

3.11. Polynômes de Taylor.

Rappel. ($n=1$).

Soit f une fonction d'une variable de classe C^{k+1} , avec $k \in \mathbb{N}$

Soit $x_0 \in D(f)$. Le polynôme de Taylor de degré k de f autour de x_0 est donné par:

$$p_k(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 \\ + \dots + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

Cas importants:

$$k=1 : p_1(x) = f(x_0) + f'(x_0)(x-x_0) \quad (\text{approximation linéaire})$$

$$k=2 : p_2(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 \\ = p_1(x) + \frac{f''(x_0)}{2!}(x-x_0)^2 \quad (\text{approximation quadratique})$$

But: Généraliser cette notion aux fonctions de n variables.

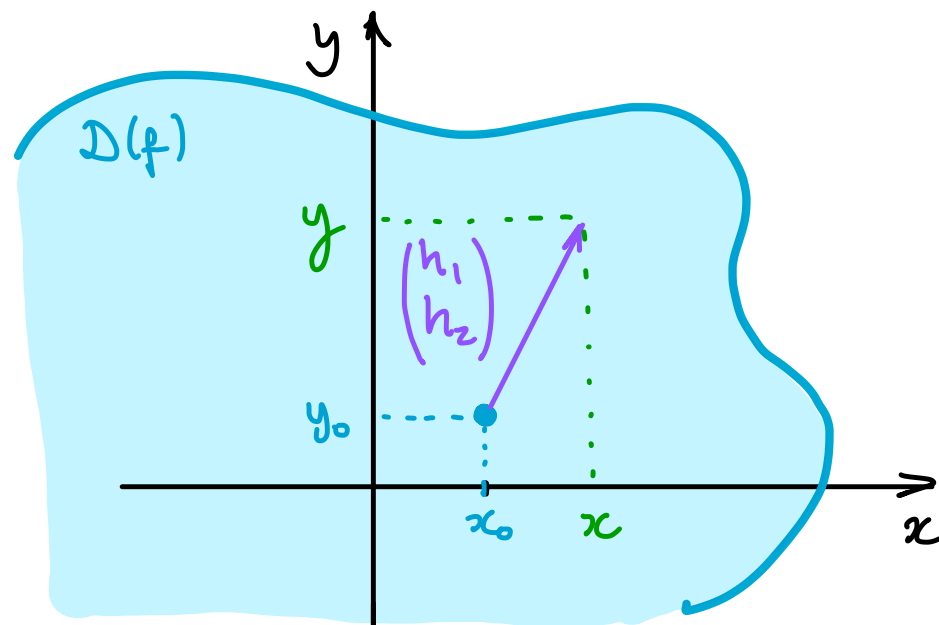
Soit $f: D(f) \rightarrow \mathbb{R}$ une fonction de deux variables de classe C^k (avec $k \geq 2$) et soit $(x_0, y_0) \in D(f)$

Soit $(x, y) \in D(f)$ un point fixé (proche de (x_0, y_0))

$$\text{On a } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 + x - x_0 \\ y_0 + y - y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

où nous avons posé

$$\vec{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$



Le segment reliant (x_0, y_0) à (x, y) peut-être paramétrée par

$$\vec{\varphi}(t) = (x_0 + t h_1, y_0 + t h_2), \text{ avec } t \in [0, 1]$$

Soit $g(t) = (f \circ \vec{\varphi})(t) = f(x_0 + t h_1, y_0 + t h_2)$ (fonction d'une variable)

Par construction: $g(0) = f(x_0, y_0)$

$$g(1) = f(x, y)$$

les polynômes de Taylor de degrés 1 et 2 de g autour de $t=0$ sont,

$$p_1(t) = g(0) + g'(0)t$$

$$p_2(t) = g(0) + g'(0)t + \frac{g''(0)}{2} t^2$$

On peut utiliser $p_1(1) = g(0) + g'(0)$ et $p_2(1) = g(0) + g'(0) + \frac{1}{2}g''(0)$ pour approximer $g(1) = f(x, y)$!

On doit donc calculer $g'(0)$ et $g''(0)$:

Comme $g(t) = (f \circ \vec{\varphi})(t) = f(x_0 + t h_1, y_0 + t h_2)$

on a

$$\begin{aligned} g'(t) &= \vec{\varphi}'(t) \cdot \vec{\nabla}_f(\vec{\varphi}(t)) = \langle \vec{\varphi}'(t), \vec{\nabla}_f(\vec{\varphi}(t)) \rangle \\ &= \langle \vec{h}, \vec{\nabla}_f(\vec{\varphi}(t)) \rangle \end{aligned}$$

$$= h_1 \frac{\partial f}{\partial x}(x_0 + t h_1, y_0 + t h_2) + h_2 \frac{\partial f}{\partial y}(x_0 + t h_1, y_0 + t h_2)$$

d'où : $g'(0) = h_1 \frac{\partial f}{\partial x}(x_0, y_0) + h_2 \frac{\partial f}{\partial y}(x_0, y_0)$

Ainsi, $p_1(1) = g(0) + g'(0)$

$$= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

et on retrouve la formule pour l'approximation linéaire.

D'autre part,

$$\begin{aligned} g''(t) &= \left(h_1 \frac{\partial f}{\partial x}(x_0 + th_1, y_0 + th_2) + h_2 \frac{\partial f}{\partial y}(x_0 + th_1, y_0 + th_2) \right)' \\ &= h_1 \frac{\partial^2 f}{\partial x^2}(x_0 + th_1, y_0 + th_2) h_1 + h_1 \frac{\partial^2 f}{\partial y \partial x}(x_0 + th_1, y_0 + th_2) h_2 \\ &\quad + h_2 \frac{\partial^2 f}{\partial x \partial y}(x_0 + th_1, y_0 + th_2) h_1 + h_2 \frac{\partial^2 f}{\partial y^2}(x_0 + th_1, y_0 + th_2) h_2 \end{aligned}$$

d'où :

$$\begin{aligned} g''(0) &= h_1 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) h_1 + h_1 \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) h_2 \\ &\quad + h_2 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) h_1 + h_2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) h_2 \end{aligned}$$

$$= (h_1 \ h_2) \underbrace{\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{pmatrix}}_{\text{matrice hessienne}} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

matrice hessienne

Si l'on note $H_f(x_0, y_0)$ la matrice hessienne, nous avons donc :

$$\begin{aligned} g''(0) &= \vec{h}^T H_f(x_0, y_0) \vec{h} \\ &= \langle \vec{h}, H_f(x_0, y_0) \vec{h} \rangle \end{aligned}$$

Ainsi,

$$\begin{aligned} p_2(1) &= g(0) + g'(0) + \frac{1}{2} g''(0) \\ &= f(x_0, y_0) + \langle \vec{\nabla} f(x_0, y_0), \vec{x} - \vec{x}_0 \rangle + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(x_0, y_0) (\vec{x} - \vec{x}_0) \rangle \end{aligned}$$

approxime $f(x, y)$.

Définition.

Soit $f: D(f) \rightarrow \mathbb{R}$ une fonction différentiable sur $D(f)$.

Soit $\vec{x}_0 = (x_0, y_0) \in D(f)$. Le polynôme de Taylor d'ordre 1 de f autour de $\vec{x}_0$ est défini par

$$p_1(\vec{x}) = f(\vec{x}_0) + \langle \vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0 \rangle, \text{ pour tout } \vec{x} \in \mathbb{R}^2$$

Si l'on suppose que f est de classe C^2 , on définit le polynôme de Taylor d'ordre 2 de f autour de $\vec{x}_0$ par:

$$p_2(\vec{x}) = f(\vec{x}_0) + \langle \vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0 \rangle + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \rangle,$$

pour tout $\vec{x} \in \mathbb{R}^2$.

$$\text{On a } p_2(\vec{x}) = p_1(\vec{x}) + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \rangle.$$

Remarques.

Ces définitions se généralisent naturellement aux fonctions de n variables $f: D(f) \rightarrow \mathbb{R}$ avec $D(f) \subset \mathbb{R}^n$:

$$p_1(\vec{x}) = f(\vec{x}_0) + \langle \vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0 \rangle,$$

$$\begin{aligned} p_2(\vec{x}) &= f(\vec{x}_0) + \langle \vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0 \rangle + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \rangle \\ &= p_1(\vec{x}) + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \rangle \end{aligned}$$

pour tout $\vec{x} \in \mathbb{R}^n$

Exemple.

Soit $f(x, y) = \sin(x+y) + \cos(x-3y)$, fonction de classe $C^\infty(\mathbb{R}^2)$

Calculer les polynômes de Taylor d'ordre 1 et 2 de f autour de l'origine.

$$\text{On a } f(0,0) = \sin(0) + \cos(0) = 1$$

$$\left. \begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \cos(x+y) - \sin(x-3y) \\ \frac{\partial f}{\partial y}(x, y) &= \cos(x+y) + 3\sin(x-3y) \end{aligned} \right\} \Rightarrow \vec{\nabla} f(0,0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\left. \begin{aligned} \frac{\partial^2 f}{\partial x^2}(x, y) &= -\sin(x+y) - \cos(x-3y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= -\sin(x+y) + 3\cos(x-3y) \\ \frac{\partial^2 f}{\partial y^2}(x, y) &= -\sin(x+y) - 9\cos(x-3y) \end{aligned} \right\} \Rightarrow H_f(0,0) = \begin{pmatrix} -1 & 3 \\ 3 & -9 \end{pmatrix}$$

$$\text{Comme } \langle \vec{\nabla} f(0,0), \begin{pmatrix} x \\ y \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = x+y$$

$$\begin{aligned} \text{et } \langle \begin{pmatrix} x \\ y \end{pmatrix}, H_f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} \rangle &= \langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} -1 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \rangle = \langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} -x+3y \\ 3x-3y \end{pmatrix} \rangle \\ &= -x^2 + 3xy + 3xy - 3y^2 = -x^2 - 3y^2 + 6xy \end{aligned}$$

Nous trouvons:

$$p_1(x,y) = f(0,0) + \langle \vec{\nabla} f(0,0), \begin{pmatrix} x \\ y \end{pmatrix} \rangle \Rightarrow p_1(x,y) = 1 + x + y$$

$$p_2(x,y) = p_1(x,y) + \frac{1}{2} \langle \begin{pmatrix} x \\ y \end{pmatrix}, H_f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} \rangle$$

$$\Rightarrow p_2(x,y) = 1 + x + y - \frac{1}{2}x^2 - \frac{3}{2}y^2 + 3xy$$

Remarque.

De manière analogue aux fonctions d'une variable, il est aussi possible d'utiliser des développements limités (en gardant les termes relevants) pour calculer des polynômes de Taylor :

Rappel: $\sin(t) = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots$

$$\cos(t) = 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots$$

$$\text{d'où } \sin(x+y) = (x+y) - \frac{1}{3!}(x+y)^3 + \dots$$

$$\cos(x-3y) = 1 - \frac{1}{2}(x-3y)^2 + \dots = 1 - \frac{1}{2}(x^2 - 6xy + 9y^2) + \dots$$

$$\Rightarrow \begin{cases} p_1(x, y) = 1 + x + y \\ p_2(x, y) = 1 + x + y - \frac{1}{2}x^2 - \frac{3}{2}y^2 + 3xy \end{cases} \quad (\text{comme avant})$$

Rappels. (Algèbre linéaire)

- Une matrice A de taille $n \times n$ est symétrique si $A^T = A$.
- Si A est une matrice symétrique de taille $n \times n$, alors elle est diagonalisable orthogonalement : $A = P D P^T$ où $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ est une matrice diagonale dont les coefficients diagonaux sont les n valeurs propres de A (pas forcément différentes) et $P = (\vec{u}_1, \dots, \vec{u}_n)$ est une matrice orthogonale telle que $A\vec{u}_j = \lambda_j \vec{u}_j$ et $\|\vec{u}_j\| = 1$.

- Soit $E = \{\vec{e}_1, \dots, \vec{e}_n\}$ la base canonique de $\mathbb{R}^n$ et $B = \{\vec{u}_1, \dots, \vec{u}_n\}$ une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de A .
la matrice $P = (\vec{u}_1, \dots, \vec{u}_n)$ est la matrice de passage de B vers E
et $P^T = P^{-1}$ est la matrice de passage de E vers B

Soit $\vec{x} \in \mathbb{R}^n$. On peut écrire :

$$\vec{x} = x_1 \vec{e}_1 + \dots + x_n \vec{e}_n \quad \text{et} \quad \vec{x} = y_1 \vec{u}_1 + \dots + y_n \vec{u}_n$$

$$\text{d'où } [\vec{x}]_E = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{x} \quad \text{et} \quad [\vec{x}]_B = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = P^T [\vec{x}]_E = P^T \vec{x}$$

- Si A est une matrice symétrique de taille $n \times n$, alors la fonction $Q: \mathbb{R}^n \rightarrow \mathbb{R}$ définie par $Q(\vec{x}) = \vec{x}^T A \vec{x}$, avec $\vec{x} \in \mathbb{R}^n$ est appelée la **forme quadratique** associée.

On dit que Q est définie positive si $Q(\vec{x}) > 0$ pour tout $\vec{x} \neq \vec{0}$

On dit que Q est définie négative si $Q(\vec{x}) < 0$ pour tout $\vec{x} \neq \vec{0}$

On dit que Q est non définie si $Q(\vec{x})$ prend à la fois des valeurs positives et négatives.

Si $A = P D P^T$, alors on peut écrire:

$$\begin{aligned} Q(\vec{x}) &= \langle \vec{x}, A \vec{x} \rangle = \vec{x}^T A \vec{x} = \vec{x}^T P D P^T \vec{x} \\ &= (P^T \vec{x})^T D (P^T \vec{x}) = \lambda_1 \underbrace{y_1^2}_{\geq 0} + \lambda_2 \underbrace{y_2^2}_{\geq 0} + \lambda_3 \underbrace{y_3^2}_{\geq 0} + \dots + \lambda_n \underbrace{y_n^2}_{\geq 0} \end{aligned}$$

- Soit A une matrice symétrique de taille $n \times n$.

Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de A .

Soit $Q(\vec{x}) = \vec{x}^T A \vec{x}$ la forme quadratique associée.

On peut montrer :

Q est définie positive $\Leftrightarrow \lambda_j > 0$ pour tout $j = 1, \dots, n$

Q est définie négative $\Leftrightarrow \lambda_j < 0$ pour tout $j = 1, \dots, n$

Q est non définie $\Leftrightarrow A$ possède des valeurs propres positives et négatives.

Fin du "rappel".

Application à la recherche de points d'extremum.

Soit $f: D(f) \rightarrow \mathbb{R}$ une fonction de n variables de classe C^2 .

Dans ce cas, $H_f(\vec{x}_0)$ est une matrice symétrique.

Si $\vec{x}_0$ est un point stationnaire de f , alors $\vec{\nabla} f(\vec{x}_0) = \vec{0}$

et le polynôme de Taylor d'ordre 2 autour de $\vec{x}_0$ devient:

$$\begin{aligned} P_2(\vec{x}) &= f(\vec{x}_0) + \frac{1}{2} \langle \vec{x} - \vec{x}_0, H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) \rangle \\ &= f(\vec{x}_0) + \frac{1}{2} Q_f(\vec{x} - \vec{x}_0) \end{aligned}$$

où Q_f est la forme quadratique associée à $H_f(\vec{x}_0)$ (la matrice hessienne de f évaluée en $\vec{x}_0$).

On distingue trois cas :

- 1) S'il existe un nombre $\delta > 0$ tel que pour tout $\vec{x} \in B(\vec{x}_0, \delta) \setminus \{\vec{x}_0\}$ on a $Q_f(\vec{x} - \vec{x}_0) > 0$, alors $\vec{x}_0$ est un point de minimum local de f .
- 2) S'il existe un nombre $\delta > 0$ tel que pour tout $\vec{x} \in B(\vec{x}_0, \delta) \setminus \{\vec{x}_0\}$ on a $Q_f(\vec{x} - \vec{x}_0) < 0$, alors $\vec{x}_0$ est un point de maximum local de f .
- 3) Si pour tout $\delta > 0$, $Q_f(\vec{x} - \vec{x}_0)$ prend des valeurs positives et négatives sur $B(\vec{x}_0, \delta) \setminus \{\vec{x}_0\}$ alors $\vec{x}_0$ n'est pas un point d'extremum local de f .

Théorème.

Soit $f: D(f) \rightarrow \mathbb{R}$ une fonction de n variables de classe C^2 .

Soit $\vec{x}_0 \in D(f)$ un point stationnaire de f (c-à-d. $\vec{\nabla} f(\vec{x}_0) = \vec{0}$).

Soient $\lambda_1, \lambda_2, \dots, \lambda_n$ les valeurs propres de $H_f(\vec{x}_0)$.

- Si $\lambda_j > 0$ pour tout $j=1, \dots, n$, alors $\vec{x}_0$ est un point de minimum local de f .
- Si $\lambda_j < 0$ pour tout $j=1, \dots, n$, alors $\vec{x}_0$ est un point de maximum local de f .
- S'il y a des valeurs propres positives et négatives, alors $\vec{x}_0$ est un point selle de f .