Método de la Potencia

Agustín Díaz

2020

Índice

Introducción

Recordemos la definición de autovalor y autovector.

Definición

Sea $A \in \mathbb{R}^{nxn}$. Si λ es un escalar y $v \neq \mathbf{0}$ es un vector tales que

$$Av = \lambda v$$

decimos que λ es un autovalor de A y v es su autovector asociado.

Introducción

Recordemos la definición de autovalor y autovector.

Definición

Sea $A \in \mathbb{R}^{nxn}$. Si λ es un escalar y $v \neq \mathbf{0}$ es un vector tales que

$$Av = \lambda v$$

decimos que λ es un autovalor de A y v es su autovector asociado.
Notemos que si v es un autovector asociado al autovalor λ,
$\forall c \in \mathbb{R}, c \neq 0, cv$ también es un autovector asociado a λ.

$$A(cv) = cAv = c\lambda v = \lambda(cv)$$

Motivación

Cuando se trabaja con matrices dispersas de gran tamaño, puede ser muy costoso calcular el autovalor dominante y su autovector asociado de manera precisa. Por esta razón, el método de la potencia tiene aplicaciones en métodos computacionales debido a su bajo costo por iteración.

Motivación

Cuando se trabaja con matrices dispersas de gran tamaño, puede ser muy costoso calcular el autovalor dominante y su autovector asociado de manera precisa. Por esta razón, el método de la potencia tiene aplicaciones en métodos computacionales debido a su bajo costo por iteración.

- Google ha desarrollado la familia de algoritmos PageRank que utilizan el método de la potencia para asignar de forma numérica la relevancia de los documentos (o páginas web) indexados por un motor de búsqueda.

Motivación

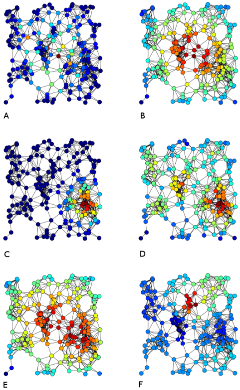
Cuando se trabaja con matrices dispersas de gran tamaño, puede ser muy costoso calcular el autovalor dominante y su autovector asociado de manera precisa. Por esta razón, el método de la potencia tiene aplicaciones en métodos computacionales debido a su bajo costo por iteración.

- Google ha desarrollado la familia de algoritmos PageRank que utilizan el método de la potencia para asignar de forma numérica la relevancia de los documentos (o páginas web) indexados por un motor de búsqueda.
- Twitter utiliza el método de la potencia en su sistema de recomendación de quién seguir.

Motivación

Centralidad de Autovalores

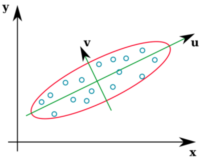
En teoría de grafos y análisis de redes, la **centralidad** en un grafo se refiere a una medida posible de un vértice en dicho grafo, que determina su importancia relativa dentro del mismo. Una manera posible de medir esto es utilizar el autovalor dominante y su autovector asociado de la matriz de adyacencia del grafo. Esto es muy utilizado en algoritmos de redes sociales.



Motivación

Matriz de Covarianza

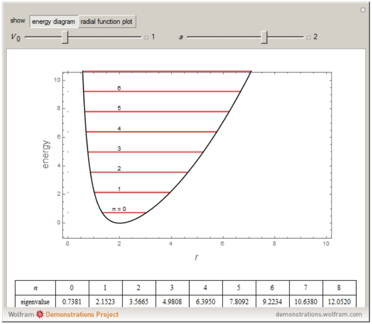
En probabilidad y estadística, la **covarianza** es un valor que indica el grado de variación conjunta de dos variables aleatorias respecto a sus medias. Obteniendo el autovalor dominante y su autovector asociado se puede determinar la tendencia de los datos.



Motivación

Niveles de energía en sistemas cuánticos

En mecánica cuántica
las ecuaciones de onda
son ecuaciones de autovalores.
En la ecuación de Shrödinger
los autovalores de la matriz
Hamiltoniana representan
los niveles de energía
del sistema, y su autovalor
mínimo determina el estado
estacionario (*ground state*).



Método de la Potencia

El método de la potencia es un método iterativo que permite estimar el radio espectral de una matriz

Teorema del Método de la Potencia

Sea la matriz $A \in \mathbb{R}^{nxn}$ diagonalizable y sus autovalores $\lambda_1, ..., \lambda_n$ repetidos según su multiplicidad y ordenados

$$|\lambda_1| > |\lambda_2| \geq ... \geq |\lambda_n| \geq 0$$

Suponiendo que existe un único autovalor de módulo máximo

Método de la Potencia

Sea $\{v_1, v_2, ..., v_n\}$ la base de autovectores correspondiente. Sea $z^{(0)}$ una estimación inicial del autovector v_1 arbitrario.

Método de la Potencia

Sea $\{v_1, v_2, ..., v_n\}$ la base de autovectores correspondiente. Sea $z^{(0)}$ una estimación inicial del autovector v_1 arbitrario.
Definimos

$$w^{(n+1)} = Az^{(n)}$$

$$z^{(n+1)} = \frac{w^{(n+1)}}{\|w^{(n+1)}\|_\infty}, \quad n \geq 0$$

Método de la Potencia

Entonces resulta

$$z^{(n)} \to \beta \frac{v_1}{\|v_1\|_\infty} \quad \text{cuando} \quad n \to \infty$$

siendo β de módulo 1.

Método de la Potencia

Entonces resulta

$$z^{(n)} \to \beta \frac{v_1}{\|v_1\|_\infty} \quad \text{cuando} \quad n \to \infty$$

siendo β de módulo 1.
Eligiendo una componente k no nula de $w^{(n-1)}$

$$\lambda^{(n)} = \frac{w_k^{(n)}}{z_k^{(n-1)}} \to \lambda_1 \quad \text{cuando} \quad n \to \infty$$

Demostración

Primero demostraremos por inducción una propiedad que nos será útil más tarde

Demostración

Primero demostraremos por inducción una propiedad que nos será útil más tarde

$$z^{(n)} = \frac{A^n z^{(0)}}{\|A^n z^{(0)}\|_\infty}, \quad n \geq 1$$

Demostración

Primero demostraremos por inducción una propiedad que nos será útil más tarde

$$z^{(n)} = \frac{A^n z^{(0)}}{\|A^n z^{(0)}\|_\infty}, \quad n \geq 1$$

Veamos que se cumple para los casos base $n = 1$ y $n = 2$

$$z^{(1)} = \frac{w^{(1)}}{\|w^{(1)}\|_\infty} = \frac{Az^{(0)}}{\|Az^{(0)}\|_\infty}$$

$$z^{(2)} = \frac{w^{(2)}}{\|w^{(2)}\|_\infty} = \frac{Az^{(1)}}{\|Az^{(1)}\|_\infty} = \frac{A\frac{Az^{(0)}}{\|Az^{(0)}\|_\infty}}{\left\|A\frac{Az^{(0)}}{\|Az^{(0)}\|_\infty}\right\|_\infty} = \frac{A^2 z^{(0)}}{\|A^2 z^{(0)}\|_\infty}$$

Demostración

Por inducción obtenemos

$$z^{(n)} = \frac{w^{(n)}}{\|w^{(n)}\|_\infty} = \frac{Az^{(n-1)}}{\|Az^{(n-1)}\|_\infty} = \frac{A\frac{A^{n-1}z^{(0)}}{\|A^{n-1}z^{(0)}\|_\infty}}{\|A\frac{A^{n-1}z^{(0)}}{\|A^{n-1}z^{(0)}\|_\infty}\|_\infty} = \frac{A^n z^{(0)}}{\|A^n z^{(0)}\|_\infty}$$

Demostración

Siendo $\{v_1, ..., v_n\}$ una base de $\mathbb{R}^n$, podemos escribir

$$z^{(0)} = \sum_{j=1}^{n} \alpha_j v_j$$

Demostración

Siendo $\{v_1, ..., v_n\}$ una base de $\mathbb{R}^n$, podemos escribir

$$z^{(0)} = \sum_{j=1}^{n} \alpha_j v_j$$

Suponiendo $\alpha_1 \neq 0$ (siempre es posible dado que $z^{(0)}$ es arbitrario)

$$\begin{aligned} A^m z^{(0)} &= A^m \sum_{j=1}^{n} \alpha_j v_j = \sum_{j=1}^{n} \alpha_j A^m v_j = \sum_{j=1}^{n} \alpha_j \lambda_j^m v_j \\ &= \lambda_1^m \left(\alpha_1 v_1 + \sum_{j=2}^{n} \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^m v_j \right) \end{aligned}$$

Demostración

Sabiendo que $|\lambda_1| > |\lambda_j|$ para todo $j = 2, ..., n$, nos queda $\left(\frac{\lambda_j}{\lambda_1}\right)^m \to 0$ cuando $m \to \infty$ para todo $j = 2, ..., n$. Entonces

$$z^{(m)} = \frac{A^m z^{(0)}}{\|A^m z^{(0)}\|_\infty} = \frac{\lambda_1^m \left(\alpha_1 v_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^m v_j\right)}{|\lambda_1|^m \left\|\alpha_1 v_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^m v_j\right\|_\infty}$$

Demostración

Sabiendo que $|\lambda_1| > |\lambda_j|$ para todo $j = 2, ..., n$, nos queda $\left(\frac{\lambda_j}{\lambda_1}\right)^m \to 0$ cuando $m \to \infty$ para todo $j = 2, ..., n$. Entonces

$$z^{(m)} = \frac{A^m z^{(0)}}{\|A^m z^{(0)}\|_\infty} = \frac{\lambda_1^m \left(\alpha_1 v_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^m v_j\right)}{|\lambda_1|^m \left\|\alpha_1 v_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^m v_j\right\|_\infty}$$

Tomando el límite

$$z^{(m)} \to \beta \frac{v_1}{\|v_1\|_\infty} \quad \text{cuando} \quad m \to \infty$$

donde

$$\beta = \left(\frac{\lambda_1}{|\lambda_1|}\right)^m \frac{\alpha_1}{|\alpha_1|}$$

tiene módulo 1.

Demostración

Tomando una componente $k \neq 0$ de $w^{(m)}$

$$\lambda^{(m)} = \frac{w_k^{(m)}}{z_k^{(m-1)}} = \frac{\left(Az^{(m-1)}\right)_k}{(z^{(m-1)})_k}$$

Demostración

Tomando una componente $k \neq 0$ de $w^{(m)}$

$$\lambda^{(m)} = \frac{w_k^{(m)}}{z_k^{(m-1)}} = \frac{\left(Az^{(m-1)}\right)_k}{(z^{(m-1)})_k}$$

Por lo tanto

$$\lambda^{(m)} \to \frac{(Av_1)_k}{(v_1)_k} = \lambda_1 \quad \text{cuando} \quad m \to \infty$$

Notas

- El método de la potencia es válido empleando cualquier norma vectorial
- Al comienzo no se sabe si la matriz tiene un autovalor dominante
- La convergencia es geométrica a razón $\left|\frac{\lambda_2}{\lambda_1}\right|$
- Con una implementación adecuada es efectivo para matrices dispersas de grandes dimensiones

Ejemplo

Sea la matriz

$$A = \begin{pmatrix} 1{,}25 & 0{,}75 \\ 0{,}75 & 1{,}25 \end{pmatrix}$$

Ejemplo

Sea la matriz

$$A = \begin{pmatrix} 1{,}25 & 0{,}75 \\ 0{,}75 & 1{,}25 \end{pmatrix}$$

Sus autovalores y autovectores asociados son

$$\lambda_1 = 2 \qquad v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 0{,}5 \qquad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Ejemplo

Sea la matriz

$$A = \begin{pmatrix} 1{,}25 & 0{,}75 \\ 0{,}75 & 1{,}25 \end{pmatrix}$$

Sus autovalores y autovectores asociados son

$$\lambda_1 = 2 \qquad v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 0{,}5 \qquad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Dado que los autovectores conforman una base de $\mathbb{R}$, podemos expresar cualquier vector z como

$$z = \alpha_1 v_1 + \alpha_2 v_2$$

Interpretación numérica

Veamos qué pasa cuando premultiplicamos a z por A varias veces

$$
\begin{aligned}
Az &= \lambda_1\alpha_1 v_1 + \lambda_2\alpha_2 v_2 = 2\alpha_1 v_1 + 0{,}5\alpha_2 v_2\\
A^2 z &= \lambda_1^2\alpha_1 v_1 + \lambda_2^2\alpha_2 v_2 = 4\alpha_1 v_1 + 0{,}25\alpha_2 v_2\\
A^3 z &= \lambda_1^3\alpha_1 v_1 + \lambda_2^3\alpha_2 v_2 = 8\alpha_1 v_1 + 0{,}0125\alpha_2 v_2
\end{aligned}
$$

Interpretación numérica

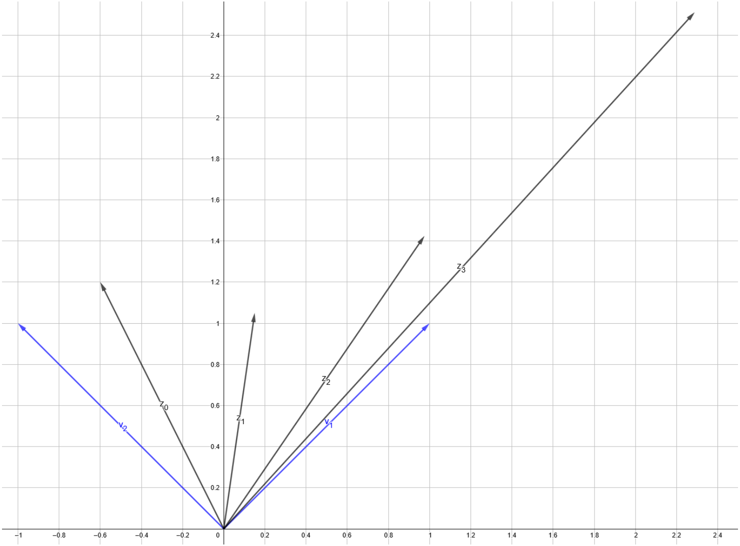
Veamos qué pasa cuando premultiplicamos a z por A varias veces

$$
\begin{aligned}
Az &= \lambda_1\alpha_1 v_1 + \lambda_2\alpha_2 v_2 = 2\alpha_1 v_1 + 0{,}5\alpha_2 v_2 \\
A^2 z &= \lambda_1^2\alpha_1 v_1 + \lambda_2^2\alpha_2 v_2 = 4\alpha_1 v_1 + 0{,}25\alpha_2 v_2 \\
A^3 z &= \lambda_1^3\alpha_1 v_1 + \lambda_2^3\alpha_2 v_2 = 8\alpha_1 v_1 + 0{,}0125\alpha_2 v_2
\end{aligned}
$$

Podemos ver que a medida que aumentamos la cantidad de iteraciones, la componente del autovector asociado al autovalor dominante tiende a ser mucho mayor que el resto.
El vector $z^{(n)} = \frac{A^n z^{(0)}}{\|A^n z^{(0)}\|_\infty}$ también mantendrá la desigualdad anterior.

Interpretación gráfica

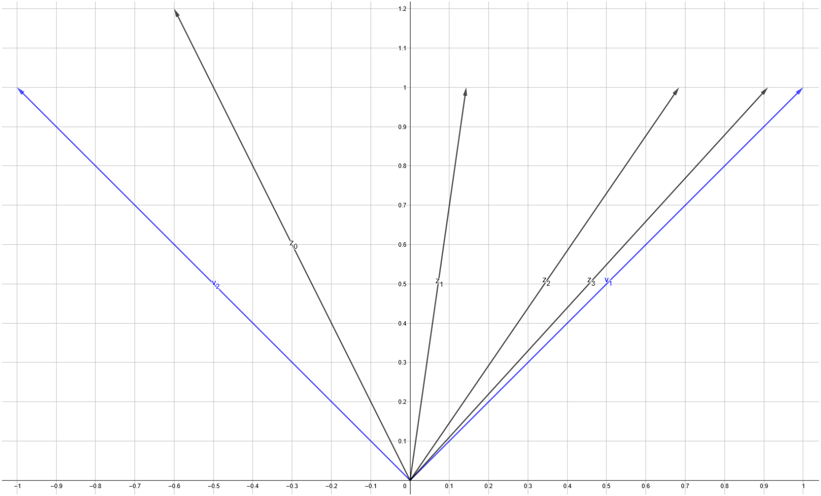
Partiendo desde $z = 0{,}3v_1 + 0{,}9v_2 = -0{,}6i + 1{,}2j$



$$z_i = A^i z$$

Interpretación gráfica

Partiendo desde $z = 0{,}3v_1 + 0{,}9v_2 = -0{,}6i + 1{,}2j$



$$z_i = z^{(i)}$$

Código

Código escrito en Scilab

```
function [lambda,z,i]=metodoDeLaPotencia(A,z,eps,maxIter)
for i=1:maxIter
    w = A * z
    znuevo = w / norm(w, 'inf')
    if norm(znuevo - z) < eps   // Regla de corte
        z = znuevo  // z representa nuestro z^(n)
        break
    end
    z = znuevo
end
// Encontramos posición de alguna componente no nula
// (en este caso la de mayor módulo)
[_, k] = max(abs(w))
w = A*z                       // Conseguimos w^(n+1)
lambda = w(k) / z(k)          // Calculamos el autovalor
endfunction
```