UNIVERSITE DE GENEVE — Section de mathématiques

Faculté des sciences — Mercredi 22 novembre 2006

Optimisation – Série 3 Corrigé

Exercice 1 a) *2 pts* Étude de $f_1(x,y) = x^2 + y^2 - xy$. Le gradient est $(2x - y, -x + 2y)$. Il vaut 0 uniquement en $(0,0)$. En $(0,0)$, la Hessienne de f_1 vaut

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Cette matrice est définie positive. Le seul minimum local possible se trouve donc en $(0,0)$. Comme $\lim\limits_{|(x,y)|\to+\infty} f_1(x,y) = +\infty$, le point $(0,0)$ est un minimum global.

b) *2 pts* Étude de $f_2(x,y) = x^4 + y^4 - \frac{3}{2}x^2y^2$. Le gradient vaut $(4x^3 - 3xy^2, -3y^2x - 4y^3)$, Il ne peut être nul qu'au point $(0,0)$. En ce point la Hessienne est nulle donc positive mais non définie. Mais, $f_2(0,0) = 0$ et $f_2(x,y) = (x^2 - y^2)^2 + \frac{1}{2}x^2y^2$ est strictement positif ailleurs qu'en $(0,0)$. La fonction $f_2(x,y)$ atteint son minimum global au point $(0,0)$.

c) *2 pts* Étude de $f_3(x,y) = 3x^2 - 3y^2 + 8xy$. Calculons le gradient de f_3, il vaut $(6x + 8y, 8x - 6y)$. Leur gradient ne peut-être nul qu'en $(0,0)$. La Hessienne est

$$\begin{bmatrix} 6 & 8 \\ 8 & -6 \end{bmatrix}$$

Elle a une valeur propre positive et une valeur propre négative. Il n'y a pas de minimum global. Le point $(0,0)$ est un point selle.

d) *6 pts* $f_4(x,y) = \exp(\sin(50x)) + \sin(60\exp(y)) + \sin(70\sin(x)) + \sin(\sin(80y)) - \sin(10(x+y)) + \frac{x^2+y^2}{4}$. Il est hors de question d'étudier cette fonction. Il y aurait un nombre bien trop important de minimum locaux. Nous allons employer Maple et Matlab.

```
>> f4=@(x,y) exp(sin(50*x))+sin(60*exp(y))+sin(70*sin(x))...
       +sin(sin(80*y))-sin(10*(x+y))+(x.^2+y.^2)/4;
>> [X,Y]=meshgrid(-20:0.04:20,-20:0.04:20);
>> mesh(X,Y,f4(X,Y)); %Regard initial
>> [X,Y]=meshgrid(-10:0.01:10,-10:0.01:10);
>> mesh(X,Y,f4(X,Y));
>> [X,Y]=meshgrid(-2.5:0.005:2.5,-2.5:0.005:2.5);
```

```
>> mesh(X,Y,f4(X,Y));
>> [X,Y]=meshgrid(-1:0.002:1,-1:0.002:1);
>> mesh(X,Y,f4(X,Y));
>> [X,Y]=meshgrid(-0.2:0.001:0.2,0:0.001:0.5);
>> mesh(X,Y,f4(X,Y));
>> [X,Y]=meshgrid(-0.1:0.005:0,0.15:0.005:0.25);
>> figure
>> mesh(X,Y,f4(X,Y));
>> [X,Y]=meshgrid(-0.05:0.005:0,0.18:0.005:0.23);
>> figure
>> mesh(X,Y,f4(X,Y));
```

On a trouvé le minimum dans le rectangle $[0.32, 0.36] \times [0.34, 4]]$. Nous pouvons maintenant utiliser Maple avec

```
> f:=exp(sin(50*x))+sin(60*exp(y))+sin(70*sin(x))
            +sin(sin(80*y))-sin(10*(x+y))+(x^2+y^2)/4;
> fsolve({diff(f,x)=0,diff(f,y)=0},{x=-0.05..0,y=0.18..0.23});
     {x = -0.02440307969, y = 0.2106124272}
> unapply(f, x, y)(-0.02440307969,0.2106124272);
                                          -3.306868653
```

Le minimum de $f4$, -3.306868653 est donc atteint en $-0.02440307969, 0.2106124272$.

Exercice 2 a) *1 pt Entre deux noeuds, quelle est la forme prise par une route de longueur minimale ? Quelle est la courbe de longueur minimale reliant deux points sur le plan ?* Sur un plan doté de la norme euclidienne, c'est un segment. Par définition de la longueur d'une courbe. En effet

$$|\gamma(1) - \gamma(0)| = \left| \int_0^1 \gamma'(t)\mathrm{d}t \right| \leq \int_0^1 |\gamma'(t)|\,\mathrm{d}t$$

b) *1 pt* Supposons qu'un point de croisement C ne soit que le point de concourrance de 2 routes, il ne sert alors qu'à relier deux autres nœuds : A et B. Dans ce cas supprimer le point de croisement et remplacer les deux routes qui y étaient connectées par le segment $[AB]$ donne un réseau de longueur inférieure.

c) *2 pts* À chaque ville est associée au moins 1 route. À chaque point de croisement, au moins 3. Or le nombre de routes vaut :

$$r = \frac{1}{2} \sum_{i \in \text{noeuds}} \text{nombre de routes reliées au nœud } i$$

Nous avons divisé par 2 car chaque route a deux extrémités et est comptée deux fois. Donc $r \geq \frac{3p+n}{2}$.

d) *3pts* Supposons qu'il y ait $N = p+n$ noeuds dans un réseau. Le réseau de longueur minimal ne peut avoir de cycle, sinon une des routes du réseau serait superflue. Un réseau sans cycle contient au moins un nœud connecté par une seule route.

En effet si ce n'était pas le cas, on pourrait construire un cycle en partant d'un nœud quelconque puis en suivant les routes non encore prises. Comme, aucun nœud n'est un cul-de-sac, on peut toujours avancer. On finit par retomber sur un nœud déjà traversé car le nombre de nœuds est fini.

On peut maintenant raisonner par récurrence sur le nombre de nœuds. Pour 2 nœuds, nous avons besoin d'une seule route. Supposons qu'un réseau de longueur minimale sur N noeuds est toujours constitué de $N-1$ routes. Soit un réseau de longueur minimal sur $N+1$ noeuds. Soit a le noeud relié par une unique route r au reste du réseau. Enlevons a et r du réseau. Le réseau restant est un réseau de longueur minimale sur N noeuds et comporte donc $N-1$ routes. On rajoute la route enlevée et nous obtenons bien nos N routes.

e) *1 pt* Nous avons l'inégalité $\frac{3p+n}{2} \leq r = n+p-1$. D'où $p \leq n-2$.

f) On regarde les villes suivantes disposées sur un carré où chaque côté a pour longueur a :

Supposons P un point de croisement dans un réseau de longueur connecté à k autre nœuds A_i. Déplacer P ne doit pas diminuer la longueur du réseau. Calculer le gradient de $\sum_{i=1}^{k} |\vec{PX_i}|$ donne comme condition de stationnarité :

$$\sum_{i=1}^{k} \frac{\vec{PX_i}}{PX_i} = 0$$

On pourra vérifier que si $k=3$, cela signifie que P doit être à l'unique point dans le triangle $X_1X_2X_3$ telle que les angles (PX_1, PX_2) (PX_2, PX_3) (PX_3, PX_1) soit tous égaux à $\frac{2\pi}{3}$ ou à $-\frac{2\pi}{3}$. On vérifie aussi que pour $k=4$, cela signifie que P doit être situé à l'intersection entre deux segments $[X_iX_j]$ et $[X_lX_m]$ avec i, j, l, m tous distincts.

i. *2 pts Trouver le réseau de longueur minimal avec $p = 0$ (facile).* S'il n'y pas de point de croisement, un réseau minimal est constitué par le choix de trois des quatre arêtes du carré. Cela donne une longueur de $3a$.

ii. *4 pts Trouver le réseau de longueur minimal avec $p = 1$ (moins facile).* Il y a deux choix. Un point de croisement connecté à trois villes, disons les villes A, C et D. Le point de croisement doit alors se situer en $(\frac{3\sqrt{2}-\sqrt{6}}{6}a, \frac{3\sqrt{2}-\sqrt{6}}{6}a)$ Ce qui donne une longueur du réseau de $\frac{a}{2}(2+\sqrt{6}+\sqrt{2}) \approx 2.93a$.

Ou, mieux, le point de croisement est connecté aux quatre villes et se situe au centre du carré (croisement des diagonales). Ce qui donne une longueur de réseau égale à $2\sqrt{2}a \approx 2.83a$ et une meilleur solution.

iii. *4 pts Quel est le meilleur réseau que vous pouvez imaginer avec $p = 2$?* C'est le cas limite, donc chaque point de croisement ne peut voir se rejoindre que 3 routes et chaque ville une seule au maximum. Les deux points de croisements sont reliés par une route sinon il y aurait une ville au moins avec deux routes. Pour respecter les contraintes d'angles $\frac{2\pi}{3}$, le seul réseau possible est de longueur $(\sqrt{3}+1)a \approx 2.73a$

iv. *2 pts Quel est le réseau optimal ?* C'est le réseau optimal obtenu avec $p = 2$.

Exercice 3 a) *5 pts* Avec la barre centrale (voir figure 3) Cela fait une longueur de $2h + (4+\pi)r$ et une surface de $\frac{\pi}{2}r^2 + 2rh$

Il faut donc résoudre le problème d'optimisation suivant :

$$\begin{aligned} \frac{\pi}{2}r^2 + 2rh &\to \max \\ 2h + (4+\pi)r &= l \end{aligned}$$

On cherche donc (r, h, λ) telle que

$$\begin{aligned} \pi r + 2h - (4+\pi)\lambda &= 0 \\ 2r - 2\lambda &= 0 \\ (4+\pi)r + 2h &= l \end{aligned}$$

L'unique point vérifiant ces conditions est $(r, h, \lambda) = (\frac{l}{8+\pi}, \frac{2l}{8+\pi}, \frac{l}{8+\pi})$.

Sans la barre centrale, voir figure 3. Cela fait une longueur de $2h + (2+\pi)r$ et une surface de $\frac{\pi}{2}r^2 + 2rh$. Il faut résoudre

$$\begin{aligned} \frac{\pi}{2}r^2 + 2rh &\to \max \\ 2h + (2+\pi)r &= l \end{aligned}$$

On cherche donc (r, h, λ) telle que

$$\begin{aligned} \pi r + 2h - (2+\pi)\lambda &= 0 \\ 2r - 2\lambda &= 0 \\ (2+\pi)r + 2h &= l \end{aligned}$$

L'unique point vérifiant ces conditions est $(r, h, \lambda) = (\frac{l}{4+\pi}, \frac{l}{4+\pi}, \frac{l}{4+\pi})$.

b) *10 pts* On veut résoudre

$$\sum_{ij} \sigma_{ij}\pi_i\pi_j \to \min$$
$$\pi_a + \pi_b + \pi_c = 1$$
$$\pi_a + 2.2\pi_b + 2.7\pi_c \geq 2$$
$$\pi_a \geq 0$$
$$\pi_b \geq 0$$
$$\pi_c \geq 0$$

On commence par remplacer π_c par $1-\pi_a-\pi_b$. La fonction à minimiser devient :

$$\sigma_{cc} + 2(\sigma_{ac} - \sigma_{cc})\pi_a + 2(\sigma_{bc} - \sigma_{cc})\pi_b + (\sigma_{aa} + \sigma_{cc} - 2\sigma_{ac})\pi_a\pi_a + (\sigma_{bb} + \sigma_{cc} - 2\sigma_{bc})\pi_b\pi_b + 2(\sigma_{ab} + \sigma_{cc} - \sigma_{bc} - \sigma_{ac})\pi_a\pi_b$$

Cela donne le problème d'optimisation équivalent suivant :

$$f(\pi_a, \pi_b) = -3.9\pi_a - 3.84\pi_b + 2{\pi_a}^2 + 2.34{\pi_b}^2 + 3.78\pi_a\pi_b \to \min$$
$$\pi_a + \pi_b \leq 1$$
$$1.7\pi_a + 0.5\pi_b \leq 0.7$$
$$\pi_a \geq 0$$
$$\pi_b \geq 0$$

Remarquons d'abord que la Hessienne de f est constante et toujours définie positive. Le domaine sur lequel on cherche à minimiser est donné sur la figure 4.
Calculons la dérivée de f.

$$\nabla f = \begin{bmatrix} -3.9 + 4\pi_a + 3.78\pi_b \\ -3.84 + 3.78\pi_a + 4.68\pi_b \end{bmatrix}$$

Le minimum sur le plan est donc situé en $(0.2727, 0.746)$ qui est en dehors du domaine de calcul.
Calculons le minimum sur le segment $[OC]$. Il y a minimum quand

$$\pi_a = 0$$
$$-3.9 + 3.78\pi_b - \lambda = 0$$
$$-3.84 + 4.68\pi_b = 0$$

donc le minimum sur la droite (OC) est situé en $(0, 0.8205)$ (point appartenant aussi segment) avec une valeur approximative de -1.5754.

Calculons le minimum sur le segment $[0A]$. Le minimum sur la droite (OA) est atteint quand

$$\begin{aligned} \pi_b &= 0 \\ -3.9 + 4\pi_a &= 0 \\ -3.84 + 3.78\pi_a - \lambda &= 0 \end{aligned}$$

Après calcul, cela donne $\pi_a = 0.9750$ qui n'appartient pas au segment $[OA]$.

Calculons le minimum sur le segment AB, il y a minimum quand

$$\begin{aligned} 4\pi_a + 3.78\pi_b - 1.7\lambda &= 3.9 \\ 3.78\pi_a + 4.68\pi_b - 0.5\lambda &= 3.84 \\ 1.7\pi_a + 0.5\pi_b &= 0.7 \end{aligned}$$

L'unique solution est $(\pi_a, \pi_b, \lambda) = (0.2417, 0.57837, -0.4394)$. C'est bien à l'intérieur du segment. La valeur atteinte est -1.7355.

Calculons le minimum sur le segment $[BC]$, il y a minimum

$$\begin{aligned} -3.9 + 4\pi_a + 3.78\pi_b - \lambda &= 0 \\ -3.84 + 3.78\pi_a + 4.68\pi_b - \lambda &= 0 \\ \pi_a + \pi_b &= 1 \end{aligned}$$

L'unique solution est $(\pi_a, \pi_b, \lambda) = (0.8571, 0.1429, 0.0686)$. C'est en dehors du segment $[BC]$.

De plus, $f(A) = -1.2667$, $f(B) = -1.6444$, $f(C) = -1.5$ et $f(O) = 0$.

Donc le risque minimum est atteint sur le segment $[AB]$ au point $(\pi_a, \pi_b, \pi_c) = (0.2417, 0.57837, 0.17996)$. Nous revenons à la fonction initiale (que nous avions transformée pour faciliter les calculs) et obtenons une variance de 0.05142.

Exercice 4 *10 pts* Soit $\psi : \mathcal{R}^{n-m} \supset U \to \omega \cap \mathcal{M}$ une paramétrisation autour de $\boldsymbol{x}$ avec $\psi(\boldsymbol{u}) = \boldsymbol{x}$. Nous considérons la fonction $f \circ \psi$. Nous savons que :

$$\begin{aligned} & \begin{cases} \nabla_{\boldsymbol{x}}(f \circ \psi)(\boldsymbol{u}^*) = 0 \\ \boldsymbol{w}^T (f \circ \psi)'' \boldsymbol{w} > 0 \quad \forall \boldsymbol{w} \in \mathbb{R}^{n-m} \end{cases} \\ \Longrightarrow \quad & f \circ \psi \text{ possède un minimum local en } \boldsymbol{u}^* \\ \Longrightarrow \quad & \begin{cases} \nabla_{\boldsymbol{x}}(f \circ \psi)(\boldsymbol{u}^*) = 0 \\ \boldsymbol{w}^T (f \circ \psi)'' \boldsymbol{w} \geq 0 \quad \forall \boldsymbol{w} \in \mathbb{R}^{n-m} \end{cases} \end{aligned}$$

Pour avoir notre résultat, il suffit de démontrer que chacune des lignes de ce résultat connu est équivalent à la ligne respective de l'implication que nous souhaitons démontrer.

D'abord, que f admette un minimum local en $\boldsymbol{x}^*$ sur $\mathcal{M}$ est équivalent à ce que $f \circ \psi$ admette un minimum local en $\boldsymbol{u}^*$ quand $\psi(\boldsymbol{u}^*) = \boldsymbol{x}^*$.

Nous avons aussi :

$$(f \circ \psi)' = f'(\psi)\psi'$$
$$(f \circ \psi)'' = f''(\psi)(\psi', \psi') + f'(\psi)\psi''$$

Donc $(f \circ \psi)'(\boldsymbol{u}) = 0$ équivaut à ce que la forme linéaire $f'(\psi(\boldsymbol{u}))$ soit nulle sur l'espace tangeant à $\mathcal{M}$ en $\boldsymbol{x}$ donc à ce que $f'(\boldsymbol{x})\boldsymbol{w} = 0$ pour tout $\boldsymbol{w}$ telle que $g'(\boldsymbol{x})\boldsymbol{w} = 0$. *I.e.* $\cap_{i=1}^m \mathrm{Ker}(g_i'(\boldsymbol{x})) \subset \mathrm{Ker}(f'(\boldsymbol{x})) \subset$, ce qui est équivalent à ce qu'il existe $\boldsymbol{\lambda}$ dans $\mathbb{R}^m$ tel que $f'(\boldsymbol{x}) - \sum_i \lambda_i g_i'(\boldsymbol{x}) = 0$.

Remarquons que pour tout i, $1 \leq i \leq m$, nous avons l'identité $g_i'(\psi(\boldsymbol{u}))(\psi''(\boldsymbol{u})(\boldsymbol{h}, \boldsymbol{h})) = g_i''(\psi(\boldsymbol{u}))(\psi'(\boldsymbol{u})\boldsymbol{h}, \psi'(\boldsymbol{u})\boldsymbol{h})$. Cela vient de $g_i \circ \psi = 0$ et de $(g_\circ \psi)'' = 0$.

Que la forme bilinéaire $(f \circ \psi)''$ en un point $\boldsymbol{u}$ tel que $(f \circ \psi)'(\boldsymbol{u}) = 0$ soit positive est équivalent à ce que $f''(\boldsymbol{x})(\boldsymbol{h}, \boldsymbol{h}) - \sum_{i=1}^m \lambda_i g_i'(\psi(\boldsymbol{u}))(\psi''(\boldsymbol{u})(\boldsymbol{h}, \boldsymbol{h})) \geq 0$ et donc à ce que $(f - \sum_{i=1}^m \lambda_i g_i)''(\boldsymbol{x})(\boldsymbol{w}, \boldsymbol{w}) \geq 0$ pour tout vecteur $\boldsymbol{w}$ tangeant à $\mathcal{M}$ en $\boldsymbol{x}$.

De même, que la forme bilinéaire $(f \circ \psi)''$ en un point $\boldsymbol{u}$ tel que $(f \circ \psi)'(\boldsymbol{u}) = 0$ soit définie positive est équivalent à ce que $f''(\boldsymbol{x})(\boldsymbol{h}, \boldsymbol{h}) - \sum_{i=1}^m \lambda_i g_i'(\psi(\boldsymbol{u}))(\psi''(\boldsymbol{u})(\boldsymbol{h}, \boldsymbol{h})) > 0$ et donc à ce que $(f - \sum_{i=1}^m \lambda_i g_i)''(\boldsymbol{x})(\boldsymbol{w}, \boldsymbol{w}) > 0$ pour tout vecteur $\boldsymbol{w}$ tangeant à $\mathcal{M}$ en $\boldsymbol{x}$.

Or f minimum local en $\boldsymbol{x}$ sur la surface $\mathcal{M}$ équivaut à ce que $f \circ \psi$ soit un minimum local en $\boldsymbol{u}$. Nous pouvons donc conclure en reliant ces équivalences aux conditions suffisantes et aux conditions nécessaire pour la minimisation sans contraintes de $f \circ \psi$.

Evaluation du cours d'optimisation

- Les exercices : Les séries d'exercices rendues en retard seront comptées comme non rendues (*i.e.*, note 1 sur 6) dans le calcul de la note finale.
- Un examen oral durant la session d'examens sur le cours.

La note finale est de : $\frac{1}{5}$(exercices.) + $\frac{4}{5}$(note examen oral).

Assistant :	Kévin Santugini
Email :	`Kevin.Santugini@math.unige.ch`
Page web :	`http://www.unige.ch/~santugin/index.php?page=enseignement`

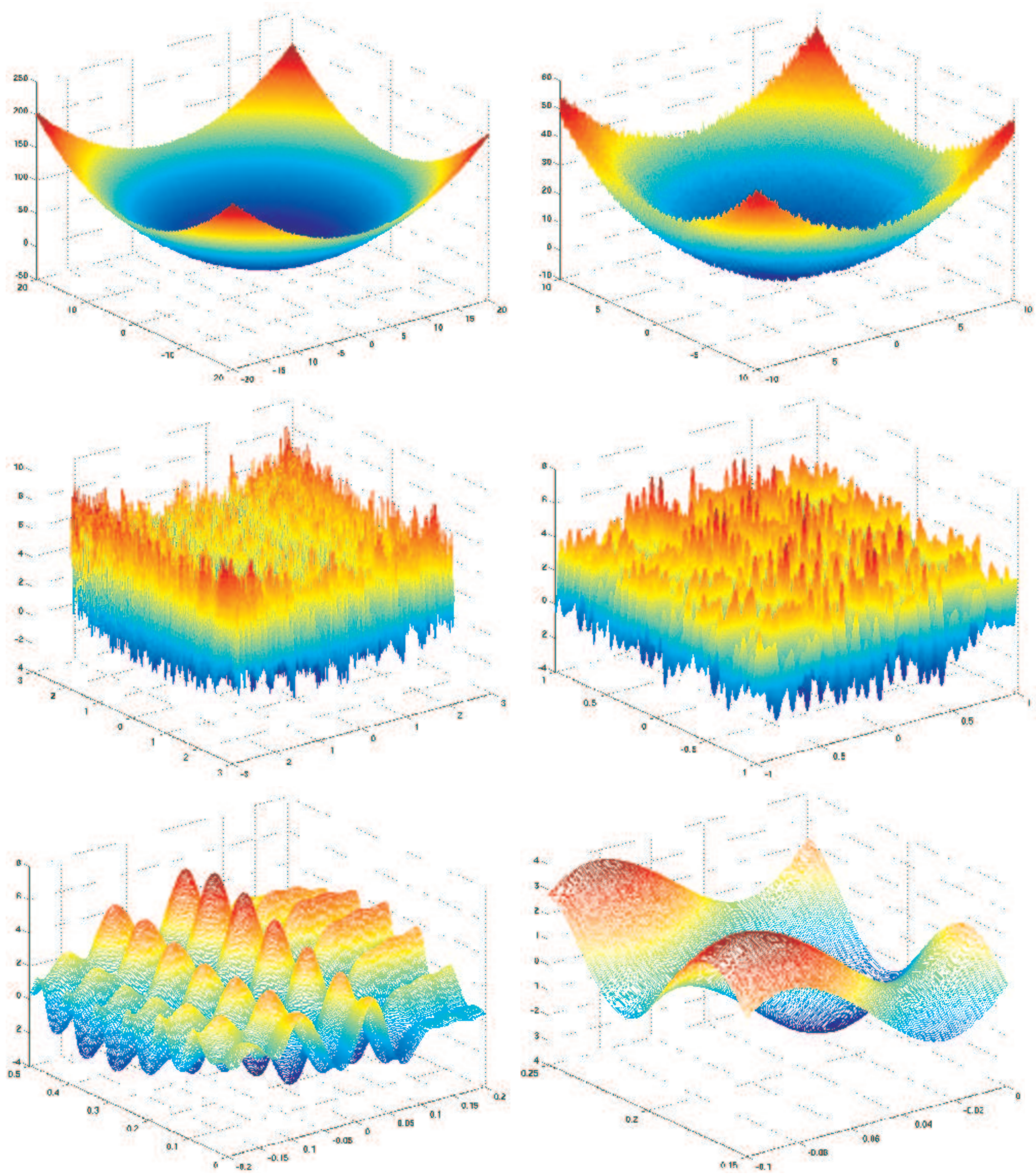


Fig. 1 – Exercice 1d, Vue de la fonction

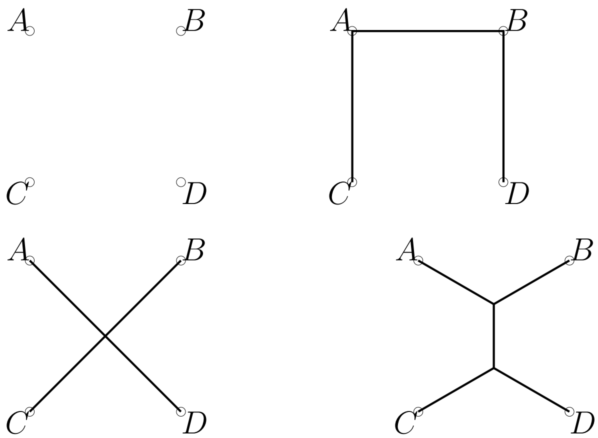
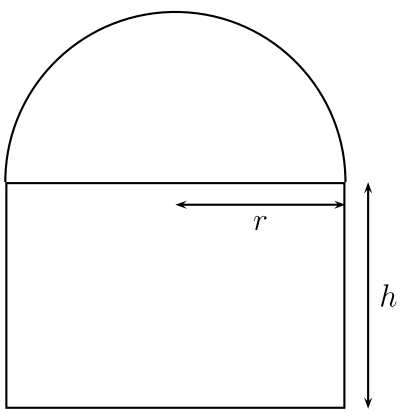


Fig. 2 – Exercice 2



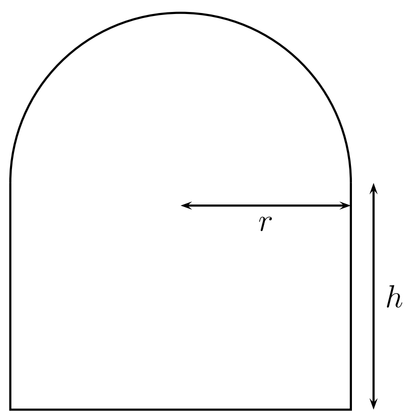


Fig. 3 – Exercice 3 : fenêtres avec et sans barre centrale

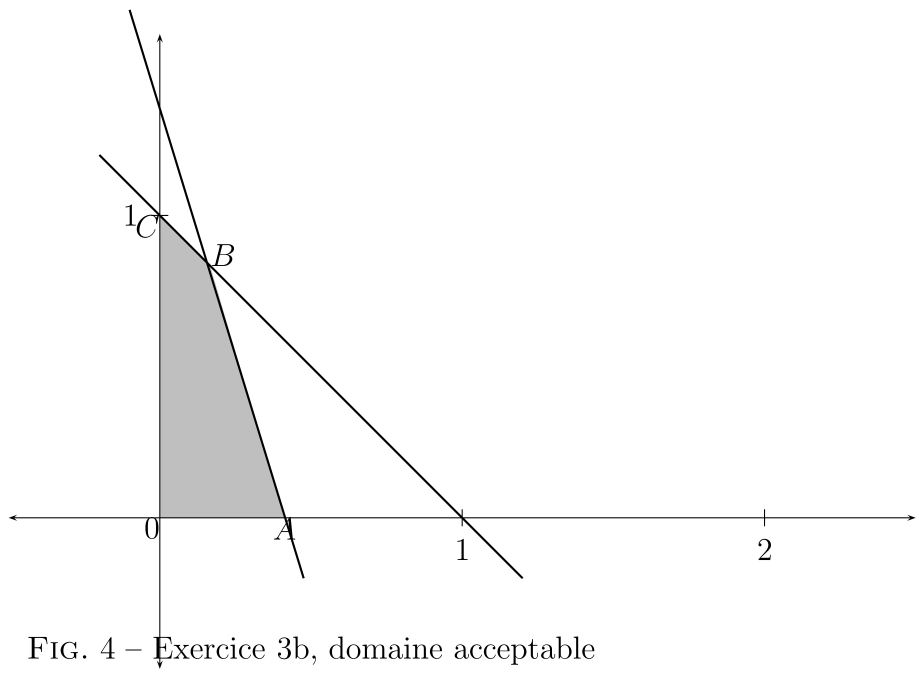


Fig. 4 – Exercice 3b, domaine acceptable