# Feuille d'exercices n$^o$9
## Corrigé

**Exercice 1**

1. a) Par récurrence sur $k$ : pour $k = 0$, c'est vrai. Si c'est vrai pour $k$, on a $Ax^{(k+1)} = Ax^{(k)} + \frac{\langle A_i, r^{(k)}\rangle}{||A_i||_2^2} Ae_i = b - r^{(k)} + \frac{\langle A_i, r^{(k)}\rangle}{||A_i||_2^2} A_i = b - r^{(k+1)}$ donc c'est vrai pour $k+1$.
b) Posons $E = \text{Vect}\{A_1, ...A_n\}$.
L'application $y \in E \to \max_{j \leq n} \left|\frac{\langle y, A_j\rangle}{||A_j||_2}\right|$ est une norme sur $E$ et $y \in E \to ||y||_2$ aussi.
Puisqu'en dimension finie toutes les normes sont équivalentes, il existe $\alpha > 0$ tel que, pour tout $y \in E$ :

$$\max_{j \leq n} \left|\frac{\langle y, A_j\rangle}{||A_j||_2}\right| \geq \alpha ||y||_2$$

Quitte à remplacer $\alpha$ par $\min(\alpha, 1/2)$, on peut supposer que $\alpha \in ]0; 1[$.
c) On procède par récurrence. Pour $k = 0$, c'est une égalité.
Si c'est vrai pour $k$, pour $k+1$ :

$$||r^{(k+1)}||_2^2 = ||r^{(k)}||_2^2 - 2\frac{\langle A_i, r^{(k)}\rangle^2}{||A_i||_2^2} + \frac{\langle A_i, r^{(k)}\rangle^2}{||A_i||_2^2} = ||r^{(k)}||_2^2 - \frac{\langle A_i, r^{(k)}\rangle^2}{||A_i||_2^2}$$

D'après la question précédente, $\left|\frac{\langle A_i, r^{(k)}\rangle}{||A_i||_2}\right| = \max_{j \leq n} \left|\frac{\langle y, A_j\rangle}{||A_j||_2}\right| \geq \alpha ||r^{(k)}||_2$ donc :

$$||r^{(k+1)}||_2^2 \leq ||r^{(k)}||_2^2 - (\alpha ||r^{(k)}||_2)^2 = (1-\alpha^2)||r^{(k)}||_2^2 \leq (1-\alpha^2)(1-\alpha^2)^k ||b||_2^2 = (1-\alpha^2)^{k+1}||b||_2^2$$

d'où le résultat pour $k+1$.

**Exercice 2**

1. On a $R_X[k] = \mathbb{E}(X[0]X[k])$. On a donc $R_X[k] \in \mathbb{R}$ pour tout $k$.
Soit $N$ comme dans la définition d'un processus à mémoire finie. On note $R_X[k] = a_k$ pour tout $k = 0, ..., N$. On sait que :

$$a_k = R_X[k] = R_X[-k]$$

Donc :

$$\begin{aligned}\hat{R}_X(e^{i\omega}) &= \sum_{n \in \mathbb{Z}} R_X[n] e^{-in\omega} \\ &= a_0 + \sum_{k=1}^{N} R_X[k] e^{-ik\omega} + \sum_{k=1}^{N} R_X[-k] e^{-i(-k)\omega} \\ &= a_0 + \sum_{k=1}^{N} a_k (e^{-ik\omega} + e^{ik\omega})\end{aligned}$$

2. On sait que $R_{h\star B} = R_B \star h \star \tilde{h}$ où $\tilde{h}[k] = h[-k]$. Puisque $B$ est un bruit blanc de variance 1, $R_B[k] = 1$ si $k = 0$ et $R_B[k] = 0$ si $k \neq 0$. Donc :

$$\begin{aligned} R_X[k] &= R_B \star h \star \tilde{h}[k] \\ &= \sum_{n\in\mathbb{Z}} R_B[n](h \star \tilde{h})[k-n] \\ &= (h \star \tilde{h})[k] \\ &= \sum_{n\in\mathbb{Z}} h[k-n]\tilde{h}[n] \\ &= \sum_{n\in\mathbb{Z}} h[k-n]h[-n] \end{aligned}$$

Si $|k| > N$, tous les termes de la somme sont nuls. En effet, on ne peut pas avoir $-n \in \{0, ..., N\}$ et $k - n \in \{0, ..., N\}$ (sinon $k = (k-n) - (-n) \in \{-N, ..., N\}$). Donc $R_X[k] = 0$ si $|k| > N$.

3. $F(e^{i\omega}) = P(e^{-i\omega})e^{iN\omega}$.

4. $P(z^{-1}) = z^{-N}(a_0 + \sum\limits_{n=1}^{N} a_n(z^{-n} + z^n)) = z^{-2N}P(z)$

Soient $\alpha_1, ..., \alpha_{2N}$ les racines de $P$ (comptées avec multiplicité).

On divise l'ensemble des $\alpha_i$ en trois ensembles $E_1, E_2, E_3$ qui contiennent respectivement les $\alpha_i$ tels que $|\alpha_i| > 1$, les $\alpha_i$ tels que $|\alpha_i| < 1$ et les $\alpha_i$ tels que $|\alpha_i| = 1$.

Comme $a_N \neq 0$ et comme $P(0) = a_N$, 0 n'est pas racine de $P$.

Pour tout $\alpha_i \in E_1$, $P(\alpha_i) = 0$ donc $P(\alpha_i^*) = 0$ (car $P$ est réel) et $P(1/\alpha_i^*) = \alpha_i^{*-2N}P(\alpha_i^*) = 0$. Donc $1/\alpha_i^*$ est aussi racine de $P$ (on peut vérifier que les multiplicités coïncident également en faisant un développement limité au voisinage de $\alpha_i$). De plus, $1/\alpha_i^* \in E_2$ (car $|1/\alpha_i^*| < 1$).

Réciproquement, si $\alpha_i \in E_2$, alors $1/\alpha_i^* \in E_1$ (et avec la même multiplicité).

Posons $z_1, ..., z_S$ les racines appartenant à $E_1$ (comptées autant de fois que leur multiplicité). D'après le raisonnement précédent, les racines appartenent à $E_2$ sont exactement $1/z_1^*, ..., 1/z_S^*$.

Pour toute racine $\alpha_i \in E_3$, $\alpha_i$ a une multiplicité paire (sinon $F(e^{i\omega})$ change de signe au voisinage de $e^{i\omega} = \alpha_i$, ce qui est impossible car on a supposé $F \geq 0$). On note $z_{S+1}, ..., z_N$ les racines contenues dans $E_3$, chaque racine apparaissant deux fois moins que sa multiplicité.

Les racines de $P$ sont donc :

$$\begin{aligned} &\{z_1, ..., z_S, 1/z_1^*, ..., 1/z_S^*, z_{S+1}, ..., z_N, z_{S+1}, ..., z_N\} \\ &\qquad = \{z_1, ..., z_S, 1/z_1^*, ..., 1/z_S^*, z_{S+1}, ..., z_N, 1/z_{S+1}^*, ..., 1/z_N^*\} \end{aligned}$$

(car pour tout $z$ tel que $|z| = 1$, $z = 1/z^*$)

Donc :

$$P(z) = a_N \prod_i (z - z_i) \prod_i (z - z_i^{*-1})$$

5.

$$\begin{aligned} -z_i^{*-1}|e^{-i\omega} - z_i|^2 &= -z_i^{*-1}(e^{-i\omega} - z_i)(e^{i\omega} - z_i^*) \\ &= -e^{i\omega} z_i^{*-1}(e^{-i\omega} - z_i)(1 - z_i^* e^{-i\omega}) \\ &= -e^{i\omega}(e^{-i\omega} - z_i)(z_i^{*-1} - e^{-i\omega}) \\ &= e^{i\omega}(e^{-i\omega} - z_i)(e^{-i\omega} - z_i^{*-1}) \end{aligned}$$

$$\begin{aligned}
F(e^{i\omega}) &= e^{iN\omega}P(e^{-i\omega}) \\
&= a_N \prod_{i=1}^{N} e^{i\omega}(e^{-i\omega} - z_i)(e^{-i\omega} - z_i^{-1*}) \\
&= a_N \prod_{i=1}^{N} - z_i^{*-1}|e^{-i\omega} - z_i|^2 \\
&= c|G(e^{-i\omega})|^2
\end{aligned}$$

si on pose $c = a_N \prod_i z_i^{*-1}$ et $G(e^{-i\omega}) = \prod_i (e^{-i\omega} - z_i)$.

Puisque $F(e^{i\omega})$ est un réel positif pour tout $\omega \in \mathbb{R}$, on doit avoir $c \in \mathbb{R}^+$. De plus, $G$ est à coefficients réels car, vue la façon dont on a construit les $z_k$ dans la question précédente et vu le fait que $P$ est à coefficients réels (ce qui implique que l'ensemble de ses racines est stable par conjugaison), l'ensemble des racines de $G$ est stable par conjugaison.
Si on pose $Q(e^{-i\omega}) = \sqrt{c}G(e^{-i\omega})$, on a le résultat voulu.

6. On pose $F = R_X$.
Soit $Q$ comme dans la question précédente tel que $F(e^{i\omega}) = |Q(e^{-i\omega})|^2$. On se contente de traiter le cas où $Q$ n'a pas de zéro sur le cercle unité.
Soit $h$ le filtre tel que $h[k] = h_k$ si $0 \leq k \leq N$ (où les $h_k$ sont les coefficients de $Q$) et $h[k] = 0$ sinon. C'est un filtre causal à réponse impulsionnelle finie.
Alors $\hat{h}(e^{i\omega}) = Q(e^{-i\omega})$.
Soit $r$ le filtre tel que $\hat{r}(e^{i\omega}) = 1/\hat{h}(e^{i\omega})$. Si $Q$ n'a pas de zéros sur le cercle unité, ce filtre est bien défini (et décroît très vite car $\hat{h}$ est $\mathcal{C}^\infty$, ce qui résoud les problèmes de convergence).
Posons $B = X \star r$ (où $X$ est le processus à mémoire finie). Alors $B$ est un bruit blanc car $\hat{R}_B(e^{i\omega}) = |\hat{r}(e^{i\omega})|^2 \hat{R}_X(e^{i\omega}) = \frac{F(e^{i\omega})}{|\hat{h}(e^{i\omega})|^2} = \frac{F(e^{i\omega})}{|Q(e^{-i\omega})|^2} = 1$.
De plus, $r \star h = \delta_0$ (car $\hat{r}\hat{h} = 1$) donc $X = X \star \delta_0 = X \star (r \star h) = B \star h$.
Note : lorsque $Q$ a des zéros sur le cercle unité, on peut adapter la démonstartion en utilisant une suite de $r$ telle que $\hat{r}$ converge (pour une bonne notion de convergence) vers $1/\hat{h}$ et en montrant que la suite des $X \star r$ associée converge au sens $L^2$ vers une variable aléatoire qui est un bruit blanc.

**Exercice 3**

1. C'est la loi des grands nombres (c'est d'ailleurs dans le cours) :

$$\log_2(1/p(X_1, ..., X_n)) = \log_2(1/p(X_1)) + ... + \log_2(1/p(X_n))$$

Comme les variables aléatoires $\log_2(1/p(X_k))$ sont indépendantes et de même loi :

$$\frac{\log_2(1/p(X_1, ..., X_n))}{n} \to E(\log_2(1/p(X))) = H$$

quand $n \to +\infty$. La convergence est par exemple à prendre au sens de la convergence presque sûrement.

Donc :

$$\mathbb{P}\left(\left|\frac{\log_2(1/p(X_1,...,X_n))}{n} - H\right| > \delta\right) \to 0 \quad \text{quand } n \to +\infty$$

En particulier, cette quantité est plus petite que $\epsilon$ pour tout $n$ assez grand. Dès qu'elle est plus petite que $\epsilon$, on note $B_1$ l'ensemble des valeurs telles que $\left|\frac{\log_2(1/p(X_1,...,X_n))}{n} - H\right| > \delta$ et $B_2$ son complémentaire.

2. Soient $\epsilon, \delta > 0$ quelconques, avec $\epsilon < q$. Pour tout $n$ assez grand, il existe $B_1$ et $B_2$ comme dans la question précédente.

La somme des probabilités des éléments de $E_{n,q} - B_1$ est au moins $q - \epsilon$. Donc :

$$q - \epsilon \leq \sum_{a \in E_{n,q} - B_1} p(a) \leq \sum_{a \in E_{n,q} - B_1} 2^{-n(H-\delta)} \leq 2^{-n(H-\delta)} \text{Card}(E_{n,q})$$

Donc $\text{Card}(E_{n,q}) \geq (q-\epsilon) 2^{n(H-\delta)}$, ce qui implique :

$$\frac{\log_2\left(\text{Card}(E_{n,q})\right)}{n} \geq \frac{\log_2(q-\epsilon)}{n} + H - \delta$$

En choisissant correctement $\delta$ en en prenant $n$ assez grand, on a l'inégalité voulue.

3. Supposons $\sum_{a \in A^n} p(a) \frac{l_n(a)}{n} < H - \eta$. Soit $B$ l'ensemble des $a \in A^n$ tels que $l_n(a)/n \geq H - \eta/2$. On doit avoir :

$$H - \eta > \sum_{a \in B} p(a) \frac{l_n(a)}{n} \geq (H - \eta/2) \sum_{a \in B} p(a)$$

Donc $\sum_{a \in B} p(a) < \frac{H-\eta}{H-\eta/2} < 1$.

Soit $q = 1 - \frac{H-\eta}{H-\eta/2}$.

La somme des probabilités des éléments de $A^n - B$ est au moins $q$ (car la somme totale des probabilités des éléments de $A^n$ vaut 1). Donc, par définition de $E_{n,q}$, $\text{Card}(A^n - B) \geq \text{Card}(E_{n,s})$.

Pour tout $a \in A^n - B$, $l_n(a) < n(H - \eta/2)$.

Or le nombre de suites de zéros et de uns de longueur strictement inférieure à $n(H-\eta/2)$ est inférieur à $2^{n(H-\eta/2)+1}$. Puisque chaque élément de $A^n - B$ est associé à un élément différent ($C_n$ est injective) :

$$\text{Card}(A^n - B) \leq 2^{n(H-\eta/2)+1} \quad \Rightarrow \quad \frac{\log_2(E_{n,q})}{n} \leq \frac{\log_2\left(\text{Card}(A^n - B)\right)}{n} \leq H - \eta/2 + \frac{1}{n}$$

D'après la question précédente, c'est impossible pour tout $n$ assez grand.

### Exercice 4

1. a)

$$\begin{aligned} H^{(k)}(y) - ||{}^tA(Ax^{(k)} - b)||_2^2 + ||Ax^{(k)} - b||_2^2 &= ||y||_2^2 + 2\langle y, {}^tA(Ax^{(k)} - b)\rangle + ||Ax^{(k)} - b||_2^2 + 2\lambda ||x^{(k)} + y||_1 \\ &= ||y||_2^2 + 2\langle Ay, Ax^{(k)} - b\rangle + ||Ax^{(k)} - b||_2^2 + 2\lambda ||x^{(k)} + y||_1 \\ &\geq ||Ay||_2^2 + 2\langle Ay, Ax^{(k)} - b\rangle + ||Ax^{(k)} - b||_2^2 + 2\lambda ||x^{(k)} + y||_1 \\ &= ||Ay + Ax^{(k)} - b||_2^2 + 2\lambda ||x^{(k)} + y||_1 \\ &= J(x^{(k)} + y) \end{aligned}$$

Pour l'inégalité, on a utilisé le fait que $|||A||| \leq 1$.
b) Il suffit de l'écrire.
c) Posons $r = {}^tA(Ax^{(k)} - b)$. Alors :

$$H^{(k)}(y) = \sum_i (y_i + r_i)^2 + 2\lambda|x_i^{(k)} + y_i|$$

Donc $H^{(k)}(y) - H^{(k)}(0) = \sum_i y_i(y_i + 2r_i) + 2\lambda(|x_i^{(k)} + y_i| - |x_i^{(k)}|)$.
On veut montrer que ce terme est négatif si $y = x^{(k+1)} - x^{(k)}$. Montrons donc que, pour ce $y$, on a pour tout $i$ :

$$y_i(y_i + 2r_i) + 2\lambda(|x_i^{(k)} + y_i| - |x_i^{(k)}|) \leq 0$$

On a $y_i = S_\lambda(x_i^{(k)} - r_i) - x_i^{(k)}$.

– Premier cas : $x_i^{(k)} - r_i \geq \lambda$. Alors $y_i = -r_i - \lambda$ et :

$$\begin{aligned}
y_i(y_i + 2r_i) + 2\lambda(|x_i^{(k)} + y_i| - |x_i^{(k)}|) &= -(r_i + \lambda)(r_i - \lambda) + 2\lambda(|x_i^{(k)} - r_i - \lambda| - |x_i^{(k)}|) \\
&= \lambda^2 - r_i^2 + 2\lambda(x_i^{(k)} - r_i - \lambda - |x_i^{(k)}|) \\
&\leq \lambda^2 - r_i^2 + 2\lambda(x_i^{(k)} - r_i - \lambda - x_i^{(k)}) \\
&= -\lambda^2 - 2\lambda r_i - r_i^2 \\
&= -(r_i + \lambda)^2 \leq 0
\end{aligned}$$

avec égalité si et seulement si $y_i = 0$.

– Deuxième cas : $x_i^{(k)} - r_i \leq -\lambda$. Même raisonnement.

– Troisième cas : $-\lambda \leq x_i^{(k)} - r_i \leq \lambda$. Alors $y_i = -x_i^{(k)}$.

$$\begin{aligned}
y_i(y_i + 2r_i) + 2\lambda(|x_i^{(k)} + y_i| - |x_i^{(k)}|) &= -x_i^{(k)}(2r_i - x_i^{(k)}) - 2\lambda|x_i^{(k)}| \\
&= -|x_i^{(k)}|(2\lambda + \text{signe}(x_i^{(k)})(2r_i - x_i^{(k)})) \\
&\leq -2|x_i^{(k)}|(\lambda + \text{signe}(x_i^{(k)})(r_i - x_i^{(k)})) \\
&\leq -2|x_i^{(k)}|(\lambda - |r_i - x_i^{(k)}|) \\
&\leq 0
\end{aligned}$$

L'égalité impose $x_i^{(k)} = 0$.

Donc le terme voulu est bien négatif, avec égalité si et seulement si $y_i = 0$ pour tout $i$.
d) Si $x^{(k+1)} \neq x^{(k)}$ :

$$\begin{aligned}
J(x^{(k+1)}) &\leq H^{(k)}(x^{(k+1)} - x^{(k)}) - ||{}^tA(Ax^{(k)} - b)||_2^2 + ||Ax^{(k)} - b||_2^2 \\
&< H^{(k)}(0) - ||{}^tA(Ax^{(k)} - b)||_2^2 + ||Ax^{(k)} - b||_2^2 \\
&= J(x^{(k)})
\end{aligned}$$

2. a) $J(x) \to +\infty$ si $||x||_1 \to +\infty$. Puisque $(J(x^{(k)}))_k$ est décroissante, $(J(x^{(k)}))_k$ est bornée et $(x^{(k)})_k$ aussi. On peut donc en extraire une sous-suite convergente $(x^{(\phi(k))})_k$, de limite $\tilde{x}$.

Puisque $J$ est continue, $J(x^{(k)}) \to J(\tilde{x})$.
D'après la question d), $J(S_\lambda(\tilde{x} - {}^tA(A\tilde{x} - b))) < J(\tilde{x})$, sauf si $S_\lambda(\tilde{x} - {}^tA(A\tilde{x} - b)) = \tilde{x}$.
Or $x^{(\phi(k)+1)} \to S_\lambda(\tilde{x} - {}^tA(A\tilde{x} - b))$ et $J(x^{(\phi(k)+1)}) \to J(\tilde{x})$. Donc $J(S_\lambda(\tilde{x} - {}^tA(A\tilde{x} - b))) = J(\tilde{x})$ et $S_\lambda(\tilde{x} - {}^tA(A\tilde{x} - b)) = \tilde{x}$.
b) D'après l'égalité obtenue à la question précédente, ${}^tA(A\tilde{x} - b) = -\lambda$ si $\tilde{x} > 0$, ${}^tA(A\tilde{x} - b) = \lambda$ si $\tilde{x} < 0$ et $-\lambda \leq {}^tA(A\tilde{x} - b) \leq \lambda$ si $\tilde{x} = 0$.
On peut en déduire que, pour tout $y$, $\lambda||\tilde{x} + y||_1 \geq \lambda||\tilde{x}||_1 + \langle y, -{}^tA(A\tilde{x} - b)\rangle$.
De plus, pour tout $y$, $||A(\tilde{x} + y) - b||^2 - ||A\tilde{x} - b||^2 \geq +2\langle y, {}^tA(A\tilde{x} - b)\rangle$.
En sommant les deux précédentes inégalités, on obtient, pour tout $y$ :

$$J(\tilde{x} + y) - J(y) \geq 2\langle y, {}^tA(A\tilde{x} - b)\rangle - 2\langle y, {}^tA(A\tilde{x} - b)\rangle = 0$$

c) Posons $R : x \to S_\lambda(x - {}^tA(Ax - b))$. C'est une fonction 1-lipschitzienne (car $|||I - {}^tAA||| \leq 1$ et $S_\lambda$ est 1-lipschitzienne).
Alors, pour tout $k$ :

$$||x^{(k+1)} - \tilde{x}||_2 = ||R(x^{(k)}) - R(\tilde{x})||_2 \leq ||x^{(k)} - \tilde{x}||$$

De plus, il existe une sous-suite de $(x^{(k)})_k$ qui converge vers $\tilde{x}$. Donc $x^{(k)} \to \tilde{x}$.