II - Généralités sur l'Optimisation

Jean-Philippe Préaux

EOAA - 2009/10

Plan

Plan

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.
Nous passons en revue les outils de calcul différentiel pour la recherche d'extrema locaux d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.
Nous passons en revue les outils de calcul différentiel pour la recherche d'extrema locaux d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$.

3 - Programmation convexe.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.
Nous passons en revue les outils de calcul différentiel pour la recherche d'extrema locaux d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$.

3 - Programmation convexe.
Nous passons en revue les diverses notions de convexité et ce qu'elles apportent en optimisation.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.
Nous passons en revue les outils de calcul différentiel pour la recherche d'extrema locaux d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$.

3 - Programmation convexe.
Nous passons en revue les diverses notions de convexité et ce qu'elles apportent en optimisation.

4 - Programmation quadratique.

Plan

1 - Conditions suffisantes d'existence d'extrema.
Nous voyons deux conditions globales - compacité du domaine, coercivité de l'application - assurant de l'existence d'extrema globaux.

2 - Recherche d'extrema locaux.
Nous passons en revue les outils de calcul différentiel pour la recherche d'extrema locaux d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$.

3 - Programmation convexe.
Nous passons en revue les diverses notions de convexité et ce qu'elles apportent en optimisation.

4 - Programmation quadratique.
Nous appliquons ces notions au cas de la programmation quadratique sans contrainte.

Sur un domaine compact $\exists$ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Sur un domaine compact $\exists$ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Preuve. Sur un compact de $\mathbb{R}^n$ une application continue réelle est bornée et atteint ses bornes.

Sur un domaine compact $\exists$ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Preuve. Sur un compact de $\mathbb{R}^n$ une application continue réelle est bornée et atteint ses bornes.

$$\exists\ \mathbf{x}_{\min} \in \mathcal{K}, \quad f(\mathbf{x}_{min}) = \inf_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \min_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) .$$

Sur un domaine compact ∃ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Preuve. Sur un compact de $\mathbb{R}^n$ une application continue réelle est bornée et atteint ses bornes.

$$\exists\ \mathbf{x}_{\min} \in \mathcal{K}, \quad f(\mathbf{x}_{min}) = \inf_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \min_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

$$\exists\ \mathbf{x}_{\max} \in \mathcal{K}, \quad f(\mathbf{x}_{max}) = \sup_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \max_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

Sur un domaine compact ∃ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Preuve. Sur un compact de $\mathbb{R}^n$ une application continue réelle est bornée et atteint ses bornes.

$$\exists\ \mathbf{x}_{\text{min}} \in \mathcal{K}, \quad f(\mathbf{x}_{min}) = \inf_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \min_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

$$\exists\ \mathbf{x}_{\text{max}} \in \mathcal{K}, \quad f(\mathbf{x}_{max}) = \sup_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \max_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

□

Sur un domaine compact $\exists$ min & max.

Théorème (Existence d'extrema sur un domaine compact.)

Si $\mathcal{K}$ est un compact (i.e. est fermé et borné) de $\mathbb{R}^n$, et $f : \mathcal{K} \longrightarrow \mathbb{R}$ est continue, alors f admet un minimum ainsi qu'un maximum global sur $\mathcal{K}$.

Preuve. Sur un compact de $\mathbb{R}^n$ une application continue réelle est bornée et atteint ses bornes.

$$\exists\ \mathbf{x}_{\min} \in \mathcal{K}, \quad f(\mathbf{x}_{min}) = \inf_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \min_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

$$\exists\ \mathbf{x}_{\max} \in \mathcal{K}, \quad f(\mathbf{x}_{max}) = \sup_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) = \max_{\mathbf{x}\in\mathcal{K}} f(\mathbf{x}) \ .$$

$\square$

N'est utile qu'en optimisation sous contrainte : un ouvert non vide de $\mathbb{R}^n$ n'est jamais compact.

Applications coercives.

Applications coercives.

Définition. Une application $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ continue est *coercive* si $\mathcal{D}$ est un fermé non borné et si :

$$\lim_{\|\mathbf{x}\| \to +\infty} f(\mathbf{x}) = +\infty$$

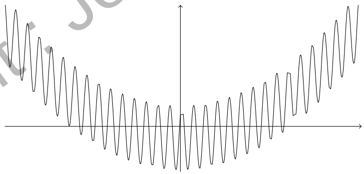
(souvent $\mathcal{D} = \mathbb{R}^n$).

Applications coercives.

Définition. Une application $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ continue est *coercive* si $\mathcal{D}$ est un fermé non borné et si :

$$\lim_{\|\mathbf{x}\| \to +\infty} f(\mathbf{x}) = +\infty$$

(souvent $\mathcal{D} = \mathbb{R}^n$).



Une application coercive $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global).

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R}$ $\Longrightarrow$ $\mathcal{K}$ est un fermé de $\mathbb{R}^n$.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies \mathcal{K}$ est un compact de $\mathbb{R}^n$.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies \mathcal{K}$ est un compact de $\mathbb{R}^n$. (Théorème précédent) $\implies$ f admet un min global $\mathbf{u}$ sur $\mathcal{K}$.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R}$ $\implies$ $\mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies$ $\mathcal{K}$ est un compact de $\mathbb{R}^n$. (Théorème précédent) $\implies$ f admet un min global $\mathbf{u}$ sur $\mathcal{K}$.
Par construction $\forall\, \mathbf{x} \in \mathcal{D} \setminus \mathcal{K},\ f(\mathbf{x}) \geqslant a \geqslant f(\mathbf{u})$

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies \mathcal{K}$ est un compact de $\mathbb{R}^n$. (Théorème précédent) $\implies f$ admet un min global $\mathbf{u}$ sur $\mathcal{K}$.
Par construction $\forall\, \mathbf{x} \in \mathcal{D} \setminus \mathcal{K},\ f(\mathbf{x}) \geqslant a \geqslant f(\mathbf{u}) \implies \forall\, \mathbf{x} \in \mathcal{D},\ f(\mathbf{x}) \geqslant f(\mathbf{u})$.

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies \mathcal{K}$ est un compact de $\mathbb{R}^n$. (Théorème précédent) $\implies f$ admet un min global $\mathbf{u}$ sur $\mathcal{K}$.
Par construction $\forall \mathbf{x} \in \mathcal{D} \setminus \mathcal{K},\ f(\mathbf{x}) \geqslant a \geqslant f(\mathbf{u}) \implies \forall \mathbf{x} \in \mathcal{D},\ f(\mathbf{x}) \geqslant f(\mathbf{u})$.
i.e. $\mathbf{u}$ est un min global de f sur $\mathcal{D}$. □

Une application coercive a un min et aucun max.

Théorème (Une application coercive a un minimum.)

Un fonction $f : \mathcal{D} \longrightarrow \mathbb{R}$ coercive admet un minimum global (et aucun maximum global). Si $-f$ est coercive, f admet un maximum global (et aucun minimum global).

Preuve. Soit $a \in \mathbb{R}$ suffisamment grand pour que $f^{-1}(]-\infty, a]) = \mathcal{K}$ soit non vide. f est continue et $]-\infty, a]$ fermé de $\mathbb{R} \implies \mathcal{K}$ est un fermé de $\mathbb{R}^n$.
De plus $\mathcal{K}$ est borné : sinon $\mathcal{K}$ contiendrait une suite $(\mathbf{x}_n)_{n\in\mathbb{N}}$ avec $\lim_{n\longrightarrow\infty} \|\mathbf{x}_n\| = +\infty$ et $\forall\, n \in \mathbb{N},\ f(x_n) \leqslant a$: contredirait f coercive.
$\implies \mathcal{K}$ est un compact de $\mathbb{R}^n$. (Théorème précédent) $\implies f$ admet un min global $\mathbf{u}$ sur $\mathcal{K}$.
Par construction $\forall\, \mathbf{x} \in \mathcal{D} \setminus \mathcal{K},\ f(\mathbf{x}) \geqslant a \geqslant f(\mathbf{u}) \implies \forall\, \mathbf{x} \in \mathcal{D},\ f(\mathbf{x}) \geqslant f(\mathbf{u})$.
i.e. $\mathbf{u}$ est un min global de f sur $\mathcal{D}$. □

Utile en programmation sous contrainte et sans contrainte ($\mathcal{D} = \mathbb{R}^n$).

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :

– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de $+$ haut degré est > 0,

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :

– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,

– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :

– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,

– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$

(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :
– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,
– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$
(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Conclusion. Une application polynomiale sur $\mathbb{R}$ admet :

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :
– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,
– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$
(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Conclusion. Une application polynomiale sur $\mathbb{R}$ admet :

Si son degré est pair :

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :
– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,
– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$
(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Conclusion. Une application polynomiale sur $\mathbb{R}$ admet :

Si son degré est pair :

• Si le coef. de son terme de plus haut degré est > 0 : un min global et aucun max global.

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :
– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de + haut degré est > 0,
– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$
(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Conclusion. Une application polynomiale sur $\mathbb{R}$ admet :

Si son degré est pair :

• Si le coef. de son terme de plus haut degré est > 0 : un min global et aucun max global.

• Si le coef. de son terme de plus haut degré est < 0 : un max global et aucun min global.

Exemple.

Une fonction polynomiale $f : \mathbb{R} \longrightarrow \mathbb{R}$ de degré > 0 :
– de degré pair est coercive sur $\mathbb{R}$ si et seulement si le coef. de son terme de $+$ haut degré est > 0,
– de degré impair (de coef. du terme de plus haut degré α) n'est jamais coercive sur $\mathbb{R}$
(est coercive sur $[c, +\infty[$ si $\alpha > 0$ et sur $]-\infty, c]$ si $\alpha < 0$.)

Conclusion. Une application polynomiale sur $\mathbb{R}$ admet :

Si son degré est pair :

• Si le coef. de son terme de plus haut degré est > 0 : un min global et aucun max global.

• Si le coef. de son terme de plus haut degré est < 0 : un max global et aucun min global.

Si son degré est impair : ni minimum ni maximum global.

Recherche d'extrema locaux.

Recherche d'extrema locaux.

- Ne s'applique qu'à une fonction différentiable (1 ou 2 fois)

Recherche d'extrema locaux.

- Ne s'applique qu'à une fonction différentiable (1 ou 2 fois)
- Ne détermine que les extrema locaux dans l'intérieur du domaine – ou sur un domaine ouvert–.

Recherche d'extrema locaux.

- Ne s'applique qu'à une fonction différentiable (1 ou 2 fois)
- Ne détermine que les extrema locaux dans l'intérieur du domaine

– ou sur un domaine ouvert–.

Plan.

Recherche d'extrema locaux.

- Ne s'applique qu'à une fonction différentiable (1 ou 2 fois)
- Ne détermine que les extrema locaux dans l'intérieur du domaine – ou sur un domaine ouvert–.

Plan.

1. Condition du 1^{er} ordre.

condition nécessaire portant sur les dérivées premières.

Recherche d'extrema locaux.

- Ne s'applique qu'à une fonction différentiable (1 ou 2 fois)
- Ne détermine que les extrema locaux dans l'intérieur du domaine – ou sur un domaine ouvert–.

Plan.

1. Condition du 1^{er} ordre.

condition nécessaire portant sur les dérivées premières.

2. Conditions du 2^{nd} ordre.

Condition nécessaire, condition suffisante, portant sur les dérivées secondes.

Vecteur gradient

Soit $\mathcal{U}$ est un ouvert de $\mathbb{R}^n$, $\mathbf{x}_0 \in \mathbb{R}^n$ et $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable en $\mathbf{x}_0$ on note :

Vecteur gradient

Soit $\mathcal{U}$ est un ouvert de $\mathbb{R}^n$, $\mathbf{x}_0 \in \mathbb{R}^n$ et $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable en $\mathbf{x}_0$ on note :

$$\nabla f(\mathbf{x}_0) \triangleq \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \in \mathbb{R}^n$$

le *vecteur gradient de f en* $\mathbf{x}_0$ (on prononce "nabla f de $\mathbf{x}_0$").

Vecteur gradient

Soit $\mathcal{U}$ est un ouvert de $\mathbb{R}^n$, $\mathbf{x}_0 \in \mathbb{R}^n$ et $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable en $\mathbf{x}_0$ on note :

$$\nabla f(\mathbf{x}_0) \triangleq \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \in \mathbb{R}^n$$

le *vecteur gradient de f en $\mathbf{x}_0$* (on prononce "nabla f de $\mathbf{x}_0$").
On a alors le
développement de Taylor-Young de f à l'ordre 1 au voisinage de $\mathbf{x}_0$:

$$\boxed{f(\mathbf{x}) = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|)\ .}$$

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *différentiable en* $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \textit{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, différentiable en $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \text{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|)\ .$$

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, différentiable en $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \text{ extremum local de } f \quad \Longrightarrow \quad \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|) .$$

• Si $\nabla f(x_0) \neq \mathbf{0}$: pour $\mathbf{x}$ suffisamment proche de $\mathbf{x}_0$, $o(\|\mathbf{x} - \mathbf{x}_0\|)$ est négligeable devant $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *différentiable en* $\mathbf{x}_0 \in \operatorname{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \text{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|) \ .$$

• Si $\nabla f(x_0) \neq \mathbf{0}$: pour $\mathbf{x}$ suffisamment proche de $\mathbf{x}_0$, $o(\|\mathbf{x} - \mathbf{x}_0\|)$ est négligeable devant $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$

$$\implies \quad f(\mathbf{x}) - f(\mathbf{x}_0) \text{ a le signe de } \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$$

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *différentiable en* $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \textit{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|) \ .$$

• Si $\nabla f(x_0) \neq \mathbf{0}$: pour $\mathbf{x}$ suffisamment proche de $\mathbf{x}_0$, $o(\|\mathbf{x} - \mathbf{x}_0\|)$ est négligeable devant $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$

$$\implies \quad f(\mathbf{x}) - f(\mathbf{x}_0) \text{ a le signe de } \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$$

Or quand $\mathbf{x}$ décrit un voisinage de $\mathbf{x}_0$, par bilinéarité du produit scalaire $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$ ne garde pas un signe constant.

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, différentiable en $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \textit{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|) \ .$$

• Si $\nabla f(x_0) \neq \mathbf{0}$: pour $\mathbf{x}$ suffisamment proche de $\mathbf{x}_0$, $o(\|\mathbf{x} - \mathbf{x}_0\|)$ est négligeable devant $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$

$$\implies \quad f(\mathbf{x}) - f(\mathbf{x}_0) \text{ a le signe de } \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$$

Or quand $\mathbf{x}$ décrit un voisinage de $\mathbf{x}_0$, par bilinéarité du produit scalaire $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$ ne garde pas un signe constant. Contradiction.

Equation d'Euler

Théorème (Equation d'Euler)

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, différentiable en $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{D})$.

$$\mathbf{x}_0 \textit{ extremum local de } f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}.$$

Preuve. On applique en $\mathbf{x}_0$ le développement de Taylor-Young (ordre 1) :

$$\underbrace{f(\mathbf{x}) - f(\mathbf{x}_0)}_{\text{de signe constant}} = \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + o(\|\mathbf{x} - \mathbf{x}_0\|) \ .$$

• Si $\nabla f(x_0) \neq \mathbf{0}$: pour $\mathbf{x}$ suffisamment proche de $\mathbf{x}_0$, $o(\|\mathbf{x} - \mathbf{x}_0\|)$ est négligeable devant $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$

$$\implies \quad f(\mathbf{x}) - f(\mathbf{x}_0) \text{ a le signe de } \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$$

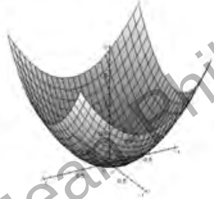
Or quand $\mathbf{x}$ décrit un voisinage de $\mathbf{x}_0$, par bilinéarité du produit scalaire $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$ ne garde pas un signe constant. Contradiction. $\nabla f(\mathbf{x}_0) = \mathbf{0}$. $\square$

$\mathbf{x}_0$ *est appelé* **point critique** *ou* **point stationnaire**.

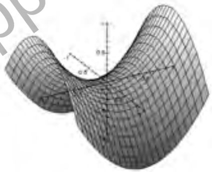
$\mathbf{x}_0$ *est appelé* **point critique** *ou* **point stationnaire**.

L'équation d'Euler est une condition nécessaire non suffisante.

max local



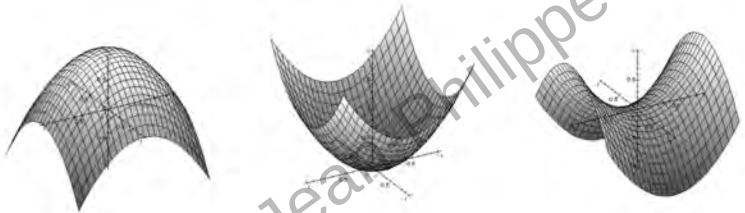
min local



point-selle

$\mathbf{x}_0$ *est appelé* **point critique** *ou* **point stationnaire**.

L'équation d'Euler est une condition nécessaire non suffisante.



max local	min local	point-selle
$1-x^2-y^2$	x^2+y^2	x^2-y^2

Matrice hessienne :

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Matrice hessienne

Si $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ est 2 fois différentiable en $\mathbf{x}_0$ on note :

Matrice hessienne

Si $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ est 2 fois différentiable en $\mathbf{x}_0$ on note :

$$\nabla^2 f(\mathbf{x}_0) \triangleq \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

la *matrice Hessienne de f en $\mathbf{x}_0$*.

Matrice hessienne

Si $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ est 2 fois différentiable en $\mathbf{x}_0$ on note :

$$\nabla^2 f(\mathbf{x}_0) \triangleq \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

la *matrice Hessienne de f en $\mathbf{x}_0$*.

C'est une matrice symétrique ; en particulier elle est diagonalisable.

Matrice hessienne

Si $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ est 2 fois différentiable en $\mathbf{x}_0$ on note :

$$\nabla^2 f(\mathbf{x}_0) \triangleq \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

la *matrice Hessienne de f en $\mathbf{x}_0$*.
C'est une matrice symétrique ; en particulier elle est diagonalisable.

On a le développement de Taylor-Young à l'ordre 2 de f en $\mathbf{x}_0$:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^\top \nabla^2 f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^2)$$

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *semi-définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\top A\mathbf{x} \geqslant 0$.

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *semi-définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\top A\mathbf{x} \geqslant 0$.
- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\mathbf{x}^\top A\mathbf{x} > 0$.

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *semi-définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\top A\mathbf{x} \geqslant 0$.
- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\mathbf{x}^\top A\mathbf{x} > 0$.
- Et de façon analogue *(semi)-définie négative*.

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *semi-définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\top A\mathbf{x} \geqslant 0$.
- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\mathbf{x}^\top A\mathbf{x} > 0$.
- Et de façon analogue *(semi)-définie négative*.

Proposition

Si $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable, alors :
– A est semi-définie positive ssi toutes ses valeurs propres sont $\geqslant 0$.
– A est définie positive ssi toutes ses valeurs propres sont > 0.

Matrice définie positive, semi-définie positive, *etc...*

Définitions.

- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *semi-définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\top A\mathbf{x} \geqslant 0$.
- Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *définie positive* si $\forall\, \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\mathbf{x}^\top A\mathbf{x} > 0$.
- Et de façon analogue *(semi)-définie négative*.

Proposition

Si $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable, alors :
– A est semi-définie positive ssi toutes ses valeurs propres sont $\geqslant 0$.
– A est définie positive ssi toutes ses valeurs propres sont > 0.

Théorème

$A \in \mathcal{M}_n(\mathbb{R})$ symétrique $\implies$ A diagonalisable.

Conditions du 2^{e} ordre

Théorème

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *et* $\mathbf{u} \in \mathrm{Int}(\mathcal{D})$, *avec* f *2 fois différentiable en* $\mathbf{u}$*.*

Conditions du 2e ordre

Théorème

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *et* $\mathbf{u} \in \mathrm{Int}(\mathcal{D})$, *avec* f *2 fois différentiable en* $\mathbf{u}$.

1. **(Condition nécessaire du 2e ordre.)**
   Si $\mathbf{u}$ *est un minimum local de* f, *alors* $\nabla f(\mathbf{u}) = \mathbf{0}$ *et* $\nabla^2 f(\mathbf{u})$ *est semi-définie positive.*

Conditions du 2^e ordre

Théorème

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, et $\mathbf{u} \in \mathrm{Int}(\mathcal{D})$, avec f 2 fois différentiable en $\mathbf{u}$.

1. **(Condition nécessaire du 2^e ordre.)**
   Si $\mathbf{u}$ est un minimum local de f, alors $\nabla f(\mathbf{u}) = \mathbf{0}$ et $\nabla^2 f(\mathbf{u})$ est semi-définie positive.
2. **(Condition suffisante du 2^e ordre.)**
   Si $\nabla f(\mathbf{u}) = \mathbf{0}$ et $\nabla^2 f(\mathbf{u})$ est définie positive, alors $\mathbf{u}$ est un minimum local strict de f.

Conditions du 2^e ordre

Théorème

Soit $f : \mathcal{D} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *et* $\mathbf{u} \in \mathrm{Int}(\mathcal{D})$, *avec* f *2 fois différentiable en* $\mathbf{u}$.

1. **(Condition nécessaire du 2^e ordre.)**
   Si $\mathbf{u}$ *est un minimum local de* f, *alors* $\nabla f(\mathbf{u}) = \mathbf{0}$ *et* $\nabla^2 f(\mathbf{u})$ *est semi-définie positive.*
2. **(Condition suffisante du 2^e ordre.)**
   Si $\nabla f(\mathbf{u}) = \mathbf{0}$ *et* $\nabla^2 f(\mathbf{u})$ *est définie positive, alors* $\mathbf{u}$ *est un minimum local strict de* f.

• La condition suffisante du 2^e ordre s'utilise pour montrer qu'un point critique est un extremum local.

• La condition nécessaire du 2^e ordre s'utilise pour montrer qu'un point critique n'est pas un extremum local.

Preuve.
1.

Preuve.

1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler).

Preuve.

1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^{\top}\,\nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Preuve.

1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.
Alors $\mathbf{y}^\top \nabla^2 f(\mathbf{u})\,\mathbf{y} = k^2 \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0 \implies \nabla^2 f(\mathbf{u})$ est semi-définie positive.

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.
Alors $\mathbf{y}^\top \nabla^2 f(\mathbf{u})\,\mathbf{y} = k^2 \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0 \implies \nabla^2 f(\mathbf{u})$ est semi-définie positive.

2.

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.
Alors $\mathbf{y}^\top \nabla^2 f(\mathbf{u})\,\mathbf{y} = k^2 \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0 \Longrightarrow \nabla^2 f(\mathbf{u})$ est semi-définie positive.

2. Puisque $\nabla f(\mathbf{u}) = \mathbf{0}$ et $\nabla^2 f(\mathbf{u})$ est définie positive la formule de Taylor-Young à l'ordre 2 s'écrit :

$$f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u}) = \underbrace{\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x}}_{>0} + o(\|\mathbf{x}\|^2)$$

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.
Alors $\mathbf{y}^\top \nabla^2 f(\mathbf{u})\,\mathbf{y} = k^2 \mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x} \geqslant 0 \implies \nabla^2 f(\mathbf{u})$ est semi-définie positive.

2.Puisque $\nabla f(\mathbf{u}) = \mathbf{0}$ et $\nabla^2 f(\mathbf{u})$ est définie positive la formule de Taylor-Young à l'ordre 2 s'écrit :

$$f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u}) = \underbrace{\mathbf{x}^\top \nabla^2 f(\mathbf{u})\,\mathbf{x}}_{>0} + o(\|\mathbf{x}\|^2)$$

Il existe un ouvert de $\mathbb{R}^n$ contenant $\mathbf{u}$ sur lequel $o(\|\mathbf{x}\|^2)$ est négligeable

Preuve.
1. Puisque $\mathbf{u}$ est un extremum local alors $\nabla f(\mathbf{u}) = \mathbf{0}$ (équation d'Euler). La formule de Taylor-Young à l'ordre 2 s'écrit :

$$\underbrace{f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u})}_{\geqslant 0} = \mathbf{x}^\top \nabla^2 f(\mathbf{u})\, \mathbf{x} + o(\|\mathbf{x}\|^2)$$

Il existe une boule ouverte $B(\mathbf{0}, r)$ dans $\mathcal{D}$ sur laquelle $o(\|\mathbf{x}\|^2)$ est négligeable $\Longrightarrow$ $\mathbf{x}^\top \nabla^2 f(\mathbf{u})\, \mathbf{x} \geqslant 0$.
Soit $\mathbf{y} \in \mathbb{R}^n$; alors $\exists\, k > 0$, $\mathbf{x} = \mathbf{y}/k$ est dans $B(\mathbf{0}, r)$.
Alors $\mathbf{y}^\top \nabla^2 f(\mathbf{u})\, \mathbf{y} = k^2 \mathbf{x}^\top \nabla^2 f(\mathbf{u})\, \mathbf{x} \geqslant 0 \implies \nabla^2 f(\mathbf{u})$ est semi-définie positive.
2.Puisque $\nabla f(\mathbf{u}) = \mathbf{0}$ et $\nabla^2 f(\mathbf{u})$ est définie positive la formule de Taylor-Young à l'ordre 2 s'écrit :

$$f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u}) = \underbrace{\mathbf{x}^\top \nabla^2 f(\mathbf{u})\, \mathbf{x}}_{>0} + o(\|\mathbf{x}\|^2)$$

Il existe un ouvert de $\mathbb{R}^n$ contenant $\mathbf{u}$ sur lequel $o(\|\mathbf{x}\|^2)$ est négligeable
$\Longrightarrow$ $f(\mathbf{x}+\mathbf{u}) - f(\mathbf{u}) \geqslant 0$: $\mathbf{u}$ est donc un minimum local (strict) de f. □

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

(iii) *$\forall\, i = 0, \ldots, n$, $c_i > 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

(iii) *$\forall\, i = 0, \ldots, n,\ c_i > 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les déterminants $\det(A_k)$ où A_k désigne $A_k = (a_{ij})_{i,j=1..k}$ sont tous > 0.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

(iii) *$\forall\, i = 0, \dots, n$, $c_i > 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les déterminants $\det(A_k)$ où A_k désigne $A_k = (a_{ij})_{i,j=1..k}$ sont tous > 0.*

(v) *Il existe une matrice M inversible tel que $M^\top M = A$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

(iii) *$\forall\, i = 0, \dots, n$, $c_i > 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les déterminants $\det(A_k)$ où A_k désigne $A_k = (a_{ij})_{i,j=1..k}$ sont tous > 0.*

(v) *Il existe une matrice M inversible tel que $M^\top M = A$.*

De plus :

(a) *Si A est définie positive alors $\forall\, i = 1, .., n$, $a_{ii} > 0$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est définie positive.*

(ii) *Toutes les valeurs propres de A sont > 0.*

(iii) *$\forall\, i = 0, \ldots, n$, $c_i > 0$, où $p_A(\lambda) = \sum_{i=0}^{n}(-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les déterminants $\det(A_k)$ où A_k désigne $A_k = (a_{ij})_{i,j=1..k}$ sont tous > 0.*

(v) *Il existe une matrice M inversible tel que $M^\top M = A$.*

De plus :

(a) *Si A est définie positive alors $\forall\, i = 1, .., n$, $a_{ii} > 0$.*

(b) *Si $A \in \mathcal{M}_2(\mathbb{R})$, A est définie positive si et seulement si $\det(A) > 0$ et $tr(A) > 0$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) *$\forall\, i = 0, \ldots, n,\ c_i \geqslant 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) *$\forall\, i = 0, \dots, n$, $c_i \geqslant 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les mineurs principaux de A sont tous $\geqslant 0$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) *$\forall\, i = 0, \ldots, n$, $c_i \geqslant 0$, où $p_A(\lambda) = \sum_{i=0}^{n}(-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les mineurs principaux de A sont tous $\geqslant 0$.*

(v) *Il existe une matrice M tel que $M^\top M = A$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) *$\forall\, i = 0, \ldots, n,\ c_i \geqslant 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les mineurs principaux de A sont tous $\geqslant 0$.*

(v) *Il existe une matrice M tel que $M^\top M = A$.*

De plus :

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) $\forall\, i = 0, \ldots, n$, $c_i \geqslant 0$, *où* $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ *est le polynôme caractéristique de A.*

(iv) *Les mineurs principaux de A sont tous $\geqslant 0$.*

(v) *Il existe une matrice M tel que $M^\top M = A$.*

De plus :

(a) *Si A est semi-définie positive alors $\forall\, i = 1, .., n$, $a_{ii} \geqslant 0$.*

Théorème

Soit $A = (a_{ij})_{i,j=1..n}$ une matrice symétrique. Les assertions suivantes sont équivalentes :

(i) *A est semi-définie positive.*

(ii) *Toutes les valeurs propres de A sont $\geqslant 0$.*

(iii) *$\forall\, i = 0, \ldots, n$, $c_i \geqslant 0$, où $p_A(\lambda) = \sum_{i=0}^{n} (-1)^i c_i \lambda_i^n$ est le polynôme caractéristique de A.*

(iv) *Les mineurs principaux de A sont tous $\geqslant 0$.*

(v) *Il existe une matrice M tel que $M^\top M = A$.*

De plus :

(a) *Si A est semi-définie positive alors $\forall\, i = 1, .., n$, $a_{ii} \geqslant 0$.*

(b) *Si $A \in \mathcal{M}_2(\mathbb{R})$, A est définie positive si et seulement si $\det(A) \geqslant 0$ et $tr(A) \geqslant 0$.*

Exemple.

Soit l'application f de classe C^∞ (*i.e.* infiniment différentiable) :

$$\begin{aligned} f : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\longrightarrow f(x,y) = x^3 + y^3 - 9xy \end{aligned}$$

Exemple.

Soit l'application f de classe C^∞ (*i.e.* infiniment différentiable) :

$$\begin{aligned} f: \quad \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\longrightarrow f(x,y) = x^3 + y^3 - 9xy \end{aligned}$$

Son vecteur gradient en un point (x, y) est :

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 9y \\ 3y^2 - 9x \end{pmatrix},$$

Exemple.

Soit l'application f de classe C^∞ (*i.e.* infiniment différentiable) :

$$f : \begin{array}{rcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longrightarrow & f(x,y) = x^3 + y^3 - 9xy \end{array}$$

Son vecteur gradient en un point (x,y) est :

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 9y \\ 3y^2 - 9x \end{pmatrix} \; ,$$

et sa matrice Hessienne :

$$\nabla^2 f(x,y) = \begin{pmatrix} 6x & -9 \\ -9 & 6y \end{pmatrix}$$

Les points critiques, solutions de $\nabla f(x, y) = 0$ sont les 2 points $(0, 0)$ et $(3, 3)$.

Les points critiques, solutions de $\nabla f(x,y) = 0$ sont les 2 points $(0,0)$ et $(3,3)$. En ces points, les matrices Hessiennes sont :

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix} \qquad \nabla^2 f(3,3) = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$$

Les points critiques, solutions de $\nabla f(x,y) = 0$ sont les 2 points $(0,0)$ et $(3,3)$. En ces points, les matrices Hessiennes sont :

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix} \qquad \nabla^2 f(3,3) = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$$

• $\nabla^2 f(0,0)$ a une trace nulle et un déterminant strictement négatif, elle n'est donc ni semi-définie positive, ni semi-définie négative : $(0,0)$ n'est pas un extremum local.

Les points critiques, solutions de $\nabla f(x,y)=0$ sont les 2 points $(0,0)$ et $(3,3)$. En ces points, les matrices Hessiennes sont :

$$\nabla^2 f(0,0)=\begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix} \qquad \nabla^2 f(3,3)=\begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$$

• $\nabla^2 f(0,0)$ a une trace nulle et un déterminant strictement négatif, elle n'est donc ni semi-définie positive, ni semi-définie négative : $(0,0)$ n'est pas un extremum local.
• $\nabla^2 f(3,3)$ a une trace et un déterminant strictement positifs : $(3,3)$ est un minimum local.

Les points critiques, solutions de $\nabla f(x,y)=0$ sont les 2 points $(0,0)$ et $(3,3)$. En ces points, les matrices Hessiennes sont :

$$\nabla^2 f(0,0)=\begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix} \qquad \nabla^2 f(3,3)=\begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$$

• $\nabla^2 f(0,0)$ a une trace nulle et un déterminant strictement négatif, elle n'est donc ni semi-définie positive, ni semi-définie négative : $(0,0)$ n'est pas un extremum local.
• $\nabla^2 f(3,3)$ a une trace et un déterminant strictement positifs : $(3,3)$ est un minimum local.
• f n'admet aucun extremum global puisque :

$$\lim_{x\to+\infty} f(x,0)=+\infty \qquad \lim_{x\to-\infty} f(x,0)=-\infty\ .$$

Programmation convexe.

Plan.

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

1.1. Définitions - Propriétés

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

1.1. Définitions - Propriétés

1.2. Caractérisations à l'ordre 1 et 2

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

1.1. Définitions - Propriétés

1.2. Caractérisations à l'ordre 1 et 2

1.3. Programmation convexe

Programmation convexe.

Plan.

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

1.1. Définitions - Propriétés

1.2. Caractérisations à l'ordre 1 et 2

1.3. Programmation convexe

2. Applications elliptiques (ou fortement convexes)

2.1. Définition - Propriétés

Programmation convexe.

Plan.

1. Applications convexes, strictement convexe

1.1. Définitions - Propriétés

1.2. Caractérisations à l'ordre 1 et 2

1.3. Programmation convexe

2. Applications elliptiques (ou fortement convexes)

2.1. Définition - Propriétés

2.2. Caractérisation à l'ordre 2

Programmation convexe.

Plan.

Ensemble convexe - Définition

Définition. Un sous-ensemble $\mathcal{C}$ de $\mathbb{R}^n$ est *convexe* si

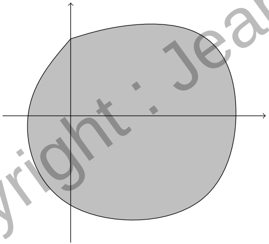
$$\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \forall\, t \in [0,1],\ t\mathbf{x} + (1-t)\mathbf{y} \in \mathcal{C}$$

Ensemble convexe - Définition

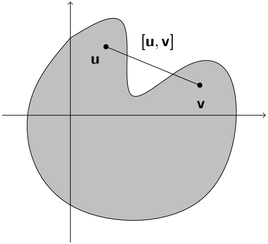
Définition. Un sous-ensemble $\mathcal{C}$ de $\mathbb{R}^n$ est *convexe* si

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \ \forall\, t \in [0,1],\ \ t\mathbf{x} + (1-t)\mathbf{y} \in \mathcal{C}$$

i.e. $\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C}$ le segment $[\mathbf{x}, \mathbf{y}]$ est inclus dans $\mathcal{C}$.



convexe



non-convexe

Ensemble convexe - Propriétés

Propriétés.

Ensemble convexe - Propriétés

Propriétés.

– Tout sous-espace affine de $\mathbb{R}^n$ (en particulier $\mathbb{R}^n$) est convexe.

Ensemble convexe - Propriétés

Propriétés.

– Tout sous-espace affine de $\mathbb{R}^n$ (en particulier $\mathbb{R}^n$) est convexe.

– Toute boule de $\mathbb{R}^n$, ouverte ou fermée, est un convexe de $\mathbb{R}^n$.

Ensemble convexe - Propriétés

Propriétés.

– Tout sous-espace affine de $\mathbb{R}^n$ (en particulier $\mathbb{R}^n$) est convexe.

– Toute boule de $\mathbb{R}^n$, ouverte ou fermée, est un convexe de $\mathbb{R}^n$.

– L'intersection de convexes de $\mathbb{R}^n$ est un convexe de $\mathbb{R}^n$.

Ensemble convexe - Propriétés

Propriétés.

– Tout sous-espace affine de $\mathbb{R}^n$ (en particulier $\mathbb{R}^n$) est convexe.

– Toute boule de $\mathbb{R}^n$, ouverte ou fermée, est un convexe de $\mathbb{R}^n$.

– L'intersection de convexes de $\mathbb{R}^n$ est un convexe de $\mathbb{R}^n$.

– Si $\mathcal{C}_1$, $\mathcal{C}_2$ sont deux convexes de $\mathbb{R}^n$, et $\lambda \in \mathbb{R}$, alors $\mathcal{C}_1 + \mathcal{C}_2$ et $\lambda\mathcal{C}_1$ sont des convexes de $\mathbb{R}^n$.

Ensemble convexe - Propriétés

Propriétés.

– Tout sous-espace affine de $\mathbb{R}^n$ (en particulier $\mathbb{R}^n$) est convexe.

– Toute boule de $\mathbb{R}^n$, ouverte ou fermée, est un convexe de $\mathbb{R}^n$.

– L'intersection de convexes de $\mathbb{R}^n$ est un convexe de $\mathbb{R}^n$.

– Si $\mathcal{C}_1$, $\mathcal{C}_2$ sont deux convexes de $\mathbb{R}^n$, et $\lambda \in \mathbb{R}$, alors $\mathcal{C}_1 + \mathcal{C}_2$ et $\lambda\mathcal{C}_1$ sont des convexes de $\mathbb{R}^n$.

– Si $\mathcal{C}_1$ est un convexe de $\mathbb{R}^p$ et $\mathcal{C}_2$ est un convexe de $\mathbb{R}^q$, leur produit cartésien $\mathcal{C}_1 \times \mathcal{C}_2 = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^p \times \mathbb{R}^q \,|\, \mathbf{x} \in \mathcal{C}_1,\, \mathbf{y} \in \mathcal{C}_2\}$ est un convexe de $\mathbb{R}^p \times \mathbb{R}^q \approx \mathbb{R}^{p+q}$.

Applications convexes - Définitions

Définitions. Soit $\mathcal{C} \subset \mathbb{R}^n$ un ensemble convexe non vide et $f : \mathcal{C} \longrightarrow \mathbb{R}$.

Applications convexes - Définitions

Définitions. Soit $\mathcal{C} \subset \mathbb{R}^n$ un ensemble convexe non vide et $f : \mathcal{C} \longrightarrow \mathbb{R}$.

- L'application f est *convexe* si :

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{C}, \ \forall t \in [0,1], \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \geqslant f(t\mathbf{x} + (1-t)\mathbf{y})$$

Applications convexes - Définitions

Définitions. Soit $\mathcal{C} \subset \mathbb{R}^n$ un ensemble convexe non vide et $f : \mathcal{C} \longrightarrow \mathbb{R}$.

- L'application f est *convexe* si :

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \forall\, t \in [0,1], \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \geqslant f(t\mathbf{x} + (1-t)\mathbf{y})$$

(*i.e.* dans $\mathbb{R}^{n+1}$ le segment joignant $(\mathbf{x}, f(\mathbf{x}))$ et $(\mathbf{y}, f(\mathbf{y}))$ reste au-dessus de la nappe représentative de la fonction.)

Applications convexes - Définitions

Définitions. Soit $\mathcal{C} \subset \mathbb{R}^n$ un ensemble convexe non vide et $f : \mathcal{C} \longrightarrow \mathbb{R}$.

- L'application f est *convexe* si :

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \forall\, t \in [0,1], \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \geqslant f(t\mathbf{x} + (1-t)\mathbf{y})$$

(*i.e.* dans $\mathbb{R}^{n+1}$ le segment joignant $(\mathbf{x}, f(\mathbf{x}))$ et $(\mathbf{y}, f(\mathbf{y}))$ reste au-dessus de la nappe représentative de la fonction.)

- L'application f est *strictement convexe* si :

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y},\ \forall\, t \in]0,1[, \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) > f(t\mathbf{x} + (1-t)\mathbf{y})$$

Applications convexes - Définitions

Définitions. Soit $\mathcal{C} \subset \mathbb{R}^n$ un ensemble convexe non vide et $f : \mathcal{C} \longrightarrow \mathbb{R}$.

- L'application f est *convexe* si :

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \forall t \in [0,1], \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \geqslant f(t\mathbf{x} + (1-t)\mathbf{y})$$

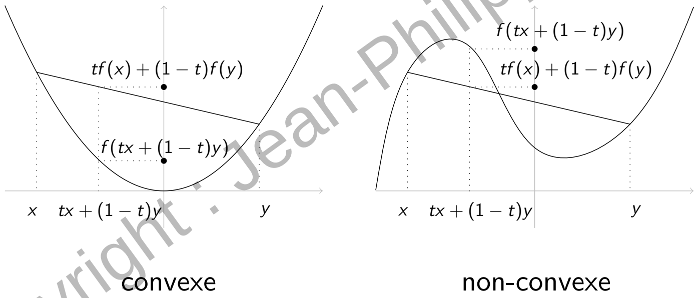
(*i.e.* dans $\mathbb{R}^{n+1}$ le segment joignant $(\mathbf{x}, f(\mathbf{x}))$ et $(\mathbf{y}, f(\mathbf{y}))$ reste au-dessus de la nappe représentative de la fonction.)

- L'application f est *strictement convexe* si :

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y},\ \forall t \in]0,1[, \quad tf(\mathbf{x}) + (1-t)f(\mathbf{y}) > f(t\mathbf{x} + (1-t)\mathbf{y})$$

(*i.e.* dans $\mathbb{R}^{n+1}$ le segment joignant $(\mathbf{x}, f(\mathbf{x}))$ et $(\mathbf{y}, f(\mathbf{y}))$ reste strictement au dessus de la nappe représentative de la fonction.)

Applications convexes



Applications convexes - Propriétés

Propriétés.

Applications convexes - Propriétés

Propriétés.

– Toute application affine, définie sur un convexe, est convexe et non strictement convexe.

Applications convexes - Propriétés

Propriétés.

– Toute application affine, définie sur un convexe, est convexe et non strictement convexe.

– La somme d'application (resp. strictement) convexes est (resp. strictement) convexe.

Applications convexes - Propriétés

Propriétés.

– Toute application affine, définie sur un convexe, est convexe et non strictement convexe.

– La somme d'application (resp. strictement) convexes est (resp. strictement) convexe.

– Si f est (resp. strictement) convexe et $\lambda \in \mathbb{R}_+$ (resp. $\lambda \in \mathbb{R}_+^*$) alors λf est (resp. strictement) convexe.

Applications convexes - Propriétés

Propriétés.

– Toute application affine, définie sur un convexe, est convexe et non strictement convexe.

– La somme d'application (resp. strictement) convexes est (resp. strictement) convexe.

– Si f est (resp. strictement) convexe et $\lambda \in \mathbb{R}_+$ (resp. $\lambda \in \mathbb{R}_+^*$) alors λf est (resp. strictement) convexe.

– Si f est (resp. strictement) convexe et $a, b \in \mathbb{R}$, $a \neq 0$, alors l'application $\mathbf{x} \longrightarrow f(a\mathbf{x} + b)$ est (resp. strictement) convexe.

Applications convexes - Propriétés

Propriétés.

– Toute application affine, définie sur un convexe, est convexe et non strictement convexe.

– La somme d'application (resp. strictement) convexes est (resp. strictement) convexe.

– Si f est (resp. strictement) convexe et $\lambda \in \mathbb{R}_+$ (resp. $\lambda \in \mathbb{R}_+^*$) alors λf est (resp. strictement) convexe.

– Si f est (resp. strictement) convexe et $a, b \in \mathbb{R}$, $a \neq 0$, alors l'application $\mathbf{x} \longrightarrow f(a\mathbf{x} + b)$ est (resp. strictement) convexe.

– Une application convexe sur $\mathcal{C}$ est continue en tout point de $\mathrm{Int}(\mathcal{C})$.

Caractérisation de la convexité.

Théorème (Caractérisations de la convexité.)

Soit $\mathcal{U}$ *un ouvert convexe de* $\mathbb{R}^n$ *et* $f : \mathcal{U} \longrightarrow \mathbb{R}$.

Caractérisation de la convexité.

Théorème (Caractérisations de la convexité.)

Soit $\mathcal{U}$ un ouvert convexe de $\mathbb{R}^n$ et $f : \mathcal{U} \longrightarrow \mathbb{R}$.

1. *Si f est différentiable sur $\mathcal{U}$, alors*
   - **a.** $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U},\ f(\mathbf{y}) \geqslant f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff f$ *est convexe sur $\mathcal{U}$,*
   - **b.** $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U},\ \mathbf{x} \neq \mathbf{y},\ f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff f$ *est strictement convexe sur $\mathcal{U}$.*

Caractérisation de la convexité.

Théorème (Caractérisations de la convexité.)

Soit $\mathcal{U}$ un ouvert convexe de $\mathbb{R}^n$ et $f : \mathcal{U} \longrightarrow \mathbb{R}$.

1. *Si f est différentiable sur $\mathcal{U}$, alors*
   - a. $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff$ *f est convexe sur $\mathcal{U}$,*
   - b. $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U}$, $\mathbf{x} \neq \mathbf{y}$, $f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff$ *f est strictement convexe sur $\mathcal{U}$.*

Géométriquement : en tout point la nappe représentative de f est au dessus de son hyperplan tangent.

Caractérisation de la convexité.

Théorème (Caractérisations de la convexité.)

Soit $\mathcal{U}$ un ouvert convexe de $\mathbb{R}^n$ et $f : \mathcal{U} \longrightarrow \mathbb{R}$.

1. *Si f est différentiable sur $\mathcal{U}$, alors*
   - **a.** $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff$ *f est convexe sur $\mathcal{U}$,*
   - **b.** $\forall \mathbf{x}, \mathbf{y} \in \mathcal{U}$, $\mathbf{x} \neq \mathbf{y}$, $f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \iff$ *f est strictement convexe sur $\mathcal{U}$.*

   Géométriquement : en tout point la nappe représentative de f est au dessus de son hyperplan tangent.
2. *Si f est 2 fois différentiable sur $\mathcal{U}$, alors*
   - **a.** $\forall \mathbf{x} \in \mathcal{U}$, $\nabla^2 f(\mathbf{x})$ *est semi-définie positive* $\iff$ *f est convexe sur $\mathcal{U}$,*
   - **b.** $\forall \mathbf{x} \in \mathcal{U}$, $\nabla^2 f(\mathbf{x})$ *définie positive* $\implies$ *f est strictement convexe.*

Programmation convexe.

On parle de programmation convexe lorsque :

Programmation convexe.

On parle de programmation convexe lorsque :

– $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une application convexe, à optimiser sur

$$\mathcal{D} = \{\mathbf{x} \in \mathcal{U} \subset \mathbb{R}^n \,|\, \varphi_i(\mathbf{x}) = 0,\ \forall\, i = 1, \ldots, p,\\ \psi_j(\mathbf{x}) \leqslant 0,\ \forall\, j = 1, \ldots, q\}$$

où :

Programmation convexe.

On parle de programmation convexe lorsque :

– $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une application convexe, à optimiser sur

$$\mathcal{D} = \{\mathbf{x} \in \mathcal{U} \subset \mathbb{R}^n \,|\, \varphi_i(\mathbf{x}) = 0,\ \forall\, i = 1, \ldots, p,\\ \psi_j(\mathbf{x}) \leqslant 0,\ \forall\, j = 1, \ldots, q\}$$

où :

– L'ensemble $\mathcal{U}$ est un sous-ensemble convexe non vide de $\mathbb{R}^n$,

– les applications $\varphi_1, \ldots, \varphi_p : \mathbb{R}^n \longrightarrow \mathbb{R}$ sont affines,

– les applications $\psi_1, \ldots, \psi_q : \mathbb{R}^n \longrightarrow \mathbb{R}$ sont convexes.

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*

Proposition (Convexité du domaine.)

– $\varphi_1, \dots, \varphi_p$ *affines,*
– $\psi_1, \dots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\Longrightarrow$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet :

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$.

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0,1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y})$

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0.$

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$, $\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$
$\implies$ $\mathcal{C}_j$ est un convexe.

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \mid \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$
$\implies$ $\mathcal{C}_j$ est un convexe.
Puisque φ_i est affine, $\{\mathbf{x} \in \mathbb{R}^n \mid \varphi_i(\mathbf{x}) = 0\}$ est un sous-espace affine

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\Longrightarrow$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$
$\Longrightarrow$ $\mathcal{C}_j$ est un convexe.
Puisque φ_i est affine, $\{\mathbf{x} \in \mathbb{R}^n \,|\, \varphi_i(\mathbf{x}) = 0\}$ est un sous-espace affine donc un convexe de $\mathbb{R}^n$.

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\Longrightarrow$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$
$\Longrightarrow$ $\mathcal{C}_j$ est un convexe.
Puisque φ_i est affine, $\{\mathbf{x} \in \mathbb{R}^n \,|\, \varphi_i(\mathbf{x}) = 0\}$ est un sous-espace affine donc un convexe de $\mathbb{R}^n$.
$\Longrightarrow$ $\mathcal{D}$ est une intersection de convexes

Proposition (Convexité du domaine.)

– $\varphi_1, \ldots, \varphi_p$ *affines,*
– $\psi_1, \ldots, \psi_q$ *convexes,*
– $\mathcal{U}$ *convexe,*
$\implies$ $\mathcal{D}$ *est un ensemble convexe de* $\mathbb{R}^n$*.*

Démonstration. ψ_j étant convexe, $\mathcal{C}_j = \{\mathbf{x} \in \mathbb{R}^n \,|\, \psi_j(\mathbf{x}) \leqslant 0\}$ est un convexe de $\mathbb{R}^n$. En effet : Soient $\mathbf{x}, \mathbf{y} \in \mathcal{C}_j$: $\psi_j(\mathbf{x}) \leqslant 0$, et $\psi_j(\mathbf{y}) \leqslant 0$. Par convexité de ψ_j, pour $t \in [0, 1]$,
$\psi_j(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant t\psi_j(\mathbf{x}) + (1-t)\psi_j(\mathbf{y}) \leqslant 0$. Ainsi $[\mathbf{x}, \mathbf{y}] \subset \mathcal{C}_j$
$\implies$ $\mathcal{C}_j$ est un convexe.
Puisque φ_i est affine, $\{\mathbf{x} \in \mathbb{R}^n \,|\, \varphi_i(\mathbf{x}) = 0\}$ est un sous-espace affine donc un convexe de $\mathbb{R}^n$.
$\implies$ $\mathcal{D}$ est une intersection de convexes $\implies$ $\mathcal{D}$ est convexe. $\square$

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ *un sous-ensemble convexe de* $\mathbb{R}^n$*,* $f : \mathcal{C} \longrightarrow \mathbb{R}$ *une application convexe et* $\mathbf{x}_0 \in \mathcal{C}$*.*

1. *Les conditions suivantes sont équivalentes :*

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ un sous-ensemble convexe de $\mathbb{R}^n$, $f : \mathcal{C} \longrightarrow \mathbb{R}$ une application convexe et $\mathbf{x}_0 \in \mathcal{C}$.

1. *Les conditions suivantes sont équivalentes :*

(i) $\mathbf{x}_0$ est un minimum local de f,

(ii) $\mathbf{x}_0$ est un minimum global de f.

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ un sous-ensemble convexe de $\mathbb{R}^n$, $f : \mathcal{C} \longrightarrow \mathbb{R}$ une application convexe et $\mathbf{x}_0 \in \mathcal{C}$.

1. *Les conditions suivantes sont équivalentes :*

(i) $\mathbf{x}_0$ *est un minimum local de* f,

(ii) $\mathbf{x}_0$ *est un minimum global de* f.

Si de plus f est différentiable en $\mathbf{x}_0 \in \mathcal{C}$, (i), (ii) sont équivalents à :

(iii) si $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{C})$, $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ un sous-ensemble convexe de $\mathbb{R}^n$, $f : \mathcal{C} \longrightarrow \mathbb{R}$ une application convexe et $\mathbf{x}_0 \in \mathcal{C}$.

1. *Les conditions suivantes sont équivalentes :*

(i) $\mathbf{x}_0$ *est un minimum local de* f,

(ii) $\mathbf{x}_0$ *est un minimum global de* f.

Si de plus f *est différentiable en* $\mathbf{x}_0 \in \mathcal{C}$, *(i), (ii) sont équivalents à :*

(iii) si $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{C})$, $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

(iv) $\forall\, \mathbf{x} \in \mathcal{C}$, $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geqslant 0$

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ un sous-ensemble convexe de $\mathbb{R}^n$, $f : \mathcal{C} \longrightarrow \mathbb{R}$ une application convexe et $\mathbf{x}_0 \in \mathcal{C}$.

1. *Les conditions suivantes sont équivalentes :*

(i) $\mathbf{x}_0$ *est un minimum local de* f,

(ii) $\mathbf{x}_0$ *est un minimum global de* f.

Si de plus f *est différentiable en* $\mathbf{x}_0 \in \mathcal{C}$, *(i), (ii) sont équivalents à :*

(iii) si $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{C})$, $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

(iv) $\forall\, \mathbf{x} \in \mathcal{C}$, $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geqslant 0$

2. *Si* f *est strictement convexe,* f *admet au plus un minimum, et un minimum de* f *est toujours strict.*

Théorème (Fondamental en programmation convexe)

Soient $\mathcal{C}$ un sous-ensemble convexe de $\mathbb{R}^n$, $f : \mathcal{C} \longrightarrow \mathbb{R}$ une application convexe et $\mathbf{x}_0 \in \mathcal{C}$.

1. *Les conditions suivantes sont équivalentes :*

(i) $\mathbf{x}_0$ *est un minimum local de* f,

(ii) $\mathbf{x}_0$ *est un minimum global de* f.

Si de plus f *est différentiable en* $\mathbf{x}_0 \in \mathcal{C}$, *(i), (ii) sont équivalents à :*

(iii) si $\mathbf{x}_0 \in \mathrm{Int}(\mathcal{C})$, $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

(iv) $\forall\, \mathbf{x} \in \mathcal{C}$, $\langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geqslant 0$

2. *Si* f *est strictement convexe,* f *admet au plus un minimum, et un minimum de* f *est toujours strict.*

Remarque. $f(x) = e^x$ est strictement convexe ($f''(x) = e^x > 0$) et n'admet aucun minimum (puisque $f'(x) = e^x \neq 0$).

Application elliptique - Définition

Définition. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^1. L'application f est *elliptique* ou encore α-*elliptique*, s'il existe un réel $\alpha > 0$, tel que :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geqslant \alpha \|\mathbf{x} - \mathbf{y}\|^2$$

Application elliptique - Définition

Définition. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^1. L'application f est *elliptique* ou encore α-*elliptique*, s'il existe un réel $\alpha > 0$, tel que :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geqslant \alpha \|\mathbf{x} - \mathbf{y}\|^2$$

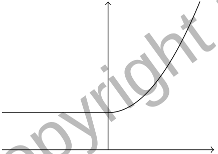
Proposition. Si f est deux fois différentiable, f est α-elliptique si et seulement si : $\forall \mathbf{x}, \mathbf{u} \in \mathbb{R}^n, \quad \mathbf{x}^\top \nabla^2 f(\mathbf{u}) \mathbf{x} \geqslant \alpha \|\mathbf{x}\|^2$.

Application elliptique - Définition

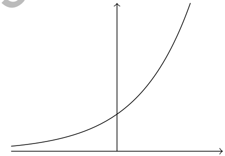
Définition. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^1. L'application f est *elliptique* ou encore α*-elliptique*, s'il existe un réel $\alpha > 0$, tel que :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geqslant \alpha \|\mathbf{x} - \mathbf{y}\|^2$$

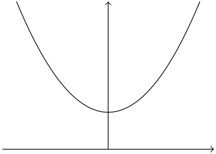
Proposition. Si f est deux fois différentiable, f est α-elliptique si et seulement si : $\forall \mathbf{x}, \mathbf{u} \in \mathbb{R}^n, \quad \mathbf{x}^\top \nabla^2 f(\mathbf{u}) \mathbf{x} \geqslant \alpha \|\mathbf{x}\|^2$.



convexe



strict. convexe



elliptique

Programmation elliptique

Théorème

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une application* α*-elliptique.*

Programmation elliptique

Théorème

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application α-elliptique. Alors f est coercive et strictement convexe.

Programmation elliptique

Théorème

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application α-elliptique. Alors f est coercive et strictement convexe. Sur un domaine convexe fermé et non vide de $\mathbb{R}^n$, elle admet un unique minimum.

Applications quadratiques

Définition. Une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est *quadratique* lorsque c'est un polynôme de degré 2.

Applications quadratiques

Définition. Une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est *quadratique* lorsque c'est un polynôme de degré 2.

Une application quadratique est de la forme :

$$f(x_1, x_2, \dots, x_n) = \underbrace{\frac{1}{2}\sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j} a_{ij}x_ix_j}_{\text{forme quadratique}} - \underbrace{\sum_{i=1}^{n} b_ix_i}_{\text{forme linéaire}} + \underbrace{c}_{constante}$$

Applications quadratiques

Définition. Une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est *quadratique* lorsque c'est un polynôme de degré 2.
Une application quadratique est de la forme :

$$f(x_1, x_2, \ldots, x_n) = \underbrace{\frac{1}{2}\sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j} a_{ij}x_ix_j}_{\text{forme quadratique}} - \underbrace{\sum_{i=1}^{n} b_ix_i}_{\text{forme linéaire}} + \underbrace{c}_{\textit{constante}}$$

En posant $A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \in \mathcal{M}_n(\mathbb{R})$, $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ et $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, on l'écrit sous *forme matricielle* :

$$f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$$

Gradient et Hessienne d'une apllication quadratique.

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$ *une application quadratique.*
Alors f *est infiniment différentiable et,*

$$\nabla f(\mathbf{x}) = A\mathbf{x} - b,$$
$$\nabla^2 f(\mathbf{x}) = A.$$

Gradient et Hessienne d'une apllication quadratique.

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$ *une application quadratique.*
Alors f *est infiniment différentiable et,*

$$\nabla f(\mathbf{x}) = A\mathbf{x} - b,$$
$$\nabla^2 f(\mathbf{x}) = A.$$

Démonstration. f polynomiale $\implies C^\infty$.

Gradient et Hessienne d'une apllication quadratique.

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$ *une application quadratique.*
Alors f *est infiniment différentiable et,*

$$\nabla f(\mathbf{x}) = A\mathbf{x} - b,$$
$$\nabla^2 f(\mathbf{x}) = A.$$

Démonstration. f polynomiale $\implies C^\infty$.
Pour tout $i = 1, 2, \ldots, n$, le calcul donne $\frac{\partial f}{\partial x_i}(\mathbf{x}) = \sum_{j=1}^n a_{ij}x_j - b_i$,
donc $\nabla f(\mathbf{x}) = A\mathbf{x} - b$.

Gradient et Hessienne d'une apllication quadratique.

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$ une application quadratique. Alors f est infiniment différentiable et,

$$\nabla f(\mathbf{x}) = A\mathbf{x} - b,$$
$$\nabla^2 f(\mathbf{x}) = A.$$

Démonstration. f polynomiale $\implies C^\infty$.
Pour tout $i = 1, 2, \ldots, n$, le calcul donne $\frac{\partial f}{\partial x_i}(\mathbf{x}) = \sum_{j=1}^n a_{ij}x_j - b_i$, donc $\nabla f(\mathbf{x}) = A\mathbf{x} - b$.
Pour tout $i, j = 1, 2, \ldots, n$, $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = a_{ij}$, et l'on obtient $\nabla^2 f(\mathbf{x}) = A$. $\square$

Convexité d'une application quadratique

Théorème

Soit l'application quadratique $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$. *Alors :*

- f *convexe* $\iff$ A *semi-définie positive*
- f *strictement convexe* $\iff$ f *elliptique* $\iff$ A *définie positive.*

Convexité d'une application quadratique

Théorème

Soit l'application quadratique $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$. *Alors :*

- f *convexe* $\iff$ A *semi-définie positive*
- f *strictement convexe* $\iff$ f *elliptique* $\iff$ A *définie positive.*

Preuve (esquisse). Utiliser la caractérisation des diverses convexités à l'ordre 2. □

Programmation quadratique

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c.$

Programmation quadratique

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$. *Si A est semi-définie positive (resp. négative), alors les propositions suivantes sont équivalentes :*

- $\mathbf{u}$ *est un minimum (resp. maximum) local de* f,
- $\mathbf{u}$ *est un minimum (resp. maximum) global de* f,
- $A\mathbf{u} = \mathbf{b}$, *i.e.* $\mathbf{u}$ *est solution du système d'équations linéaires* $A\mathbf{x} = \mathbf{b}$.

Programmation quadratique

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$. *Si* A *est semi-définie positive (resp. négative), alors les propositions suivantes sont équivalentes :*

- $\mathbf{u}$ *est un minimum (resp. maximum) local de* f,
- $\mathbf{u}$ *est un minimum (resp. maximum) global de* f,
- $A\mathbf{u} = \mathbf{b}$, *i.e.* $\mathbf{u}$ *est solution du système d'équations linéaires* $A\mathbf{x} = \mathbf{b}$.

Si A *est définie positive* f *admet un unique minimum (resp. maximum) global.*

Programmation quadratique

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$*. Si* A *est semi-définie positive (resp. négative), alors les propositions suivantes sont équivalentes :*

- $\mathbf{u}$ *est un minimum (resp. maximum) local de* f*,*
- $\mathbf{u}$ *est un minimum (resp. maximum) global de* f*,*
- $A\mathbf{u} = \mathbf{b}$*, i.e.* $\mathbf{u}$ *est solution du système d'équations linéaires* $A\mathbf{x} = \mathbf{b}$*.*

Si A *est définie positive* f *admet un unique minimum (resp. maximum) global.*

Si A *n'est pas semi-définie positive (resp. négative)* f *n'admet aucun minimum (resp. maximum) local ou global.*

Programmation quadratique

Théorème

Soit $f(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle + c$. *Si A est semi-définie positive (resp. négative), alors les propositions suivantes sont équivalentes :*

- $\mathbf{u}$ *est un minimum (resp. maximum) local de* f*,*
- $\mathbf{u}$ *est un minimum (resp. maximum) global de* f*,*
- $A\mathbf{u} = \mathbf{b}$*, i.e.* $\mathbf{u}$ *est solution du système d'équations linéaires* $A\mathbf{x} = \mathbf{b}$.

Si A est définie positive f admet un unique minimum (resp. maximum) global.

Si A n'est pas semi-définie positive (resp. négative) f n'admet aucun minimum (resp. maximum) local ou global.

Preuve (esquisse). Découle immédiatement des théorèmes précédents.

Exemple.

Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top \mathbf{x}$ avec $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ et $\mathbf{b} = (-3, 1, -2)$.

Exemple.

Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top \mathbf{x}$ avec $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ et $\mathbf{b} = (-3, 1, -2)$. Le polynôme caractéristique de A est

$$p_A(\lambda) = 8 - 12\lambda + 6\lambda^2 - \lambda^3$$

Exemple.

Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ avec $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ et $\mathbf{b} = (-3, 1, -2)$. Le polynôme caractéristique de A est

$$p_A(\lambda) = 8 - 12\lambda + 6\lambda^2 - \lambda^3$$

$\Longrightarrow$ A est définie positive

Exemple.

Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top \mathbf{x}$ avec $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ et $\mathbf{b} = (-3, 1, -2)$. Le polynôme caractéristique de A est

$$p_A(\lambda) = 8 - 12\lambda + 6\lambda^2 - \lambda^3$$

$\Longrightarrow$ A est définie positive $\Longrightarrow$ f a un unique minimum global qui est l'unique solution de $A\mathbf{x} = \mathbf{b}$.

Exemple.

Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ avec $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ et $\mathbf{b} = (-3, 1, -2)$. Le polynôme caractéristique de A est

$$p_A(\lambda) = 8 - 12\lambda + 6\lambda^2 - \lambda^3$$

$\implies$ A est définie positive $\implies$ f a un unique minimum global qui est l'unique solution de $A\mathbf{x} = \mathbf{b}$.

$$A\mathbf{x} = \mathbf{b} \iff \left\{ \begin{array}{rrrcr} 2x & -y & & = & -3 \\ -x & +2y & -z & = & 1 \\ & -y & +2z & = & -2 \end{array} \right. \implies \mathbf{x}_{\min} = \begin{pmatrix} -9/4 \\ -3/2 \\ -7/4 \end{pmatrix}$$

Exercices

Exercice 1.

Exercice 2.

Exercice 3.

Exercice 4.

Exercice 5.

Exercice 1. Déterminer les extrema locaux et globaux de l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x,y) = x^3 + y^3 + x^2 + y^2 - 1 .$$

Retour.

Exercice 1. Déterminer les extrema locaux et globaux de l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x, y) = x^3 + y^3 + x^2 + y^2 - 1 .$$

Retour.

f est C^∞. On détermine en chaque point son gradient et sa matrice Hessienne.

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 + 2x \\ 3y^2 + 2y \end{pmatrix} \qquad \nabla^2 f(x, y) = \begin{pmatrix} 6x + 2 & 0 \\ 0 & 6y + 2 \end{pmatrix}$$

Exercice 1. Déterminer les extrema locaux et globaux de l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x,y) = x^3 + y^3 + x^2 + y^2 - 1 .$$

Retour.

f est C^∞. On détermine en chaque point son gradient et sa matrice Hessienne.

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 + 2x \\ 3y^2 + 2y \end{pmatrix} \qquad \nabla^2 f(x,y) = \begin{pmatrix} 6x+2 & 0 \\ 0 & 6y+2 \end{pmatrix}$$

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 + 2x \\ 3y^2 + 2y \end{pmatrix} = 0 \quad \Longleftrightarrow \quad \begin{cases} x = 0 \text{ ou } x = -\frac{2}{3} \\ y = 0 \text{ ou } y = -\frac{2}{3} \end{cases}$$

Les points critiques sont :
$(0,0)$, $(0,-2/3)$, $(-2/3,0)$, $(-2/3,-2/3)$.

On évalue en chaque point critique la matrice Hessienne.

On évalue en chaque point critique la matrice Hessienne.

$$\nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \text{ définie positive : } (0,0) \text{ \underline{min local}.}$$

On évalue en chaque point critique la matrice Hessienne.

$$\nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$ définie positive : $(0,0)$ min local.

$$\nabla^2 f(0,-\frac{2}{3}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$ non semi-définie : n'est pas un extremum.

On évalue en chaque point critique la matrice Hessienne.

$$\nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$ définie positive : $(0,0)$ min local.

$$\nabla^2 f(0,-\frac{2}{3}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$ non semi-définie : n'est pas un extremum.

$$\nabla^2 f(-\frac{2}{3},0) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$ non semi-définie : n'est pas un extremum.

On évalue en chaque point critique la matrice Hessienne.

$$\nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$ définie positive : $(0,0)$ min local.

$$\nabla^2 f(0,-\frac{2}{3}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$ non semi-définie : n'est pas un extremum.

$$\nabla^2 f(-\frac{2}{3},0) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$ non semi-définie : n'est pas un extremum.

$$\nabla^2 f(-\frac{2}{3},-\frac{2}{3}) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$ définie négative : $(-\frac{2}{3},-\frac{2}{3})$ max local.

L'application f n'admet pas d'extremum global, car elle est surjective sur $\mathbb{R}$:

$$\lim_{x\to+\infty} f(x,0)=+\infty \quad \lim_{x\to-\infty} f(x,0)=-\infty\ .$$

Retour.

Exercice 2.

On considère l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x, y) = x^4 + y^4 - x^3 - y^3 .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?

b. Déterminer tous les extrema globaux de f.

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable et $\mathbf{u}$ un point critique de g, alors $\mathbf{u}$ est un minimum local de g si et seulement si g est convexe sur une boule ouverte centrée en $\mathbf{u}$.

d. En déduire tous les extrema locaux de f.

Retour.

$$f(x,y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?

$$f(x,y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?

Montrons que f est coercive. Formons :

$$f(x,y) - \|(x,y)\|^2 = x^2(x^2 - x - 1) + y^2(y^2 - y - 1) \ .$$

$$f(x,y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?
Montrons que f est coercive. Formons :

$$f(x,y) - \|(x,y)\|^2 = x^2(x^2 - x - 1) + y^2(y^2 - y - 1) \ .$$

Le trinôme $t^2 - t - 1$ est minoré par $-\frac{1}{4}$ et positif lorsque $t \notin]\frac{-1-\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2}[$.

$$f(x,y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?
Montrons que f est coercive. Formons :

$$f(x,y) - \|(x,y)\|^2 = x^2(x^2 - x - 1) + y^2(y^2 - y - 1) \ .$$

Le trinôme $t^2 - t - 1$ est minoré par $-\frac{1}{4}$ et positif lorsque $t \notin]\frac{-1-\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2}[$.
Ainsi lorsque x ou y est suffisamment grand, $f(x,y) \geqslant \|(x,y)\|^2$.

$$f(x, y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?
Montrons que f est coercive. Formons :

$$f(x, y) - \|(x, y)\|^2 = x^2(x^2 - x - 1) + y^2(y^2 - y - 1) \ .$$

Le trinôme $t^2 - t - 1$ est minoré par $-\frac{1}{4}$ et positif lorsque $t \notin]\frac{-1-\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2}[$.
Ainsi lorsque x ou y est suffisamment grand, $f(x, y) \geqslant \|(x, y)\|^2$.
Lorsque $\|(x, y)\|$ tend vers $+\infty$, $f(x, y) \geqslant \|(x, y)\|$ tend aussi vers $+\infty$. Donc f est coercive.

$$f(x,y) = x^4 + y^4 - x^3 - y^3 \ .$$

a. Que peut-on dire de l'existence d'extrema globaux pour f ?
Montrons que f est coercive. Formons :

$$f(x,y) - \|(x,y)\|^2 = x^2(x^2 - x - 1) + y^2(y^2 - y - 1) \ .$$

Le trinôme $t^2 - t - 1$ est minoré par $-\frac{1}{4}$ et positif lorsque $t \notin]\frac{-1-\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2}[$.
Ainsi lorsque x ou y est suffisamment grand, $f(x,y) \geqslant \|(x,y)\|^2$.
Lorsque $\|(x,y)\|$ tend vers $+\infty$, $f(x,y) \geqslant \|(x,y)\|$ tend aussi vers $+\infty$. Donc f est coercive.

On en déduit l'existence d'un minimum global et d'aucun maximum global.

b. Déterminer tous les extrema globaux de f.

b. Déterminer tous les extrema globaux de f.

On cherche les extrema locaux.

b. Déterminer tous les extrema globaux de f.

On cherche les extrema locaux. L'application f est C^∞ ; son gradient est :

$$\nabla f(x,y) = \begin{pmatrix} 4x^3 - 3x^2 \\ 4y^3 - 3y^2 \end{pmatrix}$$

b. Déterminer tous les extrema globaux de f.

On cherche les extrema locaux. L'application f est C^∞ ; son gradient est :

$$\nabla f(x,y) = \begin{pmatrix} 4x^3 - 3x^2 \\ 4y^3 - 3y^2 \end{pmatrix}$$

Ainsi f a 4 points critiques $A = (0,0)$, $B = (0, \frac{3}{4})$, $C = (\frac{3}{4}, 0)$ et $D = (\frac{3}{4}, \frac{3}{4})$.

b. Déterminer tous les extrema globaux de f.
On cherche les extrema locaux. L'application f est C^∞ ; son gradient est :

$$\nabla f(x,y) = \begin{pmatrix} 4x^3 - 3x^2 \\ 4y^3 - 3y^2 \end{pmatrix}$$

Ainsi f a 4 points critiques $A = (0,0)$, $B = (0, \frac{3}{4})$, $C = (\frac{3}{4}, 0)$ et $D = (\frac{3}{4}, \frac{3}{4})$.
Sa matrice hessienne est :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

b. Déterminer tous les extrema globaux de f.
On cherche les extrema locaux. L'application f est C^∞ ; son gradient est :

$$\nabla f(x,y) = \begin{pmatrix} 4x^3 - 3x^2 \\ 4y^3 - 3y^2 \end{pmatrix}$$

Ainsi f a 4 points critiques $A = (0,0)$, $B = (0, \frac{3}{4})$, $C = (\frac{3}{4}, 0)$ et $D = (\frac{3}{4}, \frac{3}{4})$.
Sa matrice hessienne est :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

$\nabla^2 f(D)$ est définie positive : D est un min local. En A, B, C, $\nabla^2 f(.)$ est semi-définie positive : on ne peut rien déduire...

Pour déterminer le(s) minimum(s) de f il suffit d'évaluer f en A, B, C, D.

$$f(0,0)=0\ ;\ f\left(0,\frac{3}{4}\right)=f\left(\frac{3}{4},0\right)=-\frac{3^3}{4^4}\ ;\ f\left(\frac{3}{4},\frac{3}{4}\right)=-2\frac{3^3}{4^4}\ .$$

Ainsi le minimum de f est $D=\left(\frac{3}{4},\frac{3}{4}\right)$.

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une application différentiable et* $\mathbf{u}$ *un point critique de* g*, alors* $\mathbf{u}$ *est un minimum local de* g *si et seulement si* g *est convexe sur une boule ouverte centrée en* $\mathbf{u}$.

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une application différentiable et* $\mathbf{u}$ *un point critique de* g*, alors* $\mathbf{u}$ *est un minimum local de* g *si et seulement si* g *est convexe sur une boule ouverte centrée en* $\mathbf{u}$*.*

Soit $\mathbf{u}$ un point critique de g, *i.e.* $\nabla g(\mathbf{u}) = \mathbf{0}$.

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable et $\mathbf{u}$ *un point critique de g, alors* $\mathbf{u}$ *est un minimum local de g si et seulement si g est convexe sur une boule ouverte centrée en* $\mathbf{u}$*.*

Soit $\mathbf{u}$ un point critique de g, *i.e.* $\nabla g(\mathbf{u}) = \mathbf{0}$.

$\mathbf{u}$ min local de $g \implies \exists$ une boule ouverte $\mathcal{B}$ centrée en $\mathbf{u}$ tel que,

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u})$$

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable et $\mathbf{u}$ un point critique de g, alors $\mathbf{u}$ est un minimum local de g si et seulement si g est convexe sur une boule ouverte centrée en $\mathbf{u}$.

Soit $\mathbf{u}$ un point critique de g, *i.e.* $\nabla g(\mathbf{u}) = \mathbf{0}$.

$\mathbf{u}$ min local de $g \implies \exists$ une boule ouverte $\mathcal{B}$ centrée en $\mathbf{u}$ tel que,

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u}) = g(\mathbf{u}) + \langle \nabla g(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle \ .$$

$\implies$ la restriction de g à $\mathcal{B}$ est une application convexe.

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une application différentiable et* $\mathbf{u}$ *un point critique de* g*, alors* $\mathbf{u}$ *est un minimum local de* g *si et seulement si* g *est convexe sur une boule ouverte centrée en* $\mathbf{u}$*.*

Soit $\mathbf{u}$ un point critique de g, *i.e.* $\nabla g(\mathbf{u}) = \mathbf{0}$.
$\mathbf{u}$ min local de $g \implies \exists$ une boule ouverte $\mathcal{B}$ centrée en $\mathbf{u}$ tel que,

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u}) = g(\mathbf{u}) + \langle \nabla g(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle \ .$$

$\implies$ la restriction de g à $\mathcal{B}$ est une application convexe.

Réciproquement, si g est convexe sur une boule $\mathcal{B}$ centrée en $\mathbf{u}$, alors

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u}) + \langle \nabla g(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle$$

c. Montrer le résultat :

Soit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une application différentiable et* $\mathbf{u}$ *un point critique de* g*, alors* $\mathbf{u}$ *est un minimum local de* g *si et seulement si* g *est convexe sur une boule ouverte centrée en* $\mathbf{u}$*.*

Soit $\mathbf{u}$ un point critique de g, *i.e.* $\nabla g(\mathbf{u}) = \mathbf{0}$.
$\mathbf{u}$ min local de $g \implies \exists$ une boule ouverte $\mathcal{B}$ centrée en $\mathbf{u}$ tel que,

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u}) = g(\mathbf{u}) + \langle \nabla g(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle \ .$$

$\implies$ la restriction de g à $\mathcal{B}$ est une application convexe.

Réciproquement, si g est convexe sur une boule $\mathcal{B}$ centrée en $\mathbf{u}$, alors

$$\forall \mathbf{x} \in \mathcal{B}, \quad g(\mathbf{x}) \geqslant g(\mathbf{u}) + \langle \nabla g(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle = g(\mathbf{u})$$

$\implies$ $\mathbf{u}$ est un min local de g.

d. En déduire tous les extrema locaux de f.

d. En déduire tous les extrema locaux de f.

Nous avons déterminé :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

d. En déduire tous les extrema locaux de f.
Nous avons déterminé :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

Le binôme $12t^2 - 6t$ ne garde pas un signe constant sur un voisinage de 0

d. En déduire tous les extrema locaux de f.

Nous avons déterminé :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

Le binôme $12t^2 - 6t$ ne garde pas un signe constant sur un voisinage de 0 $\implies$ sur aucun voisinage de A, B et C, la matrice hessienne ne reste semi-définie positive ou négative.

d. En déduire tous les extrema locaux de f.

Nous avons déterminé :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

Le binôme $12t^2 - 6t$ ne garde pas un signe constant sur un voisinage de 0 $\Longrightarrow$ sur aucun voisinage de A, B et C, la matrice hessienne ne reste semi-définie positive ou négative. $\Longrightarrow$ f n'est ni localement convexe ni localement concave sur un voisinage convexe de A, B ou C.

d. En déduire tous les extrema locaux de f.

Nous avons déterminé :

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 - 6x & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

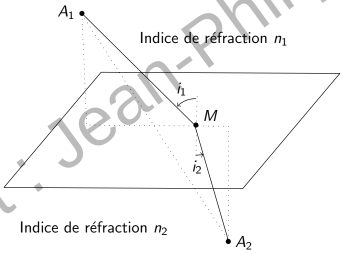
Le binôme $12t^2 - 6t$ ne garde pas un signe constant sur un voisinage de 0 $\Longrightarrow$ sur aucun voisinage de A, B et C, la matrice hessienne ne reste semi-définie positive ou négative. $\Longrightarrow$ f n'est ni localement convexe ni localement concave sur un voisinage convexe de A, B ou C.

(c) $\Longrightarrow$ ni A, ni B, ni C n'est un extremum local de f.

Retour.

Exercice 3.

Un rayon lumineux effectue un trajet spatial d'un point A_1 situé dans un milieu ayant pour indice de réfraction n_1 à un point A_2 situé dans un milieu ayant pour indice de réfraction n_2 ; les deux milieux étant séparés par un plan $\mathcal{P}$.



En appliquant le principe que la lumière parcourt le trajet le plus rapide, retrouver la loi de Descartes de réfraction de la lumière : $\boxed{n_1 \sin i_1 = n_2 \sin i_2}$.

Retour.

On rappelle que l'indice de réfraction est $n_i = \frac{c}{v_i}$ où c désigne la vitesse de propagation de la lumière dans le vide et v_i sa vitesse de propagation dans le milieu.

On rappelle que l'indice de réfraction est $n_i = \frac{c}{v_i}$ où c désigne la vitesse de propagation de la lumière dans le vide et v_i sa vitesse de propagation dans le milieu.

La lumière parcourt le trajet qui minimise le temps de parcours.

On rappelle que l'indice de réfraction est $n_i = \frac{c}{v_i}$ où c désigne la vitesse de propagation de la lumière dans le vide et v_i sa vitesse de propagation dans le milieu.
La lumière parcourt le trajet qui minimise le temps de parcours. Ce dernier est :

$$\frac{A_1M}{v_1} + \frac{A_2M}{v_2} = \frac{n_1A_1M}{c} + \frac{n_2A_2M}{c}.$$

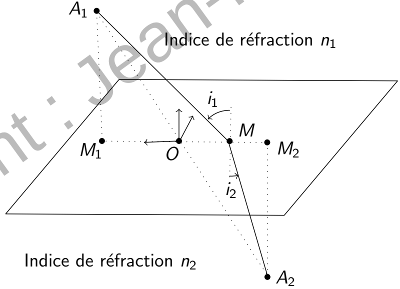
On rappelle que l'indice de réfraction est $n_i = \frac{c}{v_i}$ où c désigne la vitesse de propagation de la lumière dans le vide et v_i sa vitesse de propagation dans le milieu.
La lumière parcourt le trajet qui minimise le temps de parcours. Ce dernier est :

$$\frac{A_1M}{v_1} + \frac{A_2M}{v_2} = \frac{n_1A_1M}{c} + \frac{n_2A_2M}{c} \; .$$

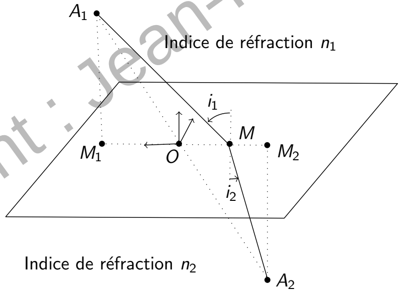
Il s'agit donc de déterminer le point M de façon à minimiser le *chemin optique* $n_1A_1M + n_2A_2M$.

$$\min_{M\in\mathcal{P}} \; n_1A_1M + n_2A_2M \; .$$

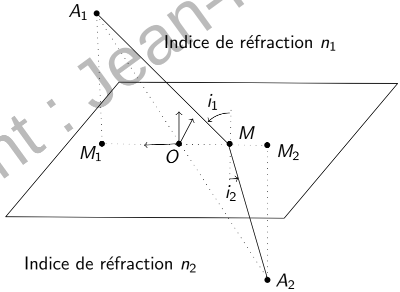
On se donne un repère orthonormé construit de la façon suivante :



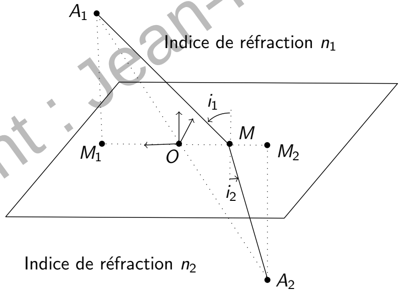
On se donne un repère orthonormé construit de la façon suivante : soit O le point d'intersection de la droite (A_1A_2) avec le plan de séparation $\mathcal{P}$.



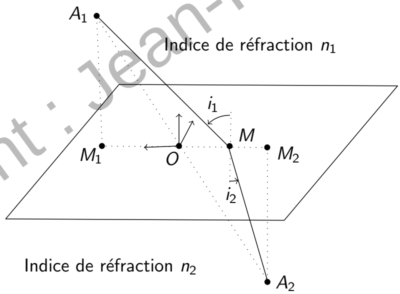
On se donne un repère orthonormé construit de la façon suivante : soit O le point d'intersection de la droite (A_1A_2) avec le plan de séparation $\mathcal{P}$. Soient M_1 et M_2 les projetés orthogonaux respectifs de A_1 et A_2 sur $\mathcal{P}$.



On se donne un repère orthonormé construit de la façon suivante : soit O le point d'intersection de la droite (A_1A_2) avec le plan de séparation $\mathcal{P}$. Soient M_1 et M_2 les projetés orthogonaux respectifs de A_1 et A_2 sur $\mathcal{P}$. Le segment $[M_1M_2]$ passe par O.



On se donne un repère orthonormé construit de la façon suivante : soit O le point d'intersection de la droite (A_1A_2) avec le plan de séparation $\mathcal{P}$. Soient M_1 et M_2 les projetés orthogonaux respectifs de A_1 et A_2 sur $\mathcal{P}$. Le segment $[M_1M_2]$ passe par O. On choisit un repère orthonormal d'origine O, tel que (Oi) est confondu avec (M_1M_2) et (Oj) est dans $\mathcal{P}$; alors k est orthogonal à $\mathcal{P}$.



Les coordonnées de M, A_1, A_2 dans ce repère sont respectivement $(x, y, 0)$, $(x_1, 0, z_1)$ et $(x_2, 0, z_2)$.

Les coordonnées de M, A_1, A_2 dans ce repère sont respectivement $(x, y, 0)$, $(x_1, 0, z_1)$ et $(x_2, 0, z_2)$. Le chemin optique s'exprime alors :

$$f(x,y) = n_1\sqrt{(x-x_1)^2+y^2+z_1^2} + n_2\sqrt{(x-x_2)^2+y^2+z_2^2}$$

et il s'agit de le minimiser.

Les coordonnées de M, A_1, A_2 dans ce repère sont respectivement $(x, y, 0)$, $(x_1, 0, z_1)$ et $(x_2, 0, z_2)$. Le chemin optique s'exprime alors :

$$f(x,y) = n_1\sqrt{(x-x_1)^2 + y^2 + z_1^2} + n_2\sqrt{(x-x_2)^2 + y^2 + z_2^2}$$

et il s'agit de le minimiser. L'application f est clairement coercive et admet donc un minimum.

Etudions ses points critiques :

Etudions ses points critiques :

$$\frac{\partial f}{\partial x}(x,y) = n_1 \frac{x - x_1}{\sqrt{(x-x_1)^2 + y^2 + z_1^2}} + n_2 \frac{x - x_2}{\sqrt{(x-x_2)^2 + y^2 + z_2^2}}$$

$$\frac{\partial f}{\partial y}(x,y) = n_1 \frac{y}{\sqrt{(x-x_1)^2 + y^2 + z_1^2}} + n_2 \frac{y}{\sqrt{(x-x_2)^2 + y^2 + z_2^2}}$$

Etudions ses points critiques :

$$\frac{\partial f}{\partial x}(x,y) = n_1 \frac{x-x_1}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{x-x_2}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$
$$\frac{\partial f}{\partial y}(x,y) = n_1 \frac{y}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{y}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$

Puisque $\frac{\partial f}{\partial y}(x,y) = 0$, on a $y = 0$.

Etudions ses points critiques :

$$\frac{\partial f}{\partial x}(x,y) = n_1 \frac{x-x_1}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{x-x_2}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$

$$\frac{\partial f}{\partial y}(x,y) = n_1 \frac{y}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{y}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$

Puisque $\frac{\partial f}{\partial y}(x,y) = 0$, on a $y = 0$.

$\Longrightarrow$ M est situé sur la droite $(M_1 M_2)$.

Etudions ses points critiques :

$$\frac{\partial f}{\partial x}(x,y) = n_1 \frac{x-x_1}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{x-x_2}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$

$$\frac{\partial f}{\partial y}(x,y) = n_1 \frac{y}{\sqrt{(x-x_1)^2+y^2+z_1^2}} + n_2 \frac{y}{\sqrt{(x-x_2)^2+y^2+z_2^2}}$$

Puisque $\frac{\partial f}{\partial y}(x,y) = 0$, on a $y = 0$.

$\Longrightarrow$ M est situé sur la droite $(M_1 M_2)$.

Puisque $\frac{\partial f}{\partial x}(x,y) = 0$, $x - x_1$ et $x - x_2$ sont de signes opposés,

Etudions ses points critiques :

$$\frac{\partial f}{\partial x}(x,y) = n_1 \frac{x - x_1}{\sqrt{(x-x_1)^2 + y^2 + z_1^2}} + n_2 \frac{x - x_2}{\sqrt{(x-x_2)^2 + y^2 + z_2^2}}$$

$$\frac{\partial f}{\partial y}(x,y) = n_1 \frac{y}{\sqrt{(x-x_1)^2 + y^2 + z_1^2}} + n_2 \frac{y}{\sqrt{(x-x_2)^2 + y^2 + z_2^2}}$$

Puisque $\frac{\partial f}{\partial y}(x,y) = 0$, on a $y = 0$.

$\Longrightarrow$ M est situé sur la droite (M_1M_2).

Puisque $\frac{\partial f}{\partial x}(x,y) = 0$, $x - x_1$ et $x - x_2$ sont de signes opposés,

$\Longrightarrow$ $x_1 \leqslant x \leqslant x_2$: M est situé sur le segment $[M_1M_2]$.

Alors au point $M(x, 0)$,

Alors au point $M(x,0)$,

$$\frac{\partial f}{\partial x}(x,0) = n_1 \frac{x - x_1}{\sqrt{(x-x_1)^2 + z_1^2}} + n_2 \frac{x - x_2}{\sqrt{(x-x_2)^2 + z_2^2}}$$

$$= n_1 \frac{M_1 M}{A_1 M} - n_2 \frac{M_2 M}{A_2 M} = 0$$

Alors au point $M(x,0)$,

$$\frac{\partial f}{\partial x}(x,0) = n_1 \frac{x - x_1}{\sqrt{(x-x_1)^2 + z_1^2}} + n_2 \frac{x - x_2}{\sqrt{(x-x_2)^2 + z_2^2}}$$

$$= n_1 \frac{M_1M}{A_1M} - n_2 \frac{M_2M}{A_2M} = 0$$

$$\Longrightarrow \quad n_1 \frac{M_1M}{A_1M} = n_2 \frac{M_2M}{A_2M}.$$

Alors au point $M(x,0)$,

$$\frac{\partial f}{\partial x}(x,0) = n_1 \frac{x-x_1}{\sqrt{(x-x_1)^2+z_1^2}} + n_2 \frac{x-x_2}{\sqrt{(x-x_2)^2+z_2^2}}$$

$$= n_1 \frac{M_1M}{A_1M} - n_2 \frac{M_2M}{A_2M} = 0$$

$$\implies \quad n_1 \frac{M_1M}{A_1M} = n_2 \frac{M_2M}{A_2M}.$$

Or pour $k=1,2$, le triangle A_kM_kM étant rectangle en M_k

$$\frac{M_kM}{A_kM} = \cos(\frac{\pi}{2} - i_k) = \sin\ i_k\ .$$

Alors au point $M(x, 0)$,

$$\begin{aligned}\frac{\partial f}{\partial x}(x,0) &= n_1 \frac{x - x_1}{\sqrt{(x-x_1)^2 + z_1^2}} + n_2 \frac{x - x_2}{\sqrt{(x-x_2)^2 + z_2^2}} \\ &= n_1 \frac{M_1 M}{A_1 M} - n_2 \frac{M_2 M}{A_2 M} = 0\end{aligned}$$

$\implies \quad n_1 \frac{M_1 M}{A_1 M} = n_2 \frac{M_2 M}{A_2 M}$.

Or pour $k = 1, 2$, le triangle $A_k M_k M$ étant rectangle en M_k

$$\frac{M_k M}{A_k M} = \cos(\frac{\pi}{2} - i_k) = \sin\ i_k \ .$$

On trouve donc qu'au minimum on a :

$$\boxed{n_1 \sin i_1 = n_2 \sin i_2} \ .$$

Retour.

Exercice 4.

Prouver le **théorème de projection convexe** :
Soit $\mathcal{C}$ un sous-ensemble convexe fermé non vide de $\mathbb{R}^n$. Donné $\mathbf{u} \in \mathbb{R}^n$ il existe un unique point $P_{\mathcal{C}}(\mathbf{u}) \in \mathcal{C}$, tel que :

$$\|P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}\| = \min_{\mathbf{v} \in \mathcal{C}} \|\mathbf{v} - \mathbf{u}\| \ .$$

On l'appelle le projeté de $\mathbf{u}$ sur $\mathcal{C}$. Il est caractérisé par :

$$\forall \mathbf{v} \in \mathcal{C}, \ \langle P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}, \mathbf{v} - P_{\mathcal{C}}(\mathbf{u}) \rangle \geqslant 0 \ .$$

De plus l'application $P_{\mathcal{C}}$ est contractante, i.e. :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \ \|P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y})\| \leqslant \|\mathbf{x} - \mathbf{y}\| \ .$$

Exercice 4.

Prouver le **théorème de projection convexe** :
Soit $\mathcal{C}$ un sous-ensemble convexe fermé non vide de $\mathbb{R}^n$. Donné $\mathbf{u} \in \mathbb{R}^n$ il existe un unique point $P_{\mathcal{C}}(\mathbf{u}) \in \mathcal{C}$, tel que :

$$\|P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}\| = \min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{v} - \mathbf{u}\| \ .$$

On l'appelle le projeté de $\mathbf{u}$ sur $\mathcal{C}$. Il est caractérisé par :

$$\forall \mathbf{v} \in \mathcal{C},\ \langle P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}, \mathbf{v} - P_{\mathcal{C}}(\mathbf{u})\rangle \geqslant 0 \ .$$

De plus l'application $P_{\mathcal{C}}$ est contractante, i.e. :

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \|P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y})\| \leqslant \|\mathbf{x} - \mathbf{y}\| \ .$$

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

b. Prouver la caractérisation donnée de $P_{\mathcal{C}}(\mathbf{u})$.

c. Utiliser cette caractérisation pour prouver que $P_{\mathcal{C}}$ est une application contractante.

Retour.

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

Le problème de minimisation $\min\limits_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|$ est équivalent au problème $\min\limits_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|^2$.

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

Le problème de minimisation $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|$ est équivalent au problème $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|^2$. Or l'application

$$f : \mathbf{x} \mapsto \|\mathbf{u}-\mathbf{x}\|^2 = \sum_{i=1}^{n}(u_i - x_i)^2 = \mathbf{x}^\top \mathrm{Id}\,\mathbf{x} - 2\mathbf{u}^\top\mathbf{x} + \|\mathbf{u}\|^2$$

est une application quadratique de matrice hessienne $2\,\mathrm{Id}$.

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

Le problème de minimisation $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|$ est équivalent au problème $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|^2$. Or l'application

$$f : \mathbf{x} \mapsto \|\mathbf{u}-\mathbf{x}\|^2 = \sum_{i=1}^{n}(u_i - x_i)^2 = \mathbf{x}^\top \mathrm{Id}\,\mathbf{x} - 2\mathbf{u}^\top \mathbf{x} + \|\mathbf{u}\|^2$$

est une application quadratique de matrice hessienne $2\,\mathrm{Id}$. f est donc une elliptique, et donc strictement convexe et coercive.

a. Prouver l'existence et l'unicité de $P_{\mathcal{C}}(\mathbf{u})$.

Le problème de minimisation $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|$ est équivalent au problème $\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{u}-\mathbf{v}\|^2$. Or l'application

$$f : \mathbf{x} \mapsto \|\mathbf{u}-\mathbf{x}\|^2 = \sum_{i=1}^{n}(u_i - x_i)^2 = \mathbf{x}^\top \mathrm{Id}\,\mathbf{x} - 2\mathbf{u}^\top\mathbf{x} + \|\mathbf{u}\|^2$$

est une application quadratique de matrice hessienne $2\,\mathrm{Id}$. f est donc une elliptique, et donc strictement convexe et coercive. Le domaine $\mathcal{C}$ étant convexe fermé et non vide elle y admet un unique minimum, $P_{\mathcal{C}}(\mathbf{u})$.

b. Prouver la caractérisation donnée de $P_{\mathcal{C}}(\mathbf{u})$.

b. Prouver la caractérisation donnée de $P_{\mathcal{C}}(\mathbf{u})$.

On est en programmation convexe, et f est différentiable. La caractérisation de $P_{\mathcal{C}}(\mathbf{u})$ est donnée par :

$$\forall \mathbf{v} \in \mathcal{C}, \quad \langle \nabla f(P_{\mathcal{C}}(\mathbf{u})), \mathbf{v} - P_{\mathcal{C}}(\mathbf{u}) \rangle \geqslant 0 \ .$$

b. Prouver la caractérisation donnée de $P_{\mathcal{C}}(\mathbf{u})$.

On est en programmation convexe, et f est différentiable. La caractérisation de $P_{\mathcal{C}}(\mathbf{u})$ est donnée par :

$$\forall \mathbf{v} \in \mathcal{C}, \quad \langle \nabla f(P_{\mathcal{C}}(\mathbf{u})), \mathbf{v} - P_{\mathcal{C}}(\mathbf{u}) \rangle \geqslant 0 \ .$$

Or f et quadratique, $\nabla f(\mathbf{x}) = 2\mathbf{x} - 2\mathbf{u}$, donc :

$$\nabla f(P_{\mathcal{C}}(\mathbf{u})) = 2\left(P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}\right) \ .$$

b. Prouver la caractérisation donnée de $P_{\mathcal{C}}(\mathbf{u})$.

On est en programmation convexe, et f est différentiable. La caractérisation de $P_{\mathcal{C}}(\mathbf{u})$ est donnée par :

$$\forall \mathbf{v} \in \mathcal{C}, \quad \langle \nabla f(P_{\mathcal{C}}(\mathbf{u})), \mathbf{v} - P_{\mathcal{C}}(\mathbf{u}) \rangle \geqslant 0 \ .$$

Or f et quadratique, $\nabla f(\mathbf{x}) = 2\mathbf{x} - 2\mathbf{u}$, donc :

$$\nabla f(P_{\mathcal{C}}(\mathbf{u})) = 2\,(P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}) \ .$$

On obtient donc :

$$\forall \mathbf{v} \in \mathcal{C}, \ 2\langle P_{\mathcal{C}}(\mathbf{u}) - \mathbf{u}, \mathbf{v} - P_{\mathcal{C}}(\mathbf{u}) \rangle \geqslant 0$$

et la caractérisation donnée en découle immédiatement.

c. Utiliser cette caractérisation pour prouver que $P_{\mathcal{C}}$ est une application contractante.

c. Utiliser cette caractérisation pour prouver que $P_{\mathcal{C}}$ est une application contractante.

En appliquant la caractérisation des points $P_{\mathcal{C}}(\mathbf{x})$ et $P_{\mathcal{C}}(\mathbf{y})$:

$$\langle P_{\mathcal{C}}(\mathbf{x}) - \mathbf{x}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant 0$$
$$\langle P_{\mathcal{C}}(\mathbf{y}) - \mathbf{y}, P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y})\rangle \geqslant 0$$

c. Utiliser cette caractérisation pour prouver que $P_{\mathcal{C}}$ est une application contractante.

En appliquant la caractérisation des points $P_{\mathcal{C}}(\mathbf{x})$ et $P_{\mathcal{C}}(\mathbf{y})$:

$$\langle P_{\mathcal{C}}(\mathbf{x}) - \mathbf{x}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant 0$$
$$\langle P_{\mathcal{C}}(\mathbf{y}) - \mathbf{y}, P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y})\rangle \geqslant 0$$

En additionnant ces deux inégalités :

$$< P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y}) - \mathbf{x} + \mathbf{y}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant 0 \ ,$$

c. Utiliser cette caractérisation pour prouver que $P_\mathcal{C}$ est une application contractante.

En appliquant la caractérisation des points $P_\mathcal{C}(\mathbf{x})$ et $P_\mathcal{C}(\mathbf{y})$:

$$\langle P_\mathcal{C}(\mathbf{x}) - \mathbf{x}, P_\mathcal{C}(\mathbf{y}) - P_\mathcal{C}(\mathbf{x})\rangle \geqslant 0$$
$$\langle P_\mathcal{C}(\mathbf{y}) - \mathbf{y}, P_\mathcal{C}(\mathbf{x}) - P_\mathcal{C}(\mathbf{y})\rangle \geqslant 0$$

En additionnant ces deux inégalités :

$$< P_\mathcal{C}(\mathbf{x}) - P_\mathcal{C}(\mathbf{y}) - \mathbf{x} + \mathbf{y}, P_\mathcal{C}(\mathbf{y}) - P_\mathcal{C}(\mathbf{x})\rangle \geqslant 0 \ ,$$

soit

$$\langle \mathbf{y} - \mathbf{x}, P_\mathcal{C}(\mathbf{y}) - P_\mathcal{C}(\mathbf{x})\rangle \geqslant \|P_\mathcal{C}(\mathbf{y}) - P_\mathcal{C}(\mathbf{x})\|^2$$

c. Utiliser cette caractérisation pour prouver que $P_{\mathcal{C}}$ est une application contractante.

En appliquant la caractérisation des points $P_{\mathcal{C}}(\mathbf{x})$ et $P_{\mathcal{C}}(\mathbf{y})$:

$$\langle P_{\mathcal{C}}(\mathbf{x}) - \mathbf{x}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant 0$$
$$\langle P_{\mathcal{C}}(\mathbf{y}) - \mathbf{y}, P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y})\rangle \geqslant 0$$

En additionnant ces deux inégalités :

$$< P_{\mathcal{C}}(\mathbf{x}) - P_{\mathcal{C}}(\mathbf{y}) - \mathbf{x} + \mathbf{y}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant 0 \ ,$$

soit

$$\langle \mathbf{y} - \mathbf{x}, P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant \|P_{\mathcal{C}}(\mathbf{y}) - P_{\mathcal{C}}(\mathbf{x})\|^2$$

et en appliquant l'inégalité de Cauchy-Schwartz :

$$\|\mathbf{y}-\mathbf{x}\|\,\|P_{\mathcal{C}}(\mathbf{y})-P_{\mathcal{C}}(\mathbf{x})\| \geqslant \langle \mathbf{y}-\mathbf{x}, P_{\mathcal{C}}(\mathbf{y})-P_{\mathcal{C}}(\mathbf{x})\rangle \geqslant \|P_{\mathcal{C}}(\mathbf{y})-P_{\mathcal{C}}(\mathbf{x})\|^2$$

dont on déduit l'inégalité recherchée.

Exercice 5

Le but de l'exercice est de prouver la **proposition** :

Soient $\mathcal{D} \subset \mathbb{R}^n$ *connexe,* $f : \mathcal{D} \longrightarrow \mathbb{R}$ *continue, et* $\mathbf{u} \in \mathcal{D}$ *un min (resp. max) local de* f*. Alors* $\mathbf{u}$ *est un min (resp. max) global de* f *ssi* $\forall\, \mathbf{x}$ *tel que* $f(\mathbf{x}) = f(\mathbf{u})$, $\mathbf{x}$ *est un min (resp. max) local de* f*.*

Exercice 5

Le but de l'exercice est de prouver la **proposition** :

Soient $\mathcal{D} \subset \mathbb{R}^n$ connexe, $f : \mathcal{D} \longrightarrow \mathbb{R}$ continue, et $\mathbf{u} \in \mathcal{D}$ un min (resp. max) local de f. Alors $\mathbf{u}$ est un min (resp. max) global de f ssi $\forall\, \mathbf{x}$ tel que $f(\mathbf{x}) = f(\mathbf{u})$, $\mathbf{x}$ est un min (resp. max) local de f.

Sans perte de généralité, quitte à changer f en $-f$, on la montrera pour $\mathbf{u}$ un min local.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D\, f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D\, f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

4. Déduire de 2 et 3 que $f^{-1}(]-\infty, u[)$ est un fermé de $\mathcal{D}$.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

Retour.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

Soit $r = v - u$, alors tout point x de la boule ouverte B de $\mathbb{R}$ centrée en v et de rayon r vérifie $x \geqslant u$,

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

Soit $r = v - u$, alors tout point x de la boule ouverte B de $\mathbb{R}$ centrée en v et de rayon r vérifie $x \geqslant u$, $\Longrightarrow$ $f^{-1}(B)$ est contenu dans $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ et contient $f^{-1}(\{v\})$.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

Soit $r = v - u$, alors tout point x de la boule ouverte B de $\mathbb{R}$ centrée en v et de rayon r vérifie $x \geqslant u$, $\Longrightarrow$ $f^{-1}(B)$ est contenu dans $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ et contient $f^{-1}(\{v\})$. Puisque B est un ouvert de $\mathbb{R}$ et f est continue, $f^{-1}(B)$ est un ouvert de $\mathcal{D}$.

Soit $u = f(\mathbf{u}) \in \mathbb{R}$.

1. Montrer que $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

Puisque $]-\infty, u[$ est un ouvert de $\mathbb{R}$ et que $f : \mathcal{D} \longrightarrow \mathbb{R}$ est continue, $f^{-1}(]-\infty, u[)$ est un ouvert de $\mathcal{D}$.

2. Montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$ pour $v > u$.

Soit $r = v - u$, alors tout point x de la boule ouverte B de $\mathbb{R}$ centrée en v et de rayon r vérifie $x \geqslant u$, $\Longrightarrow$ $f^{-1}(B)$ est contenu dans $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ et contient $f^{-1}(\{v\})$. Puisque B est un ouvert de $\mathbb{R}$ et f est continue, $f^{-1}(B)$ est un ouvert de $\mathcal{D}$. Donc $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ est un voisinage de tout point de $f^{-1}(\{v\})$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\Longrightarrow$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$,

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\implies$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$, et par définition c'est un ouvert de $\mathcal{D}$;

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\Longrightarrow$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$, et par définition c'est un ouvert de $\mathcal{D}$; $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ est donc un voisinage de $\mathbf{x}$ dans $\mathcal{D}$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\implies$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$, et par définition c'est un ouvert de $\mathcal{D}$; $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ est donc un voisinage de $\mathbf{x}$ dans $\mathcal{D}$.

4. Déduire de 2 et 3 que $f^{-1}(]-\infty, u[)$ est un fermé de $\mathcal{D}$.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D\, f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\implies$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}}\, f^{-1}(]-\infty, u[)$, et par définition c'est un ouvert de $\mathcal{D}$; $\complement_{\mathcal{D}}\, f^{-1}(]-\infty, u[)$ est donc un voisinage de $\mathbf{x}$ dans $\mathcal{D}$.

4. Déduire de 2 et 3 que $f^{-1}(]-\infty, u[)$ est un fermé de $\mathcal{D}$.

On déduit de 2 et 3 que $\complement_{\mathcal{D}}\, f^{-1}(]-\infty, u[)$ est un voisinage de tous ses points.

3. Soit $\mathbf{x} \in f^{-1}(\{u\})$; appliquer l'hypothèse que $\mathbf{x}$ est un min local pour montrer que $\complement_D f^{-1}(]-\infty, u[)$ est un voisinage de $\mathbf{x}$.

Soit $\mathbf{x} \in f^{-1}(\{u\})$; puisque $\mathbf{x}$ est un min local il existe un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ contenant $\mathbf{x}$ tel que $\forall \mathbf{y} \in \mathcal{U} \cap \mathcal{D}$, $f(\mathbf{y}) \geqslant f(\mathbf{x}) = u$. $\Longrightarrow$ $(\mathcal{U} \cap \mathcal{D}) \subset \complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$, et par définition c'est un ouvert de $\mathcal{D}$; $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ est donc un voisinage de $\mathbf{x}$ dans $\mathcal{D}$.

4. Déduire de 2 et 3 que $f^{-1}(]-\infty, u[)$ est un fermé de $\mathcal{D}$.

On déduit de 2 et 3 que $\complement_{\mathcal{D}} f^{-1}(]-\infty, u[)$ est un voisinage de tous ses points. C'est donc un ouvert de $\mathcal{D}$ et donc son complément $f^{-1}(]-\infty, u[)$ est un fermé de $\mathcal{D}$.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

On a montré en 1 et 4 que $f^{-1}(]-\infty, u[)$ est à la fois fermé et ouvert dans $\mathcal{D}$.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

On a montré en 1 et 4 que $f^{-1}(]-\infty, u[)$ est à la fois fermé et ouvert dans $\mathcal{D}$. Puisque $\mathcal{D}$ est connexe, $f^{-1}(]-\infty, u[)$ est soit $\emptyset$ soit $\mathcal{D}$.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

On a montré en 1 et 4 que $f^{-1}(]-\infty, u[)$ est à la fois fermé et ouvert dans $\mathcal{D}$. Puisque $\mathcal{D}$ est connexe, $f^{-1}(]-\infty, u[)$ est soit $\emptyset$ soit $\mathcal{D}$. Puisque $\mathbf{u} \in \mathcal{D}$ n'est pas dans $f^{-1}(]-\infty, u[)$, c'est l'ensemble vide.

5. Appliquer la connexité de $\mathcal{D}$ avec 1 et 4 pour montrer que $f^{-1}(]-\infty, u[) = \emptyset$. Conclure.

On a montré en 1 et 4 que $f^{-1}(]-\infty, u[)$ est à la fois fermé et ouvert dans $\mathcal{D}$. Puisque $\mathcal{D}$ est connexe, $f^{-1}(]-\infty, u[)$ est soit $\emptyset$ soit $\mathcal{D}$. Puisque $\mathbf{u} \in \mathcal{D}$ n'est pas dans $f^{-1}(]-\infty, u[)$, c'est l'ensemble vide. Ainsi, $\forall \mathbf{x} \in \mathcal{D}$, $f(\mathbf{x}) \geqslant f(\mathbf{u})$; $\mathbf{u}$ est donc un minimum global de f sur $\mathcal{D}$.

Retour.