Métodos Numéricos - Clase 7

Ulises Bussi- Javier Portillo

1° cuatrimestre 2020

Ajustes de curvas: Validación

Ajustes de Curvas

Validación de Ajuste

¿Cómo elegimos modelo? Dado un conjunto de datos, ¿cuál los representa mejor?.

¿Por qué? Existen muchos modelos, vamos tratar de usar el mejor.

Universidad Nacional de Quilmes

Validación de Ajuste

¿Cómo elegimos modelo? Dado un conjunto de datos, ¿cuál los representa mejor?.

¿Por qué? Existen muchos modelos, vamos tratar de usar el mejor.

Ajustes de Curvas

Validación

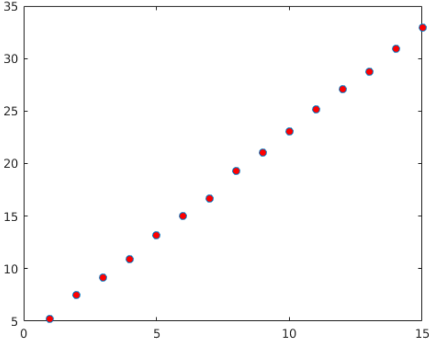
Dado un conjunto de pares ordenados $(x_i, y_i) \forall i = 1...n$ Vamos a crear dos subconjuntos disjuntos de datos $(x_i, y_i)_{\text{train}}$ y $(x_i, y_i)_{\text{validation}}$.

Realizaremos los ajustes sobre el conjunto de train. y calcularemos el r^2 sobre el otro conjunto.

Universidad Nacional de Quilmes

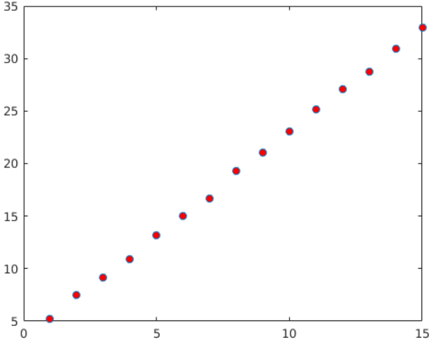
Validación: un Ejemplo

Supongamos que tenemos el conjunto de datos:



Validación: un Ejemplo

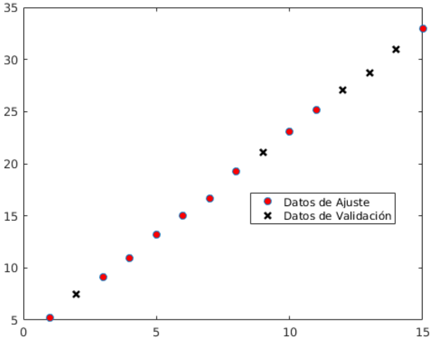
Supongamos que tenemos el conjunto de datos:



Primero debemos separar nuestro set de datos

Validación: un Ejemplo

Supongamos que tenemos el conjunto de datos:



Primero debemos separar nuestro set de datos

Validación: un Ejemplo

Realizamos el ajuste con polinomios hasta de grado 4 para ello, podemos generalizar el problema como: $A \underbrace{c}_{\text{coeficientes}} = b$

donde podemos escribir

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 & \sum x_i^7 \\ \sum x_i^4 & \sum x_i^5 & \sum x_i^6 & \sum x_i^7 & \sum x_i^8 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \sum y_i x_i^3 \\ \sum y_i x_i^4 \end{bmatrix}$$

Validación: un Ejemplo

Realizamos el ajuste con polinomios hasta de grado 4 para ello, podemos generalizar el problema como: $A \underbrace{c}_{\text{coeficientes}} = b$

donde podemos escribir

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \sum y_i x_i^3 \end{bmatrix}$$

Universidad Nacional de Quilmes

Validación: un Ejemplo

Realizamos el ajuste con polinomios hasta de grado 4 para ello, podemos generalizar el problema como: $A \underbrace{c}_{\text{coeficientes}} = b$

donde podemos escribir

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \\ & & & \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \\ \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \\ \end{bmatrix}$$

Universidad Nacional de Quilmes

Validación: un Ejemplo

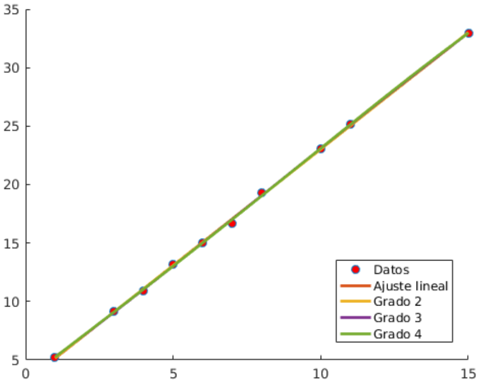
Realizamos el ajuste con polinomios hasta de grado 4 para ello, podemos generalizar el problema como: $A \underbrace{c}_{\text{coeficientes}} = b$

donde podemos escribir

$$\begin{bmatrix} n & \sum x_i & & \\ \sum x_i & \sum x_i^2 & & \\ & & & \\ & & & \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \\ \\ \end{bmatrix}$$

Validación: un Ejemplo

Una vez hallados los coeficientes para cada caso, es posible dibujar los distintos ajustes:

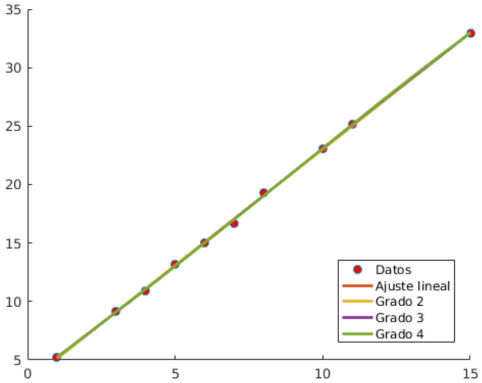


Si bien parecen todos similares miremos r^2

Universidad Nacional de Quilmes

Validación: un Ejemplo

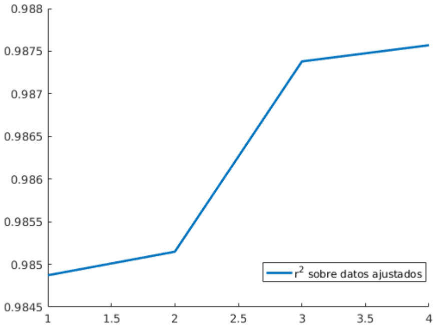
Una vez hallados los coeficientes para cada caso, es posible dibujar los distintos ajustes:



Si bien parecen todos similares miremos r^2

Validación: un Ejemplo

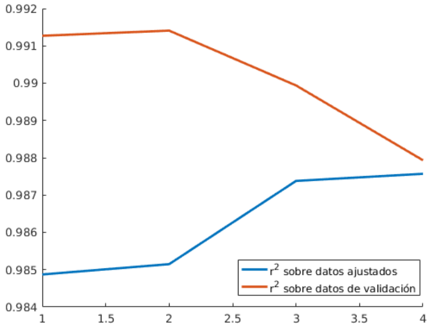
Los coeficientes de determinación



Conclusión: El mejor ajuste parece ser el cuadrático.

Validación: un Ejemplo

Los coeficientes de determinación



Conclusión: El mejor ajuste parece ser el cuadrático.