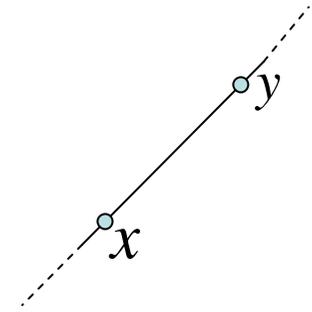


Ensembles et fonctions convexes

Ensembles convexes

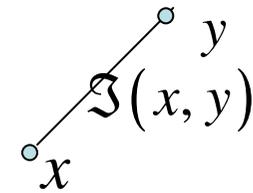
Définition: Étant donné deux points $x, y \in R^n$, la **droite** passant par x et y est définie par l'ensemble

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

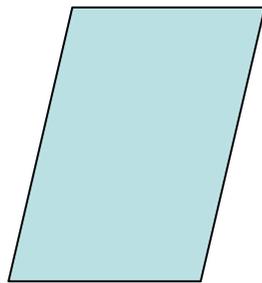
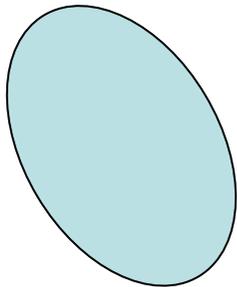


Définition: Étant donné deux points $x, y \in R^n$, le **segment de droite** reliant x et y est définie par l'ensemble

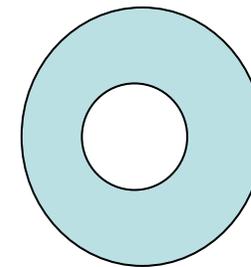
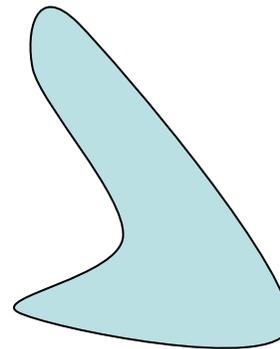
$$\mathfrak{S}(x, y) = \{z \in R^n : z = \theta x + (1 - \theta) y, \theta \in [0, 1]\}.$$



Définition: Un ensemble $X \subset R^n$ est **convexe** si pour toute paire de points $x, y \in X$, le segment de droite $\mathfrak{S}(x, y) \subset X$.



Ensembles convexes

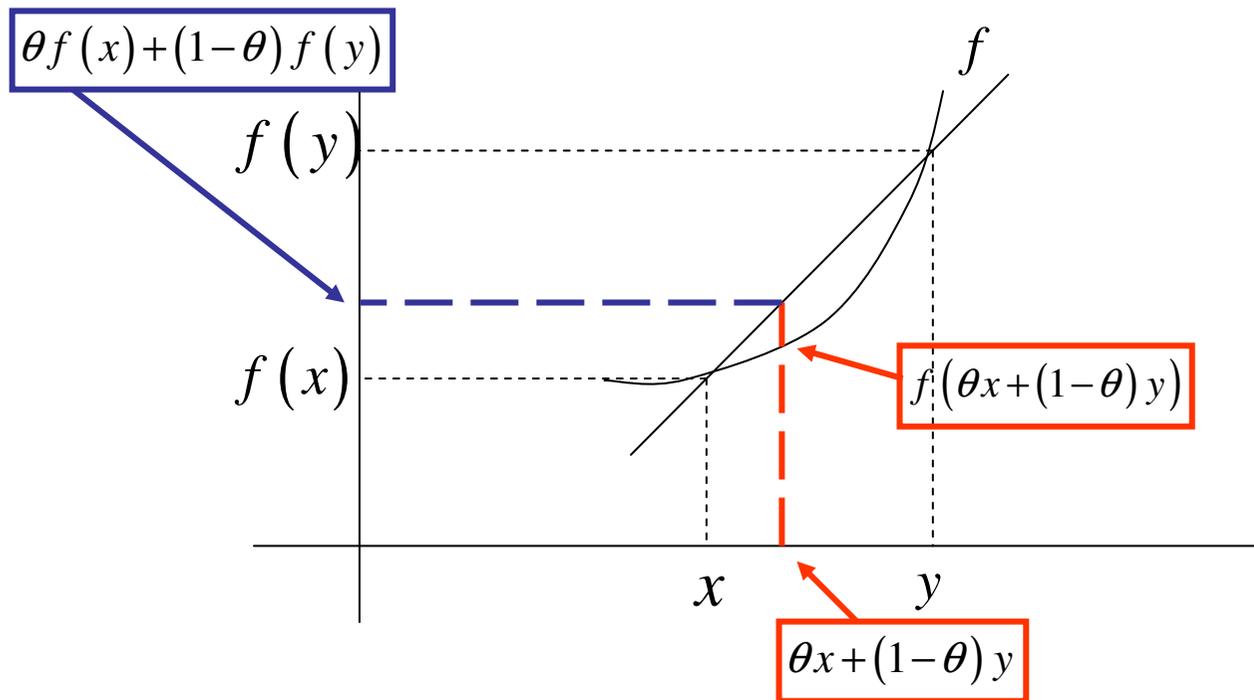


Ensembles non convexes

Fonctions convexes

Définition: Étant donné un ensemble convexe $X \subset \mathbb{R}^n$, une fonction à valeurs réelles $f : X \rightarrow \mathbb{R}^1$ est **convexe** si pour toute paire de points $x, y \in X$,

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



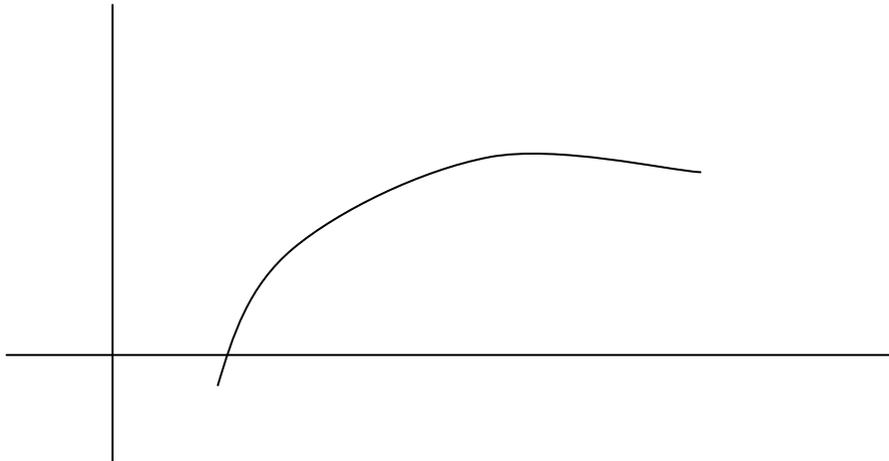
Définition: Étant donné un ensemble convexe $X \subset \mathbb{R}^n$, une fonction à valeurs réelles $f : X \rightarrow \mathbb{R}^1$ est **strictement convexe** si pour toute paire de points $x, y \in X$,

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

Définition: Étant donné un ensemble convexe $X \subset \mathbb{R}^n$, une fonction à valeurs réelles $f : X \rightarrow \mathbb{R}^1$ est **concave** si pour toute paire de points $x, y \in X$,

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

Donc f est concave si $-f$ est convexe.



Définition: Étant donné un ensemble convexe $X \subset R^n$, une fonction à valeurs réelles $f : X \rightarrow R^1$ est **strictement convexe** si pour toute paire de points $x, y \in X$,

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

Définition: Étant donné un ensemble convexe $X \subset R^n$, une fonction à valeurs réelles $f : X \rightarrow R^1$ est **concave** si pour toute paire de points $x, y \in X$,

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

Donc f est concave si $-f$ est convexe.

Définition: Étant donné un ensemble convexe $X \subset R^n$, une fonction à valeurs réelles $f : X \rightarrow R^1$ est **strictement concave** si pour toute paire de points $x, y \in X$,

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

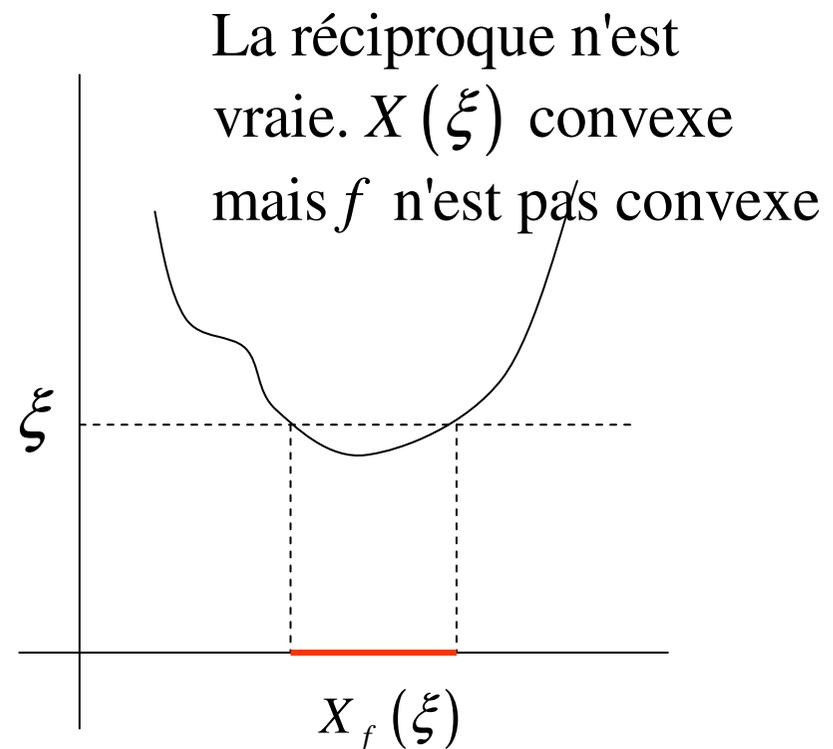
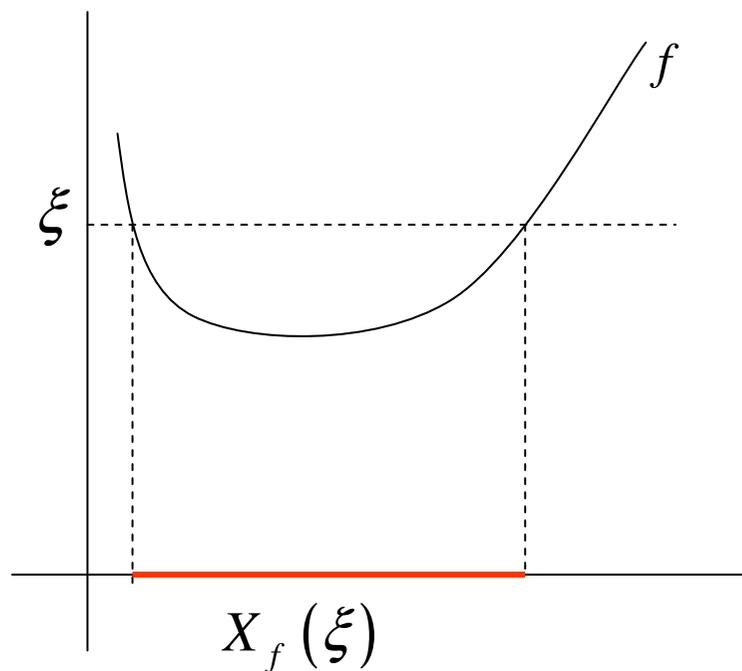
Donc f est strictement concave si $-f$ est strictement convexe.

Proposition 4.1: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Si $f : X \rightarrow \mathbb{R}^1$ est convexe, alors l'ensemble $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ est convexe pour tout $\xi \in \mathbb{R}^1$.

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

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta \xi + (1-\theta)\xi = \xi,$$

ce qui implique que $(\theta x + (1-\theta)y) \in X_f(\xi)$. Donc $X_f(\xi)$ est convexe. \square

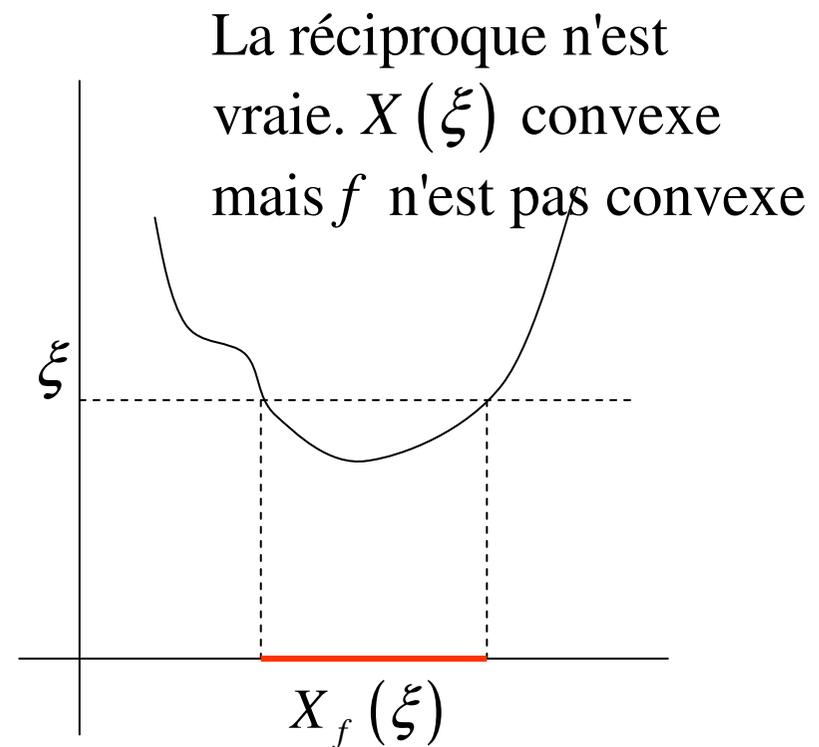
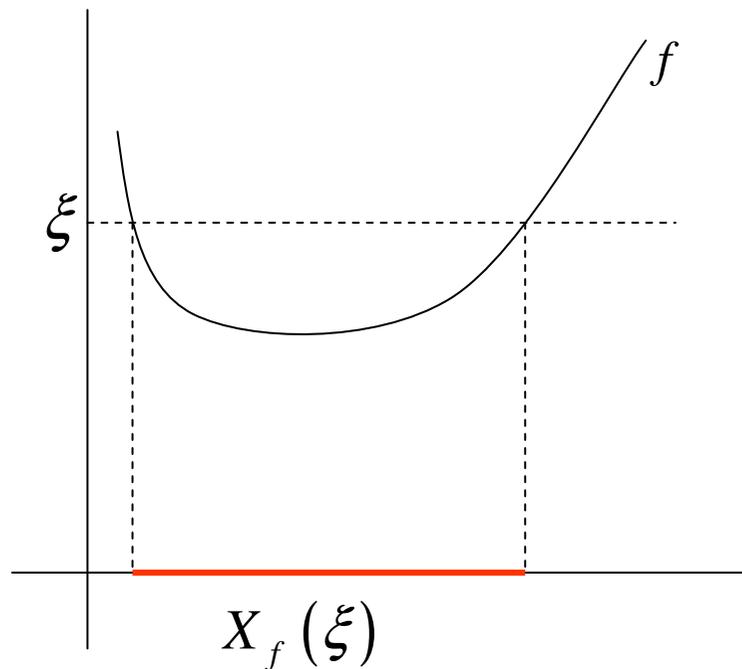


Proposition 4.1: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Si $f : X \rightarrow \mathbb{R}^1$ est convexe, alors l'ensemble $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ est convexe pour tout $\xi \in \mathbb{R}^1$.

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

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta\xi + (1-\theta)\xi = \xi,$$

ce qui implique que $(\theta x + (1-\theta)y) \in X_f(\xi)$. Donc $X_f(\xi)$ est convexe. \square

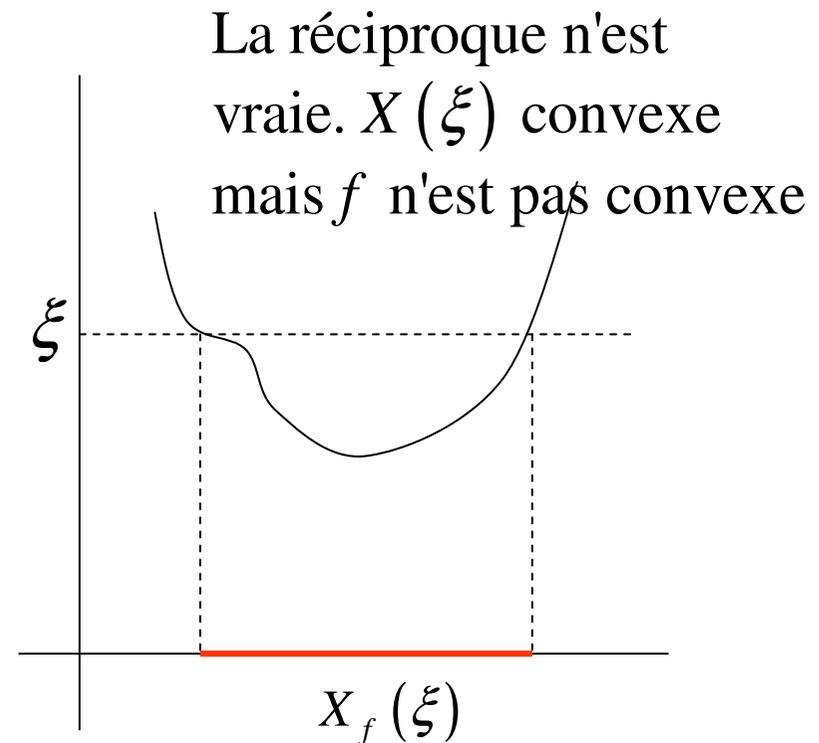
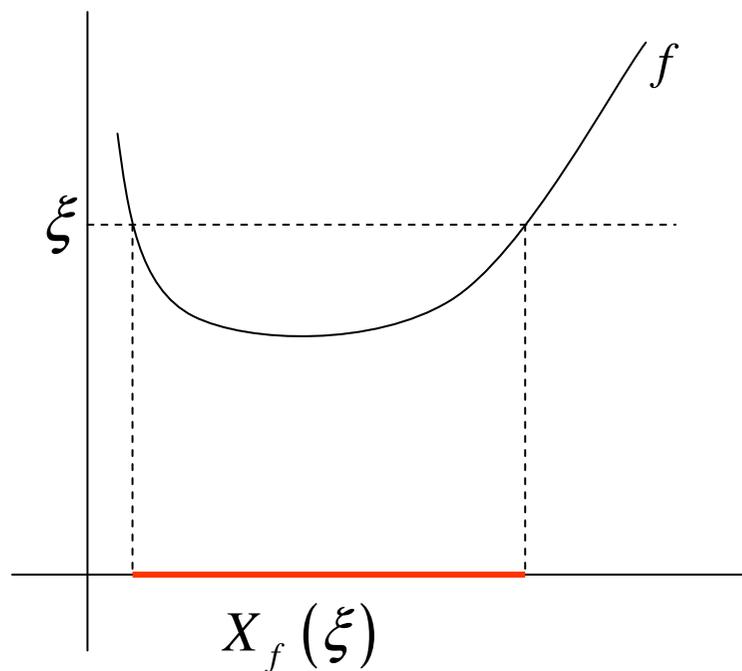


Proposition 4.1: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Si $f : X \rightarrow \mathbb{R}^1$ est convexe, alors l'ensemble $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ est convexe pour tout $\xi \in \mathbb{R}^1$.

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

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta\xi + (1-\theta)\xi = \xi,$$

ce qui implique que $(\theta x + (1-\theta)y) \in X_f(\xi)$. Donc $X_f(\xi)$ est convexe. \square

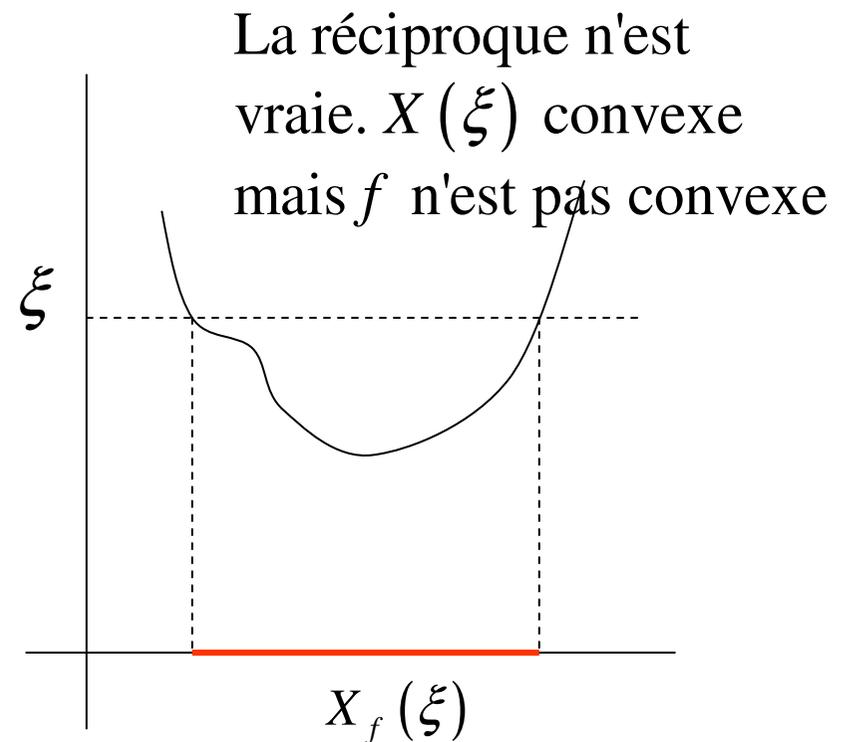
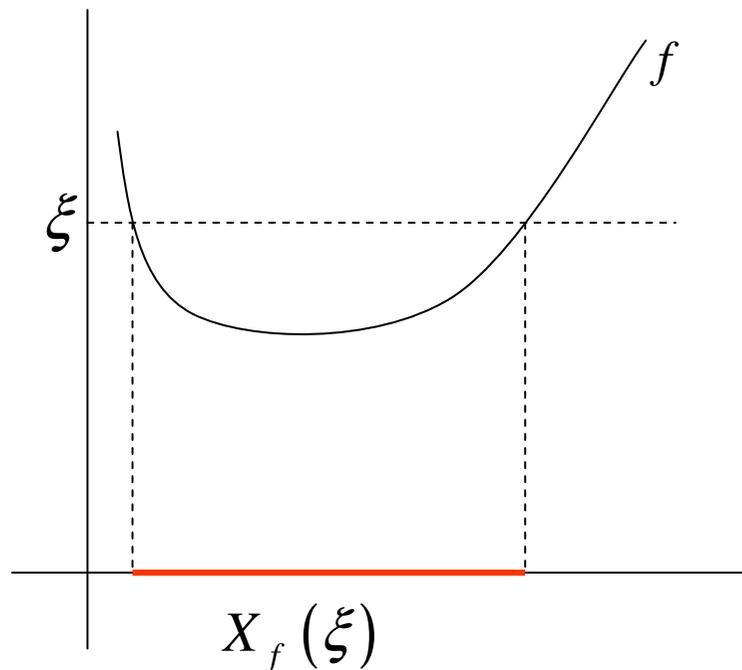


Proposition 4.1: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Si $f : X \rightarrow \mathbb{R}^1$ est convexe, alors l'ensemble $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ est convexe pour tout $\xi \in \mathbb{R}^1$.

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

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta\xi + (1-\theta)\xi = \xi,$$

ce qui implique que $(\theta x + (1-\theta)y) \in X_f(\xi)$. Donc $X_f(\xi)$ est convexe. \square



Proposition 4.2: Soit $X \subset R^n$. Si $f_i : X \rightarrow R^1$ 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 $\theta \in [0, 1]$

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

□

Proposition 4.3: Soit $X \subset R^n$ un ensemble convexe. Si $f : X \rightarrow R^1$ est convexe,

si $x^1, \dots, x^m \in X$, et si $\lambda_1, \dots, \lambda_m$ sont des scalaires non négatifs tels que $\sum_{i=1}^m \lambda_i = 1$,

$$\text{alors } f\left(\sum_{i=1}^m \lambda_i x^i\right) \leq \sum_{i=1}^m \lambda_i f(x^i).$$

Preuve. Si $m = 2$ nous retrouvons la définition de fonction convexe. Procédons par induction et supposons que le résultat est vrai pour $(m-1)$; i.e., si $\theta_1, \dots, \theta_{m-1} \geq 0$

sont tels que $\sum_{i=1}^{m-1} \theta_i = 1$, alors $\sum_{i=1}^{m-1} \theta_i x^i \in X$ et $f\left(\sum_{i=1}^{m-1} \theta_i x^i\right) \leq \sum_{i=1}^{m-1} \theta_i f(x^i)$.

Ainsi posant

$$\theta_i = \frac{\lambda_i}{\sum_{k=1}^{m-1} \lambda_k}$$

$$\sum_{i=1}^{m-1} \frac{\lambda_i}{\sum_{k=1}^{m-1} \lambda_k} = 1,$$

$$\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} \in X,$$

$$f\left(\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k}\right) \leq \sum_{i=1}^{m-1} \frac{\lambda_i f(x^i)}{\sum_{k=1}^{m-1} \lambda_k}.$$

Preuve. Si $m = 2$ nous retrouvons la définition de fonction convexe. Procédons par induction et supposons que le résultat est vrai pour $(m-1)$; i.e.,

$$\sum_{i=1}^{m-1} \frac{\lambda_i}{\sum_{k=1}^{m-1} \lambda_k} = 1,$$

$$\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} \in X,$$

$$f \left(\sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} \right) \leq \sum_{i=1}^{m-1} \frac{\lambda_i f(x^i)}{\sum_{k=1}^{m-1} \lambda_k}.$$

Par définition de convexité,

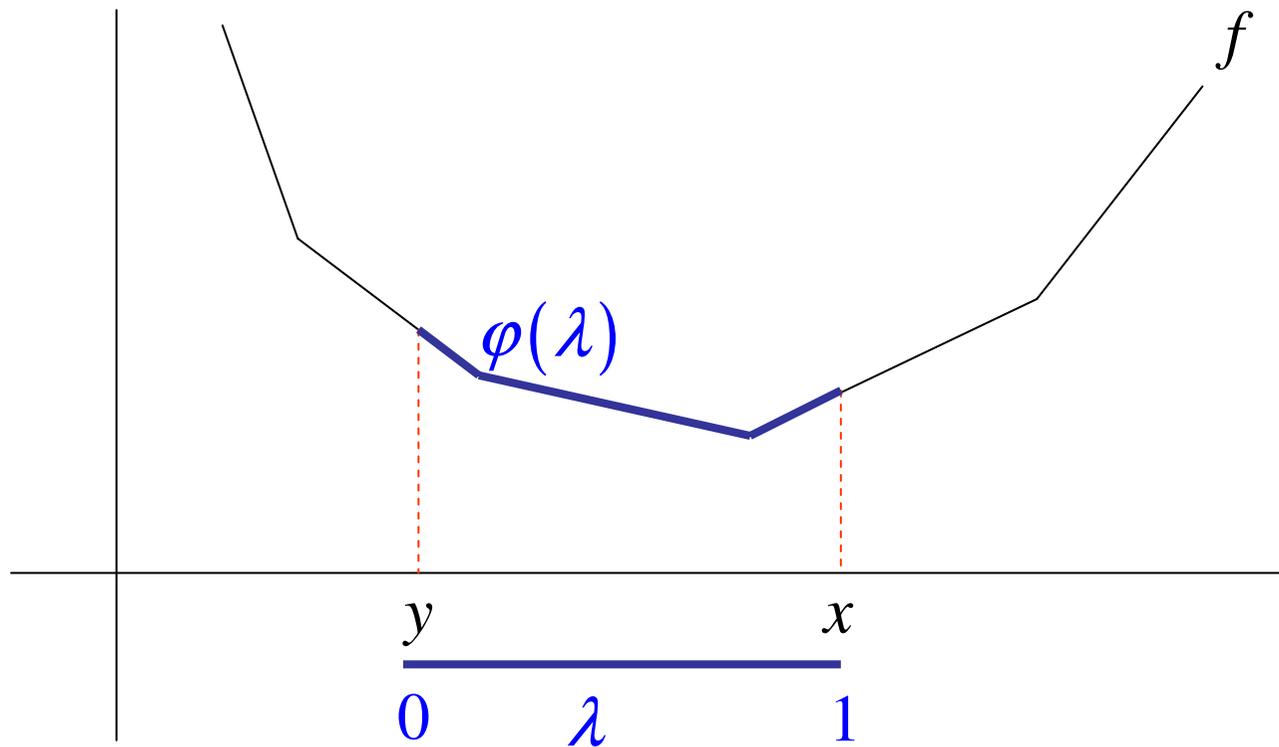
$$\sum_{k=1}^{m-1} \lambda_k \sum_{i=1}^{m-1} \frac{\lambda_i x^i}{\sum_{k=1}^{m-1} \lambda_k} + \lambda_m x^m \in X,$$

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

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

Proposition 4.4: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. $f : X \rightarrow \mathbb{R}^1$ est convexe si et seulement si $\phi(\lambda) = f(\lambda x + (1-\lambda)y)$ est une fonction convexe sur $[0,1]$ pour toute paire de points $x, y \in X$.

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



Proposition 4.4: Soit $X \subset R^n$ un ensemble convexe. $f : X \rightarrow R^1$ est convexe si et seulement si $\phi(\lambda) = f(\lambda x + (1-\lambda)y)$ est une fonction convexe sur $[0,1]$ pour toute paire de points $x, y \in X$.

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

Preuve. Pour démontrer la nécessité, considérons la paire de points $x, y \in X$. Soit $\lambda_1, \lambda_2 \in [0,1]$. Alors

$$\begin{aligned}\phi(\theta\lambda_1 + (1-\theta)\lambda_2) &= \phi(\lambda_2 + \theta(\lambda_1 - \lambda_2)) \\ &= f\left[\left(\lambda_2 + \theta(\lambda_1 - \lambda_2)\right)x + \left(1 - \lambda_2 - \theta(\lambda_1 - \lambda_2)\right)y\right] \\ &= f\left[\left(\lambda_2 + \theta(\lambda_1 - \lambda_2)\right)x + \left(1 + \theta - \theta - \lambda_2 - \theta(\lambda_1 - \lambda_2)\right)y\right] \\ &= f\left[\theta(\lambda_1 x + (1 - \lambda_1)y) + (1 - \theta)(\lambda_2 x + (1 - \lambda_2)y)\right] \\ &\leq \theta f(\lambda_1 x + (1 - \lambda_1)y) + (1 - \theta) f(\lambda_2 x + (1 - \lambda_2)y) \\ &= \theta\phi(\lambda_1) + (1 - \theta)\phi(\lambda_2).\end{aligned}$$

Donc ϕ est convexe sur $[0,1]$.

Pour démontrer la suffisance, considérons également la paire de points $x, y \in X$. La fonction $\phi(\theta) = f(\theta x + (1-\theta)y)$ étant convexe sur $[0,1]$,

$$\phi(\theta) = \phi(\theta \cdot 1 + (1-\theta)0) \leq \theta\phi(1) + (1-\theta)\phi(0)$$

pour tout $\theta \in [0,1]$.

Alors

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

pour tout $\theta \in [0,1]$. Donc f est convexe. □

La proposition suivante introduit une des propriétés qui rendent les fonctions convexes si intéressantes en programmation mathématique.

Proposition 4.5: Soit $X \subset R^n$ un ensemble convexe. Soit $f : X \rightarrow R^1$ 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 en supposant qu'il existe un point $\tilde{x} \in X$ tel

que $f(\tilde{x}) < f(\bar{x})$. Puisque f est convexe,

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

pour tout $\theta \in (0, 1]$. Or pour $\theta > 0$ suffisamment petit,

$$x(\theta) = \theta\tilde{x} + (1-\theta)\bar{x} \in B_\varepsilon(\bar{x}) \cap X.$$

Ainsi $f(x(\theta)) < f(\bar{x})$ où $x(\theta) \in B_\varepsilon(\bar{x}) \cap X$, contredisant que \bar{x} est un minimum local de f sur X . □

Proposition 4.6: (*Inégalité du gradient*). Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f : X \rightarrow \mathbb{R}^1$ une fonction de classe C^1 / X . Alors f est convexe sur X si et seulement si

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

pour toute paire de points $x, y \in X$.

Preuve. (nécessité) Soit f convexe. Alors pour toute paire de points $x, y \in X$ et pour tout $\theta \in [0, 1]$

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

ce qui s'écrit aussi sous la forme

$$f(y + \theta(x - y)) - f(y) \leq \theta(f(x) - f(y)). \quad (4.1)$$

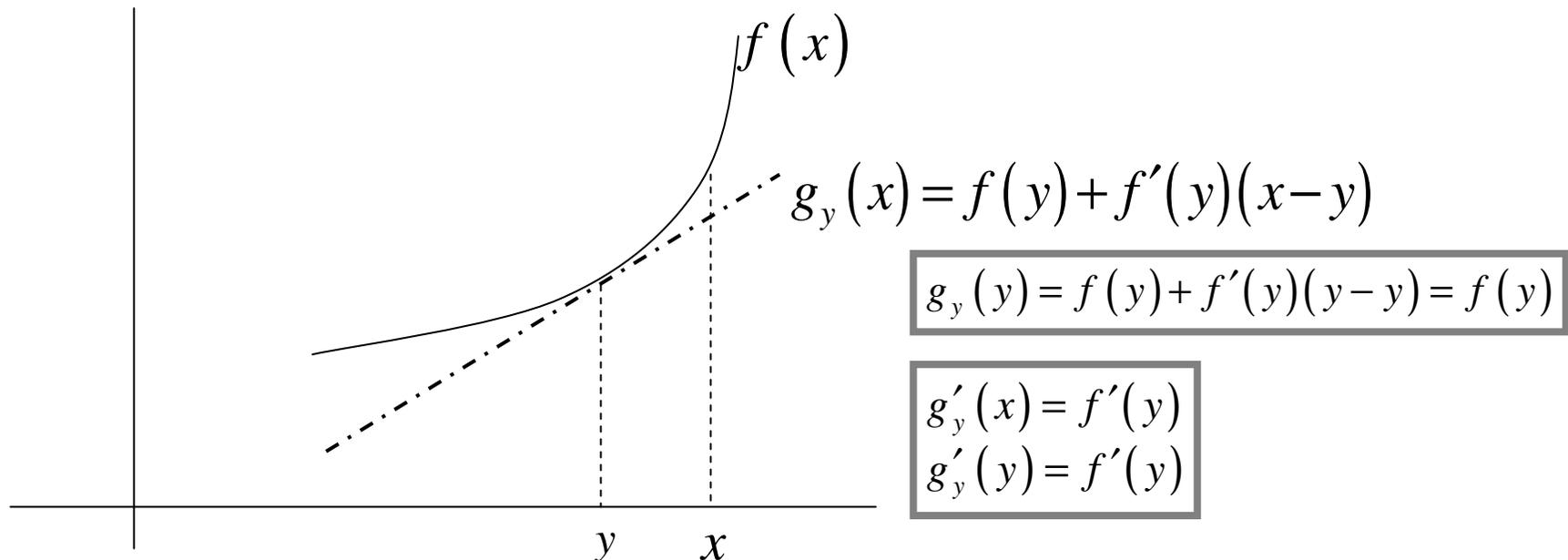
Se référant au Théorème de Taylor, il existe un $\tau \in [0, 1]$ tel que

$$f(y + \theta(x - y)) = f(y) + \nabla f \left[\tau(y + \theta(x - y)) + (1 - \tau)y \right]^T (y + \theta(x - y) - y)$$

$$f(y + \theta(x - y)) - f(y) = \theta \nabla f \left[y + \tau\theta(x - y) \right]^T (x - y). \quad (4.2)$$

La figure suivante illustre l'inégalité du gardient pour une fonction $f : R^1 \rightarrow R^1$. Dans ce cas l'inégalité du gradient devient

$$f(x) \geq f(y) + f'(y)(x - y).$$



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

$$g_y(y) = f(y) ; g'_y(y) = f'(y).$$

Ainsi la droite $g_y(x) = f(y) + f'(y)(x - y)$ prend la même valeur que f au point y , et sa pente est la même que celle de f au point y . Cette droite est donc 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 .

Proposition 4.6: (*Inégalité du gradient*). Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f : X \rightarrow \mathbb{R}^1$ une fonction de classe C^1 / X . Alors f est convexe sur X si et seulement si

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

pour toute paire de points $x, y \in X$.

Preuve. (nécessité) Soit f convexe. Alors pour toute paire de points $x, y \in X$ et pour tout $\theta \in [0, 1]$

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

ce qui s'écrit aussi sous la forme

$$f(y + \theta(x - y)) - f(y) \leq \theta(f(x) - f(y)). \quad (4.1)$$

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

$$f(y + \theta(x - y)) = f(y) + \nabla f \left[\tau(y + \theta(x - y)) + (1 - \tau)y \right]^T (y + \theta(x - y) - y)$$

$$f(y + \theta(x - y)) - f(y) = \theta \nabla f \left[y + \tau\theta(x - y) \right]^T (x - y). \quad (4.2)$$

$$\boxed{f(y + \theta(x - y)) - f(y)} \leq \theta(f(x) - f(y)). \quad (4.1)$$

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

$$f(y + \theta(x - y)) = f(y) + \nabla f \left[\tau(y + \theta(x - y)) + (1 - \tau)y \right]^T (y + \theta(x - y) - y)$$

$$\boxed{f(y + \theta(x - y)) - f(y)} = \theta \nabla f \left[y + \tau\theta(x - y) \right]^T (x - y). \quad (4.2)$$

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

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

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

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

pour tout $\theta \in (0, 1]$. Par conséquent, en prenant la limite quand $\theta \rightarrow 0$ nous obtenons

$$\lim_{\theta \rightarrow 0} \nabla f \left[y + \tau\theta(x - y) \right]^T (x - y) \leq \lim_{\theta \rightarrow 0} (f(x) - f(y))$$

ou encore

$$\nabla f \left[y + \lim_{\theta \rightarrow 0} \tau\theta(x - y) \right]^T (x - y) \leq (f(x) - f(y))$$

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

(suffisance) Démontrons maintenant que la condition est suffisante.

Puisque $f(x) \geq f(y) + \nabla f(y)^T (x - y)$ pour toute paire de points $x, y \in X$

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

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

(4.3) et (4.4) s'écrivent respectivement

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

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

Multiplions (4.3) par θ et (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:

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

Corollaire 4.7: Soit $X \subset \mathbb{R}^n$ un ensemble convexe. Soit $f : X \rightarrow \mathbb{R}^1$ une fonction convexe de classe C^1 / X . Alors si $\nabla f(x^*)^\top (x - x^*) \geq 0$ pour tout $x \in X$, 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 X , alors

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*)$$

pour tout $x \in X$. Puisque par hypothèse $\nabla f(x^*)^\top (x - x^*) \geq 0$, alors

$$f(x) - f(x^*) \geq \nabla f(x^*)^\top (x - x^*) \geq 0$$

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 les fonctions f de classe C^2 , il existe un critère pour vérifier la convexité qui est basé sur le Hessien $\nabla^2 f$.

Proposition 3.8: Soit $X \subset \mathbb{R}^n$ un ensemble convexe dont l'intérieur est non vide. Soit $f : X \rightarrow \mathbb{R}^1$ une fonction de classe C^2 / X . Alors f est convexe sur X si et seulement si son Hessien $\nabla^2 f(x)$ est une matrice semi définie positive pour tout $x \in X$.