

INTRODUCTION

Une matrice A de taille $m \times n$ représente une TRANSFORMATION LINÉAIRE de \mathbb{R}^n à \mathbb{R}^m , en considérant la fonction $\mathbf{x} \mapsto A\mathbf{x}$.

Pour des matrices carrées, cette transformation se comprend très bien en considérant les valeurs et vecteurs propres de A : c'est la théorie qui soutient notre analyse de systèmes dynamiques et chaînes de Markov, entre autres. Cette théorie dépend d'une diagonalisation de A , qui n'est pas toujours possible. Certainement ce n'est jamais possible pour une matrice non-carrée.

VALEURS SINGULIÈRES

Soit A une matrice complètement arbitraire, de taille $m \times n$. Pour un vecteur $\mathbf{x} \in \mathbb{R}^n$ avec $\|\mathbf{x}\| = 1$, on considère $\|A\mathbf{x}\|$. Cette norme mesure le changement en longueur entre \mathbf{x} et $A\mathbf{x}$. Note que

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T (A\mathbf{x}) = \mathbf{x}^T A^T A \mathbf{x}$$

Ceci suggère de considérer un vecteur propre de $A^T A$, c'est-à-dire un \mathbf{x} avec $A^T A \mathbf{x} = \lambda \mathbf{x}$.

La matrice $A^T A$ est carrée ($n \times n$) et symétrique (puisque $(A^T A)^T = A^T (A^T)^T = A^T A$). Donc on peut trouver une décomposition $A^T A = P D P^T$ et alors une base orthogonale de \mathbb{R}^n formée de vecteurs propres de $A^T A$. Dénotons cette base par $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, où $A^T A \mathbf{v}_j = \lambda_j \mathbf{v}_j$.

Théorème 12.1. *Les valeurs propres de $A^T A$ sont réelles et non-négatives.*

Preuve. Si $A^T A \mathbf{x} = \lambda \mathbf{x}$ alors $\mathbf{x}^T A^T A \mathbf{x} = \lambda \mathbf{x}^T \mathbf{x}$ et donc $\lambda = \|A\mathbf{x}\| / \|\mathbf{x}\|$. Les deux normes sont non-négatives, et le dénominateur n'est pas nul, donc $\lambda \geq 0$. \square

On considère les valeurs propres en ordre décroissant, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Si le rang de $A^T A$ est r , alors $\lambda_j = 0$ pour $j > r$: les dernières $n - r$ valeurs propres sont 0.

Les VALEURS SINGULIÈRES de A sont $\sigma_1, \sigma_2, \dots, \sigma_r$, où $\sigma_i = \sqrt{\lambda_i}$. Ce sont les racines carrées des valeurs propres non-nulles de $A^T A$.

Exemple 12.2. Soit $A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$. Calculer les valeurs singulières de A , et de A^T .

Pour A , on a $A^T A = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$. On calcul

$$\begin{aligned} \det(A^T A - \lambda I) &= \begin{vmatrix} 11 - \lambda & 1 \\ 1 & 11 - \lambda \end{vmatrix} \\ &= (\lambda - 11)(\lambda - 11) - 1 \\ &= \lambda^2 - 22\lambda + 120 \\ &= (10 - \lambda)(\lambda - 12) \end{aligned}$$

Les valeurs propres sont 12, 10. Les valeurs singulières de A^T sont alors $\sigma_1 = \sqrt{12}$ et $\sigma_2 = \sqrt{10}$.

Pour A^T on a $(A^T)(A^T)^T = AA^T = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$. On calcul

$$\begin{aligned} \det(AA^T - \lambda I) &= \begin{vmatrix} 10 - \lambda & 0 & 2 \\ 0 & 10 - \lambda & 4 \\ 2 & 4 & 2 - \lambda \end{vmatrix} \\ &= (10 - \lambda) \det \begin{bmatrix} 10 - \lambda & 4 \\ 4 & 2 - \lambda \end{bmatrix} + 2 \det \begin{bmatrix} 0 & 10 - \lambda \\ 2 & 4 \end{bmatrix} \\ &= (10 - \lambda) \left((10 - \lambda)(2 - \lambda) - 16 \right) + 2 \left((10 - \lambda)(2) - 0 \right) \\ &= (10 - \lambda) \left((10 - \lambda)(2 - \lambda) - 16 + 4 \right) \\ &= (10 - \lambda)(\lambda)(\lambda - 12) \end{aligned}$$

Les valeurs propres sont 12, 10, 0. Les valeurs singulières de A sont $\sigma_1 = \sqrt{12}$ et $\sigma_2 = \sqrt{10}$.

DÉCOMPOSITION

On observe dans l'exemple 12.2 que A et A^T avaient les mêmes valeurs singulières. Ce n'est pas par hasard...

Théorème 12.3. *Soit A une matrice de taille $m \times n$. Les matrices $A^T A$ et AA^T ont exactement les mêmes valeurs propres non-nulles.*

Preuve. Soit \mathbf{x} un vecteur propre de $A^T A$ avec $A^T A \mathbf{x} = \lambda \mathbf{x}$ et $\lambda \neq 0$. Alors

$$\begin{aligned} A(A^T A \mathbf{x}) &= A(\lambda \mathbf{x}) \\ AA^T(A \mathbf{x}) &= \lambda(A \mathbf{x}) \end{aligned}$$

Aussi, $A \mathbf{x} \neq \mathbf{0}$, car sinon on aurait $\lambda \mathbf{x} = A(A^T \mathbf{x}) = A \mathbf{0} = \mathbf{0}$. Donc $A \mathbf{x}$ est vecteur propre de AA^T avec valeur propre λ . Si \mathbf{y} est vecteur propre de AA^T , l'analyse est similaire. \square

Ce théorème est à comparer avec le théorème 10.9. C'est un théorème où la preuve est encore plus utile que le résultat. On y voit que étant donné des vecteurs propres de $A^T A$, on obtient des vecteurs propres de AA^T en multipliant par A . Il y a encore plus.

Théorème 12.4. *Soit A une matrice de taille $m \times n$. Soit \mathbf{x} et \mathbf{y} des vecteurs propres de $A^T A$ avec $A^T A \mathbf{x} = \lambda \mathbf{x}$, $A^T A \mathbf{y} = \mu \mathbf{y}$ et $\lambda, \mu \neq 0$. Si $\mathbf{x} \perp \mathbf{y}$ alors $(A \mathbf{x}) \perp (A \mathbf{y})$.*

Preuve. Premièrement, on observe que si $\lambda \neq \mu$, alors on a automatiquement $\mathbf{x} \perp \mathbf{y}$, donc c'est un hypothèse très raisonnable (le résultat s'applique aussi si $\lambda = \mu$). On note que $A \mathbf{x} \neq \mathbf{0}$, car sinon on aurait $\lambda = 0$; de la même manière $A \mathbf{y} \neq \mathbf{0}$. Être perpendiculaire veut donc dire que le produit scalaire est zéro. On calcul directement:

$$(A \mathbf{x}) \cdot (A \mathbf{y}) = (A \mathbf{x})^T (A \mathbf{y}) = \mathbf{x}^T A^T A \mathbf{y} = \lambda (\mathbf{x}^T \mathbf{y}) = \lambda (\mathbf{x} \cdot \mathbf{y}) = \lambda (0) = 0 \quad \square$$

La conséquence c'est que un ensemble orthogonal de vecteurs propres de $A^T A$ devient un ensemble orthogonal de vecteurs propres de AA^T en multipliant chaque vecteur par A — pourvu que les valeurs propres correspondants sont non-nulles (autrement on obtient des vecteurs nuls en multipliant par A). Cette observation est la base d'une DÉCOMPOSITION EN VALEURS SINGULIÈRES de A .

Théorème 12.5. Soit A une matrice de taille $m \times n$. Alors il existe des matrices U, Σ, V tel que $A = U\Sigma V^T$ avec Σ diagonale, $U^T U = I$ et $V^T V = I$. \square

Note que U est nécessairement de taille $m \times m$, V est $n \times n$ et Σ est $m \times n$. Que $U^T U = I$ signifie que les colonnes de U forment une base orthonormale de \mathbb{R}^m ; que $V^T V = I$ signifie que les colonnes de V forment une base orthonormale de \mathbb{R}^n . Comme preuve on présente l'algorithme suivant, qui *construit* les matrices requises.

Algorithme 12.6 (Décomposition en valeurs singulières).

Soit A une matrice de taille $m \times n$.

1. On calcul $A^T A$, et on forme une décomposition $A^T A = P D P^T$.
2. On identifie les valeurs propres non nulles de $A^T A$: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$.
3. Les valeurs singulières de A sont $\sigma_1, \sigma_2, \dots, \sigma_r$, où $\sigma_j = \sqrt{\lambda_j}$.
4. On pose Σ la matrice de taille $m \times n$ avec les valeurs singulières sur le diagonal.
5. On pose V la matrice P .
6. On identifie les colonnes $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ de V , correspondant aux valeurs singulières.
7. On pose $\mathbf{u}_j = A\mathbf{v}_j / \|A\mathbf{v}_j\| = A\mathbf{v}_j / \sigma_j$.
8. Si $r < m$ on étend l'ensemble $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ à une base orthonormale de \mathbb{R}^m .
9. On pose U la matrice avec les colonnes $\mathbf{u}_1, \dots, \mathbf{u}_m$.

On a alors $A = U\Sigma V^T$. \square

Il faudrait démontrer que cette algorithme fonctionne, c'est-à-dire, que les matrices U , Σ et V ont les bonnes propriétés. On voit directement que $U^T U = I$ et $V^T V = I$ car les colonnes de U forment une base orthonormale pour \mathbb{R}^m et les colonnes de V forment une base orthonormale pour \mathbb{R}^n . Considérons le produit AV . La colonne j est $A\mathbf{v}_j$, qui est $\sigma_j \mathbf{u}_j$, qui est la colonne j du produit $U\Sigma$. Donc $AV = U\Sigma$. Multipliant par V^T , on obtient $A = U\Sigma V^T$.

Les étapes à suivre pour faire une décomposition en valeurs singulières sont des étapes qu'on connaît déjà. On souligne que la dernière étape peut se faire de deux manières. Les vecteurs $\mathbf{u}_{r+1}, \dots, \mathbf{u}_m$ forment une base pour l'espace nul de AA^T . Alternativement, c'est une base pour le complément orthogonale de $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$.

Exemple 12.7. Trouver une décomposition en valeurs singulières de $A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$.

On calcul premièrement $A^T A = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$. On obtient $A^T A = P D P^T$ avec

$$A = P D P^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 12 & \\ & 10 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$$

On pose donc

$$V = P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \Sigma = \begin{bmatrix} 12 & 0 \\ 0 & 10 \\ 0 & 0 \end{bmatrix}$$

On calcule les vecteurs \mathbf{u}_1 et \mathbf{u}_2 .

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1} = \frac{1}{\sqrt{12}} \begin{bmatrix} 2/\sqrt{2} \\ 4/\sqrt{2} \\ 2/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2} = \frac{1}{\sqrt{10}} \begin{bmatrix} -4/\sqrt{2} \\ 2/\sqrt{2} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

Ici on voit que $r = 2 < m = 3$. Donc il manque le vecteur \mathbf{u}_3 . On peut procéder de deux manières. Soit on cherche un vecteur orthogonal à \mathbf{u}_1 et \mathbf{u}_2 , donc on met ces deux vecteurs comme rangées d'une matrice B et on cherche l'espace nul de cette matrice. On peut prendre des multiples de chaque vecteur pour simplifier les calculs.

$$B = \begin{bmatrix} 1 & 2 & 1 \\ -2 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 2/5 \end{bmatrix} \quad \text{base pour } \ker(B): \left\{ \begin{bmatrix} -1/5 \\ -2/5 \\ 1 \end{bmatrix} \right\}$$

Cette première méthode à une version alternative, pour ceux qui n'aiment pas lire des solutions générales dans des matrices. On forme la matrice C ayant \mathbf{u}_1 et \mathbf{u}_2 comme colonnes et on réduit la matrice $C|I$. Les rangées nulles à gauche indiquent quelles rangées à droite forment la base de $\ker(C^T) = \ker(B)$.

$$\left[\begin{array}{cc|ccc} 1 & -2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|ccc} 1 & -2 & 1 & 0 & 0 \\ 0 & 5 & -2 & 1 & 0 \\ 0 & 0 & -1/5 & -2/5 & 1 \end{array} \right] \quad \text{base: } \left\{ \begin{bmatrix} -1/5 \\ -2/5 \\ 1 \end{bmatrix} \right\}$$

Cette base est une base orthogonale, car ce n'est qu'un seul vecteur. Donc Gram-Schmidt n'est pas requis. On normalise pour obtenir

$$\mathbf{u}_3 = \frac{1}{\sqrt{30}} \begin{bmatrix} -1 \\ -2 \\ 5 \end{bmatrix}$$

Alternativement, on pourrait trouver \mathbf{u}_3 en cherchant une base orthonormale pour $\ker(AA^T)$.

$$AA^T = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 2/5 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{base pour } \ker(AA^T): \left\{ \begin{bmatrix} -1/5 \\ -2/5 \\ 1 \end{bmatrix} \right\}$$

Encore une fois c'est déjà une base orthogonale, donc aucun Gram-Schmidt. On normalise pour obtenir

$$\mathbf{u}_3 = \frac{1}{\sqrt{30}} \begin{bmatrix} -1 \\ -2 \\ 5 \end{bmatrix}$$

On a maintenant la matrice U , formée des trois colonnes \mathbf{u}_1 , \mathbf{u}_2 et \mathbf{u}_3 .

$$U = \begin{bmatrix} 1/\sqrt{6} & -2\sqrt{5} & 1/\sqrt{30} \\ 2/\sqrt{6} & 1\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & -5/\sqrt{30} \end{bmatrix}$$

Ceci donne la décomposition en valeurs singulières

$$\begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & -2\sqrt{5} & 1/\sqrt{30} \\ 2/\sqrt{6} & 1\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & -5/\sqrt{30} \end{bmatrix} \begin{bmatrix} 12 & 0 \\ 0 & 10 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T \quad \square$$

L'exemple était un peu long. Par contre on a calculé le vecteur \mathbf{u}_3 de trois manières différentes. Typiquement, une fois suffit...

La décomposition en valeurs singulières n'est pas unique. On aurait pu prendre ici $-\mathbf{u}_3$ au lieu de \mathbf{u}_3 . Aussi, on aurait pu prendre $-\mathbf{v}_1$ au lieu de \mathbf{v}_1 (ceci aurait changé \mathbf{u}_1). En générale, si on a une valeur singulière de multiplicité plus grand que un, on pourrait avoir plusieurs différentes bases orthogonales pour l'espace propre correspondant de A^TA .

Exercice 12.8. Recalculer la décomposition en valeurs singulières de l'exemple précédent en utilisant $-\mathbf{v}_1$ au lieu de \mathbf{v}_1 . Comparer votre décomposition avec celle ci-haut. (indice: les calculs sont très similaires, on peut recycler...).

On se rappelle qu'on peut calculer les valeurs singulières de A en calculant les valeurs propres de la matrice A^TA ou la matrice AA^T . On peut dire plus: on peut calculer toute la décomposition en valeurs singulières par l'une ou l'autre de ces matrices. Il s'agit de "commencer à la gauche" au lieu de "commencer à la droite". Voici la version "gauche" de l'algorithme 12.9.

Algorithme 12.9 (Décomposition en valeurs singulières, version gauche).

Soit A une matrice de taille $m \times n$.

1. On calcul AA^T , et on forme une décomposition $AA^T = PDP^T$.
2. On identifie les valeurs propres non nulles de AA^T : $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$.
3. Les valeurs singulières de A sont $\sigma_1, \sigma_2, \dots, \sigma_r$, où $\sigma_j = \sqrt{\lambda_j}$.
4. On pose Σ la matrice de taille $m \times n$ avec les valeurs singulières sur le diagonal.
5. On pose U la matrice P .
6. On identifie les colonnes $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ de U , correspondant aux valeurs singulières.
7. On pose $\mathbf{v}_j = A^T \mathbf{u}_j / \|A^T \mathbf{u}_j\| = A \mathbf{u}_j / \sigma_j$.
8. Si $r < n$ on étend l'ensemble $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ à une base orthonormale de \mathbb{R}^n .
9. On pose V la matrice avec les colonnes $\mathbf{v}_1, \dots, \mathbf{v}_n$.

On a alors $A = U\Sigma V^T$.

Exercice 12.10. Calculer une décomposition en valeurs singulières de $A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$ en utilisant la méthode "gauche". Note que puisque $m > n$, alors la première étape est plus longue et la dernière méthode est plus courte, en comparaison avec la méthode "droite".

APPROXIMATION: COMPOSANTES PRINCIPALES

La décomposition en valeurs singulières permet une représentation plus compacte d'une matrice. Voici une première observation.

Théorème 12.11. Soit $A = U\Sigma V^T$. Soit U_r la matrice formée des colonnes $\mathbf{u}_1, \dots, \mathbf{u}_r$, V_r la matrice formée des colonnes $\mathbf{v}_1, \dots, \mathbf{v}_r$, et Σ_r la matrice diagonale $r \times r$ formée des valeurs singulières.

Alors $A = U_r \Sigma_r V_r^T$.

C'est une observation basée sur le fait que des rangées nulles de Σ "annulent" les colonnes correspondant de U , et les colonnes nulles de Σ "annulent" les colonnes correspondant de V (rangées de V^T).

Exemple 12.12. On considère la décomposition ci-haut. On vérifie que

$$\begin{aligned} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{6} & -2\sqrt{5} & 1/\sqrt{30} \\ 2/\sqrt{6} & 1\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & -5/\sqrt{30} \end{bmatrix} \begin{bmatrix} 12 & 0 \\ 0 & 10 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T \\ &= \begin{bmatrix} 1/\sqrt{6} & -2\sqrt{5} \\ 2/\sqrt{6} & 1\sqrt{5} \\ 1/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} 12 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T \quad \square \end{aligned}$$

C'est une première "approximation": on simplifie le produit matriciel en enlevant les parties qui contribuent absolument rien. Ce qui est plus utile c'est d'enlever les parties qui contribuent peu. Un résultat technique est nécessaire.

Théorème 12.13. Si $A = U\Sigma V^T$ alors $A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$. □

Note que le produit $\mathbf{u}_1 \mathbf{v}_1^T$ est le produit d'une matrice $m \times 1$ par une matrice $1 \times n$, donnant une matrice $m \times n$. La matrice $\sigma_j \mathbf{u}_j \mathbf{v}_j^T$ est la j ème composante principale de A . On dénote par

$$\widehat{A}_t = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_t \mathbf{u}_t \mathbf{v}_t^T$$

la somme des t premières composantes principales. Note que $\widehat{A}_r = A$ exactement. De plus, le rang de \widehat{A}_t est t . Dans un sens technique, \widehat{A}_t est la meilleure approximation de A de rang t .

Exemple 12.14. Une expérience donne un résultat en forme matriciel $\begin{bmatrix} 1,02 & 2,03 & 4,20 \\ 0,25 & 0,51 & 1,06 \\ 1,74 & 3,46 & 7,17 \end{bmatrix}$.

On cherche à savoir le rang de cette matrice. On fait une méthode de Gauss, qui donne que le rang est 1, mais en faisant cette réduction on voit que les chiffres deviennent très petits. Est-ce que c'est peut-être de l'erreur expérimentale? Si oui, quelle est la "vraie" matrice?

On calcul les valeurs singulières: $\sigma_1 = 9,5213$, $\sigma_2 = 0,0071$ et $\sigma_3 = 0,0023$. Il semble que peut-être le rang est vraiment un, et que il y a eu de l'erreur expérimentale. On calcul

$$\widehat{A}_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = \begin{bmatrix} 1,0193 & 2,0280 & 4,2011 \\ 0,2567 & 0,5107 & 1,0580 \\ 1,7395 & 3,4610 & 7,1696 \end{bmatrix}$$

C'est la matrice de rang un qui est la meilleur approximation de A . □

Exercice 12.15. Appliquer la méthode de Gauss à la matrice de l'exemple précédent. Montrer qu'on obtient deux rangées "presque" nulles.

En continuant avec la méthode de Gauss, on multiplie ces rangées par des très grands chiffres: si ces deux rangées s'agissait des erreurs expérimentales, alors on multiplierait des erreurs par des très grands chiffres! Pour cette raison, la méthode de Gauss-Jordan n'est pas NUMÉRIQUEMENT STABLE. □

Exemple 12.16. Une satellite prend plusieurs images digitales du même secteur de la Terre, chaque image captant une différente longueur d'onde (i.e., une différente "couleur"). On voudrait savoir quelles couleurs sont les plus importantes dans l'image. Il y a c couleurs au total.

On forme une matrice A de taille $c \times 1000000$, où chaque rangée représente toutes les pixels d'une des images, donc une couleur. On calcul une décomposition en valeurs singulières de A . Si il y a quelques (disons t) valeurs singulières qui dominant, alors on pourrait être confiant que la matrice \hat{A}_t est une bonne approximation de A .

Le satellite pourrait transmettre alors \hat{A}_t au lieu de A . L'avantage, c'est que

$$\hat{A}_t = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_t \mathbf{u}_t \mathbf{v}_t^T$$

Donc on pourrait transmettre t valeurs singulières, t vecteurs de taille c (les \mathbf{u}_j) et t vecteurs de taille 1000000 (les vecteurs \mathbf{v}_j). Il faut donc transmettre $t \times (1000000 + c + 1)$ chiffres au lieu de $c \times 1000000$. \square

Exemple 12.17. On pourrait aussi considérer une image comme une matrice de taille 1000×1000 . Si une décomposition en valeurs singulières montre que les premières t valeurs singulières dominant, alors \hat{A}_t est une bonne approximation à l'image. Il faudrait transmettre t valeurs singulières et $2t$ vecteurs de taille 1000, au lieu de 1000000 valeurs. \square