Corrigé succinct X-ens PSI 2012 par Christophe Hénocq

Préambule

1. l'hypothèse $\lim_{\|x\| \longrightarrow +\infty} f(x) = +\infty$ peut être traduite par

$$\forall A > 0 \ \exists B > 0 \ \forall x \in \mathbf{R}^n \ \ \|x\| > B \Longrightarrow f(x) > A$$

En prenant $A = |f(0)\| + 1$, il existe un réel $M > 0$ tel que $\forall x \in \mathbf{R}^n \ \ \|x\| > M \Longrightarrow f(x) > |f(0)\| + 1$.

2. La boule fermée : $B = \{x \in \mathbf{R}^n \ \ \|x\| \leq M\}$ est une partie fermée et bornée de $\mathbf{R}^n$ donc compacte. La restriction de f à B est continue. Elle atteint un minimum en x^*. Soit $x \in \mathbf{R}^n$. Si $x \in B$ on a $f(x^*) \leq f(x)$. Si $x \notin B$ alors $f(x^*) \leq f(0) < |f(0)| + 1 < f(x)$. Donc f atteint un minimum global en x^*.

3. On sait qu'une fonction de classe $\mathcal{C}^1$ sur un ouvert admettant un extremum local en un point x_0 vérifie $\nabla f(x_0) = 0$. On a donc ici $\nabla f(x^*) = 0$.

Partie I

4. Soit $x \in \mathbf{R}^3$. L'inégalité de Cauchy-Schwarz donne $(b, x) \leq |(b, x)| \leq \|b\|\|x\|$.
On en déduit $g(x) = \dfrac{1}{2}(Ax, x) - (b, x) \geq \dfrac{1}{2}C\|x\|^2 - \|b\|\|x\| = \|x\|(\dfrac{1}{2}C\|x\| - \|b\|)$. Donc $\lim_{\|x\| \longrightarrow +\infty} g(x) = +\infty$.

5. L'application $x \longmapsto Ax$ est linéaire. Donc sa différentielle en un point x est $h \longmapsto Ah$. De même la différentielle de $x \longmapsto (b, x)$ est $h \longmapsto (b, h)$. On en déduit que g est de classe $\mathcal{C}^1$ sur $\mathbf{R}^3$ et que sa différentielle en un point $x \in \mathbf{R}^3$ est définie par :

$$\mathrm{d}g(x).h = \frac{1}{2}(Ah, x) + \frac{1}{2}(Ax, h) - (b, h)$$

Comme A est symétrique on a $(Ah, x) = (Ax, h)$. Donc $\mathrm{d}g(x).h = (Ax - b, h) = (\nabla g(x), h)$ pour tout $h \in \mathbf{R}^3$. On peut donc conclure $\nabla g(x) = Ax - b$.
L'existence de ce minimum a été vue dans le préambule. Pour l'unicité on montre que A est inversible. Si x est un vecteur de $\mathbf{R}^3$ tel que $Ax = 0$, on a $0 = (Ax, x) \geq C\|x\|^2$. Donc $\|x\| = 0$ et $x = 0$. Il existe donc exactement un vecteur x^* de $\mathbf{R}^3$ vérifiant $Ax^* - b = 0$, c'est-à-dire $\nabla g(x^*) = 0$.

6. On a $u_{k+1} - x^* = u_k - \alpha \nabla g(u_k) - x^* = u_k - \alpha A u_k - \alpha b - x^* = (I_3 - \alpha A)(u_k - x^*)$.

7. Considérons la suite $(v_k)_{k \in \mathbf{N}}$ définie par $v_k = u_k - x^*$ et la matrice $B = I_3 - A$. On a, pour tout k, $v_{k+1} = Bv_k$. Donc, par une récurrence immédiate $v_k = B^k v_0$. Montrons que la limite de B^k quand $k \longrightarrow +\infty$ est la matrice nulle.
Commençons par prouver que les valeurs propres de A appartiennent au segment $[C, L]$. Si x est un vecteur propre de A associé à une valeur propre λ, on a $Ax = \lambda x$ donc $(Ax, x) = \lambda \|x\|^2$. Ainsi $C\|x\|^2 \leq \lambda \|x\|^2$, et, comme $x \neq 0$, $C \leq \lambda$.
A est symétrique donc il existe une matrice orthogonale P telle que ${}^tP\,A\,P = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$. On a ensuite ${}^tP\,B\,P = \begin{pmatrix} 1 - \alpha\lambda_1 & 0 & 0 \\ 0 & 1 - \alpha\lambda_2 & 0 \\ 0 & 0 & 1 - \alpha\lambda_3 \end{pmatrix}$ et $B^k = P \begin{pmatrix} (1 - \alpha\lambda_1)^k & 0 & 0 \\ 0 & (1 - \alpha\lambda_2)^k & 0 \\ 0 & 0 & (1 - \alpha\lambda_3)^k \end{pmatrix} {}^tP$. On a pour $i \in \{1, 2, 3\}$, $1 - \alpha L \leq 1 - \alpha\lambda_i \leq 1 - \alpha C$, donc $1 - \alpha\lambda_i \in]-1, 1[$. l'application $M \longmapsto PM{}^tP$ est linéaire donc continue. On en déduit que $B^k \longrightarrow 0$ puis $u_k \longrightarrow x^*$.

Partie II

8. On a $h(x_{k+1}) = x_k^2 + 2\epsilon_k x_k t_k + \epsilon_k^2 t_k^2$. Posons $\delta_k = 2\epsilon_k x_k t_k + \epsilon_k^2 t_k^2$. Distinguons trois cas :
$\triangleright$ si $x_k = 0$ alors $\delta_k = 0$
$\triangleright$ si $x_k > 0$ alors $\epsilon_k = -1$ et $\delta_k = -t_k(2x_k - t_k) < 0$ car $t_k \in]0, 2x_k[$.
$\triangleright$ si $x_k < 0$ alors $\epsilon_k = 1$ et $\delta_k = t_k(2x_k + t_k) < 0$ car $t_k \in]0, -2x_k[$.
On a donc dans tous les cas $h(x_{k+1}) \leq h(x_k)$.

9. On a pour tout k, $v_k - v_{k+1} = \dfrac{1}{2^{k+1}}$. En additionnant on obtient $\displaystyle\sum_{k=0}^{n}(v_k - v_{k+1}) = \frac{1}{2}\sum_{k=0}^{n}\frac{1}{2^k}$. On en déduit $v_0 - v_{n+1} = 1 - \dfrac{1}{2^{n+1}}$. Finalement pour tout $n \in \mathbf{N}$, $v_n = 1 + \dfrac{1}{2^n}$. Ainsi, pour tout k, $v_k > 1$. On choisit $\epsilon_k = -1$ et $t_k = \dfrac{1}{2^{k+1}} \in]0, 2v_k[$. Il s'agit bien d'une descente de gradient pour la fonction h. Elle converge vers 1 qui n'est pas le minimum global de h puisque celui-ci est 0.

10. On procède de la même façon : $w_k - w_{k+1} = (-1)^k \left(2 + \dfrac{3}{2^{k+1}}\right)$. En additionnant lorsque k décrit $[\![0, n]\!]$, on trouve $2 - w_{n+1} = 1 - (-1)^{n+1} + 1 - \left(-\frac{1}{2}\right)^{n+1}$. Donc $w_{n+1} = (-1)^{n+1} + \left(-\frac{1}{2}\right)^{n+1}$. Finalement pour tout $k \in \mathbf{N}$, on a $w_k = (-1)^k \left(1 + \dfrac{1}{2^k}\right)$. La suite (w_k) diverge car $w_k \sim (-1)^k$. De plus w_k a même signe que $(-1)^k$. On a bien $t_k = (-1)^{k+1}$. Il reste à prouver que $2 + \dfrac{3}{2^{k+1}} < 2|w_k| = 2\left(1 + \dfrac{1}{2^k}\right)$, ce qui est vrai car $\dfrac{3}{2} < 2$. Il s'agit donc bien d'une descente de gradient pour h.

Partie III

11. Si $\nabla f(x) \neq 0$, on prend $d = -\dfrac{\nabla f(x)}{\|\nabla f(x)\|}$, donc D_x est non vide. Toujours dans l'hypothèse $\nabla f(x) \neq 0$, si $d \in D_x$, on écrit le développement limité de f en x à l'ordre 1 :

$$f(x+h) = f(x) + (\nabla f(x), h) + \|h\|\varepsilon(h) \quad \text{avec} \quad \lim_{\|h\| \longrightarrow 0} \varepsilon(h) = 0$$

En remplaçant h par td on a $f(x+td) - f(x) = t\Big((\nabla f(x), d) + \|d\|\varepsilon(td)\Big)$. Il existe $\beta > 0$ tel que pour tout $t \in]0, \beta[$ on a $(\nabla f(x), d) + \|d\|\varepsilon(td) < 0$, d'où $f(x+td) < f(x)$. Donc $T_{d,x}$ est non vide.

12. Dans le cas de la fonction particulière h, on a $x^* = 0$. Si $x_k = 0$, on a $d_k = t_k = 0$. Si $x_k \neq 0$, (d_k, t_k) est tel que $x_k d_k < 0$ et $(x_k + d_k t_k)^2 < x_k^2$. On a donc $x_k d_k < 0$ et $2x_k d_k t_k + d_k^2 t_k^2 < 0$. On en déduit que $d_k t_k$ est de même signe que $-x_k$ et que $0 < |d_k t_k| < 2|x_k|$. D'où le résultat demandé.

13. Si f est coercitive, on a pour tout k $f(x_{k+1}) = f(x_k + t_k d_k) \leq f(x_k)$. La suite $(f(x_k))$ est donc décroissante et et minorée par $f(x^*)$. Elle est donc convergente. La suite (x_k) est bornée sinon on pourrait trouver des $f(x_k)$ aussi grands que l'on veut.

14. On calcule $g(u_{k+1}) - g(u_k)$. On sait que $u_{k+1} = u_k - \alpha r_k$.
On a $g(u_{k+1}) = \dfrac{1}{2}(Au_k - \alpha A r_k, u_k - \alpha r_k) - (b, u_k) + \alpha(b, r_k) = g(u_k) - \dfrac{\alpha}{2}(Ar_k, u_k) - \dfrac{\alpha}{2}(r_k, Au_k) + \dfrac{\alpha^2}{2}(Ar_k, r_k)$.
Or A est symétrique donc $(Ar_k, u_k) = ((Au - k; r_k)$. De plus $r_k = Au_k - b$.
Ainsi $g(u_{k+1}) - g(u_k) = \dfrac{\alpha^2}{2}(Ar_k, r_k) - \alpha\|r_k\|^2$.
$\triangleright$ Si $u_k = x^*$, alors $r_r = \nabla g(u_k) = 0$ et $u_{k+1} = u_k$.
$\triangleright$ Si $u_k \neq x^*$, alors $r_r = \nabla g(u_k) \neq 0$ et $u_{k+1} = u_k - \alpha r_k$. On prend $d_k = -\dfrac{r_k}{\|r_k\|}$ et $t_k = \alpha\|r_k\|$. On a $d_k \in D_{u_k}$, $u_{k+1} = u_k + t_k d_k$. On doit vérifier que $t_k \in T_{d_k, u_k}$ c'est-à-dire $g(u_{k+1}) - g(u_k) < 0$ ou encore

$\frac{\alpha^2}{2}(Ar_k, r_k) - \alpha\|r_k\|^2 < 0$. Pour cela on diagonalise A dans une base orthonormale de $\mathbf{R}^3$. En notant x_1, x_2, x_3 les coordonnées de r_k dans cette base, $g(u_{k+1}) - g(u_k) = \frac{\alpha^2}{2}(\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2) - \alpha(x_1^2 + x_2^2 + x_3^2)$. On trouve $g(u_{k+1}) - g(u_k) = \alpha\Big(\sum_{i=1}^{3}(\frac{\alpha\lambda_i}{2} - 1)x_i^2\Big) < 0$ dès que $\alpha \in]0, \frac{2}{L}[$ (avec les notations de la question 7, on retrouve le même intervalle).

Partie III

15. On a, d'après (2), $0 \geq t_k(d_k, \nabla f(x_k)) \geq \frac{1}{m_1}(f(x_{k+1}) - f(x_k))$ Or la suite $(f(x_k))$ est décroissante et minorée par $f(x^*)$. Elle est donc convergente. La série de terme général $f(x_{k+1}) - f(x_k)$ est donc aussi convergente. On en déduit que $\sum t_k(d_k, \nabla f(x_k))$ converge. Son terme général a pour limite 0 quand $k \longrightarrow +\infty$

16. On distingue deux cas :
▷ Si $C_1 = \min(C_1, C_2|(d_k, \nabla f(x_k))|)$, alors $|(d_k, \nabla f(x_k))| = \frac{1}{t_k}|t_k(d_k, \nabla f(x_k))| \leq \frac{1}{C_1}|t_k(d_k, \nabla f(x_k))| \longrightarrow 0$.
▷ Si $C_2|(d_k, \nabla f(x_k))| = \min(C_1, C_2|(d_k, \nabla f(x_k))|)$, on a $t_k \geq C_2|(d_k, \nabla f(x_k))|$. Donc $|(d_k, \nabla f(x_k))| \leq \frac{1}{C_2}t_k$. On en déduit $(d_k, \nabla f(x_k))^2 \leq \frac{1}{C_2}t_k|(d_k, \nabla f(x_k))| \longrightarrow 0$. En considérant la racine carrée, on peut dire que $|(d_k, \nabla f(x_k))|$ est toujours majoré par le terme général d'une suite convergeant vers 0.

17. Posons $r_k = \nabla f(x_k)$. On sait que $\frac{(Br_k, r_k)}{\|Br_k\|} \longrightarrow 0$ quand $k \longrightarrow +\infty$. Diagonalisons B dans une base orthonormale. En notant $x_1, ..., x_n$ les coordonnées de r_k dans cette base, on a $\frac{(Br_k, r_k)}{\|Br_k\|} = \frac{\sum_{i=1}^{n}\lambda_i x_i^2}{\sqrt{\sum_{i=1}^{n}\lambda_i^2 x_i^2}}$ les λ_i étant les valeurs propres de B. Posons $m = \min(\lambda_i)$ et $M = \max(\lambda_i)$.
On a $\frac{(Br_k, r_k)}{\|Br_k\|} \geq \frac{m\|r_k\|^2}{M\|r_k\|} = \frac{m}{M}\|r_k\|$. On en déduite que $r_k \longrightarrow 0$.

18. L'existence d'un minimum est acquise depuis le préambule. L'unicité résulte de la stricte convexité. En effet si f atteint son minimum en deux points x_1 et x_2, elle est constante sur l'intervalle $[x_1, x_2]$. Cela contredit la stricte convexité.

19. Supposons construite une telle suite, on sait que (x_k) est bornée. Soit $(x_{\phi(k)})$une suite extraite convergeant vers l. On a $\nabla f(x_{\phi(k)}) \longrightarrow 0$ c'est-à-dire $f'(x_{\phi(k)}) \longrightarrow 0$. Donc $f'(l) = 0$. On en déduit $l = x^*$ car f' est strictement croissante et ne s'annule qu'en x^*. D'après les résultats admis $x_k \longrightarrow 0$.