Análisis Numérico

Primer cuatrimestre 2023

Cuadrados mínimos

Outline

1. Aproximación en un espacio con producto interno
2. Ejemplo 1: aproximación de funciones continuas por polinomios
3. Ejemplo 2: ajuste en L^2 por funciones de forma dadas
4. Data–fitting

Outline

1. Aproximación en un espacio con producto interno
2. Ejemplo 1: aproximación de funciones continuas por polinomios
3. Ejemplo 2: ajuste en L^2 por funciones de forma dadas
4. Data–fitting

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Algunos ejemplos:

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Algunos ejemplos:

- $V = \mathbb{R}^n$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Algunos ejemplos:

- $V = \mathbb{R}^n$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$
- $V = \mathbb{R}^n$, $\langle x, y \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots + w_n x_n y_n$ (con $w_i > 0 \, \forall i$)

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Algunos ejemplos:

- $V = \mathbb{R}^n$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$
- $V = \mathbb{R}^n$, $\langle x, y \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots + w_n x_n y_n$ (con $w_i > 0 \, \forall i$)
- $V = \mathcal{C}([a, b])$, $\langle f, g \rangle = \int_a^b f(x) g(x) \, dx$

Producto interno

Sea V un espacio vectorial. Recordemos que

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

es un *producto interno* en V si para todo $x, y, z \in V$, $\alpha \in \mathbb{R}$ tenemos

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ si $x \neq 0$

Algunos ejemplos:

- $V = \mathbb{R}^n$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$
- $V = \mathbb{R}^n$, $\langle x, y \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots + w_n x_n y_n$ (con $w_i > 0 \, \forall i$)
- $V = \mathcal{C}([a, b])$, $\langle f, g \rangle = \int_a^b f(x) g(x) \, dx$
- $V = \mathcal{C}([a, b])$, $\langle f, g \rangle = \int_a^b w(x) f(x) g(x) \, dx$ si $w(x) > 0$ en $[a, b]$

Producto interno

Desigualdad de Cauchy–Schwartz

Si $x, y \in V$ entonces

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{\frac{1}{2}} \langle y, y\rangle^{\frac{1}{2}}$$

Producto interno

Desigualdad de Cauchy–Schwartz

Si $x, y \in V$ entonces

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{\frac{1}{2}} \langle y, y\rangle^{\frac{1}{2}}$$

Norma

La función

$$\|\cdot\| : V \to \mathbb{R}, \qquad \|x\| = \langle x, x\rangle^{\frac{1}{2}}$$

es una *norma* sobre V. La llamamos *norma inducida por el producto interno* en V

Ortogonalidad

- Si $\langle x, y\rangle = 0$ decimos que x e y son *ortogonales* y escribimos $x \perp y$
- $A, B \subset V$ son ortogonales $(A \perp B)$ si $x \perp y$ para todo $x \in A, y \in B$

Problemas de aproximación

Proyección ortogonal

Dado un subespacio $\mathcal{S}$ del espacio con producto interno V, sean $x \in V, y \in S$. Entonces son equivalentes

a. $\|x - y\| = \min\{\|x - s\| : s \in \mathcal{S}\}$
b. $\langle x - y, s \rangle = 0$ para todo $s \in \mathcal{S}$

Proyección ortogonal

Dado un subespacio $\mathcal{S}$ del espacio con producto interno V, sean $x \in V, y \in S$. Entonces son equivalentes

a. $\|x - y\| = \min\{\|x - s\| : s \in \mathcal{S}\}$
b. $\langle x - y, s\rangle = 0$ para todo $s \in \mathcal{S}$

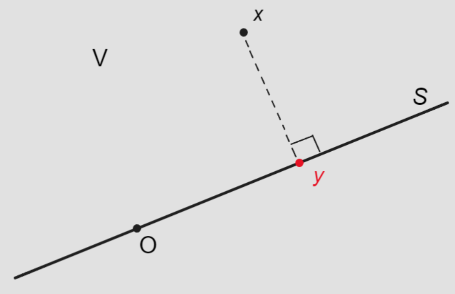
Si existe $y \in \mathcal{S}$ verificando (a) o (b) entonces es único

Proyección ortogonal

Dado un subespacio $\mathcal{S}$ del espacio con producto interno V, sean $x \in V, y \in S$. Entonces son equivalentes

- (a) $\|x - y\| = \min\{\|x - s\| : s \in \mathcal{S}\}$
- (b) $\langle x - y, s\rangle = 0$ para todo $s \in \mathcal{S}$

Si existe $y \in \mathcal{S}$ verificando (a) o (b) entonces es único



Proyección ortogonal

Si $y \in V$ verifica (a) o (b), decimos que y es la proyección ortogonal de x sobre $\mathcal{S}$ en V y anotamos $y = P_{\mathcal{S}}x$

Proyección ortogonal

Si $y \in V$ verifica (a) o (b), decimos que y es la proyección ortogonal de x sobre $\mathcal{S}$ en V y anotamos $y = P_{\mathcal{S}}x$

En lo que sigue veremos que si el subespacio $\mathcal{S}$ es de dimensión finita siempre existe la proyección ortogonal sobre $\mathcal{S}$ para cada x de V, y cómo construirla.

Proyección ortogonal

Si $y \in V$ verifica (a) o (b), decimos que y es la proyección ortogonal de x sobre $\mathcal{S}$ en V y anotamos $y = P_{\mathcal{S}}x$

En lo que sigue veremos que si el subespacio $\mathcal{S}$ es de dimensión finita siempre existe la proyección ortogonal sobre $\mathcal{S}$ para cada x de V, y cómo construirla.

Un subconjunto A de un espacio vectorial con producto interno V se dice *ortonormal* si

- $\|x\| = 1$ para todo $x \in A$, y
- $x \perp y$ para todo par $x, y \in A$ con $x \neq y$

Proyección ortogonal

Existencia de la proyección ortogonal

Sea $\mathcal{S}$ un subespacio de dimensión finita del e.v.c.p.i. V. Entonces $P_{\mathcal{S}}x$ existe y es única para cada $x \in V$. Además si $\{v_1, v_2, \ldots, v_m\}$ es una base ortonormal de $\mathcal{S}$ entonces

$$P_{\mathcal{S}}x = \sum_{i=1}^{m} \langle x, v_i \rangle v_i$$

Proyección ortogonal

Existencia de la proyección ortogonal

Sea $\mathcal{S}$ un subespacio de dimensión finita del e.v.c.p.i. V. Entonces $P_{\mathcal{S}}x$ existe y es única para cada $x \in V$. Además si $\{v_1, v_2, \ldots, v_m\}$ es una base ortonormal de $\mathcal{S}$ entonces

$$P_{\mathcal{S}}x = \sum_{i=1}^{m} \langle x, v_i \rangle v_i$$

De esta manera, podemos hallar $P_{\mathcal{S}}x \in \mathcal{S}$ que minimiza

$$\|x - P_{\mathcal{S}}x\| = \min \{\|x - s\| : s \in \mathcal{S}\}$$

si contamos con una base ortonormal de $\mathcal{S}$. Para ello podemos usar el método de Gram–Schmidt

Proyección ortogonal

Método de Gram–Schmidt

Sea $\{r_1, r_2, \dots, r_m\}$ una base $\mathcal{S}$. Entonces definimos

$$u_1 = r_1, \qquad v_1 = \frac{u_1}{\|u_1\|}$$

y para $k = 2, \dots, m$

$$u_k = r_k - \sum_{i=1}^{k-1} \langle r_k, v_i \rangle v_i, \qquad v_k = \frac{u_k}{\|u_k\|}$$

Entonces

- El conjunto $\{u_1, \dots, u_m\}$ es ortogonal
- $\{v_1, \dots, v_m\}$ es una base ortonormal de $\mathcal{S}$.

Outline

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

Problema

Dada $f \in \mathcal{C}([a,b])$ y $m \in \mathbb{N}$ hallar un polinomio p en $\mathcal{P}_m$ que minimice

$$\int_a^b w(x)\,[f(x) - p(x)]^2 \; dx$$

donde $w(x)$ es una función de peso positiva en $[a,b]$.

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

Problema

Dada $f \in \mathcal{C}([a,b])$ y $m \in \mathbb{N}$ hallar un polinomio p en $\mathcal{P}_m$ que minimice

$$\int_a^b w(x)\,[f(x) - p(x)]^2\ dx$$

donde $w(x)$ es una función de peso positiva en $[a,b]$.

Si definimos el producto interno en $V = \mathcal{C}([a,b])$ por

$$\langle g, h\rangle_w = \int_a^b w(x)g(x)h(x)\,dx$$

y $\|\cdot\|_w$ es la norma inducida, entonces podemos reescribir el problema como: hallar $p_m^* \in \mathcal{P}_m$ tal que

$$\|f - p_m^*\|_w = \min\{\|f - p\|_w : p \in \mathcal{P}_m\}$$

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

Para hallar p_m^* basta encontrar una base ortonormal $\{p_0, p_1, \ldots, p_m\}$ del subespacio $\mathcal{S} = \mathcal{P}_m$ de $V = \mathcal{C}([a,b])$ (notar que $m+1 = \dim \mathcal{P}_m$), y definir

$$p_m^* = \sum_{i=0}^{m} \langle f, p_i \rangle_w p_i$$

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

Para hallar p_m^* basta encontrar una base ortonormal $\{p_0, p_1, \ldots, p_m\}$ del subespacio $\mathcal{S} = \mathcal{P}_m$ de $V = \mathcal{C}([a,b])$ (notar que $m+1 = \dim \mathcal{P}_m$), y definir

$$p_m^* = \sum_{i=0}^{m} \langle f, p_i \rangle_w p_i$$

Para esto podemos aplicar GS a la base $\left\{1, x, x^2, \ldots, x^m\right\}$ de $\mathcal{P}_m$

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

Para hallar p_m^* basta encontrar una base ortonormal $\{p_0, p_1, \dots, p_m\}$ del subespacio $\mathcal{S} = \mathcal{P}_m$ de $V = \mathcal{C}([a,b])$ (notar que $m+1 = \dim \mathcal{P}_m$), y definir

$$p_m^* = \sum_{i=0}^{m} \langle f, p_i \rangle_w p_i$$

Para esto podemos aplicar GS a la base $\left\{1, x, x^2, \dots, x^m\right\}$ de $\mathcal{P}_m$

GS para polinomios

Aplicamos GS a la base $\left\{1, x, x^2, \dots, x^m\right\}$ de $\mathcal{P}_m$:

$$q_0 = 1, \qquad p_0 = 1/\|q_0\|_w,$$

y para $k = 1, \dots, m$

$$q_k(x) = x^k - \sum_{i=0}^{k-1} \langle x^k, p_i \rangle_w p_i, \qquad p_k = q_k/\|q_k\|_w$$

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a,b])$

GS aplicado a la base canónica de polinomios

Los polinomios q_k verifican las siguientes propiedades:

- q_k es un polinomio mónico de grado k
- $\{q_0, q_1, \ldots, q_m\}$ es una base ortogonal de $\mathcal{P}_m$
- $q_k \perp \mathcal{P}_{k-1}$ (q_k es ortogonal a todo polinomio de grado $< k$)

Ejemplo 1: aproximación polinomial en $\mathcal{C}([a, b])$

GS aplicado a la base canónica de polinomios

Los polinomios q_k verifican las siguientes propiedades:

- q_k es un polinomio mónico de grado k
- $\{q_0, q_1, \ldots, q_m\}$ es una base ortogonal de $\mathcal{P}_m$
- $q_k \perp \mathcal{P}_{k-1}$ (q_k es ortogonal a todo polinomio de grado $< k$)

Como consecuencia del teorema de Weierstrass obtenemos fácilmente el siguiente resultado

Convergencia de los polinomios minimizantes

La sucesión de polinomios minimizantes p_m^* converge a f en la norma $\|\cdot\|_w$, es decir

$$\|f - p_m^*\|_w \to 0, \qquad m \to \infty$$

Outline

Ejemplo 2: ajuste en L^2 por funciones de forma dadas

Problema

Hallar (numéricamente) la mejor aproximación en norma $L^2([1,2])$ de la función $f(x) = x^{-1}$ por una combinación lineal de las funciones $e^x, \sin x$ y $\Gamma(x)$.

Ejemplo 2: ajuste en L^2 por funciones de forma dadas

Problema

Hallar (numéricamente) la mejor aproximación en norma $L^2([1,2])$ de la función $f(x) = x^{-1}$ por una combinación lineal de las funciones $e^x, \sin x$ y $\Gamma(x)$.

Podemos utilizar el comando `quad` para calcular numéricamente las integrales involucradas en GS

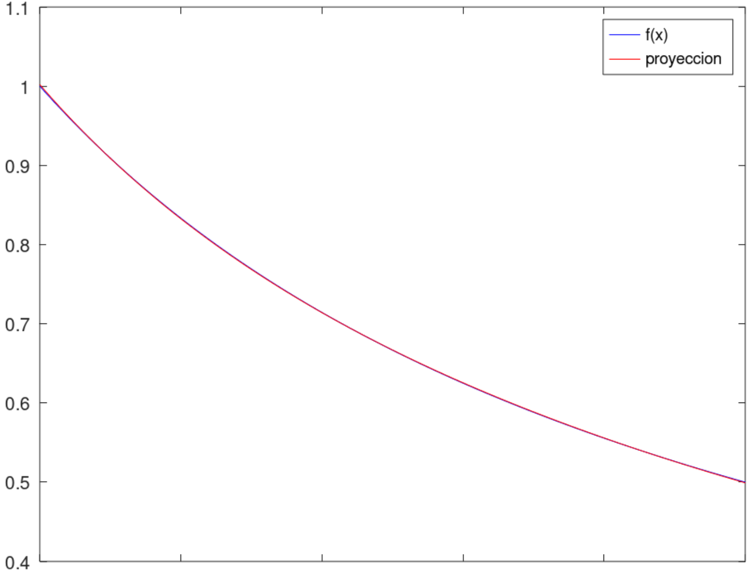
Ejemplo 2: ajuste en L^2 por funciones de forma dadas

```
# calcula la proyeccion L2([a,b]) de f sobre el subespacio
# generado por las funciones definidas en la estruc fcs
f = @(x) 1./x;
a = 1;
b = 2;
fcs = {@exp, @sin , @gamma};
# ortonormalizamos fcs
bon = {};  # base ortonormal
for i = 1:length(fcs)
  gi = @(x) fcs{i}(x);
  for k = 1:(i-1)
    integ = @(x) gi(x).*bon{k}(x);
    c = quad(integ,a,b);
    gi = @(x) gi(x) - c*bon{k}(x);
  endfor
  integ = @(x) gi(x).^2;
  n = sqrt(quad(integ,a,b));
  bon{i} = @(x) gi(x)/n;
endfor

```

Ejemplo 2: ajuste en L^2 por funciones de forma dadas

```
# proyectamos f sobre el espacio <fcs{i}: i=1:length(fcs)>
p = @(x) 0;
for i = 1:length(fcs)
  integ = @(x) f(x)*bon{i}(x);
  c = quad(integ,a,b);
  p = @(x) p(x) + c*bon{i}(x);
endfor
# graficamos f y p en [a,b]
N = 100;
x = a + [0:N]/N*(b-a);
figure
plot(x,f(x),'b');
hold on
plot(x,p(x),'r');
hold off

```

Ejemplo 2: ajuste en L^2 por funciones de forma dadas



Outline

Sistemas sobredeterminados

Problema de cuadrados mínimos

Dada $A \in \mathbb{R}^{m \times n}$ con $m > n$ y $b \in \mathbb{R}^m$, hallar $x \in \mathbb{R}^n$ tal que

$$\|b - Ax\|_2$$

sea mínimo.

Sistemas sobredeterminados

Problema de cuadrados mínimos

Dada $A \in \mathbb{R}^{m \times n}$ con $m > n$ y $b \in \mathbb{R}^m$, hallar $x \in \mathbb{R}^n$ tal que

$$\|b - Ax\|_2$$

sea mínimo.

- Como $m > n$, el sistema $Ax = b$ es sobredeterminado, y en general no tiene solución

Sistemas sobredeterminados

Problema de cuadrados mínimos

Dada $A \in \mathbb{R}^{m \times n}$ con $m > n$ y $b \in \mathbb{R}^m$, hallar $x \in \mathbb{R}^n$ tal que

$$\|b - Ax\|_2$$

sea mínimo.

- Como $m > n$, el sistema $Ax = b$ es sobredeterminado, y en general no tiene solución
- $r = b - Ax$ se llama *residuo*, y es lo que se quiere minimizar

Sistemas sobredeterminados

Problema de cuadrados mínimos

Dada $A \in \mathbb{R}^{m \times n}$ con $m > n$ y $b \in \mathbb{R}^m$, hallar $x \in \mathbb{R}^n$ tal que

$$\|b - Ax\|_2$$

sea mínimo.

- Como $m > n$, el sistema $Ax = b$ es sobredeterminado, y en general no tiene solución
- $r = b - Ax$ se llama *residuo*, y es lo que se quiere minimizar
- $\|\cdot\|_2$ es la norma euclídea, pero podría considerarse otras normas inducidas por un producto interno en $\mathbb{R}^m$, por ejemplo, norma ℓ_2 con pesos

Sistemas sobredeterminados

Problema de cuadrados mínimos

Dada $A \in \mathbb{R}^{m \times n}$ con $m > n$ y $b \in \mathbb{R}^m$, hallar $x \in \mathbb{R}^n$ tal que

$$\|b - Ax\|_2$$

sea mínimo.

- Como $m > n$, el sistema $Ax = b$ es sobredeterminado, y en general no tiene solución
- $r = b - Ax$ se llama *residuo*, y es lo que se quiere minimizar
- $\| \cdot \|_2$ es la norma euclídea, pero podría considerarse otras normas inducidas por un producto interno en $\mathbb{R}^m$, por ejemplo, norma ℓ_2 con pesos
- El problema puede plantearse en $\mathbb{C}$ casi sin cambios, solo cambiar ^{t} (traspuesta) por * (conjugada traspuesta)

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \text{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \text{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$
- Sabemos que $y \in \mathcal{S}$ resuelve (1) si y solo si
$$\langle b - y, s\rangle = 0 \qquad \forall s \in \mathcal{S}$$

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \mathrm{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$
- Sabemos que $y \in \mathcal{S}$ resuelve (1) si y solo si
$$\langle b - y, s\rangle = 0 \qquad \forall s \in \mathcal{S}$$
- Así podemos plantear: Hallar $x \in \mathbb{R}^n$ tal que
$$\langle b - Ax, Az\rangle = 0 \qquad \forall z \in \mathbb{R}^n$$

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \text{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$
- Sabemos que $y \in \mathcal{S}$ resuelve (1) si y solo si
$$\langle b - y, s\rangle = 0 \qquad \forall s \in \mathcal{S}$$
- Así podemos plantear: Hallar $x \in \mathbb{R}^n$ tal que
$$\langle b - Ax, Az\rangle = 0 \qquad \forall z \in \mathbb{R}^n \qquad \text{ó}, \quad r \perp \text{Rg}(A)$$

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \text{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$
- Sabemos que $y \in \mathcal{S}$ resuelve (1) si y solo si
$$\langle b - y, s\rangle = 0 \qquad \forall s \in \mathcal{S}$$
- Así podemos plantear: Hallar $x \in \mathbb{R}^n$ tal que
$$\langle b - Ax, Az\rangle = 0 \qquad \forall z \in \mathbb{R}^n \qquad \text{ó}, \quad r \perp \text{Rg}(A)$$
- Como $\langle b - Ax, Az\rangle = \langle A^t b - A^t Ax, z\rangle$, $x \in \mathbb{R}^n$ debe verificar
$$\langle A^t b - A^t Ax, z\rangle = 0 \qquad \forall z \in \mathbb{R}^n \tag{2}$$

Sistemas sobredeterminados

- Si definimos $\mathcal{S} = \mathrm{Rg}(A) = \{Ax : x \in \mathbb{R}^n\}$, entonces podemos resolver el problema en dos pasos:
  1. Encontrar $y \in \mathcal{S}$ tal que
  $$\|b - y\|_2 = \min\{b - s : s \in \mathcal{S}\} \tag{1}$$
  2. Hallar x tal que $Ax = y$
- Sabemos que $y \in \mathcal{S}$ resuelve (1) si y solo si
$$\langle b - y, s\rangle = 0 \qquad \forall s \in \mathcal{S}$$
- Así podemos plantear: Hallar $x \in \mathbb{R}^n$ tal que
$$\langle b - Ax, Az\rangle = 0 \qquad \forall z \in \mathbb{R}^n \qquad \text{ó}, \quad r \perp \mathrm{Rg}(A)$$
- Como $\langle b - Ax, Az\rangle = \langle A^t b - A^t Ax, z\rangle$, $x \in \mathbb{R}^n$ debe verificar
$$\langle A^t b - A^t Ax, z\rangle = 0 \qquad \forall z \in \mathbb{R}^n \tag{2}$$
- (2) es equivalente a
$$A^t Ax = A^t b$$

Sistemas sobredeterminados

Ecuaciones normales

$x \in \mathbb{R}^n$ es solución de

$$\|b - Ax\|_2 = \min\{\|b - Az\|_2 : z \in \mathbb{R}^n\}$$

si y solo si

$$A^t Ax = A^t b \tag{3}$$

Sistemas sobredeterminados

Ecuaciones normales

$x \in \mathbb{R}^n$ es solución de

$$\|b - Ax\|_2 = \min\{\|b - Az\|_2 : z \in \mathbb{R}^n\}$$

si y solo si

$$A^t A x = A^t b \tag{3}$$

- Las ecuaciones (3) se llaman *ecuaciones normales*

Sistemas sobredeterminados

Ecuaciones normales

$x \in \mathbb{R}^n$ es solución de

$$\|b - Ax\|_2 = \min\{\|b - Az\|_2 : z \in \mathbb{R}^n\}$$

si y solo si

$$A^t A x = A^t b \tag{3}$$

- Las ecuaciones (3) se llaman *ecuaciones normales*
- $A^t A$ es no singular si y solo si A es full-rank

Sistemas sobredeterminados

Ecuaciones normales

$x \in \mathbb{R}^n$ es solución de

$$\|b - Ax\|_2 = \min\{\|b - Az\|_2 : z \in \mathbb{R}^n\}$$

si y solo si

$$A^t Ax = A^t b \tag{3}$$

- Las ecuaciones (3) se llaman *ecuaciones normales*
- $A^t A$ es no singular si y solo si A es full-rank
- La solución x es única si y solo si A es full-rank

Pseudoinversa

Pseudoinversa

Si $A \in \mathbb{R}^{m \times n}$ es full-rank, la matriz

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

se llama pseudoinversa de A

Pseudoinversa

Pseudoinversa

Si $A \in \mathbb{R}^{m\times n}$ es full-rank, la matriz

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

se llama pseudoinversa de A

- Vimos que si A es full-rank, la solución (única) $x \in \mathbb{R}^n$ del problema de cuadrados mínimos es

$$x = A^{+} b$$

Tres algoritmos para resolver cuadrados mínimos

Cuadrados mínimos vía ecuaciones normales

Como A^tA es simétrica definida positiva (si A es full-rank), tiene una factorización de Cholesky R^tR con R tringular superior (usamos R^t en lugar de L para usar la notación de Octave)

Cuadrados mínimos vía ecuaciones normales

Como A^tA es simétrica definida positiva (si A es full-rank), tiene una factorización de Cholesky R^tR con R tringular superior (usamos R^t en lugar de L para usar la notación de Octave)

Algoritmo

1. Formar la matriz A^tA y el vector A^tb
2. Calcular la factorización de Cholesky $A^tA = R^tR$
3. Resolver el sistema triangular inferior $R^tw = A^tb$ para w
4. Resolver el sistema triangular superior $Rx = w$ para x

Cuadrados mínimos vía ecuaciones normales

Como A^tA es simétrica definida positiva (si A es full-rank), tiene una factorización de Cholesky R^tR con R tringular superior (usamos R^t en lugar de L para usar la notación de Octave)

Algoritmo

1. Formar la matriz A^tA y el vector A^tb
2. Calcular la factorización de Cholesky $A^tA = R^tR$
3. Resolver el sistema triangular inferior $R^tw = A^tb$ para w
4. Resolver el sistema triangular superior $Rx = w$ para x

El comando `R = chol(B)` de Octave calcula la factorización de Cholesky $B = RR^t$

Cuadrados mínimos vía factorización QR

Suponemos A full-rank. Si A tiene la factorización QR reducida $A = \widehat{Q}\widehat{R}$ entonces

$$A^t A x = A^t b \iff \widehat{R}^t \widehat{Q}^t \widehat{Q} \widehat{R} x = \widehat{R}^t \widehat{Q}^t b \iff \widehat{R} x = \widehat{Q}^t b$$

Cuadrados mínimos vía factorización QR

Suponemos A full-rank. Si A tiene la factorización QR reducida $A = \widehat{Q}\widehat{R}$ entonces

$$A^t A x = A^t b \quad \Longleftrightarrow \quad \widehat{R}^t \widehat{Q}^t \widehat{Q} \widehat{R} x = \widehat{R}^t \widehat{Q}^t b \quad \Longleftrightarrow \quad \widehat{R} x = \widehat{Q}^t b$$

Algoritmo

1. Calcular la factorización QR reducida $A = \widehat{Q}\widehat{R}$
2. Calcular el vector $\widehat{Q}^t b$
3. Resolver el sistema triangular superior $\widehat{R} x = \widehat{Q}^t b$ para x

Cuadrados mínimos vía factorización QR

Suponemos A full-rank. Si A tiene la factorización QR reducida $A = \widehat{Q}\widehat{R}$ entonces

$$A^t A x = A^t b \quad \Longleftrightarrow \quad \widehat{R}^t \widehat{Q}^t \widehat{Q}\widehat{R} x = \widehat{R}^t \widehat{Q}^t b \quad \Longleftrightarrow \quad \widehat{R} x = \widehat{Q}^t b$$

Algoritmo

1. Calcular la factorización QR reducida $A = \widehat{Q}\widehat{R}$
2. Calcular el vector $\widehat{Q}^t b$
3. Resolver el sistema triangular superior $\widehat{R}x = \widehat{Q}^t b$ para x

- Notar que el vector $y \in \mathbb{R}^m$ de (1), que es la proyección ortogonal de b sobre $\mathrm{Rg}(A) = \mathrm{Rg}(\widehat{Q})$ puede calcularse como $y = \widehat{Q}\widehat{Q}^t b$
- Como $y \in \mathrm{Rg}(A)$, el sistema $Ax = y$, o $\widehat{Q}\widehat{R}x = \widehat{Q}\widehat{Q}^t b$, tiene solución, que puede encontrarse premultiplicando por $\widehat{Q}^t$ y resolviendo el sistema triangular del algoritmo

Cuadrados mínimos vía SVD

Supongamos A full-rank

Cuadrados mínimos vía SVD

Supongamos A full-rank

- Si $A = \widehat{U}\widehat{\Sigma}V^t$ es la descomposición svd reducida de A, entonces $\mathrm{Rg}(A) = \mathrm{Rg}(\widehat{U})$, y por lo tanto, el vector y de (1) es

$$y = \widehat{U}\widehat{U}^t b$$

Tres algoritmos para resolver cuadrados mínimos

Cuadrados mínimos vía SVD

Supongamos A full-rank

- Si $A = \widehat{U}\widehat{\Sigma}V^t$ es la descomposición svd reducida de A, entonces $\mathrm{Rg}(A) = \mathrm{Rg}(\widehat{U})$, y por lo tanto, el vector y de (1) es

$$y = \widehat{U}\widehat{U}^t b$$

- Por lo tanto, basta hallar $x \in \mathbb{R}^n$ tal que

$$\widehat{U}\widehat{\Sigma}V^t = \widehat{U}\widehat{U}^t b \qquad \text{ó} \qquad \widehat{\Sigma}V^t = \widehat{U}^t b$$

Cuadrados mínimos vía SVD

Supongamos A full-rank

- Si $A = \widehat{U}\widehat{\Sigma}V^t$ es la descomposición svd reducida de A, entonces $\mathrm{Rg}(A) = \mathrm{Rg}(\widehat{U})$, y por lo tanto, el vector y de (1) es

$$y = \widehat{U}\widehat{U}^t b$$

- Por lo tanto, basta hallar $x \in \mathbb{R}^n$ tal que

$$\widehat{U}\widehat{\Sigma}V^t = \widehat{U}\widehat{U}^t b \qquad \text{ó} \qquad \widehat{\Sigma}V^t = \widehat{U}^t b$$

Algoritmo

1. Calcular la factorización SVD reducida $A = \widehat{U}\widehat{\Sigma}V^t$
2. Calcular el vector $\widehat{U}^t b$
3. Resolver el sistema diagonal $\widehat{\Sigma}w = \widehat{U}^t b$ para w
4. Definir $x = Vw$

Tres algoritmos para resolver cuadrados mínimos

Algunas observaciones

- La matriz A^tA suele ser mal condicionada, por lo que la resolución de las ecuaciones normales puede tener problemas de estabilidad respecto de los errores de redondeo
- En general el algoritmo "recomendado" es el que usa la factorización QR
- Si las columnas de A son "casi linealmente dependientes", entonces conviene usar el algoritmo basado en la SVD, aunque es el más costoso computacionalmente

Ejemplo: polinomial data-fitting

Problema: Hallar un polinomio p de grado ≤ 7 que minimice el error cuadrático

$$\sum_{i=0}^{10} |p(x_i) - y_i|^2$$

si x_i, y_i están dados por la tabla

x_i	-5	-4	-3	-2	-1	0	1	2	3	4	5
y_i	0	0	0	1	1	1	0	0	0	0	0

Comparar con el polinomio de interpolación en los 11 puntos x_i

Ejemplo: polinomial data-fitting

Problema: Hallar un polinomio p de grado ≤ 7 que minimice el error cuadrático

$$\sum_{i=0}^{10} |p(x_i) - y_i|^2$$

si x_i, y_i están dados por la tabla

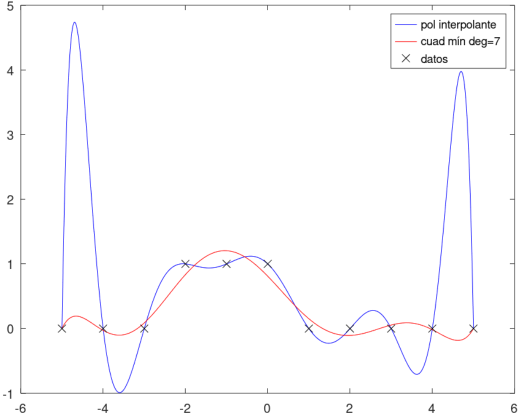
x_i	-5	-4	-3	-2	-1	0	1	2	3	4	5
y_i	0	0	0	1	1	1	0	0	0	0	0

Comparar con el polinomio de interpolación en los 11 puntos x_i

Si denotamos $a = (a_0, a_1, \dots, a_7)^t$ los coeficientes del polinomio $p(x) = \sum_{k=0}^{7} a_k x^k$ buscado, notemos que se trata de hallar el vector a que minimice $\|Aa - b\|_2$, si $A \in \mathbb{R}^{11\times 8}$ y $b \in \mathbb{R}^{11}$ están dados por

$$A = \begin{pmatrix} 1 & x_0 & x_1^2 & \cdots & x_0^7 \\ 1 & x_1 & x_1^2 & \cdots & x_1^7 \\ \vdots & \vdots & \dots & \vdots & \\ 1 & x_{10} & x_1^{10} & \cdots & x_{10}^7 \end{pmatrix} \qquad b = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{10} \end{pmatrix}$$

Ejemplo 3: polinomial data-fitting



El vector a de coeficientes calculado es

$a = (8.0995e-01, -6.5462e-01, -1.3943e-01, 1.3735e-01, 7.5540e-03, -8.8235e-03, -1.3072e-04, 1.7507e-04)$

Ejemplo 4: data-fitting con funciones generales

Problema: Similarmente al Ejemplo 2, hallar una combinación lineal

$$h(x) = a_0 e^x + a_1 \sin(x) + a_2 \Gamma(x)$$

tal que si $x_i = 1 + i/10$, $i = 0, \ldots, 10$ se minimice el error

$$E = \sum_{i=0}^{10} |f(x_i) - h(x_i)|^2$$

para $f(x) = 1/x$. Graficar $h(x)$ y comparar con el resultado del ejemplo 2

Ejemplo 4: data-fitting con funciones generales

Problema: Similarmente al Ejemplo 2, hallar una combinación lineal

$$h(x) = a_0 e^x + a_1 \sin(x) + a_2 \Gamma(x)$$

tal que si $x_i = 1 + i/10$, $i = 0, \ldots, 10$ se minimice el error

$$E = \sum_{i=0}^{10} |f(x_i) - h(x_i)|^2$$

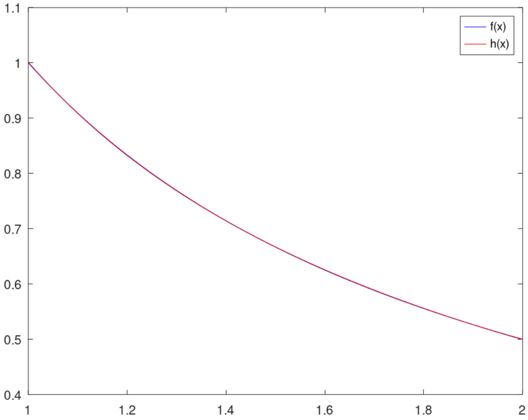
para $f(x) = 1/x$. Graficar $h(x)$ y comparar con el resultado del ejemplo 2

Si definimos la matriz $A \in \mathbb{R}^{11\times 3}$ y $b \in \mathbb{R}^{11}$ por

$$A = \begin{pmatrix} e^{x_0} & \sin(x_0) & \Gamma(x_0) \\ e^{x_1} & \sin(x_1) & \Gamma(x_1) \\ \vdots & \vdots & \vdots \\ e^{x_{10}} & \sin(x_{10}) & \Gamma(x_{10}) \end{pmatrix} \qquad b = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{10}) \end{pmatrix}$$

se trata de hallar $a = (a_0, a_1, a_2)^t$ que minimice $\|Aa - b\|_2$

Ejemplo 4: data-fitting con funciones generales



El vector a de coeficientes calculado es

$$a = (-0.107619, 0.010815, 1.284665)$$

y el error cuadrático obtenido es de $E = 2.2718e - 03$