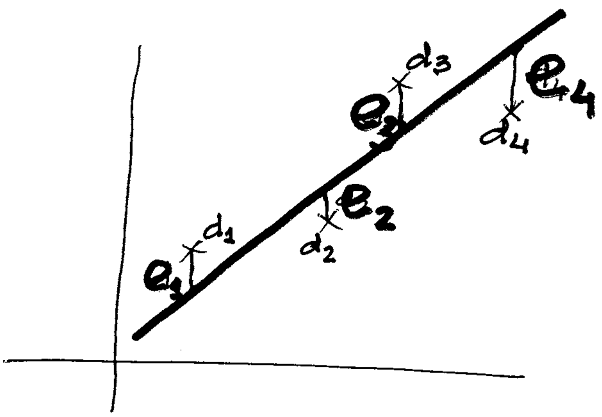
Práctica 1

*) Concepto de aproximación por mínimos cuadrados

Consideramos una serie de datos que se disponen en una nube de puntos. Nuestro objetivo es buscar una curva que aproxime tales datos.

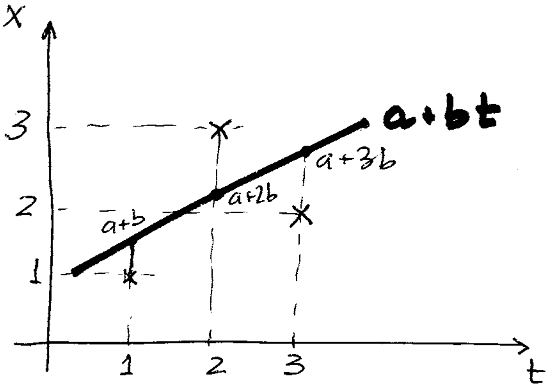
El problema que surge primero es cómo medir en cuánto nos equivocamos al tomar una curva u otra. Una posibilidad es usar el error cuadrático.



Error cuadrático $= e_1^2 + e_2^2 + e_3^2 + e_4^2 = E_2$

El método de aproximación por mínimos cuadrados consiste en buscar una curva (de una familia de curvas dada) que haga que el error cuadrático sea el menor posible.

* Ejemplo: recta de mínimos cuadrados



Datos: (1,1), (2,3), (3,2)

Familia de rectas:

$$x(t) = a + bt$$

$\downarrow \quad \downarrow$

Parámetros

$$E_2 = (1-a-b)^2 + (3-a-2b)^2 + (2-a-3b)^2$$

Al buscar el mínimo error surge un problema de cálculo de mínimos para funciones de varias variables.

"Mínimo de $F(a,b) = (1-a-b)^2 + (3-a-2b)^2 + (3-a-3b)^2$"

Para buscar el mínimo utilizamos las derivadas parciales de F:

$$\left.\begin{aligned} \frac{\partial F}{\partial a} &= 0 \\ \frac{\partial F}{\partial b} &= 0 \end{aligned}\right\}$$ Tenemos un sistema de ecuaciones fácil de resolver.

Cuando utilizamos familias de curvas correspondiente a funciones polinómicas, los sistemas obtenidos son fáciles de resolver. Otra cuestión es ajustar por funciones exponenciales, logarítmicas, logísticas, etc.

*) Mínimos cuadrados. - Linealización

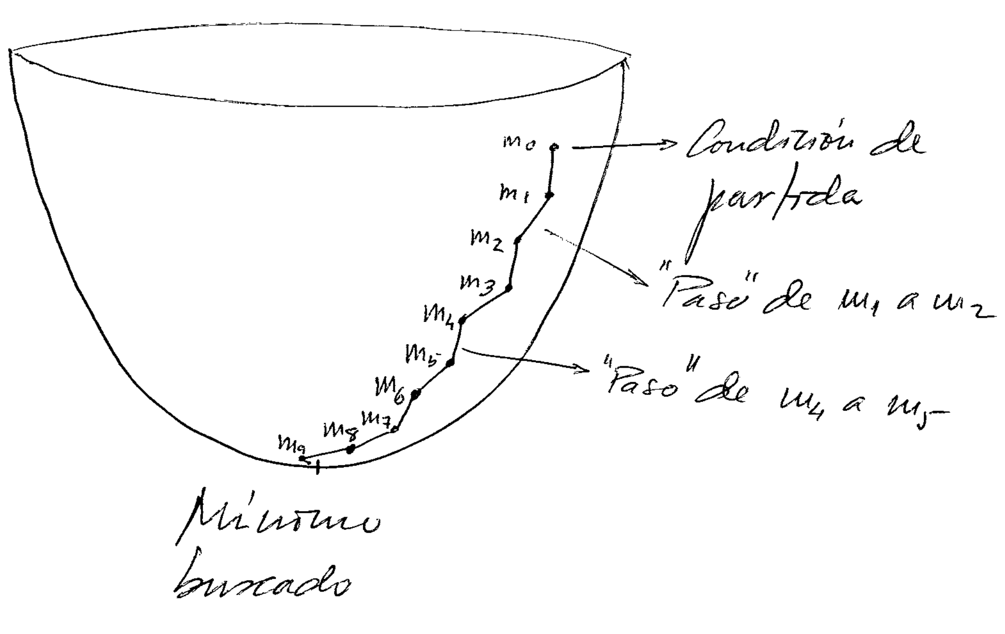
Cuando ajustamos por familias generales, una posibilidad de obtener sistemas fáciles de resolver es emplear el método de linealización. Esto lo veremos en el Tema 4.

*) Métodos numéricos de resolución de sistemas

Otra posibilidad para ajustar por familias de curvas generales es usar métodos de resolución aproximada de los sistemas obtenidos.

En la Práctica 1 seguiremos este camino: para resolver los sistemas obtenidos utilizaremos un método perteneciente a la familia de los métodos de gradiente.

Gráficamente este método consiste en:



1) Comenzamos con unos valores concretos de los parámetros (m_0)

2) "Avanzamos" "buscando" el mínimo ($m_0 \to m_1 \to m_2 \to \ldots$)

3) Obtenemos una sucesión de puntos $\{m_n\}_{n\geq 0}$ que se corresponden con sucesivas elecciones de los parámetros.

4) "Si todo va bien", nuestra sucesión estará cada vez más cerca del mínimo.

Surgen varias cuestiones:

i) ¿Dónde empezamos (m_0)?

ii) ¿Cuánto avanzamos cada vez (paso de m_n a m_{n+1})?

iii) ¿Cuándo hemos llegado al mínimo?

iv) ¿Podemos llegar al mínimo?

En la Práctica 1 intentaremos responder a estas preguntas.

*) Error medio en %

Para poder comparar ajustes, correspondientes a distintas familias, utilizaremos el "error medio en %".

Si denotamos por E_2 al error cuadrático, el "error medio en %" viene dado por la expresión:

$$\frac{\sqrt{E_2/n}}{\text{Media de los datos}} \times 100.$$

Como los datos se corresponderán con magnitudes positivas (población, altura, peso, etc) no hay problema con el caso "media = 0".

* Familia exponencial

$$x' = cx \Rightarrow x(t) = a e^{c(t-t_0)}$$

$$\begin{cases} a \equiv \text{valor en } t_0 \; (x(t_0)) \\ t_0 \equiv \text{punto medio de los datos} \\ c \equiv \text{tasa de crecimiento } \left(\frac{x'}{x}\right) \\ ac \equiv \text{pendiente en } t_0 \; (x'(t_0)) \end{cases}$$

Al ajustar con esta familia, realmente sólo necesitamos dos parámetros. Como tenemos tres (a, t_0, c) fijaremos t_0; esto es, t_0 será un valor concreto que fijaremos nosotros.

* Familia logística

$$x' = \frac{c}{a}\, x\,(a-x) \Rightarrow x(t) = \frac{a}{1+e^{-c(t-t_0)}}$$

$$\left[x' = c\,x\left(1-\frac{x}{a}\right)\right]$$

$a \equiv$ valor límite en $+\infty$ $\left(\lim_{t\to+\infty} x(t) = a\right)$ si $c>0$

$t_0 \equiv$ punto de inflexión

$\frac{c}{2} \equiv$ tasa de crecimiento en t_0 $\left(\frac{x'(t_0)}{x(t_0)}\right)$

$\frac{ac}{4} \equiv$ pendiente en t_0 $(x'(t_0))$

* Familia de Gompertz

$$x' = c\,x\,(\ln a - \ln x) \Rightarrow x(t) = a\,e^{-e^{-c(t-t_0)}}$$

$$\left[x' = c\,x \ln\frac{a}{x}\right]$$

$a \equiv$ valor límite en $+\infty$ $\left(\lim_{t\to+\infty} x(t) = a\right)$ si $c>0$

$t_0 \equiv$ punto de inflexión

$c \equiv$ tasa de crecimiento en t_0 $\left(\frac{x'(t_0)}{x(t_0)}\right)$

$\frac{ac}{e} \equiv$ pendiente en t_0 $(x'(t_0))$

$(e \approx 2'71828)$