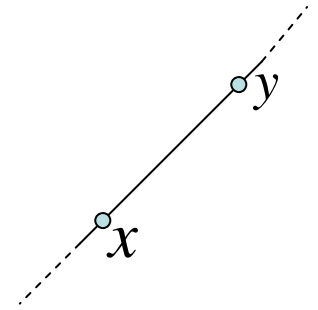


Convexity

Convex sets

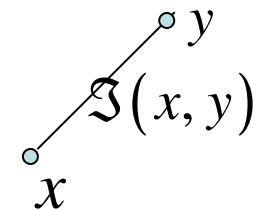
Definition: Let $x, y \in R^n$, the **straight line** through x and y is specified as follows

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



Definition: Let $x, y \in R^n$, the **line segment** linking x and y is specified as follows

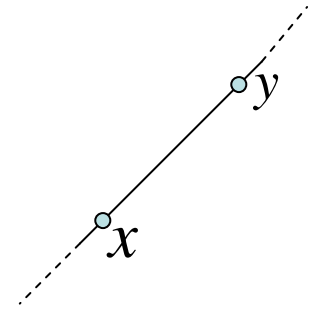
$$\mathfrak{S}(x, y) = \{z \in R^n : z = \theta x + (1 - \theta) y, \theta \in [0, 1]\}.$$



Definition: A set $X \subset R^n$ is **convex** if for any pair of points $x, y \in X$, the line segment $\mathfrak{S}(x, y) \subset X$.

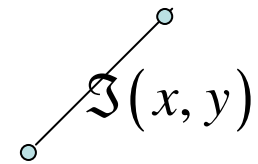
Definition: Let $x, y \in R^n$, the **straight line** through x and y is specified as follows

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

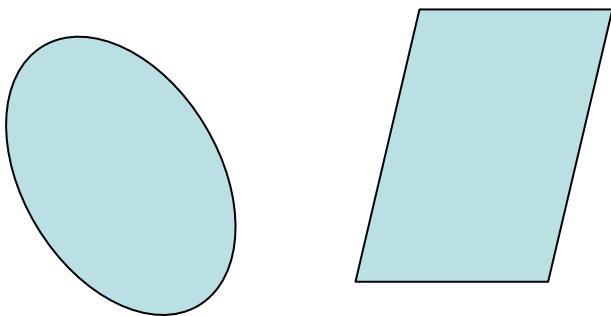


Definition: Let $x, y \in R^n$, the **line segment** linking x and y is specified as follows

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



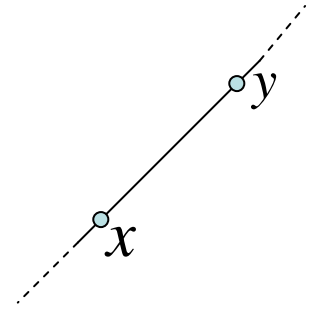
Definition: A set $X \subset R^n$ is **convex** if for any pair of points $x, y \in X$, the line segment $\mathfrak{S}(x, y) \subset X$.



Convex sets

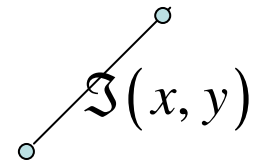
Definition: Let $x, y \in R^n$, the **straight line** through x and y is specified as follows

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

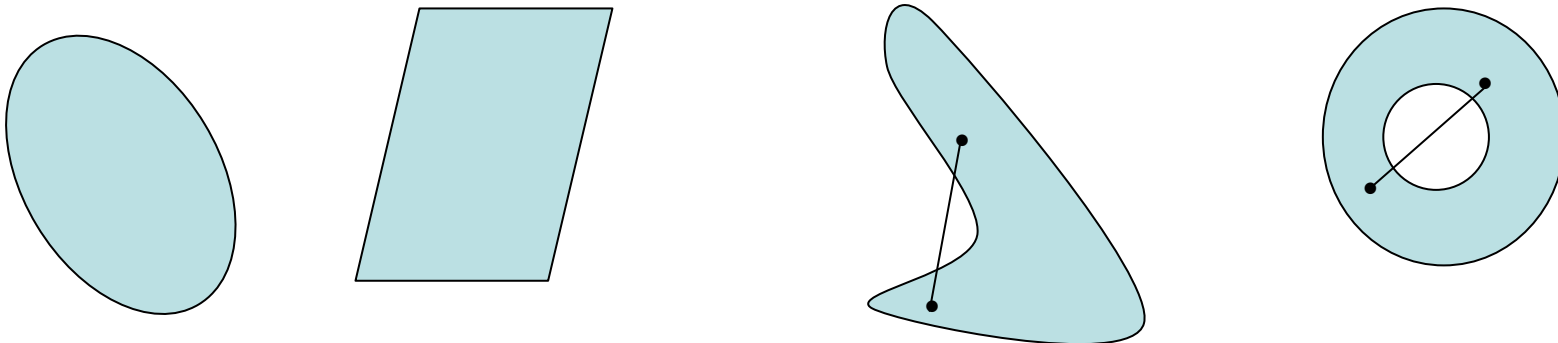


Definition: Let $x, y \in R^n$, the **line segment** linking x and y is specified as follows

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



Definition: A set $X \subset R^n$ is **convex** if for any pair of points $x, y \in X$, the line segment $\mathfrak{S}(x, y) \subset X$.



Convex sets

Non convex sets

Recall: Notation et definitions

- **Notation**

R^n vectorial space of n dimension *vectors*.

$f : R^n \rightarrow R^1$ real valued function (taking its value in R^1)

For any $x \in R^n$

$\|x\| = \sqrt{x^T x}$ is the euclidean norm and

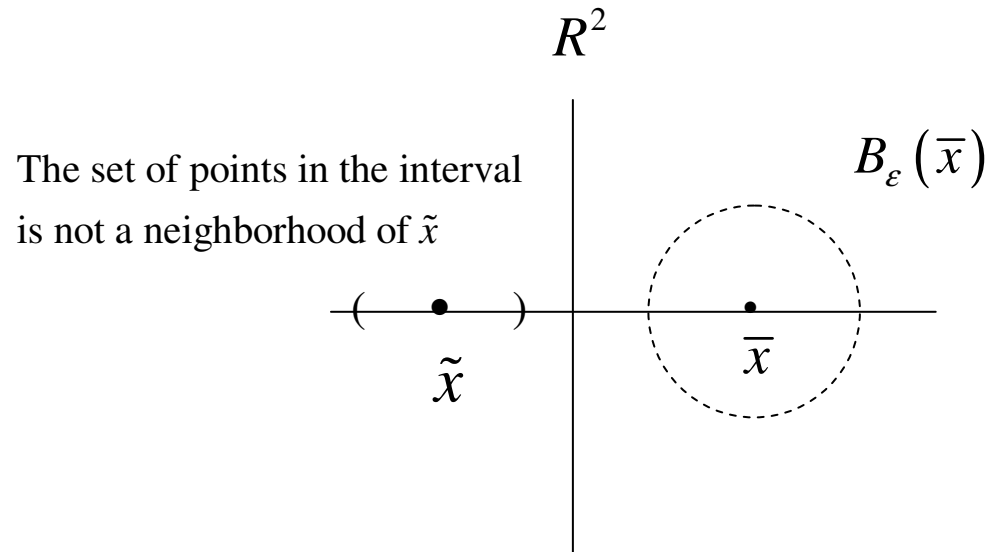
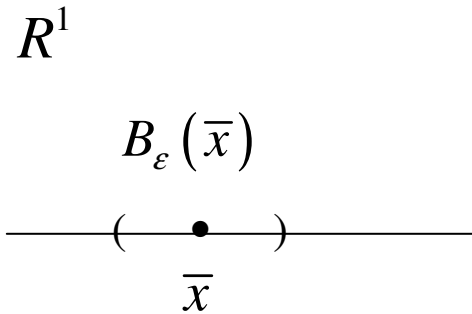
$x^T x = \sum_{j=1}^n x_j^2$ is the scalar product of x by itself

Definition. A point $\bar{x} \in R^n$ is a **local minimum** of f if there exists a neighborhood of \bar{x} denoted

$$B_\varepsilon(\bar{x}) = \left\{ x \in R^n : \|x - \bar{x}\| = \sqrt{(x - \bar{x})^T (x - \bar{x})} < \varepsilon \right\}$$

such that

$$f(\bar{x}) \leq f(x) \quad \forall x \in B_\varepsilon(\bar{x})$$

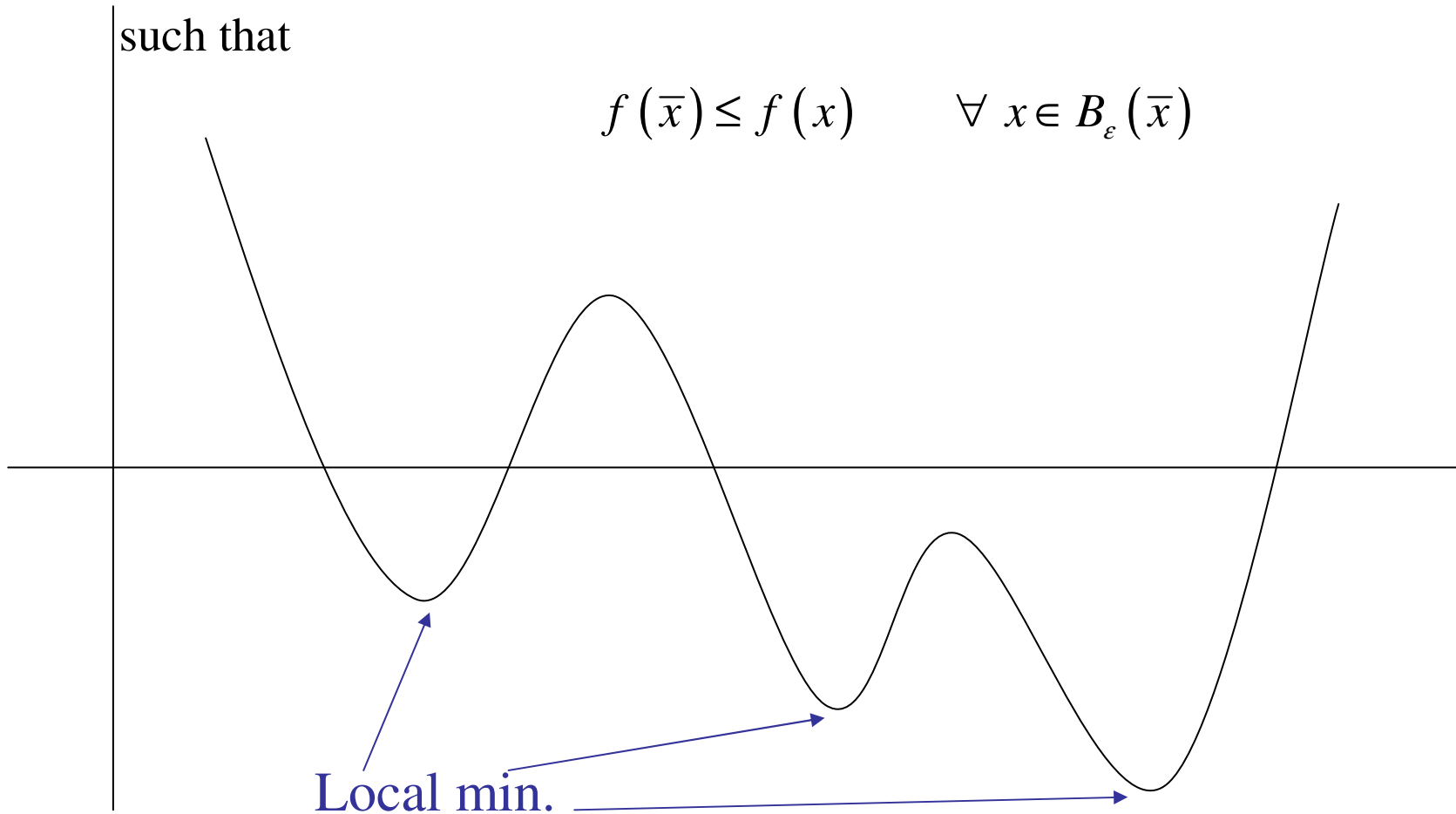


Definition. A point $\bar{x} \in R^n$ is a **local minimum** of f if there exists a neighborhood of \bar{x} denoted

$$B_\varepsilon(\bar{x}) = \left\{ x \in R^n : \|x - \bar{x}\| = \sqrt{(x - \bar{x})^T (x - \bar{x})} < \varepsilon \right\}$$

such that

$$f(\bar{x}) \leq f(x) \quad \forall x \in B_\varepsilon(\bar{x})$$



Definition. A point $\bar{x} \in R^n$ is a **local minimum** of f if there exists a neighborhood of \bar{x} denoted

$$B_\varepsilon(\bar{x}) = \left\{ x \in R^n : \|x - \bar{x}\| = \sqrt{(x - \bar{x})^T (x - \bar{x})} < \varepsilon \right\}$$

such that

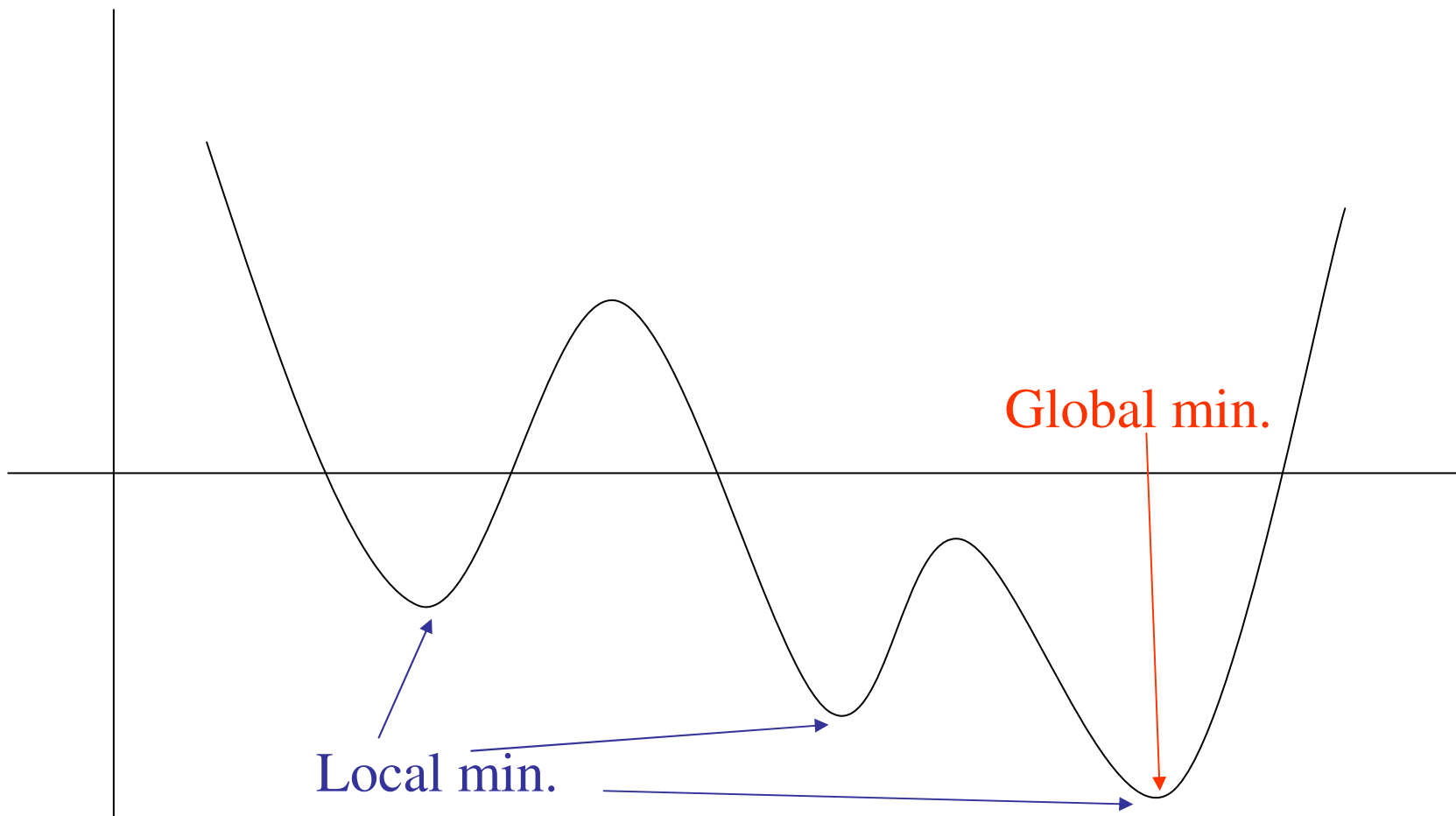
$$f(\bar{x}) \leq f(x) \quad \forall x \in B_\varepsilon(\bar{x})$$

Definition. A point $\bar{x} \in R^n$ is a **global minimum** of f if

$$f(\bar{x}) \leq f(x) \quad \forall x \in R^n$$

Definition. A point $\bar{x} \in R^n$ is a **global minimum** of f if

$$f(\bar{x}) \leq f(x) \quad \forall x \in R^n$$



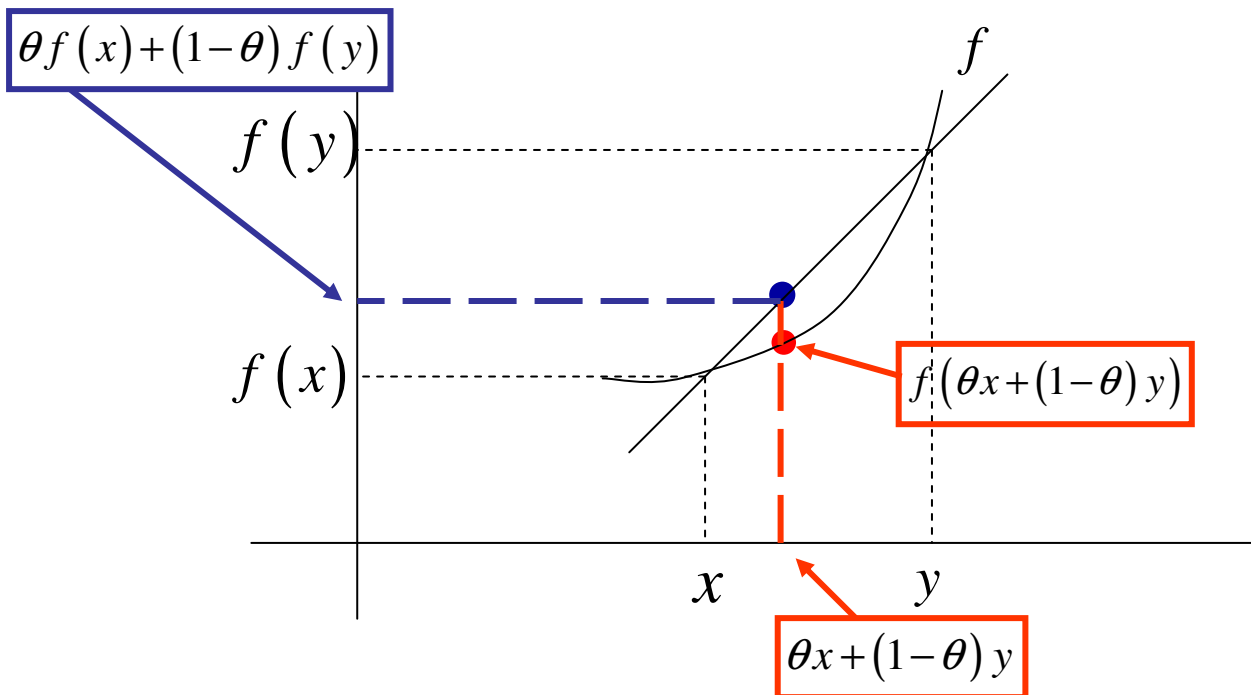
Convex function

Definition: Let $X \subset \mathbb{R}^n$ be a convex set, a real valued function $f : X \rightarrow \mathbb{R}^1$ is **convex** if for all pairs of points $x, y \in X$,

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

Fonctions convexes

Definition: Let $X \subset \mathbb{R}^n$ be a convex set, a real valued function $f : X \rightarrow \mathbb{R}^1$ is **convex** if for all pairs of points $x, y \in X$,

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


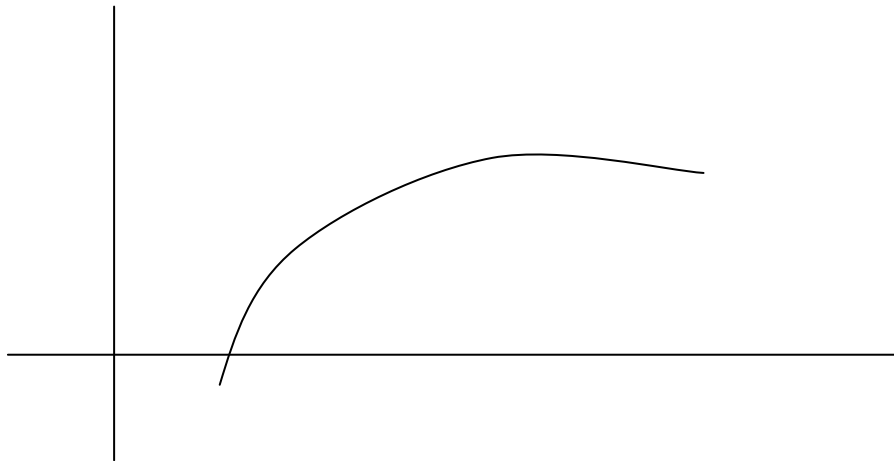
Definition: Let $X \subset \mathbb{R}^n$ be a convex set, a real valued function $f : X \rightarrow \mathbb{R}^1$ is **convex** if for all pairs of points $x, y \in X$,

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

Definition: Let $X \subset \mathbb{R}^n$ be a convex set, a real valued function $f : X \rightarrow \mathbb{R}^1$ is **concave** if for all pairs of 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.

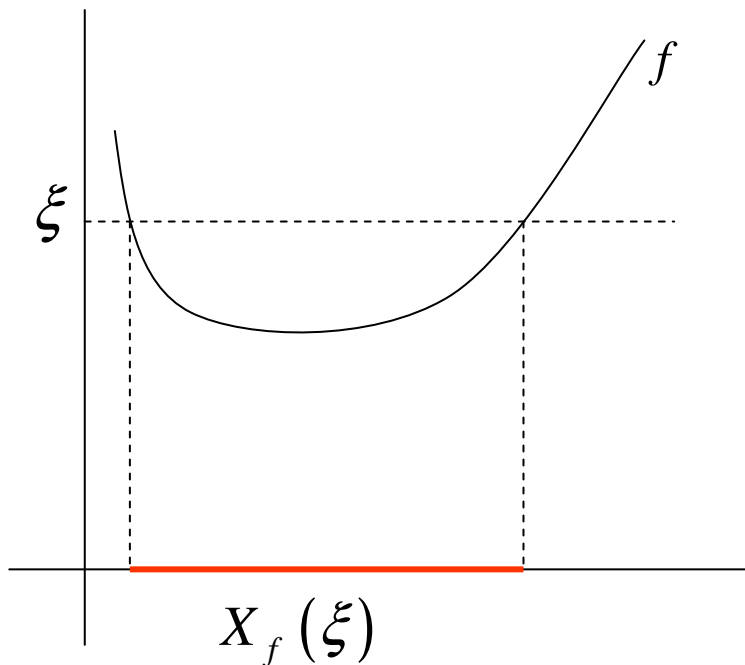


Proposition 1: Let $X \subset \mathbb{R}^n$ be a convex set. If $f : X \rightarrow \mathbb{R}^1$ is convex, then the set $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ is convex for all $\xi \in \mathbb{R}^1$.

Proof. Let x et $y \in X_f(\xi)$ and $\theta \in (0,1)$. The, since f is convex,

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

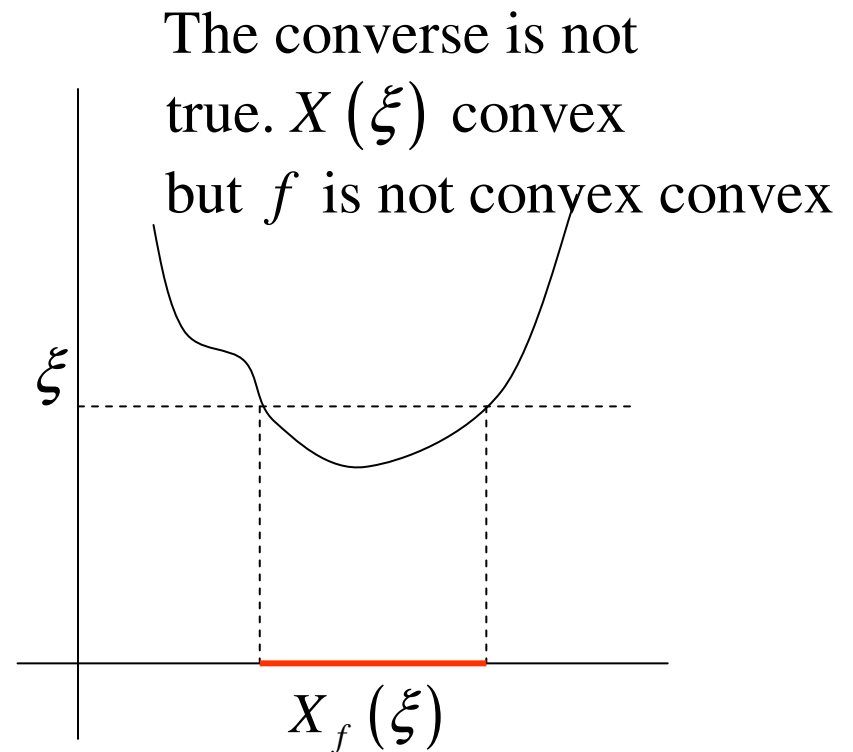
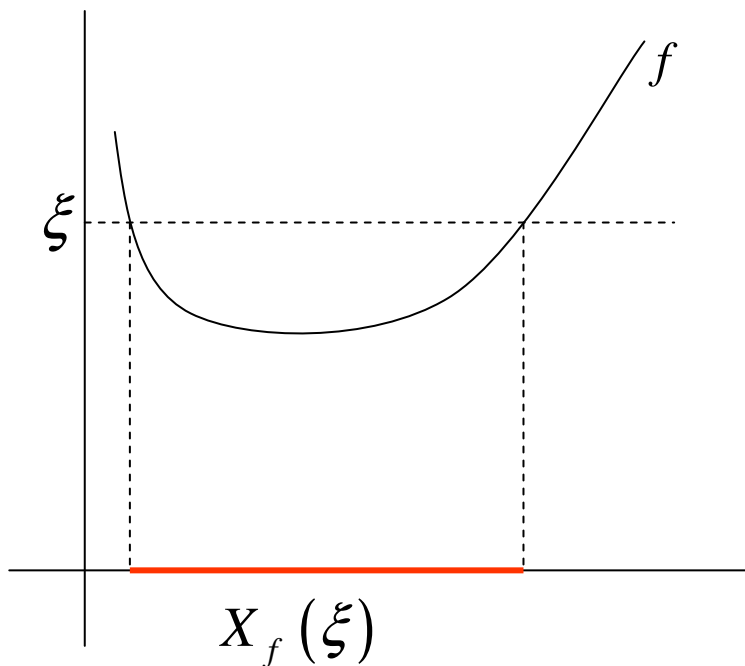
inducing that $(\theta x + (1-\theta)y) \in X_f(\xi)$. Thus $X_f(\xi)$ is convex. \square



Proposition 1: Let $X \subset \mathbb{R}^n$ be a convex set. If $f : X \rightarrow \mathbb{R}^1$ is convex, then the set $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ is convex for all $\xi \in \mathbb{R}^1$.

Proof. Let x and $y \in X_f(\xi)$ and $\theta \in (0,1)$. Then, since f is convex,

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta \xi + (1-\theta)\xi = \xi,$$
 inducing that $(\theta x + (1-\theta)y) \in X_f(\xi)$. Thus $X_f(\xi)$ is convex. \square

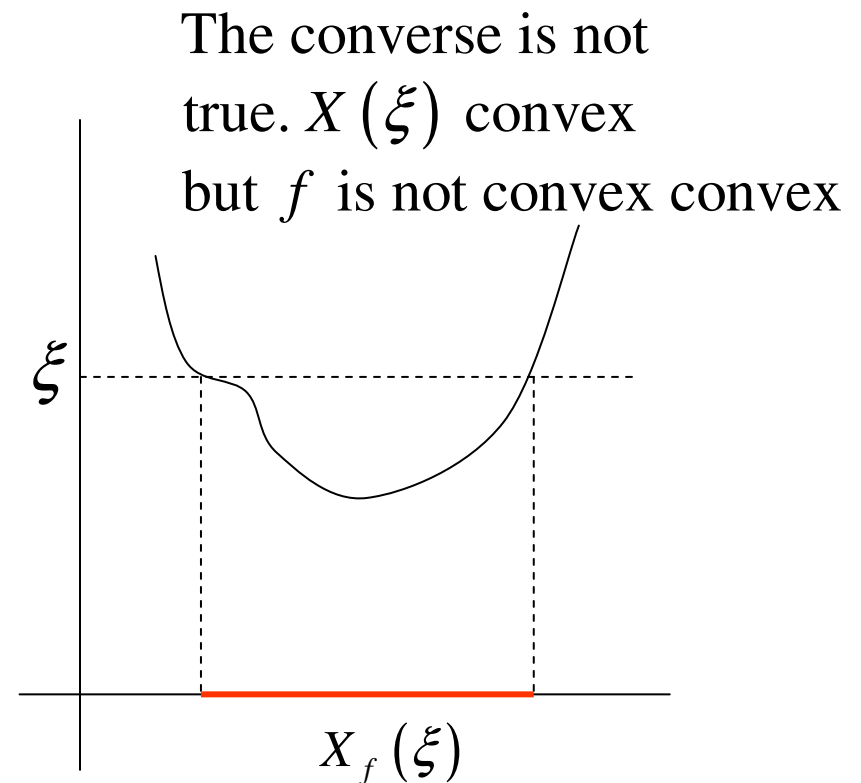
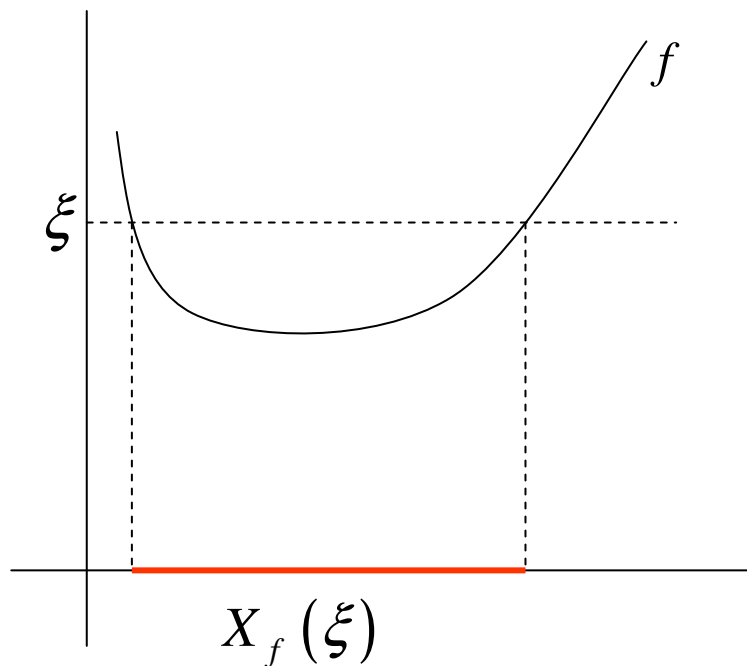


Proposition 1: Let $X \subset \mathbb{R}^n$ be a convex set. If $f : X \rightarrow \mathbb{R}^1$ is convex, then the set $X_f(\xi) = \{x \in X : f(x) \leq \xi\}$ is convex for all $\xi \in \mathbb{R}^1$.

Proof. Let x et $y \in X_f(\xi)$ and $\theta \in (0,1)$. The, since f is convex,

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

inducing that $(\theta x + (1-\theta)y) \in X_f(\xi)$. Thus $X_f(\xi)$ is convex. \square



Proposition 2: Let $X \subset R^n$. If $f_i : X \rightarrow R^1$ is convex for $i = 1, \dots, m$ and if $\lambda_1, \dots, \lambda_m$ are non negative scalars, then $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ is convex on X .

Proposition 2: Let $X \subset R^n$. If $f_i : X \rightarrow R^1$ is convex for $i = 1, \dots, m$ and if $\lambda_1, \dots, \lambda_m$ are non negative scalars, then $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ is convex on X .

Proof. For all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

$$f(\theta x + (1 - \theta)y) = \sum_{i=1}^m \lambda_i f_i(\theta x + (1 - \theta)y)$$

Proposition 2: Let $X \subset R^n$. If $f_i : X \rightarrow R^1$ is convex for $i = 1, \dots, m$ and if $\lambda_1, \dots, \lambda_m$ are non negative scalars, then $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ is convex on X .

Proof. For all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

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

Proposition 2: Let $X \subset R^n$. If $f_i : X \rightarrow R^1$ is convex for $i = 1, \dots, m$ and if $\lambda_1, \dots, \lambda_m$ are non negative scalars, then $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ is convex on X .

Proof. For all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

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

Proposition 2: Let $X \subset R^n$. If $f_i : X \rightarrow R^1$ is convex for $i = 1, \dots, m$ and if $\lambda_1, \dots, \lambda_m$ are non negative scalars, then $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ is convex on X .

Proof. For all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

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

□

Proposition 3: Let $X \subset R^n$ be a convex set. If $f : X \rightarrow R^1$ is convex, if $x^1, \dots, x^m \in X$, and if $\lambda_1, \dots, \lambda_m$ are non negative scalars such that $\sum_{i=1}^m \lambda_i = 1$,

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

Proposition 4: Let $X \subset R^n$ be a convex set. Let $f : X \rightarrow R^1$ be a convex function. If \bar{x} is a local minimum of f on X , then \bar{x} is a global minimum of f on X .

Proposition 4: Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a convex function. If \bar{x} is a local minimum of f on X , then \bar{x} is a global minimum of f on X .

Proof. For contradiction, suppose there exists another point $\tilde{x} \in X$ such that $f(\tilde{x}) < f(\bar{x})$.

Proposition 4: Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a convex function. If \bar{x} is a local minimum of f on X , then \bar{x} is a global minimum of f on X .

Proof. For contradiction, suppose there exists another point $\tilde{x} \in X$ such that

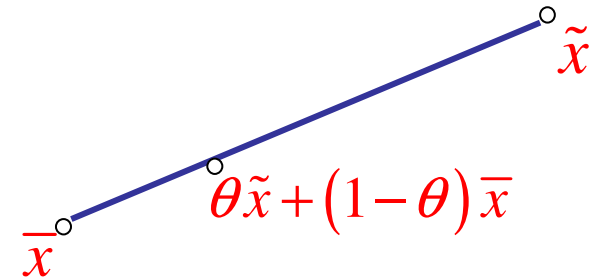
$f(\tilde{x}) < f(\bar{x})$. Since f is convex,

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

for all $\theta \in (0, 1]$. But for $\theta > 0$ sufficiently small,

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

Then $f(x(\theta)) < f(\bar{x})$ where $x(\theta) \in B_\varepsilon(\bar{x}) \cap X$, contradicting the fact that \bar{x} is a local minimum of f on X . □



Proposition 5: (*gradient inequality*). Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable function in C^1 / X . Then f is convex on X if and only if

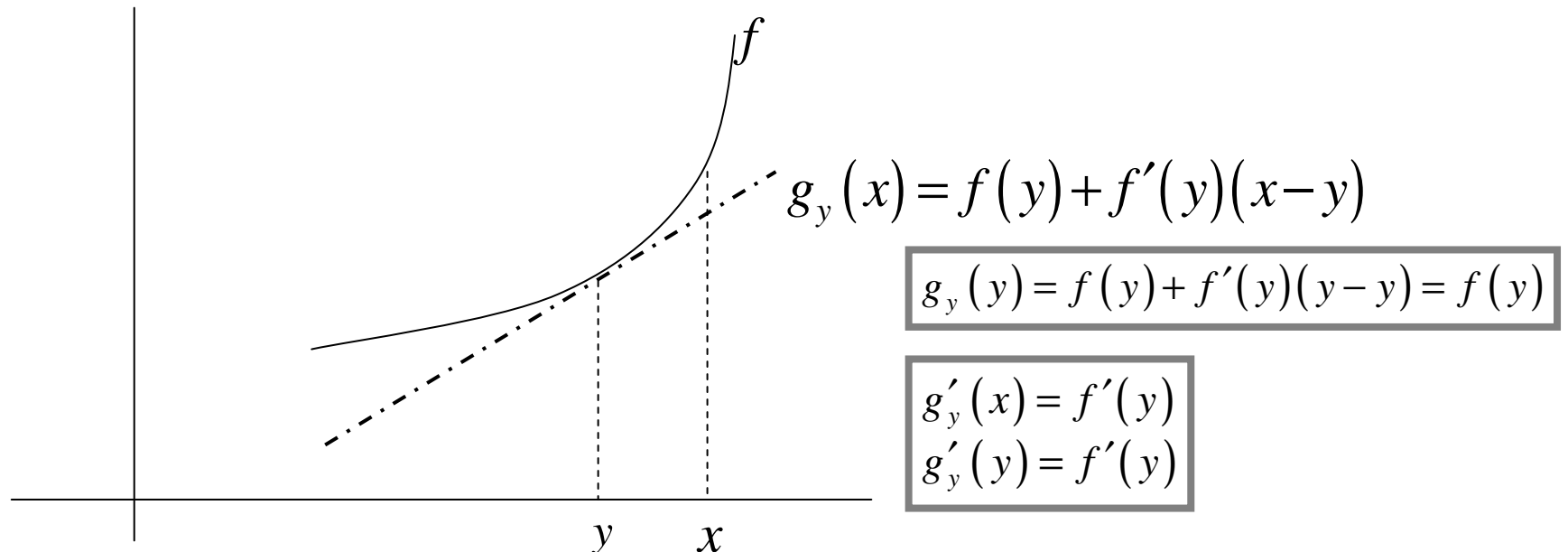
$$f(x) \geq f(y) + \nabla f(y)^T (x - y)$$

for all pairs of points $x, y \in X$.

$$\nabla f(x) = \left[\frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

Let us illustrate the gradient inequality for a function $f : \mathbb{R}^1 \rightarrow \mathbb{R}^1$.
 In this case the gradient inequality reduces to

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



The line $g_y(x) = f(y) + f'(y)(x - y)$ has the following properties:

$$g_y(y) = f(y); \quad g_y'(y) = f'(y).$$

Hence the line $g_y(x) = f(y) + f'(y)(x - y)$ takes the same value as f at point y , and its slope is the same as the slope of f at the point y .

Thus this line $g_y(x)$ is a support of f at point y .

The gradient inequality indicates that the function takes a value always larger or equal to the support function at any point y .

Proposition 5: (*gradient inequality*). Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable function in C^1 / X . Then f is convex on X if and only if

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

for all pairs of points $x, y \in X$.

Proposition 5: (*gradient inequality*). Let $X \subset R^n$ be a convex set. Let $f : X \rightarrow R^1$ be a differentiable function in C^1 / X . Then f is convex on X if and only if

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

$$\nabla f(x) = \left[\frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

for all pairs of points $x, y \in X$.

Proof. (necessity \Rightarrow) Let f be convex. Then for all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

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

or

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

Proposition 5: (*gradient inequality*). Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable function in C^1 / X . Then f is convex on X if and only if

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

$$\nabla f(x) = \left[\frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

for all pairs of points $x, y \in X$.

Proof. (necessity \Rightarrow) Let f be convex. Then for all pairs of points $x, y \in X$ and for all $\theta \in [0, 1]$

or

If $f \in C^1 / X$ (f continuously differentiable on X), then for all pair of points $r, s \in X$, there exists a point z on the line segment between r and s such that

the **Taylor development of first order** holds

$$f(s) = f(r) + \nabla f(z)^T (s - r)$$

(4.1)

Proposition 5: (*gradient inequality*). Let $X \subset R^n$ be a convex set. Let $f : X \rightarrow R^1$ be a differentiable function in C^1 / X . Then f is convex on X if and only if

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

$$\nabla f(x) = \left[\frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

for all pairs of points $x, y \in X$.

Proof. (necessity \Rightarrow) If $f \in C^1 / X$ (f continuously differentiable on X), then for all pair of points $r, s \in X$, there exists a point $z \in X$ on the line segment between r and s such that

f the **Taylor development of first order** holds

$$f(s) = f(r) + \nabla f(z)^T (s - r)$$

or

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

Referring to Taylor Theorem, there exists $\tau \in [0, 1]$ such that

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

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

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

Referring to Taylor theorem, there exists $\tau \in [0, 1]$ such that

$$f(y + \theta(x - y)) = f(y) + \nabla f \left[\tau(y + \theta(x - y)) + (1 - \tau)y \right]^T (y + \theta(x - y) - y)$$
$$f(y + \theta(x - y)) - f(y) = \theta \nabla f \left[y + \tau\theta(x - y) \right]^T (x - y). \quad (4.2)$$

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

Referring to Taylor theorem, there exists $\tau \in [0, 1]$ such that

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

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

Using relations (4.1) et (4.2), for all $\theta \in [0, 1]$

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

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

Referring to Taylor theorem, there exists $\tau \in [0, 1]$ such that

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

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

Using relations (4.1) et (4.2), for all $\theta \in [0, 1]$

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

If $\theta > 0$, then we can divide on both side by θ to obtain

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

for all $\theta \in (0, 1]$. Hence, taking the limit when $\theta \rightarrow 0$

we obtain

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

or

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

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

(sufficiency \Leftarrow) Now we show the converse.

)
)
)

(sufficiency \Leftarrow) Now we show the converse.

Since $f(x) \geq f(y) + \nabla f(y)^\top (x - y)$ for all pairs of points $x, y \in X$

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

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

(sufficiency \Leftarrow) Now we show the converse.

Since $f(x) \geq f(y) + \nabla f(y)^T (x - y)$ for all pairs of 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) can be written as

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

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

(sufficiency \Leftarrow) Now we show the converse.

Since $f(x) \geq f(y) + \nabla f(y)^T (x - y)$ for all pairs of 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) can be written as

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

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

Multiply (4.3) by θ and (4.4) by $(1 - \theta)$

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

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

(sufficiency \Leftarrow) Now we show the converse.

Since $f(x) \geq f(y) + \nabla f(y)^T (x - y)$ for all pairs of 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) can be written as

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

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

Multiply (4.3) by θ and (4.4) by $(1 - \theta)$

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

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

Summing up the two relations :

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

Corollary 6: Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable convex function of class C^1 / X . Then if $\nabla f(x^*)^T (x - x^*) \geq 0$ for all $x \in X$, then x^* is a global minimum of f on X .

Corollary 6: Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable convex function of class C^1 / X . Then if $\nabla f(x^*)^T (x - x^*) \geq 0$ for all $x \in X$, then x^* is a global minimum of f on X .

Proof. This result follows directly from the gradient inequality. Indeed, since f is convex on X , it follows that

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

for all $x \in X$. By assumption $\nabla f(x^*)^T (x - x^*) \geq 0$, and then

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

implying that $f(x) \geq f(x^*)$ for all $x \in X$. Thus x^* is a global minimum of f on X . □

Corollary 6: Let $X \subset \mathbb{R}^n$ be a convex set. Let $f : X \rightarrow \mathbb{R}^1$ be a differentiable convex function of class C^1 / X . Then if $\nabla f(x^*)^T (x - x^*) \geq 0$ for all $x \in X$, then x^* is a global minimum of f on X .

Proof. This result follows directly from the gradient inequality. Indeed, since f is convex on X , it follows that

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

for all $x \in X$. By assumption $\nabla f(x^*)^T (x - x^*) \geq 0$, and then

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

implying that $f(x) \geq f(x^*)$ for all $x \in X$. Thus x^* is a global minimum of f on X . □

Consequence: If $\nabla f(x^*) = 0$, then x^* is a global minimum of f on X .