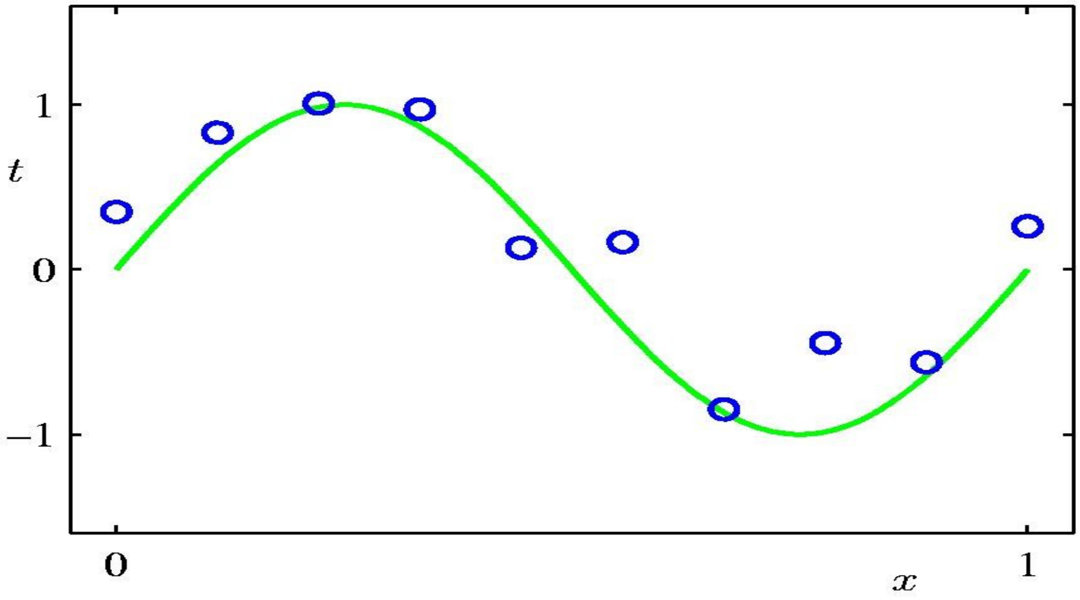
Inteligencia Artificial

Aproximación de datos polinomial por mínimos cuadrados

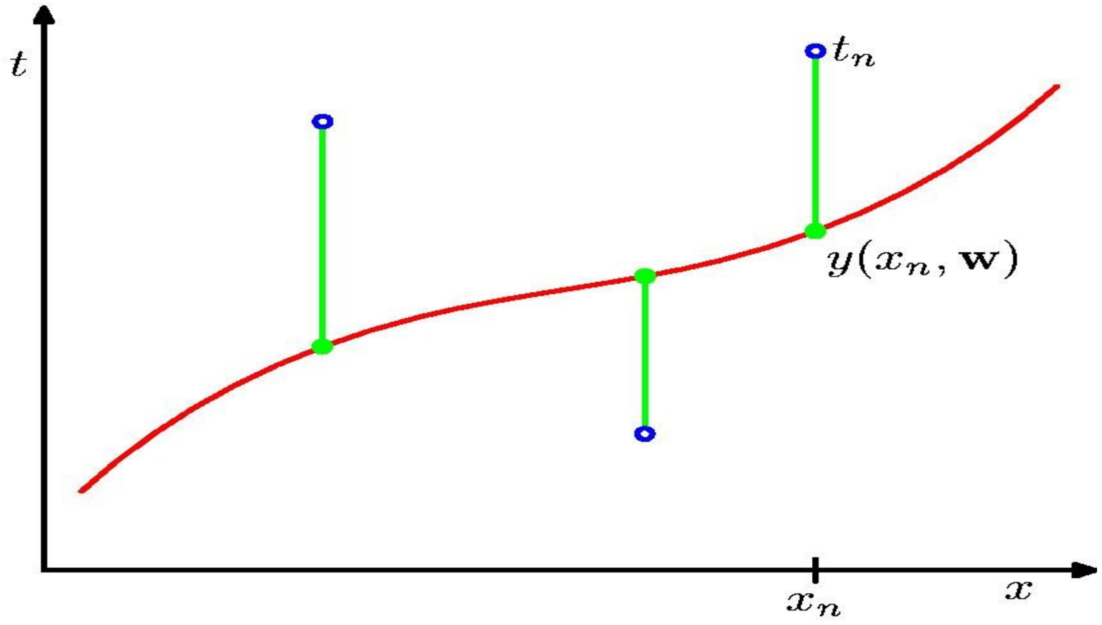
M.C. Miguelangel Fraga Aguilar
mfraga@itmorelia.edu.mx

Ajuste de datos polinomial



$$y(x,\boldsymbol{w}) = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M = \sum_{j=0}^{M} w_j x^j$$

Mínimos cuadrados



$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=0}^{N}\{y(x_n, \boldsymbol{w}) - t_n\}^2$$

Planteamiento punto por punto

$$w_0(1)+w_1x_0+w_2{x_0}^2+\ldots+w_Mx_0{}^M\approx t_0$$
$$w_0(1)+w_1x_1+w_2{x_1}^2+\ldots+w_Mx_1{}^M\approx t_1$$
$$\vdots$$
$$w_0(1)+w_1x_N+w_2{x_N}^2+\ldots+w_Mx_N{}^M\approx t_N$$

Planteamiento en forma matricial

Matriz de diseño
Matriz de funciones base
(de Vandermonde)

$$A = \begin{pmatrix} 1 & x_0 & x_0^{2} & \cdots & x_0^{M} \\ 1 & x_1 & x_1^{2} & \cdots & x_1^{M} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N & x_N^{2} & \cdots & x_N^{M} \end{pmatrix}$$

Vector de parámetros

$$\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_M \end{pmatrix}$$

Vector de objetivos

$$\boldsymbol{b} = \begin{pmatrix} t_0 \\ t_1 \\ \vdots \\ t_N \end{pmatrix}$$

$$\boxed{A\,\boldsymbol{w} \simeq \boldsymbol{b}}$$

Ecuaciones normales y pseudoinversa

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=0}^{N}\{y(x_n,\boldsymbol{w}) - t_n\}^2 = \frac{1}{2}\|A\boldsymbol{w} - \boldsymbol{b}\|_2^2$$

$$E(\boldsymbol{w}) = \frac{1}{2}(A\boldsymbol{w} - \boldsymbol{b})^T(A\boldsymbol{w} - \boldsymbol{b}) = \frac{1}{2}\left[\boldsymbol{w}^T A^T(A\boldsymbol{w} - \boldsymbol{b}) - \boldsymbol{b}^T(A\boldsymbol{w} - \boldsymbol{b})\right]$$

$$E(\boldsymbol{w}) = \frac{1}{2}(\boldsymbol{w}^T A^T A\boldsymbol{w} - 2\boldsymbol{w}^T A^T\boldsymbol{b} + \boldsymbol{b}^T\boldsymbol{b})$$

$$\boldsymbol{0} = \nabla E(\boldsymbol{w}) = A^T A\boldsymbol{w} - A^T\boldsymbol{b}$$

Ecuaciones Normales

$$A^T A\boldsymbol{w} = A^T\boldsymbol{b}$$

$$\boldsymbol{w} = (A^T A)^{-1} A^T\boldsymbol{b}$$

Pseudoinversa de Moore-Penrose

$$A^+ = (A^T A)^{-1} A^T$$

Métodos Numéricos

- Las ecuaciones normales son mal condicionadas numéricamente
- Factorizacion Q-R tiene mejor condicionamiento
- La descomposición en valores singulares tiene aún mejor condicionamiento, pero tiene un mayor costo computacional
- La forma normal de calcular la pseudoinversa es por descomposición en valores singulares

Mínimos cuadrados regularizados Regularización de Tikhonov

$$E(\boldsymbol{w}) = \frac{1}{2}\left\{\|A\boldsymbol{w} - \boldsymbol{b}\|_2^2 + \lambda^2\|\boldsymbol{w}\|_2^2\right\}$$

Equivalente a

$$E(\boldsymbol{w}) = \frac{1}{2}\left\|\begin{pmatrix} A \\ \lambda I \end{pmatrix}\boldsymbol{w} - \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{pmatrix}\right\|_2^2$$

Ecuaciones normales

$$\begin{pmatrix} A \\ \lambda I \end{pmatrix}^T \begin{pmatrix} A \\ \lambda I \end{pmatrix}\boldsymbol{w} = \begin{pmatrix} A \\ \lambda I \end{pmatrix}^T \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{pmatrix}$$

Pseudoinversa

$$\boldsymbol{w}_\lambda = \left(A^T A + \lambda^2 I\right)^{-1} A^T \boldsymbol{b}$$

- Se añade un termino a la función de costo que es un múltiplo de la magnitud del vector de pesos
- La solución numérica es por los mismos métodos que los mínimos cuadrados normales, pero con el sistema extendido

$$\begin{pmatrix} A \\ \lambda I \end{pmatrix}\boldsymbol{w} \simeq \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{pmatrix}$$