Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

M1 MMA 2024-2025
Optimisation Sous-Contraintes

Feuille de TD n°2

Cônes, projection.

Exercice 4

Soit $g : \Omega \to \mathbb{R}$ une fonction différentiable sur Ω ouvert non vide de $\mathbb{R}^n$. On considère $C = \{g \leqslant 0\} = \{x \in \mathbb{R}^n \,/\, g(x) \leqslant 0\}$ qui est supposé convexe et non vide.

Soit $y_0 \notin C$. L'objectif est de démontrer que si $x_0 = P_C(y_0)$ alors x_0 satisfait l'une des deux propositions

— $\nabla g(x_0) = 0$,

— $\exists \lambda > 0,\ y_0 = x_0 + \lambda \nabla g(x_0)$.

1. Justifier que x_0 est bien défini.

   On suppose dans la suite que $\nabla g(x_0) \neq 0$.
2. Proposer une heuristique géométrique justifiant la deuxième condition.
3. Montrer que $T_C(x_0) = \{h \in \mathbb{R}^n \,/\, \langle \nabla g(x_0),\, h\rangle \leqslant 0\}$. *Indication : on pourra raisonner par double inclusion.*
4. Déterminer $(T_C(x_0))^\star$ et en déduire la deuxième assertion.

<u>Correction</u>.

1. L'ensemble C est convexe non vide par hypothèse, et comme g est différentiable sur Ω, g est continue sur Ω donc C est fermé comme image réciproque d'un fermé par une application continue. Donc la projection de y_0 sur C est bien définie.
2. Explications détaillées avec dessins en classe. "Brièvement" : $x_0 \in \partial C$ avec ∂C ensemble de niveau 0 de g. Si $\nabla g(x_0) \neq 0_{\mathbb{R}^n}$, alors $\nabla g(x_0)$ est "orthogonal" à l'ensemble de niveau 0 de g : ∂C. Plus précisément $\nabla g(x_0)$ est la normale à la direction de l'hyperplan affine tangent à ∂C en x_0.[1]. De ce fait, comme $C = g^{-1}(]-\infty, 0])$, $x_0 + T_C(x_0)$ est le demi-espace de $\mathbb{R}^n$ contenant C et de frontière l'espace affine tangent à ∂C en x_0. D'où $T_C(x_0)^* = \mathbb{R}_- \nabla g(x_0)$. Or comme x_0 est l'unique élément de C minimisant $F : x \mapsto \|x - y_0\|_2^2$ sur C, on a par condition d'optimalité d'ordre 1 : $\nabla F(x_0) = x_0 - y_0 \in T_C(x_0)^*$, d'où il existe $\lambda \geqslant 0$ tel que $y_0 = x_0 + \lambda \nabla g(x_0)$. On ne peut avoir $\lambda = 0$ car sinon $y_0 = x_0 \in C$.

   Montrons dans la suite cette heuristique rigoureusement.
3. Commençons par écrire le développement de Taylor à l'ordre de 1 de g en x_0, ce qui nous sera utile dans le suite. Comme g est différentiable en x_0, il existe V voisinage de 0 et $\varepsilon : V \to \mathbb{R}$ une fonction de limite nulle en $0_{\mathbb{R}^n}$ tels que pour tout $h \in V$, $g(x_0 + h) = g(x_0) + \langle \nabla g(x_0),\, h\rangle + \|h\|_2\, \varepsilon(h)$.

   ($\subset$) Soit $h \in T_C(x_0)$, alors il existe $(h_k)_{k\in\mathbb{N}}$ une suite de $\mathbb{R}^n$ convergeant vers h et $(t_k)_{k\in\mathbb{N}}$ une suite de $\mathbb{R}_+^*$ convergeant vers 0 tels que pour tout $k \in \mathbb{N}$, $x_0 + t_k h_k \in C$ i.e. $g(x_0 + t_k h_k) \leqslant 0$. Comme $(t_k h_k)_{k\in\mathbb{N}}$ converge vers $0_{\mathbb{R}^n}$, il existe $k_0 \in \mathbb{N}$ tel que pour tout $k \geqslant k_0$, on a $t_k h_k \in V$ et donc
   $$g(x_0 + t_k h_k) = g(x_0) + t_k \langle \nabla g(x_0),\, h_k\rangle + t_k \|h_k\|_2\, \varepsilon(t_k h_k).$$
   Comme $y_0 \notin C$, on a $x_0 = P_C(y_0) \in \partial C$ et donc $g(x_0) = 0$ (si $g(x_0) < 0$ par continuité de g, on montre que $x_0 \in \mathring{C}$ i.e. $x_0 \notin \partial C = \overline{C} \setminus \mathring{C} = C \setminus \mathring{C}$). Ainsi pour tout $k \geqslant k_0$, on a
   $$g(x_0 + t_k h_k) \leqslant 0 \Leftrightarrow \langle \nabla g(x_0),\, h_k\rangle + \|h_k\|_2\, \varepsilon(t_k h_k) \leqslant 0,$$
   où on a simplifié par $t_k > 0$. En passant à la limite dans l'inégalité lorsque $k \to +\infty$, comme $\varepsilon(t_k h_k) \to 0$, on obtient $\langle \nabla g(x_0),\, h\rangle \leqslant 0$.

   ($\supset$) Soit $h \in \mathbb{R}^n$ tel que $\langle \nabla g(x_0),\, h\rangle < 0$. Montrons que h est une directement admissible à C en x_0 ce qui justifiera alors que $h \in T_C(x_0)$. On a vu précédemment que $g(x_0) = 0$, de plus il existe $t_0 > 0$ tel que pour tout $t \in]0, t_0[$, $th \in V$. On a donc $g(x_0 + th) = t\,(\langle \nabla g(x_0),\, h\rangle + \|h\|_2\, \varepsilon(th))$. Or $\|h\|_2\, \varepsilon(th) \to 0$ quand $t \to 0$ et comme $\langle \nabla g(x_0),\, h\rangle < 0$ il existe donc $t_1 \geqslant t_0$ tel que pour tout $t \in]0, t_1[$, $\|h\|_2\, \varepsilon(th) < -\langle \nabla g(x_0),\, h\rangle$, d'où $g(x_0 + th) < 0$. Ainsi pour tout $t \in [0, t_1[$, $x_0 + th \in C$ et donc h est une directement admissible à C.

1. Si ∇g ne s'annule pas sur ∂C, alors ∂C est une sous variété de $\mathbb{R}^n$ de dimension $n-1$ dont l'espace tangent en tout point $x \in \partial C$ est un hyperplan affine dont la normale à la direction est $\nabla g(x)$. Cf par exemple "Petit guide du calcul différentiel..." de Rouvière.

On a ainsi montré que $\{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle < 0\} \subset T_C(x_0)$. Comme $T_C(x_0)$ est fermé, on a

$$\overline{\{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle < 0\}} \subset T_C(x_0),$$

i.e.

$$\{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle \leqslant 0\} \subset T_C(x_0).$$

4. On a $T_C(x_0) = \{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle \leqslant 0\}$, donc

$$P := \{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle = 0\} \subset T_C(x_0).$$

D'où $T_C(x_0)^* \subset P^*$. On vérifie facilement que $P^* = \mathrm{Vect}(\nabla g(x_0))$ en utilisant le fait que P est un sous-espace vectoriel de $\mathbb{R}^n$.

Soit $\lambda \leqslant 0$, alors pour tout $h \in T_C(x_0) = \{h \in \mathbb{R}^n \ / \ \langle \nabla g(x_0),\, h\rangle \leqslant 0\}$, on a $\langle \lambda g(x_0),\, h\rangle = \lambda \underbrace{\langle g(x_0),\, h\rangle}_{\leqslant 0} \geqslant 0$.

D'où $\lambda \nabla g(x_0) \in T_C(x_0)^*$. Donc $\mathbb{R}_- \nabla g(x_0) \subset T_C(x_0)^* \subset \mathbb{R}\nabla g(x_0)$.

Soit $\lambda > 0$, alors pour tout $h \in T_C(x_0)$, on a $\langle \lambda g(x_0),\, h\rangle = \lambda \underbrace{\langle g(x_0),\, h\rangle}_{\leqslant 0} \leqslant 0$. D'où $\lambda \nabla g(x_0) \notin T_C(x_0)^*$.

Donc $T_C(x_0)^* \subset \mathbb{R}_- \nabla g(x_0)$. D'où $T_C(x_0)^* = \mathbb{R}_- \nabla g(x_0)$.

Or comme x_0 est l'unique élément de C minimisant $F : x \mapsto \|x - y_0\|_2^2$ sur C, on a par condition d'optimalité d'ordre 1 : $\nabla F(x_0) = x_0 - y_0 \in T_C(x_0)^*$, d'où il existe $\lambda \geqslant 0$ tel que $y_0 = x_0 + \lambda \nabla g(x_0)$. On ne peut avoir $\lambda = 0$ car sinon $y_0 = x_0 \in C$.