Ce document comporte une présentation théorique, suivie de calculs effectués en Maple. Les questions sont données dans la partie Maple.

1 Minima Locaux d'une Fonction d'une Variable

Voir la partie Maple pour un exemple. Pour trouver un minimum local d'une fonction $f(x)$ d'une variable réelle on procède en deux temps. On cherche d'abord les points du graphe $y = f(x)$ de la fonction qui ont une tangente horizontale. Comme le coefficient directeur de la tangente à un point d'abscisse x_0 est égal à $f'(x_0)$, on cherche les racines de la dérivée $f'(x)$ de la fonction à minimiser. Ensuite, pour chaque point à tangente horizontale d'abscisse x_0, on détermine s'il s'agit d'un minimum local en étudiant la dérivée seconde : si $f''(x_0) > 0$ alors $(x_0, f(x_0))$ est un minimum local.

La raison se comprend bien en considérant le développement limité de la fonction f en un point x_0

$$f(x_0+h) \quad = \quad f(x_0) + \underbrace{h\,f'(x_0)}_{=0} + \underbrace{\frac{1}{2}\,h^2\,f''(x_0)}_{\text{même signe que } f''(x_0)} + \cdots$$

Fixons l'abscisse x_0. La variable c'est h qui représente un déplacement par rapport à x_0 : si $h > 0$ on se déplace vers la droite ; si $h < 0$ on se déplace vers la gauche. Comme on cherche un minimum *local*, on peut prendre h aussi petit qu'on veut (en valeur absolue) et supposer $|h| \gg |h|^2 \gg |h|^3 \gg \cdots$ On aura donc un minimum local si le premier terme non nul du développement limité est positif, quel que soit le signe de h. Supposons $f'(x_0) = 0$. Alors le premier terme après le terme constant s'annule. Supposons de plus la dérivée seconde positive. Comme $h^2 > 0$ quel que soit $h \neq 0$, on voit que $f(x_0 + h) > f(x_0)$ pour tout h non nul suffisamment petit.

2 Minima Locaux d'une Fonction de Plusieurs Variables

On prend le cas d'une fonction $f(x, y)$ de deux variables réelles pour simplifier mais le raisonnement se généralise au cas de n variables. Dans le cadre du cours de régression linéaire, la fonction f pourrait être une *vraisemblance* et les deux variables x et y deux paramètres du modèle.

Pour trouver un minimum local (en régression linéaire, on chercherait plutôt un *maximum* de la vraisemblance), on procède en deux temps. On cherche d'abord les points du graphe $z = f(x, y)$ de la fonction (voir la partie Maple pour un exemple) qui ont un plan tangent horizontal. Pour cela, on cherche les abscisses (x_0, y_0) qui annulent les deux dérivées partielles de la fonction à minimiser

$$\frac{\partial f}{\partial x}(x_0, y_0) \quad = \quad \frac{\partial f}{\partial y}(x_0, y_0) \quad = \quad 0\,. \tag{1}$$

Ensuite, pour chaque point à plan tangent horizontal d'abscisse (x_0, y_0) on détermine s'il s'agit d'un minimum local en étudiant la matrice H des dérivées partielles secondes, dite matrice *hessienne* de f, évaluée en (x_0, y_0)

$$H(x, y) \quad = \quad \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\,\partial y} \\ \frac{\partial^2 f}{\partial x\,\partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \quad = \quad \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}.$$

Si $H(x_0, y_0)$ est définie positive alors $(x_0, y_0, f(x_0, y_0))$ est un minimum local.

La raison se comprend bien en considérant le développement limité de la fonction f en un point (x_0, y_0).

$$\begin{aligned} f(x_0 + h, y_0 + \ell) \quad &= \quad f(x_0, y_0) + \underbrace{h\, f_x(x_0, y_0) + \ell\, f_y(x_0, y_0)}_{=0} \\ &\qquad + \underbrace{\frac{1}{2}\, h^2\, f_{xx}(x_0, y_0) + h\, \ell\, f_{xy}(x_0, y_0) + \frac{1}{2}\, \ell^2\, f_{yy}(x_0, y_0)}_{\text{terme d'ordre deux}} + \cdots \end{aligned}$$

et en remarquant que le terme d'ordre deux est de la forme $v^T\, A\, v$ (à un facteur 2 près) où $v^T = (h\ \ \ell)$ et $A = H(x_0, y_0)$:

$$h^2\, f_{xx}(x_0, y_0) + 2\, h\, \ell\, f_{xy}(x_0, y_0) + \ell^2\, f_{yy}(x_0, y_0) \quad = \quad \begin{pmatrix} h & \ell \end{pmatrix}\, H(x_0, y_0) \begin{pmatrix} h \\ \ell \end{pmatrix}. \tag{2}$$

Fixons l'abscisse (x_0, y_0). La variable, c'est le couple (h, ℓ) qui représente un déplacement par rapport à (x_0, y_0). On peut voir $(h\ \ell)$ comme les coordonnées d'un vecteur dans le plan tangent au point $(x_0, y_0, f(x_0, y_0))$. Dire que ce point est un minimum local, c'est dire que, quel que soit le vecteur $(h\ \ell)$ (qu'on peut supposer très petit puisqu'on cherche un minimum *local*) on a $f(x_0 + h, y_0 + \ell) > f(x_0, y_0)$ (la surface monte dans toutes les directions). C'est donc dire que le terme d'ordre deux est strictement positif quel que soit le vecteur $(h\ \ell)$ non nul. D'après l'identité (2), c'est dire que $H(x_0, y_0)$ est définie positive.

3 Calcul du Minimum

Il suffit d'appliquer la méthode Newton (voir DM 1) pour chercher une solution du système (1), c'est-à-dire un point (x_0, y_0) qui annule le vecteur *gradient* de f.

Il se trouve que la matrice jacobienne du vecteur gradient n'est rien d'autre que la matrice hessienne de f.

On voit donc que la matrice hessienne joue un double rôle ici : elle fournit la matrice jacobienne nécessaire à la méthode de Newton et elle permet de vérifier que la solution obtenue est bien un minimum local.

```
> restart;
  Digits := 40:
```

Minima locaux d'une fonction d'une variable

```
> f := x -> 2*x**3 - 9*x**2 + 12*x;
  df := D(f);
```

$$f := x \mapsto 2\,x^3 - 9\,x^2 + 12\,x$$

$$df := x \mapsto 6\,x^2 - 18\,x + 12 \qquad \textbf{(1.1)}$$

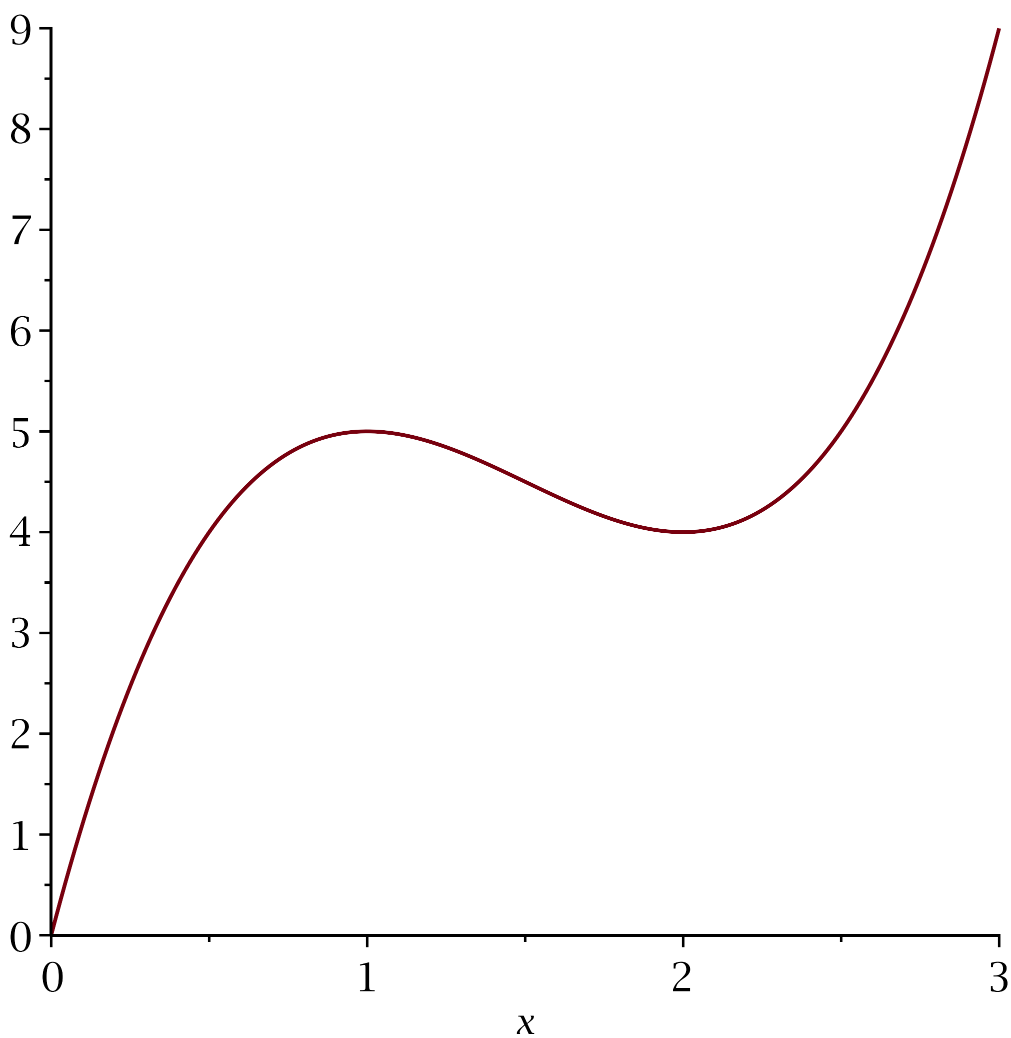
Pour trouver les minima locaux de f, on commence par chercher les points du graphe de f à tangente horizontale, c'est-à-dire ceux où la dérivée de f s'annule.

```
> solve (df(x));
```

$$2, 1 \qquad \textbf{(1.2)}$$

Le graphique confirme que x=1 et x=2 sont les abscisses de deux extrema locaux.

```
> plot (f(x), x=0..3);
```



On calcule un développement de Taylor de f(x0+h) avec x0=1.
On voit qu'il s'agit d'un maximum local au fait que la dérivée seconde de f est

négative en x0=1.

```
> taylor(f(1+h),h);
```

$$5-3\,h^2+2\,h^3 \tag{1.3}$$

On calcule un développement de Taylor de f(x0+h) avec x0=2.
On voit qu'il s'agit d'un minimum local parce que la dérivée seconde de f est positive en x0=2.

```
> taylor(f(2+h),h);
```

$$4+3\,h^2+2\,h^3 \tag{1.4}$$

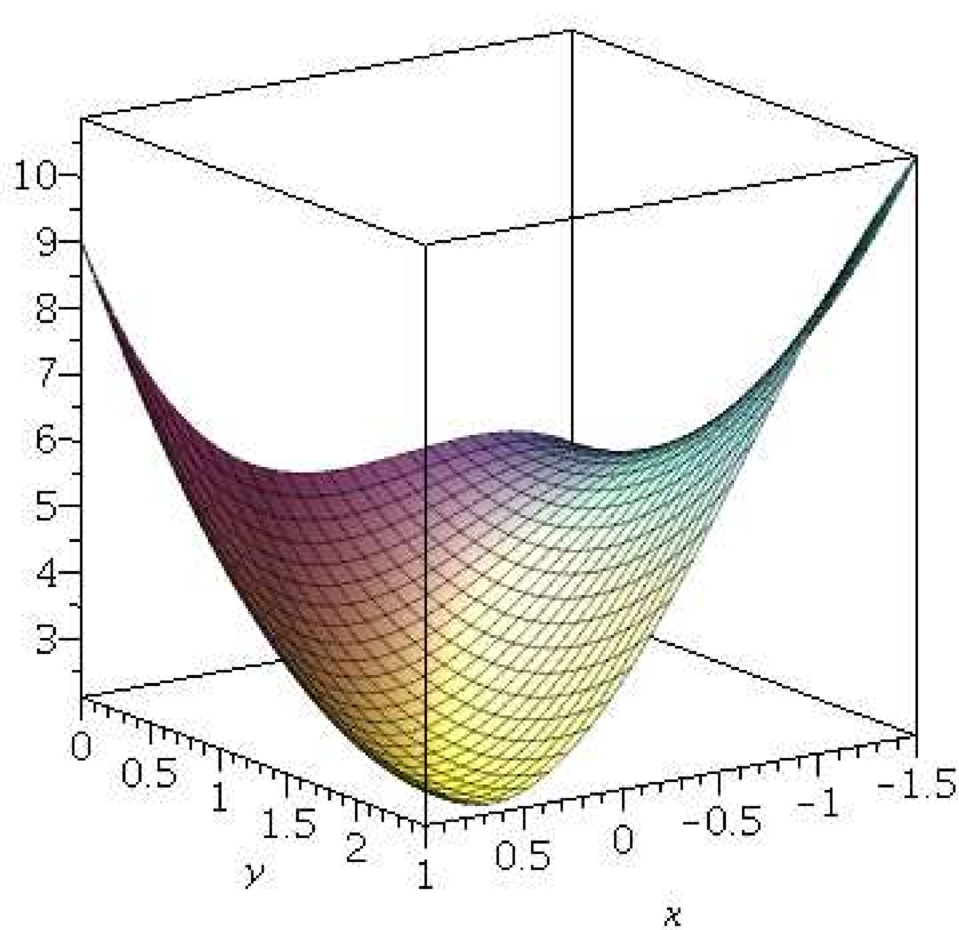
Minima locaux d'une fonction de plusieurs variables

```
> with(LinearAlgebra):
> f := x^3+2*x^2-2*x*y+y^2+x-3*y+5;
```

$$f:=x^3+2\,x^2-2\,x\,y+y^2+x-3\,y+5 \tag{2.1}$$

Le graphe de z = f(x,y) est une surface

```
> plot3d(f,x=-1.5..1,y=0..2.5);
```



Notons g1 et g2 les deux dérivées partielles de f par rapport à x et par rapport à y
On cherche les points du graphe à plan tangent horizontal, c'est-à-dire ceux où

les deux dérivées partielles s'annulent en même temps.

```
> g1 := diff(f,x);
  g2 := diff(f,y);
```

$$g1 := 3\,x^2 + 4\,x - 2\,y + 1$$

$$g2 := -2\,x + 2\,y - 3 \tag{2.2}$$

Il s'agit donc de résoudre un système de deux équations en deux variables réelles. On applique la méthode du DM 1. On calcule la matrice jacobienne de [g1,g2].

```
> J := VectorCalculus:-Jacobian([g1,g2],[x,y]);
```

$$J := \begin{bmatrix} 6\,x + 4 & -2 \\ -2 & 2 \end{bmatrix} \tag{2.3}$$

Parce que g1 et g2 sont les deux dérivées partielles de f, cette jacobienne est une matrice symétrique. C'est la matrice dite *hessienne* de f.

```
> H := VectorCalculus:-Hessian(f,[x,y]);
```

$$H := \begin{bmatrix} 6\,x + 4 & -2 \\ -2 & 2 \end{bmatrix} \tag{2.4}$$

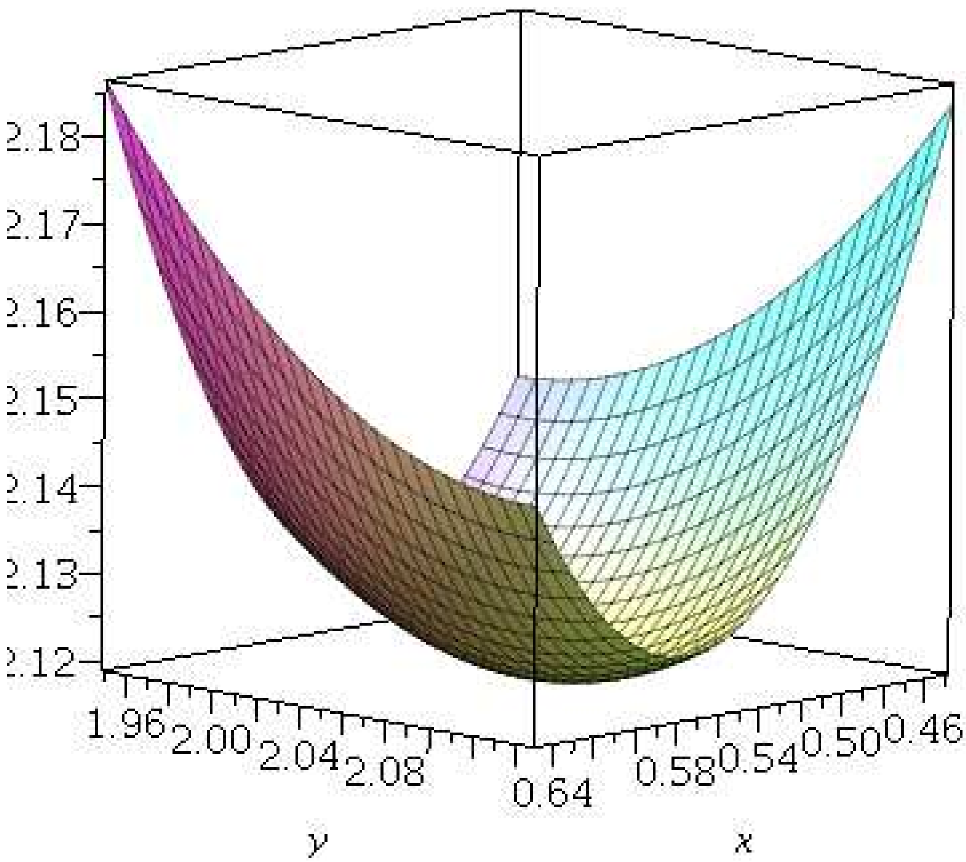
Appliquons la méthode de Newton en plusieurs variables

```
> u[0] := <0.2,3.>:
  for n from 0 to 5 do
      Hn := eval (H, {x=u[n][1],y=u[n][2]});
      mgn := eval (< -g1, -g2 >, {x=u[n][1],y=u[n][2]});
      z := LinearSolve (Hn, mgn):
      u[n+1] := u[n] + z
  end do:
> seq (u[n], n = 0 .. 6);
```

$$\begin{bmatrix} 0.2 \\ 3. \end{bmatrix}, \begin{bmatrix} 0.6625000000000000000000000000000000000000 \\ 2.162500000000000000000000000000000000000 \end{bmatrix}, \begin{bmatrix} 0.5550993723849372384937238493723849372385 \\ 2.055099372384937238493723849372384937238 \end{bmatrix}, \begin{bmatrix} 0.5486076624682115966509297811632678505327 \\ 2.048607662468211596650929781163267850533 \end{bmatrix}, \begin{bmatrix} 0.5485837706784866993782359854491430692818 \\ 2.048583770678486699378235985449143069282 \end{bmatrix}, \begin{bmatrix} 0.5485837703548635302265826888281235527008 \\ 2.048583770354863530226582688828123552701 \end{bmatrix}, \begin{bmatrix} 0.5485837703548635301672052512130868085722 \\ 2.048583770354863530167205251213086808572 \end{bmatrix} \tag{2.5}$$

Vérification graphique. Le résultat semble être un minimum.

```
> plot3d(f, x=u[n][1]-.1 .. u[n][1]+.1, y=u[n][2]-.1 .. u[n][2]
  +.1);
```



Vérification algébrique. Le point d'abscisses u[n] est bien un point à plan tangent horizontal.

```
> eval ([g1,g2], {x=u[n][1],y=u[n][2]});
```

$$[1.2\ 10^{-38}, 0.]$$ **(2.6)**

Vérification algébrique. La matrice hessienne est définie positive. Le point d'abscisses u[n] est bien un minimum local.

```
> Eigenvalues (Matrix(Hn,shape=symmetric));
```

$$\begin{bmatrix} 1.3291265207091907414227149024082626518 22 \\ 7.9623761014199904399367812305604786643 85 \end{bmatrix}$$ **(2.7)**

Questions.

Répondre en Python.
Remplacer l'appel à LinearSolve par des appels à des fonctions de plus bas niveau.

Vérifier par un algorithme étudié en cours que la matrice hessienne finale est définie positive.

Vérifier par ce même algorithme que l'autre extremum local n'est pas un minimum local (voir la section suivante pour obtenir ses coordonnées).

Refaire l'exercice qui montre qu'une matrice symétrique est définie positive si et seulement si ses valeurs propres sont strictement positives.

Résolution exacte

Dans cette section, on utilise des méthodes « avancées » qui ne sont pas au programme du cours pour calculer les extrema locaux de façon exacte. Ce sont les mêmes méthodes que celles du DM 1.

```
> B := Groebner:-Basis ([g1,g2], plex(y,x));
```

$$B := [3x^2 + 2x - 2, -2x + 2y - 3] \quad (3.1)$$

```
> readlib(realroot):
> solutions_intervalles := realroot (B[1], 10^(-Digits));
```

$$solutions\_intervalles := \left[\left[-\frac{26465810887129034144824202030303359974513}{21778071482940061661655974875633165533184}, -\frac{105863243548516136579296808121213439898051}{87112285931760246646623899502532662132736}\right], \left[\frac{5973548282584496518526776056607291476195}{10889035741470030830827987437816582766592}, \frac{47788386260675972148214208452858331809561}{87112285931760246646623899502532662132736}\right]\right] \quad (3.2)$$

```
> abscisses := [seq (evalf (solutions_intervalles[i][1]), i=1..2)
  ];
```

$$abscisses := [-1.215250437021530196833871917879753475237, 0.5485837703548635301672052512130868085701] \quad (3.3)$$

```
> ordonnees := [seq (solve (eval (B[2],x=abscisses[i])), i=1..2)
  ];
```

$$ordonnees := [0.2847495629784698031661280821202465247630, 2.048583770354863530167205251213086808570] \quad (3.4)$$

Les coordonnées (x,y) des deux extrema locaux

```
> solutions := [seq (<abscisses[i], ordonnees[i]>, i = 1 .. 2)];
```

$$solutions := \left[\begin{bmatrix} -1.215250437021530196833871917879753475237 \\ 0.2847495629784698031661280821202465247630 \end{bmatrix}, \right. \quad (3.5)$$

$$\begin{bmatrix} 0.5485837703548635301672052512130868085701 \\ 2.048583770354863530167205251213086808570 \end{bmatrix}$$

Vérification

```
> seq (eval ([g1,g2], {x=abscisses[i], y=ordonnees[i]}), i=1..2);
```

$$[1.3\ 10^{-39}, 0.], [0., 0.] \tag{3.6}$$