

Álgebra Lineal Computacional - Clase Práctica 10 - Pseudoinversa y cuadrados mínimos

Fausto Martínez

6 de Junio de 2025

En azul están remarcados los typos y errores de la primera versión, mil disculpas :(
La idea de esta clase es desarrollar en detalle las técnicas de resolución que tenemos cuando

afrontamos un sistema lineal

$$Ax = b$$

con $A \in \mathbb{C}^{m \times n}$, $x \in \mathbb{C}^n$, $b \in \mathbb{C}^m$ cuando $m > n$.

Recordemos brevemente entonces cuáles eran los casos con los que nos podíamos encontrar:

- Si $b \in \text{Im}(A)$

Por definición, esto quiere decir que existe un $x \in \mathbb{C}^n$ tal que $Ax = b$. Recordando que

$$\dim(\ker(A)) = n - \text{rk}(A)$$

tenemos que si

- $\text{rk}(A) = n$, entonces la única solución del sistema es x
- $\text{rk}(A) < n$, entonces las infinitas soluciones del sistema son $x + \ker(A)$
- Si $b \notin \text{Im}(A)$, entonces el sistema no tiene solución.

Ahora, como en la práctica suele ocurrir que $m \gg n$, entonces lo más probable es que $b \notin \text{Im}(A)$, pues

$$\dim(\text{Im}(A)) \leq n \ll m$$

Aun así, debido a que la aparición de este tipo de problemas en la práctica es muy frecuente, uno no se puede quedarse de brazos cruzados dando como respuesta que no hay solución. Tenemos que buscar algo así como *lo más cercano a una solución* que podamos.

Una posibilidad sería descartar las últimas $m - n \gg 0$ ecuaciones del sistema para llegar a uno cuadrado con los que ya sabemos trabajar, pero esto estaría muy poco justificado si todas las ecuaciones son igual de importantes. De esa manera, el enfoque que utilizamos es *minimizar por cuánto le pifiamos a todas las m ecuaciones*, lo cual, por supuesto, desarrollaremos con más detalle matemáticamente a lo largo del documento.

Pseudoinversa

Como ya vieron la clase anterior, dada una matriz $A \in \mathbb{C}^{m \times n}$, $m > n$ siempre existe una descomposición en valores singulares suya, dada por

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

donde $U \in \mathbb{C}^{m \times m}$ y $V \in \mathbb{C}^{n \times n}$ son matrices unitarias y $\Sigma \in \mathbb{R}^{m \times n}$ es una matriz “diagonal” en el sentido de que $(\Sigma)_{ij} = 0$ para $i \neq j$ y $(\Sigma)_{ii} = \sigma_i \geq 0$ con $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$.

Los números $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ son los autovalores de A^*A y las columnas v_1, v_2, \dots, v_n de V son los autovectores de la misma (llamados *vectores singulares derechos*), a la vez que las columnas de U , es decir u_1, u_2, \dots, u_m son los autovectores de AA^* (llamados *vectores singulares izquierdos*). Los números $\sigma_1, \sigma_2, \dots, \sigma_n$ los llamamos *valores singulares* de A .

Resulta entonces

$$A = \left(\begin{array}{c|c|c|c} | & | & \dots & | \\ \hline u_1 & u_2 & \dots & u_m \\ \hline | & | & \dots & | \end{array} \right)_{m \times m} \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ \hline 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}_{m \times n} \begin{pmatrix} - & v_1^* & - \\ - & v_2^* & - \\ \vdots & \vdots & \vdots \\ - & v_n^* & - \end{pmatrix}_{n \times n}$$

Como vieron en la teórica, uno puede definir la *SVD truncada de rango k* quedándose hasta el k -ésimo valor singular en la diagonal y haciendo que los demás sean 0. Si llamamos A_k a esa matriz, tenemos:

$$A_k = \left(\begin{array}{c|c|c|c} | & | & \dots & | \\ \hline u_1 & u_2 & \dots & u_m \\ \hline | & | & \dots & | \end{array} \right)_{m \times m} \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_k & \\ & & & & \ddots \\ \hline & & & & 0 \\ 0 & & \dots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & \dots & & 0 \end{pmatrix}_{m \times n} \begin{pmatrix} - & v_1^* & - \\ - & v_2^* & - \\ \vdots & \vdots & \vdots \\ - & v_n^* & - \end{pmatrix}_{n \times n}$$

$$= \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \dots + \sigma_k u_k v_k^*$$

$$= \sum_{i=1}^k \sigma_i u_i v_i^*$$

Definición

Dada una matriz $A \in \mathbb{C}^{m \times n}$, definimos su descomposición SVD truncada en k como

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^*$$

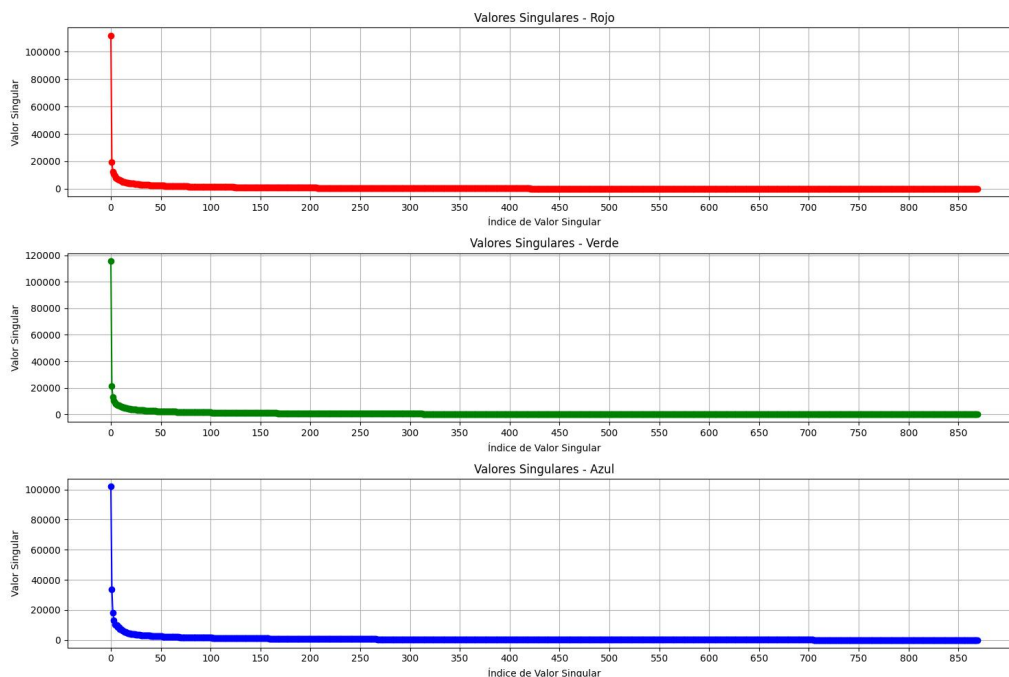
Proposición

La descomposición SVD truncada en k es la mejor aproximación de rango k que se puede hacer de A , es decir

$$A_k = \arg \min \{ \|A - B\|_2 : \text{rk}(B) = k \}$$

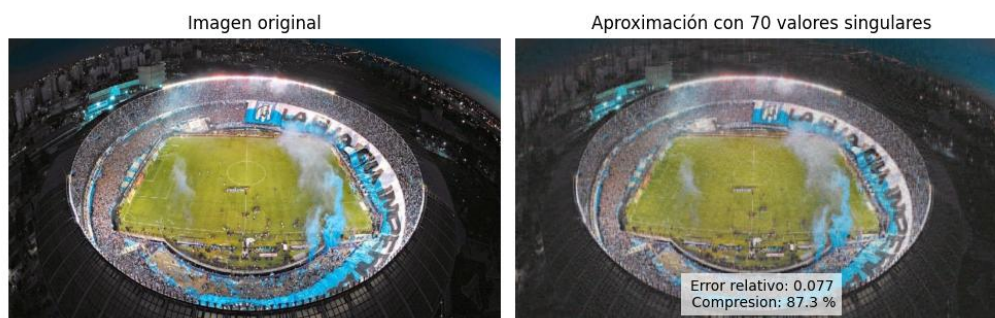
Este resultado es importante pues podemos hacer una interpretación de los valores singulares: de cierto modo el valor numérico de los mismos nos dice cuánta información guarda el mismo de la matriz. Habrán visto en el laboratorio también que este resultado puede utilizarse para comprimir imágenes.

Lo copado también es que usualmente los valores singulares decaen rapidísimo. Esto es, los primeros valores singulares tienen valores altos, y al seguir viendo los siguientes el valor decae rapidísimo. Veamos un ejemplo:



Valores singulares de las 3 matrices que codificaban una imagen en RGB

Como podrán observar en este caso, donde las matrices eran de 870×1500 , el valor de cada valor singular decae rapidísimo, y eso nos indica que efectivamente podíamos hacer buenas aproximaciones de la imagen considerando los primeros 70 valores singulares, lo cual en este ejemplo en particular hubiera hecho que en lugar de guardar casi 4 millones de datos, pasemos a tener la imagen guardada en tan solo 500 mil. Solo por completitud dejo cuál era la imagen original y cuál era su aproximación con 70 valores singulares, para que ustedes juzguen efectivamente si pueden encontrar alguna diferencia a pesar de los datos que nos ahorramos



Comparación entre la imagen original y la aproximada con SVD

Finalmente, volviendo a lo que nos competía para esta clase, veamos esta definición

Definición

Dada una matriz $A \in \mathbb{C}^{m \times n}$ y su descomposición SVD, definimos su *pseudoinversa* como

$$A^\dagger = V \Sigma^\dagger U^*$$

$$= \begin{pmatrix} | & | & \dots & | \\ v_1 & v_2 & \dots & v_n \\ | & | & \dots & | \end{pmatrix}_{n \times n} \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_s} & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}_{n \times m} \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{n \times m} \begin{pmatrix} - & u_1^* & - \\ - & u_2^* & - \\ \vdots & \vdots & \vdots \\ - & u_m^* & - \end{pmatrix}_{m \times m}$$

donde σ_s es el último valor singular positivo y por ende $\sigma_{s+1}, \dots, \sigma_n$ eran 0.

Esta matriz, por supuesto, no es tan copada como la inversa, pero cumple ciertas propiedades que hacen que se le parezca, y van a tener que demostrarlas como ejercicio de la guía:

$$\begin{aligned} AA^\dagger A &= A \\ A^\dagger AA^\dagger &= A^\dagger \\ (AA^\dagger)^* &= AA^\dagger \\ (A^\dagger A)^* &= A^\dagger A \\ A^\dagger &\text{ es única} \end{aligned}$$

Ejercicio 1

Demostrar que $AA^\dagger A = A$

Solución: Dada una matriz $A \in \mathbb{C}^{m \times n}$, su descomposición en valores singulares es

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

con $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ unitarias y $\Sigma \in \mathbb{R}^{m \times n}$ “diagonal”.

Luego tenemos

$$\begin{aligned} AA^\dagger A &= U \Sigma V^* V \Sigma^\dagger U^* U \Sigma V^* \\ &= U \Sigma \Sigma^\dagger \Sigma V^* \end{aligned}$$

podemos escribir a Σ como una matriz en bloques $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, entonces

$$\begin{aligned} &= U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^* \\ &= U \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^* \\ &= U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^* \\ &= U \Sigma V^* \\ &= A \end{aligned}$$

Cuadrados mínimos

Volvemos entonces al eje de la clase con las ecuaciones del tipo

$$Ax = b$$

con $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ y $b \in \mathbb{R}^m$.

Habíamos comentado que ante lo poco probable que es que este sistema tenga solución, debíamos encontrar algo cercano a la misma sin deshacernos de ecuaciones. Entonces, lo que se hace es buscar una solución que le pifie lo menos posible. A esta solución la llamaremos \hat{x} y va a ser

$$\hat{x} = \arg \min_{x \in \mathbb{R}^n} \{ \|Ax - b\|_2^2 \}$$

Es decir, no pudimos hacer que $Ax - b$ sea 0, pero intentamos acercarnos a eso lo máximo que podamos. Para eso entonces definamos la función de n variables:

$$\begin{aligned} f(x) &= \|Ax - b\|_2^2 \\ &= (Ax - b)^T (Ax - b) \\ &= (x^T A^T - b^T)(Ax - b) \\ &= x^T A^T Ax - 2b^T Ax + b^T b \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^m x_i x_j a_{ki} a_{kj} - 2 \sum_{i=1}^m \sum_{j=1}^n b_i a_{ij} x_j + \sum_{i=1}^m b_i^2 \end{aligned}$$

Recordemos también de Análisis I, que si buscamos minimizar esta función debemos igualar su gradiente a 0. Analicemos entonces las derivadas direccionales de la función f y tendremos:

$$\frac{\partial f}{\partial x_\ell} = \frac{\partial}{\partial x_\ell} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^m x_i a_{ki} a_{kj} x_j - \frac{\partial}{\partial x_\ell} 2 \sum_{i=1}^m \sum_{j=1}^n b_i a_{ij} x_j + \frac{\partial}{\partial x_\ell} \sum_{i=1}^m b_i^2.$$

Vamos por partes, tenemos para el último término que

$$\frac{\partial}{\partial x_\ell} \sum_{i=1}^m b_i^2 = 0, \quad \forall \ell = 1, \dots, n$$

Para el segundo término tenemos

$$\frac{\partial}{\partial x_\ell} 2 \sum_{i=1}^m \sum_{j=1}^n b_i a_{ij} x_j = 2 \sum_{i=1}^m b_i a_{i\ell} = (2A^T b)_\ell$$

Y finalmente para el primero,

$$\frac{\partial}{\partial x_\ell} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^m x_i a_{ki} a_{kj} x_j = \sum_{i=1}^n \sum_{k=1}^m x_i a_{ki} a_{k\ell} + \sum_{j=1}^n \sum_{k=1}^m x_j a_{k\ell} a_{kj} = A^T Ax + A^T Ax = (2A^T Ax)_\ell$$

Bueno, luego de comernos este garronazo de cuentas, que vale la pena hacerlo una vez en la vida (hagánlo por sus cuentas) y después para el resto de la vida simplemente creerlo, tenemos que el gradiente es

$$\nabla f(x) = 2A^T Ax - 2A^T b$$

Y por ende es igual a 0 cuando

$$A^T Ax = A^T b$$

Notemos que lo que tenemos ahora es un sistema con una matriz cuadrada $A^T A$.

Proposición

Dadas la matriz $A \in \mathbb{R}^{m \times n}$, y el vector $b \in \mathbb{R}^m$, la solución del problema de mínimos cuadrados

$$\arg \min_{x \in \mathbb{R}^n} \{\|Ax - b\|_2^2\}$$

es la solución al problema

$$A^T A x = A^T b$$

Como veremos a lo largo de esta clase, y probablemente por el resto de sus vidas, los problemas de mínimos cuadrados aparecen en innumerables aplicaciones, por lo que tenerlo asociado a un sistema lineal, para los cuales desarrollamos ya diversas maneras de solucionar, es un resultado muy groso y hay que saber apreciarlo. Para ponernos en contexto, enumero algunas de esas aplicaciones:

- Ajustar datos a una recta.
- Ajustar datos a un polinomio cualquiera.
- Reducción de la dimensionalidad de un dataset.
- En economía y finanzas, los ajustes son usados para predecir tendencias económicas o analizar riesgos.
- En Física, se utilizan ajustes para calibrar sensores, ajustar trayectorias y estimar parámetros de modelos físicos (por ejemplo, la elasticidad o resistencia de algún material o la gravedad de la Tierra)
- En los procesos de reconstrucción de imágenes.
- Y muchísimas más en Medicina, Machine Learning, y prácticamente todo lo que se imaginen.

En fin, los ajustes lineales son una herramienta prácticamente universal en la ciencia, y me pareció importante notar esto antes de seguir con las cuentas.

Vayamos entonces a ver las aplicaciones. Veamos como primer ejemplo el clásico, ajustar linealmente por una recta:

Ejercicio 2

Dado el conjunto de datos $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), (x_5, y_5)\}$, que están dados por la siguiente tabla

x	1	2	5	8	9
y	-1.7	-0.3	0.5	1.8	2.7

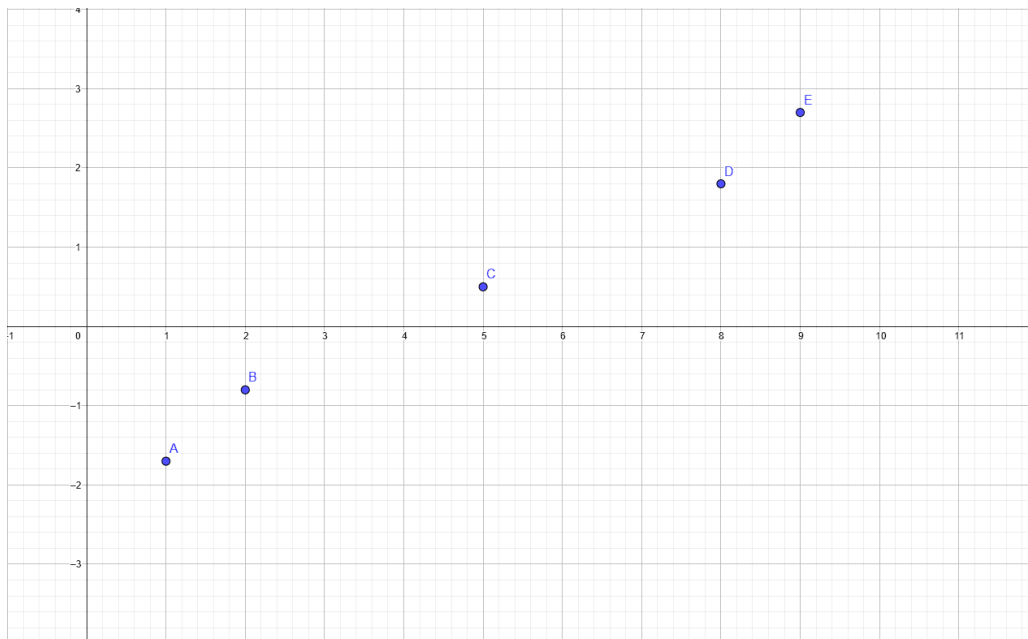
hallar la recta que mejor aproxime a los datos.

Solución: Uno podría proponer varias maneras de decidir cuál es la recta que “mejor aproxima” a los datos, pero la que más se usa es considerar los residuos entre nuestros datos y la función que queremos que los aproxime, es decir $r_i = y_i - f(x_i)$ para $i = 1, \dots, 5$.

Lo que queremos minimizar es entonces la suma de los residuos cuadrados, es decir:

$$S(\beta_1, \beta_0) = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n (y_i - \beta_1 x_i - \beta_0)^2$$

Noten que acá nuestras incógnitas son β_1 y β_0 , los parámetros de la recta



Datos del Ejercicio 1 graficados en el plano

Por supuesto, queremos escribir este problema de forma matricial. Entonces notamos que si definimos

$$A = \begin{pmatrix} | & | \\ 1 & x \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 5 \\ 1 & 8 \\ 1 & 9 \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_0 \end{pmatrix}, y = \begin{pmatrix} -1.7 \\ -0.3 \\ 0.5 \\ 1.8 \\ 2.7 \end{pmatrix}$$

Podemos reescribir lo que estamos minimizando como

$$S(\beta) = \|A\beta - y\|_2^2$$

Llegamos entonces al mismo problema que habíamos desarrollado al comienzo. Gracias a la horrible cuenta que hicimos una vez y no haremos nunca más, sabemos que la solución es la solución al sistema lineal

$$A^T A \beta = A^T y$$

Tenemos que resolver entonces

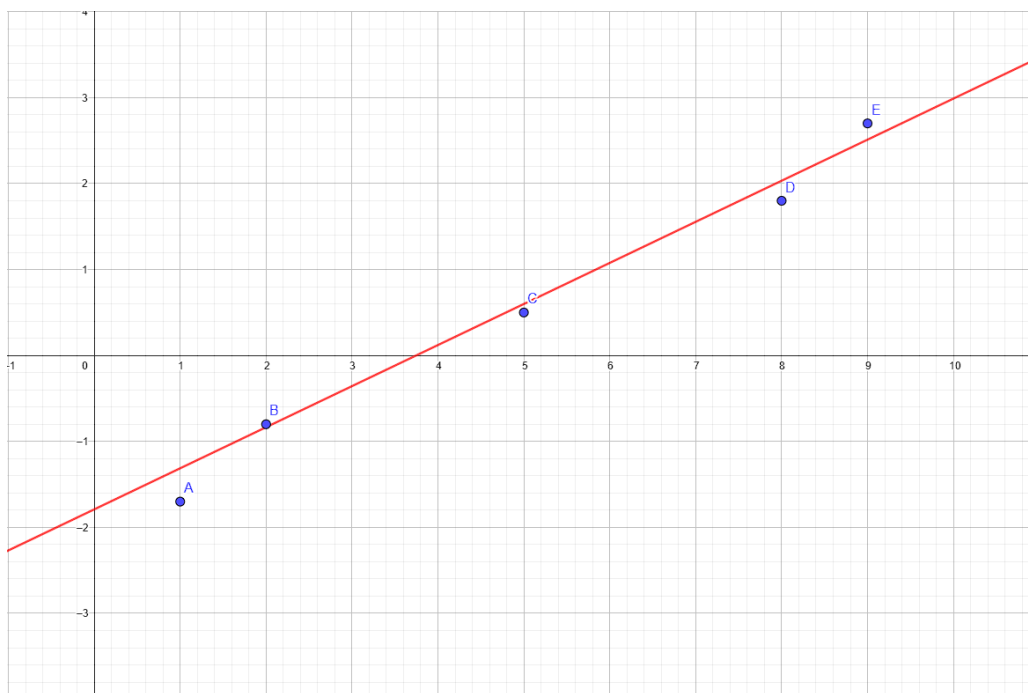
$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 5 \\ 1 & 8 \\ 1 & 9 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 8 & 9 \end{pmatrix} \begin{pmatrix} -1.7 \\ -0.3 \\ 0.5 \\ 1.8 \\ 2.7 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 25 \\ 25 & 175 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 38.9 \end{pmatrix}$$

Con lo que finalmente obtenemos que

$$\begin{aligned} \beta_1 &= 0.478 \\ \beta_0 &= -1.79 \end{aligned}$$

Es decir, que la recta que mejor aproxima a nuestros datos es la dada por la función



Recta que ajusta a los datos del Ejercicio 1

$$f(x) = 0.478x - 1.79$$

Esto puede ser útil en muchos sentidos. Como había dicho, en física puede ser que los parámetros β_1 o β_0 nos estén diciendo cuál es el valor de algo de la naturaleza que nos interesa (por ejemplo, la gravedad de la tierra, o la constante elástica de un resorte). A su vez también puede servir pues si uno le asigna algún significado a estos valores, la recta nos puede servir como un predictor. Como ejemplo, si los x_i fueran los miles de metros cuadrados de una propiedad, y los y_i fueran su precio, que ya conocemos, si ahora nos dan una propiedad nueva, de la cual desconocemos su valor, de digamos 11 mil metros cuadrados, podríamos estimar el mismo a partir de $f(11)$.

Cómputo de la solución de mínimos cuadrados

Ya vimos con todo el cuenterío que

$$\hat{x} = \arg \min_{x \in \mathbb{R}^n} \{ \|Ax - b\|_2^2 \}$$

se puede obtener como la solución del sistema lineal

$$A^T A x = A^T b.$$

El anterior es un resultado de vital importancia teórica pero que en la práctica es una manera jamás utilizada para resolver un problema de cuadrados mínimos. El motivo es simple: no podemos controlar $\text{cond}(A^T A)$.

Habrán visto en la teoría que podíamos generalizar el número de condición en norma 2 a matrices $A \in \mathbb{R}^{m \times n}$ con $m > n$ si llamamos al valor singular más grande de A como $\sigma_1(A)$ y al más pequeño como $\sigma_n(A)$, del siguiente modo:

$$\text{cond}_2(A) = \frac{\sigma_1(A)}{\sigma_n(A)}$$

cuando A era de rango completo y si no decíamos que $\text{cond}_2(A) = +\infty$.

Luego, aboquémonos al siguiente ejercicio

Ejercicio 3

Dada una matriz $A \in \mathbb{R}^{m \times n}$ con $m > n$, demostrar que

$$\text{cond}_2(A^T A) = \text{cond}_2(A)^2$$

Solución: Si A es de rango completo,

$$\text{cond}_2(A^T A) = \frac{\lambda_1(A^T A)}{\lambda_n(A^T A)} = \frac{\sigma_1(A)^2}{\sigma_n(A)^2} = \text{cond}_2(A)^2$$

donde, siguiendo la misma idea que antes, denoté a $\lambda_1(AA^T)$ como el autovalor más grande de AA^T y con $\lambda_n(AA^T)$ al autovalor más chico.

Observación

Dividir por $\lambda_n(A^T A)$ está bien definido, pues al ser A de rango completo, $A^T A$ es definida positiva y por ende inversible, como se puede ver de

$$x^t A^T A x = (Ax)^T Ax = \|Ax\|_2^2 > 0$$

pues $Ax \neq 0$ para $x \neq 0$, pues A es de rango completo.

De la definida positividad de A se desprende que todos sus autovalores son positivos, entonces se puede dividir por $\lambda_n(A^t A)$

Si A no fuera de rango completo, es decir

$$\text{rk}(A) < n$$

entonces $A^T A$ no sería inversible, pues

$$\text{rk}(A^T A) \leq \min\{\text{rk}(A^T), \text{rk}(A)\} = \text{rk}(A) < n < m.$$

Esto quiere decir que si $\text{cond}_2(A) = +\infty$, entonces también $\text{cond}_2(A^T A) = +\infty$. \square

La conclusión a la que quiero llegar con esto y que me gustaría que se lleven es que utilizar la ecuación normal, eleva al cuadrado la condición del problema, pues

$$\text{cond}_2(A^T A) = \text{cond}_2(A)^2$$

Dado que la condición es siempre un número mayor o igual que 1, y en la práctica, para una matriz A arbitraria, bastante mayor que 1, utilizar la ecuación normal hace al problema **súper inestable**, y no es un método deseable para resolver computacionalmente mínimos cuadrados.

¿Qué se hace en la práctica entonces? Una opción es utilizar la descomposición QR de A , dado que las matrices unitarias nos vienen siempre copadas pues como ya vimos, vale la siguiente proposición

Proposición

Dada una matriz $Q \in \mathbb{R}^{n \times n}$ ortogonal, vale que

- $\|Q\|_2 = 1$
- $\|Qx\|_2 = \|x\|_2$, para todo $x \in \mathbb{R}^n$

Tenemos, entonces, considerando la descomposición QR de A , dada por

$$A = QR$$

donde Q es ortogonal y R es triangular superior, que

$$\|Ax - b\|_2^2 = \|Q^T(Ax - b)\|_2^2 = \|Rx - Q^Tb\|_2^2$$

Tenemos, entonces que resolver el sistema

$$Rx = Q^Tb$$

también es una solución del problema de mínimos cuadrados.

¿Qué pasará ahora con la estabilidad?

Ejercicio 4

Dada la descomposición QR de una matriz $A \in \mathbb{R}^{m \times n}$ con $m > n$, dadas por

$$A = QR$$

con $Q \in \mathbb{R}^{m \times m}$ ortogonal y $R \in \mathbb{R}^{m \times n}$ triangular superior, demostrar que

$$\text{cond}_2(A) = \text{cond}_2(R)$$

Solución: Recordando que

$$\text{cond}_2(A) = \frac{\sigma_1(A)}{\sigma_n(A)}$$

Basta con ver que los valores singulares de $R = Q^T A$ son los mismos que los de A , y esto vale pues al considerar la descomposición en valores singulares de A dada por

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

podemos hallar fácilmente la de R como

$$R = Q^T A = Q^T U \Sigma V^T = \tilde{U} \Sigma V^T$$

lo que muestra que R tiene los mismos valores singulares que A y por ende

$$\text{cond}_2(R) = \frac{\sigma_1(R)}{\sigma_n(R)} = \frac{\sigma_1(A)}{\sigma_n(A)} = \text{cond}_2(A).$$

□

Vemos que mejoramos mucho la estabilidad del método pues ahora dejamos el número de condición igual que el de la matriz original, y aparte solo nos queda por resolver un sistema triangular que como vimos es fácil en términos de complejidad (dada ya la descomposición QR, claro).

¿Podemos mejorar aún más esto? Bueno, la pista debería estar en la definición del número de condición. Como vemos, el mismo depende claramente de qué tan grandes o chicos sean σ_1 y σ_n . Claramente no podemos hacer más chico σ_1 pues es el valor singular que más información de la matriz nos da y no nos podemos deshacer de él así nomas, pero con respecto a σ_n , ya vimos que la mayor parte de los últimos valores singulares aportaban poco y nada a la información que tenemos sobre la matriz, así que quizás sea una buena idea tirarlos, y quedarnos con la SVD Truncada. Formalicemos esto.

Dada la descomposición SVD de $A \in \mathbb{R}^{m \times n}$ con $m > n$ dada por

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

donde el último valor singular positivo es el σ_s , como venimos notando, el problema de mínimos cuadrados puede ser escrito como

$$\begin{aligned}
 \|Ax - b\|_2^2 &= \|U\Sigma V^T x - b\|_2^2 \\
 &= \|U^T(U\Sigma V^T x - b)\|_2^2 \\
 &= \|\Sigma V^T x - U^T b\|_2^2 \\
 &\text{y, si denotamos } y = V^T x, c = U^T b \\
 &= \|\Sigma y - c\|_2^2 \\
 &= \sum_{i=1}^s (\sigma_i y_i - c_i)^2 + \sum_{i=s+1}^n c_i^2
 \end{aligned}$$

Como estamos minimizando una suma de términos positivos, queremos hacer tantos 0 como sea posible. Esto es, tomar $y_i = \frac{c_i}{\sigma_i}$ para $i = 1, \dots, s$ y y_i arbitraria para $i = s+1, \dots, n$ (podríamos tomar $y_i = 0$ para $i = s+1, \dots, n$ por una cuestión de que nos quede una solución de norma pequeña). Una vez hallado el \hat{y} que minimiza esa expresión, volvemos a la variable original x haciendo $\hat{x} = V\hat{y}$. ¿Cómo se escribe todo este proceso matricialmente?

Fijense que lo que terminamos obteniendo fue para $i = 1, \dots, s$

$$\begin{aligned}
 y_i &= \frac{c_i}{\sigma_i} \\
 &= \frac{u_i^T b}{\sigma_i} \\
 &= \frac{1}{\sigma_i} u_i^T b
 \end{aligned}$$

Es decir, matricialmente

$$(\hat{y})_{i=1, \dots, s} = \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_s} \end{pmatrix}_{s \times s} \begin{pmatrix} - & u_1^T & - \\ \vdots & \vdots & \vdots \\ - & u_s^T & - \end{pmatrix}_{s \times m} b$$

Y aparte completamos con ceros, es decir:

$$\hat{y} = \begin{pmatrix} \frac{1}{\sigma_1} & & & & \\ & \frac{1}{\sigma_2} & & & \\ & & \ddots & & \\ & & & \frac{1}{\sigma_s} & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}_{n \times m} \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{n \times m} \begin{pmatrix} - & u_1^T & - \\ \vdots & \vdots & \vdots \\ - & u_s^T & - \\ \vdots & \vdots & \vdots \\ - & u_m^T & - \end{pmatrix}_{m \times m} b$$

Luego hicimos $\hat{x} = V\hat{y}$ volviendo a la variable original.

Finalmente, llegamos entonces a

$$\hat{x} = V\Sigma^\dagger U^T = A^\dagger b$$

Es decir, que **¡La pseudoinversa resuelve el problema de mínimos cuadrados!**

Y aparte, recordando que si quitamos valores singulares pequeños, no perdemos mucha información al respecto de la matriz, podemos en lugar de considerar la matriz original A que tiene condición

$$\text{cond}_2(A) = \frac{\sigma_1(A)}{\sigma_n(A)}$$

considerar a la matriz de mejor aproximación de rango k , para algún $k < n$ que nos sea pertinente para el problema, que sería

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^T$$

y tendría una condición de

$$\text{cond}_2(A_k) = \frac{\sigma_1(A)}{\sigma_k(A)} \leq \frac{\sigma_1(A)}{\sigma_n(A)} = \text{cond}_2(A)$$

Consiguiendo entonces resolver el problema de mínimos cuadrados con un condicionamiento aún mejor que el original de A .

Como comentario al margen, ya que no es un contenido del curso, el truncado de SVD es tan solo uno de montones de métodos de *regularización* que existen, que se abocan a estabilizar las soluciones y evitar la amplificación de errores por ruido en los datos.

En particular, es un método de filtrado espectral, que son los que intentan amortiguar el efecto de las componentes asociadas a valores singulares pequeños de la descomposición SVD de la matriz, utilizando un filtro Φ_i y obteniendo

$$\hat{x} = \sum_{i=1}^n \Phi_i \frac{1}{\sigma_i} u_i^T b v_i \text{ con } \Phi_i = \begin{cases} 1 & \text{para valores pequeños de } i \\ 0 & \text{para valores grandes de } i \end{cases}$$

Como conclusión, tenemos la siguiente tabla, que resume un poco todo para intentar ordenar las ideas

Método	Formación	Resolución	Complejidad Total	Estabilidad / Sensibilidad
Ecuaciones normales	$A^T A$ y $A^T b$: $\mathcal{O}(mn^2)$	$\mathcal{O}(n^3)$	$\mathcal{O}(mn^2 + n^3)$	$\text{cond}_2(A^T A) = \text{cond}_2(A)^2$
QR	$A = QR$: $\mathcal{O}(mn^2)$	$\mathcal{O}(n^2)$	$\mathcal{O}(mn^2)$	$\text{cond}_2(R) = \text{cond}_2(A)$
SVD completa	$A = U \Sigma V^T$: $\mathcal{O}(mn^2)$	$\mathcal{O}(mn)$	$\mathcal{O}(mn^2)$	$\text{cond}_2(A) = \frac{\sigma_1}{\sigma_n}$
SVD truncado en k	$A \approx U_k \Sigma_k V_k^T$: $\mathcal{O}(mnk)$	$\mathcal{O}(mk)$	$\mathcal{O}(mnk)$	$\text{cond}_2(A_k) = \frac{\sigma_1}{\sigma_k} \leq \text{cond}_2(A)$

Comparación de métodos de resolución de mínimos cuadrados. Se asume $A \in \mathbb{R}^{m \times n}$ con $m \gg n$.

Creo que no está de más decir, que a pesar de que la discusión de qué es más estable para una máquina es vital en una materia en la que nos abocamos a aprender algoritmos que ejecutará una computadora, para un ser humano resolviendo una instancia pequeña del problema de mínimos cuadrados como hicimos en el ejercicio 2, probablemente la manera más concisa de hacerlo sea precisamente como lo hicimos, utilizando las ecuaciones normales.

Interpolación polinomial

Supongamos que nos dan $n+1$ puntos $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ con $x_i \in \mathbb{R}$ todos distintos entre sí, y queremos encontrar un polinomio p de grado a lo sumo n tal que

$$p(x_i) = y_i \quad \text{para } i = 0, 1, \dots, n.$$

Este problema se conoce como *interpolación polinomial*. Lo que buscamos entonces es determinar los coeficientes c_0, c_1, \dots, c_n del polinomio

$$p(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n,$$

de manera tal que se cumpla $p(x_i) = y_i$ para cada $i = 0, \dots, n$.

Si escribimos esta condición para cada uno de los puntos, obtenemos un sistema de $n+1$ ecuaciones con $n+1$ incógnitas:

$$\begin{cases} c_0 + c_1x_0 + c_2x_0^2 + \dots + c_nx_0^n = y_0 \\ c_0 + c_1x_1 + c_2x_1^2 + \dots + c_nx_1^n = y_1 \\ \vdots \\ c_0 + c_1x_n + c_2x_n^2 + \dots + c_nx_n^n = y_n \end{cases}$$

Esto puede reescribirse como un sistema lineal de la forma

$$Ac = y,$$

donde

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}, \quad c = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix}, \quad y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

La matriz A que aparece en este sistema se conoce como la **matriz de Vandermonde** asociada a los puntos x_0, \dots, x_n . Esta matriz tiene una propiedad muy importante, que queda como ejercicio para el lector interesado: **si los x_i son todos distintos, entonces A es invertible** (ver el final del documento para más detalles). Esto significa que el sistema tiene una única solución, y por lo tanto, existe un único polinomio de grado a lo sumo n que interpola los $n+1$ puntos dados.

Conclusión: Encontrar un polinomio que interpola un conjunto de puntos equivale a resolver el sistema lineal $Ac = y$, donde A es la matriz de Vandermonde construida con las abscisas de los puntos (es decir, los x_i y y es el vector de las ordenadas (los y_i)).

Ejercicio 5

Hallar el polinomio de grado a lo sumo 2 que interpola los datos:

$$(1, 2), \quad (2, 3), \quad (3, 5)$$

Solución: En este caso, tenemos $n = 2$, y por lo tanto buscamos un polinomio $p(x) = c_0 + c_1x + c_2x^2$ tal que $p(1) = 2$, $p(2) = 3$ y $p(3) = 5$.

La matriz de Vandermonde y los vectores son:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Resolvemos el sistema $A\mathbf{c} = \mathbf{y}$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Resolviendo el sistema por cualquiera de los métodos vistos en el curso, se obtiene:

$$c_0 = 2, \quad c_1 = -\frac{1}{2}, \quad c_2 = \frac{1}{2}$$

Por lo tanto, el polinomio que interpola los datos es:

$$p(x) = 2 - \frac{1}{2}x + \frac{1}{2}x^2$$

Como mencioné antes, este método puede extenderse a cualquier conjunto de $n + 1$ puntos con abscisas distintas (x_i todos distintos), y el sistema lineal resultante siempre tiene solución única, ya que la matriz de Vandermonde es invertible en ese caso. Para ver eso, dejo como ejercicio al que le interese ver que

Ejercicio 6

Dada una matriz de Vandermonde A generada por una secuencia de n números reales $\alpha_1, \dots, \alpha_n$, es decir

$$A = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

entonces

$$\det(A) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$$

Observación

Quien demuestre esto, demostrará que la matriz de Vandermonde será invertible siempre que los x_i sean todos distintos.

A su vez, si quisiéramos un polinomio de grado k que “intente” interpolar a los $n + 1$ puntos, entonces lo que deberíamos plantear es análogo a lo anterior, pero esta vez de la matriz de Vandermonde nos quedaremos hasta la $(k + 1)$ -ésima columna, es decir, tendríamos

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^k \\ 1 & x_1 & x_1^2 & \cdots & x_1^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{pmatrix} \in \mathbb{R}^{(n+1) \times (k+1)}, \quad \mathbf{c} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_k \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Y si se fijan, al plantear el sistema

$$A\mathbf{c} = \mathbf{y}, \quad A \in \mathbb{R}^{(n+1) \times (k+1)}, \mathbf{c} \in \mathbb{R}^{(k+1)}, \mathbf{y} \in \mathbb{R}^{(n+1)}, \quad (k + 1) < (n + 1)$$

estamos en la misma situación del principio de la clase, cuando nos enfrentamos a problemas con más ecuaciones que incógnitas, y por ende este es un problema que se puede solucionar con cualquiera de los métodos de cuadrados mínimos que vimos.

Comentario final: Las consultas, dudas, sugerencias son súper bienvenidas a fmartinez@dm.uba.ar o al campus de la materia.