Optimisation

Janvier 2017 **Documents autorisés**

Toute affirmation intuitive non argumentée sera à éviter

Exercice 1. Méthode des contraintes actives.
Soit $a > 0$. On considère le problème $\mathcal{P}$ de minimisation de $f(x, y) = a\,(x-2)^2 + (y-1)^2$ sous la contrainte $|x| + |y| \leq 1$.

(1) Montrer que $\mathcal{P}$ admet une solution. Celle-ci est-elle unique ?
(2) Calculer $f(1, 0)$ et $f(0, 1)$. Sans aucun calcul, représenter précisément les contraintes, les isocontours de f, identifier graphiquement la solution, tout ceci pour $a = 1$. Même question pour $a = 1/100$. On suppose pour **le reste de l'exercice que $\mathbf{a = 1}$**.
(3) On considère le problème $\mathcal{Q}$ de minimisation de $f(x, y) = (x-2)^2 + (y-1)^2$ sous les quatre contraintes simultanées suivantes : $x + y \leq 1$, $-x + y \leq 1$, $x - y \leq 1$ et $-x - y \leq 1$. Quel est le lien entre le problème $\mathcal{P}$ et le problème $\mathcal{Q}$.
(4) On considère le problème $\mathcal{Q}$. Sans aucun calcul, construire graphiquement le chemin des itérés de la méthode des contraintes actives vue en cours, en partant du point $(-1, 0)$ et cela pour les deux choix de contraintes actives de départ : l'ensemble vide d'une part, l'unique contrainte $\{y - x - 1 = 0\}$ d'autre part.
(5) Calculer les 3 premières itérations de la méthode des contraintes actives en partant du point $(-1, 0)$ et de la contrainte active $\{y - x - 1 = 0\}$.

Exercice 2. Problèmes à contraintes d'égalité et d'inégalités.
Soit $\mathbb{R}_+^n$ l'ensemble des vecteurs de $\mathbb{R}^n$ dont les composantes sont toutes positives : $\mathbb{R}_+^n = \{x = (x_1, \ldots, x_n);\ \forall i,\ x_i \geq 0\}$. Pour tout élément de $\mathbb{R}_+^n$, on appelle appelle moyenne géométrique la quantité $(\prod_{i=1}^n x_i)^{1/n}$ et moyenne arithmétique la quantité $\frac{1}{n}\sum_{i=1}^n x_i$. On cherche à redémontrer l'inégalité dite arithmético-géométrique

$$\forall x \in \mathbb{R}_+^n;\ \left(\prod_{i=1}^n x_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n x_i.$$

(1) On considère la fonction f définie sur $\mathbb{R}_+^n \setminus \{0\}$ par $f(x) = \frac{(\prod_{i=1}^n x_i)^{1/n}}{\sum_{i=1}^n x_i}$. Montrer que si $\lambda > 0$ et $x \in \mathbb{R}_+^n \setminus \{0\}$, on a $f(\lambda x) = f(x)$. En déduire le lien entre l'inégalité arithmético-géométrique et le problème $\mathcal{P}$

$$\max_{x \in X} \left(\prod_{i=1}^n x_i\right)^{1/n} \quad \text{où } X = \left\{x \in \mathbb{R}_+^n,\ \frac{1}{n}\sum_{i=1}^n x_i = 1\right\}.$$

(2) Montrer que $\mathcal{P}$ possède au moins une solution.

(3) On désire résoudre $\mathcal{P}$ par la technique des multiplicateurs vue en cours.
(3.1) Les contraintes vérifient-elles l'une des hypothèses de qualification des contraintes vues en cours ?
(3.2) Ecrire la condition nécessaire d'optimalité au premier ordre, en faisant attention à prendre en compte les contraintes d'égalité et d'inégalité.
(3.3) Supposer dans un premier temps qu'aucune des contraintes d'inégalité n'est active. Résoudre par la technique des multiplicateurs de Lagrange le problème.
(3.4) Un point où l'une des contraintes d'inégalité est active peut-il être maximum ? En déduire l'inégalité arithmético-géométrique.

Problème. Descente par coordonnées (DPC).

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continûment différentiable, et $(e_i)_{i=1\ldots n}$ la base canonique de $\mathbb{R}^n$. On s'intéresse à la méthode DPC de résolution de $(\mathcal{P})$: $\min_x f(x)$ où les itérés sont donnés par

$$x_{k+1} = x_k + \tau_k d_k, \text{ où } \left\{ \begin{array}{l} k = L \cdot n + j, \text{ pour } j = 1, \ldots, n \text{ et } L = 0, 1, 2 \ldots \\ d_k = e_j \end{array} \right.$$

Dit différemment, la méthode réalise une recherche directionnelle le long des vecteurs de la base canonique, ces vecteurs apparaîssant de manière cyclique au cours des itérations ($j-1$ est le reste de la division Euclidienne de $k-1$ par n, L est l'indice du cycle). On suppose de plus que f vérifie l'hypothèse H suivante : pour tout x et d le problème unidimensionnel $\min_{\tau\in\mathbb{R}} f(x+\tau d)$ admet toujours une unique solution $\tau^\star$. C'est cette solution, dans le cas où $x = x_k$ et $d = d_k$, qui sera utilisée pour la mise à jour $x_{k+1} = x_k + \tau_k d_k$. Il est à noter que la question **(5)** peut être traitée indépendemment des autres.

(1) On considère le cas où $n = 2$ et $f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + x_1 - 3x_2$. En partant de $x_1 = (0, 0)$, calculer les trois premiers itérés de la méthode et la solution du problème $(\mathcal{P})$. Expliquez le bon comportement de cette méthode.

(2) Pour $n \in \mathbb{N}$, soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique définie positive, $b \in \mathbb{R}^n$ et $f(x) = \frac{1}{2}x^T Ax - x^T b$. On sait que $\forall z$, $\|z\|_A = \sqrt{z^T Az}$ définit une norme sur $\mathbb{R}^n$. On pose $x^\star = A^{-1}b$.
(2.1) Montrer qu'à l'itération k, $f(x_k + \tau d_k) = f(x_k) + \tau(Ax_k - b)^T d_k + \frac{\tau^2}{2} d_k^T Ad_k$. En déduire que $\tau_k = -\frac{(Ax_k-b)^T d_k}{d_k^T Ad_k}$
(2.2) Montrer que $f(x_k) - f(x_{k+1}) = \frac{1}{2}\|x_k - x^\star\|_A^2 - \frac{1}{2}\|x_{k+1} - x^\star\|_A^2$. En déduire que les itérés de DPC restent dans une partie bornée de $\mathbb{R}^n$. En déduire que la suite du premier itéré de chaque cycle, $(x_{Ln+1})_{L\geq 1}$, possède une sous-suite convergeant vers un certain $\bar{x} \in \mathbb{R}^n$, sous-suite que l'on notera $(x_k)_{k\in\mathcal{K}}$ dans la suite.

(3) On veut montrer que pour la quadratique de la question **(2)** on a $\nabla f(\bar{x}) = A\bar{x} - b = 0$. On suppose dans **(3)**, **pour établir une contradiction, que $\mathbf{\nabla f(\bar{x}) \neq 0}$**. On définit pour $j = 1 \ldots n$, les applications ϕ_j qui réalisent les j premières iterations dans un cycle de DPC à partir d'un x donné. En particulier $\phi_j(x_{L\cdot n+1}) = x_{L\cdot n+1+j}$.
(3.1) Montrer que la méthode DPC partant de $x_1 = \bar{x}$ rencontre au n-ième pas de son premier cycle un certain $\bar{y} = \bar{x} + \sum_{i=1}^n \bar{\tau}_i e_i$ tel que $f(\bar{y}) < f(\bar{x})$.
(3.2) Montrer que pour tout $j = 1, \ldots n$ on a $\lim_{x\to\bar{x}} \phi_j(x) = \bar{x} + \sum_{i=1}^j \bar{\tau}_i e_i$. En déduire que $\lim_{k\to+\infty, k\in\mathcal{K}} \phi_n(x_k) = \bar{y}$. Montrer que ceci est impossible, compte tenu du caractère monotone de DPC, de la définition cd $\bar{x}$ et de $f(\bar{y}) < f(\bar{x})$. Conclure la question **(3)**.

(4). On considère à présent le cas où f n'est plus nécessairement quadratique, et on suppose que les itérés restent bornés. Reprendre rapidement les questions **(2)** et **(3)** dans ce cadre plus général en précisant les changements à opérer pour montrer l'existence d'une sous-suite convergeant vers $\bar{x}$ tel que $\nabla f(\bar{x}) = 0$.

(5) On désire réaliser un algorithme par blocs, appelé DPCB. L'ensemble $\{1, \ldots n\}$ est partitionné une fois pour toutes en n_e sous-ensembles notés $(I_i)_{i=1,\ldots n_e}$: $I_i \cap I_j = 0$ si $i \neq j$ et $\cup_{i=1}^{n_e} I_i = \{1, \ldots n\}$. Dans le cas où f est quadratique de Hessienne définie positive, décrire et calculer explicitement les itérés de la méthode DPCB implantant cycliquement des minimisations sur chacun des n_e espaces vectoriels engendrés par les familles $(e_j)_{j\in I_i}$.