Corrigé du contrôle de connaissances de MDI 210

Durée : 3 h.
Documents autorisés : deux feuilles recto verso (soit quatre pages), format A4 ; dictionnaire autorisé pour les élèves dont le français n'est pas la langue maternelle.
Les calculatrices et les ordinateurs sont interdits, ainsi que tout objet permettant de communiquer avec l'extérieur.
L'épreuve est constituée de trois exercices indépendants.
Un résultat non justifié pourra ne pas être considéré comme juste. Un résultat obtenu autrement que par les méthodes indiquées dans l'énoncé et plus généralement par des méthodes autres que celles du cours pourra ne pas être considéré comme juste.
On détaillera les calculs effectués, même si cela n'est pas demandé explicitement.
Le barème, sur 40, n'est donné qu'à titre indicatif et pourra être modifié.

Exercice 1 (13 points)
On considère la matrice A définie par :

$$A = \begin{pmatrix} 0 & 0 & 12 & 0 \\ 0 & -1 & 0 & 2 \\ 12 & 0 & 7 & 0 \\ 0 & 2 & 0 & -1 \end{pmatrix}.$$

1. (7,5 points) Déterminer les valeurs propres et une base orthonormée de vecteurs propres de A à l'aide de la méthode vue en cours et que l'on identifiera ; on détaillera les calculs et on indiquera clairement, pour chaque vecteur propre, quelle valeur propre lui est associée.

Indication. Dans les calculs, les racines carrées se simplifient sauf $\sqrt{2}$.

Corrigé

On applique la méthode de Jacobi. On reprend les notations du cours.

* 1^{re} étape : $p = 1$, $q = 3$; $x = (a_{qq} - a_{pp})/2a_{pq} = 7/24$; on résout l'équation $t^2 + 2xt - 1 = 0$: la bonne racine vaut $t = 3/4$, d'où $c = 4/5$ et $s = 3/5$. Puis on applique les formules :

- pour $i \notin \{p, q\}$ et $i \notin \{p, q\}$, $b_{ij} = a_{ij}$;
- pour $i \notin \{p, q\}$, $b_{pi} = b_{ip} = c.a_{pi} - s.a_{qi}$ et $b_{qi} = b_{iq} = s.a_{pi} + c.a_{qi}$;
- $b_{pp} = a_{pp} - t.a_{pq}$ et $b_{qq} = a_{qq} + t.a_{pq}$.

À l'issue de cette étape, on obtient la matrice $\begin{pmatrix} -9 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 16 & 0 \\ 0 & 2 & 0 & -1 \end{pmatrix}$. La matrice Ω vaut

$$\Omega_1 = \begin{pmatrix} 4/5 & 0 & 3/5 & 0 \\ 0 & 1 & 0 & 0 \\ -3/5 & 0 & 4/5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

* 2^{de} étape : $p = 2$, $q = 4$; $x = (a_{qq} - a_{pp})/2a_{pq} = 0$; on résout l'équation $t^2 + 2xt - 1 = 0$: la bonne racine vaut $t = 1$, d'où $c = 1/\sqrt{2}$ et $s = 1/\sqrt{2}$. Puis on applique les formules :

- pour $i \notin \{p, q\}$ et $i \notin \{p, q\}$, $b_{ij} = a_{ij}$;
- pour $i \notin \{p, q\}$, $b_{pi} = b_{ip} = c.a_{pi} - s.a_{qi}$ et $b_{qi} = b_{iq} = s.a_{pi} + c.a_{qi}$;
- $b_{pp} = a_{pp} - t.a_{pq}$ et $b_{qq} = a_{qq} + t.a_{pq}$.

À l'issue de cette étape, on obtient la matrice $\begin{pmatrix} -9 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. La matrice Ω vaut

$$\Omega_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 0 & 1 & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}.$$

La matrice courante est diagonale : on a fini. Les valeurs propres sont –9, –3, 16 et 1. On obtient une base orthonormée de vecteurs propres grâce au produit $\Omega_1\Omega_2$:

$$\Omega_1\Omega_2 = \begin{pmatrix} 4/5 & 0 & 3/5 & 0 \\ 0 & 1 & 0 & 0 \\ -3/5 & 0 & 4/5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 0 & 1 & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 4/5 & 0 & 3/5 & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -3/5 & 0 & 4/5 & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}.$$

Les colonnes de cette matrice définissent les vecteurs propres de la base cherchée, associés aux valeurs propres dans l'ordre dans lequel elles sont énumérées ci-dessus.
Fin de corrigé

2. (1,5 points) Comment pourrait-on en déduire la matrice inverse de A, si A est inversible (si tel est le cas, il n'est pas demandé de calculer la matrice inverse).
Corrigé
Les valeurs propres étant non nulles, on constate d'abord que A est en effet inversible. Appelons D la matrice diagonale obtenue ci-dessus et posons $\Omega = \Omega_1\Omega_2$. On a donc $D = \Omega^{-1}A\Omega$, ou encore $A = \Omega D\Omega^{-1}$; d'où $A^{-1} = (\Omega D\Omega^{-1})^{-1} = \Omega D^{-1}\Omega^{-1}$, ou encore $A^{-1} = \Omega D^{-1}\Omega^{t}$, puisque Ω est une matrice orthogonale ; dans cette formule, D^{-1} s'obtient immédiatement : c'est la matrice diagonale dont les termes diagonaux sont les inverses des termes diagonaux de D.
Fin de corrigé

3. (1 point) On pose $X = (x_1, x_2, x_3, x_4)^t \in \mathbf{R}^4$. Soient a, b, c et d quatre réels. On considère la forme quadratique Q définie sur $\mathbf{R}^4$ par :

$$Q(X) = \frac{1}{2}\left(-x_2^2 + 7x_3^2 - x_4^2\right) + 12x_1x_3 + 2x_2x_4 + ax_1 + bx_2 + cx_3 + dx_4 .$$

Quels sont les quadruplets (a, b, c, d) pour lesquels Q est convexe ?
Corrigé
On remarque que $Q(X)$ s'écrit aussi $Q(X) = \frac{1}{2}X^tAX + \begin{pmatrix} a & b & c & d \end{pmatrix}\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix}^t$; la matrice hessienne de Q est donc égale à A. Les valeurs propres de A n'étant pas toutes positives ou nulles, Q n'est convexe pour aucun quadruplet (a, b, c, d).
Fin de corrigé

4. (1 point) Un élève de Télécom Paris envisage d'appliquer la méthode de Newton en partant de $X = 0$ afin de déterminer un extremum (minimum ou maximum) local de $Q(X)$ sur $\mathbf{R}^4$. Pour quels quadruplets (a, b, c, d) peut-on l'encourager dans ce projet (on justifiera la réponse) ?

Corrigé

La matrice hessienne de Q n'étant ni définie positive (présence d'au moins une valeur propre négative) ni définie négative (présence d'au moins une valeur propre positive), on ne peut encourager cet élève pour aucun quadruplet (a, b, c, d).

Fin de corrigé

5. (2 points) Un autre élève de Télécom Paris veut annuler le gradient de Q.

a. Quel système linéaire doit-il résoudre pour cela (il n'est pas demandé de résoudre ce système) ?

Corrigé

Le gradient de Q appliqué à X vaut $AX + (a, b, c, d)^t$. Annuler le gradient de Q revient donc à résoudre le système $AX = -(a, b, c, d)^t$.

Fin de corrigé

b. Que vaut le conditionnement pour la norme 2 du système linéaire de la question précédente ?

Corrigé

Puisque A est une matrice symétrique (donc normale) et inversible, le conditionnement pour la norme 2 du système linéaire précédent est donné par le rapport entre la plus grande valeur absolue des valeurs propres de A et la plus petite valeur absolue des valeurs propres de A, soit ici $16/1 = 16$.

Fin de corrigé

c. Cet élève peut-il utiliser la méthode de Cholesky pour résoudre ce système linéaire (on justifiera la réponse ; si celle-ci est positive, il n'est pas demandé d'appliquer cette méthode) ?

Corrigé

La méthode de Cholesky s'applique à des matrices symétriques définies positives, ce que n'est pas A (puisque A possède au moins une valeur propre négative). L'élève ne peut donc pas appliquer cette méthode.

Fin de corrigé

Exercice 2 (11 points)

On considère le problème (P) d'optimisation linéaire suivant :

$$(P) \qquad \begin{array}{l} \text{Maximiser } z = x_1 + 5x_2 - 4x_3 \\ \text{avec les contraintes } \left\{ \begin{array}{c} x_1 + x_3 \leq 4 \\ x_1 + x_2 - x_3 \leq 7 \\ x_1 - x_2 + 2x_3 \leq 5 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{array} \right. \end{array}$$

1. (5 points) Résoudre (P) à l'aide de l'algorithme du simplexe, en appliquant le premier critère de Dantzig.

Rappel : le premier critère de Dantzig consiste à choisir comme variable entrante la variable hors-base pourvue du plus grand coefficient (strictement positif) dans z.

Corrigé

Après avoir introduit les variables d'écart, on obtient le dictionnaire suivant :
$x_4 = 4 - x_1 - x_3$
$x_5 = 7 - x_1 - x_2 + x_3$
$x_6 = 5 - x_1 + x_2 - 2x_3$
$z = x_1 + 5x_2 - 4x_3$
La variable x_2 entre en base et x_5 en sort. Le nouveau dictionnaire est :
$x_2 = 7 - x_1 + x_3 - x_5$
$x_4 = 4 - x_1 - x_3$
$x_6 = 12 - 2x_1 - x_3 - x_5$
$z = 35 - 4x_1 + x_3 - 5x_5$
La variable x_3 entre en base et x_4 en sort. Le nouveau dictionnaire est :
$x_3 = 4 - x_1 - x_4$
$x_2 = 11 - 2\,x_1 - x_4 - x_5$
$x_6 = 8 - x_1 + x_4 - x_5$
$z = 39 - 5x_1 - x_4 - 5x_5$
Tous les coefficients dans z étant négatifs ou nuls, on a fini. Une solution optimale est $x_1 = 0$, $x_2 = 11$, $x_3 = 4$, ce qui donne à z la valeur 39.
Fin de corrigé

2. (2 points) Donner l'expression du problème dual de (P) et, à l'aide de la question 1, en donner une solution optimale ; on précisera comment on obtient cette solution optimale.
Corrigé
Le problème (P) étant sous forme standard, on obtient directement le problème dual (D) :

$$\text{Minimiser } w = 4y_1 + 7y_2 + 5y_3$$

$$\text{avec les contraintes } \begin{cases} y_1 + y_2 + y_3 \geq 1 \\ y_2 - y_3 \geq 5 \\ y_1 - y_2 + 2y_3 \geq -4 \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0 \end{cases}$$

Une solution optimale de (D) est donnée par le dernier dictionnaire de la résolution de (P), en considérant l'opposé des coefficients des variables d'écart dans z. On obtient ainsi $y_1 = 1$, $y_2 = 5$, $y_3 = 0$, ce qui donne à w la valeur 39, qui est donc le minimum de (D).
Fin de corrigé

3. (1,5 points) Les contraintes de (P) sont associées à l'utilisation de ressources dont les quantités disponibles sont actuellement égales à 4, 7 et 5 respectivement. Le propriétaire de ces ressources envisage de les vendre, au moins partiellement, à un prix unitaire égal à 2 pour chaque ressource. Quelle(s) ressource(s) pourrait-on conseiller de vendre, au moins en partie ? On justifiera la réponse ; si on conseille de vendre certaines de ces ressources, il n'est pas demandé d'indiquer quelle quantité il est conseillé de vendre.
Corrigé
On sait que les variables y_i peuvent être interprétées comme la valeur unitaire de la ressource i, c'est-à-dire aussi le prix minimum auquel on accepte de vendre une partie de la ressource i : on a intérêt à vendre si le prix d'achat proposé par unité de la ressource i est supérieur à y_i. Ici, on a intérêt à vendre une partie des ressources 1 et 3 (puisqu'on a $y_1 = 1 < 2$ et $y_3 = 0 < 2$; notons que cela est par ailleurs évident pour la ressource 3, puisqu'elle n'est pas entièrement consommée) mais non pour la ressource 2 (puisqu'on a $y_2 = 5 > 2$).

Fin de corrigé

4. (2,5 points) La fonction objectif de (P) est en fait égale à $z(a) = ax_1 + 5x_2 - 4x_3$, où a est un paramètre réel. À l'aide du théorème des écarts complémentaires, déterminer les valeurs de a pour lesquelles la solution déterminée à la question 1 reste optimale.

Corrigé

Remarquons que la solution optimale de la question 1, (0, 11, 4), reste réalisable. L'application du théorème des écarts complémentaires donne les relations suivantes :

* $y_3 = 0$ puisque la troisième contrainte primale est non saturée ;
* $y_2 - y_3 = 5$ puisque x_2 est non nulle ;
* $y_1 - y_2 + 2y_3 = -4$ puisque x_3 est non nulle.

On obtient $y_1 = 1$, $y_2 = 5$, $y_3 = 0$ (on retrouve la solution optimale de (D)). Il faut encore vérifier que les contraintes duales sont satisfaites. La première est satisfaite si et seulement si on a $y_1 + y_2 + y_3 \geq a$, c'est-à-dire $a \leq 6$. Les autres contraintes (y compris les contraintes de signe) sont satisfaites. La solution (0, 11, 4) reste donc optimale si et seulement si on a $a \leq 6$.

Fin de corrigé

Exercice 3 (16 points)

On considère la fonction f définie sur $\mathbf{R}^2$ par :

$$f(x, y) = 2x^2 + 3\exp(2y) + y - 4x\exp(y) = 2x^2 + 3e^{2y} + y - 4xe^y,$$

où e est la base des logarithmes népériens et exp la fonction exponentielle de base e.

1. (1 point) Donner l'expression du gradient et de la matrice hessienne de f.

Corrigé

À l'aide de dérivations partielles, on obtient :

$$\nabla f(x, y) = \begin{pmatrix} 4x - 4e^y \\ 6e^{2y} + 1 - 4xe^y \end{pmatrix} \text{ et } \nabla^2 f(x, y) = \begin{pmatrix} 4 & -4e^y \\ -4e^y & 12e^{2y} - 4xe^y \end{pmatrix}.$$

Fin de corrigé

2. (1,5 points) S'il en existe, donner tous les points de $\mathbf{R}^2$ constituant des minima de f (on précisera pour chacun de ces points s'il s'agit d'un minimum local ou global).

Corrigé

Soit (x, y) un minimum local de f, s'il en existe. Une condition nécessaire de minimalité locale est l'annulation du gradient. On doit donc avoir :

$$\nabla f(x, y) = \begin{pmatrix} 4x - 4e^y \\ 6e^{2y} + 1 - 4xe^y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

D'où le système $\begin{cases} 4x - 4e^y = 0 \\ 6e^{2y} + 1 - 4xe^y = 0 \end{cases}$, ou encore $\begin{cases} 4x = 4e^y \\ 2e^{2y} + 1 = 0 \end{cases}$, système qui n'admet pas de solution. La fonction f n'a donc pas de minimum (local ou global) sur $\mathbf{R}^2$.

Fin de corrigé

3. (3,5 points) Que donne l'application de la méthode de plus grande descente à pas optimal vue en cours à partir du point (1, 0) ? Que peut-on en déduire pour le minimum de f sur $\mathbf{R}^2$?

Corrigé

Au point (1, 0), on a $\nabla f(x, y) = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. On part donc dans la direction opposée au gradient de *f*, c'est-à-dire dans la direction $d = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. On s'intéresse ainsi aux points de la forme $\begin{pmatrix} 1 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -s \end{pmatrix}$, où *s* est un réel positif. Posons $h(s) = f(1, -s)$, ce qui donne $h(s) = 2 + 3e^{-2s} - s - 4e^{-s}$; on cherche à minimiser $h(s)$ pour *s* positif. Par dérivation, on obtient $h'(s) = -6e^{-2s} + 4e^{-s} - 1$, qui est un polynôme du deuxième degré en la variable $x = e^{-s}$. Cherchons les valeurs de *x* vérifiant $-6x^2 + 4x - 1 = 0$, s'il en existe. Le discriminant de cette équation, égal à 16 – 24, est négatif ; par conséquent, $h'(s)$ ne s'annule jamais et est ici toujours négatif : la fonction *h* est strictement décroissante pour $s \geq 0$. On voit de plus que *h* tend vers $-\infty$ quand *s* tend vers $+\infty$. On en déduit que *f* tend vers $-\infty$ pour les points de la forme (1, *y*) quand *y* tend vers $-\infty$. En particulier, *f* n'admet pas de minimum global (ce que montrait déjà la question 2).

Remarque : on peut s'épargner le calcul de $h'(s)$ et l'étude des variations de la fonction *h* en constatant directement que *h* tend vers $-\infty$ quand *s* tend vers $+\infty$; cela suffit pour conclure que *f* n'admet pas de minimum global.

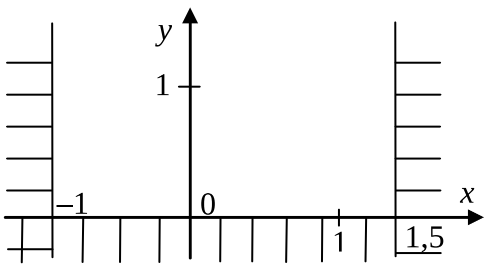
Fin de corrigé

Dans la suite, on se restreint au domaine réalisable *D* défini par $-1 \leq x \leq 1{,}5$ et $y \geq 0$.

4. (1 point) Indiquer en quels points de *D* les contraintes sont qualifiées.

Corrigé

L'allure du domaine *D* est représentée ci-dessous.



Posons $g_1(x, y) = -x - 1$, $g_2(x, y) = x - 1{,}5$ et $g_3(x, y) = -y$: les contraintes s'écrivent alors $g_1(x, y) \leq 0$, $g_2(x, y) \leq 0$ et $g_3(x, y) \leq 0$. Ces trois fonctions étant affines, elles sont convexes. De plus, *D* possède un intérieur non vide (par exemple, le point (0, 1) en fait partie). Un résultat du cours permet d'affirmer que les contraintes sont qualifiées en tout point de *D*.

Fin de corrigé

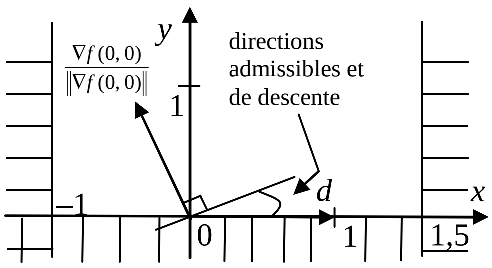
5. (2,5 points) Que donne l'application des conditions de Karush, Kuhn et Tucker au point (–1, 0) (on détaillera l'application) ?

Corrigé

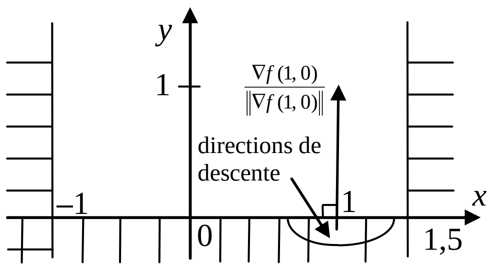
Le point (–1, 0) est bien réalisable. En (–1, 0), les deux contraintes g_1 et g_3 sont saturées. Les conditions de Karush, Kuhn et Tucker au point (–1, 0) indiquent que, pour que ce point soit un minimum de *P*, il faut qu'il existe deux réels positifs ou nuls λ_1 et λ_3 vérifiant l'égalité $\nabla f(-1, 0) = -\lambda_1 \nabla g_1(-1, 0) - \lambda_3 \nabla g_3(-1, 0)$. Or, on a : $\nabla f(-1, 0) = (-8, 11)^t$, $\nabla g_1(-1, 0) = (-1, 0)^t$, $\nabla g_3(-1, 0) = (0, -1)^t$. On obtient donc $\lambda_1 = -8$ et $\lambda_3 = 11$. Les

conditions de Karush, Kuhn et Tucker ne sont pas satisfaites, ce qui permet de conclure que le point (–1, 0) n'est pas un minimum de P.
Fin de corrigé

6. (3,5 points) Appliquer la méthode de plus grande descente (admissible) à pas optimal vue en cours à partir du point (0, 0) (on détaillera les calculs ; on pourra s'aider d'un dessin pour déterminer les directions à suivre, en justifiant cependant les choix qui seront faits).
Corrigé
Au point (0, 0), la première contrainte est saturée. Or on a $\nabla f(0,0) = \begin{pmatrix} -4 \\ 7 \end{pmatrix}$. La direction donnée par l'opposé du gradient n'est pas admissible puisqu'elle fait sortir de D ; on suit la direction qui, parmi les directions ne faisant pas sortir immédiatement du domaine, fait le plus petit angle avec l'opposé du gradient : c'est la plus grande descente admissible. Plus précisément, les directions admissibles en (0, 0) sont les directions dont la seconde composante est positive ou nulle, tandis que les directions de descente sont les directions faisant un angle supérieur à $\pi/2$ avec le gradient de f en (0, 0). Il est facile de voir, comme l'illustre le dessin ci-dessous, que cette direction admissible de plus forte descente est la direction $d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.



On cherche donc un point de la forme $\begin{pmatrix} 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} s \\ 0 \end{pmatrix}$, où s est un réel positif. Posons $h(s) = f(s, 0) = 2s^2 - 4s + 3$ et cherchons à minimiser h sur D. Pour ne pas sortir du domaine, on doit avoir $s \leq 1{,}5$. On optimise donc h avec la contrainte $s \leq 1{,}5$. Cherchons la valeur de s qui annule la dérivée de h. Celle-ci vaut : $h'(s) = 4s - 4$. Elle s'annule pour $s = 1$. On constate, en considérant les sens de variation de h, que $s = 1$ correspond bien au minimum de h. On obtient donc le point (1, 0). En ce point, le gradient de f vaut $\nabla f(1,0) = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. On constate graphiquement (voir le dessin ci-dessous) que, $\nabla f(1,0)$ étant orthogonal au bord de D, il n'y a plus de direction de descente qui soit admissible : les directions de (stricte) descente sont celles dont la seconde composante est strictement négative, tandis que les directions admissibles sont celles dont la seconde composante est positive ou nulle. La méthode s'arrête.



Fin de corrigé

7. (3 points) On admet que la fonction f est strictement convexe sur un ouvert contenant D et on considère que le domaine de définition de f est cet ouvert. Le point obtenu à la question précédente est-il un minimum de f sur D ? Si oui, on précisera s'il est local ou global et s'il est l'unique minimum (local ou global) sur D.

Corrigé

Pour vérifier la minimalité du point (1, 0), appliquons une nouvelle fois les conditions de Karush, Kuhn et Tucker. En (1, 0), seule la contrainte g_3 est saturée. L'application des conditions de Karush, Kuhn et Tucker au point (1, 0) indiquent que, pour que ce point soit un minimum de P, il faut qu'il existe un réel positif ou nul λ_3 vérifiant $\nabla f(1, 0) = -\lambda_3 \nabla g_3(1, 0)$. Or, on a $\nabla f(1, 0) = (0, 3)^t$ et $\nabla g_3(1, 0) = (0, -1)^t$. On obtient donc $\lambda_3 = 3$. Les conditions de Karush, Kuhn et Tucker sont satisfaites. De plus, les contraintes sont convexes car affines, l'intérieur de D est non vide et on a admis que f est strictement convexe sur son domaine de définition. On est donc dans un cas pour lequel on a les résultats suivants :

- tout minimum local est global ;
- la condition de Karush, Kuhn et Tucker est une condition nécessaire et suffisante de minimalité globale ;
- un point vérifiant la condition de Karush, Kuhn et Tucker est l'unique minimum global sur le domaine considéré.

Conclusion : (1, 0) est l'unique minimum global de f sur D et il n'existe pas de minimal local ailleurs dans D.

Fin de corrigé