Descomposición de Matrices

Alondra Baerzunza

- El objetivo de los sistemas de recomendación es proporcionar un conjunto de sugerencias útiles a un usuario final.
- Para generar las recomendaciones, frecuentemente se usa la similaridad entre productos, usuarios o ambos.

Modelos de Factor Latente

- Usan los métodos de reducción de dimensionalidad para llenar los valores ausentes en la matriz de ratings.
- La idea básica es explotar que filas y columnas en una matriz están altamente correlacionadas (matriz de bajo rango).

Objetivo: completar los valores ausentes usando la matriz de bajo rango.

Los modelos de factor latente intentan explicar los ratings de los usuarios a partir de inferir patrones de productos y usuarios.

Características de la Factorización de Matrices

Se busca identificar patrones entre usuarios y productos, para lograrlo se calculan los vectores latentes (a través de factorización de matrices).

- Tiene buena escalabilidad.
- Se obtienen predicciones más precisas.
- Flexibilidad.

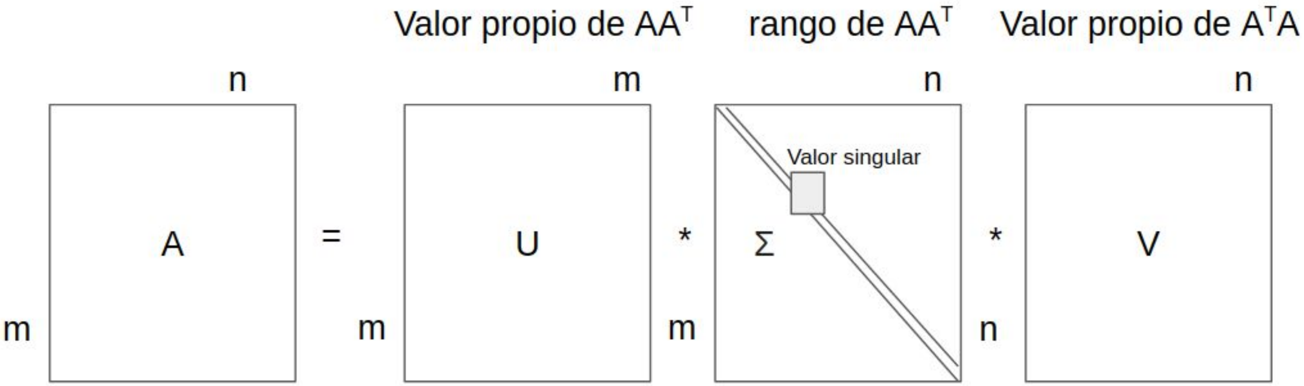
Generación de Vectores Latentes

- Usar descomposición de valores singulares (SVD).
- Análisis de componentes principales (PCA).

Vectores Latentes

Descomposición de valores singulares

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



Desventaja de SVD

- La SVD generalmente se emplea en matrices sin datos ausentes.
- Enfoques previos, usaban el promedio de los ratings de los usuarios y productos para llenar las matrices.

Vectores Latentes

La clave de los modelos de factor latente es la habilidad para estimar los vectores latentes con datos ausentes.

Partimos del supuesto que los elementos en la matriz están altamente correlacionados, por lo tanto, estas redundancias ayudan a reconstruir los valores ausentes.

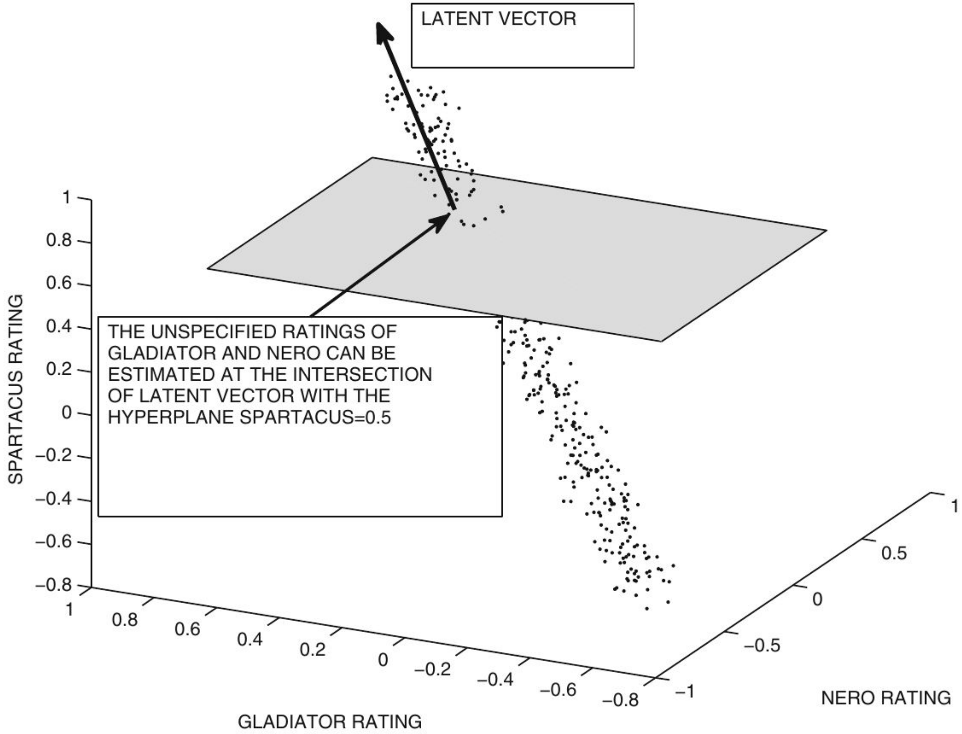
Ejemplo

Supongamos que tenemos una matriz con 3 películas, en la cual los 3 productos están positivamente correlacionados.

- Películas: Nerón, Gladiador y Espartaco.
- Los ratings son valores continuos [-1,1].

Si los ratings están positivamente correlacionados, entonces la gráfica de dispersión tridimensional de las 3 películas sería una línea unidimensional.

Ejemplo



Aprovechar las redundancias basadas en la correlación para estimar los datos faltantes para un usuario cuya única calificación especificada es un valor de 0.5 para la película Spartacus.

La intuición geométrica vista en el ejemplo anterior es útil, cuando los vectores latentes son mutuamente ortogonales.

Sin embargo, eso no siempre sucede.

Factorización de Matrices

¿Cómo calcular los vectores latentes cuando no son mutuamente ortogonales?

La factorización es una manera general de aproximación de una matriz cuando es propensa a la reducción de dimensionalidad (correlaciones entre columnas o filas).

Factorización de Matrices

Principios Básicos

Dada una matriz de ratings R de mxn, esta matriz puede ser factorizada en 2 matrices:

- Matriz U: mxk
- Matriz V: nxk

$$R \approx UV^T$$

- Cada columna de U o V se refiere como vector latente o componente latente
- Cada fila de U o V se refiere como factor latente.

Cálculo de Vectores Latentes

donde rij es la calificación que el usuario i le da al producto j y es del tamaño mxn, k es el rango, ui es el vector del usuario i y vj es el vector del producto j.

Factorización de Matrices

Cada fila $\bar{u}_i$ en la matriz U se denomina factor de usuario.

Cada fila $\bar{v}_i$ de la matriz V se denomina factor de producto.

- Objetivo final: modelar cada producto y usuario como un vector de factores.

Por lo tanto, para predecir el rating r_{ij} en R puede ser expresado como:

$$r_{ij} \approx \bar{u}_i \cdot \bar{v}_j$$

Factorización de Matrices

Dado que los factores latentes $\bar{u}_i = (u_{i1}, ..., u_{ik})$ y $\bar{v}_i = (v_{j1}, ..., v_{jk})$ pueden verse como las afinidades de los usuarios por k conceptos diferentes, la ecuación anterior, puede ser expresada como:

Ejemplo

		NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	1	1	1	0	0	0
	2	1	1	1	0	0	0
	3	1	1	1	0	0	0
BOTH	4	1	1	1	1	1	1
ROMANCE	5	- 1	- 1	- 1	1	1	1
	6	- 1	- 1	1	1	1	1
	7	- 1	- 1	- 1	1	1	1

R

		NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	1	1	1	0	0	0
	2	1	1	1	0	0	0
	3	1	1	1	0	0	0
BOTH	4	1	1	1	1	1	1
ROMANCE	5	-1	-1	-1	1	1	1
	6	-1	-1	1	1	1	1
	7	-1	-1	-1	1	1	1

R

$\approx$

	HISTORY	ROMANCE
1	1	0
2	1	0
3	1	0
4	1	1
5	-1	1
6	-1	1
7	-1	1

U

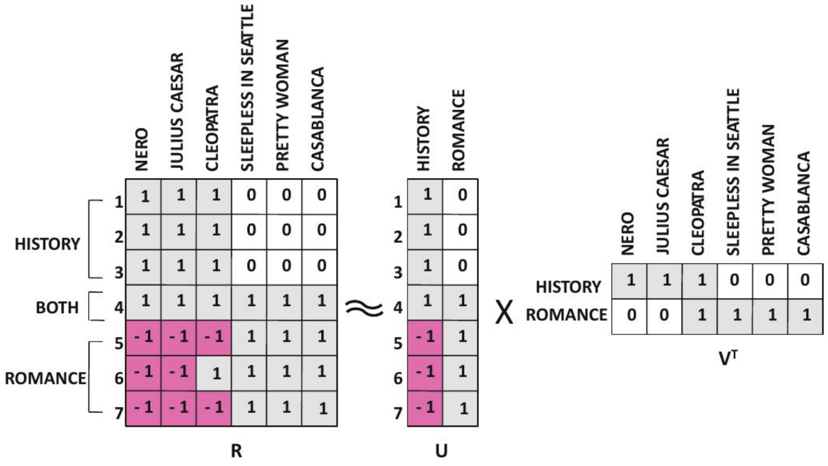
X

	NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	1	1	0	0	0
ROMANCE	0	0	1	1	1	1

V^T

Ejemplo

Los conceptos se refieren a los géneros romántico e histórico.



(afinidad del usuario i para historia) * (afinidad del producto j para historia) + (afinidad del usuario i para romance) * (afinidad del producto j para romance).

Importante

La noción de concepto (ej., romance e historia) frecuentemente no tienen una interpretación semántica.

- Un vector latente comúnmente es más parecido a esto:
- Lo importante es: un vector latente describe los patrones de correlación en la matriz R.

Principales diferencias entre métodos

Existen muchos métodos de factorización de matrices, las principales diferencias son:

- Las restricciones de las matrices U y V.
- Ortogonalidad o no negatividad de los vectores latentes.
- La naturaleza de la función objetivo.
- Minimizar la norma de Frobenius o maximizar la estimación de probabilidad.

Obtener las matrices U y V que aproximen a R

Se puede formular como un problema de optimización.

Esta función objetivo se puede ver como una **función de pérdida cuadrática**. Asimismo, métodos del descenso del gradiente pueden dar una solución óptima para esta factorización.

Obtener las matrices U y V que aproximen a R con datos ausentes

Las matrices de ratings se caracterizan tener datos ausentes, es decir, solo un conjunto de entradas tiene valores observados.

- La función objetivo anterior, no podría ser aplicada.
- No es posible calcular la norma de Frobenius con entradas faltantes.

Solución: la función objetivo debe re-escribirse usando solo las entradas observadas para aprender U y V.

Una vez aprendidos los factores latentes U y V es posible reconstruir la matriz completa usando UV T .

Obtener las matrices U y V que aproximen a R con datos ausentes

Supongamos todos los pares observados (i, j) en R los cuales denotaremos como S.

Sea $i \in (1, ..., m)$ el índice de un usuario y $j \in (1, ..., n)$ el índice de un producto.

Por lo tanto, el conjunto S de pares observados (usuario-producto) se define:

Obtener las matrices U y V que aproximen a R con datos ausentes

Una vez calculadas las matrices U y V usando las entradas observadas, es posible predecir las entradas faltantes:

La diferencias entre las entradas observadas y las entradas predichas de una entrada específica (i, j) es:

Actualización de la función objetivo

Dada una matriz R con entradas faltantes, la función objetivo se calcula sobre las entradas observadas:

Nota: la función objetivo suma el error únicamente sobre las entradas observadas.

Actualización de la función objetivo

Cabe señalar que cada uno de los términos (1), es el error cuadrado $e_{i,j}^2$ entre la entrada observada y la entrada predicha (i, j).

u_{is} y v_{js} son las variables desconocidas las cuales necesitan ser aprendidas para minimizar la función objetivo.

El algoritmo del descenso del gradiente es el método más usado para minimizar la función objetivo.

- Fácil de implementar.
- El tiempo de ejecución es relativamente rápido.