



FACULTAD de  
**CIENCIAS ECONÓMICAS**

## **VARIABLES LATENTES: NMF**

ALAN REYES-FIGUEROA

ELEMENTS OF MACHINE LEARNING

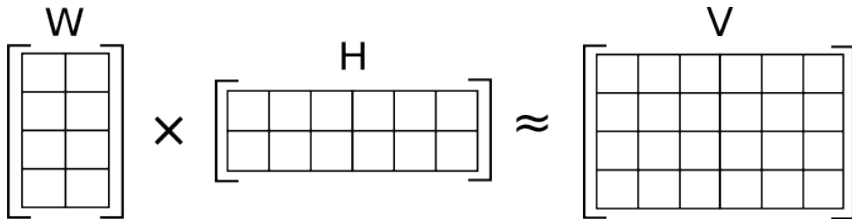
(AULA 10) 16.FEBRERO.2024

# NNMF

La Factoración de Matrices No-Negativas (NNMF) es otra estrategia para obtener variables latentes.

Sea  $\mathbb{X} \in \mathbb{R}^{n \times d}$  una matriz de datos. El objetivo es descomponer  $\mathbb{X}$  como el producto de dos matrices con entradas no-negativas  $W \in \mathbb{R}^{n \times r}$  y  $H \in \mathbb{R}^{r \times d}$

$$\mathbb{X} = WH,$$



El producto de las matrices matrices  $W$  y  $H$  se implementa de modo que cada columna de  $\mathbb{X}$  es una combinación lineal de las columnas en  $W$ , poderadas por los coeficientes proporcionados por las columnas de  $H$ . Esto es, cada columna de  $\mathbb{X}$  es

$$X_i = \mathbb{X}_i = W\mathbf{h}_i, \quad i = 1, 2, \dots, d.$$

donde  $X_i$  es el vector de la  $i$ -ésima columna de  $\mathbb{X}$  y  $\mathbf{h}_i$  la  $i$ -ésima columna de  $H$ .

Lo que se pretende es que las dimensiones de las matrices factores  $W$  y  $H$  sea baja. Esto es, definimos y valor  $1 \leq r \leq d$ , típicamente  $r \ll d$  de modo que las matrices factores  $W$  y  $H$  cumplan

$$W \in \mathbb{R}^{n \times r}, \quad H \in \mathbb{R}^{r \times d}, \quad \text{y } \text{rank}(H) = \text{rank}(W) = r.$$

El problema de encontrar estas matrices de bajo rango  $W$  y  $H$  se resuelve mediante el problema de optimización

$$\min_{W,H} \|\mathbb{X} - WH\|_F^2, \quad \text{sueto a } W \geq 0, H \geq 0. \quad (1)$$

Se requieren técnicas de optimización tipo gradiente proyectado, métodos de punto interior, u otras técnicas avanzadas para tratar este problema.

En la práctica, se utilizan esquemas de gradiente proyectado alternado, esto es, resolvemos el problema (1) en dos pasos:

1. Fijar  $W_k$ , y resolver  $H_k = \operatorname{argmin}_{H \geq 0} \|\mathbb{X} - W_k H\|_F^2$ .
2. Fijar  $H_k$ , y resolver  $W_{k+1} = \operatorname{argmin}_{W \geq 0} \|\mathbb{X} - WH_k\|_F^2$ .

## Observaciones:

- NMF tiene una estructura de clustering inherente: agrupa las columnas de  $\mathbb{X}$  en función de las componentes que encuentra en la factorización aproximada  $WH$ .
- Las columnas de  $W$  forman atributos o segmentos. Los coeficientes de  $H$  miden el grado de pertenencia a estos segmentos.
- Si además imponemos una restricción de ortogonalidad en  $H$ , es decir  $HH^T = I$ , la minimización (1) es equivalente al método de agrupamiento *k-means*.
- Cuando la restricción de ortogonalidad  $HH^T = I$  no se impone, la ortogonalidad se mantiene en gran medida, y la propiedad de agrupamiento también.
- Existen otras técnicas similares, cuya base es factorar la matriz  $\mathbb{X}$ , como el análisis de factores (FA), o el análisis semántico latente (LSA).
- Si la función de error es la divergencia Kullback-Leibler, NMF es idéntico al análisis semántico latente probabilístico LSA o LDA.

# Ejemplos

## Ejemplo 1: Recomendaciones en compras.

	John	Alice	Mary	Greg	Peter	Jennifer
Vegetables	0	1	0	1	2	2
Fruits	2	3	1	1	2	2
Sweets	1	1	1	0	1	1
Bread	0	2	3	4	1	1
Coffee	0	0	0	0	1	0

Number of purchases in category

# Ejemplos

## Ejemplo 1: Recomendaciones en compras.

	Fruits pickers	Bread eaters	Veggies
Vegetables	0	0.04	2.74
Fruits	1.93	0.15	0.47
Sweets	0.97	0	0
Bread	0	2.66	1.18
Coffee	0	0	0.59

W matrix — segment perspective

	Fruits pickers	Bread eaters	Veggies
Vegetables	0	0.04	2.74
Fruits	1.93	0.15	0.47
Sweets	0.97	0	0
Bread	0	2.66	1.18
Coffee	0	0	0.59

W matrix — category perspective

# Ejemplos

## Ejemplo 1: Recomendaciones en compras.

	John	Alice	Mary	Greg	Peter	Jennifer
Fruits pickers	1.04	1.34	0.55	0.26	0.89	0.9
Bread eaters	0	0.6	1.12	1.36	0.03	0.07
Veggies	0	0.35	0	0.34	0.77	0.69

H matrix

Pregunta: con esta información ¿Cómo hacer recomendaciones?



# Ejemplos

## Ejemplo 1: Recomendaciones en compras.

	John	Alice	Mary	Greg	Peter	Jennifer
Vegetables	0	0.98	0.04	0.98	2.11	1.9
Fruits	2	2.84	1.23	0.87	2.07	2.06
Sweets	1.01	1.31	0.54	0.26	0.86	0.87
Bread	0	2.01	2.99	4.01	0.99	1
Coffee	0	0.2	0	0.2	0.45	0.41

Reconstructed  $X$  matrix — person perspective

Rankeamos la columna correspondiente a un usuario(columna  $h_i$ ). Las recomendaciones se hacen en función de este *score* indicado en los coeficientes de la columna  $h_i$

## Ejemplo 1: Recomendaciones en compras.

	John	Alice	Mary	Greg	Peter	Jennifer
Fruits pickers	1.04	1.34	0.55	0.26	0.89	0.9
Bread eaters	0	0.6	1.12	1.36	0.03	0.07
Veggies	0	0.35	0	0.34	0.77	0.69

H matrix

Pregunta: con esta información ¿Cómo hacer recomendaciones?

# Ejemplo

## Ejemplo 1: Compras en supermercado.

	Anahí	Babá	Cadú	Didí	Edú	Fabi
Vegetales	0	1	0	1	2	2
Frutas	2	3	0	1	2	2
Dulces	1	1	0	0	1	1
Pan	0	2	3	4	1	1
Café	0	0	0	0	1	2
Carne	0	0	2	1	2	0
Lácteos	2	1	0	2	0	1
Pastas	0	3	3	1	1	0
Salsas	0	3	2	2	1	0

## Ejemplo 2: Opiniones sobre películas.

	John	Alice	Mary	Greg	Peter	Jennifer
<b>Game of Thrones</b>	5.0	NaN	1.0	NaN	NaN	NaN
<b>House of Cards</b>	NaN	NaN	NaN	7.0	NaN	6.0
<b>Friends</b>	6.0	4.0	NaN	3.0	5.0	3.0
<b>Band of Brothers</b>	2.0	9.0	NaN	NaN	NaN	9.0
<b>Breaking Bad</b>	NaN	NaN	NaN	9.0	NaN	2.0

Imdb ratings (NaN — not rated)