13. Décomposition en valeurs singulières (SVD)

MTH1008

Sébastien Le Digabel
Polytechnique Montréal

H2024

2024-04-17

v3

Plan

1. Décomposition en valeurs singulières (SVD)

2. Applications et exemples

Liens avec le livre et exercices suggérés

7.4	Décomposition en valeurs singulières	1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 24, 25

Rappels

1. Si $\mathbf{A}$ est carrée $n \times n$ avec n vecteurs propres indépendants alors $\mathbf{A}$ est diagonalisable : $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{\Lambda}$ et $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$

2. Si $\mathbf{A}$ est symétrique :
   - Elle est diagonalisable (théorème spectral)
   - Ses vecteurs propres peuvent être choisis orthonormaux
   - $\mathbf{Q}^\top\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ et $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$

3. Si $\mathbf{A}$ est symétrique et semi-définie positive alors $\mathbf{Q}^\top\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ où $\lambda_i \geq 0$ pour tout $i \in \{1, 2, \ldots, n\}$

SVD : Diagonalisation d'une matrice quelconque $m \times n$, avec des vecteurs *singuliers* orthogonaux

Théorème : Décomposition SVD

Une matrice $\mathbf{A}$ de taille $m \times n$ et de rang r peut décomposer A comme suit :

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^\top$$

avec $\mathbf{U} \in \mathbb{R}^{m \times m}$ et $\mathbf{V} \in \mathbb{R}^{n \times n}$ deux matrices **orthogonales**, et $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ une matrice “diagonale” formée par les r *valeurs singulières* $\sigma_1, \sigma_2, \ldots, \sigma_r$

Valeurs singulières

Soit $\mathbf{A} \in \mathbb{R}^{m\times n}$ de rang r. La matrice $\mathbf{A}^\top\mathbf{A} \in \mathbb{R}^{n\times n}$ est symétrique et **semi-définie positive**. Soient :

- $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ les valeurs propres de $\mathbf{A}^\top\mathbf{A}$. Les r premières sont strictement positives. Les $n-r$ suivantes sont nulles
- Les σ_i sont les *valeurs singulières* de $\mathbf{A}$ (pour $i = 1, 2, \ldots, r$)
- $\sigma_i = \sqrt{\sigma_i^2} \geq 0$: Les valeurs singulières sont toujours positives ou nulles

Vecteurs singuliers (deux ensembles : u et v)

Soient de plus :

- $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_r \in \mathbb{R}^n$ les vecteurs propres correspondant aux valeurs propres strictement positives de $\mathbf{A}^\top \mathbf{A}$
  Ils peuvent être pris comme orthonormaux, forment une base de $\mathrm{Im}(\mathbf{A}^\top)$, et sont appelés les *vecteurs singuliers* de $\mathbf{A}$
- $\mathbf{v}_{r+1}, \mathbf{v}_{r+2} \ldots, \mathbf{v}_n \in \mathbb{R}^n$ une base orthonormale de $\mathrm{Ker}(\mathbf{A})$
- $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \in \mathbb{R}^m$ les vecteurs unitaires de $\mathrm{Im}(\mathbf{A})$ définis par
  $$\mathbf{u}_i = \frac{1}{\sigma_i} \mathbf{A} \mathbf{v}_i \qquad \text{pour } i = 1, 2, \ldots, r$$
  Ils sont orthonormaux, forment une base de $\mathrm{Im}(\mathbf{A})$, et des vecteurs propres de $\mathbf{A}\mathbf{A}^\top$. Ils sont aussi appelés singuliers
- $\mathbf{u}_{r+1}, \mathbf{u}_{r+2} \ldots, \mathbf{u}_m \in \mathbb{R}^m$ une base orthonormale de $\mathrm{Ker}(\mathbf{A}^\top)$

Diagonalisation $\mathbf{A}\mathbf{V}_r = \mathbf{U}_r\boldsymbol{\Sigma}_r$

- $\mathbf{A}$ est diagonalisée avec

$$\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1 \ , \ \mathbf{A}\mathbf{v}_2 = \sigma_2\mathbf{u}_2 \ , \ \ldots \ , \ \mathbf{A}\mathbf{v}_r = \sigma_r\mathbf{u}_r$$

- Sous forme matricielle : $\mathbf{A}\mathbf{V}_r = \mathbf{U}_r\boldsymbol{\Sigma}_r$:

$$\mathbf{A}\begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

$$(m \times n)(n \times r) = (m \times r)(r \times r)$$

Complétion de U, V, et Σ

- On complète $\mathbf{U}_r$ et $\mathbf{V}_r$ avec $\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \ldots, \mathbf{u}_m$ et $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n$ pour obtenir les matrices $\mathbf{U}$ et $\mathbf{V}$ carrées et orthogonales
- Les matrices $\mathbf{U}$ et $\mathbf{V}$ contiennent alors des bases orthonormales pour les quatre sous-espaces :

$$\begin{array}{rlr} r & \text{premières colonnes de } \mathbf{V}: & \mathrm{Im}(\mathbf{A}^\top) \\ n-r & \text{dernières colonnes de } \mathbf{V}: & \mathrm{Ker}(\mathbf{A}) \\ r & \text{premières colonnes de } \mathbf{U}: & \mathrm{Im}(\mathbf{A}) \\ m-r & \text{dernières colonnes de } \mathbf{U}: & \mathrm{Ker}(\mathbf{A}^\top) \end{array}$$

- On passe aussi de $\mathbf{\Sigma}_r \in \mathbb{R}^{r\times r}$ à $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ en ajoutant $m-r$ lignes de zéros et $n-r$ colonnes de zéros

Décomposition $\mathrm{A} = \mathrm{U}\Sigma\mathrm{V}^\top$ (SVD)

On a

$$\mathbf{AV} = \mathbf{U\Sigma},\ \mathbf{A} = \mathbf{U\Sigma V}^\top,\ \text{et}\ \mathbf{\Sigma} = \mathbf{U}^\top\mathbf{AV}$$

avec les matrices

$$\mathbf{V} = \left[\begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{array}\right] \in \mathbb{R}^{n\times n}$$

$$\mathbf{U} = \left[\begin{array}{cccc} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{array}\right] \in \mathbb{R}^{m\times m}$$

$$\mathbf{\Sigma} = \left[\begin{array}{cccc|c} \sigma_1 & & \mathbf{0} & & \\ & \sigma_2 & & & \\ & & \ddots & & \mathbf{0} \\ \mathbf{0} & & & \sigma_r & \\ \hline & & \mathbf{0} & & \mathbf{0} \end{array}\right] \in \mathbb{R}^{m\times n}$$

1. Décomposition en valeurs singulières (SVD)

2. Applications et exemples

Exemples

Donner la décomposition SVD des matrices

$$\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$$

et

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$$

Conditionnement d'une matrice

- Le *conditionnement* de la matrice $\mathbf{A}$, noté $\kappa(\mathbf{A})$, mesure à quel point la solution de $\mathbf{Ax} = \mathbf{b}$ change lorsque $\mathbf{b}$ change
- C'est la *sensibilité* de $\mathbf{Ax} = \mathbf{b}$ relativement à $\mathbf{b}$
- Il est préférable d'avoir $\kappa(\mathbf{A})$ le plus petit possible afin de minimiser l'erreur numérique sur $\mathbf{x}$ lors de la résolution de $\mathbf{Ax} = \mathbf{b}$
- Avec une norme matricielle $\|.\|$, on a $\kappa(\mathbf{A}) = \|\mathbf{A}^{-1}\| \times \|\mathbf{A}\|$
- Avec les valeurs singulières, on a $\kappa(\mathbf{A}) = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}$ où $\sigma_{\min}(\mathbf{A})$ et $\sigma_{\max}(\mathbf{A})$ sont les plus petite et plus grande valeurs singulières, respectivement

Application : Approximation d'une matrice pour la compression (avec pertes)

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^\top$$

peut s'écrire

$$\mathbf{A} = \sigma_{(1)} \mathbf{u}_{(1)} \mathbf{v}_{(1)}^\top + \sigma_{(2)} \mathbf{u}_{(2)} \mathbf{v}_{(2)}^\top + \cdots + \sigma_{(r)} \mathbf{u}_{(r)} \mathbf{v}_{(r)}^\top$$

avec (1) l'indice de la plus grande valeur singulière, (2) l'indice de la deuxième plus grande valeur singulière, etc.

- Chaque matrice $\sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ est de rang 1
- Toute troncature à droite de cette somme donnera une approximation de $\mathbf{A}$

Exemple pour la compression d'images : https://timbaumann.info/svd-image-compression-demo/