Optimisation convexe

Ch. Dossal

Janvier 2017

Introduction

Ce document est un support pour le cours d'optimisation. Il n'a donc pas vocation à être complet mais vient en appui aux séances de cours.
L'objet de ce cours est de présenter des outils théoriques d'optimisation convexe ainsi que des algorithmes développés ces dernières années pour résoudre des problèmes liés plus particulièrement au traitement des images et des statistiques. Comme ce cours traite à la fois de la théorie de l'optimisation convexe et des algorithmes plus spécifiquement des algorithmes proximaux de spliting, il ne pourra pas être exhaustif. Certains choix sont faits et certaines notions ne sont donc pas approfondies et certains algorithmes ne sont pas traités. L'objectif de ce cours aura été rempli s'il permet aux personnes qui le suivent d'avoir les outils pour appronfondir seules les notions abordées et d'être capable de mettre en oeuvre d'autres algorithmes de la grande famille des algorithmes proximaux qui forme aujourd'hui un vaste continent.

L'optimisation continue, c'est à dire résoudre un problème de la forme

$$\min_{x \in E} F(x)$$

où F est une fonction à valeur dans $\mathbb{R} \cup \{+\infty\}$ et où E est un ensemble infini, par exemple un espace de Hilbert, un espace euclidien, ou un sous ensemble de tels espaces, intervient dans de nombreux domaines des mathématiques (transport, traitement d'images, statistiques ...). Dans ce cours nous nous focaliserons sur le cas particulier où F est une fonction convexe et où E est un ensemble convexe. Non pas que les autres situations ne sont pas intéressantes ou utiles mais que c'est dans ce cadre plus simple où il est possible de développer le plus facilement des résultats et des algorithmes efficaces. Nous verrons qu'en pratique, certains algorithmes d'analyse convexe peuvent s'appliquer à des fonctions F non convexes mais que le résultat de ces algorithmes ne fournit que rarement un optimum global au problème.

Ce cours se divise en deux grandes parties, la première est consacrée à l'optimisation de fonctions convexes différentiables, la seconde de l'optimisation des fonctions convexes non différentiables, dites non lisses. Avant de traiter ces deux grandes parties nous présentons quelques exemples de problèmes qui nous intéresserons ainsi que les premières définitions qui permettront de poser un cadre théorique clair dans lequel nous pourrons ensuite travailler.

Le plan de ce document est donc le suivant.

Dans une première partie nous présentons des exemples, les premières définitions et le cadre théorique que nous allons étudier. Dans une seconde partie, nous commençons par donner les définitions et les propriétés essentielles spécifiques au cas des fonctions différentiables dont nous auront besoin. Nous introduirons par exemple la notion de dualité qui est une notion essentielle en optimisation convexe.
Nous traiterons en premier lieu des problèmes sans contraintes, puis des problèmes avec contraintes et plus spécifiquement des conditions d'optimalité (Lagrange conditions KKT).
Dans un second temps nous évoquerons les algorithmes d'optimisation différentiables, en commençant par les méthodes de gradient et en passant par les méthodes d'ordre supérieure. Nous évoquerons aussi les méthodes de lissage, ou smoothing, qui permettent d'approcher des problèmes non lisses par des problèmes différentiables.

Dans une troisième partie nous traiterons de l'optimisation non différentiable. Nous introduirons les outils spécifiques à l'optimisation non lisse comme le sous différentiel, et l'opérateur proximal ainsi que des propriétés élémentaires de ces derniers. Nous étudierons ensuite les conditions d'optimalité dans le cadre non lisse. Puis nous détaillerons plusieurs algorithmes proximaux et leur cadre d'application. Nous étudierons en priorité le Forward-Backward et sa version accélérée FISTA, l'algorithme Primal-Dual, l'algorithme Douglas-Rachford ainsi que l'ADMM. Pour conclure cette partie nous verrons comment un même problème peut être réécrit de plusieurs manières et ainsi être résolu via divers algorithmes. La dualité jouera ici un rôle crucial.

1 Motivation et définitions

1.1 Pourquoi résoudre un problème d'optimisation ?

Les problèmes d'optimisation interviennent dans différents domaines des mathématiques :

- Déterminer les dimensions optimales d'un objet.
- Déterminer la fonction d'une classe donnée la plus proche d'une fonction ou d'un nuage de poinst données.
- Déterminer la stratégie optimale dans un jeu.
- Déterminer un estimateur optimal pour une classe de fonction (seuillage pour un débruitage d'images par exemple), maximum de vraisemblance ou maximum a prosteriori en statitiques.
- Déterminer le minimum d'un critère en restauration d'images.
- Transport optimal.

On peut écrire beaucoup de ces problèmes sous la forme :

$$\min_{x \in E} F(x)$$

où F est une fonction convexe définie sur un espace $\mathcal{H}$ qui sera un espace de Hilbert ou plus souvent un espace euclidien, typiquement $\mathbb{R}^n$, à valeur dans $\mathbb{R} \cup \{+\infty\}$ et où E peut être l'espace $\mathcal{H}$ tout entier, un sous espace de dernier ou un sous ensemble convexe de $\mathcal{H}$ défini par un ensemble de contraintes. Les points $x \in E$ sont dits admissibles.

Ce problème n'admet pas toujours de solution ou peut admettre des minima locaux. On peut ramener les problèmes de détermination de maxima à celle de minima. On peut aussi avoir à résoudre des problèmes de points-selle c'est à dire des problèmes de la forme :

$$\min_{x \in E_1} \max_{y \in E_2} G(x, y)$$

Nous allons nous concentrer ici sur les problèmes de minimisation de fonctions F et des contraintes E convexes. Nous verrons que ce cadre est assez général et que les techniques utilisées peuvent parfois s'appliquer à des problèmes non convexes. Nous verrons aussi que certains problèmes non convexes peuvent être convexifiés et que certains problèmes de minimisation peuvent être réécrits sous forme de problèmes de détermination de points-selle via la théorie de la dualité.
Nous allons maintenant rappeler certaines définitions et propriétés de base de calcul différentiel qui seront utiles pour la suite.

1.2 Rappels de calcul différentiel

1.2.1 Fonctions de $\mathbb{R}$ dans $\mathbb{R}$

Soit f une fonction définie de d'un intervalle ouvert $I \subset \mathbb{R}$ à valeurs réelles. On rappelle quelques définitions classiques.

- On appelle point critique de f sur un point $x \in I$ tel que $f'(x) = 0$.
- On appelle minimiseur (respectivement maximiseur) local de f sur I un point $x \in I$ tel qu'il existe $r > 0$ tel que pour tout $y \in I \cap]x - r, x + r[$, $f(y) \geqslant f(x)$ (respectivement $f(y) \leqslant f(x)$). On appelle minimum local ou maximum local la valeur $f(x)$.
- Une fonction f est convexe sur I si et seulement si

$$\forall \lambda \in [0, 1],\ \forall (x, y) \in I^2,\ f(\lambda x + (1 - \lambda) y) \leqslant \lambda f(x) + (1 - \lambda) f(y).$$

- Une fonction f est strictement convexe sur I si et seulement si

$$\forall \lambda \in [0, 1],\ \forall (x, y) \in I^2,\ f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y).$$

On rappelle les quelques propriétés suivantes :

- Si f est dérivable au point $x \in I$ alors $f(x+h) = f(x) + hf'(x) + o(h)$.
- Si f est dérivable deux fois en $x \in I$ alors
  $f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + o(h^2)$.
- Si f est dérivable sur un ouvert I si f admet un minimum ou maximum local en $x \in I$ alors $f'(x) = 0$.
- Si f est convexe et dérivable sur I, alors f' est une fonction croissante sur I. (Réciproque ?)
- Si f est convexe et deux fois différentiable sur I alors f'' est une fonction positive sur I.
- Si f est convexe, alors tout minimum local sur I est un minimum global sur I et tout minimiseur local et un minimiseur global.
- Si f est strictement convexe sur I, elle admet au plus un minimiseur.
- La fonction $x \mapsto e^x$ est strictement convexe sur $\mathbb{R}$ et n'admet pas de minimum ni de minimiseur sur $\mathbb{R}$.

1.2.2 Optimisation de fonctions de $\mathbb{R}$ dans $\mathbb{R}$

Si f est une fonction définie d'un intervalle $I \subset \mathbb{R}$ et à valeurs dans $\mathbb{R} \cup \{+\infty\}$, il existe plusieurs manières d'en calculer un minimum local ou global suivant les hypothèses que l'on formule sur f.
Si la fonction f est différentiable, un point critique x_0 est caractérisé par le fait que $f'(x_0) = 0$. Ainsi, les minimieurs locaux sont contenus dans l'ensemble des points x_0 vérifiant x_0. Si on dispose d'un tel point, en analysant le comportement de la fonction au voisinage du point, on peut préciser s'il s'agit d'un minimiseur, d'un maximiseur ou d'un point d'inflexion.

1. La méthode par dichotomie permet d'approcher de manière séquentielle une valeur x_0 telle que $g(x_0) = 0$ où g est une fonction continue en définissant une suite d'intervalles emboités dont la taille tend vers 0 et qui contiennent tous x_0.
   Dans la version classique de cet algorithme, dont il existe des variantes, on considère un intervalle borné $I = [a, b]$ tel que $g(a)g(b) < 0$. On peut supposer par exemple que $g(a) > 0$ et $g(b) < 0$. On pose $a_0 = a$ et $b_0 = b$ puis pour tout $n \geqslant 0$ on définit $m = \frac{a_n + b_n}{2}$ et si $g(m) < 0$ on pose $b_{n+1} = m$ et $a_{n+1} = a_n$ et sinon on pose $a_{n+1} = m$ et $b_{n+1} = b_n$. Le théorème des valeurs intermédiaires asssure que pour tout $n \in \mathbb{N}$, l'intervalle $I_n = [a_n, b_n]$ contient un élément x_0 tel que $g(x_0) = 0$.

2. La méthode de descente de gradient explicite à pas fixe est définie par un point de départ $u_0 \in I$ et un pas fixe h on construit alors la suite u_n définie par $u_{n+1} = u_n - hf'(u_n)$.
   On peut montrer que sous certaines hypothèses de sur la fonction f (telle que la convexité) et sur h, cette méthode converge vers un point critique de f, indépendemment de u_0.

3. Il existe une variante implicite de cete méthode où la suite $(u_n)_{n\in\mathbb{N}}$ est définie par $u_{n+1} = u_n - hf'(u_{n+1})$. Une telle méthode nécessite cependant de résoudre une équation à chaque étape. Selon la forme de la fonction f, résoudre cette équation est plus ou moins couteux. Cette méthode est plus stable que la méthode de descente de gradient explicite.

4. Si la fonction f est deux fois dérivable et que l'on sait calculer la dérivée seconde f, on peut utiliser la méthode dite de *Newton*. Cette méthode est à l'origine une méthode pour déterminer le *zéro* d'une fonction dérivable en utilisant la tangente de la fonction au point u_n ainsi pour approcher le *zéro* de la fonction g on construit une suite de la manière suivante : $u_{n+1} = u_n - \frac{g(u_n)}{g'(u_n)}$. Ainsi si on cherche un point critique d'une fonction f, on peut construire une suite $(u_n)_{n\in\mathbb{N}}$ de la manière suivante : $u_{n+1} = u_n - \frac{f'(u_n)}{f''(u_n)}$. On peut voir cette méthode comme une méthode de descente de gradient à pas adaptatif. Nous verrons par la suite, dans le cadre plus général des fonctions convexes de $\mathbb{R}^n$ dans $\mathbb{R}$ des conditions suffisantes de convergence de cette méthode.

1.2.3 Fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

On peut généraliser la notion de dérivée première et seconde aux cas des fontions définies d'un ouvert $\Omega \in \mathbb{R}^n$ et à valeur dans $\mathbb{R}$ en introduisant la notion de gradient :

Définition 1 *On dit que f est différentiable en un point $x \in \Omega$, il existe un vecteur $u \in \mathbb{R}^n$ tel que*

$$f(x+h) = f(x) + \langle h, u\rangle + o(\|h\|). \tag{1}$$

S'il existe, ce vecteur u est unique et on le note $\nabla f(x)$.

On peut noter que le gradient est un vecteur de $\mathbb{R}^n$.
Dans le cas où $n = 2$, on peut représenter la fonction f par une surface d'équation $z = f(x, y)$. Le gradient de f au point (x, y) est un vecteur qui est dirigé dans la direction de plus forte pente de la surface au point (x, y) et dont la norme est proportionnelle *à la pente locale.*

Les composantes du gradient ne sont rien d'autre que les dérivées partielles de f par rapport à chacune des variables $(x_i)_{1\leqslant i\leqslant n}$.

Définition 2 *Soit f une fonction définie d'un ouvert $\Omega \subset \mathbb{R}^n$ et à valeurs réelles, i_0 un entier compris entre 1 et n, pour tout $n-1$ uplet $(x_1, \cdots, x_{i_0-1}, x_{i_0+1}, \cdots, x_n$ on note g l'application partielle définie par*
$g(x) = f(x_1, \cdots, x_{i_0-1}, x, x_{i_0+1}, \cdots, x_n)$.
Si la fonction g est dérivable au point x, on dit que la fonction f admet une dérivée partielle d'indice i_0 au point $(x_1, \cdots, x_{i_0-1}, x, x_{i_0+1}, \cdots, x_n)$. On note $\frac{\partial f}{\partial x_{i_0}}$ cette dérivée.

Proposition 1 *Si toutes les dérivées partielles de f existent et sont continues en tout point de Ω, alors on dit que f est C^1 sur Ω et on a*

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right)^t$$

Définition 3 *Soit f une fonction définie d'un ouvert $\Omega \in \mathbb{R}^n$ dans $\mathbb{R}$ de classe C^1 sur Ω et si l'application $x \mapsto \nabla f(x)$ est elle même C^1, on dit que f est de classe C^2, il existe alors une unique forme bilinaire φ_x telle que*

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \varphi_x(h,h) + o(\|h\|^2).$$

On appelle Hessienne la matrice H représentant cette application bilinéaire. La matrice hessienne est la matrice symétrique réelle dont le terme d'indice (i,j) est

$$H_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Remarque :
Nous avons fait l'hypothèse que la fonction est C^2 et ainsi la matrice hessienne est symétrique ce que lon peut traduire que l'ordre de dérivation n'importe pas.
Exemple :
Quelle est la Hessienne au point x de l'application $x \mapsto \frac{1}{2}\|Ax - y\|^2$, où A est une matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ et $y \in \mathbb{R}^m$?

Proposition 2 *Soit f est une fonction définie de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}$ et C^2. Alors f est convexe si et seulement si en tout point x, la matrice hessienne est une matrice positive, c'est à dire que toute ses valeurs propres sont positives.*

Exercice :
Justifier que l'application définie dans l'exercice précédent est une application convexe. Est-elle nécessairement strictement convexe ?

Définition 4 *Soit f une fonction définie de E dans $\bar{\mathbb{R}} = \mathbb{R} \cup +\infty$, on appelle domaine de f et on note $dom(f) = \{x \in E$ tels que $f(x) \neq +\infty\}$.*

Définition 5 *Une fonction f définie sur E est dite coercive si*

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Définition 6 *On dire qu'une fonction f définie de E dans $\mathbb{R} \cup +\infty$ est semi-continue inférieurement (on notera sci) si pour tout $x \in E$,*

$$\liminf_{y \to x} f(y) \geqslant f(x).$$

Remarque :

- La semi-continuité inférieure est stable par somme.
- Toute fonction continue est sci.
- L'indicatrice d'un convexe fermé est sci (c'est à dire la fonction qui vaut 0 en tout point du convexe et $+\infty$ en dehors).
- Si f est sci, pour tout tout $\alpha \in \mathbb{R}$, les ensembles

  $$\{x \in E,\ f(x) \leqslant \alpha\} \text{ et } \{(x, \alpha) \in E \times \mathbb{R} \text{ tels que } f(x) \leqslant \alpha\}$$

  sont fermés.

Définition 7 *Une fonction de E dans $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ est dite propre si elle n'est pas identiquement égale à $+\infty$ et si elle ne prend pas la valeur $-\infty$.*

Proposition 3 *Si F est une fonction définie de E dans $\bar{\mathbb{R}}$, propre, coercive et semi continue inférieurement, elle est minorée.*

Démonstration :
Pour tout $r \in \mathbb{R}$, l'ensemble $H_r = \{x \in E \text{ tels que } F(x) \leqslant r\}$ est fermé car F est sci et borné car F est coercive. Donc pour tout $r \in \mathbb{R}$, l'ensemble H_r est compact.
L'ensemble $H = \displaystyle\bigcap_{r \in \mathbb{R}} H_r$ est une intersection de compacts emboités. Comme F est propre, $H = \emptyset$ donc il existe $r_0 \in \mathbb{R}$ tel que H_{r_0} est vide et donc r_0 est un minorant de F, ce qui conclut la démonstration de la proposition.

Proposition 4 *Soit F une fonction définie sur E à valeur dans $\bar{\mathbb{R}}$, convexe, propre, sci et coercive, alors F admet au moins un minimiseur.*

Comme F est propre, il existe $x_0 \in E$ tel que $F(x_0) < +\infty$ et $F(x_0) \neq -\infty$. L'ensemble $H = \{x \in E \text{ tels que } F(x) \leqslant F(x_0)\}$ est fermé car F est sci et borné car F est coercive. H est donc compact car E est de dimension finie. On a $\inf_{x \in E} F(x) = \inf_{x \in H} F(x) > -\infty$ car F est bornée inférieurement. Soit $(x_n)_{n \in \mathbb{N}}$ une suite minimisante d'éléments de H tels que $\lim_{n \to \infty} F(x_n) = \inf_{x \in H} F(x)$. Cette suite admet une sous-suite convergente vers un élément x^∞ de H. Comme F est sci on a

$$\liminf_{n \to \infty} F(x_n) \geqslant F(x^\infty).$$

Comme $\lim_{n \to \infty} F(x_n) = \inf_{x \in H} F(x)$ on en déduit que $F(x^\infty) = \inf_{x \in H} F(x) = \inf_{x \in E} F(x)$

2 Optimisation de fonctions différentiables

Dans cette partie on s'intéresse au problème générique suivant :

$$\min_{x \in E} F(x) \tag{2}$$

où F est une fonction différentiable définie sur un espace de Hilbert $\mathcal{H}$ et à valeur dans $\mathbb{R}$, convexe, coercive et que E est un ensemble convexe fermé. Comme F est différentiable, elle est sci et propre par définition. Ces hypothèses assurent que ce problème d'optimisation (2) admet au moins une solution. Si de plus F est strictement convexe alors il existe une unique solution à (2).

2.1 Propriétés des fonctions différentiables

Les fonctions convexes différentiables ont la propriété d'être minorées par leurs approximations affines :

Proposition 5 *Soit f une fonction convexe différentiable définie de $\mathbb{R}^n$ dans $\mathbb{R}$ alors pour tout couple $(x, y) \in (\mathbb{R}^n)^2$*

$$f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle \tag{3}$$

Démonstration
Par convexité on a pour tout $\theta \in]0, 1[$,

$$f(x + \theta(y - x)) \leqslant (1 - \theta)f(x) + \theta f(y).$$

ce qui peut s'écrire

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta} \leqslant f(y) - f(x)$$

en faisant tendre θ vers 0 on obtient

$$\langle \nabla f(x), y - x \rangle \leqslant f(y) - f(x)$$

On déduit de cette proposition, en l'appliquant aux points x et y le caractère monotone du gradient :

Définition 8 *Soit g une fonction définie de $\mathbb{R}^n$ dans $\mathbb{R}^n$, on dit que g est monotone si pour tout couple $(x, y) \in (\mathbb{R}^n)^2$,*

$$\langle g(x) - g(y), x - y \rangle \geqslant 0.$$

Corollaire 1 *Soit f une fonction convexe différentiable définie de $\mathbb{R}^n$ dans $\mathbb{R}$, le gradient de f est monotone.*

Il existe une notion de convexité, appelée forte convexité qui permet des contrôles locaux de la fonction encore un peu plus fins.

Définition 9 *Soit f une fonction définie de E dans $\mathbb{R} \cup \{+\infty\}$ et soit $\alpha > 0$, on dit que la fonction f est $\alpha-$fortement convexe (ou $\alpha-$convexe) si la fonction g définie par $g(x) = f(x) - \frac{\alpha}{2} \|x\|^2$ est convexe.*

On peut remarquer (exercice !) que s'il existe x_0 tel que la fonction $g_{x_0} = f(x) - \alpha \|x - x_0\|^2$ est convexe alors f est $\alpha-$convexe et bien sûr qu'une fonction fortement convexe est strictement convexe et donc convexe. Elle admet en particulier un unique minimiseur.
Par définition, si f est convexe et si y est un élément de E, la fonction $x \mapsto f(x) + \frac{1}{2\gamma} \|x - y\|$ est $\frac{1}{\gamma}$-fortement convexe.
Le lemme suivant montre que si on s'écarte du minimiseur d'une fonctionnelle fortement convexe, la valeur de la fonctionenlle croit au moins de manière quadratique. Ce lemme sera utile par la suite.

Lemma 1 *Soit f une foncion $\alpha-$fortement convexe, et soit x^* le minimiseur de f, alors pour tout $x \in E$ on a*

$$f(x^*) + \frac{\alpha}{2} \|x^* - x\|^2 \leqslant f(x) \tag{4}$$

Démonstration :
On note g, la fonction $x \mapsto f(x) - \frac{\alpha}{2} \|x\|^2$ qui est donc convexe par définition. Soit $\lambda \in]0, 1[$, par convexité de g on a

$$\begin{aligned}
g(\lambda x^* + (1-\lambda)x) &\leqslant \lambda g(x^*) + (1-\lambda) g(x) \\
f(\lambda x^* + (1-\lambda)x) - \frac{\alpha}{2} \|\lambda x^* + (1-\lambda)x\|^2 &\leqslant \lambda \left(f(x^*) - \frac{\alpha}{2} \|x^*\|^2 \right) + (1-\lambda) \left(f(x) - \frac{\alpha}{2} \|x\|^2 \right) \\
(1-\lambda)\left(f(x^*) - f(x)\right) &\leqslant \frac{\alpha}{2} \left((\lambda^2 - \lambda) \|x^*\|^2 + 2\lambda(1-\lambda)\langle x^*, x\rangle - \lambda(1-\lambda) \|x\|^2 \right) \\
f(x^*) - f(x) &\leqslant \frac{-\lambda\alpha}{2} \|x^* - x\|^2 \\
f(x^*) + \lambda \frac{\alpha}{2} \|x^* - x\|^2 &\leqslant f(x)
\end{aligned}$$

où on utilise à la troisième ligne le fait que x^* est un minimiseur. Ceci pour tout $\lambda \in]0, 1[$, ce qui conclut la preuve du Lemme.

Nous aurons aussi besoin de la définition du caractère Lipschitz d'une fonction :

Définition 10 *On dit qu'une fonction g définie de $\mathbb{R}^n$ dans $\mathbb{R}$ est $L-$Lipschitz si pour tout couple $(x, y) \in (\mathbb{R}^n)^2$ on a*

$$\|g(x) - g(y)\| \leqslant L \|x - y\|$$

Si g est 1-Lipschitz on dit que g est non expansive et si g est $L-$Lipschitz avec $L < 1$ on dit que g est L-contractante ou contractante.

Nous venons de voir que les fonctions convexes sont minorées par leurs approximations affines mais si le gradient de f est $L - Lipschitz$, on peut obtenir un majorant de f indépendemment de la convexité en effet :

Lemma 2 *Soit f une fonction de définie de $\mathbb{R}^n$ dans $\mathbb{R}$, différentiable dont le gradient est $L-$Lipschitz, alors pour tout couple $(y,z) \in (\mathbb{R}^n)^2$*

$$f(z) \leqslant f(y) + \langle \nabla f(y), z-y \rangle + \frac{L}{2} \|z-y\|^2 . \tag{5}$$

Démonstration :
Soit g une fonction dérivable définie sur $\mathbb{R}$ telle que g' est $K-$Lipschitz,

$$g(1) = g(0) + \int_0^1 g'(t)dt = g(0) + g'(0) + \int_0^1 (g'(t) - g'(0))dt \leqslant g(0) + g'(0) + \frac{K}{2}.$$

Soit $(y,z) \in (\mathbb{R}^n)^2$, pour tout $t \in [0,1]$ on pose $v = z-y$, $y_t = y + t(z-y)$ et $g(t) = f(y_t)$. La fonction g est dérivable et $g'(t) = \langle \nabla f(y_t), v\rangle$. D'après les hypothèses sur f, g' est K-Lipchitz avec $K = L\|v\|^2$.
On en déduit que pour tout $(y,z) \in (\mathbb{R}^n)^2$

$$f(z) \leqslant f(y) + \langle \nabla f(y), z-y \rangle + \frac{L}{2} \|z-y\|^2 .$$

On peut également définir une notion plus stricte que la non-expansivité, la ferme non-expansivité (firmly non expansive en anglais) qui sera utile dans la suite :

Définition 11 *On dit qu'une fonction f est fermement non-expansive si pour tout couple $(x,y) \in (\mathbb{R}^n)^2$*

$$\|f(x) - f(y)\|^2 + \|x - f(x) - y + f(y)\|^2 \leqslant \|x-y\|^2$$

l'exemple classique d'une fonction fermement non-expansive est celui d'un projecteur sur un convexe fermé mais nous verrons qu'il en existe d'autres tout aussi intéressant.

Proposition 6 *Soit f une fonction convexe définie sur $\mathbb{R}^n$ différentiable à gradient $L-$Lipschitz. Si $\gamma < \dfrac{1}{L}$, l'application $Id - \gamma\nabla f$ est fermement non-expansive et donc 1-Lipschitz.*

Nous verrons que cette propriété est cruciale pour démontrer la convergence de plusieurs méthodes d'optimisation. En effet nombre de méthodes d'optimisation reposent sur des théorèmes de points fixes associés à des opérateurs 1 Lipschitz.
Démonstration :
On se ramène à déterminer le signe d'un produit scalaire.
Soit $(x,y) \in (\mathbb{R}^n)^2$, on pose
$u = \dfrac{1}{\gamma}(\nabla f(x) - \nabla f(y))$ et $v = x - \dfrac{1}{\gamma}\nabla f(x) - (y - \dfrac{1}{\gamma}\nabla f(y))$.
On a

$$\|x-y\|^2 = \|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\langle u, v\rangle$$

Pour montrer que $Id - \gamma\nabla f$ est fermement non-expansive il suffit de montrer que le produit scalaire $\langle u, v\rangle$ est positif, ce qui découle directement du fait que $\gamma < \frac{1}{L}$ et de la co-coercivité de ∇f qui est énoncée et démontrée dans le lemme suivant.

Lemma 3 *Soit f une fonction différentiable et convexe tel qu'il existe $L > 0$ tel pour tout $(x, y) \in (\mathbb{R}^n)^2$, $\|\nabla f(x) - \nabla f(y)\| \leqslant L \|x - y\|$. On a la relation suivante :*

$$\forall (x, y) \in (\mathbb{R}^n)^2, \quad \langle \nabla f(x) - \nabla f(y), x - y \rangle \geqslant \frac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2 . \qquad (6)$$

Cette propriété est appelée co-coercivité de la fonction ∇f.

Démonstration :
On applique la majoration (119) en prenant ensuite $z = x$ et $y = x - \frac{1}{L} \nabla f(x)$ on obtient que

$$\frac{1}{2L} \|\nabla f(x)\|^2 \leqslant f(x) - f(x - \frac{1}{L} \nabla f(x)) \leqslant f(x) - f(x^\star).$$

où $x^\star$ est un minimiseur de f. On a ainsi pour tout $x \in \mathbb{R}^n$,

$$\frac{1}{2L} \|\nabla f(x)\|^2 \leqslant f(x) - f(x^\star). \qquad (7)$$

Soit $(x, y) \in (\mathbb{R}^n)^2$, on introduit ensuite une fonction $h_1(z) = f(z) - \langle \nabla f(x), z \rangle$ et $h_2(z) = f(z) - \langle \nabla f(y), z \rangle$. Ces deux fonctions sont convexes et admettent pour minimiseurs respectivement x et y. On applique l'inégalité (120) à ces deux fonctions et on déduit

$$h_1(x) \leqslant h_1(y) - \frac{1}{2L} \|\nabla h_1(y)\|^2 \text{ et } h_2(y) \leqslant h_2(x) - \frac{1}{2L} \|\nabla h_2(y)\|^2 .$$

En sommant ces deux inégalités on obtient le résultat du lemme.

Un cadre théorique classique qui assure que le gradient d'une fonction est $L-$Lipschitz est celui des fonctions deux fois différentiables : Si on est capable de borner la norme d'opérateur de la Hessienne par L alors on en déduit que le gradient de f est $L-$*Lipschitz* et on peut appliquer les résultats précédents.

2.2 Conditions d'optimalité

Les minimiseurs x^* d'une fonction f convexe et différentiable sur $\mathbb{R}^n$ sont caractérisés simplement par l'équation d'Euler :

$$\nabla f(x^*) = 0 \qquad (8)$$

Une autre manière de le formuler est de dire que tous les points critiques d'une fonction convexe sont des minima globaux.
Les difficultés commencent quand on traite le cas des problèmes contraints. C'est à dire de la forme :

$$\min_{x \in K} f(x) \qquad (9)$$

où K est un ensemble convexe fermé.
On caractérise les minimiseurs de la manière suivante :

Théorème 1 *Soit f une fonction différentiable et K un convexe fermé, x^* est une solution de (9) si et seulement si pour tout $v \in K$,*

$$\langle \nabla f(x^*), v - x^* \rangle \geqslant 0 \qquad (10)$$

2.2.1 Minimisation sous contraintes différentiables, Multiplicateurs de Lagrange, conditions KKT

Même si ce cours traite essentiellement de l'optimisation de fonctions convexes nous allons énoncer dans cette section des résultats plus généraux qui s'appliquent à des fonctions différentiables non nécessairement convexes sous des contraintes non nécessairement convexes. Ces résultats font intervenir les multiplicateurs de Lagrange et induisent des résultats plus précis dans le cas convexe. Cette théorie des multiplicateurs de Lagrange permet de caractériser les solutions de problèmes contraints quand l'ensemble des contraintes K est défini par des contraintes dites égalités ou inégalités. On va dans un premier temps énoncer une condition **nécessaire** d'optimalité d'une solution de (9) quand l'ensemble K est défini à travers des contraintes d'égalité. On énoncera ensuite des conditions d'optimialité avec des contraintes inégalité puis on spécifiera ces résultats au cas convexe.

Multiplicateurs de Lagrange On énonce ici un théorème général, dont la démonstration utilise le Théorème des fonctions implicites qui s'applique à des fonctions différentiables non nécessairement convexes et à des contraintes dites d'égalité faisant intervenir des foncions C^1 non nécessairement affines :

Théorème 2 *Soit f une fonction différentiable définie de $\mathbb{R}^n$ dans $\mathbb{R}$, et F_i m fonctions C^1 définie de $\mathbb{R}^n$ dans $\mathbb{R}$ et $K = \{x \in \mathbb{R}^n \text{ tels que } \forall i \in [\![1, m]\!],\ F_i(x) = 0.\}$. On considère le problème d'optimisation suivant*

$$\min_{x \in K} f(x) \tag{11}$$

Si u est un vecteur tel que la famille $(\nabla F_i(u))_{i \in [\![1,m]\!]}$ est une famille libre et si u est une solution du problème ci dessus alors il existe m réels $(\lambda_i)_{i \in [\![1,m]\!]}$ tels que

$$\nabla f(u) + \sum_{i=1}^{m} \lambda_i \nabla F_i(u) = 0. \tag{12}$$

Les variables $(\lambda_i)_{i \leqslant M}$ sont appelés multiplicateurs de Lagrange ou variables duales. L'équation (12) traduit le fait que u est un point critique du *Lagrangien augmenté* :

$$\mathcal{L}(u, \Lambda) = f(u) + \sum_{i=1}^{M} \lambda_i F_i(u) = f(u) + \langle \Lambda, F(u) \rangle. \tag{13}$$

En effet la relation (12) assure que $\frac{\partial \mathcal{L}(u,\Lambda)}{\partial u} = 0$ et de plus le fait que u satisfait les contraintes assure que $\frac{\partial \mathcal{L}(u,\Lambda)}{\partial \lambda_i} = F_i(u) = 0$ pour tout $i \leqslant M$.

On peut noter que cette condition n'est pas suffisante pour assurer que u est un minimiseur mais une condition nécessaire. En effet sans hypothèses plus pécises sur f et les contraintes F_i, le fait d'être un point critique du Lagrangien augmenté (13) ne garantit pas d'être un minimiseur de f.

Quelques exemples :

1. On veut estimer les dimensions optimales en terme de quantité de métal d'une boîte cylindrique. Le problème à résoudre s'exprime alors de la manière suivante :
$$\min_{hr^2=V} r^2 + rh$$
Qu'en est-il si on considère une boite de section carré ?

2. Soit A une matrice symétrique réelle. On veut déterminer la solution du problème suivant :
$$\min_{\|x\|_2=1} \langle Ax, x\rangle$$

3. Déterminer la distance entre les courbes d'équation $y = x^2$ et $y = x-2$.

Si l'ensemble convexe K des contraintes n'est pas défini par des contraintes dites *égalité* mais aussi par des contraintes inégalités, positivité de composantes, appartenance à un polytope spécifique, on peut aussi utiliser les multiplicateurs de Lagrange.

Avant d'énoncer un théorème similaire au précédent on va définir la notion de *contraintes qualifiées* dans le cas de *contraintes inégalités.* Il existe une autre définition si on veut prendre en compte des *contraintes inégalités* et des contraintes *contraintes égalités* mais pour la clarté de l'exposé nous choisissons ici de rester dans le cas simple où

$$K = \{v \in E, \, , \, F_i(u) \leqslant 0 \text{ pour } 1 \leqslant i \leqslant M\} \tag{14}$$

Définition 12 *On dit que des contraintes* (14) *sont qualifiées en $u \in K$ si et seulement s'il existe une direction $w \in E$ telle que pour toute contrainte active, c'est-à-dire pour tout i tel que $F_i(u) = 0$,*

- *ou bien $\langle F'(u), w\rangle < 0$,*
- *ou bien $\langle F_i'(u), w\rangle = 0$ et F_i est affine.*

On peut maintenant énoncer le théorème suivant qu'on ne démontrera pas ici :

Théorème 3 *On suppose que K est donné par* (14)*, que les fonctions f et $(F_i)_{1\leqslant i\leqslant M}$ sont dérivables en u et que les contraintes sont qualifiées en u. Alors si u est un minimiseur local de f sur K, il existe $(\lambda_1, \cdots, \lambda_M)$ positifs ou nuls appelés multiplicateurs de Lagrange, tels que*

$$\nabla f(u) + \sum_{i=1}^{M} \lambda_i \nabla F_i(u) = 0, \, \lambda_i \geqslant 0, \, \lambda_i = 0 \text{ si } F_i(u) < 0 \, \forall i \in [\![1, M]\!]. \tag{15}$$

La différence essentielle avec le théorème précédent et le cas des *contraintes égalités* est que cette fois ci les multiplicateurs de Lagrange sont nécessairement positifs.

Exemple

Dans une usines, deux produits u_1 et u_2 sont fabriqués et rapportent par unité e_i kilo euros, nécessitent t_1 et t_2 heures de travail machines et q_1 et q_2

tonnes de matières premières. On dispose de 10h en tout de travail machines et de 15 tonnes de matières premières.
Formaliser ce problème sous la forme d'un problème d'optimisation sous contraintes et le résoudre pour $(e_1, t_1, q_1) = (6, 2, 1)$ et $(e_2, t_2, q_2) = (5, 1, 3)$. Est il intéressant d'augmenter la quantité de matière première disponible ? Jusqu'à quel point ?

$$\begin{array}{rl} \min & -6x_1 - 5x_2 \\ 2x_2 + x_2 - 10 \leqslant 0 & \\ x_1 + 3x_2 - 15 \leqslant 0 & \\ -x_1 \leqslant 0 & \\ -x_2 \leqslant 0 & \end{array}$$

Faire un dessin pour vérifier quelles contraintes sont actives.

D'une manière générale, il n'est pas aisé de savoir quelles sont les contraintes qui seront actives. Dans des cas complexes, il existe des algorithmes qui recherchent quelles contraintes actives sotn actives et lesquelles ne le sont pas.

Dans la suite nous verrons que si les fonctions f et $(F_i)_{i\in[\![1,M]\!]}$ sont convexes, cette condition nécessaire d'optimalité est une condition nécessaire et suffisante.

Conditions de Khun et Tucker Dans ce paragaphe, nous retournons à l'analyse convexe, c'est à dire au cas d'une fonction f convexe et de contraintes convexes. On peut noter que si les fonctions $(F_i)_{i\in[\![1,M]\!]}$ sont convexes et différentiables, l'ensemble K défini par des *contraintes inégalités* en (14) est un convexe fermé. On peut observer également qu'une *contrainte d'égalité* affine peut s'exprimer comme une double *contrainte inégalité* convexe, en effet $F_i(u) = 0$ est équivalent à $F_i(x) \leqslant 0$ et $-F_i(x) \leqslant 0$.
Dans ce cas, le théorème suivant donne une condition nécessaire et suffisante de minimalité sous contrainte qui s'appuie sur le Lagragien associé au problème :

Théorème 4 *On suppose que les fonctions f et $(F_i)_{i\in[\![1,M]\!]}$ sont convexes et continues sur E et dérivables sur l'ensemble des contraintes K défini en* (14)*. On introduit le Lagrangien $\mathcal{L}$ associé*

$$\mathcal{L}(v,q) = f(v) + qF(v) \quad \forall (v,q) \in E \times (\mathbb{R}^+)^M. \tag{16}$$

Soit $u \in K$ où les contraintes sont qualifiées au sens de la définition 12. Alors u est un minimum global de f sur K si eu seulement s'il existe $p \in (\mathbb{R}^+)^M$ tel que (u,p) est un point-selle du Lagrangien $\mathcal{L}$ sur $E \times (\mathbb{R}^+)^M$ ou de manière équivalente, tel que

$$F_i(u) \leqslant 0,\ \forall i \in [\![1,M]\!],\ p \geqslant 0,\ \langle p, F(u)\rangle = 0,\ \nabla f(u) + \sum_{i=1}^{M} p_i \nabla F_i(u) = 0 \tag{17}$$

Nous ne démontrons pas ici ce théorème dont la démonstration s'appuie sur le Théorème 3. Le fait que cette condition est aussi une condition suffisante de minimalité vient du fait que les fonctions en jeu sont convexes et que les seuls points critiques d'une fonction convexes sont les minima globaux.

Dans la section suivante, nous verrons que l'algorithme d'Usawa peut permettre de résoudre ce problème d'optimisation sous contrainte dans des cas favorables.

EXEMPLES.

2.3 Méthodes du gradient

Dans le cas où f est une fonction convexe différentiable, il existe plusieurs méthodes de descente de gradient qui permettent de construire des suites minimisantes. Dans cette section, nous allons décrire les méthodes les plus classiques n'utilisant que le gradient de la fonction f pour approcher les solutions du problème suivant :

$$\min_{u \in X} f(x). \tag{18}$$

2.3.1 Descente de gradient à pas fixe.

La méthode du gradient à pas fixe est extrèmement simple à décrire. On considère un réel strictement positif γ appelé *pas*, un élement $x_0 \in X$ et pour tout $n \in \mathbb{N}$ on définit :

$$x_{n+1} = x_n - \gamma \nabla f(x_n). \tag{19}$$

Si on note T l'opérateur défini de X dans X par

$$Tx = x - \gamma \nabla f(x) = (I - \gamma \nabla f)(x)$$

la suite $(x_n)_{n \in \mathbb{N}}$ est définie simplement par $x_{n+1} = Tx_n$.

Si aucune hypothèse n'est faite sur le gradient de f, une telle méthode peut ne pas converger vers un minimiseur, en effet l'algorithme peut diverger si γ est choisi trop grand (Exemples d'oscillation et de divergence). En revanche si le gradient de f est L- Lipschitz, alors en choisissant correctement γ, cette méthode converge vers un minimiseur de f :

Théorème 5 *Soit f une fonction convexe différentiable, admettant un minimiseur dont le gradient est $L-$Lipschitz et si $\gamma < \frac{2}{L}$ alors pour tout $x_0 \in X$, la suite définie pour tout $n \in \mathbb{N}$ par $x_{n+1} = Tx_n$ converge vers un minimiseur de f.*

Démonstration

Il existe différentes preuves de cette convergence, nous en proposons une s'appuyant sur le caratère 1-Lipschitz (ou non expansif) de l'opérateur T que nous pourrons appliquer à d'autres algorithmes. Nous utiliserons ainsi le lemme suivant après avoir démontré que l'opérateur T associé à la descente de gradient explicite vérifie les hypothèses du lemme :

Lemma 4 *Soit T un opérateur défini de X dans X, 1-Lipschitz admettant un point fixe. Soit $x_0 \in X$ et $(x_n)_{n\in\mathbb{N}}$ la suite définie par $x_{n+1} = Tx_n$.*
Si $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$ alors la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un point fixe de T.

Démonstation :
Soit y un point fixe de T, comme T est 1-lipschitz, la suite $(\|x_n - y\|)_{n\in\mathbb{N}}$ est décroissante donc bornée, ainsi la suite $(x_n)_{n\in\mathbb{N}}$ est bornée.

Comme X est de dimension finie, la suite $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence $z \in X$. Comme $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$, cette valeur d'adhérence vérifie $Tz = z$ et est donc un point fixe de T.

La suite $\|z - x_n\|$ est donc décroissante et admet une sous suite qui tend vers 0, elle converge donc vers 0, ce qui conclut la démonstration.

Pour démonstrer le Théorème 5, nous allons montrer que l'opérateur T vérifie toutes les hypothèses du lemme. Le fait que T est 1-Lipschitz est l'objet du lemme suivant. Notons que si $\gamma \leqslant \frac{1}{L}$, l'opérateur T vérifie une propriété encore plus intéressante, le caractère fermement non expansif.

Lemma 5 *Soit f une fonction convexe différentiable dont le gradient est $L-$Lipschitz, pour $\gamma > 0$ on note $T = Id - \gamma\nabla f$, alors*

1. *Si $\gamma \leqslant \frac{1}{L}$, l'opérateur T est fermement non expansif c'est-à-dire :*
$$\forall (x,y) \in X^2 \quad \|Tx - Ty\|^2 + \|x - Tx - y + Ty\|^2 \leqslant \|x - y\|^2 . \tag{20}$$
2. *Si $\gamma \leqslant \frac{2}{L}$, l'opérateur T est 1-Lipschitz*

La démonstration de ce lemme renvoyée en Appendice repose sur le caractère cocoercif du gradient d'une fonction convexe différentiable.
On peut noter ensuite qu'il y équivalence entre être un point fixe de T et être minimieur de f. En effet, $Tx = x$ est équivalent à $\nabla f(x) = 0$. Or les hypothèses du Théorème 5 supposent l'existence d'un tel minimiseur donc celles d'un point fixe de T. Pour terminer la démonstration du théorème il suffit donc de montrer que $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$. Pour cela on va appliquer le Lemme 2 aux points $y = x_n$ et $z = x_{n+1}$. On a alors $z - y = \gamma\nabla f(x_n)$ et ainsi

$$f(x_{n+1}) \leqslant f(x_n) - \frac{1}{\gamma}\|x_{n+1} - x_n\|^2 + \frac{L}{2}\|x_{n+1} - x_n\|^2 \tag{21}$$

et donc si $\gamma < \frac{2}{L}$

$$f(x_{n+1}) + \left(\frac{2 - \gamma L}{2\gamma}\right)\|x_{n+1} - x_n\|^2 \leqslant f(x_n) \tag{22}$$

Comme la suite $(f(x_n))_{n\in\mathbb{N}}$ est bornée inférieurement on en déduit que $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$ ce qui conclut la preuve du théorème. On peut noter ici que la méthode de gradient à pas fixe est une méthode de descente, c'est à dire que la valeur de $f(x_n)$ décroit. Nous verrons plus tard, quand nous étudierons l'algorithme Forward-Backward qui est une généralisation de cet algorithme qu'il est possible de contrôler la décroissance de $f(x_n) - f(x^*)$ vers 0, où x^* désigne un minimiseur quelconque de f.

2.3.2 Descente de gradient à pas fixe, implicite

La descente de gradient à pas fixe implicite s'exprime aussi simplement que la méthode explicite mais pose un problème pratique de résolution de problème implicite à chaque étape.

On considère un réel strictement positif γ appelé *pas*, un élement $x_0 \in X$ et pour tout $n \in \mathbb{N}$ on définit :

$$x_{n+1} = x_n - \gamma \nabla f(x_{n+1}). \tag{23}$$

Sans hypothèse sur f cette relation ne garantit pas l'existence et l'unicité de x_{n+1} et la résolution de cette équation implicite peut être pratiquepment compliquée.

La convexité de f assure l'existence et l'unicité de x_{n+1} en effet on peut remarquer que x_{n+1} est l'unique minimiseur d'une fonction fortement convexe et coercive :

$$x_{n+1} = \arg\min_{x \in X} f(x) + \frac{1}{2} \|x_n - x\|^2 \tag{24}$$

Ainsi x_{n+1} est l'unique élement de X tel que $x_{n+1} + \gamma \nabla f(x_{n+1}) = x_n$, ce qui nous permettra d'écrire $x_{n+1} = (Id + \gamma \nabla f)^{-1}(x_n)$, soit en notant $T = (Id + \gamma \nabla f)^{-1}$, on a $x_{n+1} = Tx_n$.
Nous verrons plus tard que cette seconde définition de x_{n+1} ne faisant pas appel au gradient peut se généraliser à n'importe quelle fonction convexe propre et sci sans hypothèse de différentiabilité.

Nous allons montrer que la méthode de gradient implicite permet de construire une suite minimisante, quel que soit le choix de x_0 et quel que soit le choix de $\gamma > 0$. Cette méthode présente l'avantage par rapport à la méthode explicite de ne pas nécessiter de condition sur γ, mais a l'inconvénient de nécessiter la résolution d'un problème potentiellement difficile à chaque étape.

Théorème 6 *Soit f une fonction convexe différentiable, admettant un minimiseur alors pour tout $x_0 \in X$, la suite définie pour tout $n \in \mathbb{N}$ par $x_{n+1} = (Id + \gamma \nabla f)^{-1} x_n$ converge vers un minimiseur de f.*

La démonstration de ce Théorème est similaire à celui du théorème assurant la convergence de la méthode du gradient explicite. Elle s'appuie sur le caractère 1-Lipschitz de l'opérateur $T = (Id + \gamma \nabla f)^{-1}$. Nous allons en donner ici une démonstration dans le cas où f est différentiable mais nous verrons plus tard que cette hypothèse est superflue.

Soient $(x, y) \in E^2$, on définit

$$p = \arg\min_{z \in E} f(z) + \frac{1}{2} \|x - z\|^2 \text{ et } q = \arg\min_{z \in E} f(z) + \frac{1}{2} \|y - z\|^2. \tag{25}$$

Nous allons montrer que

$$\|p - q\|^2 + \|x - p - y + q\|^2 \leqslant \|x - y\|^2 \tag{26}$$

ce qui implique que T est fermement non expansif et donc 1-Lipschitz.
Remarquons d'abord que

$$\|x - y\|^2 - \|p - q\|^2 - \|x - p - y + q\|^2 = 2\langle p - q, x - p - (y - q)\rangle \tag{27}$$

Pour montrer le caractère non expansif de T il est donc nécessaire et suffisant de montrer la positivité du produit scalaire ci dessus.

Par définition de p et q, tous deux minimiseurs de fonctionnelles convexes différentiables :

$$\nabla f(p) = p - x \text{ et } \nabla f(q) = q - y \tag{28}$$

et en utilisant le fait que f est minorée par ses approximants affines on a pour tout $z \in E$

$$f(z) \geqslant f(p) + \langle z - p, p - x\rangle \text{ et } f(z) \geqslant f(q) + \langle z - q, q - y\rangle. \tag{29}$$

En appliquant la première inégalité avec $z = q$ et la seconde avec $z = p$ et en sommant les deux inégalités on déduit que $\langle p - q, x - p - (y - q)\rangle \geqslant 0$ ce qui assure que T est bien fermement non expansif et donc 1-Lipchitz.

Comme dans le cas du gradient explicite, les points fixes de T sont caractérisés par $\nabla f(x) = 0$ et sont les minimiseurs de f, donc par hypothèse, il existe au moins un tel point fixe. Il reste maintenant à démontrer que $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$ ce qui est direct car par définition de x_{n+1} :

$$f(x_{n+1}) + \frac{1}{2}\|x_{n+1} - x_n\|^2 \leqslant f(x_n) + \frac{1}{2}\|x_n - x_n\|^2 = f(x_n). \tag{30}$$

On déduit donc que $(f(x_n))_{n\in\mathbb{N}}$ est une suite décroissante, comme elle est minorée on en déduit que $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$, ce qui conclut la preuve du théorème.

EXEMPLES.

2.3.3 Gradient à pas optimal

L'algorithme de descente de gradient à pas optimal consiste comme son nom l'indique à adapter le pas de descente à chaque étape de manière à faire décroitre au maximum la valeur de la fonctionnelle. Ainsi on définit une suite par $x_0 \in E$ et pour tout $n \in \mathbb{N}$:

$$x_{n+1} = x_n - \gamma_n \nabla f(x_n) \text{ avec } \gamma_n = \arg\min_{\gamma>0} f(x_n - \gamma\nabla f(x_n)) \tag{31}$$

Cette méthode converge vers l'unique minimiseur de f si f est $\alpha-$fortement convexe.

A DEMONTRER (ou pas) Livre de G. Allaire p339 de l'édition française.

Cette méthode a l'avantage de converger plus rapidement que la méthode de gradient à pas fixe en terme de nombre d'étapes mais chaque étape nécessite la résolution d'un problème d'optimisation qui peut être résolu de manière exacte ou approchée.

2.4 Gradient projeté, Algorithme d'Uzawa

Si on souhaite résoudre le problème d'optimisation sous contrainte suivant :

$$\min_{x\in K} f(x) \tag{32}$$

où K est un ensemble convexe fermé et f est différentiable et à gradient $L-$Lipschitz, il est possible d'adapter la méthode du gradient à pas fixe.

Théorème 7 *Si f est convexe et differentiable, si ∇f est L-Lipschitz, si K est un convexe fermé et si $\gamma < \frac{2}{L}$ alors pour tout $x_0 \in E$ la suite définie par*

$$x_{n+1} = P_K(x_n - \gamma \nabla f(x_n)) \tag{33}$$

converge vers un minimiseur de f.

Nous ne démontrons pas ici ce résultat pour deux raisons. La première est que c'est une simple adaptation de la preuve de la convergence de la méthode de descente de gradient à pas fixe et la seconde c'est que cette méthode de gradient projeté est un cas particulier d'un algorithme plus général qui sera traité plus loin, le Forward-Backward et dont la convergence sera démontrée.

La principale difficulté dans la mise en place de cette méthode est la projection sur K qui peut ne pas être explicite dans certains cas. En revanche si K est de la forme

$$K = \prod_{i=1}^{M} [a_i, b_i] \tag{34}$$

la projection correspond juste à une tronquature. Cette remarque est la base de l'algorithme Uzawa pour résoudre certains problèms de Lagrangiens augmentés.
Comme nous l'avons vu précédemment, si f est une fonction convexe et si les $(F_i)_{i \leqslant M}$ sont des fonctions convexes, résoudre le problème suivant

$$\min_{F_i(x) \leqslant 0, \forall i \in [\![1,M]\!]} f(x) \tag{35}$$

est équivalent à déterminer un point-selle du Lagrangien suivant :

$$\sup_{q \in (\mathbb{R}^+)^M} \inf_{v \in E} \mathcal{L}(v,q) = f(v) + \langle q, F(v) \rangle \tag{36}$$

On pose $\mathcal{G}$ la fonction définie par

$$\mathcal{G}(q) = \inf_{v \in E} \mathcal{L}(v,q) = f(v) + \langle q, F(v) \rangle. \tag{37}$$

Le couple (u,p) est un point-selle de $\mathcal{L}$ si et seulement si p est solution de

$$\inf_{q \in (\mathbb{R}^+)^M} -\mathcal{G}(q). \tag{38}$$

Ce problème est un problème de minimisation d'une fonction convexe sous contraintes convexes sur lesquelles la projection est aisément calculable. On peut donc utiliser un algorithme de gradient projeté pour le résoudre si on sait calculer le gradient de $\mathcal{G}$. Sous des hypothèses assez générales, $\mathcal{G}$ est différentiable et on peut exprimer simplement son gradient.
Nous n'entrerons pas ici dans les justifications théoriques mais nous pouvons donner des arguments formels.

Pour tout $q \in (\mathbb{R}^+)^M$, la fonctionnelle $u \mapsto f(u) + qF(u)$ est convexe et admet un unique minimiseur u_q (ce qui est, avouons le, une hypothèse assez forte), ainsi $\mathcal{G}(q) = f(u_q) + \langle q, F(u_q)\rangle$. Si on dérive formellement cette relation on obtient que

$$\mathcal{G}'(q) = F(u_q) + \langle \nabla f(u_q) + \langle q, F'(u_q)\rangle, u_q'\rangle = F(u_q) \tag{39}$$

On a ainsi une expression simple du gradient de $\mathcal{G}$ au point q, il s'agit de la valeur de F au point u_q.
On en déduit ainsi l'algorithme d'Uzawa permettant de résoudre (35) qui n'est autre qu'un gradient projeté. On considère $p_0 \in (\mathbb{R}^+)^M$ et pour tout $n \in \mathbb{N}$ on définit :

$$\left\{ \begin{array}{ccc} u_n & = & \min_{u\in E} f(u) + \langle p_n, F(u)\rangle \\ p_{n+1} & = & P_{(\mathbb{R}^+)^M}(p_n + \gamma F(u_n)) \end{array} \right. \tag{40}$$

où $\gamma < \frac{2}{L}$ si ∇f est $L-$Lipschitz.
EXEMPLES.

2.5 Gradient conjugué

La méthode du gradient conjugué est une méthode itérative permettant de résoudre de manière approchée et itératives le problème inverse suivant :

$$Ax = b \tag{41}$$

où A est une matrice symétrique définie positive de $\mathbb{R}^{n\times n}$ et b un vecteur de $\mathbb{R}^n$.
L'idée de l'algorithme est de partir d'un point $x_0 \in \mathbb{R}^n$ et d'un résidu $r_0 = b - Ax_0$ et de minimiser itérativement la fonction suivante :

$$\min_{x\in K_n} \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \tag{42}$$

où $K_n = Vect(r_0, Ar_0, \cdots, A^n r_0)$ est le sous-espace de Krilov de dimension $n+1$ associé à r_0.
Pour réaliser une telle minimisation sur cette famille d'espaces emboités on construit les suites (x_n) de telle sorte que le résidu $r_{n+1} = Ax_{n+1}$ soit orthogonal à l'espace de Krilov K_n. L'algorithme proprement dit ne nécessite que l'application de l'opérateur A une fois par étape peut être décrit de la manière suivante :

Définition 13 *Soit A une matrice définie positive, $x_0 \in \mathbb{R}^n$ on définit les trois suites (x_n, r_n, p_n) de la manière suivante :*

$$p_0 = r_0 = b - Ax_0, \ \ et\ \forall k > 0, \left\{ \begin{array}{ccc} x_{k+1} & = & x_k + \alpha_k p_k \\ r_{k+1} & = & r_k - \alpha_k A p_k \\ p_{k+1} & = & r_{k+1} - \beta_k p_k. \end{array} \right. \tag{43}$$

avec

$$\alpha_k = \frac{\|r_k\|^2}{\langle Ap_k, p_k\rangle} \ \ et\ \beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2} \tag{44}$$

En pratique cette méthode converge rapidement, c'est à dire que la valeur approchée x_k est très satisfaisante pour de petites valeurs de k et d'autant plus rapide que x_0 est bien choisi. On peut montrer que la convergence vers la valeur exacte x^* est linéaire et dépend du conditionnement de A.

Nous verrons par la suite que l'inversion de systèmes faisant intervenir des matrices symétriques définies positives est un calcul qui intervient dans un grand nombre d'algorithmes d'optimisation.

2.6 Méthode de Newton

la méthode de Newton est une méthode utilisée originellement pour approcher un zero régulier d'une fonction F de classe C^2 définie de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On dit que u est un zéro régulier de F si

$$F(u) = 0 \text{ et } F'(u) \text{ est inversible.} \tag{45}$$

Cette méthode s'appuie sur un développement de Taylor au voisinage d'un zéro de F et construit une suite à partir d'un élément x_0 du voisinage de u définie pour tout $n \in \mathbb{N}$ par

$$x_{n+1} = x_n - (F'(x_n))^{-1} F(x_n). \tag{46}$$

pour que cette suite soit bien définie il faut que la matrice $F'(x_n)$ soit bien inversible ce qui n'est assuré que si le point de départ x_0 de la méthode est suffisamment proche de u.

On a la proposition suivante :

Proposition 7 *Soit F une fonction de classe C^2 de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et u un zéro régulier de F. Il existe $\varepsilon > 0$ tel que si $\|x_0 - x\| \leqslant \varepsilon$, la méthode de Newton définie par* (46) *converge vers u et il existe une constante C telle que*

$$\|x_{n+1} - u\| \leqslant C \|x_n - u\|^2 . \tag{47}$$

La démonstration de cette proposition sera proposée en exercice.

2.7 Méthodes de pénalisation des contraintes et de smooting

3 Fonctions convexes non différentiables

Dans cette partie nous nous intéresserons à des problèmes d'optimisation convexes dits non lisses c'est à dire que la fonction F que nous voulons minimiser est non différentiable. Un cas particulier de non différentiabilité qui attirera notre attention est ce lui de la somme de deux fonctions convexes dont l'une au moins est non différentiable :

$$\min_{x \in E} F(X) = \min_{x \in E} f(x) + g(x) \tag{48}$$

où g est non différentiable. Les problèmes de minimisation de fonctions différentiables f sous contraintes convexes peuvent se réécrire de cette manière via l'indicatrice d'un convexe, ainsi

$$\min_{x\in K} f(x) \Longleftrightarrow \min_{x\in E} f(x) + i_K(x) \tag{49}$$

le point de vue que nous adopterons dans cette section et les algorithmes associés Mais ce n'est pas la seule situation. Pour résoudre nombre de problèmes de traitement d'images par exemple, il est fréquent de minimiser des fonctionnelles sommes d'un terme dit *d'attache aux données* et d'un terme *de régularisation* ce dernier étant construit de manière à promouvoir un a priori sur l'image à reconstruire est souvent non différentiable, c'est le cas de la minimisation ℓ_1 qui promeut la parcimonie ou de la variation totale ($g(x) = \|\nabla x\|_1$) qui vise à respecter les contours des images. Le LASSO en statistiques est un estimateur qui se calcule aussi en minimisant un telle fonctionnelle. Nous verrons que certains problèmes de deblurring, d'inpainting par exemple peuvent se résoudre en résolvant des problèmes de la forme :

$$\min_{x\in E} \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_1 \tag{50}$$

via un éventuel changement de variable.

Une méthode classique de débruitage est la minimisation de variation totale de l'image :

$$\min_{x\in E} \frac{1}{2}\|y - x\|_2^2 + \lambda \|\nabla x\|_1 \tag{51}$$

qui fait intervenir un terme non différentiable. Nous verrons que ce problème peut se reformuler sous la forme suivante :

$$\min_{p} \frac{1}{2}\|\text{div } p - z\|_2^2 + i_{\mathcal{B}_{\ell_\infty}(\lambda)}(z) \tag{52}$$

que l'on sait traiter en utilisant plusieurs algorithmes.
Certains problèmes de débruitage avec des bruits non gaussiens, speckle ou laplaciens nécessitent de prendre en compte des termes d'attaches aux données non régulières :

$$\min_{x\in E} \frac{1}{2}\|y - x\|_1 + \lambda \|D^* x\|_1 \tag{53}$$

où D^* est une transformée bien choisie.

Il est parfois utile d'écrire F comme la somme de plusieurs fonctions, même si ces dernières sont toutes non différentiables, car certains algorithmes dits de *splitting* permettent de minimiser la somme de fonctions en alternant des opérations élémentaires utilisant chacune des fonctions prises séparément. Nous verrons le cas par exemple de l'algorithme *Douglas Rachford* et de l'ADMM (Alternating Sescend Method of Multipliers)

3.1 Définitions

Définition 14 *Soit f une fonction définie de E dans $\mathbb{R} \cup +\infty$, la sous différentielle de f au point $x \in E$ est l'opérateur multivalué qui à x associe*

$$\partial f(x) = \{u \in E \text{ tels que } \forall y \in E,\ \langle y - x, u\rangle + f(x) \leqslant f(y)\}$$

Remarques :

- La sous-différentielle d'une fonction en un point peut être vide, ou être un ensemble convexe éventuellement réduit à un singleton.

- Si f est convexe et différentiable sur E en tout point $x \in E$, la sous-différentielle est réduite au singleton $\{\nabla f(x)\}$ gradient de f au point x .
- La sous différentielle de la fonction $x \mapsto -x^2$ est vide en tout point de E alors même qu'elle est différentiable en tout point. Ici f n'est pas convexe.
- On peut montrer qu'en tout point x de l'intérieur relatif du domaine de f **convexe**, la sous différentielle $\partial f(x)$ est non vide.
- Sur le bord du domaine les choses peuvent être plus complexes. La fonction f définie de $\mathbb{R}$ dans $\mathbb{R} \cap +\infty$ par

  $$f(x) = \begin{cases} -\sqrt{x} & \text{si } x \geqslant 0 \\ +\infty & \text{si } x < 0 \end{cases}$$

  est convexe, prend une valeur finie en 0 mais n'admet pas de sous différentielle en 0.

L'importance de la sous-différentielle prend tout son sens avec la règle de Fermat suivante :

Théorème 8 *Soit f une fonction propre définie de E dans $\mathbb{R} \cup +\infty$, on a*

$$Argminf = zer\ \partial f = \{x \in E \text{ tels que } 0 \in \partial f(x)\}.$$

Démonstration

$$x \in \text{Argmin} f \Leftrightarrow \forall y \in E,\ \langle y - x, 0\rangle + f(x) \leqslant f(y) \Leftrightarrow 0 \in \partial f(x)$$

Ainsi pour minimiser une fonctionnelle on s'attachera à chercher un zéro de la sous-différentielle. Si on sait pour des raisons de semi-continuité inférieure et de coercivité qu'une fonctionnelle admet (au moins) un minimum alors on sait qu'en ce minimum, la sous différentielle est définie et que 0 appartient à la sous-différentielle.
Cette règle est à mettre en relation avec le fait bien connu qu'au minimum global d'une fonction différentiable le gradient est nul.
Terminons ce paragraphe par un résultat sur la somme de sous différentielles

Lemma 6 *Soit $(f_i)_{i\leqslant p}$ une famille de fonctions sci propres et $x \in E$. On a*

$$\partial(\sum_{i=1}^{p} f_i)(x) \supset \sum_{i=1}^{p} \partial f_i(x)$$

De plus si

$$\bigcap_{1\leqslant i\leqslant p} inter\ (dom\ (f_i)) \neq \emptyset$$

alors on a égalité.

Voir Rockafellar [**?**].
L'hypothèse assurant l'égalité n'est pas très contraignante. Dans les cas pratiques intéressants, elle est toujours vérifiée.
Remarque :
Si J est la somme d'une fonction convexe différentiable f et d'une fonction convexe g, alors sous les hypothèses du lemme 1,

$$\partial J = \nabla f + \partial g.$$

L'unicité du minimiseur est due à la stricte convexité de la la fonction $y \mapsto \frac{1}{2}\|x-y\|^2$.

L'opérateur proximal peut être vu comme la généralisation du concept de projection sur un convexe dans la mesure où si f est l'indicatrice d'un convexe fermé C alors $\text{Prox}_f(x)$ est la projection de x sur C.
Il existe dans la litérature de nombreuses fonctions pour lesquelles l'opérateur proximal a été calculé. On peut noter que si la fonction f est séparable comme $x \mapsto \|x\|_p^p$ par exemple, l'opérateur proximal se calcule composante par composante, ce qui simplifie les calculs.
Dans la suite nous dirons **qu'une fonction f est simple quand son opérateur proximal se calcule facilement.**
On peut noter que si f n'est pas convexe, l'opérateur proximal peut ne pas être défini ou être multivalué. La convexité de f est une condition suffisante à son existence et son unicité, mais elle n'est pas nécessaire.
L'opérateur proximal possède plusieurs propriétés remarquables qui en font un outil privilégié pour la minimisation de fonction convexe.

Proposition 8 *Soit f une fonction convexe sur E, $(x,p) \in E^2$,*

$$p = Prox_f(x) \Leftrightarrow \forall y \in E,\ \langle y-p, x-p\rangle + f(p) \leqslant f(y) \tag{54}$$

Ainsi $p = Prox_f(x)$ est l'unique vecteur $p \in E$ tel que $x - p \in \partial f(p)$.

Cette dernière caractérisation est très importante. Ainsi la décomposition de $x = p + z$ avec $z \in \partial f(p)$ comme la somme d'un élément p de E et d'un élément de la sous différentielle de f au point p est unique.
Cette propriété remarquable est à l'origine de la notation

$$\text{Prox}_f(x) = (Id + \partial f)^{-1}(x)$$

à laquelle on peut donner un sens même si l'application ∂f est multivaluée.
Démonstration :

- Supposons que $p = \mathrm{Prox}_f(x)$. Soit $\alpha \in]0,1[$ et notons $p_\alpha = \alpha y + (1-\alpha)p$ alors en utilisant la définition de $\mathrm{Prox}_f(x)$ puis la convexité de f on a :

$$\begin{aligned} f(p) &\leqslant f(p_\alpha) + \frac{1}{2}\|x-p_\alpha\|^2 - \frac{1}{2}\|x-p\|^2 \\ &\leqslant \alpha f(y) + (1-\alpha)f(p) - \alpha\langle y-p, x-p\rangle + \frac{\alpha^2}{2}\|y-p\|^2 \end{aligned}$$

et donc
$$\langle y-p, x-p\rangle + f(p) \leqslant f(y) + \frac{\alpha}{2}\|y-p\|^2 .$$
En faisant tendre α vers 0 on en déduit la première implication.

- Réciproquement si on suppose que $\langle y-p, x-p\rangle + f(p) \leqslant f(y)$ on a

$$f(p)+\frac{1}{2}\|x-p\|^2 \leqslant f(y)+\frac{1}{2}\|x-p\|^2+\langle x-p, p-y\rangle+\frac{1}{2}\|p-y\|^2 = f(y)+\frac{1}{2}\|x-y\|^2$$

et on en déduit que $p = \mathrm{Prox}_f(x)$ en utilisant la définition de $\mathrm{Prox}_f(x)$.

Une propriété importante des points fixes de l'opérateur proximal est la suivante.

Proposition 9 *Soit f une fonction convexe propre définie sur E.*

$$\mathit{Fix}\,\mathit{Prox}_f = \mathit{Argmin} f.$$

Démonstration :

$$x = \mathrm{Prox}_f(x) \Leftrightarrow \forall y \in E,\ \langle y-x, x-x\rangle + f(x) \leqslant f(y) \Leftrightarrow x \in \mathrm{Argmin} f$$

Cette caractérisation permet de mettre au point des algorithmes itératifs utilisant l'opérateur proximal de f pour minimiser une fonctionnelle f.
Une autre propriété essentielle de l'opérateur proximal est son caractère firmly non expansive :

Proposition 10 *Si f est une fonction convexe propre définie sur E alors les applications Prox_f et $\mathit{Id} - \mathit{Prox}_f$ sont firmly nonexpansive, c'est à dire pour tout $(x, y) \in E^2$,*

$$\|\mathit{Prox}_f(x) - \mathit{Prox}_f(y)\|^2 + \|(x - \mathit{Prox}_f(x)) - (y - \mathit{Prox}_f(y))\|^2 \leqslant \|x-y\|^2$$

En particulier ces deux opérateurs sont nonexpansive (1-Lipschitz).

Démonstration :
Soit $(x, y) \in E^2$, si on note $p = \mathrm{Prox}_f(x)$ et $q = \mathrm{Prox}_f(y)$ alors par définition de p et q on a

$$\langle q-p, x-p\rangle + f(p) \leqslant f(q) \text{ et } \langle p-q, y-q\rangle + f(q) \leqslant f(p)$$

en additionnant ces deux inégalités on obtient

$$0 \leqslant \langle p-q, (x-p)-(y-q)\rangle$$

et on conclut en observant que

$$\begin{aligned}\|x-y\|^2 &= \|p-q+(x-p)-(y-q)\|^2\\ &= \|p-q\|^2+\|(x-p)-(y-q)\|^2+2\langle p-q,(x-p)-(y-q)\rangle\end{aligned}$$

Nous aurons besoin dans la suite de définir le *reflected prox* défini de la manière suivante

Définition 15 *Si f est convexe et propre, on définit le reflected prox ou Rprox par*

$$Rprox_f = 2Prox_f - Id$$

Pour montrer que le Rprox est 1-Lipschitz nous allons avoir besoin du lemme suivant :

Lemma 7 *Soit T un opérateur défini de E dans E. On a équivalence entre*

- *T est firmly nonexpansive et*
- *$R = 2T - Id$ est nonexpansive (1-Lipschitz)*

Démonstration :
Nous allons montrer que ces deux propriétés sont équivalentes à la positivité d'un produit scalaire. Soit $(x, y) \in E^2$ et notons $u = Tx - Ty$ et $v = Tx - x - (Ty - y)$.

$$\|x-y\|^2 = \|u-v\|^2 = \|u\|^2+\|v\|^2-2\langle u,v\rangle$$

donc le fait que T soit firmly nonexpansive est équivalent à la condition $\langle u, v\rangle \geqslant 0$. De plus

$$\|Rx-Ry\|^2 = \|u+v\|^2 = \|u\|^2+\|v\|^2+2\langle u,v\rangle = \|x-y\|^2+4\langle u,v\rangle$$

donc R est nonexpansive est équivalent à $\langle u, v\rangle \geqslant 0$ ce qui conclut la démonstration.

3.2 Calculs d'opérateurs proximaux

Il existe des tables qui recensent les opérateurs proximaux associés à de nombreuses fonctions. Il n'est pas question ici d'en faire une liste exhaustive mais on peut cependant mentionner quelques résultats utiles dont une partie est laissé en exercice.

1. Si f est la fonction indicatrice d'un convexe fermé C, son opérateur proximal est la projection sur C.
2. Si $f(x) = \lambda \|x\|_1$ alors l'opérateur proximal de f est un seuillage doux de seuil λ :

$$\mathrm{Prox}_f(x)_i = \begin{cases} x_i - \lambda & \text{si } x_i > \lambda \\ x_i + \lambda & \text{si } x_i < -\lambda \\ 0 & \text{sinon} \end{cases}$$

3. D'une manière générale si f est une fonction convexe différentiable, le *prox* p au point x de $f$est défini par l'équation implicite :

$$\nabla f(p) + p = x$$

Si $f(x) = \frac{1}{2}\|Ax - b\|^2$ alors

$$p = (Id + A^t A)^{-1}(x + A^t b)$$

Selon les propriétés de A il est plus ou moins facile de calculer p. Il existe des situations où ce calcul peut être fait de manière rapide, c'est le cas par exemple si on dispose d'une transformée rapide pour diagonaliser $A^t A$. Si A est un opérateur de convolution circulaire, il est diagonalisable dans une base de Fourier et donc p se calcule rapidement. C'est évidemment aussi le cas si A est une matrice orthogonale. Si on ne dispose pas d'algoritmes rapides, on peut utiliser des algorithmes itératifs qui donneront des valeurs approchées du prox en un nombre finie d'étapes.

4. Il n'existe pas de formule générale liant le prox de f et celui de g si $f(x) = g(Ax + b)$ pour un opérateur A quelconque, cependant dans le cas particulier où $AA^t = \frac{1}{\alpha}Id$, on a la relation suivante :

$$\text{Prox}_f(x) = x - \alpha A^t(Ax + b - \text{Prox}_{\alpha^{-1}g}(Ax + b))$$

Cette relation peut permettre de calculer l'opérateur proximal d'une fonction f de la forme $f(x_1, x_2) = g(x_1 + x_2)$ à partir de celui de g en considérant la matrice $A = [Id \quad Id]$. On a alors :

$$\text{Prox}_f(x)_i = x_i - \frac{1}{2}(x_1 + x_2) + \frac{1}{2}\text{Prox}_{2g}(x_1 + x_2).$$

Cette formule peut se généraliser à une somme de k termes.

5. Si $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ l'opérateur proximal de f s'exprime simplement à partir de ceux de f_1 et de f_2 dans la mesure où les minimisations sur x_1 et x_2 sont indépendantes, ainsi

$$\text{Prox}_f(x_1, x_2) = (\text{Prox}_{f_1}(x_1), \text{Prox}_{f_2}(x_2)).$$

3.3 Algorithme du point proximal

L'algorithme du point proximal comme son nom l'indique est l'algorithme qui consiste à appliquer de manière récursive, l'opérateur proximal de la fonction γg pour minimiser g sur E. Ainsi on choisit un $\gamma > 0$ et $x_0 \in E$ on construit la suite $(x_n)_{n\in\mathbb{N}}$ définie par

$$x_{n+1} = Tx_n = \text{Prox}_{\gamma g}(x_n)$$

Une telle suite converge vers un minimiseur de g, il suffit pour cela d'appliquer le lemme 4. En effet, nous avons vu que les points fixes de l'opétareur proximal sont les minimiseurs de g, que cet opéteur était 1-Lipschitz, de plus par construction, comme pour le gradient implicite on a :

$$g(x_{n+1}) + \frac{1}{2\gamma} \|x_n - x_{n+1}\|^2 \leqslant g(x_n)$$

ce qui implique que la suite de terme général $\|x_n - x_{n+1}\|^2$ tend vers 0.
En pratique cet algorithme est rarement utilisé dans la mesure où calculer un tel opérateur proximal est souvent aussi difficile que de minimiser g directement. On utilise plutôt l'opérateur proximal comme un ingrédient dans des algorithmes plus sophistiqués.

3.4 Forward-Backward et FISTA

L'algorithme appelé *Forward-Backward* (FB) est un algorithme qui permet de résoudre le problème d'optimisation suivant :

$$\min_{x\in E} F(x) = \min_{x\in E} f(x) + g(x)$$

où f est une fonction convexe différentiable dont le gradient est $L-$Lipschitz et où g est une fonction dont on sait calculer l'opérateur proximal. Cet algorithme consiste à alterner un descente de gradient explicite sur f et un opérateur proximal sur g. L'algorithme FISTA est une accélération de FB proposée par Beck et Teboulle basée sur des idées de Nesterov utilisant un terme inertiel. L'algorithme FB a beaucoup été utilisé au début des années 2000 pour résoudre des problèmes où g était une norme ℓ_1, l'opérateur proximal de g est dans ce cas un seuillage doux et le nom donné à cet algorithme était alors ISTA pour Iterative Shrinkage Thresholding Algorithm. Beck et Teboulle ont donné le nom de FISTA (Fast Iterative Shrinkage Thresholding Algorithm) à leur algorithme rapide en référence à ISTA mais c'est une accélération de FB et qui ne s'applique donc pas qu'au cas où g est une norme ℓ_1.

3.4.1 Forwward-Backward

L'algorithme Forward backward, dans sa version simple est défini par un point de départ $x_0 \in E$ et un $\gamma > 0$ de la manière suivante:

$$\forall n \in \mathbb{N} \quad x_{n+1} = Tx_n = \text{Prox}_{\gamma g}(x_n - \gamma \nabla f(x_n)) \tag{55}$$

Nous verrons plus tard qu'il existe également une version utilisant les itérations de Krasnoselskii-Mann.
L'algorithme découle de la proposition suivante :

Proposition 11 *Soit $F = f+g$ une fonctionnelle définie de E dans $\mathbb{R}\cup+\infty$ somme de deux fonctions convexes sci propres vérifiant les hypothèses du lemme 6 et tel que f soit différentiable et soit $\gamma > 0$. On a*

$$zeros(\partial F) = Fix(Prox_{\gamma g}(Id - \gamma \nabla f)) \tag{56}$$

Démonstration :

$$\begin{aligned} 0 \in \partial F(x) &\Leftrightarrow 0 \in \partial \gamma F(x) \\ &\Leftrightarrow 0 \in \nabla \gamma f(x) + \partial \gamma g(x) \\ &\Leftrightarrow -\gamma \nabla f(x) \in \partial \gamma g(x) \\ &\Leftrightarrow x - \gamma \nabla f(x) \in (Id + \partial \gamma g)(x) \\ &\Leftrightarrow x = \text{Prox}_{\gamma g}(x - \gamma \nabla f(x)) \end{aligned}$$

Ce qui conclut la démonstration.
Le fait d'avoir un opérateur T dont les points fixes sont les minimiseurs de F ne garantit bien évidemment que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un tel minimiseur. Mais c'est bien le cas comme l'indique le théorème suivant :

Théorème 9 *Soit $F = f + g$ une fonction somme de deux fonctions convexes, sci, coercives et bornées inférieurement. On suppose que f est différentiable et à gradient $L-$Lipschitz. Soit $\gamma < \dfrac{2}{L}$, soit $x_0 \in E$ et soit $(x_n)_{n\in\mathbb{N}}$ la suite définie pour tout $n \in \mathbb{N}$ par*

$$x_{n+1} = Prox_{\gamma g}(Id - \gamma \nabla f)(x_n) \tag{57}$$

Alors la suite $(x_n)_{\in\mathbb{N}}$ converge vers un minimiseur de F.

Démonstration :
Pour démontrer cette convergence, nous allons procéder comme pour la convergence de la méthode du gradient. C'est à dire que nous allons montrer que l'opérateur T est $1-$Lipschitz et que la suite de terme général $\|x_n - x_{n+1}\|$ tend vers 0. On concluera en utilisant le lemme 4.
Le fait que l'opérateur T est $1-$Lipschitz découle directement du fait que c'est une composition d'un opérateur proximal qui est 1-Lipschitz et d'un opérateur de la forme $Id - \gamma \nabla f$ qui est 1-Lipschitz si $\gamma \leqslant \frac{2}{L}$, comme il l'a été montré dans le lemme 5.

Pour démontrer le second point, nous allons introduire le concept des *surrogate functions*. En effet on peut montrer qu'à chaque étape de l'algorithme, à chaque calcul d'un nouveau terme de la suite on fait décroitre la valeur de la fonctionnelle F. De plus cette décroissance assure qua la série de terme général $\|x_{n+1} - x_n\|^2$ est sommable et donc tend vers 0 quand n tend vers $+\infty$.

Lemma 8 *La suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_{n+1} = Prox_{\gamma g}(x_n - \gamma \nabla f(x_n))$ vérifie la relation suivante :*

$$F(x_{n+1}) + \left(\frac{1}{\gamma} - \frac{L}{2}\right) \|x_{n+1} - x_n\|^2 \leqslant F(x_n). \tag{58}$$

Démonstration :
Par définition de x_{n+1},

$$x_{n+1} = \underset{x\in E}{\text{Argmin}}\, \gamma g(x) + \frac{1}{2} \|x - x_n + \gamma \nabla f(x_n)\|^2$$

On peut observer que x_{n+1} est également le minimiseur de la fonctionnelle suivante :

$$x_{n+1} = \operatorname*{Argmin}_{x \in E} g(x) + f(x_n) + \langle \nabla f(x_n), x - x_n \rangle + \frac{1}{2\gamma} \|x - x_n\|^2 \qquad (59)$$

D'après l'inégalité (119), on a pour tout $x \in E$,

$$f(x) \leqslant f(x_n) + \langle \nabla f(x_n), x - x_n \rangle + \frac{L}{2} \|x - x_n\|^2 \qquad (60)$$

Le terme de droite est une fonction qui majore F et qui coïncide avec F au point x_n. On appelle cette fonction une *surrogate function*.
De (59) et en utilisant le fait que la fonction considérée est $\frac{1}{\gamma}$-fortement convexe on déduit que

$$g(x_{n+1}) + f(x_n) + \langle \nabla f(x_n), x_{n+1} - x_n \rangle + \frac{1}{2\gamma} \|x_{n+1} - x_n\|^2 \leqslant F(x_n) - \frac{1}{2\gamma} \|x_{n+1} - x_n\|^2$$

En spécialisant (60) à x_{n+1} et en sommant avec l'inégalité précédente on obtient

$$F(x_{n+1}) + \frac{1}{\gamma} \|x_{n+1} - x_n\|^2 \leqslant F(x_n) + \frac{L}{2} \|x_{n+1} - x_n\|^2 .$$

Ce qui conclut la preuve du lemme et du théorème.

Nous verrons dans la section suivante qu'il est possible de contrôler la vitesse de décroissance de la fonctionnelle vers son minimum.

3.4.2 Fast Iterative Shrinkage thresholding Algorithm (FISTA)

L'idée de FISTA est d'appliquer l'opérateur $T = \mathrm{Prox}_{\gamma g} \circ (id - \gamma \nabla f)$ en utilisant un terme d'inertie pour améliorer les bornes sur les vitesses de convergence.

FISTA est définie par une suite $(t_n)_{n \in \mathbb{N}}$ de nombres réels supérieurs à 1 et par un point $x_0 \in E$. Soit $(t_n)_{n \geqslant 1}$ une suite de nombres réels positifs et $x_0 \in E$, les suites $(x_n)_{n \in \mathbb{N}}$, $(y_n)_{n \in \mathbb{N}}$ et $(u_n)_{n \in \mathbb{N}}$ sont définies par $y_0 = u_0 = x_0$ et pour tout $n \geqslant 1$,

$$x_n = T(y_{n-1}) \qquad (61)$$

$$u_n = x_{n-1} + t_n (x_n - x_{n-1}) \qquad (62)$$

$$y_n = \left(1 - \frac{1}{t_{n+1}}\right) x_n + \frac{1}{t_{n+1}} u_n. \qquad (63)$$

Le point y_n peut aussi être défini à partir des points x_n et x_{n-1} par

$$y_n = x_n + \alpha_n (x_n - x_{n-1}) \text{ with } \alpha_n := \frac{t_n - 1}{t_{n+1}} \qquad (64)$$

Pour des choix adéquats de $(t_n)_{n \geqslant 1}$ la suite $(F(x_n))_{n \in \mathbb{N}}$ converge vers $F(x^*)$, i.e la suite $(w_n)_{n \in \mathbb{N}}$, définie de la manière suivante

$$w_n := F(x_n) - F(x^*) \geqslant 0 \qquad (65)$$

tend vers 0 quand n tend vers l'infinie.

Plusieures preuves utilisent la variation locale de la suite $(x_n)_{n\in\mathbb{N}}$, que nous noterons δ_n:

$$\delta_n := \frac{1}{2}\left\|x_n - x_{n-1}\right\|_2^2. \tag{66}$$

la suite $(v_n)_{n\in\mathbb{N}}$ décrivant la distance entre u_n et un minimiseur fixé x^* de F sera aussi utile:

$$v_n := \frac{1}{2}\left\|u_n - x^*\right\|_2^2. \tag{67}$$

Pour compléter ces notations nous définissons également la suite $(\rho_n)_{n\geqslant 2}$, associée à $(t_n)_{n\geqslant 1}$, dont la positivité assure la vitesse de convergence de FISTA:

$$\rho_n := t_{n-1}^2 - t_n^2 + t_n. \tag{68}$$

Avant de détailler le choix de la suite $(t_n)_{n\geqslant 1}$ et donc de la suite $(\alpha_n)_{n\in\mathbb{N}}$ fait par Nesterov puis par Beck et Teboulle et les vitesses de convergence qui en découlent, nous allons énoncer quelques résultats préliminaires. Le premier est un lemme important sur l'opérateur T et est démontré en annexe :

Lemma 9 *Soit $\gamma \in]0, \frac{1}{L}]$, où L est la constante de Lipschitz de ∇f, et soit $\bar{x} \in E$ aet $\hat{x} = T\bar{x}$. Alors*

$$\forall x \in E \quad F(\hat{x}) + \frac{\left\|\hat{x} - x\right\|^2}{2\gamma} \leqslant F(x) + \frac{\left\|x - \bar{x}\right\|^2}{2\gamma}. \tag{69}$$

Si on applique ce lemme aux points $\bar{x} = y_n$, $\hat{x} = x_{n+1}$ et $x = x_n$, on déduit que

$$F(x_{n+1}) + \frac{\left\|x_{n+1} - x_n\right\|^2}{2\gamma} \leqslant F(x_n) + \frac{\alpha_n^2\left\|x_n - x_{n-1}\right\|^2}{2\gamma}. \tag{70}$$

Cette inégalité implique que si les α_n sont des éléments de $[0, 1]$ alors pour tout $n \geqslant 1$

$$w_{n+1} + \delta_{n+1} \leqslant w_n + \delta_n \tag{71}$$

c'est à dire que la suite $(w_n + \delta_n)_{n\in\mathbb{N}}$ est décroissante.

Le théorème suivant décrit quelques propriétés des suites $(w_n)_{n\in\mathbb{N}}$ et $(\delta_n)_{n\in\mathbb{N}}$ en fonction du choix de la suite $(\alpha_n)_{n\geqslant 1}$. Nous discuterons ensuite des différents choix possibles de sur-relaxation.

Théorème 10 *Pour tout $x_0 \in E$, si la suite $(t_n)_{n\in\mathbb{N}}$ vérifie*

$$\forall n \geqslant 2 \quad t_n^2 - t_n \leqslant t_{n-1}^2 \tag{72}$$

et $t_1 = 1$, si $\gamma \leqslant \frac{1}{L}$ alors pour tout $N \geqslant 2$,

$$t_{N+1}^2 w_{N+1} + \sum_{n=1}^{N} \rho_{n+1} w_n \leqslant \frac{v_0 - v_{N+1}}{\gamma}. \tag{73}$$

Démonstration :

En appliquant le lemme 9 avec $\bar{x} = y_n$, $\hat{x} = x_{n+1}$ et $x = (1 - \frac{1}{t_{n+1}})x_n + \frac{1}{t_{n+1}}x^*$, on obtient

$$F(x_{n+1}) + \frac{\left\| \frac{1}{t_{n+1}}u_{n+1} - \frac{1}{t_{n+1}}x^* \right\|_2^2}{2\gamma}$$
$$\leqslant F\left((1 - \frac{1}{t_{n+1}})x_n + \frac{1}{t_{n+1}}x^*\right) + \frac{\left\| \frac{1}{t_{n+1}}x^* - \frac{1}{t_{n+1}}u_n \right\|_2^2}{2\gamma}$$

En utilisant la convexité de F, on a

$$F(x_{n+1}) - F(x^*) - \left(1 - \frac{1}{t_{n+1}}\right)(F(x_n) - F(x^*)) \leqslant \frac{\|u_n - x^*\|_2^2}{2\gamma t_{n+1}^2} - \frac{\|u_{n+1} - x^*\|_2^2}{2\gamma t_{n+1}^2}$$

En utilisant les définitions de w_n et v_n ces inégalités peuvent être reformulées de la manière suivante

$$t_{n+1}^2 w_{n+1} - (t_{n+1}^2 - t_{n+1})w_n \leqslant \frac{v_n - v_{n+1}}{\gamma} \tag{74}$$

Sommer ces inégalités de $n = 1$ à $n = N$ conduit à

$$t_{N+1}^2 w_{N+1} + \sum_{n=1}^{N} \rho_{n+1} w_n \leqslant \frac{v_0 - v_{N+1}}{\gamma}. \tag{75}$$

ce qui conclut la preuve du théorème 10.

Une des conséquences directes de ce théorème est ainsi l'inégalité suivante valable pour tout $n \geqslant 1$:

$$F(x_n) - F(x^*) \leqslant \frac{\|x_0 - x^*\|^2}{2t_n^2\gamma} \tag{76}$$

De ce théorème on peut déduire différents résultats suivants les choix de la suite $(t_n)_{n\in\mathbb{N}}$:

1. Si $(t_n)_{n\in\mathbb{N}}$ est la suite constante égale à 1, la suite $(\alpha_n)_{n\in\mathbb{N}}$ est la suite nulle et l'algorithme se réduit à FB. La suite $(\rho_n)_{n\geqslant 2}$ est alors la suite constante égale à 1.
   On déduit du théorème 10 que la série de terme générale $(w_n)_{n\in\mathbb{N}}$ est convergente. Comme on sait de plus que la suite $(w_n)_{n\in\mathbb{N}}$ est décroissante, on en déduit que $w_n = o(\frac{1}{n})$.

2. Le choix original de Nesterov et celui de Beck et Teboulle et celui qui optimise la borne dans (76), c'est à dire le choix de la suite $(t_n)_{n\in\mathbb{N}}$ qui maximise t_n à chaque étape en respectant la condition (72). C'est à dire le choix
$$t_{n+1} = \sqrt{t_n^2 + \frac{1}{2}} + \frac{1}{2} \tag{77}$$

On peut alors montrer que $t_n \geqslant \frac{n+1}{2}$ et que $\lim_{n\to\infty} \frac{n}{2t_n} = 1$ i.e t_n est de l'ordre de $\frac{n}{2}$. Dans ce cas l'inégalité (76) devient

$$F(x_n) - F(x^*) \leqslant \frac{2\left\|x_0 - x^*\right\|^2}{(n+1)^2\gamma} \tag{78}$$

3. On a la même vitesse de décroissance et la même borne si on utilise $t_n = \frac{n+1}{2}$, ce qui correspond au choix $\alpha_n = \frac{n-1}{n+2}$.

4. Si on choisit $t_n = \frac{n+a-1}{a}$ avec $a > 2$, l'inégalité (76) devient

$$F(x_n) - F(x^*) \leqslant \frac{a^2\left\|x_0 - x^*\right\|^2}{2(n+a-1)^2\gamma} \tag{79}$$

   ce qui est localement moins bon que le choix de Nesterov ou le choix $t_n = \frac{n+1}{2}$, en revanche dans ce cas la suite ρ_n est de l'ordre de n et on déduit du théorème que la série de terme général nw_n est sommable, ce qui permet en fait d'avoir un comportement asymptotique de la suite meilleur que ce qu'induit la borne obtenue par (78) ou (79) pour un n donné.

Plus précisément pour ce choix de paramètre on a le théorème suivant :

Théorème 11 *Si $t_n = \frac{n+a-1}{a}$ avec $a > 2$ alors*

1. *La suite de terme général nw_n est sommable et $w_n = o(\frac{1}{n^2})$*
2. *La suite de terme général $n\delta_n$ est sommable et $\delta_n = o(\frac{1}{n^2})$*
3. *La suite $(x_n)_{n\in\mathbb{N}}$ converge vers un minimiseur x^* de F*

Démonstration :
Nous ne démontrerons pas le dernier point qui est technique, mais nous pouvons noter ici, qu'en 2016, il n'existe pas de preuve de convergence de la suite $(x_n)_{n\in\mathbb{N}}$ pour le choix original de Nesterov. La preuve de convergence s'appuyant sur la sommabilité des suites de terme général nw_n et $n\delta_n$.

La sommabilité de la série de terme général nw_n est une conséquence directe du Théorème 10, car pour ce choix de t_n on a

$$\rho_n = \frac{1}{a^2}((a-2)n + a^2 - 3a + 3)$$

qui est de l'ordre de n dès que $a > 2$.

Pour démontrer la convergence de la série de terme général $n\delta_n$ on reprend l'inégalité (70) qui peut se réécrire

$$\delta_{n+1} - \alpha_n^2\delta_n \leqslant \gamma(w_n - w_{n+1})$$

Si $t_n = \frac{n+a-1}{a}$, $\alpha_n = \frac{t_n - 1}{t_{n+1}} = \frac{n-1}{n+a}$.

En multipliant les inégalités par $(n+a)^2$ et en sommant de $n = 1$ à $n = N$ on obtient

$$\sum_{n=1}^{N}(n+a)^2(\delta_{n+1} - \alpha_n^2\delta_n) \leqslant \gamma \sum_{n=1}^{N}(n+a)^2(w_n - w_{n+1}),$$

ce qui donne

$$(N+a)^2\delta_{N+1} + \sum_{n=2}^{N}((n+a-1)^2 - (n+a)^2\alpha_n^2)\delta_n \leqslant$$
$$\gamma\left((a+1)^2 w_1 - (N+a)^2 w_{N+1} + \sum_{n=2}^{N}\left((n+a)^2 - (n+a-1)^2\right) w_n\right)$$

c'est à dire

$$(N+a)^2\delta_{N+1} + \sum_{n=2}^{N} a(2n-2+a)\delta_n$$
$$\leqslant \gamma\left((a+1)^2 w_1 - (N+a)^2 w_{N+1} + \sum_{n=2}^{N}(2n+2a-1)w_n\right)$$

D'après le premier point le membre de droite de cette inégalité est borné indépendemment de N, ce qui assure que la suite $(n\delta_n)_{n\in\mathbb{N}}$ appartient à $\ell_1(\mathbb{N})$. On en déduit également que $N^2\delta_{N+1}$ est borné, mais en fait on a mieux que cela. En effet, d'après (71) la suite de terme général $w_n + \delta_n$ est décroissante, or la série de terme général $n(w_n + \delta_n)$ est convergente. Ceci implique (Exercice !) que $w_n = o(\frac{1}{n^2})$ et $\delta_n = o(\frac{1}{n^2})$ ce qui conclut la preuve du théorème.

3.5 Algorithme de Douglas Rachford

Dans cette section on souhaite résoudre le problème

$$\min_{x\in E} F(x) = \min_{x\in E} f(x) + g(x)$$

où les fonctions f et g sont convexes, propres, sci, coercives et que F est bornée inférieurement. Ici on ne fait pas d'hypothèses de différentiabilité sur f.
De plus on suppose qu'on sait calculer les opérateurs proximaux de f et de g.
Une première chose à noter est qu'il existe un minimiseur de F sous ces hypothèses.

Comme pour l'algorithme Forward-Backward, on va identifier un opérateur 1-Lipschitz dont les points fixes sont associés aux minimiseurs de F. Dans le cas de Forward-Backward ces points points fixes sont les minimiseurs de F, pour l'algorithme Douglas Rachford, le minimiseurs sont les images par un opérateurs de ces points fixes. Contrairement à FB, on ne pourra pas se contenter d'itérer l'opérateur T pour en approcher un point fixe. Cependant il existe un moyen de construire une suite qui converge vers un point fixe d'un opérateur T qui en possède au moins 1 et qui est 1-Lipschitz en utilisant un procédé de mélange. Il s'agit de lalgorithme de Krasnosel'skii-Mann que nous présentons maintenant. On peut noter que cet algorithme peut également être utilisé pour Forward-Backward.

Théorème 12 *Algorithme de Krasnosel'skii-Mann.*
Soit D un convexe fermé non vide de E et soit T, $D \to D$ un opérateur 1-Lipschitz tel que l'ensemble des points fixes de T est non vide. Soit $(\lambda_n)_{n\in\mathbb{N}}$ une suite de réels de $[0,1]$ telle que $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$, et soit $x_0 \in D$. On définit

$$\forall n \in \mathbb{N}, \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$$

alors la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un point fixe de T.

Remarque
Si E n'est pas de dimension finie on a seulement une convergence faible vers un point fixe et une convergence forte de $Tx_n - x_n$ vers 0.
Démonstration
Comme D est convexe la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie.
Soit y un point fixe de T et $n \in \mathbb{N}$,

$$\begin{aligned}
\|x_{n+1} - y\|^2 &= \|(1-\lambda_n)(x_n - y) + \lambda_n(Tx_n - y)\|^2 \\
&= (1-\lambda_n)\|x_n - y\|^2 + \lambda_n \|Tx_n - Ty\|^2 - \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2 \\
&\leqslant \|x_n - y\|^2 - \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2 .
\end{aligned}$$

On en déduit que

$$\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2 \leqslant \|x_0 - y\|^2 . \tag{80}$$

De plus

$$\begin{aligned}
\|Tx_{n+1} - x_{n+1}\| &= \|Tx_{n+1} - Tx_n + (1-\lambda_n)(Tx_n - x_n)\| \\
&\leqslant \|x_{n+1} - x_n\| + (1-\lambda_n)\|Tx_n - x_n\| \\
&\leqslant \|Tx_n - x_n\|
\end{aligned}$$

Comme $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$ et que la suite $(\|Tx_n - x_n\|)_{n\in\mathbb{N}}$ est décroissante on déduit de (80) que $(Tx_n - x_n)_{n\in\mathbb{N}}$ tend vers 0.
La suite $\|x_n - y\|$ étant décroissante on en déduit que tous les éléments de la suite $(x_n)_{n\in N}$ appartiennent à la boule fermée de centre y et de rayon $\|y - x_0\|$ qui est un compact car E est de dimension finie.
On peut donc extraire une sous suite $(x_{n_k})_{k\in\mathbb{N}}$ convergente vers un point x de E.
Comme $Tx_n - x_n$ tend vers 0, la suite $(Tx_{n_k} - x_{n_k})_{k\in\mathbb{N}}$ tend vers 0, on en déduit que la suite $(Tx_{n_k})_{k\in\mathbb{N}}$ converge également vers x. Comme cette suite converge aussi vers Tx on déduit que $x = Tx$ est un point fixe de T.

Comme x est un point fixe de T, la suite $(\|x_n - x\|)_{n\in\mathbb{N}}$ est décroissante. De plus cette suite admet une sous suite qui tend vers 0, on en déduit donc que la suite $(x_n)_{n\in\mathbb{N}}$ tend vers x point fixe de T, ce qui conclut la démonstration du théorème.

L'algorithme de Douglas-Rachford s'appuie sur la proposition suivante :

Proposition 12 *Soit $F = f + g$ une fonctionnelle définie de E dans $\mathbb{R}\cup+\infty$ somme de deux fonctions convexes sci propres vérifiant les hypothèses du lemme 6 et soit γ un réel strictement positif. On a*

$$zeros\ (\partial F) = Prox_{\gamma g}\left(Fix(Rprox_{\gamma f} Rprox_{\gamma g})\right). \tag{81}$$

Démonstration :
Sous les hypothèses du lemme 6 on

$$\begin{aligned} 0 \in \partial F(x) &\Leftrightarrow 0 \in \partial\gamma F(x) \\ &\Leftrightarrow 0 \in \partial\gamma f(x) + \partial\gamma g(x) \\ &\Leftrightarrow \exists z \in E \text{ tel que } -z \in \partial\gamma f(x) \text{ et } z \in \partial\gamma g(x) \\ &\Leftrightarrow \exists y \in E \text{ tel que } x - y \in \partial\gamma f(x) \text{ et } y - x \in \partial\gamma g(x) \end{aligned}$$

On peut écrire $x - y \in \partial\gamma f(x)$ sous la forme $2x - y \in (Id + \partial\gamma f)(x)$. De même, la relation $y - x \in \partial\gamma g(x)$ peut s'écrire $y \in (Id + \gamma\partial g)(x)$ et donc $x = \mathrm{Prox}_{\gamma g}(y)$, ce qui donne

$$0 \in \partial F(x) \Leftrightarrow \exists y \in E \text{ tel que } 2x - y \in (Id + \partial\gamma f)(x) \text{ et } x = \mathrm{Prox}_{\gamma g}(y).$$

En utilisant la définition du Rprox on obtient :

$$\begin{aligned} 0 \in \partial F(x) &\Leftrightarrow \exists y \in E \text{ tel que } x = \mathrm{Prox}_{\gamma f}(\mathrm{Rprox}_{\gamma g} y) \text{ et } x = \mathrm{Prox}_{\gamma g}(y) \\ &\Leftrightarrow \exists y \in E \text{ tel que } y = 2x - \mathrm{Rprox}_{\gamma g} y = \mathrm{Rprox}_{\gamma f}(\mathrm{Rprox}_{\gamma g} y) \text{ et } x = \mathrm{Prox}_{\gamma g}(y) \\ &\Leftrightarrow \exists y \in E \text{ tel que } y \in Fix\left(\mathrm{Rprox}_{\gamma f}\mathrm{Rprox}_{\gamma g}\right) \text{ et } x = \mathrm{Prox}_{\gamma g}(y) \end{aligned}$$

Ce qui conclut la démonstration de la Proposition.
De cette proposition on peut tirer un algorithme de minimisation de F en effet si on arrive à trouver un point fixe de $T = \mathrm{Rprox}_{\gamma f}\mathrm{Rprox}_{\gamma g}$, il suffira d'appliquer $\mathrm{Prox}_{\gamma f}$ pour obtenir un minimiseur de F. Comme cet opérateur T est 1-Lipschitz. Nous verrons qu'il est possible d'utiliser un algorithme dit de Krasnosel'skii-Mann pour en trouver un point fixe.

Théorème 13 *Soit f et g deux fonctions convexes, sci, propres, coercives et bornées inférieurement. Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite d'éléments de $[0,2]$ telle que $\sum_{n\in\mathbb{N}} \mu_n(2-\mu_n) = +\infty$, soit $\gamma > 0$, soit $x_0 \in E$ et $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ et $(z_n)_{n\in\mathbb{N}}$ les suites définie par*

$$\forall n \in \mathbb{N} \left\{ \begin{array}{rc} y_n = & Prox_{\gamma g}(x_n), \\ z_n = & Prox_{\gamma f}(2y_n - x_n), \\ x_{n+1} = & x_n + \mu_n(z_n - y_n) \end{array} \right. \tag{82}$$

Alors il existe $x \in E$ minimiseur de $(f+g)$ tel que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x.

Démonstration :
On pose $T = \text{Rprox}_{\gamma f}\text{Rprox}_{\gamma g}$. On sait que cet opérateur est 1-Lipschitz car composé de deux opérateurs 1-Lipschitz. De plus on sait que l'ensemble de ses points fixes est égal à l'ensemble des minimiseurs de $f+g$ qui est non vide d'après les hypothèses faites sur f et g. On remarque que $x_{n+1} = x_n + \frac{\mu_n}{2}(Tx_n - x_n)$. On conclut en appliquant le Théorème de Krasnosel'skii-Man.

L'algorithme de Douglas-Rachford peut être exprimé de plusieurs manières dans la littérature. Il faut être vigilent pour bien le reconnaitre. Il est fréquent par exemple que les paramètres μ_n soient fixés à 1, l'algorithme s'exprime alors de la manière suivante :

$$\forall n \in \mathbb{N} \begin{cases} y_n = & \text{Prox}_{\gamma g}(x_n), \\ z_n = & \text{Prox}_{\gamma f}(2y_n - x_n), \\ x_{n+1} = & x_n + z_n - y_n \end{cases} \tag{83}$$

Mais en se passant de la variable z_n on obtient alors la description des suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ de la manière suivante :

$$\forall n \geqslant 1 \begin{cases} x_n = & x_{n-1} + \text{Prox}_{\gamma f}(2y_{n-1} - x_{n-1}) - y_{n-1}, \\ y_n = & \text{Prox}_{\gamma g}(x_n) \end{cases} \tag{84}$$

On peut aussi introduire come variable auxilliaire $u_n = \text{Prox}_{\gamma f}(2y_{n-1} - x_{n-1})$, changer l'ordre de mise à jour des variables et réécrire l'algorithme de la manière suivante :

$$\forall n \geqslant 1 \begin{cases} u_n = & \text{Prox}_{\gamma f}(2y_{n-1} - x_{n-1}) \\ y_n = & \text{Prox}_{\gamma g}(x_{n-1} + u_n - y_{n-1}), \\ x_n = & x_{n-1} + u_n - y_{n-1} \end{cases} \tag{85}$$

Beaucoup de choix sont possibles, on peut citer pour terminer un changement de variables qui apparait quelque fois dans la littérature : $w_n = y_n - x_n$, on a alors

$$\forall n \geqslant 1 \begin{cases} u_n = & \text{Prox}_{\gamma f}(y_{n-1} + w_{n-1}) \\ y_n = & \text{Prox}_{\gamma g}(u_n - w_{n-1}), \\ w_n = & w_{n-1} + y_n - u_n \end{cases} \tag{86}$$

Toutes ces formulations sont ainsi équivalentes. Nous verrons par la suite comment il est possible d'adapter cet algorithme dans le cas où le problème à minimiser est de la forme

$$\min_{x\in E} f(x) + g(Ax)$$

où A est une application linéaire.

3.6 Alternating Direction Method of Multipliers (ADMM)

L'ADMM est un algorithme destimé à résoudre un problème d'optimisation de la forme :

$$\min_{(x_1,x_2)\in E,\, A_1x_1+A_2x_2=b} f_1(x_1) + f_2(x_2) \tag{87}$$

où A_1 et A_2 sont deux applications linéairesà valeurs dans $\mathbb{R}^m$, b un vecteur de $\mathbb{R}^m$, f_1 et f_2 deux fonctions convexes, propres et sci.
C'est un cadre assez général qui contient le cas $x_2 = x_1$ et même $x_2 = Ax_1$. Nous présentons ici l'algorithme sans en donner de preuve de convergence dans un cadre général. Nous verrons que dans le cas particulier où $A_1 = Id$, $A_2 = -Id$ et $b = 0$, c'est à dire le cas $x_1 = x_2$, cet algorithme se réduit à l'algorithme de Douglas-Rachford.

Algorithme ADMM
Soit $(x_1^0, x_2^0) \in E$, soit $\gamma > 0$ et $z^0 \in \mathbb{R}^m$. On définit les suites $(x_1^n)_{n\in\mathbb{N}}$, $(x_2^n)_{n\in\mathbb{N}}$ et $(z^n)_{n\in\mathbb{N}}$ de la manière suivante $\forall n \geqslant 1$:

$$\begin{cases} x_1^n = & \arg\min_x f_1(x) + \langle z^{n-1}, A_1 x\rangle + \frac{1}{2\gamma}\left\|A_1 x + A_2 x_2^{n-1} - b\right\|^2 \\ x_2^n = & \arg\min_y f_2(y) + \langle z^{n-1}, A_2 y\rangle + \frac{1}{2\gamma}\left\|A_1 x_1^n + A_2 y - b\right\|^2 \\ z^n & = z^{n-1} + \gamma(A_1 x_1^n + A_2 x_2^n - b) \end{cases} \tag{88}$$

Cet algorithme permet de résoudre le problème (87) au sens suivant :

Théorème 14 *Si f_1 et f_2 sont deux fonctions propres, sci, convexes et coercives alors :*

1. *La suite $(f_1(x_1^n)+f_2(x_2^n))_{n\in\mathbb{N}}$ converge vers la valeur minimale de f_1+f_2.*
2. *Les suites $(x_1^n)_{n\in\mathbb{N}}$ et $(x_2^n)_{n\in\mathbb{N}}$ convergent.*
3. *La suite $(Ax_1^n + Ax_2^n - b)_{n\in\mathbb{N}}$ tend vers 0.*

Nous ne démontrons pas ici ce résultat dont la preuve la plus simple consite à montrer que résoudre (87) est équivalent à résoudre un problème dual de (87) par un algorithme de Douglas Rachford.

On peut cependant faire plusieurs remarques sur cet algortihme :

1. Son nom vient du fait qu'il peut être vu comme une variante d'un algorithme connu sous le nom de méthode du Lagrangien augmenté. Si on remplace les mises à jours de x_1 et de x_2 par une mise à jour conjointe :

   $$(x_1^n, x_2^n) = \arg\min_{(x,y)\in E} f_1(x)+f_2(y)+\langle z^{n-1}, A_1x+A_2y\rangle+\frac{1}{2\gamma}\left\|A_1x + A_2y - b\right\|^2$$

   on retrouve la méthode du Lagrangien augmenté qui consiste à pénéliser les contraintes avec d'une part un multiplicateur de Lagrange z et un terme quadratique. Un des problème de cette méthode et qu'une telle minimisation conjointe est souvent très difficile à réaliser. L'ADMM découple ce problème en optimisant d'abord sur la première variable x_1 puis sur la variable x_2.

2. La variable z peut être interprétée comme un multiplicateur de Lagrange qui est mis à jour à chaque étape.

3. Si $A_1 = Id$, $A_2 = -Id$, $b = 0$ et $\gamma = 1$, l'algorithme peut s'exprimer à l'aide d'opérateurs proximaux, plus précisément il devient :

$$\begin{cases} x_1^n = & \text{Prox}_{f_1}(x_2^{n-1} - z^{n-1}) \\ x_2^n = & \text{Prox}_{f_2}(x_1^n + z^{n-1}) \\ z^n & = z^{n-1} + (x_1^n - x_2^n) \end{cases} \tag{89}$$

   ce qui est bien une des formes du Douglas-Rachford, voir (86).

4. Dans un cadre général, aucune des deux premières mises à jour n'est simple à effectuer. Si $A_1 = Id$ la mise à jour de x_1 revient à un calcul de prox de f, même chose si $A_2 = Id$ pour la seconde mise à jour. Si l'un des deux opérateurs n'est pas l'identité, il faut souvent utiliser un algorithme itératif pour effectuer la minimisation. On a alors des boucles internes. Chaque situation doit être étudiée au cas par cas, en fonction des fonctions f_1 et f_2 et des opérateurs A_1 et A_2. les performances de l'ADMM dépendront alors des choix qui seront faits pour réaliser ces 2 minimisations. On peut noter toutefois que les deux problèmes de minimisation peuvent être traités par FB ou FISTA si on dispose des opérateurs proximaux de f_1 et f_2. En effet le terme quadratique est différentiable et son gradient est explicite.

5. Il est possible de montrer que réaliser un ADMM est équivalent à effectuer un Douglas-Rachford sur un autre problème, appelé problème dual de (87) et cela quelque soit le choix des opérateurs A_1 et A_2 mais nous sortons ici du cadre de cours.

3.7 Reformulation de problèmes d'optimisation

Certains problèmes d'optimisation se présentent sous une forme que l'on peut traiter directement via un algorithme tel que le Forward-Backward, ou Douglas-Rachford ou l'ADMM. Par Exemple si on considère le problème du LASSO :

$$\arg\min_{x \in E} \frac{1}{2} \|y - Ax\|^2 + \lambda \|x\|_1 \tag{90}$$

On peut choisir $f(x) = \frac{1}{2} \|y - Ax\|^2$ et $g(x) = \lambda \|x\|_1$ et utiliser Forward-Backward ou FISTA.
Si on veut utiliser directement un Douglas-Rachford ou un ADMM sur ces deux fonctions, il faudra calculer un gradient implicite sur f, ce qui nécessite de résoudre une équation du type $p = (Id + A^t A)^{-1}(z)$ à chaque étape. Nous avons vu que s'il existait une transformée rapide pour diagonaliser rapidement $A^t A$ une telle étape pouvait être rapide mais si ce n'est pas le cas, la mise en place de ces deux algorithmes est moins simple mais est toujours possible avec des algorithmes itératifs.
Si on considère un problème de la forme

$$\arg\min_{x \in E} f(x) + g(Ax) \tag{91}$$

où A est un opérateur linéaire quelconque et que g n'est pas différentiable, on ne peut a priori pas utiliser ni FB, ni FISTA, ni même Douglas-Rachford ou l'ADMM directement même si on dispose de forme explicite des opérateurs proxuimaux de f et g. En revanche on peut réécrire ce problème de la manière suivante pour utiliser un ADMM

$$\arg \min_{x_2 - Ax_1 = 0} f(x_1) + g(x_2) \tag{92}$$

Dans ce cas, $A_2 = Id$ et une mise à jour de l'ADMM se calcule via un opérateur proximal et l'autre doit être traitée à part, soit en utilisant des propriétés de l'opérateur A, soit en utilisant FB ou FISTA.
On peut également réécrire ce problème de la manière suivante

$$\arg \min_{x_1, x_2} f(x_1) + g(x_2) + i_{x_2 = Ax_1}(x_1, x_2) \tag{93}$$

Si on considère $f_1(x_1, x_2) = f(x_1) + g(x_2)$ et $f_2(x_1, x_2) = i_{x_2=Ax_1}(x_1, x_2)$, on peut appliquer un Douglas-Rachford, en effet l'opérateur proximal de f_1 s'exprime simplement à partir de celui de f et de g (voir la section sur le calcul des opérateurs proximaux) et l'opérateur proximal de l'indicatrice est une projection sur un espace vectoriel. Cette projection nécessite la résolution d'un système de la forme $p = (I + A^t A)^{-1} z$ qui peut être plus ou moins facile selon l'opérateur A mais il est toujours possible de le résoudre au moins de manière approchée.
On peut noter que cette manière de réécrire le problème est également possible même si $A = Id$, on peut alors appliquer un Douglas-Rachford non pas sur f et g mais sur $f(x_1) + g(x_2)$ d'un côté et sur une indicatrice de l'autre.

4 Dualité

La notion de dualité est un terme général en mathématiques qui recouvre des objets et des concepts différents. Nous allons ici développer une notion de dualité d'une grande utilité en optimisation, la notion de dualité dite *min-max* qui permet d'associer à un problème d'optimisation dit primal, parfois difficile à résoudre pour diverses raisons à d'autres problèmes qui lui sont liés et qui permettront de le résoudre plus facilement. Cette notion de dualité n'est pas valable uniquement pour les fonctions convexes même si nous verrons qu'elle est particulièrement bien adpatée à la résolution de problèmes convexes. Cette section s'appuie essentiellement sur le travail de Jean-Charles Gilbert de l'INRIA Rocquencourt.

4.1 Dualité *min-max*

4.1.1 Fonction de couplage et problème dual

Soit f une fonction définie d'un espace euclidien E à valeurs dans $\mathbb{R} \cup \{+\infty\}$, non nécessairement convexe. On considère le problème d'optimisation suivant :

$$\inf_{x \in E} f(x) \tag{P}$$

qu'on l'on appelle problème *primal* et dont les solutions sont appelées *solutions primales*. On suppose que l'on peut écrire f comme le supremum d'une fonction φ définie de $E \times Y$ où $Y = \mathbb{R}^m$ est un espace euclidien a priori différent de E :

$$f(x) = \sup_{y \in Y} \varphi(x, y) \tag{94}$$

On appelle cette fonction φ une fonction de couplage. On peut ainsi réécrire le problème primal sous la forme

$$\inf_{x \in E} \sup_{y \in Y} \varphi(x, y) \tag{95}$$

On appelle alors problème dual associé à (P) le problème (D) obtenu en inversant l'ordre du supremum et de l'infimum

$$\sup_{y \in Y} \inf_{x \in E} \varphi(x, y) \tag{D}$$

Les solutions $\bar{y}$ de ce problème sont appelées *solutions duales*.
Le problème dual consiste donc à minimiser ce que l'on appelle la *fonction duale* δ associée à φ :

$$\delta \, : y \mapsto \delta(y) := -\inf_{x \in E} \varphi(x, y)$$

Il faut noter qu'une telle fonction n'est a priori pas explicite et que son évaluation en chaque $y \in Y$ peut nécessiter la résolution d'un problème d'optimisation. Nous verrons par la suite que pour certaines fonctions $\varphi(x, y)$ bien choisies il peut exister des formes explicites à la fonction δ. Le problème dual s'écrit ainsi

$$-\inf_{y \in Y} -\delta(y)$$

On peut donner donner ici deux exemples simples de fonctions de couplage φ :

1. On veut minimiser une fonction f sur E sous la contrainte $Kx = 0$ où K est une application de E dans $Y = \mathbb{R}^m$ pas nécessairement linéaire ou convexe. Le problème d'optimisation peut s'écrire

   $$\inf_{Kx=0} \bar{f}(x) = \inf_{x \in E} \bar{f}(x) + i_{Kx=0}(x) \tag{96}$$

   Le problème original sous contrainte peut être vu comme un problème de minimisation sans contrainte de la fonction $f(x) = \bar{f}(x) + i_{Kx=0}(x)$ et cette fonction peut s'écrire comme un suprémum d'une fonction de couplage :

   $$f(x) = \sup_{y \in Y} \bar{f}(x) + \langle y, Kx \rangle. \tag{97}$$

   Il est clair ici que la convexité de f n'intervient pas dans la construction de la fonction de couplage $\varphi(x, y) = \bar{f}(x) + \langle y, Kx \rangle$.
   On a alors par définition :

   $$\delta(y) = \inf_{x \in E} \bar{f}(x) + \langle y, Kx \rangle$$

   Si f est une fonction convexe et si K est un opérateur linéaire alors le minimiseur $\bar{x}$ de cette fonctionnelle est défini par $\nabla \bar{f}(\bar{x}) = -K^* y$.

2. Si f est convexe, propre et sci, on sait qu'elle est égale en tout point à sa biconjuguée de Fenchel f^{**}, ce qui permet d'écrire :

$$f(x) = \sup_{y \in E} \langle y, x \rangle - f^*(y) \tag{98}$$

On peut remarquer que dans cet exemple la variable duale y appartient au même espace que la variable primale x.
L'évaluation de la fonction duale δ est ici très simple dans la mesure où

$$\delta(y) = \inf_{x \in E} \langle y, x \rangle - f^*(y) = \begin{cases} -f^*(0) & \text{si } y = 0 \\ +\infty & \text{sinon} \end{cases}$$

On peut faire quelques remarques :

- La notion de problème dual dépend du choix de la fonction de couplage. Il existe la plupart du temps un grand choix de fonctions de couplage et donc autant de dualisations du problème primal. Dans le premier exemple du problème contraint, on a dualisé les contraintes $Kx = 0$ mais on aurait aussi pu décider d'écrire une partie des contraintes sous forme d'une indicatrice et l'incorporer à la fonction f et de dualiser une autre partie des contraintes. De même dans le cas des fonctions convexes, le choix de la conjuguaison pour définir φ n'est pas le seul. Nous verrons qu'on peut utiliser la conjugaison de différentes manières quand le problème est structuré (par exemple si la fonction à minimiser est la somme de deux fonctions) pour obtenir des problèmes duaux différents.

- L'intérêt du problème dual réside dans le fait qu'il doit permettre de résoudre le problème primal, il faut donc que les deux problèmes aient un lien entre eux, ce qui peut ne pas être le cas si la fonction φ est mal choisie. Nous verrons que dans les cas favorables les couples $(\bar{x}, \bar{y}) \in E \times Y$ où $\bar{x}$ est $\bar{y}$ sont des solutions respectives du problème primal et du problème dual, sont des *points-selles* de la fonction de couplage φ.

- Il est intéressant de résoudre le problème dual si ce dernier est plus facile à résoudre que le problème primal. Nous verrons que dans certains cas, on peut passer d'un problème primal contraint à un problème dual non contraint ou à un problème qui permet l'utilisation d'algorithmes efficaces alors qu'il n'est pas possible d'utiliser les mêmes algorithmes sur le problème primal.

- Certains algorithmes de résolution de (P) par dualité cherche à résoudre le problème dual (D) en générant une suite (y_k) où y_{k+1} est calculé à partir de y_k et de la solution éventuellement approchée du problème primal interne suivant :

$$\inf_{x \in E} \varphi(x, y)$$

Dans ce cas il est important que les problèmes primaux internes soient résolubles de manière explicite ou au moins en un temps raisonnable.

> Dans le cas du Lagrangien d'une fonction différentiable, les problèmes primaux internes sont différentiables et sans contraintes, contrairement au problème primal (P).

Sans aucune hypothèse sur φ il n'y a aucune raison que le problème dual permette de résoudre le problème dual. Néanmoins il existe des relations entre ces deux problèmes.

4.1.2 Relations entre les problèmes primal et dual, point-selle

La proposition suivante indique que la *valeur* du problème dual est en toute généralité inférieure à celle du problème primal et cela sans aucune hypothèse sur la fontion de couplage φ :

Proposition 13 *Dualité faible :*

$$\sup_{y\in Y}\inf_{x\in E}\varphi(x,y) \leqslant \inf_{x\in E}\sup_{y\in Y}\varphi(x,y). \tag{99}$$

La démonstration est immédiate et découle de la définition de l'infimum et du supremum.
Si on définit la *valeur* du problème primal

$$val(P) := \inf_{x\in E}\sup_{y\in Y}\varphi(x,y)$$

et la *valeur* du problème dual

$$val(D) := \sup_{y\in Y}\inf_{x\in E}\varphi(x,y),$$

on appelle saut (ou gap) de dualité la différence entre ces deux valeurs

$$\text{Saut de dualité } := val(P) - val(D).$$

On dira abusivement qu'il n'y a pas de saut de dualité quand ce dernier sera nul. Une telle situation est bien entendu la plus favorable et si ce saut est non nul, les solutions éventuelles des deux problèmes sont rarement liées. En revanche l'existence d'un point-selle de φ va permettre à la fois d'assurer l'existence de solutions aux problèmes primal et dual et de les lier :

Définition 16 *Soit X et Y deux ensembles et φ une fonction définie dans $X\times Y$ et à valeurs dans $\mathbb{R}\cup\{+\infty\}$. On dit que $(\bar{x},\bar{y})\in X\times Y$ est un point-selle de φ sur $X\times Y$ si on a pour tout $(x,y)\in X\times Y$*

$$\varphi(\bar{x},y) \leqslant \varphi(\bar{x},\bar{y}) \leqslant \varphi(x,\bar{y}).$$

Ainsi la fonction $x\mapsto\varphi(x,\bar{y})$ atteint un minimum en $\bar{x}$ et la fonction $y\mapsto\varphi(\bar{x},y)$ atteint son maximum en $\bar{y}$.

On peut faire deux remarques sur cette définition :

- Les ensembles X et Y peuvent ne pas être des espaces vectoriels, ce qui peut avoir une importance quand on veut étudier des extrema locaux de fonctions non convexes.

- Une fonction de couplage φ peut ne pas admettre de point-selle quand bien même elle serait convexe en x et concave en y. En effet les infimums et les supremums peuvent ne pas être atteints.

La proposition suivante précise le lien entre les points-selles de la fonction de couplage φ et les solutions des problèmes primal et dual.

Proposition 14 *Un couple $(\bar{x}, \bar{y}) \in E \times Y$ est un pointselle de φ sur $E \times Y$ si et seulement si $\bar{x}$ est solution du problème primal, $\bar{y}$ est solution du problème dual et on a*

$$\sup_{y \in Y} \inf_{x \in E} \varphi(x, y) = \inf_{x \in E} \sup_{y \in Y} \varphi(x, y). \tag{100}$$

Dans ces conditions, la valeur de (100) *est* $\varphi(\bar{x}, \bar{y})$.

Démonstration
D'après l'inégalité de dualité faible on a pour tout couple $(\bar{x}, \bar{Y}) \in E \times Y$

$$\inf_{x \in E} \varphi(x, \bar{y}) \leqslant \sup_{y \in Y} \inf_{x \in E} \varphi(x, y) \leqslant \inf_{x \in E} \sup_{y \in Y} \leqslant \sup_{y \in Y} \varphi(\bar{x}, y). \tag{101}$$

Si $(\bar{x}, \bar{y})$ est un point-selle de φ, les deux membres à l'extrême gauche et à l'extrême droite sont égaux et tous les deux égaux à $\varphi(\bar{x}, \bar{y})$. On en déduit qu'on a égalité partout et donc $\bar{x}$ est une solution primale, par l'égalité de droite, que $\bar{y}$ est une solution du problème dual, par l'égalité de gauche et qu'il n'y a pas de saut de dualité par l'égalité du milieu.

Réciproquement, si $\bar{x}$ est une solution du problème primal et $\bar{y}$ est une solution du problème dual et s'il n'y a pas de saut de dualité, on a égalité partout dans l'inégalité (101). Ainsi par définition de l'infimum et du supremum on a la relation suivante :

$$\varphi(\bar{x}, \bar{y}) \leqslant \sup_{y \in Y} \varphi(\bar{x}, y) \leqslant \inf_{x \in E} \varphi(x, \bar{y}) \leqslant \varphi(\bar{x}, \bar{y}).$$

Ces termes sont donc tous égaux et ainsi $(\bar{x}, \bar{y})$ est un point-selle de φ, ce qui conclut la démonstration de la proposition.

Lorsque la fonction de couplage φ a un point-selle, on peut utiliser le problème dual pour résoudre le problème primal de la manière suivante : On cherche $\bar{y} \in Y$ solution du problème dual :

$$\sup_{y \in Y} \sup_{x \in E} \varphi(x, y).$$

On considère alors le problème primal interne pour une telle valeur $\bar{y}$:

$$\inf_{x \in E} \varphi(x, \bar{y}).$$

D'après le théorème précédent, ce problème admet une solution $\bar{x}$ et c'est une solution du problème primal.
Nous verrons que dans certains cas, comme dans la dualité de Fenchel, les liens entre les solutions primales et duales apparaissent également via le sous-différentiel.

L'existence d'un point-selle de φ est un élément crucial dans l'approche par dualité. Le Théorème suivant (Théorème 8 de Applied Nonlinear Analysis de Aubin et Ekeland) donné sans démonstration précise des conditions suffisantes sur φ au delà du fait qu'elle doit être convexe-concave pour qu'elle admette un ensemble de points-selles non vide.

Théorème 15 *Soit M et N deux ensembles convexes de deux espaces vectoriels X et Y. Si on suppose que*

1. $\forall y \in N, \quad x \mapsto \varphi(x, y)$ *est convexe et sci.*
2. $\exists y_0 \in N$ *tel que* $x \mapsto \varphi(x, y_0)$ *est inf-compact.*
3. $\forall x \in M, \quad y \mapsto -\varphi(x, y)$ *est convexe et sci*
4. $\exists x_0 \in M$ *tel que* $x \mapsto -\varphi(x, y_0)$ *est inf-compact.*

alors il existe un point-selle $(\bar{x}, \bar{y}) \in M \times N$ *de* φ.

Ainsi quand la fonction de couplage admet un ensemble non vide de points-selles, au problème primal

$$\inf_{x \in E} f(x)$$

on peut associer un problème dual

$$-\sup_{y \in Y} \delta(y)$$

et un problème dit primal-dual

$$\inf_{x \in E} \sup_{y \in Y} \varphi(x, y)$$

où l'ordre du supremum et de l'infimum est indifférent. Les solutions $(\bar{x}, \bar{y})$ du problème primal-dual sont alors les solutions $\bar{x}$ du problèmes primal et les solutions $\bar{y}$ du problème dual

Dans la section suivante, nous allons proposer une méthode pour construire des fonctions de couplages φ.

4.2 Dualité par perturbation

La méthode présentée ici permet dans des cas favorables de créer des fonctions de couplages associées à des fonctions f et de se ramener ainsi au cadre de la dualité *min-max* décrite dans la sous-section précédente. Nous verrons que la convexité de f donne plus de lattitude mais elle n'est ni une condition suffisante ni une condition nécessaire à l'écriture de f comme un supremum de fonctions de couplage. Cette méthode consiste à construire une fonction de couplage appelée *Lagrangien* qui pourra correspondre au Lagrangien classique dans le cas de l'optimisation sous contrainte, mais qui peut prendre des formes plus variées.

4.3 Lagrangien associé à une perturbation

L'idée de la dualité par perturbation est de considérer une famille de problèmes perturbés *proches* du problème primal.

Considérons le problème primal suivant :

$$\inf_{x\in E} f(x) \tag{P}$$

on définit une fonction de perturbation F de $E\times P$ où $Y=\mathbb{R}^m$ est un espace vectoriel de dimension fini appelé espace des perturbations ou espace dual, à valeurs dans $\mathbb{R}\cup\{+\infty\}$: $(x,p)\mapsto F(x,p)$, telle que

$$F(x,0)=f(x),\ \forall x\in E. \tag{102}$$

On peut ainsi voir F comme une perturbation de f. On considère alors la famille de problèmes

$$\inf_{x\in E} F(x,p) \tag{P_p}$$

qui sont des perturbations de (P).
Le Lagrangien associé à la perturbation est la fonction

$$\ell\,:\,E\times Y\longrightarrow \mathbb{R}\cup\{+\infty\}$$

définie par

$$\ell(x,\lambda)=-\sup_{p\in Y}\left(\langle\lambda,p\rangle-F(x,p)\right). \tag{103}$$

Sous certaines hypothèses sur le choix de la perturbation F, on va pouvoir écrire f comme un supremum de telles fonctions de couplages. Prenons deux exemples usuels.

1. Le Lagrangien classique avec des contraintes *égalité*.
   On considère ne fonction f de la forme

   $$f(x)=\bar f(x)+i_{Kx=0}(x)$$

   où f est une fonction non nécessairement convexe et où K est une application de E dans $\mathbb{R}^m$ non nécessairement affine ou convexe. On considère la perturbation suivante :

   $$F(x,p)=\bar f(x)+i_{Kx+p=0}(x)$$

   Dans ce cas on a

   $$\begin{aligned}\ell(x,\lambda)&=-\sup_p\big(\langle\lambda,p\rangle-\bar f(x)-i_{Kx+p=0}(x)\big)\\&=\bar f(x)-\sup_p\left(-i_{Kx+p=0}+\langle\lambda,p\rangle\right)\end{aligned}$$

   Or $\sup_p(-i_{Kx+p=0}+\langle\lambda,p\rangle)=-\langle\lambda,Kx\rangle$,

   $$\ell(x,\lambda)=\bar f(x)+\langle\lambda,Kx\rangle$$

   qui correspond au Lagrangien classique et on a ainsi

   $$\sup_\lambda \ell(x,\lambda)=f(x).$$

   On peut procéder de même avec des contraintes de type *inégalité*.

2. Si f est convexe, propre et sci, on peut considérer la perturbation suivante : $F(x,p) = f(x+p)$. Un calcul direct donne alors que

$$\ell(\lambda, x) = \langle \lambda, x \rangle - f^*(\lambda). \tag{104}$$

Comme f est convexe, propre et sci, elle est égale à sa bijonguée de Fenchel et ainsi

$$\sup_{\lambda} \ell(\lambda, x) = f(x).$$

3. Si f est de la forme $f(x) = h(x) + g(Kx)$ où K est une application linéaire de E dans $\mathbb{R}^m$ et où g est une fonction convexe propre et sci alors si on considère la perturbation définie par

$$F(x,p) = h(x) + g(Kx + p)$$

on a

$$\ell(x, \lambda) = h(x) + \langle \lambda, Kx \rangle - g^*(\lambda).$$

Sous les hypotèses précédentes et sous des conditions sur les domaines de h et $g \circ K$ nous verrons qu'on a bien également

$$\sup_{\lambda} \ell(x, \lambda) = f(x).$$

On peut observer que fonction de couplage correspond exactement à la dualisation des contraintes dans le cas d'un Lagrangien classique où la fonction g est une indicatrice.
On peut remarquer également que la fonction de couplage et donc le problème dual n'est pas le même si on applique la perturbation à l'une des deux fonctions ou à la somme.

Dans la suite nous allons donner des conditions sur la perturbation F et plus précisément sur la fonction v définie de Y dans $\mathbb{R} \cup \{+\infty\}$ par

$$v(p) = \inf_{x \in E} F(x,p) \tag{105}$$

qui vont assurer l'équivalence entre le problème primal et le problème dual c'est à drire $val(P) = val(D)$.

4.3.1 Conditions assurant l'équivalence entre les problèmes primal et dual.

On peut remarquer tout d'abord que par définition de la perturbation F et de v, la valeur du problème primal $val(P) = \inf_{x \in E} f(x)$ est égale à $v(0)$. Dans le cas de la dualité par perturbation, la fonction de couplage φ est un Lagrangien ℓ défini par (103). La fonction duale δ définie génériquement dans une dualité *min-max* se définit donc par $\delta(y) = -\inf_{x \in E} \ell(x,y)$. Le problème dual s'exprime de la manière suivante à partir de la fonction δ par

$$\sup_{y \in Y} \inf_{x \in E} \ell(x,y) = -\inf_{y \in Y} \delta(y) \tag{106}$$

Les fonctions v et δ sont liées par conjugaison de Fenchel, ainsi

$$\begin{aligned} v^*(\lambda) &= \sup_{p \in Y} \left(\langle p, \lambda \rangle - v(p) \right) \\ &= \sup_{p \in Y} \left(\langle p, \lambda \rangle - \inf_{x \in E} F(x,p) \right) \\ &= \sup_{p \in Y} \sup_{x \in E} \left(\langle p, \lambda \rangle - F(x,p) \right) \\ &= \sup_{x \in E} \sup_{p \in Y} \left(\langle p, \lambda \rangle - F(x,p) \right) \\ &= \sup_{y \in Y} -\ell(x, \lambda) \\ &= \delta(\lambda) \end{aligned}$$

Ainsi le problème dual peut s'écrire via la fonction v comme une minimisation de la conjuguée de la fonction v :

$$\sup_{y \in Y} \inf_{x \in E} \ell(x,y) = -\inf_{y \in Y} v^*(y) = \sup_{y \in Y} -v^*(y). \tag{D}$$

Or ce dernier supremum peut s'écrire également comme la biconguée de v au point 0. Ainsi

$$\sup_{y \in Y} \inf_{x \in E} \ell(x,y) = v^{**}(0). \tag{D}$$

Ainsi, la valeur du problème dual est égal à la valeur de la biconjuguée de v au point 0. On peut alors interpréter l'absence de saut de dualité comme l'égalité de v et de v^{**} au point 0. On en déduit les propositions suivantes :

Proposition 15 *L'ensemble des solutions du problème dual est égal à* $\partial v^{**}(0)$.

Démonstration

$$v^*(\bar{\lambda}) = \inf_p v^*(p) \Longleftrightarrow 0 \in \partial v^*(\bar{\lambda}) \Longleftrightarrow \bar{\lambda} \in \partial v^{**}(0)$$

car v^* est convexe.

Proposition 16 *Conditions nécessaires pour un saut de dualité nul.*

1. *Si la fonction v est convexe et si 0 appartient au domaine de v alors*
   (a) $val(P) = val(D) \Longleftrightarrow v$ *est sci en 0,*
   (b) $\partial v(0) \neq \emptyset \Longleftrightarrow val(P) = val(D)$, $sol(D) \neq \emptyset$ *et* $sol(D) = \partial v(0)$.
   (c) *Si 0 appartient à l'intérieur du domaine de v alors* $val(P) = val(D)$ *et* $sol(D) = \partial v(0)$.
2. *Si la fonction v est convexe propre et sci alors* $val(P) = val(D)$.

On peut remarquer que les dernières conditions assurent que le problème dual admet un ensemble de solution non vide. Mais aucune de ces conditions n'assure que le problème primal admette un ensemble de solutions non vide. Pour cela il faut pouvoir démontrer l'existence d'un point-selle de la fonction de couplage, voir par exemple le théorème 15.

Démonstration

- Le premier point vient du fait que si une fonction est convexe, elle est égale à sa biconjuguée au point x si et seulement si elle est sci en x. Or nous avons vu plus haut que l'égalité entre $val(P)$ et $val(D)$ pouvait s'écrire $v(0) = v^{**}(0)$.

- Si $\partial v(0) \neq \emptyset$ alors $v(0) = v^{**}(0)$ et $\partial v(0) = \partial v^{**}(0)$ et il n'y a donc pas de saut de dualité. De plus d'après la proposition précédente l'ensemble des solutions du problème dual est $\partial v^{**}(0) = \partial v(0)$.
  La réciproque est directe.

- Si 0 appartient à l'intérieur du domaine de v alors v est sci en 0 et $\partial v(0) \neq \emptyset$.

- Si v est convexe, sci et propre alors les deux fonctions v et v^{**} coïncident, d'où le résultat.

4.4 Exemples de dualité

Dans cette section, nous allons passer en revue quelques exemples classiques de dualisation.

4.4.1 Dualité de Fenchel

On considère le problème primal usuel avec une fonction f s'exprimant sous la forme $f(x) = h(x) + g(Kx)$ où K est un opéreteur linéaire de E dans $Y = R^m$ et f et g sont deux fonctions non nécessairement convexes définies respectivement de E dans $\mathbb{R} \cup +\infty$ et de Y dans $\cup\{+\infty\}$.
Nous allons considérer la perturbation $F(x,p) = h(x)+g(Kx+p)$. la fonction v est ainsi définie dans ce cas par :

$$v(p) = \inf_{x \in E} (h(x) + g(Kx + p))$$

La conjuguée de cette fonction permet de définir le problème dual :

$$v^*(\lambda) = \sup_{x \in E} \sup_{p \in Y} (\langle \lambda, p \rangle - h(x) - g(Kx + p)) = h^*(-K^*) + g^*(\lambda)$$

Comme nous l'avons souligné plusieurs fois, il n'y a pas unicité du problème dual, même en prenant $K = Id$, l'expression de v^* n'est pas symétrique en h et h. Cela est dû au fait qu'on a considérer une perturbation dans g et non dans h. D'une manière générale les deux choix sont bien évidemment possibles, mais si g est composée avec un opérateur, le problème dual que l'on obtient en insérant de cette manière la perturbation permet de faire *passer* l'opérateur K de g vers f^*. Dans certains cas pratiques, particulièrement quand f est difförntiable et f^* est explicite, ce transfert peut s'avérer très utile. Le problème dual s'écrit ainsi

$$\sup_{\lambda \in Y} (-h^*(K^*x) - g^*(-\lambda)) \tag{107}$$

Rappelons que le problème primal dual s'écrit ici :

$$\sup_{y\in Y}\inf_{x\in E} h(x) + \langle \lambda, Kx\rangle - g^*(\lambda)$$

et dans ce problème où l'on recherche un point-selle, l'opérateur K est totalement découplé des deux fonctions h et g.
Le théorème de dualité faible s'applique ici et on a $val(D) \leqslant val(P)$. Pour s'assurer qu'on a une dualité forte, c'est à dire qu'on a égalité entre ces deux valeurs, il faut que la fonction v vérifie certaines propriétés énoncées dans la proposition 16. Nous énonçons ici une condition suffisante qui a cependant une portée assez large et qui est vérifiée dès lors que les domaines des fonctions h et g sont pleins.
On peut remarquer dans un premier temps que le domaine de v est de la forme suivante :

$$\text{dom } v = domg - K\text{dom } f$$

Le théorème suivant est ainsi un corollaire de la proposition 16 :

Théorème 16 *On suppose que h et g sont convexe, que f est propre et que*

$$0 \in (dom\ g - K dom\ h)^o \tag{108}$$

alors $val(P) = val(D)$ et $sol(D)$ est non vide.

Comme nous l'avons souligné pour la proposition 16, cette condition ne garantit pas l'existence d'une solution du problème primal, en revanche sous des hypothèses supplémentaires de coercivité sur f, on sait qu'une telle solution existe. Le théorème précédent assure alors que les deux problèmes primal et dual atteignent la même valeur.

4.4.2 Dualité de Lagrange

On reprend dans ce paragraphe, un problème d'optimisation sous contrainte sous la forme

$$\inf_{x\in E} f(x) = \inf_{\substack{x\in E\\ c_E(x)=0\\ c_i(x)\leqslant 0}} \bar{f}(x) = \inf_{x\in E} \bar{f}(x) + i_{\substack{c_E(x)=0\\ c_i(x)\leqslant 0}}(x). \tag{109}$$

où f est une fonction non nécessairement convexe, c_E une application de E dans $\mathbb{R}^{m_E}$ et c_I une application de E dans $\mathbb{R}^{m_I}$. On note $m = m_E + m_I$. On peut construire une fonction de perturbation de la manière suivante :

$$F(x,p) = \begin{cases} \bar{f}(x) \text{ si } c_E(x) + p_E = 0 \text{ et } c_I(x) + p_I \leqslant 0 \\ \qquad\qquad +\infty \end{cases} \tag{110}$$

Un calcul simple montre que le Lagrangien est alors de la forme

$$\ell(x,\lambda) = \begin{cases} \bar{f}(x) + \sum_i \lambda_i c_i(x) & \text{si } \lambda_i \geqslant 0,\ \forall i \in I \\ +\infty & \text{sinon} \end{cases} \tag{111}$$

Ce Lagrangien est celui qu'on obtient classiquement par la méthode des multiplicateurs de Lagrange.
Le théorème suivant précise le lien entre les points-selles de ce Lagrangien et les solutions des problèmes primal et dual dans le cas où le problème est convexe ou pas. Il reprend les résultats déjà énoncés dans la partie consacrée aux multiplicateurs de Lagrange :

Théorème 17 *1. Si $(\bar{x}, \bar{\lambda})$ est un point selle du Lagrangien 111 sur $E \times \mathbb{R}^m$ sur $E \times \mathbb{R}^m$, alors $\bar{x}$ est une solution du problème primal* (109).

2. Réciproquement, si f et c sont convexes, si $\bar{x}$ est une solution du problème primal (109), *que f et c sont différentiables en $\bar{x}$ et s'il existe un multiplicateur $\bar{\lambda}$ vérifiant les conditions KKT alors $(\bar{x}, \bar{\lambda})$ est un point-selle du Lagrangien* (111).

4.4.3 Dualité et Lagrangien augmenté

Pour les problèmes d'optimisation sous contraintes, considérer des perturbations qui conduisent à des multiplicateurs de Lagrange classiques n'est pas le seul choix. Une autre pénalisation conduisent à ce qu'on appelle un Lagrangien augmenté qui peut s'avérer particulièrement utile pour certains problèmes non convexes. Nous n'entrerons pas ici dans les détails mais nous présenterons ici brièvement la méthode dans le cas de contraintes *égalité* mais celle ci peut se géénraliser à des contraintes it inégalité.
On considère le problème contraint :

$$\inf_{x \in E} f(x) = \inf_{\substack{x \in E \\ c_E(x) = 0}} \bar{f}(x) = \inf_{x \in E} \bar{f}(x) + i_{c_E(x) = 0}(x). \tag{112}$$

et la perturbation suivante :

$$F(x,p) = \begin{cases} \bar{f}(x) + \frac{r}{2} \|p\|_2^2 & \text{si } c_E(x) + p_E = 0 \\ & +\infty \end{cases} \tag{113}$$

où r est un réel positif. Si $r = 0$, on retrouve le lagrangien classique.
La fonction $v(p)$ associée peut s'exprimer à partir de la fonction v_0 que l'on obtient avec la perturbation associée au Lagrangien classique :

$$v(p) = v_0(p) + \frac{r}{2} \|p\|_2^2 \tag{114}$$

Le fonction de couplage s'exprime alors de la manière suivante :

$$\ell(x, \lambda) = \bar{f}(x) + \langle \lambda, c_E(x) \rangle + \frac{r}{2} \|c_E(x)\|_2^2 . \tag{115}$$

Le choix de la valeur de r peut avoir un impact sur la convexité locale en 0 ou sur la convexité globale de v et sous certaines hypothèses peut assurer l'existence de points-selles locaux de la fonction de couplage.

5 Questions diverses

5.1 Convexité

1. Parmi les fonctions suivantes, dire lesquelles sont convexes.
   (a) $x \mapsto \sum_{k=1}^{n} |x_i|^2$.
   (b) $x \mapsto \left(\sum_{k=1}^{n} |x_i|^{\frac{1}{2}}\right)^2$.
   (c) $x \mapsto \|Dx\|_1$ où D est une application linéaire.
   (d) $x \mapsto x \log x$ si $x > 0$ et $+\infty$ sinon.
   (e) $x \mapsto e^{-\|x\|^2}$.
   (f) $x \mapsto 2x + cos(x)$.
   (g) $x \mapsto \begin{cases} 0 & \text{si } Ax = y \\ +\infty & \text{sinon} \end{cases}$ où A est une application linéaire et y un vecteur quelconque.
   (h) $x \mapsto \begin{cases} -\sqrt{x} & \text{si } x \geqslant 0 \\ +\infty & \text{sinon} \end{cases}$

2. Parmi les fonctions suivantes, dire lesquelles sont fortement convexes.
   (a) $x \mapsto \sum_{k=1}^{n} |x_i|^2$.
   (b) $x \mapsto \|x\|_2^4$.
   (c) $x \mapsto \|Dx\|_1$ où D est une application linéaire.
   (d) $x \mapsto x \log x$ si $x > 0$ et $+\infty$ sinon.
   (e) $x \mapsto \|Dx\|_2^2$ où D est une application linéaire non injective.
   (f) $x \mapsto \|Dx\|_2^2$ où D est une application linéaire bijective.

3. Toute fonction strictement convexe définie de $\mathbb{R}$ dans $\mathbb{R}$ admet-elle un minimum ?

4. Pour toute application linéaire A, l'application $x \mapsto \|Ax\|^2$ est coercive. Vrai ou faux.

5. Toute fonction strictement convexe est fortement convexe. Vrai ou faux.

6. Le maximum de deux fonctions convexes est-il convexe ?

7. Le minimum de deux fonctions convexes est-il convexe ?

8. La somme d'une fonction convexe et d'une fonction fortement convexe est fortement convexe. Vrai ou faux.

5.2 Optimisation différentiable

Dans les questions de cette section, on considère une fonction f convexe différentiable, à gradient L-Lipschitz.

9. La méthode du gradient explicite appliquée à f converge vers un minimiseur de f indépendemment du pas. Vrai ou faux.

10. La méthode du gradient implicite appliquée à f converge vers un minimiseur de f indépendemment du pas. Vrai ou faux.

11. Si $\gamma < \frac{2}{L}$ l'opérateur $Id - \gamma \nabla f$ est fermement non expansif. Vrai ou faux.

12. Pour tout couple $(x, y) \in E^2$, on a nécessairement

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

Vrai ou faux.

13. Si f est coercive pour tout $\gamma > 0$ la suite $(x_n)_{n\in\mathbb{N}}$ produite par la méthode de gradient explicite est nécessairement bornée. Vrai ou faux.

14. Soit g est une fonction différentiable minorée à gradient $L-$Lipschitz mais pas nécessairement convexe, soit $x_0 \in E$, $\gamma < \frac{1}{L}$ et soit $(x_n)_{n\in\mathbb{N}}$ la suite associée à la descente de gradient explicite sur g.

    (a) La suite $(g(x_n))_{n\in\mathbb{N}}$ est-elle nécessairement décroissante ?

    (b) La suite $(g(x_n))_{n\in\mathbb{N}}$ converge t-elle nécessairement ?

    (c) La suite $(x_n)_{n\in\mathbb{N}}$ converge t-elle nécessairement vers un point critique de g ?

    (d) La suite $(x_n)_{n\in\mathbb{N}}$ converge t-elle nécessairement vers un mimimum local de g ?

15. La méthode du gradient implicite est une méthode de descente. Vrai ou faux.

16. Peut-on toujours définir l'algorithme de descente de gradient implicite sur une fonction non convexe ?

17. Toute fonction g définie de $\mathbb{R}$ dans $\mathbb{R}$ de classe C^2 est à gradient Lipschitz. Vrai ou faux.

18. Appliquer la méthode du gradient explicite à la fonction $x \mapsto \|Ax - y\|^2$ où A est une application linéaire nécessite une inversion de matrice à chaque étape. Vrai ou faux.

19. Appliquer la méthode du gradient implicite à la fonction $x \mapsto \|Ax - y\|^2$ où A est une application linéaire nécessite une inversion de matrice à chaque étape. Vrai ou faux.

20. Il existe une solution explicite au problème de minimisation suivant :

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-y\|^2 + \lambda\|x\|^2$$

où A est une matrice et λ un réel positif. Vrai ou faux.

5.3 Optimisation non lisse

On suppose dans cette section que f est une fonction sci, convexe, propre et coercive.

21. Pour tout x et pour tout $\gamma > 0$, l'opérateur $T = Id+\gamma\partial f$ est monovalué. Vrai ou faux.
22. Pour tout x et pour tout $\gamma > 0$, l'opérateur $T = (Id + \gamma\partial f)^{-1}$ est monovalué et injectif. Vrai ou faux.
23. Si $f(x) = \|x\|_1$ alors $Id - \text{Prox}_f$ est l'indicatrice d'un convexe fermé. Vrai ou faux.
24. Si f est l'indicatrice d'un convexe fermé C, Prox_f est le projecteur sur C. Vrai ou faux.
25. Soit A une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ telle que $\alpha AA^t = Id$, $b \in \mathbb{R}^m$ et f et g deux foncions convexe reliées par la relation suivante $f(x) = g(Ax + b)$ alors les opérateurs proximaux de f et g sont reliés par l'une des relations suivantes, laquelle :
    (a) $\text{Prox}_f(x) = x - \alpha^{-1}A^t(Ax + b - \text{Prox}_{\alpha^{-1}g}(Ax + b))$
    (b) $\text{Prox}_f(x) = x - \alpha A^t(Ax + b - \text{Prox}_{\alpha^{-1}g}(Ax + b))$
26. Soit f et g deux fonctions convexes. On rappelle que si $f(x) = g(Ax)$ avec $\alpha AA^t = Id$ alors $\text{Prox}_f(x) = x - \alpha A^t(Ax - \text{Prox}_{\alpha^{-1}g}(Ax))$.
    Si on a maintenant $f(x_1, \cdots, x_m) = g(x_1 + \cdots + x_m)$ alors l'opérateur proximal de f s'exprime à partir de celui de g d'une des deux manières suivantes : laquelle ?
    (a)
    $$\text{Prox}_f(x_1,\cdots,x_m)_i = x_i - \frac{1}{m}\sum_{j=1}^m x_j + \frac{1}{m}\text{Prox}_{mg}\left(\sum_{j=1}^m x_j\right),\ i = 1,\cdots,m$$
    (b)
    $$\text{Prox}_f(x_1,\cdots,x_m)_i = x_i + \frac{1}{m}\sum_{j=1}^m x_j - \frac{1}{m}\text{Prox}_{mg}\left(\sum_{j=1}^m x_j\right),\ i = 1,\cdots,m$$
27. Soit λ un réel positif, si $g = \lambda f$ alors $\partial g = \lambda\partial f$. Vrai ou faux.

28. Soit λ un réel positif, si $g = \lambda f$ alors $\text{Prox}_g(x) = \lambda \text{Prox}_f(\frac{x}{\lambda})$. Vrai ou faux.

29. Soit $x_0 \in E$ tel que $\partial f(x_0)$ est réduit à un point, f est-elle nécessairement différentiable en x_0 ?

30. Si f est définie de $\mathbb{R}$ dans $\mathbb{R}$ et qu'il existe x_0 tel que $\{-1, 1\} \subset \partial f(x_0)$ alors nécessairement x_0 est un minimiseur de f. Vrai ou faux.

31. Il existe un point du domaine de la fonction $x \mapsto \begin{cases} -\sqrt{x} & \text{si } x \geqslant 0 \\ +\infty & \text{sinon} \end{cases}$ où la sous-différentielle est vide.

32. Il existe un point du domaine de la fonction $x \mapsto \begin{cases} 0 & \text{si } x \geqslant 0 \\ +\infty & \text{sinon} \end{cases}$ où la sous-différentielle est vide.

33. Si f est différentiable mais non nécessairement convexe et g convexe, l'algorithme Forward-Backward est-il toujours défini ?

34. Si la fonction f n'est pas convexe, quand il est défini, FB est-il toujours un algorithme de descente ?

35. Parmi les algorithmes suivants, dire lesquels sont des cas particuliers de Forward-Backward :

    (a) La méthode du gradient explicite.
    (b) La méthode du gradient projeté.
    (c) La méthode du gradient implicite.
    (d) La méthode du gradient à pas optimal.
    (e) L'algorithme du point proximal.
    (f) La méthode de Newton.

36. FISTA est-il un algorithme de descente ?

6 TD, TP sur Forward-Backward

Le but de ce qui suit est de mettre en place l'algorithme *Forward-Backward* en matlab. Dans un premier temps nous allons l'appliquer dans le cadre du LASSO, puis pour des fonctionelles plus variées.

6.1 FB pour le LASSO (Basis Pursuit)

Une grande variété de problèmes inverses peuvent être résolu via la miminisation de la forme

$$\min_{x \in E} \frac{1}{2} \|Ax - y\|^2 + \lambda \|x\|_1 \tag{116}$$

telle que la super-résolution, le debluring ou l'inpainting en traitement d'image. En effet si x^0 est parcimonieux c'est à dire qu'il a peu de composantes non nulles ou que l'essentiel de son énergie est concentrée sur un petit nombre de composantes, alors on peut estimer x^0 à partir $y = Ax_0$ ou même $y = Ax_0+b$, en résolvant (116) pour des petites valeurs de λ et cela pour une large classe d'opérateurs linéaires A. L'object ici n'est pas de faire une étude théorique sur les choix de A, de λ du niveau de bruit b et de la parcimonie de x pour lesquels une telle minimisation est efficace. Nous allons travailler sur des exemples où le minimiseur x^* de (116) est proche de x_0.

Dans le cas de la super-résolution, x est un vecteur parcimonieux et A est un opérateur de filtrage passe bas en Fourier. On pourra aussi regarder ce qu'il se passe pour un filtre aléatoire.

Dans le cas du debluring ou de l'inpainting, ce sont rarement les signaux x d'intérêt qui sont parcimonieux mais leurs coefficients dans une base orthonormée $z = Tx$, comme une transformée de Fourier ou une transformée en Ondelettes. Dans ce cas on cherchera à minimiser une fonctionnelle de la forme

$$\min_{x\in E} \frac{1}{2} \|Ax - y\|^2 + \lambda \|Tx\|_1 \tag{117}$$

où T est une transformée orthogonale. L'opérateur A est alors un opérateur de convolution (flou) ou de masquage dans le problème de l'inpainting.

1. Si on pose $f(x) = \frac{1}{2} \|Ax - y\|^2$, expliciter le gradient de f.

2. Si l'application linéaire A est un opérateur de masquage donner une expression plus simple de ∇f.

3. Même question si A est une convolution par un filtre symétrique réel h.

4. Si on pose $g(x) = \lambda \|x\|_1$, expliciter $\mathrm{Prox}_{\gamma g}$ pour $\gamma > 0$.

5. Si on pose $g(x) = \lambda \|Tx\|_1$, où T est une transformée orthogonale, expliciter $\mathrm{Prox}_{\gamma g}$ pour $\gamma > 0$.

6. Ecrire une fonction matlab

   ```
   function S=GenSig(N,k)
   ```

   qui génère aléatoirement un vecteur de taille N ayant k composantes non nulles et dont les composantes non nulles valent 1 ou -1 avec une probabilité $\frac{1}{2}$.

7. Ecrire une fonction matlab

   ```
   function h=Pasbas(N,n)
   ```

   qui crée un filtre passe-bas idéal de taille N conservant $2n+1$ fréquences.

8. Ecrire une fonction matlab

   ```
   function h=FiltreGaus(N,sigma2)
   ```

qui crée un filtre gaussien de taille N et de variance $sigma2$. Le filtre doit être construit de telle sorte que $\|h\|_1 = 1$.

9. Ecrire une fonction

```
function A=Filtre2Mat(h)
```

qui à un filtre h associe la matrice de la convolution circulaire par h.

10. Ecrire un script qui génère un signal parcimonieux x_0 de parcimonie k (On pourra choisir $k = 5$ et $N = 1024$ pour commencer), qui effectue la convolution par un filtre passe bas h dont le résultat est un vecteur S.

11. Ecrire une fonction

```
function x=DeconvFB(S,A,lambda,N)
```

qui prend en entrée un vecteur S, une matrice A, un réel positif $lambda$, un entier N et qui effectuer N étapes de l'algorithme *Forward-Backward* pour minimiser la fonction suivante :

$$\frac{1}{2}\|Ax - S\|_2^2 + \lambda \|x\|_1$$

et qui renvoie la dernière itération de l'algorithme. On pourra utiliser S ou le vecteur nul comme initialisation.
On tracera la courbe représentant la valeur de la foncionnelle à chaque étape.

12. Faire plusieurs tests, avec différentes valeurs de lambda, de S et de filtres h. On affichera pour chaque test sur la même figure, le vecteur $\hat{x}$ produit par l'algorithme et le vecteur $x0$ original. On pourra également comparer la valeur de la fonctionnelle en $\hat{x}$ et en $x0$.

13. Quel critère d'arrêt peut-on proposer pour l'algorithme ?

14. Ecrire une fonction similaire

```
function x=DeconvFB2(S,h,lambda,N)
```

où la matrice A est remplacée par le filtre h.

15. Vérifier que les deux fonctions produisent le même résultat.

16. La commande `dct` permet de calculer la transformée en cosinus locaux d'un vecteur. Cette transformée est une transformée orthogonale dont la transformée inverse est `idct`. Ecrire, sur le modèle de `GenSig` une fonction qui génère un signal parcimonieux dans une base de cosinus locaux.

17. A l'aide de la commande `randperm`, écrire une fonction qui crée un vecteur de masquage M de taille N et dont d composantes aléatoires valent 1 et les autres 0. Tester la fonction en prenant $N = 1024$ et $d = 512$ et utiliser ce vecteur pour masquer un vecteur parcimonieux en dct. On notera S ce vecteur masqué

18. Ecrire une fonction

```
function x=InpaintFB(S,M,lambda,N)
```

qui prend en entrée un vecteur S, un vecteur de masquage M, un réel positif *lambda*, un entier N et qui effectuer N étapes de l'algorithme *Forward-Backward* pour minimiser la fonction suivante :

$$\frac{1}{2}\|Mx - S\|_2^2 + \lambda \|Tx\|_1$$

où T est la transformée en dct, et qui renvoie la dernière itération de l'algorithme. On pourra utiliser S ou le vecteur nul comme initialisation.
On tracera la courbe représentant la valeur de la foncionnelle à chaque étape.

19. Tester la fonction pour différentes réalisations du masquage, différentes valeurs du paramètre *lambda* et différentes parcimonies, en comparant à chaque fois au vecteur objectif ?

20. Quelle est la parcimonie limite qui assure la reconstruction d'un vecteur $x0$?

21.

A Démonstration du Lemme 5

La démonstration de ce lemme s'appuie sur le caractère cocoercif du gradient des fonctions convexes différentiables :

Lemma 10 *Soit f une fonction différentiable et convexe telle qu'il existe $L > 0$ tel pour tout $(x, y) \in E^2$, $\|\nabla f(x) - \nabla f(y)\| \leqslant L\|x - y\|$. On a la relation suivante :*

$$\forall (x,y) \in E^2, \quad \langle \nabla f(x) - \nabla f(y), x - y\rangle \geqslant \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|^2. \quad (118)$$

Cette propriété est appelée co-coercivité de la fonction ∇f.

Démonstration :
Soit g une fonction dérivable définie sur $\mathbb{R}$ telle que g' est $K-$Lipschitz,

$$g(1) = g(0) + \int_0^1 g'(t)dt = g(0) + g'(0) + \int_0^1 (g'(t) - g'(0))dt \leqslant g(0) + g'(0) + \frac{K}{2}.$$

Soit $(y,z) \in E^2$, pour tout $t \in [0,1]$ on pose $v = z - y$, $y_t = y + t(z-y)$ et $g(t) = f(y_t)$. La fonction g est dérivable et $g'(t) = \langle \nabla f(y_t), v\rangle$. D'après les hypothèses sur f, g' est K-Lipchitz avec $K = L\left\|v\right\|^2$.
On en déduit que pour tout $(y,z) \in E^2$

$$f(z) \leqslant f(y) + \langle \nabla f(y), z - y\rangle + \frac{L}{2}\left\|z-y\right\|^2. \tag{119}$$

En prenant ensuite $z = x$ et $y = x - \frac{1}{L}\nabla f(x)$ on obtient que

$$\frac{1}{2L}\left\|\nabla f(x)\right\|^2 \leqslant f(x) - f(x - \frac{1}{L}\nabla f(x)) \leqslant f(x) - f(x^\star).$$

où $x^\star$ est un minimiseur de f. On a ainsi pour tout $x \in E$,

$$\frac{1}{2L}\left\|\nabla f(x)\right\|^2 \leqslant f(x) - f(x^\star). \tag{120}$$

Soit $(x,y) \in E^2$, on introduit ensuite une fonction $h_1(z) = f(z) - \langle \nabla f(x), z\rangle$ et $h_2(z) = f(z) - \langle \nabla f(y), z\rangle$. Ces deux fonctions sont convexes et admettent pour minimiseurs respectivement x et y. On applique l'inégalité (120) à ces deux fonctions et on déduit

$$h_1(x) \leqslant h_1(y) - \frac{1}{2L}\left\|\nabla h_1(y)\right\|^2 \text{ et } h_2(y) \leqslant h_2(x) - \frac{1}{2L}\left\|\nabla h_2(y)\right\|^2.$$

En sommant ces deux inégalités on obtient le résultat du lemme.

Si on note $T = Id - \gamma\nabla f$ et $S = Id - T = \gamma\nabla f$, la cocoercivité du gradient s'exprime de la manière suivante :

$$\langle Sx - Sy, x - y\rangle \geqslant \frac{1}{\gamma L}\left\|Sx - Sy\right\|^2. \tag{121}$$

On peut maintenant démontrer le Lemme 5.

Remarquons d'abord que

$$\left\|x-y\right\|^2 - \left\|Tx - Ty\right\|^2 = 2\langle Sx - Sy, x-y\rangle - \left\|Sx - Sy\right\|^2. \tag{122}$$

1. Si $\gamma \leqslant \frac{1}{L}$, on déduit de (121) et (122) que

   $$\left\|x-y\right\|^2 - \left\|Tx - Ty\right\|^2 \geqslant \left\|Sx - Sy\right\|^2 \tag{123}$$

   et donc que T est fermement non expansif.

2. Si $\gamma \leqslant \frac{2}{L}$, on déduit de (121) et (122) que

   $$\left\|x-y\right\|^2 - \left\|Tx - Ty\right\|^2 \geqslant 0 \tag{124}$$

   et donc que T est 1-Lipschitz.

A.1 Démonstration du Lemme 9

Par définition de l'opérateur proximal, $\hat{x}$ est le minimiseur de la fonction $\frac{1}{\gamma}$-fortement convexe suivante

$$z \longmapsto g(z) + f(\bar{x}) + \langle \nabla f(\bar{x}), z - \bar{x} \rangle + \frac{1}{2\gamma} \|z - \bar{x}\|^2$$

Ainsi pour tout $z \in E$

$$\begin{aligned} g(\hat{x}) + f(\bar{x}) + \langle \nabla f(\bar{x}), \hat{x} - \bar{x} \rangle + \frac{\|\bar{x} - \hat{x}\|^2}{2\gamma} + \frac{\|z - \hat{x}\|^2}{2\gamma} \\ \leqslant g(z) + f(\bar{x}) + \langle \nabla f(\bar{x}), z - \bar{x} \rangle + \frac{\|z - \bar{x}\|^2}{2\gamma}. \end{aligned}$$

Comme $\gamma \leq 1/L$, il suit que

$$g(\hat{x}) + f(\hat{x}) + \frac{1}{2\gamma} \|z - \hat{x}\|^2 \leqslant g(z) + f(\bar{x}) + \langle \nabla f(\bar{x}), z - \bar{x} \rangle + \frac{1}{2\gamma} \|z - \bar{x}\|^2 .$$

par convexité, $f(\bar{x}) + \langle \nabla f(\bar{x}), z - \bar{x} \rangle \leqslant f(z)$. On en déduit (69) est satisfaite ce qui conclut la preuve du lemme.