Outils mathématiques - Optimisation
DEA IVR
Examen du 20 décembre 2001
Tous les documents sont autorisés

1 Critères d'arrêt

Une méthode d'optimisation itérative ne va, en général, jamais atteindre un minimum « exact » (c'est-à-dire un point où le gradient de la fonction de coût est exactement égal à zéro). C'est pourquoi beaucoup de méthodes utilisent un ou plusieurs « critères d'arrêt », qui déterminent quand on stoppe les itérations et se satisfait de la solution actuelle.

- Proposez de tels critères d'arrêt qui vous paraissent plausibles.

2 Problèmes de moindres carrés linéaires sous contraintes linéaires

(a) Soient $\mathbf{x} = (x_1, x_2, x_3)^\mathsf{T}$ les inconnues du problème suivant :

$$\begin{array}{ll} \text{minimiser} & f(\mathbf{x}) = (x_1 + 2x_2 + 4x_3 - 6)^2 + (3x_1 + 8x_2 + 5x_3 - 3)^2 \\ \text{sous la contrainte} & \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \end{pmatrix} \end{array}$$

- Appliquez la méthode esquissée au §14 pour résoudre ce problème (jusqu'à obtention du système linéaire sur les inconnues et le multiplicateur de Lagrange).

(b) Considérons un problème de moindres carrés linéaires général, sous contraintes linéaires générales :

$$\begin{array}{ll} \text{minimiser} & f(\mathbf{x}) = \frac{1}{2} \left(\|\mathsf{J}\mathbf{x} + \mathbf{b}\|_2 \right)^2 \\ \text{sous les contraintes} & \mathsf{A}\mathbf{x} = \mathbf{y} \end{array}$$

Soient les dimensions des différentes matrices et vecteurs comme suit :

– Il y a n inconnues, donc $\mathbf{x}$ est de longueur n.
– La fonction de coût est composée de m résidus, donc J est de dimension $m \times n$ et $\mathbf{b}$ de longueur m.
– Il y a p contraintes, donc A est de dimension $p \times n$ et $\mathbf{y}$ de longueur p.

- Donnez la forme générale (c'est-à-dire en fonction de $\mathsf{J}, \mathbf{b}, \mathsf{A}$ et $\mathbf{y}$) du système linéaire permettant de résoudre les inconnues et les multiplicateurs de Lagrange.
- Indiquez les dimensions des différentes parties de ce système linéaire.
- Considérons le cas où les contraintes sur $\mathbf{x}$ ne sont pas linéairement indépendantes. Quel problème peut être causé par ceci, quant à la résolution du système linéaire ?

3 Initialisation pour des problèmes sous contraintes linéaires

La plupart des méthodes d'optimisation qui ont été traitées dans le cours procèdent par une recherche progressive, partant d'une solution initiale hypothétique. Considérons maintenant les problème sous contraintes. Typiquement, les méthodes appropriées calculent, au cours des itérations, des solutions satisfaisant toutes les contraintes. Ce calcul est souvent basé sur l'hypothèse que la solution précédente satisfait déjà toutes les contraintes. Or, ceci n'est pas forcément le cas pour la solution initiale (donnée par l'utilisateur).

Soit $\mathbf{x}_0$ la solution initiale, donnée par l'utilisateur, pour un problème dont les contraintes forment un système linéaire :

$$\mathsf{A}\mathbf{x} = \mathbf{y}$$

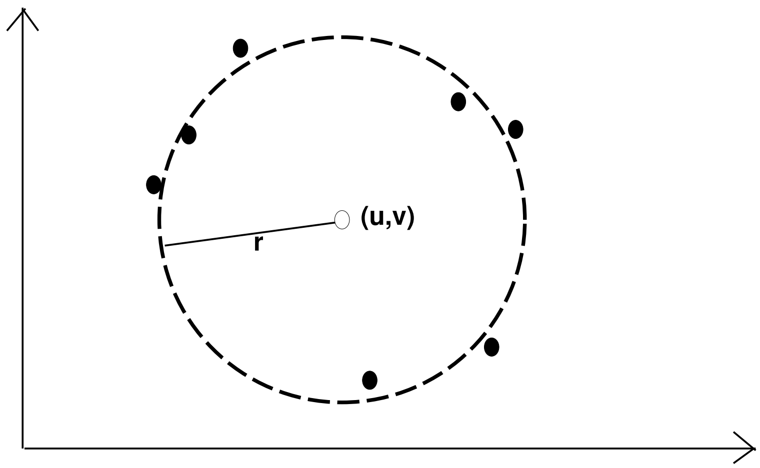
Nous nous intéressons donc au cas où $\mathbf{x}_0$ ne satisfait pas toutes ces contraintes.

- Trouvez un moyen pour calculer une solution initiale $\mathbf{x}'_0$ qui soit « proche » de $\mathbf{x}_0$ et qui satisfasse toutes les contraintes (trouver x'_0 est en effet en soi un problème d'optimisation sous contraintes . . .).

4 Ajustement d'un cercle

Considérons le problème suivant : étant donnés m points dans le plan, le but est d'estimer le « meilleur » cercle, c'est-à-dire le cercle qui passe le plus proche par les points donnés. Soient $(u_i, v_i), i = 1, ..., m$ les coordonnées des points. Soient les inconnues du problème : u et v, les coordonnées du centre du cercle, et r son rayon.

- Définissez une fonction de coût pour estimer les paramètres du cercle.



- Avec quelle méthode résoudriez-vous l'optimisation de votre fonction de coût et pourquoi ? Ecrivez explicitement tous les ingrédients pour appliquer la méthode choisie (gradients, Hessiens, Jacobiens, ...).
- Qu'est-ce que vous feriez pour améliorer le conditionnement de ce problème d'optimisation ? Quels facteurs peuvent influencer la stabilité de la solution ?

A Solution Question 1 – Critères d'arrêt

Différentes possibilités sont :

- Valeur de la fonction de coût inférieur à un seuil (qui dépend de la nature du problème d'optimisation).
- Différence entre les valeurs de la fonction de coût à deux itérations successives inférieure à un seuil.
- Norme du gradient inférieure à un seuil.
- Différence entre les valeurs des paramètres à deux itérations successives inférieure à un seuil.
- Borne supérieure au nombre d'itérations (le critère d'arrêt vraiment le plus basique ...).

B Solution Question 2 – Problèmes de moindres carrés linéaires sous contraintes linéaires

B.1 Question 2 (a)

On a :

$$\begin{aligned}
\mathsf{A} &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \\
\mathbf{b} &= \begin{pmatrix} 10 \end{pmatrix} \\
\mathbf{g}(\mathbf{x}) &= \begin{pmatrix} 20x_1 + 52x_2 + 38x_3 - 30 \\ 52x_1 + 136x_2 + 96x_3 - 72 \\ 38x_1 + 96x_2 + 82x_3 - 78 \end{pmatrix} = \begin{pmatrix} 20 & 52 & 38 \\ 52 & 136 & 96 \\ 38 & 96 & 82 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \begin{pmatrix} 30 \\ 72 \\ 78 \end{pmatrix}
\end{aligned}$$

Après introduction du multiplicateur de Lagrange λ, les conditions nécessaires pour un minimum local sont :

$$\begin{aligned}
\mathsf{A}\mathbf{x} &= \mathbf{b} \\
\mathbf{g}(\mathbf{x}) - \mathsf{A}^\mathsf{T}\lambda &= \mathbf{0}
\end{aligned}$$

Donc :

$$\begin{aligned}
\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \mathbf{x} &= \begin{pmatrix} 10 \end{pmatrix} \\
\begin{pmatrix} 20 & 52 & 38 \\ 52 & 136 & 96 \\ 38 & 96 & 82 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \begin{pmatrix} 30 \\ 72 \\ 78 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \lambda &= \mathbf{0}
\end{aligned}$$

Regroupant ces équations en une équation matricielle, on obtient :

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 20 & 52 & 38 & -1 \\ 52 & 136 & 96 & -1 \\ 38 & 96 & 82 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix} = \begin{pmatrix} 10 \\ 30 \\ 72 \\ 78 \end{pmatrix}$$

La solution est donnée par :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix} = \begin{pmatrix} \frac{196}{13} \\ -\frac{73}{13} \\ \frac{7}{13} \\ 0 \end{pmatrix}$$

B.2 Question 2 (b)

Le système linéaire et ses dimensions :

$$\begin{array}{ccc} \begin{pmatrix} \mathsf{A} & 0 \\ \mathsf{J}^\mathsf{T}\mathsf{J} & -\mathsf{A}^\mathsf{T} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \lambda \end{pmatrix} & = & \begin{pmatrix} \mathbf{y} \\ -\mathsf{J}^\mathsf{T}\mathbf{b} \end{pmatrix} \\ \begin{pmatrix} p \times n & p \times p \\ n \times n & n \times p \end{pmatrix} \begin{pmatrix} n \\ p \end{pmatrix} & = & \begin{pmatrix} p \\ n \end{pmatrix} \end{array}$$

Si les contraintes ne sont pas linéairement indépendantes, la matrice A n'est pas de rang plein. La matrice du système d'équations ci-dessus est singulière et la solution du système par inversion directe de la matrice échouera.

C Solution Question 3 – Initialisation pour des problèmes sous contraintes linéaires

Une solution est de trouver $\mathbf{x}_0'$ qui satisfasse toutes les contraintes, et qui soit le plus proche de $\mathbf{x}_0$. On peut formuler ceci comme un problème d'optimisation sous contraintes :

$$\begin{array}{ll} \text{trouver } \mathbf{x}_0' \text{ qui minimise} & f(\mathbf{x}_0') = \frac{1}{2}\left(\left(x_{0,1}' - x_{0,1}\right)^2 + \ldots + \left(x_{0,n}' - x_{0,n}\right)^2\right) \\ \text{sous les contraintes} & \mathsf{A}\mathbf{x}_0' = \mathbf{y} \end{array}$$

La fonction de coût peut aussi s'écrire comme :

$$f(\mathbf{x}_0') = \frac{1}{2}\left(\|\mathsf{J}\mathbf{x}_0' + \mathbf{b}\|_2\right)^2$$

avec :

$$\begin{aligned} \mathsf{J} &= \mathtt{I} \\ \mathbf{b} &= -\mathbf{x}_0 \end{aligned}$$

Selon la méthode des multiplicateurs de Lagrange, nous obtenons le système d'équations linéaires suivant (cf. aussi la solution de la question 2) :

$$\begin{pmatrix} \mathsf{A} & 0 \\ \mathtt{I} & -\mathsf{A}^\mathsf{T} \end{pmatrix} \begin{pmatrix} \mathbf{x}_0' \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x}_0 \end{pmatrix}$$

On peut le résoudre en deux étapes. D'abord, en utilisant la « partie basse » du système, on obtient :

$$\mathbf{x}_0' = \mathbf{x}_0 + \mathsf{A}^\mathsf{T}\lambda$$

En remplaçant ceci dans la « partie haute », on obtient :

$$\mathsf{A}\mathbf{x}_0 + \mathsf{A}\mathsf{A}^\mathsf{T}\lambda = \mathbf{y}$$

Puis :

$$\lambda = \left(\mathsf{A}\mathsf{A}^\mathsf{T}\right)^{-1}(\mathbf{y} - \mathsf{A}\mathbf{x}_0)$$

L'inversion de matrice est bien définie ici puisque $\mathsf{A}\mathsf{A}^\mathsf{T}$ est une matrice carrée, non singulière si A est de rang plein.

En remplaçant ceci dans la solution pour $\mathbf{x}'_0$, on obtient finalement :

$$\mathbf{x}'_0 = \mathbf{x}_0 + \mathsf{A}^\mathsf{T}\left(\mathsf{A}\mathsf{A}^\mathsf{T}\right)^{-1}(\mathbf{y} - \mathsf{A}\mathbf{x}_0)$$

D Solution Question 4 – Ajustement d'un cercle

Une fonction de coût plausible est :

$$\sum_{i=1}^{m}\left((u_i - u)^2 + (v_i - v)^2 - r^2\right)^2$$

Il s'agit d'un problème de moindres carrés, donc on peut appliquer par exemple la méthode de Levenberg-Marquardt pour le résoudre. Pourtant, ceci n'est pas obligatoire : l'un des atouts des méthodes de moindres carrés est qu'elles ne nécessitent pas le calcul de dérivées secondes qui, souvent, peut être lourd. Pour ce qui est de la fonction de coût proposée, ceci n'est pas vraiment un problème. Donc, des méthodes non-spécifiques aux problème de moindres carrés, peuvent aussi être utilisées ici.

Le conditionnement peut être amélioré comme il l'a été montré lors du cours pour l'exemple de l'estimation de droites. Parmi les facteurs influençant la stabilité de la solution, on peut nommer :

- le nombre de points disponibles ;
- leur répartition, surtout le degré de recouvrement du cercle : si les points ne se trouvent que sur un petit arc du cercle, la solution du rayon et du centre ne sera pas stable ;
- le bruit dans les données (la déviation des points mesurés du cercle original).