Métodos Numéricos

Trabajo práctico 2

Tu cara me suena

Resumen

El presente trabajo documenta un algoritmo de reconocimiento de rostros para sujetos pertenecientes a una base de datos. Hace uso de un Análisis de Componentes Principales (PCA según sus siglas en inglés) sobre un conjunto de imágenes digitales de rostros de sujetos. Se alega poder identificar con una tasa de error mínima una nueva imagen de un rostro de una persona perteneciente a la base de datos. Se explican los supuestos matemáticos asumidos en el algoritmo, se muestran resultados de diferentes pruebas de su efectividad y se llega a conclusiones sobre todo el proceso.

Integrante	LU	Correo electrónico
Gandini, Luciano	207/10	`gl.gandini@gmail.com`
Russo, Christian Sebastián	679/10	`christian.russo@gmail.com`
Danós, Alejandro	381/10	`adp007@gmail.com`

Palabras claves:

Reconocimiento caras. PCA. Power Method. Deflation. Autovalores. Autovectores. Matriz semi definida positiva.

Índice

1. Introducción Teórica

1.1. Metodo de la Potencia

El **método de la potencia** es una técnica iterativa que permite determinar el autovalor dominante de una matriz, es decir, el autovalor con mayor magnitud. Una ligera modificación en el método permite usarlo para determinar otros autovalores. Una propiedad útil del método de la potencia es que no solo produce un autovalor, sino también un autovector asociado.

De hecho, es frecuente que el método se aplique para calcular un autovalor para un autovector determinado por otros medios.

Para aplicar el método de la potencia supondremos que la matriz $A \in \mathbb{R}^{n\times n}$ tiene n autovalores λ_1, λ_2, ... , λ_n con un conjunto asociado de autovectores linealmente independientes $\{v_1, v_2, ..., v_n\}$. Más aún, supondremos que A tiene exactamente un autovalor con módulo máximo, λ_1, cuya magnitud es la mayor, por lo que

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_n| \geq 0$$

El procedimiento consiste en elegir un vector inicial $x \in \mathbb{R}^n$ y multiplicarlo por izquierda por la matriz A^k con $k \in \mathbb{N}$, y normalizar el vector resultante. Se puede demostrar que cuando $k \to \infty$, $\frac{A^k}{\|A^k\|}$ tiende al autovector asociado al autovalor dominante λ_1. Para más información, referirse a [1]

1.2. Deflación

"Deflation techniques involve forming a new matrix B whose eigenvalues are the same as those of A, except that the dominant eigenvalue of A is replaced by the eigenvalue 0 *in B."* - [1, p. 604]

En nuestro caso, para obtener la matriz B, le restaremos a la matriz A otra matriz formada por su autovalor dominante y el autovector asociado a éste, de la siguiente manera: $B = A^tA - \lambda_1 vv^t$. Más adelante demostraremos que esta operación cumple con lo esperado de una deflación.

2. Desarrollo

2.1. Lectura de la Entrada

2.1.1. Explicacion

2.1.2. Pseudocodigo

2.2. Armado de Matriz

2.2.1. Explicacion

2.2.2. Pseudocodigo

2.3. Metodo de la Potencia

2.3.1. Explicacion

2.3.2. Pseudocodigo

Algorithm 1 Método de la Potencia(Matrix B, Matrix x0, int niters, Matrix autovector)

```
v = x0
for i = 0 hasta niters do
   w = B*v
   autovector = w/(w.normVector())
   if (converge) then
      Continue
   else
      Break
   end if
end for
vt = transponer v
Bv = B*v
vtbv = vt * bv
lambda = vtbv.mat[0][0]/vtv.mat[0][0]
vtv = vt*v
return lambda
```

2.3.3. Ejemplo

$$A = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

Inicializando con el vector

$$x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Fase 1:

$$y^{(1)} = Ax^{(0)} =$$

$$\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

$$c_1 = 2 \text{ (componente dominante de } y^{(0)})$$

$$x^{(1)} = \frac{1}{2} y^{(1)} = \frac{1}{2} \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Fase 2:

$$y^{(2)} = Ax^{(1)} =$$

$$\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 3 \end{bmatrix}$$

$$c_2 = 3$$

$$x^{(2)} = \frac{1}{3} \begin{bmatrix} 3 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{-2}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} 1{,}0 \\ -0{,}6667 \\ 1{,}0 \end{bmatrix}$$

Fase 3:

$$y^{(3)} = Ax^{(2)} =$$

$$\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1{,}0 \\ -0{,}6667 \\ 1{,}0 \end{bmatrix} = \begin{bmatrix} 3{,}6667 \\ -3{,}3333 \\ 3{,}6667 \end{bmatrix}$$

$$c_3 = 3{,}6667$$

$$x^{(3)} = \begin{bmatrix} 1 \\ -0{,}9091 \\ 1 \end{bmatrix}$$

Fase 4:

$$y^{(4)} = Ax^{(3)} =$$

$$\begin{bmatrix} 3{,}9091 \\ -3{,}8181 \\ 3{,}9091 \end{bmatrix}$$

$$c_4 = 3{,}9091$$

$$x^{(4)} = \begin{bmatrix} 1 \\ -0{,}9767 \\ 1 \end{bmatrix}$$

Fase 5:

$$y^{(5)} = Ax^{(4)} =$$

$$\begin{bmatrix} 3{,}9767 \\ -3{,}9534 \\ 3{,}9767 \end{bmatrix}$$

$$c_5 = 3{,}9767$$

$$x^{(5)} = \begin{bmatrix} 1 \\ -0{,}9942 \\ 1 \end{bmatrix}$$

Fase 6:

$$y^{(5)} = Ax^{(4)} =$$

$$\begin{bmatrix} 3{,}9942 \\ -3{,}9883 \\ 3{,}9942 \end{bmatrix}$$

$$c_6 = 3{,}9942$$

$$x^{(6)} = \begin{bmatrix} 1 \\ -0{,}9985 \\ 1 \end{bmatrix}$$

Fase 7:

$$y^{(7)} = Ax^{(6)} =$$

$$\begin{bmatrix} 3{,}9985 \\ -3{,}9970 \\ 3{,}9985 \end{bmatrix}$$

$$c_7 = 3{,}9985$$

$$x^{(7)} = \begin{bmatrix} 1 \\ -0{,}9996 \\ 1 \end{bmatrix}$$

Fase 8:

$$y^{(8)} = Ax^{(7)} =$$

$$\begin{bmatrix} 3{,}9996 \\ -3{,}9993 \\ 3{,}9996 \end{bmatrix}$$

$$c_8 = 3{,}9996$$

$$x^{(8)} = \begin{bmatrix} 1 \\ -0{,}9999 \\ 1 \end{bmatrix}$$

Fase 9:

$$y^{(9)} = Ax^{(8)} =$$

$$\begin{bmatrix} 3{,}9999 \\ -3{,}9998 \\ 3{,}9999 \end{bmatrix}$$

$$c_9 = 3{,}9999$$

$$x^{(9)} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Fase 10:

$$y^{(10)} = Ax^{(9)} =$$

$$\begin{bmatrix} 4 \\ -4 \\ 4 \end{bmatrix}$$

$$c_1 0 = 4$$

$$x^{(10)} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = x^{(9)}$$

Entonces

$$\lambda_1 = \lim_j c_j = 4$$

Y su autovector asociado:

$$v = \lim_j x^{(j)} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Confirmamos usando usando el software MatLab que los autovectores de la matriz son $\lambda_1 = 4, \lambda_2 = 3, \lambda_3 = 1$ que es compatible con el λ_1 encontrado por el método de la potencia.

2.4. Deflacion

2.4.1. Explicacion

2.4.2. Pseudocodigo

Algorithm 2 Deflación(Matriz A, vector v, valor λ)

Require: λ autovalor de módulo máximo y v su autovector asociado
Require: $\|v\|$=1 y ortogonal al resto de los autovectores
Require: $A \in \mathbb{R}^{nxn}$
  Sea $B \in \mathbb{R}^{nxn}$
  for $i = 1$ hasta n **do**
    for $j = 1$ hasta n **do**
      $B_{ij} \leftarrow A_{ij} - \lambda v_i v_j$
    end for
  end for
  return B
Ensure: v autovector de B con autovalor asociado 0
Ensure: Si v' autovector de A con autovalor asociado λ', también lo es para B

2.4.3. Ejemplo

Volviendo al mismo ejemplo que en el Método de la Potencia, teníamos a

$$A = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

Y habíamos calculado su autovector $\lambda_1 = 4$ dominante y al autovector $v = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ asociado. Para deflación requerimos de un autovector con norma igual a 1, entonces lo dividimos por su norma:

$$v_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$B = A - \lambda_1 v_1 v_1^t = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} - 4/(\sqrt{3})^2 \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 5/3 & 1/3 & -4/3 \\ 1/3 & 2/3 & 1/3 \\ -4/3 & 1/3 & 5/3 \end{bmatrix}$$

Si calculamos los autovectores y autovalores de esta matriz, usando software específico a elección o usando lápiz y papel, obtendremos los siguientes:

Los autovalores quedarían: [0, 1, 3] y sus respectivos autovectores por columna

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

como esperábamos.

2.5. Demostraciones

El método de la potencia asume que tenemos un autovalor dominante y que todos los autovalores son mayores o iguales a 0. A^tA es simétrica por lo que tenemos una base ortonormal de autovectores y además es semi-defininda positiva, por lo que sus autovalores son positivos o 0. El problema es que no podemos asegurar que después de aplicar deflación, la matriz seguirá siendo semi-definida positiva, aunque sí simétrica. En esta sección, demostraremos los supuestos que asumimos para aplicar las técnicas en el trabajo.

Sean $A \in \mathbb{R}^{n\times m}$, $A^tA \in \mathbb{R}^{m\times m}$, $\lambda \in \mathbb{R}$, $v \in \mathbb{R}^m$ y $B = A^tA - \lambda_1 vv^t$.

Lema: la matriz A^tA y AA^t son simétricas.

Prueba:

$$(A^tA)^t = (A)^t(A^t)^t = A^tA \ (AA^t)^t = (A^t)^t(A)^t = AA^t$$

(1) A^tA simétrica y v vector.

Lema: la matriz B es simétrica

Prueba:

$$B_{ij} = (A^tA)_{ij} - \lambda v_i(v^t)_j = (A^tA)_{ij} - \lambda v_i v_j =^{(1)} (A^tA)_{ji} - \lambda v_j(v^t)_i = B_{ji}$$

(1) A^tA simétrica y v vector.

Lema: Los valores singulares de A son los mismos que los valores singulares de A^t.

Prueba: Los valores singulares de la matriz Σ son las raices de los autovalores en orden decreciente por la diagonal. Los autovalores están definidos como los valores que anulan a la función $\psi(\lambda) = det(\lambda I - A)$. En el caso de la traspuesta, sus autovalores son los que anulan a la función $\psi(\lambda) = det(\lambda I - A^t)$, pero $(\lambda I - A)^t = (\lambda I - A^t)$ y el determinante es invariante al trasponer una matriz. Entonces los autovalores son los mismos y, por ende, los valores singulares también.

Lema: Si $A \in \mathbb{R}^{nxm}, \Sigma \in \mathbb{R}^{mxn}, U \in \mathbb{C}^{mxm}, V \in \mathbb{C}^{nxn}, A = U\Sigma V^t$, entonces:

- $A^t = V\Sigma U^t$ con $A^t \in \mathbb{R}^{m\times n}$
- $AA^t = U\Lambda U^t$ con Λ la matriz con los autovalores de A y A^t en la diagonal y $AA^t \in \mathbb{R}^{n\times n}$.
- $A^tA = V\Lambda V^t$ con Λ la matriz con los autovalores de A y A^t en la diagonal $A^tA \in \mathbb{R}^{m\times m}$.

Prueba: El primero es inmediato de trasponer A y del hecho de que Σ es diagonal. Para el segundo y el tercero:

$$AA^t = U\Sigma V^tV\Sigma U^t =^{(1)} U\Sigma\Sigma U^t =^{(2)} U\Lambda U^t$$

$$A^tA = V\Sigma U^tU\Sigma V^t =^{(1)} V\Sigma\Sigma V^t =^{(2)} V\Lambda V^t$$

(1) U y V matrices ortogonales

(2) λ diagonal con valores singulares en la diagonal.

3. Tests

4. Apéndice

4.1. Enunciado

4.2. Método de compilación

4.3. Referencias bibliográficas

Referencias

[1] Richard L. Burden and J. Douglas Faires *Numerical Analysis*. 2005.