Descomposición en Valor Singular (SVD: Singular Value Decomposition)

El siguiente teorema se ha citado en este texto, obviando su demostración. El capítulo 7 puede ser de gran ayuda para leer este apéndice con mayor comprensión.

El libro gratuito "Introducción al Álgebra Lineal en Contexto" se encuentra en www.abaco.com

Se sigue aquí muy de cerca la exposición sobre el tema de Lay $[1]$.

El siguiente teorema se ha citado en el texto, mostrando sus posibilidades en la práctica, cuando se ejecutó la diagonalización de matrices simétricas, en relación con la descomposición QR.

Teorema Espectral para Matrices Simétricas

En este breve estudio seguiremos a Lay $[1]$ y Moler $[2]$.

Una matriz simétrica tiene las siguientes propiedades:

a) A tiene n autovalores reales, contando multiplicidades
b) La dimensión del Espacio Propio E(λ), para cada valor propio (autovalor) λ es igual a la multiplicidad de λ como raíz de la ecuación característica.
c) Los espacios propios E(λ) son mutuamente ortogonales, en el sentido de que vectores propios (autovectores), correspondientes a autovalores diferentes, son ortogonales.
d) Existe una base ortonormal formada por autovectores. Es decir, A es diagonalizable ortogonalmente.

Aquí comienzan las aplicaciones.

Dada A_{mxn}, la matriz $A^T A$ es una matriz simétrica de orden n (ya que

$$\left(A^T A\right)^T = A^T \left(A^T\right)^T = A^T A.$$

Sus autovalores son reales y sus autovectores pueden escogerse de tal manera que constituyan una base ortonormal de (de ahí proviene su posibilidad de diagonalización por una transformación ortogonal.

Al escoger $v_1, v_2, ..., v_n$ tal que $\|v_i\| = 1$, $1 \le i \le n$, y $\langle v_i, v_j \rangle = 0$ si $i \ne j$, $\langle v_i, v_i \rangle = 1$, tenemos que:

$$\|Av_i\|^2 = \langle Av_i, Av_i \rangle = (Av_i)^T Av_i = v_i^T A^T Av_i = v_i^T \lambda_i v_i = \lambda_i v_i^T v_i = \lambda_i$$

Donde los λ_i , $1 \leq i \leq n$, son los autovalores de $A^T A$. Como la norma es no-negativa, así serán los autovalores de $A^T A$.

En consecuencia:

Los autovalores de $A^T A$, son *siempre* no negativos.

Los *valores singulares* de la matriz A son números *no-negativos* $\sigma \geq 0$, asociados con vectores $v \neq 0$ y $u \neq 0$, con las propiedades:

(1) $Av = \sigma u$ (2) $A^T u = \sigma v$

Los cuales se denominarán *vectores singulares* (derecho e izquierdo respectivamente).

Nuestro proyecto consiste en mostrar la existencia de los mismos y cómo una adecuada escogencia de los vectores singulares, nos lleva a una descomposición, llamada SVD (Singular Value Decomposition), con numerosas aplicaciones en campos como compresión de datos, procesamiento de imágenes, estadística, regresión lineal, manejo de información, etc.

Tal descomposición, es tan util, que comenzaremos pronto a añadir hipervínculos a documentos en diferentes idiomas que tratan de manera mas o menos comprensible, sus aplicaciones.

De las condiciones dadas a los valores singulares, señaladas arriba concluimos

$$A^T Av = A^T \sigma u = \sigma A^T u = \sigma(\sigma v) = \sigma^2 v, \quad v \neq 0$$
$$AA^T u = A\sigma v = \sigma Av = \sigma(\sigma u) = \sigma^2 u, u \neq 0$$

De donde concluímos que los σ^2 deben ser autovalores de la matriz simétrica $A^T A$, los cuales, afortunadamente, son siempre no negativos, lo cual facilita nuestra definición:

Los valores singulares de una matriz *A* de dimensión *mxn, se definen como* $\sigma_1 = \sqrt{\lambda_i}, \sigma_2 = \sqrt{\lambda_2}, ..., \sigma_n = \sqrt{\lambda_n}$ donde $\lambda_1, \lambda_2, ..., \lambda_n$ son los autovalores, no necesariamente distintos, de $A^T A$. Los vectores singulares derechos, serán autovectores de $A^T A$. Los vectores singulares izquierdos, autovectores de AA^T.

Si se tienen los vectores singulares derechos $v_1, v_2, ..., v_n$ y los valores singulares σ_i, los vectores singulares izquierdos correspondientes a valores singulares σ_i diferentes de 0, se pueden calcular a partir de la ecuación $Av_i = \sigma u_i$

En base a lo señalado, tenemos que como los autovectores de $A^T A$, se pueden escoger ortogonales y de norma 1, para tal escogencia, tenemos que:

$$\|Av_i\|^2 = \langle Av_i, Av_i \rangle = (Av_i)^T AV_i = v_i^T A^T Av_i = v_i^T \lambda_i v_i = \lambda_i v_i^T v_i = \lambda_i = \sigma_i^2 \geq 0$$

$$\sigma = \|Av_i\|$$

Los valores singulares serán por lo tanto las longitudes de los vectores Av_i (recuerde que los $\{v_i\}$, $1 \leq i \leq n$, constituyen una base ortonormal de R^n.

Por lo tanto los valores singulares dan una idea de la expansión producida por la matriz *A* al aplicarse al vector base y da idea de la longitud de Av_i. En el caso de las formas cuadráticas estudiadas en el capítulo 7, los valores σ_i darían una idea de las medidas de los ejes mayor y menor de una elipse rotada y lo mismo en el caso de elipsoides en R^3.

Algunos de los autovalores de $A^T A$ pueden ser 0. De ahí lo siguiente.

Asuma que los autovalores de $A^T A$, se han ordenado de tal manera que $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0, \lambda_{r+1} = o, \lambda_{r+2} = 0, \ldots, \lambda_n = 0$, es decir que se han ordenado de manera "decreciente" los autovalores diferentes de 0 (repitiéndolos según multiplicidad). Desde el mayor (dominante), hasta el menor (subdominante). Sean $v_1, v_2, \ldots, v_r$ los correspondientes autovectores de $A^T A$. En tal caso $Av_1, Av_2, \ldots, Av_r$ es una base del espacio columna de A, *C(A).* Por lo tanto *r* es el rango de A.
Además los vectores singulares izquierdos definidos por $u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma} Av_i$, serán también una base de tal espacio columna *c(A).*

Tomando la base de R^n formada por un conjunto ortonormal de autovectores de $A^T A$; $v_1, v_2, \ldots, v_n$, tales que $Av_i = \lambda_i v_i$, es claro que como $\|Av_i\| = \sigma = \sqrt{\lambda_i}$, tenemos que $Av_{r+1} = Av_{r+2} = \ldots = Av_n = 0$.

Como los vectores $v_1, v_2, \ldots, v_n$ generan a R^n ,es claro que $Av_1, Av_2, \ldots, Av_n$, generan a $\{Ax, x \in R^n\}$=*C(A).* Como $Av_{r+1} = Av_{r+2}, = \ldots, = Av_n = 0$, es claro que $Av_1, Av_2, \ldots, Av_r$ generan también a *C(A).* La independencia de los $Av_i, 1 \leq i \leq r$, se deriva de su ortogonalidad ya que $\langle Av_i, Av_j \rangle = (Av_i)^T Av_j = v_i^T A^T Av_j = v_i^T \lambda_j v_j = \lambda_j v_i^T v_j = 0$, si $i \neq j$, por constituir los v_i una base ortonormal.

La descomposición en Valor Singular SVD

Sean los $u_1, u_2, \ldots, u_r$ valores singulares izquierdos (ej: autovectores de AA^T correspondientes a los valores singulares de A, diferentes de 0). Entonces

$$A = U \sum V^T$$

En donde $U_{mxm} = [u_1 \quad u_2 \quad \ldots \quad u_r \quad u_{r+a} \quad . \quad u_m]$, donde $[u_1 u_2 \ldots u_r]$ se ha completado a una base de R^m

$$\sum_{mxn} = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}, \quad D_r = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & \sigma_r \end{bmatrix}$$

$\sigma_1, \sigma_2, \ldots, \sigma_r$, los valores singulares de A, diferentes de 0

$$V = [v_1 \quad v_2 \quad . \quad . \quad v_n]$$

Una base ortonormal de R^n formada por vectores singulares derechos de A.

Justificación:

$$AV = [Av_1 \quad Av_2 \quad . \quad Av_r \quad 0 \quad . \quad 0] = [\sigma_1 u_1 \quad \sigma_2 u_2 \quad . \quad \sigma_r u_r \quad 0 \quad 0 \quad 0] = U\Sigma$$

Por lo tanto $A = U\Sigma V^T$, ya que *V es una matriz ortogonal.*

Bibliografía

[1]. *Lay, David C.* [1999] . "Algebra Lineal y sus Aplicaciones". 2da. Edición. Ed. Pearson. Pags. 466-486

[2]. *Moler, Cleve* [2004]. "Numerical Computing with MatLab". The MathWorks. www.mathworks.com/moler

[3]. *Strang, Gilbert* [2004]. "Introduction to Linear Algebra". 3rd. Edition. Pags. 352-362