Les différentes parties du sujet sont indépendantes.

Soient m, $n \in \mathbb{N}^*$ et $f : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}^m$ une fonction de classe C^1 définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$. On note $\nabla f(x) = \left(\dfrac{\partial f_i}{\partial x_j}(x)\right)_{1\leq i\leq m, 1\leq j\leq n} \in \mathbb{R}^{m\times n}$ son gradient au point $x \in \mathcal{O}$.
Lorsque $m = n$, on peut déterminer numériquement les zéros de f à l'aide de la méthode de Newton. Cette méthode est fondée sur l'étude de la suite

$$x_{k+1} = x_k - \nabla f(x_k)^{-1} f(x_k), \text{ pour tout } k \in \mathbb{N}$$

où $x_0 \in \mathbb{R}^n$ est donné. Ici $\nabla f(x_k)^{-1} \in \mathbb{R}^{n\times n}$ représente, s'il existe, l'inverse de la matrice $\nabla f(x_k)$.

Partie I

On suppose ici que $m = n = 1$ et que f est analytique sur un intervalle ouvert $I \subset \mathbb{R}$. Pour tout $x \in I$ tel que $f'(x) \neq 0$, on pose

$$\alpha(x) = \sup_{k\geq 2} \left| \frac{1}{f'(x)} \frac{f^{(k)}(x)}{k!} \right|^{\frac{1}{k-1}}$$

1. Soient x, y dans I tels que $f'(x) \neq 0$, $f'(y) \neq 0$ et $v = |y - x|\alpha(x) < 1 - \dfrac{\sqrt{2}}{2}$.

   (a) Montrer que
   $$\left| \frac{f'(y)}{f'(x)} - 1 \right| \leq \sum_{k=1}^{\infty} (k+1) v^k.$$
   En déduire que
   $$\left| \frac{f'(x)}{f'(y)} \right| \leq \frac{(1-v)^2}{p(v)}$$
   avec $p(v) = 1 - 4v + 2v^2$.

   (b) Montrer que
   $$\alpha(y) \leq \frac{\alpha(x)}{(1-v)p(v)}$$
   puis que
   $$\left| \frac{f(y)}{f'(y)} \right| \leq \frac{1-v}{p(v)} \left((1-v) \left| \frac{f(x)}{f'(x)} \right| + |y - x| \right).$$

2. Soit $\tilde{x} \in I$ tel que $f'(\tilde{x}) \neq 0$. On définit $r > 0$ et $0 < \alpha_0 < 1$ et on suppose que $v_0 = r\alpha(\tilde{x}) < 1 - \dfrac{\sqrt{2}}{2}$ et $\left| \dfrac{f(\tilde{x})}{f'(\tilde{x})} \right| \alpha(\tilde{x}) \leq \alpha_0$.

(a) Montrer que, pour tout $y \in]\tilde{x} - r, \tilde{x} + r[$

$$|\phi'(y)| \leq \lambda$$

où $\lambda = 2\dfrac{(1 - v_0)\alpha_0 + v_0}{p(v_0)^2}$ et $\phi : I \mapsto \mathbb{R}$ est définie par

$$\phi(y) = y - \frac{f(y)}{f'(y)}, \forall\, y \in I \text{ tel que } f'(y) \neq 0.$$

(b) On suppose que $\lambda < 1$ et que $\alpha_0 + \lambda v_0 \leq v_0$. Montrer que dans ce cas ϕ admet un unique point fixe x_* dans $]\tilde{x} - r, \tilde{x} + r[$.

Partie II

On chercher à utiliser la méthode de Newton pour déterminer les racines d'un polynôme. Soit $P \in \mathbb{C}[X]$ un polynôme de degré n :

$$P(x) = x^n - a_1 x^{n-1} + a_2 x^{n-2} + \cdots + (-1)^j a_j x^{n-j} + \cdots + (-1)^n a_n.$$

On suppose que P admet n racines que l'on note r_j, $1 \leq j \leq n$. On peut écrire P sous la forme suivante : $\forall\, x \in \mathbb{C}$,

$$P(x) = \prod_{j=1}^{n} (x - r_j).$$

On note $r = (r_j)_{1 \leq j \leq n} \in \mathbb{C}^n$, $A = (a_j)_{1 \leq j \leq n} \in \mathbb{C}^n$ et $S(r) = (s_j(r))_{1 \leq j \leq n} \in \mathbb{C}^n$ avec les fonctions symétriques s_j définies par

$$\begin{cases} s_0(r) & = 1 \\ s_j(r) & = \displaystyle\sum_{1 \leq i_1 < \cdots < i_j \leq n} r_{i_1} \cdots r_{i_j} \text{ pour } 1 \leq j \leq n. \end{cases}$$

On note aussi $s_j(\hat{r}_k)$ la fonction symétrique relative au vecteur r privé de sa k-ième composante :

$$s_j(\hat{r}_k) = s_j(r_1, \cdots, r_{k-1}, r_{k+1}, \cdots, r_n).$$

Déterminer les racines de P revient à résoudre le système non linéaire :

$$s_j(r) = a_j,\ 1 \leq j \leq n \ \Leftrightarrow\ S(r) = A.$$

On peut appliquer la méthode de Newton à ce système et considérer la suite définie par $r^0 \in \mathbb{C}^n$ puis, pour tout $m \in \mathbb{N}$, $r^{m+1} = \Psi(r^m)$ où $\Psi(r)$ est défini pour tout $r \in \mathbb{C}^n$ tel que $\nabla S(r)$ est inversible par :

$$\Psi(r) = r - \nabla S(r)^{-1}(S(r) - A). \tag{1}$$

1. Montrer que ∇S est donné par :

$$\nabla S(r) = \begin{pmatrix} s_0(\hat{r}_1) & s_0(\hat{r}_2) & \cdots & s_0(\hat{r}_n) \\ s_1(\hat{r}_1) & s_1(\hat{r}_2) & \cdots & s_1(\hat{r}_n) \\ \vdots & \vdots & \vdots & \vdots \\ s_{n-1}(\hat{r}_1) & s_{n-1}(\hat{r}_2) & \cdots & s_{n-1}(\hat{r}_n) \end{pmatrix}.$$

2. On définit $\Delta(n,r) = \det(\nabla S(r))$.

   (a) Montrer que

   $$\Delta(n,r) = C_n \prod_{1 \leq i < j \leq n} (r_i - r_j),$$

   où C_n est un réel qui ne dépend que de n.

   (b) Montrer par récurrence que $C_n = 1$ pour tout $n \in \mathbb{N}^*$.

3. Soit r un élément de $\mathbb{C}^n$ dont les coordonnées sont distinctes deux à deux. On définit le polynôme de degré $n-1$: $\forall x \in \mathbb{C}$,

$$M_j(x) = \frac{\displaystyle\sum_{k=1}^{n} (-1)^{k-1} s_{k-1}(\hat{r}_j) x^{n-k}}{\displaystyle\prod_{k \neq j} (r_j - r_k)}.$$

   (a) Montrer que, pour tout $1 \leq i, j \leq n$, $M_j(r_i) = \delta_{ij}$ où

   $$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

   (b) Montrer que

   $$V \nabla S(r) D = I_n,$$

   où V est la matrice $n \times n$ définie par $V_{ij} = (-1)^{j-1} r_i^{n-j}$, pour tout $1 \leq i, j \leq n$ et D est une matrice diagonale à déterminer. En déduire que l'inverse de ∇S est donné par

   $$(\nabla S(r)^{-1})_{ij} = \frac{(-1)^{j-1} r_i^{n-j}}{\displaystyle\prod_{k \neq i} (r_i - r_k)}$$

4. Montrer que la fonction Ψ définie par (1) est donnée par :

   $$\Psi_i(r) = r_i - \frac{P(r_i)}{\displaystyle\prod_{k \neq i} (r_i - r_k)},$$

   pour tout $r \in \mathbb{C}^n$ dont les coordonnées sont distinctes deux à deux.

Partie III

Dans le cas où le gradient de la fonction $f : \mathcal{O} \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ pour laquelle on cherche des zéros ne peut pas être calculé, on considère ici une variante de la méthode de Newton. Cette méthode, appelée méthode de Broyden, consiste à se donner $x_0 \in \mathbb{R}^n$ et une approximation $A_0 \in \mathbb{R}^{n\times n}$ du gradient de f en x_0 et à introduire pour tout $k \in \mathbb{N}$,

$$\begin{cases} u_k = -A_k^{-1} f(x_k), \\ x_{k+1} = x_k + u_k, \\ y_k = f(x_{k+1}) - f(x_k), \\ A_{k+1} = A_k + \dfrac{(y_k - A_k u_k)u_k^t}{u_k^t u_k}. \end{cases}$$

On définit la norme matricielle $|||\cdot|||$ par : $\forall\, A = (A_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$

$$|||A||| = \mathrm{tr}(A^t A)^{1/2},$$

où tr représente la trace et est défini par $\mathrm{tr}(M) = \displaystyle\sum_{i=1}^n M_{ii}$, $\forall\, M = (M_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$.

On définit de plus la norme vectorielle $\|\cdot\|_2$ par :

$$\|y\|_2 = \left(\sum_{i=1}^n y_i^2\right)^{\frac{1}{2}},\ \forall\, y = (y_i)_{1\le i\le n} \in \mathbb{R}^n$$

et la norme matricielle $|||\cdot|||_2$ induite par $\|\cdot\|_2$ c'est-à-dire :

$$|||A|||_2 = \sup_{x\in\mathbb{R}^n/\|x\|_2\neq 0} \frac{\|Ax\|_2}{\|x\|_2}.$$

1. On définit l'espace $Q(y,u) = \{B \in \mathbb{R}^{n\times n}/Bu = y\}$. On suppose qu'on est à l'étape k et que A_k, y_k et u_k sont connus. Montrer que A_{k+1} est l'unique solution du problème : trouver $B_* \in Q(y_k, u_k)$ tel que

$$|||B_* - A_k|||^2 = \min_{B\in Q(y_k,u_k)} |||B - A_k|||^2.$$

2. On suppose que f est de classe C^1 sur $\mathcal{O} \subset \mathbb{R}^n$ un ouvert convexe et qu'il existe $\sigma > 0$ tel que

$$|||\nabla f(x) - \nabla f(y)|||_2 \le \sigma \|x - y\|_2,\ \forall\, x,\, y \in \mathcal{O}.$$

Dans cette question et dans la suite de la Partie III, on se donne un élément $x_* \in \mathcal{O}$ et on note $e_k = x_k - x_*$ pour tout $k \in \mathbb{N}$.

   (a) Montrer que pour tout x, $y \in \mathcal{O}$,

$$\|f(x) - f(y) - \nabla f(x_*)(x-y)\|_2 \le \sigma \frac{\|y - x_*\|_2 + \|x - x_*\|_2}{2} \|y - x\|_2.$$

   Pour cela, on pourra utiliser la fonction $g(t) = f(y + t(x-y))$, pour $0 \le t \le 1$.

(b) Montrer que pour tout A, B dans $\mathbb{R}^{n\times n}$,

$$|||AB||| \leq |||A|||\ |||B|||_2.$$

(c) Montrer que, pour tout $k \in \mathbb{N}$,

$$|||A_{k+1} - \nabla f(x_*)||| \leq |||(A_k - \nabla f(x_*))\left(I_n - \frac{u_k u_k^t}{u_k^t u_k}\right)||| + \frac{\sigma}{2}(\|e_{k+1}\|_2 + \|e_k\|_2).$$

En déduire que, pour tout $k \in \mathbb{N}$,

$$|||A_{k+1} - \nabla f(x_*)||| \leq |||A_k - \nabla f(x_*)||| + \frac{\sigma}{2}(\|e_{k+1}\|_2 + \|e_k\|_2).$$

3. En plus des hypothèses de la question 2, on suppose que $x_* \in \mathcal{O}$ satisfait les propriétés suivantes pour un certain $\beta > 0$:

$$f(x_*) = 0,\ \nabla f(x_*) \text{ est inversible et } |||\nabla f(x_*)^{-1}||| \leq \beta.$$

Soit $r > 0$ tel que $B(x_*, r) \subset \mathcal{O}$. On considère $\delta > 0$ et $\epsilon > 0$ tels que $6\beta\delta \leq 1$, $\epsilon < r$ et $3\sigma\epsilon \leq 2\delta$ et on suppose que

$$|||A_0 - \nabla f(x_*)||| \leq \delta \text{ et } \|x_0 - x_*\|_2 \leq \epsilon.$$

(a) Soit $i \geq 1$. On suppose que pour tout $k = 0, 1 \ldots, i-1$, $x_k \in \mathcal{O}$, A_k est inversible et :

$$|||A_k - \nabla f(x_*)||| \leq \left(2 - \frac{1}{2^k}\right)\delta \text{ et } \|e_{k+1}\|_2 \leq \frac{\|e_k\|_2}{2}. \qquad (2)$$

(i) Montrer que

$$|||A_i - \nabla f(x_*)||| \leq \left(2 - \frac{1}{2^i}\right)\delta.$$

En déduire que A_i est inversible et donner une borne sur $|||A_i^{-1}|||$.

(ii) Montrer que

$$A_i e_{i+1} = (-f(x_i) + f(x_*) + \nabla f(x_*) e_i) + (A_i - \nabla f(x_*))e_i.$$

En déduire que

$$\|e_{i+1}\|_2 \leq \frac{\|e_i\|_2}{2}.$$

Conclure que (2) est vraie pour tout $k \in \mathbb{N}$.

Dans la suite, nous allons améliorer ces résultats et montrer que :

$$\lim_{k\to+\infty} \frac{\|e_{k+1}\|_2}{\|e_k\|_2} = 0. \qquad (3)$$

(b) On suppose que

$$\lim_{k\to+\infty} \frac{\|(A_k - \nabla f(x_*))u_k\|_2}{\|u_k\|_2} = 0. \tag{4}$$

Montrer que

$$\lim_{k\to+\infty} \frac{\|f(x_{k+1})\|_2}{\|u_k\|_2} = 0.$$

Montrer qu'il existe $m > 0$ tel que, pour tout $k \in \mathbb{N}$

$$\|f(x_{k+1})\|_2 \geq m\|e_{k+1}\|_2.$$

Conclure que (3) est vérifiée.

(c) Il reste donc à montrer que l'hypothèse (4) est satisfaite. Soit $B_k = A_k - \nabla f(x_*)$. Montrer que

$$|||B_{k+1}||| \leq |||B_k||| - \frac{1}{2\,|||B_k|||}\left(\frac{\|B_k u_k\|_2}{\|u_k\|_2}\right)^2 + \frac{3\sigma}{4}\|e_k\|_2.$$

En déduire que la série $\sum \frac{\|B_k u_k\|_2^2}{\|u_k\|_2^2}$ est convergente et conclure.

Partie IV

On considère maintenant une fonction f de classe $\mathcal{C}^1$ définie de $\mathcal{O} \subset \mathbb{R}^n$ dans $\mathbb{R}^m$ avec $n \leq m$. Nous allons proposer une généralisation de la méthode de Newton dans ce cadre. On dit que $L^\dagger \in \mathbb{R}^{n\times m}$ est l'inverse généralisé d'une matrice $L \in \mathbb{R}^{m\times n}$ si

$$LL^\dagger L = L,\ L^\dagger LL^\dagger = L^\dagger,\ (LL^\dagger)^t = LL^\dagger,\ (L^\dagger L)^t = L^\dagger L.$$

Soit E un sous-espace vectoriel de $\mathbb{R}^m$, on note Π_E la projection othogonale sur E.

1. Soit $L \in \mathbb{R}^{m\times n}$ une matrice donnée. On suppose qu'il existe une matrice $M \in \mathbb{R}^{n\times m}$ telle que $LML = L$ et $(LM)^t = LM$. Montrer que $LM = \Pi_{\operatorname{Im} L}$.
Soit $b \in \mathbb{R}^m$. Montrer que

$$\inf_{x\in\mathbb{R}^n} \|Lx - b\|_2$$

est atteint pour $x = Mb$.

2. Soit $L \in \mathbb{R}^{m\times n}$ une matrice telle que $\operatorname{Ker} L = \{0\}$. Dans ce cas, l'inverse généralisé de L est donné par la formule suivante :

$$L^\dagger = (L^tL)^{-1}L^t.$$

On vérifie facilement que $L^\dagger L = I_n$ et $LL^\dagger = \Pi_{\operatorname{Im} L}$. Montrer que, pour tout $A \in \mathbb{R}^{m\times n}$,

$$DL^\dagger(A) = -L^\dagger AL^\dagger + (L^tL)^{-1}A^t\Pi_{(\operatorname{Im} L)^\perp}.$$

3. On définit la méthode de Newton de la façon suivante : $x_0 \in \mathcal{O}$ étant donné, on pose, $\forall\, k \in \mathbb{N}$, $x_{k+1} = \Phi(x_k)$ où Φ est donné par

$$\Phi(x) = x - (\nabla f(x))^{\dagger} f(x),$$

pour tout $x \in \mathcal{O}$ tel que $\nabla f(x)$ est injectif. On définit $F : \mathcal{O} \mapsto \mathbb{R}$ par :

$$F(x) = \frac{1}{2}\|f(x)\|_2^2$$

où $\|\cdot\|_2$ est la norme euclidienne sur $\mathbb{R}^m$.

(a) Montrer que

$$\Phi(x) = x \Leftrightarrow \nabla F(x) = 0.$$

On suppose à partir de maintenant que f est de classe $\mathcal{C}^2$ sur $\mathcal{O}$. Pour tout $x \in \mathcal{O}$, calculer $D^2F(x) \in \mathbb{R}^{n\times n}$ défini par $D^2F(x)_{ij} = \dfrac{\partial^2 F}{\partial x_i \partial x_j}(x)$.

(b) Soit $x \in \mathcal{O}$ tel que $\nabla f(x)$ est injectif.
Calculer $\nabla\Phi(x)$. On note $\vec{0}_{\mathbb{R}^n}$ le vecteur nul de $\mathbb{R}^n$. Montrer que si $\nabla F(x) = \vec{0}_{\mathbb{R}^n}$, alors

$$\nabla\Phi(x) = -(\nabla f(x)^t \nabla f(x))^{-1} D^2 f(x)^t f(x)$$

où $(D^2 f(x)^t f(x))_{ij} = \displaystyle\sum_{k=1}^{m} \frac{\partial^2 f_k}{\partial x_i \partial x_j}(x) f_k(x)$.

(c) Soit $x_* \in \mathcal{O}$ tel que $\nabla f(x_*)$ est injectif et tel que $\nabla F(x_*) = \vec{0}_{\mathbb{R}^n}$. Montrer que si le spectre de $\nabla\Phi(x_*)$ est inclus dans $]-1, 1[$ alors x_* est un minimum local strict de F.

Fin de l'épreuve