Aproximación mínimo cuadrática (C1)

Profesora: Ángela León Mecías

Introducción

Como se conoce, en la vida real se presentan diversas situaciones en las que es necesario aproximar una función, ya sea porque la expresión con que se cuenta es muy complicada para evaluar y operar con ella, ó porque sólo se conocen valores en ciertos puntos del dominio de definición. En el tema anterior se estudió la aproximación de funciones por interpolación, que se utiliza cuando la precisión de los datos con que se cuenta es suficiente como para que se puedan considerar exactos. Ahora bien si además de contar con muchos datos, se presenta una de las siguientes situaciones:

- éstos se consideran afectados de error ó
- se conoce la expresión analítica de la función, pero es muy irregular como se muestra en la figura (1)

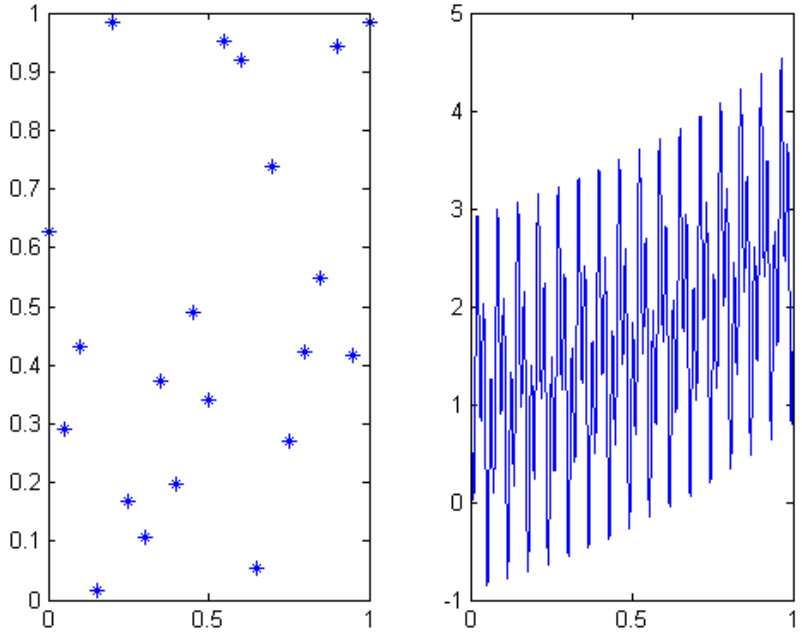

Figura 1:

entonces no tiene sentido obligar a que la función de aproximación pase exactamente por estos valores, es decir no tiene sentido aproximar la función por interpolación. Sin embargo por lo general se puede establecer la forma aproximada que tendrá la función de aproximación (un polinomio lineal como en la figura (2), uno cuadrático como en la figura (3), una función trigonométrica, entre otras).

Cuando se trata de encontrar una función que aproxime un conjunto de datos se deben considerar los siguientes aspectos,

- las características de la función f continua o discreta, que debe ser aproximada,
- el conjunto Φ donde se seleccionan las funciones de aproximación F
- la función error $f - F$ y la norma que se use para medir la calidad de la aproximación.

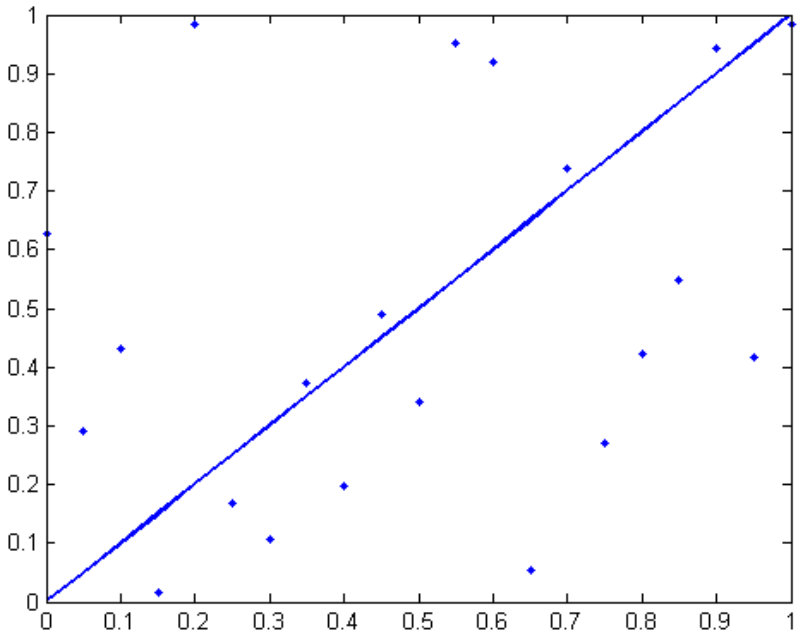

Figura 2:

Problema de aproximación mínimo cuadrática

El problema de aproximar una función se puede considerar en un marco bien general; dado un espacio lineal normado L y un subespacio Φ del espacio lineal normado L ($\Phi \subset L$). Entonces, dada $f \in L$, el problema de aproximación consiste en determinar la función $\widehat{F} \in \Phi$ que más cerca esté de f según una norma dada

$$\left\| f - \widehat{F} \right\| = \min_{F \epsilon \Phi} \|f - F\| . \tag{1}$$

Si la distancia entre las dos funciones (residual) se mide según la norma Euclideana, entonces estamos ante un problema de mínimos cuadrados (tomar otras normas define otros tipos de aproximación que no trataremos aquí, pero se pueden consultar en [], []). El problema de aproximación mínimo cuadrática se puede encontrar relacionado con diferentes aplicaciones como las que se relacionan a continuación

- Resolver $Ax = b$, $A_{m \times n}$, con $m > n$ ó $m < n$ que implica $min_x \|Ax - b\|_2$, (vista en el primer semestre)
- ajuste de curvas
- modelación estadística de datos con ruido
- modelación geodésica
- problema de optimización sin restricciones

Nosotros comenzaremos considerando el problema de mínimos cuadradros desde el punto de vista del ajuste de curvas. Supongamos que se tiene una función $f : \mathbb{R} \to \mathbb{C}$ y se quiere encontrar $\widehat{F} \in \Phi$ que mejor aproxime a f en el conjunto de funciones

$$\Phi = \left\{ F(x, c) : \mathbb{R} \to \mathbb{C}, c \in \Re^{n+1} \right\}$$

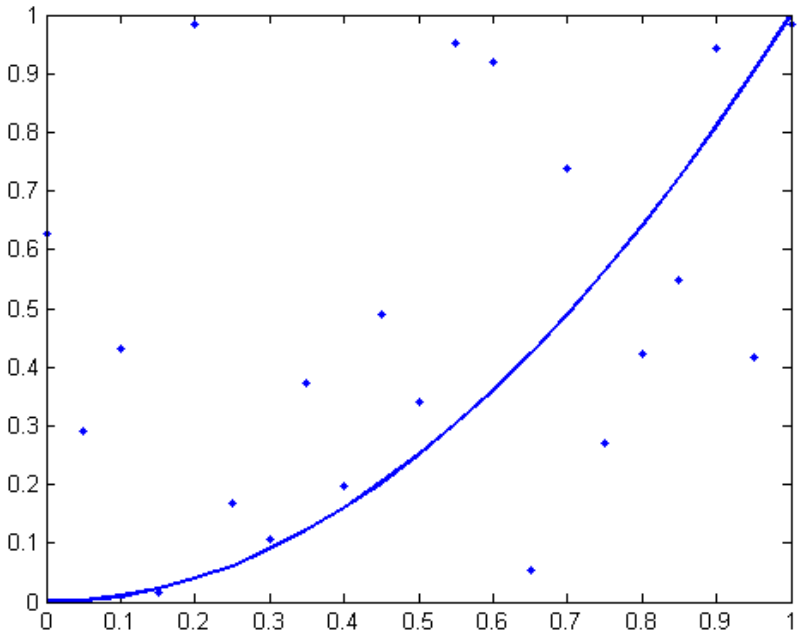

Figura 3:

Definición 1 *Dada una función f y una familia de funciones Φ, determinadas a priori, la función $\widehat{F} \in \Phi$, es la aproximación mínimo cuadrática de f si existen parámetros $c^* = (c_i^*)_{i=0,..n}$, tales que*

$$r_{min} = \left\| f - \widehat{F}(x, c^*) \right\|_2 = min_{c \in \Re^{n+1}} \|f - F(x, c)\|_2$$

En el caso de la norma discreta (ver al final del documento expresiones de las normas), se tiene

$$\left\| f - \widehat{F}(x, c^*) \right\|_2 = min_{c \in \Re^{n+1}} \left[\sum_{i=0}^{N} (f(x_i) - F(x_i, c_0, c_1, ...c_n))^2 \right]^{\frac{1}{2}}$$

El tratamiento del problema y las vías de solución están relacionados con el hecho de que

- la función $F(x, c_0, .., c_n)$, depende linealmente de los c_i
- la función $F(x, c_0, .., c_n)$, no depende linealmente de los c_i

Problema de aproximación mínimo cuadrática lineal

Supongamos que se tiene una función $f : \mathbb{R} \to \mathbb{C}$ y $F(x, c)$ es de la forma

$$F(x, c) = \sum_{j=o}^{n} c_j \varphi_j(x),$$

siendo $\{\varphi_j(x)\}_{j=o}^n$ un conjunto de funciones linealmente independientes conocidas, es decir cuando $F(x)$ depende linealmente de los coeficientes desconocidos c_j, entonces estamos ante un problema de aproximación mínimo cuadrática lineal, es decir el subespacio Φ es de dimensión finita y está generado por $\{\varphi_k\}_{k=0}^n$. Entonces el problema

a resolver es

$$\left\| f - \sum_{j=o}^{n} c_j^* \varphi_j(x) \right\|_2^2 = \min_{c \in \Re^{n+1}} \left\| f - \sum_{j=o}^{n} c_j \varphi_j(x) \right\|_2^2 \tag{2}$$

Se considera que la norma fue inducida por un producto escalar, $\|f\| = \sqrt{\langle f, f \rangle}$ y como producto escalar se tiene:

- producto escalar continuo $\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}dx$
- producto escalar discreto $\langle f, g \rangle = \sum_{i=1}^{n} f(x_i)\overline{g(x_i}$

Entonces el problema (2) se reduce a

$$\min_{c \in \Re^{n+1}} \langle f - F(x,c), f - F(x,c) \rangle$$

Note que si denotamos $g(c) = \|f - F(x,c)\|$, entonces $g(c) \geq 0$, y $(g(c))^2$ será una función monótona creciente, por tanto

$$ming(c) \Leftrightarrow min(g(c))^2$$

$$\left\| f - \widehat{F}(x, c^*) \right\|_2 = \min_{c \in \Re^{n+1}} \|f - F(x,c)\|_2 \tag{3}$$

es decir se quiere hallar una función aproximante que minimice la norma Euclideana de la función error

$$\left\| f - \widehat{F}(x, c^*) \right\|_{2,\omega} = \left(\int_a^b \omega(x)[f(x) - \widehat{F}(x,c)]^2 dx \right)^{1/2}, \text{ en el caso continuo}$$

$$\left\| f - \widehat{F}(x, c^*) \right\|_{2,\omega,M} = \left(\sum_{i=1}^{m} \omega(x_i)[f(x_i) - \widehat{F}(x_i,c)]^2 \right)^{1/2}, \text{ en el caso discreto}$$

Es importante observar que la elección de la función de peso $\omega(x)$ y los pesos $\omega(x_i)$ respectivamente, que aparecen en las expresiones anteriores, afecta a $\widehat{F}$. En el caso *continuo,* con una elección adecuada se puede forzar a que $\widehat{F}$ concuerde mejor con f en una parte de $[a,b]$ que en el resto del intervalo, veamos

- $\omega(x) = 1$ en $[a,b]$, asigna igual peso a los valores de la función error en todo $x \in [a,b]$
- $\omega(x) = 1/\sqrt{1-x^2}$ en $(-1,1)$, asigna mayor peso al error cerca de $x = -1$ y $x = 1$
- $\omega(x) = e^{-x}$ en [0,∞), asigna peso máximo al error en $x = 0$, decreciente cuando $x \longrightarrow \infty$
- $\omega(x) = e^{-x^2}$ en $(-\infty, +\infty)$, asigna peso máximo al error en $x = 0$, decreciente cuando $x \longrightarrow \pm\infty$

En el caso discreto, un valor grande $\omega_i = \omega(x_i)$ significa que al valor del error $f_i - F_i$ se le confiere mucha importancia porque f_i fue medido con gran precisión, y un valor ω_i pequeño es indicador de poca confiabilidad del valor f_i (en la terminología estadística, se dice que (x_i, f_i) es un punto mentiroso o *outlier* en este caso). Nosotros comenzaremos considerando el caso discreto con $\omega(x) = 1$. El siguiente teorema es la base para la determinación de la mejor aproximación mínimo cuadrática lineal $\widehat{F}$, tanto en el caso continuo como en el discreto.

Teorema 2 *Sean las funciones $\varphi_o, \varphi_1, ..., \varphi_n$ l.i. y que generan al subespacio Φ. Entonces existe una única función $\widehat{F}$ de la forma $\widehat{F} = \sum_{j=o}^{n} c_j^* \varphi_j$, tal que*

$$\left\| f - \widehat{F} \right\|_2^2 \leq \|f - F\|_2^2, \ \forall F = \sum_{j=o}^{n} c_j \varphi_j,$$

$\widehat{F}$ es también solución del sistema de ecuaciones lineales que se obtiene resolviendo las ecuaciones normales:

$$< f - \widehat{F}, \varphi_k >= 0, \quad 0 \leq k \leq n.$$

y viceversa.

Demostración:

Comencemos con la obtención del sistema de ecuaciones normales. Teniendo en cuenta las propiedades del producto escalar y la forma de $\widehat{F}$,

$$\begin{aligned} &< f - \widehat{F}, \varphi_k >= 0 \Leftrightarrow < \widehat{F}, \varphi_k >=< f, \varphi_k > \\ &0 \leq k \leq n \end{aligned}$$

sustituyendo $\widehat{F}$

$$< \sum_{j=o}^{n} c_j^* \varphi_j, \varphi_k >=< f, \varphi_k >$$

y teniendo en cuenta las propiedades asociativa y distributiva del producto escalar, se llega a que

$$\sum_{j=o}^{n} c_j^* < \varphi_j, \varphi_k >=< f, \varphi_k > \, , \, 0 \leq k \leq n, \tag{4}$$

lo cual constituye el sistema de ecuaciones lineales

$$\begin{array}{ll} k=0: & c_o^* < \varphi_o, \varphi_o > + c_1^* < \varphi_1, \varphi_o > + \cdots + c_n^* < \varphi_n, \varphi_o >=< f, \varphi_o > \\ k=1: & c_o^* < \varphi_o, \varphi_1 > + c_1^* < \varphi_1, \varphi_1 > + \cdots + c_n^* < \varphi_n, \varphi_1 >=< f, \varphi_1 > \\ \cdots & \cdots \\ k=n: & c_o^* < \varphi_o, \varphi_n > + c_1^* < \varphi_1, \varphi_n > + \cdots + c_n^* < \varphi_n, \varphi_n >=< f, \varphi_n > \end{array}$$

que se puede escribir en forma matricial como

$$Bc^* = h$$

$$B = \begin{bmatrix} < \varphi_o, \varphi_o > & < \varphi_1, \varphi_o > & \cdots & < \varphi_n, \varphi_o > \\ < \varphi_o, \varphi_1 > & < \varphi_1, \varphi_1 > & \cdots & < \varphi_n, \varphi_1 > \\ \cdots & \cdots & \cdots & \cdots \\ < \varphi_o, \varphi_n > & < \varphi_1, \varphi_n > & \cdots & < \varphi_n, \varphi_n > \end{bmatrix} \tag{5}$$

$$h = \begin{bmatrix} < f, \varphi_o > \\ < f, \varphi_1 > \\ \cdots \\ < f, \varphi_n > \end{bmatrix}$$

$c = (c_o^*, c_1^*, ..., c_n^*)^T$ y se conoce como sistema de las ecuaciones normales (SEN) o de Gauss. Este sistema $Bc^* = h$ también puede escribirse como $\phi^T \phi c^* = \phi^T f$, donde ϕ es una matriz de N filas y n columnas

$$\phi = \begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \cdots & \phi_n(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \ddots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_n(x_N) \end{bmatrix} \tag{6}$$

Observe que c^* es solución de $\phi^T \phi c^* = \phi^T f$ sí y sólo sí es solución del sistema lineal sobredeterminado $\phi c^* = f$.

La denominación de normales proviene del hecho que

$$< f - \widehat{F}, \phi_k >= 0 \;\; \text{equivale a que } f - \widehat{F} \perp \Phi \tag{7}$$

según la generalización del concepto de ortogonalidad, pues la distancia mínima de f al subespacio Φ está dada por la longitud del vector $f - \widehat{F}$, siendo $\widehat{F}$ su proyección ortogonal. Nótese que debido a la conmutatividad del producto escalar, la matriz B es simétrica. La matriz de los productos escalares B del SEN (5) es una matriz de Gram, y se puede demostrar que es definida positiva siempre que las funciones φ_j sean linealmente independientes.

Definición 3 *Una matriz A es definida positiva (semidefinida positiva), $A \succ 0 (\succeq 0)$ si y sólo sí $x^T Ax > 0 (x^T Ax \geq 0)$ para toda $x \neq 0 (x \in \Re^n)$*

Luego el sistema tiene solución única c^*, que define la función de mejor aproximación. Para demostrar que cualquier función $F = \sum_j c_j\varphi_j$ con al menos un $c_j \neq c_j^*$ tiene mayor distancia a f que $\widehat{F}$, planteamos la diferencia $f - F = f - \sum_j c_j\varphi_j$. Sumando y restando $\widehat{F}$

$$\begin{aligned} f - F &= (f - \widehat{F}) + (\widehat{F} - \sum_j c_j\varphi_j) \\ &= (f - \widehat{F}) + \sum_j (c_j^* - c_j)\varphi_j \ . \end{aligned}$$

Entonces,

$$\begin{aligned} \|f - F\|_2^2 &=< f - F,\ f - F > \\ &=< f - \widehat{F} + \sum_j (c_j^* - c_j)\varphi_j,\ f - \widehat{F} + \sum_j (c_j^* - c_j)\varphi_j > \\ &=< f - \widehat{F}, f - \widehat{F} > +2 < \sum_j (c_j^* - c_j)\varphi_j,\ f - \widehat{F} > \\ &+ < \sum_j (c_j^* - c_j)\varphi_j, \sum_j (c_j^* - c_j)\varphi_j > \\ &= \left\| f - \widehat{F} \right\|_2^2 + \left\| \sum_j (c_j^* - c_j)\varphi_j \right\|_2^2, \end{aligned}$$

pues teniendo en cuenta (7), el sumando que contiene el coeficiente 2 se anula, y como al menos un $c_j \neq c_j^*$, el segundo sumando en la última expresión es estrictamente positivo, y queda

$$\|f - F\|_2^2 \geq \left\| f - \widehat{F} \right\|_2^2,$$

con lo que se completa la demostración.

Lo que se acaba de demostrar es totalmente congruente y equivalente con las exigencias de optimalidad que aseguran la existencia de la solución del problema (2), veamos,

$$\left\langle \sum_{k=0}^{n} c_k\varphi_k(x) - f(x), \sum_{k=0}^{n} c_k\varphi_k(x) - f(x) \right\rangle = \sum_{k=0}^{n} c_k \sum_{j=0}^{n} c_j \left\langle \varphi_k(x), \varphi_j(x) \right\rangle \tag{8}$$

$$- 2\sum_{0}^{n} c_k Re(\langle f, \varphi_k(x) \rangle) + \langle f, f \rangle \tag{9}$$

Como el último sumando no depende de las variables con respecto a las que se está optimizando y asumiendo que $F(x, c)$ y $f(x)$ toman valores reales, pues es suficiente resolver

$$\min_{c \in R^n} \sum_{k=0}^{n} c_k \sum_{j=0}^{n} c_j \left\langle \varphi_k(x), \varphi_j(x) \right\rangle - 2\sum_{0}^{n} c_k \left\langle f, \varphi_k(x) \right\rangle \tag{10}$$

Veamos cuáles son las condiciones de optimalidad para el problema

$$\min_{c \in \Re^n} g(c) \tag{11}$$

Definición 4

- *c^* es un mínimo global de (11) si $\forall\ c \in \Re^n$; $g(c) \geq g(c^*)$*
- *Si existe una vecindad V_{c^*} de c^* tal que $\forall\ c \in V_{c^*} \cap \Re^n$; $g(c) \geq g(c^*)$, entonces c^* es un mínimo local de (11)*

Definición 5 *La función $g(x)$ es convexa si $\forall x_1, x_2 \in \Re^n$, $\alpha \in [0, 1]$ se tiene*

$$g(\alpha(x_1) + (1 - \alpha)x_2) \leq \alpha g(x_1) + (1 - \alpha)g(x_2)$$

Además si $g(c) \in \mathbb{C}^2$ entonces g es convexa si y sólo sí $\nabla^2 g(c) \succ 0$. Por otro lado la condición necesaria de mínimo local es como sigue:

Teorema 6 *Si c^* es mínimo local de (11) entonces*

- *si $g \in C^1$ entonces $\nabla g(c^*) = 0$.*
- *si $g \in C^2$ entonces $\nabla g(c^*) = 0$ y $\nabla^2 g(c^*)$ es semidefinida positiva.*

Condiciones suficientes

Teorema 7 *Si $g \in C^2$, $\nabla g(c^*) = 0$ y $\nabla^2 g(c^*)$ es definida positiva entonces c^* es un mínimo local de (11).*

Ahora bien si g es convexa, entonces la condición $\nabla g(c^*) = 0$ es suficiente para la existencia del mínimo global. Retomando el problema (10), la función $g(c) \in C^\infty$

$$g(c) = \sum_{k=0}^{n} c_k \sum_{j=0}^{n} c_j \langle \varphi_k(x), \varphi_j(x) \rangle - 2 \sum_{0}^{n} c_k \langle f, \varphi_k(x) \rangle \tag{12}$$

y $\nabla g = 2Bc - 2h$, con B la matriz de los productos escalares obtenida más arriba y h el vector de los productos escalares de f con las funciones φ_k que se dijo son l.i., por tanto $\nabla g(c) = 0$ es equivalente a resolver el sistema de ecuaciones normales lineales $Bc = h$, lo cual se demostró tiene solución única c^*, ya que precisamente al ser las φ_j l.i. la matriz B es definida positiva, con lo cual se obtiene que $\nabla^2 g = 2B$ es definida positiva y por tanto esto implica que la función g es convexa, de ahí que c^* es el único mínimo global.

Caso discreto, aproximación por polinomios

Si $F \epsilon P_n = \Phi$, entonces

$$F(x) = c_o + c_1 x + c_2 x^2 + \cdots + c_n x^n = \sum_{j=o}^{n} c_j x^j,$$

y se dispone de la función f que se quiere aproximar en la forma de una tabla de $N+1$ pares ($N >> n$, N grande):

x	x_o	x_1	$\cdots$	x_N
f(x)	f_o	f_1	$\cdots$	f_N

En este caso, pueden tomarse como funciones linealmente independientes conocidas las potencias de x:

$$\{\varphi_j(x)\}_{j=o}^{n} = \left\{x^j\right\}_{j=o}^{n}.$$

Entonces el producto escalar estará definido (con función de peso $\omega(x) = 1$), por

$$< \varphi_j, \varphi_k >= \sum_{i=o}^{N} \varphi_j(x_i)\varphi_k(x_i) = \sum_{i=o}^{N} x_i^j x_i^k = \sum_{i=o}^{N} x_i^{j+k}$$

y

$$< f, \varphi_k >= \sum_{i=o}^{N} f(x_i)\varphi_k(x_i) = \sum_{i=o}^{N} f_i x_i^k$$

obteniéndose, por ejemplo,

si $j = k = 0$ $< \varphi_o, \varphi_o >= \sum x_i^{o+o} = \sum 1 = N + 1$

si $j = k = 1$ $< \varphi_1, \varphi_1 >= \sum x_i^{1+1} = \sum x_i^2$

si $j = 0, k = 1$ $< \varphi_o, \varphi_1 >= \sum x_i^{o+1} = \sum x_i$, etc.

El sistema () de las ecuaciones normales tendrá ahora la forma

$$\begin{bmatrix} N+1 & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+1} \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum x_i^n & \sum x_i^{n+1} & \sum x_i^{n+2} & \cdots & \sum x_i^{2n} \end{bmatrix} \begin{bmatrix} c_o^* \\ c_1^* \\ c_2^* \\ \vdots \\ c_n^* \end{bmatrix} \tag{13}$$

$$= \begin{bmatrix} \sum f_i \\ \sum x_i f_i \\ \sum x_i^2 f_i \\ \vdots \\ \sum x_i^n f_i \end{bmatrix}, \tag{14}$$

y bastará resolverlo para hallar el vector c^* de los coeficientes del polinomio $\widehat{F}$ de la mejor aproximación mínimo cuadrática.

Aspectos computacionales de la aproximación por polinomios

La determinación de la matriz B y el correspondiente término independiente h

$$B = \begin{bmatrix} N+1 & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum x_i^n & \sum x_i^{n+1} & \sum x_i^{n+2} & \cdots & \sum x_i^{2n} \end{bmatrix},$$

$$h = \begin{bmatrix} \sum f_i \\ \sum x_i f_i \\ \cdots \\ \sum x_i^n f_i \end{bmatrix}$$

requieren el cálculo de todas las sumatorias que éstos contienen. Para obtener expresiones que faciliten la automatización de dicho cálculo, denotemos por X y f, respectivamente la matriz de datos de orden $(N+1)\times(n+1)$ y el vector de $N+1$ componentes siguientes:

$$X = \begin{bmatrix} 1 & x_o & x_o^2 & \cdots & x_o^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_N & x_N^2 & \cdots & x_N^n \end{bmatrix}, \qquad f = \begin{bmatrix} f_o \\ f_1 \\ \cdots \\ f_N \end{bmatrix}.$$

Se puede demostrar que entonces B y f se pueden calcular mediante:

$$B = X^T X \quad \text{y} \quad h = X^T f.$$

La matriz X es una matriz del tipo Vandermonde, y se genera fácilmente a partir del vector

$$x = (x_o, x_1, \ldots, x_N)^T.$$

Cuando la función de aproximación F es polinómica, la matriz B es desbalanceada (sus filas y columnas son de orden diverso), lo que ocasiona problemas con la propagación de los errores de redondeo, y si n es grande, resulta ser una matriz mal condicionada.

¿Qué se puede hacer con vista a obtener la solución de las ecuaciones normales con máxima precisión?

- tomar n lo menor posible, siempre que el comportamiento de F sea semejante a f
- elegir cuidadosamente el conjunto $\{\varphi_j\}_{j=o}^n$ para que además de linealmente independiente, sea ortogonal
- no generar las ecuaciones normales $X^T X c = X^T f$, sino tratar directamente la minimización de $\|f - Xc\|$ por la vía de descomponer la matriz de datos X en factores

Aproximación mínimo cuadrática y estadística

En la modelación estadística de datos con ruido la aproximación mínimo cuadrática se identifica, en el caso discreto, con el problema llamado de ajuste de datos o determinación de una función de regresión.

Aproximación por una recta (o regresión lineal)

La regresión lineal es el caso más frecuente en la práctica, que coincide con la aproximación mediante una recta, siendo $\widehat{F}(x) = c_o^* + c_1^* x$ la función aproximante a determinar con $\{\varphi_j(x)\}_{j=o}^1 = \{1, x\}$ y los datos están representados por:

$$X = \begin{bmatrix} 1 & x_o \\ 1 & x_1 \\ \cdots & \cdots \\ 1 & x_N \end{bmatrix}, \qquad f = \begin{bmatrix} f_o \\ f_1 \\ \cdots \\ f_N \end{bmatrix}.$$

luego el sistema de las ecuaciones normales tiene la forma
$$\begin{bmatrix} N+1 & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} c_o^* \\ c_1^* \end{bmatrix} = \begin{bmatrix} \sum f_i \\ \sum x_i f_i \end{bmatrix}$$

Ejemplo 8 *Dada la función f por la tabla de valores*

x	1	2	3	4
$f(x)$	3	5	10	10

,hallar la recta de mejor aproximación mínimo cuadrática.

En este caso, n=1, N=4 y el conjunto de funciones linealmente independientes es $\{\varphi_j(x)\}_{j=o}^{1} = \{1, x\}$ y evaluando en el sistema de las ecuaciones normales se obtiene

$$\begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} \begin{bmatrix} c_o^* \\ c_1^* \end{bmatrix} = \begin{bmatrix} 28 \\ 83 \end{bmatrix}.$$

Aplicando el método de Gauss para resolver el sistema, se obtiene

$$c^* = \begin{bmatrix} 1/2 \\ 13/5 \end{bmatrix} \text{ , y por tanto, } \widehat{F}(x) = \frac{1}{2} + \frac{13}{5}x.$$

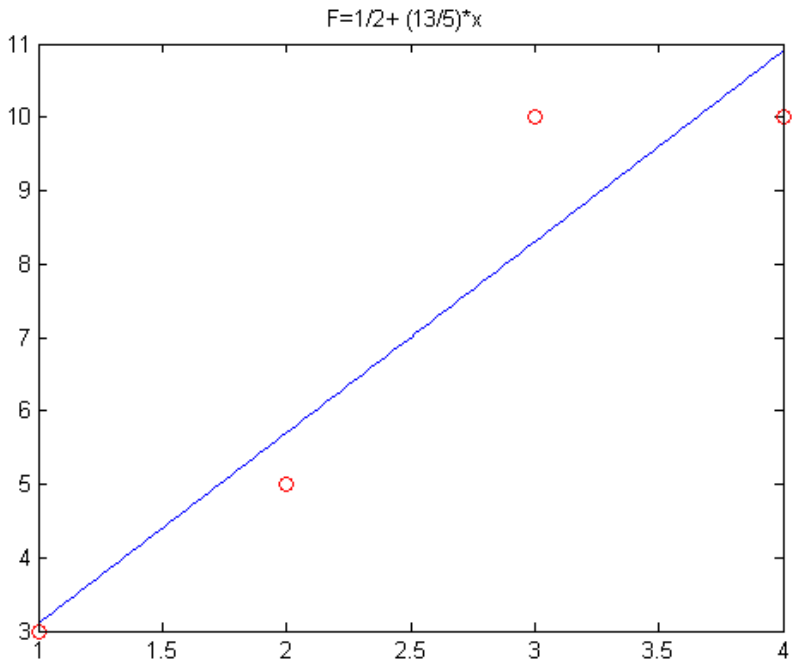

Figura 4:

Note que
$\widehat{F}(1) = \frac{1}{2} + \frac{13}{5} = 3{,}1 \neq 3 = f(1)$
$\widehat{F}(2) = \frac{1}{2} + \frac{26}{5} = 5{,}7 \neq 5 = f(2)$
$\widehat{F}(3) = \frac{1}{2} + \frac{39}{5} = 8{,}3 \neq 10 = f(3)$
$\widehat{F}(4) = \frac{1}{2} + \frac{52}{5} = 10{,}9 \neq 10 = f(4)$,
y en general, la recta no pasa por los puntos (x_i, f_i).
Caso particular: Aproximación mediante una función no lineal, linealizable

Tal es el caso cuando se aproxima por una función exponencial mínimo cuadrática de la forma $F(x) = c_o e^{c_1 x}$. La función aproximante no es intrínsecamente no lineal en este caso, ya que puede linealizarse aplicando logaritmos. Aplicando logaritmos se convierte en una recta:

$$\ln F(x) = \ln c_o + c_1 x,$$

o sea,

$$G(x) = c'_o + c_1 x,$$

donde $G(x) = \ln F(x)$ y $c'_o = \ln c_o$. Con la transformación logarítmica, se ha convertido la función aproximante F que depende en forma no lineal de c_o y c_1, en la función aproximante G, que depende de c'_o y c_1 linealmente. El problema de aproximación se convierte entonces en hallar $\widehat{G}$ tal que

$$\left\|\ln f - \widehat{G}\right\|_2^2 = \min_G \|\ln f - G\|_2^2 = \sum_{i=o}^{N} [\ln f_i - (c'_o + c_1 x_i)]^2$$

Está claro que los coeficientes $c_o^* = \exp(c_o'^*)$ y c_1^*, que se obtienen minimizando $\|\ln f - \ln F\|_2^2$, no coinciden con los que se obtendrían minimizando directamente $\|f - F\|_2^2$, los cuales son más difíciles de calcular debido a la no linealidad. Pero en la práctica, por evitar la resolución de un sistema no lineal, se aceptan como tales, pues son bastante cercanos debido a la inyectividad de la transformación logarítmica.Tomando en este caso, también $\{\varphi_j(x)\}_{j=o}^{1} = \{1, x\}$ y los datos representados por

$$X = \begin{bmatrix} 1 & x_o \\ 1 & x_1 \\ \cdots & \cdots \\ 1 & x_N \end{bmatrix}, \quad f = \begin{bmatrix} \ln f_o \\ \ln f_1 \\ \cdots \\ \ln f_N \end{bmatrix},$$

el SEN será:

$$\begin{bmatrix} N+1 & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} c_o'^* \\ c_1^* \end{bmatrix} = \begin{bmatrix} \sum \ln f_i \\ \sum x_i \ln f_i \end{bmatrix},$$

cuya solución $(c_o'^*, c_1^*)$ permite determinar finalmente $c_o^* = e^{c_o'^*}$, y definir la función de aproximación $\widehat{F}(x) = c_o^* e^{c_1^* x}$.

Aproximación lineal múltiple

Si la función empírica f depende linealmente de p variables, $f = f(x_1, x_2, \ldots, x_p)$, y se realizan $N+1$ observaciones , tendremos la tabla siguiente:

$$\begin{array}{ccccccc} \text{obs} & \text{x}_1 & \text{x}_2 & \cdots & \text{x}_p & \text{f} \\ 0 & \text{x}_{o1} & \text{x}_{o2} & \cdots & \text{x}_{op} & \text{f}_o \\ 1 & \text{x}_{11} & \text{x}_{12} & \cdots & \text{x}_{1p} & \text{f}_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \text{i} & \text{x}_{i1} & \text{x}_{i2} & \cdots & \text{x}_{ip} & \text{f}_i \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \text{N} & \text{x}_{N1} & \text{x}_{N2} & \cdots & \text{x}_{Np} & \text{f}_N \end{array},$$

Las funciones de aproximación tienen la forma

$$F(x_1, x_2, \ldots, x_p) = c_o + c_1 x_1 + c_2 x_2 + \cdots + c_p x_p,$$

y supuesto que los vectores $x_1, x_2, \ldots, x_p$ que definen las variables son linealmente independientes, puede considerarse el conjunto de funciones $\{\varphi_j(x)\}_{j=o}^{p} = \{1, x_1, x_2, \ldots, x_p\}$ y los datos representados por:

$$X = \begin{bmatrix} 1 & x_{o1} & x_{o2} & \cdots & x_{op} \\ 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Np} \end{bmatrix}, \quad f = \begin{bmatrix} f_o \\ f_1 \\ \vdots \\ f_N \end{bmatrix},$$

con lo cual, el SEN tendrá la forma:

$$\begin{bmatrix} N+1 & \sum x_{i1} & \sum x_{i2} & \cdots & \sum x_{ip} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \cdots & \sum x_{i1}x_{ip} \\ \sum x_{i2} & \sum x_{x2}x_{i1} & \sum x_{i2}^2 & \cdots & \sum x_{i2}x_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum x_{ip} & \sum x_{ip}x_{i1} & \sum x_{ip}x_{i2} & \cdots & \sum x_{ip}^2 \end{bmatrix} \begin{bmatrix} c_o^* \\ c_1^* \\ c_2^* \\ \vdots \\ c_p^* \end{bmatrix} = \begin{bmatrix} \sum f_i \\ \sum x_{i1}f_i \\ \sum x_{i2}f_i \\ \vdots \\ \sum x_{ip}f_i \end{bmatrix}.$$

Caso particular 1*:* Sistema lineal sobredeterminado.

La resolución aproximada de un sistema lineal sobredeterminado $Ac = b$, con $A_{n\times m}$, $b_{n\times 1}$, $n > m$, se puede interpretar como la aproximación del vector $b = f \,\epsilon\, \Re^n$ por la combinación lineal de las columnas de A, donde $a^{(j)}$ es la j-ésima columna de A, que minimice el residuo $r = b - Ac$:

$$\|b - A\widehat{c}\|_2 = \min_c \|b - Ac\|_2 .$$

En este caso, $\{\varphi_j\}_{j=1}^m = \left\{a^{(1)}, a^{(2)}, \ldots, a^{(m)}\right\}$, y los datos están representados por:

$$X = A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}, \quad f = b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

El SEN toma entonces la forma , $A^T Ac^* = A^T b$, donde c^* es el vector de los coeficientes desconocidos. Esto ya se vio en la asignatura MNI. En este contexto se demuestra el siguiente resultado

Teorema 9 *Sean X e Y dos espacios vectoriales de dimensión finita n y m sobre $\mathbb{R}$ y L una transformación lineal representada en dos bases X e Y por la A. Para un vector dado $b \in Y$, el vector $x \in X$ minimiza $\|Ax - b\|_2 \Longleftrightarrow A^T Ax = A^T b$*

Caso particular 2: Si la función empírica depende en forma no lineal de los coeficientes, pero es linealizable mediante la aplicación de logaritmos, obtenemos el caso lineal múltiple. Por ejemplo, si la función de aproximación es de la forma

$$F(x, y, z) = \alpha \frac{x^\beta y^\gamma}{z^\delta},$$

entonces

$$\ln F = \ln \alpha + \beta \ln x + \gamma \ln y - \delta \ln z,$$

βo sea,

$$G = \alpha' + \beta x' + \gamma y' - \delta z',$$

y tenemos una función de aproximación lineal múltiple G. Tomando en este caso $\{\varphi_j\}_{j=o}^3 = \{1, x', y', z'\}$ y los datos representados por:

$$X = \begin{bmatrix} 1 & \ln x_1 & \ln y_1 & -\ln z_1 \\ 1 & \ln x_2 & \ln y_2 & -\ln z_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \ln x_n & \ln y_n & -\ln z_n \end{bmatrix}, \quad f = \begin{bmatrix} \ln f_1 \\ \ln f_1 \\ \vdots \\ \ln f_n \end{bmatrix}$$

se obtiene el SEN para determinar $\widehat{G}$ con coeficientes α', β, γ y δ, que minimiza

$$\|\ln f - G\|_2 = \sum_{i=1}^n [\ln f(x_i, y_i, z_i) - (\alpha' + \beta \ln x_i + \gamma \ln y_i - \delta \ln z_i)]^2.$$

Problema de mínimos cuadrados no lineal

Definamos la función S como el cuadrado del error de la aproximación mínimo cuadrática:

$$S = \left\| f - \widehat{F} \right\|_2^2 = \sum_{i=o}^{N} [f(x_i) - F(x_i; c_o, c_1, \ldots, c_n)]^2, \tag{15}$$

$$S = S(c_o, c_1, \ldots, c_n).$$

que es una función continua, positiva y diferenciable de los parámetros $c_o, c_1, \ldots, c_n$ por lo menos hasta de segundo orden, $S : \mathbb{R}^n \to \mathbb{R}$. Luego el problema formulado como sigue: encontrar $c^* \in \Re^n$ tal que

$$S(c^*) = \underset{c \in \Re^n}{min} S(c)$$

es un problema de optimización sin restricciones. Entonces se aplican las condiciones de optimalidad vistas anteriormente

$$\nabla S(c) = \left(\frac{\partial S}{\partial c_o}, \frac{\partial S}{\partial c_1}, \ldots, \frac{\partial S}{\partial c_n} \right)^T = \overrightarrow{0_{\Re^{n+1}}}$$

Teniendo en cuenta (15), y derivando con respecto a los c_j, se obtiene

$$\begin{aligned} \frac{\partial S}{\partial c_j} &= \sum_{i=o}^{N} \left\{ \frac{\partial}{\partial c_j} \left([f(x_i) - F(x_i; c_o, \ldots, c_n)]^2 \right) \right\} \\ &= -2 \sum_{i=o}^{N} \left\{ [f(x_i) - F(x_i; c_o, \ldots, c_n] \frac{\partial F}{\partial c_j} \right\}, \end{aligned}$$

como habíamos visto la condición necesaria de extremo da lugar al sistema de las ecuaciones normales, que en este caso será no lineal

$$\sum_{i=o}^{N} \left\{ [f(x_i) - F(x_i; c_o^*, \ldots, c_n^*] \frac{\partial F}{\partial c_j} \right\} = 0, \quad 0 \le j \le n.$$

La solución del Sistema de Ecuaciones No Lineales (SEN) es el vector c^* que constituye el único mínimo de S y define la mejor función de aproximación mínimo cuadrática $\widehat{F}$.

Desde el punto de vista computacional, la dificultad fundamental está en la resolución del SEN, que exige el uso de métodos iterativos.

Recíprocamente, si tenemos un sistema (en general, no lineal) de n ecuaciones con m incógnitas:

$$f(x) = 0, \quad x = (x_1, x_2, \ldots, x_m)^T, \quad f : \Re^m \longrightarrow \Re^n,$$

o sea,

$$\begin{aligned} f_1(x) &= 0 \\ f_2(x) &= 0 \\ &\cdots \\ f_n(x) &= 0 \end{aligned}$$

y queremos minimizar el error residual

$$\|f(x)\|_2^2 = (f_1(x))^2 + (f_2(x))^2 + \cdots + (f_n(x))^2,$$

tenemos un problema de optimización sin restricciones.

Ejemplo 10 *Aproximar la función tabulada f:*

x	x_o	x_1	$\cdots$	x_N
$f(x)$	f_o	f_1	$\cdots$	f_N

por una función exponencial mínimo cuadrática de la forma $F(x) = c_o e^{c_1 x}$ *, sin linealizar.*

Aplicando la forma general del método de los mínimos cuadrados, definimos

$$S = \sum_{i=o}^{N} [f(x_i) - c_o e^{c_1 x_i}]^2.$$

Derivando con respecto a los c_j e igualando a cero, se obtien el SEN:

$$\sum_{i=o}^{N} [\exp(c_1^* x_i) f_i - c_o^* \exp(2c_1^* x_i)] = 0$$
$$\sum_{i=o}^{N} x_i [\exp(c_1^* x_i) f_i - c_o^* \exp(2c_1^* x_i)] = 0.$$

Nótese la no linealidad del SEN con respecto a c_o^* y c_1^*. Su resolución puede realizarse usando el método de Newton, lo que requiere definir una aproximación inicial $c^{(o)}$ que garantice la convergencia del proceso iterativo.

Formas de medir el error de la aproximación mínimo cuadrática

El error de la aproximación mínimo cuadrática está dado por

$$E = \left\| f - \widehat{F} \right\|_2 = \sqrt{< f - \widehat{F}, f - \widehat{F} >}$$

Una vez calculados los coeficientes c_j^* de la mejor aproximación $\widehat{F}$, basta sustituir en la expresión anterior para obtener el error. En el ejemplo sencillo resuelto anteriormente para la aproximación por la mejor recta mínimo cuadrática para el caso discreto, habrá que evaluar $\widehat{F}$ para las mismas abscisas, y calcular después $E = \sqrt{\sum_{i=o}^{3} [f_i - \widehat{F}_i]^2}$:

x	1	2	3	4
$f(x)$	3	5	10	10
$\widehat{F}(x)$	3.1	5.7	8.3	10.9
$(f - \widehat{F})(x)$	-0.1	-0.7	1.7	-0.9
$(f - \widehat{F})^2(x)$	.01	.49	2.89	0.81

de donde, $E = \sqrt{4{,}20} = 2{,}05$.

El valor de E depende de las componentes del vector f, pudiendo obtenerse una aproximación $\widehat{F}$ bastante buena con un valor no necesariamente pequeño para E. De ahí la existencia de otros criterios o formas de medir el error de la aproximación mínimo cuadrática, que en ciertos casos resultan más convenientes. Por ejemplo,

- suma de cuadrados de los errores: $E^2 = \sum_{i=o}^{N} (f_i - \widehat{F}_i)^2$
- desviación cuadrática media: $E/(N+1)$
- varianza: E/N
- desviación típica: $\sqrt{varianza}$
- error relativo: $E/\left\|f\right\|_2$
- otras estadísticas (la mayoría, basadas en E)

Apéndice

Definiciones de norma más usadas

a)Caso continuo : $g \in C_{[a,b]}$

$$\left\|g\right\|_2 = \sqrt{\int_a^b g(x)^2 dx} : \quad \text{norma euclideana}$$
$$\left\|g\right\|_\infty = \max_{x \in [a,b]} |g(x)| : \quad \text{norma de Chebyshev}$$

Las dos normas son caso especial de la norma en L_p :

$$\|g\|_p = \left(\int_a^b |g(x)|^p \, dx \right)^{1/p}.$$

b)Caso discreto, para funciones definidas en una malla o retícula $M = \{x_i\}_{i=1}^m$ constituida por un conjunto finito de puntos .

La correspondiente norma se define como

$$\|g\|_{p,M} = \left(\sum_{i=1}^m |g(x_i|^p \right)^{1/p}.$$

Se dice que esta norma es realmente una *seminorma* si g es continua, ya que en ese caso no se satisface el primero de los requerimientos de la definición para la función g, que puede ser cero en el conjunto M sin ser idénticamente nula.

c)Con función de peso:

Las definiciones de norma pueden generalizarse introduciendo una cierta función positiva $\omega(x)$ para $a < x < b$,llamada función de peso (weight), que en el caso discreto sería un vector $\omega = (\omega(x_1), ..., \omega(x_m))$. Se obtienen entonces las expresiones:

$$\|g\|_{p,\omega} = \left(\int_a^b \omega(x) \, |g(x)|^p \, dx \right)^{1/p}$$

$$\|g\|_{p,\omega,M} = \left(\sum_{i=1}^m \omega(x_i) \, |g(x_i)|^p \right)^{1/p}$$

d) Producto escalar

El producto escalar $< g, h >$ de g y h pertenecientes a L se define como

$$< g, h >= \int_a^b \omega(x) g(x) h(x) dx, \text{ en el caso continuo}$$

$$< g, h >= \sum_{i=1}^m \omega(x_i) g(x_i) h(x_i), \text{ en el caso discreto.}$$

Para la norma euclidiana se tiene entonces, que

$$\|g\|_2 = \sqrt{< g, g >}.$$

El producto escalar es un número real, y tiene las propiedades siguientes:

$$< g, g >\geq 0 \ \ \forall g, \text{ y } \ < g, g >= 0 \Longrightarrow g = 0$$

$$< g, h >=< h, g > \ : \text{conmutativa}$$

$$< \alpha g, h >= \alpha < g, h > \ : \text{asociativa con respecto a multiplicación por un escalar}$$

$$< g + f, h >=< g, h > + < f, h > \ : \text{distributiva}$$

La introducción del concepto de norma permite generalizar la noción de distancia entre dos elementos de un espacio. El concepto de producto escalar permite, además, hacer la extensión de otras nociones geométricas tales como ángulos y ortogonalidad. Por ejemplo, g y h ϵL se dice que son ortogonales si se cumple que $< g, h >= 0$.

Algunos resultados teóricos generales

Definición 11 *Sea M un espacio métrico, $f : M \to \mathbb{R}$, diremos que f es continua inferiormente en un punto $x^* \in M$ si para toda sucesión $\{x_n\} \subset M$ que converge a x^* se cumple*

$$f(x^*) \leq \lim \inf_{n \to \infty} f(x_n).$$

Definición 12 *Sea M un espacio de Banach, $\varphi \in (M, \mathbb{R})$ un funcional, $B \subseteq M$ convexo. Se dice que φ es fuertemente convexa en B si: $\frac{1}{2}\varphi(x) + \frac{1}{2}\varphi(y) - \varphi(\frac{x+y}{2}) \geq \gamma \|x - y\|^2$, para $x, y \in B$, $\gamma > 0$.*

Teorema 13 *Sea M un espacio de Banach, $\varphi \in (M, \mathbb{R})$ un funcional continuo inferiormente, acotado inferiormente y fuertemente convexo sobre el conjunto cerrado (convexo) $B \subseteq M$. Entonces φ alcanza su valor mínimo en $u \in B$ determinado de forma única.*

Teorema 14 *Sea H un espacio de Hilbert, $B \subseteq H$ acotado, convexo y cerrado. Entonces todo funcional $f \in H^*$(con H^*se denota el espacio dual de H que es el conjunto de las aplicaciones lineales y continuas de $H \to \mathbb{R}$ y se denota también por $\mathcal{L}(H, \mathbb{R})$) alcanza su mínimo en B.*