

Partiel du 29/10/2019
“Optimisation et programmation dynamique”
Durée 2h - Calculatrice et documents non autorisés

Exercice 1. On considère le problème

$$(\mathcal{P}) \quad \begin{array}{ll} \min & x - 2y \\ & -x^2 \leq y \leq 0 \\ & -1 \leq x \leq 0 \end{array}$$

1. Montrer que le problème admet au moins une solution.
2. Dessiner la contrainte $K := \{(x, y) \in \mathbb{R}^2, -x^2 \leq y \leq 0, -1 \leq x \leq 0\}$.
3. Montrer que, tout point $(x, y) \in K$ avec $(x, y) \neq (0, 0)$, la contrainte est qualifiée en (x, y) .
4. Trouver l'ensemble des points vérifiant les conditions nécessaires d'optimalité.
5. Déterminer la ou les solutions du problème.

Solution : Sur (9pts)

1. La contrainte $K := \{(x, y) \in \mathbb{R}^2, -x^2 \leq y \leq 0, -1 \leq x \leq 0\}$ est fermée (car définie par des inégalités larges invoquant des fonctions polynômiales donc continues). De plus la contrainte est bornée, car, si $(x, y) \in K$, alors $|x| \leq 1$ et $0 \geq y \geq -x^2 \geq -1$. Donc la contrainte est compacte. Enfin le critère $f(x, y) := x - 2y$ est continu (linéaire). Par conséquent le problème admet une solution.

2. Dessin

3. Pour étudier les conditions de qualification de la contrainte, introduisons les fonctions définissant celle-ci : soient $g_1(x, y) = -x - 1$, $g_2(x, y) = y$, $g_3(x, y) = -y - x^2$, $g_4(x, y) = x$. Soit $(x, y) \in K$ avec $(x, y) \neq (0, 0)$.

S'il n'y a aucune contrainte saturée, il n'y a rien à vérifier.

On considère maintenant le cas où une seule contrainte est saturée. Si c'est g_1 , g_2 ou g_4 , alors la contrainte est affine et non nulle, donc son gradient est non nul (respectivement égal à $(-1, 0)$, $(0, 1)$ et $(1, 0)$) et la contrainte est qualifiée. Reste le cas de la contrainte g_3 , où $\nabla g_3(x, y) = (-2x, -1)$, qui est à nouveau non nul. Donc dans tous ces cas, la contrainte est qualifiée.

Considérons le cas où deux contraintes sont saturées. Si $x = 0$, alors $y = 0$ et ce cas est exclus par hypothèse. Si $x = -1$, alors soit $y = 0$, soit $y = -1$. Dans le premier cas, les contraintes saturées sont g_1 et g_2 . Si on prend $v = (1, -1)$, alors $\langle \nabla g_1(-1, 0), (1, -1) \rangle = \langle (-1, 0), (1, -1) \rangle = -1 < 0$ et $\langle \nabla g_2(-1, 0), (1, -1) \rangle = \langle (0, 1), (1, -1) \rangle = -1 < 0$. Donc la contrainte est qualifiée en $(-1, 0)$. En $(-1, -1)$, les contraintes saturées sont g_1 et g_3 . Si on choisit $v = (1, 3)$, alors $\langle \nabla g_1(-1, -1), (1, 3) \rangle = \langle (-1, 0), (1, 3) \rangle = -1 < 0$ et $\langle \nabla g_3(-1, -1), (1, 3) \rangle = \langle (2, -1), (1, 3) \rangle = -1 < 0$. Donc la contrainte est qualifiée en $(-1, -1)$.

Reste finalement le cas où plus de 2 contraintes sont saturées. Cela ne peut arriver que si $-x^2 = y = 0$ et $x = 0$, c'est-à-dire si $(x, y) = (0, 0)$. Mais ce cas est exclu par hypothèse. Donc nous avons montré que la contrainte est qualifiée en tout point de K , sauf éventuellement en $(0, 0)$.

(on peut s'arrêter là ou bien noter que la contrainte n'est pas qualifiée en $(0, 0)$. En effet, en $(0, 0)$, il y a 3 contraintes saturées : g_2 , g_3 et g_4 , dont les gradients en $(0, 0)$ sont respectivement $(0, 1)$, $(0, -1)$ et $(1, 0)$. Comme $\nabla g_2(0, 0) = -\nabla g_3(0, 0)$, il est clair qu'il n'existe pas de vecteur $v = (v_x, v_y) \in \mathbb{R}^2$ tel que $\langle \nabla g_i(0, 0), (v_x, v_y) \rangle < 0$ pour $i = 1, 2, 3$.)

4. Nous cherchons maintenant l'ensemble des points $(x, y) \in K$ vérifiant les conditions nécessaires d'optimalité. Si (x, y) est un point de minimum non nul du problème, alors la contrainte est qualifiée en ce point et le théorème de Kuhn et Tucker affirme qu'il existe $\lambda_1, \dots, \lambda_4 \geq 0$ tels que

$$\begin{aligned} (i) \quad & 1 - \lambda_1 - 2x\lambda_3 + \lambda_4 = 0 \\ (ii) \quad & 2 + \lambda_2 - \lambda_3 = 0 \\ (iii) \quad & \lambda_1(-x - 1) = \lambda_2y = \lambda_3(-y - x^2) = \lambda_4x = 0 \end{aligned}$$

On discute à nouveau suivant le nombre de contraintes saturées. Si aucune contrainte n'est saturée, alors tous les λ_i sont nul, ce qui est impossible, puisque (i) deviendrait $1 = 0$.

Discutons à présent du cas où une seule contrainte est saturée : la seule possibilité est que ce soit g_3 (car sinon soit (i) soit (ii) conduit à une absurdité). Alors le système devient :

$$1 - 2x\lambda_3 = 2 - \lambda_3 = 0,$$

ce qui implique $\lambda_3 = 2$ et $x = 1/(2\lambda_3) = 1/4$ contredisant la contrainte $x \leq 0$.

Étudions maintenant le cas où deux contraintes sont saturées. Au vu de la discussion sur la qualification, il y a 2 cas possibles : $(x, y) = (-1, 0)$ ou $(x, y) = (0, -1)$. Dans le premier cas, les contraintes saturées sont g_1 et g_2 , avec $\lambda_3 = \lambda_4 = 0$, et le système devient

$$1 - \lambda_1 = 2 + \lambda_2 = 0,$$

ce qui est impossible puisque $\lambda_2 \geq 0$. En $(-1, -1)$, les contraintes saturées sont g_1 et g_3 , avec $\lambda_2 = \lambda_4 = 0$, et le système devient

$$1 - \lambda_1 + 2\lambda_3 = 2 - \lambda_3 = 0,$$

qui a pour solution $\lambda_3 = 2$ et $\lambda_1 = 5$. Le point $(-1, 1)$ vérifie les conditions nécessaires d'optimalité.

Enfin, nous avons vu lors de la discussion sur la qualification, que le cas où plus de 2 contraintes sont saturées ne conduit qu'au point $(0, 0)$, qu'on peut exclure puisqu'on n'a pas prouvé que la contrainte était qualifiée en ce point. (en fait, les conditions d'optimalité ne sont pas satisfaites en $(0, 0)$ car les contraintes g_2 , g_3 et g_4 étant saturées, le système devient

$$1 + \lambda_4 = 2 + \lambda_2 - \lambda_3 = 0,$$

ce qui conduit à une impossibilité puisque $\lambda_4 \geq 0$).

5. Au vu des discussions ci-dessus, nous savons que le problème admet une solution, et que cette solution est à chercher parmi les points où la contrainte est qualifiée et où les conditions nécessaires d'optimalité sont satisfaites—c'est-à-dire $(-1, -1)$ —et parmi les points où la contrainte n'est pas forcément qualifiée—c'est-à-dire $(0, 0)$. Calculons la valeur du critère en ces points : $f(-1, -1) = -1 + 2 = 1$ et $f(0, 0) = 0$. Le minimum est donc atteint en $(0, 0)$ et a pour valeur 0.

Exercice 2. On considère le problème de programmation linéaire

$$(\mathcal{P}) \quad \inf_{x \in K} \langle a, x \rangle$$

où $a \in \mathbb{R}^n$, $C \in \mathbb{R}^{J \times n}$, $d \in \mathbb{R}^J$ (n et J sont des entiers non nuls) et où

$$K := \{x \in \mathbb{R}^n, Cx = d \quad \text{et} \quad x_i \geq 0 \quad \forall i \in \{1, \dots, n\}\}.$$

On suppose que K est non vide et que le problème (\mathcal{P}) admet un minimum. On admettra l'égalité¹ :

$$\inf_{x \in K} \langle a, x \rangle = \sup_{\lambda \in \mathbb{R}_+^n, \mu \in \mathbb{R}^J} \inf_{x \in \mathbb{R}^n} \{ \langle a, x \rangle - \langle \lambda, x \rangle + \langle \mu, (Cx - d) \rangle \}. \quad (1)$$

1. Ecrire les conditions nécessaires d'optimalité du problème.
2. En utilisant l'égalité (1), démontrer que

$$\inf_{x \in K} \langle a, x \rangle = - \inf_{\mu \in \tilde{K}} \langle d, \mu \rangle$$

où

$$\tilde{K} := \{ \mu \in \mathbb{R}^J, (C^T \mu)_j + a_j \geq 0, \forall j \in \{1, \dots, n\} \}.$$

3. On suppose que $x \in K$ et $\mu \in \tilde{K}$. Montrer que $\langle a, x \rangle + \langle d, \mu \rangle \geq 0$.
4. Soit x^* un point de minimum du problème (\mathcal{P}) et $(\lambda^*, \mu^*) \in \mathbb{R}_+^n \times \mathbb{R}^J$ les multiplicateurs associés dans les conditions nécessaires d'optimalité de x^* . Montrer que μ^* est un point de minimum du problème

$$(\tilde{\mathcal{P}}) \quad \inf_{\mu \in \tilde{K}} \langle d, \mu \rangle.$$

5. On suppose que x^* est un point de minimum de (\mathcal{P}) et μ^* est une solution de $(\tilde{\mathcal{P}})$. Montrer alors que $\langle a, x^* \rangle + \langle d, \mu^* \rangle = 0$.

Solution : Sur (6pts)

1. Soit x^* un point de minimum du problème. Comme les contraintes sont affines, elles sont qualifiées en tout point. Le théorème de Kuhn-Tucker affirme qu'il existe $\lambda_1, \dots, \lambda_n \geq 0$ et $\mu_1, \dots, \mu_J \in \mathbb{R}$ tels que

$$\begin{aligned} a_i - \lambda_i e_i + (C^T \mu)_i &= 0 & \forall i = 1 \dots n, \\ \lambda_i x_i &= 0 & \forall i \in \{1, \dots, n\}. \end{aligned}$$

1. Cette égalité de dualité a été prouvée dans le cours dans un contexte un tout petit peu différent.

2. Il suffit de calculer le membre de droite de l'égalité. Notons d'abord que

$$\inf_{x \in \mathbb{R}^n} \{ \langle a, x \rangle - \langle \lambda, x \rangle + \langle \mu, (Cx - d) \rangle \} = \inf_{x \in \mathbb{R}^n} \{ \langle a - \lambda + C^T \mu, x \rangle - \langle \mu, d \rangle \}.$$

Si $a - \lambda + C^T \mu \neq 0$, alors, en considérant la suite $(x_k := -k(a - \lambda + C^T \mu))$, on a

$$\lim_{k \rightarrow \infty} \langle a - \lambda + C^T \mu, x_k \rangle = -k \|a - \lambda + C^T \mu\|^2 = -\infty.$$

Donc

$$\inf_{x \in \mathbb{R}^n} \{ \langle a, x \rangle - \langle \lambda, x \rangle + \langle \mu, (Cx - d) \rangle \} = \begin{cases} -\infty & \text{si } a - \lambda + C^T \mu \neq 0 \\ -\langle \mu, d \rangle & \text{si } a - \lambda + C^T \mu = 0 \end{cases} \quad (2)$$

Par conséquent,

$$\begin{aligned} & \sup_{\lambda \in \mathbb{R}_+^n, \mu \in \mathbb{R}^J} \inf_{x \in \mathbb{R}^n} \{ \langle a, x \rangle - \langle \lambda, x \rangle + \langle \mu, (Cx - d) \rangle \} \\ &= \sup_{\substack{\lambda \in \mathbb{R}_+^n, \mu \in \mathbb{R}^J, \\ a - \lambda + C^T \mu = 0}} -\langle \mu, d \rangle = - \inf_{\substack{\lambda \in \mathbb{R}_+^n, \mu \in \mathbb{R}^J, \\ a - \lambda + C^T \mu = 0}} \langle \mu, d \rangle. \end{aligned}$$

Pour conclure, il suffit de remarquer que l'ensemble des $\mu \in \mathbb{R}^J$ pour lesquels il existe $\lambda \in \mathbb{R}_+^n$ avec $a - \lambda + C^T \mu = 0$ est précisément l'ensemble \tilde{K} . En utilisant (1), on en déduit l'égalité

$$\inf_{x \in K} \langle a, x \rangle = - \inf_{\mu \in \tilde{K}} \langle d, \mu \rangle.$$

3. Soient $x \in K$ et $\mu \in \tilde{K}$. Alors

$$\langle d, \mu \rangle = \langle Cx, \mu \rangle = \langle x, C^T \mu \rangle = \sum_{i=1}^n x_i (C^T \mu)_i \geq \sum_{i=1}^n x_i (-a_i) = -\langle a, x \rangle,$$

où on a utilisé successivement le fait que $x \in K$ et donc $d = Cx$ (dans la première égalité) et $x_i \geq 0$ (dans l'inégalité) et le fait que $\mu \in \tilde{K}$ (dans l'inégalité). On en déduit que $\langle a, x \rangle + \langle d, \mu \rangle \geq 0$.

4. Soit x^* un point de minimum du problème (\mathcal{P}) et $(\lambda^*, \mu^*) \in \mathbb{R}_+^d \times \mathbb{R}^J$ les multiplicateurs associés dans les conditions nécessaires d'optimalité de x^* . On sait par la question 1 que

$$\begin{aligned} a - \lambda^* + C^T \mu^* &= 0, \\ \lambda_i^* x_i^* &= 0 \quad \forall i \in \{1, \dots, n\}. \end{aligned}$$

On remarque que $\mu^* \in \tilde{K}$ par la première égalité. Toujours par cette égalité, on a aussi

$$0 = \langle a - \lambda^* + C^T \mu^*, x^* \rangle = \langle a, x^* \rangle + \langle Cx^*, \mu^* \rangle = \langle a, x^* \rangle + \langle d, \mu^* \rangle,$$

où on a utilisé le fait que $\lambda_i^* x_i^* = 0$ pour tout i dans la seconde égalité et la contrainte $Cx^* = d$ dans la dernière. Par conséquent, $\langle d, \mu^* \rangle = -\langle a, x^* \rangle$. Or d'après la question précédente appliquée à tout $\mu \in \tilde{K}$ et à $x^* \in K$, on a

$$\langle d, \mu \rangle \geq -\langle a, x^* \rangle = \langle d, \mu^* \rangle.$$

Donc μ^* est un point de minimum du problème $(\tilde{\mathcal{P}})$.

5. On suppose que x^* est un point de minimum de (\mathcal{P}) et μ^* est une solution de $(\tilde{\mathcal{P}})$. Alors d'après la question 2, on a

$$\langle a, x^* \rangle = \inf_{x \in K} \langle a, x \rangle = - \inf_{\mu \in \tilde{K}} \langle d, \mu \rangle = -\langle d, \mu^* \rangle.$$

D'où le résultat.

Exercice 3. Soit A une matrice carrée $d \times d$ symétrique définie positive, $b \in \mathbb{R}^d$, $C \in \mathbb{R}^{J \times d}$, $d \in \mathbb{R}^J$ (comme toujours, les vecteurs sont assimilés à des matrices colonnes). Dans toute la suite, $\|\cdot\|$ désigne à la fois les normes euclidiennes sur \mathbb{R}^d et sur \mathbb{R}^J , et les normes matricielles associées.

On cherche à résoudre

$$\min_{Cx=d} \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle.$$

On suppose qu'il existe $x \in \mathbb{R}^d$ tel que $Cx = d$.

1. Montrer que le problème admet un unique point de minimum noté x^* .
2. Ecrire les conditions nécessaires d'optimalité. On notera $\mu^* \in \mathbb{R}^J$ un lagrangien associé au problème.

On cherche à résoudre numériquement le problème. Etant donnés des pas $\tau_1 > 0$ et $\tau_2 > 0$ et des conditions initiales $x^0 \in \mathbb{R}^d$ et $\mu^0 \in \mathbb{R}^J$, on définit par récurrence la suite

$$\begin{cases} x^{n+1} = x^n - \tau_1(Ax^n + b + C^T \mu^n) \\ \mu^{n+1} = \mu^n + \tau_1 \tau_2 (Cx^{n+1} - d) \end{cases}$$

3. Montrer que l'on peut choisir $\tau_1 > 0$ suffisamment petit pour que $\|I_d - \tau_1 A\| < 1$ (où I_d est la matrice identité de format d). On fera ce choix par la suite et on posera alors $\beta := \|I_d - \tau_1 A\|$.
4. En utilisant le fait que x^* vérifie les contraintes, montrer que l'on a

$$\|\mu^{n+1} - \mu^*\|^2 \leq \|\mu^n - \mu^*\|^2 + 2\tau_1 \tau_2 \langle (\mu^n - \mu^*), C(x^{n+1} - x^*) \rangle + (\tau_1 \tau_2)^2 \|C^T C\| \|x^{n+1} - x^*\|^2.$$

5. En utilisant les conditions d'optimalité et la définition de x^{n+1} , vérifier que

$$\tau_1 C^T (\mu^n - \mu^*) = -(x^{n+1} - x^*) + (I_d - \tau_1 A)(x^n - x^*).$$

6. En déduire que

$$\gamma \|x^{n+1} - x^*\|^2 \leq \left(\frac{\|\mu^n - \mu^*\|^2}{\tau_2} + \beta \|x^n - x^*\|^2 \right) - \left(\frac{\|\mu^{n+1} - \mu^*\|^2}{\tau_2} + \beta \|x^{n+1} - x^*\|^2 \right)$$

où $\gamma := (2 - \tau_1^2 \tau_2 \|C^T C\| - 2\beta)$ sera supposé strictement positif (grâce au fait que $\beta \in]0, 1[$ et à un choix judicieux de τ_2).

7. En déduire que (x^n) converge vers x^* .
8. Quel est la différence entre cet algorithme et l'algorithme d'Uzawa?

Solution : Sur (8pts)

1. La contrainte est définie par des inégalités larges faisant intervenir des fonctions convexes et continue. Donc elle est convexe et fermée. De plus le critère est fortement convexe, puisque sa matrice hessienne est A qui est définie positive par hypothèse. Par théorème de cours on en déduit que le problème admet un unique point de minimum noté x^* .
2. Soit $f(x) = \frac{1}{2}\langle Ax, x \rangle + \langle b, x \rangle$. Donc $\nabla f(x) = Ax + b$. La contrainte est affine, donc qualifiée en tout point. Comme x^* est le minimum du problème, le théorème de Kuhn-Tucker affirme qu'il existe $\mu^* \in \mathbb{R}^J$ tel que

$$Ax^* + b + C^T \mu^* = 0.$$

3. Par définition de la norme matricielle,

$$\|I_d - \tau_1 A\| = \sup_{x \in \mathbb{R}^d, x \neq 0} \frac{\|(I_d - \tau_1 A)x\|}{\|x\|}.$$

Or

$$\|(I_d - \tau_1 A)x\|^2 = \langle (I_d - \tau_1 A)^2 x, x \rangle \leq \theta \|x\|^2,$$

où θ est la plus grande valeur propre de la matrice symétrique $(I_d - \tau_1 A)^2$. Comme A est symétrique, on sait que les valeurs propres de $(I_d - \tau_1 A)^2$ sont les $(1 - \tau_1 \lambda_i)^2$ où les λ_i sont les valeurs propres de A . Comme A est définie positive, les λ_i sont tous strictement positifs. Donc en prenant $\tau_1 = \min_i \{1/(2\lambda_i)\} = 1/(2 \max_i(\lambda_i))$, on a

$$1 > 1 - \tau_1 \lambda_i \geq 1 - \tau_1 \max_i \lambda_i = 1/2,$$

et donc, pour ce choix de τ_1 , on a $\theta = \max_i (1 - \tau_1 \lambda_i)^2 < 1$. On en déduit que

$$\|I_d - \tau_1 A\| \leq \theta^{1/2} < 1.$$

4. Comme x^* vérifie les contraintes, on a

$$\mu^{n+1} - \mu^* = \mu^n - \mu^* + \tau_1 \tau_2 C(x^{n+1} - x^*),$$

et donc

$$\begin{aligned} \|\mu^{n+1} - \mu^*\|^2 &= \|\mu^n - \mu^*\|^2 + 2\tau_1 \tau_2 \langle (\mu^n - \mu^*), C(x^{n+1} - x^*) \rangle + (\tau_1 \tau_2)^2 \|C(x^{n+1} - x^*)\|^2 \\ &\leq \|\mu^n - \mu^*\|^2 + 2\tau_1 \tau_2 \langle (\mu^n - \mu^*), C(x^{n+1} - x^*) \rangle + (\tau_1 \tau_2)^2 \langle C^T C(x^{n+1} - x^*), (x^{n+1} - x^*) \rangle \\ &\leq \|\mu^n - \mu^*\|^2 + 2\tau_1 \tau_2 \langle (\mu^n - \mu^*), C(x^{n+1} - x^*) \rangle + (\tau_1 \tau_2)^2 \|C^T C\| \|x^{n+1} - x^*\|^2, \end{aligned}$$

ce qui est l'inégalité demandée.

5. Par la définition de x^{n+1} et les conditions d'optimalité pour x^* , on a

$$\begin{aligned} x^{n+1} - x^* - (I_d - \tau_1 A)(x^n - x^*) &= x^n - \tau_1 (Ax^n + b + C^T \mu^n) - x^* - (I_d - \tau_1 A)(x^n - x^*) \\ &= -\tau_1 (-Ax^* - C^T \mu^* + C^T \mu^n) - \tau_1 Ax^* = -\tau_1 C^T (\mu^n - \mu^*). \end{aligned}$$

En réorganisant, on obtient le résultat demandé :

$$\tau_1 C^T (\mu^n - \mu^*) = -(x^{n+1} - x^*) + (I_d - \tau_1 A)(x^n - x^*).$$

6. On utilise l'égalité obtenue dans la question précédent dans l'inégalité de la question 4 :

$$\begin{aligned}
\| \mu^{n+1} - \mu^* \|^2 &\leq \| \mu^n - \mu^* \|^2 + 2\tau_1\tau_2 \langle C^T(\mu^n - \mu^*), (x^{n+1} - x^*) \rangle + (\tau_1\tau_2)^2 \| C^T C \| \| x^{n+1} - x^* \|^2 \\
&\leq \| \mu^n - \mu^* \|^2 + 2\tau_2 \langle -(x^{n+1} - x^*) + (I_d - \tau_1 A)(x^n - x^*), (x^{n+1} - x^*) \rangle \\
&\quad + (\tau_1\tau_2)^2 \| C^T C \| \| x^{n+1} - x^* \|^2 \\
&\leq \| \mu^n - \mu^* \|^2 - 2\tau_2 \| x^{n+1} - x^* \|^2 + 2\tau_2 \| (I_d - \tau_1 A) \| \| x^n - x^* \| \| x^{n+1} - x^* \| \\
&\quad + (\tau_1\tau_2)^2 \| C^T C \| \| x^{n+1} - x^* \|^2
\end{aligned}$$

où on a utilisé l'inégalité de Cauchy-Schwarz dans la dernière inégalité. Comme $ab \leq (a^2 + b^2)/2$ et par définition de β , on en déduit que

$$\begin{aligned}
&\| \mu^{n+1} - \mu^* \|^2 \\
&\leq \| \mu^n - \mu^* \|^2 - 2\tau_2 \| x^{n+1} - x^* \|^2 + \tau_2\beta \| x^n - x^* \|^2 + \tau_2\beta \| x^{n+1} - x^* \|^2 \\
&\quad + (\tau_1\tau_2)^2 \| C^T C \| \| x^{n+1} - x^* \|^2
\end{aligned}$$

En réarrangeant, on obtient l'inégalité voulue :

$$\gamma \| x^{n+1} - x^* \|^2 \leq \left(\frac{\| \mu^n - \mu^* \|^2}{\tau_2} + \beta \| x^n - x^* \|^2 \right) - \left(\frac{\| \mu^{n+1} - \mu^* \|^2}{\tau_2} + \beta \| x^{n+1} - x^* \|^2 \right)$$

où $\gamma := (2 - \tau_1^2\tau_2 \| C^T C \| - 2\beta)$.

7. D'après la question précédente, La suite $\theta_n := \frac{\| \mu^n - \mu^* \|^2}{\tau_2} + \beta \| x^n - x^* \|^2$ est décroissante (car $\gamma \geq 0$) et minorée par 0. Donc elle converge, et donc, en utilisant cette fois que $\gamma > 0$, prouve que $\| x^{n+1} - x^* \|^2$ est majorée par la suite $\gamma^{-1}(\theta_n - \theta_{n+1})$ qui tend vers 0. Donc (x^n) converge vers x^* .
8. La différence entre cet algorithme et l'algorithme d'Uzawa est que, dans l'algorithme d'Uzawa, il faut calculer un problème de minimisation dans \mathbb{R}^d pour définir x^{n+1} , alors que dans cet algorithme, x^{n+1} est calculé par un pas de gradient, ce qui est beaucoup moins coûteux numériquement.