Optimisation Non linèaire

Par

Professeur Abdellatif El Afia

Chapitre 4

Partie 2: Optimisation avec contraintes

1. **Analyse Convexe**
    1. **Ensembles convexes**
    2. **fonctions convexes**
    3. **Hyperplan**
    4. **Théories de séparation**
    5. **Problème de programmation convexe**
2. **Conditions d’optimalité**
    1. **Multiplicateur de Lagrange**
    2. **Conditions de Karush-Kuhn-Tucker**
3. **Algorithmes primales**

Optimisation avec contraintes
Analyse Convexe
Chapitre 4

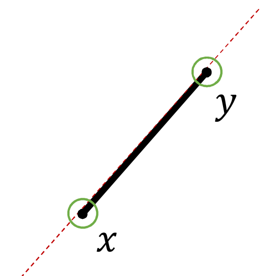
1. **Ensembles convexes**
2. **fonctions convexes**
3. **Hyperplan**
4. **Théories de séparation**

Ensemble convexes

Définition : Etant donné deux points $x, y \in \mathbb{R}^n$,

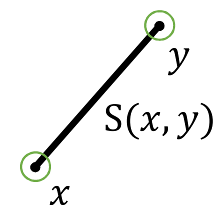
- la **droite** passante par x et y est définie par l'ensemble

$$\mathrm{D}(x, y) = \{z \in \mathbb{R}^n : z = \theta x + (1 - \theta) y, \theta \in \mathbb{R}^n\}$$

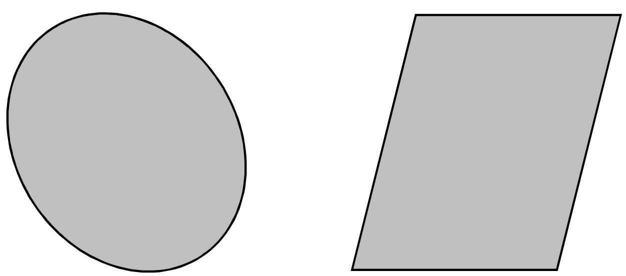


- le **segment de droite** reliant x et y est définie par l'ensemble:

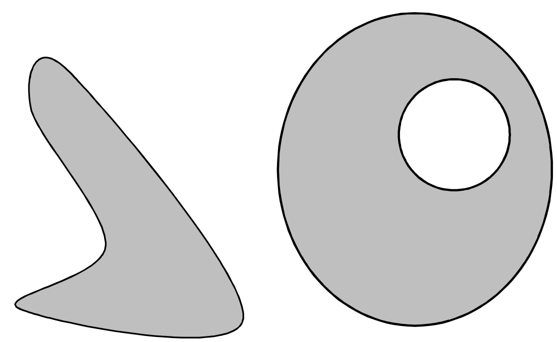
$$S(x, y) = \{z \in \mathbb{R}^n : z = \theta x + (1 - \theta) y, \theta \in [0,1]\}$$



- Un ensemble $X \subset \mathbb{R}^n$ est **convexe** si $\forall\, x, y \in X$ alors $\mathrm{S}(x, y) \subset X$.



Ensemble convexes



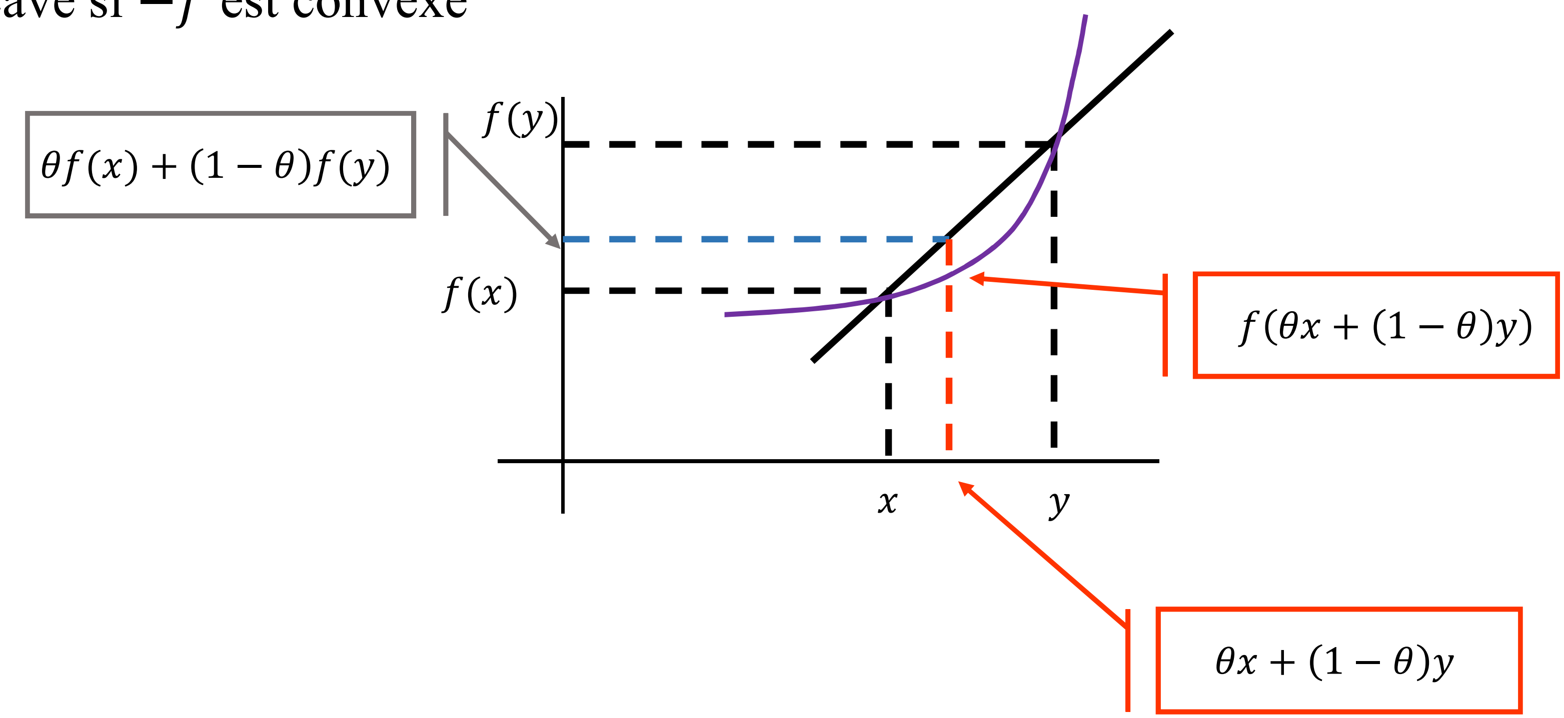
Ensemble non convexes

Fonctions convexes

Définition : Etant donné un ensemble convexe $X \subset \mathbb{R}^n$,une fonction à valeur Réelles $f: X \rightarrow \mathbb{R}$ est **convexe** si $\forall\ (x, y) \in X \times X$

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y), \qquad \forall \theta \in [0,1]$$

f est concave si $-f$ est convexe



Fonctions convexes

Définition :
Etant donné un ensemble convexe $X \subset \mathbb{R}^n$, une fonction à valeur réelles $f: X \to \mathbb{R}$ est:

strictement convexe si $\forall\ (x, y) \in X \times X$

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y), \qquad \forall \theta \in (0,1)$$

- **concave** si $\forall\ (x, y) \in X \times X$

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta)f(y), \qquad \forall \theta \in [0,1]$$

- **strictement concave** si $\forall\ (x, y) \in X \times X$

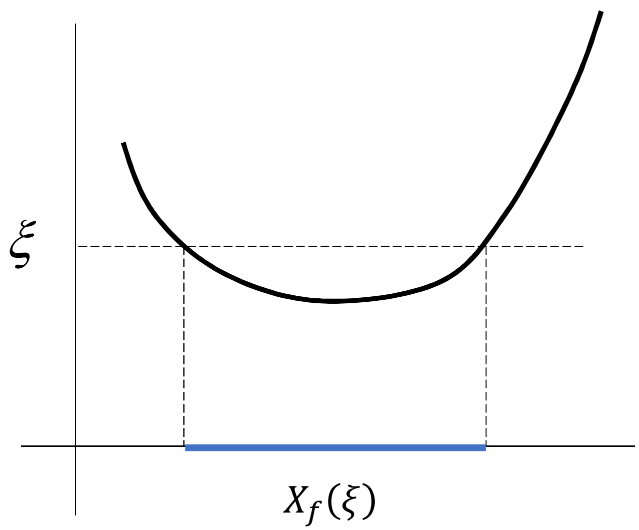
$$f(\theta x + (1-\theta)y) > \theta f(x) + (1-\theta)f(y), \qquad \forall \theta \in (0,1)$$

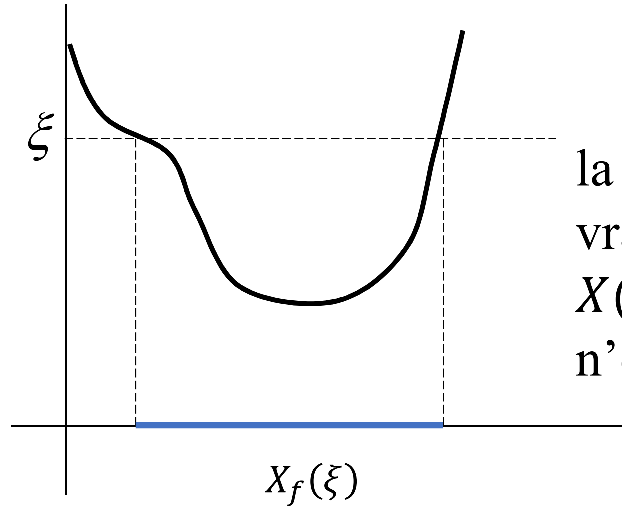
Fonctions convexes

Proposition 1 : Soit $X \subset \mathbb{R}^n$ un ensemble convexe, si $f: X \to \mathbb{R}$ est convexe, alors l'ensemble $X_f(\xi) = \{x \in X: f(x) \leq \xi\}$ est convexe pour $\forall \xi \in \mathbb{R}$

Preuve : Soient x et $y \in X_f(\xi)$ et $\theta \in (0,1)$. Alors, puisque f est convexe,

- $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta\xi + (1-\theta)\xi = \xi$
- $\Longrightarrow (\theta x + (1-\theta)y) \in X_f(\xi)$. Donc $X_f(\xi)$ est convexe





la réciproque n'est pas vraie
$X(\xi)$ convexe mais f n'est pas convexe

Fonctions convexes

Proposition 2 : soit $X \subset \mathbb{R}^n$, si $f_i : X \to \mathbb{R}$ est convexe pour $i = 1, \dots, m$ et si $\lambda_1, \dots, \lambda_m$ sont des scalaires non négatifs, alors $f(x) = \sum_{i=1}^{m} \lambda_i f_i(x)$ est convexe sur X.

Preuve :

Pour toute paire de points $x, y \in X$ et pour tout $\theta \in [0,1]$

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &= \sum_{i=1}^{m} \lambda_i f_i(\theta x + (1-\theta)y) \\
&\leq \sum_{i=1}^{m} \lambda_i [\theta f_i(x) + (1-\theta)y] \text{ puisque } f_i \text{ convexe et } \lambda_i \geq 0 \\
&= \theta \sum_{i=1}^{m} \lambda_i f_i(x) + (1-\theta) \sum_{i=1}^{m} \lambda_i f_i(y) \\
&= \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

Fonctions convexes

Proposition 3 : Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Si $f: X \to \mathbb{R}$ est convexe, Si $x^1, \dots, x^m \in X$, et si $\lambda_1, \dots, \lambda_m$ sont des scalaires non négatifs tels que $\sum_{i=1}^{m} \lambda_i = 1$, alors: $f\left(\sum_{i=1}^{m} \lambda_i x^i\right) \leq \sum_{i=1}^{m} \lambda_i f(x^i)$

Preuve :

Si $m = 2$ nous retrouvons la définition de fonction convexe.

Procédons par induction et supposons que c'est vrai pour $(m-1)$; si $\theta_1, \dots, \theta_{m-i} \geq 0$ tels que $\sum_{i=1}^{m-1} \theta_i = 1$,

alors

- $\sum_{i=1}^{m-1} \theta_i x^i \in X \Longrightarrow f\left(\sum_{(i=1)}^{m-1} \theta_i x^i\right) \leq \sum_{i=1}^{m-1} \theta_i f\left(x^i\right)$
- Posant $\theta_i = \frac{\lambda_i}{\sum_{k=1}^{m-1} \lambda_k}$ on a $\sum_{i=1}^{m-1} \frac{\lambda_i}{\sum_{k=1}^{m-1} \lambda_k} = 1$, $\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} \in X \Longrightarrow f\left(\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k}\right) \leq \sum_{i=1}^{m-1} \frac{\lambda_i f(x^i)}{\sum_{k=1}^{m-1} \lambda_k}$

Par définition de convexité,

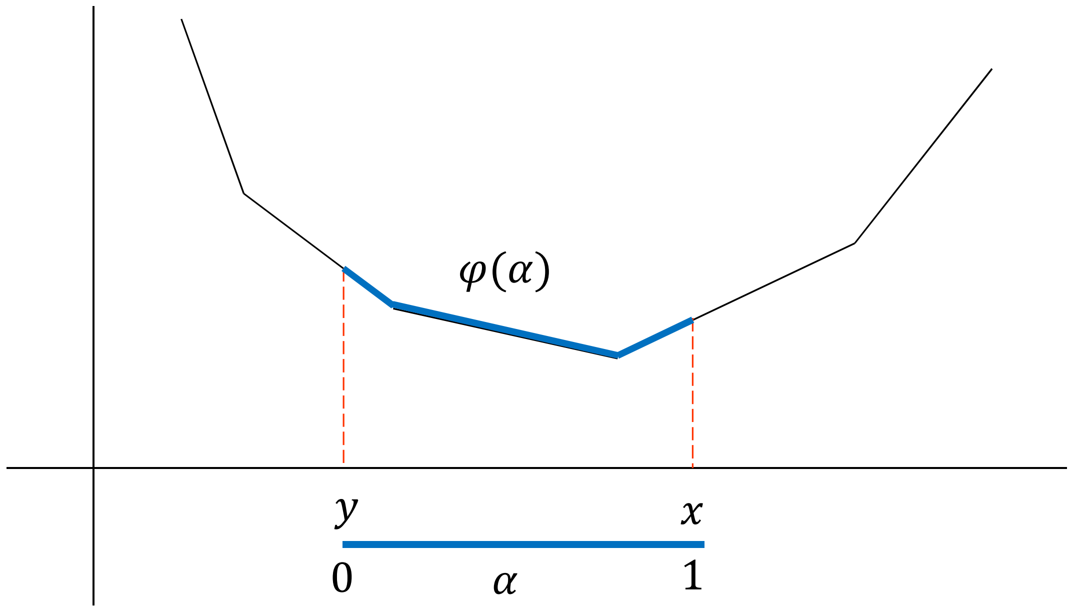
- $\sum_{k=1}^{m-1} \lambda_k \sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} + \lambda_m x^m \in X$
- $\Longrightarrow f\left(\sum_{k=1}^{m-1} \lambda_k \sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} + \lambda_m x^m\right) \leq \sum_{k=1}^{m-1} \lambda_k f\left(\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k}\right) + \lambda_m f(x^m)$
- $\Longrightarrow f\left(\sum_{(i=1)}^{m-1} \lambda_i x^i + \lambda_m x^m\right) \leq \sum_{k=1}^{m-1} \lambda_k \sum_{i=1}^{m-1} \frac{\lambda_i f(x^i)}{\sum_{k=1}^{m-1} \lambda_k} + \lambda_m f(x^m)$

Fonctions convexes

Proposition 4 :

soit $X \subset \mathbb{R}^n$ un ensemble convexe. $f: X \to \mathbb{R}$ est convexe si et seulement si $\varphi(\alpha) = f(y + \alpha(x-y))$ est une fonction convexe sur $[0,1]$ pour toute paire de points $x, y \in X$.

Note : $\varphi(\alpha)$ dénote la restriction de f sur le segment de droite entre x et y.



Fonctions convexes

Preuve :

Pour démontrer la nécessité, considérons la paire de points $x, y \in X$.

Soit $\alpha_1, \alpha_2 \in [0,1]$, alors :

$$\begin{aligned}
\varphi(\theta\alpha_1 + (1-\theta)\alpha_2) &= \varphi\big(\alpha_2 + \theta(\alpha_1 - \alpha_2)\big) \\
&= f\Big[\big(\alpha_2 + \theta(\alpha_1 - \alpha_2)\big)x + \big(1 - \alpha_2 - \theta(\alpha_1 - \alpha_2)\big)y\Big] \\
&= f\Big[\big(\alpha_2 + \theta(\alpha_1 - \alpha_2)\big)x + \big(1 + \theta - \theta - \alpha_2 - \theta(\alpha_1 - \alpha_2)\big)y\Big] \\
&= f[\theta(\alpha_1 x + (1-\alpha_1)y) + ((1-\theta)(\alpha_2 x + (1-\alpha_2)y)] \\
&\leq \theta f(\alpha_1 x + (1-\alpha_1)y) + (1-\theta) f(\alpha_2 x + (1-\alpha_2)y) \\
&\leq \theta\varphi(\alpha_1) + (1-\theta)\varphi(\alpha_2)
\end{aligned}$$

Donc φ est convexe sur $[0,1]$

Pour démontrer la suffisance, soient $x, y \in X$.

- $\varphi(\alpha) = f(\alpha x + (1-\alpha)y)$ est convexe sur $[0,1]$ alors pour $x = 1, y = 0$ on a
- $\varphi(\alpha) = \varphi(\alpha.1 + (1-\alpha)0) \leq \alpha\varphi(1) + (1-\alpha)\alpha(0)\ \forall \alpha \in [0,1]$
- $\Longrightarrow f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \quad \forall \alpha \in [0,1]$

Donc f est convexe.

Fonctions convexes

Proposition 5 :

Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f: X \to \mathbb{R}$ une fonction convexe. Si $\bar{x}$ est un minimum local de f sur X, alors $\bar{x}$ est un minimum global de f sur X.

Preuve :

la preuve se fait par contradiction:

Supposant que $\exists \tilde{x} \in X$ tel que $f(\tilde{x}) < f(\bar{x})$. Puisque f est convexe, pour tout $\theta \in (0,1]$ on a:

$$f(\theta\tilde{x} + (1-\theta)\bar{x}) \leq \theta f(\tilde{x}) + (1-\theta)f(\bar{x}) < \theta f(\bar{x}) + (1-\theta)f(\bar{x}) = f(\bar{x})$$

Or pour $\theta > 0$ suffisamment petit, $x(\theta) = \theta\tilde{x} + (1-\theta)\bar{x} \in V_\varepsilon(\bar{x}) \cap X$

Ainsi $f\big(x(\theta)\big) < f(\bar{x})$ où $x(\theta) \in V_\varepsilon(\bar{x}) \cap X$,

contredisant que $\bar{x}$ est un minimum local de f sur X.

Fonctions convexes

Proposition 6 : (**Inégalité du Gradient**). Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f: X \to \mathbb{R}$ une fonction de classe C/X. Alors f est convexe sur X si et seulement si

$$f(x) \geq f(y) + \nabla f(y)^T(x-y) \;\; \forall x, y \in X$$

Preuve (nécessité)

Soit f convexe. Alors $\forall x, y \in X$ et $\forall \theta \in [0,1] f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$

ce qui s'écrit aussi sous la forme

$$f\big(y + \theta(x-y)\big) - f(y) \leq \theta\big(f(x) - f(y)\big) \quad \textbf{(4.1)}$$

Se référant au théorème de Taylor, il existe un $\tau \in [0,1]$ tel que

- $f\big(y + \theta(x-y)\big) = f(y) + \nabla f\big[\tau\big(y + \theta(x-y)\big) + (1-\tau)y\big]^T (y + \theta(x-y) - y)$
- $\Longrightarrow f\big(y + \theta(x-y)\big) - f(y) = \theta \nabla f[y + \tau\theta(x-y)]^T(x-y) \quad \textbf{(4.2)}$

Combinant les relations **(4.1)** et **(4.2)**, pour tout $\theta \in [0,1]$

$$\theta \nabla f[y + \tau\theta(x-y)]^T(x-y) \leq \theta\big(f(x) - f(y)\big)$$

Si $\theta > 0$, nous pouvons diviser par θ de chaque côté de l'inéquation et obtenir

$$\nabla f[y + \tau\theta(x-y)]^T(x-y) \leq \big(f(x) - f(y)\big) \;\; \forall \theta \in (0,1]$$

Par conséquent, $\lim_{\theta \to 0} \nabla f[y + \tau\theta(x-y)]^T(x-y) \leq \lim_{\theta \to 0}\big(f(x) - f(y)\big)$

Ou encore $\nabla f\left[y + \lim_{\theta \to 0} \tau\theta(x-y)\right]^T (x-y) \leq \big(f(x) - f(y)\big) \Longrightarrow f(x) \geq f(y) + \nabla f(y)^T(x-y)$

Fonctions convexes

Preuve suffisance

Démontrons maintenant que la condition est suffisante.

Puisque $\forall\, x, y \in X \;\; f(x) \geq f(y) + \nabla f(y)^T(x-y)$

$$f(x) \geq f(\theta x + (1-\theta)y) + \nabla f(\theta x + (1-\theta)y)^T(x - \theta x - (1-\theta)y) \quad \mathbf{(4.3)}$$

$$f(y) \geq f(\theta x + (1-\theta)y) + \nabla f(\theta x + (1-\theta)y)^T(y - \theta x - (1-\theta)y) \quad \mathbf{(4.4)}$$

$\mathbf{(4.3)}$ et $\mathbf{(4.4)}$ s'écrivent respectivement

$$f(x) \geq f(\theta x + (1-\theta)y) + (1-\theta)\nabla f(\theta x + (1-\theta)y)^T(x-y) \quad \mathbf{(4.3)}$$

$$f(y) \geq f(\theta x + (1-\theta)y) - \theta\nabla f(\theta x + (1-\theta)y)^T(x-y) \quad \mathbf{(4.4)}$$

Multiplions $\mathbf{(4.3)}$ par θ et $\mathbf{(4.4)}$ par $(1-\theta)$

$$\theta f(x) \geq \theta f(\theta x + (1-\theta)y) + \theta(1-\theta)\nabla f(\theta x + (1-\theta)y)^T(x-y)$$

$$(1-\theta)f(y) \geq (1-\theta)f(\theta x + (1-\theta)y) - (1-\theta)\theta\nabla f(\theta x + (1-\theta)y)^T(x-y)$$

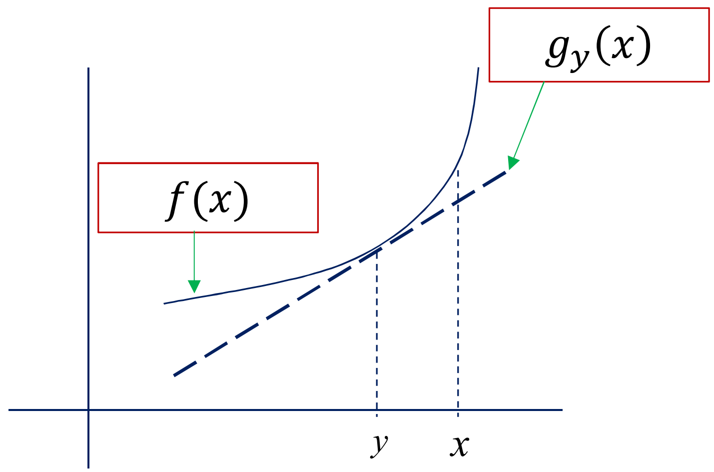
Et additionnons les deux relations résultantes :

$$\theta f(x) + (1-\theta)f(y) \geq f(\theta x + (1-\theta)y)$$

Fonctions convexes

La figure suivant illustre l'inégalité du gradient pour une fonction $f: \mathbb{R} \to \mathbb{R}$
Dans ce cas l'inégalité du gradient devient

$$f(x) \geq g_y(x) = f(y) + f'(y)(x - y)$$



La droite $g_y(x) = f(y) + f'(y)(x - y)$ a les propriétés suivantes :

- $g_y(y) = f(y),\ g_y'(y) = f'(y).$
- Cette droite est une fonction support de f au point y.
- L'inégalité du gradient indique que la fonction prend toujours une valeur plus grande ou égale à celle de la fonction support en tout point y.

Fonctions convexes

Corollaire 7 : soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f: X \to \mathbb{R}$ une fonction convexe de classe C/X. Alors si $\nabla f(x^*)^T(x - x^*) \geq 0 \ \forall x \in X$, alors x^* est un minimum global de f sur X.

Preuve :

le résultat découle directement de l'inégalité du gradient. En effet, puisque f est convexe sur ,

alors $f(x) \geq f(x^*) + \nabla f(x^*)^T(x - x^*), \forall x \in X$

Puisque par hypothèse $\nabla f(x^*)^T(x - x^*) \geq 0$, alors $f(x) - f(x^*) \geq \nabla f(x^*)^T(x - x^*) \geq 0$

D'où $f(x) \geq f(x^*)$ pour tout $x \in X$. Donc x^* est un minimum global de f sur X.

Conséquence : Si $\nabla f(x^*) = 0$, alors x^* est un minimum global de f sur X.

Pour $\forall f \in C^2/X$, il existe un critère pour vérifier la convexité qui est basé sur le Hessien $\nabla^2 f$.

Proposition 8 :

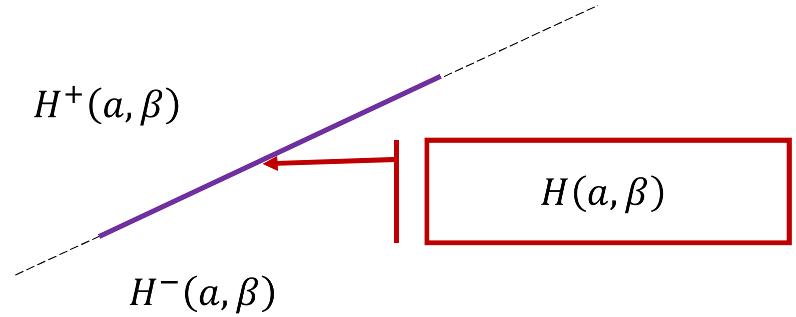
Soient $X \subset \mathbb{R}^n$ un ensemble convexe dont l'intérieur est non vide, $f: X \to \mathbb{R}, f \in C^2/X$. Alors f est convexe sur X si et seulement si son Hessien $\nabla^2 f(x)$ est une matrice semi défini positive pour tout $x \in X$.

Hyperplan

Définition :

- **L'hyperplan** spécifié par le point $a \in \mathbb{R}^n$ et le scalaire β est l'ensemble de $\mathbb{R}^n$ **:** $H(a,\beta) = \{x \in \mathbb{R}^n : a^T x = \beta\}$
- Les **demis espaces** (fermés) associés a l'hyperplan $H(a,\beta)$ sont les ensemble suivants de $\mathbb{R}^n$:

$$H^+[a,\beta] = \{x \in \mathbb{R}^n : a^T x \geq \beta\}$$
$$H^-[a,\beta] = \{x \in \mathbb{R}^n : a^T x \leq \beta\}$$



Note : il est facile de vérifier que tous ces ensembles

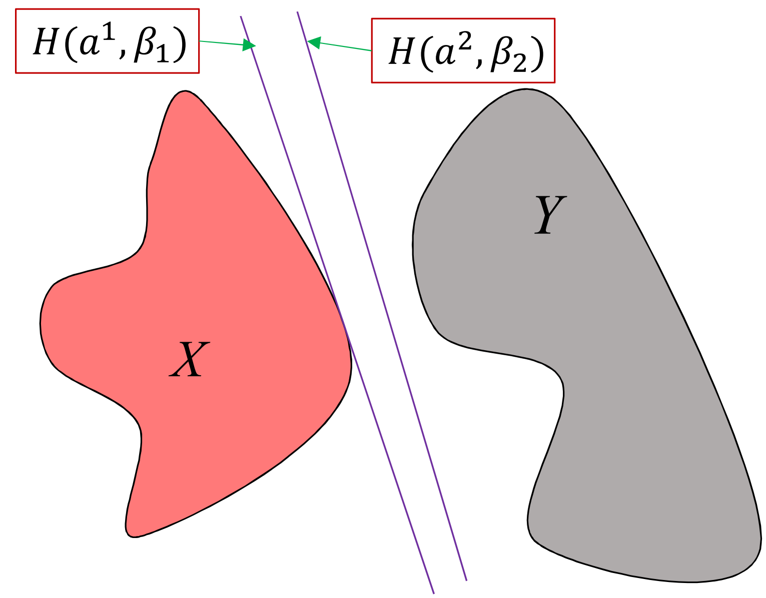
Définition Hyperplan de séparation

L'hyperplan $H(a,\beta)$ sépare **(Stricte)** deux ensembles non vides X et Y si

$a^T x \geq \beta$ $(\boldsymbol{a^T x > \beta})$ pour tout $x \in X$ (i.e., $X \subset H^+[a,\beta]$)

$a^T y \leq \beta$ $(\boldsymbol{a^T x > \beta})$ pour tout $y \in Y$ (i.e., $Y \subset H^-[a,\beta]$)

- $H(a^1,\beta_1)$ hyperplan de séparation
- $H(a^2,\beta_2)$ hyperplan de séparation stricte

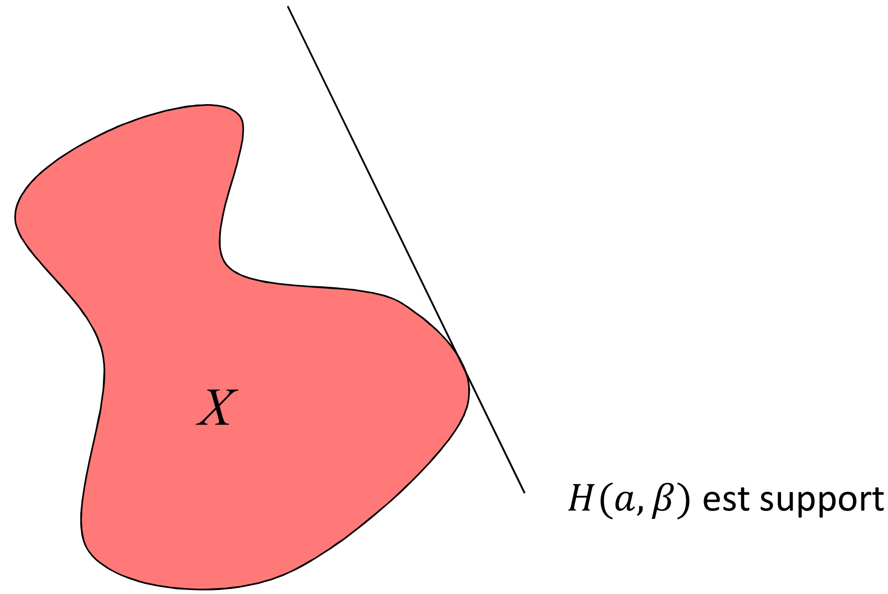


Théorie de séparation

Définition
étant donné un ensemble non vide $X \subset \mathbb{R}^n$
l'hyperplan $H(a, \beta)$ est un **support** de X si :

- $H(a, \beta) \cap \bar{X} \neq \emptyset$ et
- $X \subset H^+[a, \beta]$ ou $X \subset H^-[a, \beta]$

(où $\bar{X}$ dénote la **fermeture** de l'ensemble X)



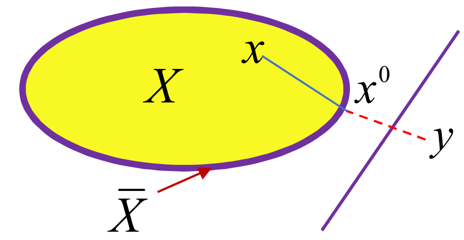
Théorème 1 : Soient les vecteurs $x, y, a \in \mathbb{R}^n$.
Si $\boldsymbol{a^T y < a^T x}$, alors pour tout $\theta \in (0,1)$ on a :

$$a^T y < a^T(\theta x + (1-\theta)y) < a^T x$$

Preuve :

$$a^T(\theta x + (1-\theta)y) < a^T(\theta x + (1-\theta)x) = a^T x$$
$$a^T(\theta x + (1-\theta)y) > a^T(\theta y + (1-\theta)y) = a^T y$$

Théorie de séparation



Théorème 2 : (Théorème de séparation)
Si $X \subset \mathbb{R}^n$ est un ensemble convexe non vide et si $y \notin \bar{X}$,
alors il existe un hyperplan qui sépare (strictement) $\bar{X}$ et y.
Preuve :
Il existe un point $x^0 \in \bar{X}$ tel que : $\boldsymbol{\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|}$
(où Il est facile de vérifier que $\bar{X}$ est un ensemble convexe, et par conséquent le segment de droite $S(x^0, x) \subset \bar{X}\ \forall x \in \bar{X}$. Donc $\forall \theta \in [0,1]$, $\|x^0 - y\| \le \|(\theta x + (1-\theta)x^0) - y\|$
Ainsi $\forall \theta \in [0,1]$

$$(x^0 - y)^T(x^0 - y) \le \left((\theta x + (1-\theta)x^0) - y\right)^T\left((\theta x + (1-\theta)x^0) - y\right)$$

$$(x^0 - y)^T(x^0 - y) \le \left((x^0 - y) + \theta(x - x^0)\right)^T\left((x^0 - y) + \theta(x - x^0)\right)$$

$$(x^0 - y)^T(x^0 - y) \le (x^0 - y)^T(x^0 - y) + 2\theta(x - x^0)^T(x^0 - y) + \theta^2(x - x^0)^T(x - x^0)$$

Par conséquent $\forall \theta \in [0,1]$, $2\theta(x - x^0)^T(x^0 - y) + \theta^2(x - x^0)^T(x - x^0) \ge 0$
Mais alors ceci implique que $(x - x^0)^T(x^0 - y) \ge 0$
$(x - x^0)^T(x^0 - y) < 0$, alors pour $\tilde{\theta} > 0$ suffisamment petit

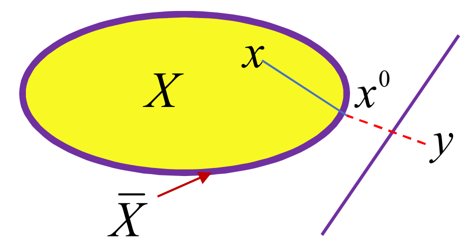
$$2\tilde{\theta}|(x - x^0)^T(x^0 - y)| > \tilde{\theta}^2|(x - x^0)^T(x - x^0)|$$

$$-2\tilde{\theta}(x - x^0)^T(x^0 - y) > \tilde{\theta}^2(x - x^0)^T(x - x^0)$$

Et alors nous aurions que $2\tilde{\theta}(x - x^0)^T(x^0 - y) + \tilde{\theta}^2(x - x^0)^T(x - x^0) < 0$

Théorie de séparation

Preuve (suite) :



Nous avons donc que $(x - x^0)^T(x^0 - y) \geq 0$

$$\Longrightarrow {x^0}^T(x^0 - y) \leq x^T(x^0 - y) \quad (5.3)$$

Puisque $y \notin \bar{X}$,

$$\|x^0 - y\|^2 = (x^0 - y)^T(x^0 - y) > 0 \Longrightarrow y^T(x^0 - y) < {x^0}^T(x^0 - y) \quad (5.4)$$

D'après le Théorème 1 avec les vecteurs $(x^0 - y), x^0, y$ et utilisant la relation 5.4

$$y^T(x^0 - y) < (\theta x^0 + (1-\theta)y)^T(x^0 - y) < {x^0}^T(x^0 - y)$$

Et avec $\theta = \frac{1}{2} \Longrightarrow\ y^T(x^0 - y) < \frac{1}{2}(x^0 + y)^T(x^0 - y) < {x^0}^T(x^0 - y)$

En utilisant maintenant la relation 5.3

$$y^T(x^0 - y) < \frac{1}{2}(x^0 + y)^T(x^0 - y) < {x^0}^T(x^0 - y) \leq x^T(x^0 - y)$$

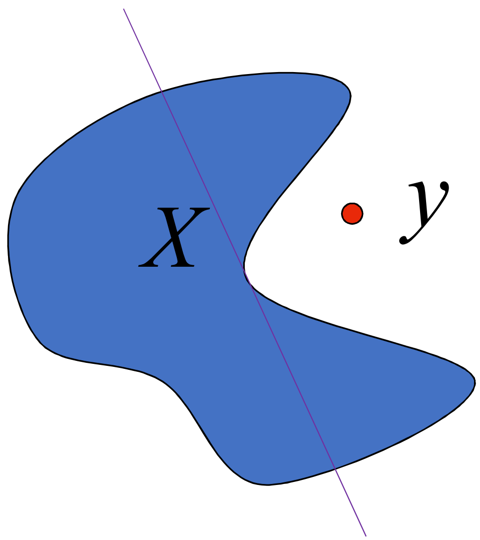
Par conséquent nous avons que

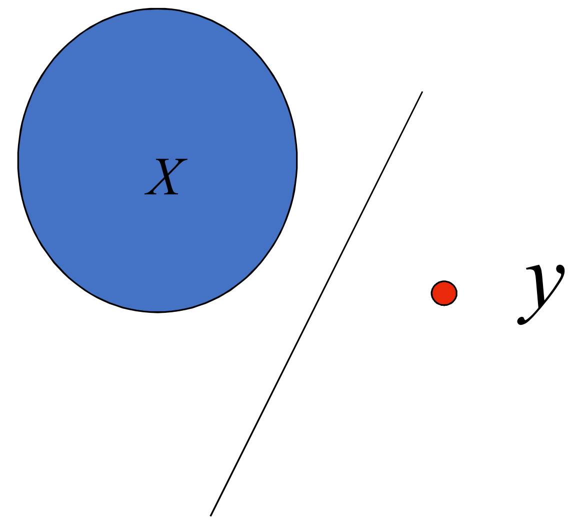
$$y^T(x^0 - y) < \frac{1}{2}(x^0 + y)^T(x^0 - y) < x^T(x^0 - y)$$

Puisque cette relation est valable pour tout $x \in \bar{X}$, il s'ensuit que l'hyperplan $H(a, \beta)$ où $a = (x^0 - y)$ et $\beta = \frac{1}{2}(x^0 + y)^T(x^0 - y)$ sépare (strictement) $\bar{X}$ et y.

Théorie de séparation

Dans le théorème 2, la convexité de X est une condition suffisante pour assurer l'existence d'un hyperplan le séparant (strictement) de $y \notin \bar{X}$





Théorie de séparation: Théorème de Farkas

Théorème 3 : Étant donné:

- la matrice $A = \begin{pmatrix} \vdots & \vdots & \vdots \\ a^1 & \cdots & a^n \\ \vdots & \vdots & \vdots \end{pmatrix}$: $\forall i \in \{1, \dots, n\}\ a^i \in \mathbb{R}^m$, $b \in \mathbb{R}^m$
- H_1: $\boldsymbol{\forall y \in \mathbb{R}^m}$ **tel que** $\boldsymbol{y^T a^i \geq 0}$ **pour** $\boldsymbol{\forall i \in \{1, \dots, n\} \Longrightarrow y^T b \geq 0}$, $\Longrightarrow \neg H_1 = \left\{\begin{matrix} \exists y \in \mathbb{R}^m, \text{tel que} \\ y^T a^i \geq 0\ \forall i = 1, \dots, n, \\ \text{et } y^T b < 0 \end{matrix}\right\}$
- H_2: $\boldsymbol{\exists x = (x_1, \dots, x_n) \geq 0}$ **tels que** $\boldsymbol{b = a^1 x_1 + \cdots + a^n x_n = Ax}$, $\Longrightarrow \neg H_2 = \left\{\begin{matrix} \nexists x_1, \dots, x_n \geq 0 \text{ tels que} \\ b = a^1 x_1 + \cdots + a^n x_n \end{matrix}\right\}$

une condition suffisante pour que H_2 soit vérifier est que H_1 soit vérifier

Preuve :

il faut démontrer que H_2: $\left\{\begin{matrix} \forall y \in \mathbb{R}^m, \\ \text{si } \forall j \in \{1, \dots, n\}:\ y^T a^j \geq 0 \\ \text{alors nécessairement } y^T b \geq 0 \end{matrix}\right\} \Rightarrow H_1$: $\left\{\begin{matrix} \exists x_1, \dots, x_n \geq 0 \text{ tels que} \\ b = a^1 x_1 + \cdots + a^n x_n \end{matrix}\right\}$

Nous allons plutôt démontrer la contraposé de l'implication :

$$\left\{\begin{matrix} \nexists x_1, \dots, x_n \geq 0 \text{ tels que} \\ b = a^1 x_1 + \cdots + a^n x_n \end{matrix}\right\} \Rightarrow \left\{\begin{matrix} \exists y \in \mathbb{R}^m, \text{tel que} \\ y^T a^i \geq 0\ \forall i = 1, \dots, n, \\ \text{et } y^T b < 0 \end{matrix}\right\}$$

Théorie de séparation: Théorème de Farkas

Considérons l'ensemble $Z = \{z \in \mathbb{R}^m : \exists x_1, \dots, x_n \geq 0 \text{ tel que } z = a^1 x_1 + \cdots + a^n x_n\}$
Il est facile de démontrer que Z est convexe. Il est aussi possible de démontrer que Z est un ensemble fermé (i.e. $Z = \bar{Z}$), soit $z^1, z^2 \in Z$

- $\Longrightarrow z^1 = a^1 x_1^1 + \cdots + a^n x_n^1,\ z^2 = a^1 x_1^2 + \cdots + a^n x_n^2$
- $\theta z^1 + (1-\theta) z^2 = a^1[\theta x_1^1 + (1-\theta) x_1^2] + \cdots + a^n[\theta x_n^1 + (1-\theta) x_n^2] \Longrightarrow \theta z^1 + (1-\theta) z^2 \in Z$

Puisque par hypothèse de la contraposé $b \notin Z = \bar{Z}$, alors par le **théorème 2**. $\exists H(p, \beta)$ qui sépare strictement Z et b : $p^T b < \beta < p^T z, \forall z \in Z$ $\mathbf{(5.5)}$

Or $0 \in Z$, ce qui implique que $\beta < 0$ et par conséquent que $p^T b < 0$ Egalement, $p^T z \geq 0\ \forall z \in Z$, en effet $\exists \tilde{z} \in Z$ tel que $p^T \tilde{z} < 0$, alors puisque Z est un cône (i.e., si $z \in Z$ alors $\forall \lambda \geq 0\ \lambda z \in Z$), nous aurions que $p^T(\lambda \tilde{z}) \xrightarrow{\lambda \to \infty} -\infty$ contredisant$\mathbf{(5.5)}$.

Comme il est facile de vérifier que $a^j \in Z, j = 1, \dots, n$

$$Z = \{z \in \mathbb{R}^m : \exists x_1, \dots, x_n \geq 0 \text{ tel que } z = a^1 x_1 + \cdots + a^n x_n\}$$

Il s'ensuit que $p^T a^j \geq 0, j = 1, \dots, n$
Nous avons donc démontré la contraposé puisque $p \in \mathbb{R}^m$ Tel que $p^T a^j \geq 0, j = 1, \dots, n$ et $p^T b < 0$

Théorie de séparation: Théorème d'alternatives

Théorème 4: Soit A une matrice $m \times n$, exactement une des deux alternatives suivantes est vérifiée :

I. Le système $Ax = b, x \geq 0$ possède une solution $x \in \mathbb{R}^n$

II. Le système $A^T y \geq 0, b^T y < 0$ possède une solution $y \in \mathbb{R}^m$

Preuve :

il est facile de vérifier que les deux alternatives ne peuvent tenir en même temps, Car autrement la relation suivante serait satisfaite : $0 > b^T y = x^T A^T y \geq 0$

Une contradiction.

Dénotons par $a_{\bullet j}, j = 1, \ldots, n,$ la j$^{\text{ième}}$colonne de A

Considérant **l'alternative** (**II**), celle-ci est vérifiée ou elle ne l'est pas.

Dans le cas où elle ne l'est pas, il s'ensuit que le système $A^T y \geq 0, b^T y < 0$ ne possède pas de solution $y \in \mathbb{R}^m$

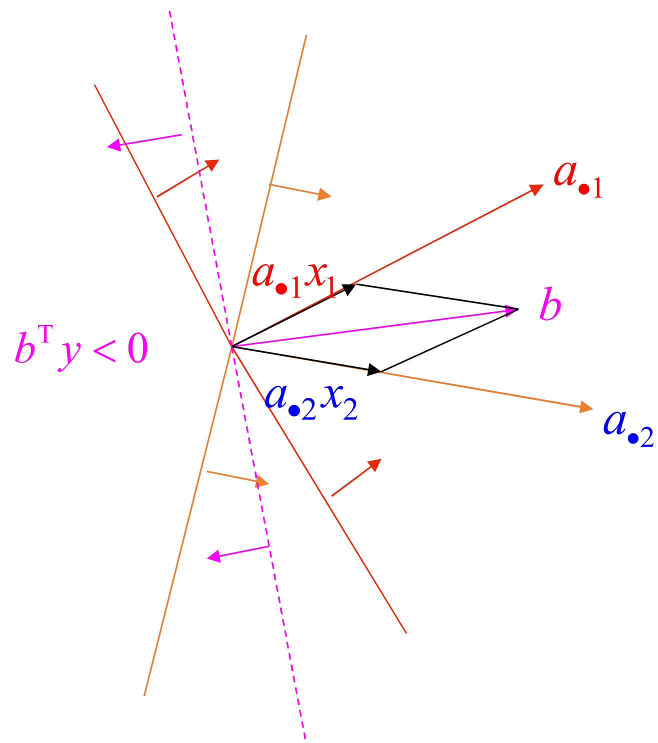
i.e., le système $a_{\bullet j}^T y \geq 0, j = 1, \ldots, n, b^T y < 0$ ne possède pas de solution $y \in \mathbb{R}^m$

si $a_{\bullet j}^T y \geq 0, j = 1, \ldots, n,$ alors nécessairement $b^T y \geq 0.$

Donc par le Théorème 11, il existe un vecteur $x \in \mathbb{R}^n, x \geq 0$ tel que $Ax = a_{\bullet 1} x_1 + \cdots + a_{\bullet \mathrm{n}} x_n = b$, et **l'alternative** (**I**) tient.

Théorie de séparation

Illustration du cas où le système $Ax = b, x \geq 0$ possède une solution



$$c^T d = \|c\|\|d\| \cos\theta$$

$a_{\bullet 1}{}^{\mathrm{T}} y \geq 0$

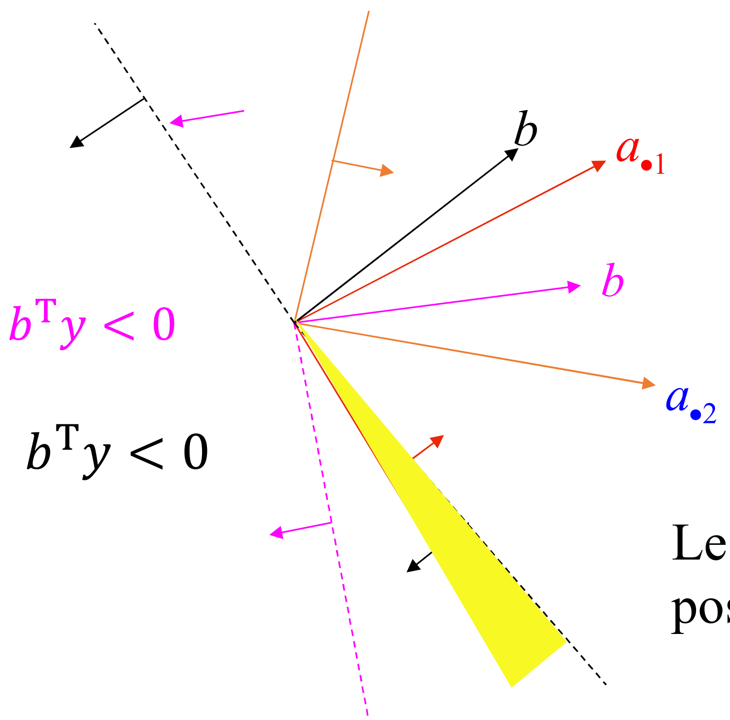
et

$a_{\bullet 2}{}^{\mathrm{T}} y \geq 0$

Le système $A^T y \geq 0, b^T y < 0$ ne possède pas de solution.

Théorie de séparation

Illustration du cas où le système $Ax = b, x \geq 0$ ne possède pas une solution



$a_{\bullet 1}{}^{\mathrm{T}}y \geq 0$

et

$a_{\bullet 2}{}^{\mathrm{T}}y \geq 0$

Le système $A^T y \geq 0, b^T y < 0$ possède une solution

Problème de programmation convexe

Définition:

Un problème de programmation convexe est un problème d'optimisation sous la forme

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

Où $f_0(x), \{f_i(x)\}_{i\in I_1}$ sont des fonctions convexes continues sur $\mathbb{R}^d$et $\{h_i(x)\}_{i\in I_2}$ sont des fonctions linéaires

Théorème:

- Le domaine réalisable, D_F, est convexe $D_F = \{x \in \mathbb{R}^d \,|\, f_i(x) \leq 0 \quad i \in I_1, h_i(x) = 0 \quad i \in I_2\}$
- L'ensemble solution de (P) est un ensemble fermé convexe
- Si x^* est la solution locale de (P) alors x^* est aussi la solution globale
- Si la fonction objectif $f_0(x)$ est strictement convexe alors la solution de (P) est unique

Problème de programmation convexe

Definition: **(Problème de programmation quadratique (QP))**

$$(QP)\begin{cases} Min & \frac{1}{2}x^TQx + c^Tx \\ s.t & \bar{A}x - \bar{b} \leq 0 \\ & Ax - b = 0 \\ & x \in \mathbb{R}^n \end{cases}$$

Où $Q \in \mathbb{R}^{n\times n}$ c $\in \mathbb{R}^n$, $\bar{A} \in \mathbb{R}^{m\times n}$, $A \in \mathbb{R}^{p\times n}$, $\bar{b} \in \mathbb{R}^m$, b $\in \mathbb{R}^p$

Si $\boldsymbol{Q}$ est semi-défini positif, alors le problème $\boldsymbol{QP}$ est une programmation convexe

Théorème:

- L'ensemble $D_F = \left\{x \in \mathbb{R}^d \middle| \bar{A}x - \bar{b} \leq 0, Ax - b = 0\right\}$ est convexe
- L'ensemble solution de (QP) est un ensemble fermé convexe
- Si x^* est la solution locale de(QP)alors x^* est aussi la solution globale
- Si $\boldsymbol{Q}$ est définie positive alors La solution de (QP) est unique

Problème de programmation convexe

Definition:

Considérons le problème de programmation convexe (P) avec la variable x divisée sous la forme $x = (x_1, x_2) \in \mathbb{R}^n$.

- $x_1^* \in \mathbb{R}^{m_1}$est appelé sa solution par rapport à x_1 s'il existe un $x_2^* \in \mathbb{R}^{n-m_1}$ tel que $x^* = (x_1^*, x_2^*)$ est sa solution
- L'ensemble de toutes les solutions par rapport à x_1 est appelé l'ensemble de solutions par rapport à x_1

Théorème:

Si le problème de programmation convexe (P) avec la variable $x = (x_1, x_2) \in \mathbb{R}^n$, alors

- son ensemble solution par rapport à x_1 est un ensemble fermé convexe
- Si $f_0(x) = F_1(x_1) + F_2(x_2)$ où est strictement convexe de la variable x_1 alors la solution de (P) par rapport à x_1 est unique lorsqu'elle a une solution