Lycée Fénelon | PCSI
Année 2025-2026 | A. Panetta

Devoir surveillé de mathématiques n°5
Samedi 24 janvier 2026 (4h00)

L'énoncé est constitué de deux problèmes et comporte 6 pages.
Le candidat attachera la plus grande importance à la clarté, la précision et la concision de la rédaction. Le soin de la copie ainsi que l'orthographe entreront également pour une part importante dans l'appréciation du travail rendu.
Les résultats doivent être encadrés.
Si un candidat repère ce qui peut lui sembler être une erreur d'énoncé, il le signalera sur sa copie et devra poursuivre sa composition en expliquant les raisons des initiatives qu'il a été amené à prendre.
Lorsqu'un raisonnement utilise le résultat d'une question précédente, il est demandé au candidat d'indiquer précisément le numéro de la question utilisée.

Les calculatrices sont interdites.

Problème 1 : Matrices productives

Partie I : Résultats théoriques

Soit n un entier naturel supérieur ou égal à 2. Pour tout $p \in \mathbb{N}^*$, on note $\mathcal{M}_{n,p}$ l'ensemble des matrices à n lignes et p colonnes à cœfficients réels.

• Pour toute matrice $M \in \mathcal{M}_{n,p}(\mathbb{R})$, la notation $M \geqslant 0$ (respectivement $M > 0$) signifie que tous les cœfficients de M sont positifs (respectivement strictement positifs).

On dit alors que M est **positive** (respectivement **strictement positive**).

• Pour toutes matrices $(M, N) \in (\mathcal{M}_{n,p}(\mathbb{R}))^2$, la notation $M \geqslant N$ (respectivement $M > N$) signifie que $M - N \geqslant 0$ (respectivement $M - N > 0$).

1. Soit $B \in \mathcal{M}_{n,n}(\mathbb{R})$.
   (a) Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Montrer que si $B \geqslant 0$ et $X \geqslant 0$, alors on a $BX \geqslant 0$.
   (b) Etablir, réciproquement, que si, pour toute matrice colonne $X \in \mathcal{M}_{n,1}(\mathbb{R})$ positive, on a $BX \geqslant 0$, alors B est positive.

Si $A \in \mathcal{M}_{n,n}(\mathbb{R})$ est une matrice carrée, on dit qu'elle est **productive** si A est positive et s'il existe une matrice colonne $P \in \mathcal{M}_{n,1}(\mathbb{R})$ positive, telle que $P > AP$.

Dans les questions 2 à 4, on considère deux matrices $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ et $P = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$ vérifiant les conditions de cette définition.

2. Montrer que $P > 0$.
3. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ appartenant à $\mathcal{M}_{n,1}(\mathbb{R})$ telle que $X \geqslant AX$.

   On pose $c = \min \left\{ \dfrac{x_j}{p_j}, j \in [\![1, n]\!]. \right\}$.

   Soit $k \in [\![1, n]\!]$ tel que $c = \dfrac{x_k}{p_k}$.

   (a) Etablir que :
   $$c \left(p_k - \sum_{j=1}^{n} a_{k,j} p_j \right) \geqslant \sum_{j=1}^{n} a_{k,j}(x_j - c p_j).$$
   (b) En déduire que $c \geqslant 0$ puis que X est positive.
   (c) On suppose dorénavant que $AX = X$.
   En utilisant l'inégalité $-X \geqslant A(-X)$, montrer que $X = 0$ et en déduire que $(I_n - A)$ est inversible.
4. (a) Soit $X \geqslant 0$. Montrer que $Y = (I_n - A)^{-1} X$ est positive. En déduire que $(I_n - A)^{-1}$ est positive.
   (b) On considère dans cette question B une matrice carrée positive telle que $I_n - B$ soit inversible et d'inverse positive.
   Soit $V = (I_n - B)^{-1} U$, où U est la matrice colonne dont toutes les composantes valent 1.
   Montrer que $V > BV$ et en déduire que B est productive.
5. Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$. Montrer que les deux assertions suivantes sont équivalentes :
   (a) $A \in \mathcal{M}_{n,n}(\mathbb{R})$ est productive ;
   (b) $A \geqslant 0, (I_n - A)$ est inversible et $(I_n - A)^{-1} \geqslant 0$.

6. Montrer que si A est une matrice productive, alors sa transposée A^T est également une matrice productive.
7. Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice positive et nilpotente. Montrer que A est productive.
   Indication : on pourra utiliser le fait que si $A \geqslant 0$, alors pour tout $k \in \mathbb{N}, A^k \geqslant 0$.

Partie II : Exemples de matrices productives

1. Soit $\lambda \in \mathbb{R}$. A quelle condition sur λ la matrice scalaire λI_n est-elle productive ?
   Plus généralement, si $D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_n \end{pmatrix}$ est une matrice diagonale, à quelle condition sur les $(d_i)_{i\in[\![1,n]\!]}$ la matrice D est-elle productive ?
2. Soit $a \in \mathbb{R}_+$. Justifier que la matrice $A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ est productive et trouver une matrice colonne $P \in \mathcal{M}_{n,1}(\mathbb{R})$ positive telle que $P > AP$.
3. Montrer que la matrice $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix}$ est productive et trouver une matrice colonne $Q \in \mathcal{M}_{n,1}(\mathbb{R})$ positive telle que $Q > BQ$.
4. Justifier que la matrice $C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ n'est pas productive.

Problème 2 : Algorithme de la descente de gradient

Notations

On note $\mathcal{C}(\mathbb{R})$ l'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$, et $\mathcal{C}^1(\mathbb{R})$ l'ensemble de celles qui sont dérivables sur $\mathbb{R}$ à dérivée continue.
On dit qu'une fonction $g : \mathbb{R} \to \mathbb{R}$ est L-Lipschitzienne, où $L > 0$ est un nombre réel, si elle satisfait

$$\forall x, y \in \mathbb{R}, |g(x) - g(y)| \leqslant L|x - y|.$$

On dit qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est convexe si elle satisfait

$$\forall x, y \in \mathbb{R}, \forall t \in [0, 1], f((1-t)x + ty) \leqslant (1-t)f(x) + tf(y). \tag{1}$$

On rappelle qu'une fonction $f \in \mathcal{C}^1(\mathbb{R})$ est convexe si et seulement si sa dérivée f' est croissante sur $\mathbb{R}$.
Dans les parties I, II et III du sujet, la suite $(x_n)_{n\in\mathbb{N}}$ est toujours définie par la relation de récurrence

$$\forall n \in \mathbb{N}, x_{n+1} := x_n - \tau f'(x_n), \tag{2}$$

déterminée par un terme initial $x_0 \in \mathbb{R}$, un pas de temps $\tau > 0$, et une fonction $f \in \mathcal{C}^1(\mathbb{R})$.

Objectif de l'énoncé

L'objet de cette composition est d'établir certaines propriétés de l'algorithme de la descente de gradient et de ses variantes. Lorsque la fonction f possède de fortes propriétés (convexité, régularité,...), la suite $(x_n)_{n\in\mathbb{N}}$ des itérées converge rapidement vers un minimiseur de f. Sous des hypothèses plus faibles, on peut parfois obtenir une convergence plus lente, ou bien avoir recours à une variante telle que la descente de gradient implicite. Comprendre le comportement fin de l'algorithme lorsque f a des propriétés plus faibles (par exemple f non-convexe) fait l'objet de recherches actuelles, et du champ d'investigation mathématique dit de l'optimisation numérique, dont les applications sont nombreuses (ingénierie, intelligence artificielle, etc).

A l'exception des préliminaires, toutes les parties sont indépendantes. Ne pas hésiter à admettre le résultat d'une question pour répondre aux suivantes.

Partie I : Préliminaires

Dans cette partie préliminaire, on établit d'abord l'existence d'un minimiseur, sous des hypothèses adéquates, puis une première propriété des itérées $(x_n)_{n\in\mathbb{N}}$ de la descente de gradient.

1. Soit $f \in \mathcal{C}(\mathbb{R})$ telle que

$$\lim_{x\to-\infty} f(x) = +\infty \quad \text{et} \quad \lim_{x\to+\infty} f(x) = +\infty. \tag{3}$$

   Montrer qu'il existe $x_* \in \mathbb{R}$ tel que $f(x_*) = \min\{f(x) | x \in \mathbb{R}\}$. On dit que x_* est un minimiseur de f.

2. On suppose dans cette question que $f \in \mathcal{C}^1(\mathbb{R})$ est convexe, et que f' est L-Lipschitzienne, pour un certain $L > 0$.
   (a) Montrer que pour tous $x, y \in \mathbb{R}$

$$|f'(x) - f'(y)|^2 \leqslant L(x-y)(f'(x) - f'(y))$$

   (b) Soient $x, y \in \mathbb{R}$ et soient $\tilde{x} := x - \tau f'(x)$ et $\tilde{y} := y - \tau f'(y)$. Montrer que

$$|\tilde{x} - \tilde{y}|^2 \leqslant |x-y|^2 - \tau(2 - \tau L)(x-y)(f'(x) - f'(y)).$$

(c) On suppose de plus que f admet un minimiseur x_*, et que $0 < \tau \leqslant 2/L$. Montrer que la suite $(|x_n - x_*|)_{n\in\mathbb{N}}$ est décroissante. (On rappelle que $(x_n)_{n\in\mathbb{N}}$ satisfait (2).)

Partie II : Convergence rapide, sous des hypothèses fortes

Dans cette partie, on montre que la suite $(x_n)_{n\in\mathbb{N}}$ définie par l'algorithme de descente de gradient converge rapidement vers un minimiseur de f, on parle de convergence *géométrique*, en faisant des hypothèses fortes sur cette fonction. Commençons par l'étude d'un exemple.

3. Dans cette question seulement, on pose $f(x) := \frac{1}{2}Lx^2$ pour tout $x \in \mathbb{R}$, où $L > 0$ est fixé.
   (a) Montrer que $x_{n+1} = (1 - \tau L)x_n$, puis exprimer directement x_n en fonction de x_0 et n.
   (b) On suppose $x_0 \neq 0$. Justifier que $x_n \to 0$ si et seulement si $0 < \tau < 2/L$.

Hypothèses : Dans la suite, on se donne $f \in \mathcal{C}^1(\mathbb{R})$ telle que f' est L-Lipschitzienne, avec $L > 0$, et on fixe τ tel que $0 < \tau \leqslant 2/L$. On suppose de plus que f est α-convexe, avec $\alpha > 0$, c'est à dire que

$$g(x) := f(x) - \frac{1}{2}\alpha x^2, \quad \text{est une fonction convexe sur}\, \mathbb{R}. \tag{4}$$

4. Justifier que $f'(x) - \alpha x$ est une fonction croissante de $x \in \mathbb{R}$. En déduire que $\alpha \leqslant L$.
5. Montrer que $f(x) \geqslant f(0) + f'(0)x + \alpha x^2/2$ pour tout $x \in \mathbb{R}$. En déduire que f admet un minimiseur sur $\mathbb{R}$.

On note $x_* \in \mathbb{R}$ un minimiseur de f, dont l'existence vient d'être établie. Les hypothèses faites permettent d'établir que les itérées $(x_n)_{n\in\mathbb{N}}$ de la descente de gradient s'en rapprochent.

6. Montrer que pour tous $x, y \in \mathbb{R}$
$$\alpha|x-y|^2 \leqslant (f'(x) - f'(y))(x-y).$$
7. En déduire que pour tous $x, y \in \mathbb{R}$, en notant $\tilde{x} := x - \tau f'(x)$ et $\tilde{y} := y - \tau f'(y)$, on a
$$|\tilde{x} - \tilde{y}|^2 \leqslant |x-y|^2(1 - \alpha\tau(2 - L\tau)).$$
8. On suppose $0 < \tau < 2/L$. Montrer que $|x_n - x_*| \leqslant \rho^n |x_0 - x_*|$, où ρ est une constante que l'on précisera, et telle que $0 \leqslant \rho < 1$. En déduire que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x_*.

Partie III : Convergence lente, sous des hypothèses faibles

Dans cette partie, on se passe de l'hypothèse très forte (4) utilisée précédemment. On montre que l'algorithme du gradient converge en valeur, c'est à dire que la suite $(f(x_n))_{n\in\mathbb{N}}$ tend vers le minimum $f(x_*)$ de la fonction f. Commençons de nouveau par l'étude d'un exemple :

$$f(x) := \frac{1}{3}x^3 \,\text{si}\, x \geqslant 0, \quad f(x) := 0 \,\text{si}\, x < 0.$$

9. Justifier que $f \in \mathcal{C}^1(\mathbb{R})$ et que f est convexe. Donner l'ensemble de ses minimiseurs.
10. On suppose dans cette question que $0 < x_0 < 1/\tau$.
    (a) Justifier que la suite $(x_n)_{n\in\mathbb{N}}$, définie par la relation de récurrence (2), est décroissante, à valeurs strictement positives, et satisfait $x_{n+1} = x_n(1 - \tau x_n)$ pour tout $n \in \mathbb{N}$.
    (b) Justifier que $x_n \to 0$ lorsque $n \to \infty$.
    (c) Montrer que $1/x_{n+1} = 1/x_n + \tau/(1 - \tau x_n)$ pour tout $n \in \mathbb{N}$.
    En déduire que $x_n \leqslant x_0/(1 + n\tau x_0)$.

11. On suppose seulement $\tau > 0$. Montrer que pour tout $x_0 \in \mathbb{R}$, la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un minimiseur de f.

Hypothèses : On se donne dans la suite $f \in \mathcal{C}^1(\mathbb{R})$. On suppose que f est convexe, admet un minimiseur $x_* \in \mathbb{R}$, et que f' est L-Lipschitzienne. On suppose également que $0 < \tau < 2/L$.

12. (a) Justifier sans démonstration que pour tous $x, y \in \mathbb{R}$

$$f(y) \geqslant f(x) + f'(x)(y-x).$$

(b) Montrer que pour tous $x, y \in \mathbb{R}$

$$f(y) \leqslant f(x) + f'(x)(y-x) + \frac{L}{2}(y-x)^2.$$

(c) Etablir que pour tout $n \in \mathbb{N}$

$$f(x_{n+1}) \leqslant f(x_n) - \frac{\tau}{2}(2-\tau L)|f'(x_n)|^2.$$

En déduire que la suite $(f(x_n))_{n\in\mathbb{N}}$ est décroissante.

13. Montrer que $0 \leqslant f(x) - f(x_*) \leqslant |x - x_*||f'(x)|$ pour tout $x \in \mathbb{R}$.
14. Montrer que pour tout $n \in \mathbb{N}$, en supposant $x_0 \neq x_*$,

$$f(x_{n+1}) \leqslant f(x_n) - \frac{\tau}{2}(2-\tau L)\frac{|f(x_n) - f(x_*)|^2}{|x_0 - x_*|^2}.$$

Indication : utiliser 2.(c)

15. En conclure que quel que soit le choix de $x_0 \in \mathbb{R}$, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(x_*)$.