TD N° 1 : Introduction et rappels

Exercice 1. Montrer que pour toute matrice $X \in \mathbb{R}^{n\times m}$, $\mathrm{Ker}(X) = \mathrm{Ker}(X^\top X)$. En déduire que les rangs suivant sont identiques : $\mathrm{rg}(X) = \mathrm{rg}(X^\top X) = \mathrm{rg}(XX^\top) = \mathrm{rang}(X^\top)$.

Correction:
Prenons $w \in \mathbb{R}^m$ tel que $Xw = 0$. On en déduit que $X^\top Xw = 0$,*i.e.*, $\mathrm{Ker}(X) \subset \mathrm{Ker}(X^\top X)$.
Réciproquement supposons que $X^\top Xw = 0$. Ainsi $w^\top X^\top Xw = 0$, c'est-à-dire que $\|Xw\|^2 = 0$, et donc que $Xw = 0$, d'où $\mathrm{Ker}(X^\top X) \subset \mathrm{Ker}(X)$, et ainsi $\boxed{\mathrm{Ker}(X) = \mathrm{Ker}(X^\top X)}$
On conclut en utilisant le théorème du rang : $\mathrm{rang}(X) + \dim(\mathrm{Ker}(X)) = m = \mathrm{rang}(X^\top X) + \dim(\mathrm{Ker}(X^\top X))$, *i.e.*, $\mathrm{rg}(X) = \mathrm{rang}(X^\top X)$.
Enfin, on obtient le dernier point en prenant le résultat ci-dessus pour la matrice $X^\top$, ainsi : $\mathrm{rang}(X^\top) + \dim(\mathrm{Ker}(X^\top)) = n = \mathrm{rang}(XX^\top) + \dim(\mathrm{Ker}(XX^\top))$, et en rappelant que $\mathrm{rang}(X) = \mathrm{rang}(X^\top)$

Exercice 2. Montrer que $\hat\beta^{(\ell_2)} \triangleq X^+y$ est une solution du problème des moindres carrées :

$$\operatorname*{arg\,min}_{\beta\in\mathbb{R}^p} \|y - X\beta\|^2 \quad , \tag{1}$$

avec $y \in \mathbb{R}^n$ et $X \in \mathbb{R}^{n\times p}$, et que de plus parmi toute les solutions c'est la solution de norme (euclidienne) minimale.

Correction:
Partons de la SVD de X, $X = \sum_{i=1}^r s_i u_i v_i^\top$ (avec $(u_1, \dots, u_n)$ base orthonormée de $\mathbb{R}^n$, et $(v_1, \dots, v_n)$ base orthonormée de $\mathbb{R}^p$) :

$$\begin{aligned}
\|X\beta - y\|^2 &= \left\| \sum_{i=1}^r s_i u_i v_i^\top \beta - \sum_{i=1}^n u_i u_i^\top y \right\|^2 \\
\|X\beta - y\|^2 &= \left\| \sum_{i=1}^r u_i (s_i v_i^\top \beta - u_i^\top y) - \sum_{i=r+1}^n u_i u_i^\top y \right\|^2 \\
\|X\beta - y\|^2 &= \left\| \sum_{i=1}^r u_i (s_i v_i^\top \beta - u_i^\top y) \right\|^2 + \left\| \sum_{i=r+1}^n u_i u_i^\top y \right\|^2 && \text{(par orthogonalité)} \\
\|X\beta - y\|^2 &= \sum_{i=1}^r \left(s_i v_i^\top \beta - u_i^\top y \right)^2 + \sum_{i=r+1}^n (u_i^\top y)^2
\end{aligned}$$

Prenons alors $\hat\beta^{(\ell_2)} = X^+y = \sum_{i=1}^r \frac{1}{s_i} v_i u_i^\top y = \sum_{i=1}^r \left(\frac{1}{s_i} u_i^\top y\right) v_i$. On remarque le premier terme du seconde membre est alors annulé. Comme toute l'ensemble de solution est affine, il s'écrit alors $\hat\beta^{(\ell_2)} + \mathrm{Ker}(X)$. Mais l'on remarque comme $\mathrm{Ker}(X) = \mathrm{vect}(v_{r+1}, \dots, v_p)$ que toute solution $\hat\beta$ peut s'écrire $\hat\beta^{(\ell_2)} + \sum_{i=r+1}^p \alpha_i v_i$, et que par orthogonalité :

$$\left\| \hat\beta^{(\ell_2)} + \sum_{i=r+1}^p \alpha_i v_i \right\|^2 = \left\| \hat\beta^{(\ell_2)} \right\|^2 + \left\| \sum_{i=r+1}^p \alpha_i v_i \right\|^2 = \left\| \hat\beta^{(\ell_2)} \right\|^2 + \sum_{i=r+1}^p \alpha_i^2 \quad . \tag{2}$$

On en déduit donc que

$$\hat\beta^{(\ell_2)} = \operatorname*{arg\,min}_{\beta\in\hat\beta^{(\ell_2)}+\mathrm{Ker}(X)} \|\beta\|^2 \tag{3}$$

Exercice 3.

1) Calculer la SVD de la matrice

$$X = \begin{bmatrix} \frac{\mathbb{1}_n}{\sqrt{n}} & \frac{\mathbb{1}_{C_1}}{\sqrt{n_1}} & \frac{\mathbb{1}_{C_2}}{\sqrt{n_2}} \end{bmatrix} \ ,$$

en prenant des vecteurs $\mathbb{1}_{C_1}, \mathbb{1}_{C_2}$ les indicatrices d'ensembles C_1, C_2 formant une partition de l'ensemble $[\![1, n]\!]$, en supposant qu'il y a n_1 (resp. n_2) observations dans la classe C_1 (resp. C_2). On notera que $\mathbb{1}_{C_1} + \mathbb{1}_{C_2} = \mathbb{1}_n$, et $\mathbb{1}_1\mathbb{1}_2 = 0 \in \mathbb{R}^n$.

2) Donner X^+, la pseudo-inverse de la matrice X.

<u>Correction</u>:

1)

$$X^\top X = \begin{bmatrix} 1 & \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}} \\ \sqrt{\frac{n_1}{n}} & 1 & 0 \\ \sqrt{\frac{n_2}{n}} & 0 & 1 \end{bmatrix} \ . \tag{4}$$

Tout d'abord X est de rang 2, donc $\lambda_0 = 0$ est valeur propre de $X^\top X$. Un vecteur propre associé est facilement trouvé comme on peut voir que $\sqrt{n}X_{:,1} = \sqrt{n_1}X_{:,2} + \sqrt{n_2}X_{:,3}$

$$v_0 = \frac{1}{\sqrt{2n}} \cdot \begin{bmatrix} -\sqrt{n} \\ \sqrt{n_1} \\ \sqrt{n_2} \end{bmatrix} \ . \tag{5}$$

De plus, on voit aussi que la seconde valeur propre est $\lambda_1 = 1$ avec comme vecteur propre associé

$$v_1 = \frac{1}{\sqrt{n}} \cdot \begin{bmatrix} 0 \\ -\sqrt{n_2} \\ \sqrt{n_1} \end{bmatrix} \ . \tag{6}$$

Enfin comme la trace de $X^\top X$ vaut 3, on en déduit que la dernière valeur propre est $\lambda_2 = 2$. En résolvant le système linéaire associé, on trouve que x_2 est donné par :

$$v_2 = \frac{1}{\sqrt{2n}} \cdot \begin{bmatrix} \sqrt{n} \\ \sqrt{n_1} \\ \sqrt{n_2} \end{bmatrix} \ . \tag{7}$$

On trouve ensuite

$$u_1 = Xv_1/\sqrt{\lambda_1} = \sqrt{\frac{n_1}{nn_2}}\mathbb{1}_{C_2} - \sqrt{\frac{n_2}{nn_1}}\mathbb{1}_{C_1}, \tag{8}$$

$$u_2 = Xv_2/\sqrt{\lambda_2} = \frac{\mathbb{1}_n}{\sqrt{n}} \ . \tag{9}$$

On déduit donc

$$X = \sqrt{\lambda_1}u_1v_1^\top + \sqrt{\lambda_2}u_2v_2^\top = u_1v_1^\top + \sqrt{2}u_2v_2^\top \tag{10}$$

2)

$$X^+ = v_1u_1^\top + \frac{1}{\sqrt{2}}v_2u_2^\top \ . \tag{11}$$

On en déduit alors

$$\begin{aligned} \hat{\beta}^{(\ell_2)} &= X^+y = v_1u_1^\top y + \frac{1}{\sqrt{2}}v_2u_2^\top y \\ &= \sqrt{\frac{n_1n_2}{n}}\left(\bar{y}_{C_2} - \bar{y}_{C_1}\right)v_1 + \sqrt{\frac{n}{2}}\bar{y}_nv_2 \ . \end{aligned}$$

Exercice 4.

1) Calculer la SVD de la matrice

$$X = \begin{bmatrix}\alpha_1 x_1 + \alpha_2 x_2 & x_1 & x_2\end{bmatrix} \ ,$$

sous la contrainte $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$, $\|x_1\| = \|x_2\| = 1$, $x_1^\top x_2 = 0$ et $(\alpha_1, \alpha_2) \in \mathbb{R}^2$.

2) Donner X^+, la pseudo-inverse de la matrice X.

<u>**Correction:**</u>

1)

$$X^\top X = \begin{bmatrix} \alpha_1^2 + \alpha_2^2 & \alpha_1 & \alpha_2 \\ \alpha_1 & 1 & 0 \\ \alpha_2 & 0 & 1 \end{bmatrix} \ . \tag{12}$$

Tout d'abord X est de rang 2, donc $\lambda_0 = 0$ est valeur propre de $X^\top X$. Un vecteur propre associé est facilement trouvé :

$$v_0 = \frac{1}{\sqrt{1+\alpha_2^2+\alpha_2^2}} \cdot \begin{bmatrix} -1 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} \ . \tag{13}$$

De plus, on voit aussi que la seconde valeur propre est $\lambda_1 = 1$ avec comme vecteur propre associé

$$v_1 = \frac{1}{\sqrt{\alpha_2^2+\alpha_2^2}} \cdot \begin{bmatrix} 0 \\ -\alpha_2 \\ \alpha_1 \end{bmatrix} \ . \tag{14}$$

Enfin comme la trace de $X^\top X$ vaut $2 + \alpha_1^2 + \alpha_2^2$, on en déduit que la dernière valeur propre est $\lambda_2 = 1 + \alpha_1^2 + \alpha_2^2$. En résolvant le système linéaire associé, on trouve que x_2 est donné par :

$$v_2 = \frac{1}{\sqrt{(\alpha_2^2+\alpha_2^2)^2+\alpha_2^2+\alpha_2^2}} \cdot \begin{bmatrix} \alpha_1^2+\alpha_2^2 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} \ . \tag{15}$$

Enfin en calculant $u_1 = Xv_1/\sqrt{\lambda_1}$ et $u_2 = Xv_2/\sqrt{\lambda_2}$ on trouve finalement

$$u_1 = \frac{\alpha_2 x_2 - \alpha_1 x_1}{\sqrt{\alpha_1^2+\alpha_2^2}}, \quad u_2 = \frac{\alpha_1 x_1 + \alpha_2 x_2}{\sqrt{\alpha_1^2+\alpha_2^2}} \ . \tag{16}$$

On déduit donc

$$X = \sqrt{\lambda_1} u_1 v_1^\top + \sqrt{\lambda_2} u_2 v_2^\top \tag{17}$$

2)

$$X^+ = \frac{1}{\sqrt{\lambda_1}} v_1 u_1^\top + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top \ . \tag{18}$$

On en déduit alors

$$\begin{aligned} \hat{\beta}^{(\ell_2)} &= X^+ y \\ &= \frac{1}{\sqrt{\lambda_1}} v_1 u_1^\top y + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top y \\ &= v_1(\alpha_2 x_2 - \alpha_1 x_1)^\top y + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top y \ . \end{aligned}$$

3) On repart de l'expression précédente en notant que $\alpha_1 = \alpha_2 = 1$

$$\hat{\beta}^{(\ell_2)} = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} (\mathbb{1}_{C_2}^\top y - \mathbb{1}_{C_1}^\top y) + \frac{1}{3\sqrt{2}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \mathbb{1}_n^\top y \ .$$

Exercice 5.

1) Calculer la SVD de la matrice

$$X = \begin{bmatrix} \mathbb{1}_n & \mathbb{1}_{C_1} & \mathbb{1}_{C_2} \end{bmatrix} \ ,$$

en prenant des vecteurs $\mathbb{1}_{C_1}, \mathbb{1}_{C_2}$ les indicatrices d'ensembles C_1, C_2 formant une partition de l'ensemble $[\![1, n]\!]$. On notera que $\mathbb{1}_{C_1} + \mathbb{1}_{C_2} = \mathbb{1}_n$, et $\mathbb{1}_1 \mathbb{1}_2 = 0 \in \mathbb{R}^n$, on supposera qu'il y a n_1 (resp. n_2) observations dans la classe C_1 (resp. C_2).

Correction:

1)

$$X^\top X = \begin{bmatrix} n & n_1 & n_2 \\ n_1 & n_1 & 0 \\ n_2 & 0 & n_2 \end{bmatrix} \ . \tag{19}$$

Tout d'abord X est de rang 2, donc $\lambda_0 = 0$ est valeur propre de $X^\top X$. Un vecteur propre associé est facilement trouvé :

$$v_0 = \frac{1}{\sqrt{3}} \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \ . \tag{20}$$

De plus, on voit aussi que la seconde valeur propre est $\lambda_1 =???$ avec comme vecteur propre associé

$$v_1 = \frac{1}{\sqrt{\alpha_2^2 + \alpha_2^2}} \cdot \begin{bmatrix} 0 \\ -\alpha_2 \\ \alpha_1 \end{bmatrix} \ . \tag{21}$$

Enfin comme la trace de $X^\top X$ vaut $2 + \alpha_1^2 + \alpha_2^2$, on en déduit que la dernière valeur propre est $\lambda_2 = 1 + \alpha_1^2 + \alpha_2^2$. En résolvant le système linéaire associé, on trouve que x_2 est donné par :

$$v_2 = \frac{1}{\sqrt{(\alpha_2^2 + \alpha_2^2)^2 + \alpha_2^2 + \alpha_2^2}} \cdot \begin{bmatrix} \alpha_1^2 + \alpha_2^2 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} \ . \tag{22}$$

Enfin en calculant $u_1 = Xv_1/\sqrt{\lambda_1}$ et $u_2 = Xv_2/\sqrt{\lambda_2}$ on trouve finalement

$$u_1 = \frac{\alpha_2 x_2 - \alpha_1 x_1}{\sqrt{\alpha_1^2 + \alpha_2^2}}, \quad u_2 = \frac{\alpha_1 x_1 + \alpha_2 x_2}{\sqrt{\alpha_1^2 + \alpha_2^2}} \ . \tag{23}$$

On déduit donc

$$X = \sqrt{\lambda_1} u_1 v_1^\top + \sqrt{\lambda_2} u_2 v_2^\top \tag{24}$$

2)

$$X^+ = \frac{1}{\sqrt{\lambda_1}} v_1 u_1^\top + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top \ . \tag{25}$$

On en déduit alors

$$\begin{aligned} \hat{\beta}^{(\ell_2)} &= X^+ y \\ &= \frac{1}{\sqrt{\lambda_1}} v_1 u_1^\top y + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top y \\ &= v_1 (\alpha_2 x_2 - \alpha_1 x_1)^\top y + \frac{1}{\sqrt{\lambda_2}} v_2 u_2^\top y \ . \end{aligned}$$

3) On repart de l'expression précédente en notant que $\alpha_1 = \alpha_2 = 1$

$$\hat{\beta}^{(\ell_2)} = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} (\mathbb{1}_{C_2}^\top y - \mathbb{1}_{C_1}^\top y) + \frac{1}{3\sqrt{2}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \mathbb{1}_n^\top y \ .$$