SVD

Singular Value Decomposition

(y otras yerbas)

¿Que es?

La SVD es un método de descomposición de matrices, nos permite descomponer **cualquier** matriz A en tres matrices de la forma:

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

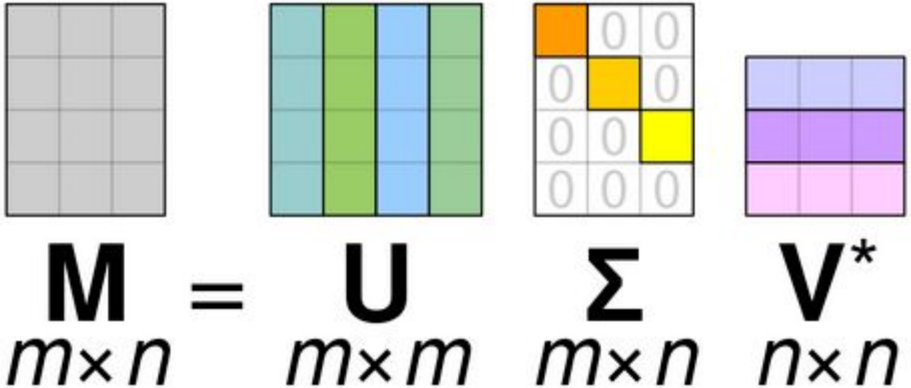
La descomposición

Siendo A una matriz de $m \times n$

U es de $m \times m$

Σ es de $m \times n$

V es de $n \times n$



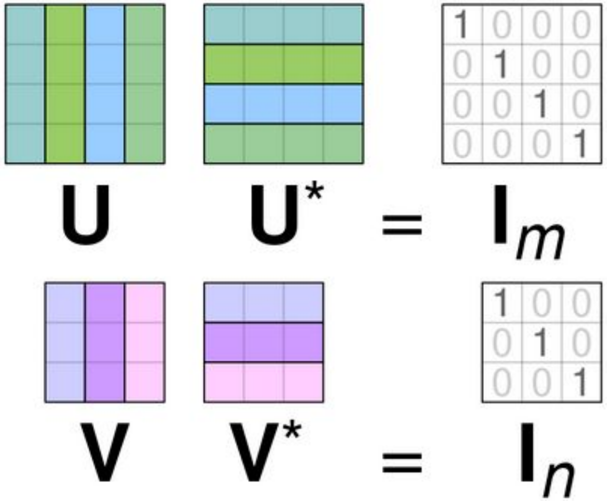
Cómo obtener la descomposición

- Las columnas de U son los autovectores de AA^T
- Las columnas de V son los autovectores de $A^T A$
- Σ es una matriz diagonal que contiene la raíz de los autovalores de $A^T A$ (o de AA^T) ordenados en forma decreciente
- Los valores de Σ son llamados *valores singulares*

Matrices Unitarias

U y V son matrices unitarias, esto significa que:

$UU^T = U^TU = I$ y que $V^TV = VV^T = I$



Esto nos permite hacer cosas como:

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

Multiplicando a ambos lados por V

$$AV = U\Sigma V^T V = U\Sigma I = U\Sigma$$

→ Si *A* contiene algún tipo de datos, se dice que las columnas de V son los *ejes principales* de los datos, y que las columnas de AV (o de $U\Sigma$) son las ***componentes principales*** de estos, más sobre esto próximamente.

SVD Reducida/Truncada

Si la matriz Σ contiene ceros en su diagonal o no es cuadrada podemos usar una representación más *compacta* de la SVD:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

SVD Reducida/Truncada

Si la matriz Σ contiene ceros en su diagonal o no es cuadrada podemos usar una representación más *compacta* de la SVD:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & \cancel{0} \\ -1 & 0 & 0 & \cancel{0} \\ 0 & 0 & 0 & \cancel{-1} \\ 0 & 0 & -1 & \cancel{0} \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

SVD Reducida/Truncada

Si la matriz Σ contiene ceros en su diagonal o no es cuadrada podemos usar una representación más *compacta* de la SVD:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & \cancel{0} \\ -1 & 0 & 0 & \cancel{0} \\ 0 & 0 & 0 & \cancel{-1} \\ 0 & 0 & -1 & \cancel{0} \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & \cancel{0} & \cancel{0} \\ 0 & \sqrt{5} & 0 & \cancel{0} & \cancel{0} \\ 0 & 0 & 2 & \cancel{0} & \cancel{0} \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \\ 0 & -1 & 0 & 0 & 0 \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{1} & \cancel{0} \\ \cancel{-\sqrt{0.8}} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{\sqrt{0.2}} \end{bmatrix}$$

SVD Reducida/Truncada

En este caso se dice que nos quedamos con la SVD ***reducida***

Notar que la matriz resultante sigue siendo la misma

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \\ 0 & -1 & 0 & 0 & 0 \end{bmatrix}$$

SVD Reducida/Truncada

¿Que pasa si no nos conformamos con eliminar los valores singulares iguales a 0 y vamos más allá?

Probemos eliminar el 2:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & \cancel{0} \\ -1 & 0 & \cancel{0} \\ 0 & 0 & \cancel{0} \\ 0 & 0 & \cancel{-1} \end{bmatrix} \begin{bmatrix} 3 & 0 & \cancel{0} \\ 0 & \sqrt{5} & \cancel{0} \\ \cancel{0} & \cancel{0} & \cancel{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \\ \cancel{0} & \cancel{-1} & \cancel{0} & \cancel{0} & \cancel{0} \end{bmatrix}$$

SVD Reducida/Truncada

En este caso la matriz resultante ya no es la misma, pero si la observamos detenidamente es muy parecida a la original.

Se dice que tenemos una aproximación de rango 2(porque nos quedamos con 2 valores singulares) de la matriz original.

Se suele llamar a esto SVD ***truncada***[1]

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ -\sqrt{0.2} & 0 & 0 & 0 & -\sqrt{0.8} \end{bmatrix}$$

1: En español no se usa mucho el término pero en inglés es un algoritmo muy usado(ver Truncated SVD).

Aproximaciones de rango r de una matriz

Es un problema de optimización que consiste en encontrar una matriz $\hat{A}$ de menor rango(cantidad de autovalores distintos de 0) que la matriz original A que minimice la diferencia*:

$$||A - \hat{A}||$$

*Habiendo definido una norma para comparar las matrices(espectral, frobenius)

Teoremon (Eckart-Young)

La SVD truncada nos da la **mejor** aproximación de rango **r** de una matriz

Para obtenerla hacemos lo mismo que antes, obtenemos la SVD, nos quedamos con los **r** valores singulares más grandes y volvemos a multiplicar para obtener la matriz

Aplicaciones

Muy lindo todo, pero, ¿para que me sirve esto?

Compresión de imágenes

Si consideramos a la imagen como una matriz de píxeles, podemos aplicarle la SVD, y quedarnos con los **r** valores singulares más grandes(aproximación de rango **r**), mientras más chico r mayor será la compresión.

Almacenar estas matrices truncadas $\mathbf{U}_r$, Σ_r y $\mathbf{V}_r$ nos puede resultar menos pesado en disco que almacenar la matriz original.

Sin embargo, es una compresión con pérdida, estamos descartando información importante para la imagen, ¿que pasa si ahora intentamos reconstruir la imagen?

Compresión de imágenes

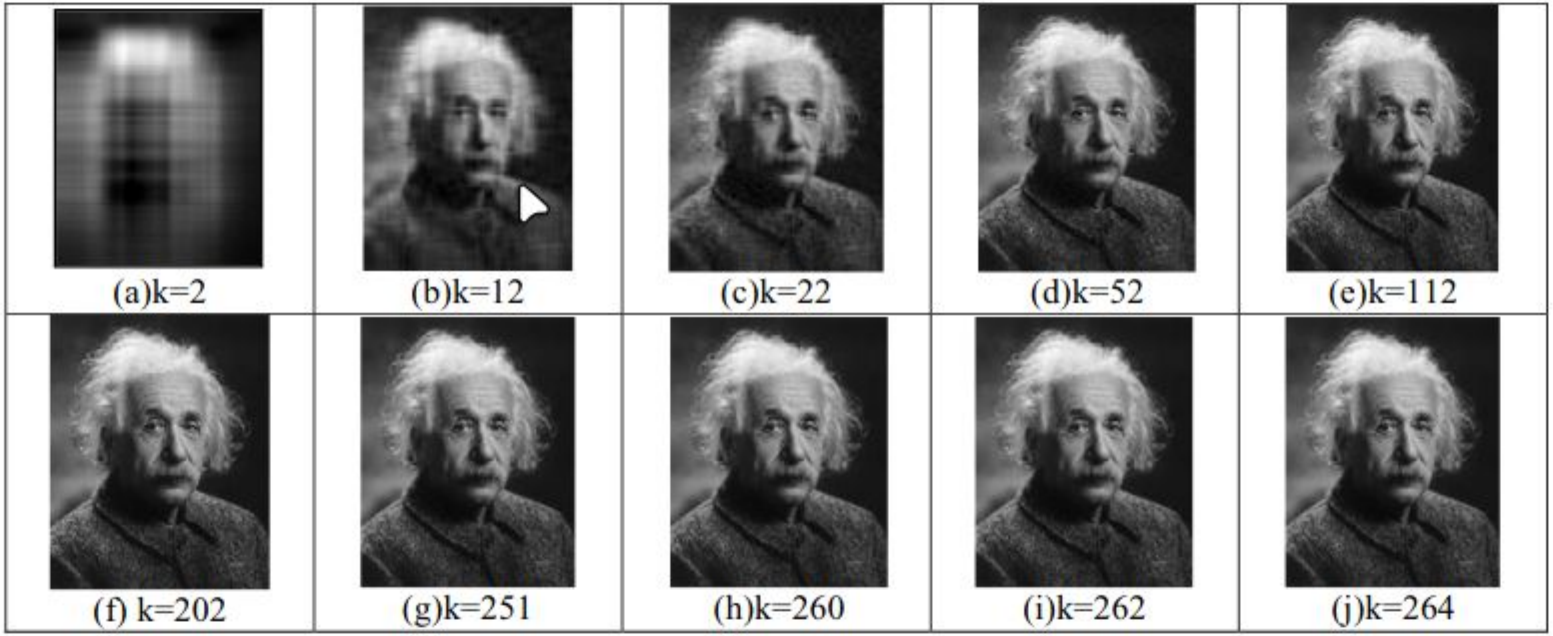


Fig 3. SVD compression for different Singular Values'k'

Compresión de imágenes

Pagina muy buena donde se muestra un ejemplo interactivo sobre esto:

http://timbaumann.info/svd-image-compression-demo/

Otras aplicaciones

- Reducción de dimensiones de un dataset
- Information retrieval, Analisis Semantico Latente(LSA)
- Sistemas de recomendación
- Procesamiento de señales
- Motores de busqueda (ver CubeSVD p.ej)

Y muchas mas

Referencias y otras lecturas

Lectura de G. Strang:

- https://www.youtube.com/watch?v=mBcLRGuAFUk

Algo de intuición sobre lo que hace la SVD:

- https://gregorygundersen.com/blog/2018/12/10/svd/

Explicación con algo de código:

- https://ethen8181.github.io/machine-learning/dim_reduct/svd.html