



# Chapitre 3

## Optimisation locale de fonctions différentiables sans contrainte



### Sujets du chapitre

- Analyse des conditions d'optimalité (équations ou inéquations satisfaites pour toute solution du problème d'optimisation).
- Dédution d'une famille d'algorithmes itératifs globalement convergents à partir de la violation des conditions d'optimalité.
- Présentation de l'algorithme de Newton modifié.
- Algorithmes spécialisés aux fonctions quadratiques.
- Problèmes de moindres carrés.
- Problèmes de très grandes dimensions.

## Introduction

Les problèmes d'optimisation les plus simples sont ceux qui consistent à minimiser (ou maximiser) une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  où la fonction  $f$  est supposée définie dans  $\mathbb{R}^n$  en entier, et de plus de classe  $\mathcal{C}^3(\mathbb{R}^n)$ . En fait, on suppose que  $f$  est suffisamment différentiable pour éviter les difficultés dans l'analyse des algorithmes de solution. De plus, seuls les *minima locaux* retiennent notre attention dans ce premier contact.

Les vecteurs de  $\mathbb{R}^n$  sont des vecteurs *colonnes*. Les gradients des fonctions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , notés  $\nabla f(x)$  sont des vecteurs *lignes*. Une motivation pour cette notation est que le gradient est en fait une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ , donc située dans le dual de  $\mathbb{R}^n$ . Bien que les espaces euclidiens soient auto-duaux, il est commode de considérer le dual comme un espace de vecteurs lignes.

La classe de tels problèmes (aussi appelés *programmes non-linéaires*) peut sembler trop simple pour être vraiment utile et/ou intéressante à étudier. Cependant, l'étude de cette classe de problèmes permet d'introduire les notions fondamentales qui seront utilisées constamment dans l'étude de problèmes plus généraux, comme par exemple les problèmes où la fonction  $f$  n'est plus assez différentiable pour permettre une analyse simple telle que nous présentons dans ce chapitre, ou encore la classe de problèmes de minimisation d'une fonction dans un domaine restreint de  $\mathbb{R}^n$ .

Comme dans les chapitres suivants, nous analysons d'abord les équations et inéquations nommées *conditions d'optimalité* qui caractérisent une solution au problème étudié, et ensuite nous déduisons des algorithmes itératifs à partir de la non-réalisation de ces conditions d'optimalité.

### 3.1 Formulation du problème

Énonçons le problème que nous résolvons dans ce chapitre. Il consiste à trouver un point  $x^*$  qui *minimise* la fonction  $f$  dans  $\mathbb{R}^n$  :

$$f(x^*) = \min_{x \in \mathbb{R}^n} f(x),$$

ou encore trouver  $x^*$  tel que

$$f(x^*) \leq f(x), \quad \forall x \in \mathbb{R}^n.$$

Comme nous l'avons expliqué au chapitre précédent, nous nous limitons à une version *locale* de ce problème, soit qu'il existe un voisinage  $V(x^*)$  de  $x^*$  tel que

$$f(x^*) \leq f(x), \quad \forall x \in V(x^*).$$

Puisque la notion de voisinage entre de manière très intime dans la définition du problème, on doit s'attendre à ce que la notion de proximité soit utilisée dans l'analyse de sa solution.

En général, l'analyse du problème de minimiser localement une fonction de  $n$  variables ne peut pas se réduire à l'analyse de la minimisation locale d'une fonction d'une seule variable. On pourrait récrire le problème ainsi :

$$f(x^*) \leq f(x^* + d), \quad \forall d \in V(0),$$

où  $d$  représente le déplacement de  $x^*$  à un point voisin, ou encore, en séparant la notion de *direction* et la notion de *distance* du déplacement,

$$f(x^*) \leq f(x^* + \theta d), \quad \forall d \in \mathbb{R}^n, \|d\| = 1, \text{ et } \forall \theta < \epsilon,$$

où  $\epsilon$  représente le diamètre d'une boule centrée en  $x^*$  contenue dans  $V(x^*)$ . Cependant, nous ne pouvons pas établir d'équivalence entre les restrictions unidimensionnelles et les problèmes originaux.

En fait, si  $x^*$  est un minimum local de  $f$ , alors 0 est un minimum local de la fonction d'une seule variable  $h_{x^*,d}(\theta) \stackrel{\text{def}}{=} f(x^* + \theta d)$ , et ce quelle que soit la valeur de  $d \in \mathbb{R}^n$  telle que  $\|d\| = 1$ . Cependant, la contraposée n'est pas vraie, comme l'illustre le programme suivant (voir figure 3.2 et l'exercice 3.2.6) :

$$\min(x^2 - y)(x^2 - 4y),$$

pour lequel l'origine est un point stationnaire correspondant à un point de selle. Pourtant, le long de toute droite passant par l'origine, cette fonction possède un minimum local à l'origine !

## 3.2 Conditions d'optimalité

Comme nous venons de le voir, en général, l'analyse du problème de minimiser localement une fonction de  $n$  variables ne peut pas se réduire à l'analyse de la minimisation locale d'une fonction d'une seule variable. Malheureusement, dans la pratique, il est très difficile de traiter systématiquement des exemples comme la fonction  $(x^2 - y)(x^2 - 4y)$ , et on se contente souvent de caractérisations issues de l'analyse de la fonction univariée  $h_{x^*,d}$ .

### 3.2.1 Analyse du problème à une seule variable

Penchons-nous sur le problème de minimiser une fonction réelle  $h(\theta)$ . Supposons sans perte de généralité que nous étudions un minimum local de  $h$  situé en  $\theta = 0$ . Exprimons  $h$  en série de Taylor autour de  $\theta = 0$  :

$$h(\theta) = h(0) + h'(0)\theta + \frac{h''(0)\theta^2}{2!} + \frac{h'''(0)\theta^3}{3!} + \dots + \frac{h^{(j)}(0)\theta^j}{j!} + \dots$$

Supposons maintenant que les  $j - 1$  premières dérivées de  $h$  au point 0 sont toutes nulles, et que la première non-nulle est  $h^{(j)}(0)$ . Alors, en utilisant le théorème de Taylor, on a

$$h(\theta) = h(0) + \frac{h^{(j)}(0)\theta^j}{j!} + R_{j+1}(\theta),$$

où  $R_{j+1}(\theta)$  est le reste d'ordre  $j + 1$ . On voit donc que 0 est un minimum local si  $\frac{h^{(j)}(0)\theta^j}{j!} + R_{j+1}(\theta)$  est positif pour des valeurs suffisamment petites de  $|\theta|$ . Détaillons ce raisonnement.

Maintenant, toujours supposant que  $h$  est suffisamment ( $j + 1$  fois) différentiable, le reste  $R_{j+1}(\theta)$  satisfait à la borne

$$\exists K_1 : |R_{j+1}(\theta)| \leq K_1|\theta^{j+1}|.$$

Plus précisément, rappelons que lorsque les fonctions sont suffisamment différentiables, le reste s'écrit

$$R_{j+1}(\theta) = \frac{h^{(j+1)}(\xi)\theta^{j+1}}{(j+1)!},$$

où comme à l'accoutumée,  $\xi = \alpha\theta$  avec  $0 \leq \alpha \leq 1$ . Puisque  $h$  est suffisamment différentiable,  $h^{(j+1)}$  est une fonction continue, et donc, bornée; en fait, quelle que soit la borne que l'on ait en tête de la forme  $|h^{(j+1)}(0)| + \delta$ , il existe un  $\epsilon_\delta > 0$  tel que si  $|\xi| < \epsilon$ , alors  $|h^{(j+1)}(\xi)| < |h^{(j+1)}(0)| + \delta$ . En d'autres termes, si une fonction  $\phi$  est continue à l'origine, alors, pas trop loin de l'origine,  $\phi(\xi)$  ne sera pas beaucoup plus grand que  $\phi(0)$ . Supposons donc maintenant  $\delta = 1$  de sorte que si  $|\xi| < \epsilon_1$ , on a la garantie que  $|h^{(j+1)}(\xi)| < |h^{(j+1)}(0)| + 1$ . Donc, on peut prendre  $K_1 = \frac{1}{(j+1)!}(|h^{(j+1)}(0)| + 1)$ .

Intéressons-nous maintenant au signe de la quantité  $\frac{h^{(j)}(0)\theta^j}{j!} + R_{j+1}(\theta)$  qui peut se récrire

$$\frac{\theta^j}{j!} \left( h^{(j)}(0) + \frac{h^{(j+1)}(\xi)\theta}{(j+1)} \right). \quad (1)$$

Puisque nous avons fait l'hypothèse que les  $j - 1$  premières dérivées de  $h$  étaient nulles et que la  $j^{\text{ième}}$  est non-nulle, vérifions que le signe de  $(h^{(j)}(0) + \frac{h^{(j+1)}(\xi)\theta}{(j+1)})$  est donné par le signe de  $h^{(j)}(0)$  lorsque  $\theta$  est assez petit, et précisons ce que nous entendons par *assez petit*.

Pour ce faire, argumentons que si  $h^{(j)}(0)$  domine  $\frac{h^{(j+1)}(\xi)\theta}{(j+1)}$  en valeur absolue, ce dernier terme ne peut pas faire changer le signe de la somme, donné par le signe du premier terme. Écrivons donc

$$|h^{(j)}(0)| > \left| \frac{h^{(j+1)}(\xi)\theta}{(j+1)} \right| = \left| \frac{h^{(j+1)}(\xi)}{(j+1)} \right| |\theta|.$$

Donc, si

$$|\theta| < \frac{(j+1)|h^{(j)}(0)|}{|h^{(j+1)}(\xi)|},$$

la somme des deux termes est du même signe que le premier terme  $h^{(j)}(0)$ . Utilisons maintenant la borne que nous avons trouvée, qui assure que si  $\theta < \epsilon_1$ ,  $|h^{(j+1)}(\xi)| < |h^{(j)}(0)| + 1$  pour écrire que si

$$|\theta| < \frac{(j+1)|h^{(j)}(0)|}{|h^{(j+1)}(0)| + 1},$$

alors forcément

$$|\theta| < \frac{(j+1)|h^{(j)}(0)|}{|h^{(j+1)}(\xi)|},$$

à condition bien sûr que  $|\theta| < \epsilon_1$ . Par conséquent, le signe de la somme est donné par le signe de  $h^{(j)}(0)$  dès que

$$|\theta| < \min \left( \epsilon_1, \frac{(j+1)|h^{(j)}(0)|}{|h^{(j+1)}(0)| + 1} \right).$$

Dans ces conditions, le signe que nous voulions étudier, de l'équation (1), est le même que le signe de  $h^{(j)}(0)\theta^j$ .

Maintenant, observons que  $j$  doit être pair, car autrement,  $h^{(j)}(0)\theta^j$  serait négatif si  $\theta$  était de signe opposé à  $h^{(j)}(0)$ . Nous venons donc de déduire que la première dérivée de  $h$  en un minimum local est *d'ordre pair*. Si  $j$  est pair, alors  $h^{(j)}(0)\theta^j$  est positif à condition que  $h^{(j)}(0)$  soit positif. Nous venons donc de déduire que cette dérivée d'ordre pair doit avoir une valeur positive.

**Théorème 3.2.1** *Une condition nécessaire et suffisante pour qu'une fonction analytique  $h : \mathbb{R} \rightarrow \mathbb{R}$  possède un minimum local en  $\theta$  est que la première dérivée non nulle de  $h$  en  $\theta$  soit d'ordre pair, et de valeur positive.*

**\*Exercice 3.2.1** [Extension des conditions d'optimalité] Le théorème 2.2.1 concerne la minimisation sans aucune restriction sur  $\theta$ . Considérons le problème  $\min_{a \leq \theta \leq b} h(\theta)$ . Si un point  $\theta_0$  est dans l'intérieur de l'intervalle  $[a, b]$ , alors le théorème s'applique directement ; lorsque  $\theta_0 = a$  ou  $\theta_0 = b$ , il suffit que  $h$  soit croissante (en  $a$ ) ou décroissante (en  $b$ ) pour avoir un minimum local sous contrainte : en effet, on ne compare  $a$  qu'avec des valeurs  $\theta \geq a$ . Obtenez une condition nécessaire et suffisante pour que  $a$  ou  $b$  soient des minima locaux de  $h$ . L'exercice consiste donc à fournir les conditions encadrées et démontrer :

- $\theta^*$  est min local de la fonction  $h$  si et seulement si
- $\theta^* = a$  et conditions pour que  $h$  soit croissante ou
  - $a < \theta^* < b$  et la première dérivée non nulle de  $h$  est d'ordre pair et de signe positif  
ou
  - $\theta^* = b$  et conditions pour que  $h$  soit décroissante.

### 3.2.2 Conditions d'optimalité pour un problème à $n$ variables

Comme nous l'avons exprimé plus haut, on se contente fréquemment de conditions d'optimalités déduites de l'analyse de problèmes univariés. Nous explorons dans cette section les implications du théorème 3.2.1 lorsqu'appliqué aux fonctions  $h_{x^*,d}$ , quelle que soit la valeur de  $d \in \mathbb{R}^n$ .

**Exercice 3.2.2** [Dérivées de  $h$ ] Obtenez les expressions des deux premières dérivées de  $h_{x,d}(\theta)$  par rapport à  $\theta$ .

Il est bon de se familiariser une bonne fois pour toutes avec la dérivation de fonctions vectorielles. Assurez-vous maintenant d'avoir fait l'exercice A.1.3.

Dans tous les cas, le théorème 3.2.1 nous informe que  $h'_{x^*,d}(0)$  doit être nulle, et ce pour tout  $d$  dans  $\mathbb{R}^n$ . Voyons ce que ça implique.

$$h'_{x^*,d}(0) = \nabla f(x^*)d.$$

En fait,  $h'_{x^*,d}(0)$  représente la dérivée directionnelle de  $f$  en  $x^*$  dans la direction  $d$ . Si  $\nabla f(x^*)d$  doit s'annuler pour tout  $d \in \mathbb{R}^n$ , en particulier, il doit s'annuler pour  $d = \nabla f(x^*)^t$ , et donc  $\|\nabla f(x^*)\| = 0$ , et  $\nabla f(x^*) = 0$ , ce qui constitue la condition d'optimalité de premier ordre.

Remarquons que la recherche de minima locaux de  $f$  peut débiter par la recherche de racines de l'équation  $\nabla f(x) = 0$ ; comme dans le cas d'une fonction réelle d'une seule variable réelle, de telles racines sont nommées *points stationnaires* pour le problème  $\min f(x)$ . Nous discuterons plus loin des gains que nous pouvons effectuer en sachant que l'équation  $\nabla f(x) = 0$  provient d'un problème d'optimisation.

En plus que  $h'(0)$  soit nulle, le théorème nous informe qu'il faut que la première dérivée non nulle de  $h$  soit d'ordre pair. Interprétons cette notion pour la *seconde* dérivée de  $h_{x^*,d}$ .

$$h''_{x^*,d}(0) = d^t \nabla^2 f(x^*)d.$$

C'est la dérivée seconde directionnelle de  $f$  en  $x^*$  dans la direction  $d$ . Si  $x^*$  est un minimum local de  $f$ , d'après le théorème 3.2.1, cette dérivée est soit nulle, ou encore positive. Par ailleurs, si cette dérivée est positive quel que soit  $d \in \mathbb{R}^n$ , d'après le théorème 3.2.1, nous sommes assurés que  $x^*$  est un minimum local de  $f$ . Dans tous les cas, si cette dérivée est négative,  $x^*$  n'est certainement pas un minimum local de  $f$ .

Il découle des propriétés (définitions) des matrices symétriques que  $d^t \nabla^2 f(x^*)d$  est positive (non-négative) quel que soit  $d \in \mathbb{R}^n$  si et seulement si la matrice symétrique  $\nabla^2 f(x^*)$  est une matrice définie positive (semi-définie positive).

**Définition 3.2.1** Une matrice symétrique  $n \times n$   $Q$  est dite définie positive et notée  $Q > 0$  si quel que soit  $v \in \mathbb{R}^n$ ,  $v \neq 0$ ,  $v^t Q v > 0$ ; elle est semi-définie positive et notée  $Q \geq 0$  si  $\forall v \in \mathbb{R}^n$ ,  $v \neq 0$ ,  $v^t Q v \geq 0$ .

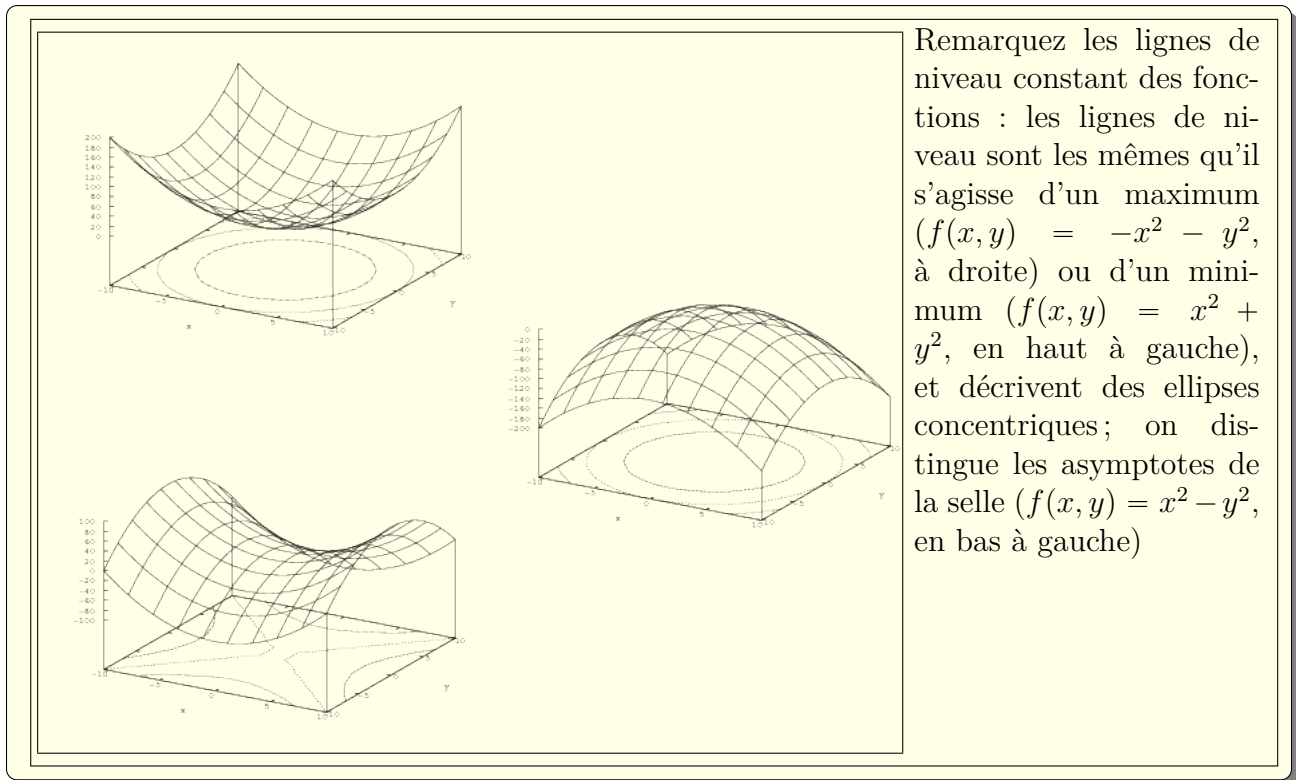


FIGURE 3.1 – Points stationnaires



**Définition 3.2.2** Une matrice  $O$  est dite orthogonale si  $O^t O = \Delta$ , où  $\Delta$  est une matrice diagonale ; elle est dite orthonormale si la matrice  $\Delta$  est l'identité ; dans ce cas, on constate que  $O^{-1} = O^t$ .

**Théorème 3.2.2** Soit  $Q$  une matrice définie positive. Alors,

- les valeurs propres de  $Q$  sont toutes positives ;
- il existe une matrice orthonormale  $O$  telle que  $O^t Q O = \Delta$ , la matrice  $\Delta$  étant une matrice diagonale dont les éléments sont les valeurs propres de  $Q$ , et donc positifs.

Si  $Q$  est semi-définie positive, les valeurs propres sont non-négatives, et la matrice diagonale  $\Delta$  comporte des éléments non-négatifs.

Assurez-vous maintenant d'avoir compris l'exercice A.2.1

Avant de poursuivre, voici un énoncé du théorème de Taylor multidimensionnel. D'abord, pour une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , nous énonçons un développement d'ordre 2.

**Théorème 3.2.3** Soit une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^3$  au voisinage d'un point  $x_0$  ; on peut approcher la valeur de la fonction en un point voisin  $x_0 + d$  de  $x_0$  par une fonction polynomiale de degré 2

$$f(x_0 + d) \approx \mathcal{P}(d) = f(x_0) + \nabla f(x_0)d + \frac{1}{2}d^t \nabla^2 f(x_0)d.$$

De plus, on peut estimer l'erreur commise par un reste :

$$f(x_0 + d) - \mathcal{P}(d) \sim \mathcal{O}(\|d\|^3).$$

Si  $f$  est seulement de classe  $\mathcal{C}^2$ , mais toutefois  $\nabla^2 f$  est une fonction Lipschitz, on a la borne :

$$f(x_0 + d) - \mathcal{P}(d) \sim o(\|d\|^2).$$

Maintenant, pour une fonction  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , nous énonçons un développement d'ordre un.

**Théorème 3.2.4** Soit une fonction  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  de classe  $\mathcal{C}^2$  au voisinage d'un point  $x_0$  ; on peut approcher la valeur de la fonction en un point voisin  $x_0 + d$  de  $x_0$  par une fonction polynomiale de degré un

$$F(x_0 + d) \approx \mathcal{P}(d) = F(x_0) + \nabla F(x_0)d.$$

De plus, on peut estimer l'erreur commise par un reste :

$$\|F(x_0 + d) - \mathcal{P}(d)\| \sim \mathcal{O}(\|d\|^2).$$

Si  $F$  est seulement de classe  $\mathcal{C}^1$ , mais toutefois  $\nabla F$  est une fonction Lipschitz, on a la borne :

$$\|F(x_0 + d) - \mathcal{P}(d)\| \sim o(\|d\|).$$

Nous limitons ces développements afin de pouvoir en exprimer les termes par des expressions ne comportant que des scalaires, vecteurs et matrices.

**Remarque 3.2.1** *Le théorème de la moyenne 2.3.3 n'est pas valide pour des fonctions de  $F(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Par conséquent, l'expression de l'erreur du développement de Taylor du théorème 2.3.4 non plus. Cependant, une version plus faible du théorème de la moyenne stipule que  $\|F(b) - F(a)\| \sim \mathcal{O}(\|b - a\|)$ . Ainsi, la forme du théorème 2.6.1 peut être utilisée.*

Pour des problèmes à  $n$  variables, il est difficile de fournir soit des interprétations, soit des formules algébriques pour les dérivées d'ordre supérieur à deux de la fonction  $h_{x^*,d}$ . Par conséquent, nous nous contenterons de la caractérisation (incomplète) suivante des minima locaux de  $f$  :

**Théorème 3.2.5** *Soit une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$ ; si  $x^*$  est un minimum local de  $f$ , alors  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice semi-définie positive; à l'opposé, si  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice définie positive, alors  $x^*$  est un minimum local de  $f$ .*

**Preuve** La condition nécessaire découle de la discussion précédente concernant les deux premières dérivées de la fonction  $h_{x^*,d}$  en 0.

Nous présentons maintenant une preuve en supposant que  $f$  est de classe  $\mathcal{C}^3$ , laissant en exercice la preuve dans le cas où  $f$  est seulement de classe  $\mathcal{C}^2$ .

Si  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice définie positive, alors, puisque  $f$  est de classe  $\mathcal{C}^3$ , on peut écrire un développement de Taylor avec reste de la forme, où  $d$  est unitaire,  $\|d\| = 1$  :

$$\begin{aligned} f(x^* + \theta d) - f(x^*) &= \theta \nabla f(x^*) d + \frac{1}{2} \theta^2 d^t \nabla^2 f(x^*) d + \mathcal{O}(\theta^3) \\ &= \frac{1}{2} \theta^2 d^t \nabla^2 f(x^*) d + \mathcal{O}(\theta^3); \end{aligned}$$

Puisque  $\nabla^2 f(x^*)$  est définie positive, il existe une constante positive  $K$  telle que  $d^t \nabla^2 f(x^*) d > K$ , d'où

$$f(x^* + \theta d) - f(x^*) > K \frac{1}{2} \theta^2 + \mathcal{O}(\theta^3) = \theta^2 \left( \frac{K}{2} + \mathcal{O}(\theta) \right).$$

Cette dernière quantité est positive si  $\theta$  est assez petit.  $\square$

**Corollaire 3.2.1** *Soit une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ; si  $x^*$  est un maximum local de  $f$ , alors  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice semi-définie négative; à l'opposé, si  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice définie négative, alors  $x^*$  est un maximum local de  $f$ .*

**Corollaire 3.2.2** *Si  $x^*$  est tel que  $\theta^* = 0$  satisfait aux conditions suffisantes pour le problème  $\min_{\theta} h_{x^*,d}(\theta)$ , alors  $x^*$  satisfait aux conditions suffisantes pour  $\min_x f(x)$ .*

Remarquons que le théorème 3.2.5 n'est pas aussi fort que le théorème 3.2.1. En effet, il ne permet pas de distinguer le problème  $\min x^4 + y^4$  du problème  $\min x^4 - y^4$  car en 0, point stationnaire des deux problèmes, le hessien de chacun des problèmes s'annule, et il faudrait examiner leur dérivées d'ordre supérieur.

L'analyse pour des fonctions  $\mathcal{C}^2$  utilise une propriété des matrices.

**Théorème 3.2.6** *Les valeurs propres d'une matrice  $M$  sont des fonctions continues des éléments de la matrice  $M$*

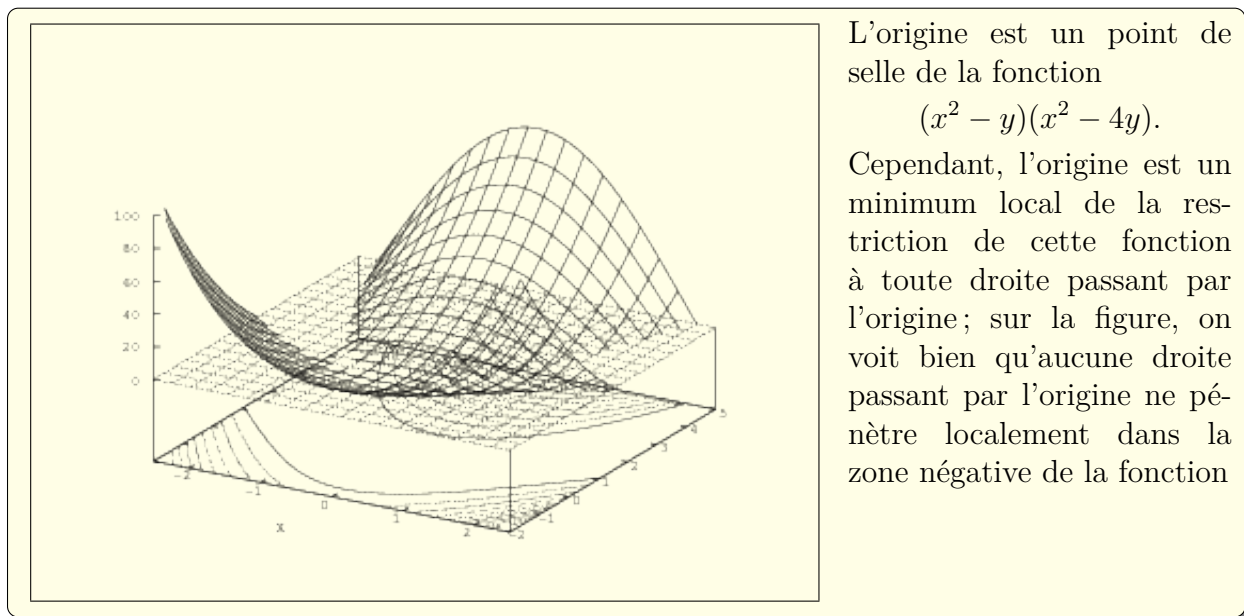
Remarquons que la continuité des valeurs propres n'implique pas, en général, la continuité des vecteurs propres associés sauf si les valeurs propres sont distinctes.

**Exercice 3.2.3** [*Minima forts*] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$ ; si  $x^*$  est un minimum local de  $f$ , montrez qu'un point qui satisfait aux conditions suffisantes de second ordre est un minimum local strict et isolé. Vous pouvez utiliser le théorème 3.2.6 pour assurer que  $\nabla^2 f(x^*)$  demeure définie positive dans un voisinage approprié de  $x^*$

**Exercice 3.2.4** [*Points stationnaires*] Considérez la fonction  $f(x) = (x_1 + x_2)^2 + (2(x_1^2 + x_2^2 - 1) - \frac{1}{3})^2$ .

- a) Vérifiez que l'origine ainsi que le point  $(-\sqrt{\frac{7}{12}}, \sqrt{\frac{7}{12}})$  constituent des points stationnaires de  $f$ .
- b) Déterminez la nature de chacun de ces points stationnaires.

**Exercice 3.2.5** [*Géométrie*] Considérez deux droites d'équations paramétriques  $x(t_1) = o_1 + t_1 d_1$  et  $x(t_2) = o_2 + t_2 d_2$ . On cherche à trouver les points sur les droites les plus proches, permettant de donner la distance entre les droites. On minimise donc  $f(t) = \|x_1(t_1) - x_2(t_2)\|^2$  par rapport à  $t$ , de dimension deux. Utilisez les conditions d'optimalité pour trouver une expression pour  $t^* \in \mathbb{R}^2$ , et vérifiez que c'est bel et bien un minimum local.



L'origine est un point de selle de la fonction

$$(x^2 - y)(x^2 - 4y).$$

Cependant, l'origine est un minimum local de la restriction de cette fonction à toute droite passant par l'origine; sur la figure, on voit bien qu'aucune droite passant par l'origine ne pénètre localement dans la zone négative de la fonction

FIGURE 3.2 – Point de selle

**Exercice 3.2.6** [*Point de selle*] Nous avons présenté plus haut (figure 3.2) la fonction  $(x^2 - y)(x^2 - 4y)$  pour laquelle l'origine n'est qu'un point de selle, mais constitue un minimum local de la fonction objectif le long de toute droite passant par l'origine. Cet exercice approfondi l'étude de ce problème.

- Vérifiez que l'origine est un point stationnaire de la fonction  $f(x, y) = (x^2 - y)(x^2 - 4y)$  :  $\nabla f(0, 0) = 0$ .
- Considérez la fonction réelle  $h_d(\theta) \stackrel{\text{def}}{=} f(\theta d_x, \theta d_y)$ . Vérifiez que quel que soit  $d$  avec  $d_y \neq 0$ ,  $h'_d(0) = 0$  et  $h''_d(0) > 0$ .
- Pour un vecteur  $d$  quelconque, identifiez une valeur  $\epsilon_d > 0$  pour laquelle  $h_d(\theta) \geq 0$  pour toute valeur de  $\theta$  satisfaisant  $|\theta| \leq \epsilon_d$ .
- Vérifiez que  $f(x, y) \leq 0$  pour tous les points satisfaisant  $x^2 = 2y$ .

**Exercice 3.2.7** [*Moindres carrés*] Dans cet exercice, nous explorons un problème d'ajustement de fonction. Les valeurs suivantes de  $x$  et  $y$  sont données :

x	0	1	2	3
y	0	0	3	9

On soupçonne qu'une relation d'ordre 3 relie  $x$  à  $y$ , et on envisage un modèle du type  $y = a_0 + a_3x^3$ .

- a) Formulez le problème d'identifier les valeurs de  $a_0$  et  $a_3$  qui minimisent la somme des carrés des différences entre le modèle et les données.
- b) Exprimez les conditions d'optimalité (nécessaires et suffisantes, de premier et second ordre) du problème en a).
- c) Nous allons traiter une simplification du problème en supposant que  $a_0 = 0$ , donc que la relation recherchée est purement cubique (nouveau problème avec seulement  $a_3$ ). On analyse ici le problème de dimension un qui ne comporte que la variable  $a_3$ . Cette simplification semble raisonnable puisque pour  $x = 0$ ,  $y = 0$ .
  - i) calculez la valeur de  $a_3$  qui minimise la somme des carrés des différences entre le modèle simplifié et les données ;
  - ii) est-ce que les conditions suffisantes d'optimalité sont satisfaites pour la solution du problème de dimension un en i) ?
- d) Revenons maintenant au modèle de dimension deux en a) pour calculer les valeurs de  $a_0$  et  $a_3$ . Calculez les valeurs de  $a_0$  et  $a_3$  du modèle avec l'algorithme du gradient conjugué. Est-ce que les conditions suffisantes d'optimalité sont satisfaites ?

### 3.3 Dédution d'une classe d'algorithmes itératifs

Dans cette section, nous construisons une famille d'algorithmes itératifs pour calculer un minimum local d'une fonction  $f$ . La base de notre étude est la suivante : si un point  $x$  ne satisfait pas aux conditions d'optimalité, il est possible de produire un autre point,  $\bar{x}$  pour lequel  $f(\bar{x}) < f(x)$ . Cette famille d'algorithmes porte le nom de *méthodes de descente*. La disponibilité de méthodes de descentes pour les problèmes d'optimisation en fait des problèmes privilégiés, permettant d'obtenir des propriétés de convergence plus fortes que pour des problèmes d'équations non-linéaires  $F(x) = 0$  lorsque  $F \neq \nabla f$ , pour une certaine fonction réelle  $f$ . En effet, à moins d'accepter des hypothèses assez fortes sur une application  $F$ , il est difficile de construire un algorithme dont la convergence vers un zéro soit assurée. De tels algorithmes prennent la forme d'itérations de points fixes, et seront explorés plus loin. Pour l'instant, nous exploitons la structure d'optimisation de nos problèmes pour construire une classe d'algorithmes assurés de converger vers un point stationnaire quel que soit le point de départ  $x_0$  utilisé.

Supposons que nous disposions d'un point  $x$  qui n'est pas un minimum local de la fonction  $f$ . Alors, au moins une des deux conditions suivantes n'est pas vérifiée :

1.  $\nabla f(x) = 0$ , ou
2.  $\nabla^2 f(x)$  est une matrice semi-définie positive.

Dans les deux sous-section suivante, nous exploitons la première des conditions précédentes.

### 3.3.1 Recherche de points stationnaires

Supposons pour l'instant que nous nous contentons de calculer un point stationnaire, satisfaisant  $\nabla f(x) = 0$ ; donc, nous ne nous préoccupons pas de la condition 2 ci haut. En fait, cette seconde condition sera étudiée plus loin, à la section 3.3.3.

Donc,  $\nabla f(x) \neq 0$ . En se référant au théorème 3.2.1, on constate que l'ensemble des  $d \in \mathbb{R}^n$  pour lesquels  $h'_{x,d} < 0$  satisfait à la relation

$$\nabla f(x)d < 0,$$

et cet ensemble est nommé  $DED(x)$ , c'est-à-dire le *Demi Espace de Diminution* de la fonction  $f$  en  $x$ . La famille des algorithmes de descente utilisera donc des directions de déplacement  $d \in DED(x)$ . En fait, nous n'avons pas encore spécifié la forme générale des algorithmes que nous étudierons ci-après. Il s'agit d'algorithmes de descente qui ont tous la même forme, consistant à calculer d'abord une direction dans le  $DED(x)$ , et ensuite à minimiser (approximativement) la fonction  $h_{x,d}(\theta)$  pour passer au point suivant  $x^+ = x + \theta d$ . Nommons ce passage de  $x$  à  $x^+ = A(x)$ , de telle sorte que  $x^+ = A(x)$ . Les propriétés élémentaires de  $d$  et de la minimisation (approximative) de  $h_{x,d}$  nous assurent que  $f(x^+) = f(A(x)) < f(x)$ . Cependant, en analysant l'application  $A$  qui engendre une suite  $x_{k+1} = A(x_k)$ , il faudra introduire d'autres notions pour garantir que la suite  $\{x_k\}$  possède bel et bien des points d'accumulation qui sont stationnaires (qui satisfont  $\nabla f(x) = 0$ ). Avant d'approfondir cette étude, énonçons le schéma algorithmique que nous analyserons, et observons un exemple simple d'un algorithme typique.

Examinons un exemple : minimisons la fonction quadratique  $f(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$  en utilisant le choix  $d = -\nabla f(x)^t$  comme direction de descente. Puisqu'il s'agit d'une fonction quadratique, on peut calculer le pas  $\theta^*$  exactement : notons  $f(x) = 0.5x^t Q x + cx$ , où dans l'exemple,  $Q = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix}$  et  $c = (2 \quad -2)$ ; alors,  $\nabla f(x) = x^t Q + c$ . Les conditions d'optimalité pour  $\min_{\theta > 0} f(x + \theta d)$  sont données par  $\nabla f(x + \theta d)d = 0$ , ce qui se simplifie à  $\nabla f(x)d + \theta d^t Q d = 0$ , ou encore

$$\theta^* = -\frac{\nabla f(x)d}{d^t Q d}.$$

Pour notre exemple, on a de plus  $d = -\nabla f(x)^t$  de sorte que

$$\theta^* = \frac{\|\nabla f(x)\|^2}{\nabla f(x)Q\nabla f(x)^t}.$$

<div style="border: 1px solid black; display: inline-block; padding: 2px 5px; margin-bottom: 5px;">Itération de descente</div> <pre style="margin: 0; padding: 5px;"> { Donnés : <math>x_0</math>, <math>\epsilon &gt; 0</math> et <math>0 &lt; M &lt; \infty</math> } terminé <math>\leftarrow</math> faux borné <math>\leftarrow</math> vrai <math>x \leftarrow x_0</math> <b>tantque</b> ( <math>\neg</math> terminé <math>\wedge</math> borné )   <math>d \leftarrow</math> direction(<math>f, x</math>)   { <math>h'_{x,d} &lt; 0.</math> }   <math>\theta^* \leftarrow</math> arg min <math>h_{x,d}(\theta)</math>   <math>x \leftarrow x + \theta^* d</math>   terminé <math>\leftarrow</math> <math>\ \nabla f(x)\  \leq \epsilon</math>   borné <math>\leftarrow</math> <math>\ x\  &lt; M</math> </pre>
---

Algorithme 3.1: Descente simplifiée.

Appliquons maintenant l'algorithme à partir du point initial  $(10, 5)^t$ . En ce point,  $\nabla f(10, 5) = \begin{pmatrix} 32 & -12 \end{pmatrix}$  et le calcul de  $\theta$  donne  $\theta^* = 0.197297$ . Le prochain point engendré est donc le point  $\begin{pmatrix} 10 \\ 5 \end{pmatrix} + 0.197297 \begin{pmatrix} -32 \\ 12 \end{pmatrix} = \begin{pmatrix} 3.686486 \\ 7.367568 \end{pmatrix}$ . La suite des itérations est donnée dans le tableau 3.1. On peut observer la nature cyclique de cet algorithme, connu sous le nom d'algorithme du gradient, ou encore de la pente la plus forte. À toutes les itérations paires les déplacements sont parallèles et les valeurs du pas  $\theta^*$  sont les mêmes. Cette propriété sera exploitée à la section 3.6.2.

Iter	$\begin{pmatrix} x \\ y \end{pmatrix}$	$f(x, y)$	$\theta$	$\nabla f(x, y)$
0	$\begin{pmatrix} 10.000000 \\ 5.000000 \end{pmatrix}$	135.000000	0.197297	(32.00000  -12.00000)
1	$\begin{pmatrix} 3.686486 \\ 7.367568 \end{pmatrix}$	19.778378	1.073529	( 2.01081  5.362162)
2	$\begin{pmatrix} 1.527822 \\ 1.611129 \end{pmatrix}$	2.174566	0.197297	( 4.889030  -1.833386)
3	$\begin{pmatrix} 0.563229 \\ 1.972851 \end{pmatrix}$	-0.514983	1.073529	( 0.307216  0.819243)
4	$\begin{pmatrix} 0.233424 \\ 1.093370 \end{pmatrix}$	-0.925898	0.197297	( 0.746957  -0.280109)

TABLE 3.1 – Itérations de l'algorithme du gradient sur la fonction  $f(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$ .

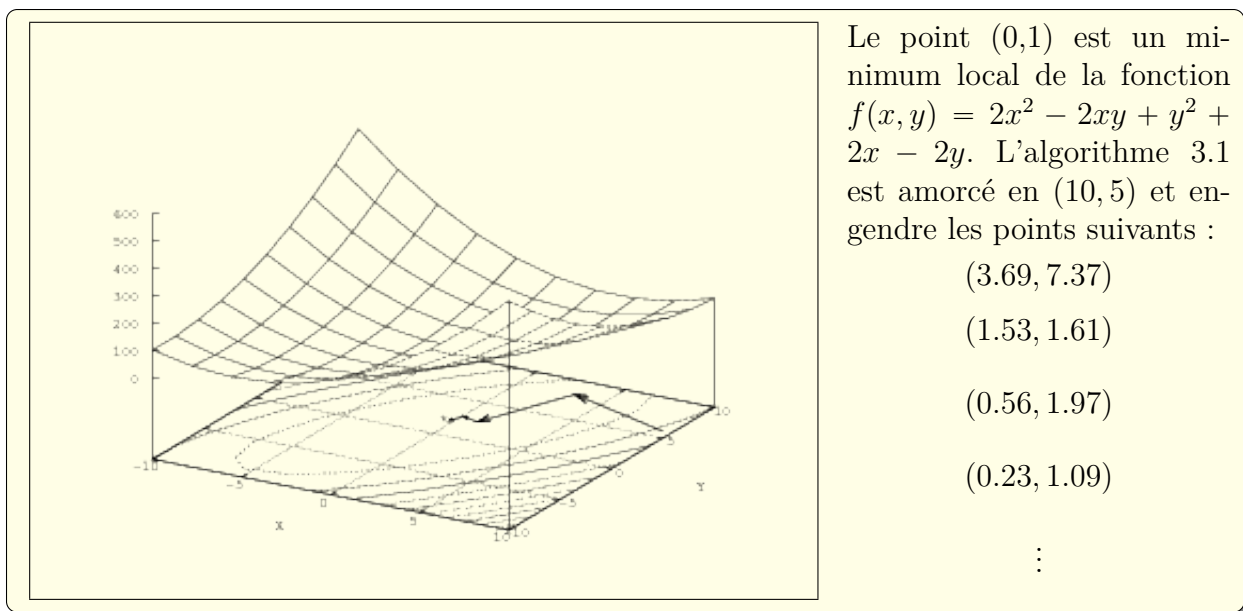


FIGURE 3.3 – Exemple de l'algorithme

**Exercice 3.3.1** [Algorithme du gradient]

- a) Pour la fonction  $f(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$ , vérifiez que si à une certaine étape  $k$  l'algorithme atteint le point  $\begin{pmatrix} 0 \\ 1 - 1/5^k \end{pmatrix}$ , alors à l'étape  $k + 2$ , il aura atteint le point



$$\begin{pmatrix} 0 \\ 1 - 1/5^{k+1} \end{pmatrix}.$$

b) Pour la fonction  $x^2 + 2y^2 + 4x + 4y$ , vérifiez par induction que l'algorithme du gradient amorcé à l'origine engendre la suite  $\begin{pmatrix} \frac{2}{3^k} - 2 \\ (-\frac{1}{3})^k - 1 \end{pmatrix}$ . Déduisez-en le minimum de la fonction.

### 3.3.2 Conditions pour la convergence

Notre but est maintenant de produire une suite de points  $\{x_k\}$  dont les points d'accumulation sont stationnaires, et ce quel que soit le point de départ  $x_0$  de la suite. Nous utiliserons pour ce faire le fait que nous tentons de minimiser une fonction  $f$ , ce qui nous aide à évaluer la qualité relative de deux points consécutifs de la suite.

Plusieurs auteurs ont proposé des schémas d'analyse pour des algorithmes de descente. Le moins restrictif de ceux-ci a été proposé par Zangwill. Simplifié pour notre contexte de minimisation sans contrainte, il s'énonce ainsi :

Si l'application  $A$  qui décrit le passage de  $x_k$  à  $x_{k+1}$  est continue, et si  $f(x_{k+1}) < f(x_k)$ , alors les points d'accumulation de la suite  $\{x_k\}$ , s'il y en a, sont stationnaires pour le problème  $\min f(x)$ .

Bien qu'étant une caractérisation favorite de l'auteur, nous ne l'approfondirons pas maintenant, car nous donnerons la préférence à un schéma d'analyse moins abstrait, plus facilement applicable directement, et dû à Wolfe.

Pour énoncer ce schéma d'analyse, nous avons besoin des conditions suivantes qui caractérisent les directions  $d$  et les pas  $\theta$  conjointement qui assurent que les points d'accumulation de la suite engendrée par un algorithme de descente les utilisant sont tous stationnaires.

**Définition 3.3.1** Une direction  $d$  est considérée suffisamment descendante s'il existe deux constantes positives  $\gamma_0$  et  $\gamma_1$  indépendantes de  $x$  telles que  $d$  satisfasse aux inégalités suivantes :

$$\begin{aligned} \nabla f(x)d &\leq -\gamma_0 \|\nabla f(x)\|^2; \\ \|d\| &\leq \gamma_1 \|\nabla f(x)\|. \end{aligned}$$

La définition précédente nous assure que toute direction  $d$  utilisée par un algorithme de descente est un vecteur assez "long", et fait un angle assez aigu avec l'opposé de  $\nabla f$ . Il est clair que  $d = -\nabla f(x)^t$  satisfait à la définition précédente.

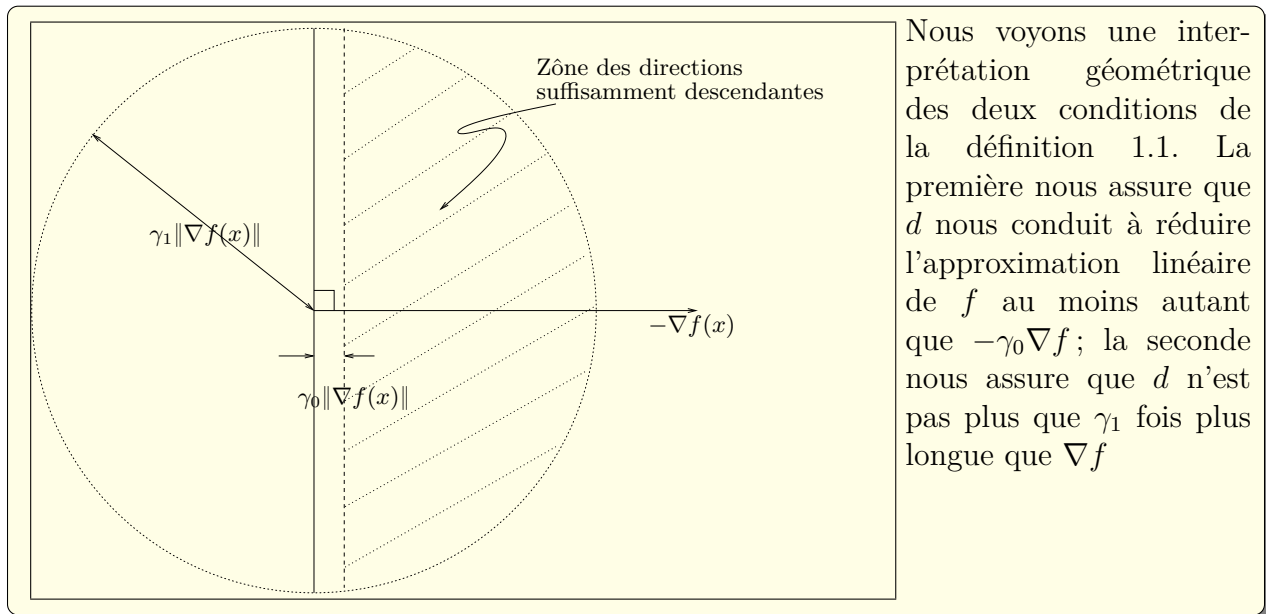


FIGURE 3.4 – Direction suffisamment descendante

**Exercice 3.3.2** [Direction de l'opposé du gradient] Justifiez que la direction  $d = -\nabla f(x)^t$  satisfait aux conditions de la définition 3.3.1; pour quelles valeurs de  $\gamma_0$  et  $\gamma_1$  ?

**Exercice 3.3.3** [Géométrie des directions] Quel est l'angle le plus grand entre une direction  $d$  et l'opposé du gradient tel que la direction satisfasse la définition 3.3.1? L'angle est fonction de  $\gamma_0$  et  $\gamma_1$ .

Une fois qu'une direction est obtenue, nous allons spécifier une longueur de déplacement admissible. On utilise aussi l'expression "pas de déplacement". Mentionnons d'entrée que la fonction  $h_{x,d}(\theta) = f(x + \theta d)$  est décroissante à l'origine, car  $h'_{x,d}(0) = \nabla f(x)d$  qui est suffisamment négatif selon la définition 3.3.1.

**Définition 3.3.2** Un pas  $\theta$  est dit admissible pour une direction suffisamment descendante  $d$  lorsqu'il satisfait aux deux inégalités suivantes, nommées critère d'Armijo et de Wolfe respectivement :

$$f(x + \theta d) - f(x) \leq \tau_0 \theta \nabla f(x)d, \quad \tau_0 \in ]0, \frac{1}{2}[, \quad (\text{Armijo})$$

On voit que les pas qui satisfont aux conditions de la définition 3.3.2 constituent un ensemble d'intervalles dans lesquels  $f$  est assez petite alors que ses dérivées directionnelles sont assez grandes

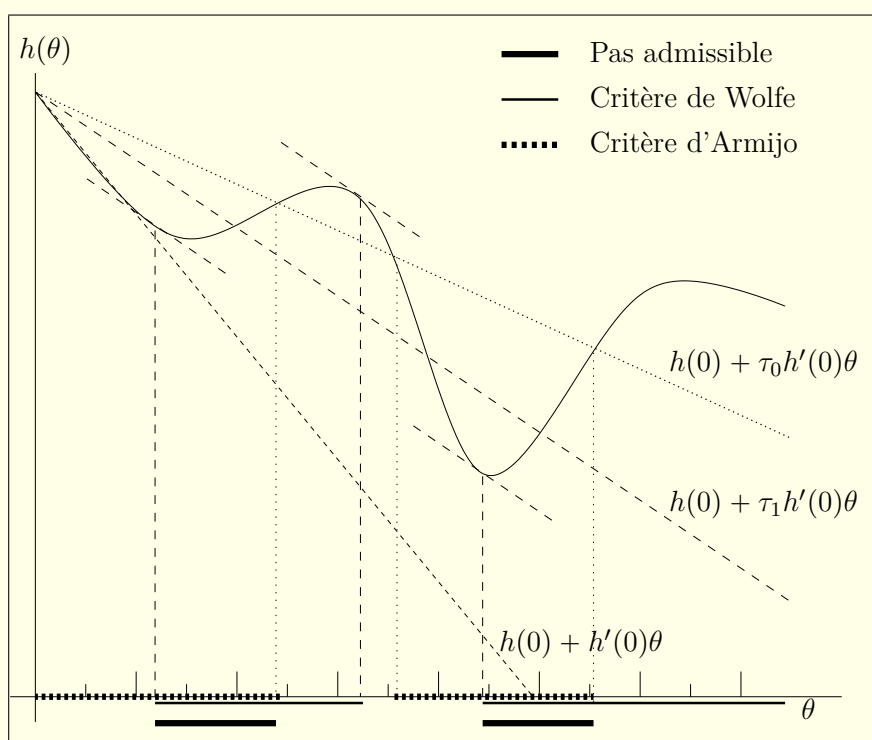


FIGURE 3.5 – Pas admissible

$$\nabla f(x + \theta d)d \geq \tau_1 \nabla f(x)d, \quad \tau_1 \in ]\tau_0, 1[, \quad (\text{Wolfe})$$

où  $\tau_0$  et  $\tau_1$  sont des paramètres fixés pour déterminer l'admissibilité d'un pas. L'ensemble des valeurs admissibles de  $\theta$  sera noté  $\Theta$ .

**Remarque 4** En utilisant la fonction  $h_{x,d}(\theta)$ , on peut exprimer le critère de Armijo comme :

$$h_{x,d}(\theta) - h_{x,d}(0) \leq \tau_0 \theta h'_{x,d}(0)$$

et le critère de Wolfe comme

$$h'_{x,d}(\theta) \geq \tau_1 h'_{x,d}(0).$$

**Exercice 3.3.4** [Pas admissible] Suite de l'exercice 2.9.1 page 87. Considérez la fonction  $h(\theta) = 1 - 12\theta + 7.5\theta^2 - \theta^3$ .

- Pour quelles valeurs de  $\tau_0$  le critère d'Armijo est-il satisfait pour le minimum local situé dans l'intervalle  $[0, 3]$  ?
- Explicitez le(s) intervalle(s) de pas  $\theta$  admissible(s) pour des valeurs de  $\tau_0 = 0.25$  et  $\tau_1$  de votre choix.

L'existence de pas admissibles est cruciale pour l'obtention d'un algorithme bien défini, et fait l'objet du

**Théorème 3.3.1** Soit  $f \in C^2(\mathbb{R}^n)$ , et  $d$  satisfaisant à la définition 3.3.1. Alors, si  $h_{x,d}(\theta)$  est bornée inférieurement, l'intérieur de l'ensemble des pas admissibles  $\Theta$  est non-vide.

Non seulement existe-t-il des pas admissibles, mais puisque l'intérieur de  $\Theta$  est non vide, on peut espérer développer un algorithme qui identifie un élément de  $\Theta$  en un nombre fini de calculs. La preuve de ce théorème sera présentée un peu plus loin.

Avec ces deux définitions, nous sommes en mesure de démontrer le théorème de convergence globale suivant, sachant qu'un algorithme utilisant des directions suffisamment descendantes et des pas admissibles est bien défini.

**Théorème 3.3.2** Soit un algorithme de descente appliqué au problème

$$\min_{x \in \mathbb{R}^n} f(x), \quad f \in C^1(\mathbb{R}^n);$$

supposons qu'à chaque itération la direction utilisée  $d_k$  est une direction suffisamment descendante, et pour lequel le pas utilisé dans cette direction est un pas admissible ; alors, tous les points d'accumulation de la suite  $\{x_k\}$  engendrée par l'algorithme sont des points stationnaires pour le problème  $\min f(x)$ .

**Preuve** Supposons que la suite  $\{x_k\}$  possède au moins un point d'accumulation  $\bar{x}$ . Alors, par l'admissibilité des pas  $\theta_k$  utilisés dans l'algorithme,  $f(x_k) > f(\bar{x})$ , et ce pour tout  $k$ . Donc,  $f(x_k)$  est une suite de nombres réels monotone décroissante, et bornée inférieurement par  $f(\bar{x})$ . Par le théorème de Cauchy,  $f(x_{k+1}) - f(x_k)$  converge vers zéro. Donc,

$$\lim_{k \rightarrow \infty} f(x_{k+1}) - f(x_k) = \lim_{k \rightarrow \infty} f(x_k + \theta_k d_k) - f(x_k) = 0,$$

et en utilisant la condition d'Armijo, ceci entraîne que

$$\lim_{k \rightarrow \infty} \tau_0 \theta_k \nabla f(x_k) d_k = 0.$$

Alors, la première des conditions de la définition 3.3.1 nous assure que  $\tau_0\theta_k\nabla f(x_k)d_k \leq -\tau_0\theta_k\gamma_0\|\nabla f(x_k)\|^2 < 0$ . Ainsi, il se peut que  $\theta_k$  converge vers 0, ou encore que  $\|\nabla f(x_k)\|$  converge vers 0. Supposons que  $\theta_k$  converge vers 0. Alors, la seconde condition de la définition 3.3.1 nous assure que  $0 < \|\theta_k d_k\| \leq \gamma_1\theta_k\|\nabla f(x_k)\|$ , et donc que  $\|\theta_k d_k\|$  converge vers zéro. Donc, par la condition de Wolfe, on a

$$\lim_{k \rightarrow \infty} \nabla f(x_k + \theta_k d_k)d_k = \lim_{k \rightarrow \infty} \nabla f(x_k)d_k \geq \lim_{k \rightarrow \infty} \tau_1 \nabla f(x_k)d_k,$$

ce qui entraîne que  $\lim_{k \rightarrow \infty} \nabla f(x_k)d_k = 0$  puisque  $\tau_1 < 1$ .

Maintenant, puisque  $\nabla f(x_k)d_k < -\gamma_0\|\nabla f(x_k)\|^2$ , d'après la définition 3.3.1, on retrouve

$$0 = \lim_{k \rightarrow \infty} \nabla f(x_k)d_k \leq -\gamma_0 \lim_{k \rightarrow \infty} \|\nabla f(x_k)\|^2,$$

et donc que  $\lim_{k \rightarrow \infty} \|\nabla f(x_k)\| = 0$ . Finalement, puisque  $\nabla f$  est une fonction continue,  $\nabla f(\bar{x}) = 0$ .  $\square$

Un autre variante utilise le *procédé* d'Armijo qui consiste à rechercher un pas  $\theta$  qui satisfait à la première condition d'admissibilité parmi les puissances de  $\frac{1}{2}$ ; la première de ces puissances est éloignée de zéro par le procédé plutôt que par la condition d'admissibilité.

**Théorème 3.3.3** *Soit un algorithme de descente appliqué au problème*

$$\min_{x \in \mathbb{R}^n} f(x), \quad f \in \mathcal{C}^1(\mathbb{R}^n);$$

*supposons qu'à chaque itération la direction utilisée  $d_k$  est une direction suffisamment descendante, et pour lequel le pas utilisé dans cette direction est la première valeur de la suite*

$$\{1, 1/2, 1/4, \dots, (1/2)^m \dots\}$$

*qui satisfasse au critère d'Armijo; alors, tous les points d'accumulation de la suite  $\{x_k\}$  engendrée par l'algorithme sont des points stationnaires pour le problème  $\min f(x)$ .*

**Preuve** Imitons la preuve du théorème 3.3.2. Il est clair que nous aurons encore

$$\lim_{k \rightarrow \infty} \tau_0\theta_k\nabla f(x_k)d_k = 0,$$

et que nous devons justifier que si  $\theta_k$  converge vers zéro (pour une certaine sous-suite), alors nécessairement  $\nabla f(x_k)d_k$  converge vers 0. Supposons que  $\nabla f(x_k)d_k$  ne converge pas vers zéro. Alors,  $\theta_k$  converge vers zéro, et donc, à partir d'un certain rang, on ne prend plus le pas unitaire. Donc,

$$f(x_k + 2\theta_k d_k) - f(x) > 2\tau_0\theta_k\nabla f(x_k)d_k. \quad (3.1)$$

En dénotant

$$p_k = \frac{d_k}{\|d_k\|}, \quad \bar{\theta}_k = 2\theta_k\|d_k\|,$$

on a toujours que  $\bar{\theta}_k$  converge vers zéro. Puisque  $p_k$  est normalisé, la suite  $p_k$  possède des points d'accumulation, et on peut extraire une sous-suite telle que  $p_k$  converge vers  $\bar{p}$ . De la relation (3.1), on tire que

$$\frac{f(x_k + \bar{\theta}_k p_k) - f(x_k)}{\bar{\theta}_k} > \tau_0 \nabla f(x_k) p_k, \quad (3.2)$$

et en prenant la limite dans (3.2), on retrouve

$$0 \leq (1 - \tau_0) \nabla f(\bar{x}) \bar{p},$$

et puisque  $\tau_0 < 1$ ,

$$0 \leq \nabla f(\bar{x}) \bar{p},$$

ce qui contredit le fait que les directions  $d_k$  sont suffisamment descendantes, puisque  $\bar{p}$  est la normalisation de la limite d'une sous-suite de directions  $d_k$ .  $\square$

Si l'objectif  $f$  est supposé de classe  $\mathcal{C}^2$ , une preuve plus simple permet de conclure.

**Corollaire 3.3.1** *Soit un algorithme de descente appliqué au problème*

$$\min_{x \in \mathbb{R}^n} f(x), \quad f \in \mathcal{C}^2(\mathbb{R}^n);$$

*supposons qu'à chaque itération la direction utilisée  $d_k$  est une direction suffisamment descendante, et pour lequel le pas utilisé dans cette direction est la première valeur de la suite  $\{1, 1/2, 1/4, \dots, (1/2)^m \dots\}$  qui satisfasse au critère d'Armijo; alors, tous les points d'accumulation de la suite  $\{x_k\}$  engendrée par l'algorithme sont des points stationnaires pour le problème  $\min f(x)$ .*

**Preuve** Imitons la preuve du théorème 3.3.3. Le début est semblable, mais on réexprime le critère d'Armijo en utilisant un développement de Taylor d'ordre deux :  $f(x + \theta d) = f(x) + \theta \nabla f(x) d + \frac{\theta^2}{2} d^t \nabla^2 f(\tilde{x}) d$  pour obtenir

$$\theta_k (1 - \tau_0) \nabla f(x_k) d_k + \frac{1}{2} \theta_k^2 d_k^t \nabla^2 f(\tilde{x}) d_k < 0.$$

Maintenant, selon la définition 3.3.1, on a

$$\nabla f(x_k) d_k < \frac{-\gamma_0}{\gamma_1^2} \|d_k\|^2,$$

et puisque  $f$  est de classe  $\mathcal{C}^2(\mathbb{R}^n)$ , il existe une constante positive  $A$  telle que, dès que  $x_k$  est suffisamment proche d'un point stationnaire  $\bar{x}$ ,

$$\frac{1}{2}d_k^t \nabla^2 f(\tilde{x}) d_k < A \|d_k\|^2.$$

En combinant ces deux dernières inégalités dans la réécriture du critère d'Armijo, on a que

$$\theta_k(1 - \tau_0) \frac{-\gamma_0}{\gamma_1^2} \|d_k\|^2 + \theta_k^2 A \|d_k\|^2 < 0$$

entraîne la condition d'Armijo. Mais  $\theta_k > 0$ , et donc dès que  $\theta_k < \frac{(1-\tau_0)\gamma_0}{A\gamma_1^2}$ , le critère d'Armijo est satisfait, et donc si  $\nabla f(x_k)d_k$  ne converge pas vers zéro,  $\theta_k$  non plus puisqu'il satisfait  $\theta_k > \frac{(1-\tau_0)\gamma_0}{2A\gamma_1^2}$  par construction.  $\square$

### Preuve (du théorème 3.3.1)

Il est clair de la discussion précédente qu'il existe des valeurs de  $\theta$  qui satisfont au critère d'Armijo. Vérifions que parmi celles-ci, au moins une satisfait au critère de Wolfe.

Considérons la plus petite valeur de  $\theta$  pour laquelle le critère d'Armijo est satisfait avec égalité :  $\bar{\theta} = \inf_{\theta} : f(x + \theta d) - f(x) > \tau_0 \theta \nabla f(x)d$ . Si le critère d'Armijo est satisfait pour tout  $\theta > 0$ , c'est que la fonction  $f$  n'est pas bornée inférieurement dans la direction  $d$ .

En  $\bar{\theta}$ , la dérivée directionnelle de  $f$  dans la direction  $d$  est plus grande ou égale à  $\tau_0 \nabla f(x)d$  puisqu'immédiatement après ce point, la fonction devient plus grande que la droite de pente  $\tau_0 \nabla f(x)d$ . Or, cette dérivée directionnelle est justement donnée par l'expression  $\nabla f(x + \theta d)d$ . On a donc

$$\nabla f(x + \theta d)d \geq \tau_0 \nabla f(x)d > \tau_1 \nabla f(x)d$$

puisque  $\tau_1 > \tau_0$ . Comme la dernière inégalité est stricte, l'intérieur de  $\Theta$  n'est pas vide.  $\square$

En fait, il est instructif d'observer ces relations pour une fonction  $h$  quadratique.

**Exercice 3.3.5** [*Armijo avec  $h$  quadratique*] Dans cet exercice, nous allons examiner la fonction  $h(\theta) = \frac{a}{2}\theta^2 + b\theta + c$ . Commençons en a) et b) par se poser la question sur les valeurs intéressantes des coefficients  $a, b$  et  $c$ .

- Si on sait que la fonction  $h$  est en fait la fonction  $h_{x,d}(\theta) = f(x + \theta d)$  pour une direction  $d$  suffisamment descendante, que peut-on dire des coefficients  $a, b$  et  $c$  ?
- Si on sait que la fonction  $h$  est bornée inférieurement, que peut-on dire des coefficients  $a, b$  et  $c$  ?
- Supposons maintenant que la fonction  $h$  possède les caractéristiques de a) et b), i. e. qu'elle provient d'une fonction  $f$ , selon une direction suffisamment descendante, et qu'elle est bornée inférieurement. Justifiez que  $h$  possède un seul minimum local, et ce pour une valeur  $\theta^* > 0$ .

- d) Pour quel ensemble de valeurs de  $\tau_0$  le minimum local de  $h$  satisfait-il au critère d'Armijo ?
- e) Pour quel ensemble de valeurs de  $\tau_1$  le minimum local de  $h$  satisfait-il au critère de Wolfe ?

### Direction de la pente la plus forte

Des conditions de descente suffisante, il devrait maintenant être clair que la direction dite de *la pente la plus forte*,  $d = -\nabla f(x)^t$  serait une direction candidat privilégiée. Pour mieux qualifier la nature optimale de cette direction, revenons aux conditions d'optimalité présentées plus haut. Puisque nous ne cherchons seulement qu'un point stationnaire, tentons de trouver  $d$  qui minimise  $h'_{x,d}(0) = \nabla f(x)d$ . Évidemment, un tel  $d$  serait non-borné, égal à  $-\infty \nabla f(x)$ . Pour se sortir d'une impasse apparente, considérons l'inégalité de Cauchy-Schwartz :

$$-\|d\|\|\nabla f\| \leq \nabla f(x)d \leq \|d\|\|\nabla f\|;$$

Rappelons que cette inégalité découle de l'identité reliant le produit scalaire de deux vecteurs à leurs normes et l'angle  $\theta$  entre eux :  $v^t u = \|v\|\|u\| \cos \theta$ . Alors, on voit bien que les valeurs de  $d$  qui atteignent la borne inférieure des précédentes inégalités sont précisément les multiples positifs de  $d = -\nabla f(x)^t$ .

Malheureusement, cette direction  $d = -\nabla f(x)^t$  conduit à un algorithme qui converge très lentement.

**Exercice 3.3.6** [*Vitesse de convergence de l'algorithme du gradient*] On peut montrer que l'ordre de convergence de l'algorithme de la pente la plus forte est linéaire, et que le taux de convergence est donné par le rapport  $\frac{1-r}{1+r}$  où  $r = \lambda_{min}/\lambda_{max}$  et  $\lambda_{min}$  et  $\lambda_{max}$  sont les valeurs propres extrêmes de la matrice  $\nabla^2 f(x^*)$ . Considérez la fonction quadratique  $f(x, y) = 10x^2 + y^2$ , et vérifiez numériquement que ce taux est atteint si l'on amorce l'algorithme du gradient en  $\begin{pmatrix} 0.1 \\ 1 \end{pmatrix}$ . Remarquons que ce taux peut être arbitrairement proche de un si le rapport des valeurs propres extrêmes est proche de zéro.

Parmi les fonctions proposées au fil des années pour tester des algorithmes de descente, la fonction de Rosenbrock est certainement la plus fameuse. Il s'agit de la fonction  $f(x) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$ , et un point de départ usuel est  $\begin{pmatrix} -1.2 \\ 1 \end{pmatrix}$ . Les codes contemporains



obtiennent la solution  $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  en quelques dizaines d'itérations. Or, la méthode utilisant la direction de la pente la plus forte requiert plusieurs centaines, voire milliers d'itérations. Il est très rare qu'on ne puisse pas faire beaucoup mieux que d'utiliser cette méthode connue sous plusieurs noms dont "méthode du gradient", "descente de gradient", "méthode de la pente la plus forte", "steepest descent" en anglais.

### Direction de Newton

Plutôt que de "minimiser"  $h'_{x,d}(0)$ , minimisons par rapport à  $d$  la fonction  $h'_{x,d}(0) + \frac{1}{2}h''_{x,d}(0)$ . Un examen sommaire des expressions impliquées permet de reformuler cette minimisation par rapport à  $d$  comme

$$\min_d \nabla f(x)d + \frac{1}{2}d^t \nabla^2 f(x)d,$$

dont la solution, en autant que  $\nabla^2 f(x)$  soit une matrice définie positive (et donc inversible), peut s'écrire

$$d_N = -\nabla^2 f(x)^{-1} \nabla f(x)^t.$$

Le fait qu'elle satisfasse aux critères de la définition 3.3.1 est démontré à la section 3.4.

Cette direction  $d_N$  est importante, et correspond à la direction dite de Newton. Pour comprendre cette terminologie de *direction de Newton*, rappelons la méthode de Newton pour la solution de systèmes d'équations non-linéaires  $F(x) = 0$  :

$$x_{k+1} \leftarrow x_k - \nabla F(x_k)^{-1} F(x_k).$$

Pour les problèmes de minimisation, nous retrouvons les identités  $\nabla f \equiv F^t$  et  $\nabla^2 f \equiv \nabla F$ . Selon cette notation, la direction de Newton serait  $x_{k+1} - x_k = -\nabla F(x_k)^{-1} F(x_k)$ .

Nous proposons à la section 3.4.3 une analyse qui illustre que cette direction de Newton conduit, au moins localement, à un algorithme efficace.

**Exercice 3.3.7** [*Méthode de Newton*] Considérez à nouveau (exercice 3.2.4 page 108) la fonction  $f(x) = (x_1 + x_2)^2 + (2(x_1^2 + x_2^2) - 1) - \frac{1}{3}$ . Vérifiez que la méthode de Newton peut être appliquée à partir de  $x^0 = (\sqrt{7/6}, 0)^t$  et que le prochain point, en utilisant  $\theta = 1$  est situé sur la droite reliant les deux points stationnaires.

### Une famille de directions suffisamment descendantes

La direction de Newton peut être vue comme une transformation linéaire de la direction de la plus forte pente. La matrice représentant cette transformation linéaire pour la direction de Newton est l'inverse de la matrice hessienne de  $f$ .

Considérons le cas général d'une direction  $\bar{d} \stackrel{\text{def}}{=} -B\nabla f(x)^t$ . En supposant que  $B$  est une matrice symétrique définie positive, il est facile de vérifier que  $\bar{d}$  satisfait aux conditions de la définition 3.3.1. En effet, on peut écrire

$$\nabla f(x)\bar{d} = -\nabla f(x)B\nabla f(x)^t,$$

et si  $\lambda_{\min} > 0$  dénote la valeur propre la plus petite de  $B$ , on a

$$\nabla f(x)B\nabla f(x)^t \geq \lambda_{\min}\|\nabla f(x)\|^2.$$

Ceci assure que

$$\nabla f(x)\bar{d} \leq -\lambda_{\min}\|\nabla f(x)\|^2.$$

Pour vérifier la seconde inégalité de la définition 3.3.1, nous utilisons plutôt la plus grande valeur propre de  $B$ ,  $\lambda_{\max}$  pour écrire

$$\|\bar{d}\| \leq \lambda_{\max}\|\nabla f(x)\|.$$

Enfin, remarquons que si  $B$  est obtenue par l'inversion d'une matrice symétrique (inversible)  $H$ , qui approche probablement le hessien de  $f$ , alors  $B$  est définie positive si  $H$  l'est aussi. En effet, pour un vecteur propre  $d$  associé à la valeur propre  $\lambda$  de  $H$ , on a  $Hd = \lambda d$ , et donc  $Bd = H^{-1}d = \frac{1}{\lambda}d$  et donc  $d$  est vecteur propre de  $B$  associé à la valeur propre  $\frac{1}{\lambda}$ .

**Exercice 3.3.8** [*Direction de descente*] Complétez l'analyse des directions  $d = -H^{-1}\nabla f(x)^t$  en fournissant des expressions pour  $\gamma_0$  et  $\gamma_1$ .

**Exercice 3.3.9** [*Algorithme de descente*] Considérez la fonction  $f(x) = x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + e^{x_1+x_2}$ .

- a) Écrivez la condition d'optimalité de premier ordre pour cette fonction. Est-ce que cette condition est suffisante pour ce problème? En d'autres termes, est-ce que tous les points stationnaires de cette fonction sont des minima locaux?
- b) Vérifiez que l'origine n'est pas un point stationnaire de  $f$ , et identifiez une direction de descente (suffisamment descendante) pour  $f$  à l'origine.
- c) Appliquez le procédé d'Armijo (la première valeur parmi  $1, \frac{1}{2}, \frac{1}{4}, \dots$ ) à la fonction  $h_d(\theta) = f(\theta d)$  pour la direction que vous avez proposée en b). Utilisez une valeur de  $\tau_0 = 0.1$  dans le critère d'Armijo.

**Exercice 3.3.10** [*Newton—descente*]

- a) Pour minimiser une fonction de  $n$  variables, si  $\nabla^2 f(x)^{-1}$  possède des valeurs propres comprises entre 1 et 100 quel que soit  $x \in \mathbb{R}^n$ , justifiez que la direction de Newton,  $d_N = -\nabla^2 f(x)^{-1} \nabla f(x)^t$  est suffisamment descendante en exhibant des valeurs de  $\gamma_0$  et  $\gamma_1$ .
- b) Si  $\nabla^2 f(x)$  possède des valeurs propres négatives, est-il vrai que  $d_N$  est forcément *non* descendante ? Si oui, justifiez ; sinon, donnez un exemple où  $d_N$  est malgré tout descendante, et donnez un exemple où  $d_N$  n'est pas descendante.

**3.3.3 Recherche de minima locaux faibles**

Si la condition que  $\nabla^2 f(x)$  soit semi-définie positive n'est pas vérifiée, même si  $x$  est un point stationnaire, il n'en est pas un minimum local pour autant. En fait il est essentiel que  $\nabla^2 f(x)$  soit semi-définie positive pour qu'un point stationnaire  $x$  soit un minimum local, mais ce n'est toutefois pas suffisant. Cependant, comme nous en avons discuté à la section 3.2, nous nous contentons de ces conditions.

Dans la suite de cette sous-section, nous développons un schéma algorithmique qui assure que les points d'accumulation des algorithmes, en plus d'être stationnaires, possèdent un hessien  $\nabla^2 f(x)$  semi-défini positif. Ce schéma algorithmique nous permet également d'introduire une technique qui n'utilise pas un pas admissible dans une direction suffisamment descendante, mais plutôt une restriction ensembliste sur la direction, dans laquelle on effectue un pas unitaire. Ce type d'algorithme généralise au cas multi-dimensionnel l'algorithme de région de confiance de la section 2.10.

L'idée sous-jacente aux méthodes de *région de confiance* est que le développement de Taylor d'ordre un ou deux utilisé pour calculer une direction constitue une bonne approximation de  $f$  dans un certain voisinage du point courant. Le développement de Taylor

$$q_x(d) = f(x) + \nabla f(x)d + \frac{1}{2}d^t \nabla^2 f(x)d$$

sera nommé *fonction modèle* de  $f$ , et la direction sera calculée en résolvant un problème de la forme

$$\min_{\|d\| \leq \Delta_k} q_{x_k}(d).$$

Une partie de la motivation vient de la formulation suivante d'une condition nécessaire d'optimalité.

**Proposition 3.3.4** *Soit  $f$  une fonction de classe  $\mathcal{C}^2$  et  $x^*$  un minimum local de  $f$ . Alors,  $d^* = 0$  est un minimum local de la fonction  $q_{x^*}(d) = f(x^*) + \nabla f(x^*)d + \frac{1}{2}d^t \nabla^2 f(x^*)d$ .*

**Preuve** Nous savons que  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*) \geq 0$ . Par conséquent,  $\nabla f(x^*)d + \frac{1}{2}d^t \nabla^2 f(x^*)d \geq 0$  pour tout  $d \neq 0$ .  $\square$

La valeur de  $\Delta_k$  est ajustée à l'aide de la quantité  $r_k$  ainsi définie :

$$\begin{aligned} \mathcal{D}f_{x_k} &\stackrel{\text{def}}{=} f(x_k) - f(x_{k+1}) \\ \mathcal{D}q_{x_k} &\stackrel{\text{def}}{=} f(x_k) - q_{x_k}(d_k) = q_{x_k}(0) - q_{x_k}(d_k) \\ r_k &\stackrel{\text{def}}{=} \frac{\mathcal{D}f_{x_k}}{\mathcal{D}q_{x_k}}. \end{aligned}$$

Ainsi,  $r_k$  représente le rapport entre la réelle diminution de  $f$  en comparaison avec la diminution prévue par la fonction modèle  $q$ . Pour la méthode de Newton, on s'attend à ce que  $r_k$  s'approche de l'unité lorsque  $x_k$  converge vers un minimum local de  $f$ .

Pour formuler un algorithme utilisant ces quantités, remarquons que si  $r_k$  est petit, il est vraisemblable que  $\Delta_k$  soit trop grand, alors que si  $r_k$  s'approche de un, il faudrait augmenter  $\Delta_k$  pour que l'algorithme se comporte exactement comme l'itération de Newton asymptotiquement. Ceci conduit aux règles suivantes, où des choix typiques des paramètres  $0 < \eta_1 < \eta_2 < 1$  et  $0 < \gamma_1 < 1 < \gamma_2$  sont  $\eta_1 = 0.25$ ,  $\eta_2 = 0.75$ ,  $\gamma_1 = 0.25$  et  $\gamma_2 = 2$ .

- si  $r_k < \eta_1$  alors  $\Delta_{k+1} \leftarrow \gamma_1 \Delta_k$  ;
- si  $r_k > \eta_2$  alors
  - si  $\|d_k\| = \Delta_k$  alors  $\Delta_{k+1} \leftarrow \gamma_2 \Delta_k$
  - sinon  $\Delta_{k+1} \leftarrow \Delta_k$  ;
- si  $r_k > \eta_1$  alors  $x_{k+1} \leftarrow x_k + d_k$
- sinon  $x_{k+1} \leftarrow x_k$ .

Cet algorithme possède des propriétés de convergence globales très fortes, au prix de sous-problèmes plus difficiles à résoudre, la minimisation *globale*  $\min_{\|d\| \leq \Delta_k} q_{x_k}(d)$ . Nous présenterons à la section 3.8.2 une implantation simple pour ce calcul.

**Théorème 3.3.5** *Si les points engendrés par l'algorithme demeurent dans un compact  $K$ , avec  $f$  de classe  $\mathcal{C}^2$  dans  $K$ , et  $\nabla^2 f$  uniformément borné dans  $K$ , alors il existe un point d'accumulation  $\bar{x}$  de la suite  $\{x_k\}$ , et tout point d'accumulation satisfait que  $\nabla f(\bar{x}) = 0$  et  $\nabla^2 f(\bar{x})$  est une matrice semi-définie positive.*

**Preuve** Selon que  $\inf \Delta_k = 0$  dans l'algorithme, il existe une sous-suite notée  $\{x_k\}$  qui converge vers  $\bar{x}$  pour laquelle soit

1.  $r_k \geq \eta_1$ , et donc  $\inf \Delta_k > 0$ , ou
2.  $r_k < \eta_1$ , entraînant  $\Delta_{k+1} \rightarrow 0$  et  $\|d_k\| \rightarrow 0$ .

**cas 1**

Newton—région de confiance( $x, \Delta, \epsilon, f$ )

```

{ Donnés :  $x$  ; }
{ un critère d'arrêt,  $\epsilon > 0$  ; }
{ une fonction deux fois différentiable  $f$  ; }
{ une taille initiale  $\Delta$  de la région de confiance. }
{ les paramètres d'ajustement de la région de confiance }
{  $0 < \eta_1 < \eta_2 < 1$  et  $0 < \gamma_1 < 1 < \gamma_2$ . }
répéter
   $d_R \leftarrow \arg \min_{\|d\| \leq \Delta} q_x(d)$ 
  { le calcul approché de  $d_R$  peut par exemple être effectué par l'algorithme 3.5 }
   $r \leftarrow \frac{f(x) - f(x + d_R)}{q(0) - q(d_R)}$ 
  si ( $r < \eta_1$ ) alors  $\Delta \leftarrow \gamma_1 \Delta$ 
  sinon
     $x \leftarrow x + d_R$ 
    si ( $r > \eta_2$ ) alors
      si ( $\|d_R\| = \Delta$ ) alors
         $\Delta \leftarrow \gamma_2 \Delta$ 
  jusqu'à ( $\|\nabla f(x)\| < \epsilon$ )
  Résultat  $\leftarrow x$ 

```

Algorithme 3.2: Newton avec région de confiance.

Ce cas devrait se produire le plus souvent car on espère que  $d_k$  sera localement identique à la direction de Newton, et donc que la borne  $\|d_k\| \leq \Delta_k$  ne sera pas saturée. Puisque  $f(\bar{x})$  est une quantité bornée (les  $x_k$  sont contenus dans un compact), on a que la somme  $\sum \mathcal{D}f_{x_k} < \infty$ , ce qui entraîne que  $\mathcal{D}f_{x_k} \rightarrow 0$ , et donc, puisque  $r_k \geq \eta_1$ ,  $\mathcal{D}q_{x_k} \rightarrow 0$ .

Maintenant, en définissant  $\bar{q}(d) \stackrel{\text{def}}{=} f(\bar{x}) + \nabla f(\bar{x})d + \frac{1}{2}d^t \nabla^2 f(\bar{x})d$ , et  $0 < \bar{\Delta} < \inf \Delta_k$ , considérons  $\bar{d}$  la solution de  $\min_{\|d\| \leq \bar{\Delta}} \bar{q}(d)$ . Dénotons par  $\hat{x} \stackrel{\text{def}}{=} \bar{x} + \bar{d}$ . Observons que si  $k$  est assez grand,  $\|\hat{x} - x_k\| \leq \Delta_k$ , ce qui entraîne que

$$q_{x_k}(\hat{x} - x_k) \geq q_{x_k}(d_k) = f(x_k) - \mathcal{D}q_{x_k}.$$

En passant à la limite dans les dernières relations, on retrouve

$$\bar{q}(\bar{d}) \geq f(\bar{x}) = \bar{q}(0),$$

ce qui nous indique que 0 est solution de  $\min_{\|d\| \leq \bar{\Delta}} \bar{q}(d)$ , et puisque  $0 < \bar{\Delta}$ , ceci entraîne que  $\nabla f(\bar{x}) = 0$  et  $\nabla^2 f(\bar{x})$  est une matrice semi-définie positive, condition nécessaire de la proposition 3.3.4.

### cas 2

Séparons cette portion de la preuve en deux : établissons d'abord une contradiction au fait que  $\nabla f(\bar{x}) \neq 0$ , et ensuite établissons une contradiction au fait que  $\nabla^2 f(\bar{x})$  possède une valeur propre négative.

1. Supposons que  $\nabla f(\bar{x}) \neq 0$ . à partir d'un  $x_k$  quelconque, considérons un déplacement le long de la direction opposée au gradient  $\hat{d}_k = -\nabla f(x_k)^t$ . Alors, en un point le long de cette direction,

$$q_{x_k}(\alpha \hat{d}_k) = f(x_k) - \alpha \|\nabla f(x_k)\|^2 + \frac{\alpha^2}{2} \hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k$$

est minimisé (si  $\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k$  est positif) pour  $\alpha^* = \|\nabla f(x_k)\|^2 / (\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k)$ , et  $q^* = q_{x_k}(\alpha^* \hat{d}_k) = \frac{1}{2} \|\nabla f(x_k)\|^4 / (\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k)$ . Par ailleurs, pour demeurer dans la région de confiance, on doit avoir  $\alpha \leq \Delta_k / \|\hat{d}_k\| = \alpha_{max}$ , qui est également le minimum lorsque  $\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k$  est négatif. Puisque nous sommes dans le cas où  $\Delta_k \rightarrow 0$ ,  $\mathcal{D}q_{x_k} \geq q_{x_k}(0) - q_{x_k}(\alpha_{max} \hat{d}_k)$ . Si  $\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k$  est positif,  $q_{x_k}(\alpha \hat{d}_k)$  est convexe en  $\alpha$ , et on obtient  $\mathcal{D}q_{x_k} \geq (q_{x_k}(0) - q^*) \frac{\Delta_k}{\alpha^*} = \frac{\Delta_k}{2} \|\nabla f(x_k)\|$ . On retrouve cette dernière inégalité directement si  $\hat{d}_k^t \nabla^2 f(x_k) \hat{d}_k$  est négatif.

Un développement de Taylor de  $f$  près de  $x_k$  donne

$$f(x_k + d_k) = q_{x_k}(d_k) + \mathcal{O}(\|d_k\|^3), \quad (**)$$

et donc

$$\mathcal{D}f_{x_k} = \mathcal{D}q_{x_k} + \mathcal{O}(\|d_k\|^3).$$

Or,  $\|d_k\| < \Delta_k$  et on vient de voir que  $\mathcal{D}q_{x_k} \geq \frac{\Delta_k}{2} \|\nabla f(x_k)\| \sim \mathcal{O}(\Delta_k)$ . Par conséquent,  $r_k = 1 + \mathcal{O}(\|d_k\|)$ , contredisant l'hypothèse que  $r_k < \eta_1$ .

2. Supposons maintenant qu'il existe une direction de courbure négative unitaire  $\bar{d}$  telle que

$$\bar{d}^t \nabla^2 f(\bar{x}) \bar{d} = -\delta < 0,$$

et en un point de la sous-suite, effectuons un déplacement de longueur  $\Delta_k$  dans la direction  $\sigma \bar{d}$  avec  $\sigma = \pm 1$  choisi pour que  $\nabla f(x_k) \bar{d} \leq 0$ . Encore une fois, on écrit

$$\mathcal{D}q_{x_k} \geq q_{x_k}(0) - q_{x_k}(\Delta_k \sigma \bar{d}) \geq -\frac{1}{2} \Delta_k^2 \bar{d}^t \nabla^2 f(x_k) \bar{d} = \frac{1}{2} \Delta_k^2 \delta + \mathcal{O}(\Delta_k^3).$$

Ceci implique  $r_k = 1 + \mathcal{O}(\|d_k\|)$ , contredisant encore une fois l'hypothèse que  $r_k < \eta_1$ .

□

La preuve du théorème est importante car elle suggère des critères plus faibles que la minimisation globale de  $q_x(d)$  dans la région  $\|d\| \leq \Delta$ . En particulier, si l'on veut garantir que les points d'accumulation sont stationnaires, il suffit que  $q$  décroisse autant qu'un pas optimal dans la direction de la plus forte pente  $-\nabla f(x)^t$ ; un tel pas est habituellement nommé *point de Cauchy*. Pour la recherche de points stationnaires, il n'est pas nécessaire que le modèle  $q$  soit défini à partir de  $\nabla^2 f$ , et une matrice  $M_k$  peut fort bien être utilisée.

**Corollaire 3.3.2** *Si l'algorithme 3.2 utilise un modèle  $q(d) = f(x) + \nabla f(x)d + \frac{1}{2}d^t M_x d$  avec des matrices  $M_x$  uniformément bornées, et si  $q$  est minimisée approximativement au moins pour obtenir la réduction du point de Cauchy, alors les points d'accumulation sont stationnaires.*

**\*Exercice 3.3.11** [Preuve du corollaire 3.3.2] Adaptez la preuve du théorème 3.3.5; indice : le cas 2 n'a pas besoin de modification, et le cas 1 doit tenir compte du fait que  $d$  n'est plus un minimum global, mais seulement une réduction au moins aussi bonne que le point de Cauchy.

D'un autre côté, pour garantir que les points d'accumulation satisfont aux conditions nécessaires de second ordre, il faut bien utiliser  $\nabla^2 f$ , mais il suffit d'assurer que  $q_x$  décroît au moins autant qu'une direction de courbure négative associée à la valeur propre la plus petite, et que le point de Cauchy.

**Corollaire 3.3.3** *Si l'algorithme 3.2 utilise un modèle  $q(d) = f(x) + \nabla f(x)d + \frac{1}{2}d^t \nabla^2 f(x)d$ , et si  $q$  est minimisée approximativement au moins pour obtenir la réduction du point de Cauchy et une réduction donnée par la direction de courbure la plus négative, alors les points d'accumulation satisfont aux conditions nécessaires de second ordre.*

**\*Exercice 3.3.12** [Preuve du corollaire 3.3.3] Adaptez la preuve du théorème 3.3.5 ; indice : le cas 2 n'a pas besoin de modification, et le cas 1 doit tenir compte du fait que  $d$  n'est plus un minimum global, mais seulement une réduction au moins aussi bonne que le point de Cauchy et la direction de courbure la plus négative ; par le corollaire 3.3.2, les points d'accumulation sont stationnaires.

Si un point d'accumulation satisfait aux conditions suffisantes de second ordre, alors la suite engendrée par l'algorithme converge vers ce point, et ce avec un ordre asymptotique quadratique, comme nous le démontrerons un peu plus loin.

Dans la suite de ce chapitre, nous approfondirons surtout l'étude de méthodes utilisant des directions suffisamment descendantes et des pas admissibles. Une des raisons est que la solution des sous-problèmes du calcul de  $d_k$  dans les méthodes de région de confiance nécessite des techniques d'optimisation avec contraintes d'inégalités, que nous étudierons au chapitre 6. Nous y reviendrons à ce moment.

## 3.4 Itération de Newton modifiée pour l'optimisation

Nous avons vu à la section 3.3.2 que la direction de Newton est une direction suffisamment descendante si  $\nabla^2 f(x)$  est une matrice définie positive. Qu'en est-t-il sinon, ou encore si au fil des itérations, cette matrice se rapproche d'une matrice singulière ? Il en est que la méthode de Newton n'est peut-être plus globalement convergente. Par ailleurs, l'utilisation d'un pas  $\theta$  est susceptible de détruire les propriétés de convergence locale quadratique de l'itération de Newton, dont la preuve utilisait toujours un plein pas  $\theta \equiv 1$ . Nous modifions dans cette section la méthode de Newton pour la rendre globalement convergente, et, en supposant que toute la suite  $\{x_k\}$  converge vers un point stationnaire  $x^*$  pour lequel  $\nabla^2 f(x^*)$  est une matrice définie positive, la modification conserve la propriété de convergence locale quadratique.

### 3.4.1 Modification de la direction de Newton

Pour éviter les problèmes associés à une matrice hessienne  $\nabla^2 f(x_k)$  qui ne serait pas uniformément définie positive, nous allons la remplacer par une approximation  $\bar{H}_k$ , uniformément définie positive, et donc inversible, d'inverse  $\bar{B}_k$ , également définie positive. Conceptuellement, il suffit de prendre

$$\bar{H}_k = \nabla^2 f(x) + \max(-\lambda_{\min} + \epsilon, 0)I$$

où  $\lambda_{\min}$  est la plus petite valeur propre de  $\nabla^2 f(x)$  et  $\epsilon$  est une petite quantité positive. En effet, pour un tel choix,  $\bar{H}_k$  sera égal au hessien de  $f$  si celui-ci est une matrice définie



positive, et se verra ajouter un multiple de l'identité adéquat pour le rendre défini positif sinon.

Si  $x_k$  est proche d'un point stationnaire  $x^*$  pour lequel  $\nabla^2 f(x^*)$  est une matrice définie positive, alors  $\nabla^2 f(x_k)$  est également une matrice définie positive, et donc  $\bar{H}_k = \nabla^2 f(x_k)$ . Autrement, la valeur propre la plus petite de  $\bar{H}_k$  vaut au moins  $\epsilon$ , et donc est uniformément éloignée de zéro, garantissant que la direction satisfait aux conditions de la définition 3.3.1.

Il serait prématuré de présenter ici un algorithme de calcul pour  $\bar{B}_k$ . Nous reportons l'étude de l'implantation de la direction de Newton modifiée à la section qui traite du cas particulier de fonctions objectif quadratiques. En effet, une direction  $\bar{d} = -\bar{B}_k \nabla f(x)^t$  pour une matrice  $\bar{B}_k = \bar{H}_k^{-1}$  définie positive est solution du sous-problème quadratique suivant :

$$\min_d \frac{1}{2} d^t \bar{H}_k d + \nabla f(x) d.$$

**Exercice 3.4.1** [*Propriété de descente de la méthode de Newton*] Trouvez la plus grande boule centrée sur l'origine pour laquelle le hessien de la fonction  $f(x) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + x_1^4$  est une matrice définie positive. Si l'on démarre en un point  $x^0 = (x, x)$ , pour quelles valeurs dans cette boule l'algorithme de Newton (avec  $\theta = 1$ ) converge-t-il ?

### 3.4.2 Analyse du pas $\theta$ pour la direction de Newton

Il apparaît qu'un pas  $\theta = 1$  est admissible pour la direction de Newton lorsque  $x_k$  devient assez proche de  $x^*$ , un point stationnaire pour lequel  $\nabla^2 f$  est défini positif. En fait, ceci découle d'un résultat beaucoup plus général qui stipule que si une direction est de longueur "optimale" pour l'approximation de Taylor d'ordre deux de  $f$ , alors un pas unitaire est localement admissible.

**Théorème 3.4.1** *Soit  $f$  une fonction de classe  $C^3$ ,  $d$  une direction suffisamment descendante, telle que le minimum de la fonction quadratique*

$$h_{x,d}(\theta) \stackrel{\text{def}}{=} f(x) + \theta \nabla f(x) d + \frac{\theta^2}{2} d^t \nabla^2 f(x) d$$

*soit atteint pour  $\theta = 1$ . Alors, le pas  $\theta = 1$  est admissible pour la fonction  $f$  dans la direction  $d$  lorsque  $\|d\|$  est suffisamment petite.*

**Preuve** Puisque  $\theta = 1$  minimise  $h_{x,d}(\theta)$ , alors  $h'_{x,d}(1) = 0$ . Donc,

$$\nabla f(x) d + d^t \nabla^2 f(x) d = 0.$$

Écrivons maintenant la condition d'Armijo :

$$f(x + d) - f(x) - \tau_0 \nabla f(x)d < 0$$

en exprimant  $f(x + d) = f(x) + \nabla f(x)d + \frac{1}{2}d^t \nabla^2 f(x)d + R_3(d)$  :

$$f(x + d) - f(x) - \tau_0 \nabla f(x)d = (1 - \tau_0) \nabla f(x)d + \frac{1}{2}d^t \nabla^2 f(x)d + \mathcal{O}(\|d\|^3) \quad (3.3)$$

$$= \left(\frac{1}{2} - \tau_0\right) \nabla f(x)d + \mathcal{O}(\|d\|^3), \quad (3.4)$$

$$(3.5)$$

la dernière égalité provenant de la condition  $h'_{x,d}(1) = 0$ . Maintenant, selon la définition 3.3.1,  $\nabla f(x)d < -\frac{\gamma_0}{\gamma_1} \|d\|^2$  et donc la condition d'Armijo est impliquée par l'inégalité suivante :

$$\left(\tau_0 - \frac{1}{2}\right) \frac{\gamma_0}{\gamma_1} \|d\|^2 + \mathcal{O}(\|d\|^3) < 0$$

qui est vérifiée lorsque  $\|d\|$  est suffisamment petite.

Quant au critère de Wolfe, il suffit d'écrire que  $\nabla f(x + d) = \nabla f(x) + d^t \nabla^2 f(x) + R_2(d)$  pour réexprimer le critère de Wolfe comme

$$(1 - \tau_1) \nabla f(x)d + d^t \nabla^2 f(x)d + dR_2(d) = -\tau_1 \nabla f(x)d + dR_2(d),$$

qui demeure positif lorsque  $d$  est assez petite car il existe une constante positive  $\lambda$  telle que  $dR_2(d) \leq \lambda \|d\|^3$  lorsque  $\|d\|$  est assez petit alors que  $-\nabla f(x)d > \frac{\gamma_0}{\gamma_1} \|d\|^2$ .  $\square$

**Corollaire 3.4.1** *Si  $\nabla^2 f(x)$  est une matrice définie positive, et si la direction de Newton  $d_N = -\nabla^2 f(x)^{-1} \nabla f(x)^t$  est une direction de descente suffisamment petite (au voisinage d'un minimum local de  $f$ ), alors le pas  $\theta = 1$  est admissible pour  $f$  dans la direction  $d_N$ .*

### 3.4.3 Convergence de la méthode de Newton modifiée

Nous avons maintenant tous les ingrédients pour démontrer la convergence globale ainsi que la convergence locale quadratique de la méthode de Newton modifiée pour l'optimisation. En fait, nous allons démontrer la convergence d'une version approchée de la méthode de Newton.

Rappelons que l'itération de Newton modifiée s'écrit :

$$\begin{aligned} x_{k+1} &\leftarrow x_k - \theta_k d_k; \\ d_k &\leftarrow -\bar{H}_k^{-1} \nabla f(x_k)^t. \end{aligned}$$

On voit donc que  $d_k$  satisfait au système

$$d_k^t \bar{H}_k + \nabla f(x_k) = r_k,$$

où, dans la méthode classique,  $r_k \equiv 0$ . Cette dernière formulation constitue en fait un sous-problème linéaire, dont la solution  $d_k$  n'est qu'approximative, à  $r_k$  près, et donc exacte pour la méthode de Newton usuelle, où  $r_k \equiv 0$ . L'analyse à venir n'utilisera pas cette propriété que  $r_k \equiv 0$ , et nous pourrons déduire les propriétés requises par la suite  $r_k$  pour qu'une méthode se comporte "bien".

**Théorème 3.4.2** *Soit une fonction  $f$  de classe  $\mathcal{C}^3$ . Les points d'accumulation de la méthode de Newton modifiée pour minimiser la fonction  $f$  sont tous des points stationnaires de la fonction  $f$ . De plus, si un des points d'accumulation satisfait à la condition suffisante de second ordre, alors toute la suite  $\{x_k\}$  y converge, le pas  $\theta = 1$  est éventuellement toujours admissible et la vitesse de convergence asymptotique est quadratique*

**Preuve** Puisque la méthode de Newton modifiée assure que les directions sont suffisamment descendantes, et que les pas utilisés sont admissibles, tous les points d'accumulation sont stationnaires.

Supposons maintenant (nous justifierons plus loin cette supposition) que toute la suite  $\{x_k\}$  converge vers un point limite  $x^*$  qui satisfait aux conditions suffisantes de second ordre. Alors,  $\nabla^2 f(x^*)$  est une matrice définie positive et donc, asymptotiquement,  $\bar{B}_k \equiv \nabla^2 f(x_k)$ . Par conséquent, les hypothèses du théorème 3.3.2 sont satisfaites, et le pas  $\theta = 1$  est admissible asymptotiquement comme le montre le théorème 3.4.1. Enfin, complétons la preuve de convergence quadratique.

D'abord, exprimons  $\nabla f(x^*)$  en utilisant une série de Taylor autour du point  $x_k$  en notant  $\epsilon_k \stackrel{\text{def}}{=} (x_k - x^*)$  :

$$\begin{aligned} 0 = \nabla f(x^*) &= \nabla f(x_k) - \epsilon_k^t \nabla^2 f(x_k) + \mathcal{O}(\|\epsilon_k\|^2) \\ &= (\nabla f(x_k) + d_k^t \nabla^2 f(x_k)) - (d_k + \epsilon_k)^t \nabla^2 f(x_k) + \mathcal{O}(\|\epsilon_k\|^2) \\ &= r_k - (\epsilon_{k+1})^t \nabla^2 f(x_k) + \mathcal{O}(\|\epsilon_k\|^2). \end{aligned}$$

De cette dernière équation, on tire que  $\epsilon_{k+1}$  est solution du système linéaire

$$(\epsilon_{k+1})^t \nabla^2 f(x_k) = -r_k + \mathcal{O}(\|\epsilon_k\|^2).$$

Utilisons maintenant les faits que

- $\nabla^2 f(x^*)$  est une matrice inversible, et

- $\nabla^2 f$  est une fonction Lipschitz, c'est-à-dire que  $\|\nabla^2 f(x_k) - \nabla^2 f(x^*)\| \leq L \cdot \|x_k - x^*\|$ , pour conclure que

$$\epsilon_{k+1} \sim \mathcal{O}(\|r_k\| + \|\epsilon_k\|^2) \sim \mathcal{O}(\max(\|r_k\|, \|\epsilon_k\|^2)).$$

On voit donc que si  $\|r_k\| \sim \mathcal{O}(\|\epsilon_k\|^2)$ , la convergence quadratique locale de l'itération de Newton modifiée est assurée. En particulier, pour le cas  $r_k \equiv 0$  que nous traitons maintenant, la convergence quadratique locale de l'itération de Newton modifiée est assurée.

Finalement, si un des points d'accumulation, disons  $\hat{x}$  satisfait aux conditions suffisantes de second ordre, alors, il existe une sous-suite de  $\{x_k\}$ , notée  $\{x_{k_i}\}$  telle que, pour des valeurs assez grande de  $k_i$ , les raisonnements locaux précédents s'appliquent au point  $x_{k_i}$ . Par conséquent, à partir d'un tel  $k_i$ , le reste de la suite converge vers  $\hat{x}$ .  $\square$

Comme nous l'avons vu dans la preuve, il n'est pas plus difficile de démontrer la convergence quadratique de l'itération de Newton si les sous-problèmes linéaires ne sont résolus qu'approximativement. Cependant, l'admissibilité du pas unitaire n'est plus assurée par le théorème 3.4.1. Si l'on choisit  $\|r_k\| \sim \|\epsilon_k\|^2$ , on conserve la propriété car la preuve du théorème 3.4.1 s'adapte facilement.

**\*Exercice 3.4.2** [Admissibilité du pas unitaire] Justifiez les affirmations suivantes. Si  $\nabla^2 f(x^*)$  est une matrice définie positive, et si  $x_k$  est proche de  $x^*$ , alors

- $\|\epsilon_k\| \sim \|\nabla f(x_k)\|$ ; il s'agit de montrer que  $\|\epsilon_k\| \sim \mathcal{O}(\|\nabla f(x_k)\|)$  et  $\|\nabla f(x_k)\| \sim \mathcal{O}(\|\epsilon_k\|)$
- si  $\|r_k\| \sim \mathcal{O}(\|\epsilon_k\|^2)$ , alors un pas unitaire demeure localement admissible. Modifiez légèrement la preuve du théorème 3.4.1 pour y incorporer  $r_k$ , puis constater que le même raisonnement demeure valable.
- sous quelles conditions sur  $\|r_k\|$  l'adaptation de la preuve du théorème 3.4.1 en b) demeure-t-elle valable ?

### 3.4.4 Convergence de la méthode de Newton — régions de confiance

**Théorème 3.4.3** *Si un point d'accumulation de l'algorithme 3.2 satisfait aux conditions suffisantes de second ordre, alors toute la suite engendrée y converge, et  $r_k \rightarrow 1$ , la suite  $\Delta_k$  demeure éloignée de 0 et l'ordre de convergence est superlinéaire.*

**Preuve** Pour fin de contradiction, supposons qu'une sous-suite se produit avec le cas 2 du théorème 3.3.5. En un point  $x_k$ , considérons un déplacement dans la direction de Newton, bien définie au voisinage du point d'accumulation satisfaisant aux conditions suffisantes d'optimalité. Si la direction de Newton  $d_N = (\nabla^2 f(x_k)^{-1})\nabla f(x_k)^t \leq \Delta_k$ , alors  $d_R = d_N$  et  $\mathcal{D}q_{x_k} = \frac{1}{2}d_R^t \nabla^2 f(x_k)d_R \geq \frac{1}{2}\lambda_k \|d_R\|^2$ , avec  $\lambda_k$  la plus petite valeur propre de  $\nabla^2 f(x_k)$ . Si au contraire la direction de Newton  $d_N > \Delta_k$ , alors l'optimalité de  $q$  nous permet d'établir

que  $\mathcal{D}q_{x_k} \geq \frac{1}{2}\lambda_k\|d_R\|^2$  également. Comme  $\mathcal{D}f_{x_k} = \mathcal{D}q_{x_k} + \mathcal{O}(\|d_R\|^3)$ , on obtient que  $r_k \rightarrow 1$ , contredisant que la sous-suite provient du cas 2.

Donc,  $\Delta_k$  est forcément éloigné de 0, et comme les directions de Newton proche du point d'accumulation satisfaisant aux conditions de second ordre s'approchent de 0, l'algorithme 3.2 se ramène à une itération de Newton pure, d'où le résultat de convergence superlinéaire.  $\square$

### 3.5 Méthodes quasi-Newton

Il est parfois difficile, ou pas pratique de manipuler les matrices hessiennes des fonctions objectifs à minimiser. Au chapitre 2, nous avons vu la technique de la sécante permettant d'atteindre un ordre de convergence superlinéaire (l'ordre dans ce cas est donné par le nombre d'or  $\sim 1.618$ ). La pente  $m_k$  de la sécante à la fonction  $g = f'$  satisfaisait alors la relation  $m_k(x_k - x_{k-1}) = g(x_k) - g(x_{k-1})$ , ou encore, dénotant  $y_k = g(x_k) - g(x_{k-1})$  et  $s_k = x_k - x_{k-1}$ ,  $m_k s_k = y_k$ . L'itération de sécante prenait alors la forme  $x_{k+1} = x_k - g(x_k)/m_k$ .

En plusieurs dimensions, l'information relative à la pente du vecteur  $g = \nabla f$  se représente comme une matrice que nous notons  $H$ . Si on conserve les notations  $y_k = g(x_k) - g(x_{k-1})$  et  $s_k = x_k - x_{k-1}$ , l'équation définissant la pente devient

$$s_k^t H_k = y_k. \quad (3.6)$$

Remarquons que  $H_k$  n'est pas uniquement déterminée par cette équation, nommée *équation quasi-Newton*. On se propose d'utiliser la matrice  $H_k$  en remplacement de la matrice  $\nabla^2 f(x_k)$  dans une méthode itérative de la forme  $x_{k+1} = x_k - \theta_k H_k^{-1} \nabla f(x_k)^t$ .

Puisque l'équation quasi-Newton ne détermine pas complètement  $H_k$ , plusieurs formules ont été proposées. En fait, il est raisonnable de *construire* une matrice que l'on espère s'approcher de  $\nabla^2 f$  au fil des itérations, donc la matrice  $H_k$  sera en fait obtenue à partir d'une mise-à-jour de la matrice  $H_{k-1}$  en tenant compte de l'équation (3.6) qu'elle doit satisfaire. Broyden a introduit une première forme de mise-à-jour en raisonnant que  $H_k z = H_{k-1} z$  si  $z^t s_k = 0$ , que l'effet de la mise-à-jour ne devrait pas avoir d'influence sur le complément orthogonal de la direction  $s_k$ . Ce raisonnement conduit à la formule de mise-à-jour suivante :

$$H_k = H_{k-1} + \frac{s_k(y_k - s_k^t H_{k-1})}{s_k^t s_k}. \quad (3.7)$$

Remarquons dès maintenant que  $B_k = H_k^{-1}$  obéit à une mise-à-jour similaire :

$$B_k = B_{k-1} + \frac{B_{k-1}^t s_k (s_k^t - y_k B_{k-1})}{y_k B_{k-1} s_k}. \quad (3.8)$$

On obtient la formule pour  $B_k$  à l'aide de la formule de Sherman et Morrison. Ces formules ne conduisent pas à des matrices  $B_k$  ou  $H_k$  symétriques, et donc, bien qu'utiles pour des problèmes d'équations non-linéaires ( $F(x) = 0$ ), on leur préfère des formules qui assurent que

la suite des matrices  $H_k$  demeurent symétriques. La formule suivante réalise cette contrainte supplémentaire :

$$H_k = H_{k-1} + \frac{(y_k - s_k H_{k-1})^t (y_k - s_k H_{k-1})}{(y_k - s_k H_{k-1}) s_k}. \quad (3.9)$$

Cette formulation, connue sous le nom de *symétrique de rang un*, échoue lorsque  $(y_k - H_{k-1} s_k)^t s_k = 0$ . Pour éviter cette possibilité, on peut développer une formule de rang deux.

À ce point, introduisons une autre contrainte que devraient respecter les matrices  $H_k$  : pour engendrer des directions de descente, les matrices devraient être définies positives. La formule la plus répandue est étiquetée BFGS, des initiales de chercheurs Broyden, Fletcher, Goldfarb, Shanno qui l'ont proposé simultanément et indépendamment les uns des autres :

$$H_k = H_{k-1} + \frac{y_k^t y_k}{y_k s_k} - \frac{H_{k-1} s_k s_k^t H_{k-1}}{s_k^t H_{k-1} s_k}. \quad (3.10)$$

La formule pour  $B_k = H_k^{-1}$  s'écrit :

$$B_k = \left( I - \frac{s_k y_k}{y_k s_k} \right) B_{k-1} \left( I - \frac{s_k y_k}{y_k s_k} \right) + \frac{s_k s_k^t}{y_k s_k}. \quad (3.11)$$

Remarquons qu'une mise-à-jour, bien que de rang deux, n'implique qu'une seule paire  $y_k, s_k$ . La mise-à-jour est de rang deux puisque  $H_k$  ou  $B_k$  se voient ajouter deux matrices de rang un chacune. En développant les détails, il est concevable d'exprimer  $H_k$  à partir de  $H_0$  et d'une liste de paires  $y_i, s_i$  pour  $i$  allant de un à  $k$ . Nous verrons plus tard que l'algorithme L-BFGS exploite subtilement cette possibilité.

### 3.5.1 Convergence globale

Les propriétés de convergence globale des méthodes quasi-Newtoniennes ne sont pas encore bien comprises. Ces méthodes ont certainement été utilisées avec grand succès sur une grande variété d'applications, mais une preuve de convergence globale semble toujours nous échapper.

Une modification mineure, une mise-à-jour prudente, permet cependant de justifier que les matrices  $B_k$  demeurent *uniformément* définies positives : Li et Fukushima[26] montrent en effet qu'il suffit de conserver  $B_k = B_{k-1}$  si  $y_k s_k < \epsilon \|g_k\|^\alpha \|s_k\|^2$ , pour des constantes positives  $\epsilon$  et  $\alpha$ .

### 3.5.2 Vitesse de convergence

On peut montrer que si un point d'accumulation de la suite  $x_k$  satisfait aux conditions suffisantes, alors la méthode BFGS y converge superlinéairement. On a également l'estimé  $\|x_{k+n} - x^*\| \sim \mathcal{O}(\|x_k - x^*\|^2)$ .

## 3.6 Fonctions objectif quadratiques

Examinons maintenant deux algorithmes, une méthode dite directe, et une autre dite itérative pour solutionner un problème de minimiser une fonction quadratique  $q(x) \stackrel{\text{def}}{=} \frac{1}{2}x^t Qx + cx$  sans contrainte;  $c$  est un vecteur ligne. Il est clair que la matrice  $Q$ , qui est symétrique, doit être au moins semi-définie positive, sans quoi la fonction  $q$  n'est pas bornée inférieurement. En fait, si la matrice  $Q$  n'est pas définie positive, il est également possible que  $q$  ne soit pas bornée inférieurement. Par exemple, considérons la fonction  $q(x, y) = x^2$ . Pour cette fonction, la matrice  $Q = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$  est semi-définie positive, et l'axe des  $y$  en entier est solution du problème. Cependant, la fonction  $q(x, y) = x^2 + y$  n'est pas bornée inférieurement, bien qu'elle partage le même hessien que la fonction  $q(x, y) = x^2$ .

Supposons donc que la matrice  $Q$  qui définit la forme quadratique  $q$  est une matrice définie positive. Alors, la solution au problème  $\min_x q(x)$  peut s'écrire  $x^* = -Q^{-1}c^t$ . Nous présentons ci-après deux techniques pour calculer  $x^*$ . La première est une technique d'algèbre linéaire numérique, une méthode directe de solution d'un système linéaire symétrique défini positif alors que la seconde est un algorithme d'optimisation itératif spécialisé aux formes quadratiques.

### 3.6.1 Décomposition de Cholesky

Plaçons-nous dans le contexte de résoudre un système d'équations linéaires  $Qx + c^t = 0$ , où  $Q$  est une matrice définie positive. Il est connu que toute matrice définie positive peut s'exprimer comme  $Q = LL^t$ , où  $L$  est une matrice triangulaire inférieure (de l'anglais, *Lower triangular*). Cette expression de  $Q$  est connue sous le nom de décomposition de Cholesky de  $Q$ . Une fois  $Q$  ainsi exprimée, la solution du système linéaire  $Qx + c^t = 0$  est facile, puisqu'elle se réduit à deux résolutions de systèmes triangulaires, que l'on peut obtenir par les méthodes de substitution ascendantes ou descendantes :

$$\begin{aligned} Lv &= -c^t \\ L^t x &= v. \end{aligned}$$

Il est avantageux d'utiliser une formulation de la forme  $Q = LDL^t$ , où  $D$  est une matrice diagonale car cette forme évite d'extraire les racines carrées au sein de l'algorithme. Notons que si  $Q$  est une matrice définie positive, les racines carrées de l'algorithme sont toutes bien définies.

### Formulations récursives

On peut calculer la décomposition de Cholesky en analysant l'équation  $Q = LL^t$  ; débutons par une expression récursive :

$$Q = \left( \begin{array}{c|c} Q_{11} & Q_{1\cdot} \\ \hline Q_{\cdot 1} & \hat{Q} \end{array} \right) = \left( \begin{array}{c|c} L_{11} & 0 \\ \hline L_{\cdot 1} & \hat{L} \end{array} \right) \left( \begin{array}{c|c} L_{11} & L_{\cdot 1}^t \\ \hline 0 & \hat{L}^t \end{array} \right) = \left( \begin{array}{c|c} L_{11}^2 & L_{11}L_{\cdot 1}^t \\ \hline L_{\cdot 1}L_{11} & L_{\cdot 1}L_{\cdot 1}^t + \hat{L}\hat{L}^t \end{array} \right).$$

La notation d'indices avec un point (".") est utilisée pour désigner une ligne ou une colonne d'une matrice. On déduit que (la division d'un vecteur par un scalaire étant interprétée naturellement)

- $L_{11} = \sqrt{Q_{11}}$  ;
- $L_{\cdot 1} = \frac{Q_{\cdot 1}}{L_{11}}$  ;
- $Q_2 = \hat{Q} - L_{\cdot 1}L_{\cdot 1}^t$  .

On applique maintenant le même processus, les mêmes formules à la nouvelle sous-matrice  $n - 1 \times n - 1$   $Q_2$ , et ainsi de suite récursivement.

Si on adopte plutôt la formulation  $Q = LDL^t$ ,

$$Q = \left( \begin{array}{c|c} Q_{11} & Q_{1\cdot} \\ \hline Q_{\cdot 1} & \hat{Q} \end{array} \right) = \left( \begin{array}{c|c} L_{11} & 0 \\ \hline L_{\cdot 1} & \hat{L} \end{array} \right) \left( \begin{array}{c|c} d_1 & 0 \\ \hline 0 & \hat{D}_2 \end{array} \right) \left( \begin{array}{c|c} L_{11} & L_{\cdot 1}^t \\ \hline 0 & \hat{L}^t \end{array} \right) = \left( \begin{array}{c|c} d_1 & d_1L_{\cdot 1}^t \\ \hline L_{\cdot 1}d_1 & L_{\cdot 1}L_{\cdot 1}^td_1 + \hat{L}\hat{D}_2\hat{L}^t \end{array} \right).$$

Dans ce cas, on retrouve les formules suivantes (notons que les matrices, dénotées  $L$  dans les deux algorithmes, diffèrent entre elles!) :

- $d_1 = Q_{11}$  ;
- $L_{\cdot 1} = \frac{Q_{\cdot 1}}{d_1}$  ;
- $Q_2 = \hat{Q} - L_{\cdot 1}L_{\cdot 1}^td_1$  .

Cette dernière formulation est préférable, car elle évite d'extraire la racine carrée de  $Q_{11}$ . Pour une matrice  $Q$  définie positive, les éléments  $D_{ii}$  sont tous positifs.

### Formulations séquentielles

On peut également exprimer séquentiellement les calculs, ligne par ligne ou colonne par colonne, et obtenir :

- $L_{11} = \sqrt{Q_{11}}$  ;
- $L_{21} = \frac{Q_{21}}{L_{11}}$  ;
- $L_{22} = \sqrt{Q_{22} - L_{21}^2}$  ;
- $\vdots$
- $L_{ii} = \sqrt{Q_{ii} - \sum_{k=1}^{i-1} L_{ik}^2}$  ;
- $L_{ij} = \frac{Q_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{jk}}{L_{jj}}$ ,  $j < i$ ,



que l'on peut résumer dans ce cas d'algorithme ligne par ligne par

$$L_{ij} = \frac{Q_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{jk}}{L_{jj}}, i = 1 : n, j = 1 : i.$$

La formulation colonne par colonne s'écrit alors :

$$L_{ji} = \frac{Q_{ji} - \sum_{k=1}^{i-1} L_{ik}L_{jk}}{L_{ii}}, i = 1 : n, j = i : n.$$

La complexité de calculs de cette décomposition est  $\mathcal{O}(n^3)$ .

**\*Exercice 3.6.1** [*Formulation sans racine carrée*] Développez des formules itératives ligne par ligne et colonne par colonne pour calculer  $Q = LDL^t$  pour une matrice  $Q$  définie positive ; dans ce cas,  $D$  est une matrice diagonale positive.

### 3.6.2 Algorithme du gradient conjugué

Toujours pour trouver la racine d'équations linéaires issues de matrices  $Q$  définies positives, construisons l'algorithme du gradient conjugué à partir d'observations élémentaires. En fait, il est plus commode pour cette interprétation de considérer que nous tentons de minimiser la fonction  $q(x) = \frac{1}{2}x^tQx + cx$ .

Pour un point  $x$  arbitraire, exprimons le minimum de  $q$  le long d'un axe de direction  $v$  originant en  $x$  :

$$\min_{\theta} q(x + \theta v) = \min_{\theta} \frac{1}{2}(x + \theta v)^tQ(x + \theta v) + c(x + \theta v).$$

En utilisant les conditions d'optimalité de la section 3.2, on retrouve que

$$((x + \theta v)^tQ + c)v = 0,$$

d'où l'on tire que les points minima  $y = x + \theta v$  satisfont tous la relation  $(y^tQ + c)v = 0$ , ou encore  $y^t(Qv) = -cv$ . Par conséquent, en notant  $a = (Qv)$  et  $b = -cv$ , on voit que les minima  $y$  satisfont à la relation  $y^ta = b$ , et donc appartiennent tous au même hyperplan, qui ne dépend pas de  $x$  ! De plus, pour se déplacer d'un point à un autre sur cet hyperplan, il suffit d'utiliser un déplacement  $d$  satisfaisant  $d^ta = 0$ , ou, remplaçant  $a$  par sa valeur,  $d^tQv = 0$ . Deux vecteurs  $v$  et  $d$  qui satisfont à une telle relation sont nommés *conjugués* par rapport à la matrice  $Q$ . Ces observations formalisent l'intuition fournie par l'exemple du tableau 3.1 où nous observons que les directions des itérations paires étaient parallèles, et utilisaient les mêmes pas  $\theta^*$ .

D'après les précédents développements, si l'on pouvait restreindre un algorithme à n'utiliser que des directions conjuguées après sa première étape, il se comporterait comme un algorithme en  $n - 1$  variables. Si on pouvait restreindre l'algorithme restreint à n'utiliser que des directions conjuguées aux deux premières directions de l'algorithme, il se comporterait comme un algorithme en  $n - 2$  variables, et ainsi de suite. Les algorithmes de directions conjuguées constituent justement une réalisation de ce "rêve".

### Algorithme des directions conjuguées

Supposons que nous ayons à notre disposition  $n$  vecteurs linéairement indépendants  $v_1, \dots, v_n$ . Il est possible les convertir en  $n$  vecteurs  $u_1, \dots, u_n$  mutuellement conjugués par rapport à une matrice  $Q$  symétrique et définie positive. En effet, il suffit d'appliquer le procédé d'orthogonalisation de Gram-Schmidt aux vecteurs  $v$ . Une fois ces vecteurs  $u$  en main, il suffit d'appliquer l'algorithme

$$\begin{aligned} x_0 &\leftarrow \text{donné} \\ x_i &\leftarrow x_{i-1} + \left( \arg \min_{\theta} (q(x_{i-1} + \theta u_i)) \right) u_i \end{aligned}$$

pour identifier  $x_n$ , la solution au problème  $\min q(x)$ . En effet, ce procédé séquentiel assure qu'à une étape donnée, la suite des itérations sont toutes dans l'intersection des hyperplans de directions conjuguées aux itérations précédentes.

### Algorithme du gradient conjugué

Lorsque les vecteurs  $v$  utilisés pour produire les vecteurs mutuellement conjugués  $u$  proviennent de gradients de la fonction  $q$ , forme quadratique associée à la matrice  $Q$ , justement la matrice par rapport à laquelle les vecteurs  $u$  sont conjugués, alors le procédé de Gram-Schmidt se réduit à la récurrence de deux termes suivante :

$$d_k = -\nabla q(x_k)^t + \beta_{k-1} d_{k-1},$$

où

$$\beta_k \stackrel{\text{def}}{=} \frac{\nabla q(x_{k+1}) Q d_k}{d_k^t Q d_k}.$$

**Exercice 3.6.2** [Rappels de Gram-Schmidt] Quelles sont les formules utilisées dans le procédé d'orthogonalisation de Gram-Schmidt ?

**\*Exercice 3.6.3** [*Simplification de Gramm-Schmidt*] Vérifiez que le procédé de Gramm-Schmidt se simplifie bien aux formules du gradient conjugué lorsque les directions linéairement indépendantes utilisées pour produire les directions conjuguées sont les gradients des points  $x_k$ .

L'algorithme du gradient conjugué est donc exprimé comme suit :

$$\begin{aligned}d_k &= -\nabla q(x_k)^t + \beta_{k-1}d_{k-1} \\ \theta_k &= \frac{-\nabla q(x_k)d_k}{d_k^t Q d_k} \\ x_{k+1} &= x_k + \theta_k d_k \\ \beta_k &= \frac{\nabla q(x_{k+1})Q d_k}{d_k^t Q d_k}.\end{aligned}$$

On retrouve la formule utilisée dans l'exemple du tableau 3.1 pour le calcul du pas  $\theta^*$ .

**Exercice 3.6.4** [*Application de l'algorithme du gradient conjugué*] Appliquez l'algorithme du gradient conjugué à la fonction  $q(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$  à partir du point  $x_0 = (10, 5)^t$ ; comparez vos résultats avec l'exemple du tableau 3.1.

**\*Exercice 3.6.5** [*Formules équivalentes du gradient conjugué*] Démontrez que les formules suivantes constituent bel et bien une implantation de l'algorithme du gradient conjugué :

$$\begin{aligned}d_k &= -\nabla q(x_k)^t + \beta_{k-1}d_{k-1} \\ \theta_k &= \frac{\|\nabla q(x_k)\|^2}{d_k^t Q d_k} \\ x_{k+1} &= x_k + \theta_k d_k \\ \beta_k &= \frac{\|\nabla q(x_{k+1})\|^2}{\|\nabla q(x_k)\|^2}.\end{aligned}$$

**Exercice 3.6.6** [Programme quadratique] Suite de l'exercice 3.2.7, page 109.

- a) Calculez les valeurs de  $a_0$  et  $a_3$  du modèle en avec l'algorithme du gradient conjugué.
- b) Calculez les valeurs de  $a_0$  et  $a_3$  du modèle en à l'aide de la décomposition de Cholesky.

## 3.7 Problèmes de moindres carrés

Les moindres carrés constituent une classe de problèmes d'optimisation tellement fréquente et utilisée qu'il est justifié de leur consacrer une section. Le contexte est celui d'un système d'équations  $F(x) = 0$  où  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Si  $m > n$ , en général, il n'y a pas de solution et en minimisant  $f(x) = \|F(x)\|^2$ , on cherche à satisfaire au mieux nos équations. Au contraire, si  $m < n$ , il y a une infinité de solutions, et on choisit celle de norme minimale en minimisant encore une fois  $\|F(x)\|^2$ . Même dans le cas où  $m = n$ , on peut minimiser la norme de  $F$  pour espérer obtenir la propriété de convergence globale.

On choisit de minimiser la norme au carré pour s'assurer que la fonction à minimiser est différentiable, ce qui n'est pas le cas pour une norme à l'origine.

**Exercice 3.7.1** [Normes] Montrez qu'il est équivalent de minimiser  $\|F(x)\|$  et  $\|F(x)\|^2$ .

**Exercice 3.7.2** [Somme de normes] On sait que la lumière emprunte toujours le chemin le plus court, ce qui explique que lors d'une réflexion miroir, l'angle de réflexion égale l'angle d'incidence. Cet exercice obtient le résultat dans une situation très simplifiée. On se place en dimension deux, et considère deux points  $P = (P_x, P_y)^t$  et  $Q = (Q_x, Q_y)^t$  avec  $P_y > 0$  et  $Q_y > 0$ . On cherche à identifier le point  $X = (x, 0)^t$  qui minimise la somme des distances de  $P$  à  $X$  et de  $X$  à  $Q$ . On cherche donc à résoudre  $\min_{x \in \mathbb{R}} f(x) = \|P - X\| + \|Q - X\|$  où seul  $X$  dépend de  $x$ . (voir figure 3.6).

- a) En utilisant les conditions d'optimalité, montrez que pour la solution optimale, l'angle entre le segment  $\overline{PX}$  et l'axe des  $x$  (thetaP) est égal à l'angle entre le segment  $\overline{QX}$  et l'axe des  $x$  (thetaQ). La fonction norme n'est pas différentiable à l'origine, mais puisque  $P_y > 0$ ,  $Q_y > 0$  et  $X_y = 0$ , la solution recherchée aura des normes strictement positives, donc la fonction  $f$  est différentiable aux points d'intérêts.
- b) Vérifiez qu'il n'est pas équivalent de minimiser  $\|P - X\|^2 + \|Q - X\|^2$ .

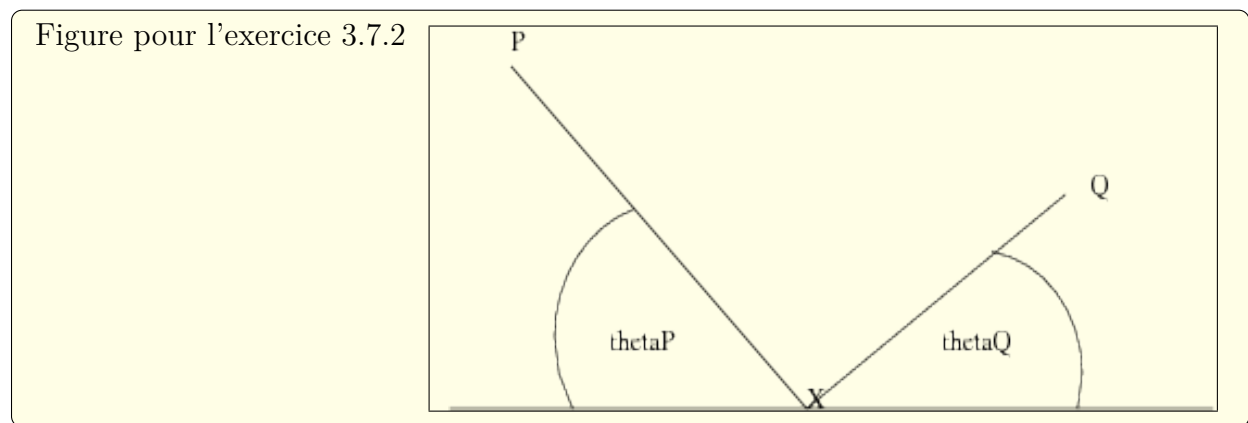


FIGURE 3.6 – Lumière

### 3.7.1 Méthodes de descente pour les systèmes d'équations

Sans hypothèses additionnelles, il est impossible de résoudre un problème de trouver la racine de  $F$ ,  $F(x) = 0$  par un algorithme globalement convergent utilisant une méthode de descente. Cependant, sous certaines hypothèses ( $F$  est une application monotone), on peut appliquer un algorithme de minimisation au problème  $\min \|F(x)\|^2$ , qui sera alors globalement convergent. Cependant, dès que  $\nabla F$  n'est pas inversible en un  $x$  qui n'est pas une racine de  $F$ , la fonction  $\|F(x)\|^2$  peut posséder un minimum local en  $x$  si  $F(x) \in \ker(\nabla F(x))$ , et donc un algorithme de descente échoue à résoudre le problème.

**Exercice 3.7.3** [Fausses racines] Vérifiez à l'aide d'un exemple de fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  qu'il est possible que  $f^2$  possède un minimum local en  $x$  tel que  $f(x) \neq 0$ .

### 3.7.2 Méthodes spécialisées aux moindres carrés

Nous avons étudié les expressions des dérivées de la fonction  $f(x) = \|F(x)\|^2$  à l'exercice A.1.3.

$$\nabla f(x) = F(x)^t \nabla F(x) = \sum_{i=1}^m F_i(x) \nabla F_i(x).$$

$$\nabla^2 f(x) = \nabla F(x)^t \nabla F(x) + \sum_{i=1}^m F_i(x) \nabla^2 F_i(x).$$

La matrice Hessienne comporte donc deux termes, le premier ne faisant intervenir que la matrice Jacobienne de  $F$ . Si on utilise les deux termes, on peut appliquer une variante de

la méthode de Newton à la fonction  $f$ . Cependant, le second terme comporte une somme de matrices Hessiennes des composantes de  $F$ , possiblement fastidieux à manipuler. C'est pourquoi on peut envisager une approximation de la méthode de Newton qui n'utilisera que le premier terme  $\nabla F(x)^t \nabla F(x)$ . C'est la méthode nommée Gauss-Newton. Comme la matrice  $\nabla F(x)^t \nabla F(x)$  n'est inversible que si  $m \geq n$ , et encore, on régularise souvent les équations en leur ajoutant un multiple de l'identité, idée connue sous le nom de méthode Levenberg-Marquardt.

### Gauss-Newton

Donc, l'algorithme de Gauss-Newton consiste en l'itération

$$x^+ = x - (\nabla F(x)^t \nabla F(x))^{-1} \nabla F(x)^t F(x),$$

la condition bien sûr que  $(\nabla F(x)^t \nabla F(x))$  soit inversible.

### Levenberg-Marquardt

L'algorithme de Levenberg-Marquardt consiste en l'itération

$$x^+ = x - (\nabla F(x)^t \nabla F(x) + \lambda I)^{-1} \nabla F(x)^t F(x),$$

où  $\lambda > 0$  est un paramètre de régularisation. Plus  $\lambda$  est grand, plus l'itération ressemble à un algorithme de descente du gradient alors que de petites valeurs de  $\lambda$  font que l'algorithme se rapproche de celui de Gauss-Newton.

## 3.8 Mises en œuvre

Nous présentons maintenant les étapes permettant de coder les variantes de méthodes introduites précédemment.

### 3.8.1 Mise en œuvre de la direction de Newton modifiée

En adaptant adéquatement les algorithmes de la section précédente, nous pourrions proposer des algorithmes pour calculer la direction de Newton modifiée utilisée à la section 3.4 pour construire un algorithme globalement convergent possédant la propriété de convergence locale quadratique. L'algorithme 3.3 résume les modifications; une implantation comportera en plus une précaution qui limite le nombre total d'itérations.

```

Newton modifié( $x, \epsilon, f$ )
{
  Donnés :  $x$  ;
  un critère d'arrêt,  $\epsilon > 0$  ;
  une fonction deux fois différentiable  $f$  ;
  répéter
  |  $d_N \leftarrow$  Direction_Newton_Modifiée( $\nabla f(x), \nabla^2 f(x)$ )
  |  $\theta \leftarrow$  Pas_admissible( $x, d_N, f$ )
  |  $x \leftarrow x + \theta d_N$ 
  jusqu'à (  $\|\nabla f(x)\| < \epsilon$  )
  Résultat  $\leftarrow x$ 

```

Algorithme 3.3: Newton modifié.

### \* Adaptation de la décomposition de Cholesky

Nous avons vu plus haut comment calculer la décomposition de Cholesky d'une matrice symétrique définie positive. Si la matrice d'intérêt n'est pas définie positive, l'algorithme ne pourra pas terminer, car certaines racines carrées qu'il doit obtenir porteront sur des nombres négatifs.

En modifiant la procédure de calcul de la décomposition de Cholesky d'une matrice symétrique  $Q$  pour remplacer *toute racine carrée d'un nombre négatif* par un petit nombre positif ( $\epsilon$ ), nous obtenons

$$Q = LL^t + E,$$

où  $E$  représente l'“erreur” de la décomposition due à la modification de la procédure de calcul. Ainsi, si  $Q$  est véritablement définie positive,  $E \equiv 0$  alors que si  $Q$  possède une ou des valeurs propres négatives, ou encore très petites,  $E$  assurera que  $LL^t$  est une matrice uniformément définie positive.

Cette idée simple n'est toutefois pas suffisante pour obtenir un algorithme performant. En effet, comme le calcul des facteurs de Cholesky fait intervenir une division par les éléments diagonaux, s'ils valent  $\epsilon$ , les facteurs de Cholesky deviendront énormes. Gill *et al* [18] ont proposé une stratégie beaucoup plus subtile et efficace. Il s'agit de déceler que  $Q$  n'est pas définie positive bien avant de rencontrer un pivot négatif. On peut récrire les formules de Cholesky comme suit, en dénotant  $Q_j$  la sous-matrice qu'il reste à traiter à l'étape  $j$ , et donc  $Q = Q_1$ ,

$$Q_j = \begin{pmatrix} \alpha_j & q_j^t \\ q_j & \hat{Q}_j \end{pmatrix}, \quad (3.12)$$

on peut calculer  $L_{jj} = \sqrt{\alpha_j + \delta_j}$ ,  $L_{ij} = (q_j)_i / L_{ii}$  pour  $i = j + 1, \dots, n$ , et

$$Q_{j+1} = \hat{Q}_j - \frac{q_j q_j^t}{\alpha_j + \delta_j}. \quad (3.13)$$

$\delta_j$  constitue la perturbation diagonale que l'on applique. La stratégie simple introduite ci-haut s'écrit  $\delta_j = -\min(\alpha_j - \epsilon, 0)$ . La stratégie préconisée dans Gill *et al* choisit  $\delta_j$  comme étant le plus petit nombre non-négatif tel que

$$0 \leq \frac{\|q_j\|_\infty^2}{\alpha_j + \delta_j} \leq \beta^2; \quad (3.14)$$

puisque  $\|q_j\|_\infty^2$  est positif,  $\alpha_j + \delta_j$  sera  $\geq 0$  et  $\geq \frac{\|q_j\|_\infty^2}{\beta^2}$ . En fait, si  $\alpha_j < 0$ , Gill *et al* recommandent  $\delta_j = -2\alpha_j$  pour conserver l'information de  $\alpha_j$  à un signe près, et finalement, on retient que  $\delta_j = \max(\epsilon, |\alpha_j|, \frac{\|q_j\|_\infty^2}{\beta^2}) - \alpha_j$ . On voit bien que si  $\alpha_j$  est suffisamment positive,  $|\alpha_j|$  devrait dominer dans le max, et alors  $\delta_j = 0$ .

$\beta$  est une borne a priori choisie pour minimiser un pire cas sur la norme  $\|E\|$  tout en assurant que  $E = 0$  si  $A$  est suffisamment définie positive. Si on dénote  $\epsilon_M$  comme étant la précision machine, Gill *et al* obtiennent

$$\beta^2 = \max\{\gamma, \xi/\sqrt{n^2 - 1}, \epsilon_M\}, \quad (3.15)$$

où  $\gamma = \max_i |Q_{ii}|$  et  $\xi = \max_{j < i} |Q_{ij}|$ . La portion  $\beta^2 \geq \gamma$  assure que  $E = 0$  lorsque  $Q$  est définie positive.

Bien entendu, ces choix de  $\delta_j$  à chaque pas de l'algorithme de décomposition de Cholesky influencera le comportement global de la modification de l'itération de Newton.

Un des objectifs de toute méthode de modification de la décomposition de Cholesky est d'obtenir une perturbation (la matrice  $E$ ) la plus petite possible. Dans cet esprit, il est avantageux de procéder à des permutations des lignes—colonnes de la sous-matrice  $\hat{Q}_j$  pour ramener dans le coin supérieur gauche le plus grand élément de la diagonale qui reste à traiter. De plus, comme discuté précédemment, nous utiliserons plutôt la forme avec la matrice diagonale  $D$ , et ainsi nous calculons  $P^tQP = LDL^t + E$ , où  $P$  est la matrice de permutation ramenant l'élément diagonal dans le coin supérieur gauche.

**Exercice 3.8.1** [*Direction de Newton modifiée*] Considérez la fonction  $f(x) = x_1^4 + x_1x_2 + (1 + x_2)^2$ .

- Montrez qu'à partir de l'origine, ni  $d_N$  ni  $-d_N$  ne constitue une direction de descente.
- Calculez la direction de Newton modifiée à l'aide de la décomposition de Cholesky modifiée.
- Vérifiez que la direction obtenue est bien une direction de descente.



**Exercice 3.8.2** [Implantation de la décomposition de Cholesky modifiée] Complétez les détails de la décomposition  $P^tQP = LDL^t + E$ ; implantez le calcul de  $L, D$  et  $P$ , et fournissez les outils pour calculer une direction de Newton modifiée.

### Adaptation du gradient conjugué

Nous adaptons maintenant l'algorithme du gradient conjugué pour résoudre les sous-problèmes linéaires issus de la méthode de Newton, même si la matrice  $\nabla^2 f(x)$  n'est pas définie positive. L'idée est la suivante : nous allons engendrer une suite  $0 = \delta_0, \delta_1, \dots$  par l'algorithme itératif du gradient conjugué appliqué au problème

$$\min_d q(d) \stackrel{\text{def}}{=} \nabla f(x_k)d + \frac{1}{2}d^t \nabla^2 f(x_k)d = g_k d + \frac{1}{2}d^t H_k d.$$

Pour ces sous-problèmes, nous noterons  $p_i$  la suite des directions conjuguées, et  $\vartheta_i$  les pas utilisés à l'intérieur de l'algorithme du gradient conjugué, de sorte que  $\delta_{i+1} = \delta_i + \vartheta_i p_i$ .  $\delta_i$  dénote la suite qui convergerait vers  $d_N$  si  $\nabla^2 f(x_k)$  était définie positive. Nous allons interrompre l'algorithme du gradient conjugué aussitôt qu'il devient capable de déceler que  $H_k$  n'est pas définie positive. Pour ce faire, remarquons que

$$\delta_l = \sum_{i=0}^l \vartheta_i p_i,$$

et, puisque les directions  $p_i$  sont mutuellement conjuguées par rapport à la matrice  $H_k$ ,

$$\delta_l^t H_k \delta_l = \sum_{i=0}^l \vartheta_i (p_i^t H_k p_i).$$

Par conséquent, il est certain que  $\delta_l^t H_k \delta_l > 0$  si tous les  $p_j$  qui entrent dans la somme satisfont tous  $p_j^t H_k p_j > 0$ . Nous interrompons donc l'algorithme du gradient conjugué aussitôt qu'une des directions  $p_{\bar{j}}$  ne satisfait pas  $p_{\bar{j}}^t H_k p_{\bar{j}} > \epsilon \|p_{\bar{j}}\|$ . Si jamais  $\bar{j} = 0$ , c'est que la direction  $p_0 = -\nabla f(x_k)^t$  est une direction de courbure négative, mais c'est néanmoins une direction de descente forte, et nous l'utiliserons :  $d = p_0$ . Autrement,  $\bar{j} > 0$  et nous utiliserons

$$d = \sum_{i=0}^{\bar{j}-1} \vartheta_i p_i,$$

et la définition de  $d$  assure que c'est une direction suffisamment descendante.

**\*Exercice 3.8.3** [*Direction descendante*] Démontrez que la direction obtenue par le gradient conjugué interrompu tel que décrit ci-haut est suffisamment descendante.

Finalement, on peut implanter une stratégie de résolution inexacte des problèmes linéaires en terminant l'algorithme du gradient conjugué aussitôt que  $\|\nabla q\| < \|\nabla f\|^2$ , ce qui ne nuit pas à la convergence locale quadratique de la méthode (voir section 3.4.3). L'algorithme 3.4 incorpore ces stratégies. Une mise en garde s'impose à ce moment : l'algorithme 3.4 est décrit *mathématiquement*, mais il faudra, dans son implantation, prévoir des "sorties de secours" car si la matrice  $Q$  est le moins mal conditionnée, l'algorithme mathématique ne terminera jamais, ne réussissant pas à atteindre la précision voulue avec des calculs en virgule flottante.

**Exercice 3.8.4** [*Direction de Newton modifiée*] Considérez à nouveau la fonction  $f(x) = x_1^4 + x_1x_2 + (1 + x_2)^2$ .

- a) à partir de l'origine, calculez la direction de Newton modifiée à l'aide de l'algorithme du gradient conjugué.
- b) Vérifiez que la direction obtenue est bien une direction de descente.

**Exercice 3.8.5** [*Implantation de la direction à l'aide du gradient conjugué*] Implantez l'algorithme 3.4. Remarquez que la matrice  $Q$  n'est utilisée que dans des expressions produit avec un vecteur. Ainsi, il n'est pas nécessaire de connaître la matrice  $Q$ , mais il suffit d'être capable d'évaluer  $\nabla_t(\nabla_x f(x + tv)) = v^t \nabla^2 f(x + tv)$  en  $t = 0$ . On peut utiliser des différences finies, ou un outil de différentiation automatique pour cette évaluation.

## 3.8.2 Mise en œuvre d'algorithmes de régions de confiance

### Adaptation du gradient conjugué

Il est remarquable que la technique de la section précédente peut également être utilisée pour résoudre approximativement le sous-problème introduit dans la section 3.3.3. Rappelons qu'il s'agit de résoudre un problème quadratique  $\min_{\|d\| \leq \Delta} q(d)$ . Une légère modification de l'algorithme 3.4 nous fournit l'algorithme 3.5. Plutôt que d'interrompre les itérations lorsqu'une direction de courbure négative est rencontrée, il suffit de suivre cette direction jusqu'à la frontière de la région de confiance. La mise en garde concernant les critères d'arrêt en précision finie s'appliquent encore ici.

Direction\_Newton\_Modifiée(c,Q)

```

{ Version du gradient conjugué ;  $\xi$  contrôle l'ordre de convergence }
{  $q(\delta) \stackrel{\text{def}}{=} (1/2)\delta^t Q \delta + c\delta$  }
 $\delta \leftarrow \vec{0}$ 
 $p \leftarrow -\nabla q(\delta)^t$ 
si ( $(p^t Q p < \epsilon_2 \|p\|^2)$ ) alors
  | Résultat  $\leftarrow p$ 
sinon
  répéter
    |  $\vartheta \leftarrow \frac{-\nabla q(\delta)p}{p^t Q p}$ 
    |  $\delta \leftarrow \delta + \vartheta p$ 
    | assez_précis  $\leftarrow (\|\nabla q(\delta)\| < \min(\epsilon_1, \|\nabla q(0)\|^{1+\xi}))$ 
    | si ( $\neg$  assez_précis) alors
      |  $\beta \leftarrow \frac{\nabla q(\delta) Q p}{p^t Q p}$ 
      |  $p \leftarrow -\nabla q(\delta)^t + \beta p$ 
      | courbure_négative  $\leftarrow (p^t Q p < \epsilon_2 \|p\|^2)$ 
    | jusqu'à ( assez_précis  $\vee$  courbure_négative)
  | Résultat  $\leftarrow \delta$ 

```

Algorithme 3.4: GC—direction de Newton modifiée

```

Région_de_confiance(c,Q,Δ,||·||)
{ Version du gradient conjugué; ξ contrôle l'ordre de convergence }
{  $q(\delta) \stackrel{\text{def}}{=} (1/2)\delta^t Q \delta + c\delta$  }
 $\delta \leftarrow \vec{0}$ 
 $p \leftarrow -\nabla q(\delta)^t$ 
répéter
   $\alpha \leftarrow \max\{\alpha > 0 : \|\delta + \alpha p\| \leq \Delta\}$ 
  sortie  $\leftarrow p^t Q p \leq 0$ 
  assez_précis  $\leftarrow (\|\nabla q(\delta)\| < \min(\epsilon_1, \|\nabla q(0)\|^{1+\xi}))$ 
  si ( $\neg$  sortie) alors
     $\vartheta \leftarrow \frac{-\nabla q(\delta)p}{p^t Q p}$ 
    sortie  $\leftarrow \vartheta > \alpha$ 
    si ( $\neg$  sortie) alors
       $\delta \leftarrow \delta + \vartheta p$ 
       $\beta \leftarrow \frac{\nabla q(\delta)Qp}{p^t Q p}$ 
       $p \leftarrow -\nabla q(\delta)^t + \beta p$ 
jusqu'à ( sortie  $\vee$  assez_précis)
si (sortie) alors  $\delta \leftarrow \delta + \alpha p$ 
Résultat  $\leftarrow \delta$ 

```

Algorithme 3.5: GC—Région de confiance

### 3.8.3 Mise en œuvre d'un pas admissible

Nous avons déjà vu que le procédé d'Armijo conduit à un algorithme globalement convergent. Cependant, il peut parfois être avantageux d'utiliser un algorithme plus efficace pour identifier un pas admissible, donc satisfaisant aux conditions de la définition 3.3.2 :

$$f(x + \theta d) - f(x) < \tau_0 \theta \nabla f(x) d, \quad \tau_0 \in ]0, \frac{1}{2}[,$$

$$\nabla f(x + \theta d) d > \tau_1 \nabla f(x) d, \quad \tau_1 \in ]\tau_0, 1[.$$

Un tel pas est une approximation d'un minimum local de la fonction  $h_{x_k, d}(\theta) = f(x_k + \theta d)$ .

Comme nous l'avons vu au chapitre précédent, le type d'algorithme le plus avantageux combine la réduction d'intervalle à une technique d'interpolation polynômiale.

```

Pas_admissible(x,d,f)
{ Version de l'algorithme d'Armijo }
{ 0 < τ₀ < 0.5. }
θ ← 1
tantque (f(x + θd) - f(x) > θτ₀∇f(x)d)
  | θ ← θ/2
  | Résultat ← θ

```

Algorithme 3.6: Armijo

### Interpolation polynômiale avec garde-fous

Tout comme dans le cas multidimensionnel, où la méthode d'interpolation polynômiale (Newton) a été modifiée pour lui procurer la propriété de convergence globale, nous combinons maintenant un algorithme de type réduction d'intervalle à une méthode d'interpolation polynômiale pour rendre la méthode modifiée globalement convergente. Il s'agit donc de la même démarche qu'aux sections 2.9 et 3.4, mais les outils de modification sont un peu différents.

Afin d'utiliser les algorithmes robustes du chapitre 2, introduisons la fonction

$$\phi(\theta) = h_{x,d}(\theta) - h_{x,d}(0) - \tau_0 \theta h'_{x,d}(0).$$

La fonction  $\phi$  vaut 0 en  $\theta = 0$ , et tout point  $\theta$  tel que  $\phi(\theta) < 0$  satisfait au critère d'Armijo. De plus, un point stationnaire de  $\phi$  satisfait  $\phi'(\theta) = h'_{x,d}(\theta) - \tau_0 h'_{x,d}(0) = 0$ , et donc, puisque  $\tau_0 < \tau_1$  et  $h'_{x,d}(0) < 0$ , satisfait au critère de Wolfe. Par conséquent, la recherche linéaire consiste à approcher un point stationnaire de la fonction  $\phi$  possédant une valeur négative  $\phi(\theta) < 0$ .

Après avoir identifié un intervalle initial à l'aide de l'algorithme 2.1, il est avantageux de le réduire en utilisant un polynôme d'interpolation cubique. En effet, le seul cas défavorable correspond maintenant à un point qui satisfait aux critères de Armijo et Wolfe, c'est-à-dire que  $\phi(\theta) < 0$  et  $\phi'(\theta) \approx 0$ .

Pour être véritablement assuré d'obtenir un pas admissible, cependant, il faut raffiner les algorithmes du chapitre 2 afin de garantir que la valeur obtenue satisfait bel et bien à la condition que  $\phi(\theta) < 0$ .

## 3.9 Problèmes de très grande taille

Lorsque les problèmes à résoudre sont de très grandes dimensions (un million et plus de variables), il est impossible de manipuler les matrices des dérivées secondes à moins de problèmes présentant une structure favorable et des matrices Hessiennes creuses. Nous discutons ici brièvement des stratégies évitant l'utilisation de matrices en tant que telles.

Même si le but est d'éviter l'utilisation explicite des matrices, il est hors de question de se ramener à la méthode de la pente la plus forte, dont la vitesse de convergence est beaucoup trop faible. Trois familles de stratégies sont couramment utilisées :

- l'exploitation de la structure éparsée des matrices.
- gradient conjugué non-linéaire ;
- méthode de Newton utilisant seulement des *produits Hessien-vecteur*.

Nous ne traiterons pas des matrices creuses dans cette section, mais donnerons quelques détails des deux autres stratégies.

### 3.9.1 Gradient conjugué non-linéaire

Le gradient conjugué que nous avons examiné pour la solution de problèmes quadratiques peut se généraliser à la minimisation de fonctions non-quadratiques. Plusieurs chercheurs ont proposé des formules, toutes équivalentes pour les fonctions quadratiques, pour définir un algorithme de gradient conjugué non-linéaire. Certaines variantes se sont avérées plus efficace et robustes.

Les méthodes quasi-Newton sont également équivalentes au gradient conjugué linéaire appliqué à des fonctions quadratiques. Une idée ingénieuse pour les adapter aux problèmes de très grande taille consiste à se limiter à quelques mises-à-jour sans constituer explicitement la matrice  $B_k$  ou  $H_k$ . Ces variantes sont nommés méthodes quasi-Newton à mémoire limitée, et peuvent être vue comme des améliorations des algorithmes de gradient conjugué non-linéaire. Rappelons les formules équivalentes du gradient conjugué, en remplaçant la fonction quadratique  $q$  par une fonction  $f$  :

$$d_k = -\nabla f(x_k)^t + \beta_{k-1}d_{k-1} \quad (3.16)$$

$$\theta_k = \text{Pas de déplacement admissible} \quad (3.17)$$

$$x_{k+1} = x_k + \theta_k d_k \quad (3.18)$$

$$\beta_k = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}. \quad (3.19)$$

Les variantes du gradient conjugué non-linéaire diffèrent par la formule de calcul du paramètre  $\beta_k$ . Pour alléger les notations, nous utiliserons  $g_k = \nabla f(x_k)$ ,  $y_k = g_k - g_{k-1}$ . Pour des fonctions quadratiques avec un calcul exact de  $\theta_k$ , toutes ces formules deviennent équivalentes. Pour des fonctions non-linéaires, avec l'utilisation de pas admissibles possiblement peu précis (valeurs de  $\tau_0$  proche de 0 et  $\tau_1$  proches de 1), la formule de Hager et Zhang semble plus robuste et performante en pratique.

### Formule de Fletcher et Reeves

La formule de Fletcher et Reeves est justement la formule (3.19) que l'on peut récrire

$$\beta_k = \frac{g_{k+1}g_{k+1}^t}{g_k g_k^t}$$

### Formule de Polak et Ribière

Dans cette formule, en fait, on ajoute le terme  $-g_{k+1}g_k^t$  au numérateur ; pour une fonction quadratique avec un pas de déplacement  $\theta$  optimal, ce terme additionnel vaut 0.

$$\beta_k = \frac{g_{k+1}y_{k+1}^t}{g_k g_k^t}$$

### Formule de Hestenes et Stiefel

$$\beta_k = \frac{g_{k+1}y_{k+1}^t}{y_{k+1}d_k}$$

### Formule de Hager et Zhang

$$\beta_k = \frac{1}{y_{k+1}d_k} \left( y_{k+1} - 2d_k^t \frac{\|y_{k+1}\|^2}{y_{k+1}d_k} \right) g_{k+1}^t$$

**Exercice 3.9.1** [Équivalence des formules pour fonctions quadratiques] Lorsque  $f$  est une fonction quadratique, il est possible de définir  $\theta$  exactement à l'aide d'une formule explicite, et nous sommes dans le royaume du gradient conjugué linéaire tel que présenté à la section 3.6.2. Dans ce cas,  $g_k \perp g_{k-1}$  et  $g_k \perp d_{k-1}$ . Utilisez ces propriétés pour démontrer que toutes ces formules sont équivalentes.

### Méthode L-BFGS

Cette méthode signifie **L**imited memory **BFGS** et propose une mise-à-jour des matrices quasi-Newtoniennes à partir de quelques paires  $y_i, s_i$  appliquées à une matrice de mise-à-l'échelle diagonale  $H_0$  plutôt que manipuler explicitement des matrices  $n \times n$ .

Cette implantation de la célèbre formule BFGS consiste à conserver  $m$  paires de quantités  $y_i = g_i - g_{i-1}$  et  $s_i = x_i - x_{i-1}$  pour reconstruire un produit  $Hg_k^t$  correspondant à la direction quasi-Newton. La matrice  $H$  n'est jamais formée explicitement, mais les calculs repris en accumulant dans deux boucles les résultats de produits scalaires appropriés.

Dans l'algorithme 3.7, on met à jour une matrice diagonale  $H_k^0$  en utilisant les  $m$  derniers points. Le choix de la matrice à mettre-à-jour a un impact sur l'efficacité de la méthode. En effet, deux études distinctes arrivent à cette conclusion [27] et [16]. Un choix décommandé est d'utiliser un multiple de l'identité  $\gamma_k I$ . Le choix de  $\gamma_k = \frac{y_k^t s_k}{\|y_k\|^2}$  est recommandé dans [27] et [16]. Remarquons ici que l'option 'gc' de la commande optim de scilab est basée sur l'article [16].

```

L-BFGS( $g_k$ )
{ Calcule le produit  $Hg_k^t$  à partir de  $m$  paires  $y_i$  et  $s_i$  }
{  $r_i \stackrel{\text{def}}{=} \frac{1}{y_i^t s_i}$  }
 $q \leftarrow g_k$ 
pour ( $i = k \xrightarrow{-1} k - (m - 1)$ )
   $a_i \leftarrow r_i q^t s_i$ 
   $q \leftarrow q - a_i y_i$ 
 $z \leftarrow H_k^0 q^t$ 
pour ( $i = k - (m - 1) \xrightarrow{+1} k$ )
   $b \leftarrow r_i y_i^t z$ 
   $z \leftarrow z + s_i (a_i - b)$ 
Résultat  $\leftarrow z$ 

```

Algorithme 3.7: L-BFGS



### 3.9.2 Newton tronqué

Lorsqu'on utilise le gradient conjugué pour résoudre les équations de Newton, nul besoin de disposer de la matrice Hessienne explicitement : seuls les produits Hessien–vecteur sont utilisés au sein de l'algorithme, que ce soit dans la variante avec recherche linéaire ou celle avec région de confiance. Pour des problèmes de grande taille, mieux vaut éviter de constituer la matrice Hessienne.

En d'autres termes, dans les algorithmes 3.4 et 3.5, à chaque fois qu'on utilise la matrice  $Q$ , c'est au sein d'une expression impliquant le produit  $Qp$ . Voyons comment calculer directement les produits hessien–vecteur sans calculer explicitement la matrice  $\nabla^2 f(x)$ . En anglais, on parle de *matrix free Newton methods*.

#### Différences finies

Une manière de calculer le produit  $\nabla^2 f(x)v$  est d'évaluer  $\frac{\partial(\nabla f(x+tv)}{\partial t}$  en  $t = 0$ . Or, une dérivée  $\frac{d\phi(t)}{dt}$  peut s'approcher par  $\frac{\phi(t+\epsilon)-\phi(t)}{\epsilon}$  pour une valeur de  $\epsilon$  assez petite. Choisissons donc  $\epsilon \approx 10^{-8}$ , environ la racine carrée de la précision machine et écrivons  $\nabla^2 f(x)v \approx \frac{\nabla f(x)-\nabla f(x+\epsilon v)}{\epsilon}$ .

#### Différentiation automatique

Puisque l'utilisation des dérivées, gradients et hessiens, est omni présente dans les développements de conditions d'optimalité et des algorithmes, penchons-nous sur l'obtention de ces dérivées.

Le calcul d'une fonction  $f$  est en fait constitué d'une séquence d'opérations élémentaires (+, −, \*, ÷) et de fonctions usuelles (sin, cos, ln, etc). En appliquant les règles de dérivation à une telle séquence d'opérations on peut obtenir une séquence calculant les dérivées (gradients et hessiens) de la fonction  $f$ . Un outil qui automatise cette opération se nomme *différentiation automatique*. Contrairement aux différences finies, l'obtention des dérivées par la différentiation automatique n'implique pas d'approximation ou de valeur  $\epsilon$ , mais ces calculs sont exacts, du moins aussi exacts que les calculs de la fonction  $f$  elle même. De plus, il existe une variante qui permet d'obtenir  $\nabla f(x)$  à un coût de calcul proportionnel au coût de calcul de  $f$ , et ce peu importe la dimension  $n$  des paramètres  $x$ .

## 3.10 Résumé

Dans ce chapitre, nous avons développé des outils d'analyse de problèmes d'optimisation différentiable sans contrainte qui conduisent à l'énoncé de conditions nécessaires, légèrement différentes des conditions suffisantes d'optimalité : un point  $x^*$  qui est un minimum local de  $f$  satisfait nécessairement  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice semi-définie positive. Par

ailleurs, pour avoir l'assurance qu'un point est bien un minimum local de  $f$ , il faut exiger de plus que la matrice  $\nabla^2 f(x^*)$  soit définie positive. L'écart vient du fait que nous n'utilisons que les dérivées d'ordre un et deux de  $f$ , alors que ce sont possiblement ses dérivées d'ordre supérieur qui décrivent adéquatement son comportement près de  $x^*$ .

Ensuite, nous avons construit des schémas algorithmiques itératifs abstraits qui possèdent la propriété de convergence globale, c'est-à-dire que peu importe le point de départ des itérations, les points d'accumulation de la suite engendrée par l'algorithmes satisfont à une condition d'optimalité. L'existence d'algorithmes d'une telle famille, nommés algorithmes de descente, provient de la structure du problème d'optimisation ; la solution de systèmes d'équations quelconques ne peut pas, en général, être obtenue par des algorithmes de descente.

Nous avons ensuite modifié l'itération de Newton en assurant une propriété de descente suffisante de la direction, et en utilisant un pas admissible dans la direction. Ces modifications préservent la convergence locale quadratique de la méthode, mais la dotent en plus de la propriété de convergence globale.

L'implantation de la méthode de Newton modifiée est réalisée en adaptant judicieusement les algorithmes de minimisation d'une fonction quadratique, et nous avons illustré ce point par l'analyse de deux techniques : la décomposition de Choleski, et l'algorithme du gradient conjugué.

Finalement, nous avons étudié quelques algorithmes unidimensionnels, que nous avons combinés en un algorithme robuste tout en étant efficace, en appliquant encore une fois le principe de modification d'un algorithme d'interpolation polynômiale, comme l'algorithme de Newton, pour le rendre globalement convergent.

## 3.11 Extensions et références

L'optimisation différentiable sans contrainte ainsi que les problèmes de racines d'équations non-linéaires ont été beaucoup étudiées. Il est évidemment impossible de rendre justice à la richesse de méthodes, d'idées publiées sur ce sujet dans un texte d'introduction aussi succinct ! Le livre de Dennis et Schnabel [12] est consacré dans son entier à de tels problèmes et le livre sur les méthodes de région de confiance de Conn, Gould et Toint [6] consacre 1000 pages à cette seule approche.

Les méthodes de Newton modifiées efficaces ont été popularisées assez tard, vers 1972 [29, 17]. Gill, Murray et leurs collaborateurs sont à l'origine des méthodes utilisant la décomposition de Cholesky ; avant leurs travaux, les méthodes de Newton modifiées étaient considérées trop délicates à mettre en œuvre pour être compétitives. L'implantation à l'aide du gradient conjugué modifié ainsi que la notion de méthode de Newton inexacte sont dues à Dembo, Eisenstat et Steihaug [9] alors que l'implantation à l'aide du gradient conjugué se retrouve dans [10].

Les méthodes de région de confiance originent dans les travaux de Levenberg[25] et

Marquardt[28] pour des problèmes de moindres carrés. Goldfeld, Quandt et Trotter[19] ont développé ces idées pour des problèmes de programmation non-linéaire, et Steihaug[35] et Toint[36] sont à l'origine de l'utilisation du gradient conjugué pour résoudre les sous-problèmes quadratiques.

Les méthodes quasi-Newtonniennes ont connu un développement remarquable dans les années 1970. L'article de Dennis et Moré[11] présente un tour d'horizon complet. Plus récemment, des chercheurs ont étudié des variantes de l'équation quasi-Newton, et les progrès récents sont bien présentés dans [37].

L'algorithme du gradient conjugué est dû à Hestenes et Stiefel [23]. La dérivation présentée à la section 3.6.2 est inspirée d'un article de Beale [1]. De nombreux travaux, dont ceux de Polak et Ribière [31], ainsi que ceux de Fletcher et Reeves [15] ont porté sur la généralisation de l'algorithme du gradient conjugué à des fonctions non-quadratiques.

**Exercice 3.11.1** [*Optimisation dans  $\mathbb{R}^n$* ] La fonction de Rosenbrock est fameuse pour les tests de méthodes d'optimisation dans  $\mathbb{R}^n$ . La version originale de cette fonction pour  $n = 2$  est la suivante :

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Son minimum est en  $x = (1, 1)^t$  et le point de départ habituel est  $x_0 = (-1, 1.2)^t$ .

$$f(x_0) = 8, \nabla f(x_0) = (76, 40) \text{ and } \nabla^2 f(x_0) = \begin{pmatrix} 722 & 400 \\ 400 & 200 \end{pmatrix}$$

- Vérifiez que  $x^* = (1, 1)^t$  est un minimum local satisfaisant aux conditions suffisantes d'optimalité.
- Vérifiez que  $\nabla^2 f(x_0)$  n'est pas définie positive.
- Malgré que  $\nabla^2 f(x_0)$  ne soit pas définie positive, il est possible que la direction de Newton (non modifiée) soit tout de même une direction de descente. Vérifiez que la direction de Newton est en effet une direction de descente dans ce cas particulier.
- Utilisez l'algorithme gradient conjugué pour calculer la direction de Newton modifiée en utilisant le test d'arrêt "courbure\_négative".
- Vérifiez que le pas de déplacement  $\theta = 1$  est admissible pour la direction de Newton modifiée, la direction de Newton non modifiée, mais pas pour la direction  $d = -\nabla f(x_0)$ . Prenez  $\tau_0 = 0.1$  et  $\tau_1 = 0.9$ .

**Exercice 3.11.2** [*Projet de programmation*] Implantez quelques variantes de l'algorithme de Newton modifié : comparez les résultats d'une variante avec une recherche de pas admissible simple, l'algorithme d'Armijo qui prend la première puissance de  $\frac{1}{2}$  satisfaisant

au critère d'Armijo, une recherche linéaire plus sophistiquée combinant les idées des algorithmes 2.1 et 2.4, et enfin une version avec régions de confiance. Vous pouvez utiliser la modification issue de la décomposition de Cholesky, ou encore l'algorithme du gradient conjugué (comparez le comportement de la version région de confiance avec celle de la recherche linéaire dans ce cas). Votre programme utilisera trois fonctions qui retournent respectivement  $f(x)$ ,  $\nabla f(x)$  et  $\nabla^2 f(x)$ . Testez votre programme (entre autres) avec la fonction de Rosenbrock :  $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$  à partir du point de départ  $(-1.2, 1.0)^t$ . Cette fonction a longtemps constitué un défi pour les programmes de minimisation. Les modifications de l'algorithme de Newton traitent habituellement cette fonction facilement.

Le point  $(1,1)$  est un minimum local de la fonction de Rosenbrock. L'algorithme de Newton modifié produit les itérations tracées en rouge alors que l'algorithme de descente du gradient celles en noir. On peut voir que le gradient zigzagge de part et d'autre de la vallée, puis "saute par dessus" plutôt que de se glisser dans la vallée, mais malgré tout, n'arrive pas à atteindre la solution en un nombre d'itérations raisonnable car il zigzage encore en se rapprochant lentement du minimum.

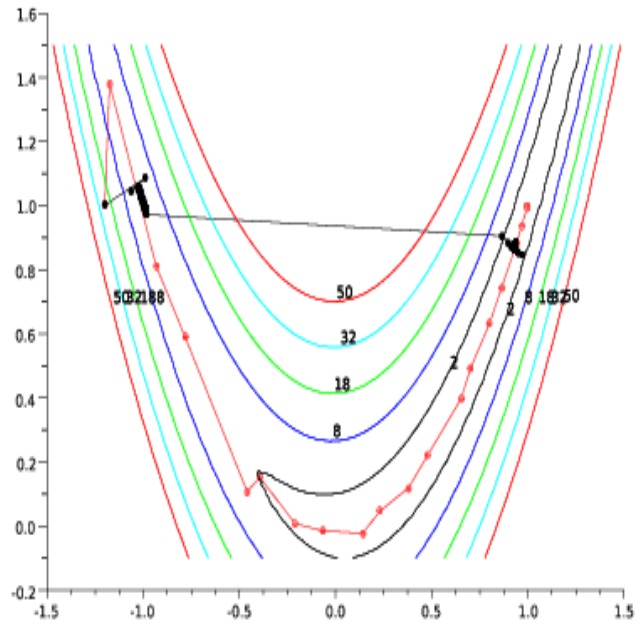
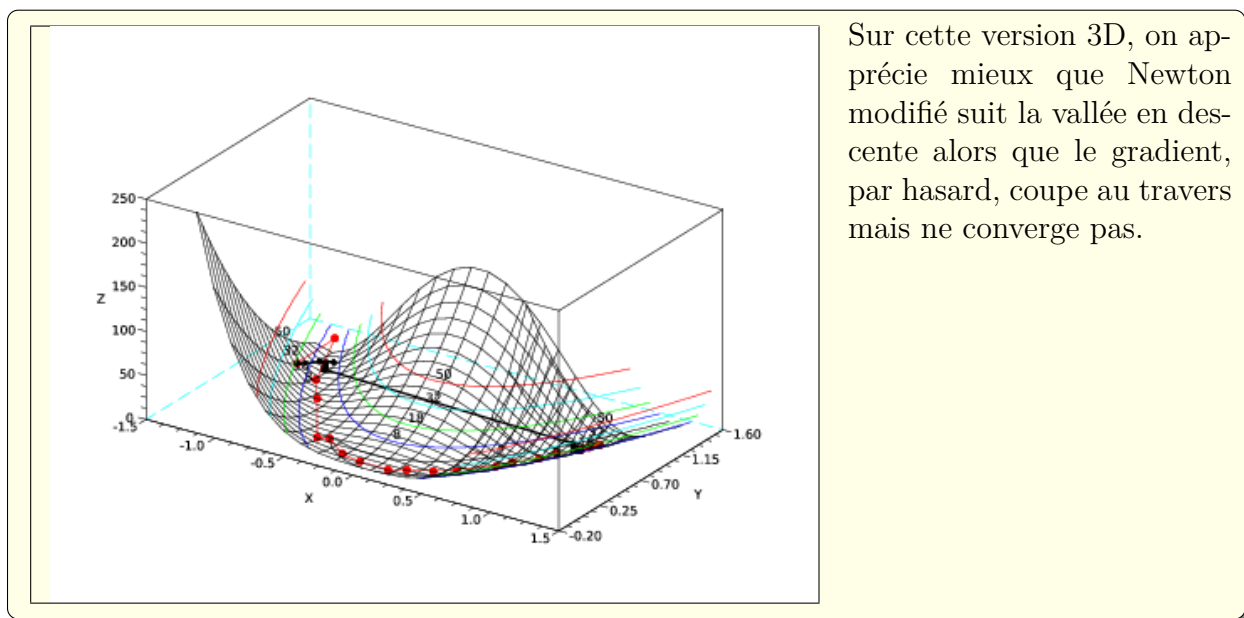


FIGURE 3.7 – Exemple des algorithmes de Newton et du gradient



Sur cette version 3D, on apprécie mieux que Newton modifié suit la vallée en descente alors que le gradient, par hasard, coupe au travers mais ne converge pas.

FIGURE 3.8 – Exemple des algorithmes de Newton et du gradient

## 3.12 Tous les exercices du chapitre

**\*Exercice (3.2.1, page 103)** [*Extension des conditions d'optimalité*] Le théorème 2.2.1 concerne la minimisation sans aucune restriction sur  $\theta$ . Considérons le problème  $\min_{a \leq \theta \leq b} h(\theta)$ . Si un point  $\theta_0$  est dans l'intérieur de l'intervalle  $[a, b]$ , alors le théorème s'applique directement ; lorsque  $\theta_0 = a$  ou  $\theta_0 = b$ , il suffit que  $h$  soit croissante (en  $a$ ) ou décroissante (en  $b$ ) pour avoir un minimum local sous contrainte : en effet, on ne compare  $a$  qu'avec des valeurs  $\theta \geq a$ . Obtenez une condition nécessaire et suffisante pour que  $a$  ou  $b$  soient des minima locaux de  $h$ . L'exercice consiste donc à fournir les conditions encadrées et démontrer :

- $\theta^*$  est min local de la fonction  $h$  si et seulement si
- $\theta^* = a$  et conditions pour que  $h$  soit croissante ou
  - $a < \theta^* < b$  et la première dérivée non nulle de  $h$  est d'ordre pair et de signe positif ou
  - $\theta^* = b$  et conditions pour que  $h$  soit décroissante.

**Exercice (3.2.2, page 104)** [*Dérivées de  $h$* ] Obtenez les expressions des deux premières dérivées de  $h_{x,d}(\theta)$  par rapport à  $\theta$ .

**Exercice (3.2.3, page 108)** [*Minima forts*] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$  ; si  $x^*$  est un minimum local de  $f$ , montrez qu'un point qui satisfait aux conditions suffisantes de second ordre est un minimum local strict et isolé. Vous pouvez utiliser le théorème 3.2.6 pour assurer que  $\nabla^2 f(x^*)$  demeure définie positive dans un voisinage approprié de  $x^*$

**Exercice (3.2.4, page 108)** [*Points stationnaires*] Considérez la fonction  $f(x) = (x_1 + x_2)^2 + (2(x_1^2 + x_2^2 - 1) - \frac{1}{3})^2$ .

- Vérifiez que l'origine ainsi que le point  $(-\sqrt{\frac{7}{12}}, \sqrt{\frac{7}{12}})$  constituent des points stationnaires de  $f$ .
- Déterminez la nature de chacun de ces points stationnaires.

**Exercice (3.2.5, page 108)** [*Géométrie*] Considérez deux droites d'équations paramétriques  $x(t_1) = o_1 + t_1 d_1$  et  $x(t_2) = o_2 + t_2 d_2$ . On cherche à trouver les points sur les droites les plus proches, permettant de donner la distance entre les droites. On minimise donc  $f(t) = \|x_1(t_1) - x_2(t_2)\|^2$  par rapport à  $t$ , de dimension deux. Utilisez les conditions d'optimalité pour trouver une expression pour  $t^* \in \mathbb{R}^2$ , et vérifiez que c'est bel et bien un minimum local.

**Exercice (3.2.6, page 109)** [*Point de selle*] Nous avons présenté plus haut (figure 3.2) la fonction  $(x^2 - y)(x^2 - 4y)$  pour laquelle l'origine n'est qu'un point de selle, mais constitue un minimum local de la fonction objectif le long de toute droite passant par l'origine. Cet exercice approfondi l'étude de ce problème.

- Vérifiez que l'origine est un point stationnaire de la fonction  $f(x, y) = (x^2 - y)(x^2 - 4y)$  :  $\nabla f(0, 0) = 0$ .
- Considérez la fonction réelle  $h_d(\theta) \stackrel{\text{def}}{=} f(\theta d_x, \theta d_y)$ . Vérifiez que quel que soit  $d$  avec  $d_y \neq 0$ ,  $h'_d(0) = 0$  et  $h''_d(0) > 0$ .
- Pour un vecteur  $d$  quelconque, identifiez une valeur  $\epsilon_d > 0$  pour laquelle  $h_d(\theta) \geq 0$  pour toute valeur de  $\theta$  satisfaisant  $|\theta| \leq \epsilon_d$ .
- Vérifiez que  $f(x, y) \leq 0$  pour tous les points satisfaisant  $x^2 = 2y$ .

**Exercice (3.2.7, page 109)** [*Moindres carrés*] Dans cet exercice, nous explorons un problème d'ajustement de fonction. Les valeurs suivantes de  $x$  et  $y$  sont données :

x	0	1	2	3
y	0	0	3	9

On soupçonne qu'une relation d'ordre 3 relie  $x$  à  $y$ , et on envisage un modèle du type  $y = a_0 + a_3 x^3$ .

- Formulez le problème d'identifier les valeurs de  $a_0$  et  $a_3$  qui minimisent la somme des carrés des différences entre le modèle et les données.
- Exprimez les conditions d'optimalité (nécessaires et suffisantes, de premier et second ordre) du problème en a).

- c) Nous allons traiter une simplification du problème en supposant que  $a_0 = 0$ , donc que la relation recherchée est purement cubique (nouveau problème avec seulement  $a_3$ ). On analyse ici le problème de dimension un qui ne comporte que la variable  $a_3$ . Cette simplification semble raisonnable puisque pour  $x = 0$ ,  $y = 0$ .
- i) calculez la valeur de  $a_3$  qui minimise la somme des carrés des différences entre le modèle simplifié et les données;
  - ii) est-ce que les conditions suffisantes d'optimalité sont satisfaites pour la solution du problème de dimension un en i) ?
- d) Revenons maintenant au modèle de dimension deux en a) pour calculer les valeurs de  $a_0$  et  $a_3$ . Calculez les valeurs de  $a_0$  et  $a_3$  du modèle avec l'algorithme du gradient conjugué. Est-ce que les conditions suffisantes d'optimalité sont satisfaites ?

**Exercice (3.3.1, page 113)** [Algorithme du gradient]

- a) Pour la fonction  $f(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$ , vérifiez que si à une certaine étape  $k$  l'algorithme atteint le point  $\begin{pmatrix} 0 \\ 1 - 1/5^k \end{pmatrix}$ , alors à l'étape  $k + 2$ , il aura atteint le point  $\begin{pmatrix} 0 \\ 1 - 1/5^{k+1} \end{pmatrix}$ .
- b) Pour la fonction  $x^2 + 2y^2 + 4x + 4y$ , vérifiez par induction que l'algorithme du gradient amorcé à l'origine engendre la suite  $\begin{pmatrix} \frac{2}{3^k} - 2 \\ (-\frac{1}{3})^k - 1 \end{pmatrix}$ . Déduisez-en le minimum de la fonction.

**Exercice (3.3.2, page 115)** [Direction de l'opposé du gradient] Justifiez que la direction  $d = -\nabla f(x)^t$  satisfait aux conditions de la définition 3.3.1 ; pour quelles valeurs de  $\gamma_0$  et  $\gamma_1$  ?

**Exercice (3.3.3, page 115)** [Géométrie des directions] Quel est l'angle le plus grand entre une direction  $d$  et l'opposé du gradient tel que la direction satisfasse la définition 3.3.1 ? L'angle est fonction de  $\gamma_0$  et  $\gamma_1$ .

**Exercice (3.3.4, page 117)** [Pas admissible] Suite de l'exercice 2.9.1 page 87. Considérez la fonction  $h(\theta) = 1 - 12\theta + 7.5\theta^2 - \theta^3$ .

- a) Pour quelles valeurs de  $\tau_0$  le critère d'Armijo est-il satisfait pour le minimum local situé dans l'intervalle  $[0, 3]$  ?
- b) Explicitez le(s) intervalle(s) de pas  $\theta$  admissible(s) pour des valeurs de  $\tau_0 = 0.25$  et  $\tau_1$  de votre choix.

**Exercice (3.3.5, page 120)** [Armijo avec  $h$  quadratique] Dans cet exercice, nous allons examiner la fonction  $h(\theta) = \frac{a}{2}\theta^2 + b\theta + c$ . Commençons en a) et b) par se poser la question sur les valeurs intéressantes des coefficients  $a, b$  et  $c$ .



- a) Si on sait que la fonction  $h$  est en fait la fonction  $h_{x,d}(\theta) = f(x + \theta d)$  pour une direction  $d$  suffisamment descendante, que peut-on dire des coefficients  $a, b$  et  $c$  ?
- b) Si on sait que la fonction  $h$  est bornée inférieurement, que peut-on dire des coefficients  $a, b$  et  $c$  ?
- c) Supposons maintenant que la fonction  $h$  possède les caractéristiques de a) et b), i. e. qu'elle provient d'une fonction  $f$ , selon une direction suffisamment descendante, et qu'elle est bornée inférieurement. Justifiez que  $h$  possède un seul minimum local, et ce pour une valeur  $\theta^* > 0$ .
- d) Pour quel ensemble de valeurs de  $\tau_0$  le minimum local de  $h$  satisfait-il au critère d'Armijo ?
- e) Pour quel ensemble de valeurs de  $\tau_1$  le minimum local de  $h$  satisfait-il au critère de Wolfe ?

**Exercice (3.3.6, page 121)** [*Vitesse de convergence de l'algorithme du gradient*] On peut montrer que l'ordre de convergence de l'algorithme de la pente la plus forte est linéaire, et que le taux de convergence est donné par le rapport  $\frac{1-r}{1+r}$  où  $r = \lambda_{\min}/\lambda_{\max}$  et  $\lambda_{\min}$  et  $\lambda_{\max}$  sont les valeurs propres extrêmes de la matrice  $\nabla^2 f(x^*)$ . Considérez la fonction quadratique  $f(x, y) = 10x^2 + y^2$ , et vérifiez numériquement que ce taux est atteint si l'on amorce l'algorithme du gradient en  $\begin{pmatrix} 0.1 \\ 1 \end{pmatrix}$ . Remarquons que ce taux peut être arbitrairement proche de un si le rapport des valeurs propres extrêmes est proche de zéro.

**Exercice (3.3.7, page 122)** [*Méthode de Newton*] Considérez à nouveau (exercice 3.2.4 page 108) la fonction  $f(x) = (x_1 + x_2)^2 + (2(x_1^2 + x_2^2) - 1) - \frac{1}{3})^2$ . Vérifiez que la méthode de Newton peut être appliquée à partir de  $x^0 = (\sqrt{7/6}, 0)^t$  et que le prochain point, en utilisant  $\theta = 1$  est situé sur la droite reliant les deux points stationnaires.

**Exercice (3.3.8, page 123)** [*Direction de descente*] Complétez l'analyse des directions  $d = -H^{-1}\nabla f(x)^t$  en fournissant des expressions pour  $\gamma_0$  et  $\gamma_1$ .

**Exercice (3.3.9, page 123)** [*Algorithme de descente*] Considérez la fonction  $f(x) = x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + e^{x_1+x_2}$ .

- a) Écrivez la condition d'optimalité de premier ordre pour cette fonction. Est-ce que cette condition est suffisante pour ce problème ? En d'autres termes, est-ce que tous les points stationnaires de cette fonction sont des minima locaux ?
- b) Vérifiez que l'origine n'est pas un point stationnaire de  $f$ , et identifiez une direction de descente (suffisamment descendante) pour  $f$  à l'origine.
- c) Appliquez le procédé d'Armijo (la première valeur parmi  $1, \frac{1}{2}, \frac{1}{4}, \dots$ ) à la fonction  $h_d(\theta) = f(x + \theta d)$  pour la direction que vous avez proposée en b). Utilisez une valeur de  $\tau_0 = 0.1$  dans le critère d'Armijo.

**Exercice (3.3.10, page 124)** [*Newton—descente*]

- a) Pour minimiser une fonction de  $n$  variables, si  $\nabla^2 f(x)^{-1}$  possède des valeurs propres comprises entre 1 et 100 quel que soit  $x \in \mathbb{R}^n$ , justifiez que la direction de Newton,  $d_N = -\nabla^2 f(x)^{-1} \nabla f(x)^t$  est suffisamment descendante en exhibant des valeurs de  $\gamma_0$  et  $\gamma_1$ .
- b) Si  $\nabla^2 f(x)$  possède des valeurs propres négatives, est-il vrai que  $d_N$  est forcément *non* descendante ? Si oui, justifiez ; sinon, donnez un exemple où  $d_N$  est malgré tout descendante, et donnez un exemple où  $d_N$  n'est pas descendante.

**\*Exercice (3.3.11, page 128)** [Preuve du corollaire 3.3.2] Adaptez la preuve du théorème 3.3.5 ; indice : le cas 2 n'a pas besoin de modification, et le cas 1 doit tenir compte du fait que  $d$  n'est plus un minimum global, mais seulement une réduction au moins aussi bonne que le point de Cauchy.

**\*Exercice (3.3.12, page 129)** [Preuve du corollaire 3.3.3] Adaptez la preuve du théorème 3.3.5 ; indice : le cas 2 n'a pas besoin de modification, et le cas 1 doit tenir compte du fait que  $d$  n'est plus un minimum global, mais seulement une réduction au moins aussi bonne que le point de Cauchy et la direction de courbure la plus négative ; par le corollaire 3.3.2, les points d'accumulation sont stationnaires.

**Exercice (3.4.1, page 130)** [Propriété de descente de la méthode de Newton] Trouvez la plus grande boule centrée sur l'origine pour laquelle le hessien de la fonction  $f(x) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + x_1^4$  est une matrice définie positive. Si l'on démarre en un point  $x^0 = (x, x)$ , pour quelles valeurs dans cette boule l'algorithme de Newton (avec  $\theta = 1$ ) converge-t-il ?

**\*Exercice (3.4.2, page 133)** [Admissibilité du pas unitaire] Justifiez les affirmations suivantes. Si  $\nabla^2 f(x^*)$  est une matrice définie positive, et si  $x_k$  est proche de  $x^*$ , alors

- a)  $\|\epsilon_k\| \sim \|\nabla f(x_k)\|$  ; il s'agit de montrer que  $\|\epsilon_k\| \sim \mathcal{O}(\|\nabla f(x_k)\|)$  et  $\|\nabla f(x_k)\| \sim \mathcal{O}(\|\epsilon_k\|)$
- b) si  $\|r_k\| \sim \mathcal{O}(\|\epsilon_k\|^2)$ , alors un pas unitaire demeure localement admissible. Modifiez légèrement la preuve du théorème 3.4.1 pour y incorporer  $r_k$ , puis constater que le même raisonnement demeure valable.
- c) sous quelles conditions sur  $\|r_k\|$  l'adaptation de la preuve du théorème 3.4.1 en b) demeure-t-elle valable ?

**\*Exercice (3.6.1, page 138)** [Formulation sans racine carrée] Développez des formules itératives ligne par ligne et colonne par colonne pour calculer  $Q = LDL^t$  pour une matrice  $Q$  définie positive ; dans ce cas,  $D$  est une matrice diagonale positive.

**Exercice (3.6.2, page 139)** [Rappels de Gramm-Schmidt] Quelles sont les formules utilisées dans le procédé d'orthogonalisation de Gramm-Schmidt ?

**\*Exercice (3.6.3, page 140)** [*Simplification de Gramm-Schmidt*] Vérifiez que le procédé de Gramm-Schmidt se simplifie bien aux formules du gradient conjugué lorsque les directions linéairement indépendantes utilisées pour produire les directions conjuguées sont les gradients des points  $x_k$ .

**Exercice (3.6.4, page 140)** [*Application de l'algorithme du gradient conjugué*] Appliquez l'algorithme du gradient conjugué à la fonction  $q(x, y) = 2x^2 - 2xy + y^2 + 2x - 2y$  à partir du point  $x_0 = (10, 5)^t$ ; comparez vos résultats avec l'exemple du tableau 3.1.

**\*Exercice (3.6.5, page 140)** [*Formules équivalentes du gradient conjugué*] Démontrez que les formules suivantes constituent bel et bien une implantation de l'algorithme du gradient conjugué :

$$\begin{aligned}d_k &= -\nabla q(x_k)^t + \beta_{k-1}d_{k-1} \\ \theta_k &= \frac{\|\nabla q(x_k)\|^2}{d_k^t Q d_k} \\ x_{k+1} &= x_k + \theta_k d_k \\ \beta_k &= \frac{\|\nabla q(x_{k+1})\|^2}{\|\nabla q(x_k)\|^2}.\end{aligned}$$

**Exercice (3.6.6, page 141)** [*Programme quadratique*] Suite de l'exercice 3.2.7, page 109.

- Calculez les valeurs de  $a_0$  et  $a_3$  du modèle en avec l'algorithme du gradient conjugué.
- Calculez les valeurs de  $a_0$  et  $a_3$  du modèle en à l'aide de la décomposition de Cholesky.

**Exercice (3.7.1, page 141)** [*Normes*] Montrez qu'il est équivalent de minimiser  $\|F(x)\|$  et  $\|F(x)\|^2$ .

**Exercice (3.7.2, page 141)** [*Somme de normes*] On sait que la lumière emprunte toujours le chemin le plus court, ce qui explique que lors d'une réflexion miroir, l'angle de réflexion égale l'angle d'incidence. Cet exercice obtient le résultat dans une situation très simplifiée. On se place en dimension deux, et considère deux points  $P = (P_x, P_y)^t$  et  $Q = (Q_x, Q_y)^t$  avec  $P_y > 0$  et  $Q_y > 0$ . On cherche à identifier le point  $X = (x, 0)^t$  qui minimise la somme des distances de  $P$  à  $X$  et de  $X$  à  $Q$ . On cherche donc à résoudre  $\min_{x \in \mathbb{R}} f(x) = \|P - X\| + \|Q - X\|$  où seul  $X$  dépend de  $x$ . (voir figure 3.6).

- En utilisant les conditions d'optimalité, montrez que pour la solution optimale, l'angle entre le segment  $\overline{PX}$  et l'axe des  $x$  (thetaP) est égal à l'angle entre le segment  $\overline{QX}$  et l'axe des  $x$  (thetaQ). La fonction norme n'est pas différentiable à l'origine, mais puisque  $P_y > 0$ ,  $Q_y > 0$  et  $X_y = 0$ , la solution recherchée aura des normes strictement positives, donc la fonction  $f$  est différentiable aux points d'intérêts.
- Vérifiez qu'il n'est pas équivalent de minimiser  $\|P - X\|^2 + \|Q - X\|^2$ .

**Exercice (3.7.3, page 142)** [*Fausses racines*] Vérifiez à l'aide d'un exemple de fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  qu'il est possible que  $f^2$  possède un minimum local en  $x$  tel que  $f(x) \neq 0$ .

**Exercice (3.8.1, page 145)** [*Direction de Newton modifiée*] Considérez la fonction  $f(x) = x_1^4 + x_1x_2 + (1 + x_2)^2$ .

- Montrez qu'à partir de l'origine, ni  $d_N$  ni  $-d_N$  ne constitue une direction de descente.
- Calculez la direction de Newton modifiée à l'aide de la décomposition de Cholesky modifiée.
- Vérifiez que la direction obtenue est bien une direction de descente.

**Exercice (3.8.2, page 146)** [*Implantation de la décomposition de Cholesky modifiée*] Complétez les détails de la décomposition  $P^tQP = LDL^t + E$ ; implantez le calcul de  $L$ ,  $D$  et  $P$ , et fournissez les outils pour calculer une direction de Newton modifiée.

**\*Exercice (3.8.3, page 147)** [*Direction descendante*] Démontrez que la direction obtenue par le gradient conjugué interrompu tel que décrit ci-haut est suffisamment descendante.

**Exercice (3.8.4, page 147)** [*Direction de Newton modifiée*] Considérez à nouveau la fonction  $f(x) = x_1^4 + x_1x_2 + (1 + x_2)^2$ .

- à partir de l'origine, calculez la direction de Newton modifiée à l'aide de l'algorithme du gradient conjugué.
- Vérifiez que la direction obtenue est bien une direction de descente.

**Exercice (3.8.5, page 147)** [*Implantation de la direction à l'aide du gradient conjugué*] Implantez l'algorithme 3.4. Remarquez que la matrice  $Q$  n'est utilisée que dans des expressions produit avec un vecteur. Ainsi, il n'est pas nécessaire de connaître la matrice  $Q$ , mais il suffit d'être capable d'évaluer  $\nabla_t(\nabla_x f(x + tv)) = v^t \nabla^2 f(x + tv)$  en  $t = 0$ . On peut utiliser des différences finies, ou un outil de différentiation automatique pour cette évaluation.

**Exercice (3.9.1, page 153)** [*Équivalence des formules pour fonctions quadratiques*] Lorsque  $f$  est une fonction quadratique, il est possible de définir  $\theta$  exactement à l'aide d'une formule explicite, et nous sommes dans le royaume du gradient conjugué linéaire tel que présenté à la section 3.6.2. Dans ce cas,  $g_k \perp g_{k-1}$  et  $g_k \perp d_{k-1}$ . Utilisez ces propriétés pour démontrer que toutes ces formules sont équivalentes.

**Exercice (3.11.1, page 156)** [*Optimisation dans  $\mathbb{R}^n$* ] La fonction de Rosenbrock est fameuse pour les tests de méthodes d'optimisation dans  $\mathbb{R}^n$ . La version originale de cette fonction pour  $n = 2$  est la suivante :

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Son minimum est en  $x = (1, 1)^t$  et le point de départ habituel est  $x_0 = (-1, 1.2)^t$ .  $f(x_0) = 8$ ,  $\nabla f(x_0) = (76, 40)$  and  $\nabla^2 f(x_0) = \begin{pmatrix} 722 & 400 \\ 400 & 200 \end{pmatrix}$

- a) Vérifiez que  $x^* = (1, 1)^t$  est un minimum local satisfaisant aux conditions suffisantes d'optimalité.
- b) Vérifiez que  $\nabla^2 f(x_0)$  n'est pas définie positive.
- c) Malgré que  $\nabla^2 f(x_0)$  ne soit pas définie positive, il est possible que la direction de Newton (non modifiée) soit tout de même une direction de descente. Vérifiez que la direction de Newton est en effet une direction de descente dans ce cas particulier.
- d) Utilisez l'algorithme gradient conjugué pour calculer la direction de Newton modifiée en utilisant le test d'arrêt "courbure\_négative".
- e) Vérifiez que le pas de déplacement  $\theta = 1$  est admissible pour la direction de Newton modifiée, la direction de Newton non modifiée, mais pas pour la direction  $d = -\nabla f(x_0)$ . Prenez  $\tau_0 = 0.1$  et  $\tau_1 = 0.9$ .

**Exercice (3.11.2, page 156)** [*Projet de programmation*] Implantez quelques variantes de l'algorithme de Newton modifié : comparez les résultats d'une variante avec une recherche de pas admissible simple, l'algorithme d'Armijo qui prend la première puissance de  $\frac{1}{2}$  satisfaisant au critère d'Armijo, une recherche linéaire plus sophistiquée combinant les idées des algorithmes 2.1 et 2.4, et enfin une version avec régions de confiance. Vous pouvez utiliser la modification issue de la décomposition de Cholesky, ou encore l'algorithme du gradient conjugué (comparez le comportement de la version région de confiance avec celle de la recherche linéaire dans ce cas). Votre programme utilisera trois fonctions qui retournent respectivement  $f(x)$ ,  $\nabla f(x)$  et  $\nabla^2 f(x)$ . Testez votre programme (entre autres) avec la fonction de Rosenbrock :  $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$  à partir du point de départ  $(-1.2, 1.0)^t$ . Cette fonction a longtemps constitué un défi pour les programmes de minimisation. Les modifications de l'algorithme de Newton traitent habituellement cette fonction facilement.

## Exercices supplémentaires de la Partie II

Voici quelques exercices utilisés dans des examens au fil des années.

1. [Formulations] Soit  $f$  une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$  ; pour un point  $p \in \mathbb{R}^{n+1}$ , on désire calculer le point  $\langle\langle x, f(x) \rangle\rangle \in \mathbb{R}^{n+1}$  le plus proche de  $p$ .
  - (a) Dans le cas particulier où  $n = 1$  et la fonction  $y = f(x) = ax + b$  est linéaire, le problème consiste à calculer le point d'une droite le plus proche d'une cible. Ça ressemble au cas traité dans les notes pour une droite paramétrique. Formulez le problème de trouver le point de la droite le plus proche de l'origine  $x = 0, y = 0$ . Puisque seule  $x$  est variable, la forme du problème est  $\min_{x \in \mathbb{R}^1} \phi(x)$ .
  - (b) Obtenez l'expression pour  $\phi'(x)$  et  $\phi''(x)$ .
  - (c) Écrivez les conditions d'optimalité du problème  $\min_{x \in \mathbb{R}^1} \phi(x)$  ; insistez sur les différences entre les conditions nécessaires et les conditions suffisantes.
  - (d) En utilisant les conditions d'optimalité, obtenez une expression pour le point  $\langle\langle x, y = f(x) \rangle\rangle$  le plus proche de l'origine. Expliquez pourquoi les conditions suffisantes sont toujours satisfaites pour ce problème.
  - (e) Si la fonction  $f$  n'était pas linéaire, mais deux fois continûment différentiable, quelles seraient les conditions sur  $f$  pour que les conditions suffisantes soient satisfaites en une solution  $x^*$  ?
  - (f) formulez le problème général pour  $x \in \mathbb{R}^n$ ,  $f(x)$  quelconque (pas forcément linéaire) et la cible  $p \in \mathbb{R}^{n+1}$  (pas forcément l'origine) comme un problème de minimisation ; la seule variable sujette à minimisation est le vecteur  $x \in \mathbb{R}^n$  ; le problème prend donc la forme  $\min_{x \in \mathbb{R}^n} \psi(x)$  ; fournissez l'expression pour  $\psi(x)$ .
  - (g) Obtenez une expression pour le gradient  $\nabla \psi(x)$ .

## 2. [Optimisation générale]

Considérez les deux fonctions

$$\begin{aligned} f_1(x) &= \frac{x^3}{3} - \frac{3x^2}{2} - 4x \\ f_2(x) &= e^{f_1(x)} \end{aligned}$$

Pour chaque fonction, répondez aux questions.

- La fonction est-elle bornée inférieurement ? Si oui fournissez  $f^* = \inf f(x)$  et indiquez si  $f$  atteint son inf en fournissant  $x^*$  tel que  $f(x^*) = \min f(x)$ . Si vous répondez qu'elle n'est pas bornée ou encore n'atteint pas son inf, justifiez.
- La fonction est-elle bornée supérieurement ? Si oui fournissez  $f^* = \sup f(x)$  et indiquez si  $f$  atteint son sup en fournissant  $x^*$  tel que  $f(x^*) = \max f(x)$ . Si vous répondez qu'elle n'est pas bornée ou encore n'atteint pas son sup, justifiez.
- Calculez tous les points stationnaires de chaque fonction. Il n'est pas difficile de se convaincre que  $f_2$  possède les mêmes points stationnaires que  $f_1$  qui en possède deux.
- Caractérissez les points stationnaires en (c) (min locaux, max locaux, indéterminés).

## 3. [Optimisation en une variable]

Considérez la fonction  $h(t) = f_1(t)$  du numéro deux.

- À partir de  $x_0 = 0$ , appliquez l'algorithme `trouve_intervalle` en utilisant  $\delta = 3$ .
- À partir de l'intervalle  $[0, 6]$ , effectuez 2 itérations de l'algorithme `bisection`.
- Que se passe-t-il lorsqu'on tente d'appliquer l'algorithme de Newton à partir de  $t = 0$  ?
- Que se passe-t-il lorsqu'on tente d'appliquer l'algorithme de Newton à partir de  $t = \frac{3}{2}$  ?
- À partir de l'intervalle initial  $[0, 6]$ , appliquez l'algorithme de bisection—Newton jusqu'à ce que le calcul de Newton soit choisi ; considérez que le point courant est  $t = 0$  ; utilisez  $\gamma = 0.9$  dans l'algorithme.
- À partir de  $t = 0$  en utilisant une taille de région de confiance  $\Delta_0 = 2$ , effectuez deux itérations de l'algorithme de régions de confiance.

## 4. [Pas (longueur) de déplacement admissible]

Reprenons la fonction  $f_1$  du numéro deux. Nous renommons cette fonction  $h(\theta) = f_1(\theta)$ .

- Justifiez que la fonction pourrait être issue d'un algorithme de descente, i.e. que l'on pourrait avoir  $h(\theta) = h_{x_0,d}(\theta) = f(x_0 + \theta d)$  pour une certaine direction de descente  $d$  en un certain point  $x_0$  d'une certaine fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

- (b) Pour quel ensemble de valeurs de  $\tau_0$  le minimum local positif  $\theta^*$  satisfait-il au critère d'Armijo  $h(\theta^*) \leq h(0) + \tau_0 \theta h'(0)$  ?
- (c) Pour quel ensemble de valeurs de  $\tau_1$  le minimum local positif satisfait-il au critère de Wolfe  $h'(\theta^*) \geq \tau_1 h'(0)$  ?
- (d) Quel est l'intervalle des valeurs qui satisfont au critère d'Armijo pour  $\tau_0 = \frac{1}{4}$  ?
- (e) Quel est l'intervalle des valeurs qui satisfont au critère de Wolfe pour  $\tau_1 = \frac{9}{16}$  ?

5. [Optimisation dans  $\mathbb{R}^n$ ]

Considérez la fonction suivante.

$$f(x) = \frac{1}{2} \left( (x_1 - 1)^2 + (x_2 - 3)^2 + (x_1 x_2 - 3x_1)^2 \right).$$

Elle possède un minimum en  $x = (1, 3)^t$ . Cette fonction ressemble beaucoup à celle traitée en exercice le 19 février. Attention, cependant, il y a une différence. Une copie de la solution faite en classe sans adaptation vaudra 0.

- (a) Vérifiez que  $x^* = (1, 3)^t$  est un minimum local satisfaisant aux conditions suffisante d'optimalité.
- (b) Considérons un point de départ  $x_0 = (0, 1)^t$ . Fournissez une direction descendante en ce point  $x_0$ .

## 6. [Conditions d'optimalité]

Nous avons vu au chapitre un qu'il faut incorporer des bornes  $x \geq 0$  si cela décrit bien la situation plutôt que de tronquer à zéro les composantes négatives de la solution. Considérons la fonction 2D

$$q(x) = \frac{(x_1 + x_2 + 3)^2}{10} + \frac{(x_1 - x_2 + 1)^2}{2}.$$

- (a) Écrivez les conditions nécessaires et suffisantes d'optimalité pour le problème  $\min_{x \in \mathbb{R}^2} q(x)$ .
- (b) Justifiez que la solution est unique et obtenez-la ; vous trouverez  $x_1^* < 0$  et  $x_2^* < 0$ .
- (c) Les conditions suffisantes sont-elles satisfaites pour la solution ?
- (d) Nous allons maintenant vérifier que  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  n'est pas une solution pour le problème  $\min_{x \in \mathbb{R}^2, x \geq 0} q(x)$ . Pour ce faire, nous allons faire des exercices sur les conditions d'optimalité 1D.
  - i. Utilisez les conditions d'optimalité pour obtenir la solution  $\begin{pmatrix} 0 \\ x_2^* \end{pmatrix}$  du problème 1D  $\min_{x_2} q\begin{pmatrix} 0 \\ x_2 \end{pmatrix}$  et vérifiez que la condition suffisante est satisfaite. Est-ce que  $x_2^* \geq 0$  ? Si oui,  $\hat{x} = \begin{pmatrix} 0 \\ x_2^* \end{pmatrix}$  satisfait aux contraintes du problème 2D.



- ii. Utilisez les conditions d'optimalité pour obtenir la solution  $\begin{pmatrix} x_1^* \\ 0 \end{pmatrix}$  du problème 1D  $\min_{x_1} q\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$  et vérifiez que la condition suffisante est satisfaite. Est-ce que  $x_1^* \geq 0$ ? Si oui,  $\bar{x} = \begin{pmatrix} x_1^* \\ 0 \end{pmatrix}$  satisfait aux contraintes du problème 2D.
- iii. Une des solutions 1D,  $\hat{x}$  ou  $\bar{x}$  satisfait aux contraintes du problème. Nommez-la  $\tilde{x}$  et vérifiez que  $q\begin{pmatrix} 0 \\ 0 \end{pmatrix} > q(\tilde{x})$ , ce qui confirme que l'origine n'est pas une solution optimale. Cela confirme que tronquer à zéro les composantes négatives ne fournit généralement pas une solution optimale.

### 7. [Géométrie et conditions d'optimalité]

Considérez deux droites d'équations paramétriques  $x(t_1) = o_1 + t_1 d_1$  et  $x(t_2) = o_2 + t_2 d_2$ . On cherche à trouver les points sur les droites les plus proches, permettant de donner la distance entre les droites. On minimise donc  $f(t) = \frac{1}{2} \|x_1(t_1) - x_2(t_2)\|^2$  par rapport à  $t$ , de dimension deux et les origines  $o_1, o_2$  et directions  $d_1, d_2$  sont dans  $\mathbb{R}^n$ . On suppose que les droites ne sont pas parallèles, donc que  $d_1$  et  $d_2$  sont des vecteurs linéairement indépendants.

- (a) Écrivez sous forme vectorielle-matricielle le problème, donc le problème s'exprimera avec un vecteur  $t$  de dimension 2.
- (b) Utilisez les conditions d'optimalité pour trouver une expression pour  $t^* \in \mathbb{R}^2$ , et vérifiez que c'est bel et bien un minimum local.
- (c) Si  $o_1 = o_2$ , vérifiez que  $t^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .
- (d) Si les droites sont dans  $\mathbb{R}^2$ , vérifiez que  $f(t^*) = 0$ , i.e. deux droites non-parallèles de  $\mathbb{R}^2$  se croisent toujours. Indice : si  $A$  et  $B$  sont deux matrices inversibles,  $(AB)^{-1} = B^{-1}A^{-1}$ .

### 8. [Algorithme en une variable]

Considérez la fonction  $h(\theta) = \frac{\theta^5}{5} + \frac{\theta^4}{4}$ . Pour cet exercice, portez attention au cas où on aboutit avec une dérivée nulle soit dans `Trouve_intervalle`, soit dans `Bissection-Newton`.

- (a) Vérifiez que  $\theta = 0$  et  $\theta = -1$  sont des points stationnaires de  $h$ ; identifiez ceux qui satisfont aussi à la condition nécessaire d'ordre 2 pour un minimum local.
- (b) Utilisez la condition *nécessaire et suffisante* (d'ordre  $p$ ) pour déterminer la nature de ces points.
- (c) À partir de  $\theta_0 = 1$ , appliquez l'algorithme `Trouve_intervalle` en utilisant des sauts  $\delta = \frac{1}{2}$  (évidemment multipliés par  $-h'(\theta_0)$ ).
- (d) À partir de l'intervalle  $[-\frac{1}{2}, \frac{1}{2}]$ , appliquez l'algorithme de `Bissection-Newton` ( $\gamma = 0.9$ ) deux fois, une fois en supposant que  $\theta = -\frac{1}{2}$  est le dernier point (on tente Newton à partir de  $-\frac{1}{2}$ ) et une fois en supposant que c'est  $\theta = \frac{1}{2}$  qui est le dernier point; effectuez suffisamment d'itérations pour que le calcul de Newton soit retenu.

- (e) Si on applique l'algorithme de Newton pur à  $h'(\theta) = \theta^4 + \theta^3 = 0$  au voisinage de la racine  $\theta^* = 0$ , analysez sa vitesse de convergence.

9. [Optimisation en une variable]

Considérez la fonction  $h(\theta) = (-9\theta + 6\theta^2 - \theta^3)^2$ . Note :  $h(0) = 16$ ,  $h'(0) = -72$ ,  $h''(0) = 258$ ,  $h(2.5) = 11.39$ ,  $h'(2.5) = 15.1875$ ,  $h''(2.5) = -10.125$ .

- (a) Vérifiez que  $\theta = 1$ ,  $\theta = 3$  et  $\theta = 4$  sont tous des points stationnaires de  $h$ ; identifiez ceux qui satisfont aussi à la condition nécessaire d'ordre 2 pour un minimum local.
- (b) Parmi ces points, utilisez les conditions suffisantes d'ordre 2 pour déterminer la nature de deux de ces points. Utilisez la condition *nécessaire et suffisante* (d'ordre  $p$ ) pour déterminer la nature du point pour lequel aucune condition suffisante d'ordre 2 (ni pour un min, ni pour un max) n'est satisfaite.
- (c) Appliquez l'algorithme de bisection pour réduire l'intervalle  $[0, 2.5]$  qui contient un minimum local à un intervalle de longueur  $\leq 1^2$ ;
- (d) Effectuez *une* itération de Newton à partir de  $\theta = 0$ ; cette itération se rapproche-t-elle du minimum local en  $\theta = 1$ ? Effectuez ensuite *une* itération de Newton à partir de  $\theta = 2.5$ ; cette itération se rapproche-t-elle de  $\theta = 1$ ?
- (e) Vérifiez qu'en appliquant l'algorithme bisection–Newton (algorithme 2.4 des notes) (avec  $\gamma = 0.9$ ) à partir de l'intervalle  $[\theta_0 = 0, \theta_1 = 2.5]$  (on tente une itération de Newton à partir de  $\theta_1 = 2.5$ ), on passe à la méthode de Newton après seulement une utilisation de la bisection.

10. [Optimisation générale —  $4 \times 5 = 20$  points]

Considérez les trois fonctions

$$\begin{aligned} h_1(x) &= \frac{x^3}{3} - \frac{x^2}{2} = x^2 \left( \frac{x}{3} - \frac{1}{2} \right) \\ h_2(x) &= e^{h_1(x)} \\ h_3(x) &= (h_1(x))^2 \end{aligned}$$

Pour chacune des 3 fonctions, répondez aux questions. Remarquons que  $h_1(0) = h_1(\frac{3}{2}) = 0$ .

- (a) La fonction est-elle bornée inférieurement? Si oui fournissez  $h^* = \inf h(x)$  et indiquez si  $h$  atteint son inf en fournissant  $x^*$  tel que  $h(x^*) = \min h(x)$ . Si vous répondez qu'elle n'est pas bornée ou encore n'atteint pas son inf, justifiez.
- (b) La fonction est-elle bornée supérieurement? Si oui fournissez  $h^* = \sup h(x)$  et indiquez si  $h$  atteint son sup en fournissant  $x^*$  tel que  $h(x^*) = \max h(x)$ . Si vous répondez qu'elle n'est pas bornée ou encore n'atteint pas son sup, justifiez.

---

2. Donc  $\theta = 1$  est un minimum local et en (b), c'est l'utilisation des conditions d'optimalité qui est le plus important, pas la nature du point

- (c) Calculez tous les points stationnaires de chaque fonction. Il n'est pas difficile de se convaincre que  $h_2$  possède les mêmes points stationnaires que  $h_1$  qui en possède deux et  $h_3$  trois.
- (d) Caractérissez les points stationnaires en (c) (min locaux, max locaux, indéterminés).

11. [Algorithmes scalaires —  $5 \times 5 = 25$  points]

Considérons la fonction  $h_1$  ci haut.

- (a) Puisque  $h(1) < 0$ , justifiez qu'il existe un minimum local dans l'intervalle  $[0, \frac{3}{2}]$ .
- (b) Appliquez l'algorithme de Fibonacci pour réduire l'intervalle  $[0, \frac{3}{2}]$  à un intervalle de longueur  $\leq \frac{1}{3}$ .
- (c) Effectuez *une* itération de Newton à partir de  $\theta = 0$ ; cette itération se rapproche-t-elle du minimum local dans l'intervalle  $[0, \frac{3}{2}]$  ?
- (d) Effectuez *une* itération de Newton à partir de  $\theta = \frac{3}{2}$ ; cette itération se rapproche-t-elle du minimum local dans l'intervalle  $[0, \frac{3}{2}]$  ?
- (e) Vérifiez qu'en appliquant l'algorithme bisection–Newton (algorithme 1.4 des notes que vous avez codé dans le devoir 1) (avec  $\gamma = 0.9$ ) à partir de l'intervalle  $[\theta_1 = 0, \theta_0 = \frac{3}{2}]$  (on tente une itération de Newton à partir de  $\theta_1 = 0$ ), on passe à la méthode de Newton après seulement une utilisation de la bisection.

12. [Orthogonalité et optimalité] Ces exercices explorent les relations d'orthogonalité reliées aux notions d'optimalité.

- (a) Considérez une droite paramétrique de  $\mathbb{R}^n$ ,  $x(\theta) = o + \theta d$  ainsi qu'une cible  $x_0$ . Exprimez le problème d'identifier le point de la droite le plus proche de  $x_0$  comme un problème de minimisation. Obtenez l'expression de la solution  $\theta^*$  et vérifiez que  $(x(\theta^*) - x_0) \perp d$ .
- (b) Considérez maintenant un plan paramétrique de  $\mathbb{R}^n$ ,  $x(\theta) = o + \theta_1 d_1 + \theta_2 d_2$ . En constituant la matrice  $D = [d_1 d_2]$  et le vecteur  $\theta = (\theta_1, \theta_2)^t$  exprimez le problème d'identifier le point du plan le plus proche de  $x_0$  comme un problème de minimisation. Obtenez l'expression de la solution  $\theta^* \in \mathbb{R}^2$  et vérifiez que  $(x(\theta^*) - x_0) \perp d_1$  et  $(x(\theta^*) - x_0) \perp d_2$ .
- (c) Rappelons que l'algorithme du gradient conjugué linéaire est décrit par les équations :

$$\begin{aligned} d_k &= -\nabla q(x_k)^t + \beta_{k-1} d_{k-1} \\ \theta_k &= \frac{-\nabla q(x_k) d_k}{d_k^t Q d_k} \\ x_{k+1} &= x_k + \theta_k d_k \\ \beta_k &= \frac{\nabla q(x_{k+1}) Q d_k}{d_k^t Q d_k}. \end{aligned}$$

Démontrez que  $\nabla q(x_{k+1}) \perp d_k$ .

- (d) Dans l'algorithme du gradient conjugué, on peut montrer que  $\nabla q(x_{k+1}) \perp \nabla q(x_k)$ . Utilisez cette propriété (sans la démontrer, prenez-la pour acquis) pour montrer que les formules suivantes pour  $\theta_k$  et  $\beta_k$  sont équivalentes aux précédentes.

$$\theta_k = \frac{\|\nabla q(x_k)\|^2}{d_k^t Q d_k}$$
$$\beta_k = \frac{\|\nabla q(x_{k+1})\|^2}{\|\nabla q(x_k)\|^2}.$$

Observez que  $\theta_k Q d_k = Q x_{k+1} - Q x_k = \nabla q(x_{k+1})^t - \nabla q(x_k)^t$ . On remarque que la formule résultante pour  $\beta_k$  n'utilise plus explicitement la matrice  $Q$ , mais seulement  $\nabla q$ .