Descomposición Espectral

Análisis Numérico

Joaquin Cavieres G.

Descomposición espectral de matrices

Revisión de valores propios y vectores propios (*Eigenvalues* y *Eigenvectors*)

Definición 1. Una matriz $\boldsymbol{A}$ de dimensión $d \times d$ tiene un valor propio λ si hay un vector d-dimensional $\boldsymbol{u} \neq 0$ para cada $\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$. Este $\boldsymbol{u}$ es un vector propio correspondiente a λ.

En otras palabras, la transformación lineal de $\boldsymbol{A}$ "mapea" el vector $\boldsymbol{u}$ en la misma dirección. Así, cualquier transformación lineal necesariamente tiene puntos fijos direccionales de este tipo. Los siguientes puntos nos ayudan a comprender esto:

Si λ es un valor propio de $\boldsymbol{A}$, entonces:

- Existe un $\boldsymbol{u} \neq 0$ con $\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$
- Existe un $\boldsymbol{u} \neq 0$ con $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{u} = 0$. Aquí $\boldsymbol{I}$ es la matriz identidad
- $(\boldsymbol{A} - \lambda\boldsymbol{I})$ es singular (no es invertible)
- La determinante de $(\boldsymbol{A} - \lambda\boldsymbol{I})$ es igual a 0 (Condición suficiente y necesaria para un Kernel no trivial)

La determinante de $(\boldsymbol{A} - \lambda\boldsymbol{I})$ es un polinomial de grado d en λ, por tanto, tiene d raices (y algunas de ellas pueden ser raicces complejas). Esto implica la existencia de valores propios. Cuando ($\boldsymbol{A}$ tiene un valores real y además es simétrica, entonces los valores propios también son reales.

Vamos a denotar por $E(\lambda)$ al subespacio generado por todos los vectores propios asociados a λ, y el conjunto de valores propios de $\boldsymbol{A}$ será denotado como $spec(\boldsymbol{A})$ y llamado como el *spectrum* de $\boldsymbol{A}$).

Observaciones importantes:

- De acuerdo al teorema fundamental del álgebra, los valores propios siempre existen y podrían ser potencialmente números complejos
- $\boldsymbol{A}$ es invertible, si y solo si, $0 \notin spec(\boldsymbol{A})$

Ejemplo 1: Considere la siguiente matriz:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Valores propios: $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = (1-\lambda)^2 - 2^2 = (1-\lambda+2)(1-\lambda-2) = -(3-\lambda)(1+\lambda)$

De lo anterior tenemos obtenemos dos distintos valores propios $\lambda_1 = 3$ y $\lambda_2 = -1$.

Vectores propios: Para $\lambda_1 = 3$ entonces tenemos:

$$\boldsymbol{A} - 3\boldsymbol{I} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$$

por lo tanto, $E(\lambda_1 = 3) = \text{span}\left\{ \left(\begin{array}{c} 1 \\ 1 \end{array} \right) \right\}$

Para $\lambda_2 = -1$:

$$A - (-1I) = \left(\begin{array}{cc} 2 & 2 \\ 2 & 2 \end{array} \right)$$

así $E(\lambda_2 = -1) = \text{span}\left\{ \left(\begin{array}{c} 1 \\ -1 \end{array} \right) \right\}$

Calculo de valores propios y vectores propios en `R`

```r
# Matriz A
A = matrix(c(1,2,2,1), nrow = 2, byrow = TRUE)
A
```

```
##      [,1] [,2]
## [1,]    1    2
## [2,]    2    1
```

```r
sp_descomp = eigen(x = A) # Descomposición espectral

sp_descomp$values         # Valores propios
```

```
## [1]  3 -1
```

```r
sp_descomp$vectors        # Vectores propios
```

```
##           [,1]       [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Podemos verificar que mediante el calculo de los valores propios:

```r
lambda_1 = sp_descomp$values[1]
lambda_2 = sp_descomp$values[2]

vec_1 = sp_descomp$vectors[,1]
```

```r
vec_2 = sp_descomp$vectors[,2]

(A %*% vec_1 == lambda_1*vec_1) & (A %*% vec_2 == lambda_2*vec_2)
```

```
##      [,1]
## [1,] TRUE
## [2,] TRUE
```

Teorema espectral

Se considera a $\boldsymbol{A}$ una matriz simétrica con distintos valores propios $\lambda_1, \lambda_2, \cdots, \lambda_k$, además, $E(\lambda_i)$ es el espacio propio (*eigenspace*) de $\boldsymbol{A}$ y correspondiente a los valores propios λ_i. $P(\lambda_i) : \mathbb{R}^n \longrightarrow E(\lambda_i)$ corresponde a la proyección ortogonal de $\mathbb{R}^n$ sobre $E(\lambda_i)$, por consecuencia, las siguientes sentencias son verdaderas:

a) Hay una descomposición de suma directa $\mathbb{R}^n = \bigoplus_{i=1}^k E(\lambda_i)$
b) Si $B(\lambda_i) := \bigoplus_{i\neq j}^k E(\lambda_i)$ entonces $E(\lambda_i)^\perp = B(\lambda_i)$
c) La proyección ortogonal satisface que $P(\lambda_i)P(\lambda_j) = \delta_{ij}P(\lambda_i)$ para $1 \leq i, j \leq k$. En este caso $1 \leq i, j \leq k$ denota el delta Kronneker
d) La matriz identidad puede ser factorizada como $\boldsymbol{I} = \sum_{i=i}^k P(\lambda_i)$
e) La matriz $\boldsymbol{A}$ puede ser expresada como $\boldsymbol{A} = \sum_{i=i}^k \lambda_i P(\lambda_i)$.

Observaciones especiales:

1) En palabras simples, en el primer enunciado se quiere decir que existe una base de $\mathbb{R}^n$ de vectores propios consistentes de la matriz $\boldsymbol{A}$.

2) La base del punto anterior puede ser ortonormal usando el proceso de Gram-Schmidt dentro de cada espacio propio.

Finalmente, como consecuencia del teorema espectral, podemos decir que existe una matriz orgotonal $\boldsymbol{Q} \in SO(n)$, por ejemplo $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$ y $det(\boldsymbol{Q}) = \boldsymbol{I}$ tal que:

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1},$$

donde $\boldsymbol{D}$ es la matriz diagonal que contiene a los valores propios en $\boldsymbol{A}$. La matriz $\boldsymbol{Q}$ es construida ordenando los vectores propios de $\boldsymbol{A}$ como vectores columna.

Ejemplo 2: Considere la siguiente matriz:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Los puntos a), b) y c) los tenemos ya cubiertos con los ejemplos anteriores, por tanto debemos calcular entonces $P(\lambda_1 = 3) + P(\lambda_2 = -1)$ como expresa el punto d) como sigue:

$$\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}$$

Finalmente, para el punto e) hacemos:

$$\frac{3}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \boldsymbol{A}$$

Calculos en `R`: Primero construimos una función para la proyección.

```
proj = function(v, x) {

  inner_product <- as.numeric(v %*% x)

  inner_product*v
}
```

Luego las bases canónicas:

```
proj_1 = cbind(matrix(data = proj(v = sp_descomp$vectors[, 1], x = c(1, 0)), ncol = 1),
               matrix(data = proj(v = sp_descomp$vectors[, 1], x = c(0, 1)), ncol = 1))

proj_2 = cbind(matrix(data = proj(v = sp_descomp$vectors[, 2], x = c(1, 0)), ncol = 1),
               matrix(data = proj(v = sp_descomp$vectors[, 2], x = c(0, 1)), ncol = 1))

proj_1
```

```
##      [,1] [,2]
## [1,]  0.5  0.5
## [2,]  0.5  0.5
```

```
proj_2
```

```
##      [,1] [,2]
## [1,]  0.5 -0.5
## [2,] -0.5  0.5
```

Luego,

```
proj_1 + proj_2
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]    0    1
```

```
sp_descomp$values[1]*proj_1 + sp_descomp$values[2]*proj_2
```

```
##      [,1] [,2]
## [1,]    1    2
## [2,]    2    1
```

En donde la matriz $\boldsymbol{A}$ esta dada por:

$$\boldsymbol{Q} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

que da como resultado:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

pero en `R`:

```
Q = sp_descomp$vectors
Q
```

```
##           [,1]       [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

```
D = t(Q) %*% A %*% Q
D
```

```
##      [,1] [,2]
## [1,]    3    0
## [2,]    0   -1
```

Descomposición en valores singulares (SVD)

Para cualquier matriz real simétrica $\boldsymbol{A}$ de dimensión $m \times n$, y el rango de esta matriz $\boldsymbol{A}$ es r, entonces podemos diagonalizar $\boldsymbol{A}$ usando de vectores singulares. Estos dos vectores singulares son $\boldsymbol{u}$ y $\boldsymbol{v}$. Los vectores $\boldsymbol{u}$ estan en $\mathbb{R}^m$ y los vectores $\boldsymbol{v}$ estan en $\mathbb{R}^n$. Ellos deberían ser las columnas de una matriz $\boldsymbol{U}$ de dimensión $m \times m$ y las columnas de una matriz $\boldsymbol{V}$ de dimensión $n \times n$.

Se dice que $\boldsymbol{A}$ es diagonalizada cuando:

$$\boldsymbol{A}\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1, \;\; \boldsymbol{A}\boldsymbol{v}_2 = \sigma_2 \boldsymbol{u}_2 \;\; ,......, \;\; \boldsymbol{A}\boldsymbol{v}_r = \sigma_r \boldsymbol{u}_r$$

y los valores singulares σ_1 a σ_r son positivos. Ya que los vectores $\boldsymbol{u}$ son ortonormales, la matriz $\boldsymbol{U}$ con aquellas r columnas, es $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$. Para $\boldsymbol{V}$, con vectores $\boldsymbol{v}$ ortonormales también, se tiene que $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$. Por tanto, las ecuaciones $\boldsymbol{A}\boldsymbol{v}_i = \sigma_i\boldsymbol{u}_i$ nos indican que columna por columna $\boldsymbol{A}\boldsymbol{V}_r = \boldsymbol{U}_r\boldsymbol{\Sigma}_r$:

$$\boldsymbol{A}\boldsymbol{V}_r = \boldsymbol{U}_r\boldsymbol{\Sigma}_r \qquad \mathbb{A}\begin{bmatrix} \boldsymbol{v}_1 & \cdot & \cdot & \boldsymbol{v}_r \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \cdot & \cdot & \boldsymbol{u}_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \cdot & \\ & & \sigma_r \end{bmatrix}$$

A través de otros calculos matriciales adicionales, finalmente llegamos a la expresión final de la SVD:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{u}_1\boldsymbol{\sigma}_1\boldsymbol{v}_1^T + ... + \boldsymbol{u}_r\boldsymbol{\sigma}_r\boldsymbol{v}_r^T$$

En escencia, la SVD hace:

- Produce bases ortonormales de vectores $\boldsymbol{v}$ y $\boldsymbol{u}$ en subespacios fundamentales.
- Usando aquellas bases, $\boldsymbol{A}$ se convierte en una matriz diagonal $\boldsymbol{\Sigma}$ y $\boldsymbol{A}\boldsymbol{v}_i = \boldsymbol{\sigma}_i\boldsymbol{u}_i$, en donde $\boldsymbol{\sigma}_i$ es un valores singular.
- Las dos bases de diagonalización para hacer $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ generalmente contiene una amplia información.

¿Que sucede si estamos lidiando con una matriz $\boldsymbol{A}$ que no es cuadrada, por ejemplo, de dimensión $m \times n$ con $m \leq n$?

$$\boldsymbol{A}_{m\times n} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La solución compacta en tales casos podría ser $\boldsymbol{A}^T\boldsymbol{A}$ o $\boldsymbol{A}\boldsymbol{A}^T$. La descomposición de valores propios de estas matrices tienen buenas representaciones de $\boldsymbol{A}$.

Relación entre $\boldsymbol{A}^T\boldsymbol{A}$ y $\boldsymbol{A}\boldsymbol{A}^T$

Lemma: Si $\boldsymbol{A}^T\boldsymbol{A}$ y $\boldsymbol{A}\boldsymbol{A}^T$ son matrices semidefinidas positivas,

Proof: Consideramos $\boldsymbol{A}^T\boldsymbol{A}$ (lo otro es similar). Primero, esta es simetrica:

$$(\boldsymbol{A}^T\boldsymbol{A})_{ij} = \sum_k (\boldsymbol{A}^T)_{ik}\boldsymbol{A}_{ik} = \sum_k \boldsymbol{A}_{ki}\boldsymbol{A}_{kj} = \sum_k (\boldsymbol{A}^T)_{jk}\boldsymbol{A}_{ki} = (\boldsymbol{A}^T\boldsymbol{A})_{ji}$$

Luego, $\boldsymbol{A}^T\boldsymbol{A} \approx 0$ ya que para cualquier $\boldsymbol{z} \in \mathbb{R}^n$, tenemos que $\boldsymbol{z}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{z} = ||\boldsymbol{A}\boldsymbol{z}||^2 \geq 0$

¿Cual deberiamos usar entre $\boldsymbol{A}^T\boldsymbol{A}$ o $\boldsymbol{A}\boldsymbol{A}^T$ idealmente con la que tenga menor tamaño, en este caso $\boldsymbol{A}\boldsymbol{A}^T$ ya que calcular los valores propios es costoso computacionalmente de igual manera.

En terminos prácticos, la descomposición de valores singulares también fue una de las técnicas más valoradas en la solución ganadora del concurso de mejora del sistema de recomendación de Netflix (sistema de recomendación de películas o series).

Ejemplo 1

Se tiene la siguiente matriz:

$$A = \begin{bmatrix} 4 & -5 & -1 \\ 7 & -2 & 3 \\ -1 & 4 & -3 \\ 8 & 2 & 6 \end{bmatrix}$$

```
A = as.matrix(data.frame(c(4,7,-1,8), c(-5,-2,4,2), c(-1,3,-3,6)))
A
```

```
##      c.4..7...1..8. c..5...2..4..2. c..1..3...3..6.
## [1,]              4              -5              -1
## [2,]              7              -2               3
## [3,]             -1               4              -3
## [4,]              8               2               6
```

La descompisición de valores singulares de la matriz A se hace a través de la función `svd()`:

```
svd_A <- svd(A)
svd_A
```

```
## $d
## [1] 13.161210  6.999892  3.432793
##
## $u
##            [,1]       [,2]        [,3]
## [1,] -0.2816569  0.7303849 -0.42412326
## [2,] -0.5912537  0.1463017 -0.18371213
## [3,]  0.2247823 -0.4040717 -0.88586638
## [4,] -0.7214994 -0.5309048  0.04012567
##
## $v
##            [,1]        [,2]       [,3]
## [1,] -0.8557101  0.01464091 -0.5172483
## [2,]  0.1555269 -0.94610374 -0.2840759
## [3,] -0.4935297 -0.32353262  0.8073135
```

Así, la matriz A es factorizada como:

$$U = \begin{bmatrix} 0{,}281657 & -0{,}730385 & -0{,}424123 & 0{,}455332 \\ 0{,}591254 & -0{,}146302 & -0{,}183712 & -0{,}771534 \\ -0{,}224782 & 0{,}404072 & -0{,}885866 & -0{,}0379443 \\ 0{,}721499 & 0{,}530905 & 0{,}0401257 & 0{,}442683 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 13{,}1612 & 0 & 0 \\ 0 & 6{,}99989 & 0 \\ 0 & 0 & 3{,}43279 \\ 0 & 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} 0{,}85571 & -0{,}0146409 & -0{,}517248 \\ -0{,}155527 & 0{,}946104 & -0{,}284076 \\ 0{,}49353 & 0{,}323533 & 0{,}807314 \end{bmatrix}$$