelo ideal
Muestras reales
Modelo cuadrático
Modelo cuadrático condicionado
Resistencia de kΩ
Temperatura en °C
25
20
15
10
5
0
10
20
30
40
50
60
70
80
90
100

Dado un conjunto de puntos $\{x_1(i), x_2(i), y(i)\}$ es necesario obtener un modelo $f(x_1, x_2)$ tal que satisfaga la medida de desempeño.



$$\begin{Bmatrix} x_1(1) & x_2(1) & y(1) \\ x_1(2) & x_2(2) & y(2) \\ \vdots & \vdots & \vdots \\ x_1(i) & x_2(i) & y(i) \\ \vdots & \vdots & \vdots \\ x_1(n) & x_2(n) & y(n) \end{Bmatrix}$$

$$\hat{y} = f(x_1\,, x_2)$$

$$f(x_1, x_2) = a_0 + a_1x_1 + a_2x_2$$

$$\widehat{y(i)} = f\big(x_1(i), x_2(i)\big)$$

$$f\big(x_1(i), x_2(i)\big) = a_0 + a_1x_1(i) + a_2x_2(i)$$

$$\varepsilon(i) = y(i) - \widehat{y(i)} = y(i) - f\big(x_1(i), x_2(i)\big)$$

$$SSE = \sum \varepsilon(i)^2 = \sum \Big(y(i) - f\big(x_1(i), x_2(i)\big)\Big)^2$$

$$MSE = \frac{\sqrt{SSE}}{n} \qquad R^2 = 1 - \frac{SSE}{\sigma_y^2}$$

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2$$

$$SSE = \sum \varepsilon(i)^2 = \sum \Big(y(i) - f\big(x_1(i), x_2(i)\big)\Big)^2$$

$$SSE = \sum \big(y(i) - a_0 - a_1 x_1(i) - a_2 x_2(i)\big)^2$$

$$SSE = \sum [y(i)^2 - 2a_0 y(i) - 2a_1 x_1(i) y(i) - 2a_2 x_2(i) y(i) + a_0^2 + 2a_0 a_1 x_1(i) + \cdots$$
$$\ldots + 2a_0 a_2 x_2(i) + a_1^2 x_1^2(i) + 2a_1 a_2 x_1(i) x_2(i) + a_2^2 x_2^2(i)]$$

$$\frac{\partial SSE}{\partial a_0} = -2\sum y(i) + 2na_0 + 2a_1 \sum x_1(i) + 2a_2 \sum x_2(i)$$

$$\frac{\partial SSE}{\partial a_1} = -2\sum x_1(i) y(i) + 2a_0 \sum x_1(i) + 2a_1 \sum x_1(i)^2 + 2a_2 \sum x_1(i) x_2(i)$$

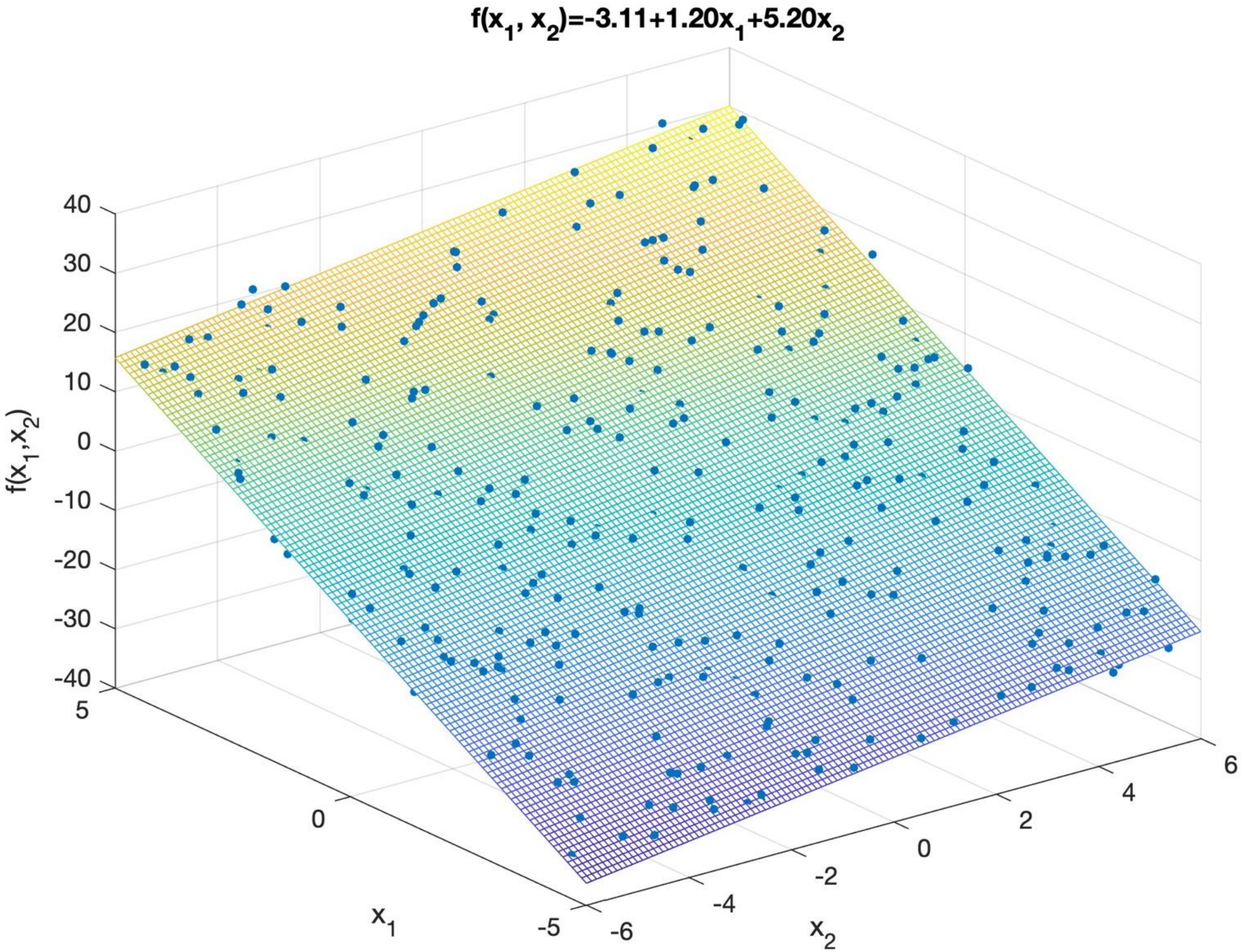
$$\frac{\partial SSE}{\partial a_2} = -2\sum x_2(i) y(i) + 2a_0 \sum x_2(i) + 2a_1 \sum x_1(i) x_2(i) + 2a_2 \sum x_2(i)^2$$

$$na_0 + a_1 \sum x_1(i) + a_2 \sum x_2(i) = \sum y(i)$$

$$a_0 \sum x_1(i) + a_1 \sum x_1(i)^2 + a_2 \sum x_1(i) x_2(i) = \sum x_1(i) y(i)$$

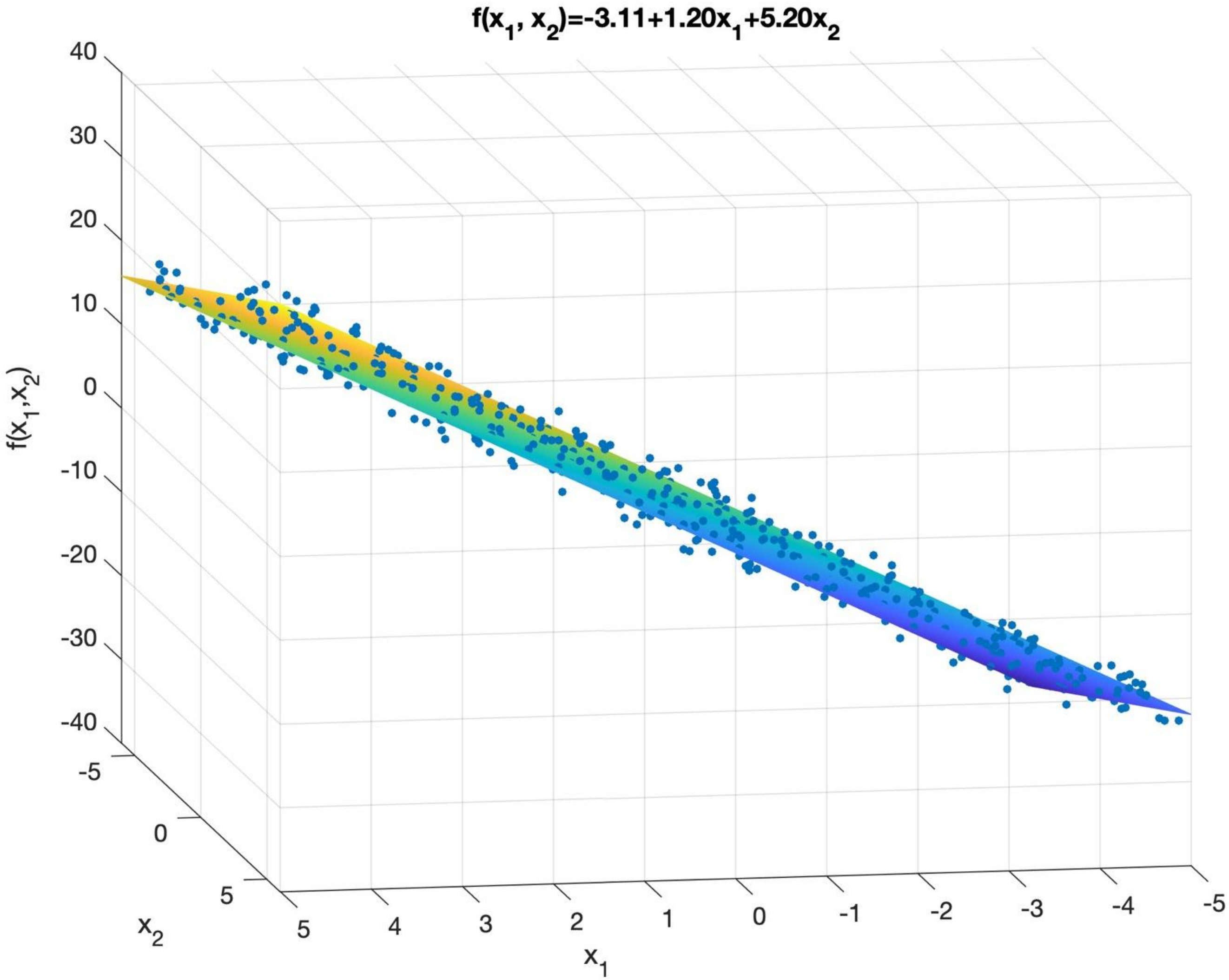
$$a_0 \sum x_2(i) + a_1 \sum x_1(i) x_2(i) + a_2 \sum x_2(i)^2 = \sum x_2(i) y(i)$$

$$\begin{bmatrix} n & \sum x_1(i) & \sum x_2(i) \\ \sum x_1(i) & \sum x_1(i)^2 & \sum x_1(i) x_2(i) \\ \sum x_2(i) & \sum x_1(i) x_2(i) & \sum x_2(i)^2 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y(i) \\ \sum x_1(i) y(i) \\ \sum x_2(i) y(i) \end{bmatrix}$$

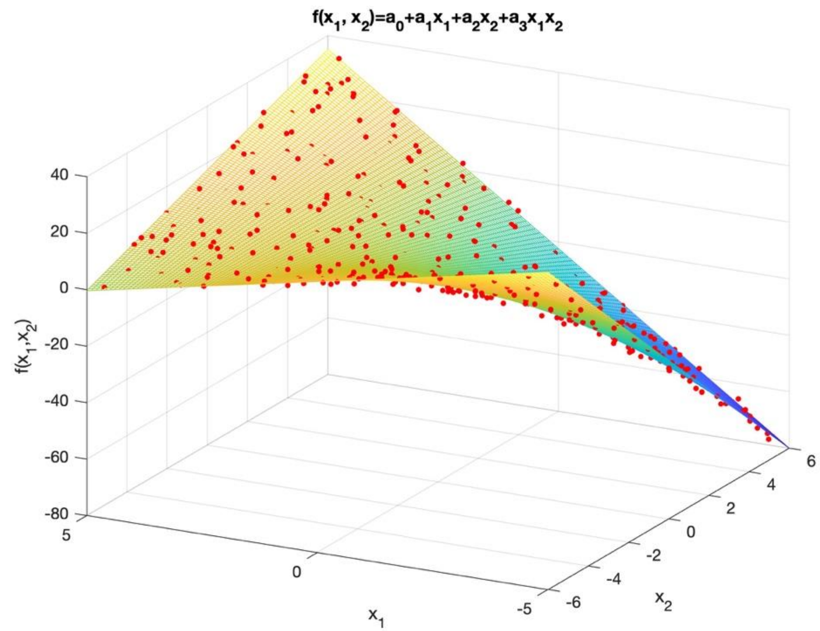


f(x1, x2)=-3.11+1.20x1+5.20x2
40
30
20
10
0
-10
-20
-30
-40
f(x1,x2)
5
0
-5
x1
-6
-4
-2
0
2
4
6
x2

Ajuste de polinomios y funcional / Regresión



Dado un conjunto de puntos $\{x_1(i), x_2(i), y(i)\}$ es necesario obtener un modelo $f(x_1, x_2)$ tal que satisfaga la medida de desempeño.



$$\begin{Bmatrix} x_1(1) & x_2(1) & y(1) \\ x_1(2) & x_2(2) & y(2) \\ \vdots & \vdots & \vdots \\ x_1(i) & x_2(i) & y(i) \\ \vdots & \vdots & \vdots \\ x_1(n) & x_2(n) & y(n) \end{Bmatrix}$$

$$\hat{y} = f(x_1\,, x_2)$$

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + \boldsymbol{a_3 x_1 x_2}$$

$$\widehat{y(i)} = f\big(x_1(i), x_2(i)\big)$$

$$f\big(x_1(i), x_2(i)\big) = a_0 + a_1 x_1(i) + a_2 x_2(i) + \boldsymbol{a_3 x_1(i) x_2(i)}$$

$$\varepsilon(i) = y(i) - \widehat{y(i)} = y(i) - f\big(x_1(i), x_2(i)\big)$$

$$SSE = \sum \varepsilon(i)^2 = \sum \Big(y(i) - f\big(x_1(i), x_2(i)\big)\Big)^2$$

$$MSE = \frac{\sqrt{SSE}}{n}$$

$$R^2 = 1 - \frac{SSE}{\sigma_y^2}$$

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2$$

$$SSE = \sum \varepsilon(i)^2 = \sum \Big(y(i) - f\big(x_1(i), x_2(i)\big)\Big)^2$$

$$SSE = \sum \big(y(i) - a_0 - a_1 x_1(i) - a_2 x_2(i) - a_3 x_1(i) x_2(i)\big)^2$$

$$SSE = \sum [y(i)^2 - 2a_0 y(i) - 2a_1 x_1(i) y(i) - 2a_2 x_2(i) y(i) - 2a_3 x_1(i) x_2(i) y(i) + a_0^2 + \cdots$$

$$\ldots + 2a_0 a_1 x_1(i) + 2a_0 a_2 x_2(i) + 2a_0 a_3 x_1(i) x_2(i) + a_1^2 x_1^2(i) + 2a_1 a_2 x_1(i) x_2(i) + \cdots$$

$$\ldots + 2a_1 a_3 x_1^2(i) x_2(i) + a_2^2 x_2^2(i) + 2a_2 a_3 x_1(i) x_2^2(i) + a_3^2 x_1^2(i) x_2^2(i)]$$

$$\frac{\partial SSE}{\partial a_0} = -2\sum y(i) + 2n a_0 + 2a_1 \sum x_1(i) + 2a_2 \sum x_2(i) + 2a_3 \sum x_1(i) x_2(i)$$

$$\frac{\partial SSE}{\partial a_1} = -2\sum x_1(i) y(i) + 2a_0 \sum x_1(i) + 2a_1 \sum x_1(i)^2 + 2a_2 \sum x_1(i) x_2(i) + 2a_3 \sum x_1(i)^2 x_2(i)$$

$$\frac{\partial SSE}{\partial a_2} = -2\sum x_2(i) y(i) + 2a_0 \sum x_2(i) + 2a_1 \sum x_1(i) x_2(i) + 2a_2 \sum x_2(i)^2 + 2a_3 \sum x_1(i) x_2(i)^2$$

$$\frac{\partial SSE}{\partial a_3} = -2\sum x_1(i) x_2(i) y(i) + 2a_0 \sum x_1(i) x_2(i) + 2a_1 \sum x_1(i)^2 x_2(i) + 2a_2 \sum x_1(i) x_2(i)^2 + \cdots$$

$$\ldots + 2a_3 \sum x_1(i)^2 x_2(i)^2$$

$$na_0 + a_1 \sum x_1(i) + a_2 \sum x_2(i) + a_3 \sum x_1(i)x_2(i) = \sum y(i)$$

$$a_0 \sum x_1(i) + a_1 \sum x_1(i)^2 + a_2 \sum x_1(i)x_2(i) + a_3 \sum x_1(i)^2 x_2(i) = \sum x_1(i)y(i)$$

$$a_0 \sum x_2(i) + a_1 \sum x_1(i)x_2(i) + a_2 \sum x_2(i)^2 + a_3 \sum x_1(i)x_2(i)^2 = \sum x_2(i)y(i)$$

$$a_0 \sum x_1(i)x_2(i) + a_1 \sum x_1(i)^2 x_2(i) + a_2 \sum x_1(i)x_2(i)^2 + a_3 \sum x_1(i)^2 x_2(i)^2 = \sum x_1(i)x_2(i)y(i)$$

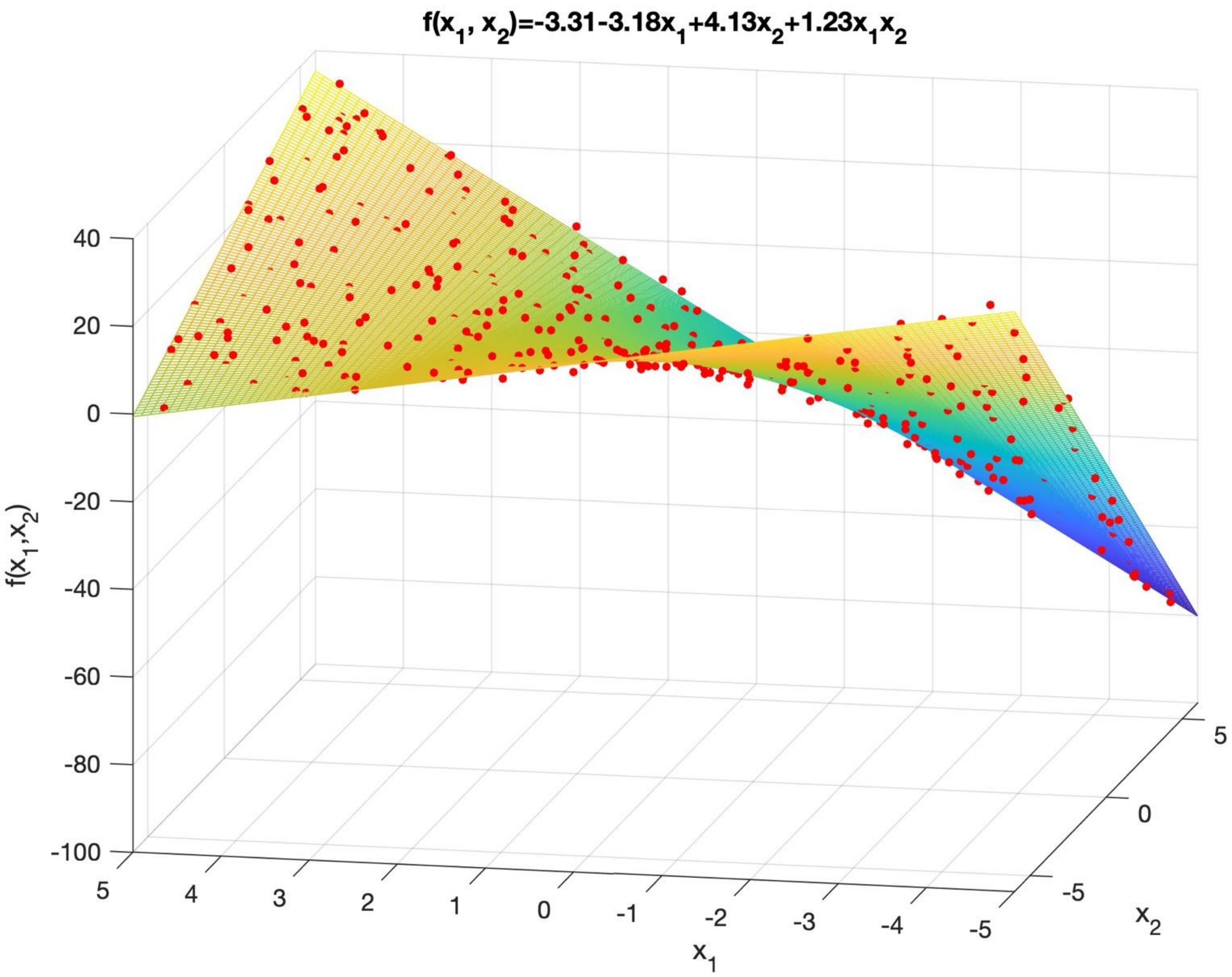
$$\begin{bmatrix} n & \sum x_1(i) & \sum x_2(i) & \sum x_1(i)x_2(i) \\ \sum x_1(i) & \sum x_1(i)^2 & \sum x_1(i)x_2(i) & \sum x_1(i)^2 x_2(i) \\ \sum x_2(i) & \sum x_1(i)x_2(i) & \sum x_2(i)^2 & \sum x_1(i)x_2(i)^2 \\ \sum x_1(i)x_2(i) & \sum x_1(i)^2 x_2(i) & \sum x_1(i)x_2(i)^2 & \sum x_1(i)^2 x_2(i)^2 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y(i) \\ \sum x_1(i)y(i) \\ \sum x_2(i)y(i) \\ \sum x_1(i)x_2(i)y(i) \end{bmatrix}$$

$$\begin{bmatrix} X_L^T X_L & 0 & h_L^T \\ 0 & X_R^T X_R & -h_R^T \\ h_L & -h_R & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ \lambda \end{bmatrix} = \begin{bmatrix} X_L^T y_L \\ X_R^T y_R \\ 0 \end{bmatrix}$$

$$h_L a - h_R b = 0$$

$$h_L = \left[1, x_{th}, x_{th}^2, \dots, x_{th}^{p_L}\right]$$

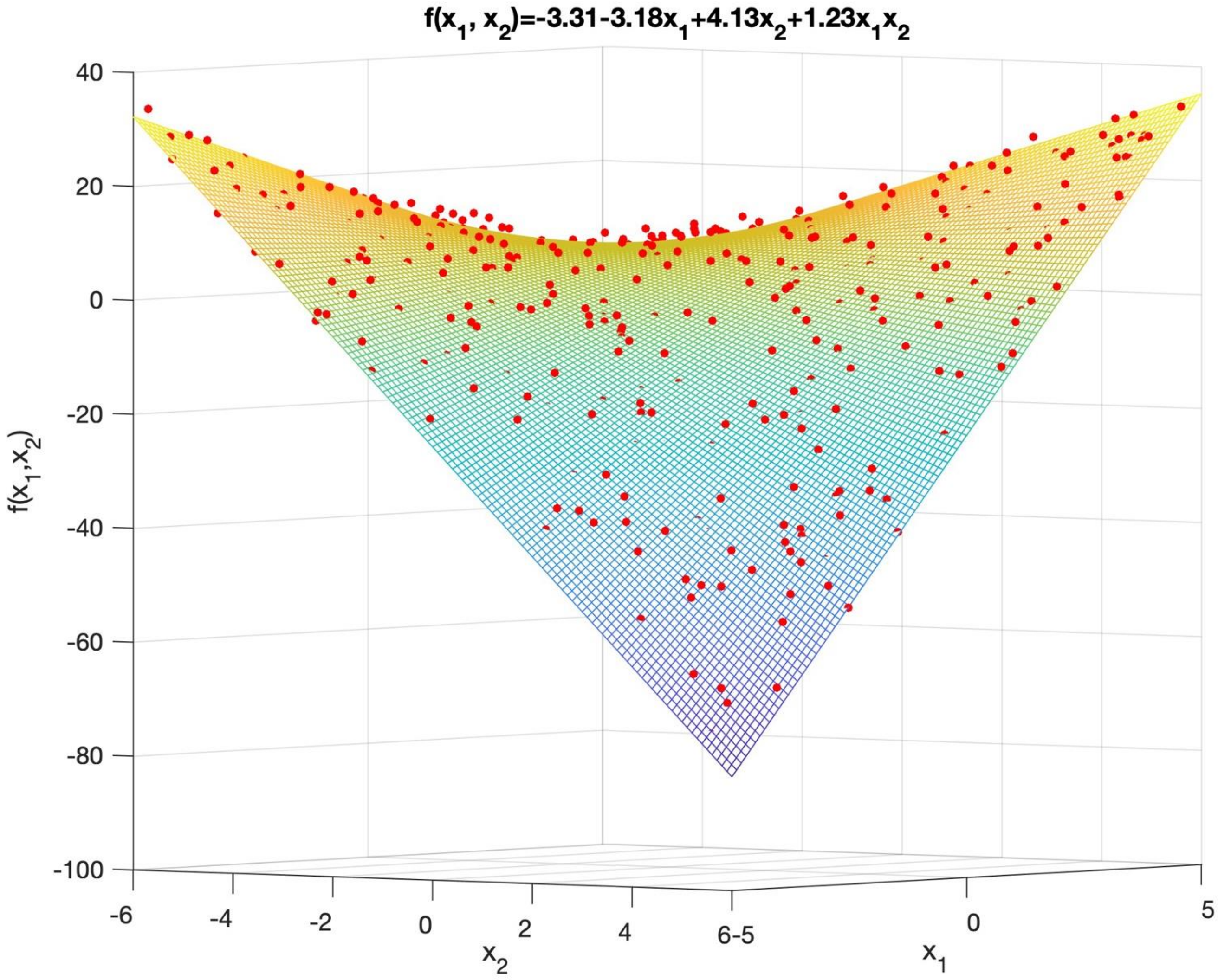
$$h_R = \left[1, x_{th}, x_{th}^2, \dots, x_{th}^{p_R}\right]$$

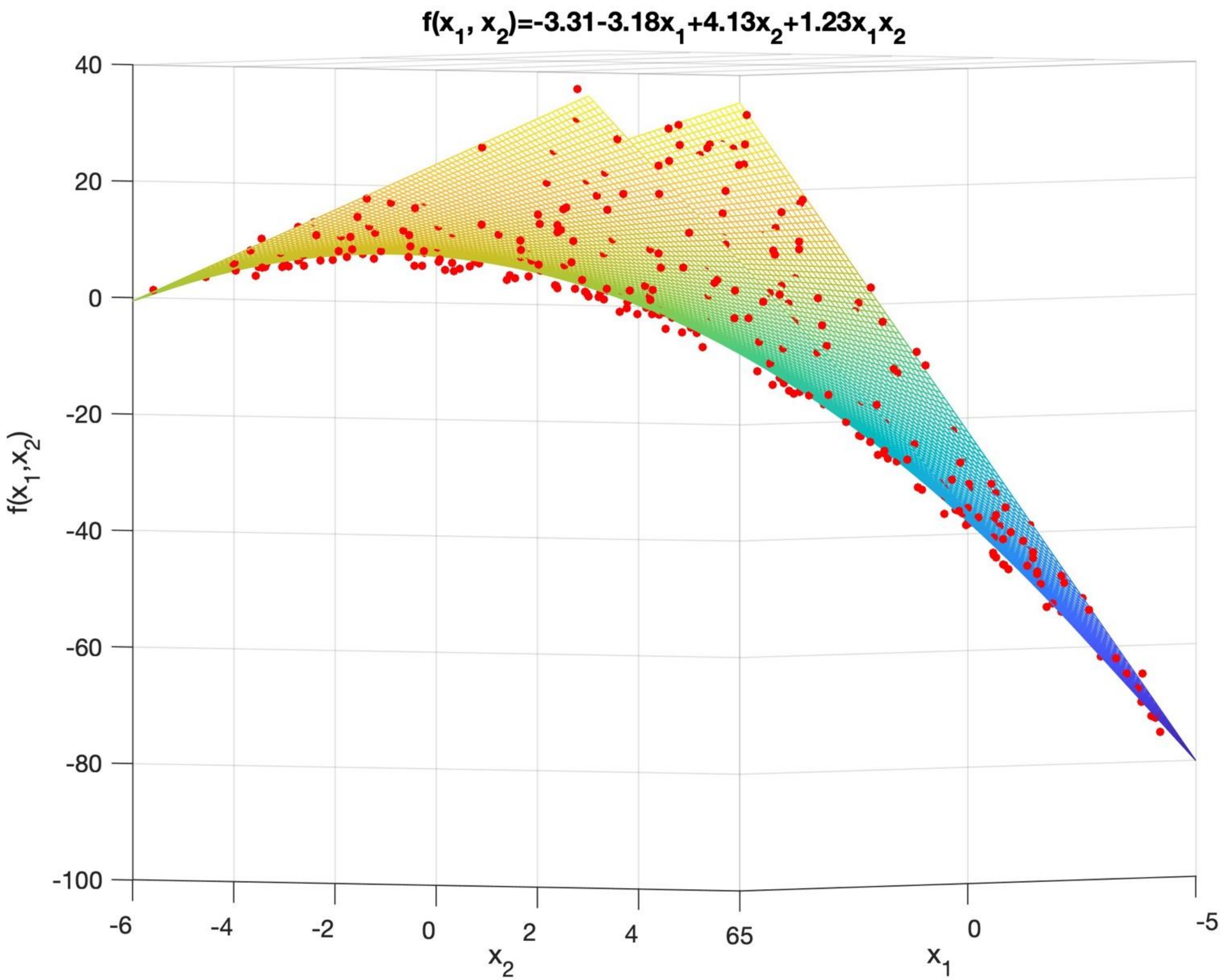


f(x_1, x_2)=-3.31-3.18x_1+4.13x_2+1.23x_1x_2
40
20
0
-20
-40
-60
-80
-100
f(x_1,x_2)
5
4
3
2
1
0
-1
-2
-3
-4
-5
x_1
5
0
-5
x_2

Ajuste de polinomios y funcional / Regresión

UNIVERSIDAD DE GUANAJUATO





f(x1, x2)=-3.31-3.18x1+4.13x2+1.23x1x2
40
20
0
-20
-40
-60
-80
-100
f(x1,x2)
-6
-4
-2
0
2
4
6
5
0
-5
x2
x1

La interpolación consisten en obtener nuevos puntos partiendo del conocimiento de un determinado conjunto de puntos. Los algoritmos de interpolación permite estimar valores intermedios entre datos precisos.

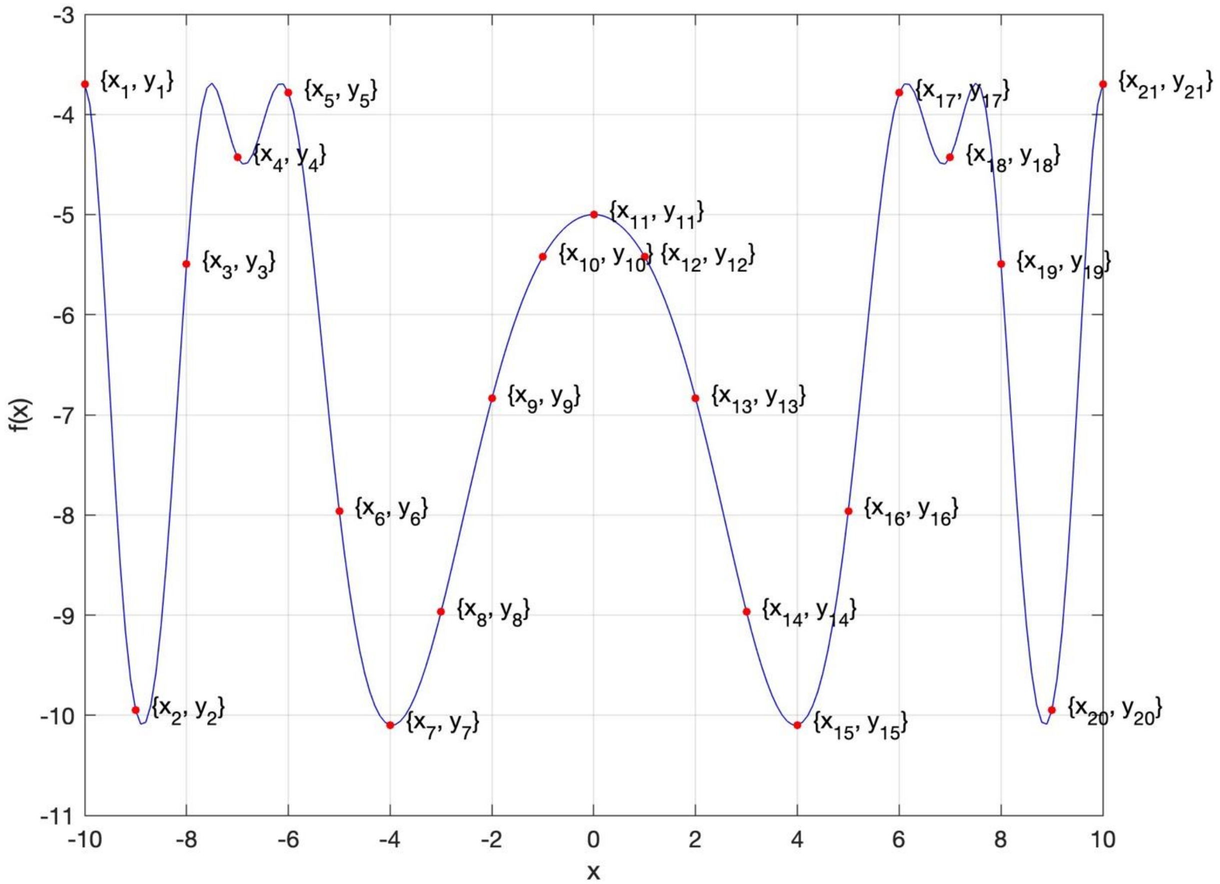
En ingenierías y algunas ciencias es frecuente disponer de un número de puntos obtenidos por muestreo o a partir de un experimento y pretender construir una función que los ajuste.

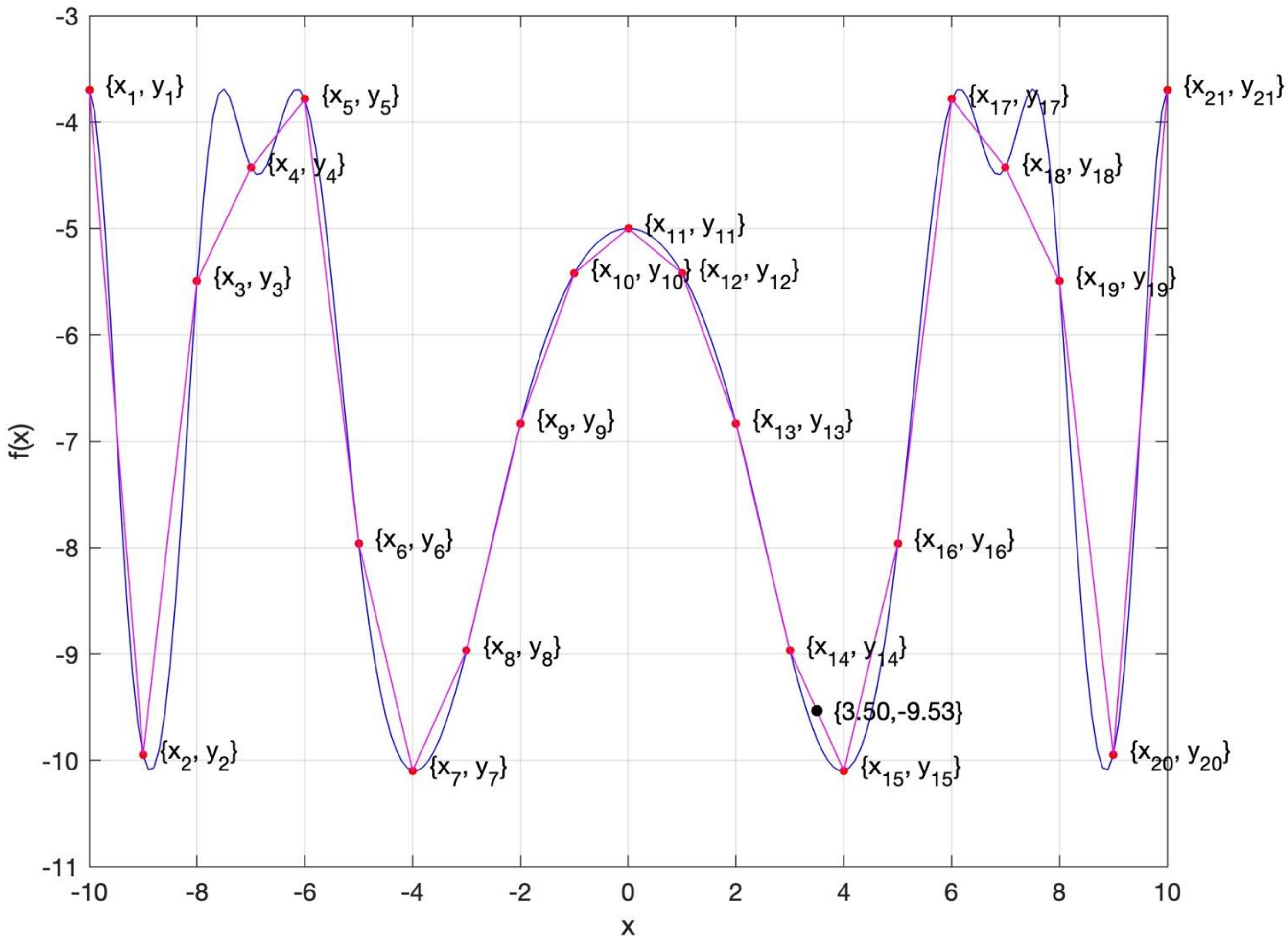
La interpolación es una forma reducida de estimar los valores de funciones complejas, reduciendo el costo computacional debido a la reducción de valores considerados, esto nos permite construir una función más simple. En general, no es posible obtener los mismos valores si evaluamos la función obtenida que la función original, esto debido a que depende de las características del problema, la selección de los puntos muestreados y del método de interpolación.

La interpolación es un método numérico que permite estimar un valor desconocido entre un conjunto de valores conocidos.

Dado un conjunto de puntos $\{x_i, y_i\}$ es posible estimar un valor deseado intermedio considerando los puntos conocidos.

$$\begin{Bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_i & y_i \\ \vdots & \vdots \\ x_n & y_n \end{Bmatrix}$$





$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{\hat{f}(x) - f(x_0)}{x - x_0}$$

$$\hat{f}(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)$$

$$f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1)$$

$$f_2(x) = a_0 + a_1 x + a_2 x^2$$

$$a_0 = b_0 - b_1 x_0 + b_2 x_0 x_1$$

$$a_1 = b_1 - b_2 x_0 - b_2 x_1$$

$$a_2 = b_2$$

$$b_0 = f(x_0)$$

$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$$b_2 = \frac{\dfrac{f(x_2) - f(x_1)}{x_2 - x_1} - \dfrac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_1}$$

Interpolación de polinomios de Newton

$$f_n(x) = f(x_0) + (x - x_0)f[x_1, x_0] + (x - x_0)(x - x_1)f[x_2, x_1, x_0] + \cdots$$

$$\ldots + (x - x_0)(x - x_1)\ldots(x - x_n)f[x_n, x_{n-1}, \ldots, x_0]$$

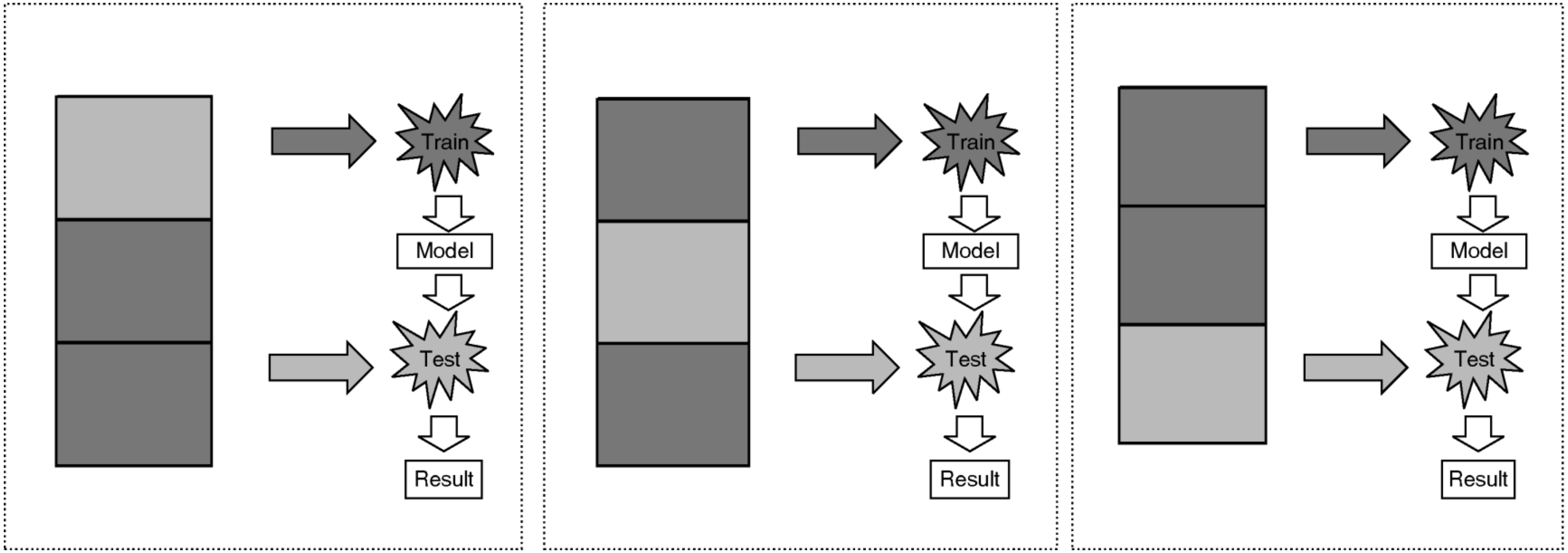
$$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{x_i - x_j}$$

$$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{x_i - x_k}$$

Interpolación de polinomios de Lagrange

$$f_n(x) = \sum_{i=0}^{n} L_i(x)f(x_i)$$

$$L_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$



Proceso para validación cruzada

Validation method	Pros	Cons
Resubstitution Validation	Simple	Over-fitting
Hold-out Validation	Independent training and test	Reduced data for training and testing; Large variance
k-fold cross validation	Accurate performance estimation	Small samples of performance estimation; Overlapped training data; Elevated Type I error for comparison; Underestimated performance variance or overestimated degree of freedom for comparison
Leave-One-Out cross-validation	Unbiased performance estimation	Very large variance
Repeated *k*-fold cross-validation	Large number of performance estimates	Overlapped training and test data between each round; Underestimated performance variance or overestimated degree of freedom for comparison

Métricas de desempeño

UNIVERSIDAD DE GUANAJUATO

Una matriz de confusión es una herramienta que permite visualizar el desempeño de una clasificador al mostrar el número la predicción para cada una de las muestras.

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

TP (True Positives): Número de muestras positivas reales que fueron clasificadas correctamente como positivas.
TN (True Negatives): Número de muestras negativas reales que fueron clasificadas correctamente como negativas.
FP (False Positives): Número de muestras negativas reales que fueron clasificadas incorrectamente como positivas. Es conocido también como error tipo 1 que es cuando un clasificador identifica algo como positivo cuando no lo es.
FN (False Negatives): Número de muestras positivas reales que fueron clasificadas incorrectamente como negativas. Es conocido como error tipo 2, es decir cuando el clasificador detecta algo como positivo cuando en realidad lo es.
Error de rechazo: es la medida de cuántas veces el clasificador no puedo tomar una decisión definitica sobre una observación. En cierto contexto, es preferible que el clasificador rechace hacer una clasificación en lugar de hacer una incorrecta. Esto es común en aplicaciones críticas donde un error de clasificación puede tener consecuencias significativas.

Exactitud (Accuracy): Es la proporción de predicciones correctas entre el número de muestras evaluadas

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

$$Exactitud = \frac{TP + TN}{TP + FN + FP + TN}$$

Precisión (Precision): Es la proporción de predicciones positivas que fueron realmente correctas.

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

$$Precisión = \frac{TP}{TP + FP}$$

Sensibilidad o Tasa de Verdaderos Positivos (Recall o True Positive Rate): Es la proporción de positivos reales que fueron correctamente clasificados.

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

$$Sensibilidad = \frac{TP}{TP + FN}$$

Especificidad o Tasa de Verdaderos Negativos (Specificity o True Negative Rate): Es la proporción de negativos correctamente clasificados.

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

$$Especificidad = \frac{TN}{FP + TN}$$

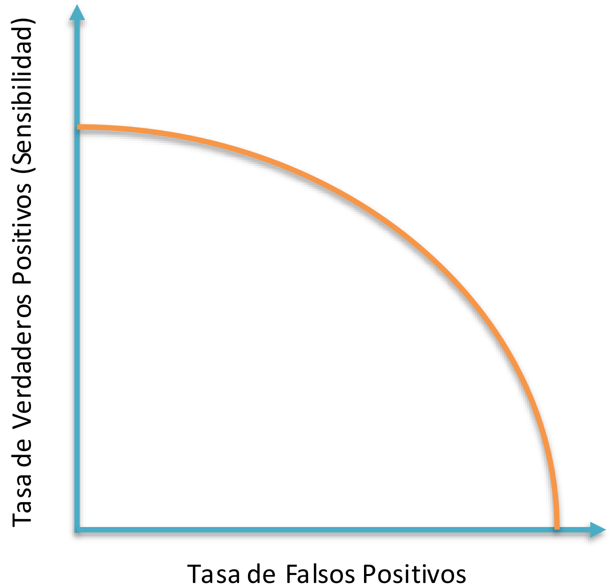
Puntaje F1 (F1 Score): Es la medida armónica de la precisión y la sensibilidad. Ofrece un balance entre la precisión y la sensibilidad.

		Predicciones	
		Positiva	Negativa
Reales	Positiva	Verdaderos positivos (TP – True Positives)	Falsos negativos (FN - False Negatives)
	Negativa	Falsos positivos (FP – False Positives)	Verdaderos negativos (TN – True Negatives)

$$Puntaje\ F1 = \frac{2 \times Precisión \times Sensibilidad}{Precisión + Sensibilidad}$$

Curva ROC (Receiver Operating Characteristic): Es la gráfica que muestra el desempeño de un clasificador a medida que se varía el umbral de discriminación. Es últil para visualizar la tasa de verdaderos positivos frente a la tasa de falsos positivos.

Área bajo la Curva ROC (Area Under the ROC Curve o AUC-ROC): Es un valor numérico que cuantifica el rendimiento general de un clasificador. Un valor de 1 representa un clasificador perfecto, miestras que un valor de 0.5 sugiere que el clasificador no es mejor que un proceso de asignación aleatoria.



UNIVERSIDAD DE GUANAJUATO