Algèbre linéaire

Dans cette partie, on présente quelques résultats d'algèbre linéaire utiles dans le cadre de ce cours. Pour plus d'information, vous pouvez vous référer au cours MAT-1200, à Deisenroth, Faisal, et Ong (2020) (en anglais) et à Grifone (2024) (en français).

1 Quelques propriétés matricielles

Notons $M_{n,m}(\mathbb{R})$, l'ensemble des matrices à n lignes et m colonnes dont les entrées appartiennent à $\mathbb{R}$. Notons $M_n(\mathbb{R})$, l'ensemble des matrices carrées de taille n, i.e. à n lignes et n colonnes dont les entrées appartiennent à $\mathbb{R}$. Soient M, N et P des matrices appartenant à $M_{n,m}(\mathbb{R})$. Soient A et B des matrices appartenant à $M_n(\mathbb{R})$. Notons I_n la matrice identité de taille n, i.e. qui contient des 1 sur le diagonale et des 0 sur les éléments hors de la diagonale. Soient u et v appartenant à $\mathbb{R}^n$, i.e. des vecteurs colonnes de taille n.

Propriétés de l'inverse de matrices

Supposons que les matrices A et B soient inversibles. Alors le produit matriciel AB est inversible et est donné par :

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Preuve

Posons $C = AB$ et $D = B^{-1}A^{-1}$. Alors

$$\begin{aligned} CD &= ABB^{-1}A^{-1} \\ &= AA^{-1} \\ &= I_n \end{aligned}$$

De la même façon, on trouve que $DC = I_n$. Ainsi, AB est inversible et son inverse est donné par $B^{-1}A^{-1}$.

Propriétés du déterminant de matrices

Considérant les matrices définies en début de section, on a :

1. $\det(A^\top) = \det(A)$,
2. $\det(AB) = \det(A)\det(B)$,
3. $\det(A^{-1}) = 1/\det(A)$.

Preuve

Les preuves des propriétés 1 et 2 sont techniques et sont omises, mais peuvent être trouvées, par exemple, ici. Pour ce qui est de la troisième propriété, par définition, on a $AA^{-1} = I_n$. Le déterminant de I_n est égale à 1 (produit des éléments sur la diagonale). Donc $\det(AA^{-1}) = 1$. Or, d'après la deuxième propriété, $\det(AA^{-1}) = \det(A)\det(A^{-1})$. On a donc bien $\det(A^{-1}) = 1/\det(A)$.

Propriétés de la trace de matrices

Considérant les matrices définies en début de section, on a :

1. $\operatorname{tr}(A) = \operatorname{tr}(A^\top)$,
2. $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$,
3. $\operatorname{tr}(MN^\top) = \operatorname{tr}(N^\top M)$.

Preuve

Pour une matrice carré A, notons a_{ij}, l'élément de la matrice A à la ligne i et à la colonne j. La trace de A est donnée par la somme des éléments diagonaux, i.e. $\operatorname{tr}(A) = \sum_{i=1}^n a_{ii}$.

1. La transposition ne changeant pas les éléments diagonaux, le résultat est direct.
2. Notons $C = A + B$. Comme A et B sont des matrices carrées, C est une matrice carrée. On a $c_{ij} = a_{ij} + b_{ij}$ pour tout $i, j = 1, \ldots, n$. Donc

$$\operatorname{tr}(A + B) = \operatorname{tr}(C) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n a_{ii} + b_{ii} = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = \operatorname{tr}(A) + \operatorname{tr}(B).$$

3. Les matrices $MN^\top$ et $N^\top M$ sont carrées, de dimension respectives $n \times n$ et $m \times m$, on peut donc bien calculer leur trace. Notons $C = MN^\top$ et $D = N^\top M$.

$$\operatorname{tr}(MN^\top) = \operatorname{tr}(C) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{j=1}^m m_{ij} n_{ji} = \sum_{j=1}^m \sum_{i=1}^n n_{ji} m_{ij} = \sum_{j=1}^m d_{jj} = \operatorname{tr}(D) = \operatorname{tr}(N^\top M).$$

Définition

1. Soit A une matrice symétrique appartenant à $M_n(\mathbb{R})$. A est **définie positive** si $u^\top Au > 0$ pour tout $u \in \mathbb{R}^n$ tel que $u \neq 0$.
2. Soit A appartenant à $M_n(\mathbb{R})$. A est **orthogonal** si $A^\top A = AA^\top = I_n$.

2 Valeurs et vecteurs propres

Définition

Soit A appartenant à $M_n(\mathbb{R})$. On dit que $\lambda \in \mathbb{R}$ est une **valeur propre** de A s'il existe un vecteur $u \in \mathbb{R}^n$ non nul tel que

$$Au = \lambda u. \tag{1}$$

Le vecteur u est appelé **vecteur propre** de A correspondant à la valeur propre λ. L'ensemble des nombres réels λ satisfaisant Équation 1 est appelé **spectre** de la matrice A et noté $\mathrm{sp}(A)$.

Propriété des vecteurs propres

1. Si u est un vecteur propre de A correspondant à une valeur propre λ, alors le vecteur cu, $c \in \mathbb{R}^\star$ est également un vecteur propre de A correspondant à λ.
2. Si A est symétrique et u_1 et u_2 sont des vecteurs propres correspondant à des valeurs propres différentes de A, alors u_1 et u_2 sont orthogonaux, i.e. $u_1^\top u_2 = 0$.

Preuve

1. Soit $c \in \mathbb{R}^\star$ et u un vecteur propre de A associé à la valeur propre λ. On a :

   $$A(cu) = cAu = c\lambda u = \lambda(cu).$$

   Donc, le vecteur cu est aussi vecteur propre de A associé à la valeur propre λ.
2. Soient λ_1 et λ_2, les valeurs propres associées à u_1 et u_2, tel que $\lambda_1 \neq \lambda_2$. On a $Au_1 = \lambda_1 u_1$ et $Au_2 = \lambda_2 u_2$. Ensuite

   $$\lambda_1 u_1^\top u_2 = u_1^\top Au_2 = \lambda_2 u_1^\top u_2.$$

   Cela implique que $(\lambda_1 - \lambda_2)u_1^\top u_2 = 0$. Or, $\lambda_1 \neq \lambda_2$. Donc, nécessairement, $u_1^\top u_2 = 0$.

Cette deuxième propriété nous sera utile lorque l'on s'intéressera à la réduction de dimension et, en particulier, à l'analyse en composantes principales.

Caractérisation de matrices avec ses éléments propres

1. Si A est symétrique, alors **toutes** ses valeurs propres sont réelles.
2. Si A est définie positive, alors **toutes** ses valeurs propres sont strictement positives.

Preuve

1. Considérons le cas plus général où A est une matrice hermitenne. La matrice A est égale la transposé de son conjugué, noté A^*. Notons λ une valeur propre associée à un vecteur propre u, éventuellement complexe. On a :

$$\overline{u}^\top A u = \overline{u}^\top \lambda u = \lambda \overline{u}^\top u, \tag{2}$$

$$\overline{u}^\top A u = \overline{u}^\top A^* u = \overline{Au}^\top u = \overline{\lambda}\overline{u}^\top u. \tag{3}$$

   Cela implique que $(\lambda - \overline{\lambda})\overline{u}^\top u = 0$. Comme $u \neq 0$, on a $\lambda = \overline{\lambda}$. Donc $\lambda \in \mathbb{R}$.
2. Considérons u, vecteur propre de A associé à la valeur propre λ. On a que $u^\top A u = \lambda u^\top u$. Or, comme $u \neq 0$, $u^\top u \neq 0$. Donc

$$\lambda = \frac{u^\top A u}{u^\top u}.$$

   Comme A est définie postive, $u^\top A u > 0$ pour tout vecteur u non nul. On en déduit que $\lambda > 0$.

3 Diagonalisation de matrices

Définition

Soit A appartenant à $M_n(\mathbb{R})$. On dit que A est **diagonalisable** s'il existe une matrice P appartenant à $M_n(\mathbb{R})$ non-singulière et une matrice diagonale D appartenant à $M_n(\mathbb{R})$ telles que

$$P^{-1}AP = D \iff A = PDP^{-1}.$$

Théorème de décomposition spectrale

Soit A une matrice symétrique appartenant à $M_n(\mathbb{R})$ et $\lambda_1, \dots, \lambda_n$, ses n valeurs propres. Alors, il existe une matrice orthogonal P appartenant à $M_n(\mathbb{R})$ telle que

$$A = P\Lambda P^\top, \quad \text{où} \quad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Si A admet n valeurs propres positives distinctes, alors on peut prendre P comme étant la matrice dont la k-ième colonne est le vecteur propre normé correspondant à la k-ième valeur propre λ_k de A.

Soit deux matrices symétriques, A et B, comment déterminer le vecteur u tel que $u^\top Au$ soit maximal, sachant que $u^\top Bu = 1$? Il suffit de prendre u comme le vecteur propre de $B^{-1}A$ associé à λ la valeur propre maximale de $B^{-1}A$. On obtient ainsi

$$u^\top Au = u^\top \lambda Mu = \lambda U^\top Mu = \lambda.$$

Caractérisation du déterminant et de la trace de matrices avec ses éléments propres

Si A a comme valeurs propres (réelles, mais pas forcément distinctes) $\lambda_1, \dots, \lambda_n$, alors

1. $\det(A) = \prod_{i=1}^n \lambda_i$
2. $\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i$.

Preuve

En utilisant le théorème de décomposition spectrale, il existe une matrice P inversible tel que $A = P\Lambda P^{-1}$, où Λ est une matrice diagonale contenant les valeurs propres. On a donc, pour le déterminant,

$$\det(A) = \det(P\Lambda P^{-1}) = \det(P)\det(\Lambda)\det(P^{-1}) = \det(P)\det(\Lambda)\det(P)^{-1} = \det(\lambda) = \prod_{i=1}^n \lambda_i,$$

et, pour la trace,

$$\operatorname{tr}(A) = \operatorname{tr}(P\Lambda P^{-1}) = \operatorname{tr}(P^{-1}P\Lambda) = \operatorname{tr}(\Lambda) = \sum_{i=1}^n \lambda_i.$$

Deisenroth, Marc Peter, A. Aldo Faisal, et Cheng Soon Ong. 2020. *Mathematics for Machine Learning*. 1 éd. Cambridge University Press. https://doi.org/10.1017/9781108679930.

Grifone, Joseph. 2024. *Algèbre Linéaire*. 7e edition. Toulouse : CEPADUES.