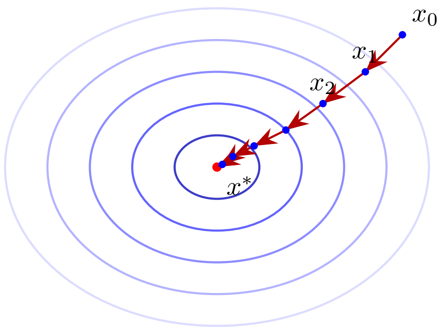
Chapitre 1.

Algorithmes d'optimisation



Descente de gradient
Minimisation de $f(x)$ par itérations successives

Considérons un problème de la forme

$$\min_{x \in \mathbb{R}^n} f(x), \tag{P}$$

avec $f : \mathbb{R}^n \to \mathbb{R}$. L'objectif est de construire une suite $(x_k)_k$ telle que $f(x_k) \to f^* := \min_{\mathbb{R}^n} f$ quand $k \to \infty$ et, si possible, $x_k \to x^* \in \operatorname{argmin}_{\mathbb{R}^n} f$ dans un sens à préciser. Une façon de générer une telle suite est de considérer

$$x_{k+1} = x_k + t_k \mathbf{d}_k,$$

où $\mathbf{d}_k \in \mathbb{R}^n$ est un vecteur de $\mathbb{R}^n$ et $t_k \geq 0$ est un pas de déplacement. Dans un premier temps, on s'intéressera donc à $(t_k, \mathbf{d}_k)$.

1.1 Directions de descente

Définition 1. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1. On dit qu'un vecteur non nul $\mathbf{d}$ de $\mathbb{R}^n$ est une direction de descente de f en x si

$$f^{'}(x; \mathbf{d}) = \nabla f(x)^T \mathbf{d} < 0.$$

Il se trouve qu'en suivant des directions de descente avec un pas suffisamment petit, on arrive à décroître la fonction objectif. On a le résultat suivant :

Lemme 1. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 et $\mathbf{d}$ une direction de descente de f en $x \in \mathbb{R}^n$. Alors, il existe $\eta > 0$ tel que

$$f(x + t\mathbf{d}) < f(x),$$

pour tout $t \in (0, \eta]$.

Preuve. Soit $x \in \mathbb{R}^n$ et $\mathbf{d}$ une direction de descente en x. On a $f(x+t\mathbf{d}) = f(x) + t\langle \nabla f(x), \mathbf{d}\rangle + o(t\|\mathbf{d}\|)$. On sait que pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $\frac{|o(t\|\mathbf{d}\|)|}{t} \leq \varepsilon$ pour tout $t \in (0, \eta]$. En prenant $\varepsilon = -\langle \nabla f(x), \mathbf{d}\rangle$ qui est strictement positif puisque $\mathbf{d}$ est une direction de descente en x, on obtient que

$$o(t\|\mathbf{d}\|)/t \leq |o(t\|\mathbf{d}\|)|/t \leq -\langle \nabla f(x), \mathbf{d}\rangle,$$

donc $f(x + t\mathbf{d}) < f(x), \forall t \in (0, \eta]$. □

D'une façon générale, une méthode de descente s'écrit comme suit :

Méthode de descente

- Initialisation : On se donne un $x_0 \in \mathbb{R}^n$.
- Pour $k \geq 0$
  - Choisir une direction de descente $\mathbf{d}_k$.
  - Choisir un pas t_k tel que : $f(x_k + t_k\mathbf{d}_k) < f(x_k)$.
  - Prendre $x_{k+1} = x_k + t_k\mathbf{d}_k$.
  - Si un critère d'arrêt est vérifié, renvoyer x_{k+1}.

Remarque 1.

- Le choix de la direction de descente donne lieu à des méthodes différentes de descente (e.g., gradient, gradient conjugué, Newton, quasi-Newton etc).
- Il existe plusieurs critères d'arrêt ; le plus utilisé est $\|\nabla f(x_{k+1})\| \leq \varepsilon$ pour une tolérance $\varepsilon > 0$ donnée.
- Le procédé permettant de trouver le pas t_k s'appelle recherche linéaire (line search en anglais). L'appellation vient du fait que trouver un tel t_k revient à minimiser la restriction de f sur la ligne $x_k + t\mathbf{d}_k$, i.e., la fonction $g(t) = f(x_k + t\mathbf{d}_k)$. On discute ici quelques choix populaires.

  1. **Pas constant** : dans ce cas $t_k = \bar{t}$ pour tout k. Bien que ce soit un choix simple, un pas trop petit peut donner lieu à une convergence lente de l'algorithme tandis qu'avec un pas trop grand on peut perdre la décroissance de la fonction objectif (cf. TP).
  2. **Recherche linéaire exacte** : dans ce cas on cherche à minimiser exactement la fonction $g(t)$ définie ci-dessus. On choisit donc $t_k \in \operatorname{argmin}_{t\geq 0} f(x_k + t\mathbf{d}_k)$. Il se trouve qu'en général, on ne peut pas trouver le minimiseur exact de g.
  3. **Rebroussement (ou Backtracking)** : cette approche permet d'éviter une recherche linéaire exacte en choisissant convenablement un pas assurant la décroissance de la fonction objectif. L'approche requiert trois paramètres $s > 0$ et $\alpha, \beta \in (0, 1)$. Dans un premier temps on prend $t_k = s$ et tant que

     $$f(x_k) - f(x_k + t_k\mathbf{d}_k) < -\alpha t_k \nabla f(x_k)^T \mathbf{d}_k,$$

     (i.e., $f(x_{k+1}) > f(x_k)$) on met à jour t_k par $t_k\beta$. Dans ce cas, le pas choisi serait $t_k = s\beta^{i_k}$ où i_k est le plus petit entier tel que

     $$f(x_k) - f(x_k + t_k\mathbf{d}_k) \geq -\alpha s\beta^{i_k} \nabla f(x_k)^T \mathbf{d}_k. \tag{1.1}$$

     Comme $\mathbf{d}_k$ est une direction de descente de f en x_k, l'inégalité (1.1) assure la décroissance de la suite $(f(x_k))_k$.

Il s'avère que (1.1) est toujours vérifiée pour des pas t_k suffisamment petits.

Lemme 2. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 et $\mathbf{d}$ une direction de descente de f en $x \in \mathbb{R}^n$, et soit $\alpha \in (0, 1)$. Alors, il existe $\eta > 0$ tel que

$$-\alpha t\langle \nabla f(x), \mathbf{d}\rangle < f(x) - f(x + t\mathbf{d}),$$

pour tout $t \in (0, \eta]$.

Preuve. Soit $x \in \mathbb{R}^n$ et $\mathbf{d}$ une direction de descente en x. On a $f(x+t\mathbf{d}) = f(x)+t\langle \nabla f(x), \mathbf{d}\rangle + o(t\|\mathbf{d}\|)$. Par définition :

$$\forall \varepsilon > 0, \exists \eta > 0 : \frac{|o(t\|\mathbf{d}\|)|}{t} \leq \varepsilon, \forall t \in (0, \eta].$$

En prenant $\varepsilon = -(\alpha-1)\langle \nabla f(x), \mathbf{d}\rangle$ qui est strictement positif puisque $\mathbf{d}$ est une direction de descente en x, on obtient que

$$o(t\|\mathbf{d}\|)/t \leq |o(t\|\mathbf{d}\|)|/t \leq (\alpha-1)\langle \nabla f(x), \mathbf{d}\rangle,$$

soit

$$o(t\|\mathbf{d}\|) + t\langle \nabla f(x), \mathbf{d}\rangle < \alpha t\langle \nabla f(x), \mathbf{d}\rangle$$

donc $-\alpha t\langle \nabla f(x), \mathbf{d}\rangle < f(x) - f(x+t\mathbf{d}), \forall t \in (0, \eta]$. □

L'exemple suivant présente le cas d'une recherche linéaire exacte pour une fonction quadratique.

Exemple 1. On considère $f(x) = x^T Ax + 2b^T x + c$ avec $A \in \mathcal{S}_n^{++}, b \in \mathbb{R}^n, c \in \mathbb{R}$. Soit $x \in \mathbb{R}^n$ et $\mathbf{d} \in \mathbb{R}^n \setminus \{0\}$ une direction de descente de f en x. Effectuer une recherche linéaire exacte revient à résoudre

$$\min_{t \geq 0} f(x + t\mathbf{d}) := g(t).$$

On a $g'(t) = \mathbf{d}^T \nabla f(x + t\mathbf{d})$ et comme $\nabla f(x) = 2(Ax + b)$, on trouve que $g'(t) = 2(\mathbf{d}^T A\mathbf{d})t + 2\mathbf{d}^T \nabla f(x)$. Donc

$$g^{'}(t) = 0 \Leftrightarrow t = -\frac{\mathbf{d}^T \nabla f(x)}{2\mathbf{d}^T A\mathbf{d}} > 0.$$

Comme $g^{''}(t) = 2(\mathbf{d}^T A\mathbf{d}) > 0$ alors le pas donné par la méthode de recherche linéaire exacte est $t^* = -\frac{\mathbf{d}^T \nabla f(x)}{2\mathbf{d}^T A\mathbf{d}}$.

1.2 Descente de gradient

Comme son nom l'indique, le choix de la direction de descente dans la méthode du gradient est $\mathbf{d}_k = -\nabla f(x_k)$. Ce choix est bien une direction de descente au sens de la Définition 1. En effet, pour $\nabla f(x_k) \neq 0$:

$$f^{'}(x_k; -\nabla f(x_k)) = -\nabla f(x_k)^T \nabla f(x_k) = -\|\nabla f(x_k)\|^2 < 0.$$

Au-delà du fait que $\mathbf{d} = -\nabla f(x_k)$ est une direction de descente, le vecteur unitaire $\mathbf{d}^* = \mathbf{d}/\|\mathbf{d}\|$ est celui de la plus forte pente, i.e., il minimise la dérivée directionnelle $f'(x_k; \mathbf{d})$ parmi tous les vecteurs unitaires $\mathbf{d}$. Plus formellement, on a le résultat suivant.

Proposition 1. Soit $f \in C^1(\mathbb{R}^n)$ et $x \in \mathbb{R}^n$ tel que $\nabla f(x) \neq 0$. Alors une solution du problème

$$\min_{\mathbf{d} \in \mathbb{R}^n:\ \|\mathbf{d}\|=1} f'(x; \mathbf{d}), \tag{1.2}$$

est $\mathbf{d}^* = -\nabla f(x)/\|\nabla f(x)\|$.

Preuve. Le problème (1.2) s'écrit aussi

$$\min_{\mathbf{d}:\|\mathbf{d}\|=1} \langle \nabla f(x), \mathbf{d}\rangle,$$

et grâce à l'inégalité de Cauchy-Schwarz, on a

$$\langle \nabla f(x), \mathbf{d}\rangle \geq -\|\mathbf{d}\|\|\nabla f(x)\| = -\|\nabla f(x)\|,$$

donc $-\|\nabla f(x)\|$ est une borne inférieure de la valeur optimale de (1.2). Cette valeur est atteinte en particulier en $\mathbf{d}^*$. En effet

$$f'(x;\mathbf{d}^*) = \langle \nabla f(x), \frac{-\nabla f(x)}{\|\nabla f(x)\|}\rangle = -\|\nabla f(x)\|,$$

d'où le résultat. □

On présente ci-dessous les étapes de la méthode, avec comme critère d'arrêt $\|\nabla f(x_{k+1})\| \leq \varepsilon$ pour une tolérance ε donnée.

Méthode du gradient

- Initialisation : On se donne un $x_0 \in \mathbb{R}^n$.
- Pour $k \geq 0$
  - Choisir un pas t_k par recherche linéaire sur $g(t) = f(x_k - t\nabla f(x_k))$.
  - Prendre $x_{k+1} = x_k - t_k \nabla f(x_k)$.
  - Si $\|\nabla f(x_{k+1})\| \leq \varepsilon$, s'arrêter et renvoyer x_{k+1}.

1.3 Analyse de convergence

Avant de commencer notre analyse, rappelons qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dans $C_L^{1,1}(\mathbb{R}^n)$ si elle est de classe C^1 et que son gradient $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ est L-Lipschitz, i.e.,

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x-y\| \text{ pour tout } x, y \in \mathbb{R}^n.$$

Exemple 2.
- On vérifie facilement que $f(x) = a^T x + b \in C_0^{1,1}(\mathbb{R}^n)$.
- Une fonction quadratique $f(x) = x^T Ax + 2b^T x + c$ avec $A \in \mathcal{S}_n, b \in \mathbb{R}^n$ et $c \in \mathbb{R}$ est $2\|A\|$-Lipschitz. En effet, comme $\nabla f(x) = 2Ax + 2b$ on a

$$\|\nabla f(x) - \nabla f(y)\| = 2\|(Ax+b) - (Ay+b)\| \leq 2\|A\|\|x-y\|.$$

Si la fonction est C^2, alors la régularité Lipschitz du gradient est équivalente au fait que la Hessienne est bornée :

Théorème 1. Soit $f \in C^2(\mathbb{R}^n)$. Alors

$$f \in C_L^{1,1}(\mathbb{R}^n) \Leftrightarrow \|\nabla^2 f(x)\| \leq L \text{ pour tout } x \in \mathbb{R}^n.$$

Le résultat suivant joue un rôle important dans la suite.

Lemme 3 (Lemme de descente). Soit $f \in C_L^{1,1}(\mathbb{R}^n)$. Alors pour tout $x, y \in \mathbb{R}^n$

$$f(y) \leq f(x) + \nabla f(x)^T (y-x) + \frac{L}{2}\|x-y\|^2.$$

Preuve. Soient $x, y \in \mathbb{R}^n$. Par le théorème fondamental de l'analyse

$$f(y) = f(x) + \int_0^1 \langle \nabla f(x + t(y-x)), y-x\rangle \mathbf{d}t,$$

donc

$$f(y) - f(x) = \langle \nabla f(x), y-x\rangle + \int_0^1 \langle \nabla f(x+t(y-x)) - \nabla f(x), y-x\rangle \mathbf{d}t.$$

Il s'ensuit que

$$\begin{aligned}\Big|f(y)-f(x)-\langle\nabla f(x),y-x\rangle\Big| &= \Big|\int_0^1 \langle\nabla f(x+t(y-x))-\nabla f(x),y-x\rangle \mathbf{d}t\Big|,\\ &\le \int_0^1 \Big|\langle\nabla f(x+t(y-x))-\nabla f(x),y-x\rangle\Big|\mathbf{d}t,\\ &\underset{\le}{\text{C.S}} \int_0^1 \|\nabla f(x+t(y-x))-\nabla f(x)\|\|y-x\|\mathbf{d}t,\\ &\underset{\le}{f\in C_L^{1,1}} \int_0^1 Lt\|y-x\|^2\mathbf{d}t = \frac{L}{2}\|x-y\|^2.\end{aligned} \tag{1.3}$$

□

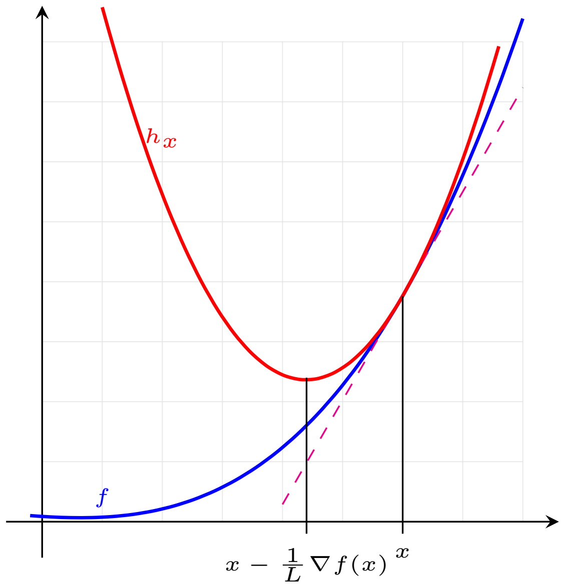


Figure 1.1: La fonction quadratique h_x majore f et la touche en x.

Fixons $x \in \mathbb{R}^n$ et considérons $h_x(y) = f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|x-y\|^2$ qui est quadratique en y et majore f : $f(y) \le h_x(y)$ pour tout $y \in \mathbb{R}^n$. De plus, le minimum de h_x est atteint en $x^* = x - \frac{1}{L}\nabla f(x)$ (voir Fig. 1.1).

Comme conséquence du Lemme 3, on a

Lemme 4. Soit $f \in C_L^{1,1}(\mathbb{R}^n)$. Alors pour tout $x \in \mathbb{R}^n$ et $t > 0$

$$f(x) - f(x - t\nabla f(x)) \ge t\left(1 - \frac{tL}{2}\right)\|\nabla f(x)\|^2. \tag{1.4}$$

Preuve. On applique le Lemme de descente avec $y = x - t\nabla f(x)$ pour un $x \in \mathbb{R}^n$. On a donc

$$f(x - t\nabla f(x)) \le f(x) - t\|\nabla f(x)\|^2 + \frac{Lt^2}{2}\|\nabla f(x)\|^2,$$

d'où le résultat en réarrangeant. □

Si nous optons pour une stratégie avec un pas constant, i.e., $t_k = t^* \in (0, 2/L)$ pour tout k, on a alors d'après le Lemme 4

$$f(x_k) - f(x_{k+1}) \ge t^*\left(1 - \frac{t^*L}{2}\right)\|\nabla f(x_k)\|^2.$$

La fonction $t \mapsto t\left(1 - \frac{tL}{2}\right)$ sur $(0, 2/L)$ atteint un maximum en $t^* = 1/L$. Pour ce choix on a $x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k)$ et

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2L}\|\nabla f(x_k)\|^2.$$

Pour une recherche linéaire exacte, i.e., $x_{k+1} = x_k - t_k \nabla f(x_k)$ avec

$$t_k \in \operatorname*{argmin}_{t\geq 0} f\Big(x_k - t\nabla f(x_k)\Big),$$

on remarque que $f(x_k - t_k\nabla f(x_k)) \leq f\Big(x_k - \frac{1}{L}\nabla f(x_k)\Big)$ soit

$$f(x_k) - f(x_{k+1}) \geq f(x_k) - f\Big(x_k - \frac{1}{L}\nabla f(x_k)\Big) \geq \frac{1}{2L}\|\nabla f(x_k)\|^2.$$

Dans le cadre de backtracking, on cherche pour $\alpha \in (0,1)$, un pas t_k suffisamment petit tel que

$$f(x_k) - f(x_{k+1}) \geq \alpha t_k \|\nabla f(x_k)\|^2.$$

Dans ce cas deux options se présentent : prendre $t_k = s$ où $s > 0$ est la valeur initiale du pas, soit prendre t_k avec la méthode de backtracking comme décrit dans la Remarque 1, et dans ce cas, le choix t_k/β ne serait pas admissible, i.e.,

$$f(x_k) - f\Big(x_k - \frac{t_k}{\beta}\nabla f(x_k)\Big) < \frac{\alpha t_k}{\beta}\|\nabla f(x_k)\|^2. \tag{1.5}$$

En appliquant (1.4) avec $x = x_k$ et $t = \frac{t_k}{\beta}$ on obtient

$$f(x_k) - f\Big(x_k - \frac{t_k}{\beta}\nabla f(x_k)\Big) \geq \frac{t_k}{\beta}\left(1 - \frac{t_k L}{2\beta}\right)\|\nabla f(x_k)\|^2. \tag{1.6}$$

Les équations (1.5)–(1.6) impliquent que $\frac{t_k}{\beta}\left(1 - \frac{t_k L}{2\beta}\right) < \frac{\alpha t_k}{\beta}$, soit $t_k > \frac{2(\alpha-1)\beta}{L}$. Finalement, pour la méthode de backtracking, on a

$$f(x_k) - f(x_k - t_k\nabla f(x_k)) \geq \alpha \min\left(s, \frac{2(\alpha-1)\beta}{L}\right)\|\nabla f(x_k)\|^2.$$

On récapitule cette discussion dans le résultat suivant :

Proposition 2. Soit $f \in C_L^{1,1}(\mathbb{R}^n)$ et $(x_k)_k$ la suite générée par la méthode du gradient pour résoudre

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

avec pas constant, recherche linéaire exacte ou backtracking de paramètres $s > 0, \alpha, \beta \in (0,1)$. Alors

$$f(x_k) - f(x_{k+1}) \geq M\|\nabla f(x_k)\|^2, \tag{1.7}$$

avec

$$M = \begin{cases} t^*\left(1 - \frac{t^* L}{2}\right) & \text{pas constant,} \\ \frac{1}{2L} & \text{recherche linéaire exacte,} \\ \alpha \min\left(s, \frac{2(\alpha-1)\beta}{L}\right) & \text{backtracking.} \end{cases}$$

On démontre maintenant que $\|\nabla f(x_k)\| \to 0$ quand $k \to \infty$.

Théorème 2. Soit $f \in C_L^{1,1}(\mathbb{R}^n)$ et $(x_k)_k$ la suite générée par la méthode du gradient pour résoudre

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

avec pas constant, recherche linéaire exacte ou backtracking de paramètres $s > 0, \alpha, \beta \in (0,1)$. Supposons que f est minorée, i.e., il existe $c \in \mathbb{R}$ tel que $f(x) \geq c$ pour tout $x \in \mathbb{R}^n$. Alors

1. La suite $(f(x_k))_k$ est décroissante, avec $f(x_{k+1}) < f(x_k)$ sauf si $\|\nabla f(x_k)\| = 0$.
2. $\|\nabla f(x_k)\| \to 0$ quand $k \to \infty$.

Preuve. 1. On a grâce à (1.7)

$$f(x_k) - f(x_{k+1}) \geq M\|\nabla f(x_k)\|^2 \geq 0, \tag{1.8}$$

pour $M > 0$. Donc $(f(x_k))_k$ est décroissante et $f(x_k) = f(x_{k+1})$ ne peut avoir lieu que si $\nabla f(x_k) = 0$.

2. Comme la suite $(f(x_k))_k$ est décroissante et f minorée, alors elle est convergente, i.e., $f(x_k) - f(x_{k+1}) \underset{k\to\infty}{\to} 0$. Cela implique avec (1.8) que $\|\nabla f(x_k)\| \to 0$ quand $k \to \infty$.

□

On termine cette section avec le résultat suivant permettant d'avoir des taux de convergence de la norme du gradient.

Théorème 3. Sous les mêmes hypothèses, on a pour tout $n \geq 0$

$$\min_{k=0,\dots,n} \|\nabla f(x_k)\| \leq \sqrt{\frac{f(x_0) - f^*}{M(n+1)}}, \tag{1.9}$$

où f^* est la limite de la suite $(f(x_k))_k$ et la constante M est donnée dans Proposition 2.

Preuve. En sommant de 0 à n dans l'inégalité (1.7), on obtient

$$f(x_0) - f(x_{n+1}) = \sum_{k=0}^{n}(f(x_k) - f(x_{k+1})) \geq M \sum_{k=0}^{n} \|\nabla f(x_k)\|^2.$$

Par définition, $f^* \leq f(x_{n+1})$, et par suite

$$f(x_0) - f^* \geq M \sum_{k=0}^{n} \|\nabla f(x_k)\|^2 \geq M \min_{k=0,\dots,n} \|\nabla f(x_k)\|^2 \sum_{k=0}^{n} 1,$$

soit

$$f(x_0) - f^* \geq M(n+1) \min_{k=0,\dots,n} \|\nabla f(x_k)\|^2.$$

Finalement $\min_{k=0,\dots,n} \|\nabla f(x_k)\| \leq \sqrt{\frac{f(x_0)-f^*}{M(n+1)}}$, comme voulu. □

Remarque 2. On note $\nabla f_n^* = \min_{k=0,\dots,n} \|\nabla f(x_k)\|$, on vient de démontrer que $\nabla f_n^* = \mathcal{O}(1/\sqrt{n})$. Au-delà du fait qu'il s'agisse d'un taux de convergence peu satisfaisant, le résultat ne donne aucune information sur les quantités $f(x_k) - f^*$ et $\|x_k - x^*\|$. Nous verrons plus tard que dans le cas convexe, le taux démontré précédemment peut être amélioré en $\mathcal{O}(1/n)$ et des taux de convergence de la fonction objectif et des itérées $(x_k)_k$ peuvent être démontrés.

1.4 Méthode de Newton

Avant de présenter la méthode de Newton comme une méthode de descente, rappelons la méthode de Newton pour trouver les zéros d'une fonction réelle. Soit donc $f : \mathbb{R} \to \mathbb{R}$, on cherche à résoudre l'équation $f(t) = 0$. L'idée est de remplacer f par son approximation de premier ordre en un point t_0, i.e., résoudre $\tilde{f}(t) = f(t_0) + (t - t_0)f'(t_0) = 0$ au lieu de $f(t) = 0$. Si $f'(t_0) \neq 0$, on obtient $t = t_0 - \frac{f(t_0)}{f'(t_0)}$. Les itérations de Newton s'écrivent

$$t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}.$$

Maintenant, soit $g : \mathbb{R}^n \to \mathbb{R}^n$. On s'intéresse au système $g(x) = 0$. De la même manière, en approchant g à l'ordre un autour d'un point $x_0 \in \mathbb{R}^n$, i.e., $g(x_0) + \nabla g(x_0)(x - x_0) = 0$. On obtient

$$\tilde{g}(x) := g(x_0) + \nabla g(x_0)(x - x_0) = 0 \Rightarrow x = x_0 - \Big(\nabla g(x_0)\Big)^{-1} g(x_0).$$

L'équation ci-dessus suppose évidemment que $\det\Big(\nabla g(x_0)\Big) \neq 0$. De même, on peut construire une suite $(x_k)_k$ qui approchera, a priori, le zéro de g, en définissant

$$x_{k+1} = x_k - \Big(\nabla g(x_k)\Big)^{-1} g(x_k). \tag{1.10}$$

Maintenant, revenons à notre problème d'optimisation

$$\min_{x \in \mathbb{R}^n} \; f(x). \tag{1.11}$$

On s'intéresse ici au système $\nabla f(x) = 0$. Rappelons que lorsque f est convexe, la condition $\nabla f(x^*) = 0$ est nécessaire et suffisante pour dire que x^* minimise f. On applique (1.10) avec $g = \nabla f$, pour trouver

$$x_{k+1} = x_k - \Big(\nabla^2 f(x_k)\Big)^{-1} \nabla f(x_k). \tag{1.12}$$

Ceci est une façon de faire le lien entre Newton et Newton...

Une autre façon de retrouver (1.12) est la suivante. Supposons que f est C^2 et que $\nabla^2 f(x) \in \mathcal{S}^n_{++}$ pour tout $x \in \mathbb{R}^n$. On cherchera donc à minimiser l'approximation d'ordre deux de f autour d'un certain $x_0 \in \mathbb{R}^n$, i.e., remplacer f par

$$\tilde{f}(x) = f(x_0) + (x - x_0)^T \nabla f(x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0) \tag{1.13}$$

dans le problème (1.11). Donc étant donnée une itérée x_k, x_{k+1} est obtenue comme suit

$$x_{k+1} = \operatorname*{argmin}_{x \in \mathbb{R}^n} \Big\{ f(x_k) + (x - x_k)^T \nabla f(x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k) \Big\}. \tag{1.14}$$

L'unique minimiseur de (1.14) est en fait l'unique point stationnaire, et on a

$$\nabla \tilde{f}(x_{k+1}) = 0 \Leftrightarrow \nabla f(x_k) + \nabla^2 f(x_k)(x_{k+1} - x_k) = 0,$$

ce qui donne

$$x_{k+1} = x_k - \Big(\nabla^2 f(x_k)\Big)^{-1} \nabla f(x_k). \tag{1.15}$$

La méthode de Newton est donc une méthode de descente avec comme direction $\mathbf{d}_k = -\Big(\nabla^2 f(x_k)\Big)^{-1} \nabla f(x_k)$.

Méthode de Newton

- Initialisation : On se donne un $x_0 \in \mathbb{R}^n$ et une tolérance $\varepsilon > 0$.
- Pour $k \geq 0$
  - On calcule la direction de Newton $\mathbf{d}_k$ comme solution du système linéaire
    $$\nabla^2 f(x_k)\mathbf{d}_k = -\nabla f(x_k).$$
  - Prendre $x_{k+1} = x_k + \mathbf{d}_k$.
  - Si $\|\nabla f(x_{k+1})\| \leq \varepsilon$, s'arrêter et renvoyer x_{k+1}.

Il se trouve que sous certaines conditions de régularité sur la Hessienne, on peut obtenir autour de la solution optimale un taux de convergence quadratique.

Théorème 4. Soit $f \in C^2(\mathbb{R}^n)$ et supposons que :

1. il existe $m > 0$ tel que $\nabla^2 f(x) \succeq mI_n$ pour tout $x \in \mathbb{R}^n$,
2. il existe $L > 0$ tel que
$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\|$$
pour tout $x, y \in \mathbb{R}^n$.

Alors
$$\|x_{k+1} - x^*\| \leq \frac{L}{2m}\|x_k - x^*\|^2,$$
où $(x_k)_k$ est la suite générée par la méthode de Newton et x^* l'unique minimiseur de f sur $\mathbb{R}^n$. De plus, si $\|x_0 - x^*\| \leq m/L$, alors
$$\|x_k - x^*\| \leq \frac{2m}{L}\left(\frac{1}{2}\right)^{2^k}, \quad k \geq 0.$$

Preuve. On a pour tout $k \geq 0$
$$\begin{aligned}
x_{k+1} - x^* &= x_k - \left(\nabla^2 f\left(x_k\right)\right)^{-1} \nabla f(x_k) - x^* \\
&= x_k - x^* + \left(\nabla^2 f(x_k)\right)^{-1} \left(\nabla f(x^*) - \nabla f(x_k)\right) \\
&= x_k - x^* + \left(\nabla^2 f\left(x_k\right)\right)^{-1} \int_0^1 \left[\nabla^2 f\left(x_k + t\left(x^* - x_k\right)\right)\right] \left(x^* - x_k\right) \mathbf{d}t \qquad (1.16) \\
&= \left(\nabla^2 f\left(x_k\right)\right)^{-1} \int_0^1 \left[\nabla^2 f\left(x_k + t\left(x^* - x_k\right)\right) - \nabla^2 f(x_k)\right] \left(x^* - x_k\right)\mathbf{d}t.
\end{aligned}$$

Comme $\nabla^2 f(x) \succeq mI_n$, il vient que $\| \left(\nabla^2 f\left(x_k\right)\right)^{-1} \| \leq 1/m$. En utilisant la Lipschitz-continuité de la Hessienne, on obtient que
$$\begin{aligned}
\|x_{k+1} - x^*\| &\leq \| \left(\nabla^2 f\left(x_k\right)\right)^{-1} \| \left\| \int_0^1 \left[\nabla^2 f\left(x_k + t\left(x^* - x_k\right)\right) - \nabla^2 f(x_k)\right] \left(x^* - x_k\right)\mathbf{d}t \right\| \\
&\leq \frac{1}{m} \int_0^1 \|\nabla^2 f\left(x_k + t\left(x^* - x_k\right)\right) - \nabla^2 f(x_k)\| \|x^* - x_k\| \mathbf{d}t \qquad (1.17) \\
&\leq \frac{1}{m} \int_0^1 Lt\|x^* - x_k\|^2 \mathbf{d}t = \frac{L}{2m}\|x^* - x_k\|^2.
\end{aligned}$$

On conclut par récurrence. □

1.5 Méthode de gradient projeté

1.5.1 Cas général

On s'intéresse ici à des problèmes du type
$$\min_{x \in C} f(x), \qquad (1.18)$$
où $C \subset \mathbb{R}^n$ est un convexe fermé et $f : C \to \mathbb{R}$ de classe C^1. Il se trouve qu'on peut caractériser les points stationnaires de (1.18) en termes de l'opérateur de projection orthogonale.

Théorème 5. Soit $f : C \to \mathbb{R}$ de classe C^1 où C est un convexe fermé et $s > 0$. Alors x^* est un point stationnaire de (1.18) si et seulement si

$$x^* = P_C(x^* - s\nabla f(x^*)). \tag{1.19}$$

Preuve. Laissée en exercice. □

L'équation (1.19) affirme que x^* est point fixe de l'application $x \mapsto P_C(x - s\nabla f(x))$. Cela suggère des algorithmes du type point fixe pour résoudre (1.19), et donc (1.18).

Méthode du gradient projeté
On se donne une tolérance $\varepsilon > 0$.

- Initialisation : On se donne un $x_0 \in C$.
- Pour $k \geq 0$
  - Choisir un pas t_k par recherche linéaire.
  - Prendre $x_{k+1} = P_C(x_k - t_k\nabla f(x_k))$.
  - Si $\|x_k - x_{k+1}\| \leq \varepsilon$, renvoyer x_{k+1}.

On démontre un résultat similaire au Lemme 4.

Lemme 5. Soit $f \in C_L^{1,1}(\mathbb{R}^n)$. Alors pour tout $x \in \mathbb{R}^n$ et $t > 0$

$$f(x) - f(P_C(x - t\nabla f(x))) \geq t\left(1 - \frac{tL}{2}\right)\left\|\frac{1}{t}(x - P_C(x - t\nabla f(x)))\right\|^2. \tag{1.20}$$

Preuve. Il suffit d'appliquer le lemme de descente Lemme 3 avec $y = P_C(x - t\nabla f(x))$ et utiliser le fait que

$$\langle x - t\nabla f(x) - y, x - y\rangle \leq 0.$$

□

Remarque 3. Quand $C = \mathbb{R}^n$, la méthode du gradient projeté n'est rien d'autre que la méthode de descente de gradient classique présentée en Section 1.2.

Notation. Par la suite, on note, pour $M > 0$

$$G_M(x) = M\left(x - P_C\left(x - \frac{1}{M}\nabla f(x)\right)\right).$$

On note que pour $C = \mathbb{R}^n$, $G_M(x) = \nabla f(x)$. L'opérateur servira comme une mesure d'optimalité. Avec cette notation, le Lemme 5 se réécrit

$$f(x) - f(P_C(x - t\nabla f(x))) \geq t\left(1 - \frac{tL}{2}\right)\|G_{1/t}(x)\|^2.$$

On peut démontrer sans difficultés un résultat similaire à la Proposition 2 et au Théorème 2 pour la méthode du gradient projeté ; on changera $\nabla f(x)$ par $G_M(x)$.

1.5.2 Cas convexe

On s'intéresse ici à des problèmes du type

$$\min_{x \in C} f(x), \tag{1.21}$$

où $C \subset \mathbb{R}^n$ est un convexe fermé et $f \in C_L^{1,1}(C)$ est convexe. Dans ce cas, on peut démontrer à la fois la convergence de $(f(x_k))_k$ vers f^* ainsi que celle de la suite $(x_k)_k$.

Théorème 6. Soit $f \in C_L^{1,1}(C)$ une fonction convexe avec C un convexe fermé de $\mathbb{R}^n$. Soit $(x_k)_k$ la suite générée par la méthode du gradient projeté avec un pas $t_k = t^* \in (0, 1/L]$. Supposons que $\operatorname{argmin}_C f \neq \emptyset$. Alors

1. Pour tout $k \geq 0$ et $x^* \in \operatorname{argmin}_C f$
$$f(x_k) - f^* \leq \frac{\|x_0 - x^*\|^2}{2t^*k}.$$
2. $x_k \to x^* \in \operatorname{argmin}_C f$ lorsque $k \to \infty$.

Remarque 4.
- **Amélioration du taux de convergence :** Dans le cas convexe, on obtient un taux de convergence $f(x_k) - f^* = \mathcal{O}(1/k)$ pour la fonction objectif, ce qui est significativement meilleur que le taux $\|\nabla f(x_k)\| = \mathcal{O}(1/\sqrt{k})$ obtenu dans le cas général non-convexe (voir Théorème 2). De plus, on obtient ici une information directe sur la décroissance de la fonction objectif, et non seulement sur son gradient.
- **Choix optimal du pas :** Le pas optimal est $t^* = 1/L$, qui maximise la vitesse de convergence tout en garantissant la décroissance. Ce choix donne
$$f(x_k) - f^* \leq \frac{L\|x_0 - x^*\|^2}{2k}.$$
Notons que contrairement au cas non contraint où on pouvait prendre $t^* \in (0, 2/L)$, ici la contrainte $t^* \leq 1/L$ assure que la projection reste bien définie et que l'algorithme converge.
- **Convergence des itérées :** Le point (2) du théorème garantit la convergence de la suite $(x_k)_k$ elle-même vers un minimiseur, et pas seulement la convergence de $(f(x_k))_k$ vers f^*. Ceci est une conséquence directe de la convexité de f et de la structure du problème contraint. Si $\operatorname{argmin}_C f$ est un singleton, alors la convergence est vers l'unique minimiseur. Sinon, la suite $(x_k)_k$ converge vers un élément de $\operatorname{argmin}_C f$, qui dépend du point initial x_0.
- **Comparaison avec le cas non contraint :** Lorsque $C = \mathbb{R}^n$, la méthode du gradient projeté se réduit à la méthode du gradient classique, et on retrouve $G_M(x) = \nabla f(x)$. Le théorème ci-dessus généralise donc les résultats de convergence de la Section 1.2 au cas contraint, avec les mêmes garanties de convergence sous l'hypothèse de convexité.
- **Dépendance en la condition initiale :** La borne sur $f(x_k) - f^*$ dépend de $\|x_0 - x^*\|^2$, c'est-à-dire de la distance entre le point initial et un minimiseur. Cela signifie qu'un bon choix de x_0 (proche d'un minimiseur) peut accélérer significativement la convergence. En pratique, on ne connaît évidemment pas x^*, mais cette borne théorique donne une estimation du nombre d'itérations nécessaires pour atteindre une précision donnée.
- **Complexité algorithmique :** Pour obtenir une solution ε-optimale, i.e., $f(x_k) - f^* \leq \varepsilon$, il suffit de prendre
$$k \geq \frac{\|x_0 - x^*\|^2}{2t^*\varepsilon} = \frac{L\|x_0 - x^*\|^2}{2\varepsilon}.$$
Le nombre d'itérations nécessaires est donc linéaire en $1/\varepsilon$, ce qui caractérise les méthodes de gradient du premier ordre pour les fonctions convexes lisses.