On a

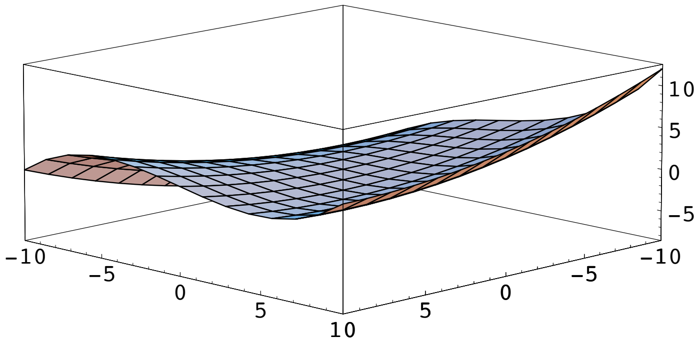
$$\det\left(H_f(3,3)\right) = 8, \quad D_1^2 f(3,3) = 2$$

donc $(3,3)$ correspond à un minimum local strict.

On a

$$\det\left(H_f(-1,-1)\right) = -8$$

donc $(-1,-1)$ n'est pas un extremum.



5.3 Application: la régression linéaire

Il s'agit d'une application de l'étude des extrema des fonctions dites "quadratiques".

5.3.1 Généralités

Une fonction de deux variables réelles[2] du type

$$f(x,y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c$$

où les coefficients des variables x, y sont des réels et où les a_{ij} ne sont pas tous nuls est appelée *fonction quadratique*. On les a déjà rencontrées précédemment dans le cadre de ce cours car elles sont intimement liées aux coniques et aux quadriques: dans le plan muni d'un repère, une équation du type $f(x,y) = 0$ est l'équation cartésienne d'une conique (en toute généralité) et dans l'espace muni d'un repère, une équation du type $z = f(x,y)$ est l'équation d'une quadrique[3].

Dans ce cas simple, le système permettant de déterminer les points stationnaires de f s'écrit

$$\left\{\begin{array}{l} D_x f(x,y) = 2a_{11}x + 2a_{12}y + 2b_1 = 0 \\ D_y f(x,y) = 2a_{22}y + 2a_{12}x + 2b_2 = 0 \end{array}\right.$$

et la matrice hessienne

$$H_f = \begin{pmatrix} D_x^2 f(x,y) & D_x D_y f(x,y) \\ D_x D_y f(x,y) & D_y^2 f(x,y) \end{pmatrix} = 2 \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

est constante. En un point stationnaire (x_0, y_0), on a ainsi

$$f(x,y) = f(x_0,y_0) + \frac{1}{2}(x - x_0 \quad y - y_0) H_f \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \quad \forall x, y \in \mathbb{R}.$$

On en déduit donc la propriété suivante (dans CE cas de fonction f).

[2] cela s'étend à plus de deux variables réelles

[3] de type 'paraboloïde'. Il s'agit en effet d'un cas particulier d'une équation du type $F(x,y,z) = 0$ où F est une fonction quadratique de trois variables réelles.

Propriété 5.3.1 *a) Si le déterminant de la matrice hessienne est non nul, alors il existe un point stationnaire unique.*

b) Un extremum est toujours global.

□

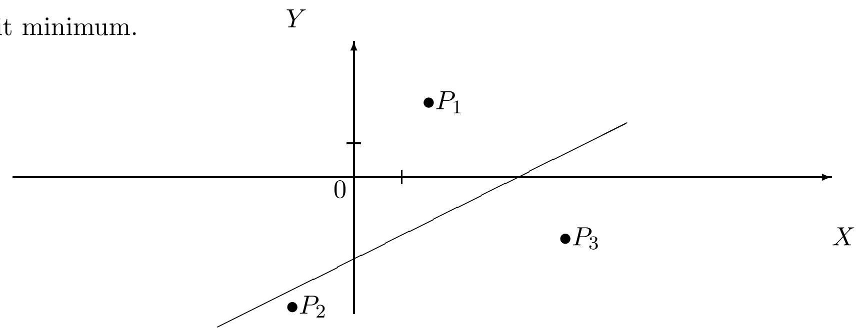
5.3.2 La régression linéaire

En guise d'application, examinons le problème de *la régression linéaire.*

Dans le plan muni d'un repère orthonormé, on donne un nombre fini de points distincts $P_j(j = 1, \ldots, J)$ par leurs coordonnées cartésiennes $(x_j, y_j)(j = 1, \ldots, J)$. Dans le cas où les x_j ne sont pas tous égaux, on demande de chercher la droite[4] d (dite *droite de régression*) d'équation cartésienne $y = mx + b$ telle que

$$E(m, b) = \sum_{j=1}^{J} (y_j - (mx_j + b))^2$$

soit minimum.



Comme $|y_j - (mx_j + b)|$ est la valeur absolue de la différence entre l'ordonnée du point P_j et celle du point de d dont l'abscisse est x_j, l'expression $E(m, b)$ mesure d'une certaine manière la façon dont les ponts P_j sont répartis autour de[5] d. La fonction E définie ci-dessus s'appelle *l'erreur quadratique totale.*

Le problème est donc celui de la recherche des extema (libres) éventuels d'une fonction quadratique, notée ici E et dont les variables sont notées m et b.

Dans ce cas particulier[6], l'expression de E permet de conclure immédiatement à l'existence d'un minimum (que l'on sait donc aussi être global). La recherche de l'équation cartésienne de d (correspondant à la recherche du minimum de E) consiste donc en la détermination de ou des solutions du système caractérisant les points stationnaires.

En utilisant les notations de la section précédente, les coefficients de $m^2, b^2, 2mb$ dans l'expression de E, sont respectivement

$$a_{11} = \sum_{j=1}^{J} x_j^2, \quad a_{22} = J, \quad a_{12} = \sum_{j=1}^{J} x_j$$

[4]existence et unicité aussi

[5]remarquons que dans le cas où les points ont tous la même abscisse r, la droite verticale d'équation $x = r$ contient tous les points P_j

[6]bien sûr si l'on s'attache à la méthode générale, c'est l'examen de la matrice hessienne qui permet de conclure au fait que l'on a affaire à un minimum

ce qui correspond à la matrice hessienne

$$H_E = 2 \begin{pmatrix} \sum_{j=1}^{J} x_j^2 & \sum_{j=1}^{J} x_j \\ \sum_{j=1}^{J} x_j & J \end{pmatrix}.$$

En utilisant les hypothèses sur les abscisses x_j et la propriété ci-dessous, on en déduit *qu'il existe un point stationnaire unique et que celui-ci donne lieu à un minimum strict global.*

Propriété. Par récurrence (sur J), montrons que, quel que soit $J \in \mathbb{N}_0$ et quels que soient les réels x_j $(j = 1, \ldots, J)$, on a

$$\left(\sum_{j=1}^{J} x_j\right)^2 \leq J \sum_{j=1}^{J} x_j^2.$$

De plus, on a l'égalité si et seulement si tous les x_j sont égaux.

De fait, pour $J = 1$ et aussi pour $J = 2$ c'est clair. Supposons à présent la propriété correcte pour J. On obtient successivement

$$\begin{aligned}
\left(\sum_{j=1}^{J+1} x_j\right)^2 = \left(\sum_{j=1}^{J} x_j + x_{J+1}\right)^2 &= \left(\sum_{j=1}^{J} x_j\right)^2 + x_{J+1}^2 + 2x_{J+1}\left(\sum_{j=1}^{J} x_j\right) \\
&\leq J\sum_{j=1}^{J} x_j^2 + x_{J+1}^2 + 2x_{J+1}\left(\sum_{j=1}^{J} x_j\right) \\
&\leq J\sum_{j=1}^{J} x_j^2 + x_{J+1}^2 + \left(\sum_{j=1}^{J}(x_j^2 + x_{J+1}^2)\right) \\
&= (J+1)\sum_{j=1}^{J+1} x_j^2
\end{aligned}$$

Le cas de l'égalité se démontre en affinant un peu la preuve ci-dessus.□

5.4 Extrema dans des ensembles bornés fermés

5.5 Extrema liés: la méthode des multiplicateurs de Lagrange