

DATOS MASIVOS II

DESCOMPOSICIÓN DE VALORES SINGULARES

Blanca Vázquez-Gómez y Gibran Fuentes-Pineda

11 de agosto de 2022

- La descomposición de valores singulares (SVD) fue propuesta por Beltrani (1873) y por Jordan (1874).
- Fue hasta 1907 por Smith y en 1912 por Weyl que fue planteado como una generalización de la descomposición propia de matrices.
- En la década de los 60 y 70, cuando consigue popularidad para el tratamiento de imágenes.

Dada una matriz A de $m \times n$, una descomposición de valores singulares de A es una **factorización** del tipo:

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

Dónde:

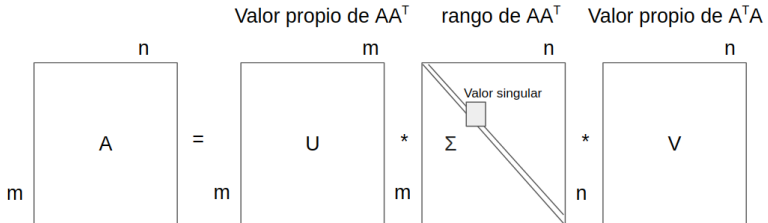
U es una matriz ortogonal de tamaño $m \times m$

Σ es una matriz diagonal de tamaño $m \times n$

V es una matriz ortogonal de tamaño $n \times n$

DESCOMPOSICIÓN DE VALORES SINGULARES

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



DESCOMPOSICIÓN DE VALORES SINGULARES

- Los elementos de la matriz Σ se les conoce como los **valores singulares de A**, se denotan por (σ_i)
- Los vectores u_i, \dots, u_m que conforman la matriz U se les conoce como **vectores singulares de A por la izquierda**.
- Los vectores v_i, \dots, 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
- El término $\sigma_i u_i v_i$ se conoce también como tripleta singular.

Compresión de imágenes

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

- 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

RELACIÓN DE COMPRESIÓN

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

$$\begin{aligned} r &= (n + m + 1)k / nm \\ r &= (480 + 640 + 1)50 / 480 \cdot 640 \\ r &= 0.18 \end{aligned}$$

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

TRUNCAMIENTO DE SVD

- En la práctica puede ser muy costoso el cómputo de la versión completa de SVD
- En la versión truncada, solo se calculan las k vectores columna de \mathbf{U} y los k vectores fila de \mathbf{V} asociados a los k valores singulares más grandes en Σ
 - \mathbf{U}_k es de tamaño $m \times k$
 - \mathbf{V}_k es de tamaño $k \times n$
 - Σ_k es de tamaño $k \times k$
- La versión truncada es una aproximación de SVD

$$\tilde{\mathbf{A}} \approx \mathbf{U}_k \Sigma_k \mathbf{V}_k^T$$