



10.12 Décomposition en valeurs singulières

On arrive à la fin de notre développement de la théorie pour pouvoir démontrer cette décomposition d'une matrice quelconque en une matrice qui est un produit de trois matrices : une matrice orthogonale, une matrice presque diagonale et puis une matrice orthogonale, donc je répertorie l'ordre : inversible, presque diagonale, inversible.

Notes

Summary



0m 03s

Théorème. Soit $A \in M_{m \times n}(\mathbb{R})$ de rang r . Alors il existe $\Sigma \in M_{m \times n}(\mathbb{R})$ t.q. $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, $D \in M_{r \times r}(\mathbb{R})$ diagonale, $D = \text{diag}(\sigma_1, \dots, \sigma_r)$, $\sigma_1 \geq \dots \geq \sigma_r > 0$ les r valeurs singulières non nulles de A , et $U \in M_{m \times m}(\mathbb{R})$, $V \in M_{n \times n}(\mathbb{R})$ orthogonales t.q. $A = U\Sigma V$.

10.12 Décomposition en valeurs singulières



Donc voilà le théorème de décomposition en valeurs singulières. Donc je me donne une matrice $m \times n$ de rang r . Alors il existe une matrice Σ telle que: Σ est de la forme [voir écran], où D est une matrice diagonale et à part ça, on a des valeurs zéro. D est en fait la matrice diagonale : $\sigma_1, \dots, \sigma_r$, où ce sont les r valeurs singulières non nulles de la matrice A . Et puis il existe une matrice U de taille $m \times m$ et une matrice V de taille $n \times n$, orthogonales telles que A est le produit de $U\Sigma V$. Donc ça, c'est la décomposition en valeurs singulières. En anglais, "singular value decomposition". Donc c'est presque le mieux qu'on puisse espérer. A n'est même pas forcément une matrice carrée et même si elle est carrée, on sait qu'on ne peut pas forcément la diagonaliser mais on peut toujours faire cette décomposition-là. C'est ça qui à beaucoup d'applications. Maintenant, j'aimerais faire la preuve de ce théorème, et je commence ici.

Notes

Summary



0m 29s

Théorème. Soit $A \in M_{m \times n}(\mathbb{R})$ de rang r . Alors il existe $\Sigma \in M_{m \times n}(\mathbb{R})$ t.q. $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, $D \in M_{r \times r}(\mathbb{R})$ diagonale, $D = \text{diag}(\sigma_1, \dots, \sigma_r)$, $\sigma_1 \geq \dots \geq \sigma_r > 0$ les r valeurs singulières non nulles de A , et $U \in M_{m \times m}(\mathbb{R})$, $V \in M_{n \times n}(\mathbb{R})$ orthogonales t.q. $A = U\Sigma V$.

Preuve Soit $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ l'application linéaire t.q. $[\varphi]_{C_2 C_1} = A$ où C_1, C_2 sont les bases canoniques de $\mathbb{R}^n, \mathbb{R}^m$ resp.

Par le théorème de l'existence des bases compatibles, il existe une base orthonormée $B_1 = (v_1, \dots, v_n)$ de $\mathbb{R}^n$ t.q. $(\varphi(v_1), \dots, \varphi(v_r))$ soit une base orthogonale de $\text{im}(\varphi)$.

10.12 Décomposition en valeurs singulières



Donc soit φ , l'application de $\mathbb{R}^n$ dans $\mathbb{R}^m$, l'application linéaire dont la matrice est la matrice A , c'est-à-dire telle que la matrice de φ par rapport aux bases canoniques C_1 et C_2 , est égal à la matrice A où C_1 et C_2 sont les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement. Alors par le théorème de l'existence des bases compatibles, je sais qu'il existe une base orthonormée $B_1 = (v_1, \dots, v_n)$ de $\mathbb{R}^n$ telle que $(\varphi(v_1), \dots, \varphi(v_r))$, soit une base orthogonale de l'image de φ . C'est ce que dit le théorème de l'existence des bases compatibles. Maintenant je reprends ça dans la prochaine slide. Ça, c'est une base orthogonale, ce n'est pas une base orthonormée.

Notes

Summary



On normalise cette base: posons $u_i = \frac{\varphi(v_i)}{\sigma_i}$, un vecteur de norme 1.

Donc $(u_1, \dots, u_r)$ est une base orthonormée de $\text{Im} \varphi$.

On complète cette base, pour former une base orthonormée de $\mathbb{R}^m$:

$$B_2 = (u_1, \dots, u_r, u_{r+1}, \dots, u_m).$$

10.12 Décomposition en valeurs singulières



Donc je vais faire ça. On normalise cette base. Donc posons $u_i = 1/\sigma_i \varphi(v_i)$, où σ_i est en fait la norme de $\varphi(v_i)$. On a déjà utilisé ça plusieurs fois. C'est un vecteur de norme un. Donc maintenant $(u_1, \dots, u_r)$ est une base orthonormée de l'image. Ensuite, je vais étendre ça à une base orthonormée de $\mathbb{R}^m$. On complète cette base pour former une base orthonormée de $\mathbb{R}^m$. J'appelle ça $B_2 = (u_1, \dots, u_r, u_{r+1}, \dots, u_m)$.

Notes

Summary



2m 55s

$\mathcal{B}_1 = (v_1, \dots, v_n)$ base orthonormée de $\mathbb{R}^n$, $u_i = \frac{\phi(v_i)}{\sigma_i}$, $(u_1, \dots, u_r)$ base orthonormée de $\text{im}(\phi)$, $\mathcal{B}_2 = (u_1, \dots, u_r, u_{r+1}, \dots, u_m)$ base orthonormée de $\mathbb{R}^m$.

$$\begin{aligned} \varphi(v_i) &= \sigma_i u_i & i \leq r, & \quad \varphi(v_j) = 0 & j \geq r+1 \\ [\varphi]_{\mathcal{B}_2 \mathcal{B}_1} &= \begin{pmatrix} \sigma_1 & 0 & & 0 \\ 0 & \sigma_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \end{aligned}$$

10.12 Décomposition en valeurs singulières

Maintenant, on a la base $\mathcal{B}_1$, une base orthonormée de $\mathbb{R}^n$. On a $(u_1, \dots, u_r)$, une base orthonormée de l'image et $(u_1, \dots, u_m)$, une base orthonormée de $\mathbb{R}^m$. Comme j'ai cette égalité-là, j'ai $\varphi(v_i) = \sigma_i u_i$. Et ça, c'est pour tout i plus petit ou égal à r . Et les $\varphi(v_j) = 0$ pour tous les j plus grands ou égaux à $r+1$, parce qu'ils corresponderaient à la valeur singulière 0. Donc maintenant, si j'écris φ par rapport à ces deux bases-là, $\mathcal{B}_1$ et $\mathcal{B}_2$, ça a une très jolie forme, donc j'aurais σ_1 parce que le premier vecteur est envoyé à σ_1 fois le premier vecteur à droite, donc zéro. Ensuite le deuxième vecteur est envoyé à σ_2 fois le deuxième vecteur, donc une colonne comme ça et ainsi de suite jusqu'à σ_r , toujours le long de la diagonale et après le reste de la matrice est 0. Voilà, c'est une matrice qui a justement ici, dans ce coin-là, une matrice diagonale avec des valeurs singulières le long de la diagonale.

Notes

Summary



$B_1 = (v_1, \dots, v_n)$ base orthonormée de $\mathbb{R}^n$, $u_i = \frac{\phi(v_i)}{\sigma_i}$, $(u_1, \dots, u_r)$ base orthonormée de $\text{im}(\phi)$, $B_2 = (u_1, \dots, u_r, u_{r+1}, \dots, u_m)$ base orthonormée de $\mathbb{R}^m$.

$$\varphi(v_i) = \sigma_i u_i, \quad i \leq r, \quad \varphi(v_j) = 0 \quad j \geq r+1$$

$$[\varphi]_{B_2 B_1} = \begin{pmatrix} \sigma_1 & 0 & & 0 \\ 0 & \sigma_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

La matrice $A = [\varphi]_{C_2 C_1}$.

Donc $A = [\varphi]_{C_2 C_1} = [id]_{C_2 B_2} [\varphi]_{B_2 B_1} [id]_{B_1 C_1}$

Posons $U = [id]_{C_2 B_2}$, $V = [id]_{B_1 C_1}$. U et V sont orthogonales car leurs colonnes forment des bases orthonormées.

10.12 Décomposition en valeurs singulières

La matrice A originale était la matrice de φ par rapport aux bases canoniques. Pour faire le lien entre cette matrice-là et celle-là, je dois juste faire des changements de base. Donc on a que A , qui est cette matrice, est égale à : Ici, je mets au milieu la matrice φ par rapport aux bases B_1 et B_2 . Comme ici, je veux commencer avec quelque chose qui prend la base C_1 , je prends la base C_1 écrite en termes de la base de B_1 , ensuite cette application me redonne un vecteur en termes de B_2 , donc je prends la base B_2 et je l'exprime en termes de C_2 . Maintenant posons U égal à la matrice ici à gauche et V égal à la matrice ici. Ces deux matrices sont des matrices orthogonales parce que ce sont des matrices de passage entre des bases orthonormées. Je répète : U et V sont orthogonales car leurs colonnes forment des bases orthonormées. Ici, c'était dans $\mathbb{R}^m$ et ici dans $\mathbb{R}^n$, respectivement.

Notes

Summary



5m 20s

$\mathcal{B}_1 = (v_1, \dots, v_n)$ base orthonormée de $\mathbb{R}^n$, $u_i = \frac{\phi(v_i)}{\sigma_i}$, $(u_1, \dots, u_r)$ base orthonormée de $\text{im}(\phi)$, $\mathcal{B}_2 = (u_1, \dots, u_r, u_{r+1}, \dots, u_m)$ base orthonormée de $\mathbb{R}^m$.

$$\varphi(v_i) = \sigma_i u_i, \quad i \leq r, \quad \varphi(v_j) = 0 \quad j \geq r+1$$

$$[\varphi]_{\mathcal{B}_2 \mathcal{B}_1} = \begin{pmatrix} \sigma_1 & 0 & & 0 \\ 0 & \sigma_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{La matrice } A = [\varphi]_{\mathcal{C}_2 \mathcal{C}_1}.$$

Donc $A = [\varphi]_{\mathcal{C}_2 \mathcal{C}_1} = [id]_{\mathcal{C}_2 \mathcal{B}_2} [\varphi]_{\mathcal{B}_2 \mathcal{B}_1} [id]_{\mathcal{B}_1 \mathcal{C}_1}$

Posons $U = [id]_{\mathcal{C}_2 \mathcal{B}_2}$, $V = [id]_{\mathcal{B}_1 \mathcal{C}_1}$. U et V sont orthogonales car leurs colonnes forment des bases orthonormées de $\mathbb{R}^m$, $\mathbb{R}^n$ resp.

Posons $\Sigma = \begin{pmatrix} \sigma_1 & & 0 & \\ & \sigma_r & & \\ 0 & & \ddots & \\ & & & \sigma_r & \\ 0 & & & & 0 \end{pmatrix}$. On a $A = U \Sigma V$.

10.12 Décomposition en valeurs singulières



Et posons, Σ , la matrice dont la diagonale est $\sigma_1, \sigma_2, \dots, \sigma_r$, et puis on a que notre matrice A est égale à $U \Sigma V$. Voilà la décomposition de cette matrice en un produit de matrices orthogonale, donc inversible, une matrice presque diagonale et orthogonale donc inversible. C'est la preuve de cette décomposition. Maintenant dans le dernier paragraphe, nous allons appliquer ça à un exemple.

Notes

Summary

