

Vimos la clase pasada:

Teorema: Dada $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$

son equivalentes:

1) $a \in \mathbb{R}^m$ minimiza $\|Aa - b\|$

2) a es solución de $A^t A a = A^t b$

Además, si las columnas de A son li, la solución a de $A^t A a = A^t b$ existe y es única. Notar $A^t A \in \mathbb{R}^{m \times m}$.

Si m es grande la matriz $A^t A$ puede ser enorme y resolver

$$A^t A x = A^t b \text{ muy costoso.}$$

En este caso hacemos la descomposición QR de $A^t A$.

Descomposición QR:

Dada $A \in \mathbb{R}^{n \times m}$

$$A = (a_1 | \dots | a_m) \text{ con } \{a_1, \dots, a_m\} \text{ li}$$

$$a_i \in \mathbb{R}^n$$

Usando Gram Schmidt obtenemos

$$Q = (q_1 | \dots | q_m) \text{ donde}$$

$$\langle a_1 \rangle = \langle q_1 \rangle$$

$$\langle a_1, a_2 \rangle = \langle q_1, q_2 \rangle$$

:

$$\langle a_1, \dots, a_k \rangle = \langle q_1, \dots, q_k \rangle$$

Esto dice que

$$a_1 = r_{11} q_1 \quad r_{11} \neq 0$$

$$a_2 = r_{12} q_1 + r_{22} q_2 \quad r_{22} \neq 0$$

:

$$a_k = r_{1k} q_1 + r_{2k} q_2 + \dots + r_{kk} q_k$$

$$a_m = \sum_{i=1}^m r_{im} q_i \quad r_{kk} \neq 0$$

Matri ci' almente,

$$(a_1 | \dots | a_m) = (q_1 | \dots | q_m) \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ 0 & r_{22} & \cdots & r_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{mm} \end{pmatrix}$$

$\underbrace{}_{A \in \mathbb{R}^{n \times m}}$
 $\underbrace{}_{Q \in \mathbb{R}^{m \times m}}$
 $\underbrace{\phantom{\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ 0 & r_{22} & \cdots & r_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{mm} \end{pmatrix}}}_{R \in \mathbb{R}^{m \times m}}$

Donde Q es una matriz con columnas que forman un conjunto ortonormal y R es triangular superior

Entonces hacemos lo siguiente

$$A = Q \cdot R$$

$\wedge \quad \wedge \quad \wedge$

$R^{n \times m} \quad R^{m \times m} \quad R^{m \times m}$

Q tiene columnas
ortonormales con
 $Q^t \cdot Q = Id$

γ R es triangular invertible
 $(\text{los } r_{ij} \neq 0)$.

$$\Rightarrow A^t A = R^t Q^t Q R = R^t R$$

$$\Leftrightarrow A^t A a = A^t b$$



$$R^t R a = R^t Q^t b \quad (R^t \text{ invertible})$$



$$R a = Q^t b$$

con R
triangular sup.

Nota: $A^t A$ crucial pero mala para
cómputo. Matrices ortogonales y triangulares
"son las buenas".

Ej de parcial:

En un bosque se ha identificado
una especie de planta exótica.
A lo largo de los últimos años se ha
registrado cómo progresó el nro. de indi-

viduos de dicha especie, se tienen los
datos

ano	$\stackrel{=}{x_1}$	$\stackrel{=}{x_2}$	$\stackrel{=}{x_3}$	$\stackrel{=}{x_4}$	$\stackrel{=}{x_5}$	$\stackrel{=}{x_6}$	$\stackrel{=}{x_7}$
cant	216	219	287	303	439	556	681
indiv	"	"	"	"	"	"	"
	y_1	y_2	y_3	y_4	y_5	y_6	y_7

a) Hallar el poli de grado menor o igual a dos q' mejor ajuste a los datos en el sentido de los cuadrados mínimos.

Buscamos $P(x) = a_0 + a_1 x + a_2 x^2$:

$$\sum_{i=1}^7 \underbrace{(P(x_i) - y_i)^2}_{(a_0 + a_1 x_i + a_2 x_i^2 - y_i)^2} \text{ sea m\'in.}$$

Es decir en

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{pmatrix} \quad y \quad b = \begin{pmatrix} 216 \\ 219 \\ 287 \\ 303 \\ 439 \\ 556 \\ 681 \end{pmatrix}$$

$$a_0, a_1, a_2: \| A \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} - b \|_2^2 \text{ sea m\'in.}$$

y vimos que como las columnas de A son \vec{b}_i (chequear), encontrar $\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ que haga $\|A\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} - \vec{b}\|_2^2$ sea mím equivalente a encontrar $\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ solución de

$$\underbrace{A^T A}_{3 \times 3} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \underbrace{A^T \vec{b}}_{3 \times 1}$$

Ej: Programar y resolver en Python.

Si después me piden estimar cuántos individuos habrá en 100 años según este modelo hace $P(100)$.

- b) Ajustar una función de la forma $g(x) = d_0 e^{d_1 x}$ aplicando cuadrados mínimos sobre la función transformada convenientemente.

Podríamos plantear la minimización
(en función de d_0 y d_1) de

$$\sum_{i=1}^7 (g(x_i) - y_i)^2 = \sum_{i=1}^7 (d_0 e^{d_1 x_i} - y_i)^2$$

Pero vemos que el parámetro d_1
no es lineal dentro del cuadrado
(como lo era para $P(x) = a_0 + a_1 x + a_2 x^2$)

Entonces no podemos usar la técnica
anterior en donde reescribimos el
problema de minimización equiva-
lente $\|A \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} - b\|_2$.

Para buscar una aproximación
y poder usar los resultados anteriores
"linealizamos" nuestro problema
tomando \ln . (logaritmo neperiano).
No estamos resolviendo el problema

de minimizar $\sum_{i=1}^7 (\ln(b e^{d_i x_i}) - \ln y_i)^2$

pero resolvemos uno similar:

Buscamos b, d , que minimice

$$\sum_{i=1}^7 \left(\ln(b e^{d_i x_i}) - \ln y_i \right)^2$$

Nota: Dados $b > a > 1$, por el teorema del valor medio, existe $c \in (a, b)$:

$$\ln b - \ln a = \underbrace{\frac{1}{c}}_{\ln} (b-a)$$
$$(\ln x)' \Big|_{x=c}$$

$$\text{como } c > 1 \Rightarrow \frac{1}{c} < 1$$

$$\Rightarrow |\ln b - \ln a| < |b-a|$$

$$\Rightarrow (\ln b - \ln a)^2 < (b-a)^2$$

De esta manera, si los y_i 's de los datos son grandes (como en el ejemplo) es de esperar que si usamos el d_0, d_1 encontrados con la linearización para el problema original, el error sea peor.

Volvemos al planteo de la linearización:

Buscamos d_0, d_1 que minimice

$$\sum_{i=1}^7 \left(\ln(d_0 e^{d_1 x_i}) - \ln y_i \right)^2, \text{ es}$$

decir buscamos d_0, d_1 que haga mínimo

$$\sum_{i=1}^7 \left(\underbrace{\ln(d_0)}_{a_0} + \underbrace{d_1}_{a_1} x_i - \underbrace{\ln y_i}_{\tilde{y}_i} \right)^2$$

lo llamamos lo llamamos lo llamamos
 a_0 a_1 \tilde{y}_i

este problema equivale a

$$\min_{a_0, a_1} \| A \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} - b \|_2$$

donde

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} \quad y \quad b = \begin{pmatrix} \ln 216 \\ \ln 219 \\ \ln 287 \\ \ln 303 \\ \ln 439 \\ \ln 556 \\ \ln 681 \end{pmatrix}$$

→ plantemos

$$A^t \cdot A \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = A^t b$$

Una vez que resolvemos y encontramos a_0, a_1 nos volvemos para atrás usando

$$\ln d_0 = a_0 \Rightarrow d_0 = e^{a_0}$$

$$d_1 = a_1$$

c) Para el mismo problema podemos encontrar un polinomio de grados 3, grados 4, etc que mejor aproxime en el sentido de los cuadrados mínimos?

En general, de de una tabla de datos

x	x_1	x_2	...	x_m
y	y_1	y_2	...	y_m

se busca un polinomio de grados menores o igual a m que mejor aproxime a mis datos en el sentido de los cuadrados mínimos.

Buscamos a_0, a_1, \dots, a_m ($m+1$ incógnitas)

de manera que si $P(x) = a_0 + a_1 x + \dots + a_m x^m$

$$\sum_{i=1}^n (P(x_i) - y_i)^2$$

sea mínimo.

Planteamos

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ \vdots & x_2 & x_2^2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^m \end{pmatrix} \in \mathbb{R}^{n \times m}$$

$$y \quad b = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

El problema es equivalente a
buscar un mínimo de

$$\| A \begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix} - b \|_2$$

y vemos que si las columnas de A
son l.i \Rightarrow la única solución viene
dada por $A^t A \begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix} = A^t b$.

La pregunta es: ¿Cuándo puedes garantizar que las columnas son li?

Para empezar como $A \in \mathbb{R}^{n \times m}$, se tienen $m+1$ columnas que son vectores de \mathbb{R}^n .

Si queremos que haya posibilidad de que sean li, debe ser $m+1 \leq n$

Cuando $m+1 = n$

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{m-1} \\ \vdots & x_2 & x_2^2 & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{m-1} \end{pmatrix}$$

se conoce
como

matriz de Vandermonde y se nota $V[x_1, x_2, \dots, x_m]$.

Veamos que en este caso cuando $V[x_1, x_2, \dots, x_m]$ es una matriz inversible. Para eso calculamos su determinante por

Inducción.

$$\det V[x_1, x_2] = \det \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} = x_2 - x_1$$

Vamos a usar que hacer operaciones de columna $C_i + \alpha C_j \rightarrow C_i$ no cambia el determinante de una matriz y que el determinante de multiplicar una columna por un escalar es igual a ese escalar por el determinante.

Es decir, vale que:

- $\det (C_1 | C_2 | \dots | C_i + \alpha C_j | \dots | C_n) = \det (C_1 | C_2 | \dots | C_i | \dots | C_n)$
- $\det (C_1 | \alpha C_2 | \dots | C_n) = \det (C_1 | \dots | C_2 | \dots | C_n)$

Y lo mismo por filas.

$$\det V[x_1, x_2, x_3] = \det \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1 x_2 \\ 1 & x_3 - x_1 & x_3^2 - x_1 x_3 \end{pmatrix}$$

$$C_3 - x_1 C_2 \rightarrow C_3$$

$$C_2 - x_1 C_1 \rightarrow C_2$$

$$= \det \begin{pmatrix} x_2 - x_1 & x_2^2 - x_1 x_2 \\ x_3 - x_1 & x_3^2 - x_1 x_3 \end{pmatrix}$$

$$= \det \begin{pmatrix} \cancel{x_2 - x_1} & (x_2 - x_1) x_2 \\ x_3 - x_1 & (x_3 - x_1) x_3 \end{pmatrix}$$

$$= (x_2 - x_1) (x_3 - x_1) \cdot \det \begin{pmatrix} 1 & x_2 \\ 1 & x_3 \end{pmatrix}$$

$$= (x_2 - x_1) (x_3 - x_1) (x_3 - x_2)$$

Usando la misma idea para el caso general se prueba que

$$\det V[x_1 \dots x_n] = \prod_{m > j > i \geq 1} (x_j - x_i)$$

Con lo cual se demuestra que esta matriz es invertible si y solo si $x_i \neq x_j \forall i \neq j$. Es decir los x_i 's de la fórmula son todos distintos.

Volvamos al caso general con $m \leq n-1$.

Vemos que $n-m = m+1$ y $x_i \neq x_j \forall i \neq j$ entonces $V[x_1 \dots x_m]$ es invertible \Rightarrow el rango de la matriz es m , es decir, sus columnas son li.

Con lo cual se demuestra que con un subconjunto de $m < n-1$ columnas también serán li.

En conclusión, desde una fórmula de m datos con $x_i \neq x_j \forall i \neq j$, existe un único polinomio de grado menor o igual a m ($\leq m-1$)

que mejor aproxima a la tabla en el sentido de los cuadrados mínimos.

OBS:

- Si los x_i están muy cerca entre sí produce oscilaciones muy altas
- Si m es muy grande cercano a $n-1$ también produce oscilaciones especialmente cerca de los extremos.
(Fenómeno de Runge).

Polinomio interpolador: Con $x_i \neq x_j \forall i \neq j$

Para $m = n-1$ se sabe que existe un único polinomio de grado menor o igual a $n-1$ que satisface

$$P(x_i) = y_i \quad i=1, \dots, n.$$

Eso decir, este polinomio pasa

por cada punto (x_i, y_i) .

Se lo llama polinomio interpolador de Lagrange.

Podemos encontrarlo como antes planteando $A = V[x_1 \dots x_n]$ y resolviendo

$$A^t A \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Si no, podemos usar la base de Lagrange:

Para cada x_j con $j=1 \dots n$,

definimos $l_j(x)$ el polinomio monico que satisface:

$$l_j(x_i) = 0 \quad \text{si } i \neq j \quad \begin{matrix} \text{Usamos} \\ x_i \neq x_j \end{matrix}$$

$$l_j(x_j) = 1,$$

$$\text{es decir } l_j(x) = \frac{\prod_{i \neq j} (x - x_i)}{\prod_{i \neq j} (x_j - x_i)}$$

$$\Rightarrow P(x) = \sum_{j=1}^n y_j \ell_j(x)$$

Veamos que P es único. Supongamos
existe otro q de grado menor o igual a

$$m-1 : q(x_j) = y_j \quad \forall j=1 \dots n$$

$$\Rightarrow (P - q)(x_j) = 0 \quad \forall j=1 \dots n \quad \textcircled{*}$$

y $P - q$ tiene grado menor o igual
a $m-1 \Rightarrow$ tiene a lo sumo $m-1$
raíces, lo cual se contradice con $\textcircled{*}$.

Error de Interpolación: Sean $f \in C^m[a,b]$
y P el poli interpolador de grado $\leq m-1$
en los puntos x_1, \dots, x_m . Para cada $x \in [a,b]$
existe $\xi = \xi(x) \in [a,b]$:

$$E(x) = f(x) - P(x) = \frac{f^{(m)}(\xi)}{m!} \prod_{i=1}^m (x-x_i)$$