Identificación de SISTemas

Descomposición en Valores Singulares
(SVD: Singular Value Decomposition)

La SVD y sus Aplicaciones

La **Descomposición en Valores Singulares** (o SVD por sus siglas en inglés) de una matriz es la descomposición en tres factores. Dos de ellos son matrices ortogonales, lo cual le confiere un interés especial. La razón es que las transformaciones ortogonales preservan normas y ángulos; en particular, preservan las longitudes de los vectores de error que son inevitables en los cálculos numéricos.

Este método de factorización tiene interés teórico y práctico a la vez. Entre sus muchas aplicaciones está la estimación más fiable del rango de una matriz en forma **numérica**.

Los **Métodos de Subespacio** están basados en la SVD y tienen diferentes aplicaciones tales como identificación de sistemas, identificación ciega, recepción de señales de radar, etc.

Matrices Ortogonales

Una matriz $\mathbf{A} \in \mathbf{R}^{n\times n}$ es **ortogonal** si y solo si

$$\boxed{\mathbf{A}^T \cdot \mathbf{A} = \mathbf{I}} \qquad \text{es decir} \qquad \boxed{\mathbf{A}^T = \mathbf{A}^{-1}}$$

En forma equivalente, si la matriz $\mathbf{A} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbf{R}^{n\times n}$ es ortogonal sus columnas forman un conjunto **ortonormal** de vectores; es decir, cumplen con

$$a_i^T \cdot a_j = \delta_{i\,j} = \begin{cases} 1 & si \quad i = j \\ 0 & si \quad i \neq j \end{cases} \qquad i, j = 1, \ldots, n$$

Análogamente, las filas de **A** también son ortonormales.

Descomposición en Valores Singulares (SVD)

Sea una matriz $\mathbf{A} \in \mathbf{R}^{m\times n}$, puede descomponerse o escribirse como

$$\boxed{\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T}$$

denominada *Descomposición en Valores Singulares* de A, donde

$$\mathbf{U} = \begin{bmatrix} u_1 & u_2 & \dots & u_m \end{bmatrix} \in \mathbf{R}^{m\times m}$$

$$\mathbf{V} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \in \mathbf{R}^{n\times n}$$

son matrices **ortogonales** y $\mathbf{\Sigma} \in \mathbf{R}^{m\times n}$ es una matriz **diagonal** tal que

$$
\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & \varnothing & \\ & & \ddots & & \\ & \varnothing & & \ddots & \\ & & & & \sigma_p \\ - & - & - & - & - \\ & & \varnothing & & \end{bmatrix} \qquad para \quad m > n
$$

$$
\boldsymbol{\Sigma} = \left[\begin{array}{ccccc|c} \sigma_1 & & & & & \\ & \sigma_2 & & \varnothing & & \\ & & \ddots & & & \varnothing \\ & \varnothing & & \ddots & & \\ & & & & \sigma_p & \end{array}\right] \qquad para \quad m < n
$$

donde $p = \min(m,n)$ y $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$ son los **valores singulares** de **A**.

Las columnas de la matriz **U** $(u_1, u_2, \ldots, u_m)$ son los **vectores singulares izquierdos** de **A**.

Las columnas de la matriz **V** $(v_1, v_2, \ldots, v_n)$ son los **vectores singulares derechos** de **A**.

En muchos casos basta con hacer una SVD de tamaño reducido (o "**economy size**" SVD o también "thin" SVD), que para una matriz $\mathbf{A} \in \mathbf{R}^{m \times n}$ resulta

$$\mathbf{A} = \begin{bmatrix} \mathbf{U}_1 & \vdots & \mathbf{U}_2 \end{bmatrix}_{m \times m} \cdot \begin{bmatrix} \mathbf{\Sigma}^* \\ \cdots \\ \varnothing \end{bmatrix}_{m \times n} \cdot \begin{bmatrix} \mathbf{V}^T \end{bmatrix}_{n \times n} \qquad para \quad m > n$$

$$= \mathbf{U}_1 \cdot \mathbf{\Sigma}^* \cdot \mathbf{V}^T$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{U} \end{bmatrix}_{m \times m} \cdot \begin{bmatrix} \mathbf{\Sigma}^* & \vdots & \varnothing \end{bmatrix}_{m \times n} \cdot \begin{bmatrix} \mathbf{V}_1^T \\ \cdots \\ \mathbf{V}_2^T \end{bmatrix}_{n \times n} \qquad para \quad m < n$$

$$= \mathbf{U} \cdot \mathbf{\Sigma}^* \cdot \mathbf{V}_1^T$$

donde las matrices $\mathbf{U}_1$ y $\mathbf{V}_1^{\ T} \in \mathbf{R}^{m\times n}$, $\mathbf{U} \in \mathbf{R}^{m\times m}$, $\mathbf{V}^T \in \mathbf{R}^{n\times n}$ y

$$\boldsymbol{\Sigma}^* = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & \varnothing & \\ & & \ddots & & \\ & \varnothing & & \ddots & \\ & & & & \sigma_p \end{bmatrix} \qquad con \quad p = \min(m,n)$$

Cuando la matriz **A** es tal que $m >> n$ (o $m << n$), la matriz **U** (o $\mathbf{V}^T$) resultante puede ser bastante grande y la mayoría de sus columnas (o filas) estarán multiplicadas por filas (o columnas) de ceros en la matriz **Σ**. En este caso, la SVD de tamaño reducido ahorra **tiempo de ejecución** y **espacio en memoria**, ya que produce matrices $\boldsymbol{\Sigma}^*$ y $\mathbf{U}_1$ (o $\mathbf{V}_1^{\ T}$) de menores dimensiones.

Propiedades de la SVD

1) Dadas la matriz $\mathbf{A}_{mxn}$ y su SVD, sus valores singulares (σ_i) y sus vectores singulares izquierdos y derechos (u_i y v_i) cumplen con

$$\begin{aligned} \mathbf{A} \cdot v_i &= \sigma_i \cdot u_i \\ \mathbf{A}^T \cdot u_i &= \sigma_i \cdot v_i \end{aligned} \qquad i = 1, \ldots, \min(m, n)$$

2) Si la matriz $\mathbf{A}_{mxn}$ tiene rango r, con $r \leq p = \min(m, n)$ entonces

- A tiene r valores singulares no nulos

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0$$

- el espacio nulo de **A** está dado por

$$\text{null}(\mathbf{A}) = \text{span}\{v_{r+1}, v_{r+2}, \ldots, v_n\}$$

- el espacio de columnas de **A** es

$$\operatorname{ran}(\mathbf{A}) = \operatorname{span}\{u_1, u_2, \ldots, u_r\}$$

- y la SVD de **A** puede expresarse como

$$\boxed{\mathbf{A} = \sum_{i=1}^{r} \sigma_i \cdot u_i \cdot v_i^{\,T}}$$

3) Definición de la *Norma-2* de una matriz, inducida por la norma-2 (Euclídea) de vectores:

$$\boxed{\left\|\mathbf{A}\right\|_2 \overset{\Delta}{=} \sup_{x} \left(\frac{\left\|\mathbf{A}x\right\|_2}{\left\|x\right\|_2} \right) \qquad \begin{array}{l} x \in \mathbf{R}^{n}, x \neq 0 \\ \mathbf{A} \in \mathbf{R}^{m \times n} \end{array}}$$

Aplicada a la SVD de **A**:

$$\left\|\mathbf{A}\right\|_2 = \sigma_1 \qquad \text{(máximo valor singular)}$$

Definición de la *Norma de Frobenius* de una matriz:

$$\|\mathbf{A}\|_F \overset{\Delta}{=} \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i\,j}^2} \qquad \mathbf{A} \in \mathbf{R}^{m\times n}$$

Aplicada a la SVD de **A**:

$$\|\mathbf{A}\|_F^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_p^2 \qquad con \quad p = \min(m,n)$$

Mínimo valor singular:

$$\min_x \left(\frac{\|\mathbf{A}x\|_2}{\|x\|_2} \right) = \sigma_p \qquad \begin{array}{l} x \in \mathbf{R}^n, x \neq 0 \\ p = \min(m,n) \end{array}$$

4) Definición del *Número de Condición* de una matriz:

$$\boxed{K(\mathbf{A}) \overset{\Delta}{=} \|\mathbf{A}\|_p \cdot \|\mathbf{A}^{-1}\|_p \qquad\qquad \mathbf{A} \in \mathbf{R}^{n\times n}}$$

Si consideramos la norma-2 resulta

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \longrightarrow \mathbf{A}^{-1} = \mathbf{V} \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{U}^T$$

$$\|\mathbf{A}\|_2 = \sigma_1 \qquad\qquad \|\mathbf{A}^{-1}\|_2 = \frac{1}{\sigma_n}$$

luego

$$K(\mathbf{A}) = \frac{\sigma_1}{\sigma_n}$$

Si el valor de $K(\mathbf{A})$ es muy grande la matriz **A** está “muy próxima” a ser singular.

5) Sea la matriz $\mathbf{A} \in \mathbf{R}^{m \times n}$ con rango *r*, queremos hallar una matriz $\mathbf{B} \in \mathbf{R}^{m \times n}$ con rango $k < r$ tal que la distancia en norma-2 $\|\mathbf{A} - \mathbf{B}\|_2$ sea mínima.

Puede probarse que si

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \cdot u_i \cdot v_i^T$$

es la SVD de A, entonces

$$\mathbf{B} = \sum_{i=1}^{k} \sigma_i \cdot u_i \cdot v_i^T$$

Además el error que se comete en la aproximación es mínimo y vale

$$\|\mathbf{A} - \mathbf{B}\|_2 = \sigma_{k+1}$$

Referencias:

[1] G. Nakos y D. Joyner. *Álgebra Lineal con Aplicaciones*. Thomson, 1999.

[2] G.H. Golub y C.F. Van Loan. *Matrix Computations, 3ra ed*. Johns Hopkins, 1996.

[3] *Using MATLAB v6*. The MathWorks.