Optimisation Non linèaire

Par

Professeur Abdellatif El Afia

Chapitre 5

Optimisation avec contraintes
Conditions d'optimalité

1. **Multiplicateur de Lagrange**
2. **Conditions de Karush-Kuhn-Tucker**
3. **Condition K-K-T suffisantes**
4. **Condition K-K-T nécessaires**
5. Problème de programmation convexe

Conditions d'optimalité: Multiplicateur de Lagrange

Considérons le problème de programmation mathématique suivant:

$$(\boldsymbol{P_1})\begin{cases} Min & f(x) \\ s.t & f_i(x) = 0 \;\; i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{cases}$$

les fonctions $f: \mathbb{R}^n \to \mathbb{R}$, $f_i: \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$

Le **lagrangien** associé au problème $(\boldsymbol{P_1})$ est obtenu comme suit en associant un **multiplicateur de Lagrange** λ_i à chaque fonction de contrainte f_i : $\boldsymbol{L(\lambda, x) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)}$

- Sans faire d'hypothèse particulière sur X ou sur les fonctions f et f_i, nous pouvons obtenir des conditions suffisantes très générales pour qu'un point x^* soit une solution optimale globale problème $(\boldsymbol{P_1})$

Théorème 1 : Supposons que le lagrangien, $\boldsymbol{L(\lambda, x)}$, associé au problème $(\boldsymbol{P_1})$ possède un minimum global x^* sur X lorsque $\lambda = \lambda^*$. Si $\forall i \in \{1, \dots, m\}$: $f_i(x^*) = 0$, alors x^*est une solution optimale globale de $(\boldsymbol{P_1})$

Preuve : la preuve se fait par contradiction en supposant que x^* n'est pas une solution optimale de $(\boldsymbol{P_1})$.

- Alors $\exists \bar{x}$ tel que $\forall i \in \{1, \dots, m\}$ $f_i(x^*) = 0$, et $f(\bar{x}) < f(x^*)$,
- Par conséquent, $\forall \lambda$ $\sum_{i=1}^{m} \lambda_i f_i(\bar{x}) = \sum_{i=1}^{m} \lambda_i f_i(x^*) = 0$ et ainsi $f(\bar{x}) + \sum_{i=1}^{m} \lambda_i f_i(\bar{x}) < f(x^*) + \sum_{i=1}^{m} \lambda_i f_i(x^*)$.
- En prennent $\lambda = \lambda^*$ la relation précédente contredit le fait que x^* est un minimum global du lagrangien sur X lorsque $\lambda = \lambda^*$

Conditions d'optimalité: Multiplicateur de Lagrange

Exemple

$$(\boldsymbol{P_1})\begin{cases} Min & f(x,y) = x^2 + 2y^2 \\ s.t & f_1(x,y) = x + y - b = 0 \\ & (x,y) \in \mathbb{R}^2 \end{cases}$$

Lagrangien correspondant $L(\lambda, x, y) = (x^2 + 2y^2) + \lambda(x + y - b)$
L convexe en (x,y). Donc minimum atteint lorsque $\nabla_{x,y} L(\lambda, x, y) = 0$, Cherchons (x,y)

- $\nabla_{x,y} L(\lambda, x, y) = \begin{pmatrix} 2x + \lambda \\ 4y + \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} \lambda = -2x \\ \lambda = -4y \end{cases} \Leftrightarrow \boldsymbol{x = 2y}$
- $f_1(x,y) = x + y - b = 2y + y - b = 3y - b = 0 \Leftrightarrow \boldsymbol{y = \frac{b}{3}}$

Donc :

- $(x,y) = \left(\frac{2}{3}b, \frac{1}{3}b\right)$
- $\lambda = -2x = -\frac{4}{3}b$

Conditions d'optimalité: Multiplicateur de Lagrange

Considérons maintenant le problème de programmation mathématique suivant

$$(P_2)\begin{cases} Min & f(x) \\ s.t & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{cases}$$

Théorème 2 : Supposons que le lagrangien associé au problème (P_2): $L(\lambda, x) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$

Possède un minimum global x^* sur X lorsque le vecteur de multiplicateurs $\lambda = \lambda^*$.

Si $\forall i = 1, \dots, m$, $f_i(x^*) \leq 0$, $\lambda_i^* \geq 0$ et $\lambda_i^* f_i(x^*) = 0$ **alors** x^*est une solution optimale globale (P_2)

Preuve : La preuve se fait par contradiction,

Supposons que x^* n'est pas une solution optimale de (P_2). Alors il existe un $\bar{x}$ tel que $f_i(\bar{x}) \leq 0$ pour tout $i = 1, \dots, m$ et $f(\bar{x}) < f(x^*)$

Par conséquent, pour $\lambda = \lambda^* \geq 0, \sum_{i=1}^{m} \lambda_i^* f_i(\bar{x}) \leq 0$ et $\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$

Et ainsi $f(\bar{x}) + \sum_{i=1}^{m} \lambda_i' f_i(\bar{x}) < f(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*)$

La relation précédente contredit le fait que x^* est un minimum global du lagrangien sur X lorsque $\lambda = \lambda^*$.

Conditions d'optimalité: Multiplicateur de Lagrange

Exemple: Si $f_i(x^*) \leq 0$, $\lambda_i^* \geq 0$ et $\lambda_i^* f_i(x^*) = 0$

$$(\boldsymbol{P_2}) \begin{cases} Min & f(x) = x^2 \\ s.t & f_1(x) = 2x + 5 \leq 0 \\ & x \in \mathbb{R} \end{cases}$$

Lagrangien correspondant $L(\lambda, x) = x^2 + \lambda(2x + 5)$

L convexe en x. Donc minimum atteint lorsque $\nabla_x L(\lambda, x) = 0$, Cherchons x

- $\nabla_x L(\lambda, x) = 2x + 2\lambda = 0 \Leftrightarrow x = -\lambda$
- $\lambda_1^* f_1(x^*) = \lambda(2x + 5) = 0$
- $Si\ \lambda \neq 0 \Longrightarrow 2x + 5 = 0 \Longrightarrow -2\lambda + 5 = 0 \Leftrightarrow \lambda = \frac{5}{2}$
- Si $\lambda = 0 \Longrightarrow x = -\lambda = 0 \Longrightarrow 2x + 5 = 5 > 0$ n'est pas une solution réalisable

Donc $x = -\lambda = -\frac{5}{2}$

on note, $\lambda = \frac{5}{2} \geq 0$

$2x + 5 = 0 \Rightarrow (2x + 5)\lambda = 0$

Conditions d'optimalité:Conditions de Karush-Kuhn-Tucker (KKT)

Pour obtenir des conditions plus facilement vérifiables, il faut poser des hypothèses sur X et sur les fonctions f et f_i

Si f et f_i sont **différentiables et convexes** dans le problème $(\boldsymbol{P_2})$

$$(\boldsymbol{P_2})\begin{cases} Min & f(x) \\ s.t & f_i(x) \leq 0 \quad i = 1, \ldots, m \\ & x \in \mathbb{R}^n \end{cases}$$

alors le lagrangien $L(\lambda, x) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$ est aussi une fonction convexe si $\lambda \geq 0$ puisque:

- $\lambda_i \geq 0$ et $f_i(x)$ convexe $\Rightarrow \lambda_i f_i(x)$ convexe
- $f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$ somme de fonctions convexes.

et il s'ensuit qu'il possède un minimum global en x^* lorsque $\lambda = \lambda^* \geq 0$ si $\nabla_x L(\lambda^*, x) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x) = 0$

Conditions d'optimalité : Conditions de Karush-Kuhn-Tucker (KKT)

Alors le théorème 2 peut s'écrire de la forme suivant:
Si les conditions de (K-K-T) est vérifier :

$$\textbf{(K−K−T)}\begin{cases} \nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0 \\ f_i(x^*) \leq 0, i = 1, \dots, m \\ \lambda_i^* \geq 0, i = 1, \dots, m \\ \lambda_i^* f_i(x^*) = 0, i = 1, \dots, m \end{cases}$$

alors x^* est une solution optimal globale

Définition: Une contrainte $f_i(x)$ est dite active au point x^* si $f_i(x^*) = 0$

Remarque: $\boldsymbol{\lambda_i^* f_i(x^*) = 0}$

- Si $\lambda_i^* > 0$ alors la contrainte associée $f_i(x)$ est active au point x^*
- Si $f_i(x) < 0$ alors $\lambda_i^* = 0$

Conditions d'optimalité : Conditions de Karush-Kuhn-Tucker (KKT)

Exemple:

$$\begin{cases} Min & f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ s.t & x_1^2 + x_2^2 \leq 5 \\ & 3x_1 + x_2 \leq 6 \\ & x = (x_1, x_2) \in \mathbb{R}^2 \end{cases}$$

- $L(\lambda, x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 + \lambda_1(x_1^2 + x_2^2 - 5) + \lambda_2(3x_1 + x_2 - 6)$
- $\nabla_x L(\lambda, x) = \begin{pmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 0 \Longrightarrow \begin{matrix} 4x_1 + 2x_2 - 10 + 2\lambda_1 x_1 = 0 \\ 2x_1 + 2x_2 - 10 + 2\lambda_1 x_2 = 0 \end{matrix}$
- $\begin{cases} \lambda_1(x_1^2 + x_2^2 - 5) = 0 \\ \lambda_2(3x_1 + x_2 - 6) = 0 \end{cases} \Longrightarrow Cas: \begin{cases} x_1^2 + x_2^2 - 5 = 0 \\ 3x_1 + x_2 - 6 < 0 \Rightarrow \lambda_2 = 0 \end{cases}$
- $\begin{cases} x_1^2 + x_2^2 - 5 \leq 0 \\ 3x_1 + x_2 - 6 \leq 0 \end{cases} \Longrightarrow \begin{matrix} x_1^2 + x_2^2 - 5 = 0 \\ 3x_1 + x_2 - 6 < 0 \end{matrix}$
- $\lambda_1, \lambda_2 \geq 0 \Longrightarrow \lambda_1 \geq 0, \lambda_2 = 0$

Conditions d’optimalité : Conditions de Karush-Kuhn-Tucker (KKT)

Exemple:

Nous pouvons faire différentes hypothèses sur quelle contrainte est active dans le but d’identifier des valeurs de $x_1, x_2, \lambda_1, \lambda_2$ satisfaisant les conditions KKT.

Supposons que la première est active et que la seconde ne l’est pas :

$$x_1^2 + x_2^2 - 5 = 0$$
$$3x_1 + x_2 - 6 < 0 \Rightarrow \lambda_2 = 0$$

Nous retrouvons alors le système avec 3 équations et 3 inconnus suivant :

$$\begin{aligned} 4x_1 + 2x_2 - 10 + 2\lambda_1 x_1 &= 0 \\ 2x_1 + 2x_2 - 10 + 2\lambda_1 x_2 &= 0 \\ x_1^2 + x_2^2 &= 5 \end{aligned}$$

Nous pouvons vérifier que $x_1 = 1, x_2 = 2, \lambda_1 = 1$ satisfont le système précédent.

Donc $x_1 = 1, x_2 = 2, \lambda_1 = 1, \lambda_2 = 0$ Satisfont les conditions KKT

Conditions d'optimalité : Condition K-K-T suffisantes

Théorème 3 :
Supposons que les fonctions f et f_i sont différentiables et convexes.
Si les conditions de KKT sont vérifiées à x^*, **alors** x^* est un minimum global du problème $(\boldsymbol{P_2})$

Preuve :
Le lagrangien étant **convexe** si $\lambda^* \geq 0$ (démontrer précédemment) alors par **l'inégalité du gradient** il s'ensuit que $\forall x \in \mathbb{R}^n$

$$f(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) \geq f(x^*) + \underbrace{\sum_{(i=1)}^{m} \lambda_i^* f_i(x^*)}_{0} + \underbrace{\left[\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)\right]^T}_{0} (x - x^*)$$

Alors $\forall x \in \mathbb{R}^n$:

$$f(x^*) - f(x) \leq \sum_{i=1}^{m} \lambda_i^* f_i(x) \leq 0 \Longrightarrow f(x^*) \leq f(x)$$

Conditions d'optimalité : Condition K-K-T **nécessaires**

Pour l'analyse de la nécessité des conditions KKT nous allons exploiter le théorème d'alternative

Supposons que x^*est une solution **locale** du problème $(\boldsymbol{P_2})$

$$(\boldsymbol{P_2}) \begin{cases} Min & f(x) \\ s.t & f_i(x) \leq 0 \;\; i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{cases}$$

Notation : Dénotons l'ensemble des contraintes actives

$$\boldsymbol{A(x^*) = \{i: f_i(x^*) = 0\} = \{i_1, \dots, i_k\} \subset \{1, \dots, m\}}$$

Hypothèse à vérifier : Supposons que nous pouvons démontrer que $\nexists d \in \mathbb{R}^n$ tel que

$$\begin{cases} \nabla f(x^*)^T d < 0 \\ \nabla f_i(x^*)^T d \leq 0, \forall i \in A(x^*) \end{cases} \Leftrightarrow (S_1) \begin{cases} \nabla f(x^*)^T d < 0 \\ \left[\nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*)\right]^T d \leq 0 \end{cases}$$

Ainsi le système (S_1) ne possède pas de solution. Appliquons maintenant le théorème d'alternative en prenons $A^T = \left[\nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*)\right]^T$et b $= \nabla f(x^*)$, On a

- $(S_1) \begin{cases} b^T d < 0 \\ -A^T d \geq 0 \end{cases}$ ne possède pas une solution
- $(S_2) \begin{cases} -A\lambda^* = b \\ \lambda^* = [\lambda^*_{i_1}, \dots, \lambda^*_{i_k}] \geq 0 \end{cases} \Leftrightarrow \begin{cases} -\sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*) \\ \lambda^* = [\lambda^*_{i_1}, \dots, \lambda^*_{i_k}] \geq 0 \end{cases}$

Conditions d'optimalité : Condition K-K-T **nécessaires**

(S_2) s'écrit également sous la forme:

- $\boldsymbol{\nabla f(x^*) + \sum_{i\in A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0}$
- $\boldsymbol{\lambda_i^* f_i(x^*) = 0,\ \forall i \in A(x^*)}$
- $\boldsymbol{f_i(x^*) \leq 0, \qquad i = 1, \dots, n}$
- $\boldsymbol{\lambda^* = \left[\lambda_{i_1}^*, \dots, \lambda_{i_k}^*\right] \geq 0}$

Posons $\lambda_i^* = 0$ pour tout $i \notin A(x^*)$. Alors :

- $\sum_{i\notin A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0,\ \lambda_i^* f_i(x^*) = 0,$ pour tout $i \notin A(x^*)$

Par conséquent nous retrouvons les conditions KKT

- $\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$
- $\lambda_i^* f_i(x^*) = 0, \qquad i = 1, \dots, n$
- $f_i(x^*) \leq 0, \qquad i = 1, \dots, n$
- $\lambda_i^* \geq 0, \qquad i = 1, \dots,$

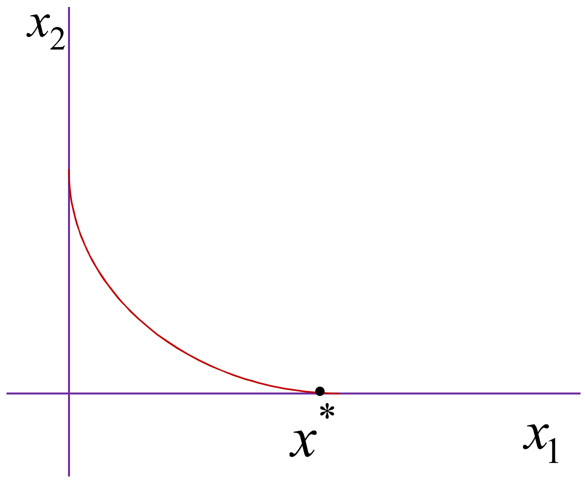
Malheureusement, notre hypothèse que nous pouvons démontrer que $\nexists d \in \mathbb{R}^n$ tel que

$$\nabla f_i(x^*)^T d \leq 0, \qquad i \in A(x^*)$$
$$\nabla f(x^*)^T d < 0$$

Ne l'est pas nécessairement pour toute solution locale x^* de tout problème tel que l'illustre l'exemple suivant.

Conditions d'optimalité : Condition K-K-T **nécessaires**

$$\begin{cases} Min & f(x_1,x_2) = -x_1 \\ s.t & f_1(x_1,x_2) = (x_1-1)^3 + x_2 \leq 0 \\ & f_2(x_1,x_2) = -x_1 \leq 0 \\ & f_3(x_1,x_2) = -x_2 \leq 0 \end{cases}$$



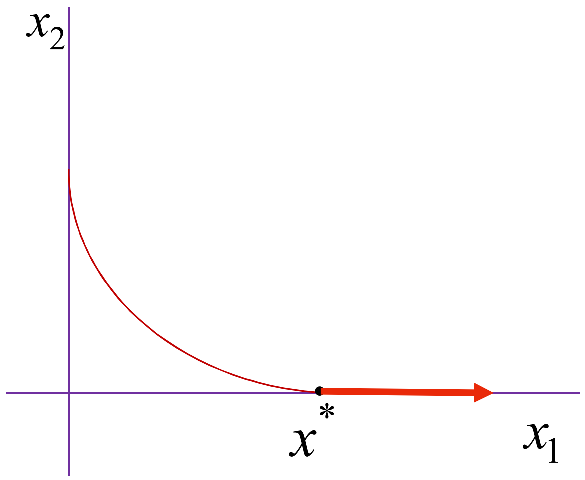
L'ensemble des solutions réalisables de ce problème est représenté par la région en-dessous de la courbe de $f_1(x_1,x_2)$ au-dessus de l'axe des x_1 et à droite de l'axe des x_2

Il est facile de vérifier que $x^* = [1{,}0]^T$ est une solution optimale globale de ce problème.

De plus $A(x^*) = \{1{,}3\}$

Conditions d'optimalité : Condition K-K-T **nécessaires**

$$\begin{cases} Min & f(x_1,x_2) = -x_1 \\ s.t & f_1(x_1,x_2) = (x_1-1)^3 + x_2 \leq 0 \\ & f_2(x_1,x_2) = -x_1 \leq 0 \\ & f_3(x_1,x_2) = -x_2 \leq 0 \end{cases}$$



Or $\nabla f(x^*) = [-1,0]^T$
$\nabla f_1(x) = [3(x_1-1)^2, 1]^T$ et $\nabla f_3(x^*) = [0,1]^T$

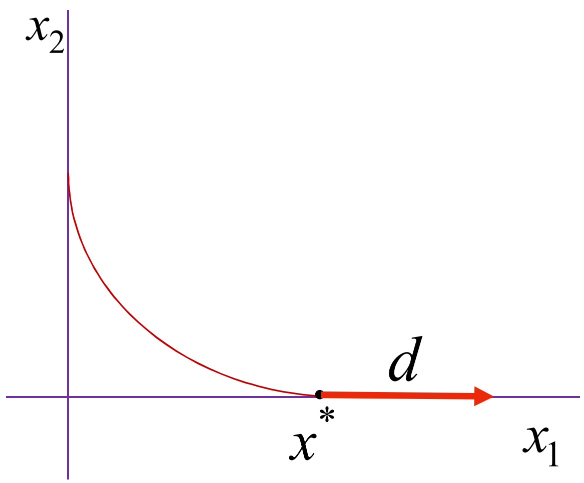
Le système :

$$\nabla f(x^*)^T d = -d_1 < 0$$
$$\nabla f_1(x^*)^T d = \quad d_2 \leq 0$$
$$\nabla f_3(x^*)^T d = -d_2 \leq 0$$

Possède une solution $d = [1,0]$

Conditions d'optimalité : Condition K-K-T **nécessaires**

Notons qu'au point x^* la direction $d = [1,0]$ pointe directement à l'extérieur du domaine réalisable.
Nous allons donc imposer certaines restrictions sur les contraintes des problèmes considérés pour éliminer de telles situations.



Restrictions sur les fonctions de contraintes

Notation : D_R dénote le domaine réalisable du problème $(\boldsymbol{P_2})$

$$\delta D_R = \{x \in D_R : \exists i, 1 \leq i \leq m, \text{tel que } f_i(x) = 0\}$$

Définition : étant donné un point $\bar{x} \in \delta D_R$, On dit que $f_1, \dots, f_m$ **satisfont les restrictions sur les fonctions de contraintes** au point $\bar{x}$ si pour tout vecteur $\tilde{d}$ solution du système $\nabla f_i(\bar{x})^T d \leq 0, i \in A(\bar{x})$, il existe une fonction différentiable $\alpha : [0,1] \to D_R$ tel que $\alpha(0) = \bar{x}$ et $\alpha'(0) = \sigma \tilde{d}, \sigma > 0$.

Conditions d'optimalité : Restrictions sur les fonctions de contraintes

Interprétation :

Se référant à la notion de direction de descente, lorsque

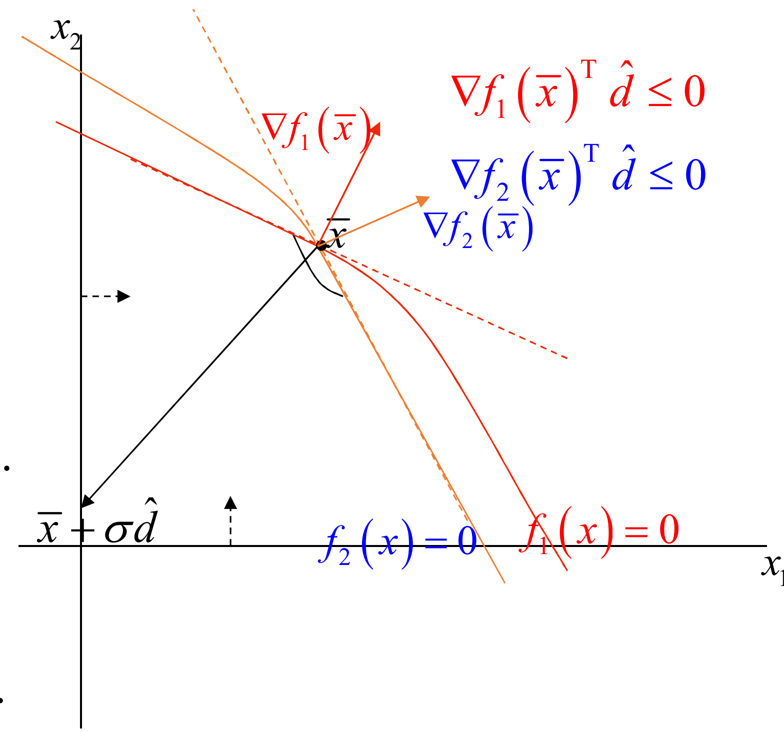
$$\boldsymbol{\nabla f_i(\bar{x})^T d < 0}$$

alors pour un faible déplacement $\tau > 0$ dans la direction d,

$$\boldsymbol{f_i(\bar{x} + \tau d) < f_i(\bar{x}) = 0}$$

Ainsi, ce déplacement nous garde dans le domaine réalisable par rapport à cette contrainte. Les restrictions sur les fonctions de contraintes prolongent en quelque sorte cette propriété même si $\nabla f_i(\bar{x})^T d = 0$ puisque la fonction α prend ses valeurs dans le domaine réalisable D_R.

Interprétation géométrique :



$$\alpha(\lambda) = \bar{x} + \lambda \sigma \hat{d}$$
$$\alpha(0) = \bar{x}$$
$$\alpha'(0) = \sigma \hat{d}$$

Conditions d'optimalité : Condition K-K-T **nécessaires**

Théorème 8 : (Nécessité des conditions K-K-T) : Soit $x^* \in \bar{X}$ une solution optimale locale du problème $(\boldsymbol{P_2})$ où $\bar{X}$ est ouvert. Supposons de plus que si $x^* \in \delta D_R$, alors $f_1, \dots, f_m$ satisfont les restrictions sur les fonctions de contraintes au point x^*.
Alors il existe un vecteur de multiplicateurs $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$ tel que :

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m$$

Preuve :
Si x^* est un point intérieur du domaine réalisable D_R (i.e.,$f_i(x^*) < 0$) pour tout i), il suffit de prendre $\lambda_i^* = 0$ pour tout $i = 1, \dots, m$.
En effet dans ce cas si $\nabla f(x^*)$ prenait une valeur différente de 0, il suffirait de considérer la direction $d = -\nabla f(x^*)$ qui serait une direction de descente de f à x^*.
Ainsi, il existerait un $\tau > 0$ suffisamment petit pour que
$(x^* + \tau d) \in B_\tau(x^*) \cap D_R$ avec $f(x^* + \tau d) < f(x^*)$, une contradiction.

Conditions d'optimalité : Condition K-K-T **nécessaires**

Soit $\boldsymbol{x^*} \in \boldsymbol{\delta} D_R$, démontrons que l'hypothèse que nous pouvons démontrer qu'il n'existe pas de vecteur $d \in \mathbb{R}^n$ tel que

$$\nabla f_i(x^*)^T d \leq 0, \qquad i \in A(x^*)$$
$$\nabla f(x^*)^T d < 0$$

est effectivement vérifiée sous **les hypothèses du théorème**.

En effet, pour fin de contradiction, Supposons qu'un tel vecteur $\tilde{d}$ existerait, puisque $f_1, \dots, f_m$ satisfont les restrictions sur les fonctions de contraintes au point x^*, alors il existe une fonction différentiable

$$\alpha: [0,1] \to D_R \text{ telle que } \alpha(0) = x^* \text{ et } \alpha'(0) = \sigma \tilde{d}, \sigma > 0$$

Mais ainsi,

$$\lim_{\theta \to 0} \frac{f\big(\alpha(\theta)\big) - f(x^*)}{\theta} = \nabla f(x^*)^T \alpha'(0) = \sigma \nabla f(x^*)^T \tilde{d} < 0$$

Ce qui implique l'existence d'un $\hat{\theta} \in [0,1]$ assez petit pour que $\alpha\big(\hat{\theta}\big) \in B_\varepsilon(x^*)$ tel que $f\left(\alpha\big(\hat{\theta}\big)\right) < f(x^*)$ **une contradiction puisque $\boldsymbol{\alpha\big(\hat{\theta}\big) \in D_R}$.**

Le reste de la preuve se fait comme précédemment lorsque nous supposions que l'hypothèse était vérifiée.

Conditions d'optimalité : Condition K-K-T **nécessaires**

Théorème 8 : (Nécessité des conditions K-K-T) : Soit $x^* \in \bar{X}$ une solution optimale locale du problème $(\boldsymbol{P_2})$ où $\bar{X}$ est ouvert. Supposons de plus que si $\nabla f_i(x^*), i \in A(x^*)$ sont linéairement indépendante
Alors il existe un vecteur de multiplicateurs $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$ tel que :

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m$$

Problème de programmation convexe

Considérons le problème de programmation convexe

$$Primal\ \ (P)\begin{cases} Min & f_0(x) \\ s.t & f_i(x) \leq 0 \quad i \in I_1 \\ & h_i(x) = a_i^T x + b_i = 0 \quad i \in I_2 \\ & x \in \mathbb{R}^n \end{cases}$$

où $f_0(x) \in \mathcal{C}^2, \forall i \in I_1\ f_i(x) \in \mathcal{C}^2$et convexe dans $\mathbb{R}^d$,

On part de l'estimation de sa valeur optimale

$$p^* = inf\{f_0(x) | x \in D_F\} = \inf_{x \in D_F} f_0(x)$$

où $D_F = \{x \in \mathbb{R}^d \middle| f_i(x) \leq 0 \quad i \in I_1, h_i(x) = 0 \ \ i \in I_2\}$

Introduisons la fonction lagrangienne

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i \in I_1} \lambda_i f_i(x) + \sum_{i \in I_2} \mu_i h_i(x)$$

où $\lambda = (\lambda_1, \dots, \lambda_m)^T$et $\mu = (\mu_1, \dots, \mu_p)^T$ sont des multiplers lagrangiens.

Problème de programmation convexe: : Conditions d'optimalité

Définition(Conditions de Karush-Kuhn-Tuker (KKT))

Considérons le problème principal de programmation convexe (P). x^* est dit satisfaire les conditions KKT s'il existe les multiplicateurs $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T$ et $\mu^* = \left(\mu_1^*, \dots, \mu_p^*\right)^T$ correspondant respectivement aux contraintes du problème primal (P) , telles que la fonction lagrangienne (P)

$$L(x,\lambda,\mu) = f_0(x) + \sum_{i\in I_1} \lambda_i f_i(x) + \sum_{i\in I_2} \mu_i h_i(x) \Longrightarrow \nabla_x L(x,\lambda,\mu) = \nabla f_0(x) + \sum_{i\in I_1} \lambda_i \nabla f_i(x) + \sum_{i\in I_2} \mu_i \nabla h_i(x)$$

Satisfait

KKT Conditions:
$$\begin{cases} \nabla_x L(x^*,\lambda^*,\mu^*) = \nabla f_0(x^*) + \sum_{i\in I_1} \lambda_i^* \nabla f_i(x^*) + \sum_{i\in I_2} \mu_i^* \nabla h_i(x^*) = 0 \\ f_i(x^*) \le 0 \quad i \in I_1 : |I_1| = m \\ h_i(x^*) = a_i^T x^* + b_i = 0 \;\; i \in I_2 : \; |I_2| = p \\ \lambda_i^* f_i(x^*) = 0 \qquad i \in I_1 \\ \lambda_i^* \ge 0 \qquad i \in I_1 \end{cases}$$

Problème de programmation convexe: Conditions d'optimalité

Définition (**Condition de Slater**)
On dit que le problème primal de programmation convexe (P) satisfait la condition de Slater s'il existe une solution réalisable x telle que:

$$\begin{cases} f_i(x) < 0 \quad i \in I_1 \\ h_i(x) = a_i^T x + b_i = 0 \quad i \in I_2 \end{cases}$$

Ou lorsque les premières contraintes d'inégalité de k sont des contraintes linéaires, il existe une solution réalisable x telle que :

$$\begin{cases} f_i(x) = a_i^T x + b_i \leq 0 \quad i \in I_1^k \\ f_i(x) < 0 \quad i \in I_1^{m-k} \\ h_i(x) = a_i^T x + b_i = 0 \quad i \in I_2 \end{cases}$$

Théorème:
Considérons le problème primal de programmation convexe (P) satisfaisant la condition de Slater. Si x^* est une solution alors x^* satisfait aux conditions KKT.

Théorème:
Considérons le problème primal de programmation convexe (P) satisfaisant la condition de Slater. Alors pour sa solution x^*, c'est une condition néssecaire et suffissante que x^*satisfait aux conditions KKT