

## ML 10 Méthode de la puissance itérée et SVD

Nous montrons comment adapter la méthode de la puissance itérée au calcul de la plus grande valeur singulière et de ses deux vecteurs singuliers associés. Nous considérons une matrice de données rectangulaire  $\mathbf{A} \in \mathbb{R}^{N \times M}$ . Nous notons  $\mathbf{U} \in \mathbb{R}^{N \times N}$  et  $\mathbf{V} \in \mathbb{R}^{M \times M}$  les deux matrices orthogonales qui contiennent les vecteurs singuliers en colonnes. Nous notons enfin  $\mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots) \in \mathbb{R}^{N \times M}$  la matrice qui contient sur sa diagonale principale les valeurs singulières, avec  $r$  le rang de  $\mathbf{A}$ . Nous avons donc :

$$\begin{aligned} \mathbf{A} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T && r \text{ est le rang de } \mathbf{A}. \\ \mathbf{U} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N] && \text{avec } \mathbf{u}_i \in \mathbb{R}^N, \text{ base orthonormale de } \mathbb{R}^N \\ \mathbf{V} &= [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M] && \text{avec } \mathbf{v}_j \in \mathbb{R}^M, \text{ base orthonormale de } \mathbb{R}^M \\ \mathbf{A} &= \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^T \end{aligned}$$

Nous allons montrer que la séquence ci-dessous converge vers  $\sigma_1$ ,  $\mathbf{u}$  et  $\mathbf{v}$  tels que  $\mathbf{A}\mathbf{v} = \sigma_1\mathbf{u}$ .

$$\begin{aligned} \mathbf{w}^{(k)} &= \mathbf{A}\mathbf{v}^{(k-1)} && \alpha_k = \|\mathbf{w}^{(k)}\|^{-1} && \mathbf{u}^{(k)} = \alpha_k \mathbf{w}^{(k)} \\ \mathbf{z}^{(k)} &= \mathbf{A}^T \mathbf{u}^{(k)} && \beta_k = \|\mathbf{z}^{(k)}\|^{-1} && \mathbf{v}^{(k)} = \beta_k \mathbf{z}^{(k)} \\ err &= \|\mathbf{A}\mathbf{v}^{(k)} - \beta_k \mathbf{u}^{(k)}\| && \sigma_1 = \beta_k \end{aligned}$$

Nous décomposons  $\mathbf{v}^{(0)} \in \mathbb{R}^M$  dans la base orthonormée formée par les vecteurs  $\mathbf{v}_j$  :

$$\mathbf{v}^{(0)} = \sum_{j=1}^M y_j \mathbf{v}_j \quad \text{avec } y_j = \mathbf{v}_j^T \mathbf{v}^{(0)}$$

Partant de ce  $\mathbf{v}^{(0)}$ , nous calculons les deux premières itérations :

$$\begin{aligned} \mathbf{w}^{(1)} &= \mathbf{A}\mathbf{v}^{(0)} = \sum_{j=1}^r \sigma_j y_j \mathbf{u}_j && \mathbf{u}^{(1)} &= \alpha_1 \sum_{j=1}^r \sigma_j y_j \mathbf{u}_j \\ \mathbf{z}^{(1)} &= \mathbf{A}^T \mathbf{u}^{(1)} = \alpha_1 \sum_{j=1}^r \sigma_j^2 y_j \mathbf{v}_j && \mathbf{v}^{(1)} &= \alpha_1 \beta_1 \sum_{j=1}^r \sigma_j^2 y_j \mathbf{v}_j \\ \mathbf{w}^{(2)} &= \mathbf{A}\mathbf{v}^{(1)} = \alpha_1 \beta_1 \sum_{j=1}^r \sigma_j^3 y_j \mathbf{u}_j && \mathbf{u}^{(2)} &= \alpha_2 \alpha_1 \beta_1 \sum_{j=1}^r \sigma_j^3 y_j \mathbf{u}_j \\ \mathbf{z}^{(2)} &= \mathbf{A}^T \mathbf{u}^{(2)} = \alpha_2 \alpha_1 \beta_1 \sum_{j=1}^r \sigma_j^4 y_j \mathbf{v}_j && \mathbf{v}^{(2)} &= \alpha_2 \beta_2 \alpha_1 \beta_1 \sum_{j=1}^r \sigma_j^4 y_j \mathbf{v}_j \end{aligned}$$

Par induction, nous avons :

$$\mathbf{v}^{(\mathbf{k})} = \delta_{2k} \sum_{j=1}^r \sigma_j^{2k} y_j \mathbf{v}_j$$

$$\text{Avec } \delta_{2k} = \alpha_k \beta_k \alpha_{k-1} \beta_{k-1} \dots \alpha_1 \beta_1$$

$$\mathbf{u}^{(\mathbf{k})} = \delta_{2k-1} \sum_{j=1}^r \sigma_j^{2k-1} y_j \mathbf{u}_j$$

$$\text{Avec } \delta_{2k-1} = \alpha_{k+1} \alpha_k \beta_k \alpha_{k-1} \beta_{k-1} \dots \alpha_1 \beta_1$$

Par ailleurs, nous avons :

$$\begin{aligned} & 1 \\ &= \{ \text{A chaque itération, la norme de } \mathbf{u}^{(\mathbf{k})} \text{ est maintenue égale à 1.} \} \\ & \quad \|\mathbf{u}^{(\mathbf{k})}\|^2 \\ &= \{ \text{Par définition de } \mathbf{u}^{(\mathbf{k})} \text{ et car } \|\mathbf{u}_j\| = 1 \} \\ & \quad \delta_{2k-1}^2 \sum_{j=1}^r \sigma_j^{4k-2} y_j^2 \end{aligned}$$

De même :

$$\begin{aligned} & 1 \\ &= \\ & \quad \|\mathbf{v}^{(\mathbf{k})}\|^2 \\ &= \\ & \quad \delta_{2k}^2 \sum_{j=1}^r \sigma_j^{4k} y_j^2 \end{aligned}$$

Donc :

$$\frac{\|\mathbf{u}^{(\mathbf{k}+1)}\|^2}{\|\mathbf{v}^{(\mathbf{k})}\|^2} = 1 = \sigma_1^2 \left( \frac{\delta_{2k+1}^2}{\delta_{2k}^2} \right) \left( \frac{\sum_{j=1}^{n_1} y_j^2 + \sum_{j=n_1}^r \left( \frac{\sigma_j}{\sigma_1} \right)^{4k+2} y_j^2}{\sum_{j=1}^{n_1} y_j^2 + \sum_{j=n_1}^r \left( \frac{\sigma_j}{\sigma_1} \right)^{4k} y_j^2} \right)$$

Ci-dessus,  $n_1$  est la multiplicité de  $\sigma_1$ , c'est-à-dire le nombre de vecteurs singuliers indépendants associés à cette valeur singulière. En pratique, sur de "vraies" données, cette multiplicité sera presque toujours égale à 1.

Nous avons ainsi montré que :

$$\lim_{k \rightarrow \infty} \frac{\delta_{2k+1}}{\delta_{2k}} = \sigma_1$$

Et :

$$\begin{aligned} & \lim_{k \rightarrow \infty} \|\mathbf{A}\mathbf{v}^{(\mathbf{k})} - \sigma_1 \mathbf{u}^{(\mathbf{k})}\| \\ &= \left\{ \mathbf{A}\mathbf{v}^{(\mathbf{k})} = \frac{\delta_{2k+1}}{\delta_{2k}} \mathbf{u}^{(\mathbf{k}+1)} \right\} \\ & 0 \end{aligned}$$