

CHAPITRE 3

Etude des extrema d'une fonction

1. Extrema : Rappels sur les fonctions d'une variable

Dans cette section on veut généraliser à plusieurs variable la discussion suivante concernant les fonctions d'une variable :

Soit f une fonction définie sur un intervalle I de \mathbf{R} ; on désire connaître les points x de I où $f(x)$ prend une valeur maximale ou minimale (on veut déterminer les *extrema* de f). Pour cela

- On commence par calculer les valeurs de f aux extrémités de I , $f(a)$, $f(b)$ au moins quand ces valeurs sont définies.
- on étudie alors f à l'intérieur de I : dans l'intervalle "ouvert" $]a, b[$.

On suppose que f est 2 fois dérivable.

PROPOSITION 3.1. *Si f admet un extremum au point x dans $]a, b[$ alors*

$$f'(x) = 0.$$

Cette proposition nous amène à trouver les solutions de l'équation

$$f'(x) = 0.$$

Les solutions sont appelés point *critiques* ou *stationnaires*. On regarde alors la dérivée seconde en un tel point

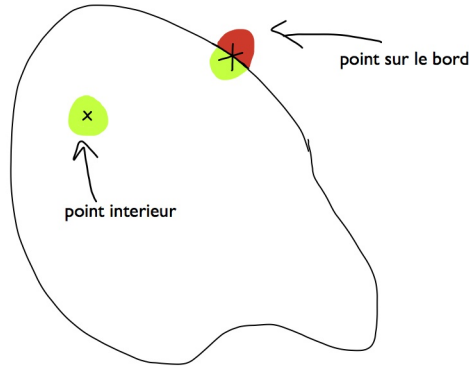
THÉORÈME 3.1. *Soit $x \in]a, b[$ tel que $f'(x) = 0$ alors, si*

- si $f''(x) < 0$, il existe un intervalle ouvert $I_x =]a_x, b_x[$ contenant x tel que f restreinte à I_x prend sa valeur maximale en x .
- si $f''(x) > 0$, il existe un intervalle ouvert $I_x =]a_x, b_x[$ contenant x tel que f restreinte à I_x prend sa valeur minimale en x .
- Si $f''(x) = 0$ plusieurs choses sont possible. En tout cas, on dit que f a un point d'inflexion en x .

Dans les deux premiers cas on dit que f admet un *extremum* (un *minimum* ou un *maximum*) local en x . Evidemment les extrema locaux sont des candidats à être des *extrema globaux* (sur I tout entier). A priori pour les point d'inflexions, tels que $f''(x) = 0$, on ne peut rien dire a priori et une étude plus approfondie est nécessaire :

EXEMPLE 1.0.1. les fonctions f suivantes admettent un point d'inflexion en $x = 0$

- $f(x) = x^4$: minimum en 0.
- $f(x) = -x^4$: maximum en 0
- $f(x) = x^3$ ni minimum ni maximum (pas même local), f est strictement croissante.



2. Cas des fonctions de deux variables ; notion de point critique

On va généraliser la discussion précédente aux fonction à deux variables. On se donne f définie sur un domaine D de \mathbf{R}^2 et on désire déterminer les $\vec{x} = (x, y)$ où $f(\vec{x})$ prend des valeurs extrêmes.

On va commencer par donner l'analogie des extrémités de I .

DÉFINITION 3.1. *On définit l'intérieur de D , D° comme l'ensemble des \vec{x} dans D tels qu'il existe une boule $B(\vec{x}, r)$ de rayon > 0 entièrement contenue dans D . On définit le bord de D , ∂D comme étant le complémentaire de D° (les points de D qui ne sont pas dans l'intérieur).*

Pour déterminer les extrema de f sur D , on procède ainsi

- (1) On étudie les extrema de f sur le bord ∂D : le bord est en général la réunion d'une courbe et de point isolés. L'étude de f le long du bord se ramène la plupart du temps à étudier les extrema d'une fonction d'une variable.
- (2) On étudie les extrema dans l'intérieur D° : on déterminera *a priori* des extrema locaux.
- (3) On compare les différent extrema qu'on a trouvé.

Comme on l'a dit la première étape se ramène à l'étude des variations d'une fonction d'une seule variable.

On passe donc à la deuxième étape : étudier les extrema de f contenus dans l'intérieur D° de D . Pour ce faire, on aura besoin de calculer les dérivées partielles de f et on supposera que f est différentiable à l'ordre au moins 3 sur D .

PROPOSITION 3.2. *Soit $\vec{x}_0 = (x_0, y_0)$ un point de l'intérieur de D . Si \vec{x}_0 est un extremum de f sur D° alors*

$$\nabla f(\vec{x}_0) = \left(\frac{\partial f}{\partial x}(\vec{x}_0), \frac{\partial f}{\partial y}(\vec{x}_0) \right) = (0, 0).$$

DÉFINITION 3.2. Un point \vec{x}_0 tel que

$$\nabla f(\vec{x}_0) = \left(\frac{\partial f}{\partial x}(\vec{x}_0), \frac{\partial f}{\partial y}(\vec{x}_0) \right) = (0, 0)$$

est appelé point critique pour f .

Preuve: Soit $\vec{x}_0 = (x_0, y_0)$ un point de l'intérieur et $B(\vec{x}_0, r)$, $r > 0$ une boule centrée en \vec{x}_0 et contenue dans D . Considérons la fonction d'une variable

$$f_1 : x \mapsto f(x, y_0);$$

f_1 est définie sur l'intervalle $]x_0 - r, x_0 + r[$ et par hypothèse admet un extremum en x_0 . On a donc

$$f_1'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0) = 0.$$

De même, la fonction d'une variable

$$f_2 : y \mapsto f(x_0, y);$$

f_2 est définie sur l'intervalle $]y_0 - r, y_0 + r[$ et par hypothèse admet un extremum en y_0 . On a donc

$$f_2'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

□

3. Le Hessian et la formule d'approximation quadratique

Soit $\vec{x}_0 \in D^\circ$, le Hessian de f en \vec{x}_0 est défini par la formule suivante

$$\text{Hess}(f)(\vec{x}_0) := \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2(\vec{x}_0) - \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2}(\vec{x}_0).$$

On a alors le résultat suivant

THÉORÈME 3.2. Soit f définie sur $B(\vec{x}_0, r)$, $r > 0$, différentiable à l'ordre 3 au moins. Supposons que \vec{x}_0 soit un point critique pour f alors si

- Si $\text{Hess}(f)(\vec{x}_0) < 0$, et si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$ alors il existe $0 < r' < r$ tel que f restreinte à la boule $B(\vec{x}_0, r')$ prend sa valeur maximale en \vec{x}_0 . On dit que \vec{x}_0 est un maximum local pour f .
- Si $\text{Hess}(f)(\vec{x}_0) < 0$, et si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$, il existe $0 < r' < r$ tel que f restreinte à la boule $B(\vec{x}_0, r')$ prend sa valeur minimale en \vec{x}_0 . On dit que \vec{x}_0 est un minimum local pour f .
- Si $\text{Hess}(f)(\vec{x}_0) > 0$, f n'est ni un maximum, ni un minimum local en \vec{x}_0 . On dit que \vec{x}_0 est un point selle, ou un point col de f .
- Si $\text{Hess}(f)(\vec{x}_0) = 0$ plusieurs configurations sont possibles. En tous cas le point \vec{x}_0 sera dit dégénéré.

REMARQUE 3.1. Dans le cas où $\text{Hess}(f)(\vec{x}_0) < 0$, pourquoi considérer $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$ plutôt que $\frac{\partial^2 f}{\partial y^2}(\vec{x}_0)$? En fait cela ne fait aucune différence : si

$$\text{Hess}(f)(\vec{x}_0) := \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2(\vec{x}_0) - \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) < 0$$

alors nécessairement

$$\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) > 0$$

et les dérivées d'ordre 2 en x et en y sont les même.

3.1. Exemples. Dans cette section on va donner des exemples caractéristiques des situations précédentes.

3.1.1. *Cas d'un minimum local.* Soit

$$f(x, y) = x^2 + y^2.$$

On a

$$\nabla f(x, y) = (2x, 2y)$$

et donc le seul point critique est le point $(0, 0)$ et $f(0, 0) = 0$. On a

$$\frac{\partial^2 f}{\partial^2 x}(x, y) = 2, \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0, \frac{\partial^2 f}{\partial^2 y}(x, y) = 2$$

et donc

$$\text{Hess}(f)(0, 0) = -4.$$

On a $\frac{\partial^2 f}{\partial^2 x}(0, 0) = 2 > 0$ donc f a un minimum local en $(0, 0)$: c'est même un minimum absolu et il est facile de voir cela directement : pour tout (x, y)

$$f(x, y) = x^2 + y^2 \geq 0 = f(0, 0).$$

3.1.2. *Cas d'un maximum local.* Il suffit de considérer l'opposé du cas précédent Soit

$$f(x, y) = -(x^2 + y^2).$$

La discussion précédente s'applique et on voit bien que $(0, 0)$ est un maximum absolu :

$$f(x, y) = -(x^2 + y^2) \leq 0 = f(0, 0).$$

3.1.3. *Cas d'un point col.* Soit

$$f(x, y) = x^2 - y^2 = (x - y)(x + y).$$

On a

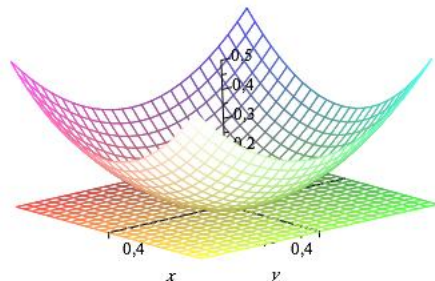
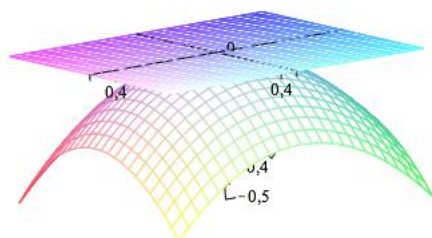
$$\nabla f(x, y) = (2x, -2y)$$

et donc le seul point critique est le point $(0, 0)$ et $f(0, 0) = 0$. On a

$$\frac{\partial^2 f}{\partial^2 x}(x, y) = 2, \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0, \frac{\partial^2 f}{\partial^2 y}(x, y) = -2$$

et donc

$$\text{Hess}(f)(0, 0) = 4 > 0.$$

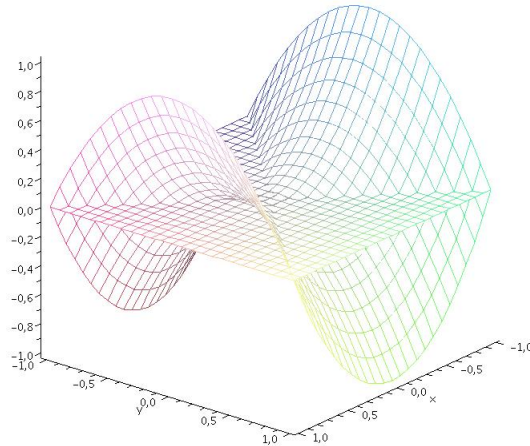
FIGURE 1. Minimum $f(x, y) = x^2 + y^2$ FIGURE 2. Maximum $f(x, y) = -x^2 - y^2$

On est dans le cas d'un point col : si $y = x$ ou bien $y = -x$, $f(x, y) = 0$, la fonction f prend donc la valeur $f(0, 0)$ le long des deux droites d'équations

$$x - y = 0 \text{ ou bien } x + y = 0$$

et

$$\begin{aligned} f(x, y) &> 0 \text{ si } x - y \text{ et } x + y \text{ sont de même signe} \\ f(x, y) &< 0 \text{ si } x - y \text{ et } x + y \text{ sont de signe opposé.} \end{aligned}$$

FIGURE 3. Col $f(x, y) = x^2 - y^2$

3.1.4. Cas d'un point dégénéré.

$$f(x, y) = x^2 + y^3.$$

On a

$$\nabla f(x, y) = (2x, 3y^2)$$

et donc le seul point critique est le point $(0, 0)$ et $f(0, 0) = 0$. On a

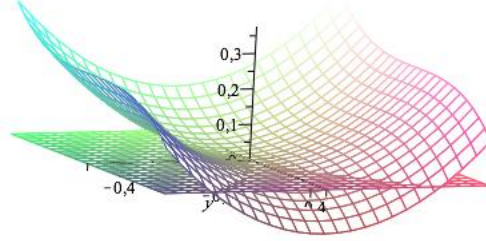
$$\frac{\partial^2 f}{\partial^2 x}(x, y) = 2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0, \quad \frac{\partial^2 f}{\partial^2 y}(x, y) = 6y$$

et donc

$$\text{Hess}(f)(0, 0) = 0.$$

Le point est donc dégénéré et comme le montre la figure, ce n'est ni un maximum, ni un minimum, ni un col. En fait, dans le cas des points dégénérés, on peut obtenir toute une variété de situations différentes. Dans ce cours, on ne les classifera pas de manière systématique.

3.2. La formule d'approximation quadratique. On va donner une idée de la preuve de ce théorème : pour cela nous aurons besoin de la formule suivante que nous admettrons

FIGURE 4. Point dégénéré $f(x, y) = x^2 + y^3$.

THÉORÈME 3.3 (Formule d'approximation quadratique). *Soit f une fonction différentiable à l'ordre 3 sur une boule $B(\vec{x}_0, r)$. Écrivons*

$$\vec{x}_0 = (x_0, y_0), \quad (x, y) = \vec{x}_0 + \vec{h} = (x_0 + h_1, y_0 + h_2)$$

$$\begin{aligned} f(x, y) = f(\vec{x}_0 + \vec{h}) &= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0) \cdot h_1 + \frac{\partial f}{\partial y}(\vec{x}_0) \cdot h_2 \\ &+ \frac{1}{2} \left(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \cdot h_1^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot h_1 h_2 + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0) \cdot h_2^2 \right) + (h_1^2 + h_2^2) \varepsilon(\vec{h}) \end{aligned}$$

En particulier si \vec{x}_0 est un point critique, la formule devient

$$\begin{aligned} f(x, y) &= f(x_0 + h_1, y_0 + h_2) \\ &= f(\vec{x}_0) + \frac{1}{2} \left(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \cdot h_1^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot h_1 h_2 + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0) \cdot h_2^2 \right) + (h_1^2 + h_2^2) \varepsilon(\vec{h}) \\ &= f(\vec{x}_0) + Q(h_1, h_2) + (h_1^2 + h_2^2) \varepsilon(\vec{h}) \end{aligned}$$

avec

$$Q(U, V) = \frac{1}{2} \left(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \cdot U^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot UV + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0) \cdot V^2 \right) = aU^2 + bUV + cV^2$$

avec

$$a = \frac{1}{2} \frac{\partial^2 f}{\partial^2 x}(\vec{x}_0), \quad b = \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0), \quad c = \frac{1}{2} \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0).$$

Considérons le polynôme quadratique

$$P(X) = aX^2 + bX + c;$$

mettant V en facteur, on a

$$Q(U, V) = V^2(a(U/V)^2 + b(U/V) + c) = V^2P(U/V)$$

et donc si $h_2 \neq 0$, on a

$$f(x_0 + h_1, y_0 + h_2) = f(\vec{x}_0) + h_2^2 P(h_1/h_2) + (h_1^2 + h_2^2)\varepsilon(\vec{h});$$

cela nous ramène à étudier le signe de $P(X)$. Le discriminant de ce polynôme est $\Delta = b^2 - 4ac$ et vaut précisément

$$\Delta = \text{Hess}(f)(\vec{x}_0).$$

4. Extrema locaux

Supposons que le Hésien est négatif; comme

$$\Delta = b^2 - 4ac < 0, \text{ on voit que } a, c \neq 0 \text{ et que } ac > 0; \text{ de plus}$$

le polynôme $P(X)$ ne s'annule pas sur \mathbf{R} et reste de signe constant. Ce signe est donné par le signe du premier coefficient a c.a.d. le signe de $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$.

Supposons alors que $a = \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$, alors pour $h_2 \neq 0$, la partie quadratique vaut

$$Q(h_1, h_2) = h_2 P(h_1/h_2) < 0$$

et si $h_2 = 0$, la partie quadratique vaut

$$ah_1^2 \leq 0$$

avec égalité ssi $h_1 = 0$. Ainsi compte-tenu du fait que le terme $(h_1^2 + h_2^2)\varepsilon(\vec{h})$ est un terme d'erreur quand (h_1, h_2) est proche de zero, on voit que

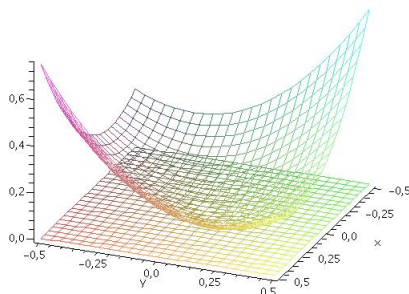
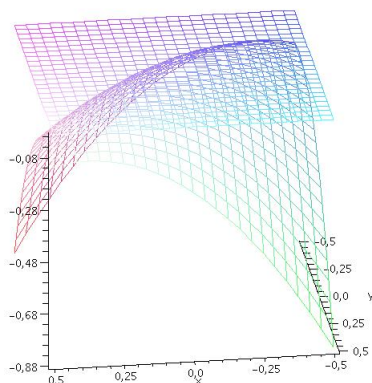
$$f(x, y) - f(\vec{x}_0) = ah_1^2 + bh_1h_2 + ch_2^2 + \text{Erreur}$$

a tendance à être négatif quand (x, y) est proche de (x_0, y_0) : $f(x_0, y_0)$ est maximal au voisinage de \vec{x}_0 : \vec{x}_0 est un maximum local.

Le même raisonnement s'applique quand $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$ et fournit un minimum local.

5. Etude du point col

Supposons que le Hésien est positif.

FIGURE 5. Minimum $f(x, y) = x^2 - xy + y^2$ FIGURE 6. maximum $f(x, y) = -(x^2 - 3xy/2 + y^2)$

5.1. Cas où P est de degré 2. Supposons d'abord que $\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \neq 0$ (ie. P est un "vrai" polynôme de degré 2) alors le polynôme $P(X)$ a deux racines réelles distinctes

$$X_+ = \frac{-b + \sqrt{\Delta}}{2a} = \frac{-\frac{\partial^2 f}{\partial x \partial y} + \sqrt{\text{Hess}(f)}}{\frac{\partial^2 f}{\partial^2 x}},$$

$$X_- = \frac{-b - \sqrt{\Delta}}{2a} = \frac{-\frac{\partial^2 f}{\partial x \partial y} - \sqrt{\text{Hess}(f)}}{\frac{\partial^2 f}{\partial^2 x}}$$

et s'écrit

$$P(X) = a(X - X_+)(X - X_-).$$

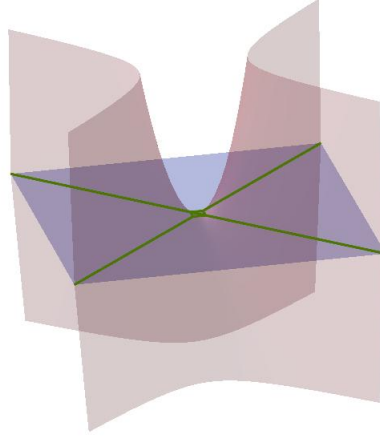


FIGURE 7. Un point col et ses droites isotropes

On en déduit que le terme quadratique vaut

$$Q(h_1, h_2) = h_2^2 P(h_1/h_2) = a(h_1 - X_+ h_2)(h_1 - X_- h_2) = \frac{1}{2} \frac{\partial^2 f}{\partial^2 x}(\vec{x}_0)(h_1 - X_+ h_2)(h_1 - X_- h_2);$$

Ainsi on voit que la différence

$$f(x, y) - f(\vec{x}_0) = \frac{1}{2} \frac{\partial^2 f}{\partial^2 x}(\vec{x}_0)(h_1 - X_+ h_2)(h_1 - X_- h_2) + \text{Erreur.}$$

à tendance à être très petite quand (h_1, h_2) est proche de $(0, 0)$ et varie le long d'une des deux droites d'équation

$$h_1 - X_+ h_2 = 0, \quad h_1 - X_- h_2 = 0;$$

en revanche quand (h_1, h_2) est en dehors de ces droites et proche de $(0, 0)$, la différence change de signe en fonction du quadrant (défini par ces deux droites) dans lequel (h_1, h_2) se trouve. Rappelons que $h_1 = x - x_0$, $h_2 = y - y_0$ et les équations de ces droites s'écrivent

$$x - x_0 - X_+(y - y_0) = 0, \quad x - x_0 - X_-(y - y_0) = 0.$$

5.2. Cas où P est de degré 1. Si maintenant $a = \frac{1}{2} \frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) = 0$, P est un polynôme de degré 1 et la formule précédente s'écrit

$$\begin{aligned} f(x, y) - f(\vec{x}_0) &= b.h_1 h_2 + c.h_2^2 + \text{Erreur} \\ &= \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0)(h_1 - X h_2)h_2 + \text{Erreur}, \quad X = -\frac{1}{2} \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0) / \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0). \end{aligned}$$

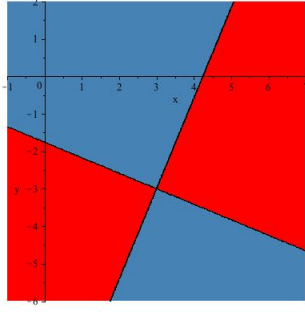


FIGURE 8. Un point col et ses droites isotropes vue du dessus

A nouveau on obtient deux droites caractéristiques, le long desquelles $f(x, y) - f(\vec{x}_0)$ est très petit, et en dehors desquelles le signe de la différence d'équation

$$x - x_0 - X(y - y_0), \quad y - y_0 = 0.$$

DÉFINITION 3.3. *Supposons que $\vec{x}_0 = (x_0, y_0)$ soit un point col : ie. $\nabla f(\vec{x}_0) = 0$, $\text{Hess}(f)(\vec{x}_0) > 0$. Les droites d'équation*

$$x - x_0 - X_+(y - y_0) = 0, \quad x - x_0 - X_-(y - y_0) = 0$$

(si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \neq 0$) ou d'équation

$$x - x_0 - X(y - y_0), \quad y - y_0 = 0$$

(si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \neq 0$) sont appelées les droites caractéristiques (ou isotropes) associées au point col \vec{x}_0 .

6. Cas des points dégénérés (facultatif)

Si $\text{Hess}(f)(\vec{x}_0) = 0$ le discriminant de P est nul.

Il y a plusieurs possibilités : si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \neq 0$ alors P est un "vrai" polynôme de degré 2 et il admet une racine double $X_0 = -b/2a$. Ainsi

$$f(x, y) - f(\vec{x}_0) = \frac{1}{2}(h_1 - X_0 h_2)^2 + \text{Erreur}, \quad X_0 = -\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) / \frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$$

Ainsi quand (x, y) varie au voisinage de \vec{x}_0 le long de la droite d'équation

$$x - x_0 - X_0(y - y_0) = 0$$

la différence $f(x, y) - f(\vec{x}_0)$ a tendance à être très petite. En dehors de cette droite (encore appelée droite caractéristique ou isotrope) on voit que au voisinage de \vec{x}_0 , cette différence a tendance à être du même signe que $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$. Ainsi le graphe de f ressemble près de \vec{x}_0 , soit à une "ligne de crête" (si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$), soit à une "crevasse" (si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$). Il y a bien d'autres configurations possibles : P peut être un polynôme

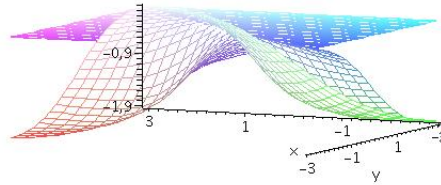


FIGURE 9. Crête $f(x, y) = -(x - y)^2 / (1 + x^2 + y^2)$

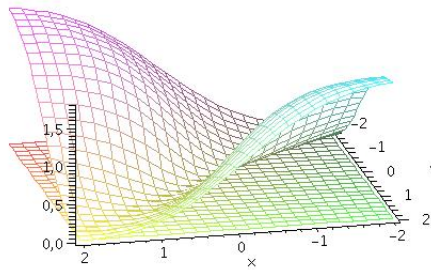
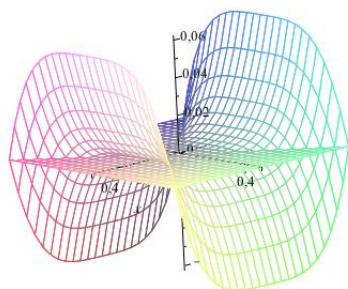
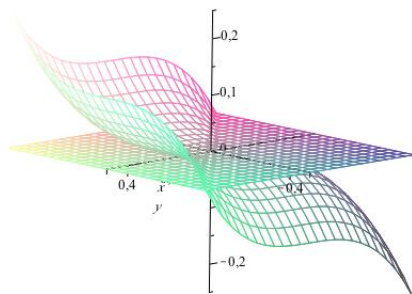


FIGURE 10. Crevasse $f(x, y) = (x - y)^2 / (1 + x^2 + y^2)$

constant, nul ou non nul. Dans le dernier cas on obtient a nouveau une droite isotrope (verticale) d'équation

$$y - y_0 = 0,$$

et une configuration de ligne de crête ou de crevasse.

FIGURE 11. Col $f(x, y) = x^4 - y^4$ FIGURE 12. Point dégénéré $f(x, y) = x^3 + y^3$

Enfin si P est le polynôme nul, la formule d'approximation quadratique prend la forme

$$f(x, y) = f(x_0 + h_1, y_0 + h_2) = f(\vec{x}_0) + (h_1^2 + h_2^2)\varepsilon(\vec{h})$$

et il est nécessaire de regarder de plus près la forme du terme d'erreur $(h_1^2 + h_2^2)\varepsilon(\vec{h})$: on peut obtenir un extrema local ($f(x, y) = x^4 + y^4$), un point col ($f(x, y) = x^4 - y^4$) ou quelque chose de tout à fait différent.