

Importancia de Aeropuertos por Vuelos

Análisis Numérico II - 2021

FAMAF

Facultad de Matemática,
Astronomía, Física y
Computación



UNC

Universidad
Nacional
de Córdoba

28 de noviembre de 2021

¿Qué es PageRank?

- Es un algoritmo empleado por Google para clasificar los sitios web en los resultados de búsqueda.
- Fue desarrollado por Larry Page y Sergey Brin los fundadores de Google.
- El modelo es un algoritmo de iteración que se basa en la aplicación del método de las potencias y el teorema de Perron-Frobenius.

- Sea P_1, P_2, \dots, P_n páginas de la web.
- Sea un vector de clasificación r , con componentes positivas.
- El rango de una página web se define,

$$r(P_i) = \sum_{j=1}^n a_{ij} r(P_j)$$

con,

$$a_{ij} = \text{Prob}(P_j \rightarrow P_i) = \begin{cases} \frac{1}{|P_j|} & \text{si } P_j \rightarrow P_i \\ 0 & \text{en otro caso.} \end{cases}$$

Define una matriz de transición A .

Método de las potencias

$$r^{(k+1)} = Ar^{(k)} \quad k = 0, 1, 2 \dots$$

Teorema (Teorema de Perron-Frobenius)

Si la matriz A es irreducible entonces, λ es el autovalor dominante y su correspondiente autovector r tiene entradas estrictamente positivas.

- Es una base de datos que contiene más de 10000 aeropuertos, estaciones de tren y terminales de ferry en todo el mundo.
- Tiene registrado 7698 aeropuertos más populares de todo el mundo.
- Durante 2014, OpenFlights registró 67663 rutas entre 3321 aeropuertos en 548 aerolíneas en todo el mundo.

Implementación

```
Proyecto > Proyecto.py > ...
1
2 import numpy as np
3
4 #A continuación se presenta los datos de OpenFlights
5
6 datos_aeropuertos = np.genfromtxt("../\\Proyecto\\airports.dat.txt", delimiter=";", dtype="str", encoding="Latin-1")
7
8 datos_rutas = np.genfromtxt("../\\Proyecto\\routes.dat.txt", delimiter=";", dtype="str", encoding="Latin-1")
9
10 #####
11
```

Los datos de datos_aeropuertos:

airports.dat.txt X

Proyecto > airports.dat.txt

```
1 1, "Goroka Airport", "Goroka", "Papua New Guinea", "GKA", "AYGA", -6.081689834590001, 145.391998291, 5282, 10, "U", "Pacific/Pc
2 2, "Madang Airport", "Madang", "Papua New Guinea", "MAG", "AYMD", -5.20707988739, 145.789001465, 20, 10, "U", "Pacific/Port_Mor
3 3, "Mount Hagen Kagamuga Airport", "Mount Hagen", "Papua New Guinea", "HGU", "AYMH", -5.826789855957031, 144.29600524902344
4 4, "Nadzab Airport", "Nadzab", "Papua New Guinea", "LAE", "AYNZ", -6.569803, 146.725977, 239, 10, "U", "Pacific/Port_Moresby", "
5 5, "Port Moresby Jacksons International Airport", "Port Moresby", "Papua New Guinea", "POM", "AYPY", -9.443380355834961, 14
6 6, "Wewak International Airport", "Wewak", "Papua New Guinea", "WVK", "AYWK", -3.58383011818, 143.669006348, 19, 10, "U", "Paci
7 7, "Narsarsuaq Airport", "Narsarsuaq", "Greenland", "UAK", "BGBW", 61.1604995728, -45.4259986877, 112, -3, "E", "America/Godt
8 8, "Godthaab / Nuuk Airport", "Godthaab", "Greenland", "GOH", "BGGH", 64.19090271, -51.6781005859, 283, -3, "E", "America/Godth
9 9, "Kangerlussuaq Airport", "Sondrestrom", "Greenland", "SFJ", "BGSF", 67.0122218992, -50.7116031647, 165, -3, "E", "America/Go
10 10, "Thule Air Base", "Thule", "Greenland", "THU", "BGTL", 76.5311965942, -68.7032012939, 251, -4, "E", "America/Thule", "airpor
```

Implementación

Los datos de `datos_rutas`:

```
Proyecto > routes.dat.txt
1  2B,410,AER,2965,KZN,2990,,0,CR2
2  2B,410,ASF,2966,KZN,2990,,0,CR2
3  2B,410,ASF,2966,MRV,2962,,0,CR2
4  2B,410,CEK,2968,KZN,2990,,0,CR2
5  2B,410,CEK,2968,OVV,4078,,0,CR2
6  2B,410,DME,4025,KZN,2990,,0,CR2
7  2B,410,DME,4025,NBC,5969,,0,CR2
8  2B,410,DME,4025,TGK,\N,,0,CR2
9  2B,410,DME,4025,UUA,5160,,0,CR2
10 2B,410,EGO,6156,KGD,2952,,0,CR2
```

Se determina una nueva matriz llamada:

```
11
12 rutas = datos_rutas[:, [2, 4]]
13
```

Se determina la cantidad exacta de aeropuertos.

```
13
14 aeropuertos = np.unique(rutas)
15
```

Implementación

Se arma una matriz de índices.

```
15
16 rutas_indice = rutas.copy()
17 for idx in range(len(aeropuertos)):
18     rutas_indice = np.where(rutas_indice == aeropuertos[idx], idx, rutas_indice)
19
```

Matriz de enlaces

```
23
24 #Matriz de enlaces
25
26 m = len(aeropuertos)
27
28 matriz_enlaces = np.zeros((m, m))
29
30 m_0, n_0 = rutas_indice.shape
31
32 for i in range(m_0):
33     matriz_enlaces[rutas_indice[i,1], rutas_indice[i, 0]] = 1
34
35
36 for j in range(m):
37     w = np.sum(matriz_enlaces[:, j])
38     if w != 0:
39
40         matriz_enlaces[:, j] = matriz_enlaces[:, j] / float(w)
41
```


Implementación

Vector inicial.

```
43
44 N = len(aeropuertos)
45 vector_inicial = (1/N) * np.ones(N)
46
```

Se aplica el método de las potencias para obtener el autovalor proporcional y el autovector correspondiente.

```
62
63 #Aplicación del método de las potencias
64
65 aeropuertos_rangos, autovalor_proporcional = autpotencias(matriz_enlaces, vector_inicial)
66
```

Se puede construir un top 10 de aeropuertos con las mejores clasificaciones.

```
80
81 top_diez_rangos = aeropuertos_rangos[:10]
82
83 top_diez = aeropuertos[:10]
84
```

Implementación

Top 10 de los aeropuertos con mejor clasificación en el mundo de acuerdo a los datos de OpenFlights registrados en 2014.

Puesto	Clasificación	Nombre	Ciudad	País	Código
1	0.15286348047817616	Frankfurt am Main Airport	Frankfurt	Germany	FRA
2	0.15033062074702938	Istanbul Airport	Istanbul	Turkey	IST
3	0.1498447170651457	Charles de Gaulle International Airport	Paris	France	CDG
4	0.14840692278730913	Amsterdam Airport Schiphol	Amsterdam	Netherlands	AMS
5	0.13811057624030693	Hartsfield Jackson Atlanta International Airport	Atlanta	United States	ATL
6	0.13260623857449214	Beijing Capital International Airport	Beijing	China	PEK
7	0.1291760432135052	Chicago O'Hare International Airport	Chicago	United States	ORD
8	0.12313806895225195	Domodedovo International Airport	Moscow	Russia	DME
9	0.12198150053731358	Munich Airport	Munich	Germany	MUC
10	0.11714838354587995	Dubai International Airport	Dubai	United Arab Emirates	DXB

El aeropuerto con menor clasificación es:

Puesto	Clasificación	Nombre	Ciudad	País	Código
3425	0.0	São Felix do Xingu Airport	Sao Felix do Xingu	Brazil	SXX

De los resultados obtenidos se puede analizar que:

- Los aeropuertos obtenidos son de ciudades o países que son considerados importantes en el mundo.
- Los rangos del top 10 son altos ya que tienen más enlaces de llegada que de salida.
- Además, los aeropuertos con enlaces de llegada también tienen rangos altos.
- La última clasificación posee un rango nulo. Su clasificación se vio afectada por qué tuvo más enlaces de salida que de llegada.

En cuanto al método PageRank:

- Para poder aplicar el algoritmo se necesita del teorema de Perron-Frobenius.
- No se pudo comprobar si la matriz de enlaces era irreducible.
- El método de PageRank cambió drásticamente los motores de búsqueda.
- Se dejó de actualizar en 2013.



Airport, airline and route data.

<https://openflights.org/data.html>



PageRank, Wikipedia.

<https://es.wikipedia.org/wiki/PageRank#:~:text=PageRank20es20una20marca20registrada,por20un20motor20de20bC3BAsgueda..>



Damián Fernández, Autovalores y autovectores.

[https://colab.research.google.com/drive/1aCMRbrRXpisjDQYI49rFdhtMvoMPqrOM?authuser=1.](https://colab.research.google.com/drive/1aCMRbrRXpisjDQYI49rFdhtMvoMPqrOM?authuser=1)



The importance of Perron-Frobenius Theorem in ranking problems.
Peretti, Alberto. Department of Economics, University of
Verona. 2014.



The Perron–Frobenius theorem and the ranking of football teams.
Keener, James P. SIAM review 1993.