

LAB ONE (MD)

AnaGSanjuanM

2023-02-22

Es la estructura de red que AMAZON ocupa para hacer recomendaciones en base a tus compras actuales

```
M = as.matrix(
  read.csv("https://raw.githubusercontent.com/PABalland/ON/master/amz.csv" ,
    sep = ",",
    header = T,
    row.names = 1))
```

Mostrar matriz y conteo de renglones y columnas

M

```
##      Tie Book Surfboard Short Water
## Pierre  0    0          1      1     1
## Ron     1    1          0      0     1
## Andrea  1    1          0      0     1
## David   0    0          1      1     1
## Cesar   1    1          0      0     1
## Paula   1    1          1      1     1
```

dim(M)

```
## [1] 6 5
```

Ver sólo el vector de clientes que compran el producto “corbata” (producto de la columna 1)

```
M[1:6, 1]
```

```
## Pierre   Ron Andrea David  Cesar  Paula
##      0      1      1      0      1      1
```

Ver sólo un fragmento de la base original, en relación a los dos productos que me interesan Puede ser útil para explorar clientes y productos

```
M[1:6, 1:2]
```

```
##           Tie Book
## Pierre    0    0
## Ron       1    1
## Andrea    1    1
## David     0    0
## Cesar     1    1
## Paula     1    1
```

Computar grado de centralidad

```
rowSums (M)
```

```
## Pierre    Ron Andrea David Cesar Paula
##         3     3     3     3     3     5
```

Función útil para enfocarnos en productos y no sólo en clientes

```
t(M)
```

```
##           Pierre Ron Andrea David Cesar Paula
## Tie           0  1     1     0     1     1
## Book          0  1     1     0     1     1
## Surfboard     1  0     0     1     0     1
## Short         1  0     0     1     0     1
## Water         1  1     1     1     1     1
```

Suma de matrices

```
M + M
```

```
##           Tie Book Surfboard Short Water
## Pierre    0    0         2     2     2
## Ron       2    2         0     0     2
## Andrea    2    2         0     0     2
## David     0    0         2     2     2
## Cesar     2    2         0     0     2
## Paula     2    2         2     2     2
```

Crear nuevo un objeto llamado x

```
x = M + M
```

Elevar matriz al cuadrado

```
M^2
```

```
##           Tie Book Surfboard Short Water
## Pierre    0    0         1     1     1
## Ron       1    1         0     0     1
## Andrea    1    1         0     0     1
## David     0    0         1     1     1
## Cesar     1    1         0     0     1
## Paula     1    1         1     1     1
```

Multiplicación de matrices idénticas

```
M * M
```

```
##           Tie Book Surfboard Short Water
## Pierre    0    0          1    1    1
## Ron       1    1          0    0    1
## Andrea    1    1          0    0    1
## David     0    0          1    1    1
## Cesar     1    1          0    0    1
## Paula     1    1          1    1    1
```

Checar tamaño de la primera matriz -dim- Solo si el número de columnas de la primera matriz es igual al número de renglones de la segunda matriz

```
dim(M)
```

```
## [1] 6 5
```

Matriz transpuesta. Esta es la segunda matriz

```
t(M)
```

```
##           Pierre Ron Andrea David Cesar Paula
## Tie           0    1      1    0      1    1
## Book          0    1      1    0      1    1
## Surfboard     1    0      0    1      0    1
## Short         1    0      0    1      0    1
## Water         1    1      1    1      1    1
```

Checar tamaño de la segunda matriz -dim- Sí se cumple con la regla de matrices

```
dim(t(M))
```

```
## [1] 5 6
```

Multiplicación de matrices después de checar tamaños Los signos de porcentaje indican que es una multiplicación de matrices con diferente orden

```
M %*% t(M)
```

```
##           Pierre Ron Andrea David Cesar Paula
## Pierre     3    1      1    3      1    3
## Ron        1    3      3    1      3    3
## Andrea     1    3      3    1      3    3
## David      3    1      1    3      1    3
## Cesar      1    3      3    1      3    3
## Paula      3    3      3    3      3    5
```

Multiplicación inversa con énfasis en los productos El algoritmo de recomendación de AMAZON, hace básicamente lo anterior. Para tener el número de veces que 2 productos han sido comprados por el mismo cliente Identificar productos que van frecuentemente juntos (co ocurrencias par)

```
t(M) %*% M
```

```
##           Tie Book Surfboard Short Water
## Tie           4   4           1   1   4
## Book          4   4           1   1   4
## Surfboard     1   1           3   3   3
## Short         1   1           3   3   3
## Water         4   4           3   3   6
```

Matriz base para visualizar que las co ocurrencias son reales Solo tres clientes (Piere, David y Paula) empataron en la compra del short y agua

```
M
```

```
##           Tie Book Surfboard Short Water
## Pierre     0   0           1   1   1
## Ron        1   1           0   0   1
## Andrea     1   1           0   0   1
## David      0   0           1   1   1
## Cesar      1   1           0   0   1
## Paula      1   1           1   1   1
```

Matriz de similitud de productos (de gustos) matriz de co ocurrencia de productos

```
P = t(M) %*% M
```

Para visualizar la matriz de similitud de productos El “error” de la matriz es que son productos idénticos pero no pueden co ocurrir cuatro veces cada producto Debe tener ceros porque se calcula el nivel de atracción

```
P
```

```
##           Tie Book Surfboard Short Water
## Tie           4   4           1   1   4
## Book          4   4           1   1   4
## Surfboard     1   1           3   3   3
## Short         1   1           3   3   3
## Water         4   4           3   3   6
```

No es una matriz identidad (como en la correlación)

```
diag (P) = 0
```

Es la misma matriz que la anterior pero la diagonal es cero en el mismo producto Ahora si es una matriz de co ocurrencias

```
P
```

```
##           Tie Book Surfboard Short Water
## Tie           0   4           1   1   4
## Book          4   0           1   1   4
## Surfboard     1   1           0   3   3
## Short         1   1           3   0   3
## Water         4   4           3   3   0
```

Similaridad de clientes - matriz de co ocurrencia de clientes Invertimos la multiplicación de matrices

```
C = M %*% t(M)
```

Para visualizar la matriz de similaridad de clientes El “error” de la matriz es que son clientes idénticos pero no puede co ocurrir seis veces cada cliente Debe tener ceros porque se calcula el nivel de atracción

C

```
##      Pierre Ron Andrea David Cesar Paula
## Pierre      3  1      1      3      1      3
## Ron         1  3      3      1      3      3
## Andrea      1  3      3      1      3      3
## David       3  1      1      3      1      3
## Cesar       1  3      3      1      3      3
## Paula       3  3      3      3      3      5
```

Matriz base para visualizar que las co ocurrencias son reales Pierre y Ron empataron en una compra

M

```
##      Tie Book Surfboard Short Water
## Pierre    0   0          1      1      1
## Ron       1   1          0      0      1
## Andrea    1   1          0      0      1
## David     0   0          1      1      1
## Cesar     1   1          0      0      1
## Paula     1   1          1      1      1
```

No es una matriz identidad (como en la correlación)

```
diag (C) = 0
```

Es la misma matriz que la anterior pero la diagonal es cero en el mismo cliente Ahora si es una matriz de co ocurrencias

C

```
##      Pierre Ron Andrea David Cesar Paula
## Pierre     0   1      1      3      1      3
## Ron        1   0      3      1      3      3
## Andrea     1   3      0      1      3      3
## David      3   1      1      0      1      3
## Cesar      1   3      3      1      0      3
## Paula      3   3      3      3      3      0
```