

ML 11 Projecteurs orthogonaux et SVD

Nous introduisons la notion de projecteur et nous montrons son lien avec la décomposition en valeurs singulières (SVD). Ainsi, nous comprendrons mieux la géométrie du SVD et nous serons plus tard en mesure de construire un algorithme qui est une extension de la méthode de la puissance itérée pour le calcul des k premières valeurs singulières et vecteurs singuliers.

1 Projecteur

Un projecteur est une matrice carrée idempotente : $\mathbf{P}^2 = \mathbf{P}$. Rappelons que l'espace des colonnes de \mathbf{P} formé de l'ensemble des vecteurs qui s'expriment comme une combinaison linéaire des colonnes de \mathbf{P} , c'est-à-dire tous les vecteurs \mathbf{v} qui s'écrivent sous la forme $\mathbf{P}\mathbf{x}$. L'espace des colonnes de \mathbf{P} est également appelé l'image de \mathbf{P} et se note $Im(\mathbf{P})$. Un vecteur \mathbf{v} qui appartient à l'image d'un projecteur \mathbf{P} est tel que $\mathbf{P}\mathbf{v} = \mathbf{v}$:

$$\begin{aligned} & \mathbf{v} \text{ appartient à l'image de } \mathbf{P} \\ = & \{ \text{Par définition de l'image d'une transformation linéaire.} \} \\ & \mathbf{v} = \mathbf{P}\mathbf{x} \\ = & \{ \text{Multiplication des deux membres de l'égalité par } \mathbf{P} \} \\ & \mathbf{P}\mathbf{v} = \mathbf{P}^2\mathbf{x} \\ = & \{ \text{Idempotence d'un projecteur.} \} \\ & \mathbf{P}\mathbf{v} = \mathbf{P}\mathbf{x} \\ = & \{ \mathbf{v} = \mathbf{P}\mathbf{x} \} \\ & \mathbf{P}\mathbf{v} = \mathbf{v} \end{aligned}$$

On appelle noyau d'une transformation linéaire, l'ensemble des vecteurs que cette dernière transforme en $\mathbf{0}$, c'est-à-dire les vecteurs qui montrent une dépendance linéaire entre des colonnes de la représentation matricielle de la transformation. Dans le cas d'un projecteur \mathbf{P} , la projection sur l'image de \mathbf{P} d'un vecteur \mathbf{v} se fait selon la direction $\mathbf{P}\mathbf{v} - \mathbf{v}$ qui appartient au noyau de \mathbf{P} noté $Ker(\mathbf{P})$ (voir la figure 1) :

$$\mathbf{P}(\mathbf{P}\mathbf{v} - \mathbf{v}) = \mathbf{P}^2\mathbf{v} - \mathbf{P}\mathbf{v} = \mathbf{P}\mathbf{v} - \mathbf{P}\mathbf{v} = \mathbf{0}$$

Le projecteur complémentaire à \mathbf{P} est $\mathbf{I} - \mathbf{P}$, il projecte sur le noyau de \mathbf{P} . C'est-à-dire que $Im(\mathbf{I} - \mathbf{P}) = Ker(\mathbf{P})$. En effet :

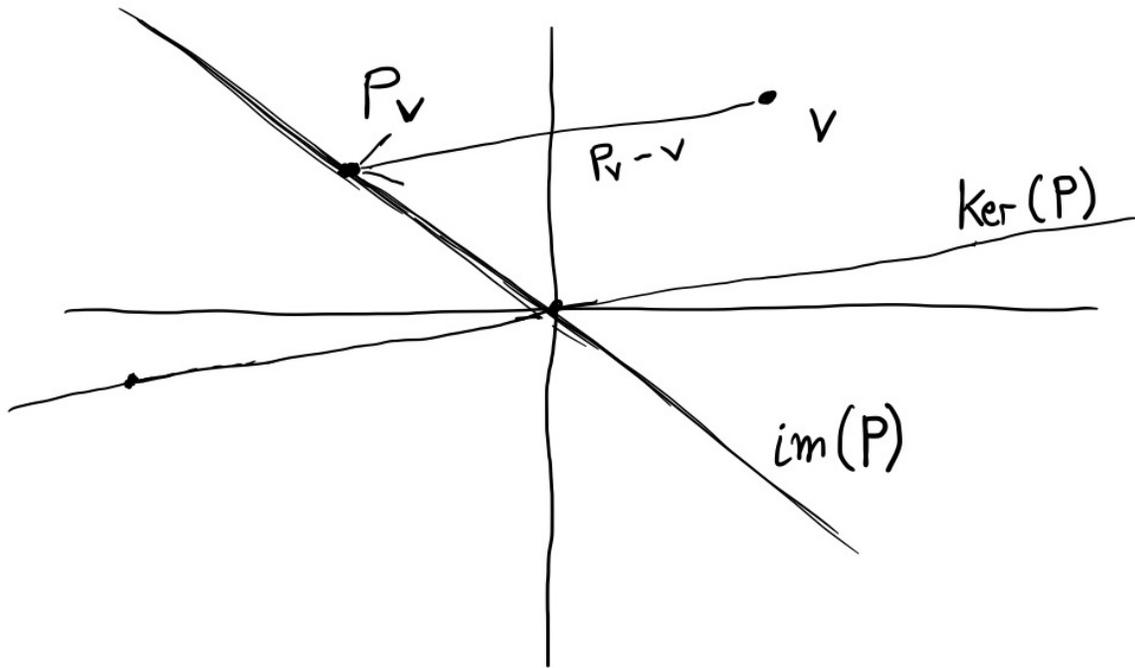


Figure 1: Schéma d'un projecteur

$$\begin{aligned}
 & \mathbf{v} \in \text{Ker}(\mathbf{P}) \\
 & = \{\text{Par définition du noyau}\} \\
 & \quad \mathbf{P}\mathbf{v} = \mathbf{0} \\
 & \Rightarrow \\
 & \quad (\mathbf{I} - \mathbf{P})\mathbf{v} = \mathbf{v} \\
 & = \{\text{Par définition de l'image}\} \\
 & \quad \mathbf{v} \in \text{Im}(\mathbf{I} - \mathbf{P})
 \end{aligned}$$

Donc, $\text{Ker}(\mathbf{P}) \subseteq \text{Im}(\mathbf{I} - \mathbf{P})$. Par ailleurs, un élément de l'image de $\mathbf{I} - \mathbf{P}$ peut se noter $(\mathbf{I} - \mathbf{P})\mathbf{v}$, ce qui est égal à $\mathbf{v} - \mathbf{P}\mathbf{v}$ et appartient donc au noyau de \mathbf{P} . Donc, nous avons aussi $\text{Im}(\mathbf{I} - \mathbf{P}) \subseteq \text{Ker}(\mathbf{P})$. Ainsi : $\text{Ker}(\mathbf{P}) = \text{Im}(\mathbf{I} - \mathbf{P})$ (de même : $\text{Ker}(\mathbf{I} - \mathbf{P}) = \text{Im}(\mathbf{P})$).

Donc, un projecteur \mathbf{P} sépare l'espace initial en deux sous-espaces $S_1 = \text{Im}(\mathbf{P})$ et $S_2 = \text{Ker}(\mathbf{P})$. \mathbf{P} projette sur S_1 **parallèlement** à (ou le long de) S_2 .

2 Projecteur orthogonal

Un projecteur est dit orthogonal quand S_1 et S_2 sont orthogonaux. Montrons qu'une projection \mathbf{P} est orthogonale si et seulement si $\mathbf{P} = \mathbf{P}^T$. Montrons d'abord que $\mathbf{P} = \mathbf{P}^T \Rightarrow$ orthogonalité.

$$\begin{aligned}
& (\mathbf{P}\mathbf{x})^T(\mathbf{I} - \mathbf{P})\mathbf{y} \quad \text{pour tout } \mathbf{x} \text{ et } \mathbf{y} \\
& = \{\text{Propriété de la transposition.}\} \\
& \quad \mathbf{x}^T \mathbf{P}^T (\mathbf{I} - \mathbf{P})\mathbf{y} \\
& = \{\mathbf{P} = \mathbf{P}^T\} \\
& \quad \mathbf{x}(\mathbf{P} - \mathbf{P}^2)\mathbf{y} \\
& = \{\mathbf{P} \text{ est un projecteur donc } \mathbf{P} = \mathbf{P}^2\} \\
& 0
\end{aligned}$$

Montrons ensuite que orthogonalité $\Rightarrow \mathbf{P} = \mathbf{P}^T$. \mathbf{P} projette sur $S1$ le long de $S2$ avec $S1 \perp S2$. Soit $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ une base orthonormale de $S1$ et $\{\mathbf{q}_{n+1}, \mathbf{q}_{n+2}, \dots, \mathbf{q}_m\}$ une base de $S2$. Soit Q la matrice orthogonale dont la jème colonne est \mathbf{q}_j . Nous avons : $\mathbf{P}Q = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n, 0, \dots, 0)$. Donc, $Q^T \mathbf{P} Q = \text{diag}(1, 1, \dots, 1, 0, \dots, 0) = \Sigma$. Ainsi, $\mathbf{P} = Q \Sigma Q^T = \mathbf{P}^T$.

Etant donné un vecteur unitaire \mathbf{q} . Le projecteur orthogonal sur l'espace à une dimension défini par \mathbf{q} est la matrice de rang 1 $\mathbf{P}_q = \mathbf{q}\mathbf{q}^T$. Le projecteur complémentaire est $\mathbf{P}_{\perp q} = \mathbf{I} - \mathbf{P}_q$ de rang $m - 1$ (si \mathbf{q} est de dimension m).

Pour toute matrice $Q \in \mathbb{R}^{M \times N}$ dont les colonnes \mathbf{q}_j sont orthonormales, $\mathbf{P} = Q Q^T = \sum \mathbf{q}_j \mathbf{q}_j^T$ est un projecteur orthogonal sur $Im(Q)$.

Rappelons la décomposition en valeurs singulière d'une matrice $\mathbf{A} \in \mathbb{R}^{M \times N}$ de rang r : $\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T$ avec $\mathbf{U} \in \mathbb{R}^{M \times M}$, $\mathbf{V} \in \mathbb{R}^{N \times N}$ des matrices aux colonnes orthonormales et $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{M \times N}$. Nous pouvons partitionner le SVD en blocs pour obtenir la forme réduite présentée sur la figure 2.

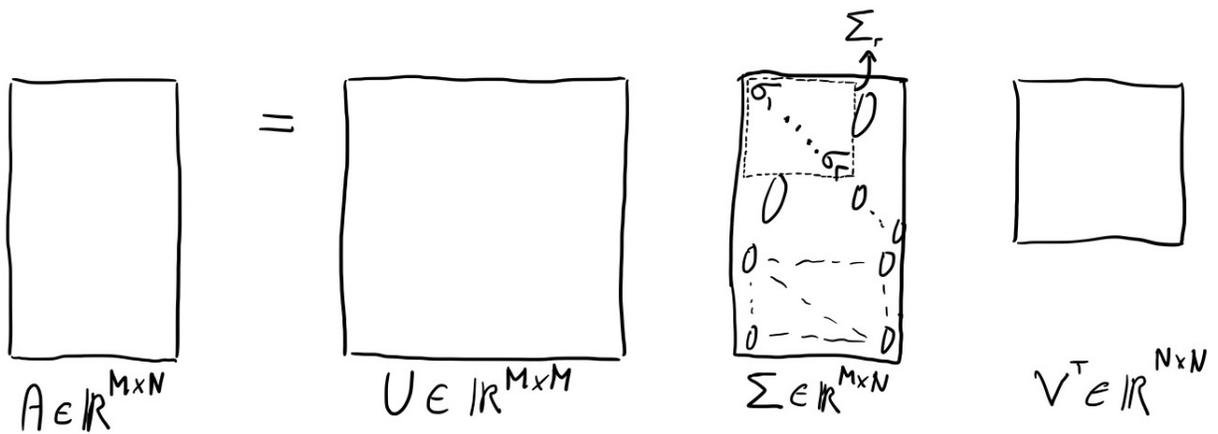
Nous avons les propriétés fondamentales suivantes :

$$\begin{array}{ll}
Im(\mathbf{A}) = Im(\mathbf{U}_r) & \mathbf{U}_r \mathbf{U}_r^T : \text{est un projecteur orthogonal sur } Im(\mathbf{A}) \\
Ker(\mathbf{A}) = Im(\tilde{\mathbf{V}}_r) & \tilde{\mathbf{V}}_r \tilde{\mathbf{V}}_r^T : \text{est un projecteur orthogonal sur } Ker(\mathbf{A}) \\
Im(\mathbf{A}^T) = Im(\mathbf{V}_r) & \mathbf{V}_r \mathbf{V}_r^T : \text{est un projecteur orthogonal sur } Im(\mathbf{A}^T) \\
Ker(\mathbf{A}^T) = Im(\tilde{\mathbf{U}}_r) & \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T : \text{est un projecteur orthogonal sur } Ker(\mathbf{A}^T)
\end{array}$$

Montrons par exemple que $Ker(\mathbf{A}) = Im(\tilde{\mathbf{V}}_r)$, c'est-à-dire que les $N - r$ dernières colonnes de $\tilde{\mathbf{V}}_r$ forment une base orthonormale pour le noyau de \mathbf{A} , c'est-à-dire pour l'ensemble des vecteurs \mathbf{x} tels que $\mathbf{A}\mathbf{x} = 0$.

Considérons $\mathbf{x} \in Im(\tilde{\mathbf{V}}_r)$, ce vecteur s'écrit sous la forme $\mathbf{x} = \sum_{i=r+1}^N w_i \mathbf{v}_i$ avec $w_i \in \mathbb{R}$ et \mathbf{v}_i la i -ème colonne de $\tilde{\mathbf{V}}_r$. Nous devons montrer que $\mathbf{A}\mathbf{x} = 0$.

$$\begin{aligned}
& \mathbf{A}\mathbf{x} \\
& = \{\text{Par définition de } \mathbf{x}.\} \\
& \quad \mathbf{A} \sum_{i=r+1}^N w_i \mathbf{v}_i \\
& = \{\text{Par linéarité.}\} \\
& \quad \sum_{i=r+1}^N w_i (\mathbf{A}\mathbf{v}_i)
\end{aligned}$$



$$A = \begin{bmatrix} U_r & \tilde{U}_r \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ \tilde{V}_r^T \end{bmatrix}$$

$U_r \in \mathbb{R}^{M \times r} \quad \tilde{U}_r \in \mathbb{R}^{M \times M-r} \quad V_r \in \mathbb{R}^{N \times r} \quad \tilde{V}_r \in \mathbb{R}^{N \times N-r}$

Figure 2: SVD réduit

Il suffit donc de montrer que chaque $\mathbf{A}\mathbf{v}_i$ est nul.

$$\begin{aligned}
& \mathbf{A}\mathbf{v}_i \\
= & \{\text{SVD de } \mathbf{A}\} \\
& (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)\mathbf{v}_i \\
= & \{\text{Les colonnes de } \mathbf{V} \text{ sont orthonormales et } \mathbf{v}_i \text{ est une colonne de } \mathbf{V}. \\
& \text{Donc } \mathbf{V}^T\mathbf{v}_i = \mathbf{e}_i \text{ le vecteur avec des 0 partout sauf le 1 en position } i.\} \\
& \mathbf{U}\mathbf{\Sigma}\mathbf{e}_i \\
= & \{\mathbf{\Sigma}\mathbf{e}_i = 0 \text{ car à partir de la } (r+1)\text{ème ligne, toutes les lignes de } \mathbf{\Sigma} \text{ sont nulles et } i \geq r + 1.\} \\
& 0
\end{aligned}$$

3 Projection sur une base quelconque

Nous cherchons à projeter un vecteur \mathbf{v} sur l'espace des colonnes d'une matrice \mathbf{A} . Notons le résultat de cette projection $\mathbf{y} \in \text{Im}(\mathbf{A})$. $\mathbf{y} - \mathbf{v}$ est orthogonal à $\text{Im}(\mathbf{A})$ (c'est-à-dire qu'il appartient à $\text{Ker}(\mathbf{A})$). Autrement dit, en posant $\mathbf{y} = \mathbf{A}\mathbf{x}$, nous avons :

$$\begin{aligned}
& \mathbf{a}_j^T(\mathbf{y} - \mathbf{v}) = 0, \forall j \\
= & \{\text{Posons } \mathbf{y} = \mathbf{A}\mathbf{x}.\} \\
& \mathbf{a}_j^T(\mathbf{A}\mathbf{x} - \mathbf{v}) = 0, \forall j \\
= & \{\text{Linéarité de la multiplication matricielle.}\} \\
& \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{v}) = \mathbf{0} \\
= & \\
& \mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{v} \\
= & \{\text{Si } \mathbf{A}^T\mathbf{A} \text{ est inversible.}\} \\
& \mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{v}
\end{aligned}$$

Ainsi, la projection $\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{v}$. Nous trouvons donc que $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ est un projecteur orthogonal sur l'espace des colonnes de \mathbf{A} . En fait, il est simple de vérifier que ce projecteur est $\mathbf{U}_r\mathbf{U}_r^T$:

$$\begin{aligned}
\mathbf{A} &= \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T \\
\mathbf{A}^T &= \mathbf{V}_r\mathbf{\Sigma}_r\mathbf{U}_r^T \\
\mathbf{A}^T\mathbf{A} &= \mathbf{V}_r\mathbf{\Sigma}_r^2\mathbf{V}_r^T \\
(\mathbf{A}^T\mathbf{A})^{-1} &= \mathbf{V}_r\mathbf{\Sigma}_r^{-2}\mathbf{V}_r^T \\
(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T &= \mathbf{V}_r\mathbf{\Sigma}_r^{-1}\mathbf{U}_r^T \\
\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T &= \mathbf{U}_r\mathbf{U}_r^T
\end{aligned}$$