LIÈGE université Sciences

Année académique 2021-2022

MATH 2014 Compléments de Mathématiques

Françoise Bastin
Version 10 septembre 2021 (V1 : 070921)

Table des matières

1 Première partie **5**

1.1 Rappels concernant le calcul intégral . . . 5

1.2 Intégrales paramétriques . . . 5

1.3 Transformation de Fourier des fonctions intégrables sur $\mathbb{R}$ . . . 6

1.3.1 Définition et interprétation . . . 6

1.3.2 Exemples . . . 7

1.3.3 Premières propriétés . . . 8

1.3.4 Dérivation et transformation de Fourier . . . 9

1.3.5 Intégration et transformation de Fourier . . . 9

1.3.6 Théorème de transfert et de Fourier . . . 9

1.4 Produit de composition . . . 10

2 Seconde partie **13**

2.1 Le développement limité de Taylor . . . 13

2.2 Extrema : définitions et premières propriétés . . . 13

2.3 Compléments de calcul matriciel . . . 13

2.4 Retour aux extrema : cas des extrema dits « libres » . . . 15

2.5 Une application la régression linéaire . . . 16

2.6 Les extrema liés (extrema « sous contrainte ») . . . 16

2.7 Les extrema sur des ensembles fermés bornés . . . 17

Chapitre 1

Première partie

1.1 Rappels concernant le calcul intégral

Voir les notes du cours de Math2007 et Math0009 et aussi le début du podcast enregistré le 06/09/20 (disponible dans MyUliege)

1.2 Intégrales paramétriques

Les intégrales paramétriques (et leur dérivation) représentent un outil très important de l'analyse. Elles apparaissent notamment dans le cadre de l'*analyse de Fourier* et du *produit de convolution de fonctions*.

Dans le présent cours, nous ne considérerons ces intégrales que dans le cas de l'intégration à une variable et dans le cas d'un seul paramètre réel.

Un podcast illustre cette matière (enregistrement le 6 septembre 2020)

Comme leur nom l'indique déjà, les ≪ intégrales paramétriques ≫ sont des fonctions qui sont définies par des intégrales, comme suit

$$\lambda \ \mapsto \ \int_A f(x,\lambda)\ dx.$$

Dès le départ, il convient de fixer un cadre qui donne sens à cette définition :
• A est un sous-ensemble de $\mathbb{R}$ (le plus souvent un intervalle) ; le paramètre λ varie dans un sous-ensemble de $\mathbb{R}$ également, noté par exemple Λ, et qui sera considéré ouvert lorsqu'il sera question de dérivation
• f est une fonction définie sur le produit cartésien $A \times \Lambda$ et, pour tout $\lambda \in \Lambda$, la fonction $x \mapsto f(x,\lambda)$ est intégrable sur A.

Enonçons à présent le ≪ Théorème de dérivation des intégrales paramétriques ≫ dans le cas d'une seule dérivation.

Théorème 1.2.1 (Dérivation des intégrales paramétriques) *On reprend les notations et les hypothèses naturelles ci-dessus. En outre, on suppose que*
• *quel que soit $x \in A$, la fonction $\lambda \mapsto f(x,\lambda)$ appartient à $C_1(\Lambda)$*
• *quel que soit $\lambda \in \Lambda$, la fonction $x \mapsto f(x,\lambda)$ est intégrable sur A*
• *(*) quel que soit l'ensemble borné fermé K inclus dans l'ouvert Λ, il existe une fonction g_K intégrable sur A telle que*

$$|D_\lambda f(x,\lambda)| \le g_K(x), \quad \forall x \in A,\ \forall \lambda \in K$$

Sous ces conditions, la fonction $\lambda \mapsto \int_A f(x,\lambda)\ dx$ appartient à $C_1(\Lambda)$ et ≪ on peut dériver sous le signe intégral ≫, ce qui signifie que

$$D_\lambda \int_A f(x,\lambda)\ dx \ = \ \int_A D_\lambda f(x,\lambda)\ dx, \qquad \forall \lambda \in \Lambda.$$

Remarque 1.2.2 *1) Lorsque $f \in C_1(A' \times \Lambda)$ avec A' ouvert contenant A et lorsque A est borné et fermé, l'hypothèse (*) est automatiquement satisfaite.*

2) Pour n dérivations (n naturel strictement plus grand que 1), on demande que $\lambda \mapsto f(x,\lambda)$ appartienne à $C_n(\Lambda)$, que les dérivées (par rapport à λ) jusqu'à l'ordre $n-1$ soient intégrables sur A et que la majoration qui apparaît dans (*) fasse intervenir $D^n_\lambda f(x,\lambda)$. On obtient alors que la fonction $\lambda \mapsto \int_A f(x,\lambda)\, dx$ appartient à $C_n(\Lambda)$ et que la dérivation s'effectue encore de la même manière (jusqu'à l'ordre n).

Exemple(s) 1.2.3 *Un exemple très utile pour la suite...*

Pour tout réel $a>0$, la fonction $x \mapsto e^{-ax^2}$ est intégrable sur $\mathbb{R}$ et on a

$$\int_0^{+\infty} e^{-ax^2}\, dx \;=\; \frac{1}{2}\sqrt{\frac{\pi}{a}}$$

Cette égalité est appelée formule de Poisson et elle est fondamentale en statistiques notamment (intégration de gaussiennes, loi normale, ...) Pour la preuve de ce résultat, voir les notes du cours Math0009.

On a en fait la généralisation suivante : *pour tout $a>0$ et tout $b \in \mathbb{R}$, la fonction $x \mapsto \cos(bx)\, e^{-ax^2}$ est intégrable sur $\mathbb{R}$ et on a*

$$\boxed{\int_0^{+\infty} \cos(bx)\, e^{-ax^2}\, dx \;=\; \frac{1}{2}\sqrt{\frac{\pi}{a}}\, e^{-b^2/(4a)}}$$

Puisque les valeurs de la fonction cosinus sont dans $[-1,1]$ et que la fonction $x \mapsto e^{-ax^2}$ est intégrable sur $\mathbb{R}$ (quand $a>0$), il est clair que $x \mapsto \cos(bx)\, e^{-ax^2}$ est intégrable sur $\mathbb{R}$ quels que soient $a>0$ et $b \in \mathbb{R}$. Cela étant cette égalité peut être démontrée en ayant recours au théorème de dérivation des intégrales paramétriques.

On fixe $a>0$ et on considère a fonction $(x,b) \mapsto \cos(bx)\, e^{-ax^2}$ avec $A = [0,+\infty[$, $\Lambda = \mathbb{R}$ et $(x,b) \in A \times \Lambda$. Notes à compléter (voir cours enseigné).

1.3 Transformation de Fourier des fonctions intégrables sur $\mathbb{R}$

1.3.1 Définition et interprétation

Donnons tout d'abord l'introduction suivante, récupérée dans wikipedia (extraits)

[...] La transformation de Fourier associe à une fonction intégrable sur $\mathbb{R}$ et à valeurs réelles ou complexes, une autre fonction sur $\mathbb{R}$ appelée transformée de Fourier dont la variable indépendante peut s'interpréter en physique comme la fréquence ou la pulsation.

[...] La transformée de Fourier représente une fonction par la densité spectrale dont elle provient, en tant que moyenne de fonctions trigonométriques de toutes fréquences.

Lorsqu'une fonction représente un phénomène physique, comme l'état du champ électromagnétique ou du champ acoustique en un point, on l'appelle signal et sa transformée de Fourier s'appelle son spectre.

Pour illustrer cette introduction, voir la fin de cette sous-section.

Définition 1.3.1 *Soit f une fonction intégrable sur $\mathbb{R}$. La transformée de Fourier « négative » de f est la fonction*

$$\mathcal{F}^- f \;:\; y \in \mathbb{R} \mapsto \int_{\mathbb{R}} e^{-ixy}\, f(x)\, dx$$

et la transformée de Fourier « positive » de f est la fonction

$$\mathcal{F}^+ f \;:\; y \in \mathbb{R} \mapsto \int_{\mathbb{R}} e^{ixy}\, f(x)\, dx.$$

Pour la valeur en y de ces fonctions, on utilise les notations

$$\mathcal{F}^-_y f, \quad \mathcal{F}^+_y f$$

ou encore

$$\left(\mathcal{F}^- f\right)(y), \quad \left(\mathcal{F}^+ f\right)(y).$$

Notons que cette définition a bien un sens puisque pour tout y, la fonction $x \mapsto e^{ixy} f(x)$ (resp. $x \mapsto e^{-ixy} f(x)$) est intégrable sur $\mathbb{R}$ puisqu'en module c'est $|f|$.

Remarque 1.3.2 *Pour toute fonction f intégrable sur $\mathbb{R}$, on a*

$$\mathcal{F}_y^- f \;=\; \mathcal{F}_{-y}^+ f \quad \forall y \in \mathbb{R}.$$

Preuve. Il s'agit d'une simple écriture différente pour l'expression

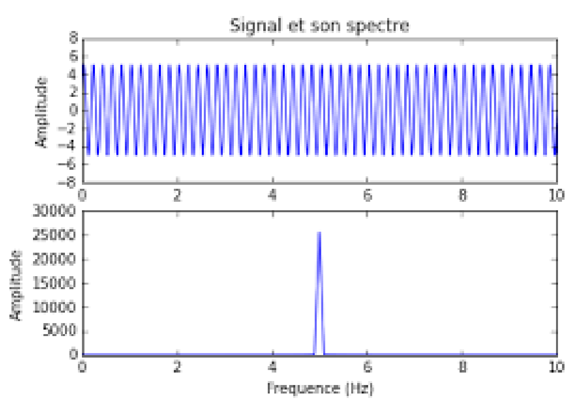
$$\int_{\mathbb{R}} e^{-ixy} f(x)\, dx \;=\; \int_{\mathbb{R}} e^{ix(-y)} f(x)\, dx.$$

$\square$

Pour illustrer un peu l'introduction qui met l'accent sur l'aspect fréquentiel de la transformation de Fourier, notons que, pour tout $r > 0$ (indispensable de recouper le cosinus, lequel n'est pas intégrable sur $\mathbb{R}$)

$$\mathcal{F}_y^{\pm}\left(\cos \chi_{[-r,r]}\right) \;=\; \begin{cases} \dfrac{\sin(r(y+1))}{y+1} + \dfrac{\sin(r(y-1))}{y-1} & \text{si } y \neq 1,\ y \neq -1 \\ \dfrac{\sin(2r)}{2} + r & \text{si } y = 1 \text{ et si } y = -1 \end{cases}$$

La plus grande valeur de cette transformée de $x \mapsto \cos(\mathbf{1}x) = \cos(-\mathbf{1}x)$ est donc en -1 et 1, c'est-à-dire qu'il y a un ≪ pic ≫ en ces points. En théorie des distributions où l'on peut vraiment prendre la transformée de Fourier de la distribution définie par le cosinus, on trouve effectivement les distributions de Dirac.



1.3.2 Exemples

Exemple(s) 1.3.3 *Transformation de Fourier de $\chi_{[a,b]}$ $(a, b \in \mathbb{R})$ et de $e^{-a|.|}$ $(a > 0)$:*

$$\mathcal{F}_y^{\pm}\chi_{[a,b]} \;=\; \begin{cases} \dfrac{e^{\pm iby} - e^{\pm iay}}{\pm iy} & si\ y \neq 0 \\ b - a & si\ y = 0 \end{cases}$$

$$\mathcal{F}_y^{\pm} e^{-a|.|} \;=\; \frac{2a}{y^2 + a^2}.$$

En particulier, pour tout $r >$, on a

$$\mathcal{F}_y^{\pm}\chi_{[-r,r]} \;=\; \begin{cases} \dfrac{2\sin(ry)}{y} & si\ y \neq 0 \\ 2r & si\ y = 0. \end{cases}$$

Preuve. Ce sont des calculs immédiats d'intégrales simples. □

Le premier exemple montre que la transformée de Fourier d'une fonction intégrable n'est pas nécessairement intégrable. Lorsqu'on aura besoin de cette propriété, on devra donc donner des conditions suffisantes pour avoir cette intégrabilité. Dans la suite, nous donnons les deux résultats les plus courants, bien utiles plus loin dans le cours.

Par ailleurs, ces deux exemples (ainsi que celui qui suit) conduisent à des transformées qui sont continues sur $\mathbb{R}$ et ont une limite nulle à l'infini. Ces propriétés sont en fait vraies pour toute transformée de Fourier de fonction intégrable, comme démontré dans la suite.

Voici alors un exemple qui se révèlera fondamental dans la preuve du théorème de Fourier.

Exemple(s) 1.3.4 (Le cas des fonctions gaussiennes) *Pour tout $a > 0$ on définit la gaussienne g_a par*

$$g_a(x) \;=\; e^{-ax^2},\; x \in \mathbb{R}.$$

Cette fonction est intégrable et on a

$$\mathcal{F}^{\pm}_x g_a \;=\; \left(\frac{\pi}{a}\right)^{1/2} g_{1/(4a)}(x) \quad \textit{pour tout } x$$

On en déduit que

$$\mathcal{F}^{\mp}_x \mathcal{F}^{\pm} g_a \;=\; 2\pi\, g_a(x) \quad \textit{pour tout } x$$

Preuve. L'intégrabilité des gaussiennes est claire. Cela étant, on a directement (on se réfère à une intégrale « remarquable » obtenue par le théorème de dérivation des intégrales paramétriques), quel que soit $y \in \mathbb{R}$,

$$\begin{aligned} \mathcal{F}^{\pm}_y g_a &= \int_{\mathbb{R}} e^{\pm ixy}\, e^{-ax^2}\, dx \\ &= 2\int_0^{+\infty} \cos(xy)\, e^{-ax^2}\, dx \\ &= \sqrt{\frac{\pi}{a}}\, e^{-y^2/(4a)} \\ &= \sqrt{\frac{\pi}{a}}\, g_{(4a)^{-1}}(y). \end{aligned}$$

On vient donc de voir que la transformée de Fourier d'une gaussienne est un multiple d'une gaussienne. Reprenons alors la transformée de Fourier de cette transformée de Fourier, en utilisant l'expression générale qui vient d'être trouvée. Quel que soit $x \in \mathbb{R}$, on a successivement [1]

$$\mathcal{F}^{\mp}_x \mathcal{F}^{\pm} g_a \;=\; \sqrt{\frac{\pi}{a}}\, \mathcal{F}^{\mp}_x g_{(4a)^{-1}} \;=\; \sqrt{\frac{\pi}{a}}\sqrt{\frac{\pi}{(4a)^{-1}}}\, g_{(4(4a)^{-1})^{-1}}(x) \;=\; 2\pi\, g_a(x)$$

et on conclut. □

1.3.3 Premières propriétés

Propriété(s) 1.3.5 *(1) La transformation de Fourier est un opérateur linéaire.*

(2) La transformée de Fourier d'une fonction intégrable f est une fonction bornée sur $\mathbb{R}$ car on a

$$|\mathcal{F}_y f| \;\leq\; \int_{\mathbb{R}} |f(x)|\, dx \quad \forall y \in \mathbb{R}$$

(3) La transformée de Fourier d'une fonction intégrable f est une fonction continue [2] sur $\mathbb{R}$.

(4) (Riemann-Lebesgue) La transformée de Fourier d'une fonction intégrable f tend vers 0 à l'infini.

1. Ici, les deux transformées de Fourier $\pm$ sont les mêmes car une gaussienne est paire ; cependant, pour le théorème de Fourier en toute généralité, cela ne sera pas toujours le cas et il importera de respecter l'alternance des « signes »

2. et même uniformément continue

Preuve. (1) C'est immédiat vu la linéarité de l'intégration.

(2) C'est immédiat car le module d'une exponentielle imaginaire pur est égal à 1 :

$$|\mathcal{F}_y f| \;=\; \left|\int_{\mathbb{R}} e^{\pm ixy}\, f(x)\; dx\right| \;\leq \int_{\mathbb{R}} \left|e^{\pm ixy}\right|\, |f(x)|\;\; dx \;=\; \int_{\mathbb{R}} |f(x)|\; dx \quad \forall y \in \mathbb{R}$$

(3) Admis

(4) Admis □

1.3.4 Dérivation et transformation de Fourier

Propriété(s) 1.3.6 *(1) Si $f \in C_L(\mathbb{R})$ et si $D^\alpha f$ est intégrable sur $\mathbb{R}$ quel que soit α tel que $\alpha \leq L$ alors*

$$\mathcal{F}_y^{\pm} D^\alpha f \;=\; (\mp iy)^\alpha\, \mathcal{F}_y^{\pm} f \quad \forall y \in \mathbb{R}$$

(2) Si les fonctions $x \mapsto x^\alpha\, f(x)$ $(\alpha \leq L)$ sont intégrables, alors $\mathcal{F}^{\pm} f \in C_L(\mathbb{R})$ et, quel que soit α tel que $|\alpha| \leq L$, on a

$$D^\alpha \mathcal{F}_y^{\pm} f \;=\; (\pm i)^\alpha \int_{\mathbb{R}} x^\alpha\, e^{\pm ixy}\, f(x)\; dx \quad \forall y \in \mathbb{R}.$$

Preuve. (1) Pour $L = 1$, cela résulte directement d'une intégration par parties. De fait, les deux limites

$$\lim_{x\to+\infty} \left(e^{\pm ixy}\, f(x)\right), \quad \lim_{x\to-\infty} \left(e^{\pm ixy}\, f(x)\right)$$

sont nulles car elles existent et les fonctions $x \mapsto e^{\pm ixy}\, f(x)$ sont intégrables à l'infini. Le cas général s'effectue de même, en répétant la manoeuvre précédente.

(2) L'expression

$$\mathcal{F}_y^{\pm} f \;=\; \int_{\mathbb{R}} e^{\pm ixy}\, f(x)\; dx$$

est une intégrale paramétrique. Dans le cas présent, vu les hypothèses données, celles du théorème de dérivation sont clairement satisfaites comme on le vérifie tout de suite et ainsi on obtient la dérivabilité et l'expression des dérivées en permutant dérivée et intégrale. □.

1.3.5 Intégration et transformation de Fourier

On a vu précédemment que la transformée de Fourier d'une fonction intégrable n'est pas nécessairement intégrable. Donnons ici une condition suffisante pour qu'elle le soit.

Propriété(s) 1.3.7 *Si $f \in C_2(\mathbb{R})$ est tel que f, Df et $D^2 f$ soient intégrables sur $\mathbb{R}$, alors $\mathcal{F}^{\pm} f$ est intégrable.*

Preuve. Notons tout d'abord que l'on a la continuité de la fonction dont on veut montrer l'intégrabilité. Cela étant, on obtient l'inégrabilité tout de suite car

$$|x|^2\, \left|\mathcal{F}_x^{\pm} f\right| \;=\; \left|x^2\, \mathcal{F}_x^{\pm} f\right| \;=\; \left|\mathcal{F}_x^{\pm} D^2 f\right| \;\leq\; C \quad \forall x$$

donc, pour tout $x \neq 0$, on a

$$\left|\mathcal{F}_x^{\pm} f\right| \;\leq\; \frac{C}{x^2}$$

et on conclut étant donné l'intégrabilité de la fonction $x \mapsto 1/x^2$ à l'infini. □

1.3.6 Théorème de transfert et de Fourier

Propriété(s) 1.3.8 (Transfert) *Si f et g sont intégrables alors*

$$\int_{\mathbb{R}} \mathcal{F}_x^{\pm} f\;\, g(x)\; dx \;=\; \int_{\mathbb{R}} f(x)\;\, \mathcal{F}_x^{\pm} g\; dx$$

Preuve. Il est clair que les deux membres de l'égalité ont un sens car le produit d'une fonction intégrable par une fonction bornée est intégrable.

Cela étant, une simple permutation de l'ordre d'intégration permet de conclure. De fait, la fonction $(x,y) \mapsto |f(y)|\,|g(x)|$ est intégrable sur $\mathbb{R}^2$ et on a

$$\begin{aligned}\int_{\mathbb{R}} \mathcal{F}_x^{\pm} f\; g(x)\; dx &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\pm ixy} f(y)\; dy \right)\; g(x)\; dx \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\pm ixy} g(x)\; dx \right)\; f(y)\; dy \\ &= \int_{\mathbb{R}} \mathcal{F}_y^{\pm} g\; f(y)\; dy\end{aligned}$$

□

Théorème 1.3.9 (Fourier) *Si f est intégrable et de transformée de Fourier intégrable, on a*

$$\mathcal{F}^{\mp}\mathcal{F}^{\pm} f \;=\; 2\pi\, f$$

en tout point où f est continu.

Preuve. Admis □

Notons ici qu'il est indispensable d'être capable d'exprimer cette égalité en un réel quelconque en explicitant les intégrales qui ≪ se cachent ≫ dans les transformées de Fourier.

1.4 Produit de composition

Nous présentons ici l'essentiel, avec pour but la présentation d'un autre outil de l'analyse du signal que la transformation de Fourier ET un lien fondamental entre ces deux outils.

Définition 1.4.1 *Soient f et g deux fonctions définies sur $\mathbb{R}$; si $y \in \mathbb{R}$ et si la fonction $x \mapsto f(x)\,g(y-x)$ est intégrable, son intégrale est notée*

$$(f * g)(y) \;=\; \int_{\mathbb{R}} f(x)\, g(y-x)\; dx.$$

Lorsque la fonction $x \mapsto f(x)\,g(y-x)$ est intégrable pour tout y, le produit de convolution (ou de composition) de f et g est la fonction

$$y \mapsto (f * g)(y).$$

Propriété(s) 1.4.2 *Le produit de composition est commutatif et linéaire sur chacun des facteurs.*

Preuve. C'est immédiat par changement de variable (linéaire) : on a

$$(f * g)(y) \;=\; \int_{\mathbb{R}} f(x)\, g(y-x)\; dx \;=\; \int_{\mathbb{R}} f(y-t)\, g(t)\; dt \;=\; (g * f)(y).$$

La linéarité est immédiate étant donné la linéarité de l'intégrale. □

Interprétation 1.4.3 Voir cours et aussi le descriptif par exemple à l'adresse

https ://fr.wikipedia.org/wiki/Produit_de_convolution

Propriété(s) 1.4.4 (Produit de composition et transformée de Fourier) *Si f et g sont intégrables alors*

$$\mathcal{F}^{\pm}(f * g) \;=\; \mathcal{F}^{\pm} f\; \mathcal{F}^{\pm} g$$

c'est-à-dire que la transformée de Fourier du produit de composition de deux fonctions intégrables est égale au produit des transformées de Fourier des deux fonctions.

Preuve. On admettra (même s'il n'est pas difficile de le démontrer) que la fonction de deux variables $(y,t) \mapsto f(t)\,g(y-t)$ est intégrable sur $\mathbb{R}^2$. Il s'ensuit que $f * g$ existe et est une fonction intégrable. On a alors directement, quel que soit $x \in \mathbb{R}$,

$$\begin{aligned}
\mathcal{F}_x^{\pm}(f*g) &= \int_{\mathbb{R}} e^{\pm ixy}\,(f*g)(y)\;dy \\
&= \int_{\mathbb{R}} e^{\pm ixy}\left(\int_{\mathbb{R}} f(t)\,g(y-t)\;dt\right)\;dy \\
&= \int_{\mathbb{R}} f(t)\left(\int_{\mathbb{R}} e^{\pm ixy}\,g(y-t)\;dy\right)\;dt \\
&= \int_{\mathbb{R}} f(t)\left(\int_{\mathbb{R}} e^{\pm ix(t+u)}\,g(u)\;du\right)\;dt \\
&= \int_{\mathbb{R}} e^{\pm ixt}\,f(t)\left(\int_{\mathbb{R}} e^{\pm ixu}\,g(u)\;du\right)\;dt \\
&= \left(\int_{\mathbb{R}} e^{\pm ixu}\,g(u)\;du\right)\times\left(\int_{\mathbb{R}} e^{\pm ixt}\,f(t)\;dt\right) \\
&= \mathcal{F}_x^{\pm} g \times \mathcal{F}_x^{\pm} f.
\end{aligned}$$

□

Chapitre 2

Seconde partie

2.1 Le développement limité de Taylor

Nous renvoyons ici à des pages extraites d'un ancien syllabus, en annexe.

A propos du développement de Taylor à plusieurs variables, il est important de noter qu'il est aussi valable dans des ouverts qui ne sont pas des produits cartésiens d'intervalles ouverts de la droite réelle. Ceci est déjà annoncé dans les pages en annexe (extraites d'un ancien syllabus). Un énoncé plus général est donc donné ci-dessous ; la preuve s'effectue de la même manière que dans le cas d'un produit cartésien d'intervalles ouverts.

Théorème 2.1.1 (Développement limité de Taylor, ouvert quelconque) *Soit Ω un ouvert de $\mathbb{R}^2$ et soit $f \in C_1(\Omega)$. Pour tous points (x_0, y_0), (x, y) de cet ouvert qui sont tels que les points du segment qui les relie appartiennent encore à l'ouvert, il existe $t_0 \in]0, 1[$ tel que*

$$f(x, y) \;=\; f(x_0, y_0) + (x - x_0)D_1 f(u_0, v_0) + (y - y_0)D_2 f(u_0, v_0)$$

où on a posé $u_0 = x_0 + t_0(x - x_0)$ et $v_0 = y_0 + t_0(y - y_0)$.

Soit Ω un ouvert de $\mathbb{R}^2$ et soit $f \in C_2(\Omega)$. Pour tous points (x_0, y_0), (x, y) de cet ouvert qui sont tels que les points du segment qui les relie appartiennent encore à l'ouvert, il existe $t_0 \in]0, 1[$ tel que

$$\begin{aligned} f(x, y) \;&=\; f(x_0, y_0) + (x - x_0)D_1 f(x_0, y_0) + (y - y_0)D_2 f(x_0, y_0) \\ &\quad + \frac{(x - x_0)^2}{2} D_1^2 f(u_0, v_0) + (x - x_0)(y - y_0)D_1 D_2 f(u_0, v_0) + \frac{(x - y_0)^2}{2} D_2^2 f(u_0, v_0) \end{aligned}$$

où on a posé $u_0 = x_0 + t_0(x - x_0)$ et $v_0 = y_0 + t_0(y - y_0)$.

2.2 Extrema : définitions et premières propriétés

Nous renvoyons ici à des pages extraites d'un ancien syllabus, en annexe.

2.3 Compléments de calcul matriciel

Soit une matrice réelle symétrique

$$A \;=\; \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

On a les résultats suivants.

Propriété(s) 2.3.1

1. *Les valeurs propres λ_1, λ_2 de A sont réelles.*
2. *La matrice A est toujours diagonalisable par une matrice S telle que $\widetilde{S}S = I$ (I désigne la matrice identité).*

3. *Les deux valeurs propres sont strictement positives (resp. négatives) si et seulement si* $\widetilde{X}AX > 0$ *(resp.* $\widetilde{X}AX < 0$*) pour tout vecteur colonne non nul* X*.*

   Si $\lambda_1\lambda_2 < 0$ *alors* $\widetilde{X}AX$ *peut changer de signe en fonction de* X*.*

4. *Les deux valeurs propres sont strictement positives (resp. négatives) si et seulement si le déterminant de* A *est strictement positif et* $a > 0$ *(resp.* $a < 0$*).*

   On a $\lambda_1\lambda_2 < 0$ *si et seulement si le déterminant de* A *est strictement négatif.*

Preuve (1) C'est immédiat par simple calcul : les valeurs propres d'une matrice sont les zéros de son polynôme caractéristique, lequel est ici

$$\lambda \mapsto \det \begin{pmatrix} a-\lambda & c \\ c & b-\lambda \end{pmatrix}$$

ou encore, en développant

$$\begin{aligned} \det \begin{pmatrix} a-\lambda & c \\ c & b-\lambda \end{pmatrix} &= (a-\lambda)(b-\lambda) - c^2 \\ &= \lambda^2 - (a+b)\lambda + ab - c^2. \end{aligned}$$

Le réalisant (discriminant) de ce polynôme est égal à

$$(a+b)^2 - 4(ab-c^2) = a^2+2ab+b^2-4ab+4c^2 = a^2-2ab+b^2+4c^2 = (a-b)^2+4c^2;$$

comme a, b, c sont des réels et que le carré d'un réel est toujours positif, ce réalisant est un réel positif. On peut donc conclure que le polynôme a deux zéros réels.

(2) Soit

$$X_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

un vecteur propre réel de valeur propre λ_1. Quitte à le multiplier par une constante non nulle, on peut supposer que

$$x_1^2 + y_1^2 = \widetilde{X_1}X_1 = 1.$$

Cela étant, définissons S par

$$S = \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}.$$

On obtient immédiatement

$$\widetilde{S}S = \begin{pmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{pmatrix} \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Montrons alors que

$$S^{-1}AS = \widetilde{S}AS$$

est une matrice diagonale. La première colonne de S est le vecteur propre X_1 ; notons Z la seconde colonne de S et remarquons que X_1 et Z sont des vecteurs colonnes orthogonaux, c'est-à-dire $\widetilde{X_1}Z = 0 = \widetilde{Z}X_1$. Cela étant, les colonnes de la matrice AS sont les vecteurs AX_1 et AZ ; on peut donc écrire

$$AS = \begin{pmatrix} AX_1 & AZ \end{pmatrix}.$$

On obtient ainsi

$$\widetilde{S}AS = \begin{pmatrix} \widetilde{X_1} \\ \widetilde{Z} \end{pmatrix} \begin{pmatrix} AX_1 & AZ \end{pmatrix} = \begin{pmatrix} \widetilde{X_1}AX_1 & \widetilde{X_1}AZ \\ \widetilde{Z}AX_1 & \widetilde{Z}AZ \end{pmatrix}.$$

Comme X_1 est vecteur propre de A de valeur propre λ_1 et tel que $\widetilde{X_1}X_1 = 1$, on a $AX_1 = \lambda_1 X_1$ et

$$\widetilde{S}AS = \begin{pmatrix} \lambda_1 & \widetilde{X_1}AZ \\ \lambda_1\widetilde{Z}X_1 & \widetilde{Z}AZ \end{pmatrix}.$$

Mais $\widetilde{Z}X_1 = 0$ donc

$$\widetilde{S}AS \;=\; \begin{pmatrix} \lambda_1 & \widetilde{X_1}AZ \\ 0 & \widetilde{Z}AZ \end{pmatrix}.$$

On sait aussi que A est symétrique, c'est-à-dire $\widetilde{A} = A$. Ainsi,

$$\widetilde{\widetilde{S}AS} \;=\; \widetilde{S}\widetilde{A}S \;=\; \widetilde{S}AS.$$

La matrice

$$\widetilde{S}AS \;=\; \begin{pmatrix} \lambda_1 & \widetilde{X_1}AZ \\ 0 & \widetilde{Z}AZ \end{pmatrix}$$

est donc symétrique, c'est-à-dire que ses éléments situés de part et d'autre de la diagonale principale sont égaux. Il s'ensuit que $\widetilde{X_1}AZ = 0$ donc que $\widetilde{S}AS$ est une matrice diagonale.

Pour (3) et (4) : voir cours. A COMPLETER

□

Remarque :
une matrice S telle que $\widetilde{S}S = I$ est appelée *matrice orthogonale*. On a alors $S^{-1} = \widetilde{S}$ et $S\widetilde{S} = I$. Les égalités $\widetilde{S}S = I$ et $S\widetilde{S} = I$ signifient que les colonnes de S ont des vecteurs orthonormés, de même que les lignes.

2.4 Retour aux extrema : cas des extrema dits « libres »

Nous renvoyons ici à des pages extraites d'un ancien syllabus, en annexe. Complétons un peu : grâce aux résultats concernant les matrices réelles symétriques, nous avons les outils pour démontrer le résultat donnant des conditions suffisantes pour qu'un point stationnaire soit un extremum (Propriété 5.2.4 des notes extraites de l'ancien syllabus).

Propriété(s) 2.4.1 *Soient un ouvert Ω de $\mathbb{R}^2$, une fonction $f \in C_2(\Omega)$, à valeurs réelles et un point (x_0, y_0) appartenant à Ω, stationnaire pour f.*

(1) Si $\det H_f(x_0, y_0) > 0$ et si $D_1^2 f(x_0, y_0)$ est un réel strictement positif, alors le point (x_0, y_0) est un minimum local strict de f dans Ω.

(2) Si $\det H_f(x_0, y_0) > 0$ et si $D_1^2 f(x_0, y_0)$ est un réel strictement négatif, alors le point (x_0, y_0) est un maximum local strict de f dans Ω.

(3) Si $\det H_f(x_0, y_0) < 0$ alors le point (x_0, y_0) n'est pas un extremum de f dans Ω.

Preuve. Pour démontrer ce résultat, il suffit de reprendre le développement limité de Taylor de f à l'ordre 2 au point stationnaire (x_0, y_0), d'utiliser une propriété de la continuité qui permet d'avoir une inégalité autour d'un point lorsque on l'a seulement au point et d'utiliser les résultats de la section 2.3.

Comme (x_0, y_0) appartient à un ouvert, on sait qu'il existe $\varepsilon > 0$ tel que si (x, y) est tel que $|x - x_0| < \varepsilon$ et $|y - y_0| < \varepsilon$, alors $(x, y) \in \Omega$ (cela signifie que si (x, y) est voisin du point (x_0, y_0) de l'ouvert alors il est aussi dans cet ouvert). De plus, un tel point est donc dans un carré centré en (x_0, y_0), donc le segment qui le relie à (x_0, y_0) est aussi inclus dans le carré, donc dans l'ouvert.

Cela étant, par le développement limité de Taylor, on sait alors qu'il existe un point (u_0, v_0) sur le segment qui relie (x_0, y_0) et (x, y) tel que

$$f(x,y) \;=\; f(x_0,y_0) + \frac{(x-x_0)^2}{2} D_1^2 f(u_0,v_0) + (x-x_0)(y-y_0) D_1 D_2 f(u_0,v_0) + \frac{(x-y_0)^2}{2} D_2^2 f(u_0,v_0)$$

(les termes avec les dérivées premières sont nuls puisque (x_0, y_0) est un point stationnaire). Notons X le vecteur colonne

$$\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

On obtient ainsi

$$f(x,y) \;=\; f(x_0,y_0) + \frac{1}{2}\widetilde{X} H_f(u_0,v_0) X$$

où H_f est la matrice hessienne de f.

Maintenant, on voudrait bien sûr utiliser les hypothèses sur la matrice hessienne, lesquelles font intervenir le point stationnaire et non (u_0, v_0) ! C'est donc le moment de se rappeler une propriété fort importante des fonctions continues. Comme f est deux fois continûment dérivable, les dérivées secondes sont continues, donc aussi le déterminant de sa matrice hessienne.

Les hypothèses de (1) impliquent donc l'existence de $\eta > 0$, que l'on peut bien sûr prendre plus petit que ε, tel que l'on ait encore $\det H_f(s,t) > 0$ et $D_1^2 f(s,t) > 0$ pour autant que $|s - x_0| < \eta$ et $|t - y_0| < \eta$.

On va alors conclure facilement en utilisant les point (3) et (4) de la Propriété 2.3.1. De fait, en prenant (x, y) tel que $|x - x_0| < \eta$ et $|y - y_0| < \eta$, on a aussi $|u_0 - x_0| < \eta$ et $|v_0 - y_0| < \eta$. Dès lors $\det H_f(u_0, v_0) > 0$ et $D_1^2 f(u_0, v_0) > 0$, donc les deux valeurs propres de la matrice $H_f(u_0, v_0)$ sont strictement positives et on en déduit finalement

$$\widetilde{X} H_f(u_0, v_0) X > 0$$

pour autant que $(x, y) \neq (x_0, y_0)$. On obtient ainsi

$$f(x,y) \; = \; f(x_0, y_0) + \frac{1}{2} \widetilde{X} H_f(u_0, v_0) X > f(x_0, y_0)$$

et on en conclut que (x_0, y_0) est bien un minimum local strict de f dans Ω.

Bien sûr la preuve s'adapte tout de suite au cas des hypothèses (2).

Enfin, si le déterminant de la matrice hessienne de f en (x_0, y_0) est négatif, il le reste dans un voisinage de ce point, les valeurs propres de $H_f(u_0, v_0)$ sont donc de signe opposé et l'on applique à nouveau (3) et (4) de a propriété 2.3.1 pour finalement conclure que comme $\widetilde{X} H_f(u_0, v_0) X$ peut changer de signe, le point (x_0, y_0) n'est pas un extremum de f dans Ω. □

2.5 Une application la régression linéaire

Nous renvoyons ici à des pages extraites d'un ancien syllabus, en annexe.

2.6 Les extrema liés (extrema « sous contrainte »)

En plus du cours enseigné, en guise d'introduction, on peut aller voir les pages en annexe (obtenues via wikipedia).

Considérons tout d'abord le cas d'une fonction f à valeurs réelles, de deux variables, définie et continûment dérivable dans un ouvert Ω pour laquelle on doit rechercher les éventuels extrema, mais pas dans Ω tout entier, mais seulement parmi les points de Ω qui sont situé sur une courbe, par exemple d'équation $g(x, y) = 0$, où g est une fonction continûment dérivable dans Ω, à valeurs réelles. On dit que l'on cherche à extrémaliser f sous la contrainte $g = 0$.

On va montrer le résultat suivant.

Propriété(s) 2.6.1 *Si le point (x_0, y_0) de Ω est un extremum local de f sous la contrainte $g = 0$ alors il existe $\lambda \in \mathbb{R}$ tel que (x_0, y_0, λ) vérifie*

$$\begin{cases} D_1 f(x_0, y_0) = \lambda D_1 g(x_0, y_0) \\ D_2 f(x_0, y_0) = \lambda D_2 g(x_0, y_0) \\ g(x_0, y_0) = 0 \end{cases}$$

Preuve. Voir plus loin □

Remarquer que le point (x_0, y_0, λ) est alors un point stationnaire pour la fonction (appelée « Lagrangien »)

$$L \; : \; (x, y, \lambda) \mapsto f(x, y) - \lambda g(x, y),$$

car

$$\begin{cases} D_1 L(x_0, y_0, \lambda) \; = \; D_1 f(x_0, y_0) - \lambda D_1 g(x_0, y_0) \; = \; 0 \\ D_2 L(x_0, y_0, \lambda) \; = \; D_2 f(x_0, y_0) - \lambda D_2 g(x_0, y_0) \; = \; 0 \\ D_3 L(x_0, y_0, \lambda) \; = \; -g(x_0, y_0) \; = \; 0 \end{cases}$$

Rechercher un extremum sous contrainte consiste donc à tout d'abord s'assurer qu'il existe, ensuite à résoudre le sytème d'équations ci-dessus et ensuite à chercher le ou les extrema en évaluant la fonction en chacune des solutions trouvées. Rappelons que si l'ensemble $\{(x,y) \in \Omega \ : \ g(x,y)=0\}$ est fermé borné et que f est continu sur cet ensemble, alors les extrema existent.

Venons-en alors à la justification de la propriété 2.6.1. Un résultat dont la preuve sort du cadre de ce cours affirme ce qui suit (on peut se convaincre que cela peut être vrai en prenant des exemples concrets, cf cours).

Proposition 2.6.2 *Soit $g \in C_1(\Omega)$, à valeurs réelles et soit $(x_0,y_0) \in \Omega$ tel que $(D_1g(x_0,y_0), D_2g(x_0,y_0)) \neq (0,0)$. Alors il existe un intervalle ouvert I de $\mathbb{R}$, un réel $t_0 \in I$, des fonctions g_1, g_2, à valeurs réelles, appartenant à $C_1(I)$, un voisinage V de (x_0,y_0) tels que $g_1(t_0)=x_0$, $g_2(t_0)=y_0$ et*

$$g(x,y)=0 \ (x,y) \in V \quad \Leftrightarrow \quad \exists t \in I \ : \left\{ \begin{array}{l} x = g_1(t) \\ y = g_2(t) \end{array} \right.$$

De plus $(Dg_1(t_0), Dg_2(t_0)) \neq (0,0)$.

Cela étant, rechercher les extrema locaux de f sous la contrainte $g=0$ revient donc à chercher les extrema de la fonction

$$F(t) \ = \ t \mapsto f\left(g_1(t), g_2(t)\right), \quad t \in I$$

On retombe donc dans le cas d'extrema libres pour une fonction d'une seule variable. Si maintenant (x_0,y_0) est extremum de f sous la contrainte $g=0$, alors t_0 est extremum de F donc est un point stationnaire pour celle-ci car F est continûment dérivable dans I (comme composée de fonction continûment dérivables). Comme

$$DF(t) \ = \ (D_1f)(g_1(t), g_2(t))\, Dg_1(t) \ + \ (D_2f)(g_1(t), g_2(t))\, Dg_2(t)$$

on obtient l'égalité (*)

$$0 = DF(t_0) \ = \ (D_1f)(x_0,y_0)\, Dg_1(t_0) \ + \ (D_2f)(x_0,y_0)\, Dg_2(t_0).$$

Mais on a aussi

$$g\left(g_1(t), g_2(t)\right) \ = \ 0 \quad \forall t \in I;$$

on en déduit que (**)

$$0 = (D_1g)(x_0,y_0)\, Dg_1(t_0) \ + \ (D_2g)(x_0,y_0)\, Dg_2(t_0).$$

Les deux égalités (*) et (**) peuvent se réécrire

$$\left\{ \begin{array}{lcl} (D_1f)(x_0,y_0)\, Dg_1(t_0) \ + \ (D_2f)(x_0,y_0)\, Dg_2(t_0) & = & 0 \\ (D_1g)(x_0,y_0)\, Dg_1(t_0) \ + \ (D_2g)(x_0,y_0)\, Dg_2(t_0) & = & 0 \end{array} \right.$$

ou encore, sous forme matricielle

$$\left(\begin{array}{cc} D_1f(x_0,y_0) & D_2f(x_0,y_0) \\ D_1g(x_0,y_0) & D_2g(x_0,y_0) \end{array} \right) \left(\begin{array}{c} Dg_1(t_0) \\ Dg_2(t_0) \end{array} \right) = \left(\begin{array}{c} 0 \\ 0 \end{array} \right)$$

Cela étant, comme $(Dg_1(t_0), Dg_2(t_0)) \neq (0,0)$, on en déduit que la matrice

$$\left(\begin{array}{cc} D_1f(x_0,y_0) & D_2f(x_0,y_0) \\ D_1g(x_0,y_0) & D_2g(x_0,y_0) \end{array} \right)$$

n'est pas inversible ou, de manière équivalente, que son déterminant est nul. Maintenant, comme la seconde ligne de la matrice n'est pas nulle, la première en est un multiple donc il existe $\lambda \in \mathbb{R}$ tel que

$$\left\{ \begin{array}{l} D_1f(x_0,y_0) = \lambda D_1g(x_0,y_0) \\ D_2f(x_0,y_0) = \lambda D_2g(x_0,y_0) \end{array} \right.$$

Comme $g(x_0,y_0)=0$, on a donc bien obtenu le sytème annoncé dans la propriété 2.6.1

2.7 Les extrema sur des ensembles fermés bornés

Voir cours et répétitions.