FONCTIONS NUMERIQUES DE DEUX VARIABLES

L'objectif de ce chapitre est de construire pour les fonctions de deux variables des outils analogues à ceux des fonctions d'une variable (limites, continuité, dérivées, …) en particulier pour déterminer un maximum ou un minimum. Pour cela, il faut définir des limites, et donc quelque chose qui remplace la valeur absolue.

I – Topologie sur $\mathbb{R}^2$

Pour définir une topologie sur $\mathbb{R}^2$, on va s'appuyer sur des interprétations géométriques dans le plan. En effet lorsque l'on définit un repère $(O,\vec{i},\vec{j})$ dans le plan, on crée une bijection entre $\mathbb{R}^2$ et l'ensemble des points du plan ou l'ensemble des vecteurs du plan : $(x,y)\mapsto M$ ou $\overrightarrow{OM}$ où M a pour coordonnées (x,y).
On identifiera le couple (x,y), le point M et le vecteur $U=\overrightarrow{OM}$. Chaque élément de $\mathbb{R}^2$ pourra donc être interprété soit comme un point ou soit comme un vecteur.
Si A et B sont deux points de $\mathbb{R}^2$, le vecteur $\overrightarrow{AB}$ est donc égal à $B-A$.

1) Droites et segments

<u>Définition</u> : Si A et U sont deux éléments de $\mathbb{R}^2$ et si $U\neq 0$, on appelle droite affine passant par A et de vecteur directeur U l'ensemble des points M de $\mathbb{R}^2$ tels que $\overrightarrow{AM}$ soit colinéaire au vecteur U : $d_{A,U}=\{M\in\mathbb{R}^2 \,/\, \exists t\in\mathbb{R}\quad M=A+tU\}$.
L'application de $\mathbb{R}$ dans $d_{A,U}$: $t\mapsto A+tU$ est une bijection appelée paramétrage de $d_{A,U}$.
Si A et B sont deux points distincts de $\mathbb{R}^2$, la droite (AB) est $d_{A,B-A}$.
On a : $(AB)=\{M\in\mathbb{R}^2 \,/\, \exists t\in\mathbb{R}\quad M=(1-t)A+tB\}$.

La bijection est évidente car $U\neq 0$.
Il existe bien sûr une infinité de paramétrages d'une même droite.
On a : $d_{A,U}=d_{B,V}$ si et seulement si $B-A$ et V sont colinéaires à U.

<u>Théorème</u> : Une partie D de $\mathbb{R}^2$ est une droite affine si et seulement si il existe des réels a, b et c avec $(a,b)\neq(0,0)$ tels que : $D=\left\{(x,y)\in\mathbb{R}^2 \,/\, ax+by+c=0\right\}$.
L'équation $ax+by+c=0$ est appelée équation cartésienne de D.

<u>Démonstration</u> : Si $D=\left\{(x,y)\in\mathbb{R}^2 \,/\, ax+by+c=0\right\}$, l'équation équivaut à $M=A+tU$

- Si $b\neq 0$, avec : $A=\left(0,-\dfrac{c}{b}\right)$, $U=(b,-a)$ et $t=\dfrac{x}{b}$.
- Si $a\neq 0$, avec : $A=\left(-\dfrac{c}{a},0\right)$, $U=(b,-a)$ et $t=-\dfrac{y}{a}$.

Réciproquement, si $D=d_{A,U}$ avec $A=(x_A,y_A)$ et $U=(\alpha,\beta)$ où $U\neq(0,0)$, on a :

$M\in D$ si et seulement si $\exists t\in\mathbb{R}\quad M=A+tU$, donc si $\exists t\in\mathbb{R}\quad \begin{cases} x-x_A=t\alpha \\ y-y_A=t\beta \end{cases}$, donc si $\beta(x-x_A)-\alpha(y-y_A)=0$, donc une équation de la forme $ax+by+c=0$ avec $(a,b)\neq(0,0)$ car $U\neq(0,0)$.

Définition : Si A et B sont deux éléments de $\mathbb{R}^2$, on appelle segment $[A,B]$ l'ensemble des points M de $\mathbb{R}^2$ défini par : $[A,B]=\{M\in\mathbb{R}^2 / \exists t\in[0,1] \quad M=(1-t)A+tB\}$.
L'application de $[0,1]$ dans $[A,B]$: $t\mapsto(1-t)A+tB$ est une bijection appelée paramétrage du segment $[A,B]$.

Remarque : On peut écrire indifféremment $(1-t)A+tB$ ou $tA+(1-t)B$, ce qui revient à changer t en $1-t$. Donc : $[A,B]=[B,A]$.

2) **Produit scalaire usuel**

Définition : On appelle produit scalaire usuel sur $\mathbb{R}^2$ l'application qui à tous vecteurs $U=(x,y)$ et $V=(x',y')$ de $\mathbb{R}^2$ associe le réel $<U,V>=xx'+yy'$.

Ce produit scalaire a été étudié dans le secondaire.

Propriétés :
- Pour tous vecteurs U et V de $\mathbb{R}^2$: $<U,V>=<V,U>$ (Symétrie).
- Bilinéarité : $U\mapsto<U,V>$ et $V\mapsto<U,V>$ sont des formes linéaires.
- Pour tout vecteur U de $\mathbb{R}^2$: $<U,U>\ \geq 0$.
- Pour tout vecteur U de $\mathbb{R}^2$: $<U,U>\ =0\Leftrightarrow U=0$.

Les démonstrations sont évidentes. La bilinéarité signifie :
- pour tout α réel et tous U_1, U_2 et V de $\mathbb{R}^2$: $<\alpha U_1+U_2,V>=\alpha<U_1,V>+<U_2,V>$.
- pour tout α réel et tous U, V_1 et V_2 de $\mathbb{R}^2$: $<\alpha U_1+U_2,V>=\alpha<U_1,V>+<U_2,V>$.

On dit que $(U,V)\mapsto<U,V>=xx'+yy'$ est une forme bilinéaire, définie et positive.

En fait, toute application de $\mathbb{R}^2$ dans $\mathbb{R}$ qui vérifie ces propriétés est appelée produit scalaire. Et $(U,V)\mapsto<U,V>=xx'+yy'$ n'en est qu'un exemple. En deuxième année, vous verrez d'autres produits scalaires.

Conséquence : (Inégalité de Cauchy-Schwarz)

Pour tous vecteurs U et V de $\mathbb{R}^2$: $\left|<U,V>\right|\leq\sqrt{<U,U><V,V>}$.

Il y a égalité si et seulement si U et V sont colinéaires.

Démonstration : $\forall\lambda\in\mathbb{R} \quad <\lambda U-V,\lambda U-V>\geq 0$.

Donc $\forall\lambda\in\mathbb{R} \quad \lambda^2<U,U>-2\lambda<U,V>+<V,V>\geq 0$.

Si $U\neq 0$, alors $<U,U>\neq 0$, donc c'est un polynôme du second degré de signe constant. Donc son discriminant est négatif ou nul.

Donc $4(<U,V>)^2\ -4<U,U><V,V>\leq 0$. Donc $(<U,V>)^2\ \leq<U,U><V,V>$.

Donc : $\left|<U,V>\right|\leq\sqrt{<U,U><V,V>}$.

Il y a égalité si et seulement si le discriminant est nul, donc si le polynôme admet une racine double, donc s'il existe un réel λ tel que $<\lambda U-V,\lambda U-V>=0$, donc tel que $\lambda U-V=0$, donc tel que $V=\lambda U$.

Si $U=0$, alors les deux membres sont nuls et l'inégalité est vérifiée pour tout V.

Il y a donc égalité si et seulement si $U=0$ ou $\exists\lambda\in\mathbb{R} \quad V=\lambda U$, donc si et seulement si U et V sont colinéaires.

3) **Norme euclidienne**

Définition : On appelle norme euclidienne sur $\mathbb{R}^2$ l'application qui à tout élément $U=(x,y)$ de $\mathbb{R}^2$ associe le réel : $\|U\|=\sqrt{<U,U>}=\sqrt{x^2+y^2}$.

Elle est bien sûr définie car pour tout U de $\mathbb{R}^2$: $<U,U>\ \geq 0$.

L'inégalité de Schwarz devient :

<u>Inégalité de Cauchy-Schwarz</u> : Pour tous U et V de $\mathbb{R}^2$: $|<U,V>| \leq \|U\| \times \|V\|$.

Des propriétés du produit scalaire, on peut déduire les propriétés suivantes :

<u>Propriétés</u> :
- Pour tout vecteur U de $\mathbb{R}^2$: $\|U\| \geq 0$ et $\|U\| = 0 \Leftrightarrow U = 0$.
- Pour tout vecteur U de $\mathbb{R}^2$: $\forall \lambda \in \mathbb{R} \quad \|\lambda U\| = |\lambda| \times \|U\|$.
- Pour tous vecteurs U et V de $\mathbb{R}^2$: $\|U+V\| \leq \|U\| + \|V\|$ (inégalité triangulaire).

L'inégalité triangulaire est conséquence de l'inégalité de Schwarz :
$(\|U+V\|)^2 = <U+V, U+V> = (\|U\|)^2 + 2<U,V> + (\|V\|)^2$.

Or $<U,V> \leq |<U,V>| \leq \|U\| \times \|V\|$. Donc : $(\|U+V\|)^2 \leq (\|U\| + \|V\|)^2$.

Les termes sont positifs. Donc : $\|U+V\| \leq \|U\| + \|V\|$.

4) **<u>Distance euclidienne</u>**

<u>Définition</u> : On appelle distance euclidienne sur $\mathbb{R}^2$ l'application qui à tous points $A = (x, y)$ et $B = (x', y')$ de $\mathbb{R}^2$ associe le réel :

$$d(A,B) = \|\overrightarrow{AB}\| = \|B - A\| = \sqrt{(x-x')^2 + (y-y')^2}.$$

Des propriétés de la norme, on peut déduire les propriétés suivantes :

<u>Propriétés</u> :
- Pour tous points A et B de $\mathbb{R}^2$: $d(A,B) \geq 0$.
- Pour tous points A et B de $\mathbb{R}^2$: $d(A,B) = 0 \Leftrightarrow A = B$.
- Pour tous points A et B de $\mathbb{R}^2$: $d(A,B) = d(B,A)$.
- Pour tous points A, B et C de $\mathbb{R}^2$: $d(A,C) \leq d(A,B) + d(B,C)$.

La troisième propriété vient du fait que $\overrightarrow{BA} = -\overrightarrow{AB}$, donc $\|\overrightarrow{BA}\| = \|\overrightarrow{AB}\|$.

La quatrième propriété vient de la relation de Chasles $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ et de l'inégalité triangulaire. D'ailleurs elle s'appelle aussi « inégalité triangulaire ».
On en déduit une autre propriété :

<u>Propriété</u> : Pour tous points A, B et C de $\mathbb{R}^2$: $|d(A,B) - d(B,C)| \leq d(A,C)$.

En effet : $d(A,B) \leq d(A,C) + d(C,B)$ et $d(B,C) \leq d(B,A) + d(A,C)$.

Donc : $d(A,B) - d(B,C) \leq d(A,C)$ et $d(B,C) - d(A,B) \leq d(A,C)$.

Donc : $|d(A,B) - d(B,C)| \leq d(A,C)$.

5) **<u>Boules</u>**

<u>Définition</u> : Soit A un point de $\mathbb{R}^2$ et un réel $r > 0$:
- L'ensemble $B(A,r) = \{M \in \mathbb{R}^2 / d(A,M) < r\}$ est appelé boule ouverte de centre A et de rayon r.
- L'ensemble $B_f(A,r) = \{M \in \mathbb{R}^2 / d(A,M) \leq r\}$ est appelé boule fermée de centre A et de rayon r.

Si $A = (a,b)$, on obtient l'ensemble des points $M = (x, y)$ de $\mathbb{R}^2$ tels que $(x-a)^2 + (y-b)^2 < r^2$ (ou $\leq r^2$), c'est-à-dire un disque ouvert (ou fermé) du plan.

<u>Propriétés</u> :

a) L'intersection de deux boules $B(A,r_1)$ et $B(A,r_2)$ de même centre A est la boule $B(A,r)$ de centre A et de rayon $r = \mathrm{Min}(r_1, r_2)$.

b) Si A_1 et A_2 sont deux points distincts de $\mathbb{R}^2$, il existe un réel $r > 0$ tel que $B(A_1,r) \cap B(A_2,r) = \varnothing$.

c) Pour tout point M d'une boule <u>ouverte</u> $B(A,r)$, il existe un réel $\varepsilon > 0$ tel que la boule ouverte $B(M,\varepsilon)$ soit entièrement contenue dans $B(A,r)$.

<u>Démonstration</u> : La propriété a) est évidente.

Pour la propriété b), il suffit de prendre $r = \dfrac{1}{3} d(A_1, A_2)$.

En effet, pour tout point $M \in B(A_1,r)$ et tout point $N \in B(A_2,r)$, on a : $d(A_1,A_2) \le d(A_1,M) + d(M,N) + d(N,A_2) \le 2r + d(M,N)$. Or $d(A_1,A_2) = 3r$.
Donc : $d(M,N) \ge r > 0$. Donc $M \ne N$. Donc $B(A_1,r) \cap B(A_2,r) = \varnothing$.

Pour la propriété c), il suffit de prendre $\varepsilon = \dfrac{1}{2}[r - d(A,M)]$ (>0 car $d(A,M) < r$).

Pour tout $N \in B(M,\varepsilon)$, on a : $d(A,N) \le d(A,M) + d(M,N) < d(A,M) + \varepsilon$.
Or $d(A,M) = r - 2\varepsilon$. Donc : $d(A,N) < r - \varepsilon < r$. Donc $N \in B(A,r)$.
On peut remarquer que cette propriété est fausse pour les boules fermées. En effet, si M appartient à la frontière de $B(A,r)$, c'est-à-dire si $d(A,M) = r$, toute boule de centre M et de rayon ε contient des points extérieurs à $B(A,r)$, par exemple le point N défini par $\overrightarrow{MN} = \dfrac{\varepsilon}{2r}\overrightarrow{AM}$ et donc n'est pas contenue dans $B(A,r)$.

6) <u>Parties ouvertes et fermées</u>

Par analogie avec la propriété c) précédente, on généralise la notion d'« ouvert ».

<u>Définition</u> : Une partie D de $\mathbb{R}^2$ est ouverte si $D = \varnothing$ ou si pour tout $M \in D$, il existe une boule ouverte de centre M contenue dans D.

C'est-à-dire : $D = \varnothing$ ou : $\forall M \in D \quad \exists \varepsilon \in]0,+\infty[\quad B(M,\varepsilon) \subset D$.

<u>Exemples</u> :

- et $\mathbb{R}^2$ sont des parties ouvertes de $\mathbb{R}^2$.
- Les boules ouvertes sont des parties ouvertes de $\mathbb{R}^2$.
- Les demi-plans $\{(x,y) \in \mathbb{R}^2 / ax + by + c > 0\}$ et $\{(x,y) \in \mathbb{R}^2 / ax + by + c < 0\}$ sont des parties ouvertes de $\mathbb{R}^2$. Montrons le pour $D = \{(x,y) \in \mathbb{R}^2 / ax + by + c > 0\}$.
  Soit $M \in D$, donc $ax_M + by_M + c > 0$. Soit $B(M,\varepsilon)$ une boule ouverte de centre M.
  Pour tout $N \in B(M,\varepsilon)$, on a : $|x_N - x_M| \le d(M,N) < \varepsilon$ et $|y_N - y_M| \le d(M,N) < \varepsilon$.
  Pour tout réel z, $|z| = Max(z,-z)$, donc $-z \le |z|$, donc $z \ge -|z|$.
  Donc : $(ax_N + by_N + c) - (ax_M + by_M + c) \ge -|a(x_N - x_M) + b(y_N - y_M)|$.
  Or : $|a(x_N - x_M) + b(y_N - y_M)| \le |a||x_N - x_M| + |b||y_N - y_M| < \varepsilon(|a| + |b|)$.
  Donc : $ax_N + by_N + c \ge (ax_M + by_M + c) - \varepsilon(|a| + |b|)$.
  Donc si $\varepsilon = \dfrac{ax_M + by_M + c}{2(|a| + |b|)}$, alors $ax_N + by_N + c \ge \dfrac{1}{2}(ax_M + by_M + c) > 0$.
  Donc il existe $\varepsilon > 0$ tel que tout $N \in B(M,\varepsilon)$ appartienne à D : $B(M,\varepsilon) \subset D$.
  Donc $D = \{(x,y) \in \mathbb{R}^2 / ax + by + c > 0\}$ est une partie ouverte de $\mathbb{R}^2$.

- Mais les boules fermées ne sont pas des parties ouvertes de $\mathbb{R}^2$.
  Par contre, si D est le complémentaire d'une boule fermée $B_f(A,r)$, c'est-à-dire si $D=\{M\in\mathbb{R}^2 / d(A,M)>r\}$, on peut montrer que D est une partie ouverte.
  En effet, pour tout $M\in D$ on peut poser $\varepsilon=\frac{1}{2}[d(A,M)-r]$ (donc >0).
  Alors : $\forall N\in B(M,\varepsilon)\quad d(A,M)\le d(A,N)+d(N,M)<d(A,N)+\varepsilon$.
  Donc $d(A,N)\ge d(A,M)-\varepsilon$. Or : $d(A,M)=r+2\varepsilon$. Donc : $d(A,N)\ge r+\varepsilon>r$.
  Donc $N\in D$. Donc $\forall M\in D\quad \exists\varepsilon\in]0,+\infty[\quad B(M,\varepsilon)\subset D$.
  Donc D est une partie ouverte de $\mathbb{R}^2$.

On en déduit une généralisation de la notion de « fermé ».

Définition : Une partie D de $\mathbb{R}^2$ est fermée si son complémentaire $\overline{D}=\mathbb{R}^2-D$ est une partie ouverte de $\mathbb{R}^2$.

Exemples :

- Les boules fermées sont des parties fermées de $\mathbb{R}^2$.
- $\varnothing$ et $\mathbb{R}^2$ sont des parties à la fois ouvertes et fermées de $\mathbb{R}^2$.
- Les demi-plans $\{(x,y)\in\mathbb{R}^2 / ax+by+c\ge 0\}$ et $\{(x,y)\in\mathbb{R}^2 / ax+by+c\le 0\}$ sont des parties fermées de $\mathbb{R}^2$.

Propriétés :
a) La réunion d'un nombre fini ou infini de parties ouvertes est une partie ouverte.
b) L'intersection d'un nombre fini de parties ouvertes est une partie ouverte.

Démonstration :

a) Soit $(D_i)_{i\in I}$ une famille d'ouverts de $\mathbb{R}^2$ et $D=\bigcup_{i\in I} D_i$. Si $D=\varnothing$, alors D est un ouvert. Sinon, soit $M\in D$. Donc il existe $i\in I$ tel que $M\in D_i$. Comme D_i est un ouvert de $\mathbb{R}^2$, il existe un réel $\varepsilon>0$ tel que $B(M,\varepsilon)\subset D_i$. Or $D_i\subset D$. Donc $B(M,\varepsilon)\subset D$. Donc D est un ouvert.

b) Soit $(D_i)_{1\le i\le n}$ une famille d'ouverts de $\mathbb{R}^2$ et $D=\bigcap_{i=1}^{n} D_i$. Soit $M\in D$. Donc pour tout $i\in[\![1,n]\!]$, on a $M\in D_i$. Comme, pour tout $i\in[\![1,n]\!]$, D_i est un ouvert de $\mathbb{R}^2$, il existe un réel $\varepsilon_i>0$ tel que $B(M,\varepsilon_i)\subset D_i$. Donc $\bigcap_{i=1}^{n} B(M,\varepsilon_i)\subset D$. Or on a vu que $\bigcap_{i=1}^{n} B(M,\varepsilon_i)=B(M,\varepsilon)$ si $\varepsilon=\mathrm{Inf}(\varepsilon_1,\dots,\varepsilon_n)$. Comme les ε_i sont en nombre fini, ε est l'un des ε_i et donc $\varepsilon>0$. Donc $B(M,\varepsilon)\subset D$. Donc D est un ouvert.

Remarque : Par contre, dans le cas d'une intersection d'un nombre infini d'ouverts, le minimum ε ne serait pas forcément atteint et on pourrait avoir $\varepsilon=0$. Donc une intersection d'un nombre infini d'ouverts n'est pas forcément un ouvert.

Conséquences :
a) La réunion d'un nombre fini de parties fermées est une partie fermée.
b) L'intersection d'un nombre fini ou infini de parties fermées est une partie fermée.

C'est évident par passage au complémentaire.

7) Parties bornées

Définition : Une partie D de $\mathbb{R}^2$ est bornée s'il existe un réel K tel que $\forall M\in D\quad \|M\|\le K$.

Or $\|M\| = d(O,M)$, donc cela signifie qu'il existe une boule $B_f(O,K)$ contenant *D*. Il n'y a pas unicité de *K*, donc la boule peut être ouverte ou fermée.
Plus généralement, s'il existe une boule $B(A,r)$ telle que $D \subset B(A,r)$, alors pour tout point *M* de *D*, on a $d(O,M) \leq d(O,A) + d(A,M)$, donc $\|M\| \leq d(O,A) + r$.

<u>Théorème</u> : Une partie *D* de $\mathbb{R}^2$ est bornée si et seulement si il existe une boule contenant *D*.

<u>Exemples</u> :

- Les boules ouvertes ou fermées sont des parties bornées de $\mathbb{R}^2$.
- Tout segment $[A,B]$ est une partie bornée de $\mathbb{R}^2$.
- Les droites et les demi-plans ne sont pas des parties bornées de $\mathbb{R}^2$.

<u>Propriétés</u> :
a) L'intersection d'un nombre fini ou infini de parties bornées est une partie bornée.
b) La réunion d'un nombre <u>fini</u> de parties bornées est une partie bornée.

L'intersection est contenue dans n'importe quelle partie, donc dans une boule.

Si $D = \bigcup_{i=1}^{n} D_i$ et si pour tout $i \in [\![1,n]\!]$, les parties D_i sont bornées, donc s'il existe $r_i > 0$ tel que $D_i \subset B(O,r_i)$, alors $D \subset B(O,r)$ avec $r = \text{Max}(r_1,...,r_n)$.

8) Parties convexes

<u>Définition</u> : Une partie *D* de $\mathbb{R}^2$ est convexe si : $\forall (M,N) \in D^2 \quad [M,N] \subset D$.
Cela revient à dire que : $\forall (M,N) \in D^2 \quad \forall t \in [0,1] \quad (1-t)M + tN \in D$.

<u>Exemples</u> :

- De manière évidente, les segments et les droites sont convexes.
- Les boules ouvertes ou fermées sont convexes.
  En effet, si $D = B(A,r)$ et $(M,N) \in D^2$, on a : $\|\overrightarrow{AM}\| < r$ et $\|\overrightarrow{AN}\| < r$.
  Soit $t \in [0,1]$ et $P = (1-t)M + tN$. Donc : $\overrightarrow{AP} = (1-t)\overrightarrow{AM} + t\overrightarrow{AN}$.
  Donc : $\|\overrightarrow{AP}\| \leq (1-t)\|\overrightarrow{AM}\| + t\|\overrightarrow{AN}\| < r$. Donc $P \in D$.
- Les demi-plans ouverts ou fermés sont convexes.
  Soit $D = \{(x,y) \in \mathbb{R}^2 / ax + by + c > 0\}$ et $(M,N) \in D^2$, on a $ax_M + by_M + c > 0$ et $ax_N + by_N + c > 0$. Soit $t \in [0,1]$ et $P = (1-t)M + tN$ donc $\begin{cases} x_P = (1-t)x_M + tx_N \\ y_P = (1-t)y_M + ty_N \end{cases}$
  Donc $ax_P + by_P + c = (1-t)(ax_M + by_M + c) + t(ax_N + by_N + c)$.
  Si $t \in]0,1[$, tous les termes sont strictement positifs, donc $ax_P + by_P + c > 0$ et donc $P \in D$. Si $t = 0$, alors $P = M$ donc $P \in D$, et si $t = 1$, $P = N$ donc $P \in D$.
- Si *f* est une fonction convexe, alors $D = \{(x,y) \in \mathbb{R}^2 / y \geq f(x)\}$ est convexe.
  En effet, si $(M,N) \in D^2$, on a : $y_M \geq f(x_M)$ et $y_N \geq f(x_N)$.
  Soit $t \in [0,1]$ et $P = (1-t)M + tN$. Donc $\begin{cases} x_P = (1-t)x_M + tx_N \\ y_P = (1-t)y_M + ty_N \end{cases}$
  Or la fonction est convexe, donc : $f(x_P) \leq (1-t)f(x_M) + tf(x_N)$.
  Donc $f(x_P) \leq (1-t)y_M + ty_N$ car $\begin{cases} 1-t \geq 0 \\ t \geq 0 \end{cases}$. Donc $f(x_P) \leq y_P$. Donc $P \in D$.

Il y a même équivalence entre « la fonction f est convexe » et « l'ensemble $D = \{(x, y) \in \mathbb{R}^2 / y \geq f(x)\}$ est convexe ».

Propriété :
L'intersection d'un nombre fini ou infini de parties bornées est une partie bornée.

Mais c'est faux pour la réunion.

II – Fonctions définies sur $\mathbb{R}^2$

1) Graphe d'une fonction

Définition : Si f est une fonction définie sur une partie D non vide de $\mathbb{R}^2$, on appelle graphe de f l'ensemble $\{(x, y, z) \in \mathbb{R}^3 / (x, y) \in D \text{ et } z = f(x, y)\}$.

Le graphe d'une fonction de 2 variables est donc une partie de $\mathbb{R}^3$.

Exemple 1 : Si $f(x, y) = ax + by + c$ avec $(a,b) \neq (0,0)$, son graphe est un plan de $\mathbb{R}^3$.

Exemple 2 : Si $f(x, y) = x^2 + y^2$, son graphe est un paraboloïde de $\mathbb{R}^3$.

Exemple 3 : Si $f(x, y) = x^2 - y^2$, son graphe est un hyperboloïde de $\mathbb{R}^3$.

2) Courbes de niveau

L'exemple le plus courant est celui des cartes IGN où l'on trouve des lignes joignant les points de même altitude en montagne.

Définition : Etant donné un réel k, on appelle courbe de niveau k d'une fonction f de deux variables l'ensemble $\mathscr{L}_k = \{(x, y) \in D_f \,/\, f(x, y) = k\}$.

Ces courbes de niveau sont représentées dans le plan.
Bien sûr, deux courbes de niveaux différents ne se coupent pas.
Et certaines courbes de niveau peuvent être vides.

Exemple 1 : Si $f(x, y) = ax + by + c$, alors $\mathscr{L}_k = \{(x, y) \in \mathbb{R}^2 / ax + by + c - k = 0\}$. On obtient une famille de droites parallèles si $(a,b) \neq (0,0)$.

Exemple 2 : Si $f(x, y) = x^2 + y^2$, alors $\mathscr{L}_k = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 = k\}$. On obtient une famille de cercles concentriques pour les réels $k \geq 0$ et $\varnothing$ si $k < 0$.

Exemple 3 : Si $f(x, y) = xy$, alors $\mathscr{L}_k = \{(x, y) \in \mathbb{R}^2 / xy = k\}$on obtient une famille d'hyperboles équilatères.

Exemple 4 : Si $f(x, y) = \ln(x - y + 2) - \ln(x + y)$, alors $(x, y) \in D_f \Leftrightarrow \begin{cases} x - y + 2 > 0 \\ x + y > 0 \end{cases}$

et $\mathscr{L}_k = \left\{(x, y) \in \mathbb{R}^2 / y = \dfrac{x(1 - e^k) + 2}{1 + e^k}\right\}$. On obtient encore une famille de droites.

Elles ne sont pas parallèles, mais passent toutes par le point $(-1,1)$ qui n'appartient pas à l'ensemble de définition.

III – Continuité et limites sur $\mathbb{R}^2$

1) Limite en un point

Une fonction f est définie au voisinage d'un point $A \in \mathbb{R}^2$ si pour tout $r > 0$, la boule $B(A, r)$ contient au moins un point de l'ensemble de définition D_f. On élimine ainsi les points isolés. La fonction peut ou non être définie en A.

Pour tout $M = (x, y)$, on notera $f(x, y) = f(M)$ pour simplifier.

Définition : Soit f une fonction définie au voisinage de $A \in \mathbb{R}^2$. On dit que la fonction f admet une limite ℓ en A si : $\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall M \in B(A, \alpha) \cap D_f \quad |f(M) - \ell| \leq \varepsilon$.

C'est une extension de la notion de limite dans $\mathbb{R}$, la distance remplaçant la valeur absolue (qui est la distance dans $\mathbb{R}$).

Les propriétés sont analogues à celles des limites de fonctions d'une variable.

<u>Théorème</u> : Si une fonction f admet une limite en A, cette limite est unique.

<u>Démonstration</u> : On fait un raisonnement par l'absurde.

Supposons que f admette deux limites distinctes ℓ_1 et ℓ_2 en A.

Puisque $\ell_1 \neq \ell_2$, $\varepsilon = \frac{1}{3}|\ell_1 - \ell_2|$ est strictement positif.

Donc : $\exists \alpha_1 > 0 \quad \forall M \in B(A,\alpha_1) \cap D \quad |f(M) - \ell_1| \leq \varepsilon$.

Et : $\exists \alpha_2 > 0 \quad \forall M \in B(A,\alpha_2) \cap D_f \quad |f(M) - \ell_2| \leq \varepsilon$.

Si $\alpha = \text{Min}(\alpha_1, \alpha_2)$, $B(A,\alpha) = B(A,\alpha_1) \cap B(A,\alpha_2)$, et donc :

$\forall M \in B(u_0,\alpha) \cap D \quad |f(M) - \ell_1| \leq \varepsilon$ et $|f(M) - \ell_2| \leq \varepsilon$.

Or, pour tout M : $|\ell_1 - \ell_2| = |[\ell_1 - f(M)] + [f(M) - \ell_2]|$ et $|a+b| \leq |a| + |b|$.

Donc : $|\ell_1 - \ell_2| \leq 2\varepsilon$, ce qui est absurde puisque $|\ell_1 - \ell_2| = 3\varepsilon$ et $\varepsilon > 0$.

Donc l'hypothèse $\ell_1 \neq \ell_2$ est fausse. Il y a unicité de la limite.

On peut remarquer que c'est la même démonstration que pour les fonctions d'une variable en remplaçant les intervalles par des boules.

<u>Notation</u> : On note $\lim\limits_{M \to A} f(M) = \ell$ ou $\lim\limits_{A} f(M) = \ell$.

Les théorèmes d'opérations sur les limites finies sont les mêmes que pour les fonctions d'une variable. Les démonstrations sont analogues en remplaçant les intervalles par des boules.

<u>Théorème</u> : Soient f et g deux fonction définie sur une partie D de $\mathbb{R}^2$.
Si $\lim\limits_{M \to A} f(M) = \ell$ et $\lim\limits_{M \to A} g(M) = \ell'$, alors : $\lim\limits_{M \to A} [f(M) + g(M)] = \ell + \ell'$

$$\lim_{M \to A} f(M)g(M) = \ell\ell' \qquad \lim_{M \to A} \frac{f(M)}{g(M)} = \frac{\ell}{\ell'} \text{ si } \ell' \neq 0$$

Pour la composition des limites, il s'agit dans un sens de composer une fonction de deux variables avec une fonction d'une variable, et dans l'autre avec deux fonctions d'une variable.

<u>Théorème</u> : Si $\lim\limits_{M \to A} f(M) = \ell$ et si φ est une fonction définie sur une partie Δ de $\mathbb{R}$ telle que $\Delta \cap f(D)$ contienne au moins un intervalle centré en ℓ et privé de ℓ, et telle que $\lim\limits_{x \to \ell} \varphi(x) = \ell'$. Alors : $\lim\limits_{M \to A} (\varphi \circ f)(M) = \ell'$.

Là aussi, la démonstration est analogue à celle du cas d'une variable.

<u>Démonstration</u> : Soit $\varepsilon > 0$.

$\lim\limits_{x \to \ell} \varphi(x) = \ell'$. Donc : $\exists \eta > 0 \quad \forall x \in [\ell - \eta, \ell + \eta] \cap \Delta \quad |\varphi(x) - \ell'| \leq \varepsilon$

$\lim\limits_{M \to A} f(M) = \ell$. Donc : $\exists \alpha > 0 \quad \forall M \in B(A,\alpha) \cap D \quad |f(M) - \ell| \leq \eta$.

Or $|f(M) - \ell| \leq \eta \Leftrightarrow f(M) \in [\ell - \eta, \ell + \eta]$.

Donc : $\exists \alpha > 0 \quad \forall M \in B(A,\alpha) \cap D \quad |(\varphi \circ f)(M) - \ell'| \leq \varepsilon$.

Ceci est vrai pour tout $\varepsilon > 0$. Donc $\lim\limits_{M \to A} (\varphi \circ f)(M) = \ell'$.

<u>Exemple</u> : $g(x,y) = \dfrac{\ln(1 + x^2 + y^2)}{x^2 + y^2}$ en $(0,0)$.

On remarque que $g = \varphi \circ f$ avec $f(x,y) = x^2 + y^2$ et $\varphi(x) = \dfrac{\ln(1+x)}{x}$.

Ici : $D = \mathbb{R}^2$, $f(D) = [0,+\infty[$, $\Delta =]-1,0[\cup]0,+\infty[$ et $A = (0,0)$.

On a : $\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall M \in B(A,\alpha) \cap D \quad |f(M)| \leq \varepsilon$.

En effet, il suffit de prendre $\alpha = \sqrt{\varepsilon}$ puisque $M \in B(A,\alpha) \Leftrightarrow \sqrt{x^2+y^2} \leq \alpha$.

Donc $\lim\limits_{M \to A} f(M) = 0$. D'autre part : $\lim\limits_{x\to 0} \varphi(x) = \lim\limits_{x \to 0} \dfrac{\ln(1+x)}{x} = 1$.

Donc par composition des limites : $\lim\limits_{M \to A} g(M) = 1$.

<u>Théorème</u> : Soient φ et ψ deux fonctions définies sur une partie Δ de $\mathbb{R}$ contenant au moins un intervalle centré en t_0 et privé de t_0, et telles que $\lim\limits_{t \to t_0} \varphi(t) = x_0$ et $\lim\limits_{t \to t_0} \psi(t) = y_0$. Si $A = (x_0, y_0)$ et $\lim\limits_{M \to A} f(M) = \ell$, alors : $\lim\limits_{t \to t_0} f[\varphi(t), \psi(t)] = \ell$.

<u>Démonstration</u> : Soit $\varepsilon > 0$.

$\lim\limits_{M \to A} f(M) = \ell$. Donc : $\exists \eta > 0 \quad \forall M \in B(A,\eta) \cap D \quad |f(M) - \ell| \leq \varepsilon$.

$\lim\limits_{t \to t_0} \varphi(t) = x_0$. Donc : $\exists \alpha_1 > 0 \quad \forall t \in [t_0 - \alpha_1, t_0 + \alpha_1] \cap \Delta \quad |\varphi(t) - x_0| \leq \dfrac{\eta}{\sqrt{2}}$.

$\lim\limits_{t \to t_0} \psi(t) = y_0$. Donc : $\exists \alpha_2 > 0 \quad \forall t \in [t_0 - \alpha_2, t_0 + \alpha_2] \cap \Delta \quad |\psi(t) - y_0| \leq \dfrac{\eta}{\sqrt{2}}$.

Soit $\alpha = \mathrm{Min}(\alpha_1, \alpha_2)$ qui est strictement positif.

Alors $\forall t \in [t_0 - \alpha, t_0 + \alpha] \cap \Delta \quad |\varphi(t) - x_0| \leq \dfrac{\eta}{\sqrt{2}}$ et $|\psi(t) - y_0| \leq \dfrac{\eta}{\sqrt{2}}$

Donc $\forall t \in [t_0 - \alpha, t_0 + \alpha] \cap \Delta \quad \sqrt{[\varphi(t) - x_0]^2 + [\psi(t) - y_0]^2} \leq \eta$.

Cela signifie que, si $x = \varphi(t)$ et $y = \psi(t)$, le point (x,y) appartient à $B(A,\eta) \cap D$.

Donc : $\exists \alpha > 0 \quad \forall t \in [t_0 - \alpha, t_0 + \alpha] \cap \Delta \quad |f[\varphi(t), \psi(t)] - \ell| \leq \varepsilon$.

Ceci est vrai pour tout $\varepsilon > 0$. Donc $\lim\limits_{t \to t_0} f[\varphi(t), \psi(t)] = \ell$.

Les théorèmes « de comparaison » s'étendent aux fonctions de plusieurs variables.

<u>Théorème</u> : Soient *f* et *g* deux fonctions définies sur une partie *D* de $\mathbb{R}^2$, telles que pour tout $M \in D$, on a $f(M) \leq g(M)$. Si $\lim\limits_{M \to A} f(M) = \ell$ et $\lim\limits_{M \to A} g(M) = \ell'$, alors : $\ell \leq \ell'$.

L'inégalité sur les limites reste large même si l'inégalité sur les fonctions est stricte.

<u>Théorème (d'encadrement)</u> : Soient *f*, *g* et *h* trois fonctions définies sur une partie *D* de $\mathbb{R}^2$, telles que $\forall M \in D \quad f(M) \leq g(M) \leq h(M)$. Si $\lim\limits_{M \to A} f(M) = \lim\limits_{M \to A} h(M) = \ell$, alors la fonction *g* admet une limite en *A* et $\lim\limits_{M \to A} g(M) = \ell$.

Les démonstrations sont analogues à celles des fonctions d'une variable.

2) <u>Continuité en un point</u>

<u>Définition</u> : Soit *D* une partie ouverte non vide de $\mathbb{R}^2$ et *A* un point de *D*. Une fonction *f* définie sur *D* est continue en *A* si $\lim\limits_{M \to A} f(M) = f(A)$, c'est-à-dire si :

$$\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall M \in B(A,\alpha) \cap D \quad |f(M) - f(A)| \leq \varepsilon.$$

<u>*Exemple*</u> : $f(x,y) = \dfrac{xy}{\sqrt{x^2+y^2}}$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$. Soit $A = (0,0)$.

$\forall(x,y) \quad |f(x,y)| = \dfrac{|x||y|}{\sqrt{x^2+y^2}}$. Or $\begin{cases} |x| \le \sqrt{x^2+y^2} \\ |y| \le \sqrt{x^2+y^2} \end{cases}$ donc $|f(x,y)| \le \sqrt{x^2+y^2}$.

Donc : $\forall \varepsilon > 0 \quad \exists \alpha = \varepsilon \quad \forall M \in B(A,\alpha) \cap D \quad |f(M) - f(A)| \le \varepsilon$.

Donc la fonction *f* est continue en $A = (0,0)$.

En réalité, on peut remarquer qu'il suffit de savoir que *f* admet une limite ℓ en *A*, donc que : $\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall M \in B(A,\alpha) \cap D \quad |f(M) - \ell| \le \varepsilon$.

En effet, pour tout α, $A \in B(A,\alpha) \cap D$ et donc pour tout $\varepsilon > 0$: $|f(A) - \ell| \le \varepsilon$.

Donc $|f(A) - \ell|$ est un nombre positif ou nul qui inférieur ou égal à tout élément de $]0,+\infty[$. Donc $|f(A) - \ell| = 0$. Donc $\ell = f(A)$.

Si la fonction *f* est définie en *A*, sa seule limite possible est $\ell = f(A)$.

<u>Définition</u> : Soit *D* une partie ouverte non vide de $\mathbb{R}^2$ ne contenant pas *A* mais contenant au moins une boule ouverte de centre *A* privée de *A*. Soit *f* une fonction définie sur *D* qui admet en *A* une limite $\lim_{M \to A} f(M) = \ell$. La fonction $\tilde{f}$ définie sur $D \cup \{A\}$ par : $\forall M \in D \quad \tilde{f}(M) = f(M)$ et $\tilde{f}(A) = \ell$ est appelée prolongement par continuité de *f* en *A*.

Par construction, elle est continue en *A*.

<u>Théorème</u> : Si *f* et *g* sont deux fonctions définies sur une partie *D* ouverte non vide de $\mathbb{R}^2$ et si elles sont continues en un point *A* de *D*, alors :
- leur somme $f + g$ est continue en *A*.
- leur produit fg est continue en *A*.
- leur quotient $\dfrac{f}{g}$ est continue en *A* à condition que $g(A) \ne 0$.

Ces propriétés découlent directement des opérations sur les limites.

<u>Théorème</u> : Soit *f* une fonction définie sur une partie *D* ouverte non vide de $\mathbb{R}^2$ et continue en un point *A* de *D*. Soit φ une fonction définie sur $f(D)$ et continue en $f(A)$. Alors la fonction $\varphi \circ f$ est continue en *A*.

<u>Théorème</u> : Si φ et ψ sont deux fonctions définies sur une partie Δ de $\mathbb{R}$ et continues en $t_0 \in \Delta$, et si *f* est une fonction définie sur une partie *D* contenant $\varphi(\Delta) \times \psi(\Delta)$ et continue en $A = (\varphi(t_0), \psi(t_0))$, alors la fonction $t \mapsto f[\varphi(t), \psi(t)]$ est continue en t_0.

Ces deux théorèmes correspondent aux théorèmes de composition des limites.

<u>Conséquence</u> : Si *f* est continue en un point *A* d'un ouvert de $\mathbb{R}^2$, alors pour tout vecteur *U* de $\mathbb{R}^2$, la fonction f_U définie par $f_U(t) = f(A + tU)$ est continue en 0.

En effet, si $U = (\alpha, \beta)$, alors $f_U(t) = f(x_A + t\alpha, y_A + t\beta)$. C'est donc la composée de *f* avec deux fonctions affines, donc continues.

Cela revient à étudier la continuité de la restriction de *f* à la droite $d_{A,U}$.

Donc inversement, si l'on trouve une droite passant par *A* où la restriction de *f* n'est pas continue en *A*, alors la fonction *f* n'est pas continue en *A*.

<u>*Exemple*</u> : $f(x,y) = \dfrac{2xy}{x^2+y^2}$ si $(x,y) \ne (0,0)$ et $f(0,0) = 0$. Soit $A = (0,0)$.

On remarque que : $\forall x \in \mathbb{R}^* \quad f(x,x) = 1$. Donc $\lim_{x \to 0} f(x,x) \ne f(0,0)$.

Donc *f* n'est pas continue en $A = (0,0)$.

<u>Remarque</u> : Attention ! La continuité de <u>toutes</u> les fonctions f_U ne prouve pas la continuité de *f*.

<u>*Exemple*</u> : $f(x,y)=1$ si $0<|y|<x^2$ et $f(x,y)=0$ sinon. Soit $A=(0,0)$.

On a donc $f(x,y)=0$ si $y=0$ ou si $x^2\leq|y|$. En particulier $f(A)=0$.

Soit $U=(\alpha,\beta)$. Donc : $\forall t\in\mathbb{R} \quad f_U(t)=f(t\alpha,t\beta)$.

Si $\beta=0$ ou $\alpha=0$, alors : $\forall t\in\mathbb{R} \quad f_U(t)=0$. Donc f_U est continue en *A*.

Si $\beta\neq 0$ et $\alpha\neq 0$, alors : $f_U(t)=0$ si $t=0$ ou $t^2\alpha^2\leq|t||\beta|$, donc si $|t|\leq\dfrac{|\beta|}{\alpha^2}$.

Donc $\lim\limits_{t\to 0} f_U(t)=0$. Donc f_U est continue en *A*.

Donc toutes les fonctions f_U sont continues en *A*. Et pourtant dans toute boule $B(A,r)$, on trouve des points $M\left(x,\dfrac{1}{2}x^2\right)$ pour lesquels $f(M)=1$ même quand *x* tend vers 0. Donc *f* n'est pas continue en *A*.

3) Continuité sur une partie

<u>Définition</u> : Soit *D* une partie ouverte non vide de $\mathbb{R}^2$. Une fonction *f* définie sur *D* est continue sur *D* si elle est continue en tout point de *D*.

<u>*Exemple*</u> : $p(x,y)=x$. Montrons qu'elle est continue en tout point $A=(x_0,y_0)$ de $\mathbb{R}^2$.

$|p(M)-p(A)|=|x-x_0|$, donc $|p(M)-p(A)|\leq d(A,M)$.

Donc : $\forall\varepsilon>0 \quad \exists\alpha=\varepsilon \quad \forall M\in B(A,\alpha)\cap D \quad |p(M)-p(A)|\leq\varepsilon$.

Donc la fonction *p* est continue en *A*.

La démonstration est analogue pour la fonction $(x,y)\mapsto y$.

Les projections $(x,y)\mapsto x$ et $(x,y)\mapsto y$ sont continues sur $\mathbb{R}^2$.

Des opérations sur les fonctions continues, on en déduit :

Les fonctions polynômes sont continues sur $\mathbb{R}^2$.
Les fonctions rationnelles sont continues sur leur ensemble de définition.
Si *f* est continue, alors $|f|$, $\sqrt{f}$, $\sin f$, $\cos f$, $\ln f$ et e^f sont continues sur leur ensemble de définition.

Comme pour les fonctions d'une variable, il y a des propriétés des images de parties :

<u>Théorème</u> : Si *f* est une fonction continue sur une partie ouverte de $\mathbb{R}^2$, l'image réciproque d'un intervalle ouvert (respectivement fermé) de $\mathbb{R}$ est une partie ouverte (respectivement fermée) de $\mathbb{R}^2$.

<u>Démonstration</u> : Soit *f* une fonction continue sur un ouvert *D* de de $\mathbb{R}^2$.

- Soit *I* un intervalle ouvert de $\mathbb{R}$ et $A\in f^{-1}(I)$. Donc $f(A)\in I$.
  Comme *I* est ouvert, il existe $\varepsilon>0$ tel que $]f(A)-\varepsilon,f(A)+\varepsilon[\subset I$. Or *f* est continue en *A*. Donc il existe $\alpha>0$ tel que $\forall M\in B(A,\alpha)\cap D \quad |f(M)-f(A)|<\varepsilon$.
  Donc il existe $\alpha>0$ tel que $\forall M\in B(A,\alpha)\cap D \quad f(M)\in I$, donc $M\in f^{-1}(I)$.
  Donc il existe $\alpha>0$ tel que $B(A,\alpha)\cap D\subset f^{-1}(I)$.
  De plus *D* et $B(A,\alpha)$ sont des ouverts, donc $B(A,\alpha)\cap D$ est un ouvert qui contient *A*, donc qui contient au moins une boule $B(A,r)$.
  Donc pour tout *A*, il existe $r>0$ tel que $B(A,r)\subset f^{-1}(I)$. Donc $f^{-1}(I)$ est ouvert.
- Soit $I=[a,b]$ un intervalle fermé de $\mathbb{R}$ et $\Delta=\overline{f^{-1}(I)}$. Soit $M\in D$.

Le point M appartient à Δ si et seulement si $M \notin f^{-1}(I)$, donc si $f(M) \notin I$, donc si $f(M) \in]-\infty, a[\cup]b, +\infty[$.

Donc $\Delta = f^{-1}(]-\infty, a[\cup]b, +\infty[) = f^{-1}(]-\infty, a[) \cup f^{-1}(]b, +\infty[)$. C'est la réunion de deux ouverts d'après ce qui précède, donc c'est un ouvert. Donc $f^{-1}(I)$ est fermé.

Remarque : On peut ainsi étudier si les demi-plans sont ouverts ou fermés plus facilement que ce que l'on a vu.

On a aussi un théorème analogue à celui des fonctions continues sur un segment $[a, b]$.

Théorème : Toute fonction continue sur une partie fermée et bornée de $\mathbb{R}^2$ admet un maximum et un minimum.

IV – Calcul différentiel sur $\mathbb{R}^2$

1) Dérivées partielles d'ordre 1

Définition : Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie sur une partie D ouverte et non vide, et soit $A = (x_0, y_0)$ appartenant à D. On appelle applications partielles de f en A les deux applications $f_1 : x \mapsto f(x, y_0)$ et $f_2 : x \mapsto f(x_0, y)$ définies respectivement sur $D_1 = \{x \in \mathbb{R} / (x, y_0) \in D\}$ et $D_2 = \{y \in \mathbb{R} / (x_0, y) \in D\}$.

On peut remarquer que, puisque D est un ouvert, f_1 est définie au voisinage de x_0 et f_2 au voisinage de y_0.

Définition : Si la fonction $f_1 : x \mapsto f(x, y_0)$ est dérivable en x_0, le nombre dérivé de f_1 en x_0 est appelé dérivée partielle d'ordre 1 de f par rapport à x en A et se note $\frac{\partial f}{\partial x}(A)$. Donc : $\frac{\partial f}{\partial x}(A) = \lim\limits_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$.

Si la fonction $f_2 : x \mapsto f(x_0, y)$ est dérivable en y_0, le nombre dérivé de f_2 en y_0 est appelé dérivée partielle d'ordre 1 de f par rapport à x en A et se note $\frac{\partial f}{\partial y}(A)$.

Donc : $\frac{\partial f}{\partial y}(A) = \lim\limits_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}$.

Exemple : $f(x, y) = x^2 + 3xy + 2y^2$ en $O = (0,0)$ et en $A = (a, b)$.

$$\frac{\partial f}{\partial x}(O) = \lim_{x \to 0} \frac{f(x,0) - f(0,0)}{x - 0} = \lim_{x \to 0} \frac{x^2}{x} = \lim_{x \to 0} x = 0.$$

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

$$\frac{\partial f}{\partial x}(A) = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x - a} = \lim_{x \to a} \frac{x^2 + 3bx - a^2 - 3ab}{x - a} = \lim_{x \to a}(x + a + 3b) = 2a + 3b.$$

$$\frac{\partial f}{\partial y}(A) = \lim_{y \to b} \frac{f(a, y) - f(a,b)}{y - b} = \lim_{y \to b} \frac{3ay + 2y^2 - 3ab - 2b^2}{y - b} = \lim_{y \to b}(3a + 2y + 2b) = 3a + 4b.$$

Définition : Si la fonction f admet en tout point M d'un ouvert D de $\mathbb{R}^2$ des dérivées partielles d'ordre 1, on appelle fonction dérivée partielle d'ordre 1 de f par rapport à x la fonction $\frac{\partial f}{\partial x} : M \mapsto \frac{\partial f}{\partial x}(M)$ et fonction dérivée partielle d'ordre 1 de f par rapport à y la fonction $\frac{\partial f}{\partial y} : M \mapsto \frac{\partial f}{\partial y}(M)$.

Cela revient à considérer l'une des variables comme paramètre et à dériver par rapport à l'autre.

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

C'est une fonction définie sur $\mathbb{R}^2$ qui est un ouvert non vide.

On dérive par rapport à x en considérant y comme constant : $\dfrac{\partial f}{\partial x}(x,y) = 3x^2 - 4xy$.

De la même manière on dérive par rapport à y : $\dfrac{\partial f}{\partial y}(x,y) = 3y^2 - 2x^2 + 1$.

Définition : Si la fonction f admet en un point A d'un ouvert D de $\mathbb{R}^2$ des dérivées partielles d'ordre 1, on appelle gradient de f en A, noté $\nabla f(A)$ ou ∇f_A, le vecteur de $\mathbb{R}^2$ défini par : $\nabla f(A) = \nabla f_A = \left(\dfrac{\partial f}{\partial x}(A), \dfrac{\partial f}{\partial y}(A)\right)$ (on lit « nabla de f en A »)

Dans l'exemple précédent, si $A = (1,1)$: $\nabla f(A) = (-1,2)$.

On verra plus tard une interprétation de ce vecteur.

Attention ! Une fonction peut admettre des dérivées partielles d'ordre 1 en un point sans être continue en ce point.

Exemple : $f(x,y) = \dfrac{2xy}{x^2 + y^2}$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$.

On a vu qu'elle n'est pas continue en $(0,0)$.

Pourtant les fonctions $f_1 : x \mapsto f(x,0) = 0$ et $f_2 : y \mapsto f(0,y) = 0$ sont dérivables en 0, donc la fonction f admet des dérivées partielles en $(0,0)$: $\dfrac{\partial f}{\partial x}(O) = \dfrac{\partial f}{\partial y}(O) = 0$.

2) Opérations sur les dérivées partielles

Les opérations algébriques sur les dérivées partielles sont analogues à celles des fonctions d'une variable avec des démonstrations identiques.

Théorème : Si f et g sont deux fonctions définies sur un ouvert non vide D de $\mathbb{R}^2$ qui admettent des dérivées partielles d'ordre 1 par rapport à x et y, et si $\lambda \in \mathbb{R}$, alors :

- $f + g$ admet des dérivées partielles d'ordre 1 par rapport à x et y :

$$\frac{\partial(f+g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \qquad \frac{\partial(f+g)}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial g}{\partial y}$$

- λf admet des dérivées partielles d'ordre 1 par rapport à x et y :

$$\frac{\partial(\lambda f)}{\partial x} = \lambda \frac{\partial f}{\partial x} \qquad \frac{\partial(\lambda f)}{\partial y} = \lambda \frac{\partial f}{\partial y}$$

- fg admet des dérivées partielles d'ordre 1 par rapport à x et y :

$$\frac{\partial(fg)}{\partial x} = \frac{\partial f}{\partial x} g + f \frac{\partial g}{\partial x} \qquad \frac{\partial(fg)}{\partial y} = \frac{\partial f}{\partial y} g + f \frac{\partial g}{\partial y}$$

- $\dfrac{f}{g}$ admet des dérivées partielles d'ordre 1 par rapport à x et y si g ne s'annule pas :

$$\frac{\partial\left(\dfrac{f}{g}\right)}{\partial x} = \frac{\dfrac{\partial f}{\partial x} g - f \dfrac{\partial g}{\partial x}}{g^2} \qquad \frac{\partial\left(\dfrac{f}{g}\right)}{\partial y} = \frac{\dfrac{\partial f}{\partial y} g - f \dfrac{\partial g}{\partial y}}{g^2}$$

Les démonstrations sont analogues à celles des fonctions d'une variable.

Conséquence : Les fonctions polynômes admettent des dérivées partielles d'ordre 1 en tout point de $\mathbb{R}^2$ et les fonctions rationnelles admettent des dérivées partielles d'ordre 1 en tout point de leur ensemble de définition.

Comme pour la continuité, il y a deux sens de composition, mais ici, on ne considère que la composition à gauche.

Théorème : Si f est une fonction définie sur un ouvert non vide D de $\mathbb{R}^2$ qui admet des dérivées partielles d'ordre 1 par rapport à x et y, et si g est une fonction d'une variable dérivable sur $f(D)$, alors la fonction $g \circ f$ admet des dérivées partielles d'ordre 1 par rapport à x et y : $\dfrac{\partial(g \circ f)}{\partial x} = (g' \circ f) \times \dfrac{\partial f}{\partial x}$ et $\dfrac{\partial(g \circ f)}{\partial y} = (g' \circ f) \times \dfrac{\partial f}{\partial y}$.

Exemple : $h(x, y) = \ln(1 + x^4 + x^2 y^4 + 2y^2)$.

$\dfrac{\partial h}{\partial x}(x, y) = \dfrac{4x^3 + 2xy^4}{1 + x^4 + x^2 y^4 + 2y^2}$ et $\dfrac{\partial h}{\partial y}(x, y) = \dfrac{4x^2 y^3 + 4y}{1 + x^4 + x^2 y^4 + 2y^2}$.

3) Développement limité d'ordre 1

Définition : Une fonction f définie sur une partie ouverte non vide D de $\mathbb{R}^2$ admet en un point $A = (x_0, y_0)$ de D un développement limité d'ordre 1 s'il existe des réels a et b et une fonction ε définie sur une boule B de centre (0,0) vérifiant $\forall (h,k) \in B \quad (x_0 + h, y_0 + k) \in D$, tels que l'on ait $\lim\limits_{(h,k)\to(0,0)} \varepsilon(h,k) = 0$ et :

$$\forall (h,k) \in B \quad f(x_0 + h, y_0 + k) = f(x_0, y_0) + ah + bk + \sqrt{h^2 + k^2}\,\varepsilon(h,k).$$

On peut remarquer que dans ce cas : $\lim\limits_{(h,k)\to(0,0)} f(x_0 + h, y_0 + k) = f(x_0, y_0)$.

Théorème : Si une fonction admet un développement limité d'ordre 1 en un point A, elle est continue en A.

De plus, pour $k = 0$, on a : $f(x_0 + h, y_0) = f(x_0, y_0) + ah + |h|\varepsilon(h,0)$.

Donc $\lim\limits_{h\to 0} \dfrac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = a$. Donc f admet une dérivée partielle d'ordre 1 en $A = (x_0, y_0)$ par rapport à x et $\dfrac{\partial f}{\partial x}(A) = a$.

Le résultat est identique pour b en étudiant le cas $h = 0$.

Théorème : Si une fonction f définie sur une partie ouverte non vide D de $\mathbb{R}^2$ admet en un point $A = (x_0, y_0)$ de D un développement limité d'ordre 1, alors elle admet en $A = (x_0, y_0)$ des dérivées partielles d'ordre 1 par rapport à x et à y et :

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0) + \sqrt{h^2 + k^2}\,\varepsilon(h,k)$$

où ε est une fonction qui vérifie $\lim\limits_{(h,k)\to(0,0)} \varepsilon(h,k) = 0$.

La quantité $\sqrt{h^2 + k^2}\,\varepsilon(h,k)$ est encore notée $o\left(\sqrt{h^2 + k^2}\right)$.

Le développement limité peut s'écrire : $f(A + H) = f(A) + \langle \nabla f(A), H \rangle + \|H\|\varepsilon(H)$.

En effet si $H = (h,k)$, alors $\|H\| = \sqrt{h^2 + k^2}$.

Attention ! La réciproque est fausse : une fonction peut admettre des dérivées partielles d'ordre 1 en un point sans avoir de développement limité d'ordre 1.

En effet, on a vu que l'existence de dérivées partielles ne garantit pas la continuité.

4) Fonctions de classe C^1

<u>Définition</u> : Une fonction f est de classe C^1 sur une partie D ouverte non vide de $\mathbb{R}^2$ si f admet des dérivées partielles d'ordre 1 sur D et si ces dérivées partielles sont continues sur D.

<u>Exemple</u> : la fonction définie par $f(x,y)=x^3+y^3-2x^2y+1$ est de classe C^1 sur $\mathbb{R}^2$.

<u>Théorème</u> : Si f est une fonction de classe C^1 sur une partie ouverte non vide D de $\mathbb{R}^2$, alors f admet en tout point de D un développement limité d'ordre 1.

En principe, dans le programme, le théorème est admis.

<u>Démonstration</u> : Soit f une fonction de classe C^1 sur un ouvert D de $\mathbb{R}^2$ et $(x_0,y_0)\in D$.

$$f(x_0+h,y_0+k)-f(x_0,y_0)=f(x_0+h,y_0+k)-f(x_0,y_0+k)+f(x_0,y_0+k)-f(x_0,y_0)$$

La fonction $x\mapsto f(x,y_0+k)$ est de classe C^1, donc admet un DL_1 :

$$f(x_0+h,y_0+k)=f(x_0,y_0+k)+h\frac{\partial f}{\partial x}(x_0,y_0+k)+h\varepsilon_1(h,k) \text{ avec } \lim_{h\to 0}\varepsilon_1(h,k)=0$$

La fonction $y\mapsto f(x_0,y)$ est de classe C^1, donc admet un DL_1 :

$$f(x_0,y_0+k)=f(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)+k\varepsilon_2(h,k) \text{ avec } \lim_{k\to 0}\varepsilon_2(h,k)=0.$$

Donc :

$$f(x_0+h,y_0+k)=f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0+k)+h\varepsilon_1(h,k)++k\frac{\partial f}{\partial y}(x_0,y_0)+k\varepsilon_2(h,k)$$

La dérivée partielle $\frac{\partial f}{\partial x}$ est continue. Donc $\frac{\partial f}{\partial x}(x_0,y_0+k)=\frac{\partial f}{\partial x}(x_0,y_0)+\varepsilon_3(h,k)$ avec $\lim_{k\to 0}\varepsilon_3(h,k)=0$. Donc :

$$f(x_0+h,y_0+k)=f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)+h\varepsilon_1(h,k)+h\varepsilon_3(h,k)+k\varepsilon_2(h,k)$$

Il reste à montrer que : $\displaystyle\lim_{(h,k)\to(0,0)}\frac{h\varepsilon_1(h,k)+h\varepsilon_3(h,k)+k\varepsilon_2(h,k)}{\sqrt{h^2+k^2}}=0$. Or :

$$\left|\frac{h\varepsilon_1(h,k)+h\varepsilon_3(h,k)+k\varepsilon_2(h,k)}{\sqrt{h^2+k^2}}\right|\le\frac{|h|}{\sqrt{h^2+k^2}}[|\varepsilon_1(h,k)|+|\varepsilon_3(h,k)|]+\frac{|k|}{\sqrt{h^2+k^2}}|\varepsilon_2(h,k)|$$

Donc : $\left|\dfrac{h\varepsilon_1(h,k)+h\varepsilon_3(h,k)+k\varepsilon_2(h,k)}{\sqrt{h^2+k^2}}\right|\le|\varepsilon_1(h,k)|+|\varepsilon_3(h,k)|+|\varepsilon_2(h,k)|$.

On en déduit le résultat.

<u>Exemple</u> : $f(x,y)=x^2+3y+xy$. La fonction f est de classe C^1 sur $\mathbb{R}^2$, donc admet un développement limité d'ordre 1 par exemple au point $(-1,1)$:

$$f(-1+h,1+k)=f(-1,1)+h\frac{\partial f}{\partial x}(-1,1)+k\frac{\partial f}{\partial y}(-1,1)+\sqrt{h^2+k^2}\,\varepsilon(h,k)$$

$f(-1,1)=3$, $\frac{\partial f}{\partial x}(x,y)=2x+y$ donc $\frac{\partial f}{\partial x}(-1,1)=-1$, $\frac{\partial f}{\partial y}(x,y)=3+x$ donc $\frac{\partial f}{\partial x}(-1,1)=2$

Donc : $f(-1+h,1+k)=3-h+2k+\sqrt{h^2+k^2}\,\varepsilon(h,k)$ avec $\displaystyle\lim_{(h,k)\to(0,0)}\varepsilon(h,k)=0$.

On peut le vérifier directement : $f(-1+h,1+k)=(-1+h)^2+3(1+k)+(-1+h)(1+k)$

$f(-1+h,1+k)=3-h+2k+h^2+hk$ donc $\varepsilon(h,k)=\dfrac{h^2+hk}{\sqrt{h^2+k^2}}$.

Or : $|\varepsilon(h,k)|\leq \dfrac{|h|}{\sqrt{h^2+k^2}}(|h|+|k|)\leq 2\sqrt{h^2+k^2}$. Donc $\lim\limits_{(h,k)\to(0,0)}\varepsilon(h,k)=0$.

<u>Théorème</u> : Si φ et ψ sont deux fonctions dérivables sur un intervalle I de $\mathbb{R}$ et si f est une fonction de classe C^1 sur un ouvert D contenant $\varphi(I)\times\psi(I)$, alors la fonction g définie par $g(t)=f[\varphi(t),\psi(t)]$ est dérivable sur I et :

$$g'(t)=\frac{\partial f}{\partial x}[\varphi(t),\psi(t)]\times\varphi'(t)+\frac{\partial f}{\partial y}[\varphi(t),\psi(t)]\times\psi'(t).$$

<u>Démonstration</u> : Pour montrer que g est dérivable en $t_0\in I$, on montre qu'elle admet un développement limité d'ordre 1.

La fonction f est de classe C^1 en $A=(\varphi(t_0),\psi(t_0))$. Donc, avec $\lim\limits_{(\alpha,\beta)\to(00)}\varepsilon(\alpha,\beta)=0$:

$$f(\varphi(t_0)+\alpha,\psi(t_0)+\beta)=f(A)+\alpha\frac{\partial f}{\partial x}(A)+\beta\frac{\partial f}{\partial y}(A)+\sqrt{\alpha^2+\beta^2}\,\varepsilon(\alpha,\beta).$$

Les fonctions φ et ψ sont dérivables en t_0, donc admettent un $DL_1(t_0)$:

$\varphi(t_0+h)=\varphi(t_0)+h\varphi'(t_0)+h\varepsilon_1(h)$ avec $\lim\limits_{h\to 0}\varepsilon_1(h)=0$.

$\psi(t_0+h)=\psi(t_0)+h\psi'(t_0)+h\varepsilon_2(h)$ avec $\lim\limits_{h\to 0}\varepsilon_2(h)=0$.

Donc $\alpha=h\varphi'(t_0)+h\varepsilon_1(h)$ et $\beta=h\psi'(t_0)+h\varepsilon_2(h)$ tendent vers 0 quand h tend vers 0.

Donc par composition $\varepsilon(\alpha,\beta)$ tend vers 0 quand h tend vers 0.

De plus $\sqrt{\alpha^2+\beta^2}=|h|\sqrt{[\varphi'(t_0)+\varepsilon_1(h)]^2+[\psi'(t_0)+\varepsilon_2(h)]^2}$

Et : $g(t_0+h)=f(\varphi(t_0+h),\psi(t_0+h))=f(\varphi(t_0)+\alpha,\psi(t_0)+\beta)$.

Donc : $g(t_0+h)=f(A)+\alpha\dfrac{\partial f}{\partial x}(A)+\beta\dfrac{\partial f}{\partial y}(A)+\sqrt{\alpha^2+\beta^2}\,\varepsilon(\alpha,\beta)$.

Donc : $g(t_0+h)=g(t_0)+h\left[\varphi'(t_0)\dfrac{\partial f}{\partial x}(A)+\psi'(t_0)\dfrac{\partial f}{\partial y}(A)\right]+h\varepsilon_3(h)$ avec $\lim\limits_{h\to 0}\varepsilon_3(h)=0$.

Donc g admet un $DL_1(t_0)$. Donc g est dérivable en t_0 et :

$$g'(t_0)=\varphi'(t_0)\frac{\partial f}{\partial x}(A)+\psi'(t_0)\frac{\partial f}{\partial y}(A)=\varphi'(t_0)\frac{\partial f}{\partial x}(\varphi(t_0),\psi(t_0))+\psi'(t_0)\frac{\partial f}{\partial y}(\varphi(t_0),\psi(t_0)).$$

<u>*Exemple*</u> : On peut à l'aide de ce théorème retrouver la dérivée du produit uv, du quotient $\dfrac{u}{v}$ ou de la puissance u^v où u et v sont des fonctions d'une variable :

- Si $f(x,y)=xy$, alors $g(t)=u(t)v(t)=f[u(t),v(t)]$.

  Or $\dfrac{\partial f}{\partial x}(x,y)=y$ et $\dfrac{\partial f}{\partial y}(x,y)=x$. Donc $g'(t)=v(t)\times u'(t)+u(t)\times v'(t)$.

- Si $f(x,y)=\dfrac{x}{y}$, alors $g(t)=\dfrac{u(t)}{v(t)}=f[u(t),v(t)]$.

  Or $\dfrac{\partial f}{\partial x}(x,y)=\dfrac{1}{y}$ et $\dfrac{\partial f}{\partial y}(x,y)=-\dfrac{x}{y^2}$.

Donc $g'(t) = \dfrac{1}{v(t)} \times u'(t) - u(t) \times \dfrac{v'(t)}{v^2(t)} = \dfrac{u'(t)v(t) - v'(t)u(t)}{v^2(t)}$.

- Si $f(x, y) = x^y = e^{y \ln x}$, alors $g(t) = [u(t)]^{v(t)} = f[u(t), v(t)]$.
  Or $\dfrac{\partial f}{\partial x}(x, y) = yx^{y-1}$ et $\dfrac{\partial f}{\partial y}(x, y) = (\ln x)x^y$.
  Donc : $g'(t) = v(t)[u(t)]^{v(t)-1} \times u'(t) + \ln[u(t)] \times [u(t)]^{v(t)} \times v'(t)$.
  Donc : $g'(t) = [v(t)u'(t) + v'(t)\ln u(t)][u(t)]^{v(t)-1}$.

<u>Théorème-Définition</u> : Si *f* est une fonction de classe C^1 sur un ouvert non vide *D* de $\mathbb{R}^2$, alors pour tout point *A* de *D* et tout vecteur unitaire *U* de $\mathbb{R}^2$ (donc tel que $\|U\| = 1$), la fonction *g* définie par $g(t) = f(A + tU)$ est dérivable en 0. Son nombre dérivé en 0 est appelé dérivée de *f* en *A* dans la direction *U* et noté $f'_U(A)$. On a donc : $f'_U(A) = < \nabla f(A), U >$.

<u>Démonstration</u> : On note $A = (x_0, y_0)$ et $U = (\alpha, \beta)$ avec $\sqrt{\alpha^2 + \beta^2} = 1$.
Donc : $g(t) = f(A + tU) = f(x_0 + t\alpha, y_0 + t\beta)$.
Or la fonction *f* est de classe C^1 et les fonctions affines sont dérivables.
Donc *g* est dérivable en 0 : $g'(0) = \alpha \dfrac{\partial f}{\partial x}(A) + \beta \dfrac{\partial f}{\partial y}(A) = < \nabla f(A), U >$.
On peut en déduire une propriété du gradient en un point d'une courbe de niveau.

<u>Théorème (admis)</u> : Si *f* est une fonction de classe C^1 sur un ouvert non vide *D* de $\mathbb{R}^2$ et si le gradient de *f* ne s'annule pas, alors en tout point *A* de la courbe de niveau $\mathscr{L}_k$ d'équation $f(x, y) = k$, le gradient $\nabla f(A)$ est normal à la courbe, c'est-à-dire orthogonal à sa tangente.

<u>*Exemple 1*</u> : $f(x, y) = x^2 + y^2$. La courbe de niveau $\mathscr{L}_k$ a pour équation $x^2 + y^2 = k$.
C'est donc un cercle si $k > 0$. Soit $A = (x_0, y_0)$ un point de $\mathscr{L}_k$, donc $x_0^2 + y_0^2 = k$.
La fonction *f* est polynômiale, donc de classe C^1 sur $\mathbb{R}^2$.
$\dfrac{\partial f}{\partial x}(x, y) = 2x$ et $\dfrac{\partial f}{\partial y}(x, y) = 2y$. Donc $\nabla f(A) = (2x_0, 2y_0) = 2\overrightarrow{OA}$.
Le gradient est non nul car $x_0^2 + y_0^2 > 0$.
On retrouve la propriété des cercles : la droite (*OA*) est orthogonale à la tangente en *A*.

<u>*Exemple 2*</u> : $f(x, y) = xy$. La courbe de niveau $\mathscr{L}_k$ a pour équation $xy = k$. C'est une hyperbole équilatère si $k \neq 0$. Soit $A = (x_0, y_0)$ un point de $\mathscr{L}_k$, donc $x_0 y_0 = k$.
$\mathscr{L}_k$ est la courbe représentative de la fonction *g* définie par $g(x) = \dfrac{k}{x}$.
Donc la tangente en *A* à $\mathscr{L}_k$ a pour équation : $y = -\dfrac{k}{x_0^2}(x - x_0) + \dfrac{k}{x_0} = -\dfrac{y_0}{x_0}x + \dfrac{2k}{x_0}$,
c'est-à-dire d'équation $y_0 x + x_0 y - 2k = 0$.
La tangente en *A* à $\mathscr{L}_k$ a donc pour vecteur directeur : $U = (-x_0, y_0)$.
La fonction *f* est polynômiale, donc de classe C^1 sur $\mathbb{R}^2$.
$\dfrac{\partial f}{\partial x}(x, y) = y$ et $\dfrac{\partial f}{\partial y}(x, y) = x$. Donc le gradient de *f* en *A* est : $\nabla f(A) = (y_0, x_0)$. Il n'est pas nul puisque $x_0 y_0 \neq 0$. Et $< \nabla f(A), U > = -x_0 y_0 + x_0 y_0 = 0$.
Donc le gradient en *A* est bien orthogonal à la tangente en *A*.

5) Dérivées partielles d'ordre 2

C'est le même principe que les dérivées successives des fonctions d'une variable, mais il faut faire attention à l'ordre de dérivation.

Définition : Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie sur une partie D ouverte et non vide, et soit $A = (x_0, y_0)$ appartenant à D. En A, la fonction f est dérivable deux fois par rapport à x si elle admet une dérivée partielle $\frac{\partial f}{\partial x}$ par rapport à x qui admet elle-même une dérivée partielle par rapport à x en u_0. Elle est notée $\frac{\partial^2 f}{\partial x^2}$.

On définit de même si elle existe $\frac{\partial^2 f}{\partial x \partial y}$ dérivée partielle de $\frac{\partial f}{\partial y}$ par rapport à x.

La dérivée partielle de $\frac{\partial f}{\partial x}$ par rapport à y est notée $\frac{\partial^2 f}{\partial y \partial x}$ si elle existe.

La dérivée partielle de $\frac{\partial f}{\partial y}$ par rapport à y est notée $\frac{\partial^2 f}{\partial y^2}$ si elle existe.

Exemple : On reprend $f(x, y) = x^3 + y^3 - 2x^2 y + 1$.

On a vu que : $\frac{\partial f}{\partial x}(x, y) = 3x^2 - 4xy$ et $\frac{\partial f}{\partial y}(x, y) = 3y^2 - 2x^2$.

Ces fonctions admettent aussi des dérivées partielles :

$\frac{\partial^2 f}{\partial x^2}(x, y) = 6x - 4y$ $\quad$ $\frac{\partial^2 f}{\partial x \partial y}(x, y) = -4x$ $\quad$ $\frac{\partial^2 f}{\partial y \partial x}(x, y) = -4x$ $\quad$ $\frac{\partial^2 f}{\partial y^2}(x, y) = 6y$

On constate sur cet exemple que $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$. On va voir que ce n'est pas un hasard.

Définition : Une fonction f est de classe C^2 sur une partie D ouverte non vide de $\mathbb{R}^2$ si f admet des dérivées partielles d'ordre 1 et si ces dérivées sont de classe C^1 sur D.

On se limitera à l'ordre 2, mais bien sûr, on pourrait continuer à dériver. Les notations deviennent plus compliquées.

Théorème de Schwarz : Si une fonction f est de classe C^2 sur une partie D ouverte non vide de $\mathbb{R}^2$, alors : $\forall (x, y) \in D \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$.

En principe, dans le programme, le théorème est admis.

Démonstration : Soit $(x_0, y_0) \in D$. On évalue de deux manières différentes :

$\delta(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) - f(x_0 + h, y_0) + f(x_0, y_0)$.

- D'abord, posons $\varphi(x) = f(x, y_0 + k) - f(x, y_0)$. Donc $\delta(h, k) = \varphi(x_0 + h) - \varphi(x_0)$

φ est de classe C^1, donc a un DL_1 : $\delta(h, k) = h\varphi'(x_0) + h\varepsilon_1(h, k)$ avec $\lim_{h \to 0} \varepsilon_1(h, k) = 0$

$\varphi'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0 + k) - \frac{\partial f}{\partial x}(x_0, y_0) = k \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) + k\varepsilon_2(h, k)$ avec $\lim_{k \to 0} \varepsilon_2(h, k) = 0$

car la dérivée partielle $\frac{\partial f}{\partial x}$ est de classe C^1, donc a un DL_1. Donc :

$\delta(h, k) = h\varphi'(x_0) + h\varepsilon_1(h, k) = hk \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) + hk\varepsilon_2(h, k) + h\varepsilon_1(h, k)$.

- D'autre part, posons $\psi(y)=f(x_0+h,y)-f(x_0,y)$. Donc $\delta(h,k)=\psi(y_0+k)-\psi(y_0)$

On raisonne de même et on trouve avec $\lim\limits_{k\to 0}\varepsilon_3(h,k)=0$ et $\lim\limits_{h\to 0}\varepsilon_4(h,k)=0$:

$$\delta(h,k)=k\psi'(y_0)+k\varepsilon_3(h,k)=hk\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)+hk\varepsilon_4(h,k)+k\varepsilon_3(h,k)$$

- En comparant les deux expressions :

$$hk\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)+hk\varepsilon_2(h,k)+h\varepsilon_1(h,k)=hk\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)+hk\varepsilon_4(h,k)+k\varepsilon_3(h,k)$$

Donc : $\dfrac{\partial^2 f}{\partial y\partial x}(x_0,y_0)+\varepsilon_2(h,k)+\dfrac{\varepsilon_1(h,k)}{k}=\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)+\varepsilon_4(h,k)+\dfrac{\varepsilon_3(h,k)}{h}$

A gauche, si on fait tendre d'abord *h* vers 0, puis *k* vers 0, la limite est $\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)$.

A droite, si on fait tendre d'abord *k* vers 0, puis *h* vers 0, la limite est $\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)$.

Comme les fonctions $\dfrac{\partial^2 f}{\partial x\partial y}$ et $\dfrac{\partial^2 f}{\partial y\partial x}$ sont continues, ce sont aussi les limites quand (h,k) tend vers $(0,0)$. On a donc le résultat par unicité de la limite.

On admettra le résultat suivant :

<u>Théorème</u> : Si *f* est une fonction de classe C^2 sur une partie ouverte non vide *D* de $\mathbb{R}^2$, alors elle admet en tout point $A=(x_0,y_0)$ de *D* un développement limité d'ordre 2 :

$$f(x_0+h,y_0+k)=f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)$$
$$+\frac{1}{2}\left[h^2\frac{\partial^2 f}{\partial x^2}(x_0,y_0)+2hk\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)+k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right]+(h^2+k^2)\varepsilon(h,k)$$

où ε est une fonction qui vérifie $\lim\limits_{(h,k)\to(0,0)}\varepsilon(h,k)=0$.

<u>Exemple</u> : On reprend $f(x,y)=x^3+y^3-2x^2y+1$.

On a vu que : $\dfrac{\partial f}{\partial x}(x,y)=3x^2-4xy$ $\qquad\dfrac{\partial f}{\partial y}(x,y)=3y^2-2x^2$.

$\dfrac{\partial^2 f}{\partial x^2}(x,y)=6x-4y\qquad\dfrac{\partial^2 f}{\partial x\partial y}(x,y)=-4x\qquad\dfrac{\partial^2 f}{\partial y\partial x}(x,y)=-4x\qquad\dfrac{\partial^2 f}{\partial y^2}(x,y)=6y$.

Donc par exemple, au point $(1,1)$:

$f(1+h,1+k)=1-h+k+h^2-4hk+3k^2+(h^2+k^2)\varepsilon(h,k)$ avec $\lim\limits_{(h,k)\to(0,0)}\varepsilon(h,k)=0$

Le calcul direct donne : $f(1+h,1+k)=(1+h)^3+(1+k)^3-2(1+h)^2(1+k)+1$.

Donc : $f(1+h,1+k)=1-h+k+h^2-4hk+3k^2+h^3+k^3-2h^2k$.

Donc : $\varepsilon(h,k)=\dfrac{h^3+k^3-2h^2k}{h^2+k^2}$ qui tend vers 0 quand (h,k) tend vers $(0,0)$ car le numérateur est d'ordre supérieur au dénominateur :

$|\varepsilon(h,k)|\le\dfrac{|h|^3+|k|^3+2|h|^2|k|}{h^2+k^2}\le 4\sqrt{h^2+k^2}$ en majorant $|h|$ et $|k|$ par $\sqrt{h^2+k^2}$.

V – Recherche d'extrema sur $\mathbb{R}^2$

En économie, un problème très important qui nécessite l'utilisation des mathématiques est celui de l'optimisation d'une quantité qui dépend de plusieurs paramètres (dans notre cas 2). Ces paramètres peuvent être libres ou bien être soumis à des contraintes. On se limitera ici à des cas simples car les optimisations sous contrainte peuvent être très compliquées.

1) Extremum

Définition : Une fonction f définie sur une partie ouverte non vide D de $\mathbb{R}^2$ admet un maximum local en $A=(x_0,y_0)$ de D s'il existe une boule ouverte $B(A,\alpha)$ incluse dans D telle que : $\forall M\in B(A,\alpha)\quad f(M)\leq f(A)$.
Le maximum est strict si l'égalité n'est vraie que pour $M=A$.
Le maximum est absolu lorsque l'inégalité est vraie en tout point M de D.

De même :

Définition : Une fonction f définie sur une partie ouverte non vide D de $\mathbb{R}^2$ admet un minimum local en $A=(x_0,y_0)$ de D s'il existe une boule ouverte $B(A,\alpha)$ incluse dans D telle que : $\forall M\in B(A,\alpha)\quad f(M)\geq f(A)$.
Le minimum est strict si l'égalité n'est vraie que pour $M=A$.
Le minimum est absolu lorsque l'inégalité est vraie sur D.

Et bien sûr :

Définition : Un extremum local (absolu) est un minimum local (absolu) ou un maximum local (absolu).

Exemple : $f(x,y)=x^2+y^2-2x-4y+1$.

On remarque que : $f(x,y)=(x-1)^2+(y-2)^2-4$.

Donc : $\forall(x,y)\in\mathbb{R}^2\quad f(x,y)\geq -4$. Donc : $\forall(x,y)\in\mathbb{R}^2\quad f(x,y)\geq f(1,2)$.
La fonction admet donc un minimum absolu au point $(1,2)$.
La méthode la plus naturelle pour savoir si une fonction admet en un point $A=(x_0,y_0)$ de D un extremum local est d'étudier le signe de $f(M)-f(A)$.
Mais le plus souvent, comme pour les fonctions d'une variable, on peut établir des critères de recherche plus pratiques.

2) Cas des fonctions de classe C^1

Théorème : Soit f une fonction de classe C^1 sur une partie ouverte non vide D de $\mathbb{R}^2$. Si elle admet un extremum local au point $A=(x_0,y_0)$ de D, alors ses dérivées partielles d'ordre 1 s'annulent en A : $\dfrac{\partial f}{\partial x}(A)=0$ et $\dfrac{\partial f}{\partial y}(A)=0$.

Démonstration : On se ramène à deux fonctions d'une variable.
Supposons qu'il s'agisse d'un maximum local. Donc il existe une boule ouverte $B(a,\alpha)\subset D$ telle que : $\forall M\in B(A,\alpha)\quad f(M)\leq f(A)$.
Donc : $\forall x\in]x_0-\alpha,x_0+\alpha[\quad f(x,y_0)\leq f(x_0,y_0)$. Et donc la fonction $x\mapsto f(x,y_0)$ admet un maximum local en x_0. La fonction f est de classe C^1 sur $B(A,\alpha)$, et donc la fonction $x\mapsto f(x,y_0)$ est dérivable, de dérivée $x\mapsto\dfrac{\partial f}{\partial x}(x,y_0)$ qui est continue. Donc la fonction $x\mapsto f(x,y_0)$ est de classe C^1 sur $]x_0-\alpha,x_0+\alpha[$ et y admet un maximum local en x_0. Donc sa dérivée s'annule en x_0. Donc $\dfrac{\partial f}{\partial x}(x_0,y_0)=0$.

Le raisonnement est identique pour la dérivée partielle par rapport à *y*.

Remarque : Il faut faire très attention car ce théorème ne donne qu'une condition nécessaire d'extremum : les dérivées partielles peuvent être nulles en un point qui n'est pas un extremum.

Exemple : $f(x,y)=x^2-y^2$. Donc $\frac{\partial f}{\partial x}(x,y)=2x$ et $\frac{\partial f}{\partial y}(x,y)=-2y$.

Le seul point où les deux dérivées s'annulent est $(0,0)$.

Or pour tout réel *x* : $f(x,0)=x^2$, donc $f(x,0)\geq f(0,0)$.

Et pour tout réel *x* : $f(0,y)=-y^2$, donc $f(0,y)\leq f(0,0)$.

Donc $(0,0)$ n'est ni un maximum ni un minimum, puisque dans toute boule, il y a à la fois des points de la forme $(x,0)$ et de la forme $(0,y)$.

Définition : Si *f* est une fonction de classe C^1 sur une partie ouverte non vide *D* de $\mathbb{R}^2$, on appelle point critique tout point *A* de *D* où ses dérivées partielles d'ordre 1 s'annulent : $\frac{\partial f}{\partial x}(A)=0$ et $\frac{\partial f}{\partial y}(A)=0$.

Les extremums de la fonction sur un ouvert sont à chercher parmi les points critiques.

3) Cas des fonctions de classe C^2

Notations de MONGE : Si *f* est une fonction de classe C^2 sur une partie ouverte non vide *D* de $\mathbb{R}^2$, on pose : $p=\frac{\partial f}{\partial x}$, $q=\frac{\partial f}{\partial y}$, $r=\frac{\partial^2 f}{\partial x^2}$, $s=\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x}$, $t=\frac{\partial^2 f}{\partial y^2}$.

Les extremums de la fonction sur un ouvert sont à chercher parmi les points critiques.
Le théorème suivant va permettre de savoir si certains points critiques sont ou nopn des extremums.

Théorème : On suppose que la fonction *f* est de classe C^2 sur un ouvert non vide *D* de $\mathbb{R}^2$. Pour chaque point critique $A=(x_0,y_0)$:

- Si $rt-s^2>0$, alors *f* admet un extremum local en *A* : si $r>0$, c'est un minimum, et si $r<0$, c'est un maximum.
- Si $rt-s^2<0$, alors *f* n'admet pas d'extremum local en *A* (point-col ou point-selle).
- Si $rt-s^2=0$, on ne peut pas conclure : il faut étudier « à la main » le signe de $f(M)-f(A)=f(x,y)-f(x_0,y_0)$.

Démonstration : Soit $A=(x_0,y_0)$ un point critique de *f*.

Et pour abréger, notons $r_0=r(x_0,y_0)$, $s_0=s(x_0,y_0)$ et $t_0=t(x_0,y_0)$.

La fonction *f* est de classe C^2, donc admet un développement limité d'ordre 2 en u_0 :

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{1}{2}(h^2r_0+2hks_0+k^2t_0)+(h^2+k^2)\varepsilon(h,k)$$

où ε est une fonction qui vérifie $\lim\limits_{(h,k)\to(0,0)}\varepsilon(h,k)=0$.

La fonction *f* admet un extremum local en u_0 si $f(x_0+h,y_0+k)-f(x_0,y_0)$ garde un signe constant au voisinage de $(0,0)$.

- Si $r_0=s_0=t_0=0$, on ne peut pas conclure car en général, on ne connaît pas ε.
- Sinon, le terme $(h^2+k^2)\varepsilon(h,k)$ est négligeable devant le terme précédent, et on admet que $f(x_0+h,y_0+k)-f(x_0,y_0)$ est du signe de $\delta(h,k)=h^2r_0+2hks_0+k^2t_0$ si ce terme n'est pas nul.

- Si $r_0 = t_0 = 0$, on a donc $s_0 \neq 0$ et $\delta(h,k) = 2hks_0$ ne garde pas un signe constant puisque par exemple $\delta(h,h)$ et $\delta(h,-h)$ sont de signes contraires. Or dans chaque boule de centre $(0,0)$, il existe des points de la forme (h,h) et $(h,-h)$. Donc *f* n'admet pas d'extremum local.
- Si $r_0 \neq 0$, on se ramène à un polynôme du second degré. Pour $k \neq 0$: $\delta(h,k) = k^2\left[\left(\dfrac{h}{k}\right)^2 r_0 + 2\left(\dfrac{h}{k}\right)s_0 + t_0\right] = k^2 P\left(\dfrac{h}{k}\right)$ où P est le polynôme défini par $P(\alpha) = \alpha^2 r_0 + 2\alpha s_0 + t_0$. Il a pour discriminant $\Delta = 4(s_0^2 - r_0t_0)$. Donc :
  * Si $r_0t_0 - s_0^2 < 0$, alors $\Delta > 0$. Le polynôme P possède deux racines α_1 et α_2. Donc : $P(\alpha) = r_0(\alpha - \alpha_1)(\alpha - \alpha_2)$ et donc $\delta(h,k) = r_0(h - \alpha_1 k)(h - \alpha_2 k)$ ce qui ne garde pas un signe constant. Par exemple $\delta(h,0) = r_0h^2$ et $\delta\left(\dfrac{\alpha_1+\alpha_2}{2}k, k\right) = -r_0\left(\dfrac{\alpha_1+\alpha_2}{2}\right)^2 k^2$ sont de signes contraires. Or dans chaque boule de centre $(0,0)$, il existe de tels points. La fonction *f* n'admet donc pas d'extremum en *A*.
  * Si $r_0t_0 - s_0^2 = 0$, alors $\Delta = 0$. Le polynôme P possède une racine double α_1. Donc : $P(\alpha) = r_0(\alpha - \alpha_1)^2$ et donc $\delta(h,k) = r_0(h - \alpha_1 k)^2$. Pour tout réel *k*, on a $\delta(\alpha_1 k, k) = \sqrt{\alpha_1^2 k^2 + k^2}\,\varepsilon(\alpha_1 k, k)$ dont on ne connaît pas le signe. Donc on ne peut pas conclure pour l'existence d'un extremum.
  * Si $r_0t_0 - s_0^2 > 0$, alors $\Delta < 0$. Donc le polynôme P ne s'annule pas et garde le signe de r_0. Donc $\delta(h,k)$ et $f(x_0+h, y_0+k) - f(x_0, y_0)$ sont du signe de r_0 pour tout (h,k). La fonction *f* admet donc un extremum en *A* : si $r_0 > 0$, c'est un minimum et si $r_0 < 0$, c'est un maximum.
- Si $t_0 \neq 0$, pour $h \neq 0$: $\delta(h,k) = h^2\left[\left(\dfrac{k}{h}\right)^2 t_0 + 2\left(\dfrac{k}{h}\right)s_0 + r_0\right] = h^2 Q\left(\dfrac{h}{k}\right)$ où Q est le polynôme défini par $Q(u) = u^2 t_0 + 2us_0 + r_0$. Il a le même discriminant que *P*. La discussion est identique et on peut remarquer que dans le cas $r_0t_0 - s_0^2 > 0$, Q garde le signe de t_0. Mais puisque $r_0t_0 > s_0^2 \geq 0$, r_0 et t_0 sont de même signe.

<u>*Exemple*</u> : $f(x,y) = x^3 + y^3 + 3xy$.

Donc $p(x,y) = \dfrac{\partial f}{\partial x}(x,y) = 3x^2 + 3y$ et $q(x,y) = \dfrac{\partial f}{\partial y}(x,y) = 3y^2 + 3x$.

Les points critiques sont donc les solutions de $\begin{cases} x^2 + y = 0 \\ y^2 + x = 0 \end{cases}$, donc de $\begin{cases} y = -x^2 \\ x^4 + x = 0 \end{cases}$.

On trouve donc deux points critiques : $(0,0)$ et $(-1,-1)$.

$r(x,y) = \dfrac{\partial^2 f}{\partial x^2}(x,y) = 6x \qquad s(x,y) = \dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 3 \qquad t(x,y) = \dfrac{\partial^2 f}{\partial y^2}(x,y) = 6y$

Pour $(0,0)$: $rt - s^2 = -9$, donc $rt - s^2 < 0$. Pas d'extremum.

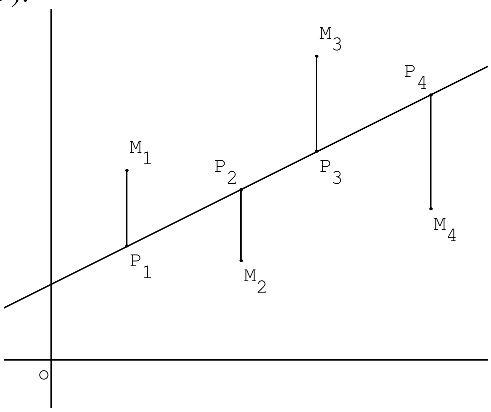
Pour $(-1,-1)$: $rt - s^2 = 27$, donc $rt - s^2 > 0$ et $r < 0$. C'est un maximum local.

4) Application : Méthode des moindres carrés

On suppose qu'on a une série statistique double équipondérée, c'est-à-dire un nuage de points $M_i(x_i, y_i)$ pour $1 \le i \le n$.

Il s'agit sur cette série observée de trouver une courbe mathématique qui « approche » le mieux ce nuage : on dit que l'on fait un « ajustement ».
La forme du nuage ou bien d'autres considérations poussent à choisir un type de modèle mathématique : l'ajustement peut être linéaire ou affine s'il s'agit d'une droite, ou logarithmique, ou exponentiel, ou … Beaucoup d'ajustements se ramènent par changement de variable à un ajustement linéaire. Ici, on se limite à celui-là.
Le problème est que, une fois que l'on a choisi le type de modèle, celui-ci dépend d'un certain nombre de paramètres. Dans le cas d'un ajustement linéaire, l'équation de la droite (*D*) est de la forme $y = ax + b$, donc il reste à déterminer *a* et *b*.

Le problème est donc de trouver une droite (*D*), donc des réels *a* et *b* qui minimisent les distances des points du nuage à leur projeté sur (*D*).



On ne peut pas espérer minimiser toutes les distances à la fois, donc on cherche à minimiser la moyenne quadratique de ces distances, ce qui revient à minimiser la somme des carrés des distances. De plus pour des raisons de simplification, on ne considère pas les projetés orthogonaux des points sur (*D*), mais les projetés P_i des points M_i sur (*D*) parallèlement à l'axe des ordonnées, ce qui donnera la droite de régression de *y* en *x* en minimisant la somme des $M_iP_i^2$.

Cette somme dépend évidemment de la droite (*D*), donc des coefficients *a* et *b*. C'est une fonction de deux variables : $f(a,b) = \sum_{i=1}^{n} M_iP_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$.

Donc : $f(a,b) = \sum_{i=1}^{n} y_i^2 + a^2 \sum_{i=1}^{n} x_i^2 - 2a \sum_{i=1}^{n} x_i y_i + 2ab \sum_{i=1}^{n} x_i - 2b \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} b^2$.

Or $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ donc $\sum_{i=1}^{n} x_i = n\bar{x}$, et $\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2}$, donc $\sum_{i=1}^{n} x_i^2 = n(\sigma_x^2 + \bar{x}^2)$.

De même : $\sum_{i=1}^{n} y_i = n\bar{y}$, et $\sum_{i=1}^{n} y_i^2 = n(\sigma_x^2 + \bar{y}^2)$.

La covariance de la série double est $\sigma_{xy} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y}$, donc $\sum_{i=1}^{n} x_i y_i = n(\sigma_{xy} + \bar{x}\,\bar{y})$

Donc : $f(a,b) = n(\sigma_y^2 + \bar{y}^2) + a^2 n(\sigma_x^2 + \bar{x}^2) - 2an(\sigma_{xy} + \overline{x}\,\overline{y}) + 2abn\bar{x} - 2bn\bar{y} + nb^2$.

C'est un polynôme, donc la fonction *f* est de classe C^2 sur $\mathbb{R}^2$.

$p(a,b) = 2na(\sigma_x^2 + \bar{x}^2) - 2n(\sigma_{xy} + \bar{x}\,\bar{y}) + 2nb\bar{x} = 2n[a(\sigma_x^2 + \bar{x}^2) - (\sigma_{xy} + \bar{x}\,\bar{y}) + b\bar{x}]$

$q(a,b) = 2na\bar{x} - 2n\bar{y} + 2nb = 2n(a\bar{x} - \bar{y} + b)$

Les points critiques sont donc les couples (a,b) solutions du système :

Donc : $\begin{cases} a(\sigma_x^2 + \bar{x}^2) + \bar{x}(\bar{y} - a\bar{x}) = \sigma_{xy} + \bar{x}\,\bar{y} \\ b = \bar{y} - a\bar{x} \end{cases}$. Donc : $\begin{cases} a\sigma_x^2 = \sigma_{xy} \\ b = \bar{y} - a\bar{x} \end{cases}$

On obtient donc un seul point critique (a_0, b_0) avec $a_0 = \dfrac{\sigma_{xy}}{\sigma_x^2}$ et $b_0 = \overline{y} - a_0\overline{x}$.

Il reste à examiner si ce point critique correspond à un minimum. On calcule les dérivées partielles d'ordre 2 :

$r(a,b) = 2n(\sigma_x^2 + \overline{x}^2)$ $\quad s(a,b) = 2n\overline{x}$ $\quad t(a,b) = 2n$

Donc $rt - s^2 = 4n^2\sigma_x^2$. Donc $rt - s^2 > 0$ et $r > 0$.

Donc il s'agit d'un minimum.

Et : $f(a_0,b_0) = \sum_{i=1}^{n}[(y_i - \overline{y}) - a_0(x_i - \overline{x})]^2 = n\sigma_y^2 - 2n\dfrac{\sigma_{xy}}{\sigma_x^2}\sigma_{xy} + n\dfrac{\sigma_{xy}^2}{\sigma_x^4}\sigma_x^2$

Donc : $f(a_0,b_0) = n\sigma_y^2\left(1 - \dfrac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2}\right) = n\sigma_y^2(1-\rho^2)$ où $\rho = \dfrac{\sigma_{xy}}{\sigma_x\sigma_y}$ est le coefficient de corrélation linéaire de la série statistique double.

On démontre même qu'il s'agit d'un minimum absolu car on peut écrire :

$f(a,b) = n(\sigma_y^2 + \overline{y}^2) + a^2 n(\sigma_x^2 + \overline{x}^2) - 2an(\sigma_{xy} + \overline{x}\,\overline{y}) + 2abn\overline{x} - 2bn\overline{y} + nb^2$.

Donc $f(a,b) = n\sigma_y^2 + a^2 n\sigma_x^2 - 2an\sigma_{xy} + n(\overline{y} - a\overline{x} - b)^2$.

Donc $f(a,b) = n\sigma_x^2\left(a - \dfrac{\sigma_{xy}}{\sigma_x^2}\right)^2 + n\sigma_y^2\left(1 - \dfrac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2}\right) + n(\overline{y} - a\overline{x} - b)^2$.

Donc $f(a,b) - f(a_0,b_0) = n\sigma_x^2\left(a - \dfrac{\sigma_{xy}}{\sigma_x^2}\right)^2 + n(\overline{y} - a\overline{x} - b)^2$.

Donc pour tous a et b : $f(a,b) \geq f(a_0,b_0)$.

La droite qui « approche le mieux » le nuage de points est la droite de régression de y en x d'équation $y = ax + b$ avec $a = \dfrac{\sigma_{xy}}{\sigma_x^2}$ et $b = \overline{y} - a\overline{x}$.

Elle approchera d'autant mieux le nuage de points que $|\rho|$ sera voisin de 1.

Remarque : On a traité le problème en privilégiant la direction de l'axe des ordonnées. Il existe une autre droite de régression de x en y en prenant les projetés des points M_i parallèlement à l'axe des abscisses. Les deux droites de régression sont en général distinctes.