Méthode du gradient, lagrangien et algorithme d'Uwaza

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il est amené à prendre

L'objet du problème est l'étude de méthodes analytiques (méthodes du gradient, du Lagrangien) pour résoudre l'équation linéaire $A.x = b$ où A est une matrice symétrique positive, inversible, b un vecteur donné de $\mathbb{R}^n$ et x un vecteur inconnu de $\mathbb{R}^n$ ou d'un sous-espace vectoriel F de $\mathbb{R}^n$.

Dans tout le problème, l'entier n est un entier naturel supérieur ou égal à 2 ($n \geq 2$); la base canonique de $\mathbb{R}^n$ est notée $e_1, e_2, \dots, e_n$; le produit scalaire de deux vecteurs x et y de $\mathbb{R}^n$ est noté $(x|y)$.

La norme d'un vecteur x est notée $\|x\|$.

Les matrices considérées sont réelles; l'espace vectoriel des matrices carrées réelles d'ordre n est noté $\mathcal{M}_n(\mathbb{R})$.

Il est admis que l'application qui, à une matrice M de $\mathcal{M}_n(\mathbb{R})$, associe la borne supérieure $N(M)$ des normes des images par M des vecteurs unitaires de $\mathbb{R}^n$ est une norme : $N(M) = \sup_{\|x\|=1} \|M.x\|$.

Une matrice symétrique A est dite positive lorsque, pour tout vecteur x de $\mathbb{R}^n$, le produit scalaire des vecteurs $A.x$ et x est positif ou nul ($(A.x|x) \geq 0$).

Une suite de vecteurs $(x_n)_{n \in \mathbb{N}}$ est dite convergente vers un vecteur x lorsque $\lim_{n \rightarrow +\infty} \|x_n - x\| = 0$.

On confondra dans la suite du texte matrice et endomorphisme canoniquement associé.

Partie 1

Le but de cette partie est la résolution de l'équation $A.x = b$ où A est une matrice carrée d'ordre n symétrique positive et inversible, b un vecteur donné de $\mathbb{R}^n$ et x un vecteur inconnu.

Résultats préliminaires

Soit M une matrice carrée symétrique d'ordre n .

1. Démontrer qu'il existe un plus grand réel p et un plus petit réel q tels que, pour tout vecteur x de $\mathbb{R}^n$, le produit scalaire $(M.x|x)$ vérifie l'encadrement suivant : $p\|x\|^2 \leq (M.x|x) \leq q\|x\|^2$. Préciser ces deux réels p et q en fonction des valeurs propres de la matrice M .
2. Montrer que, pour que cette matrice M soit inversible et positive, il faut et il suffit que toutes ses valeurs propres soient strictement positives.

3. Démontrer que la norme $N(M)$ d'une matrice M symétrique est égale à la plus grande valeur absolue des valeurs propres λ_i ($1 \leq i \leq n$) de la matrice M : $N(M) = \sup_{(1 \leq i \leq n)} |\lambda_i|$.

Etant donné la matrice carrée, d'ordre n , symétrique positive A et le vecteur b , soit α un réel strictement positif strictement majoré par $2/\lambda_n$ ($0 < \alpha < 2/\lambda_n$) où λ_n est la plus grande valeur propre de la matrice A ; soit $(x^k)_{k \in \mathbb{N}}$ la suite définie par un premier vecteur x^0 choisi arbitrairement dans $\mathbb{R}^n$ et par la relation de récurrence suivante : pour tout entier naturel k , $x^{k+1} = x^k + \alpha(b - Ax^k)$.

Etude de la suite $(x^k)_{k \in \mathbb{N}}$

4. Démontrer que la suite $(x^k)_{k \in \mathbb{N}}$ est une suite convergente de limite le vecteur z de l'espace $\mathbb{R}^n$, solution de l'équation $Ax = b$.

Soit f la fonction réelle, définie dans $\mathbb{R}^n$, par la relation : $f(x) = \frac{1}{2}(Ax|x) - (b|x)$.

Minimum de f

5. Calcul préparatoire : démontrer que l'expression $f(x+u) - f(x)$ se calcule en fonction des expressions $(A.u|u)$, $(A.x|u)$ et $(b|u)$.
6. Démontrer que la fonction $f : x \mapsto f(x)$ est de classe C^1 .
7. Exprimer le vecteur $\nabla f(x)$ au moyen de la matrice A et des vecteurs x et b .
Etant donné deux vecteurs x et u de $\mathbb{R}^n$, soit $I(x, u)$ l'expression suivante :
 $I(x, u) = f(x+u) - f(x) - (\nabla f(x)|u)$.
8. Démontrer que, pour tout vecteur x donné, il existe deux constantes positives ou nulles r et s telles que, pour tout vecteur u , $I(x, u)$ vérifie la relation suivante : $r\|u\|^2 \leq I(x, u) \leq s\|u\|^2$.
9. Démontrer que, pour que la fonction f admette en z un minimum, il faut et il suffit que le vecteur z vérifie la relation $Az = b$.

Recherche du minimum de f :

Soit α un réel compris strictement entre 0 et $2/\lambda_n$ ($0 < \alpha < 2/\lambda_n$).

10. Etant donné un vecteur x de $\mathbb{R}^n$, déterminer le signe de l'expression suivante $f(x - \alpha \nabla f(x)) - f(x)$.
11. Proposer, à partir de ce résultat, une méthode pour construire une suite de vecteurs $(y^k)_{k \in \mathbb{N}}$ qui converge vers le vecteur z en lequel la fonction f atteint son minimum ; la justification de la convergence n'est pas demandée.

Partie 2

Le but de cette partie est de rechercher un vecteur x appartenant à un sous-espace vectoriel F de $\mathbb{R}^n$ qui vérifie l'équation $Ax = b$ où A est une matrice carrée d'ordre n symétrique positive et inversible. Le sous-espace vectoriel F de $\mathbb{R}^n$ est supposé être le noyau d'une matrice B appartenant à $\mathcal{M}_n(\mathbb{R})$; ce noyau est supposé différent de tout l'espace $\mathbb{R}^n$ ($\ker B \neq \mathbb{R}^n$). L'équivalence, établie dans la première partie, entre d'une part résoudre l'équation $Ax = b$ et d'autre part chercher le vecteur z rendant minimum la fonction f définie sur $\mathbb{R}^n$ par la relation suivante $f(x) = \frac{1}{2}(Ax|x) - (b|x)$, conduit à se poser le problème suivant : Soit

B une matrice appartenant à $\mathcal{M}_n(\mathbb{R})$ dont le noyau F est différent de $\mathbb{R}^n$; rechercher un vecteur $\bar{x}$ appartenant à F rendant minimum la restriction de la fonction f au sous-espace vectoriel F .

Existence du minimum de la fonction f dans F :

12. Démontrer que la fonction f possède la propriété suivante : pour tout réel c , il existe un réel ρ tel que, pour tout vecteur x de F tel que $\|x\| \geq \rho$ alors $f(x) \geq c$.
13. En déduire que, si y est un point de F , il existe un réel r tel que pour tout vecteur x vérifiant ($\|x\| \geq r$) on ait $f(x) \geq f(y)$.
14. Démontrer à l'aide du résultat précédent qu'il existe au moins un vecteur $\bar{x}$ du sous-espace vectoriel F en lequel la restriction de la fonction f à ce sous espace F atteint un minimum.
15. Démontrer que la fonction f est strictement convexe c'est-à-dire : pour tout couple $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ de vecteurs et tout réel λ appartenant à l'intervalle ouvert $]0, 1[$, les valeurs prises par la fonction f vérifient la relation suivante $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$, où l'inégalité est stricte si et seulement si les vecteurs x et y sont différents. En déduire qu'il existe un seul vecteur $\bar{x}$ en lequel la fonction f atteint son minimum dans F .

Propriétés du point $\bar{x}$:

16. Démontrer que, pour qu'un vecteur y de F rende minimum la restriction de la fonction f au sous-espace vectoriel F , il faut et il suffit que le vecteur $Ay - b$ soit orthogonal à ce sous-espace F de $\mathbb{R}^n$.
17. Démontrer que la valeur prise par la fonction f au point $\bar{x}$, en lequel elle atteint son minimum dans F , est donnée par la relation suivante : $f(\bar{x}) = -\frac{1}{2}(A\bar{x}|\bar{x}) = -\frac{1}{2}(b|\bar{x})$.

Le Lagrangien L

Soit L la fonction définie sur l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$ par la relation suivante : $L(x, y) = f(x) + (y|Bx)$.

Un point (x^*, y^*) de l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$ est dit point selle de la fonction L , s'il possède la propriété suivante : quel que soit le point (x, y) de l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$, les valeurs prises par la fonction L aux points (x^*, y) , (x^*, y^*) et (x, y^*) vérifient la double inégalité suivante : $L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*)$.

Propriétés du Lagrangien et de ses points selles

18. Etablir dans $\overline{\mathbb{R}}$ l'inégalité suivante : $\sup_{y \in \mathbb{R}^n} \left(\inf_{x \in \mathbb{R}^n} L(x, y) \right) \leq \inf_{x \in \mathbb{R}^n} \left(\sup_{y \in \mathbb{R}^n} L(x, y) \right)$.
On pourra montrer d'abord que si $x \notin F = \ker B$ alors $\sup_{y \in \mathbb{R}^n} L(x, y) = +\infty$ puis que si $x \in F$ alors $\sup_{y \in \mathbb{R}^n} L(x, y) = f(x)$
Il est supposé dans toute la suite qu'il existe un point selle (x^*, y^*) de la fonction L .
19. Démontrer que la valeur prise par la fonction L en un point selle (x^*, y^*) vérifie les égalités suivantes : $L(x^*, y^*) = \sup_{y \in \mathbb{R}^n} \left(\inf_{x \in \mathbb{R}^n} L(x, y) \right) = \inf_{x \in \mathbb{R}^n} \left(\sup_{y \in \mathbb{R}^n} L(x, y) \right)$.
20. Démontrer, pour tout point (x_1, y_1) de $\mathbb{R}^n \times \mathbb{R}^n$, les équivalences suivantes :

$$\forall y \in \mathbb{R}^n, L(x_1, y) \leq L(x_1, y_1) \Leftrightarrow Bx_1 = 0$$

$$\forall x \in \mathbb{R}^n, L(x_1, y_1) \leq L(x, y_1) \Leftrightarrow Ax_1 + {}^t By_1 = b$$

21. Soient x_1 un vecteur du sous-espace vectoriel F et y_1 un vecteur de $\mathbb{R}^n$. Démontrer qu'une condition nécessaire et suffisante pour que le couple (x_1, y_1) soit un point selle du Lagrangien L est que le vecteur x_1 réalise le minimum de la restriction de la fonction f à F et que les vecteurs x_1 et y_1 vérifient la relation suivante : $Ax_1 + {}^tBy_1 = b$. La suite logique est la recherche d'un point selle du Lagrangien L .

Algorithme d'Uzawa

Soit toujours (x^*, y^*) un point selle, supposé exister ; étant donné un vecteur y^0 arbitraire de $\mathbb{R}^n$, une suite $(\rho_m)_{m \in \mathbb{N}}$ de réels, qui seront précisés plus loin, soient $(x^m)_{m \in \mathbb{N}}$ et $(y^m)_{m \in \mathbb{N}}$ les deux suites de vecteurs définies par les conditions suivantes :

- Pour tout entier naturel m , le vecteur x^m est le vecteur qui rend minimum la fonction $x \mapsto L(x, y^m)$.
- Pour tout entier naturel m , le vecteur y^{m+1} est défini par la relation suivante : $y^{m+1} = y^m + \rho_m Bx^m$.

22. Démontrer que les conditions énoncées permettent de déterminer tous les termes de ces deux suites $(x^m)_{m \in \mathbb{N}}$ et $(y^m)_{m \in \mathbb{N}}$ et que les vecteurs de ces suites vérifient, pour tout entier naturel m , les relations suivantes :

$$A(x^m - x^*) + {}^tB(y^m - y^*) = 0,$$

$$y^{m+1} - y^* = y^m - y^* + \rho_m B(x^m - x^*) \text{ où } x^* \text{ et } y^* \text{ sont les deux vecteurs d'un point selle de } L.$$

23. En déduire l'égalité ci-dessous :

$$\|y^{m+1} - y^*\|^2 = \|y^m - y^*\|^2 - 2\rho_m(A(x^m - x^*)|(x^m - x^*)) + (\rho_m)^2\|B(x^m - x^*)\|^2.$$

Convergence de la suite numérique de terme général $\|y^m - y^*\|^2$

24. **Un résultat préliminaire.** Démontrer l'existence d'une matrice carrée d'ordre n symétrique positive inversible, notée $A^{1/2}$, telle que : $(A^{1/2})^2 = A$.

Soit C la matrice définie par la relation suivante :

$$C = A^{-1/2} \cdot {}^tB \cdot B \cdot A^{-1/2}, \text{ où la matrice } A^{-1/2} \text{ est la matrice inverse de la matrice } A^{1/2}.$$

25. Démontrer que la matrice C est une matrice symétrique positive. Etablir qu'il existe une constante ν telle que, pour tout vecteur u de $\mathbb{R}^n$, l'inégalité ci-dessous soit vraie :

$$\|Bu\|^2 \leq \nu(Au|u).$$

Soient α et β deux réels tels que le segment $[\alpha, \beta]$ soit contenu dans l'intervalle ouvert $]0, 2/\nu[$, ($0 < \alpha < \beta < 2/\nu$). La suite des réels ρ_m est supposée vérifier pour tout entier naturel m l'inégalité suivante : $\alpha \leq \rho_m \leq \beta$.

26. Démontrer que la suite de terme général $\|y^m - y^*\|^2, m \in \mathbb{N}$ est monotone décroissante ; utiliser, pour simplifier, la suite $(u^m)_{m \in \mathbb{N}}$ dont le terme général est défini par la relation suivante : $u^m = x^m - x^*$.

27. En déduire la convergence et la limite de la suite $(x^m)_{m \in \mathbb{N}}$.

FIN DE L'ÉNONCÉ