# Éléments de correction pour les premières questions théoriques du TP/TD 2

## 1. Retour en dimension un

**(a-TD)** Sachant qu'ici le gradient en $x_k$ est $f'(x_k)$, les deux méthodes peuvent être vues sous la forme $x_{k+1} = x_k - \alpha_k f'(x_k)$. Pour la méthode de Newton : $x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$ on a $\alpha_k = \frac{1}{f''(x)}$, et pour la méthode de la sécante $x_{k+1} = x_k - \frac{(x_{k-1}-x_k)f'(x_k)}{f'(x_{k-1})-f'(x_k)}$ donc $\alpha_k = \frac{x_{k-1}-x_k}{f'(x_{k-1})-f'(x_k)}$ (dans ce cas-là, il faut avoir deux points de départ $x_0 \neq x_1$ pour que cela fonctionne).

Pour que la méthode soit vraiment une direction de descente, il faut que le pas soit positif, et que l'on ait $f(x_k + \alpha_k f'(x_k)) < f(x_k)$.

On peut essayer d'aller plus loin. Pour que le pas soit positif, une condition suffisante est que la fonction soit (un peu plus que) strictement convexe : si $f''(x) > 0$ pour tout $x$, alors le pas de la méthode de Newton est strictement positif. Si $f'$ est strictement croissante, alors le pas de la méthode de la sécante est strictement positif (dans le cas où on a toujours $x_k \neq x_{k-1}$, ce que l'on peut supposer vrai, car si à une étape on a $x_k = x_{k-1}$, c'est qu'on a eu $f'(x_{k-1}) = 0$, et dans ce cas on arrête l'algorithme, puisqu'on a trouvé un point critique et qu'on ne peut pas calculer le point suivant, le dénominateur étant nul).

Pour que la deuxième condition soit satisfaite, il existe des conditions suffisantes, par exemple que $x_k$ soit suffisamment proche d'un minimum $x_*$ où $f''(x_*) > 0$, mais c'est plus compliqué à montrer (on ne s'attend pas à cette réponse).

**(b-TD)** On a $x_{k+1} = x_k - 2\alpha x_k + \alpha x_k^3 = x_k(1-2\alpha+\alpha x_k^2)$. Pour $\alpha \in ]0,1[$, on a si $x_0^2$ est suffisamment petit avec $\rho = |1-2\alpha+\alpha x_0^2| < 1$, que $|x_k|$ est décroissante (donc $f(x_k)$ aussi) et que $|x_{k+1}| \leqslant \rho|x_k|$ donc la suite converge au moins linéairement vers 0.

Le taux de convergence linéaire est donné par $\lim_{k\to\infty} \frac{|x_{k+1}-0|}{|x_k-0|} = \lim_{k\to\infty} |1-2\alpha+\alpha x_k^2| = |1-2\alpha|$.

La convergence est donc superlinéaire quand ce taux est nul, c'est à dire pour $\alpha = \frac{1}{2}$. Essayons d'obtenir son ordre : on a dans ce cas-là $x_{k+1} = \frac{1}{2}x_k^3$, donc on obtient que $\lim_{k\to\infty} \frac{|x_{k+1}-0|}{|x_k-0|^3} = \frac{1}{2}$. Comme la suite converge vers 0, on dit que l'ordre de convergence est 3. Bien remarquer que le fait que $\frac{1}{2} < 1$ n'est aucunement important, on aurait pu avoir 12 ça aurait très bien allé.

Pour le cas $\alpha = 1$ c'est un peu plus subtil : on a $x_{k+1} = -x_k(1-x_k^2)$ de sorte que $|x_k|$ est décroissante dès que $|x_0^2| < 2$. Elle converge donc vers une limite $\ell$. En passant à la limite dans la relation de récurrence, on obtient que $\ell = \ell|1-\ell^2|$. Comme $\ell \geqslant 0$ on obtient que $\ell \in \{0, \sqrt{2}\}$. Comme on a démarré avec $|x_0| < \sqrt{2}$, on est assuré de la convergence vers 0.

Par contre dans ce cas là, on a $\lim_{k\to\infty} \frac{|x_{k+1}|}{|x_k|} = 1$, donc la convergence n'est pas linéaire.

Enfin dès que $\alpha > 1$, on obtient $|x_{k+1}| > |x_k|$ dès que $|x_k|$ est assez petit (par exemple dès que $|x_k| \leqslant \varepsilon$). Donc, à moins d'être constante égale à 0 à partir d'un certain rang, la suite $x_k$ ne converge pas vers 0 : par l'absurde, il existerait un $N_0$ tel que pour tout $k \geqslant N_0$, on ait $|x_k| \leqslant \varepsilon$, ce qui donnerait que la suite $|x_k|$ serait croissante à partir du rang $N_0$.

On peut montrer (mais ce n'est pas demandé) que le cas où la suite stationne à zéro n'arrive presque jamais : si $k_0$ est le premier endroit où $x_k = 0$, alors $x_{k-1} = \pm\sqrt{2-\frac{1}{\alpha}}$, puis $x_{k-2}$ n'a qu'au plus trois solutions (on résout une équation de degré 3), ainsi de suite, si on remontait en arrière, on n'a qu'un nombre fini de solutions pour $x_0$, et ce pour chaque $k_0$. L'ensemble des points à partir duquel la suite peut stationner à 0 est donc dénombrable, et donc de mesure nulle.

**(e*-TD)** En supposant que $0 < |x_0| < \sqrt{2}$, la suite $(x_k^2)_{k\in\mathbb{N}}$ est décroissante d'après la question **(b)** et on a

$$\frac{1}{x_{n+1}^2} - \frac{1}{x_n^2} = \frac{1}{x_n^2}\left(\frac{1}{(1-x_n^2)^2} - 1\right) = \frac{1}{x_n^2}\frac{2x_n^2 - x_n^4}{(1-x_n^2)^2} = \frac{2-x_n^2}{(1-x_n^2)^2} \to 2.$$

En faisant la somme de $n = 0$ à $N-1$, les termes se téléscopent, et en divisant par $N$, on obtient

$$\frac{1}{Nx_N^2} = \frac{1}{N}\left(\frac{1}{x_0^2} + \sum_{n=0}^{N} a_n\right),$$

où $a_n = \frac{2-x_n^2}{(1-x_n^2)^2} \to 2$. On peut donc montrer (théorème des moyennes de Cesáro) que $Nx_N^2 \to 2$, autrement dit que $x_N \sim \frac{1}{\sqrt{2N}}$, et ce quelque soit $x_0 \in ]-\sqrt{2}, \sqrt{2}[\setminus\{0\}$.

**(f-TD)** On a $f(x) = 1 - \frac{1}{1+x^2}$. Donc $x_{k+1} = x_k - \alpha\frac{2x_k}{(1+x_k^2)^2} = x_k(1 - \frac{2\alpha}{(1+x^2)^2})$.

Si $\alpha \in ]0,1]$ et $x_0 \neq 0$, alors la suite $(|x_k|)_{k\in\mathbb{N}}$ est décroissante, et converge donc vers une limite $\ell$ telle que $\ell = \ell(1 - \frac{2\alpha}{(1+\ell^2)^2})$ donc $\ell = 0$ ou (si $\alpha > \frac{1}{2}$). Si on prend $x_0$ suffisamment petit, la suite $x_k$ converge donc vers 0.

Si $\alpha > 1$, lorsque $x_k \neq 0$ pour tout $k$, alors la suite ne peut pas converger vers 0 pour les mêmes raisons que dans la question **(b)**.

Encore une fois, pour $k_0$ fixé, on peut montrer (mais ce n'est pas demandé) que pour une suite telle que $x_{k_0} = 0$, il n'y a qu'un nombre fini de cas possibles pour $x_0$.

## 2. Méthode de gradient à pas fixe, cas test en dimension 2.

**(b-TD)** La méthode par différences finies à droite permet de faire moins d'évaluations de fonctions : $n+1$ au lieu de $2n$. La méthode des différences finies centrées par contre est plus précise : on peut obtenir une erreur de l'ordre de $\eta^{\frac{2}{3}}$ au lieu de $\sqrt{\eta}$ pour les différences finies à droite (cf. TD-TP-1). En pratique on ne cherchera pas à obtenir des localisations pour le minimum de la fonction avec une précision aussi forte : si on est proche d'un minimum à $\sqrt{\eta}$ près, alors la norme du gradient est de l'ordre de $\eta$ (Taylor à l'ordre 1 pour le gradient), bien plus précis que la précision qu'on peut avoir sur la dérivée.

**(d-TD)** On fait un développement limité : $f_a(0) = ax_0^2 + x_1^2 + o(\|x\|^2)$. Ça nous donne que le gradient est nul et que la Hessienne est $\begin{pmatrix} 2a & 0 \\ 0 & 2 \end{pmatrix}$. Ses valeurs propres sont donc 2 et $2a$. On sait que le pas fixe qui convient le mieux (si on démarre suffisamment proche de zéro, la Hessienne est proche de celle en zéro) est de l'ordre de $\frac{2}{2+2a} = \frac{1}{1+a}$, et en tout cas il doit être strictement plus petit que $\frac{1}{\max\{1,a\}}$.