Optimisation sous contraintes

1 Gradient projeté

Soit A une matrice inversible de $M_N(\mathbb{R})$, et $b \in \mathbb{R}^N$. On cherche $x \in \mathbb{R}^N$ minimisant :

$$J(x) = \|Ax - b\|^2$$

sous les contraintes $x_i \geq 0$ $(1 \leq i \leq N)$.

On note $0 < \lambda_1 \leq \cdots \leq \lambda_n$ les valeurs propres de A^*A.

1. Cacluler ∇J et $\nabla^2 J$.
2. Montrer que J est elliptique, et que ∇J est lipshitzienne (préciser les constantes associées, $\alpha = 2\lambda_1$ et $L = 2\lambda_N$).
3. Expliciter l'algorithme du gradient projeté associé au problème de minimisation considéré ici. En déduire que si $\tau < \frac{\lambda_1}{2\lambda_N^2}$, l'algorithme convergence.
4. Ecrire une fonction *x=minimise(A,b,τ,r)* qui retourne le résultat x de l'algorithme à l'itération r.
5. Utiliser la fonction précédente pour trouver la solution du problème quand :
$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} -3 \\ 4 \end{pmatrix}$$
6. Vérifier à la main la solution obtenue à la question précédente en appliquant les relations de Kuhn et Tucker.

2 Gradient projeté (bis)

On considère dans $\mathbb{R}^3$ l'ensemble :

$$C = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3,\ x_1^2 + x_2^2 \leq 1\right\}$$

1. (a) Quelle est la forme géométrique de C ? Montrer que C est convexe.
   (b) Expliciter la projection orthogonale P_C sur C.
   (c) Ecrire une fonction *z=projection(x)* qui en entrée prend un vecteur x de $\mathbb{R}^3$ et en sortie renvoie le vecteur z (de même taille que que x) projection de x sur C.
2. Soit A une matrice réelle 3×3 inversible, et $b \in \mathbb{R}^3$. On considère la fonction $J : \mathbb{R}^3 \to \mathbb{R}$ définie par $J(x) = \|Ax - b\|^2$, et le problème (P) : *minimiser* $J(x)$ *sous la contrainte* $x \in C$

   (a) Montrer que (P) admet une unique solution.

(b) Expliciter l'algorithme du gradient projeté associé à (P). Pour quelles valeurs du pas est-on assuré de la convergence ?

(c) Ecrire une fonction *x=minimise(A,b,τ,r)*, faisant appel à la fonction *projection*.

(d) Calculer alors numériquement le minimum de J sur C quand :

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$$

(e) Trouver, dans ce cas particulier, la valeur empirique à 10^{-4} près du pas critique τ_c au-delà duquel la convergence ne se produit plus.

3 Plus petite valeur propre

Soit A une matrice symmétrique positive d'ordre N, et $J(x) = \langle Ax, x\rangle$. On note $0 \leq \lambda_1 \leq \cdots \leq \lambda_N$ les valeurs propres de A.

On cherche les points de minimum de J sur la sphère unité.

1. Justifier pourquoi les λ_i sont dans $\mathbb{R}^+$.
2. Ecrire le problème d'optimisation considéré.
3. En utlisant le théorème des extrema liés, montrer ques l'ensemble des solutions du problème est constitué des vecteurs propres de A de norme 1 associés à la plus petite valeur propre.
4. Résoudre le problème $\max_{\|x\|=1} \|Ax\|$ (on pourra le réécrire sous la forme $\max_{\|x\|^2=1} \|Ax\|^2$) avec A matrice inversible. En déduire le lien entre les valeurs propores de A^*A et la norme de A.

4 Fonction de pénalisation

Dans la méthode de pénalisation, la fonction de pénalisation doit être C^1, ce qui impose des resctrictions quand à son choix.

Soit g une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$. Soit g^+ la fonction définie par :

$$g^+(x) = \max(g(x), 0)$$

1. Montrer que g^+ est convexe, puis que $(g^+)^2$ est convexe (on pourra supposer g deux fois dérivable).
2. Montrer que la fonction $s \mapsto (s^+)^2$ de $\mathbb{R}$ dans $\mathbb{R}$ est dérivable.
3. En déduire que si g est dérivable, alors $(g^+)^2$ est aussi dérivable. Calculer sa différentielle en fonction de celle de g.
4. Proposer des fonctions de pénalisation pour les contraintes suivantes : $x \leq 0$, $g(x) = 0$, $g(x) \leq 0$.
5. La composée de 2 fonctions convexe est-elle convexe ?

5 Contrainte active

On veut résoudre numériquement le problème de minimiser la fonction :

$$f(x_1, x_2) = x_1^2 + x_2^2 - 14x_1 - 6x_2 - 7$$

sous les contraintes :

$$x_1 + x_2 \leq 2 \quad \text{et} \quad x_1 + 2x_2 \leq 3$$

1. Calculer la différentielle et la Hessienne de f et des contraintes. La fonction f est elle convexe ? elliptique ? à gradient Lipshitz ?
2. Résoudre d'abord numériquement le problème d'optimisation sans contrainte, par une méthode de votre choix.
3. Implémenter une méthode de pénalisation afin d'obtenir une indication du minimum et des contraintes actives. Quelle est la contrainte qui semble active ?
4. Implémenter la méthode de Newton-Lagrange, qui consiste à faire une méthode de Newton sur le système d'optimalité (extrema liés) (on suppose la contrainte active partout). Voir le poly de cours page 28.
5. Vérifier le résultat précédent en utilisant la méthode d'Uzawa (voir le polycopié de cours page 40)

6 Plus grande valeur propre

Soit A une matrice symmétrique positive d'ordre N, et $J(x) = -\langle Ax, x\rangle$. On note $0 \leq \lambda_1 \leq \cdots \leq \lambda_N$ les valeurs propres de A.

On cherche les points de minimum de J dans la boule unité.

1. Justifier pourquoi les λ_i sont dans $\mathbb{R}^+$.
2. Ecrire le problème d'optimisation considéré.
3. En utlisant les relations de Kuhn et Tucker, montrer ques l'ensemble des solutions du problème est constitué des vecteurs propres de A de norme 1 associés à la plus grande valeur propre.

7 Calcul d'opérateurs proximaux

On rappelle que :

$$y = (I + h\partial F)^{-1}(x) = \text{prox}_h^F(x) = \underset{u}{\text{argmin}} \left\{ \frac{\|u - x\|^2}{2h} + F(u) \right\}$$

Calculer les opérateurs proximaux suivants :

1. $$F(u) = \frac{1}{2}\|u\|_2^2$$

   Alors
   $$\mathrm{prox}_h^F(x) = \frac{x}{1+h}$$

2. $$F(u) = \frac{1}{2}\|u-f\|_2^2$$

   Alors
   $$\mathrm{prox}_h^F(x) = \frac{x+hf}{1+h}$$

3. $$F(u) = \frac{1}{2}\|Ku-f\|_2^2$$

   Alors
   $$\mathrm{prox}_h^F(x) = (Id + hK^*K)^{-1}(x + hK^*f)$$

4. $$F(u) = \|u\|_1$$

   Alors
   $$\mathrm{prox}_h^F(u) = ST(u,h)$$

   avec $ST(u,h)$ le seuillage doux de u de paramètre h, i.e. $ST(u,h) = u-h$ si $u > h$, $ST(u,h) = u+h$ si $u < -h$, et $ST(u,h) = 0$ si $|u| \leq h$.

5. $$F(u) = \|\nabla u\|_1$$

   On a $y = \mathrm{prox}_h^F(x)$ si et seulement si $y = x - h\,div(z)$ avec z solution de
   $$\min_{\|z\|_\infty \leq 1} \|div(z) + x/h\|_2^2$$

   On peut facilement résoudre ce problème avec un algo de gradient projeté. Indication : $H(z) = \|div(z)+x/h\|_2^2$. Alors $\nabla H(z) = -2\nabla(div(z)+x/h)$. Et si $C = \{z \ , \ \|z\|_\infty \leq 1\}$, alors l'opérateur proximal def $\chi_C$est la projection orthogonale sur C.