

ACÀMICA

¡Bienvenidas/os a Data Science!



Agenda

¿Cómo anduvieron?

Debate: Algoritmos que cautivan II

Hands-On: Notebook 38

Puesta en común Notebook 38

Break

Explicación: SVD

Hands-On: Notebook 39

Cierre



¿Cómo anduvieron?



Algoritmos que cautivan



**“La libertad es lo
contrario a la
compulsión”**

- Byung-Chul Han -



Sistemas de recomendación

La idea es usar es la tecnología juegue a nuestro favor.

Usemos las herramientas que nos ofrece en función de que nos ayude a **alcanzar nuestros objetivos.**

Teniendo **en claro** que esto también puede ser una forma de **gobierno de nuestras voluntades**



Siendo la **VOLUNTAD** una facultad humana que intersecta la

INTELIGENCIA RACIONAL

+


INTELIGENCIA EMOCIONAL

¿Cómo puedo ser dueño de elegir lo que capta mi atención?



1. Estar en el MOMENTO PRESENTE



A person is walking away from the viewer through a vast field of tall, golden-brown grass. The sun is low on the horizon, creating a warm, orange glow that fills the sky and the field. The sky is filled with soft, white clouds. The overall mood is peaceful and contemplative.

EL PODER PARA CREAR UN
MEJOR FUTURO ESTÁ
CONTENIDO EN EL MOMENTO
PRESENTE: CREAS UN BUEN
FUTURO CREANDO UN BUEN
PRESENTE.

ECKHART TOLLE



2. Aplicar:

LA AUTENTICIDAD



3. FILTRAR CON LA PREGUNTA CONSTANTE:

Esto: “¿Me lleva a mi
mejor versión?”



Hands-on INTELIGENCIA EMOCIONAL Caso de Salvador Valdés de la Fuente





Emprender una segunda etapa profesional en el mundo de hoy



Acá...

Follow

Jan 31 · 7 min...



por Salvador Valdés de la Fuente

Egresado de Data Science en Acámica



**Tip número 1: préstale
atención a tus sensaciones y
analízalas. La respuesta
siempre está dentro tuyo.**



Tip número 2: Confía en ti.
Vas a tener miedos, puedes equivocarte. Pero dale lugar a tu instinto y a eso que te mueve y te entusiasma, y síguelo con convicción.



**Tip número 3: préstale
atención a tus sensaciones y
analízalas.**

**La respuesta siempre está
dentro tuyo.**



Hands-on training

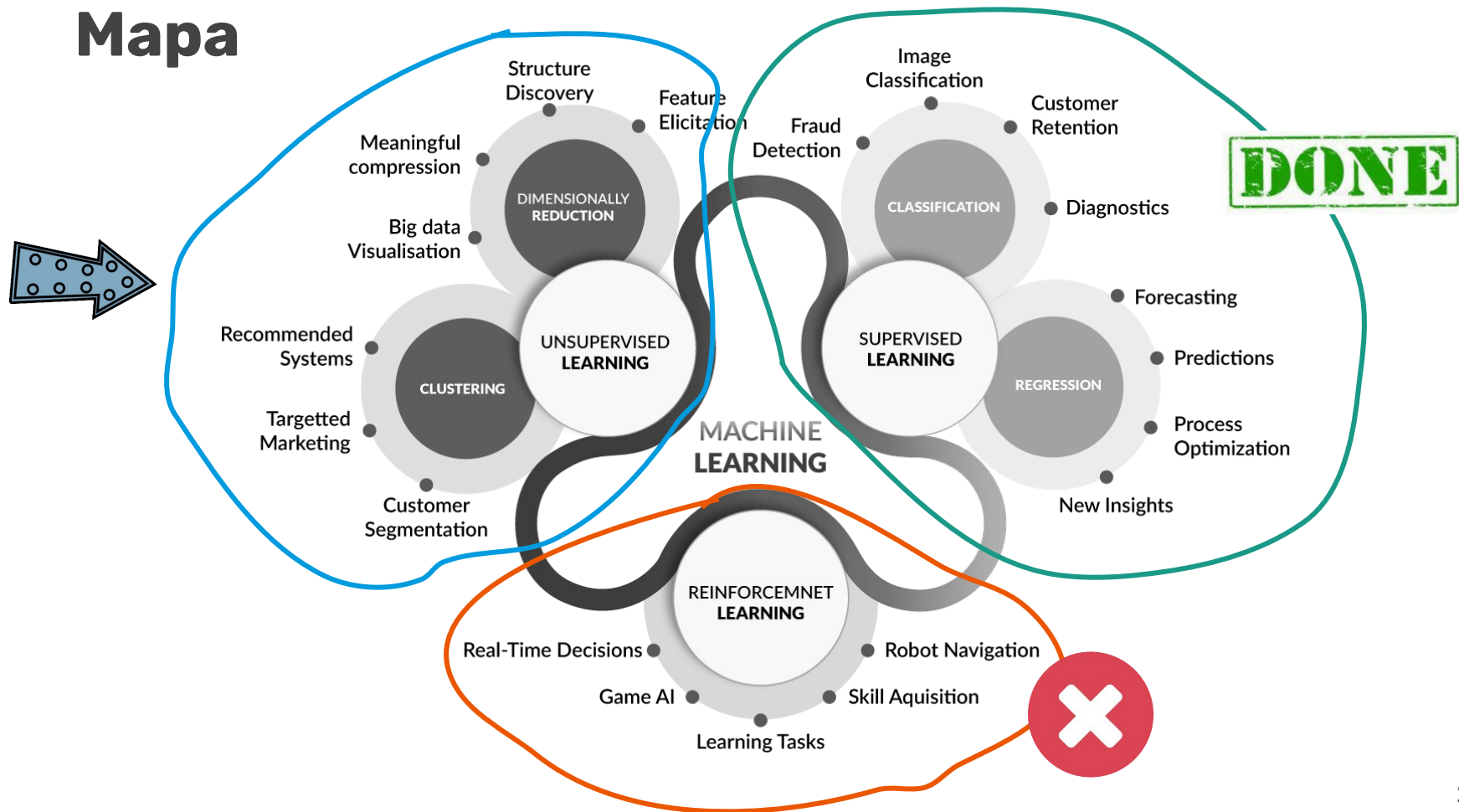
DS_Encuentro_38_Metricas.ipynb



Repaso: Aprendizaje no supervisado



Mapa



Solo datos

Llamamos **Aprendizaje No Supervisado** a los métodos para trabajar con datos (instancias) que no tienen asociados una etiqueta (una clase o un valor).

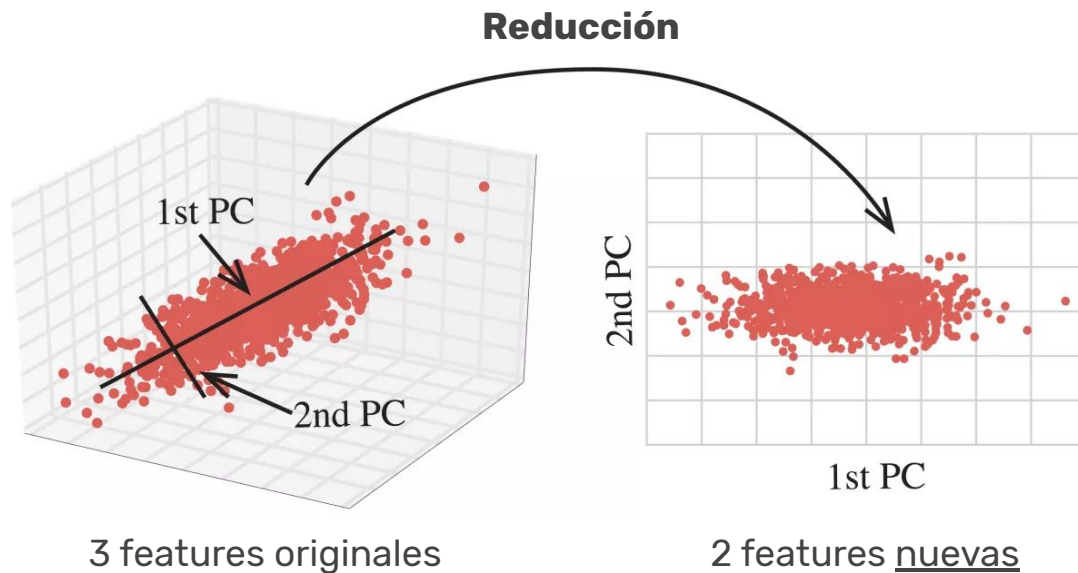
Los objetivos principales en Aprendizaje No Supervisado son:

- Clustering
- Reducción de dimensionalidad

- Clustering
- Reducción de dimensionalidad

Aprendizaje No Supervisado • Reducción de la dimensionalidad

Buscamos reducir la cantidad de features de un dataset, pero reteniendo la mayor cantidad de “información” posible.



Aprendizaje No Supervisado • Reducción de la dimensionalidad

¿Para qué sirve?

Reducir la cantidad de features en un dataset puede servir para:

- Reducir el input en un modelo de regresión o clasificación
- Compresión de archivos
- Visualización
- Detectar features relevantes en datasets
- Muchísimas mas cosas

Aprendizaje No Supervisado • Reducción de la dimensionalidad

¿Para qué sirve?

Reducir la cantidad de features en un dataset puede servir para:

- Reducir el input en un modelo de regresión o clasificación
- Compresión de archivos
- Visualización
- Detectar features relevantes en datasets
- Muchísimas mas cosas

¿Cómo se hace?

Algunos de los métodos de reducción de dimensionalidad son:

- PCA: Principal Component Analysis (usa SVD)
- MDS: Multidimensional scaling
- t-SNE: t-distributed Stochastic Neighbor Embedding
- Auto-Encoders (Se hace con Redes Neuronales)
- LDA: Linear Discriminant Analysis (si hay etiquetas de clases)

Aprendizaje No Supervisado • Reducción de la dimensionalidad

¿Para qué sirve?

Reducir la cantidad de features en un dataset puede servir para:

- Reducir el input en un modelo de regresión o clasificación
- Compresión de archivos
- Visualización
- Detectar features relevantes en datasets
- Muchísimas mas cosas

¿Cómo se hace?

Algunos de los métodos de reducción de dimensionalidad son:

- **PCA: Principal Component Analysis (usa SVD)**
- MDS: Multidimensional scaling
- t-SNE: t-distributed Stochastic Neighbor Embedding
- Auto-Encoders (Se hace con Redes Neuronales)
- LDA: Linear Discriminant Analysis (si hay etiquetas de clases)

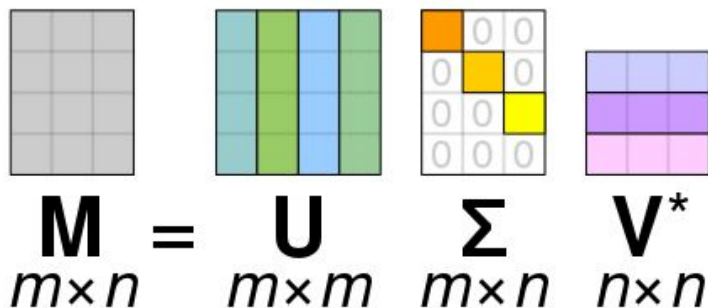
Aprendizaje No Supervisado

SVD (Singular Value Decomposition)



SVD • Definición

Es un método de álgebra lineal que nos permite representar cualquier matriz en términos de la multiplicación de otras 3 matrices.



The diagram illustrates the SVD decomposition of a matrix M into three matrices U , Σ , and V^* . Each matrix is represented by a grid of colored squares:

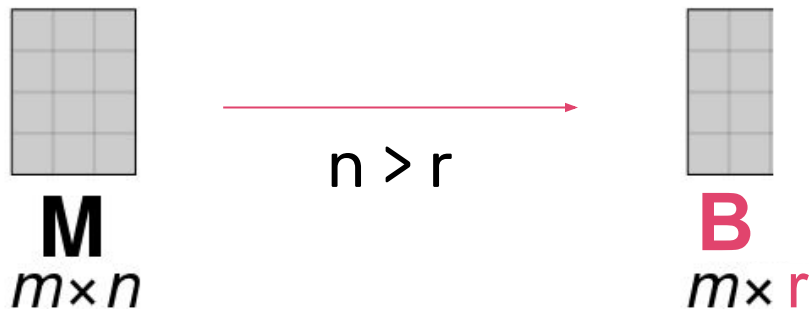
- M is a 4×4 matrix represented by a gray grid.
- U is a 4×4 matrix represented by a grid with four vertical columns of different colors: teal, green, blue, and light green.
- Σ is a 4×4 matrix represented by a grid with a diagonal of colored squares (orange, yellow, yellow, and light yellow) and zeros elsewhere.
- V^* is a 4×4 matrix represented by a grid with four horizontal rows of different colors: light blue, blue, purple, and pink.

The equation is shown as:

$$\begin{matrix} \text{Grid} \\ \mathbf{M} \\ m \times n \end{matrix} = \begin{matrix} \text{Grid} \\ \mathbf{U} \\ m \times m \end{matrix} \begin{matrix} \text{Grid} \\ \mathbf{\Sigma} \\ m \times n \end{matrix} \begin{matrix} \text{Grid} \\ \mathbf{V}^* \\ n \times n \end{matrix}$$

SVD • ¿Para qué sirve?

Para MUCHAS COSAS. Es parte del corazón de muchos algoritmos numéricos (solución sis. lineal, pseudoinversa, etc.). En este contexto vamos a usarlo para “reducir” adecuadamente la matriz M (pasar de tener muchos features a tener menos, pero que sean buenos).



SVD • Álgebra

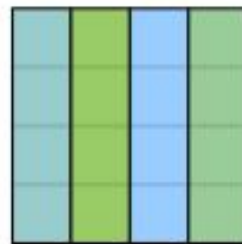
Se puede demostrar que a toda matriz M la podemos escribir como :

Matriz de Datos
(m instancias,
 n features)



$$\mathbf{M}_{m \times n}$$

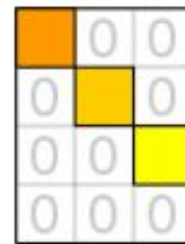
Matriz de
vectores
singulares por
izquierda



$$\mathbf{U}_{m \times m}$$

Matriz
Unitaria

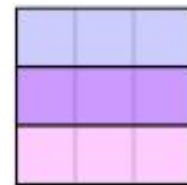
Matriz de los
valores
singulares



$$\mathbf{\Sigma}_{m \times n}$$

Matriz
Diagonal

Matriz de
vectores
singulares por
derecha



$$\mathbf{V}^*_{n \times n}$$

Matriz
Unitaria

¿Y qué tiene que ver esto con todo lo que venimos hablando?



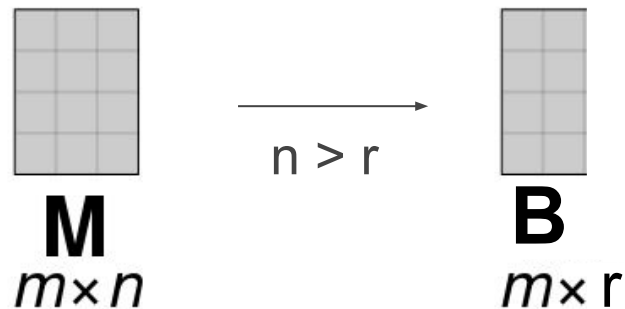
Aprendizaje No Supervisado

SVD truncado



SVD truncado

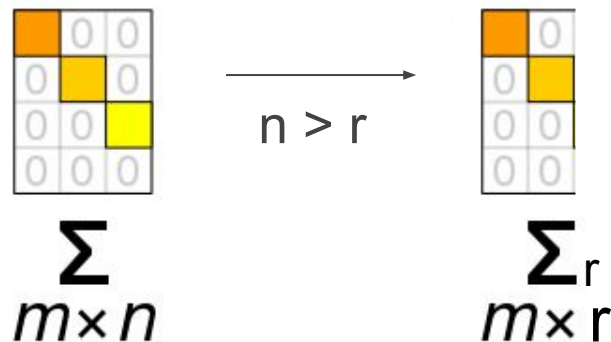
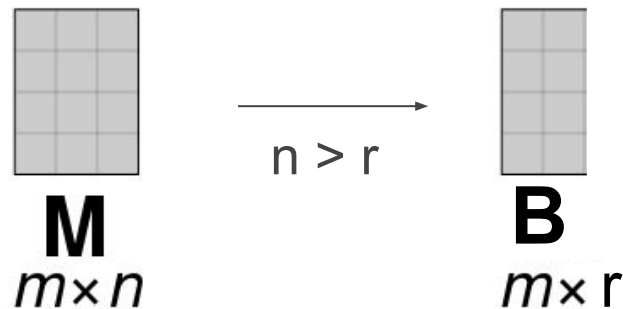
Objetivo: queremos una nueva matriz B que reemplace a M , que tenga menos columnas (menos features).



SVD truncado

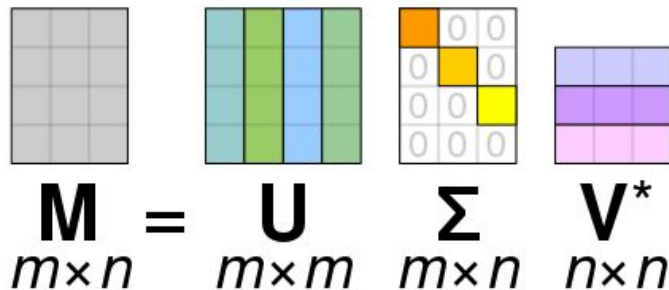
Objetivo: queremos una nueva matriz B que reemplace a M , que tenga menos columnas (menos features).

Idea de cómo lograrlo: si tomamos solo los r valores principales (elementos en la diagonal de Sigma) de valor más grande, podemos construir una matriz B que sea una “buena” reducción de M .



SVD truncado

Matriz completa: es la M original, tiene toda la información.



The diagram illustrates the SVD decomposition of a matrix M into three components: U , Σ , and V^* . Each component is represented by a grid of colored squares.

- M (size $m \times n$): A 4x4 grid of gray squares.
- U (size $m \times m$): A 4x4 grid with columns colored light blue, green, light blue, and green.
- Σ (size $m \times n$): A 4x4 grid with a diagonal of colored squares (orange, yellow, yellow, orange) and zeros elsewhere.
- V^* (size $n \times n$): A 4x4 grid with rows colored light blue, purple, purple, and pink.

$$\begin{matrix} \text{4x4 grid} \\ \mathbf{M} \\ m \times n \end{matrix} = \begin{matrix} \text{4x4 grid} \\ \mathbf{U} \\ m \times m \end{matrix} \begin{matrix} \text{4x4 grid} \\ \mathbf{\Sigma} \\ m \times n \end{matrix} \begin{matrix} \text{4x4 grid} \\ \mathbf{V}^* \\ n \times n \end{matrix}$$

SVD truncado

Matriz completa: es la M original, tiene toda la información.

$$\begin{matrix} \text{4x4} & & \text{4x4} & & \text{4x4} & & \text{4x4} \\ \mathbf{M} & = & \mathbf{U} & & \mathbf{\Sigma} & & \mathbf{V}^* \\ m \times n & & m \times m & & m \times n & & n \times n \end{matrix}$$

Matriz truncada: perdimos información. Pero si tomamos un valor de r adecuado, \hat{M} moño es muy parecida a M . Construimos una matriz B mas chica que M , esta es la matriz con la que vamos a trabajar.

$\downarrow n > r$

$$\begin{matrix} \text{4x4} & & \text{4x4} & & \text{4x3} & & \text{3x4} \\ \hat{\mathbf{M}} & = & \mathbf{U} & & \mathbf{\Sigma}_r & & \mathbf{V}_r^* \\ m \times n & & m \times m & & m \times r & & r \times n \end{matrix}$$

SVD truncado

Parecidas




Diagram illustrating the full SVD decomposition of matrix M into three matrices: U , Σ , and V^* .

Matrix M is shown as a 4x4 grid. Matrix U is shown as a 4x4 grid with columns colored green, blue, green, and blue. Matrix Σ is shown as a 4x4 grid with diagonal elements colored orange, yellow, and yellow, and zeros elsewhere. Matrix V^* is shown as a 4x4 grid with rows colored blue, purple, purple, and pink.

$$\begin{matrix} M \\ m \times n \end{matrix} = \begin{matrix} U \\ m \times m \end{matrix} \begin{matrix} \Sigma \\ m \times n \end{matrix} \begin{matrix} V^* \\ n \times n \end{matrix}$$

$n > r$





Diagram illustrating the truncated SVD decomposition of matrix M into three matrices: U , Σ_r , and V_r^* .

Matrix \tilde{M} is shown as a 4x4 grid. Matrix U is shown as a 4x4 grid with columns colored green, blue, green, and blue. Matrix Σ_r is shown as a 4x3 grid with diagonal elements colored orange, yellow, and yellow, and zeros elsewhere. Matrix V_r^* is shown as a 4x3 grid with rows colored blue, purple, purple, and pink.

$$\begin{matrix} \tilde{M} \\ m \times n \end{matrix} = \begin{matrix} U \\ m \times m \end{matrix} \begin{matrix} \Sigma_r \\ m \times r \end{matrix} \begin{matrix} V_r^* \\ r \times n \end{matrix}$$

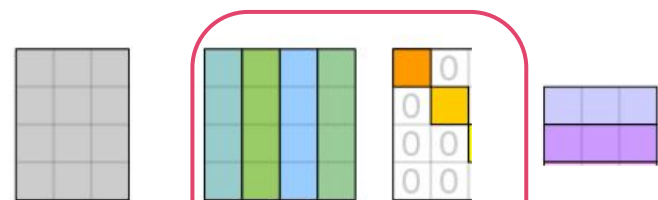
SVD truncado

Parecidas



$$\mathbf{M}_{m \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n} \mathbf{V}^*_{n \times n}$$

$n > r$



$$\tilde{\mathbf{M}}_{m \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_r_{m \times r} \mathbf{V}_r^*_{r \times n}$$

$\mathbf{B}_{m \times r}$

SVD truncado

$$\hat{\mathbf{M}}_{m \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_r_{m \times r} \mathbf{V}_r^*_{r \times n}$$

B
 $m \times r$



Matriz con la que vamos a trabajar
en vez de M, tiene la misma
información que M moño.

Esta matriz funciona
como un diccionario
para pasar del mundo
de B al mundo de M.



Aprendizaje No Supervisado

Ejemplo conceptual SVD



SVD • Ejemplo 1

Tenemos un dataset de 7 usuarios y 5 películas. Cada usuario puso un valor entre 0 a 5 a cada película.

SVD • Ejemplo 1

Tenemos un dataset de 7 usuarios y 5 películas. Cada usuario puso un valor entre 0 a 5 a cada película.

$$\mathbf{M}_{7 \times 5} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}$$

SVD • Ejemplo 1

Tenemos un dataset de 7 usuarios y 5 películas. Cada usuario puso un valor entre 0 a 5 a cada película.

$$\mathbf{M}_{7 \times 5} = \begin{array}{c} \begin{array}{ccccc} & \text{Matrix} & \text{Alien} & \text{Avatar} & \text{Titanic} & \text{Amelie} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} & \begin{array}{l} \text{Usuario 1} \\ \text{Usuario 2} \\ \dots \end{array} & \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{Ciencia Ficción} \\ \\ \end{array} \\ \\ & \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} \\ \text{Románticas} \end{array} \end{array}$$

SVD • Ejemplo 1

Buscamos una matriz B más con menos columnas que M. Proponemos usar un valor de $r=2$ es decir que B será de 7×2 . Veamos como quedaría:

The diagram illustrates the SVD decomposition of a matrix M into three components: U , Σ , and V^* . The matrix M is shown as a 7×7 grid. The matrix U is a 7×7 grid with columns colored green, blue, and green. The matrix Σ is a 7×7 grid with diagonal elements colored orange, yellow, and yellow. The matrix V^* is a 7×7 grid with rows colored purple, purple, and purple. An arrow points to the right, showing the reconstruction of M as \hat{M} using only the first two singular values and vectors. The matrix \hat{M} is a 7×7 grid. The matrix U is a 7×7 grid with columns colored green, blue, and green. The matrix Σ_r is a 7×2 grid with diagonal elements colored orange and yellow. The matrix V_r^* is a 2×7 grid with rows colored purple and purple. The entire reconstruction equation is enclosed in a red rounded rectangle, and the matrix B is labeled below it as 7×2 .

$$\begin{matrix} \text{Grid } 7 \times 7 \\ \mathbf{M} \\ m \times n \end{matrix} = \begin{matrix} \text{Grid } 7 \times 7 \\ \mathbf{U} \\ m \times m \end{matrix} \begin{matrix} \text{Grid } 7 \times 7 \\ \mathbf{\Sigma} \\ m \times n \end{matrix} \begin{matrix} \text{Grid } 7 \times 7 \\ \mathbf{V}^* \\ n \times n \end{matrix}$$
$$\begin{matrix} \text{Grid } 7 \times 7 \\ \hat{\mathbf{M}} \\ m \times n \end{matrix} = \begin{matrix} \text{Grid } 7 \times 7 \\ \mathbf{U} \\ m \times m \end{matrix} \begin{matrix} \text{Grid } 7 \times 2 \\ \mathbf{\Sigma}_r \\ m \times r \end{matrix} \begin{matrix} \text{Grid } 2 \times 7 \\ \mathbf{V}_r^* \\ r \times n \end{matrix}$$

$\mathbf{B}_{7 \times 2}$

Esta vez usaremos solo los 2 valores singulares más grandes de Sigma.

SVD • Ejemplo 1

$$U_r = \begin{bmatrix} 0.13 & 0.02 \\ 0.41 & 0.07 \\ 0.55 & 0.09 \\ 0.68 & 0.11 \\ 0.15 & -0.59 \\ 0.07 & -0.73 \\ 0.07 & -0.29 \end{bmatrix} \quad \Sigma_r = \begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix} \quad V_r^* = \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & -0.12 & -0.69 & -0.69 \end{bmatrix}$$

SVD • Ejemplo 1

$$U_r = \begin{bmatrix} \boxed{0.13} & \boxed{0.02} \\ \boxed{0.41} & \boxed{0.07} \\ \boxed{0.55} & \boxed{0.09} \\ \boxed{0.68} & \boxed{0.11} \\ \boxed{0.15} & \boxed{-0.59} \\ \boxed{0.07} & \boxed{-0.73} \\ \boxed{0.07} & \boxed{-0.29} \end{bmatrix}$$

$$\Sigma_r = \begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix}$$

$$V_r^* = \begin{bmatrix} \begin{matrix} \text{Ciencia Ficción} \\ \text{Matrix} & \text{Alien} & \text{Avatar} \end{matrix} & \begin{matrix} \text{Románticas} \\ \text{Titanic} & \text{Amelie} \end{matrix} \\ \hline \begin{matrix} 0.56 & 0.59 & 0.56 \\ 0.12 & -0.02 & -0.12 \end{matrix} & \begin{matrix} 0.09 & 0.09 \\ -0.69 & -0.69 \end{matrix} \end{bmatrix}$$

Pesos: X Z

SVD • Ejemplo 1

$$U_r = \begin{bmatrix} \boxed{0.13} & \boxed{0.02} \\ \boxed{0.41} & \boxed{0.07} \\ \boxed{0.55} & \boxed{0.09} \\ \boxed{0.68} & \boxed{0.11} \\ \boxed{0.15} & \boxed{-0.59} \\ \boxed{0.07} & \boxed{-0.73} \\ \boxed{0.07} & \boxed{-0.29} \end{bmatrix} \quad \Sigma_r = \begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix} \quad V_r^* = \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & -0.12 & -0.69 & -0.69 \end{bmatrix}$$

Ciencia Ficción			Románticas	
Matrix	Alien	Avatar	Titanic	Amelie

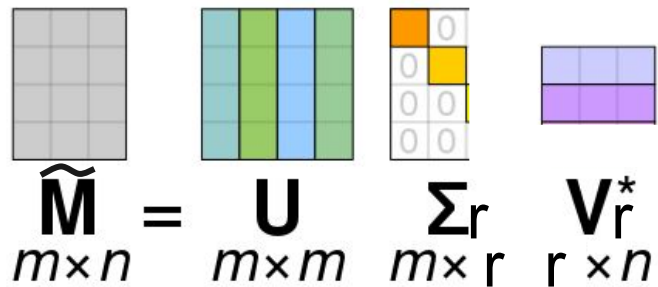
Pesos: X Z

- Ahora cada Usuario estará identificado por dos features X y Z. Notemos que los primeros 4 usuarios tienen un valor alto de X y bajo de Z. En los otros 3, se da al revés.
- Los features encontrados corresponden a los géneros.

SVD • Ejemplo 1

Pasamos de identificar a cada usuario con un puntaje al género de las películas en lugar de a las películas en sí, pasamos de 5 a 2 features.

Cuanta información perdemos por usar B en lugar de M?



The diagram illustrates the SVD decomposition of a matrix \hat{M} into three components: U , Σ_r , and V_r^* .

- \hat{M} is represented by a 4x4 grid of gray squares, with dimensions $m \times n$ indicated below.
- U is represented by a 4x4 grid of colored squares (teal, green, blue, green), with dimensions $m \times m$ indicated below.
- Σ_r is represented by a 4x4 grid of squares, with dimensions $m \times r$ indicated below. The top-left 2x2 subgrid contains colored squares (orange, white, white, yellow), and the rest are white.
- V_r^* is represented by a 4x2 grid of colored squares (light blue, purple), with dimensions $r \times n$ indicated below.

The equation is shown as:

$$\hat{M}_{m \times n} = U_{m \times m} \Sigma_r_{m \times r} V_r^*_{r \times n}$$


SVD • Ejemplo 1

Pasamos de identificar a cada usuario con un puntaje al género de las películas en lugar de a las películas en sí, pasamos de 5 a 2 features.

Cuanta información perdemos por usar B en lugar de M?

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}$$

$$\tilde{M} = \begin{bmatrix} 0.92 & 0.95 & 0.92 & 0.01 & 0.01 \\ 2.91 & 3.01 & 2.91 & -0.01 & -0.01 \\ 3.90 & 4.04 & 3.90 & 0.01 & 0.01 \\ 4.82 & 5.00 & 4.82 & 0.03 & 0.03 \\ 0.70 & 0.53 & 0.70 & 4.11 & 4.11 \\ -0.69 & 1.34 & -0.69 & 4.78 & 4.78 \\ 0.32 & 0.23 & 0.32 & 2.01 & 2.01 \end{bmatrix}$$



$$\tilde{M}_{m \times n} = U_{m \times m} \Sigma_r_{m \times r} V_r^*_{r \times n}$$

Estamos muy cerca!!

Hiperparámetro r

SVD • Hiperparámetro r

¿Cómo podríamos elegir el valor de r ?

SVD • Hiperparámetro r

¿Cómo podríamos elegir el **valor de r**?

Una posibilidad es **mirar la distancia entre M y M moño.**

$$\|M - \tilde{M}\|_F = \sqrt{\sum_{ij} (M_{ij} - \tilde{M}_{ij})^2}$$

El método de SVD nos GARANTIZA que elegimos los mejores r vectores (combinaciones de features) para minimizar esta norma!

SVD • Hiperparámetro r

¿Cómo podríamos elegir el **valor de r**?

Una posibilidad es **mirar la distancia entre M y M moño.**

$$\|M - \tilde{M}\|_F = \sqrt{\sum_{ij} (M_{ij} - \tilde{M}_{ij})^2}$$

El método de SVD nos GARANTIZA que elegimos los mejores r vectores (combinaciones de features) para minimizar esta norma!

Full-Rank Dog



Rank 200 Dog



Rank 100 Dog



Rank 50 Dog



Rank 30 Dog



Rank 20 Dog



Rank 10 Dog



Rank 3 Dog



SVD • Hiperparámetro r

¿Cómo podríamos elegir el **valor de r** ?

Una posibilidad es **mirar la distancia entre M y M moño.**

Otra posibilidad es **tener algún criterio sobre el peso relativo** de los valores singulares seleccionados respecto a la suma de todos. (Es más costoso, hay que calcular todos los valores singulares)

¿Y no hay algo un poco más visual?

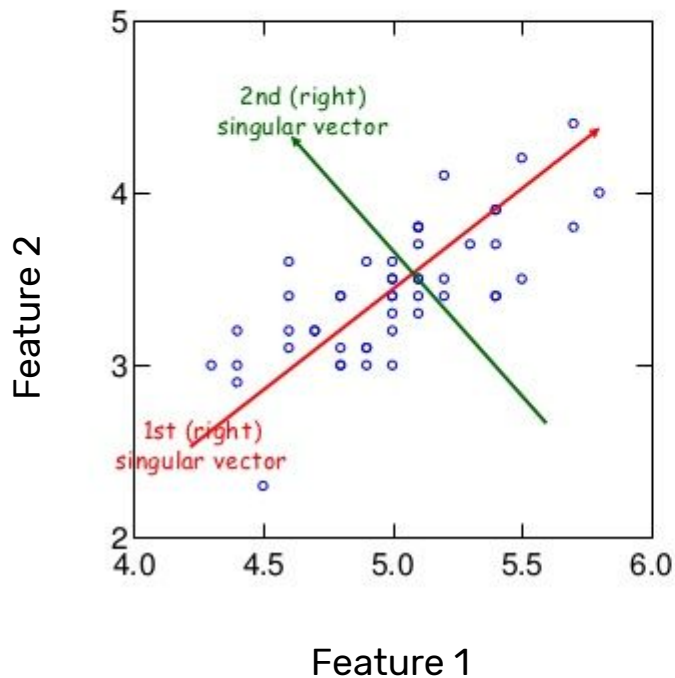


Aprendizaje No Supervisado

Representación gráfica SVD

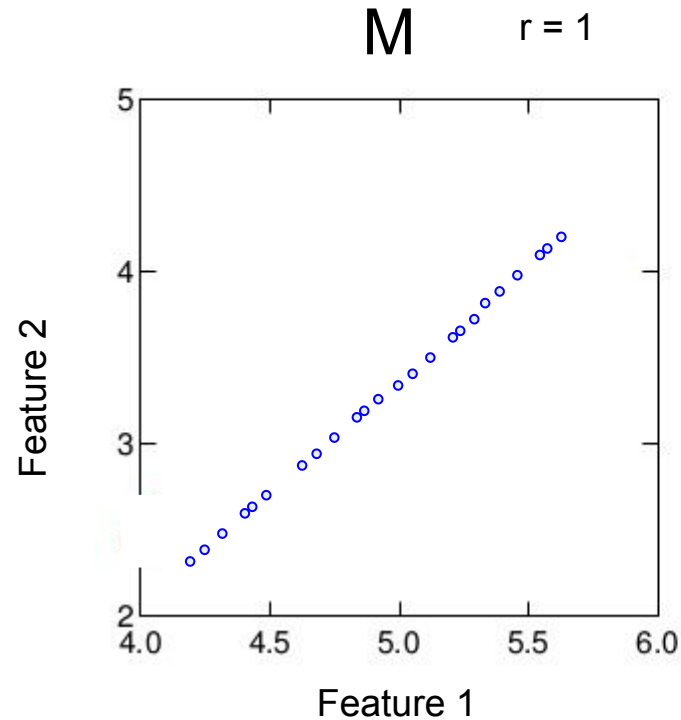
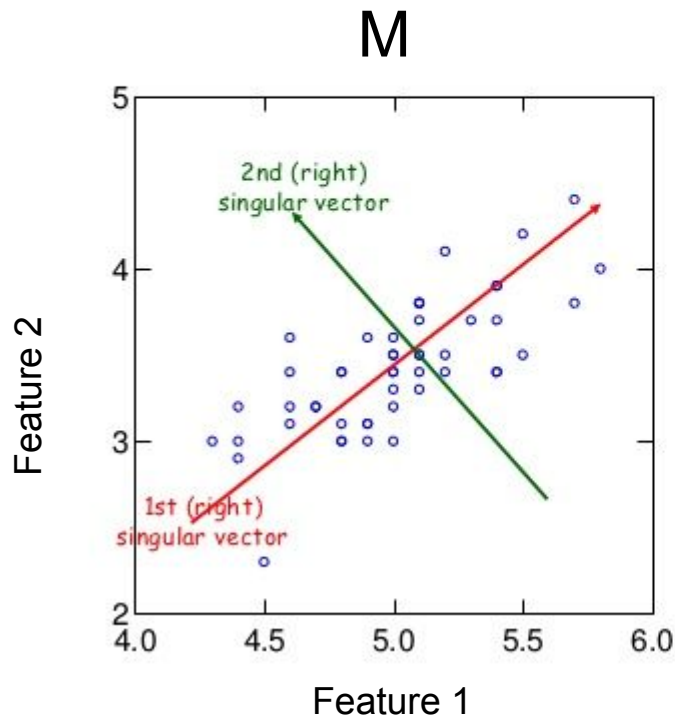


SVD • Representación gráfica



- El espacio original tiene 2 coordenadas, 2 features. Esto sirve para definir la posición de todas las instancias del dataset (cada punto azul).
- SVD nos da dos nuevos vectores, el 1er y 2do vector singular. Si usamos ambos como coordenadas, podemos definir perfecto la posición de cada punto.
- Veamos qué pasa si ahora sólo usamos el primer vector singular para definir los puntos.

SVD • Representación gráfica



A close-up photograph of a white ceramic cup filled with a latte. The surface of the milk is decorated with intricate latte art, featuring a central heart shape surrounded by concentric, wavy lines. The cup is placed on a matching white saucer. In the background, a white napkin and a silver fork are visible, though they are out of focus. The overall lighting is soft and even, highlighting the textures of the coffee and the smooth surface of the cup.

¡BREAK!



Hands-on training



Hands-on training



DS_Encuentro_38_SVD.ipynb

ACÀMICA