

Optimisation numérique  
Méthode de la région de confiance I  
Chapitre 4.1 et 4.2\*

LONGCHAMP Claude

15 octobre 2009

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Optimisation sans contrainte de fonctions continues . . . . .	2
1.2	Aperçu de la méthode . . . . .	2
1.3	Taille de la région de confiance . . . . .	2
<b>2</b>	<b>Méthode de la région de confiance</b>	<b>3</b>
2.1	Optimisation de fonctions modèles quadratiques . . . . .	3
2.2	Aperçu de l'algorithme . . . . .	4
<b>3</b>	<b>Le point de Cauchy</b>	<b>4</b>
3.1	Amélioration sur le point de Cauchy . . . . .	5
<b>4</b>	<b>Solutions presque exactes du sous-problème</b>	<b>9</b>
4.1	Caractériser les solutions exactes . . . . .	9
4.2	Calculer près des solutions exactes . . . . .	9

---

\*NOCEDAL Jorge and WRIGHT Stephen J., *Numerical Optimization*, Springer Series in Operations Research, Springer 1999.

# 1 Introduction

## 1.1 Optimisation sans contrainte de fonctions continues

Soient  $x \in \mathbb{R}^n$  un vecteur de variables et  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  la fonction lisse de coût à optimiser. Alors le problème d'optimisation sans contrainte revient à calculer  $\min_{x \in \mathbb{R}^n} f(x)$ .

## 1.2 Aperçu de la méthode

L'idée de la méthode de la région de confiance est d'approcher  $f(x_k)$  par une fonction plus simple  $m_k$  dans une région  $R_k$  de confiance

$$R_k = \{x_k + p; \|p\| \leq \Delta\}$$

pour un  $\Delta$  fixé. Cette région de confiance doit être suffisamment petite pour que

$$m_k(x_k + p) \cong f(x_k + p)$$

Au lieu de résoudre l'équation:

$$f(x_{k+1}) = \min_{\|p\| \leq \Delta} f(x_k + p) \quad (1)$$

on résoud  $m_k(x_{k+1}) = \min_{\|p\| \leq \Delta} m_k(x_k + p)$  pour trouver  $x_{k+1}$ . Si la différence entre  $f(x_{k+1})$  et  $m_k(x_{k+1})$  est trop grande, on diminue le  $\Delta$  (et donc la région de confiance) et on résoud (1) à nouveau. Un avantage de cette méthode est que toutes les directions sont prises en compte. Par contre, nous devons faire attention à ne pas trop nous éloigner de  $x_k$ , car la fonction  $m_k$  n'approche  $f$  que sur une région proche de  $x_k$ .

## 1.3 Taille de la région de confiance

La taille de la région de confiance est essentielle quant à l'efficacité de chaque pas. En effet, une région de confiance "trop petite" peut être à l'origine de calculs inutiles, et peut, de plus, faire converger la solution vers un point qui ne correspond pas au minimum de la fonction objectif  $f$ . Si, par contre, la région de confiance est "trop grande", le point qui minimise la fonction modèle  $m_k$  peut être assez éloigné du point qui minimise la fonction objectif, dans ce cas, on doit réduire la taille de la région et réessayer.

En pratique, la taille de la région de confiance est déterminée en se basant sur la performance de l'algorithme au cours des dernières itérations. Si

le modèle est bon, c'est-à-dire si la fonction modèle est capable de décrire correctement le comportement de la fonction objectif, la taille de la région de confiance peut être augmentée afin de déterminer des pas plus performants. Dans le cas contraire, on réduit la taille de la région de confiance et on réessaie.

## 2 Méthode de la région de confiance

### 2.1 Optimisation de fonctions modèles quadratiques

La fonction objectif  $f$  est remplacée par une fonction modèle quadratique  $m_k$  dans une certaine région autour d'un point  $x_k$  donné. Les deux premiers termes intervenant dans la fonction modèle  $m_k$  de chaque itéré  $x_k$  sont identiques aux deux premiers termes de la série de Taylor de  $f$  autour de  $x_k$ . On a donc:

$$m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p \quad (2)$$

où  $f_k = f(x_k)$ ,  $\nabla f_k$  est le gradient de  $f$  au point  $x_k$  et  $B_k$  est une matrice symétrique représentant le hessien de  $f$  à l'itéré  $x_k$  ou une approximation de ce dernier.

Lorsque le hessien exact est utilisé, la fonction modèle est construite en utilisant les trois premiers termes de la série de Taylor de  $f$  en  $x_k$ . Dans ce cas, la méthode est appelée *méthode à région de confiance de Newton*. Dans ce chapitre, aucune indication ne sera donné quant à l'éventuelle exactitude du hessien.

Les régions de confiance ajoutent au problème d'optimisation initiale une contrainte sur la longueur du pas. On cherche la solution de chaque sous-problème de la forme:

$$\min_{p \in \mathbb{R}^n} m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p \quad \text{sous les contraintes} \quad \|p\| \leq \Delta_k \quad (3)$$

où  $\Delta_k > 0$  est le rayon de la région de confiance. Dans ce chapitre, on définit la norme par la norme euclidienne. L'utilisation de la norme euclidienne permet de prouver que la solution  $p_k^*$  du problème (3) minimise la fonction  $m_k$  à l'intérieur de la région de confiance définie autour de  $x_k$ .

## 2.2 Aperçu de l'algorithme

Pour un point  $x_k$  et un rayon  $\Delta_k$  donnés, on détermine l'efficacité de la fonction modèle par le rapport suivant:

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \quad (4)$$

Ce rapport est utilisé comme critère d'actualisation du rayon  $\Delta_k$  de la région de confiance. On remarque que le dénominateur est forcément positif, car  $p_k$  est la solution de (3), ce qui veut dire que  $m_k$  décroît. Si  $\rho_k < 0$ ,  $f(x_k + p_k) > f(x_k)$ . Le pas doit alors être rejeté. Si  $\rho_k$  est positif, on accepte le pas et le rayon de la région de confiance est actualisé selon les valeurs prise par  $\rho_k$ . Si  $\rho_k$  est proche de 1, les fonctions modèles et objectifs sont en bon accord au cours de ce pas. Il est alors plus intéressant d'augmenter le rayon de la région de confiance pour la prochaine itération afin d'obtenir des pas plus performants. Pour finir, si  $\rho_k$  est proche de 0,  $m_k$  ne représente pas correctement  $f$ . On diminue donc le rayon de confiance.

En pratique, pour effectuer (4) on a besoin d'évaluer (3):

## 3 Le point de Cauchy

De manière similaire aux méthodes de recherches par lignes, la détermination de pas optimaux n'est pas une condition nécessaire pour obtenir une convergence globale. Bien que, en principe, on recherche une solution optimale du sous-problème (3), il suffit de trouver une solution approchée  $p_k$  à l'intérieur de la région de confiance, qui produise une "réduction suffisante" de la fonction modèle. Cette réduction peut être obtenue par la méthode du point de Cauchy  $p_k^c$ :

### ALGORITHME 3.1

1. Trouver un vecteur  $p_k^s$ , solution du problème (3) linéarisé, soit

$$p_k^s = \arg \min_{p \in \mathbb{R}^n} (f_k + \nabla f_k^T p) \quad \text{sous les contraintes} \quad \|p\| \leq \Delta_k$$

2. Calculer le scalaire  $\tau_k > 0$  qui minimise  $m_k(\tau p_k^s)$ , soit

$$\tau_k = \arg \min_{\tau > 0} m_k(\tau p_k^s) \quad \text{sous les contraintes} \quad \|\tau p_k^s\| \leq \Delta_k$$

3. Finalement, on obtient le point de Cauchy, soit

$$p_k^c = \tau_k p_k^s$$

## Calcul pratique du point de Cauchy

On trouve facilement la solution de  $p_k^s$  par:

$$p_k^s = \frac{-\Delta_k}{\|\nabla f_k\|} \nabla f_k$$

La détermination de  $\tau_k$  dépend du signe de  $\nabla f_k^T B_k \nabla f_k$ . Si  $\nabla f_k^T B_k \nabla f_k \leq 0$ , la fonction  $m_k(\tau p_k^s)$  décroît de façon monotone avec  $\tau$ , et ce, même si  $\nabla f_k \neq 0$ . Le pas  $\tau_k$  est donc la plus grande valeur qui satisfasse le rayon de confiance.  $\tau_k$  est alors égal à 1. Si  $\nabla f_k^T B_k \nabla f_k > 0$ , la fonction  $m_k(\tau p_k^s)$  est convexe et quadratique en  $\tau$ . Le pas  $\tau_k$  est alors soit la valeur qui minimise la fonction quadratique  $\frac{\|\nabla f_k\|^3}{\Delta_k \nabla f_k^T B_k \nabla f_k}$ , soit la valeur limite 1 déterminé précédemment.

### 3.1 Amélioration sur le point de Cauchy

Bien que le point de Cauchy  $p_k^s$  fournisse une réduction suffisante de la fonction modèle  $m_k$  pour produire une convergence globale et que le coût du calcul est assez petit, il est intéressant de chercher une meilleure solution approximative de (3). En effet, le point de Cauchy est défini comme le point minimisant  $m_k$  le long de la plus grande pente  $-\nabla f_k$ . C'est tout simplement la mise en oeuvre de la méthode de la plus forte descente avec un choix particulier de la longueur de chaque pas. On considère trois méthodes pour trouver une solution approchée de (3). Dans ce chapitre, on va se focaliser sur le travail interne d'une seule itération. On enlève alors l'indice  $k$  de  $\Delta_k$ ,  $p_k$  et  $m_k$  pour simplifier la notation. Le sous-problème (3) devient donc

$$\min_{p \in \mathbb{R}^n} m(p) = f + g^T p + \frac{1}{2} p^T B p \quad \text{sous les contraintes} \quad \|p\| \leq \Delta \quad (5)$$

où  $g$  est le gradient de  $f$ . On dénote la solution par  $p^*(\Delta)$  pour montrer la dépendance de  $\Delta$ .

### METHODE DE DOGLEG

On commence par examiner l'effet du rayon  $\Delta$  de la région de confiance de la solution  $p^*(\Delta)$  du sous-problème (5). Si  $B$  est définie positive, on a toujours noté que le minimiseur sans contrainte de  $m$  est le pas intégral  $p^B = -B^{-1}g$ . Quand ce point est admissible pour (5), c'est évidemment une solution, on a donc:

$$p^*(\Delta) = p^B$$

Quand  $\Delta$  est minuscule, la restriction  $\|p\| \leq \Delta$  garantit que le terme quadratique de  $m$  a un petit effet sur la solution (5). La solution  $p(\Delta)$  est approximativement la même solution que nous obtiendrions en minimisant la fonction linéaire  $f + g^T p$  sur  $\|p\| \leq \Delta$ , on a donc:

$$p^*(\Delta) \cong -\Delta \frac{g}{\|g\|} \quad \text{quand } \Delta \text{ est petit.}$$

Pour des valeurs intermédiaires de  $\Delta$ , la solution  $p^*(\Delta)$  suit une trajectoire courbe. La méthode de Dogleg trouve une solution approchée en remplaçant cette trajectoire courbe par un chemin constitué de deux segments. Le premier va de l'origine vers le minimiseur sans contrainte le long de la direction de la plus grande pente définie par:

$$p^U = -\frac{g^T g}{g^T B g} g$$

Le second segment va de  $p^U$  vers  $p^B$ . Formellement, on écrit cette trajectoire par  $\tilde{p}(\tau)$  pour  $\tau \in [0, 2]$ :

$$\tilde{p}(\tau) = \begin{cases} \tau p^U & \text{si } 0 \leq \tau \leq 1 \\ p^U + (\tau - 1)(p^B - p^U) & \text{si } 1 \leq \tau \leq 2 \end{cases}$$

**Lemme 1.** *Soit  $B$  définie positive, Alors:*

1.  $\|\tilde{p}(\tau)\|$  est une fonction croissante de  $\tau$ .
2.  $m(\tilde{p}(\tau))$  est une fonction décroissante de  $\tau$ .

Par ce lemme, il en suit que le chemin  $\tilde{p}(\tau)$  intersecte la limite de  $\|p\| = \Delta$  de la région de confiance à au plus un point si  $\|p^B\| \geq \Delta$  et sinon il n'y a pas d'intersection.

Puisque  $m$  est décroissante sur le chemin, la valeur choisie de  $p$  sera  $p^B$  si  $\|p^B\| \leq \Delta$ , sinon elle sera sur le point d'intersection entre le bord de la région de confiance et le chemin de Dogleg. Dans ce cas, on calcule la valeur appropriée de  $\tau$  en résolvant l'équation quadratique:

$$\|p^U + (\tau - 1)(p^B - p^U)\|^2 = \Delta^2$$

## METHODE DE MINIMISATION DU SOUS-ESPACE EN 2 DIMENSIONS

Lorsque  $B$  est définie positive, la méthode de Dogleg peut être légèrement plus sophistiquée par l'élargissement de la recherche de  $p$  par l'ensemble du sous-espace à deux dimensions engendré par les vecteurs  $p^U$  et  $p^B$ . Le sous-problème (5) est alors remplacé par:

$$\min_p m(p) = f + g^T p + \frac{1}{2} p^T B p \text{ sous les contraintes } \|p\| \leq \Delta \text{ et } p \in \text{span}(g, B^{-1}g) \quad (6)$$

Cette stratégie est, en fait, une extension de la méthode de Dogleg puisque le chemin de Dogleg réside dans le  $\text{span}[g, B^{-1}g]$ . Un avantage certain de cette méthode est qu'elle peut être modifiée pour gérer le cas où  $B$  serait indéfinie. Si  $B$  a des valeurs propres négatives, le sous-espace de deux dimensions de (6) est remplacé par

$$\text{span}[g, (B + \alpha I)^{-1}g] \text{ pour } \alpha \in (-\lambda_1, -2\lambda_1) \quad (7)$$

où  $\lambda_1$  est la plus négative des valeurs propres. Si  $\|(B + \alpha I)^{-1}g\| \leq \Delta$ , on supprime la recherche du sous-espace de (6) et (7), et, à la place, on définit le pas par  $p = -(B + \alpha I)^{-1}g + \nu$  ou  $\nu$  est un vecteur qui satisfait  $\nu^T (B + \alpha I)^{-1}g \leq 0$ . Dans le cas contraire, si  $B$  n'a pas de valeur propre négative, on prend le point de Cauchy  $p^C$  comme solution de (5).

## METHODE DE STEIHAUG

Les deux méthodes présentées ci-dessus nécessitent la solution d'un système linéaire contenant  $B$  ou  $B + \alpha I$ . Si  $B$  est grande, cette opération peut coûter cher. On est alors poussé à considérer d'autres techniques pour trouver une solution approchée de (5). On aimerait une méthode qui n'exige pas de solution exacte d'un système linéaire mais qui produise quand même une amélioration du point de Cauchy. Steihaug se base sur la méthode des gradients conjugués. Il propose une amélioration de cette méthode pour résoudre le sous-problème de la région de confiance. A chaque itération, on teste tout d'abord la courbure de la fonction modèle  $m$  dans la direction  $-\nabla f_k$ . Si la fonction  $f$  n'est pas convexe dans cette direction, on suit cette direction jusqu'au bord de la région de confiance et on arrête les itérations. De plus, dès qu'un itéré est à l'extérieur de la région de confiance, on suit la dernière direction calculée jusqu'au bord de la région de confiance et on arrête l'algorithme:

### ALGORITHME 3.2

Given  $\epsilon \leq 0$

```

Set  $p_0 = 0, r_0 = 0, d_0 = -r_0;$ 
if  $\|r_0\| \leq \epsilon$ 
    return  $p = p_0;$ 
for  $j = 0, 1, 2, \dots$ 
    if  $d_j^T B d_j \leq 0$ 
        Trouver  $\tau$  t.q  $p = p_j + \tau_j$  minimise  $m(p)$  dans (5) et satisfasse
 $\|p\| = \Delta;$ 
        return  $p;$ 
    Calculer  $\alpha_j = \frac{r_j^T r_j}{d_j^T B d_j};$ 
    Calculer  $p_{j+1} = p_j + \alpha_j d_j$ 
    if  $\|p_{j+1}\| \geq \Delta$ 
        Trouver  $\tau \geq 0$  t.q  $p = p_j + \tau_j$  satisfasse  $\|p\| = \Delta;$ 
        return  $p;$ 
    Calculer  $r_{j+1} = r_j + \alpha_j B d_j$ 
    if  $\|r_{j+1}\| < \epsilon \|r_0\|$ 
        return  $p = p_{j+1};$ 
    Calculer  $\beta_{j+1} = \frac{r_{j+1}^T r_{j+1}}{r_j^T r_j};$ 
    Calculer  $d_{j+1} = -r_{j+1} + \beta_{j+1} d_j;$ 
end for

```

L'initialisation de  $p_0 = 0$  est cruciale car après la première itération, on a  $p_1 = -\frac{g^T g}{g^T B g} g$  qui est exactement le point de Cauchy. Une autre propriété importante de cet algorithme est que chaque itéré  $p_j$  est plus grand dans la norme que son prédécesseur car, dans ce cas, c'est acceptable de stopper l'itération dès que la limite de la région de confiance est atteinte. Cette propriété découle du théorème suivant:

**Théorème 2.** *La série de vecteurs générés par l'algorithme précédent satisfait:*

$$0 = \|p_0\| < \dots < \|p_j\| < \|p_{j+1}\| < \dots < \|p\| = \Delta$$

On remarque donc que, par cette méthode, on se déplace du point de Cauchy  $p^C$  vers le point integral  $p^B$  jusqu'à ce qu'intervienne la région de confiance.



## 4 Solutions presque exactes du sous-problème

### 4.1 Caractériser les solutions exactes

Les méthodes vues ci-dessus ont des avantages de faibles coût et de convergence globale car elles génèrent toutes un point au moins aussi bon que le point de Cauchy. Si le problème est relativement petit, il peut être intéressant d'exploiter le modèle plus pleinement en cherchant une approximation plus proche de la solution du sous-problème (5). Essentiellement, on voit que la solution  $p^*$  des problèmes de la région de confiance satisfait la forme

$$(B + \lambda I)p^* = -g \quad \text{pour } \lambda > 0$$

Par le théorème suivant, on obtient une vérification de la solution du problème (5)

**Théorème 3.** *Le vecteur  $p^*$  est une solution globale du problème de la région de confiance*

$$\min_{p \in \mathbb{R}^n} m(p) = f + g^T p + \frac{1}{2} p^T B p \quad \text{sous les contraintes } \|p\| \leq \Delta \quad (8)$$

$\iff p^*$  est admissible et il existe  $\lambda > 0$  t.q les conditions suivantes sont satisfaites:

1.  $(B + \lambda I)p^* = -g$
2.  $\lambda(\Delta - \|p^*\|) = 0$
3.  $(B + \lambda I)$  est semi-définie positive

La condition 2. est une condition complémentaire qui dit que soit  $\lambda$  ou  $\Delta - \|p^*\|$  doit être égal à 0.

### 4.2 Calculer près des solutions exactes

La caractérisation du Théorème 3. suggère un algorithme pour trouver la solution  $p$  du problème (8):

- Si  $\lambda = 0$ : Par 1.  $Bp = -g$  et par 3.  $B$  est semi-définie positive. On obtient alors  $p = -B^{-1}g$  avec  $\|p\| \leq \Delta$ .
- Sinon: On défini  $p(\lambda) = -(B + \lambda I)^{-1}g$  pour un  $\lambda$  suffisamment large pour que  $B + \lambda I$  soit défini positive et pour un  $\lambda > 0$  t.q  $\|p(\lambda)\| = \Delta$ .

Pour voir si une valeur  $\lambda$  existe pour toutes les propriétés, on fait appel aux valeurs propres de  $B$  et on les utilise pour évaluer les propriétés de  $\|p(\lambda)\|$ . Comme  $B$  est symétrique, il existe une matrice orthogonale  $Q$  et une matrice diagonale  $\Lambda$  t.q  $B = Q\Lambda Q^T$  où  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  et  $\lambda_1 \leq \dots \leq \lambda_n$  sont les valeurs propres de  $B$ . On a alors  $B + \lambda I = Q(\Lambda + \lambda I)Q^T$  et pour  $\lambda \neq \lambda_j$  pour  $j = 1, \dots, n$ , on a

$$p(\lambda) = -Q(\Lambda + \lambda I)^{-1}Q^T g = -\sum_{j=1}^n \frac{q_j^T g}{\lambda_j + \lambda} q_j \quad (9)$$

où  $q_j$  est la  $j$ -ème colonne de  $Q$ . De plus, par orthogonalité de  $q_1, \dots, q_n$ , on a

$$\|p(\lambda)\|^2 = \sum_{j=1}^n \frac{(q_j^T g)^2}{(\lambda_j + \lambda)^2} \quad (10)$$

Grâce à (10), on peut voir: Si  $\lambda > -\lambda_1$ :  $\lambda_j + \lambda > 0$  pour tout  $j$ .  $\|p(\lambda)\|$  est alors continue et monotone décroissante sur l'intervalle  $(-\lambda_1, \infty)$ . On a donc

$$\lim_{\lambda \rightarrow \infty} \|p(\lambda)\| = 0. \quad (11)$$

De plus, si  $q_1^T g \neq 0$ , on a

$$\lim_{\lambda \rightarrow \lambda_1} \|p(\lambda)\| = \infty. \quad (12)$$

Il est alors évident que le graphe de  $\|p(\lambda)\|$  atteint la valeur de  $\|p(\lambda)\| = \Delta$  a exactement un point dans l'intervalle  $(-\lambda_1, \infty)$ .

Dans le cas où  $B$  est définie positive et  $\|B^{-1}g\| \leq \Delta$ , la valeur  $\lambda = 0$  satisfait les 3 conditions du Théorème 3. Il n'est donc pas nécessaire d'effectuer une recherche. Si  $B$  est définie positive et  $\|B^{-1}g\| > \Delta$ , il existe un  $\lambda > 0$  pour lequel  $\|p(\lambda)\| = \Delta$ . On cherche donc une solution de  $\|p(\lambda)\| = \Delta$  dans l'intervalle  $(0, \infty)$ . Dans le cas où  $B$  est indéfinie et  $q_1^T g \neq 0$ , (11) et (12) garantisse qu'on peut trouver une solution dans l'intervalle  $(-\lambda_1, \infty)$ . Nous pourrions utiliser la méthode de Newton pour trouver la valeur  $\lambda > -\lambda_1$  qui résoud

$$\phi_1(\lambda) = \|p(\lambda)\| - \Delta = 0 \quad (13)$$

Le désavantage de cette approche est que si l'on considère la forme de  $\|p(\lambda)\|$  quand  $\lambda$  est plus grand que  $-\lambda_1$  mais très proche de  $-\lambda_1$ , on a

$$\phi_1(\lambda) \cong \frac{C_1}{\lambda + \lambda_1} + C_2$$

où  $C_1 > 0$  et  $C_2$  sont des constantes. Pour ces  $\lambda$ , la fonction  $\phi_1(\lambda)$  est fortement non-linéaire; la méthode de Newton converge alors très mal.

On peut obtenir une meilleure convergence en reformulant (13). On applique la méthode de Newton pour chercher  $\phi_2(\lambda) = 0$ ,

$$\phi_2(\lambda) = \frac{1}{\Delta} - \frac{1}{\|p(\lambda)\|}$$

Alors, pour  $\lambda$  très proche de  $-\lambda_1$ , on a de (9):

$$\phi_2(\lambda) \cong \frac{1}{\Delta} - \frac{\lambda + \lambda_1}{C_3}$$

pour  $C_3 > 0$ . Dans ce cas,  $\phi_2$  est presque linéaire. La méthode de Newton sera alors plus performante.