Notes de cours
Algèbre Linéaire et Analyse Matricielle
3BIM INSA Lyon
2014-2015
v19112014

Ces notes de cours sont inspirées des livres de Algèbre des matrices de J Fresnel [1] et Numerical Linear Algebra de LN Trefethen et D Bau III [2].

3 Décomposition en valeurs singulières

On a vu que pour certaines matrices carrées, on pouvait faire une décomposition en valeurs propres:

$$A = XDX^{-1}, \tag{1}$$

où D est une matrice diagonale de valeurs propres et X est une matrice inversible de vecteurs propres. Quand on fait le produit matrice-vecteur $Ax = (XDX^{-1})x$, on prend x, on l'exprime dans la base donnée par les vecteurs propres ($X^{-1}x$), on multiplie les éléments de ce vecteur par les valeurs propres D une à une, et on refait le changement de base inverse en multipliant par X. Si on a un système linéaire $Ax = b$, on peut effectuer les changements de base $\hat{x} = X^{-1}x$, $\hat{b} = X^{-1}b$ et on obtient le système $D\hat{x} = \hat{b}$. *Limitations: pour effectuer cette transformation, la matrice A doit être carrée et diagonalisable.*

L'idée de la décomposition en valeurs singulière est similaire à la décomposition en valeurs propres, mais fonctionne pour n'importe quelle matrice A de taille $m \times n$: on factorise A en produit de trois matrices:

$$A = U\Sigma V^* \tag{2}$$

avec U une matrice $m \times m$ unitaire, V une matrice $n \times n$ unitaire et Σ une matrice $m \times n$ diagonale avec coefficients réels et positifs.

3.1 Interprétation géométrique

L'image d'une sphère unité dans $\mathbb{R}^n$ par une matrice A $m \times n$ est une hyperellipse dans $\mathbb{R}^m$. Soit les longueurs des axes principaux $\{\sigma_1, \sigma_2, ..., \sigma_m\}$ et les directions des axes principaux $\{u_1, u_2, ..., u_m\}$. On a $\sigma_i \geq 0$ et $u_i \in \mathbb{R}^m$, avec les l'ensemble $\{u_i\}$ orthonormal (on peut prendre $||u_i|| = 1$). Les vecteurs $\sigma_i u_i$ sont alors les semi-axes principaux de l'hyperellipse. Si $\operatorname{rang} A = r$, alors exactement r valeurs de σ_i seront non-nulles. Si $m \geq n$, alors au plus n axes seront de longueurs positives.

3.2 Décomposition en valeurs singulières réduite

Soit S la boule (sphère) unité dans $\mathbb{R}^n$ (ou $\mathbb{C}^n$), et soit une matrice A $m \times n$ avec $m \geq n$ à coefficients dans $\mathbb{R}$ (ou $\mathbb{C}$). La matrice A définit une application de $\mathbb{R}^n$ dans $\mathbb{R}^m$. On suppose que rang $A = n$. L'image de la boule S par A, notée AS est une hyperellipse dans $\mathbb{R}^m$.

On définit les n **valeurs singulières de** A comme les longueurs σ_i des n semi-axes principaux de AS. Par convention, on numérote les valeurs singulières par ordre décroissant: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$. On définit les n **vecteurs singuliers à gauche de** A ou les n **vecteurs de sortie de** A, $\{u_1, u_2, \ldots, u_n\}$ orientés dans les direction des semi-axes principaux, où u_i est la direction de du semi-axe de longueur σ_i. On définit les n **vecteurs singuliers à droite de** A ou les n **vecteurs d'entrée de** A, $\{v_1, v_2, \ldots, v_n\}$, $v_i \in S$ qui sont les pré-images des semi-axes principaux: $Av_i = \sigma_i u_i$ pour $i = 1, \ldots, n$. Sous forme matricielle:

$$A \left[\begin{array}{c|c|c|c} \\ v_1 & v_2 & \cdots & v_n \\ \\ \end{array} \right] = \left[\begin{array}{c|c|c|c} \\ \\ u_1 & u_2 & \cdots & u_n \\ \\ \\ \end{array} \right] \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix},$$

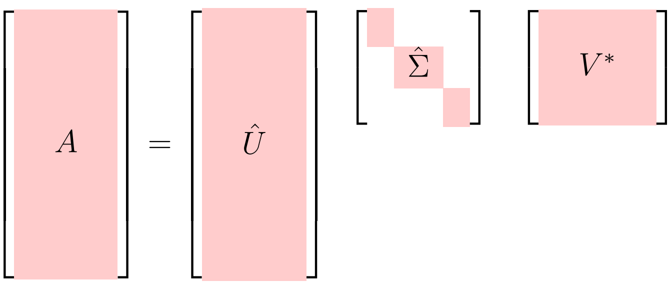
ou bien

$$AV = \hat{U}\hat{\Sigma}. \tag{3}$$

La matrice $\hat{\Sigma}$ est une matrice $n \times n$, diagonale avec coefficients positifs et réels. $\hat{U}$ est une matrice $m \times n$ avec colonnes orthonormales et V est une matrice $n \times n$ avec colonnes orthonormales. V est unitaire et on peut réécrire l'équation

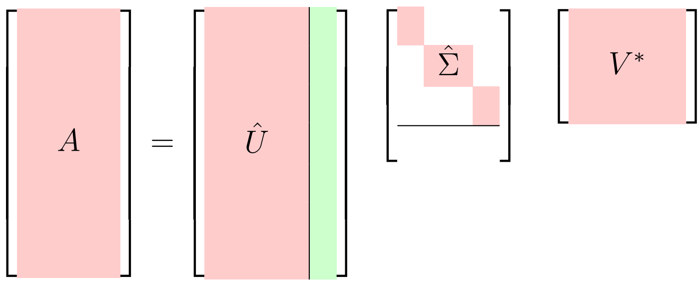
$$A = \hat{U}\hat{\Sigma}V^*. \tag{4}$$

Cette factorisation est appelée **décomposition en valeurs singulières réduite**.



3.3 Décomposition en valeurs singulières complète

$\hat{U}$ est une matrice $m \times n$ et, sauf si $m = n$, les colonnes de $\hat{U}$ ne forment pas une base de $\mathbb{C}^m$. En ajoutant $m - n$ colonnes orthonormales manquantes à $\hat{U}$, on peut en faire une matrice unitaire, que l'on appellera U. Si $\hat{U}$ est remplacée par U dans la factorisation, la matrice $\hat{\Sigma}$ doit être augmentée, en ajoutant $m-n$ lignes de zéros. On obtient alors une **décomposition en valeurs singulières complète**:



ou encore

$$A = U\Sigma V^*. \tag{5}$$

Si rang $A = r < n$, la factorisation se fait de la même façon, mais avec r vecteurs singuliers à gauche. La matrice V aura besoin de $n-r$ vecteurs additionnels.

Définition 1. Une **décomposition en valeurs singulières** (SVD – **Singular Value Decomposition**) est une factorisation $A = U\Sigma V^*$ où

- $U \in \mathbb{C}^{m\times m}$ est unitaire
- $V \in \mathbb{C}^{n\times n}$ est unitaire
- $\Sigma \in \mathbb{C}^{m\times n}$ est diagonale et les coefficients $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p$, où $p = \min(m,n)$.

$U\Sigma V^* S$ est une ellipse: V^* préserve la sphère ($||V^*x|| = ||x||$), Σ étire la sphère dans chaque direction de la base canonique et U préserve l'ellipse (une matrice unitaire ne fait que des rotations ou des réflections).

Théorème 3.1. existence et unicité de la SVD. *Toute matrice $A \in \mathbb{C}^{n\times n}$ possède une SVD. Les valeurs singulières $\{\sigma_i\}$ sont déterminées de façon unique. Si A est carrée ($m = n$) et les valeurs singulières σ_j sont distinctes, les vecteurs d'entrée et de sorties v_j et u_j sont déterminés de façon unique à un facteur complexe unité près.*

Théorème 3.2. *Les valeurs singulières d'une matrice A sont les racines carrées des valeurs propres non-nulles de A^*A et AA^**

Preuve. $A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^* U^* U\Sigma V^* = V\Sigma^*\Sigma V^*$. La matrice A^*A est donc semblable à $\Sigma^*\Sigma$, ce qui implique qu'elles ont les mêmes valeurs propres. Les valeurs propres de $\Sigma^*\Sigma$ sont $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$.

Théorème 3.3. *Les colonnes de U sont les vecteurs propres orthogonaux de AA^* et les colonnes de V sont les vecteurs propres orthogonaux de A^*A à unité complexe près.*

Éléments de la preuve. Les matrices AA^* et A^*A sont **hermitiennes**, c.-à-d. qu'elles sont auto-adjointes, ou égales à leurs adjointes. Les matrices hermitiennes sont diagonalisables, leurs valeurs propres sont réelles et positives et les vecteurs propres forment un ensemble orthogonal.

Exemples Calculs de SVD à la main.

Exemple 1.

$$A = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}. \tag{6}$$

i) Calcul des valeurs singulières. Les valeurs propres de

$$A^*A = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}. \tag{7}$$

sont 9 et 4, d'où les valeurs singulières $\sigma_1 = 3$ et $\sigma_2 = 2$.

ii) Calcul de V et U. Les vecteurs propres de AA^* sont $x_1 = {}^t(1,0)$ et $x_2 = {}^t(0,1)$. Les vecteurs propres de A^*A sont $y_1 = {}^t(1,0)$ et $y_2 = {}^t(0,1)$. Ces vecteurs propres déterminent les colonnes de U et V a constante près: $v_i = \mu_i y_i$ et $u_i = \mu_i' x_i$. On a, par la décomposition en valeurs singulières $Av_1 = \sigma_1 u_1$, ou $A\mu_1 y_1 = \sigma\mu_1' x_1$. On peut choisir μ_1 et μ_1' de façon, par exemple, a minimiser le nombre de signe négatifs dans les matrices U et V. Ici, comme les vecteurs propres son réels, les facteurs $\mu = \pm 1$. En prenant $\mu_1' = 1$, on a $u_1 = x_1$ et $A\mu y_1 = \sigma u_1$:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \mu_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3\mu_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3\mu_1 \\ 0 \end{bmatrix} \tag{8}$$

d'où $\mu_1 = 1$ et $v_1 = y_1$. Pour v_2 et u_2 on a $A\mu y_2 = \sigma_2 \mu_2' x_2$. En prenant $\mu_2' = 1$ on obtient

$$A = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \mu_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2\mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \tag{9}$$

d'où $\mu_2 = -1$, et $v_2 = {}^t(0,-1)$. La décomposition en valeurs singulières est donc

$$A = U\Sigma V^*, \quad \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{10}$$

Exemple 2.

$$A = \begin{bmatrix} 0 & 3 \\ -2 & 0 \\ 0 & 0 \end{bmatrix}. \tag{11}$$

La matrice est de taille 3×2 et est de rang 2, on cherchera donc 2 valeurs singulières.
i) Calcul des valeurs singulières. La matrice A^*A est

$$\begin{bmatrix} 0 & -2 & 0 \\ 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ -2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \tag{12}$$

Les valeurs propres sont 9 et 4, donc $\sigma_1 = 3$ et $\sigma_2 = 2$.

ii) Calcul de V et U. Les vecteurs propres de AA^* (une matrice 3 par 3) sont $x_1 = {}^t(1, 0, 0)$, $x_2 = {}^t(0, 1, 0)$ et $x_3 = {}^t(0, 0, 1)$. Les vecteurs propres de A^*A sont $y_1 = {}^t(0, 1)$ et $y_2 = {}^t(1, 0)$. On choisit les constantes μ de façon à avoir des signes positifs dans U: $\mu_1' = 1$ et $\mu_2' = 1$. On a alors $A\mu_1 y_1 = \sigma u_1$:

$$\begin{bmatrix} 0 & 3 \\ -2 & 0 \\ 0 & 0 \end{bmatrix} \mu_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3\mu 1 \\ 0 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{13}$$

d'où $\mu_1 = 1$. De la même façon, $A\mu_2 y_2 = \sigma_2 x_2$ implique que $\mu_2 = -1$. La décomposition est donc

$$A = U\Sigma V^*, \quad \begin{bmatrix} 0 & 3 \\ -2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{14}$$

References

[1] Jean Fresnel. *Algebre des matrices*. Hermann, 2013.

[2] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*, volume 50. Siam, 1997.