TP 1: Minimisation de fonctionnelles quadratiques convexes

Motivation Dans ce TP, on s'intéresse à la minimisation sur $\Omega = \mathbb{R}^N$ de fonctionnelles de la forme suivante:

$$f(x) = \frac{1}{2}\langle x|Qx\rangle + \langle b|x\rangle$$

où Q est une matrice symétrique définie positive et b est un vecteur colonne. Ces fonctionnelles apparaissent dans de nombreuses applications, et notamment dans la méthode des moindres carrés. On rappelle les définitions suivantes:

(i) Une matrice symétrique Q est dite *positive* (ce qu'on note $Q \succeq 0$) si et seulement $\forall x \in \mathbb{R}^d, \langle x|Qx\rangle \geq 0$.

(ii) Q est dite *définie positive* si $\forall x \in \mathbb{R}^d \setminus \{0\}, \langle x|Qx\rangle > 0$.

Les résultats des Exercices 1 et 2 forment la base de ce qu'il faut savoir pour ce cours: il faut être capable de les retrouver très rapidement!

Exercice 1. *Convexité*. Dans cet premier exercice, nous étudions la convexité de f. On suppose que Q est symétrique, mais pas nécessairement positive.

1. Montrer que $f(x) = \frac{1}{2}\sum_{1\leq i,j\leq N} Q_{ij}x_ix_j + \sum_{1\leq i\leq N} b_ix_i$, $\nabla f(x) = Qx + b$ et $\mathrm{D}^2 f(x) = Q$.
2. En déduire que f est convexe si et seulement si Q est positive (on pourra utiliser la caractérisation de la convexité utilisant ∇f).
3. Démontrer les égalités suivantes (on utilisera la symétrie de Q):

$$\begin{aligned} f((1-t)x+ty) - (1-t)f(x) - tf(y) &= -\frac{t(1-t)}{2}\langle Qx|x\rangle - \frac{t(1-t)}{2}\langle Qy|y\rangle + t(1-t)\langle Qx|y\rangle \\ &= -\frac{t(1-t)}{2}\langle Q(x-y)|x-y\rangle \end{aligned}$$

   En déduire que f est strictement convexe si et seulement si Q est définie positive.
4. Soit Q une matrice définie positive.
   (a) Montrer que la fonction $q_Q : x \mapsto \langle x|Qx\rangle$ atteint son minimum m sur l'ensemble $K = \{x \in \mathbb{R}^N \mid \|x\| = 1\}$, puis que $m > 0$.
   (b) En déduire que $\forall x \in K, \langle x|Qx\rangle \geq m$, puis que $\forall x \in \mathbb{R}^d, \langle x|Qx\rangle \geq m\|x\|^2$.
   (c) Démontrer que la fonction f admet un minimiseur sur $\mathbb{R}^d$.

Dans les exercices 2 et 3, on suppose que Q est symétrique définie positive

Exercice 2. *Caractérisation du minimiseur*. Montrer que f admet un unique minimiseur x^* sur $\mathbb{R}^N$, caractérisé par l'équation $Qx^* + b = 0$.

Exercice 3. *Descente de gradient à pas optimal*. On considère l'algorithme de descente de gradient à pas optimal pour une fonction f, dont les itérées sont données par $x^{(0)} \in \mathbb{R}^N$ puis:

$$\begin{cases} d^{(k)} = -\nabla f(x^{(k)}) \\ t^{(k)} = \arg\min_{t \in \mathbb{R}} f(x^{(k)} + t d^{(k)}) \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases}$$

On s'intéresse au cas $f(x) = \frac{1}{2}\langle Qx | x\rangle + \langle b | x\rangle$.

1. Soit $x \in \mathbb{R}^N$ et $v \in \mathbb{R}^N$. Montrer qu'il existe un unique t^* minimisant $t \in \mathbb{R} \mapsto f(x + tv)$ et donner son expression.
2. En déduire que les itérées de la méthode de descente de gradient vérifient

$$\begin{cases} d^{(k)} = -(Qx^{(k)} + b) \\ t^{(k)} = \frac{\langle d^{(k)} | d^{(k)} \rangle}{\langle d^{(k)} | Q d^{(k)} \rangle} \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases}$$

Exercice 4. *Application aux moindres carrés*. Dans cette question, on s'intéresse au problème de minimisation suivant,

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \left\| Ax - y \right\|^2,$$

où A est une matrice de taille $M \times N$.

1. Démontrer que $f(x) := \frac{1}{2} \left\| Ax - y \right\|^2 = \frac{1}{2}\langle Qx | x\rangle + \langle b | x\rangle + c$, où $Q \in \mathcal{M}_N(\mathbb{R})$, $b \in \mathbb{R}^N$ et $c \in \mathbb{R}$ sont à déterminer.
2. En déduire l'expression de $\nabla f(x)$ et $\mathrm{D}^2 f(x)$ en fonction de A et y.
3. Montrer que Q est symétrique, $Q \succeq 0$, et que $x^* \in \arg\min f$ si et seulement si $A^T A x^* = A^T y$, où A^T est la transposée de A.
4. On suppose que A est injective (i.e. son noyau est vide). Démontrer que Q est définie positive, puis que le problème d'optimisation admet une unique solution x^*.