Anexo 1.

DESCOMPOSICIÓN EN VALORES SINGULARES

La descomposición en valores singulares (SVD) es una de las descomposiciones de matrices más útiles, particularmente para cálculos numéricos. Su aplicación más común es la resolución de sistemas de ecuaciones sobre-determinados.

Dada una matriz cuadrada A, la SVD es una factorización de A como $\mathrm{A} = \mathrm{UDV}^{\mathrm{T}}$, donde U y V son matrices ortogonales y D es una matriz diagonal con elementos σ_i no negativos, es decir $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. La descomposición se puede llevar a cabo de tal modo que los elementos de D estén ordenados en modo decreciente, y asumimos que esto se da siempre (la orden de MATLAB ® que realiza la descomposición SVD los ordena en dicho modo). Así, decir "la columna de V correspondiente al autovalor más pequeño" equivale a decir "la última columna de V".

La SVD también existe para matrices no cuadradas. Los casos de mayor interés se dan cuando A tiene mayor número de filas que de columnas. En general, sea A es una matriz de $m \times n$ con $m \geq n$. Se puede factorizar A como $\mathrm{A} = \mathrm{UDV}^{\mathrm{T}}$ donde U es una matriz $m \times n$ con columnas ortogonales, D es una matriz diagonal $n \times n$ y V una matriz ortogonal $n \times n$. El hecho de que U tenga columnas ortogonales significa que $\mathrm{U}^{\mathrm{T}}\mathrm{U} = \mathrm{I}_{n \times n}$. Además, U tiene la propiedad de que $\|\mathrm{U}\mathbf{x}\| = \|\mathbf{x}\|$ para cualquier vector $\mathbf{x}$, como se puede ver fácilmente. Por otro lado, $\mathrm{U}^{\mathrm{T}}\mathrm{U}$ no es en general la matriz identidad a menos que $m = n$.

También se puede definir la descomposición SVD para matrices con más columnas que filas, pero en general no son casos de interés para nosotros. En caso de que fuese necesaria la descomposición en valores singulares de una matriz A con $m < n$, se puede extender A añadiendo filas de ceros hasta obtener una matriz cuadrada, y después realizar la descomposición SVD de la matriz resultante.

Propiedades de la SVD

Propiedad 1. Los valores singulares dan información sobre la singularidad de la matriz . Una matriz cuadrada A es no singular si y sólo si todos sus valores singulares son diferentes de cero. Es más, los σ_i también nos dicen cómo de cerca está A de ser una matriz singular. La relación

$$C = \frac{\sigma_1}{\sigma_n}$$

llamada número de condición, mide el grado de singularidad de A. Cuando $1/C$ es comparable con la precisión aritmética del aparato, la matriz A está mal condicionada y, a efectos prácticos, se puede considerar singular.

Propiedad 2. Si A es una matriz rectangular, el número de elementos σ_i diferentes de cero es igual al rango de A.

Propiedad 3. Si A es cuadrada y no singular, su inversa se puede escribir como

$$A^{-1} = VD^{-1}U^{T}$$

Sea A singular o no, la pseudoinversa de A, A^{+} se puede escribir como

$$A^{+} = VD_o^{-1}U^{T}$$

con D_o^{-1} igual a D^{-1} para todos los valores singulares distintos de cero, y cero en otro caso. Si A es no singular, entonces $D_o^{-1} = D^{-1}$ y $A^{+} = A^{-1}$.

Propiedad 4. Las columnas de U correspondientes a valores singulares distintos de cero abarcan el rango de A, y las columnas de V correspondientes a los valores singulares nulos el espacio nulo de A.

Propiedad 5. Los cuadrados de los valores singulares distintos de cero son los autovalores tanto de la matriz $A^T A$ de dimensiones $n \times n$ como de la matriz AA^T de dimensiones $m \times m$. Las columnas de U son autovectores de AA^T y las de V autovectores de $A^T A$. Además, $A\mathbf{u}_k = \sigma_k \mathbf{v}_k$ y $A^T \mathbf{v}_k = \sigma_k \mathbf{u}_k$, donde $\mathbf{u}_k$ y $\mathbf{v}_k$ son las columnas de U y V correspondientes a σ_k.

Se resumen a continuación las aplicaciones de la SVD usadas en este trabajo.

<u>Mínimos Cuadrados</u>

Supongamos que se debe resolver el sistema lineal de m ecuaciones,

$$A\mathbf{x} = \mathbf{b}$$

para el vector de incógnitas n-dimensional $\mathbf{x}$. La matriz de $m \times n$ A contiene los coeficientes de las ecuaciones y el vector m-dimensional $\mathbf{b}$ los datos. Si no todos los elementos de $\mathbf{b}$ son nulos, la solución se pueden encontrar multiplicando por ambos lados de la expresión por A^T con lo que se tendría

$$A^T A\mathbf{x} = A^T \mathbf{b}$$

De aquí se obtiene que la solución viene dada por

$$\mathbf{x} = \left(\mathrm{A}^{\mathsf{T}}\mathrm{A}\right)^{+}\mathrm{A}^{\mathsf{T}}\mathbf{b}$$

Esta solución se conoce como la óptima en el sentido de los mínimos cuadrados.

A menudo es buena idea calcular la pseudoinversa de $\mathrm{A}^{\mathsf{T}}\mathrm{A}$ a través de la SVD. En el caso de más ecuaciones que incógnitas es muy probable que la pseudoinversa coincida con la *inversa* de $\mathrm{A}^{\mathsf{T}}\mathrm{A}$.

Sistemas homogéneos

Supongamos que se resolver el siguiente sistema homogéneo de m ecuaciones lineales y n incógnitas

$$\mathrm{A}\mathbf{x} = \mathbf{0}$$

con $m \geq n-1$ y $rank(\mathrm{A}) = n-1$. Ignorando la solución trivial $\mathbf{x} = \mathbf{0}$, una solución única salvo en factor de escala se puede encontrar fácilmente a través de la SVD. Esta solución es simplemente proporcional al autovector correspondiente al único autovalor cero de $\mathrm{A}^{\mathsf{T}}\mathrm{A}$ (todos los demás autovalores serán estrictamente positivos pues $rank(\mathrm{A}) = n-1$). Esto se demuestra como sigue.

Puesto que la norma de la solución al sistema de ecuaciones homogéneo es arbitraria, buscamos una solución de norma unidad en el sentido de los mínimos cuadrados. Por lo tanto, queremos minimizar

$$\left\|\mathrm{A}\mathbf{x}\right\|^{2} = (\mathrm{A}\mathbf{x})^{\mathsf{T}}\mathrm{A}\mathbf{x} = \mathbf{x}^{\mathsf{T}}\mathrm{A}^{\mathsf{T}}\mathrm{A}\mathbf{x}$$

sujeto a la restricción

$$\mathbf{x}^{\mathsf{T}}\mathbf{x} = 1$$

Introduciendo el multiplicador de Lagrange λ esto es equivalente a minimizar el *Lagrangiano*

$$\zeta(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathrm{A}^{\mathsf{T}}\mathrm{A}\mathbf{x} - \lambda(\mathbf{x}^{\mathsf{T}}\mathbf{x} - 1)$$

Igualando a cero la derivada del Lagrangiano con respecto a $\mathbf{x}$ da

$$\mathbf{x}^{\mathsf{T}}\mathrm{A}^{\mathsf{T}}\mathbf{x} - \lambda\mathbf{x} = 0$$

Esta ecuación nos dice que λ es una autovalor de $\mathrm{A}^{\mathsf{T}}\mathrm{A}$, y que la solución, $\mathbf{x} = \mathbf{e}_{\lambda}$, es el correspondiente autovector. Sustituyendo $\mathbf{x}$ por $\mathbf{e}_{\lambda}$, y $\mathrm{A}^{\mathsf{T}}\mathrm{A}\mathbf{e}_{\lambda}$ por $\lambda\mathbf{e}_{\lambda}$, en el Lagrangiano queda

$$\zeta(\mathbf{e}_{\lambda}) = \lambda$$

Por lo tanto, el mínimo se alcanza para $\lambda = 0$, el mínimo autovalor de $\mathrm{A}^{\mathsf{T}}\mathrm{A}$. Pero de las ***Propiedades 4*** y ***5***, se obtiene que la solución podría haber sido establecida equivalentemente como *la columna de* V *correspondiente al único valor singular nulo de* A (el núcleo de A).