# III.2. - Optimisation sous contraintes

Jean-Philippe Préaux

EOAA - 2009/10

• Nous avons vu durant le dernier cours comment procéder en optimisation sous contraintes égalitaires.

Copyright : Jean-Philippe Préaux

• Nous avons vu durant le dernier cours comment procéder en optimisation sous contraintes égalitaires. L'équation d'Euler se généralise par les conditions de Lagrange.

Copyright : Jean-Philippe Préaux

• Nous avons vu durant le dernier cours comment procéder en optimisation sous contraintes égalitaires. L'équation d'Euler se généralise par les conditions de Lagrange.

• Nous voyons ici comment procéder dans le cas général de contraintes égalitaires et/ou inégalitaires.

Copyright . Jean-Philippe Préaux

• Nous avons vu durant le dernier cours comment procéder en optimisation sous contraintes égalitaires. L'équation d'Euler se généralise par les conditions de Lagrange.

• Nous voyons ici comment procéder dans le cas général de contraintes égalitaires et/ou inégalitaires. Les conditions de Lagrange se généralisent par les conditions de Karush-Kuhn-Tucker (ou KKT).

Copyright - Jean-Philippe Préaux

Conditions de Karush-Kuhn-Tucker
Le théorème de Karush-Kuhn-Tucker
Notation lagrangienne
Prise en compte de la convexité
Programmation convexe
Qualification de contraintes affines
Programmation quadratique sous contraintes
Théorème de programmation quadratique
Exemple
Conditions du second ordre.
Le cône tangent
Condition suffisante du $2^e$ ordre
Le sous-cône tangent
Condition nécessaire du $2^e$ ordre
Points-selles du lagrangien
Exercices
Exercice 1
Exercice 2
Exercice 3

Copyright : Jean-Philippe Préaux

# Cadre d'application

Soit $\mathcal{U}$ un ouvert non vide de $\mathbb{R}^n$ (souvent, $\mathcal{U} = \mathbb{R}^n$) ; soit $f : \mathcal{U} \longrightarrow \mathbb{R}$ une application différentiable. Soient $\varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q : \mathcal{U} \longrightarrow \mathbb{R}$ des applications de classe $C^1$ $(p, q \geqslant 0)$.

Copyright : Jean-Philippe Préaux

# Cadre d'application

Soit $\mathcal{U}$ un ouvert non vide de $\mathbb{R}^n$ (souvent, $\mathcal{U} = \mathbb{R}^n$) ; soit $f : \mathcal{U} \longrightarrow \mathbb{R}$ une application différentiable. Soient $\varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q : \mathcal{U} \longrightarrow \mathbb{R}$ des applications de classe $C^1$ $(p, q \geqslant 0)$. Et soit :

$$\mathcal{D} = \{\mathbf{x} \in \mathcal{U} \mid \underbrace{\varphi_i(\mathbf{x}) = 0, \forall i = 1, \ldots, p}_{\text{contraintes égalitaires}} ; \underbrace{\psi_j(\mathbf{x}) \leqslant 0, \forall j = 1, \ldots, q}_{\text{contraintes inégalitaires}}\}$$

Copyright - Jean-Philippe Préaux

# Cadre d'application

Soit $\mathcal{U}$ un ouvert non vide de $\mathbb{R}^n$ (souvent, $\mathcal{U} = \mathbb{R}^n$) ; soit $f : \mathcal{U} \longrightarrow \mathbb{R}$ une application différentiable. Soient $\varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q : \mathcal{U} \longrightarrow \mathbb{R}$ des applications de classe $C^1$ $(p, q \geqslant 0)$. Et soit :

$$\mathcal{D} = \{\mathbf{x} \in \mathcal{U} \mid \underbrace{\varphi_i(\mathbf{x}) = 0, \forall i = 1, \ldots, p}_{\text{contraintes égalitaires}} ; \underbrace{\psi_j(\mathbf{x}) \leqslant 0, \forall j = 1, \ldots, q}_{\text{contraintes inégalitaires}}\}$$

Comment résoudre :

$$\min_{\mathbf{x} \in \mathcal{D}} f(\mathbf{x}) \ ?$$

Copyright - Jean-Philippe Préaux

# Hypothèse de qualification des contraintes

Pour énoncer le théorème nous avons encore besoin d'une hypothèse de qualification des contraintes en $\mathbf{u} \in \mathcal{D}$ :

$(QC)_{\mathbf{u}}$ : $\{\nabla\varphi_1(\mathbf{u}), \ldots, \nabla\varphi_p(\mathbf{u})\} \cup \{\nabla\psi_j(\mathbf{u}) \mid \psi_j(\mathbf{u}) = 0\}$ est une famille de vecteurs linéairement indépendante.

Copyright : Jean-Philippe Préaux

# Théorème de Karush-Kuhn-Tucker.

## Théorème (Conditions nécessaires de Karush-Kuhn-Tucker.)

*Sous les hypothèses énoncées ci-dessus, si* $\mathbf{u} \in \mathcal{D}$ *est un minimum local de* $f$ *sur* $\mathcal{D}$*, et si l'hypothèse* $(QC)_{\mathbf{u}}$ *est vérifiée, alors* $\exists\,!\,\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q$ *tels que :*

(i) $\nabla f(\mathbf{u}) + \sum_{i=1}^{p} \lambda_i \nabla \varphi_i(\mathbf{u}) + \sum_{j=1}^{q} \mu_j \nabla \psi_j(\mathbf{u}) = \mathbf{0}$

(ii) $\forall j = 1, \ldots, q,\ \mu_j \geq 0$

(iii) $\forall j = 1, \ldots, q,\ \mu_j = 0$ *si* $\psi_j(\mathbf{u}) < 0$.

Copyright . Jean-Philippe Préaux

# Théorème de Karush-Kuhn-Tucker.

## Théorème (Conditions nécessaires de Karush-Kuhn-Tucker.)

*Sous les hypothèses énoncées ci-dessus, si* $\mathbf{u} \in \mathcal{D}$ *est un minimum local de* $f$ *sur* $\mathcal{D}$*, et si l'hypothèse* $(QC)_{\mathbf{u}}$ *est vérifiée, alors* $\exists\,!\,\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q$ *tels que :*

(i) $\nabla f(\mathbf{u}) + \sum_{i=1}^{p} \lambda_i \nabla \varphi_i(\mathbf{u}) + \sum_{j=1}^{q} \mu_j \nabla \psi_j(\mathbf{u}) = \mathbf{0}$

(ii) $\forall j = 1, \ldots, q,\ \mu_j \geq 0$

(iii) $\forall j = 1, \ldots, q,\ \mu_j = 0$ *si* $\psi_j(\mathbf{u}) < 0$.

La condition (*iii*) peut aussi s'écrire :

$$\sum_{j=1}^{q} \mu_j \psi_j(\mathbf{u}) = \mathbf{0}\ .$$

Copyright : Jean-Philippe Préaux

# Esquisse de preuve.

Copyright : Jean-Philippe Préaux

# Esquisse de preuve.

On se ramène à un problème d'optimisation à contrainte égalitaire en ajoutant $q$ variables $\mathbf{y} = (y_1, \ldots, y_q)$ :

$$\begin{cases} \min\limits_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) \\ \varphi_i(\mathbf{x}) = 0 & \forall\, i = 1, \ldots, p \\ \psi_j(\mathbf{x}) + y_j^2 = 0 & \forall\, j = 1, \ldots, q \end{cases}$$

Copyright : Jean-Philippe Préaux

# Esquisse de preuve.

On se ramène à un problème d'optimisation à contrainte égalitaire en ajoutant $q$ variables $\mathbf{y} = (y_1, \ldots, y_q)$ :

$$\begin{cases} \min\limits_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) \\ \varphi_i(\mathbf{x}) = 0 & \forall\, i = 1, \ldots, p \\ \psi_j(\mathbf{x}) + y_j^2 = 0 & \forall\, j = 1, \ldots, q \end{cases}$$

son domaine $\mathcal{D}'$ est inclus dans $\mathcal{D} \times \mathbb{R}^q \subset \mathbb{R}^{n+q}$, et il a pour minimum local $(\mathbf{u}, \mathbf{y})$ si et seulement si $\mathbf{u}$ est un minimum local de $f$ sur $\mathcal{D}$.

Copyright : Jean-Philippe Préaux

# Esquisse de preuve.

On se ramène à un problème d'optimisation à contrainte égalitaire en ajoutant $q$ variables $\mathbf{y} = (y_1, \ldots, y_q)$ :

$$\begin{cases} \min\limits_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) \\ \varphi_i(\mathbf{x}) = 0 & \forall\, i = 1, \ldots, p \\ \psi_j(\mathbf{x}) + y_j^2 = 0 & \forall\, j = 1, \ldots, q \end{cases}$$

son domaine $\mathcal{D}'$ est inclus dans $\mathcal{D} \times \mathbb{R}^q \subset \mathbb{R}^{n+q}$, et il a pour minimum local $(\mathbf{u}, \mathbf{y})$ si et seulement si $\mathbf{u}$ est un minimum local de $f$ sur $\mathcal{D}$. On lui applique les conditions de Lagrange. En $\mathbf{u} \in \mathcal{D}$, l'hypothèse $(QC)_{\mathbf{u}}$ nous assure que l'hypothèse de qualification des contraintes en $(\mathbf{u}, \mathbf{y})$ est satisfaite.

Copyright : Jean-Philippe Préaux

Notons $\mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu)$ le lagrangien de ce problème et $\mathcal{L}(\mathbf{u}, \lambda, \mu)$ le lagrangien du problème obtenu en prenant $y_1 = \cdots = y_q = 0$.

Copyright : Jean-Philippe Préaux

Notons $\mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu)$ le lagrangien de ce problème et $\mathcal{L}(\mathbf{u}, \lambda, \mu)$ le lagrangien du problème obtenu en prenant $y_1 = \cdots = y_q = 0$.
On obtient $\exists\, \lambda_1, \ldots, \lambda_p,\ \mu_1, \ldots, \mu_q$, tels que :

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \nabla f(\mathbf{u}) \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{i=1}^{p} \lambda_i \begin{pmatrix} \nabla \varphi_i(\mathbf{u}) \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{j=1}^{q} \mu_j \begin{pmatrix} \nabla \psi_j(\mathbf{u}) \\ \hline 0 \\ 2y_j \\ 0 \end{pmatrix} = \mathbf{0}$$

Copyright : Jean-Philippe Préaux

Notons $\mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu)$ le lagrangien de ce problème et $\mathcal{L}(\mathbf{u}, \lambda, \mu)$ le lagrangien du problème obtenu en prenant $y_1 = \cdots = y_q = 0$.
On obtient $\exists\, \lambda_1, \ldots, \lambda_p,\ \mu_1, \ldots, \mu_q$, tels que :

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \nabla f(\mathbf{u}) \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{i=1}^{p} \lambda_i \begin{pmatrix} \nabla \varphi_i(\mathbf{u}) \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{j=1}^{q} \mu_j \begin{pmatrix} \nabla \psi_j(\mathbf{u}) \\ \hline 0 \\ 2y_j \\ 0 \end{pmatrix} = \mathbf{0}$$

En particulier on obtient la condition $(i)$, et de plus $\forall\, j = 1, \ldots, q,\ \mu_j y_j = 0$, c'est à dire $\mu_j = 0$ dès que $\psi_j(\mathbf{u}) < 0$, c'est la condition $(iii)$.

Copyright : Jean-Philippe Préaux

Pour montrer la condition (*ii*), on se restreint au cas où les applications sont deux fois différentiables. On applique la condition nécessaire à l'ordre 2.

Copyright : Jean-Philippe Préaux

Pour montrer la condition (*ii*), on se restreint au cas où les applications sont deux fois différentiables. On applique la condition nécessaire à l'ordre 2.
On a la matrice :

$$\nabla_x^2 \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \boxed{\nabla_x^2 \mathcal{L}(\mathbf{u}, \lambda, \mu)} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{matrix} & \begin{matrix} 2\mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2\mu_q \end{matrix} \end{pmatrix}$$

Copyright : Jean-Philippe Préaux

Pour montrer la condition $(ii)$, on se restreint au cas où les applications sont deux fois différentiables. On applique la condition nécessaire à l'ordre 2.
On a la matrice :

$$\nabla^2_x \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \boxed{\nabla^2_x \mathcal{L}(\mathbf{u}, \lambda, \mu)} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{matrix} & \begin{matrix} 2\mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2\mu_q \end{matrix} \end{pmatrix}$$

Or puisque $\mu_j y_j = 0$, $y_j = 0$ dès que $\mu_j \neq 0$. En particulier le vecteur $\mathbf{e}_j$ de $\mathbb{R}^{n+q}$ dont la $(n+j)$-ème coordonnée est 1 et toutes les autres sont nulles est dans l'espace tangent à $\mathcal{D}'$ en $(\mathbf{u}, \mathbf{y})$.

Copyright : Jean-Philippe Préaux

Pour montrer la condition (*ii*), on se restreint au cas où les applications sont deux fois différentiables. On applique la condition nécessaire à l'ordre 2.
On a la matrice :

$$\nabla_x^2 \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \boxed{\nabla_x^2 \mathcal{L}(\mathbf{u}, \lambda, \mu)} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{matrix} & \begin{matrix} 2\mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2\mu_q \end{matrix} \end{pmatrix}$$

Or puisque $\mu_j y_j = 0$, $y_j = 0$ dès que $\mu_j \neq 0$. En particulier le vecteur $\mathbf{e}_j$ de $\mathbb{R}^{n+q}$ dont la $(n+j)$-ème coordonnée est 1 et toutes les autres sont nulles est dans l'espace tangent à $\mathcal{D}'$ en $(\mathbf{u}, \mathbf{y})$. Puisque $\mathbf{e}_j^\top \nabla_{\mathbf{x}}^2 \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) \mathbf{e}_j = 2\mu_j$, la condition nécessaire du second ordre implique que $\mu_j \geqslant 0$. On obtient donc la condition (*ii*).

Copyright : Jean-Philippe Préaux

Pour montrer la condition $(ii)$, on se restreint au cas où les applications sont deux fois différentiables. On applique la condition nécessaire à l'ordre 2.
On a la matrice :

$$\nabla^2_x \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) = \begin{pmatrix} \boxed{\nabla^2_x \mathcal{L}(\mathbf{u}, \lambda, \mu)} & & & & 0 & \cdots & 0 \\ & & & & \vdots & & \vdots \\ & & & & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 2\mu_1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 2\mu_q \end{pmatrix}$$

Or puisque $\mu_j y_j = 0$, $y_j = 0$ dès que $\mu_j \neq 0$. En particulier le vecteur $\mathbf{e}_j$ de $\mathbb{R}^{n+q}$ dont la $(n+j)$-ème coordonnée est 1 et toutes les autres sont nulles est dans l'espace tangent à $\mathcal{D}'$ en $(\mathbf{u}, \mathbf{y})$. Puisque $\mathbf{e}_j^\top \nabla^2_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \mathbf{y}, \lambda, \mu) \mathbf{e}_j = 2\mu_j$, la condition nécessaire du second ordre implique que $\mu_j \geqslant 0$. On obtient donc la condition $(ii)$.
L'unicité des multiplicateurs est une conséquence immédiate de l'hypothèse de qualification des contraintes. □

Copyright : Jean-Philippe Préaux

# Remarques.

Copyright : Jean-Philippe Préaux

# Remarques.

- Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.

Copyright : Jean-Philippe Préaux

# Remarques.

- Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.
- Pour un problème de maximum il suffit de changer la condition $(ii)$ en $(ii')$ : $\forall\, j = 1, \ldots, q\,, \mu_j \leqslant 0$.

Copyright : Jean-Philippe Préaux

# Remarques.

- Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.
- Pour un problème de maximum il suffit de changer la condition $(ii)$ en $(ii')$ : $\forall\, j = 1, \ldots, q\,, \mu_j \leqslant 0$.
- Une contrainte inégalitaire $\psi_j$ est dite *insaturée* ou *inactive* en $\mathbf{u}$ si $\psi_j(\mathbf{u}) < 0$, et sinon elle est dite *saturée* ou *active*.

Copyright : Jean Philippe Préaux

# Remarques.

• Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.

• Pour un problème de maximum il suffit de changer la condition $(ii)$ en $(ii')$ : $\forall\, j = 1, \ldots, q\,, \mu_j \leqslant 0$.

• Une contrainte inégalitaire $\psi_j$ est dite *insaturée* ou *inactive* en $\mathbf{u}$ si $\psi_j(\mathbf{u}) < 0$, et sinon elle est dite *saturée* ou *active*.

• Dans le cas d'une contrainte insaturée $\psi_j(\mathbf{u}) \neq 0$, le coefficient de Lagrange-KKT correspondant est nul : $\mu_j = 0$, c'est-à-dire que cette contrainte ne compte pas.

Copyright : Jean Philippe Préaux

# Remarques.

• Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.

• Pour un problème de maximum il suffit de changer la condition $(ii)$ en $(ii')$ : $\forall\, j = 1, \ldots, q\,, \mu_j \leqslant 0$.

• Une contrainte inégalitaire $\psi_j$ est dite *insaturée* ou *inactive* en $\mathbf{u}$ si $\psi_j(\mathbf{u}) < 0$, et sinon elle est dite *saturée* ou *active*.

• Dans le cas d'une contrainte insaturée $\psi_j(\mathbf{u}) \neq 0$, le coefficient de Lagrange-KKT correspondant est nul : $\mu_j = 0$, c'est-à-dire que cette contrainte ne compte pas. Lorsque toutes les contraintes inégalitaires sont insaturées en $\mathbf{u}$ on retrouve les conditions de Lagrange.

Copyright - Jean Philippe Préaux

# Remarques.

• Les $\lambda_i, \mu_j$ sont appelés les *multiplicateurs de Lagrange-KKT* ou *multiplicateurs de Lagrange généralisés*.

• Pour un problème de maximum il suffit de changer la condition $(ii)$ en $(ii')$ : $\forall\, j = 1, \ldots, q\,, \mu_j \leqslant 0$.

• Une contrainte inégalitaire $\psi_j$ est dite *insaturée* ou *inactive* en $\mathbf{u}$ si $\psi_j(\mathbf{u}) < 0$, et sinon elle est dite *saturée* ou *active*.

• Dans le cas d'une contrainte insaturée $\psi_j(\mathbf{u}) \neq 0$, le coefficient de Lagrange-KKT correspondant est nul : $\mu_j = 0$, c'est-à-dire que cette contrainte ne compte pas. Lorsque toutes les contraintes inégalitaires sont insaturées en $\mathbf{u}$ on retrouve les conditions de Lagrange.

C'était prévisible, l'ensemble des points de $\mathbb{R}^n$ où toutes les contraintes inégalitaires sont insaturées est un ouvert $\mathcal{U}$ et l'on est dans le cadre d'application du théorème de Lagrange, son domaine n'étant plus défini sur $\mathcal{U}$ que par les contraintes égalitaires.

Copyright - Jean Philippe Préaux

# Notation Lagrangienne.

Comme pour un problème d'optimisation sous contraintes égalitaires, on parle du lagrangien d'un problème d'optimisation sous contraintes égalitaires et inégalitaires.

Copyright : Jean-Philippe Préaux

# Notation Lagrangienne.

Comme pour un problème d'optimisation sous contraintes égalitaires, on parle du lagrangien d'un problème d'optimisation sous contraintes égalitaires et inégalitaires. C'est l'application :

$$\mathcal{L} \quad : \quad \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q \longmapsto \mathbb{R}$$

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i \varphi_i(\mathbf{x}) + \sum_{j=1}^{q} \mu_j \psi_j(\mathbf{x}) \ .$$

Copyright : Jean-Philippe Préaux

# Notation Lagrangienne.

Comme pour un problème d'optimisation sous contraintes égalitaires, on parle du lagrangien d'un problème d'optimisation sous contraintes égalitaires et inégalitaires. C'est l'application :

$$\mathcal{L} \quad : \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q \longmapsto \mathbb{R}$$

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i \varphi_i(\mathbf{x}) + \sum_{j=1}^{q} \mu_j \psi_j(\mathbf{x}) \ .$$

Lorsque toutes les applications sont différentiables, le vecteur gradient du lagrangien en $\mathbf{u} \in \mathbb{R}^n$ est :

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \lambda, \mu) = \nabla f(\mathbf{u}) + \sum_{i=1}^{p} \lambda_i \nabla \varphi_i(\mathbf{u}) + \sum_{j=1}^{q} \mu_j \nabla \psi_j(\mathbf{u}) \ .$$

Copyright : Jean-Philippe Préaux

La condition $(i)$ de KKT s'écrit alors, sous les hypothèses adéquates :

$$\boxed{\exists\,!\,\lambda \in \mathbb{R}^p, \mu \in (\mathbb{R}_+)^q,\ \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \lambda, \mu) = \mathbf{0}\ .}$$

Copyright : Jean-Philippe Préaux

La condition $(i)$ de KKT s'écrit alors, sous les hypothèses adéquates :

$$\boxed{\exists\,!\,\lambda \in \mathbb{R}^p, \mu \in (\mathbb{R}_+)^q,\ \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{u},\lambda,\mu) = \mathbf{0}\ .}$$

On définit aussi, lorsque les applications sont deux fois différentiables, la matrice Hessienne du lagrangien en $\mathbf{u} \in \mathbb{R}^n$, par :

$$\nabla^2_{\mathbf{x}}\mathcal{L}(\mathbf{u},\lambda,\mu) = \nabla^2 f(\mathbf{u}) + \sum_{i=1}^{p} \lambda_i \nabla^2 \varphi_i(\mathbf{u}) + \sum_{j=1}^{q} \mu_j \nabla^2 \psi_j(\mathbf{u})\ .$$

Copyright : Jean-Philippe Préaux

# Prise en compte de la convexité.

Dans le cas de la programmation convexe, c'est-à-dire si de plus :

- $\mathcal{U}$ est un ouvert convexe de $\mathbb{R}^n$ (notamment lorsque $\mathcal{U} = \mathbb{R}^n$),
- $f$ est convexe,
- $\varphi_i$, $i = 1, \ldots, p$, est linéaire,
- $\psi_j$, $j = 1, \ldots, q$ est convexe,

Copyright : Jean-Philippe Préaux

# Prise en compte de la convexité.

Dans le cas de la programmation convexe, c'est-à-dire si de plus :

– $\mathcal{U}$ est un ouvert convexe de $\mathbb{R}^n$ (notamment lorsque $\mathcal{U} = \mathbb{R}^n$),

– $f$ est convexe,

– $\varphi_i$, $i = 1, \ldots, p$, est linéaire,

– $\psi_j$, $j = 1, \ldots, q$ est convexe,

Théorème (Suffisance des conditions KKT en programmation convexe.)

*En programmation convexe, les conditions $(i)$, $(ii)$ et $(iii)$ de KKT en $\mathbf{u} \in \mathcal{D}$ sont aussi suffisantes pour que $\mathbf{u}$ soit un minimum global de $f$.*

Copyright : Jean-Philippe Préaux

# Qualification de contraintes affines

Copyright : Jean-Philippe Préaux

# Qualification de contraintes affines

Proposition (Qualification de contraintes affines.)

*Si en $\mathbf{u} \in \mathcal{D}$, toutes les contraintes égalitaires et inégalitaires actives sont affines alors on peut se passer de l'hypothèse $(QC)_{\mathbf{u}}$ dans le théorème, à ceci près que les multiplicateurs de Lagrange-KKT ne sont plus nécessairement uniques.*

Copyright : Jean-Philippe Préaux

# Qualification de contraintes affines

## Proposition (Qualification de contraintes affines.)

*Si en* $\mathbf{u} \in \mathcal{D}$, *toutes les contraintes égalitaires et inégalitaires actives sont affines alors on peut se passer de l'hypothèse* $(QC)_{\mathbf{u}}$ *dans le théorème, à ceci près que les multiplicateurs de Lagrange-KKT ne sont plus nécessairement uniques.*

**Démonstration.** La preuve procède de la même façon que pour le cas sous contraintes égalitaires. □

Copyright : Jean-Philippe Préaux

# Qualification de contraintes affines

Copyright : Jean-Philippe Préaux

## Proposition (Qualification de contraintes affines.)

*Si en $\mathbf{u} \in \mathcal{D}$, toutes les contraintes égalitaires et inégalitaires actives sont affines alors on peut se passer de l'hypothèse $(QC)_{\mathbf{u}}$ dans le théorème, à ceci près que les multiplicateurs de Lagrange-KKT ne sont plus nécessairement uniques.*

**Démonstration.** La preuve procède de la même façon que pour le cas sous contraintes égalitaires. □

Nous noterons cette nouvelle hypothèse de qualification des contraintes :

$(\overline{QC}_{\mathbf{u}})$ : Toutes les contraintes égalitaires ainsi que toutes les contraintes inégalitaires actives en $\mathbf{u}$ sont affines.

# Programmation quadratique.

Copyright : Jean-Philippe Préaux

# Programmation quadratique.

Théorème (Programmation quadratique sous contraintes.)

*Soit $f$ une application quadratique, $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ où $A$ est une matrice symétrique, et soit $\mathcal{D}$ un domaine défini par des contraintes égalitaires et inégalitaires affines.*

Copyright : Jean-Philippe Préaux

# Programmation quadratique.

## Théorème (Programmation quadratique sous contraintes.)

*Soit $f$ une application quadratique, $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ où $A$ est une matrice symétrique, et soit $\mathcal{D}$ un domaine défini par des contraintes égalitaires et inégalitaires affines.*
*Alors, si $A$ est semi-définie positive (resp. négative) un minimum (resp. maximum) global de $f$ sur $\mathcal{D}$, s'il existe, est caractérisé par les conditions $(i)$, $(ii)$ (resp. $(ii')$), $(iii)$ de KKT.*

Copyright : Jean-Philippe Préaux

# Programmation quadratique.

## Théorème (Programmation quadratique sous contraintes.)

*Soit $f$ une application quadratique, $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ où $A$ est une matrice symétrique, et soit $\mathcal{D}$ un domaine défini par des contraintes égalitaires et inégalitaires affines.*
*Alors, si $A$ est semi-définie positive (resp. négative) un minimum (resp. maximum) global de $f$ sur $\mathcal{D}$, s'il existe, est caractérisé par les conditions $(i)$, $(ii)$ (resp. $(ii')$), $(iii)$ de KKT.*
*Si de plus $A$ est définie positive (resp. négative), $f$ admet un unique minimum (resp. maximum) global sur $\mathcal{D}$.*

Copyright: Jean-Philippe Préaux

# Programmation quadratique.

## Théorème (Programmation quadratique sous contraintes.)

*Soit $f$ une application quadratique, $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ où $A$ est une matrice symétrique, et soit $\mathcal{D}$ un domaine défini par des contraintes égalitaires et inégalitaires affines.*
*Alors, si $A$ est semi-définie positive (resp. négative) un minimum (resp. maximum) global de $f$ sur $\mathcal{D}$, s'il existe, est caractérisé par les conditions $(i)$, $(ii)$ (resp. $(ii')$), $(iii)$ de KKT.*
*Si de plus $A$ est définie positive (resp. négative), $f$ admet un unique minimum (resp. maximum) global sur $\mathcal{D}$.*

**Preuve.** Sous ces hypothèses $f$ (resp. $-f$) est convexe, voire elliptique. De plus les contraintes étant affines, on vérifie une hypothèse de qualification des contraintes. Les conditions KKT sont alors nécessaires et suffisantes à l'existence d'un extremum. □

Copyright: Jean-Philippe Préaux

# Exemple.

Résoudre le problème de programmation quadratique suivant.

$$\min\ x^2+y^2+z^2 \quad \text{sous les contraintes} \quad \begin{cases} x+y+z &= 3 \\ 2x-y+z &\leqslant 5 \end{cases}$$

Copyright : Jean-Philippe Préaux

# Exemple.

Résoudre le problème de programmation quadratique suivant.

$$\min\ x^2+y^2+z^2 \quad \text{sous les contraintes} \quad \begin{cases} x+y+z & = & 3 \\ 2x-y+z & \leqslant & 5 \end{cases}$$

C'est équivalent au problème consistant à déterminer dans $\mathbb{R}^3$ la distance d'un demi-plan (défini par les contraintes) à l'origine.

Copyright : Jean-Philippe Préaux

Les contraintes étant affines, on peut appliquer telles quelles les conditions (KKT) :

Copyright : Jean-Philippe Préaux

Les contraintes étant affines, on peut appliquer telles quelles les conditions (KKT) : en un minimum $\mathbf{u}$, $\exists\, \lambda \in \mathbb{R}, \mu \in \mathbb{R}_+$ tels que $\nabla f(\mathbf{u}) + \lambda \nabla \varphi(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 0$ :

$$(i) \quad \begin{cases} 2x + \lambda + 2\mu = 0 \\ 2y + \lambda - \mu = 0 \\ 2z + \lambda + \mu = 0 \end{cases}$$

$$(ii) \quad \mu(2x - y + z - 5) = 0$$

$$(iii) \quad \mu \geqslant 0$$

Copyright : Jean-Philippe Préaux

Les contraintes étant affines, on peut appliquer telles quelles les conditions (KKT) : en un minimum $\mathbf{u}$, $\exists\, \lambda \in \mathbb{R}, \mu \in \mathbb{R}_+$ tels que $\nabla f(\mathbf{u}) + \lambda \nabla \varphi(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 0$ :

$$(i) \quad \left\{ \begin{array}{rcl} 2x + \lambda + 2\mu & = & 0 \\ 2y + \lambda - \mu & = & 0 \\ 2z + \lambda + \mu & = & 0 \end{array} \right.$$

$$(ii) \quad \mu(2x - y + z - 5) = 0$$

$$(iii) \quad \mu \geqslant 0$$

Supposons que $\mu \neq 0 \Longrightarrow 2x - y + z = 5$

$$(i) \Longrightarrow \left\{ \begin{array}{l} x = (-\lambda - 2\mu)/2 \\ y = (-\lambda + \mu)/2 \\ z = (-\lambda - \mu)/2 \end{array} \right.$$

Copyright : Jean-Philippe Préaux

$$x + y + z = 3 \implies -\lambda - 2\mu - \lambda + \mu - \lambda - \mu = 6 \implies 3\lambda + 2\mu = -6 \quad (a)$$

Copyright : Jean-Philippe Préaux

$x + y + z = 3 \implies -\lambda - 2\mu - \lambda + \mu - \lambda - \mu = 6 \implies 3\lambda + 2\mu = -6$ (a)
$2x - y + z = 5 \implies 2(-\lambda - 2\mu) - (-\lambda + \mu) + (-\lambda - \mu) = 10 \implies$
$\lambda + 3\mu = -5$ (b)

Copyright : Jean-Philippe Préaux

$x + y + z = 3 \implies -\lambda - 2\mu - \lambda + \mu - \lambda - \mu = 6 \implies 3\lambda + 2\mu = -6$ (a)
$2x - y + z = 5 \implies 2(-\lambda - 2\mu) - (-\lambda + \mu) + (-\lambda - \mu) = 10 \implies$
$\lambda + 3\mu = -5$ (b)
En formant (a)$-3$(b) on obtient $2\mu - 9\mu = -6 + 15$, soit $\mu = -9/7 < 0$ ce qui contredit (*iii*) !
Ainsi $\mu = 0$.

Copyright : Jean-Philippe Préaux

$x+y+z=3 \implies -\lambda-2\mu-\lambda+\mu-\lambda-\mu=6 \implies 3\lambda+2\mu=-6$ (a)
$2x-y+z=5 \implies 2(-\lambda-2\mu)-(-\lambda+\mu)+(-\lambda-\mu)=10 \implies$
$\lambda+3\mu=-5$ (b)
En formant (a)$-3$(b) on obtient $2\mu-9\mu=-6+15$, soit $\mu=-9/7<0$ ce qui contredit $(iii)$ !
Ainsi $\mu=0$. Donc $(i)$ devient :

$$\begin{cases} 2x+\lambda=0 \\ 2y+\lambda=0 \\ 2z+\lambda=0 \end{cases} \implies 2(x+y+z)+3\lambda=0 \implies 2\times 3+3\lambda=0 \implies \lambda=-2$$

Copyright : Jean-Philippe Préaux

$x+y+z=3 \implies -\lambda-2\mu-\lambda+\mu-\lambda-\mu=6 \implies 3\lambda+2\mu=-6$ (a)
$2x-y+z=5 \implies 2(-\lambda-2\mu)-(-\lambda+\mu)+(-\lambda-\mu)=10 \implies$
$\lambda+3\mu=-5$ (b)
En formant (a)$-3$(b) on obtient $2\mu-9\mu=-6+15$, soit $\mu=-9/7<0$ ce qui contredit $(iii)$ !
Ainsi $\mu=0$. Donc $(i)$ devient :

$$\begin{cases} 2x+\lambda=0 \\ 2y+\lambda=0 \\ 2z+\lambda=0 \end{cases} \implies 2(x+y+z)+3\lambda=0 \implies 2\times 3+3\lambda=0 \implies \lambda=-2$$

$\implies x=y=z=1$. On obtient pour solution $\boxed{\mathbf{u}=(1,1,1)}$.

Copyright : Jean-Philippe Préaux

$x+y+z=3 \implies -\lambda-2\mu-\lambda+\mu-\lambda-\mu=6 \implies 3\lambda+2\mu=-6$ (a)
$2x-y+z=5 \implies 2(-\lambda-2\mu)-(-\lambda+\mu)+(-\lambda-\mu)=10 \implies$
$\lambda+3\mu=-5$ (b)
En formant (a)$-3$(b) on obtient $2\mu-9\mu=-6+15$, soit $\mu=-9/7<0$ ce qui contredit $(iii)$ !
Ainsi $\mu=0$. Donc $(i)$ devient :

$$\begin{cases} 2x+\lambda=0 \\ 2y+\lambda=0 \\ 2z+\lambda=0 \end{cases} \implies 2(x+y+z)+3\lambda=0 \implies 2\times 3+3\lambda=0 \implies \lambda=-2$$

$\implies x=y=z=1$. On obtient pour solution $\boxed{\mathbf{u}=(1,1,1)}$.

Pour conclure que $\mathbf{u}$ est le minimum global de $f$ sur $\mathcal{D}$, on a plusieurs possibilités :
– $f$ est quadratique sous contraintes affines, de matrice $A=\mathrm{Id}$ définie positive.
– $f$ est (fortement) convexe sous contraintes affines.
– $f$ est coercive.

Copyright : Jean-Philippe Préaux

# Cône tangent

Copyright : Jean-Philippe Préaux

# Cône tangent

**Définitions.**

- En $\mathbf{u} \in \mathcal{D}$, l'*ensemble des indices de contraintes actives*, $J(\mathbf{u}) \subset \{1, \ldots, q\}$, est l'ensemble des indices $j$ pour lesquels la $j$-ème contrainte inégalitaire est active en $\mathbf{u}$ :

$$J(\mathbf{u}) = \left\{ j \in \{1, \ldots, q\} \,|\, \psi_j(\mathbf{u}) = 0 \right\} .$$

Copyright : Jean-Philippe Préaux

# Cône tangent

**Définitions.**

• En $\mathbf{u} \in \mathcal{D}$, l'*ensemble des indices de contraintes actives*, $J(\mathbf{u}) \subset \{1, \ldots, q\}$, est l'ensemble des indices $j$ pour lesquels la $j$-ème contrainte inégalitaire est active en $\mathbf{u}$ :

$$J(\mathbf{u}) = \left\{ j \in \{1, \ldots, q\} \,|\, \psi_j(\mathbf{u}) = 0 \right\} \ .$$

• Le *cône tangent* à $\mathcal{D}$ en $\mathbf{u}$, noté $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ est le sous-ensemble de $\mathbb{R}^n$ :

$$\mathcal{C}_{\mathbf{u}}\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^n \,|\, \langle \nabla\varphi_i(\mathbf{u}), \mathbf{d} \rangle = 0,\ i = 1, \ldots, q,$$
$$\langle \nabla\psi_j(\mathbf{u}), \mathbf{d} \rangle \leqslant 0,\ j \in J(\mathbf{u})\}$$

Copyright : Jean-Philippe Préaux

Le cône tangent généralise en cas de contraintes inégalitaires la notion d'espace tangent dans le sens qui suit.

Copyright : Jean-Philippe Préaux

Le cône tangent généralise en cas de contraintes inégalitaires la notion d'espace tangent dans le sens qui suit.

Proposition (Le cône tangent en tant qu'espace tangent.)

*Si en* $\mathbf{u} \in \mathcal{D}$ *une des hypothèses de qualification des contraintes* $(QC)_{\mathbf{u}}$, $(\overline{QC})_{\mathbf{u}}$ *est vérifiée, alors le cône tangent* $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ *est l'ensemble des directions* $\mathbf{d} \in \mathbb{R}^n$ *pour lesquelles soit* $\mathbf{d} = \mathbf{0}$ *soit il existe une suite* $(\mathbf{u}_k)_{k\in\mathbb{N}}$ *dans* $\mathcal{D}$, *non stationnaire, tendant vers* $\mathbf{u}$, *avec :*

$$\mathbf{u}_k = \mathbf{u} + \frac{\|\mathbf{u}_k - \mathbf{u}\|}{\|\mathbf{d}\|}\mathbf{d} + o(\|\mathbf{u}_k - \mathbf{u}\|)$$

Copyright : Jean-Philippe Préaux

Le cône tangent généralise en cas de contraintes inégalitaires la notion d'espace tangent dans le sens qui suit.

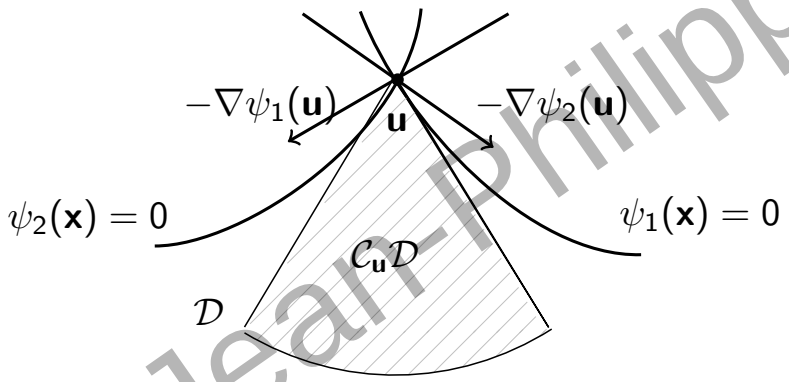
**Proposition (Le cône tangent en tant qu'espace tangent.)**

*Si en* $\mathbf{u} \in \mathcal{D}$ *une des hypothèses de qualification des contraintes* $(QC)_{\mathbf{u}}$, $(\overline{QC})_{\mathbf{u}}$ *est vérifiée, alors le cône tangent* $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ *est l'ensemble des directions* $\mathbf{d} \in \mathbb{R}^n$ *pour lesquelles soit* $\mathbf{d} = \mathbf{0}$ *soit il existe une suite* $(\mathbf{u}_k)_{k\in\mathbb{N}}$ *dans* $\mathcal{D}$, *non stationnaire, tendant vers* $\mathbf{u}$, *avec :*

$$\mathbf{u}_k = \mathbf{u} + \frac{\|\mathbf{u}_k - \mathbf{u}\|}{\|\mathbf{d}\|}\mathbf{d} + o(\|\mathbf{u}_k - \mathbf{u}\|)$$

Ce n'est plus un espace vectoriel, mais une intersection de sous-espaces vectoriels et de demi-espaces.

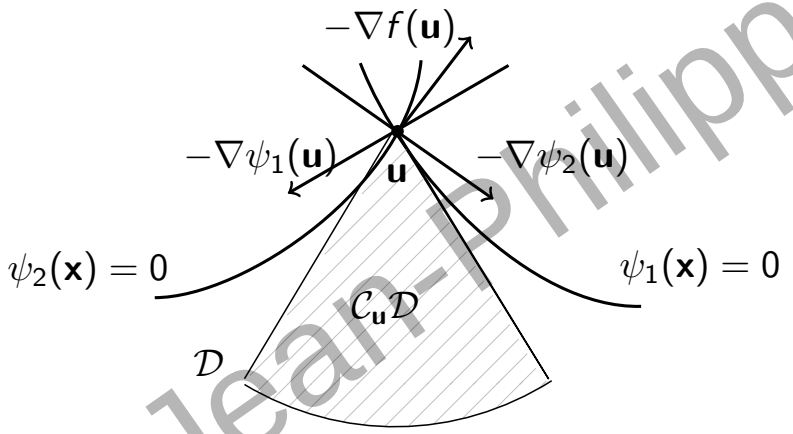
Copyright : Jean-Philippe Préaux




Le cône tangent $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ à $\mathcal{D}$ en $\mathbf{u}$. C'est une intersection de demi-espaces (pour chaque contrainte inégalitaire active) et de sous-espaces vectoriels (pour chaque contrainte égalitaire).

Copyright : Jean-Philippe Préaux

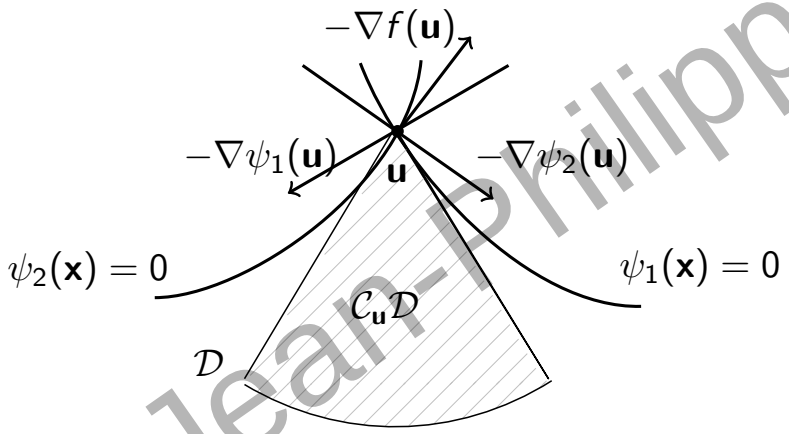
# Interprétation géométrique des conditions KKT.




On a représenté la direction de plus grande pente $-\nabla f(\mathbf{u})$. Si $\mathbf{u}$ est un minimum local elle se trouve dans le cône polaire de $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ (délimité ici par $\mathbf{u}$, $\nabla \psi_1(\mathbf{u})$ et $\nabla \psi_2(\mathbf{u})$).

Copyright : Jean-Philippe Préaux

# Interprétation géométrique des conditions KKT.




On a représenté la direction de plus grande pente $-\nabla f(\mathbf{u})$. Si $\mathbf{u}$ est un minimum local elle se trouve dans le cône polaire de $\mathcal{C}_{\mathbf{u}}\mathcal{D}$ (délimité ici par $\mathbf{u}$, $\nabla\psi_1(\mathbf{u})$ et $\nabla\psi_2(\mathbf{u})$). Les conditions nécessaires de KKT, s'expriment géométriquement : en un min (resp. max) local $\mathbf{u}$, la direction de plus grande décroissance (resp. accroissement) de $f$, $-\nabla f(\mathbf{u})$ (resp. $\nabla f(\mathbf{u})$) est dans le *cône polaire* de $\mathcal{C}_{\mathbf{u}}\mathcal{D}$, c'est à dire $\{\mathbf{c}\in\mathbb{R}^n \,|\, \forall\, \mathbf{d}\in\mathcal{C}_{\mathbf{u}}\mathcal{D},\ \langle\mathbf{d},\mathbf{c}\rangle\leqslant 0\}$.

Copyright : Jean-Philippe Préaux

# Condition suffisante du 2e ordre

## Théorème (Condition suffisante du 2e ordre.)

*Soit $\mathcal{U} \subset \mathbb{R}^n$ un ouvert, on suppose que $f, \varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q$ sont deux fois différentiables sur $\mathcal{U}$, et que $\mathbf{u} \in \mathcal{D}$ vérifie une des hypothèses de qualification des contraintes.*

Copyright : Jean-Philippe Préaux

# Condition suffisante du 2e ordre

## Théorème (Condition suffisante du 2e ordre.)

*Soit $\mathcal{U} \subset \mathbb{R}^n$ un ouvert, on suppose que $f, \varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q$ sont deux fois différentiables sur $\mathcal{U}$, et que $\mathbf{u} \in \mathcal{D}$ vérifie une des hypothèses de qualification des contraintes.*
*Si $\mathbf{u} \in \mathcal{D}$ vérifie les conditions $(i)$, $(ii)$, et $(iii)$ de KKT, en particulier si il existe $\lambda \in \mathbb{R}^p$, $\mu \in (\mathbb{R}_+)^q$ tels que :*

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \lambda, \mu) = 0$$

*et si de plus $\forall\, \mathbf{d} \neq \mathbf{0} \in \mathcal{C}_{\mathbf{u}}\mathcal{D}$, :*

$$\langle \nabla^2_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \lambda, \mu)\, \mathbf{d},\, \mathbf{d} \rangle > 0 \ .$$

*alors $\mathbf{u}$ est un minimum local strict de $f$ sur $\mathcal{D}$.*

Copyright : Jean-Philippe Préaux

# Le sous-cône tangent.

Afin d'énoncer une condition nécessaire du second ordre, nous devons nous restreindre à un sous-ensemble du cône tangent.

Copyright : Jean-Philippe Préaux

# Le sous-cône tangent.

Afin d'énoncer une condition nécessaire du second ordre, nous devons nous restreindre à un sous-ensemble du cône tangent.

**Définition.** Soit $\mathbf{u} \in \mathcal{D}$, un point vérifiant une hypothèse de qualification des contraintes ainsi que les conditions $(i)$, $(ii)$ et $(iii)$ de KKT.

- Notons :

$$J^+(\mathbf{u}) = \Big\{ j \in \{1, \ldots, q\} \,|\, \psi_j(\mathbf{u}) = 0 \text{ et } \mu_j > 0 \Big\}$$

C'est l'ensemble des indices des contraintes *fortement actives* en $\mathbf{u}$.
Notons que, plus simplement, $J^+(\mathbf{u}) = \Big\{ j \in \{1, \ldots, q\} \,|\, \mu_j \neq 0 \Big\}$.

Copyright : Jean-Philippe Préaux

# Le sous-cône tangent.

Afin d'énoncer une condition nécessaire du second ordre, nous devons nous restreindre à un sous-ensemble du cône tangent.

**Définition.** Soit $\mathbf{u} \in \mathcal{D}$, un point vérifiant une hypothèse de qualification des contraintes ainsi que les conditions $(i)$, $(ii)$ et $(iii)$ de KKT.

- Notons :

$$J^+(\mathbf{u}) = \Big\{ j \in \{1, \ldots, q\} \,|\, \psi_j(\mathbf{u}) = 0 \text{ et } \mu_j > 0 \Big\}$$

C'est l'ensemble des indices des contraintes *fortement actives* en $\mathbf{u}$. Notons que, plus simplement, $J^+(\mathbf{u}) = \Big\{ j \in \{1, \ldots, q\} \,|\, \mu_j \neq 0 \Big\}$.

- Notons

$$\mathcal{C}_{\mathbf{u}}^+\mathcal{D} = \Big\{ \mathbf{d} \in \mathcal{C}_{\mathbf{u}}\mathcal{D} \,|\, \langle \nabla\psi_j(\mathbf{u}), \mathbf{d} \rangle = 0,\, j \in J^+(\mathbf{u}), \Big\} \ .$$

Copyright : Jean-Philippe Préaux

# Condition nécessaire du 2e ordre.

## Théorème (Condition nécessaire du second ordre)

*Soit $\mathcal{U} \subset \mathbb{R}^n$ un ouvert, on suppose que $f, \varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q$ sont deux fois différentiables sur $\mathcal{U}$, et que $\mathbf{u} \in \mathcal{D}$ vérifie une des hypothèses de qualification des contraintes.*

Copyright : Jean-Philippe Préaux

# Condition nécessaire du 2e ordre.

## Théorème (Condition nécessaire du second ordre)

*Soit $\mathcal{U} \subset \mathbb{R}^n$ un ouvert, on suppose que $f, \varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_q$ sont deux fois différentiables sur $\mathcal{U}$, et que $\mathbf{u} \in \mathcal{D}$ vérifie une des hypothèses de qualification des contraintes.*
*Si $\mathbf{u}$ est un minimum local de $f$ sur $\mathcal{D}$, les conditions $(i)$, $(ii)$, $(iii)$ de KKT sont satisfaites, en particulier $\exists\, \lambda \in \mathbb{R}^p, \mu \in (\mathbb{R}_+)^q$ tels que :*

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{u}, \lambda, \mu) = 0$$

*et de plus, $\forall\, \mathbf{d} \in \mathcal{C}_{\mathbf{u}}^{+}\mathcal{D}$, :*

$$\langle \nabla_{\mathbf{x}}^2 \mathcal{L}(\mathbf{u}, \lambda, \mu)\, \mathbf{d},\, \mathbf{d} \rangle \geqslant 0 \ .$$

*(i.e. $\nabla_{\mathbf{x}}^2 \mathcal{L}(\mathbf{u}, \lambda, \mu)$ est semi-définie positive sur $\mathcal{C}_{\mathbf{u}}\mathcal{D}$).*

Copyright : Jean-Philippe Préaux

On se rappelle qu'au problème d'optimisation sous contrainte :

$$\begin{aligned} &\min_{\mathbf{x}} \quad f(\mathbf{x}) \\ &\varphi_i(\mathbf{x}) = 0,\ \forall i = 1, \ldots, p, \qquad (P) \\ &\psi_j(\mathbf{x}) \leqslant 0,\ \forall j = 1, \ldots, q, \end{aligned}$$

Copyright : Jean-Philippe Préaux

On se rappelle qu'au problème d'optimisation sous contrainte :

$$\begin{aligned} &\min_{\mathbf{x}} \quad f(\mathbf{x}) \\ &\varphi_i(\mathbf{x}) = 0,\ \forall i = 1, \ldots, p, \qquad (P) \\ &\psi_j(\mathbf{x}) \leqslant 0,\ \forall j = 1, \ldots, q, \end{aligned}$$

on associe le lagrangien du problème :

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q \longmapsto \mathbb{R}$$

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i \varphi_i(\mathbf{x}) + \sum_{j=1}^{q} \mu_j \psi_j(\mathbf{x})$$

(en notant $\lambda = (\lambda_1, \ldots, \lambda_p)$ et $\mu = (\mu_1, \ldots, \mu_q)$).

Copyright : Jean-Philippe Préaux

# Point-selle du lagrangien

**Définition.** On appelle *point-selle du lagrangien* tout triplet $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q$ vérifiant :

$$\forall\, (\mathbf{x}, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q,$$
$$\mathcal{L}(\mathbf{x}^*, \lambda, \mu) \leqslant \mathcal{L}(\mathbf{x}^*, \lambda^*, \mu^*) \leqslant \mathcal{L}(\mathbf{x}, \lambda^*, \mu^*)\ .$$

Copyright : Jean-Philippe Préaux

# Point-selle du lagrangien

**Définition.** On appelle *point-selle du lagrangien* tout triplet $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q$ vérifiant :

$$\forall\, (\mathbf{x}, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q,$$
$$\mathcal{L}(\mathbf{x}^*, \lambda, \mu) \leqslant \mathcal{L}(\mathbf{x}^*, \lambda^*, \mu^*) \leqslant \mathcal{L}(\mathbf{x}, \lambda^*, \mu^*)\ .$$

c'est-à-dire que :

- $\mathbf{x}^*$ est un minimum de $\mathbf{x} \mapsto \mathcal{L}(\mathbf{x}, \lambda^*, \mu^*)$, et
- $(\lambda^*, \mu^*)$ est un maximum de $(\lambda, \mu) \mapsto \mathcal{L}(\mathbf{x}^*, \lambda, \mu)$.

Copyright : Jean-Philippe Préaux

# point-selle et solution à $(P)$

Théorème

*Si* $(\mathbf{x}^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q$ *est un point-selle du lagrangien du problème* $(P)$*, alors* $\mathbf{x}^*$ *est solution du problème* $(P)$*.*

Copyright : Jean-Philippe Préaux

# point-selle et solution à $(P)$

## Théorème

*Si* $(\mathbf{x}^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q$ *est un point-selle du lagrangien du problème* $(P)$, *alors* $\mathbf{x}^*$ *est solution du problème* $(P)$.

## Théorème (Cas où la réciproque est vraie.)

*Supposons que* $f$ *est dérivable et convexe,* $\varphi_1, \ldots, \varphi_p$ *sont* $C^1$ *et affines, et* $\psi_1, \ldots, \psi_q$ *sont* $C^1$ *et convexes. Soit* $\mathbf{x}^*$ *un point du domaine admissible vérifiant une hypothèse de qualification des contraintes.*

Copyright : Jean-Philippe Préaux

# point-selle et solution à $(P)$

## Théorème

*Si* $(\mathbf{x}^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}_+)^q$ *est un point-selle du lagrangien du problème* $(P)$*, alors* $\mathbf{x}^*$ *est solution du problème* $(P)$*.*

## Théorème (Cas où la réciproque est vraie.)

*Supposons que* $f$ *est dérivable et convexe,* $\varphi_1, \ldots, \varphi_p$ *sont* $C^1$ *et affines, et* $\psi_1, \ldots, \psi_q$ *sont* $C^1$ *et convexes. Soit* $\mathbf{x}^*$ *un point du domaine admissible vérifiant une hypothèse de qualification des contraintes. Si* $\mathbf{x}^*$ *est solution du problème* $(P)$*, alors il existe* $(\lambda^*, \mu^*) \in \mathbb{R}^p \times (\mathbb{R}_+)^q$ *tel que* $(\mathbf{x}^*, \lambda^*, \mu^*)$ *soit un point-selle du lagrangien.*

Copyright : Jean-Philippe Préaux

**Exemple.** Le problème de minimisation suivant :

$$\min_{x \geqslant 0} x$$

est un problème de programmation convexe qui a pour solution évidente $x = 0$.

Copyright : Jean-Philippe Préaux

**Exemple.** Le problème de minimisation suivant :

$$\min_{x \geqslant 0} x$$

est un problème de programmation convexe qui a pour solution évidente $x = 0$.

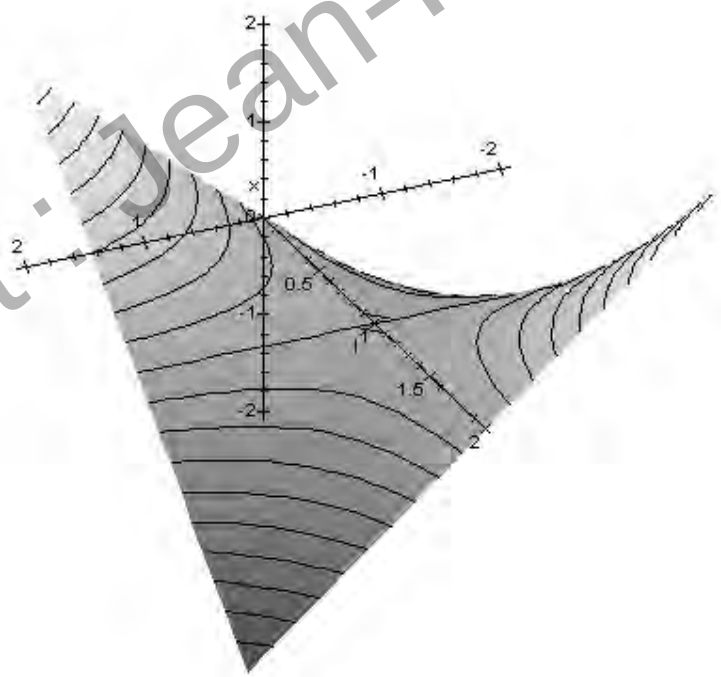
Son lagrangien est $\mathcal{L}(x, \mu) = x - \mu x$ qui admet un unique point-selle sur $\mathbb{R} \times \mathbb{R}_+$ en $x = 0$, $\mu = 1$, qui fournit le minimum et le multiplicateur de Lagrange associé.

Copyright : Jean-Philippe Préaux

**Exemple.** Le problème de minimisation suivant :

$$\min_{x \geqslant 0} x$$

est un problème de programmation convexe qui a pour solution évidente $x = 0$. Son lagrangien est $\mathcal{L}(x, \mu) = x - \mu x$ qui admet un unique point-selle sur $\mathbb{R} \times \mathbb{R}_+$ en $x = 0$, $\mu = 1$, qui fournit le minimum et le multiplicateur de Lagrange associé.




Copyright : Jean-Philippe Préaux

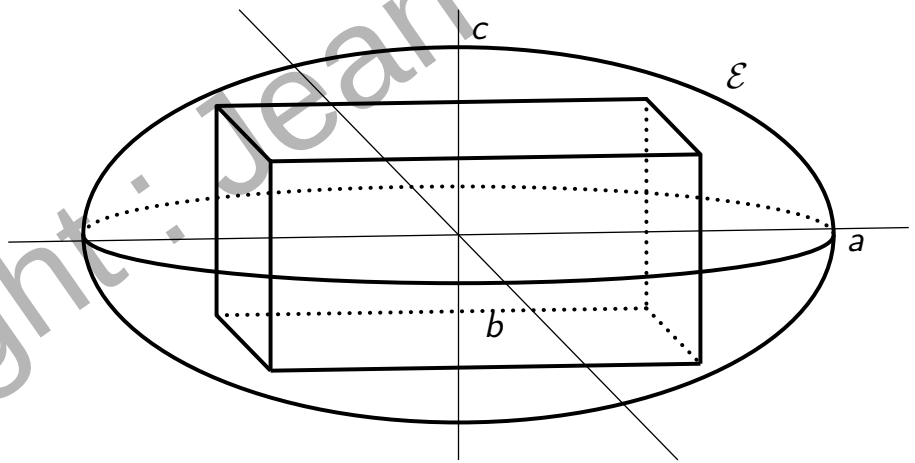
# Exercices

- Exercice 1.
- Exercice 2.
- Exercice 3.

Copyright : Jean-Philippe Préaux

# Exercice 1.

(Problème de Kepler.)
Inscrire dans l'ellipsoïde $\mathcal{E} = \{(x, y, z) \in \mathbb{R}^3 \mid x^2/a^2 + y^2/b^2 + z^2/c^2 = 1\}$ le parrallépipède de volume maximal dont les arêtes sont parallèles aux axes.




Copyright : Jean-Philippe Préaux

Retour.

Le problème se formule :

$$\begin{cases} \max\ xyz \\ x^2/a^2 + y^2/b^2 + z^2/c^2 = 1 \\ x, y, z \geqslant 0 \end{cases}$$

Copyright : Jean-Philippe Préaux

Le problème se formule :

$$\begin{cases} \max\ xyz \\ x^2/a^2 + y^2/b^2 + z^2/c^2 = 1 \\ x, y, z \geqslant 0 \end{cases}$$

On calcule :

$$\nabla f(\mathbf{x}) = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix} \; ; \; \nabla \varphi(\mathbf{x}) = \begin{pmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{pmatrix} \neq 0 \text{ sur } \mathcal{E} \text{ si } xyz \neq 0.$$

Copyright : Jean-Philippe Préaux

Le problème se formule :

$$\begin{cases} \max\ xyz \\ x^2/a^2 + y^2/b^2 + z^2/c^2 = 1 \\ x, y, z \geqslant 0 \end{cases}$$

On calcule :

$$\nabla f(\mathbf{x}) = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix} \ ; \ \nabla \varphi(\mathbf{x}) = \begin{pmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{pmatrix} \neq 0 \text{ sur } \mathcal{E} \text{ si } xyz \neq 0.$$

On vérifie en tout point $\mathbf{u}$ de $\mathcal{E}$ l'hypothèse $(QC)_{\mathbf{u}}$. On peut appliquer KKT.

Copyright : Jean-Philippe Préaux

Le problème se formule :

$$\begin{cases} \max\ xyz \\ x^2/a^2 + y^2/b^2 + z^2/c^2 = 1 \\ x, y, z \geqslant 0 \end{cases}$$

On calcule :

$$\nabla f(\mathbf{x}) = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix} \ ; \ \nabla \varphi(\mathbf{x}) = \begin{pmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{pmatrix} \neq 0 \text{ sur } \mathcal{E} \text{ si } xyz \neq 0.$$

On vérifie en tout point $\mathbf{u}$ de $\mathcal{E}$ l'hypothèse $(QC)_{\mathbf{u}}$. On peut appliquer KKT.
Les contraintes inégalitaires étant toutes insaturées car $x, y, z > 0$,
$\Longrightarrow \ \mu_1 = \mu_2 = \mu_3 = 0.$

Copyright : Jean-Philippe Préaux

Condition nécessaire ($KKT$) :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

Copyright : Jean-Philippe Préaux

Condition nécessaire ($KKT$) :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0$

Copyright : Jean-Philippe Préaux

Condition nécessaire ($KKT$) :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \Longrightarrow$ $yz = -2\lambda/(3x)$

Copyright : Jean-Philippe Préaux

Condition nécessaire $(KKT)$ :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \implies yz = -2\lambda/(3x) \implies -2\lambda/(3x) + \lambda 2x/a^2 = 0$

Copyright : Jean-Philippe Préaux

Condition nécessaire $(KKT)$ :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \implies yz = -2\lambda/(3x) \implies -2\lambda/(3x) + \lambda 2x/a^2 = 0 \implies 2\lambda(3x^2 - a^2)/(3a^2x) = 0.$

Copyright : Jean-Philippe Préaux

Condition nécessaire ($KKT$) :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \implies yz = -2\lambda/(3x) \implies -2\lambda/(3x) + \lambda 2x/a^2 = 0 \implies 2\lambda(3x^2 - a^2)/(3a^2x) = 0$. Or $\lambda = 0$ est impossible car autrement $xyz = 0$. Donc $3x^2 = a^2$.

Copyright : Jean-Philippe Préaux

Condition nécessaire ($KKT$) :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \implies yz = -2\lambda/(3x) \implies -2\lambda/(3x) + \lambda 2x/a^2 = 0 \implies 2\lambda(3x^2 - a^2)/(3a^2x) = 0$. Or $\lambda = 0$ est impossible car autrement $xyz = 0$. Donc $3x^2 = a^2$. De la même façon $3y^2 = b^2$ et $3z^2 = c^2$. Donc :

$$\boxed{x = \frac{a}{\sqrt{3}} \quad ; \quad y = \frac{b}{\sqrt{3}} \quad ; \quad z = \frac{c}{\sqrt{3}}}$$

Copyright : Jean-Philippe Préaux

Condition nécessaire $(KKT)$ :

$$\nabla f(\mathbf{x}) + \lambda \nabla \varphi(\mathbf{x}) = 0$$

$$\begin{cases} yz + \lambda 2x/a^2 = 0 \\ xz + \lambda 2y/b^2 = 0 \\ xy + \lambda 2z/c^2 = 0 \end{cases}$$

On multiplie la première ligne par $x$, la deuxième par $y$ et la dernière par $z$, puis on somme : on obtient $3xyz + 2\lambda = 0 \implies yz = -2\lambda/(3x) \implies -2\lambda/(3x) + \lambda 2x/a^2 = 0 \implies 2\lambda(3x^2 - a^2)/(3a^2x) = 0$. Or $\lambda = 0$ est impossible car autrement $xyz = 0$. Donc $3x^2 = a^2$. De la même façon $3y^2 = b^2$ et $3z^2 = c^2$. Donc :

$$\boxed{x = \frac{a}{\sqrt{3}} \quad ; \quad y = \frac{b}{\sqrt{3}} \quad ; \quad z = \frac{c}{\sqrt{3}}}$$

Et par suite, par compacité : $\boxed{\text{Vol}_{\max} = abc/3\sqrt{3}}$.

Copyright : Jean-Philippe Préaux

# Exercice 2.

(Problème de Tartaglia)
Décomposer le nombre 8 en deux parties positives $p_1$, $p_2$ de sorte que le produit de leur produit par leur différence soit maximal.

Copyright : Jean-Philippe Préaux

Retour.

# Exercice 2.

(Problème de Tartaglia)
Décomposer le nombre 8 en deux parties positives $p_1$, $p_2$ de sorte que le produit de leur produit par leur différence soit maximal.

Le problème se formule :

$$\begin{cases} \max \ p_1 p_2 \times (p_2 - p_1) = p_1 p_2^2 - p_1^2 p_2 \\ p_1 + p_2 = 8 \\ p_1, p_2 \geqslant 0 \end{cases}$$

Copyright : Jean-Philippe Préaux

Retour.

$$\mathbf{u} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \; ; \; \nabla f(\mathbf{u}) = \begin{pmatrix} p_2^2 - 2p_1p_2 \\ 2p_1p_2 - p_1^2 \end{pmatrix}$$

Copyright : Jean-Philippe Préaux

$$\mathbf{u} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \; ; \; \nabla f(\mathbf{u}) = \begin{pmatrix} p_2^2 - 2p_1p_2 \\ 2p_1p_2 - p_1^2 \end{pmatrix}$$

$$\nabla \varphi(\mathbf{u}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \; \nabla \psi_1(\mathbf{u}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \; ; \; \nabla \psi_2(\mathbf{u}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Copyright : Jean-Philippe Préaux

$$\mathbf{u} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \; ; \; \nabla f(\mathbf{u}) = \begin{pmatrix} p_2^2 - 2p_1p_2 \\ 2p_1p_2 - p_1^2 \end{pmatrix}$$

$$\nabla \varphi(\mathbf{u}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \; \nabla \psi_1(\mathbf{u}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \; ; \; \nabla \psi_2(\mathbf{u}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Les contraintes étant affines on peut appliquer les conditions (KKT).

Copyright : Jean-Philippe Préaux

$$\mathbf{u} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \; ; \; \nabla f(\mathbf{u}) = \begin{pmatrix} p_2^2 - 2p_1p_2 \\ 2p_1p_2 - p_1^2 \end{pmatrix}$$

$$\nabla\varphi(\mathbf{u}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \; \nabla\psi_1(\mathbf{u}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \; ; \; \nabla\psi_2(\mathbf{u}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Les contraintes étant affines on peut appliquer les conditions (KKT). Clairement les contraintes inégalitaires sont insaturées : $p_1, p_2 > 0 \implies \mu_1 = \mu_2 = 0$.

Copyright : Jean-Philippe Préaux

$$\mathbf{u} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \; ; \; \nabla f(\mathbf{u}) = \begin{pmatrix} p_2^2 - 2p_1p_2 \\ 2p_1p_2 - p_1^2 \end{pmatrix}$$

$$\nabla \varphi(\mathbf{u}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \; \nabla \psi_1(\mathbf{u}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \; ; \; \nabla \psi_2(\mathbf{u}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Les contraintes étant affines on peut appliquer les conditions (KKT). Clairement les contraintes inégalitaires sont insaturées : $p_1, p_2 > 0 \implies \mu_1 = \mu_2 = 0$. On obtient :

$$\begin{cases} p_2^2 - 2p_1p_2 + \lambda = 0 & (l_1) \\ 2p_1p_2 - p_1^2 + \lambda = 0 & (l_2) \end{cases}$$

Copyright : Jean-Philippe Préaux

En formant $(l_1) - (l_2)$ on obtient
$p_1^2 + p_2^2 - 4p_1p_2 = (p_1 + p_2)^2 - 6p_1p_2 = 0.$

Copyright : Jean-Philippe Préaux

En formant $(l_1) - (l_2)$ on obtient
$p_1^2 + p_2^2 - 4p_1p_2 = (p_1 + p_2)^2 - 6p_1p_2 = 0$. Puisque $p_1 + p_2 = 8$, on a $p_1p_2 = \frac{32}{3}$. Donc $p_1$, $p_2$ sont les racines du polynôme $x^2 - 8x + \frac{32}{3}$.

Copyright : Jean-Philippe Préaux

En formant $(l_1) - (l_2)$ on obtient $p_1^2 + p_2^2 - 4p_1p_2 = (p_1 + p_2)^2 - 6p_1p_2 = 0$. Puisque $p_1 + p_2 = 8$, on a $p_1p_2 = \frac{32}{3}$. Donc $p_1$, $p_2$ sont les racines du polynôme $x^2 - 8x + \frac{32}{3}$. On trouve $(\Delta = \frac{8^2}{3})$ :

$$p_1 = 4 - \frac{4}{\sqrt{3}} \; ; \quad p_2 = 4 + \frac{4}{\sqrt{3}} \; .$$

Copyright : Jean-Philippe Préaux

En formant $(l_1) - (l_2)$ on obtient
$p_1^2 + p_2^2 - 4p_1p_2 = (p_1 + p_2)^2 - 6p_1p_2 = 0$. Puisque $p_1 + p_2 = 8$, on a $p_1p_2 = \frac{32}{3}$. Donc $p_1$, $p_2$ sont les racines du polynôme $x^2 - 8x + \frac{32}{3}$. On trouve $(\Delta = \frac{8^2}{3})$ :

$$p_1 = 4 - \frac{4}{\sqrt{3}} \; ; \quad p_2 = 4 + \frac{4}{\sqrt{3}} \; .$$

*Question. Par compacité il existe aussi un minimum global que les conditions de Lagrange doivent déterminer !*

Copyright : Jean-Philippe Préaux

En formant $(l_1) - (l_2)$ on obtient
$p_1^2 + p_2^2 - 4p_1p_2 = (p_1 + p_2)^2 - 6p_1p_2 = 0$. Puisque $p_1 + p_2 = 8$, on a $p_1p_2 = \frac{32}{3}$. Donc $p_1$, $p_2$ sont les racines du polynôme $x^2 - 8x + \frac{32}{3}$. On trouve $(\Delta = \frac{8^2}{3})$ :

$$p_1 = 4 - \frac{4}{\sqrt{3}} \; ; \quad p_2 = 4 + \frac{4}{\sqrt{3}} \; .$$

*Question. Par compacité il existe aussi un minimum global que les conditions de Lagrange doivent déterminer !*
*Réponse : c'est $(p_2, p_1)$.*

Retour.

Copyright : Jean-Philippe Préaux

# Exercice 3.

Soit $A$ une matrice symétrique réelle. Justifier que le problème :

$$\max_{\|\mathbf{x}\| \leqslant 1} \mathbf{x}^\top A\mathbf{x}$$

admet une solution $\mathbf{u}$. Que représente $\mathbf{u}$ pour la matrice $A$ ?

Retour.

Copyright : Jean-Philippe Préaux

# Exercice 3.

Soit $A$ une matrice symétrique réelle. Justifier que le problème :

$$\max_{\|\mathbf{x}\| \leqslant 1} \mathbf{x}^\top A\mathbf{x}$$

admet une solution $\mathbf{u}$. Que représente $\mathbf{u}$ pour la matrice $A$ ?

Le domaine $\mathcal{D} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leqslant 1\}$ est un compact. L'application $f(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x}$ est quadratique et donc continue. Ainsi $f$ admet (au moins) un maximum sur $\mathcal{D}$.

Retour.

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$.

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0}$$

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0} \quad \Longrightarrow \quad A\mathbf{u} = -\mu\mathbf{u}\ .$$

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0} \quad \Longrightarrow \quad A\mathbf{u} = -\mu\mathbf{u}\ .$$

Ainsi :
– si $\mu \neq 0$, $\mathbf{u}$ est un vecteur propre de $A$ associé à la valeur propre $(-\mu) > 0$.

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0} \quad \Longrightarrow \quad A\mathbf{u} = -\mu\mathbf{u}\ .$$

Ainsi :
– si $\mu \neq 0$, $\mathbf{u}$ est un vecteur propre de $A$ associé à la valeur propre $(-\mu) > 0$. Dans ce cas la contrainte est saturée, $\|u\| = 1$, et $f(\mathbf{u}) = -\mu\,\mathbf{u}^\top\mathbf{u} = -\mu\,\|\mathbf{u}\|^2 = -\mu$.

Copyright : Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0} \quad \Longrightarrow \quad A\mathbf{u} = -\mu\mathbf{u}\ .$$

Ainsi :
– si $\mu \neq 0$, $\mathbf{u}$ est un vecteur propre de $A$ associé à la valeur propre $(-\mu) > 0$. Dans ce cas la contrainte est saturée, $\|u\| = 1$, et $f(\mathbf{u}) = -\mu\,\mathbf{u}^\top\mathbf{u} = -\mu\,\|\mathbf{u}\|^2 = -\mu$.
– si $\mu = 0$, $\mathbf{u}$ est un vecteur de $\ker A$.

Copyright - Jean-Philippe Préaux

Appliquons les conditions de KKT (où c'est possible). La contrainte est $\psi(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 - 1$. En un maximum $\mathbf{u} \neq 0$, il existe $\mu \leqslant 0$, tel que :

$$\nabla f(\mathbf{u}) + \mu \nabla \psi(\mathbf{u}) = 2\,A\mathbf{u} + 2\,\mu\,\mathrm{Id}\,\mathbf{u} = \mathbf{0} \quad \Longrightarrow \quad A\mathbf{u} = -\mu\mathbf{u}\ .$$

Ainsi :
– si $\mu \neq 0$, $\mathbf{u}$ est un vecteur propre de $A$ associé à la valeur propre $(-\mu) > 0$. Dans ce cas la contrainte est saturée, $\|u\| = 1$, et $f(\mathbf{u}) = -\mu\,\mathbf{u}^\top\mathbf{u} = -\mu\,\|\mathbf{u}\|^2 = -\mu$.
– si $\mu = 0$, $\mathbf{u}$ est un vecteur de $\ker A$.

On peut conclure : si $A$ a une valeur propre $> 0$, $\mathbf{u}$ est un vecteur propre unitaire associé à la plus grande valeur propre de $A$. Sinon $\mathbf{u}$ est n'importe quel élément du noyau de $A$.

Copyright : Jean-Philippe Préaux

Retour.