Cálculo Numérico

Cuadrados Mínimos

Nazareno Faillace

25/06

Departamento de Matemática - FCEyN - UBA

Problema A

En general, podemos encontrar dos tipos de problemas en Cuadrados Mínimos. Empecemos con el primero:

- Dados $w_0, \dots, w_n$ constantes positivas, $\phi_1, \dots, \phi_m$ funciones conocidas y un conjunto de datos $\{(x_i, y_i)\}_{i=0}^{n}$, queremos hallar los coeficientes $\{a_i\}_{i=1}^{m}$ que minimicen:

$$\sum_{i=0}^{n} w_i(a_1\phi_1(x_i) + \cdots + a_m\phi_m(x_i) - y_i)^2$$

Por ejemplo, si $\phi_1(x) = 1$, $\phi_2(x) = x$ y $\phi_3(x) = x^2$, entonces estamos buscando el polinomio de grado a lo sumo 2 que mejor aproxima a los datos en el sentido de cuadrados mínimos:

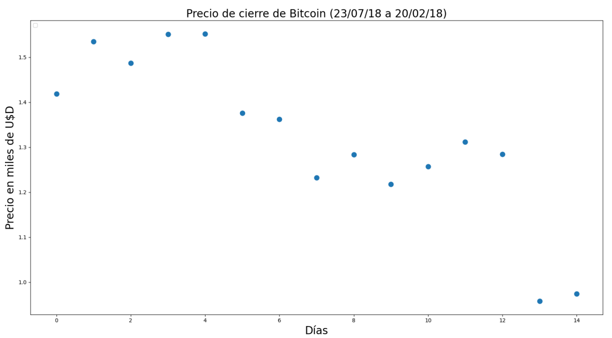
$$\sum_{i=0}^{n} w_i(a_1 + a_2 x_i + a_3 x_i^2 - y_i)^2$$

En la guía de ejercicios, a menos que se indique lo contrario, se puede suponer que $w_i = 1, \ i = 0, \dots, n$

Motivación

Supongamos que tenemos los datos de la cotización de Bitcoin en miles de dólares, desde el 06/02/18 al 20/02/18:

Día (x_i)	0	1	...	14
Valor en miles de U$D ($y_i$)	1,41873	1,53477	...	0.97452



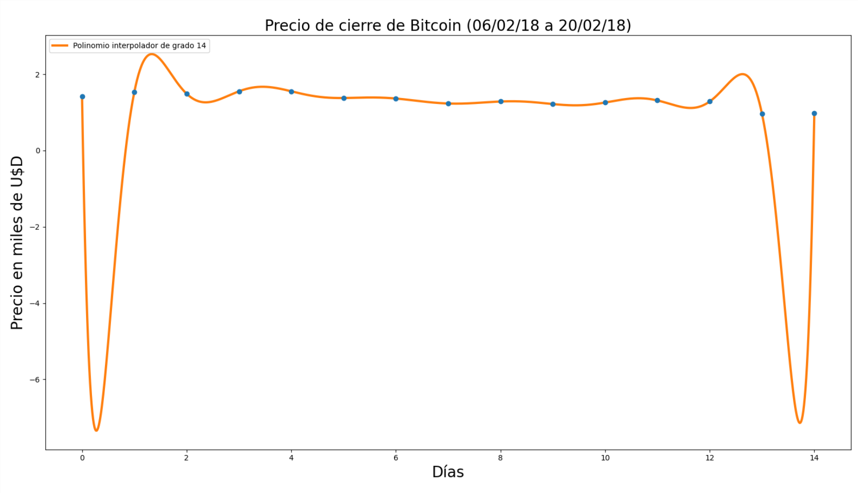
A uno le podría interesar estudiar la evolución del valor de la criptomoneda durante ese periodo de tiempo, por ejemplo, para armar un modelo muy sencillo que intente predecir su comportamiento futuro.

Motivación

¿Usamos el polinomio interpolador?

Motivación

¿Usamos el polinomio interpolador?



Motivación

Claramente no nos sirve: tenemos un polinomio de grado 16 que oscila mucho.
(¡Imagínense si tuviéramos los datos de la cotización de todo el año!)

Motivación

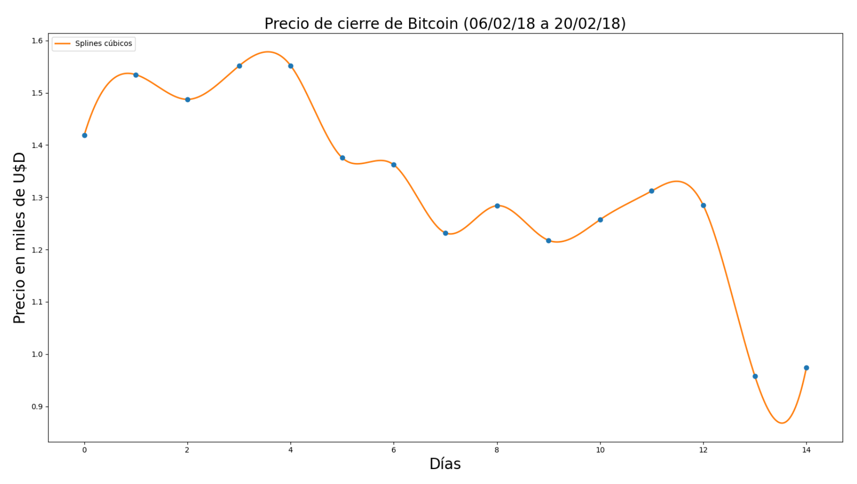
Claramente no nos sirve: tenemos un polinomio de grado 16 que oscila mucho.
(¡Imagínense si tuviéramos los datos de la cotización de todo el año!)

Claramente no nos sirve: tenemos un polinomio de grado 16 que oscila mucho. (¡Imagínense si tuviéramos los datos de la cotización de todo el año!)

Veamos qué pasa

Motivación

Usando splines cúbicos

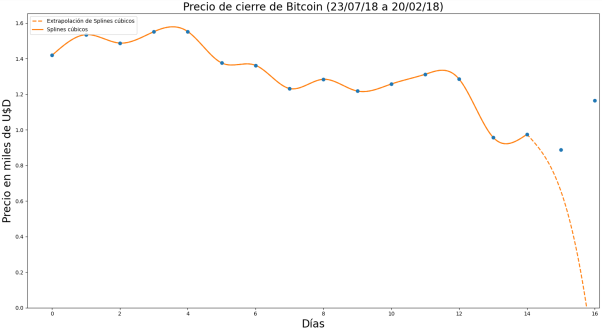


Motivación

Si bien el gráfico es mucho más sensato, esto tampoco nos sirve mucho para nuestro análisis. En general, en este tipo de problemas, interpolar causa lo que se conoce como sobreajuste ("*overfitting*"). Esto significa que, si bien nuestro modelo ajusta perfectamente los datos que tenemos, no tiene capacidad de generalización. En otras palabras, si quisiéramos extrapolar a datos por fuera de $\{(x_i, y_i)\}_{i=0}^{14}$, no nos sería de mucha ayuda:

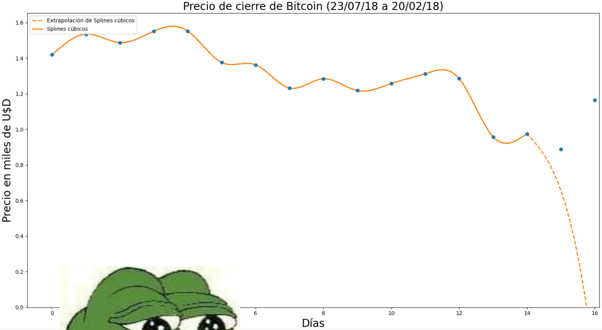
Motivación

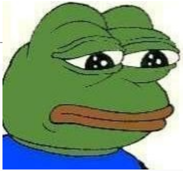
Si bien el gráfico es mucho más sensato, esto tampoco nos sirve mucho para nuestro análisis. En general, en este tipo de problemas, interpolar causa lo que se conoce como sobreajuste (“*overfitting*”). Esto significa que, si bien nuestro modelo ajusta perfectamente los datos que tenemos, no tiene capacidad de generalización. En otras palabras, si quisiéramos extrapolar a datos por fuera de $\{(x_i, y_i)\}_{i=0}^{14}$, no nos sería de mucha ayuda:



Motivación

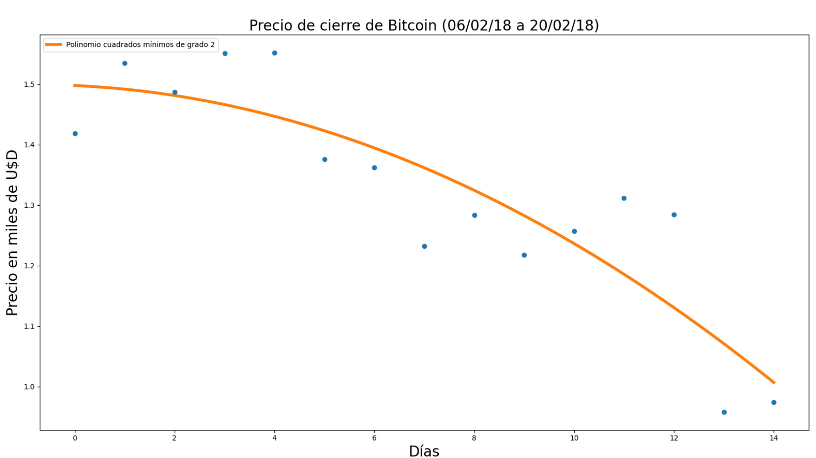
Si bien el gráfico es mucho más sensato, esto tampoco nos sirve mucho para nuestro análisis. En general, en este tipo de problemas, interpolar causa lo que se conoce como sobreajuste ("*overfitting*"). Esto significa que, si bien nuestro modelo ajusta perfectamente los datos que tenemos, no tiene capacidad de generalización. En otras palabras, si quisiéramos extrapolar a datos por fuera de $\{(x_i, y_i)\}_{i=0}^{14}$, no nos sería de mucha ayuda:



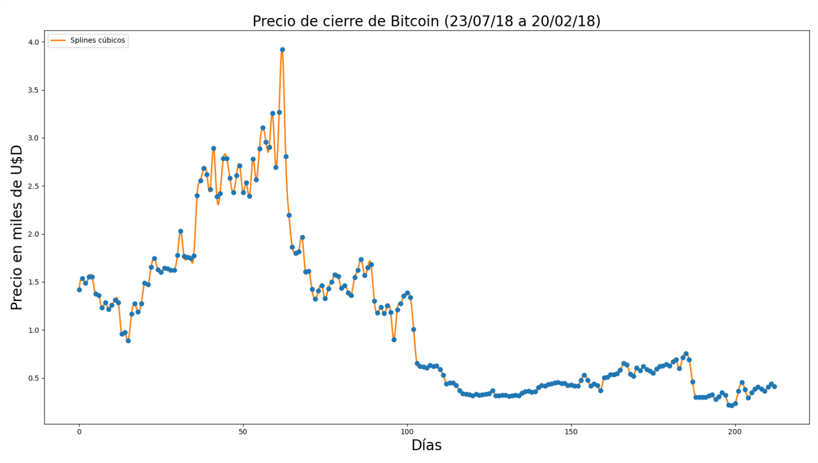


Motivación

Usando el polinomio de grado 2 que mejor aproxima en sentido de cuadrados mínimos:

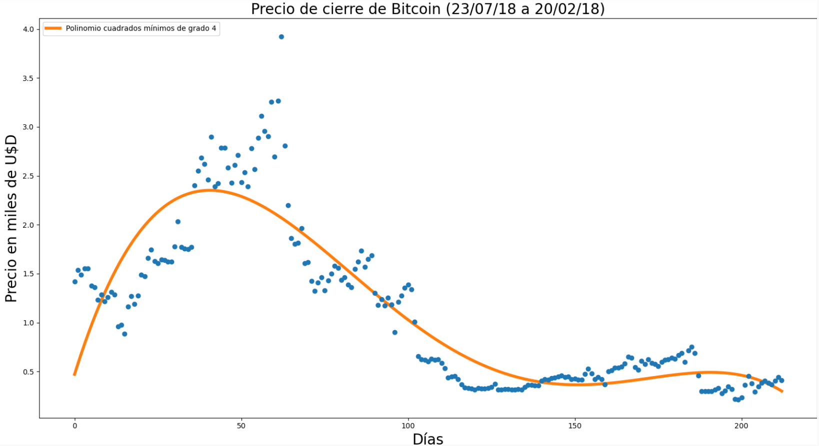


Motivación



Precio de cierre de Bitcoin (23/07/18 a 20/02/18)
Splines cúbicos
Precio en miles de U$D
Días
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
50
100
150
200

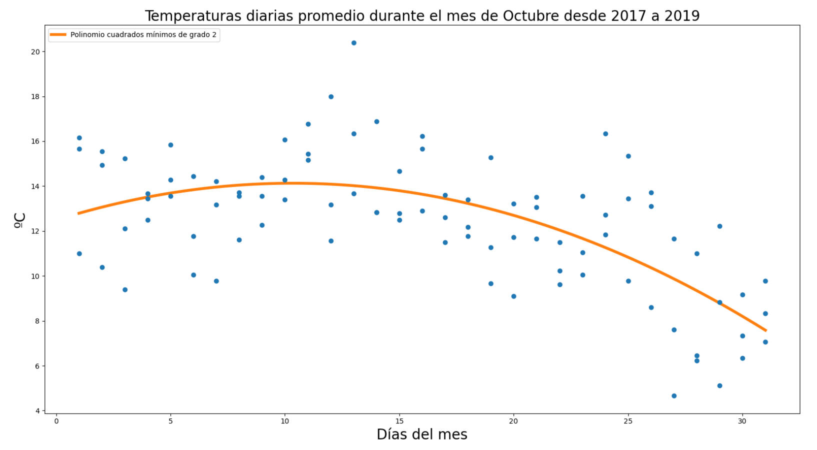
Motivación



Cuando tenemos muchos datos, es muy probable que para analizarlos nos convenga buscar una función que mejor los aproxime en sentido de cuadrados mínimos.

Motivación

A veces podemos tener muchos datos para el mismo x, por lo que en esos casos ni siquiera existe un polinomio interpolador



Problema A

- Dadas $\phi_1, \ldots, \phi_m$ funciones conocidas y un conjunto de datos $\{(x_i, y_i)\}_{i=1}^n$, queremos hallar los coeficientes $\{a_i\}_{i=1}^m$ que minimicen:

$$\sum_{i=1}^{n} (a_1\phi_1(x_i) + \cdots + a_m\phi_m(x_i) - y_i)^2$$

Utilizando la siguiente notación:

$$\underbrace{\begin{bmatrix} \phi_1(x_1) & \ldots & \phi_m(x_1) \\ \phi_1(x_2) & \ldots & \phi_m(x_2) \\ \vdots & \ddots & \vdots \\ \phi_1(x_n) & \ldots & \phi_m(x_n) \end{bmatrix}}_{A\in\mathbb{R}^{n\times m}} \quad \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}}_{a\in\mathbb{R}^m} \quad \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{b\in\mathbb{R}^n}$$

Como $n > m$, $Aa = b$ podría no tener solución. Por eso nos interesa buscar a^* tal que $\|Aa^* - b\|$ sea mínimo. Para eso contamos con el siguiente teorema:

Problema A

Teorema

Sean $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, $x^* \in \mathbb{R}^m$, son equivalentes:

1. x^* minimiza $\|Ax - b\|$
2. x^* es solución de $A^t Ax = A^t b$

Además, si las columnas de A son linealmente independientes, la solución de $A^t Ax = A^t b$ existe y es única.

Ejemplo 1: Aproximar los datos de la siguiente tabla con un modelo de la forma $f(x) \sim a_1 \dfrac{x}{x^2+1} + a_2 2^{-x}$ en el sentido de cuadrados mínimos:

x	-1	0	1	2
y	1	0,2	$-0{,}4$	$-0{,}3$

Queremos calcular la aproximación en cuadrados mínimos, siendo:

$$\phi_1(x) = \frac{x}{x^2+1} \qquad \phi_2(x) = 2^{-x}$$

Tenemos lo siguiente:

$$A = \begin{bmatrix} \phi_1(-1) & \phi_2(-1) \\ \phi_1(0) & \phi_2(0) \\ \phi_1(1) & \phi_2(1) \\ \phi_1(2) & \phi_2(2) \end{bmatrix} = \begin{bmatrix} -0{,}5 & 2 \\ 0 & 1 \\ 0{,}5 & 0{,}5 \\ 0{,}4 & 0{,}25 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 0{,}2 \\ -0{,}4 \\ -0{,}3 \end{bmatrix}$$

Problema A - Ejemplo 1

Como las columnas de A son linealmente independientes, basta calcular la solución de $A^tAa = A^tb$:

$$A^tA = \begin{bmatrix} 0{,}66 & -0{,}65 \\ -0{,}65 & 5{,}3125 \end{bmatrix} \qquad A^tb = \begin{bmatrix} -0{,}82 \\ 1{,}925 \end{bmatrix}$$

Para resolver el sistema podemos utilizar cualquiera de las herramientas que hemos estudiado hasta ahora, por ejemplo:

- como A^tA es diagonal dominante, podemos usar Jacobi o Gauss-Seidel
- como A^tA es simétrica definida positiva, podemos usar la descomposición de Cholesky
- como A^tA es inversible, podemos usar descomposición QR
- si queremos resolver el ejercicio en papel, triangulamos

De cualquier manera, la solución del sistema $A^tAa = A^tb$ es:

$$a^* = (a_1^*, a_2^*) = (-1{,}00689096, 0{,}23915687)$$

Entonces, la función que buscamos es $-1{,}00689096 \cdot \dfrac{x}{x^2+1} + 0{,}23915687 \cdot 2^{-x}$

Problema A - Ejemplo 2

Ejemplo 2: Aproximar los datos de la siguiente tabla con un modelo de la forma $f(x) \sim 2^{a_1 + a_2 x + a_3 x^2}$ en el sentido de cuadrados mínimos para la función $log_2(f(x))$:

x	0	1	2	3
y	1,9	2	2,4	3,5

Observación:

" ... en el sentido de cuadrados mínimos $\underline{\text{para } log_2(f(x))}$"

Lo que nosotros vamos a calcular:

$$a_1, a_2, a_3 \text{ que minimizan } \sum_{i=1}^{4} (log_2(2^{a_1 + a_2 x_i + a_3 x_i^2}) - log_2(y_i))^2$$

Podría no ser lo mismo que:

$$a_1, a_2, a_3 \text{ que minimizan } \sum_{i=1}^{4} (2^{a_1 + a_2 x_i + a_3 x_i^2} - y_i)^2$$

Pero, como vamos a ver después, funciona bastante bien.

Problema A - Ejemplo 2

En este caso, como nos piden aproximar en el sentido de cuadrados mínimos para la función $log_2(f(x))$, aplicamos log_2 a nuestro modelo y a los datos y:

x	0	1	2	3
$\hat{y} = log_2(y)$	$log_2(1{,}9)$	$log_2(2)$	$log_2(2{,}4)$	$log_2(3{,}5)$

Entonces ahora se trata de hallar la función de la forma $log_2(2^{a_1+a_2x+a_3x^2}) = a_1 + a_2x + a_3x^2$. Tenemos que $\phi_1(x) = 1$, $\phi_2(x) = x$ y $\phi_3(x) = x^2$.

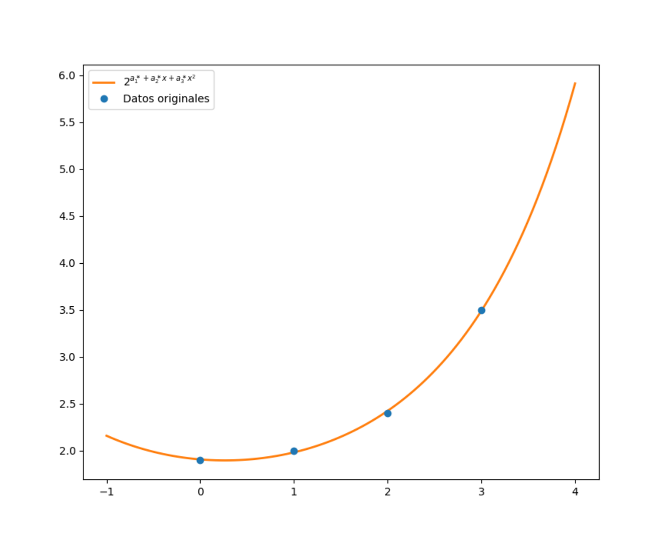
Sean:

$$A = \begin{bmatrix} \phi_1(0) & \phi_2(0) & \phi_3(0) \\ \phi_1(1) & \phi_2(1) & \phi_3(1) \\ \phi_1(2) & \phi_2(2) & \phi_3(2) \\ \phi_1(3) & \phi_2(3) & \phi_3(3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \qquad b = \begin{bmatrix} log_2(1{,}9) \\ log_2(2) \\ log_2(2{,}4) \\ log_2(3{,}5) \end{bmatrix}$$

Como las columnas de A son l.i., la solución a^* del sistema $A^tAa = A^tb$ existe y es única. Resolviéndolo, nos queda que:

$$a_1^* = 0{,}93045983 \quad a_2^* = -0{,}06169882 \quad a_3^* = 0{,}11748012$$

¡Y funciona bien!




SVD
Let's go. In and out. 20 minute adventure.

Recordar

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

- $rango(A)$ =número de filas l.i.=número de columnas l.i.= $dim(Im(A))$
- $Nu(A) = \{x \in \mathbb{R}^n : Ax = 0\}$
- $dim(Nu(A)) = n - rango(A)$
- $Nu(A) = Nu(A^T A) \quad (\Rightarrow rango(A) = rango(A^T A))$
- $A^T A \in \mathbb{R}^{n \times n}$ es simétrica definida positiva, por lo tanto todos sus autovalores son reales y no negativos. Lo mismo ocurre para $AA^T \in \mathbb{R}^{m \times m}$.
- $\lambda \neq 0$, λ es autovalor de $A^T A \iff \lambda$ es autovalor de AA^T:
  - $\lambda \neq 0$ autovalor de $A^T A$, luego:

$$A^T A v = \lambda v \Rightarrow (AA^T) A v = \lambda A v$$

  - $\lambda \neq 0$ autovalor de AA^T, luego:

$$AA^T v = \lambda v \Rightarrow (A^T A) A^T v = \lambda A^T v$$

Valores singulares

Definición

Sean $A \in \mathbb{R}^{m \times n}$ y $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ los autovalores de $A^T A$ ordenados de manera decreciente:

$$\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 \geq 0$$

decimos que $\sigma_i = \sqrt{\sigma_i^2}$ es el i-ésimo valor singular de A.

En otras palabras: los valores singulares de A son las raíces cuadradas de los autovalores de $A^T A$, ordenados de manera decreciente.

Descomposición en valores singulares

Definición

Sea $A \in \mathbb{R}^{m\times n}$, $m \geq n$, una descomposición en valores singulares de A es una factorización:

$$A = U\Sigma V^T$$

tales que $U \in \mathbb{R}^{m\times m}, \; V \in \mathbb{R}^{n\times n}$ ortogonales (i.e $UU^T = I, \; VV^T = I$) y $\Sigma \in \mathbb{R}^{m\times n}$ diagonal de la forma:

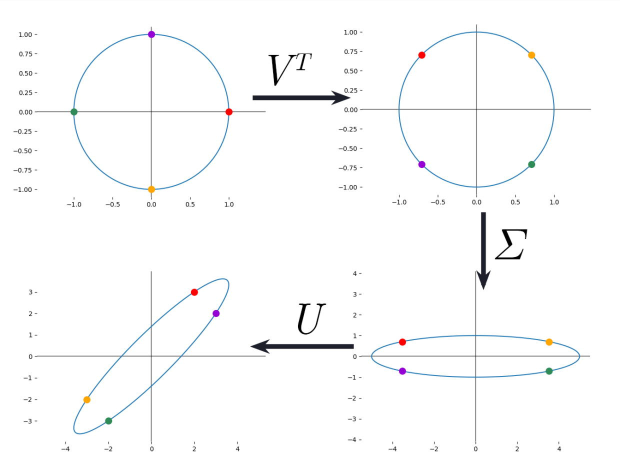
$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & & \sigma_n \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

Teorema

Sea $A \in \mathbb{R}^{m\times n}$, A admite una descomposición en valores singulares.

$$A = \underbrace{U}_{\substack{\text{rotación o}\\ \text{reflexión}}} \quad \underbrace{\Sigma}_{\text{reescalamiento}} \quad \underbrace{V^T}_{\substack{\text{rotación o}\\ \text{reflexión}}}$$

$$\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} = \underbrace{\begin{pmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}}_{\substack{U \\ \text{reflexión de } \frac{5}{8}\pi}} \underbrace{\begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}}_{\substack{\Sigma \\ \text{reescalamiento}}} \underbrace{\begin{pmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}}_{\substack{V^T \\ \text{rotación de } \frac{3}{4}\pi}}$$



Ejemplo: hallar la descomposición SVD para la siguiente matriz $A \in \mathbb{R}^{3\times 2}$:

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix}$$

Solución: comenzamos obteniendo $\Sigma \in \mathbb{R}^{3\times 2}$ a partir de los valores singulares de A. Para esto, calculamos los autovalores de $A^T A$:

$$A^T A = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} \to \sigma_1^2 = 25,\ \sigma_2^2 = 9$$

Descomposición en valores singulares

Ejemplo: hallar la descomposición SVD para la siguiente matriz $A \in \mathbb{R}^{3\times 2}$:

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix}$$

Solución: comenzamos obteniendo $\Sigma \in \mathbb{R}^{3\times 2}$ a partir de los valores singulares de A. Para esto, calculamos los autovalores de $A^T A$:

$$A^T A = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} \rightarrow \sigma_1^2 = 25,\ \sigma_2^2 = 9$$

Por lo tanto, los valores singulares de A son:

$$\sigma_1 = 5,\ \sigma_2 = 3$$

Ejemplo: hallar la descomposición SVD para la siguiente matriz $A \in \mathbb{R}^{3\times 2}$:

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix}$$

Solución: comenzamos obteniendo $\Sigma \in \mathbb{R}^{3\times 2}$ a partir de los valores singulares de A. Para esto, calculamos los autovalores de $A^T A$:

$$A^T A = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} \to \sigma_1^2 = 25,\ \sigma_2^2 = 9$$

Por lo tanto, los valores singulares de A son:

$$\sigma_1 = 5,\ \sigma_2 = 3$$

Luego:

$$\Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

El siguiente paso es construir $V \in \mathbb{R}^{2\times 2}$ ortogonal. Como $A^T A$ es simétrica y definida positiva, existe una base ortogonal de autovectores asociados a los autovalores $\sigma_1^2 = 25$, $\sigma_2^2 = 9$. Así que calculamos dichos autovectores y los normalizamos para obtener la b.o.n.:

Para $\sigma_1^2 = 25$:

$$\begin{pmatrix} 25-17 & -8 \\ -8 & 25-17 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{\text{Normalizando}} v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Para $\sigma_2^2 = 9$:

$$\begin{pmatrix} 9-17 & -8 \\ -8 & 9-17 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xrightarrow{\text{Normalizando}} v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Luego, $V = (v_1|v_2)$:

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Finalmente, construimos $U \in \mathbb{R}^{3\times 3}$.

Como $A = U\Sigma V^T$ y $V^{-1} = V^T$, entonces $AV = U\Sigma$:

$$\begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \frac{5}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{5}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{4}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 5u_{11} & 3u_{12} \\ 5u_{21} & 3u_{22} \\ 5u_{31} & 3u_{32} \end{pmatrix}$$

$$\Rightarrow U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & u_{13} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & u_{23} \\ 0 & \frac{4}{3\sqrt{2}} & u_{33} \end{pmatrix}$$

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & u_{13} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & u_{23} \\ 0 & \frac{4}{3\sqrt{2}} & u_{33} \end{pmatrix}$$

Para hallar la última columna de U basta hallar un vector que sea ortogonal a su primera y a su segunda columna:

$$\begin{cases} \frac{1}{\sqrt{2}} u_{13} + \frac{1}{\sqrt{2}} u_{23} + 0 u_{33} = 0 \\ \\ \frac{1}{3\sqrt{2}} u_{13} + \left(-\frac{1}{3\sqrt{2}}\right) u_{23} + \frac{4}{3\sqrt{2}} u_{33} = 0 \end{cases}$$

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & u_{13} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & u_{23} \\ 0 & \frac{4}{3\sqrt{2}} & u_{33} \end{pmatrix}$$

Para hallar la última columna de U basta hallar un vector que sea ortogonal a su primera y a su segunda columna:

$$\begin{cases} \frac{1}{\sqrt{2}} u_{13} + \frac{1}{\sqrt{2}} u_{23} + 0 u_{33} = 0 \\ \\ \frac{1}{3\sqrt{2}} u_{13} + \left(-\frac{1}{3\sqrt{2}}\right) u_{23} + \frac{4}{3\sqrt{2}} u_{33} = 0 \end{cases}$$

De la primera ecuación, se tiene que $u_{23} = -u_{13}$. Por lo tanto la segunda queda como:

$$\frac{1}{3\sqrt{2}} u_{13} + \left(-\frac{1}{3\sqrt{2}}\right)(-u_{13}) + \frac{4}{3\sqrt{2}} u_{33} = 0 \Rightarrow u_{33} = -\frac{1}{2} u_{13}$$

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & u_{13} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & u_{23} \\ 0 & \frac{4}{3\sqrt{2}} & u_{33} \end{pmatrix}$$

Para hallar la última columna de U basta hallar un vector que sea ortogonal a su primera y a su segunda columna:

$$\begin{cases} \frac{1}{\sqrt{2}}u_{13} + \frac{1}{\sqrt{2}}u_{23} + 0u_{33} = 0 \\ \\ \frac{1}{3\sqrt{2}}u_{13} + \left(-\frac{1}{3\sqrt{2}}\right)u_{23} + \frac{4}{3\sqrt{2}}u_{33} = 0 \end{cases}$$

De la primera ecuación, se tiene que $u_{23} = -u_{13}$. Por lo tanto la segunda queda como:

$$\frac{1}{3\sqrt{2}}u_{13} + \left(-\frac{1}{3\sqrt{2}}\right)(-u_{13}) + \frac{4}{3\sqrt{2}}u_{33} = 0 \Rightarrow u_{33} = -\frac{1}{2}u_{13}$$

Luego las soluciones son vectores del tipo $u_{13} \cdot (1, -1, -\frac{1}{2})$ para $u_{13} \in \mathbb{R}$. Como queremos que las columnas de U tengan norma igual a 1:

$$\|u_{13} \cdot \left(1, -1, -\frac{1}{2}\right)\| = 1 \iff |u_{13}| = \frac{2}{3}$$

Podemos tomar $\left(\frac{2}{3}, -\frac{2}{3}, -\frac{1}{3}\right)$ como la última columna de U.

Obtuvimos la descomposición SVD de A:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Eliminando las filas nulas de Σ y las correspondientes columnas de U, se obtiene la versión reducida de la descomposición SVD:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{4}{3\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Al final hicimos muchas cuentas

Además, de yapa, obtuvimos:

1. La descomposición SVD de A^T:

$$A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

2. La pseudoinversa de A: $A^\dagger = V\Sigma^\dagger U^T$ donde:

$$\Sigma^\dagger = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & & & \frac{1}{\sigma_n} \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

En el Ejercicio 7 de la guía de ejercicios se ve cómo aplicar $A^\dagger$ al problema de cuadrados mínimos: bajo ciertas hipótesis, el vector a^* que minimiza $\|Aa - b\|$ es $a^* = A^\dagger b$.

Algunos resultados extra

Resultados interesantes

1. $\sigma_1 = \|A\|_2 = \sqrt{\rho(A^T A)}$
2. Si A es inversible,

$$A = U\Sigma V^T \Rightarrow A^{-1} = V\Sigma^{-1}U^T \Rightarrow \|A^{-1}\|_2 = \frac{1}{\sigma_n}$$

   Entonces,

$$cond_2(A) = \frac{\sigma_1}{\sigma_n}$$