# Fiche Méthodes - ST2CXO

*Un recueil de méthodes pour accompagner les révisions de Convex Optimization*

Dоryan Denis

2025-2026

# Table des matières

# 1 Convexité des fonctions

**Acquis d'Apprentissage.**

- Savoir calculer le gradient et la matrice Hessienne d'une fonction de plusieurs variables.
- Savoir caractériser la convexité et la convexité stricte d'une fonction à partir de sa Hessienne.
- Connaître les deux techniques : critère de Sylvester et étude des valeurs propres.

## 1.1 Calculer le gradient et la matrice Hessienne d'une fonction

**Méthode.**
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$.

1. **Gradient** $\nabla f(x)$ : vecteur colonne des dérivées partielles premières,

$$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}.$$

2. **Hessienne** $\nabla^2 f(x)$ : matrice $n \times n$ des dérivées partielles secondes,

$$\left(\nabla^2 f(x)\right)_{ij} = \frac{\partial^2 f}{\partial x_i \, \partial x_j} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

   Par le théorème de Schwarz, cette matrice est **symétrique**.

3. Vérifier si $\nabla^2 f$ est **constante** (elle ne dépend pas de $x$) : c'est souvent le cas pour les fonctions polynomiales du second degré, ce qui simplifie grandement l'étude.

**Remarque.**
Pour une fonction quadratique $f(x) = x^T Q x + a^T x + b$ avec $Q$ symétrique, on a directement $\nabla^2 f(x) = 2Q$ (matrice constante). En particulier, pour une fonction de la forme $f(x, y, z) = \alpha x^2 + \beta y^2 + \gamma z^2 +$ termes linéaires $+ c$, la Hessienne est diagonale avec $\nabla^2 f = \mathrm{diag}(2\alpha,\, 2\beta,\, 2\gamma)$.

**Exemple.** Application – Annale 2025, Exercice 2
Calculer le gradient et la Hessienne de

$$f(x, y, z) = x^2 + 3y^2 + 10z^2 - 4x - 24y - 20z + 25, \quad (x, y, z) \in \mathbb{R}^3.$$

**Solution.**
On dérive terme à terme par rapport à chaque variable :

$$\nabla f(x, y, z) = \begin{pmatrix} 2x - 4 \\ 6y - 24 \\ 20z - 20 \end{pmatrix}.$$

On dérive à nouveau pour obtenir la Hessienne :

$$\nabla^2 f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 20 \end{pmatrix}.$$

La Hessienne est diagonale et constante (elle ne dépend pas de $(x, y, z)$).

## 1.2 Déterminer la convexité d'une fonction via sa Hessienne

**Propriété.** Caractérisation par la Hessienne
Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^2$. Alors :

- $f$ est **convexe** sur $\mathbb{R}^n$ si et seulement si $\nabla^2 f(x)$ est **semi-définie positive** pour tout $x$.
- $f$ est **strictement convexe** sur $\mathbb{R}^n$ si $\nabla^2 f(x)$ est **définie positive** pour tout $x$.

**Méthode.** Technique 1 : valeurs propres (matrice diagonale ou symétrique)
Si $\nabla^2 f$ est **diagonale**, ses valeurs propres sont les éléments diagonaux.
Dans le cas général, on calcule les valeurs propres $\lambda_1, \ldots, \lambda_n$ de $\nabla^2 f$ et on conclut :

1. Si toutes les valeurs propres sont $> 0$ : $\nabla^2 f$ est **définie positive** $\Rightarrow$ $f$ est **strictement convexe**.
2. Si toutes les valeurs propres sont $\geq 0$ (avec au moins un $= 0$) : $\nabla^2 f$ est **semi-définie positive** $\Rightarrow$ $f$ est **convexe** (mais pas strictement).
3. Si toutes les valeurs propres sont $< 0$ : $\nabla^2 f$ est **définie négative** $\Rightarrow$ $f$ est **strictement concave**.
4. Si toutes les valeurs propres sont $\leq 0$ (avec au moins un $= 0$) : $\nabla^2 f$ est **semi-définie négative** $\Rightarrow$ $f$ est **concave** (mais pas strictement).
5. Si les valeurs propres sont **de signes mixtes** (certaines $> 0$, d'autres $< 0$) : $f$ n'est **ni convexe ni concave**.

Si $\nabla^2 f$ est constante, la conclusion vaut pour tout $x \in \mathbb{R}^n$.

**Remarque.** Calcul des valeurs propres d'une matrice non diagonale
Lorsque $\nabla^2 f$ n'est pas diagonale, les valeurs propres $\lambda$ se calculent en résolvant l'**équation caractéristique** :

$$\det\big(\nabla^2 f - \lambda I\big) = 0.$$

En pratique, pour une matrice $n \times n$ :

1. Former la matrice $\nabla^2 f - \lambda I$ en soustrayant $\lambda$ sur chaque élément diagonal.
2. Calculer son déterminant : on obtient un polynôme en $\lambda$ de degré $n$, appelé **polynôme caractéristique**.
3. Résoudre ce polynôme pour trouver les $n$ valeurs propres $\lambda_1, \ldots, \lambda_n$ (réelles car la Hessienne est symétrique).

**Exemple en dimension 2 :** pour $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$, l'équation caractéristique est :

$$\det(A - \lambda I) = (a - \lambda)(d - \lambda) - b^2 = 0$$

$$\Leftrightarrow \quad \lambda^2 - (a + d)\lambda + (ad - b^2) = 0.$$

On résout ce trinôme du second degré pour obtenir $\lambda_1$ et $\lambda_2$.

**Remarque pratique :** en dimension 3 et au-delà, le calcul du polynôme caractéristique devient fastidieux. C'est pourquoi le **critère de Sylvester** (méthode 1.2) est souvent préféré pour les cas stricts en dimension 3.

**Exemple.** Application n°1 – Annale 2025, Exercice 2
La fonction $f(x, y, z) = x^2 + 3y^2 + 10z^2 - 4x - 24y - 20z + 25$ est-elle convexe ? Strictement convexe ?

**Solution.**
On a calculé $\nabla^2 f = \text{diag}(2,\, 6,\, 20)$. La matrice est diagonale, donc les valeurs propres sont directement :

$$\lambda_1 = 2 > 0, \quad \lambda_2 = 6 > 0, \quad \lambda_3 = 20 > 0.$$

Toutes les valeurs propres sont strictement positives, et la Hessienne est constante.
Conclusion : $\nabla^2 f$ est définie positive pour tout $(x, y, z) \in \mathbb{R}^3$, donc $\boxed{f \text{ est strictement convexe sur } \mathbb{R}^3}$.

**Méthode.** Technique 2 : critère de Sylvester (matrice symétrique)
Le critère de Sylvester s'applique **uniquement aux matrices symétriques**. Il consiste à calculer les **mineurs principaux dominants** $m_k$ (déterminants des sous-matrices $k \times k$ en haut à gauche) :

1. $m_1 = a_{11}$
2. $m_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$
3. $m_3 = \det(\nabla^2 f)$ (en dimension 3)

**Conclusion :**

- Si tous les $m_k > 0$ : $\nabla^2 f$ est **définie positive** $\Rightarrow$ $f$ est **strictement convexe**.
- Si les $m_k$ alternent en signe ($m_1 < 0$, $m_2 > 0$, $m_3 < 0$, $\ldots$) : $\nabla^2 f$ est **définie négative** $\Rightarrow$ $f$ est **strictement concave**.
- Sinon : Sylvester **ne permet pas de conclure** $\Rightarrow$ utiliser la méthode des valeurs propres.

**Remarque.**
Deux points importants :

- **Condition d'application :** la matrice doit être symétrique. C'est toujours le cas de la Hessienne $\nabla^2 f$ (par le théorème de Schwarz), donc ce critère est systématiquement applicable dans ce cours.
- **Limite du critère :** Sylvester ne fonctionne que pour les cas **stricts** (définie positive ou négative). Pour conclure à la convexité simple ou à la concavité simple (matrices semi-définies), les mineurs dominants ne suffisent pas, il faut alors calculer les **valeurs propres** (tous $\geq 0 \Rightarrow$ convexe ; tous $\leq 0 \Rightarrow$ concave).

**Exemple.** Application n°2 – Annale 2024, Exercice 2
La fonction $f(x, y, z) = x^2 + y^2 + z^2 - xy - yz$ est-elle convexe ? Strictement convexe ?

**Solution.**
On calcule le gradient et la Hessienne :

$$\nabla f = \begin{pmatrix} 2x - y \\ 2y - x - z \\ 2z - y \end{pmatrix}, \qquad \nabla^2 f = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

La Hessienne est symétrique et constante. On applique le critère de Sylvester :

$$m_1 = 2 > 0,$$

$$m_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 4 - 1 = 3 > 0,$$

$$m_3 = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 2(4 - 1) + 1(-2 - 0) = 6 - 2 = 4 > 0.$$

Tous les mineurs sont strictement positifs.
Conclusion : $\nabla^2 f$ est définie positive pour tout $(x, y, z) \in \mathbb{R}^3$, donc $\boxed{f \text{ est strictement convexe sur } \mathbb{R}^3}$.

**Récap.**
Pour déterminer la convexité d'une fonction $f$ de classe $\mathcal{C}^2$ :

- Calculer $\nabla^2 f$ (Hessienne), toujours symétrique par le théorème de Schwarz.
- Si $\nabla^2 f$ est **diagonale** $\to$ valeurs propres = éléments diagonaux $\to$ conclure directement avec le tableau ci-dessous.
- Si $\nabla^2 f$ n'est **pas diagonale** :
  - cas **strict** recherché $\to$ critère de **Sylvester** (mineurs $m_1, m_2, m_3$) : rapide et suffisant.
  - cas **semi-défini** possible $\to$ calculer les **valeurs propres** via $\det(\nabla^2 f - \lambda I) = 0$ : seule méthode fiable.
- Si $\nabla^2 f$ est **constante**, la conclusion est valable sur tout $\mathbb{R}^n$.

| **Valeurs propres** | **Nature de $f$** |
|---|---|
| Toutes $> 0$ | Strictement convexe |
| Toutes $\geq 0$ | Convexe (pas strictement) |
| Toutes $< 0$ | Strictement concave |
| Toutes $\leq 0$ | Concave (pas strictement) |
| Signes mixtes | Ni convexe ni concave |

# 2

# Optimisation sans contrainte

**Acquis d'Apprentissage.**

- Savoir trouver les points stationnaires d'une fonction de plusieurs variables.
- Savoir conclure sur l'existence, la globalité et l'unicité d'un minimum.
- Savoir démontrer les deux théorèmes de cours les plus souvent demandés à l'examen.

## 2.1 Trouver les points stationnaires d'une fonction

**Méthode.**
Un point $\bar{x} \in \mathbb{R}^n$ est un **point stationnaire** de $f$ si $\nabla f(\bar{x}) = 0_{\mathbb{R}^n}$.

1. Calculer $\nabla f(x)$ (voir méthode 1.1).
2. Résoudre le système $\nabla f(x) = 0$, c'est-à-dire :

$$\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0, \quad \dots, \quad \frac{\partial f}{\partial x_n} = 0.$$

3. Résoudre le système par substitution ou par combinaisons linéaires jusqu'à trouver la (ou les) solution(s).
4. **Conclure** en indiquant le nombre de points stationnaires trouvés.

**Remarque.**
Si la Hessienne $\nabla^2 f$ est **diagonale**, les équations du système $\nabla f = 0$ sont **découplées** : chaque variable se résout indépendamment. C'est le cas le plus simple et le plus fréquent à l'examen.

**Exemple.** Application n°1 – Annale 2025, Exercice 2
Trouver les points stationnaires de

$$f(x, y, z) = x^2 + 3y^2 + 10z^2 - 4x - 24y - 20z + 25.$$

**Solution.**
On résout $\nabla f(x, y, z) = 0$ :

$$\begin{pmatrix} 2x-4 \\ 6y-24 \\ 20z-20 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 2x-4=0 \\ 6y-24=0 \\ 20z-20=0 \end{cases} \iff \begin{cases} x=2 \\ y=4 \\ z=1. \end{cases}$$

Les équations étant découplées (Hessienne diagonale), chaque variable se résout directement.
Conclusion : $f$ admet **un unique point stationnaire** $(\bar{x}, \bar{y}, \bar{z}) = \boxed{(2,\, 4,\, 1)}$.

**Exemple.** Application n°2 – Annale 2024, Exercice 2
Trouver les points stationnaires de

$$f(x, y, z) = x^2 + y^2 + z^2 - xy - yz.$$

**Solution.**
On résout $\nabla f(x, y, z) = 0$ :

$$\begin{cases} (1)\ 2x - y = 0 \\ (2)\ 2y - x - z = 0 \\ (3)\ 2z - y = 0. \end{cases}$$

De (1) : $y = 2x$. De (3) : $y = 2z$, donc $z = \frac{y}{2} = x$.
On substitue dans (2) : $2(2x) - x - x = 4x - 2x = 2x = 0 \Rightarrow x = 0$.
On en déduit $y = 0$ et $z = 0$.
Conclusion : $f$ admet **un unique point stationnaire** $\boxed{(0, 0, 0)}$.

## 2.2 Conclure sur l'existence et l'unicité d'un minimum global

**Propriété.** Minimum d'une fonction strictement convexe
Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable et **strictement convexe**.

- Si $f$ admet un point stationnaire $\bar{x}$ (i.e. $\nabla f(\bar{x}) = 0$), alors $\bar{x}$ est l'**unique minimum global** de $f$ sur $\mathbb{R}^n$ (et aussi minimum local).
- Le minimum est **unique** car $f$ est strictement convexe.
- Le minimum est **global** car pour une fonction convexe, tout minimum local est aussi global (voir méthode 2.3).

**Méthode.**
Pour conclure sur la nature des points stationnaires d'une fonction sans contrainte, deux cas se présentent selon que la Hessienne est constante ou non.

**Cas 1 : $\nabla^2 f$ constante (fonction globalement convexe/concave)**

1. Vérifier la convexité de $f$ sur $\mathbb{R}^n$ via $\nabla^2 f$ (chapitre 1) : la conclusion vaut partout.
2. Trouver le(s) point(s) stationnaire(s) (méthode 2.1).
3. Conclure directement selon la nature de $f$ :
   - $f$ **strictement convexe** $\to$ le point stationnaire est l'**unique minimum local, global et unique**.
   - $f$ **strictement concave** $\to$ le point stationnaire est l'**unique maximum local, global et unique**.
   - $f$ **convexe** (non stricte) $\to$ les points stationnaires sont des minima globaux, mais **pas nécessairement uniques**.

**Cas 2 : $\nabla^2 f$ non constante (fonction ni globalement convexe ni concave)**

1. Trouver **tous** les points stationnaires en résolvant $\nabla f = 0$ (le système peut être non linéaire $\to$ plusieurs solutions possibles).
2. Pour **chaque** point stationnaire $\bar{x}$, évaluer $\nabla^2 f(\bar{x})$ et étudier le signe de ses valeurs propres :
   - Toutes $> 0 \to \bar{x}$ est un **minimum local**.
   - Toutes $< 0 \to \bar{x}$ est un **maximum local**.
   - Signes mixtes $\to \bar{x}$ est un **point selle** (ni min ni max).
   - Au moins un $= 0 \to$ **test non concluant**, analyse complémentaire nécessaire.
3. Un minimum local n'est **pas nécessairement global** : sans convexité globale de $f$, on ne peut pas conclure à la globalité.

**Remarque.**
Le signal qui distingue les deux cas est simple : si $\nabla^2 f$ contient des termes en $x$, $y$ ou $z$, elle n'est pas constante $\to$ on est dans le Cas 2. Si $\nabla^2 f$ ne dépend d'aucune variable (ce qui est toujours le cas pour les fonctions polynomiales de degré $\leq 2$), on est dans le Cas 1.

**Exemple.** Application n°1 – Annale 2025, Exercice 2 (Cas 1)
Analyser l'existence d'un minimum local, global et unique de

$$f(x, y, z) = x^2 + 3y^2 + 10z^2 - 4x - 24y - 20z + 25.$$

**Solution.**
D'après la méthode 1.2 (Application n°1), $f$ est **strictement convexe** sur $\mathbb{R}^3$ ($\nabla^2 f = \operatorname{diag}(2, 6, 20)$, constante, définie positive).
D'après la méthode 2.1 (Application n°1), $f$ admet un unique point stationnaire $(2, 4, 1)$.
Conclusion : Comme $f$ est strictement convexe sur $\mathbb{R}^3$,

$$\boxed{(2,\ 4,\ 1) \text{ est l'unique minimum local, global et unique de } f \text{ sur } \mathbb{R}^3.}$$

**Exemple.** Application n°2 – Cas 2 : plusieurs points stationnaires
Analyser les points stationnaires de

$$f(x, y) = x^3 - 3x + y^2 - 4y.$$

**Solution.**
**1. Gradient et points stationnaires :**

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 - 3 \\ 2y - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} x^2 = 1 \\ y = 2 \end{cases} \iff \begin{cases} x = 1 \text{ ou } x = -1 \\ y = 2. \end{cases}$$

On trouve **deux points stationnaires** : $A = (1, 2)$ et $B = (-1, 2)$.

**2. Hessienne :**

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

**En $A = (1, 2)$ :**

$$\nabla^2 f(1, 2) = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \lambda_1 = 6 > 0,\ \lambda_2 = 2 > 0 \Rightarrow \text{définie positive}$$

$\Rightarrow A = (1, 2)$ est un **minimum local**.

**En $B = (-1, 2)$ :**

$$\nabla^2 f(-1, 2) = \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \lambda_1 = -6 < 0,\ \lambda_2 = 2 > 0 \Rightarrow \text{signes mixtes}$$

$\Rightarrow B = (-1, 2)$ est un **point selle** (ni minimum ni maximum).

Conclusion : $f$ n'est pas globalement convexe (Hessienne non constante). $\boxed{A = (1, 2) \text{ est un minimum local}}$ (non global), et $B = (-1, 2)$ est un point selle.

# 3
# Optimisation sous contrainte : théorème de Karush-Kuhn-Tucker

---

**Acquis d'Apprentissage.**

- Savoir vérifier les hypothèses du théorème KKT (convexité + condition de Slater).
- Savoir écrire le système KKT associé à un problème avec contraintes inégalité.
- Savoir résoudre le système KKT par disjonction de cas et conclure.

## 3.1 Vérifier les hypothèses du théorème KKT

**Propriété.** Théorème de Karush-Kuhn-Tucker
Considérons le problème convexe avec $m$ contraintes inégalité :

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{sous} \quad g_i(x) \leq 0,\ i = 1,\dots,m.$$

Supposons que :

- $f$ et les $g_i$ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$,
- $f$ et les $g_i$ sont **convexes**,
- la **condition de Slater** est satisfaite.

Alors $x^*$ est solution du problème **si et seulement si** il existe des multiplicateurs $\lambda_i^* \geq 0$ tels que le **système KKT** soit satisfait. La solution est un **minimum global**.

**Méthode.**
Avant d'appliquer le théorème KKT, vérifier **dans l'ordre** :

1. **Régularité :** $f$ et $g_i$ sont-elles de classe $\mathcal{C}^1$ ? (Presque toujours vrai pour les fonctions polynomiales.)
2. **Convexité de $f$ :** calculer $\nabla^2 f$ et vérifier qu'elle est définie positive (méthodes du chapitre 1). Préciser si $f$ est convexe ou strictement convexe.
3. **Convexité des contraintes $g_i$ :** une contrainte **linéaire** $g(x) = a^T x + b$ est toujours convexe (et affine, donc convexe mais pas strictement). Le vérifier explicitement.
4. **Condition de Slater :** exhiber un point $\hat{x} \in \mathbb{R}^n$ tel que $g_i(\hat{x}) < 0$ pour tout $i$. Il suffit de trouver **un exemple concret**, pas besoin de justification longue.

Si toutes ces conditions sont remplies, on peut appliquer le théorème KKT.

**Remarque.**
La condition de Slater exige un point **strictement** à l'intérieur du domaine faisable ($g_i < 0$, inégalité stricte). Pour une contrainte du type $g(x) = a^T x + b \leq 0$, il suffit de trouver un point où $a^T x + b < 0$, ce qui est toujours possible pour une contrainte non triviale. Choisir un point simple (avec des grandes valeurs ou des zéros) pour que le calcul soit immédiat.

**Exemple.** Application – Annale 2025, Exercice 3
Vérifier les hypothèses du théorème KKT pour le problème :

$$\min f(x,y,z) = (x-3)^2 + (y-2)^2 + (z-1)^2 \quad \text{sous} \quad x + 2y - 3z \leq -2.$$

**Solution.**
On pose $g(x,y,z) = x + 2y - 3z + 2$, de sorte que la contrainte s'écrit $g(x,y,z) \leq 0$.
**1. Régularité :** $f$ et $g$ sont des fonctions polynomiales, donc de classe $\mathcal{C}^\infty \subset \mathcal{C}^1$. ✓

**2. Convexité de $f$ :**

$$\nabla^2 f = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2I_3.$$

La Hessienne est diagonale et constante, avec $\lambda_1 = \lambda_2 = \lambda_3 = 2 > 0$. Donc $f$ est **strictement convexe** sur $\mathbb{R}^3$. ✓

**3. Convexité de $g$ :** $g(x, y, z) = x + 2y - 3z + 2$ est une fonction **affine** (linéaire + constante), donc convexe. ✓

**4. Condition de Slater :** on cherche $(\hat{x}, \hat{y}, \hat{z})$ tel que $g(\hat{x}, \hat{y}, \hat{z}) < 0$.
Choisissons $\hat{x} = -10,\ \hat{y} = 0,\ \hat{z} = 6$ :

$$g(-10, 0, 6) = -10 + 0 - 18 + 2 = -26 < 0. \checkmark$$

Toutes les hypothèses sont vérifiées : on peut appliquer le théorème KKT.

### 3.2 Écrire le système KKT

**Méthode.**
Une fois les hypothèses vérifiées, le système KKT pour le problème

$$\min f(x) \quad \text{sous} \quad g_i(x) \leq 0,\ i = 1, \dots, m$$

s'écrit en introduisant un multiplicateur $\lambda_i^* \geq 0$ pour chaque contrainte $g_i$ :

1. **Stationnarité :**
$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0_{\mathbb{R}^n}$$
(un vecteur-équation $\Rightarrow$ $n$ équations scalaires).
2. **Complémentarité :**
$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m.$$
3. **Faisabilité primale :**
$$g_i(x^*) \leq 0, \quad i = 1, \dots, m.$$
4. **Faisabilité duale :**
$$\lambda_i^* \geq 0, \quad i = 1, \dots, m.$$

**Remarque importante :** avec $m$ contraintes, on a $m$ multiplicateurs et donc $m + 1$ cas à étudier dans la disjonction (méthode 3.3). Pour $m = 1$, il y a exactement **2 cas** : $\lambda = 0$ ou $\lambda > 0$.

**Exemple.** Application – Annale 2025, Exercice 3
Écrire le système KKT du problème :

$$\min f(x, y, z) = (x-3)^2 + (y-2)^2 + (z-1)^2 \quad \text{sous} \quad g(x, y, z) = x + 2y - 3z + 2 \leq 0.$$

**Solution.**
On calcule les gradients :

$$\nabla f(x, y, z) = \begin{pmatrix} 2(x-3) \\ 2(y-2) \\ 2(z-1) \end{pmatrix}, \qquad \nabla g(x, y, z) = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}.$$

Le système KKT s'écrit ($m = 1$ contrainte, un seul multiplicateur $\lambda \geq 0$) :

$$\begin{pmatrix} 2(x-3) \\ 2(y-2) \\ 2(z-1) \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
$$\lambda\,(x + 2y - 3z + 2) = 0$$
$$x + 2y - 3z + 2 \leq 0$$

$$\lambda \geq 0$$

Ce qui donne les équations numérotées :

$$\begin{aligned}
&(1) \quad 2(x-3)+\lambda=0 \\
&(2) \quad 2(y-2)+2\lambda=0 \\
&(3) \quad 2(z-1)-3\lambda=0 \\
&(4) \quad \lambda\,(x+2y-3z+2)=0 \\
&(5) \quad x+2y-3z+2 \leq 0 \\
&(6) \quad \lambda \geq 0
\end{aligned}$$

### 3.3 Résoudre le système KKT par disjonction de cas

**Méthode.**
La condition de complémentarité $\lambda_i^* g_i(x^*) = 0$ impose, pour chaque contrainte $i$ :

$$\text{soit } \lambda_i^* = 0 \quad \text{(contrainte inactive)}, \quad \text{soit } g_i(x^*) = 0 \quad \text{(contrainte active)}.$$

Avec $m$ contraintes, on étudie $2^m$ **cas**. Pour $m = 1$, il y a donc 2 cas à étudier :

1. **Cas 1 :** $\lambda = 0$
   - Simplifier les équations de stationnarité avec $\lambda = 0$.
   - Résoudre pour trouver un candidat $(\bar{x}, \bar{y}, \bar{z})$.
   - **Vérifier la faisabilité primale** (5) : calculer $g(\bar{x}, \bar{y}, \bar{z})$ et vérifier $\leq 0$.
   - Si (5) n'est pas satisfaite : <u>ce cas est rejeté</u>.
2. **Cas 2 :** $\lambda > 0$
   - La complémentarité (4) impose $g(x^*) = 0$ (contrainte active).
   - Exprimer $x, y, z$ en fonction de $\lambda$ depuis les équations de stationnarité.
   - Substituer dans $g(x^*) = 0$ pour trouver la valeur de $\lambda$.
   - Vérifier que $\lambda > 0$ (cohérence avec l'hypothèse du cas).
   - Vérifier la faisabilité primale (5) : avec $g = 0$, on a $0 \leq 0$, automatiquement satisfait.

**Conclusion :** le candidat retenu (celui dont toutes les conditions sont satisfaites) est l'**unique minimum global** du problème, car le problème est convexe et la solution KKT est à la fois nécessaire et suffisante.

**Exemple.** Application n°1 – Annale 2025, Exercice 3
Résoudre le système KKT du problème :

$$\min f(x,y,z) = (x-3)^2 + (y-2)^2 + (z-1)^2 \quad \text{sous} \quad x+2y-3z+2 \leq 0$$

dont le système KKT a été écrit à la méthode 3.2.

**Solution.**
**Cas 1 :** $\lambda = 0$

Les équations (1), (2), (3) donnent :

$$(1) \Rightarrow x = 3, \qquad (2) \Rightarrow y = 2, \qquad (3) \Rightarrow z = 1.$$

On vérifie la faisabilité primale (5) :

$$g(3,2,1) = 3+4-3+2 = 6 > 0.$$

La contrainte (5) n'est **pas satisfaite**. <u>Ce cas est rejeté.</u>

**Cas 2 :** $\lambda > 0$

La complémentarité (4) impose $g(x^*, y^*, z^*) = 0$, c'est-à-dire la contrainte est **active**.

Des équations (1), (2), (3) on tire $x$, $y$, $z$ en fonction de $\lambda$ :

$$(1) \Rightarrow x = 3 - \frac{\lambda}{2}, \qquad (2) \Rightarrow y = 2 - \lambda, \qquad (3) \Rightarrow z = 1 + \frac{3\lambda}{2}.$$

On substitue dans $g(x^*, y^*, z^*) = 0$ :

$$\left(3 - \frac{\lambda}{2}\right) + 2\,(2-\lambda) - 3\left(1 + \frac{3\lambda}{2}\right) + 2 = 0.$$

On développe :

$$3 - \frac{\lambda}{2} + 4 - 2\lambda - 3 - \frac{9\lambda}{2} + 2 = 0$$

$$6 - \frac{\lambda + 4\lambda + 9\lambda}{2} = 0 \;\Rightarrow\; 6 - 7\lambda = 0 \;\Rightarrow\; \boxed{\lambda = \frac{6}{7} > 0.}\checkmark$$

On calcule la solution :

$$x^* = 3 - \frac{3}{7} = \frac{18}{7}, \qquad y^* = 2 - \frac{6}{7} = \frac{8}{7}, \qquad z^* = 1 + \frac{9}{7} = \frac{16}{7}.$$

On vérifie la faisabilité primale (5) : $g = 0 \leq 0$. ✓

Toutes les conditions KKT sont satisfaites par $\left(\frac{18}{7}, \frac{8}{7}, \frac{16}{7}\right)$ et $\lambda^* = \frac{6}{7}$.

Conclusion : le problème étant convexe (hypothèses vérifiées à la méthode 3.1), le point

$$\boxed{(x^*, y^*, z^*) = \left(\frac{18}{7}, \frac{8}{7}, \frac{16}{7}\right)}$$

est l'**unique minimum local et global** du problème.

**Exemple.** Application n°2 – Annale 2024, Exercice 3
Résoudre le problème :

$$\min f(x,y,z) = x^2 + y^2 + z^2 + 1 \quad \text{sous} \quad x + y + z \leq -3.$$

**Solution.**
On pose $g(x,y,z) = x+y+z+3$. Les hypothèses KKT sont satisfaites : $f$ est strictement convexe ($\nabla^2 f = 2I_3$), $g$ est affine donc convexe, et Slater est vérifié avec par exemple $(0,0,-5)$ : $g(0,0,-5) = -2 < 0$.
Les gradients sont :

$$\nabla f = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}, \qquad \nabla g = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Le système KKT donne :

$$\begin{array}{ll} (1) & 2x + \lambda = 0 \\ (2) & 2y + \lambda = 0 \\ (3) & 2z + \lambda = 0 \\ (4) & \lambda(x+y+z+3) = 0 \\ (5) & x+y+z+3 \leq 0 \\ (6) & \lambda \geq 0 \end{array}$$

**Cas 1 :** $\lambda = 0$
$(1), (2), (3) \Rightarrow x = y = z = 0.$
Vérification de (5) : $0+0+0+3 = 3 \not\leq 0$. Cas rejeté.

**Cas 2 :** $\lambda > 0$
$(4) \Rightarrow x + y + z + 3 = 0.$

$(1), (2), (3) \Rightarrow x = y = z = -\dfrac{\lambda}{2}$.
On substitue dans $g = 0$ :

$$3 \times \left(-\frac{\lambda}{2}\right) + 3 = 0 \;\Rightarrow\; -\frac{3\lambda}{2} = -3 \;\Rightarrow\; \boxed{\lambda = 2 > 0.} \checkmark$$

Donc $x^* = y^* = z^* = -1$.
Vérification (5) : $-1 - 1 - 1 + 3 = 0 \leq 0$. ✓

Conclusion : $\boxed{(x^*, y^*, z^*) = (-1,\ -1,\ -1)}$ est l'**unique minimum global** du problème.

**Récap.**
Schéma de résolution d'un problème KKT avec 1 contrainte inégalité :

1. **Hypothèses** : $f$ convexe, $g$ convexe, Slater vérifié → KKT applicable.
2. **Calcul** : gradients $\nabla f$ et $\nabla g$.
3. **Système KKT** : stationnarité + complémentarité + faisabilité primale + dualité.
4. **Cas 1 ($\lambda = 0$)** : résoudre, vérifier (5) → souvent rejeté car le minimum libre ne satisfait pas la contrainte.
5. **Cas 2 ($\lambda > 0$)** : contrainte active ($g = 0$), exprimer $x, y, z$ en $\lambda$, substituer, trouver $\lambda$, vérifier $\lambda > 0$.
6. **Conclusion** : le candidat retenu est l'unique minimum global.

# Régression polynomiale

**Acquis d'Apprentissage.**

- Savoir modéliser un problème de régression polynomiale comme un problème d'optimisation sans contrainte.
- Savoir identifier les variables d'optimisation, la fonction objectif et la dimension du problème.

## 4.1 Formuler un problème de régression polynomiale

**Définition.** Régression polynomiale
Étant donné un nuage de $n$ points $P_1(x_1, y_1), \dots, P_n(x_n, y_n)$ dans $\mathbb{R}^2$, la **régression polynomiale de degré** $\leq d$ consiste à déterminer les coefficients d'un polynôme

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d$$

de façon à minimiser la **somme des carrés des résidus** entre les valeurs prédites $P(x_i)$ et les valeurs observées $y_i$.

**Méthode.**
Pour mettre en forme un problème de régression polynomiale de degré $\leq d$ :

1. **Identifier les variables d'optimisation :** ce sont les $d+1$ coefficients du polynôme, notés $(a_0, a_1, \dots, a_d) \in \mathbb{R}^{d+1}$.
2. **Écrire la fonction objectif :** c'est la somme des carrés des écarts entre chaque valeur observée $y_i$ et la valeur prédite $P(x_i)$ :
$$f(a_0, \dots, a_d) = \sum_{i=1}^{n} \big(y_i - P(x_i)\big)^2 = \sum_{i=1}^{n} \big(y_i - a_0 - a_1 x_i - \cdots - a_d x_i^d\big)^2.$$
3. **Formuler le problème :** il s'agit d'un problème d'optimisation **sans contrainte** en dimension $d+1$ :
$$\min_{(a_0, a_1, \dots, a_d) \in \mathbb{R}^{d+1}} f(a_0, a_1, \dots, a_d).$$
4. **Justifier la résolution (si demandé) :** la fonction $f$ est une somme de carrés de fonctions affines en $(a_0, \dots, a_d)$, donc $f$ est **convexe** (et même strictement convexe sous condition de non-dégénérescence des points). Un minimum global existe et est unique ; il s'obtient en résolvant $\nabla f = 0$.

**Remarque.**
À l'examen, **on ne demande jamais de résoudre** ce problème — uniquement de l'écrire. Les points sont attribués pour :

- la bonne identification des variables (les coefficients, pas les $x_i$),
- la forme correcte de la fonction objectif (somme de carrés),
- la mention explicite que c'est un problème **sans contrainte** en dimension $d+1$.

Ne pas confondre $n$ (nombre de points donnés, fixé) et $d+1$ (nombre d'inconnues à optimiser).

**Exemple.** Application – Annale 2025, Exercice 4
On dispose de $n$ points $P_1(x_1, y_1), \dots, P_n(x_n, y_n)$ dans $\mathbb{R}^2$. Écrire le problème d'optimisation consistant à trouver le polynôme de degré $\leq 4$ qui approche au mieux ce nuage de points.

**Solution.**
Un polynôme de degré $\leq 4$ s'écrit :

$$P(x) = a + bx + cx^2 + dx^3 + ex^4, \qquad (a, b, c, d, e) \in \mathbb{R}^5.$$

Les **variables d'optimisation** sont les 5 coefficients $(a, b, c, d, e)$.

La **fonction objectif** est la somme des carrés des résidus :

$$\begin{aligned} f(a, b, c, d, e) &= \sum_{i=1}^{n} \left(y_i - a - bx_i - cx_i^2 - dx_i^3 - ex_i^4\right)^2 \\ &= (y_1 - a - bx_1 - cx_1^2 - dx_1^3 - ex_1^4)^2 \\ &\quad + \cdots \\ &\quad + (y_n - a - bx_n - cx_n^2 - dx_n^3 - ex_n^4)^2. \end{aligned}$$

Le problème est :

$$\boxed{\min_{(a,\,b,\,c,\,d,\,e)\,\in\,\mathbb{R}^5} f(a, b, c, d, e).}$$

C'est un problème d'optimisation **sans contrainte** en **dimension 5**.

**Récap.**
Régression polynomiale de degré $\leq d$ sur $n$ points :

| | |
|---|---|
| **Variables** | $d + 1$ coefficients $(a_0, \ldots, a_d) \in \mathbb{R}^{d+1}$ |
| **Objectif** | $\min \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i - \cdots - a_d x_i^d)^2$ |
| **Type** | Sans contrainte, dimension $d + 1$ |
| **Convexité** | $f$ est convexe (somme de carrés de fonctions affines) |

# 5

# Démonstrations de cours

**Acquis d'Apprentissage.**

- Savoir restituer entièrement les démonstrations des résultats fondamentaux du cours.
- Connaître précisément les hypothèses sous lesquelles chaque résultat est valable.
- Repérer les démonstrations qui sont systématiquement demandées en question de cours (Exercice 1, 4 pts).

## 5.1 Démontrer que les hyperplans et les demi-espaces sont des ensembles convexes

**Propriété.** Convexité des hyperplans et demi-espaces
Soit $p \in \mathbb{R}^n$ un vecteur non nul fixé et $\alpha \in \mathbb{R}$. Alors :

- l'hyperplan $H = \{x \in \mathbb{R}^n : p^T x = \alpha\}$ est convexe ;
- le demi-espace $H = \{x \in \mathbb{R}^n : p^T x \leq \alpha\}$ (ou $p^T x \geq \alpha$) est convexe.

**Méthode.** Démonstration
**1. Cas de l'hyperplan.**
Soient $x_1, x_2 \in H$, c'est-à-dire $p^T x_1 = \alpha$ et $p^T x_2 = \alpha$. Soit $\lambda \in [0,1]$ et considérons le point $x = \lambda x_1 + (1-\lambda)x_2$. Par linéarité du produit scalaire :

$$p^T x = p^T\big(\lambda x_1 + (1-\lambda)x_2\big) = \lambda\, p^T x_1 + (1-\lambda)\, p^T x_2 = \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

Donc $x \in H$. <u>$H$ est donc convexe.</u>

**2. Cas du demi-espace $\{x : p^T x \leq \alpha\}$.**
Soient $x_1, x_2 \in H$, c'est-à-dire $p^T x_1 \leq \alpha$ et $p^T x_2 \leq \alpha$. Soit $\lambda \in [0,1]$ et $x = \lambda x_1 + (1-\lambda)x_2$. Comme $\lambda \geq 0$ et $(1-\lambda) \geq 0$, on peut multiplier les inégalités sans changer leur sens :

$$\lambda\, p^T x_1 \leq \lambda\alpha, \qquad (1-\lambda)\, p^T x_2 \leq (1-\lambda)\alpha.$$

En additionnant ces deux inégalités :

$$p^T x = \lambda\, p^T x_1 + (1-\lambda)\, p^T x_2 \leq \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

Donc $x \in H$. <u>$H$ est donc convexe.</u>
Le raisonnement est rigoureusement identique pour $\{x : p^T x \geq \alpha\}$ (il suffit d'inverser le sens des inégalités).
□

## 5.2 Démontrer que l'intersection non vide d'ensembles convexes est convexe

**Propriété.** Convexité de l'intersection
Soit $(C_i)_{i\in I}$ une famille (finie ou infinie) d'ensembles convexes de $\mathbb{R}^n$, telle que $C = \bigcap_{i\in I} C_i \neq \emptyset$. Alors $C$ est un ensemble convexe.

**Attention :** l'union d'ensembles convexes n'est en général **pas** convexe (voir remarque ci-dessous).

**Méthode.** Démonstration
Soient $x, y \in C$. Par définition de l'intersection, $x \in C_i$ et $y \in C_i$ pour **tout** $i \in I$. Soit $\lambda \in [0,1]$. Comme chaque $C_i$ est convexe, on a, pour tout $i \in I$ :

$$\lambda x + (1-\lambda)y \in C_i.$$

Ceci étant vrai pour tout $i \in I$, on en déduit :

$$\lambda x + (1-\lambda)y \in \bigcap_{i \in I} C_i = C.$$

$C$ est donc stable par combinaison convexe, c'est-à-dire convexe. □

Le raisonnement ne dépend pas du nombre (fini ou infini) d'ensembles dans la famille : il s'applique aussi bien à une intersection de deux ensembles qu'à une intersection infinie.

**Remarque.** L'union d'ensembles convexes n'est pas convexe en général
**Contre-exemple :** soient $C_1 = \{0\}$ et $C_2 = \{1\}$ dans $\mathbb{R}$, deux singletons (donc convexes, trivialement). Leur union $C_1 \cup C_2 = \{0,1\}$ n'est **pas convexe** : en prenant $x_1 = 0 \in C_1 \cup C_2$ et $x_2 = 1 \in C_1 \cup C_2$ avec $\lambda = \frac{1}{2}$, le point milieu $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$ n'appartient pas à $C_1 \cup C_2$.

Plus généralement, dès que deux ensembles convexes sont disjoints (ou que le segment reliant un point de l'un à un point de l'autre sort des deux ensembles), leur union n'est pas convexe.

## 5.3 Démontrer que l'épigraphe de f est convexe si et seulement si f est convexe

**Propriété.** Caractérisation de la convexité par l'épigraphe
Soit $f : C \to \mathbb{R}$ une fonction définie sur un ensemble convexe $C \subset \mathbb{R}^n$. Rappelons que l'**épigraphe** de $f$ est défini par :

$$\mathrm{epi}(f) = \{(x, r) \in C \times \mathbb{R} \;:\; r \geq f(x)\}.$$

Alors $\mathrm{epi}(f)$ est un ensemble convexe **si et seulement si** $f$ est une fonction convexe sur $C$.

**Méthode.** Démonstration
**Sens direct ($\mathrm{epi}(f)$ convexe $\Rightarrow$ $f$ convexe) :**
Soient $x, y \in C$. Les points $(x, f(x))$ et $(y, f(y))$ appartiennent tous deux à $\mathrm{epi}(f)$ (l'inégalité $r \geq f(x)$ étant satisfaite avec égalité). Soit $\lambda \in [0,1]$. Par convexité de $\mathrm{epi}(f)$, la combinaison convexe de ces deux points appartient encore à $\mathrm{epi}(f)$ :

$$\lambda\,(x, f(x)) + (1-\lambda)\,(y, f(y)) = \big(\lambda x + (1-\lambda)y,\ \lambda f(x) + (1-\lambda)f(y)\big) \in \mathrm{epi}(f).$$

Par définition de $\mathrm{epi}(f)$, cela signifie exactement :

$$\lambda f(x) + (1-\lambda)f(y) \geq f\big(\lambda x + (1-\lambda)y\big).$$

C'est précisément l'inégalité de convexité de $f$. □

**Sens réciproque ($f$ convexe $\Rightarrow$ $\mathrm{epi}(f)$ convexe) :**
Soient $(x, r_1), (y, r_2) \in \mathrm{epi}(f)$, c'est-à-dire $r_1 \geq f(x)$ et $r_2 \geq f(y)$. Soit $\lambda \in [0,1]$. On veut montrer que $\lambda(x, r_1) + (1-\lambda)(y, r_2) = \big(\lambda x + (1-\lambda)y,\ \lambda r_1 + (1-\lambda)r_2\big)$ appartient à $\mathrm{epi}(f)$, c'est-à-dire :

$$\lambda r_1 + (1-\lambda)r_2 \geq f\big(\lambda x + (1-\lambda)y\big).$$

Comme $\lambda \geq 0$ et $(1-\lambda) \geq 0$, les inégalités $r_1 \geq f(x)$ et $r_2 \geq f(y)$ se combinent ainsi :

$$\lambda r_1 + (1-\lambda)r_2 \;\geq\; \lambda f(x) + (1-\lambda)f(y).$$

Par convexité de $f$ :

$$\lambda f(x) + (1-\lambda)f(y) \;\geq\; f\big(\lambda x + (1-\lambda)y\big).$$

En combinant les deux inégalités précédentes :

$$\lambda r_1 + (1-\lambda)r_2 \;\geq\; f\big(\lambda x + (1-\lambda)y\big),$$

ce qui montre que le point combiné appartient à $\mathrm{epi}(f)$. $\mathrm{epi}(f)$ est donc convexe. □

**Remarque.**
Cette équivalence est très utile en pratique : elle permet de visualiser géométriquement la convexité d'une fonction comme la convexité de la région **au-dessus** de son graphe. Elle sert aussi de point de départ pour démontrer d'autres propriétés de stabilité de la convexité (par exemple, le fait que le maximum ponctuel d'une famille de fonctions convexes est convexe, car $\mathrm{epi}(\max_\alpha f_\alpha) = \bigcap_\alpha \mathrm{epi}(f_\alpha)$, intersection d'ensembles convexes).

## 5.4 Démontrer que la convexité forte implique la convexité stricte, qui implique la convexité

**Propriété.** Hiérarchie des convexités
Soit $f : \mathbb{R}^n \to \mathbb{R}$. Alors :

$$f \text{ fortement convexe} \implies f \text{ strictement convexe} \implies f \text{ convexe},$$

mais aucune des deux implications réciproques n'est vraie en général.

**Méthode.** Démonstration
**1. Convexité forte ⇒ convexité stricte.**
Par définition, $f$ est $\alpha$-fortement convexe ($\alpha > 0$) si la fonction $h(x) = f(x) - \frac{\alpha}{2}\|x\|^2$ est convexe. Pour tout $x, y \in \mathbb{R}^n$ et tout $\lambda \in (0,1)$, la convexité de $h$ donne :

$$h(\lambda x + (1-\lambda)y) \leq \lambda h(x) + (1-\lambda)h(y),$$

soit :

$$f(\lambda x + (1-\lambda)y) - \frac{\alpha}{2}\|\lambda x + (1-\lambda)y\|^2 \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\alpha}{2}\left[\lambda\|x\|^2 + (1-\lambda)\|y\|^2\right].$$

On réorganise :

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\alpha}{2}\underbrace{\left[\lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \|\lambda x + (1-\lambda)y\|^2\right]}_{(\star)}.$$

On calcule explicitement le terme $(\star)$ :

$$\begin{aligned}(\star) &= \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \lambda^2\|x\|^2 - 2\lambda(1-\lambda)\, x\cdot y - (1-\lambda)^2\|y\|^2 \\ &= \lambda(1-\lambda)\|x\|^2 + \lambda(1-\lambda)\|y\|^2 - 2\lambda(1-\lambda)\, x\cdot y \\ &= \lambda(1-\lambda)\big(\|x\|^2 + \|y\|^2 - 2x\cdot y\big) \\ &= \lambda(1-\lambda)\|x-y\|^2.\end{aligned}$$

Donc, pour $x \neq y$ et $\lambda \in (0,1)$, on a $(\star) = \lambda(1-\lambda)\|x-y\|^2 > 0$. On en déduit :

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\alpha}{2}\lambda(1-\lambda)\|x-y\|^2 < \lambda f(x) + (1-\lambda)f(y)$$

pour tout $x \neq y$ et $\lambda \in (0,1)$. C'est exactement la définition de la stricte convexité.

**2. Convexité stricte ⇒ convexité.**
C'est immédiat : l'inégalité stricte $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$ implique trivialement l'inégalité large $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$, qui est la définition de la convexité. □

**Remarque.** Les réciproques sont fausses
**Convexe n'implique pas strictement convexe :** $f(x) = ax$ est une fonction affine, donc convexe (et même concave), mais l'inégalité de convexité est toujours une égalité : $f$ n'est pas strictement convexe.

**Strictement convexe n'implique pas fortement convexe :** $g(x) = x^4$ est strictement convexe sur $\mathbb{R}$ (car $g''(x) = 12x^2 > 0$ pour $x \neq 0$), mais elle n'est pas fortement convexe : $g''(0) = 0$, donc aucun $\alpha > 0$ ne peut vérifier $g''(x) \geq \alpha$ pour tout $x$ (la courbure s'annule en $0$).

### 5.5 Démontrer la caractérisation du premier ordre de la convexité

**Propriété.** Caractérisation du premier ordre
Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable sur un domaine ouvert. Alors $f$ est convexe si et seulement si :

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \qquad \forall x, y \in \mathrm{dom}(f).$$

Géométriquement : le graphe de $f$ est partout au-dessus de chacune de ses tangentes.

**Méthode.** Démonstration
**Sens direct ($f$ convexe $\Rightarrow$ inégalité) :**
Par convexité, pour tout $\lambda \in [0, 1]$ :

$$f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda)f(x).$$

On réécrit le terme de gauche $f(x + \lambda(y - x))$, ce qui donne :

$$f(x + \lambda(y - x)) \leq f(x) + \lambda\big(f(y) - f(x)\big).$$

On réarrange et on divise par $\lambda \in (0, 1]$ :

$$f(y) - f(x) \geq \frac{f(x + \lambda(y - x)) - f(x)}{\lambda}.$$

En faisant tendre $\lambda \to 0^+$, le membre de droite tend vers la dérivée directionnelle de $f$ en $x$ dans la direction $(y - x)$, c'est-à-dire $\nabla f(x)^T (y - x)$. On obtient :

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

**Sens réciproque (inégalité $\Rightarrow$ $f$ convexe) :**
Fixons $x, y \in \mathrm{dom}(f)$, $\lambda \in [0, 1]$, et posons $z = \lambda x + (1 - \lambda)y$. On applique l'inégalité de l'hypothèse deux fois, en $x$ et en $y$, avec le point $z$ :

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

On multiplie la première par $\lambda$, la seconde par $(1 - \lambda)$, puis on additionne :

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(z) + \nabla f(z)^T \big(\lambda(x - z) + (1 - \lambda)(y - z)\big).$$

Or $\lambda(x - z) + (1 - \lambda)(y - z) = \lambda x + (1 - \lambda)y - z = z - z = 0$ (par définition de $z$), donc le terme en $\nabla f(z)^T (\cdots)$ s'annule. On obtient :

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(z) = f(\lambda x + (1 - \lambda)y),$$

ce qui est exactement la définition de la convexité. $\square$

### 5.6 Démontrer la caractérisation du second ordre de la convexité

**Propriété.** Caractérisation du second ordre
Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable sur un domaine ouvert. Alors $f$ est convexe si et seulement si sa Hessienne $\nabla^2 f(x)$ est **semi-définie positive** pour tout $x \in \mathrm{dom}(f)$.

**Méthode.** Démonstration (cas unidimensionnel $n = 1$, puis généralisation)
**(ii) $\Rightarrow$ (iii) : la condition du premier ordre implique $f'' \geq 0$.**
Soient $x, y \in \mathrm{dom}(f)$ avec $y > x$. La condition du premier ordre (méthode précédente) appliquée dans les deux sens donne :

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

En réorganisant la seconde inégalité ($f(y) \leq f(x) - f'(y)(x - y) = f(x) + f'(y)(y - x)$) et en combinant avec la première :

$$f'(x)(y - x) \leq f(y) - f(x) \leq f'(y)(y - x).$$

On divise par $(y-x)^2 > 0$ :
$$\frac{f'(y)-f'(x)}{y-x} \geq 0, \qquad \forall x \neq y.$$
En faisant tendre $y \to x$, on obtient $f''(x) \geq 0$ pour tout $x \in \mathrm{dom}(f)$.

**(iii) ⇒ (ii) : $f'' \geq 0$ implique la condition du premier ordre.**
On suppose $f''(x) \geq 0$ pour tout $x$. D'après la formule de Taylor avec reste de Lagrange, pour $x, y \in \mathrm{dom}(f)$, il existe $z$ entre $x$ et $y$ tel que :

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2} f''(z)(y-x)^2.$$

Comme $f''(z) \geq 0$, le dernier terme est $\geq 0$, donc :

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

**Généralisation à $\mathbb{R}^n$ :**
On utilise le théorème de convexité le long des droites : $f$ est convexe sur $\mathbb{R}^n$ si et seulement si, pour tout $x_0 \in \mathrm{dom}(f)$ et tout $v \in \mathbb{R}^n$, la fonction univariée $g(t) = f(x_0 + tv)$ est convexe. On applique alors le résultat unidimensionnel précédent à $g$, en remarquant que $g''(t) = v^T \nabla^2 f(x_0 + tv)\, v$. La condition $g''(t) \geq 0$ pour tout $t$ et tout $v$ équivaut exactement à : $\nabla^2 f(x)$ est semi-définie positive pour tout $x$. $\square$

## 5.7 Démontrer la condition nécessaire d'optimalité du premier ordre

**Propriété.** Condition nécessaire d'optimalité (sans hypothèse de convexité)
Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable. Si $x^*$ est un **minimum local** de $f$, alors $\nabla f(x^*) = 0$.

**Méthode.** Démonstration
Soit $\varepsilon > 0$ assez petit et $v \in \mathbb{R}^n$ un vecteur fixé. Comme $x^*$ est un minimum local de $f$, il existe une boule $B$ autour de $x^*$ telle que :
$$f(x^*) \leq f(x^* + \varepsilon v), \qquad \text{pour } x^* + \varepsilon v \in B.$$
Le développement de Taylor à l'ordre 1 de $f$ en $x^*$ s'écrit :

$$f(x^* + \varepsilon v) = f(x^*) + \varepsilon\, \nabla f(x^*)^T v + o(\varepsilon).$$

En combinant les deux relations :
$$0 \leq \varepsilon\, \nabla f(x^*)^T v + o(\varepsilon).$$
On divise par $\varepsilon > 0$ et on fait tendre $\varepsilon \to 0^+$ :

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

On applique exactement le même raisonnement en remplaçant $v$ par $-v$ :

$$\nabla f(x^*)^T(-v) \geq 0 \quad \Longrightarrow \quad \nabla f(x^*)^T v \leq 0.$$

Les deux inégalités $\nabla f(x^*)^T v \geq 0$ et $\nabla f(x^*)^T v \leq 0$ donnent :

$$\nabla f(x^*)^T v = 0.$$

Ceci étant vrai pour **tout** vecteur $v \in \mathbb{R}^n$ (en particulier pour $v = e_1, \ldots, e_n$ les vecteurs de la base canonique), on conclut :
$$\nabla f(x^*) = 0. \quad \square$$

**Remarque.**
La réciproque est **fausse en général** : $\nabla f(x^*) = 0$ n'implique pas que $x^*$ soit un minimum local. Exemples : $f(x) = -x^2$ vérifie $f'(0) = 0$ mais $x = 0$ est un **maximum** ; $f(x) = x^3$ vérifie $f'(0) = 0$ mais $x = 0$ est un **point selle** (ni minimum ni maximum). C'est uniquement sous hypothèse de **convexité** que la réciproque devient vraie (méthode suivante).

## 5.8 Démontrer que pour une fonction convexe, minimum local équivaut à minimum global équivaut à gradient nul

**Propriété.** Cas convexe
Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable et **convexe**. Alors :

$$x^* \text{ est un minimum local de } f \iff \nabla f(x^*) = 0 \iff x^* \text{ est un minimum global de } f.$$

**Méthode.** Démonstration (annale 2024, Ex.1)
**1. Minimum local $\Rightarrow \nabla f(x^*) = 0$.**
C'est la condition nécessaire d'optimalité du premier ordre (méthode précédente), valable pour toute fonction différentiable, donc en particulier pour $f$ convexe.

**2. Minimum local $\Rightarrow$ minimum global (grâce à la convexité).**
Supposons que $x^*$ soit un minimum local : il existe une boule $B$ autour de $x^*$ telle que $f(x^*) \leq f(x)$ pour tout $x \in B$. Soit $x \in \mathbb{R}^n$ quelconque. Il existe $\lambda \in (0, 1]$ assez petit pour que $\lambda x + (1 - \lambda)x^*$ soit dans $B$. Comme $x^*$ est minimum local :

$$f(x^*) \leq f\big(\lambda x + (1 - \lambda)x^*\big).$$

Par convexité de $f$ :

$$f\big(\lambda x + (1 - \lambda)x^*\big) \leq \lambda f(x) + (1 - \lambda)f(x^*).$$

En combinant les deux inégalités :

$$f(x^*) \leq \lambda f(x) + (1 - \lambda)f(x^*) \implies \lambda f(x^*) \leq \lambda f(x) \implies f(x^*) \leq f(x).$$

Ceci étant vrai pour tout $x \in \mathbb{R}^n$, $x^*$ est un minimum global.

**3. $\nabla f(x^*) = 0 \Rightarrow$ minimum global (condition suffisante).**
Supposons $\nabla f(x^*) = 0$. Par la caractérisation du premier ordre de la convexité (méthode 5.2), $f$ est au-dessus de sa tangente en tout point :

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*) = f(x^*) + 0 = f(x^*), \qquad \forall x \in \mathbb{R}^n.$$

Donc $x^*$ est un minimum global, ce qui implique en particulier que c'est aussi un minimum local. □

**Remarque.**
C'est **la démonstration de l'Exercice 1 de l'annale 2024** (4 pts). Il faut maîtriser les trois étapes : condition nécessaire (gradient nul, valable sans convexité), passage local → global **grâce à la convexité**, puis condition suffisante via l'inégalité du premier ordre.

## 5.9 Démontrer que l'ensemble des minimiseurs d'une fonction convexe est convexe

**Propriété.** Convexité de l'ensemble des minimiseurs
Soit $f : \mathbb{R}^n \to \mathbb{R}$ convexe. L'ensemble

$$\operatorname{argmin} f = \{x \in \mathbb{R}^n \ : \ f(x) = \min_{\mathbb{R}^n} f\}$$

est un **ensemble convexe** (il n'est pas nécessairement réduit à un singleton).

**Méthode.** Démonstration
Soient $x_1, x_2 \in \operatorname{argmin} f$, c'est-à-dire $f(x_1) = f(x_2) = \min f$. Soit $\lambda \in [0,1]$ et considérons le point $x = \lambda x_1 + (1-\lambda)x_2$ sur le segment reliant $x_1$ à $x_2$. Par convexité de $f$ :

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \le \lambda f(x_1) + (1-\lambda) f(x_2) = \min f.$$

Mais par définition de $\min f$, on a aussi $f(x) \ge \min f$ pour tout $x$. On en déduit :

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) = \min f,$$

c'est-à-dire que $\lambda x_1 + (1-\lambda)x_2$ est lui-même un minimiseur de $f$. L'ensemble $\operatorname{argmin} f$ est donc stable par combinaison convexe, ce qui signifie qu'il est convexe. □

## 5.10 Démontrer l'unicité du minimum d'une fonction strictement convexe

**Propriété.** Unicité du minimum pour $f$ strictement convexe
Le minimum sur $\mathbb{R}^n$ d'une fonction **strictement convexe** $f$ est **unique** (s'il existe).

**Méthode.** Démonstration par l'absurde (annale 2025, Ex.1)
On raisonne par l'absurde :

1. Supposons qu'il existe **deux minimiseurs distincts** $x_1 \neq x_2$ tels que
$$f(x_1) = f(x_2) = \min_{\mathbb{R}^n} f.$$
2. Considérons le point milieu $\bar{x} = \dfrac{x_1 + x_2}{2}$.
3. Par **stricte convexité** de $f$ (avec $\lambda = \frac{1}{2}$ et $x_1 \neq x_2$) :
$$f(\bar{x}) = f\left(\frac{x_1 + x_2}{2}\right) < \frac{f(x_1)}{2} + \frac{f(x_2)}{2} = \frac{\min f + \min f}{2} = \min f.$$
4. On obtient $f(\bar{x}) < \min f$ : **contradiction**, car $\min f$ est par définition la plus petite valeur prise par $f$ sur $\mathbb{R}^n$.
5. Conclusion : l'hypothèse de départ est fausse, donc le minimiseur est unique. □

**Remarque.**
La clé de cette démonstration est l'utilisation de la **stricte** convexité (et non la convexité simple) avec $\lambda = \frac{1}{2}$ exactement au point milieu, ce qui produit une inégalité **stricte** contredisant la définition du minimum. Avec une convexité simple (non stricte), on n'obtiendrait qu'une inégalité large, sans contradiction : c'est cohérent avec le résultat précédent montrant que l'ensemble des minimiseurs peut être infini dans le cas convexe non strict.

## 5.11 Démontrer le théorème de Karush-Kuhn-Tucker (condition suffisante)

**Propriété.** Théorème KKT – rappel des hypothèses et de la conclusion
Soit le problème convexe :

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sous} \quad g_i(x) \le 0,\ i = 1, \dots, m,$$

avec $f$ et les $g_i$ convexes et de classe $\mathcal{C}^1$. Si $(x^*, \lambda^*)$ vérifie les conditions KKT (stationnarité, complémentarité, faisabilité primale, faisabilité duale), alors $x^*$ est un **minimum global** du problème.

**Méthode.** Démonstration – suffisance
Soit $C = \{x \in \mathbb{R}^n : g_i(x) \le 0,\ i = 1, \dots, m\}$ l'ensemble faisable. Comme chaque $g_i$ est convexe, l'ensemble de sous-niveau $\{x : g_i(x) \le 0\}$ est convexe, et $C$ (intersection de tels ensembles) est donc **convexe**.

Notons $A = \{i : \lambda_i^* > 0\}$ l'ensemble des contraintes **actives**. Par complémentarité ($\lambda_i^* g_i(x^*) = 0$) : pour

$i \in A$, $g_i(x^*) = 0$ ; pour $i \notin A$, $\lambda_i^* = 0$. La condition de stationnarité se réduit donc à :

$$\nabla f(x^*) + \sum_{i \in A} \lambda_i^* \nabla g_i(x^*) = 0.$$

Soit $x \in C$ quelconque. Comme $C$ est convexe, pour tout $t \in [0, 1]$ : $x^* + t(x - x^*) \in C$, donc en particulier pour $i \in A$ :

$$g_i\big(x^* + t(x - x^*)\big) \le 0 = g_i(x^*), \qquad \forall t \in [0, 1].$$

On en déduit, pour $t \in (0, 1]$ :

$$\frac{g_i\big(x^* + t(x - x^*)\big) - g_i(x^*)}{t} \le 0,$$

et en faisant tendre $t \to 0^+$ :

$$\nabla g_i(x^*)^T (x - x^*) \le 0, \qquad \forall i \in A.$$

Comme $\lambda_i^* > 0$ pour $i \in A$, on multiplie et on somme :

$$\sum_{i \in A} \lambda_i^* \, \nabla g_i(x^*)^T (x - x^*) \le 0.$$

D'après la stationnarité, $\nabla f(x^*) = -\sum_{i \in A} \lambda_i^* \nabla g_i(x^*)$, donc :

$$\nabla f(x^*)^T (x - x^*) = -\sum_{i \in A} \lambda_i^* \nabla g_i(x^*)^T (x - x^*) \ge 0.$$

Enfin, par convexité de $f$ (caractérisation du premier ordre, méthode 5.2) :

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

Ceci étant vrai pour **tout** $x \in C$, on conclut que $x^*$ est un **minimum global** de $f$ sur $C$. $\square$

**Remarque.** Sur la nécessité des conditions KKT

La **nécessité** (existence des multiplicateurs $\lambda_i^*$ pour toute solution optimale $x^*$, sous condition de Slater) repose sur le **lemme de Farkas** : pour une matrice $A$ $(m \times n)$ et $b \in \mathbb{R}^m$, exactement une des deux propositions suivantes est vraie :

1. $\exists\, x \ge 0$ tel que $Ax = b$,
2. $\exists\, \lambda \in \mathbb{R}^m$ tel que $A^T \lambda \le 0$ et $b^T \lambda > 0$.

Ce lemme permet d'établir la nécessité dans le cas de contraintes affines $Ax \le b$, puis le résultat se généralise aux contraintes convexes sous condition de Slater. Cette partie de la démonstration est plus technique et moins probable comme question de cours isolée ; **la démonstration de la suffisance ci-dessus est la plus susceptible d'être demandée à l'examen**.