

# Curso de Métodos Numéricos

DEMAT, Universidad de Guanajuato

## Clase 11: Método de mínimos cuadrados pesados y ajuste de modelos

- Mínimos cuadrados pesados.
- Relación con regresión lineal.
- Ajuste de modelos polinomiales.

---

**MAT-251**

Dr. Joaquín Peña Acevedo  
CIMAT A.C.

**e-mail:** [joaquin@cimat.mx](mailto:joaquin@cimat.mx)

# Mínimos cuadrados pesados (I)

Se puede escribir el error

$$E(a, b) = \sum_{i=1}^m w_i (ax_i + b - y_i)^2$$

introduciendo las matrices y vectores

$$\mathbf{W} = \text{diag}(w_1, \dots, w_m), \quad \mathbf{A} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

$$\begin{aligned} E(\mathbf{c}) &= (\mathbf{Ac} - \mathbf{y})^\top \mathbf{W} (\mathbf{Ac} - \mathbf{y}) = (\mathbf{c}^\top \mathbf{A}^\top - \mathbf{y}^\top) \mathbf{W} (\mathbf{Ac} - \mathbf{y}) \\ &= \mathbf{c}^\top \mathbf{A}^\top \mathbf{W} \mathbf{Ac} - 2\mathbf{y}^\top \mathbf{W} \mathbf{Ac} + \mathbf{y}^\top \mathbf{W} \mathbf{y}. \end{aligned}$$

Por tanto, la solución de mínimos cuadrados está dada al resolver el sistema

$$\mathbf{A}^\top \mathbf{W} \mathbf{Ac} = \mathbf{A}^\top \mathbf{W} \mathbf{y}.$$

El problema es como fijar los pesos  $w_1, \dots, w_m$  asociados a cada ecuación. El método de mínimos cuadrados pesados hace la siguiente propuesta:

## Mínimos cuadrados pesados (II)

- 1 Inicializar  $k = 0$ , los pesos  $w_i = 1$  y formar la matriz  $\mathbf{W}^{(0)}$ .
- 2 Calcular la solución del problema de mínimos cuadrados

$$\mathbf{A}^\top \mathbf{W}^{(k)} \mathbf{A} \mathbf{c}^{(k)} = \mathbf{A}^\top \mathbf{W}^{(k)} \mathbf{y}.$$

- 3 Calcular el vector de error  $\boldsymbol{\epsilon} = \mathbf{A} \mathbf{c}^{(k)} - \mathbf{y}$ .
- 4 Si  $\|\mathbf{c}^{(k)} - \mathbf{c}^{(k-1)}\|$  o el RMSE  $\sqrt{\frac{\boldsymbol{\epsilon}^\top \mathbf{W} \boldsymbol{\epsilon}}{m}}$  es menor que la tolerancia dada, terminar el algoritmo devolviendo la solución  $\mathbf{c}^{(k)}$ .
- 5 En caso contrario, modificar el valor del peso  $w_i$  de acuerdo a la magnitud de la componente  $i$ -ésima del vector de error,  $\epsilon_i = (\mathbf{A} \mathbf{c})_i - y_i$ , de modo que si  $\epsilon_i$  es grande, el peso  $w_i$  debe ser pequeño, y viceversa. Construir la matriz  $\mathbf{W}^{(k+1)}$ . Hacer  $k = k + 1$  y volver al paso 2.

## Ejemplo de mínimos cuadrados pesados (I)

Consideramos el problema de ajustar una recta a un conjunto de puntos  $\{(x_1, y_1), \dots, (x_m, y_m)\}$  en el plano aplicando el método de mínimos cuadrados pesados.

Una vez que se tiene una solución  $\mathbf{c}^{(k)}$  en la iteración  $k$ , actualizamos el valor de los pesos  $w_i^{(k+1)}$  se actualiza para la siguiente iteración mediante:

$$w_i^{(k+1)} = \exp(-5\epsilon_i^2), \quad \text{con } \epsilon_i = (\mathbf{A}\mathbf{c}^{(k)})_i - y_i.$$

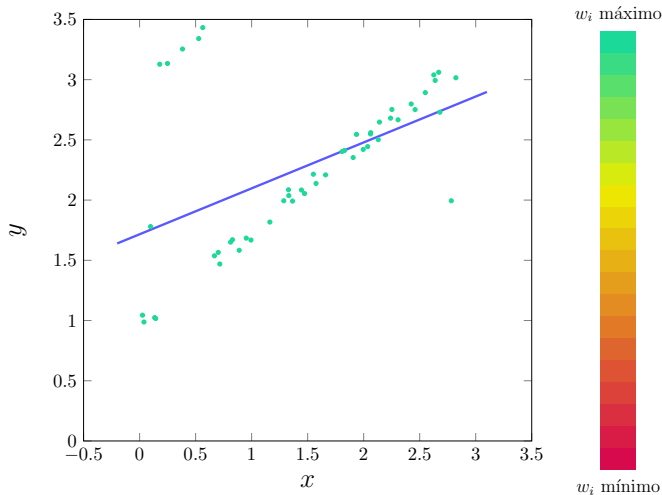
La tolerancia usada es  $\tau = \sqrt{\epsilon_m}$ .

En las siguientes gráficas se muestra el resultado de aplicar el método de mínimos cuadrados pesados.

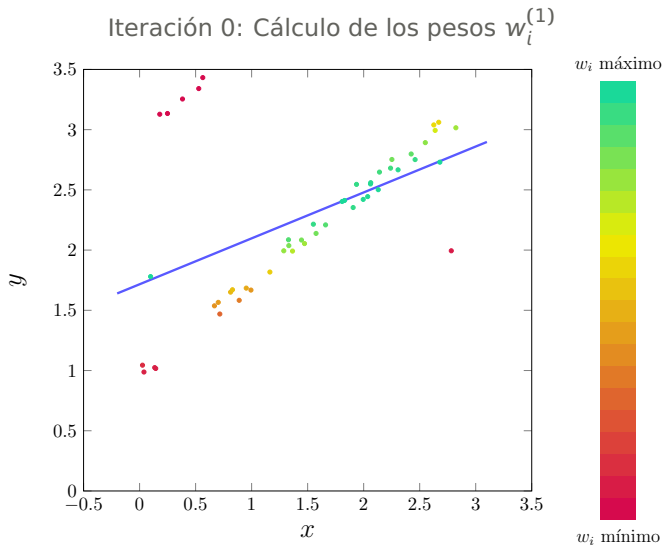
Los puntos son coloreados de acuerdo al valor del peso  $w_i$  que tienen asociado.

## Ejemplo de mínimos cuadrados pesados (II)

Iteración 0: Cálculo de la solución  $\mathbf{A}^\top \mathbf{W}^{(0)} \mathbf{A} \mathbf{c}^{(0)} = \mathbf{A}^\top \mathbf{W}^{(0)} \mathbf{y}$  con  $w_i^{(0)} = 1$

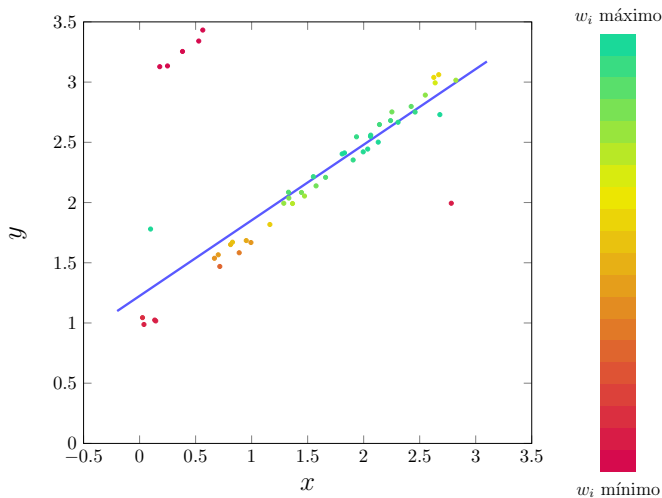


## Ejemplo de mínimos cuadrados pesados (III)



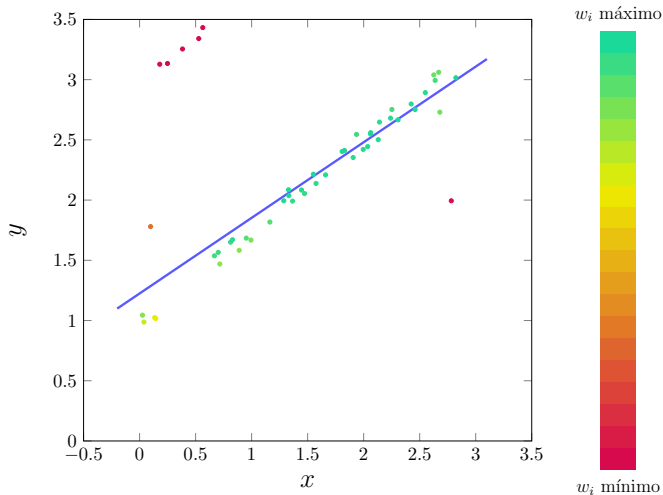
## Ejemplo de mínimos cuadrados pesados (IV)

Iteración 1: Cálculo de la solución  $\mathbf{A}^\top \mathbf{W}^{(1)} \mathbf{A} \mathbf{c}^{(1)} = \mathbf{A}^\top \mathbf{W}^{(1)} \mathbf{y}$



## Ejemplo de mínimos cuadrados pesados (V)

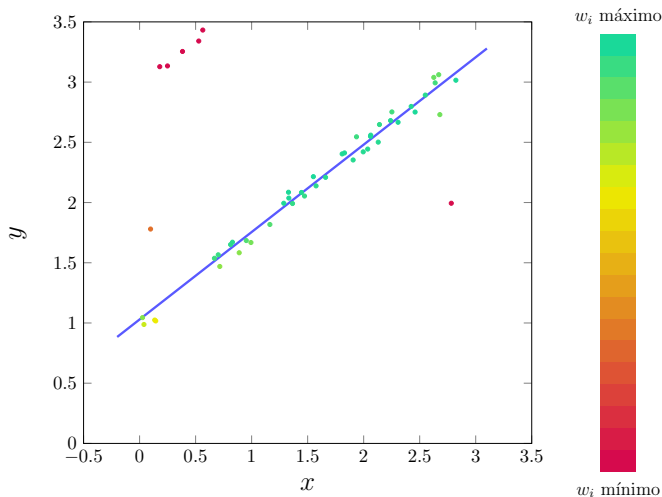
Iteración 1: Cálculo de los valores los pesos  $w_i^{(2)}$





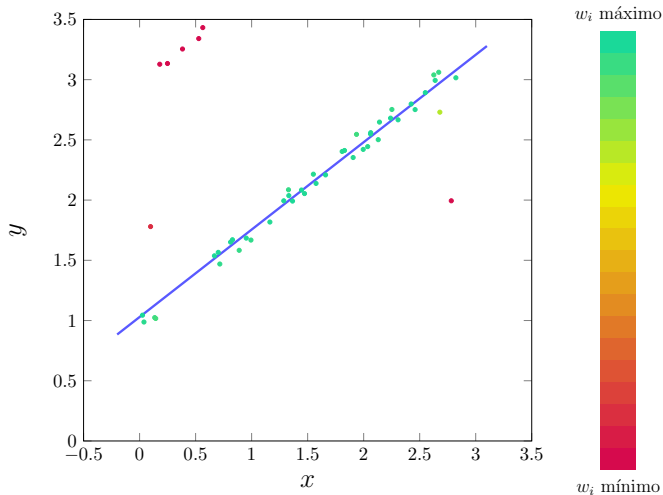
## Ejemplo de mínimos cuadrados pesados (VI)

Iteración 2: Cálculo de la solución  $\mathbf{A}^T \mathbf{W}^{(2)} \mathbf{A} \mathbf{c}^{(2)} = \mathbf{A}^T \mathbf{W}^{(2)} \mathbf{y}$



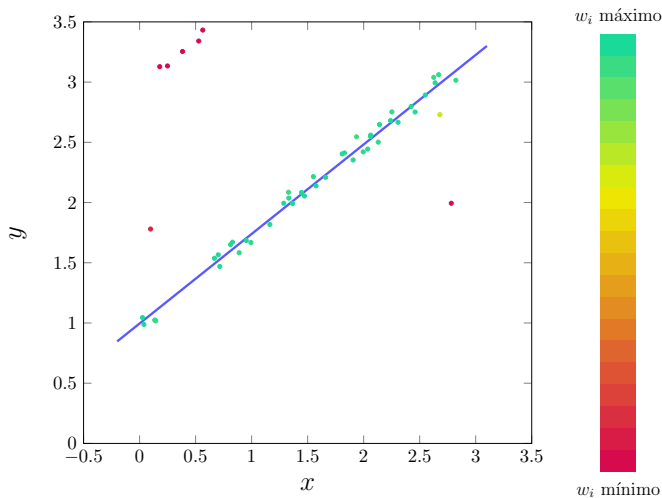
## Ejemplo de mínimos cuadrados pesados (VII)

Iteración 2: Actualización de los valores los pesos  $w_i^{(3)}$



## Ejemplo de mínimos cuadrados pesados (VIII)

Iteración 10: Cálculo de la solución  $\mathbf{A}^\top \mathbf{W}^{(10)} \mathbf{A} \mathbf{c}^{(10)} = \mathbf{A}^\top \mathbf{W}^{(10)} \mathbf{y}$



## Ejemplo de mínimos cuadrados pesados (IX)

- El método de mínimos cuadrados pesados no siempre funciona.
- El método depende de la solución en la primera iteración y de la manera en la que se asignan los pesos.

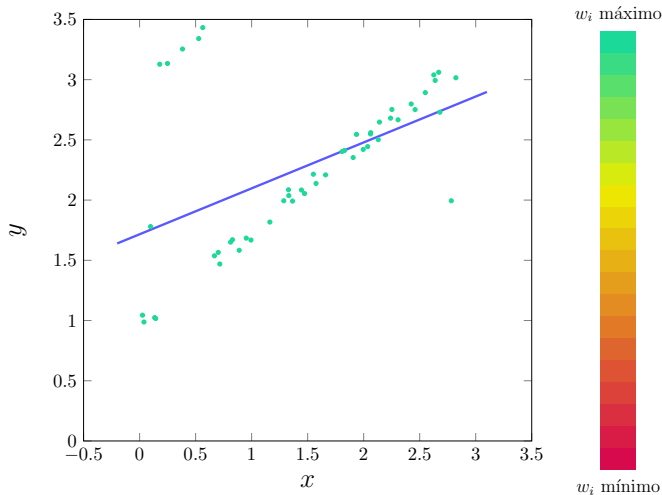
Por ejemplo, si se cambia la manera de calcular los pesos de la siguiente manera:

$$w_i^{(k+1)} = \exp(-150\epsilon_i^2), \quad \text{con } \epsilon_i = (\mathbf{A}\mathbf{c}^{(k)})_i - y_i,$$

se obtiene los siguientes resultados.

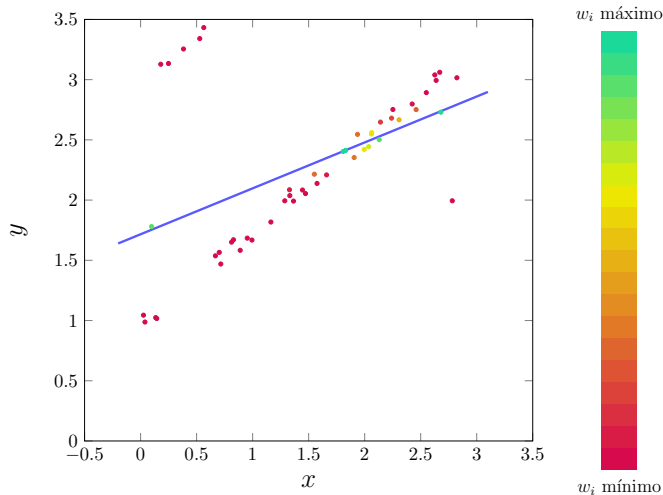
# Ejemplo de mínimos cuadrados pesados (X)

Iteración 0: Cálculo de la solución  $\mathbf{A}^T \mathbf{W}^{(0)} \mathbf{A} \mathbf{c}^{(0)} = \mathbf{A}^T \mathbf{W}^{(0)} \mathbf{y}$  con  $w_i^{(0)} = 1$



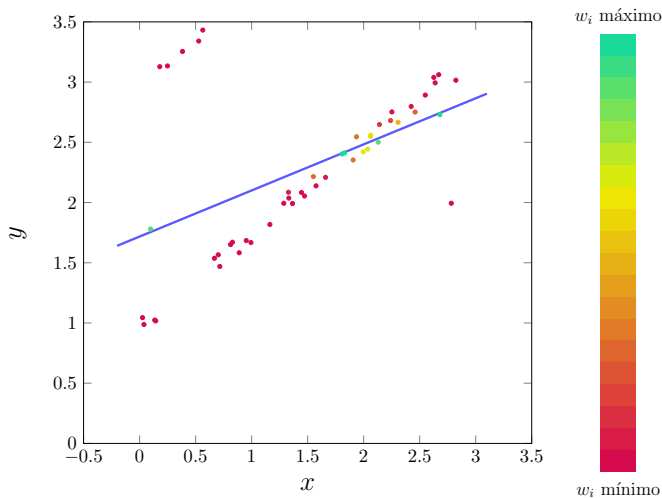
# Ejemplo de mínimos cuadrados pesados (XI)

Iteración 0: Actualización de los valores de pesos  $w_i^{(1)}$



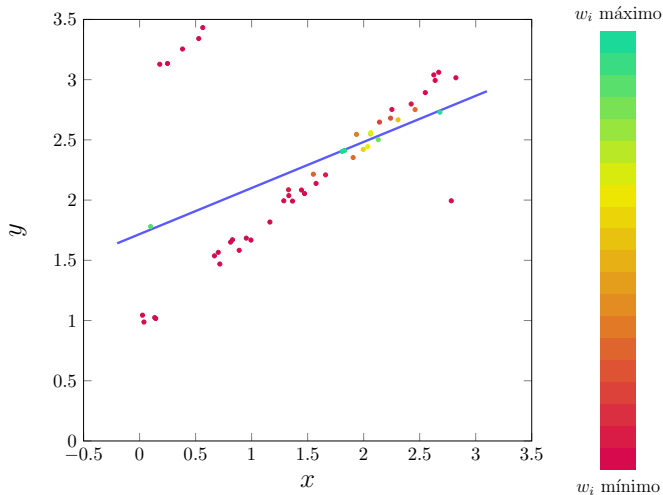
## Ejemplo de mínimos cuadrados pesados (XII)

Iteración 1: Cálculo de la solución  $\mathbf{A}^\top \mathbf{W}^{(1)} \mathbf{A} \mathbf{c}^{(1)} = \mathbf{A}^\top \mathbf{W}^{(1)} \mathbf{y}$



## Ejemplo de mínimos cuadrados pesados (XIII)

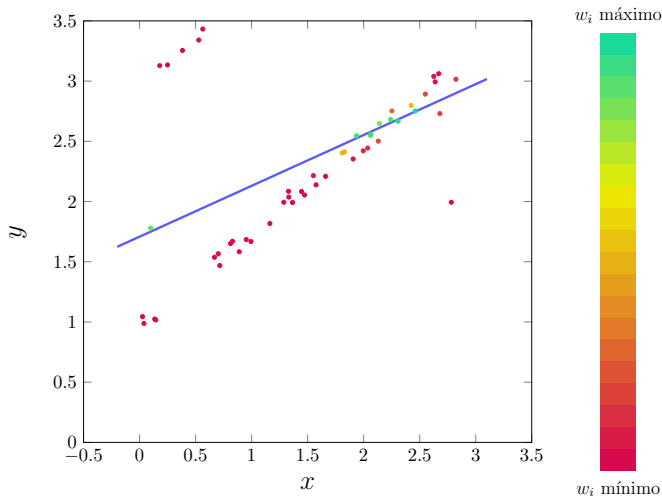
Iteración 1: Actualización de los valores los pesos  $w_i^{(2)}$





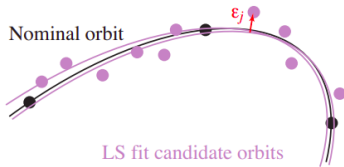
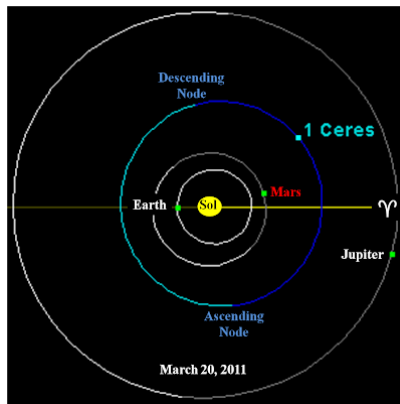
## Ejemplo de mínimos cuadrados pesados (XIV)

Iteración 118: Cálculo de la solución  $\mathbf{A}^T \mathbf{W}^{(118)} \mathbf{A} \mathbf{c}^{(118)} = \mathbf{A}^T \mathbf{W}^{(118)} \mathbf{y}$



# Referencias históricas (I)

- En enero de 1801 un astrónomo descubre el asteroide Ceres y hace 22 observaciones de su posición en 44 días, hasta que el sol impide verlo. Se empiezan a hacer predicciones de cuando y donde reaparecerá.



Estimated orbit errors

44 equations in 6 unknown parameters.

## Referencias históricas (II)

- En septiembre de 1801 Gauss publica sus estimaciones y en diciembre se confirma que las predicciones de Gauss fueron correctas (dentro de cierta tolerancia), pero Gauss no revela su método.
- En 1805 Legendre publica un libro sobre estimación de órbitas de cometas, y en el apéndice se describe el método de mínimos cuadrados.
- En 1809 Gauss publica su método para la estimación de órbitas de cuerpos que orbitan alrededor del sol y declara que él desarrolló el método de mínimos cuadrados desde 1795.
- En 1812 Laplace fue el primero deducir el método de mínimos cuadrados en términos de reducir la esperanza del error.

# Mínimos cuadrados y regresión lineal (I)

El método de *regresión* es una técnica para estudiar la relación entre un conjunto de variables dependientes o de respuesta  $\mathbf{y}$  en función de otro conjunto de independientes o covariables  $\mathbf{x}$ . Fue propuesto alrededor de 1885 por Francis Galton, aplicando mínimos cuadrados, aun cuando éste fue propuesto 80 años antes.

Galton estudió la relación de las estaturas de los hijos con respecto a la estatura promedio del padre y la madre. Sus datos formaban una colección de 928 muestras de las estaturas promedio  $x_i$  de los padres de varias familias y sus hijos  $y_i$ , tomando las mediciones en pulgadas.

En el conjunto de datos  $\{(x_1, y_1), \dots, (x_{928}, y_{928})\}$  las alturas están discretizadas, por lo que varios datos pueden corresponder al mismo punto en el plano. En la siguiente gráfica los puntos se colorean de acuerdo a la frecuencia.

## Mínimos cuadrados y regresión lineal (II)

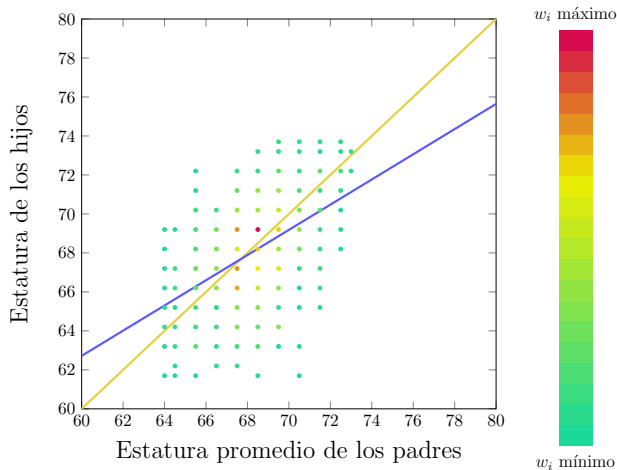


Figure: Relación entre la estatura promedio de los padres y la estura de sus hijos.

La línea amarilla en la Figura 1 representa la recta a 45 grados. Si los hijos tuvieran una estatura similar a los padres, esa es la recta que se hubiera obtenido. En cambio, se obtiene la recta azul que lo indica es que los padres de baja estatura suelen tener hijos más altos que ellos, mientras que los padres más altos suelen tener hijos más bajos. A esto hace las estatura de los hijos no se alejen demasiado de la media de la estura de los padres.

Galton se refirió a esto como una *regresión hacia la media* de los datos, de ahí se adoptó ese término para esta forma de analizar datos.

### Proposición 1

Sea  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , con  $m > n$ . La matriz  $\mathbf{A}^T \mathbf{A}$  es definida positiva si y sólo si las columnas de  $\mathbf{A}$  son linealmente independientes, es decir, si  $\text{rank}(\mathbf{A}) = n$ .

Para modelos estadísticos lineales en la que se tiene una matriz  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , con  $m > n$ , que relaciona a las variables desconocidas  $\mathbf{x} \in \mathbb{R}^n$  con un vector de observaciones  $\mathbf{b} \in \mathbb{R}^m$ , mediante la relación

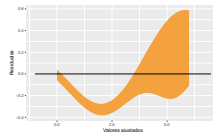
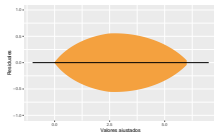
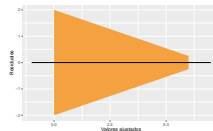
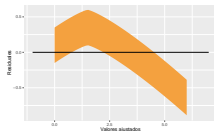
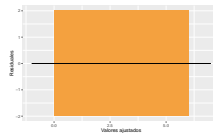
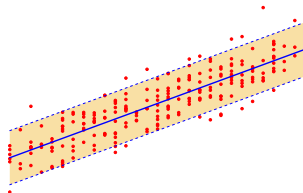
$$\mathbf{Ax} = \mathbf{b} + \boldsymbol{\epsilon}, \quad (1)$$

donde  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)^T$  es un vector de errores aleatorios independientes, con media cero y varianza  $\sigma^2$ :

$$E(\boldsymbol{\epsilon}) = \mathbf{0}, \quad \text{var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}. \quad (2)$$

# Mínimos cuadrados y regresión lineal (V)

Hay que analizar los residuales para saber si el modelo hace un buen ajuste de los datos.





## Mínimos cuadrados y regresión lineal (VI)

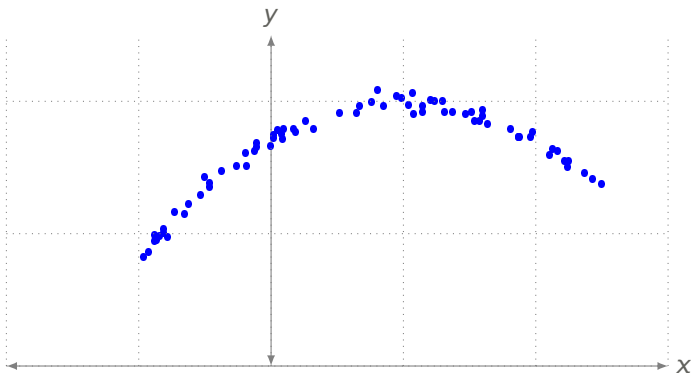


*"Essentially, all  
models are wrong,  
but some are  
useful"*

George Edward Pelham Box FRS  
(1919 – 2013)

# Ajuste de modelos polinomiales (I)

Tenemos un conjunto de datos  $\{(x_k, y_k)\}_{k=1}^m$  que no se pueden explicar por un modelo lineal, por ejemplo:



## Ajuste de modelos polinomiales (II)

Proponemos un modelo cuadrático  $ax^2 + bx + c$ , de modo que hay que determinar los parámetros  $a, b, c$  que minimicen los errores

$$\epsilon_i = ax_i^2 + bx_i + c - y_i \quad i = 1, 2, \dots, m.$$

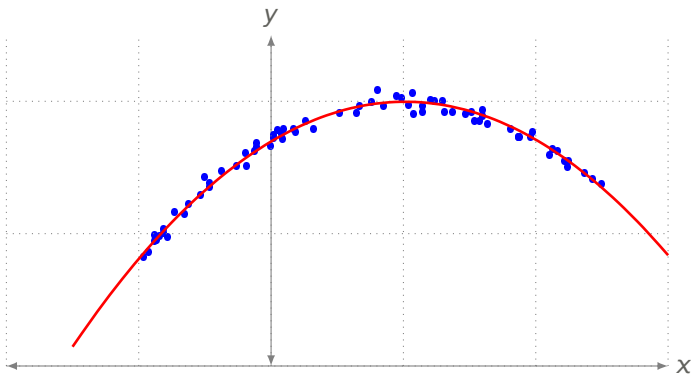
$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{pmatrix} = \begin{pmatrix} ax_1^2 + bx_1 + c \\ ax_2^2 + bx_2 + c \\ \vdots \\ ax_m^2 + bx_m + c \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_m^2 & x_m & 1 \end{bmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

$$\boldsymbol{\epsilon} = \mathbf{A}\mathbf{x} - \mathbf{y}$$

$$\min_{\mathbf{x}} \sum_{i=1}^m e_i^2 = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} = (\mathbf{x}^T \mathbf{A}^T - \mathbf{y}^T)(\mathbf{A}\mathbf{x} - \mathbf{y}) = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{y}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{y}$$

Si  $\mathbf{A}^T \mathbf{A}$  es definida positiva, el mínimo se obtiene al resolver el sistema  $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}$ :  $a = -0.29521, b = 0.59151, c = 1.69452$ .

## Ajuste de modelos polinomiales (III)



## Ajuste de modelos polinomiales (IV)

En general, se puede ajustar un modelo polinomial de grado  $n$  a un conjunto de datos  $\{(x_k, y_k)\}_{k=1}^m$ , de modo que minimicen los errores

$$\epsilon_i = \sum_{j=0}^n a_j x_i^j - y_i \quad i = 1, 2, \dots, m.$$

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^n a_j x_1^j \\ \sum_{j=0}^n a_j x_2^j \\ \vdots \\ \sum_{j=0}^n a_j x_m^j \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{bmatrix} x_1^n & \cdots & x_1 & 1 \\ x_2^n & \cdots & x_2 & 1 \\ \vdots & & \vdots & \vdots \\ x_m^n & \cdots & x_m & 1 \end{bmatrix} \begin{pmatrix} a_n \\ \vdots \\ a_1 \\ a_0 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Es decir,

$$\boldsymbol{\epsilon} = \mathbf{A}\mathbf{x} - \mathbf{y}$$

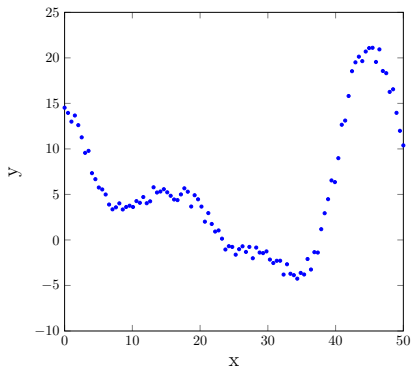
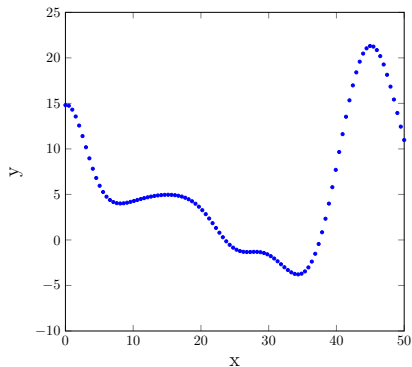
Así, se puede plantear el problema de optimización

$$\min_{\mathbf{x}} \sum_{i=1}^m e_i^2 = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} = (\mathbf{x}^T \mathbf{A}^T - \mathbf{y}^T)(\mathbf{A}\mathbf{x} - \mathbf{y}) = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{y}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{y}$$

Si  $\mathbf{A}^T \mathbf{A}$  es definida positiva, el mínimo se obtiene al resolver el sistema  $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}$ .

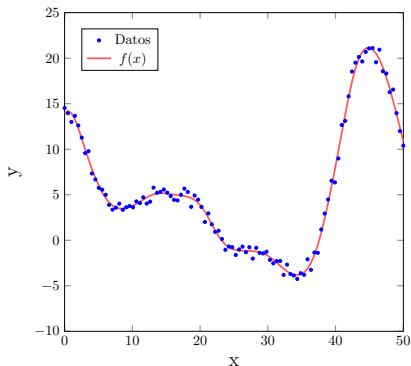
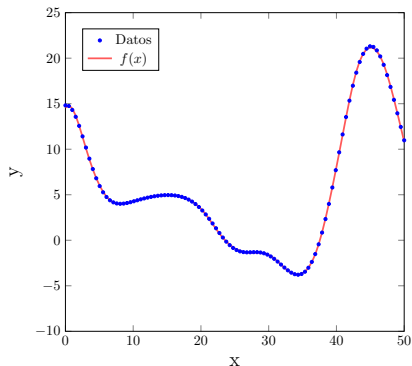
# Ajuste de datos 1D (I)

- Tenemos un conjunto de datos  $\{(x_k, y_k)\}_{k=1}^m$ .
- Queremos encontrar una función  $f(x)$  cuya gráfica pase cerca de los puntos dados.



# Ajuste de datos 1D (II)

- Tenemos un conjunto de datos  $\{(x_k, y_k)\}_{k=1}^m$ .
- Queremos encontrar una función  $f(x)$  cuya gráfica pase cerca de los puntos dados.





Primero hay que decir a que nos referimos con "cerca". Lo podemos definir en el sentido de mínimos cuadrados:

$$\min_f \frac{1}{2} \sum_{k=1}^m [f(x_k) - y_k]^2 \quad (3)$$

Observaciones:

- Si  $x_k \neq x_j$  para  $k \neq j$ , hay una infinidad de funciones cuya gráfica pasa por los puntos.
- Si existen  $k \neq j$ , tales que  $x_k = x_j$  y  $y_k \neq y_j$ , entonces no existe una función cuya gráfica pase por los puntos dados.

Para simplificar el problema, podemos restringir el espacio de solución a cierto tipo de funciones:

$$f(x) = \sum_{j=1}^n \alpha_j \phi_j(x). \quad (4)$$

Sustituyendo en (3), tenemos

$$E = \frac{1}{2} \sum_{k=1}^m [f(x_k) - y_k]^2 = \frac{1}{2} \sum_{k=1}^m \left[ \sum_{j=1}^n \alpha_j \phi_j(x_k) - y_k \right]^2.$$

Derivando respecto a  $\alpha_i$ , se tiene

$$\frac{\partial E}{\partial \alpha_i} = \sum_{k=1}^m \left[ \sum_{j=1}^n \alpha_j \phi_j(x_k) - y_k \right] \phi_i(x_k).$$

Para cada  $i$ , queremos los valores  $\alpha_j$  que hacen  $\frac{\partial E}{\partial \alpha_i} = 0$ . Esto es

$$\sum_{j=1}^n \left[ \sum_{k=1}^m \phi_j(x_k) \phi_i(x_k) \right] \alpha_j = \sum_{k=1}^m \phi_i(x_k) y_k.$$

Para obtener directamente la expresión de la solución hay que notar que para cada  $i$ , se tiene que el error está dado por

$$\epsilon_i = f(x_i) - y_i = \sum_{j=1}^n \alpha_j \phi_j(x_i) - y_i$$

Si

$$\Phi = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \cdots & \phi_n(x_m) \end{bmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

el vector de errores se expresa como

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} = \boldsymbol{\Phi}\boldsymbol{\alpha} - \mathbf{y}$$

Queremos minimizar

$$E = \|\boldsymbol{\epsilon}\|_2^2 = \boldsymbol{\epsilon}^\top \boldsymbol{\epsilon} = (\boldsymbol{\Phi}\boldsymbol{\alpha} - \mathbf{y})^\top (\boldsymbol{\Phi}\boldsymbol{\alpha} - \mathbf{y})$$

Calculamos el gradiente de  $E$  y lo igualamos a cero. Con esto obtenemos que el sistema lineal que hay que resolver para obtener la solución de mínimos cuadrados:

$$\boldsymbol{\Phi}^\top \boldsymbol{\Phi}\boldsymbol{\alpha} = \boldsymbol{\Phi}^\top \mathbf{y}.$$

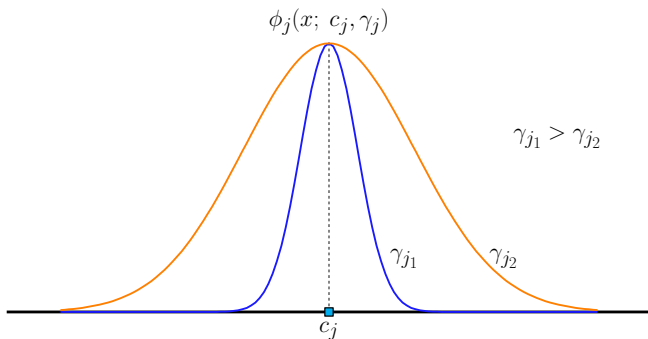
Sólo falta es seleccionar las funciones  $\phi_j(x)$ .

## Ejemplo 1. (I)

Por ejemplo, seleccionamos las funciones:

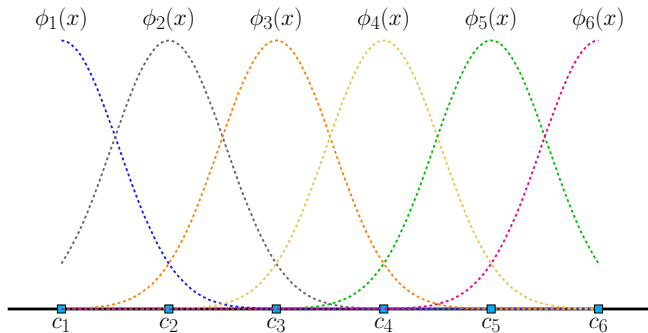
$$\phi_j(x) = \exp(-\gamma(x - c_j)^2)$$

El parámetro  $\gamma_j$  determina el ancho de la función, es decir, que tan rápido decrece.



## Ejemplo 1. (II)

El punto  $c_j$  determina el centro de la función y esto permite distribuir las funciones en un intervalo.

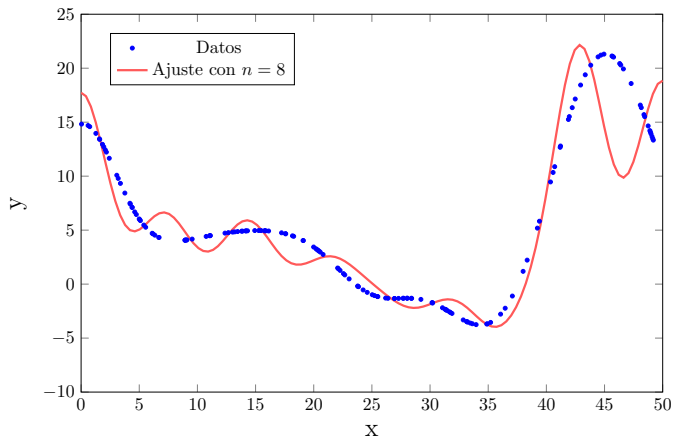


## Ejemplo 1. (III)

### Ejemplos:

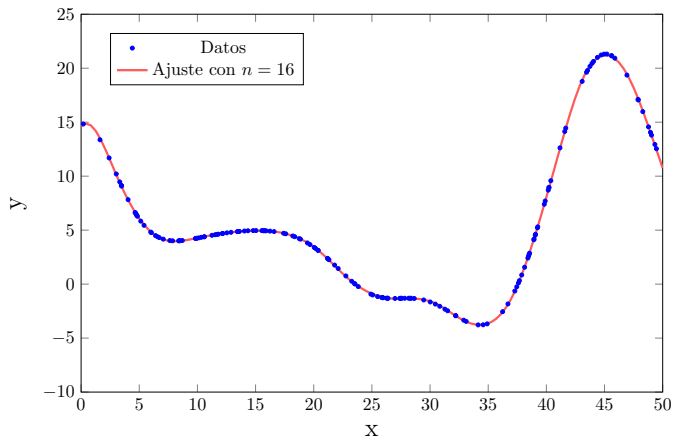
Para  $m = 120$  puntos en los que  $x_i \in [0, 50]$  y usando  $n = 8$  y  $n = 16$  funciones Gaussianas  $\phi_j(x)$  distribuidas con uniformemente en ese intervalo, se obtiene el siguiente resultado cuando no hay ruido en los datos:

## Ejemplo 1. (IV)



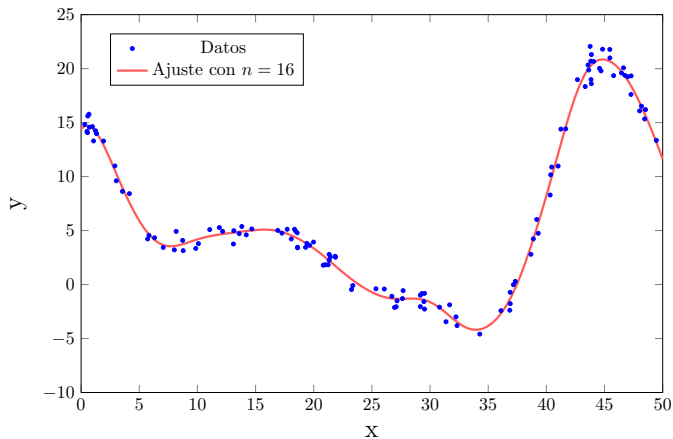


## Ejemplo 1. (V)



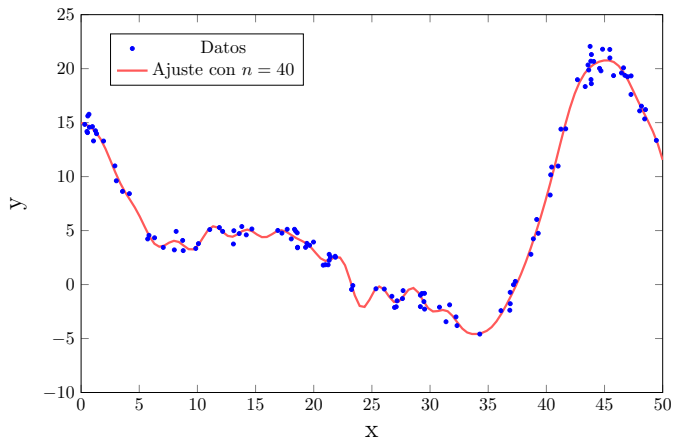
## Ejemplo 1. (VI)

Para datos con ruido se obtiene el siguiente resultado.



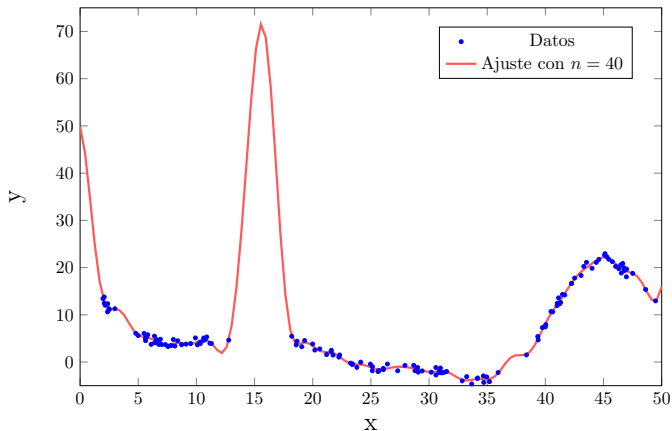
## Ejemplo 1. (VII)

Con los mismos datos, pero aumentando el número de funciones base:



## Ejemplo 1. (VIII)

Con otro conjunto de datos pueden ocurrir resultados no deseados:



## Ejemplo 1. (IX)

- Es necesario que haya datos alrededor de cada función base para que pueda contribuir en la solución.
- Para evitar que la solución tenga un mal comportamiento en los “espacios vacíos” hay que aplicar técnicas de regularización, para sesgar la solución aunque se incremente el error en el ajuste.

## Ejemplo 2 (I)

Considere un espacio cuadrado en el que hay 9 lámparas localizadas en las posiciones  $\mathbf{c}_1, \dots, \mathbf{c}_9$ .

- Las lámparas pueden estar a diferentes alturas del piso.
- La iluminación  $I_i(\mathbf{x})$  en un punto  $\mathbf{x}$  del piso debida a la  $i$ -ésima lámpara es directamente proporcional a la potencia  $p_i$  de la lámpara e inversamente proporcional a la distancia entre los puntos  $\mathbf{c}_i$  y  $\mathbf{x}$ :

$$I_i(\mathbf{x}) \propto \frac{p_i}{\|\mathbf{x} - \mathbf{c}_i\|}.$$

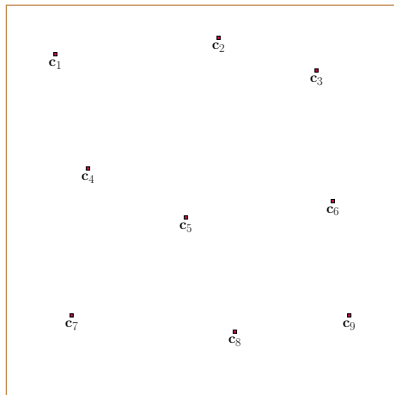
- Suponemos que hay una relación lineal en la contribución a la iluminación  $I(\mathbf{x})$  en un punto  $\mathbf{x}$  de cada lámpara, de modo que

$$I(\mathbf{x}) = I_1(\mathbf{x}) + I_2(\mathbf{x}) + \dots + I_9(\mathbf{x}).$$

## Ejemplo 2 (II)

¿De qué potencia deben ser las lámparas para que el piso quede iluminado de manera uniforme?

Distribución de las lámparas



$c_1$	=	(2.0,	21.0,	2.9)
$c_2$	=	(13.0,	22.0,	2.4)
$c_3$	=	(19.0,	20.0,	2.3)
$c_4$	=	(5.0,	14.0,	2.1)
$c_5$	=	(11.0,	11.0,	2.7)
$c_6$	=	(20.0,	12.0,	2.1)
$c_7$	=	(4.0,	5.0,	2.2)
$c_8$	=	(14.0,	4.0,	2.8)
$c_9$	=	(21.0,	5.0,	2.7)

## Ejemplo 2 (III)

Suponemos que la constante de proporcionalidad es 1. Entonces

$$I_i(\mathbf{x}; p_i) = \frac{p_i}{\|\mathbf{x} - \mathbf{c}_i\|}$$

Necesitamos escribir las relaciones entre las potencias  $p_1, \dots, p_9$  y la iluminación en el piso.

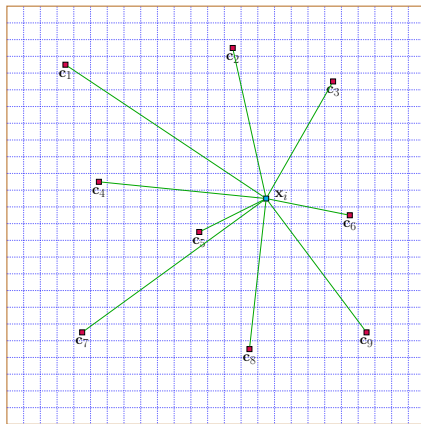
Para esto, elegimos una colección de puntos  $\mathbf{x}_i$  sobre el piso. Por conveniencia, elegimos estos puntos sobre una retícula uniforme, con 25 celdas en cada dirección y de modo que  $\mathbf{x}_i$  es el centro de la retícula.

Lo que se quiere es relacionar los 625 valores de la iluminación  $I(\mathbf{x}_1), I(\mathbf{x}_2), \dots, I(\mathbf{x}_{625})$  con los valores de las potencias  $p_1, \dots, p_9$ .



## Ejemplo 2 (IV)

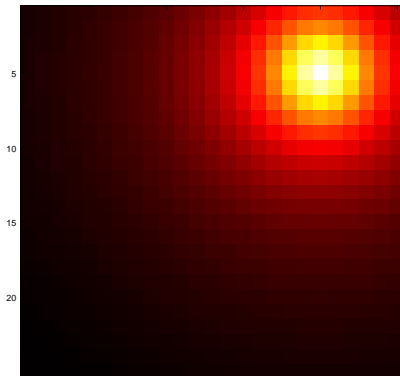
El centro de cada celda lo identificamos como el punto  $\mathbf{x}_i$ . Conviene usar una retícula para tener una muestra de puntos distribuidas en toda la región.



## Ejemplo 2 (V)

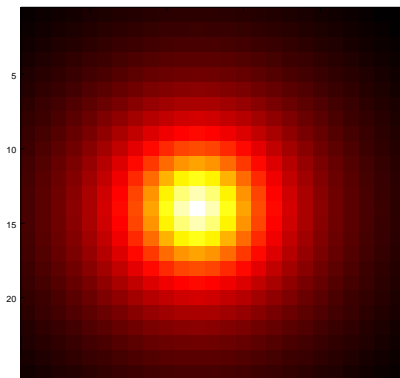
Notemos lo siguiente:

Por ejemplo, la iluminación en el piso debida a la lámpara 3 cuando la potencia es unitaria,  $p_3 = 1$ , de modo que los valores de  $I_3(\mathbf{x}; 1)$  en el cuadrado se ve como



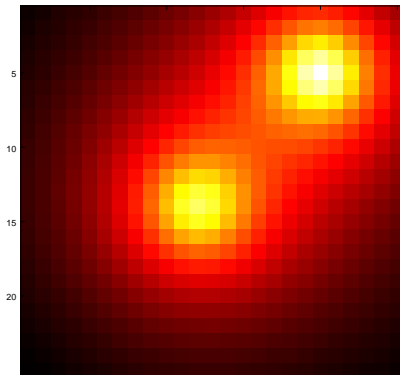
## Ejemplo 2 (VI)

La iluminación en el piso debida a la lámpara 5 cuando la potencia es unitaria,  $p_5 = 1$ , de modo que los valores de  $I_5(\mathbf{x}; 1)$  en el cuadrado se ve como



## Ejemplo 2 (VII)

Por linealidad que suponemos, la iluminación obtenida cuando las lámparas 3 y 5 tienen una potencia unitaria es:



## Ejemplo 2 (VIII)

De este modo podemos escribir que

$$I(\mathbf{x}_i) = p_1 I_1(\mathbf{x}_i; 1) + p_2 I_2(\mathbf{x}_i; 1) + \dots + p_9 I_9(\mathbf{x}_i; 1),$$

para  $i = 1, 2, \dots, 625$ .

Entonces para obtener una iluminación uniforme, por ejemplo, que sea igual a 1 en todos los puntos, entonces se debería cumplir que

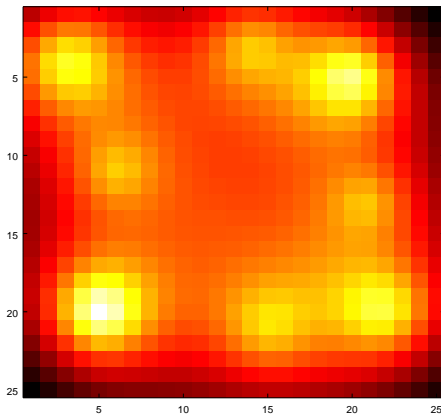
$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{bmatrix} \frac{1}{\|\mathbf{x}_1 - \mathbf{c}_1\|} & \frac{1}{\|\mathbf{x}_1 - \mathbf{c}_2\|} & \cdots & \frac{1}{\|\mathbf{x}_1 - \mathbf{c}_9\|} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{\|\mathbf{x}_{625} - \mathbf{c}_1\|} & \frac{1}{\|\mathbf{x}_{625} - \mathbf{c}_2\|} & \cdots & \frac{1}{\|\mathbf{x}_{625} - \mathbf{c}_9\|} \end{bmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_9 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{625} \end{pmatrix}$$

$$\mathbf{1} = \mathbf{A}\mathbf{p} + \boldsymbol{\epsilon}$$

Calculando la solución  $\hat{\mathbf{p}}$  de minimizar la suma de los cuadrados de los errores,  $\boldsymbol{\epsilon}^T \boldsymbol{\epsilon}$ , al graficar el resultado  $\mathbf{A}\hat{\mathbf{p}}$  se obtiene:

## Ejemplo 2 (IX)

Gráfica de  $\mathbf{A}\hat{\mathbf{p}}$ , donde  $\hat{\mathbf{p}}$  es la solución de mínimos cuadrados:



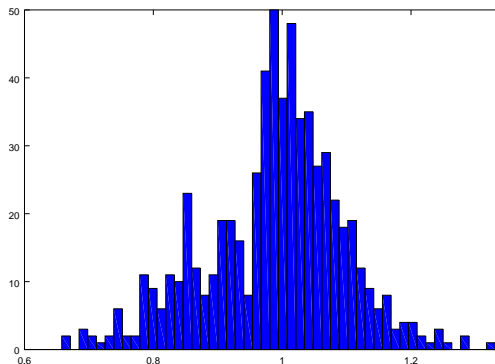
## Ejemplo 2 (X)

Se ve que no se consigue que la iluminación en el piso sea uniforme. Las componentes del vector solución  $\hat{\mathbf{p}}$  del problema de mínimos cuadrados son:

$$\begin{aligned}\hat{p}_1 &= 1.973286 \\ \hat{p}_2 &= 0.896566 \\ \hat{p}_3 &= 1.449870 \\ \hat{p}_4 &= 0.609335 \\ \hat{p}_5 &= 0.097823 \\ \hat{p}_6 &= 0.606314 \\ \hat{p}_7 &= 1.717131 \\ \hat{p}_8 &= 1.146629 \\ \hat{p}_9 &= 1.640506\end{aligned}$$

El siguiente histograma muestra las frecuencias de las componentes del vector  $\mathbf{A}\hat{\mathbf{p}}$ . Se ve que los valores están centrados alrededor de 1, pero hay variabilidad.

## Ejemplo 2 (XI)



Al menos el resultado muestra que todas las celdas están iluminadas y que tiene sentido la solución porque todas sus componentes son no negativas.

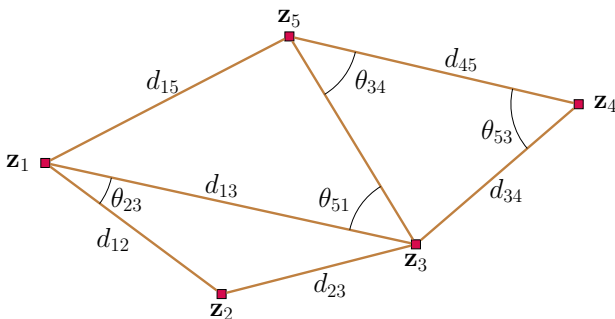


## Ejemplo sobre estimación de áreas territoriales (I)

Supongamos que se tienen las coordenadas geográficas de algunos puntos  $\mathbf{w}_i$ . También algunas longitudes entre estos puntos o ángulos entre segmentos que unen a estos puntos.

En las mediciones puede haber errores, de modo que las verdaderas coordenadas  $\mathbf{z}_i$  se relacionan con los datos mediante

$$\mathbf{z}_i = \mathbf{w}_i + \boldsymbol{\epsilon}_i \quad i = 1, \dots, m.$$



## Ejemplo sobre estimación de áreas territoriales (II)

A través de identidades trigonométricas se puede establecer algunas relaciones entre los datos. Por ejemplo

$$\cos \theta_{ij} = \frac{(\mathbf{z}_i - \mathbf{z}_k)^T (\mathbf{z}_j - \mathbf{z}_k)}{l_{ik} l_{jk}}$$

- Hay que construir más ecuaciones que incógnitas. Si  $\mathbf{z}_i$  tiene la información de latitud y longitud, hay que construir más de  $2m$  ecuaciones. Si  $\mathbf{z}_i$  se expresa en términos de latitud, longitud y altitud, se necesitan más de  $3m$  ecuaciones.
- Las ecuaciones no son lineales.
- A partir de todas estas ecuaciones se puede plantear un problema de mínimos cuadrados no lineal para determinar las posiciones  $\mathbf{z}_i$ .
- Hay métodos iterativos para resolver problemas de mínimos cuadrados no lineales.

## Ejemplo sobre estimación de áreas territoriales (III)

En 1974, la *National Geodetic Survey* en EEUU inicio el proyecto de reajustar la base *North American Datum*, la cual es una red de puntos de referencia cuyas longitudes, latitudes y, en algunos casos, las altitudes, deberían ser conocidas con una precisión en el orden de centímetros, ya que en algunos sitios el error era de 10 metros.

La base de datos consta de 200,000 puntos descritos por su longitud y latitud, más 500,000 puntos descritos por su longitud, latitud y altitud. El problema estuvo formado por aproximadamente 2.5 millones de ecuaciones con 400,000 incógnitas.

La expectativa era terminar el proyecto en 1983, considerando que se iba a requerir 700 horas de tiempo de cómputo.