Programme de colle fictif n° 23. **Résumé du cours de calcul différentiel.**

Les points suivis d'une (*) sont à **savoir** et doivent être considérés comme des « questions de cours ».

Table des matières

1 Fonctions de classe C^1 sur un ouvert de $\mathbb{R}^n$	1
1.1 Graphe, lignes de niveau et applications partielles	1
1.2 Dérivées partielles et gradient	1
1.3 Développement limité à l'ordre 1 d'une fonction de classe C^1 . .	1
1.4 Différentielle	1
1.5 The « chain's rule »	1
2 Calcul différentiel à l'ordre 2	1
2.1 Définition des dérivées partielles secondes	1
2.2 Fonctions de classe C^2	1
2.3 Matrice Hessienne	1
2.4 Développement limité du second ordre (seconde lecture, pas forcément utile)	1
3 Extrema	1
3.1 Généralités	1
3.2 Condition nécessaire du premier ordre	1
3.3 Conditions du second ordre	1
4 Compléments	1
4.1 Passer des extrema locaux aux extrema globaux	1
4.2 Une condition suffisante de minimum globale	1

1 Fonctions de classe C^1 sur un ouvert de $\mathbb{R}^n$

Cadre : $\mathcal{O}$ ouvert non vide de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$

1.1 Graphe, lignes de niveau et applications partielles

• **Graphe d'une fonction de plusieurs variables.** Il s'agit de généraliser la notion de graphe d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$. Soient $A \subset \mathbb{R}^n$ et $f : A \rightarrow \mathbb{R}$. Le graphe de f , noté G_f , est l'ensemble des points $(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ tels que $f(x_1, \dots, x_n) = x_{n+1}$. On a vu quelques exemples pour $n = 2$.

• **Lignes de niveaux.** C'est un outil qui permet de mieux comprendre le graphe d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Si λ est un réel, la ligne de niveau λ de f est :

$$L_\lambda(f) = \{m \in \mathbb{R}^n \mid f(m) = \lambda\}$$

Elle peut se visualiser pour $n = 2$ sur le graphe de f comme l'intersection de G_f et de l'hyperplan plan $z = \lambda$ de $\mathbb{R}^{n+1}$ mais c'est une partie de $\mathbb{R}^n$ et non de $\mathbb{R}^{n+1}$. Lorsque f est continue, $L_\lambda(f)$ est un fermé de $\mathbb{R}^n$ comme image réciproque d'un fermé de $\mathbb{R}$.

• **Applications partielles en un point.**

Soient $f : \rightarrow \mathbb{R}$ et $m_0 = (a_1, \dots, a_n)$ un point de . Les applications partielles de f en m_0 sont les n applications à valeurs réelles :

$$\begin{aligned} f_1 : t \mapsto f_1(t) &= f(t, a_1, a_2, \dots, a_n) \\ f_2 : t \mapsto f_2(t) &= f(a_1, t, a_2, \dots, a_n) \\ &\vdots \\ f_n : t \mapsto f_n(t) &= f(a_1, \dots, a_{n-1}, t) \end{aligned}$$

Chaque f_i est définie au moins sur un intervalle ouvert de $\mathbb{R}$ contenant a_i puisque est un ouvert de $\mathbb{R}^n$.

1.2 Dérivées partielles et gradient

• **Dérivées partielles d'ordre 1 en un point.**

Soient $f : \rightarrow \mathbb{R}$ et $m_0 = (a_1, \dots, a_n)$ un point de . Soit $i \in \{1, \dots, n\}$. Si la fonction partielle f_i en m_0 est dérivable en a_i , la dérivée partielle de f par

rapport à la i -ème variable en m_0 est :

$$\partial_i f(m_0) = f'_i(a_i).$$

On note aussi $\frac{\partial f}{\partial x_i}(m_0)$ à la place de $\partial_i f(m_0)$.

• **Gradient.** Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $m_0 = (a_1, \dots, a_n)$ un point de $\mathbb{R}^n$. On suppose que f admet en m_0 des dérivées partielles d'ordre 1. Le gradient de f en m_0 est le vecteur :

$$\nabla f(m_0) = (\partial_1 f(m_0), \dots, \partial_n f(m_0)) \in \mathbb{R}^n.$$

• Lorsque f admet en tout point m de $\mathbb{R}^n$ une dérivée partielle par rapport à la i -ème variable pour tout $i \in \{1, \dots, n\}$, on dispose donc de n fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$:

$$\partial_i f = \left(\begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R} \\ m & \mapsto & \partial_i f(m) \end{array} \right).$$

Sous les mêmes conditions on a encore une fonction $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie, pour $m \in \mathbb{R}^n$, par $\nabla f(m) = (\partial_1 f(m), \dots, \partial_n f(m))$.

• (*) L'existence des dérivées partielles en m n'implique pas la continuité en m .

• **Fonctions de classe C^1 .** Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On dit que f est de classe C^1 sur $\mathcal{O}$ lorsque :

1. f admet en tout point m de $\mathcal{O}$ des dérivées partielles $\partial_i f(m)$ pour $i \in \{1, \dots, n\}$.
2. Chacune des fonctions $\partial_i f$ est continue sur $\mathcal{O}$.

• **Exemples fondamentaux.** Les applications coordonnées sont de classe C^1 .

• **Opération sur les fonctions de classe C^1 .**

THÉORÈME.

1. On note $C^1(\mathcal{O}, \mathbb{R})$ l'ensemble des applications de classes C^1 de $\mathcal{O}$ à valeurs dans $\mathbb{R}$. C'est un sous-espace vectoriel de $\mathcal{F}(\mathcal{O}, \mathbb{R})$ donc toute combinaison linéaire de deux fonctions C^1 l'est encore.
2. Le produit, le quotient (si...) de deux fonctions C^1 est encore une fonction C^1 .

3. (*) Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions de classe C^1 . Alors $\varphi \circ f$ est de classe C^1 . De plus pour tout m dans $\mathcal{O}$ on a :

$$\partial_i (\varphi \circ f)(m) = \varphi' \circ f(m) \times \partial_i f(m)$$

COROLLAIRE. Les fonctions polynomiales, les fractions rationnelles sont des fonctions de classe C^1 .

1.3 Développement limité à l'ordre 1 d'une fonction de classe C^1

THÉORÈME. Si f est une fonction de classe C^1 sur $\mathcal{O}$ alors elle admet en tout point m_0 de $\mathcal{O}$ un développement limité à l'ordre 1 c'est à dire qu'il existe une fonction $\varepsilon : \mathcal{O} \rightarrow \mathbb{R}$ nulle et continue en m_0 telle que pour tout $v \in U$:

$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v \rangle + \|v\| \varepsilon(v)$$

Autre expression du développement limité : il existe une fonction $\varepsilon_1 : U \rightarrow \mathbb{R}$ nulle et continue en 0 où $U = \{v \in \mathbb{R}^n \mid m_0 + v \in \mathcal{O}\}$ telle que pour tout $v \in U$:

$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v \rangle + \|v\| \varepsilon_1(v)$$

COROLLAIRE. Si f est de classe C^1 sur $\mathcal{O}$ alors elle est continue.

1.4 Différentielle

Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^1 . La différentielle de f en $m_0 \in \mathcal{O}$ est la forme linéaire :

$$df(m_0) = \left(\begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R} \\ v & \mapsto & \langle \nabla f(m_0), v \rangle \end{array} \right)$$

On a donc, pour tout $v = (v_1, \dots, v_n) \in \mathbb{R}^n$:

$$df(m_0) \cdot v = \sum_{i=1}^n \partial_i f(m_0) v_i.$$

La différentielle de f est alors l'application $df : \mathcal{O} \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ qui à un point m associe la différentielle de f en m .

On peut écrire $df = \sum_{i=1}^n \partial_i f dx_i$ où les dx_i sont les fonctions coordonnées sur $\mathbb{R}^n$.

1.5 The « chain's rule »

• Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$. On considère une courbe paramétrée $c : I \rightarrow \mathcal{O}$ de classe C^1 où I est un intervalle de $\mathbb{R}$.

THÉORÈME. Avec les notations ci-dessus, si f est de classe C^1 alors $g = f \circ c$ est de classe C^1 sur I avec :

$$g'(t) = \langle \nabla f(c(t)), c'(t) \rangle$$

• **Un cas particulier essentiel.** Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$. Pour m_0 dans $\mathcal{O}$ et v dans $\mathbb{R}^n$, puisque $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$, $m_0 + tv$ est encore dans $\mathcal{O}$ pour $|t|$ petit.

Il existe donc un intervalle ouvert I de $\mathbb{R}$ centré en 0 tel que l'application :

$$g_v : t \mapsto f(m_0 + tv)$$

soit correctement définie.

THÉORÈME. Avec les notations ci-dessus, si f est de classe C^1 alors g_v est de classe C^1 sur I avec :

$$g'_v(t) = \langle \nabla f(m_0 + tv), v \rangle$$

Interprétation géométrique du gradient. On a donc (notation ci-dessus), $g'_v(0) = \langle \nabla f(m_0), v \rangle$. Ainsi, $|g'_v(0)|$ est maximal lorsque v et $\nabla f(m_0)$ sont colinéaires (par CSB) : « la fonction f varie le plus en quittant m_0 dans la direction de $\nabla f(m_0)$ ».

COROLLAIRE. Lorsque Ω est un ouvert convexe de $\mathbb{R}^n$, alors f est constante sur Ω si et seulement si pour tout $m \in \Omega$ on a $\nabla f(m) = 0$.

• Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^1 . Soient Ω un ouvert non vide de $\mathbb{R}^p$. Pour tout $i \in \{1, \dots, n\}$, on considère des fonctions $g_i : \Omega \rightarrow \mathbb{R}$ de classe C^1 ainsi que $g = (g_1, \dots, g_n) : \Omega \rightarrow \mathbb{R}^n$. On suppose que $\text{Im } g \subset \mathcal{O}$.

THÉORÈME. La fonction $h = f \circ g : \Omega \rightarrow \mathbb{R}$ est de classe C^1 avec, pour tout $i \in \{1, \dots, p\}$ et $m \in \Omega$:

$$\partial_i h(m) = \sum_{j=1}^p \partial_j f(g(m)) \times \partial_j g_i(m)$$

Commentaire. Cette formule se résume très bien par la relation suivante :

$$\nabla h(m) = \nabla f(g(m)) J_g(m),$$

où $J_g(m) \in \mathcal{M}_{n,p}(\mathbb{R})$ est la matrice jacobienne de g i.e. la matrice dont les lignes sont les gradients des fonctions g_i en m .

2 Calcul différentiel à l'ordre 2

2.1 Définition des dérivées partielles secondes

• Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ qui est de classe C^1 . On a donc n fonctions de plusieurs variables :

$$\partial_i f : m \mapsto \partial_i f(m)$$

Lorsque la fonction $\partial_i f : \mathcal{O} \rightarrow \mathbb{R}$ admet en m une dérivée partielle par rapport à la j -ième variable on note :

$$\partial_{j,i}^2 f(m) = \partial_j (\partial_i f)(m)$$

On parle de dérivée partielle seconde. . .

• Lorsque f admet en tout m de $\mathcal{O}$ des dérivées partielles seconde on dispose de n^2 applications :

$$\partial_{i,j}^2 f : \mathcal{O} \rightarrow \mathbb{R}$$

2.2 Fonctions de classe C^2

• Une fonction $f : \mathcal{O} \rightarrow \mathbb{R}$ est dite de classe C^2 lorsque qu'elle admet des dérivées partielles secondes en tout point de $\mathcal{O}$ et que les fonctions dérivées partielles secondes sont continues sur $\mathcal{O}$.

• Opérations sur les fonctions de classe C^2 . On retient notamment que les fonctions polynomiales et les fractions rationnelles sont de classe C^2 .

• LEMME DE SCHWARZ. Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 . Alors pour tout i et j dans $\{1, \dots, n\}$ on a l'égalité de fonctions sur $\mathcal{O}$:

$$\partial_{i,j}^2 f = \partial_{j,i}^2 f$$

2.3 Matrice Hessienne

Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 . Soit m dans $\mathcal{O}$. La matrice hessienne de f en m est :

$$H_f(m) = [\partial_{i,j}^2 f(m)]$$

Selon le lemme de Schwarz, c'est une matrice symétrique. On notera q_m « la forme quadratique associée » à la matrice $H_f(m)$ i.e. pour tout X dans $\mathcal{M}_{n,1}(\mathbb{R})$ on a :

$$q_m(X) = {}^t X H_f(m) f(m) X$$

2.4 Développement limité du second ordre (seconde lecture, pas forcément utile)

• THÉORÈME Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 . Alors f admet un développement limité à l'ordre 2 en tout point m_0 de $\mathcal{O}$. Pour tout v tel que $m_0 + v \in \mathcal{O}$ on a :

$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v \rangle + \frac{1}{2} q_{m_0}(v) + \|v\|^2 \varepsilon(v)$$

où ε est nulle et continue en 0

Autre expression. Le programme confond le gradient avec sa colonne des coordonnées dans cet énoncé, qui dans ce cadre se re-écrit :

$$f(m_0 + v) = f(m_0) + \nabla f(m_0)^T X + \frac{1}{2} X^T H_f(m_0) X + o(\|v\|^2),$$

où X est la colonne des coordonnées de v .

3 Extrema

3.1 Généralités

- Le vocabulaire : extrema locaux ou relatifs, globaux, extrema locaux stricts. . .
- Rappels dans le cas $n = 1$.

THÉORÈME DE ROLLE A. Soient I un intervalle de $\mathbb{R}$, $f : I \rightarrow \mathbb{R}$, $a \in I$, **non borne de I** . $\boxed{\text{Si}}$ f admet en a un extremum relatif alors $f'(a) = 0$.

3.2 Condition nécessaire du premier ordre

THÉORÈME Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ une fonction de classe C^1 . Si f admet en m_0 un extremum local alors $\nabla f(m_0) = 0$.

Vocabulaire. Les points m de $\mathcal{O}$ où le gradient s'annule s'appellent **les points critiques** de f .

3.3 Conditions du second ordre

THÉORÈME. (**Condition nécessaire du second ordre.**) Soit $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. Si f admet un minimum local en m_0 alors la matrice hessienne

de f en m_0 est positive i.e. si q_{m_0} est la forme quadratique associée à $H_f(m_0)$ on a $q_{m_0}(v) \geq 0$ pour tout v dans $\mathbb{R}^n$.

THÉORÈME. (**Condition suffisante du second ordre.**) Soient $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. On suppose que :

- m_0 est un point critique de f .
- La matrice hessienne $H_f(m_0)$ appartient à $S_n^{++}(\mathbb{R})$ i.e. admet des valeurs propres strictement positives.

Alors f admet un minimum local strict en m_0 .

COROLLAIRE. Soient $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. On suppose que m_0 est un point critique de f .

1. Si les valeurs propres de $\nabla^2 f(m_0)$ sont strictement positives alors f admet en m_0 un minimum local strict.
2. Si les valeurs propres de $\nabla^2 f(m_0)$ sont strictement négative alors f admet en m_0 un maximum local strict.
3. Si $\nabla^2 f(m_0)$ admet une valeur propre strictement négative et une valeur propre strictement positive, alors f n'admet pas en m_0 d'extremum relatif.

Procédure pour la recherche d'extrema sur un ouvert.

1. On vérifie que l'on travaille sur un ouvert et on justifie du caractère C^1 ou C^2 de la fonction.
2. On cherche les points critiques en résolvant le système donné par $\nabla f(m) = 0$.
3. • On regarde en chacun des points critiques la hessienne de f . **Une** idée est de regarder ses valeurs propres. La condition nécessaire du second ordre permet d'éliminer les mauvais candidats. • On peut aussi faire une réciproque en étudiant (localement) le signe de $f(m) - f(m_0)$ ou de $f(m+h) - f(m)$.

Bien noter que tout cela ne donne qu'un résultat local ; il faut utiliser d'autres arguments pour avoir des résultats **globaux** : voir plus bas.

Le cas où $n = 2$

On calcule le déterminant de la hessienne en m_0 :

- Si ce déterminant est strictement positif, il y a un minimum local strict lorsque la trace est positive et un maximum local strict lorsque la trace est négative.
- Si ce déterminant est strictement négatif, il n'y a pas d'extremum local : on a un point selle (ou col).
- Dans les autres cas, on ne peut rien dire : on peut regarder la différence $f(m) - f(m_0)$...

4 Compléments

4.1 Passer des extrema locaux aux extrema globaux

Je vais présenter une situation modèle qui peut s'adapter à divers cas. On considère une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, disons de classe C^1 qui admet un unique point critique m_0 .

On suppose que pour tout $v \in \mathbb{R}^n$ on a $f(v) \geq \varphi(\|v\|)$ où $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}$ vérifie : $\varphi(t) \xrightarrow{t \rightarrow +\infty} +\infty$.

Notons $\alpha = f(m_0)$. Comme $\varphi(t) \xrightarrow{t \rightarrow +\infty} +\infty$, il existe $R > 0$ tel que pour tout $t \geq R$ on ait $\varphi(t) \geq 2|\alpha|$. Ainsi pour tout $v \in \mathbb{R}^n$ tel que $\|v\| \geq R$ on a :

$$f(v) \geq 2|\alpha|.$$

La restriction g de f à la boule fermée $B_f(0, R)$ est continue sur le fermé borné $B_f(0, R)$ de l'espace vectoriel $\mathbb{R}^n$ de dimension finie : elle admet donc un minimum global en un point p_0 de $B_f(0, R)$. Ainsi, pour tout $m \in B_f(0, R)$ on a :

$$f(m) = g(m) \geq g(p_0) = f(p_0).$$

De plus, pour tout m hors de la boule $B_f(0, R)$, on a :

$$f(m) \geq 2|\alpha| \geq f(m_0),$$

et comme $m_0 \in B_f(0, R)$, il vient $f(m) \geq f(m_0)$.

Ainsi f admet en m_0 un minimum global sur $\mathbb{R}^n$. Comme c'est un ouvert et que f est de classe C^1 , p_0 est un point critique de f donc $p_0 = m_0$, par unicité du point critique.

En conclusion f admet en m_0 un minimum global.

4.2 Une condition suffisante de minimum globale

Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{O} \rightarrow \mathbb{R}$ de classe C^2 . Pour m_0 dans $\mathcal{O}$ et v dans $\mathbb{R}^n$, puisque $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$, $m_0 + tv$ est encore dans $\mathcal{O}$ pour $|t|$ petit.

Il existe donc un intervalle ouvert I de $\mathbb{R}$ centré en 0 tel que l'application :

$$g_v : t \mapsto f(m_0 + tv)$$

soit correctement définie.

Alors g_v est de classe C^2 sur I et $g''(t) = X^T H_f(m_0 + tv) X$ où X est la colonne des coordonnées de v .

Démontrons ce dernier résultat. On sait que g_v est de classe C^1 avec :

$$g'_v(t) = \langle \nabla f(m_0 + tv), v \rangle = \sum_{i=1}^n v_i \underbrace{\partial_i f(m_0 + tv)}_{=h_i(t)}.$$

Chaque fonction h_i est de classe C^1 avec :

$$h'_i(t) = \langle \nabla \partial_i f(m_0 + tv), v \rangle = \sum_{j=1}^n v_j \partial_{j,i}^2 f(m_0 + tv).$$

Par somme g'_v est donc C^1 avec : $g'_v(t) = \sum_{i=1}^n \sum_{j=1}^n v_i v_j \partial_{j,i}^2 f(m_0 + tv)$.

Une intégration par parties (ou Taylor-intégral à l'ordre 1) donne :

$$g(1) = g(0) + g'(0) + \int_0^1 (1-t)g''(t) dt.$$

Cela se traduit par :

$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v \rangle + \int_0^1 (1-t) X^T H_f(m_0 + tv) X dt.$$

Ainsi, lorsque $\mathcal{O}$ est convexe, que m_0 est un point critique de f et lorsque la matrice hessienne $H_f(*)$ est dans $S_n^{++}(\mathbb{R})$ pour tout $*$ dans $\mathcal{O}$, on peut dire que f admet en un point critique m_0 un minimum globale strict.

Les semaine suivantes : de vrais jolis problèmes et exercices à faire!!!!!!!!!!!!