Datos Masivos II

Descomposición de valores singulares

Blanca Vázquez-Gómez y Gibran Fuentes-Pineda

Septiembre 2020

Un poco de historia

- La descomposición de valores singulares (SVD) fue propuesta por Beltrani (1873) y por Jordan (1874).
- Su generalización en el contexto de ecuaciones integrales fue por Smith (1907) y por Weyl (1912).
- En la década de los 60 y 70, cuando consigue popularidad para el tratamiento de imágenes (métodos numéricos).

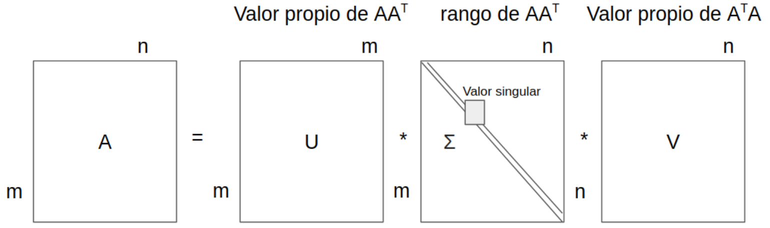
- Una matriz $A \in C^{m,m}$ es unitaria si sus columnas forman una base ortonormal de vectores C^m
- Una matriz $A \in R^{m,m}$ es ortogonal si sus columnas forman una base ortogonal de vectores R^m
- La inversa de una matriz unitaria es igual a su transpuesta

Dada una matriz A de m x n, una descomposición de valores singulares de A es una factorización del tipo:

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

Dónde:
U es una matriz ortogonal de tamaño m x m
Σ es una matriz diagonal de tamaño m x n
V es una matriz ortogonal de tamaño n x n

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


Valor propio de AA^T
rango de AA^T
Valor propio de A^TA
n
m
n
n
A
=
U
*
Σ
Valor singular
*
V
m
m
m
n

- Los elementos de la matriz Σ se les conoce como los valores singulares de A, se denotan por (σ_i)
- Los vectores $u_i, .., u_m$ que conforman la matriz U se les conoce como vectores singulares de A por la izquierda.
- Los vectores $v_i, .., v_m$ que conforman la matriz V se les conoce como vectores singulares de A por la derecha.
- Las matrices U y V son unitarias

Propiedades de los SVD

- Toda matriz de tamaño m x n puede ser descompuesta en valores singulares
- El rango de A es la cantidad de valores singulares de A distintos de 0.
- El valor absoluto del determinante de A será igual a $(\sigma_1, ..., \sigma_n)$

- Todo elemento de la diagonal σ_i cumple:

$$Au_i = \sigma_i v_i$$
$$A^T v_i = \sigma_i u_i$$

Siendo u_i la columna U correspondiente a σ_i
Siendo v_i la columna V correspondiente a σ_i

Propiedades de los SVD

Si $A \in C^{nxn}$ es invertible y sean $\sigma_i \geq .... \geq \sigma_n$ los valores singulares. Por lo tanto, los valores singulares de A^{-1} son:

$$\frac{1}{\sigma_n} \geq ... \geq \frac{1}{\sigma_1}$$

- Sea $A = U\Sigma V^T$ una descomposición de A en valores singulares y rango $(A) = r$, entonces:

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

u y v son los vectores columna de U y V respectivamente.

- La descomposición se puede representar como la suma de r matrices de rango 1.
- El término $\sigma_i u_i v_i$ se conoce también como tripleta singular.
- El rango de la matriz A, nos da el máximo número de tripletas necesarias.

- Encontrar la descomposición de valores singulares de A

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

Paso 1: para obtener las matrices U, Σ y V, es necesario calcular:

$$AA^T$$
$$A^TA$$

Posteriormente, se deben calcular los valores propios para cada una de estas matrices.

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

Por lo tanto:

$$AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

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

Por lo tanto:

$$A^T A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

Cálculo de la matriz Σ

Dada la matriz AA^T calcular los valores propios:

$$AA^T = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

Recordemos, cómo se hace la descomposición propia de una matriz:

$$\begin{gathered} Av = \lambda v \\ Av = (\lambda I)v \\ Av - (\lambda I)v = 0 \\ A - (\lambda I)v = 0 \end{gathered}$$

Cálculo del determinante de AA^T

$$\left| \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\left| \begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix} \right| = 0$$

$$|A - (\lambda I)| = (17-\lambda)(17-\lambda) - 64 = 0$$
$$|A - (\lambda I)| = 289 - 17\lambda - 17\lambda + \lambda^2 - 64 = 0$$
$$|A - (\lambda I)| = \lambda^2 - 34\lambda + 225 = 0$$

Usamos la función cuadrática para resolver la ecuación:

$$\lambda^2 - 34\lambda + 225 = 0$$

$$\lambda = \frac{-(-34) \pm \sqrt{(-34)^2 - 4(1 \cdot 225)}}{2 \cdot 1}$$

Los valores propios quedan:

$$\lambda_1 = \frac{34 + 16}{2} = 25$$

$$\lambda_2 = \frac{34 - 16}{2} = 9$$

Los valores singulares de la matriz Σ son las raíces cuadradas de los valores propios:

$$\sigma_i = \sqrt{\lambda_i}$$

Entonces:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{25} = 5$$
$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{9} = 3$$

Dado que:

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

La matriz Σ es una matriz diagonal de tamaño de A:

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

Para obtener U, es necesario calcular los vectores propios de la matriz de AA^T:

Para $\lambda_1 = 25$

$$\begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 25 \begin{bmatrix} x \\ y \end{bmatrix}$$

Se obtiene el sistema de ecuaciones:

$$\begin{aligned} 17x + 8y &= 25x \\ 8x + 17y &= 25y \end{aligned}$$

Calculando el vector propio de λ_1

Resolvemos el sistema de ecuaciones:

$$17x + 8y = 25x$$
$$8x + 17y = 25y$$

Obteniendo un posible vector propio de:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Calculamos el vector unitario: $1^2 + 1^2 = 2$

Calculando el vector propio de λ_1

El vector propio cuando $\lambda_1 = 25$ es:

$$\begin{bmatrix} 1/\sqrt{2} \\ 1\sqrt{2} \end{bmatrix}$$

Calculando el vector propio de λ_2

Para $\lambda_2 = 9$

$$\begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 9 \begin{bmatrix} x \\ y \end{bmatrix}$$

Se obtiene el sistema de ecuaciones:

$$17x + 8y = 9x$$
$$8x + 17y = 9y$$

Calculando el vector propio de λ_2

Resolvemos el sistema de ecuaciones:

$$17x + 8y = 9x$$
$$8x + 17y = 9y$$

Obteniendo un posible vector propio de:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Calculamos el vector unitario: $1^2 + -1^2 = 2$

Calculando el vector propio de λ_2

El vector propio cuando $\lambda_2 = 9$ es:

$$\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

2do resultado: matriz U

Los vectores propios para $\lambda_1 = 25$ y $\lambda_2 = 9$ son:

$$u_1 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, u_2 \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Por lo tanto,

$$U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Recordemos la matriz A^TA

Por lo tanto:

$$A^TA = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

Dada la matriz A^TA calcular el determinante:

$$\left| \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right| = 0$$

$$\left| \begin{bmatrix} 13-\lambda & 12 & 2 \\ 12 & 13-\lambda & -2 \\ 2 & -2 & 8-\lambda \end{bmatrix} \right| = 0$$

Dada la matriz A^TA calcular el determinante:

$$\left|\begin{bmatrix} 13-\lambda & 12 & 2 \\ 12 & 13-\lambda & -2 \\ 2 & -2 & 8-\lambda \end{bmatrix}\right| = 0$$

$|A-(\lambda I)| = 13-\lambda((13-\lambda)(8-\lambda)-4)-12(12(8-\lambda)-(-4))-2(2(13-\lambda)-(-24))$
$|A-(\lambda I)| = -x^3+34\lambda^2-225\lambda$
Factorizamos la ecuación y se obtiene: $-\lambda(\lambda^2-34\lambda+225)=0$

Dada la ecuación:

$$\lambda^2 - 34\lambda + 225 = 0$$

se calculan los valores propios usando la función cuadrática:

Los valores propios quedan:

$$\lambda_1 = \frac{34 + 16}{2} = 25$$

$$\lambda_2 = \frac{34 - 16}{2} = 9$$

$$\lambda_3 = 0$$

3er resultado: matriz V^T

Los vectores propios para $\lambda_1 = 25$, y $\lambda_2 = 9$ y $\lambda_3 = 0$ son:

$$v_1 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} v_2, \begin{bmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix} v_3, \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix}$$

Por lo tanto,

$$V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

Dada la matriz A,

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

La descomposición de valores singulares es:

$$A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

- SVD es la descomposición de una matriz en 3 matrices: U, Σ, V
- Σ es una matriz diagonal compuesta de los valores singulares
    - Los valores singulares son los valores importantes en una matriz, nos indica la fuerza de cada concepto.
- Los vectores propios de una matriz son direcciones de máxima dispersión o varianza de datos
- Es un método de reducción de dimensionalidad

Compresión de imágenes

- Trasmitir / enviar imágenes por medios electrónicos
- *¿Cuál es la cantidad mínima de información que se debe almacenar?* (Sin perder información valiosa y que ahorre espacio).

Supongamos que una matriz A de mxn representa los tonos de gris de la imagen de tamaño mxn pixeles.

Cada pixel tiene un número asociado $(0 - 255)$ indicando el tono del gris.

Por lo tanto, la matriz A en cada posición i, j describirá el valor del tono gris (pixel) de la imagen.

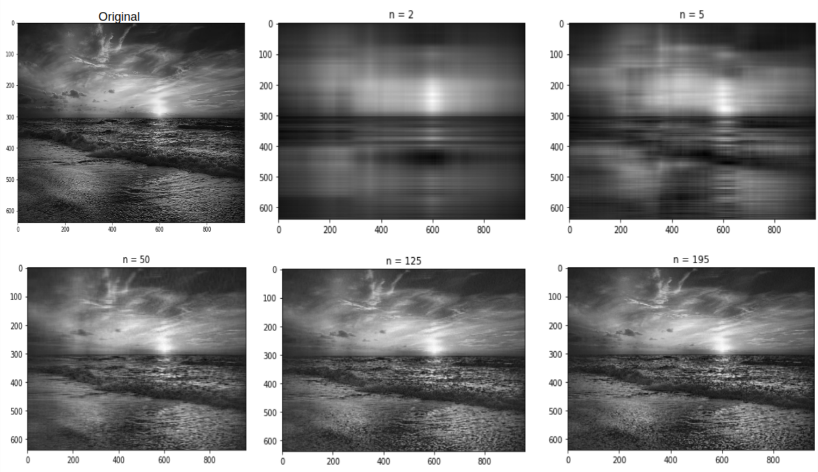
$$\begin{bmatrix} 23. & 23. & 23. & \dots & 19. & 19. & 19. \\ 23. & 23. & 23. & \dots & 19. & 19. & 19. \\ 23. & 23. & 23. & \dots & 19. & 19. & 19. \\ \dots & & & & & & \\ 39. & 29. & 24. & \dots & 45. & 47. & 40. \\ 25. & 25. & 29. & \dots & 7. & 14. & 17. \\ 23. & 27. & 32. & \dots & 21. & 22. & 19. \end{bmatrix}$$

Dada esta matriz, es posible calcular la SVD.

Compresión de imágenes

- Al generar la SVD, observaremos que los valores singulares más pequeños (Σ) provienen de las partes de la imagen con menor interés.
- Entre menor número de valores singulares, la imagen puede ser poco nítida.

Compresión de imágenes



Ventajas en la compresión de imágenes

- Una imagen contiene información redundante, que puede ser eliminada sin afectar a la información importante.
- Observamos que es posible descomponer una imagen en una matriz y calcular sus SVD.
- A partir de los valores singulares, es posible reconstruir la imagen original
- Mientras mayor sea k, mayor será la calidad y menor la compresión.

Para aproximar cuanto espacio se reduce una imagen usando SVD, se usa la siguiente relación de compresión:

$$r = (n + m + 1)k \;/\; nm$$

Dónde:
n y m es el tamaño de la imagen
k es el número de valores singulares a usar

Una imagen de 480 x 640 pixeles está compuesta de 307,200 puntos $\approx$ 0.3 MB, calcular la relación de compresión, si únicamente usamos 50 valores singulares para reconstruir la imagen:

$$r = (n + m + 1)k \; / \; nm$$
$$r = (480 + 640 + 1)50 \; / \; 480*640$$
$$r = 0.18$$

Una imagen reconstruida con 50k, únicamente requiere el 18 % de la información original $\approx$ 0.05 MB

Resolución de mínimos cuadrados

- Se requiere resolver $Ax = b$
- Se multiplica en ambos lados de la ecuación por A^T y se obtiene: $A^TA = A^Tb$
- Se calcula SVD de $A = U\Sigma V^T$

Cálculo de la inversa

- La inversa de una matriz A, puede calcular de la forma: $A^{-1} = VS^{-1}U^T$

Recordemos que: U y V son matrices ortogonales y S una matriz diagonal

Reducción de dimensionalidad en datos dispersos

- Los datos dispersos se refiere cuando tenemos filas de datos, con muchos valores en 0

$$\begin{bmatrix} 0. & 0. & 0. & ... & 3 & 3 & 3 \\ 0. & 0. & 0. & ... & 3 & 3 & 3 \\ 0. & 0. & 0. & ... & 3 & 3 & 3 \\ ... & & & & & & \\ 0. & 0. & 2. & ... & 0. & 1. & 1. \\ 0. & 2. & 3. & ... & 0. & 1. & 2 \end{bmatrix}$$

Reducción de dimensionalidad en datos dispersos
Ejemplos de información con datos dispersos, en donde es posible aplicar SVD:

- Sistemas de puntajes (películas, canciones)
- Bolsa de palabras
- One-hot encoding

Otras aplicaciones de la SVD

- Reducción de ruido
- Mínimos cuadrados
- Cálculos estadísticos
- Procesamiento de señales