IFT 3515
Fonctions à plusieurs variables
Optimisation avec contraintes
Conditions d'optimalité

Fabian Bastin
DIRO
Université de Montréal

Multiplicateurs de Lagrange : contraintes d'égalité

Considérons le problème de programmation mathématique suivant

$$
\begin{aligned}
&\min_{x \in \mathcal{X}} \ f(x) \\
&\text{t.q. } g_i(x) = 0, \ \ i = 1, \ldots, m.
\end{aligned}
$$

où $\mathcal{X} \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$.

Le Lagrangien associé à ce problème est obtenu en associant un multiplicateur de Lagrange λ_i à chaque fonction de contrainte g_i :

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

Sans faire d'hypothèse particulière sur $\mathcal{X}$ ou sur les fonctions f et g_i, nous pouvons obtenir des conditions très générales pour qu'un point x^* soit une solution optimale du problème.

Optimalité

Théorème

Supposons que le Lagrangien associé au problème

$$\min_{x \in \mathcal{X}} f(x)$$
$$t.q.\ g_i(x) = 0,\ \ i = 1, \ldots, m.$$

possède un minimum local $x^* \in \mathcal{X}$ *lorsque le vecteur de multiplicateurs* λ *vaut* λ^*. *Si* $g_i(x^*) = 0$, $i = 1, \ldots, m$, *alors* x^* *est un minimum local de* $f(x)$.

Optimalité

Démonstration.

La preuve se fait par contradiction en supposant que x^* n'est pas un minimum local de $f(x)$. Alors $\forall \epsilon > 0$, $\exists \overline{x} \in \mathcal{B}(x^*, \epsilon)$ tel que $g_i(\overline{x}) = 0$, $i = 1, \ldots, m$ et $f(\overline{x}) < f(x^*)$.

Par conséquent, pour tout λ,

$$\sum_{i=1}^{m} \lambda_i g_i(x^*) = \sum_{i=1}^{m} \lambda_i g_i(\overline{x}) = 0.$$

Dès lors,

$$f(\overline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\overline{x}) < f(x^*) + \sum_{i=1}^{m} \lambda_i g_i(x^*).$$

En prenant $\lambda = \lambda^*$, la relation précédente contredit le fait que est un minimum local du Lagrangien lorsque $\lambda = \lambda^*$. $\square$

Multiplicateurs de Lagrange : contraintes d'inégalité

Considérons le problème de programmation mathématique suivant

$$\min_{x \in \mathcal{X}} f(x)$$
$$\text{t.q. } g_i(x) \leq 0, \; i = 1, \ldots, m.$$

où $\mathcal{X} \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$.

Théorème

Supposons que le Lagrangien associé au problème

$$\min_{x \in \mathcal{X}} f(x)$$
$$t.q. \; g_i(x) \leq 0, \; i = 1, \ldots, m.$$

possède un minimum local $x^* \in \mathcal{X}$ lorsque le vecteur de multiplicateurs λ vaut λ^*. Si $g_i(x^*) = 0$, $\lambda_i^* \geq 0$, et $\lambda_i^* g_i(x^*) = 0$, $i = 1, \ldots, m$ alors x^* est un minimum local de $f(x)$.

Multiplicateurs de Lagrange : contraintes d'inégalité

Démonstration.

Comme précédemment, la preuve se fait par contradiction en supposant que x^* n'est pas un minimum local de $f(x)$. Alors $\forall \epsilon > 0$, $\exists \, \overline{x} \in \mathcal{B}(x^*, \epsilon)$ tel que $g_i(\overline{x}) \leq 0$, $i = 1, \ldots, m$ et $f(\overline{x}) < f(x^*)$. Par conséquent, pour $\lambda = \lambda^* \geq 0$,

$$\sum_{i=1}^{m} \lambda_i g_i(\overline{x}) \leq 0 \text{ et } \sum_{i=1}^{m} \lambda_i g_i(x^*) = 0.$$

Dès lors,

$$f(\overline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\overline{x}) < f(x^*) + \sum_{i=1}^{m} \lambda_i g_i(x^*).$$

La relation précédente contredit le fait que x^* est un minimum local du Lagrangien lorsque $\lambda = \lambda^*$. $\square$

Généralisation

Considérons à présent un problème ayant des contraintes d'égalité et d'inégalité :

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & f(x) \\ \text{t.q. } & g_i(x) \leq 0, \; i = 1, \ldots, m \\ & h_j(x) = 0, \; j = 1, \ldots, r \end{aligned}$$

Nous définissons le Lagrangien comme

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{r} \mu_j h_j(x)$$

et la fonction duale lagrangienne

$$\mathcal{L}(\lambda, \mu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

Problème dual

Le problème dual est

$$\max_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^r} \mathcal{L}(\lambda, \mu)$$
$$\text{tel que } \lambda \geq 0.$$

Propriétés importantes :

- Le problème dual est toujours convexe, i.e. $\mathcal{L}$ est toujours concave (même si la problème primal n'est pas convexe).
- Les valeurs optimales (globales) primale et duale, f^* et $\mathcal{L}^*$, satisfont toujours la dualité faible : $f^* \geq \mathcal{L}^*$.
- Dualité forte : sous certaines conditions (qualifications de contraintes), $f^* = \mathcal{L}^*$.

Saut de dualité

Étant donné une solution primale réalisable x et une solution duale réalisable (λ, μ), la quantité $f(x) - \mathcal{L}(\lambda, \mu)$ est appelé le saut de dualité entre x et (λ, μ). Notons que

$$f(x) - f^* \leq f(x) - \mathcal{L}(\lambda, \mu)$$

de sorte que si le saut de dualité est nul, alors x est optimal primal (et similairement, λ et μ sont optimaux duaux).

D'un point de vue algorithmique, si la dualité forte tient, ceci fournit un critère d'arrêt : si $f(x) - \mathcal{L}(\lambda, \mu) \leq \epsilon$, nous avons alors la garantie que $f(x) - f^* \leq \epsilon$.

Saut de dualité : cas local

Désignons l'ensemble réalisable par

$$\mathcal{X} = \{x \,|\, g_i(x), i = 1, \ldots, m, h_j(x), j = 1, \ldots, r\}.$$

Considérons x^* un minimum local de $f(\cdot)$, i.e.

$$\exists \epsilon > 0 \text{ t.q. } \forall x \in \mathcal{B}(x^*, \epsilon) \cap \mathcal{X}, f(x^*) \leq f(x).$$

Nous pouvons également définir la fonction duale lagrangienne restreinte à la boule $\mathcal{B}(x^*, \epsilon)$:

$$\mathcal{L}_{\mathcal{B}(x^*,\epsilon)}(\lambda, \mu) = \min_{x \in \mathcal{B}(x^*,\epsilon)} L(x, \lambda, \mu).$$

Dans ce cas, la dualité faible tient toujours localement :

$$\mathcal{L}^*_{\mathcal{B}(x^*,\epsilon)} \leq f(x^*).$$

Sous certaines conditions, la dualité fort tiens également :

$$\mathcal{L}^*_{\mathcal{B}(x^*,\epsilon)} = f(x^*).$$

Saut de dualité : cas local

Remarquons cependant que

$$\min_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \leq \min_{x \in \mathcal{B}(x^*, \epsilon)} L(x, \lambda, \mu)$$

et donc

$$\mathcal{L}^* \leq \mathcal{L}^*_{\mathcal{B}(x^*, \epsilon)}.$$

Dès lors, si x^* est un minimum local et si la dualité forte tient localement,

$$\mathcal{L}^* \leq f(x^*),$$

l'inégalité pouvant être stricte.

Conditions de Karush-Kuhn-Tucker (KKT)

Étant donné le problème général

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \; & f(x) \\
\text{t.q.} \; & g_i(x) \leq 0, \;\; i = 1, \ldots, m \\
& h_j(x) = 0, \;\; j = 1, \ldots, r
\end{aligned}$$

les conditions de Karush-Kuhn-Tucker, ou conditions KKT, sont :

$$\begin{aligned}
\nabla_x L(x, \lambda, \mu) &= 0 && \text{(stationarité)} \\
\lambda_i g_i(x) &= 0 && \text{(écarts de complémentarités)} \\
g_i(x) \leq 0, \; h_j(x) &= 0 \;\forall i, j && \text{(faisabilité primale)} \\
\lambda_i &\geq 0 \;\forall i && \text{(faisabilité duale)}
\end{aligned}$$

Nécessité

Soit x^*, (λ^*, μ^*), des solutions globales primale et duale, avec un saut de dualité nul (la dualité forte tient). Alors

$$\begin{aligned}
f(x^*) &= \mathcal{L}(\lambda^*, \mu^*) \\
&= \min_{x \in \mathbb{R}^n} \left(f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) + \sum_{i=1}^{r} \mu_i^* h_i(x) \right) \\
&\leq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{i=1}^{r} \mu_i^* h_i(x^*) \\
&\leq f(x^*)
\end{aligned}$$

Dès lors, x^* est un minimum (global) de $L(x, \lambda^*, \mu^*)$ sur $\mathbb{R}^n$.
Par conséquent, $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$. Nous retrouvons les conditions de stationarité.
Nous devons aussi avoir $\sum_{i=1}^{m} \lambda_i^* g_i(x^*) = 0$ puisque $\sum_{i=1}^{m} \lambda_i^* g_i(x^*) \geq 0$. Ceci implique que pour tout i, $\lambda_i^* g_i(x^*) = 0$.
Nous retrouvons les conditions de complémentarité.

Nécessité : cas local

Le résultat tient encore pour la minimisation locale en considerant la fonction duale lagrangienne.

Nécessité

Théorème (Nécessité des conditions KKT)

Si x^*, (λ^*, μ^*) *sont des solutions primale et duale avec un saut de dualité nul (i.e. la dualité forte tient), alors* x^*, (λ^*, μ^*) *satisfont les conditions KKT.*

Dès lors, l'hypothèse de dualité forte apparaît importante. Elle sera garantie sous certaines conditions.

- Programme linéaire. La dualité forte tient toujours.
- Programme convexe. Condition de Slater : il existe un x tel que $g_i(x) < 0$, $i = 1, \ldots, m$ et $h_i(x) = 0$, $i = 1, \ldots, r$.
- Programme non-convexe. Hypothèse de qualification de contraintes. La plus courante, mais aussi la plus forte, est la condition d'indépendance linéaire des gradients à la solution.

Nécessité (cas non convexe)

Théorème (Nécessité des conditions KKT)

Si x^* est une solution locale de

$$\begin{aligned} \min_{x \in \mathcal{X}} \ & f(x) \\ t.q.\ & g_i(x) \leq 0, \ i = 1, \ldots, m \\ & h_i(x) = 0, \ i = 1, \ldots, r, \end{aligned}$$

où les fonctions f, g_i et h_i, $i = 1, \ldots, m$, sont continûment différentiables, et qu'une condition de qualification de contrainte tient en x^*. Alors, il existe un vecteur de multiplicateurs de Lagrange (λ^*, μ^*) tel que les conditions KKT sont satisfaites en (x^*, λ^*, μ^*).

Démonstration.

Preuve : technique ! Voir par exemple Nocedal & Wright, "Numerical Optimization", Section 12.4.

Suffisance des conditions KKT

S'il existent x^*, (λ^*, μ^*) satisfaisant les conditions KKT, alors

$$\begin{aligned} L(\lambda^*, \mu^*) &= f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{i=1}^{r} \mu_i^* h_i(x^*) \\ &= f(x^*) \end{aligned}$$

Dès lors, le saut de dualité est nul (dualité forte).

Dans le cas convexe, cela implique que x^* et (λ^*, μ^*) sont des solutions globales primale et duale, respectivement.

Dans la cas non-convexe, x^* est un minimum local, pas nécessairement global, voire un point-selle.

Contraintes linéaires

Revenons à la méthode de projection. Un autre cas important de contraintes relativement faciles à traiter sont les contraintes d'égalité

$$Ax = b$$

avec $A \in \mathbb{R}^{m \times n}$ de rang plein.
Nous pouvons généraliser le problème de projection de y sur l'ensemble

$$\mathcal{X} = \{x \mid Ax = b\}$$

en considérant la norme-2 ou la norme engendré par une matrice H définie positive

$$\min_x \frac{1}{2}\|y - x\|_H \text{ tel que } Ax = b.$$

Nous allons résoudre ce problème en utilisant les conditions d'optimalité de Karush-Kuhn-Tucker (KKT).

Projection sur contraintes linéaires

Considérons le problème

$$\min_x \frac{1}{2}\|y - x\|_H \text{ tel que } Ax = b.$$

Le problème est convexe, et satisfait la condition de Slater s'il existe au moins un point réalisable.

Les conditions KKT s'écrivent

$$\begin{aligned} \nabla_x \frac{1}{2}\langle y - x, H(y - x)\rangle + \nabla_x (Ax - b)^T \lambda &= 0 \\ Ax - b &= 0 \end{aligned}$$

ou

$$\begin{aligned} H(y - x) + A^T \lambda &= 0 \\ Ax &= b \end{aligned}$$

Projection sur contraintes linéaires

Le problème peut être réorganisé comme

$$
\begin{aligned}
Hx - A^T\lambda &= Hy \\
Ax + 0\lambda &= b
\end{aligned}
$$

donnant lieu au système linéaire

$$
\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ -\lambda \end{pmatrix} = \begin{pmatrix} Hy \\ b \end{pmatrix}
$$

Si H est inversible, ce qui sera le cas dans nos problèmes comme nous prendrons H définie positive, nous pouvons résoudre le système en isolant x et en le susbtituant. Tout d'abord, nous avons

$$
x = H^{-1}Hy + H^{-1}A^T\lambda = y + H^{-1}A^T\lambda
$$

et donc

$$
A\left(y + H^{-1}A^T\lambda\right) = b
$$

Projection sur contraintes linéaires

On en tire

$$AH^{-1}A^T\lambda = b - Ay$$

puis, une fois λ déterminé

$$Hx = Hy + A^T\lambda$$

Application

Considérons le problème

$$\min_x \; f(x)$$
$$\text{t.q. } Ax = b$$

J. B. Rosen, "The Gradient Projection Method for Nonlinear Programming. Part I. Linear Constraints", Journal of the Society for Industrial and Applied Mathematics, 8(1), pp. 181-217, 1960.

Supposons que nous avons un point de départ x_0 tel que $Ax_0 = b$. Nous souhaitons garder tous les itérés réalisables, i.e. $Ax_k = b$.

Étant donné la direction de recherche d_k, nous devons dès lors avoir $Ad_k = 0$

Projection de la plus forte pente

Considérons $d_k = -\nabla f(x_k)$.

De ce qui précède, en prenant $H = I$, et en notant $\overline{d}_k$ la projection de d_k,

$$\begin{aligned}
\overline{d}_k &= d_k + A^T \lambda \\
&= d_k - A^T (AA^T)^{-1} A d_k \\
&= \left(I - A^T (AA^T)^{-1} A \right) d_k \\
&= -\left(I - A^T (AA^T)^{-1} A \right) \nabla f(x_k)
\end{aligned}$$

Algorithme

- **Étape 0** Soit x_0. Poser $k = 0$.
- **Étape 1** Calcul de la direction de recherche :

$$\overline{d}_k = -\left(I - A^T(AA^T)^{-1}A\right)\nabla f(x_k).$$

  Si $\|\overline{d}_k\| = 0$, arrêt : x_k est optimal. Sinon, aller à l'étape 2.
- **Étape 2** Résoudre (approximativement)

$$\min_{\alpha \geq 0} f(x_k + \alpha \overline{d}_k)$$

  Soit α_k la solution. Poser $x_{k+1} = x_k + \alpha_k \overline{d}_k$, et $k := k + 1$. Retour à l'étape 1.

Remarques

Nous avons ici directement projeté la direction de recherche. Il est possible de montrer que $\overline{d}_k$ est solution du problème

$$\begin{aligned} &\min_d \nabla f(x_k)^T d \\ &\text{t.q. } Ad = 0, \ \|d\| = 1. \end{aligned}$$

L'article de Rosen considère le cas plus général de contraintes sous la forme

$$\begin{aligned} a_i x &= b_i, \ i = 1, \dots, m_1 \\ a_i x &\leq b_i, \ i = m_1, \dots, m \end{aligned}$$

L'algorithme doit alors considérer quelles sont les contraintes actives.

Ensemble actif

Définition (Ensemble actif)

L'ensemble actif $\mathcal{A}(x)$ *du problème d'optimisation*

$$\begin{aligned} \min_{x \in \mathcal{X}} \; & f(x) \\ t.q. \; & g_i(x) \leq 0, \; i \in \mathcal{I} \\ & h_i(x) = 0, \; i \in \mathcal{E}, \end{aligned}$$

en un point réalisable x *est l'ensemble des indices des contraintes d'égalité et l'ensemble des indices* i *des contraintes d'inégalité telles que* $g_i(x) = 0$*, c'est-à-dire*

$$\mathcal{A}(x) = \mathcal{E} U \{ i \mid g_i(x) = 0 \}$$

LICQ

Définition (LICQ)

Étant donné le point x et l'ensemble actif $\mathcal{A}(x)$, nous disons que la qualification de contraintes d'indépendance linéaire (linear independence constraint qualification – LICQ) tient si l'ensemble des contraintes actives $\nabla_x c_i(x), i \in \mathcal{A}(x)$ est linéairement indépendant.

Lemme

Soit x^* un point du problème de programmation nonlinéaire où la LICQ tient et soit $d \in \mathbb{R}^n$ un vecteur tel que

$$\begin{aligned} d &\neq 0 \\ d^T \nabla g_i(x^*) &= 0, \;\; i \in \mathcal{E} \\ d^T \nabla g_i(x^*) &\leq 0, \;\; i \in \mathcal{A}(x^*) \cap \mathcal{I} \end{aligned}$$

LICQ

Lemme (suite)

Alors pour $\epsilon > 0$ *assez petit, il existe un chemin* $x(\cdot) \in C^2\left((-\epsilon, \epsilon), \mathbb{R}^n\right)$ *tel que*

$$
\begin{aligned}
x(0) &= x^* \\
\frac{d}{dt}x(0) &= d \\
g_i(x(t)) &= td^T \nabla g_i(x^*), \;\; i \in \mathcal{A}(x^*), \;\; t \in (-\epsilon, \epsilon)
\end{aligned}
$$

de sorte que

$$
\begin{aligned}
g_i(x(t)) &= 0, \;\; i \in \mathcal{E}, \;\; t \in (-\epsilon, \epsilon) \\
g_i(x(t)) &\leq 0, \;\; i \in \mathcal{I}, \;\; t \in [0, \epsilon)
\end{aligned}
$$

Preuve

Soit $\ell = \#\mathcal{A}$. Puisque la LICQ tient, il est possible de choisir $Z \in \mathbb{R}^{(n-\ell)\times n}$ telle que

$$\begin{pmatrix} \nabla g_i(x^*)^T, \ i \in \mathcal{A}(x^*) \\ Z \end{pmatrix}$$

soit une matrice nonsingulière.

Soit une fonction $h : \mathbb{R}^n \times \mathbb{R}$ définie par

$$h(x,t) = \begin{pmatrix} \left(g_i(x), \ i \in \mathcal{A}(x^*)\right) - t\left(\nabla g_i(x^*)^T, \ i \in \mathcal{A}(x^*)\right) d \\ Z(x - x^* - td) \end{pmatrix}$$

La matrice jacobienne de h en $(x^*, 0)$ vaut

$$Dh(x^*, 0) = \begin{pmatrix} D_x h(x^*, 0) & D_t h(x^*, 0) \end{pmatrix}$$

Rappel : matrice jacobienne

Soit

$$F : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}.$$

La matrice jacobienne de F est

$$DF = \begin{pmatrix} \nabla_x^T f_1 \\ \vdots \\ \nabla_x^T f_m \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Si $n = m$, le jacobien de F est le déterminant de DF.

Preuve

Les éléments de

$$Dh(x^*, 0) = \begin{pmatrix} D_x h(x^*, 0) & D_t h(x^*, 0) \end{pmatrix}$$

sont

$$D_x h(x^*, 0) = \begin{pmatrix} \left(\nabla g_i(x^*)^T, \ i \in \mathcal{A}(x^*)\right) \\ Z \end{pmatrix}$$

et

$$D_t h(x^*, 0) = -\begin{pmatrix} \left(\nabla g_i(x^*)^T, \ i \in \mathcal{A}(x^*)\right) d \\ Zd \end{pmatrix} = -D_x h(x^*, 0) d$$

Puisque $D_x h(x^*, 0)$ est non singulier, le théorème des fonctions implicites implique que pour $\delta > 0$ assez petit, il existe une fonction unique $x \in C^2 : (-\delta, \delta) \to \mathbb{R}^n$ et un voisinage $\mathcal{V}(x^*)$ tel que pour $x \in \mathcal{V}(x^*)$, $t \in (-\delta, \delta)$,

$$h(x, t) = 0 \Leftrightarrow x = x(t)$$

Rappel : théorème des fonctions implicites

Soit $F \in C^k : \Omega \subseteq \mathcal{R}^n \times \mathcal{R}^p \to \mathcal{R}^p$, notée

$$F(x, y) = (F_1(x, y), F_2(x, y), \ldots, F_p(x, y))$$

S'il existe (a, b) tel que $F(a, b) = 0$ et pour lequel le jacobien de $D_y F$ ($D_y F$ est inversible) est non nul, alors

1. il existe un voisinage A de a dans $\mathcal{R}^n$ et un voisinage B de b dans $\mathcal{R}^p$, une fonction $f \in C^k(A, B)$, tels que $A \times B \subseteq \Omega$ et que $\forall (x, y) \in A \times B$,

$$F(x, y) = 0 \Leftrightarrow y = f(x).$$

2. la matrice jacobienne de f par rapport à x s'écrit

$$D_x f = -\left(D_y F(x, f(x))\right)^{-1} \left(D_x F(x, f(x))\right).$$

Preuve

En particulier, nous avons $h(x^*, 0) = 0$, et donc $x(0) = x^*$. De plus, $h(x(t), t) = 0$ donne, par définition de h,

$$g_i(x(t)) = td^T \nabla g(x^*)$$

pour tout $i \in \mathcal{A}(x^*)$ et $t \in (-\delta, \delta)$. Les conditions du lemme sur d impliquent dès lors que

$$\begin{aligned} &g_i(x(t)) = 0, \; i \in \mathcal{E} \\ &g_i(x(t)) \leq 0, \; i \in \mathcal{A}(x^*) \cap \mathcal{I}, \; t \in [0, \delta) \end{aligned}$$

D'autre part, puisque $g_i(x^*) < 0$, $i \notin \mathcal{A}(x^*)$, la continuité de $x(t)$ implique qu'il existe $\epsilon \in (0, \delta)$ tel que

$$g_j(x^*) < 0, \; j \in \mathcal{I} \setminus \mathcal{A}(x^*), \; t \in (-\epsilon, \epsilon).$$

Finalement, du théorème des fonctions implicites, nous pouvons tirer que

$$\frac{d}{dt} x(0) = -(D_x h(x^*, 0))^{-1} D_t h(x^*, 0) = d$$

Retour au cône tangent

Il est également possible de montre que sous la LICQ, le cône tangent à l'ensemble réalisable $\mathcal{X}$ en x peut se réécrire comme

$$T_{\mathcal{X}}(x) = \{d \mid d^T \nabla g_i(x) = 0, i \in \mathcal{E},\ d^T \nabla g_i(x) \leq 0, i \in \mathcal{A}(x) \cap \mathcal{I}\}$$

Dès lors,

$$\frac{d}{dt}x(0) \in T_{\mathcal{X}}(x^*).$$

Nous allons utiliser ce lemme pour introduire la notion de chemin réalisable, lequel nous permettra de davantage caractériser les conditions prévalentes à une solution du problème d'optimisation.

Chemin de sortie réalisable

Définition

Soit $x^* \in \mathbb{R}^n$ un point réalisable pour le PNL et définissons $x \in C^2\left((-\epsilon, \epsilon), \mathbb{R}^n\right)$ un chemin tel que

$$
\begin{aligned}
x(0) &= x^* \\
d &:= \frac{d}{dt}x(0) \neq 0 \\
g_i(x(t)) &= 0,\ i \in \mathcal{E},\ t \in (-\epsilon, \epsilon) \\
g_i(x(t)) &\leq 0,\ i \in \mathcal{I},\ t \in [0, \epsilon)
\end{aligned}
$$

$x(t)$ est un chemin de sortie réalisable à partir de x^* et le vecteur tangent $d = \frac{d}{dt}x(0)$ est une direction de sortie réalisable à partir de x^*.

Nous pouvons imaginer que $x(t)$ est un morceau lisse de trajectoire d'une particule passant à travers x^* au temps $t = 0$ avec une vitesse non nulle d et qui se déplace dans le domaine réalisable.