Décomposition en valeurs singulières

SVD

AUTEUR·RICE

B. Charlier, N. Verzelen, J. Salmon

1 La décomposition spectrale

Théorème spectral

Une matrice symétrique $S \in \mathbb{R}^{n\times n}$ est diagonalisable en base orthonormée : il existe $\lambda_1 \geq \cdots \geq \lambda_n$ et une matrice orthogonale $U \in \mathbb{R}^{n\times n}$ telles que

$$S = U\ \mathrm{diag}(\lambda_1, \ldots, \lambda_n)U^\top \qquad \text{ou} \qquad SU = U\ \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Si $U = [u_1, \ldots, u_n]$, alors

$$S = \sum_{i=1}^{n} \lambda_i u_i u_i^\top, \qquad Su_i = \lambda_i u_i.$$

Une matrice U est orthogonale si $U^\top U = UU^\top = I_n$, ou encore $\langle u_i, u_j \rangle = \delta_{ij}$.

Les λ_i sont les **valeurs propres** et les u_i les **vecteurs propres** associés.

2 La décomposition en valeurs singulières

Théorème

Pour toute matrice $X \in \mathbb{R}^{n\times p}$, il existe des matrices orthogonales $U \in \mathbb{R}^{n\times n}$ et $V \in \mathbb{R}^{p\times p}$ telles que

$$U^\top X V = \Sigma = \mathrm{diag}(s_1, \ldots, s_{\min(n,p)}),$$

avec $s_1 \geq \cdots \geq s_{\min(n,p)} \geq 0$. Ainsi,

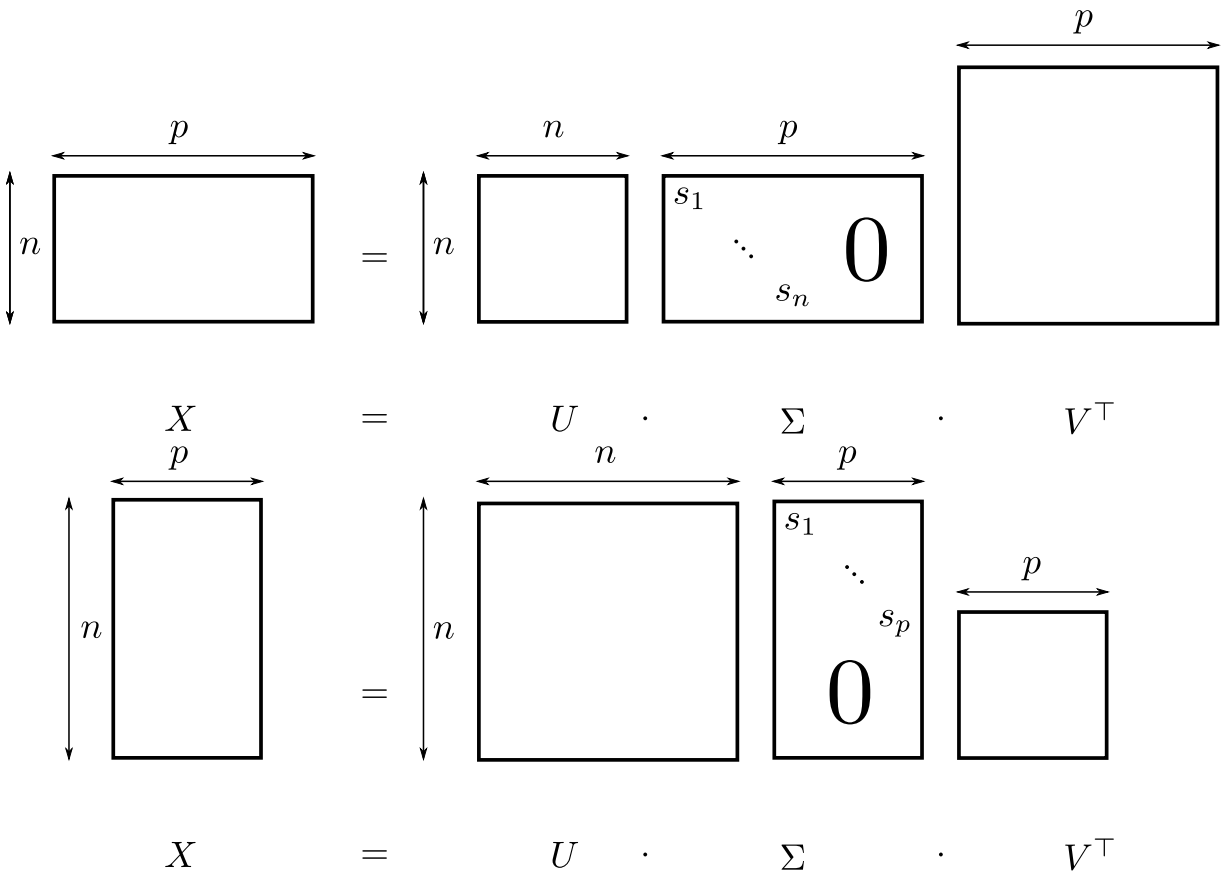
$$\boxed{X = U\Sigma V^\top}.$$

On écrit $U = [u_1, \dots, u_n]$ et $V = [v_1, \dots, v_p]$, avec $\langle u_i, u_j \rangle = \delta_{ij}$ et $\langle v_i, v_j \rangle = \delta_{ij}$.

Démonstration : diagonaliser $X^\top X$ (Golub et van Loan 1996).

Golub, G. H., et C. F. van Loan. 1996. *Matrix computations*. Third. Johns Hopkins University Press.

3 SVD : visualisation



4 SVD : propriété variationnelle

Les s_j sont les **valeurs singulières** de X ; les u_j (resp. v_j) sont les **vecteurs singuliers à gauche** (resp. **à droite**).

Propriété variationnelle de la plus grande valeur singulière

$$s_1 = \max_{u \in \mathbb{R}^n,\ v \in \mathbb{R}^p,\ \|u\| = \|v\| = 1} u^\top X v.$$

Le lagrangien est

$$\mathcal{L}(u, v) = u^\top X v - \lambda_1(\|u\|^2 - 1) - \lambda_2(\|v\|^2 - 1).$$

Les conditions du premier ordre donnent

$$Xv = 2\lambda_1 u, \qquad X^\top u = 2\lambda_2 v,$$

puis

$$X^\top X v = \alpha v, \qquad X X^\top u = \alpha u,$$

avec $\alpha = 4\lambda_1\lambda_2$.

5 SVD réduite

Partons de $X = U\Sigma V^\top$. En ne gardant que les $r = \mathbf{rang}(X)$ valeurs singulières non nulles,

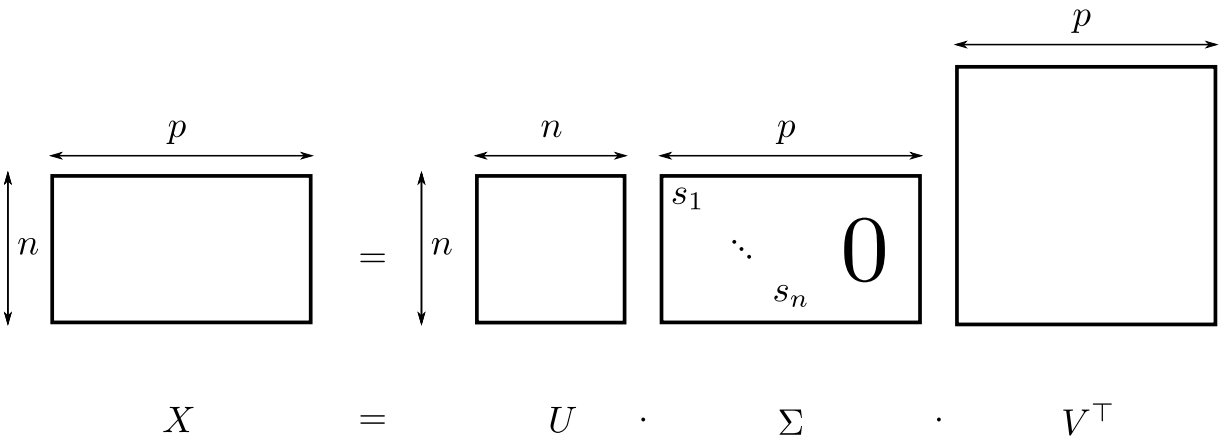
$$\boxed{X = \sum_{i=1}^{r} s_i u_i v_i^\top = U_r \,\mathrm{diag}(s_1, \ldots, s_r) V_r^\top}.$$

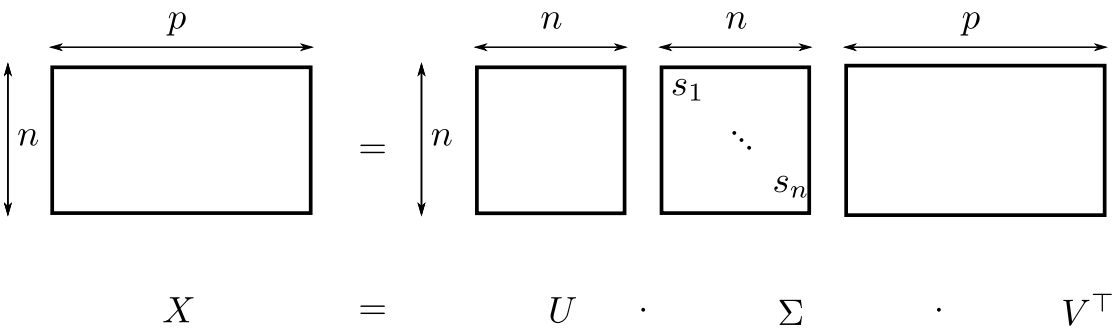
Les matrices $u_i v_i^\top$ sont de rang 1. Les u_i et les v_i sont orthonormés et engendrent respectivement les espaces des colonnes et des lignes de X :

$$\mathrm{Vect}(x_1, \ldots, x_p) = \mathrm{Vect}(u_1, \ldots, u_r).$$

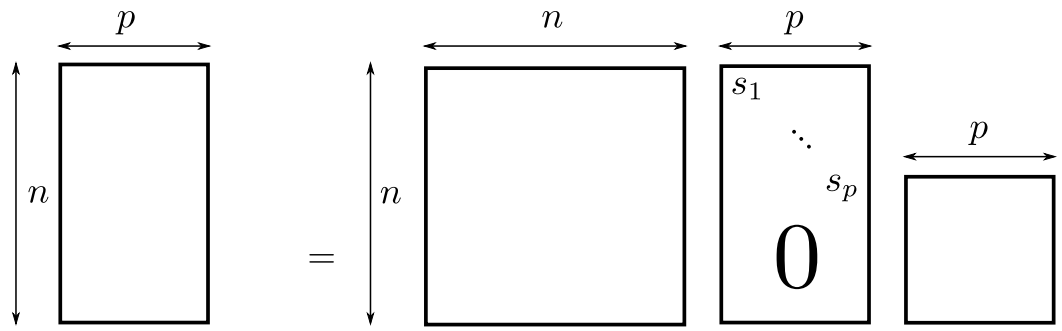
6 SVD réduite : visualisation

Les deux premières vues correspondent à une matrice large ;

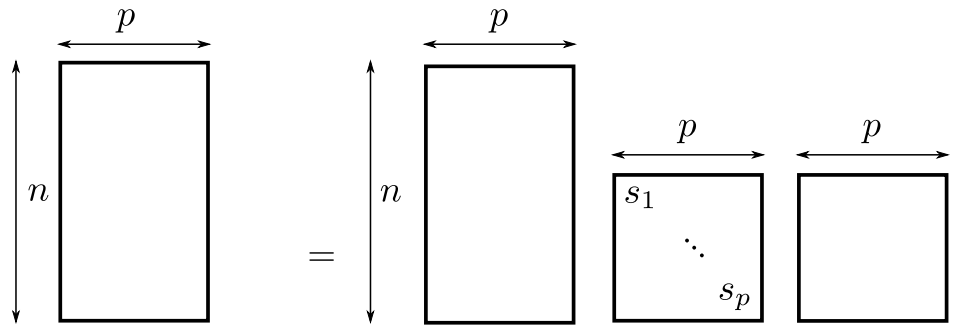




Les deux suivantes à une matrice haute.



$$X = U \cdot \Sigma \cdot V^\top$$



$$X = U \cdot \Sigma \cdot V^\top$$

7 SVD et meilleure approximation

ⓘ Théorème (meilleure approximation de rang k)

Soit $X = \sum_{i=1}^{r} s_i u_i v_i^\top$ la SVD compacte de $X \in \mathbb{R}^{n \times p}$ et $X_k = \sum_{i=1}^{k} s_i u_i v_i^\top$. Alors

$$\min_{\operatorname{rank}(Z)=k} \|X - Z\|_2 = \|X - X_k\|_2 = s_{k+1}.$$

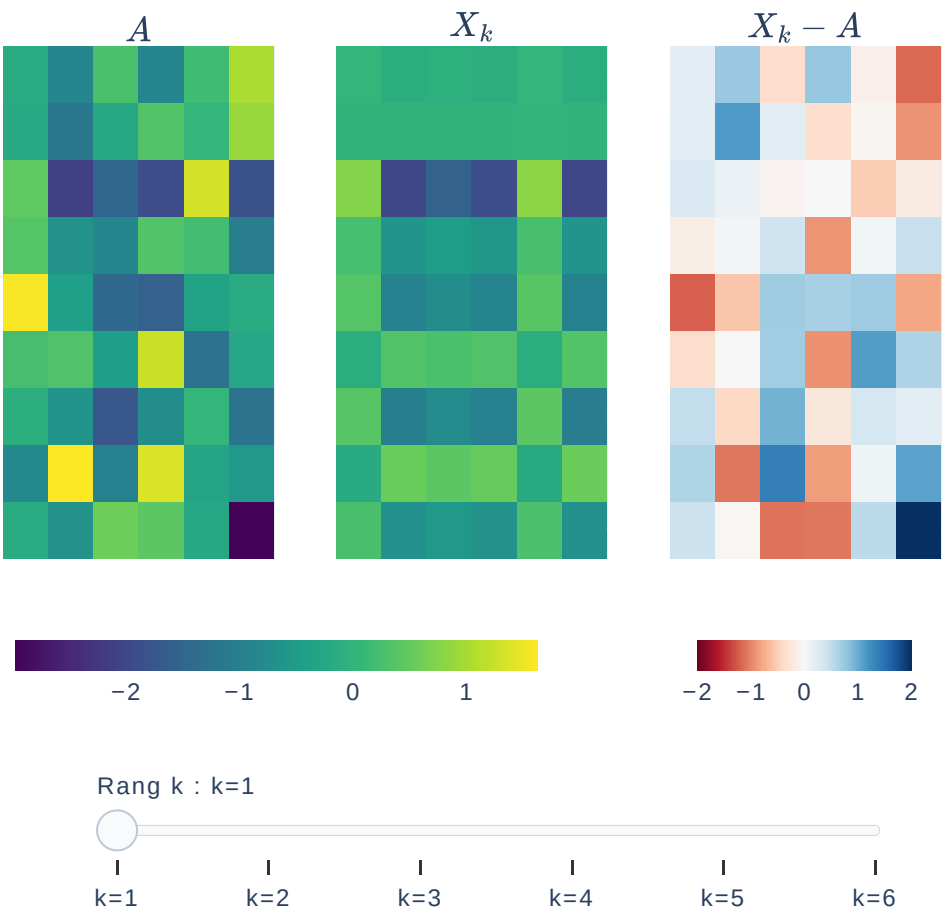
La norme spectrale est définie par

$$\|X\|_2 = \sup_{\|u\|=1} \|Xu\| = s_1(X).$$

Ce résultat est crucial pour l'analyse en composantes principales (ACP).

8 SVD et meilleure approximation : illustration

Approximation de rang k d'une matrice gaussienne $A \in \mathbb{R}^{9 \times 6}$:


A
X_k
$X_k - A$
−2 −1 0 1
−2 −1 0 1 2
Rang k : k=1
k=1 k=2 k=3 k=4 k=5 k=6