OPTIMISATION ET CONTRÔLE

Grégoire Allaire
Département de mathématiques appliquées
Ecole Polytechnique

- **Optimisation:** leçons 1 à 6.
- **Contrôle:** leçons 7 à 10.
- **Site web Moodle du cours:**
  `https://moodle.polytechnique.fr/course/view.php?id=9673`
- **Page du cours sur mon site web:**
  `http://www.cmap.polytechnique.fr/~allaire/cours_map435.html`
- **Me contacter:** `gregoire.allaire@polytechnique.fr`

Plan de cette leçon

1. Chapitre 1: introduction, motivation et exemples.
   - Optimisation.
   - Contrôle.
2. Chapitre 2: aspects théoriques de l'optimisation.
   - Définitions et notations.
   - Résultats d'existence.

Innombrables applications de l'optimisation et du contrôle:

- Sciences de l'ingénieur: industries de haute technologie (aéronautique, énergie, transport, communications), science des données, automatique, finance...
- Autres sciences: physique, mécanique, économie, informatique, sciences sociales...
- Domaines nouveaux ou inattendus: biologie, gestion, marketing, internet et réseaux sociaux...
- Chaque année de nouveaux problèmes arrivent !

Et toujours, trois étapes des mathématiques appliquées:

1. Modélisation.
2. Analyse mathématique du modèle.
3. Algorithmes de résolution numérique.

Qu'est ce que l'optimisation ?

- Pour améliorer un système, on modélise sa performance par une fonction $J : V \to \mathbb{R}$, où V est un espace de Hilbert qui représente les paramètres ou variables du modèle.
- Il peut y avoir des contraintes sur les paramètres, représentées par un sous-ensemble non-vide $K \subset V$.

Un problème d'optimisation s'écrit

$$\min_{v \in K \subset V} J(v).$$

(On minimise toujours: pour maximiser, considérer $\min -J(v)$.)

Remarque. Parfois les contraintes sont représentées par une fonction $C : V \to \mathbb{R}^M$ et on considère

$$\min_{v \in V \text{ tel que } C(v) \leq 0} J(v).$$

Pourquoi un espace de Hilbert ?

Est-il vraiment nécessaire d'étudier l'optimisation dans un espace de Hilbert de dimension infinie ?

On pourrait croire que non:

- beaucoup de problèmes d'optimisation sont posés en dimension finie, $V = \mathbb{R}^n$,
- après "discrétisation" tout se ramène à la dimension finie,
- c'est beaucoup plus simple en dimension finie.

Néanmoins, la dimension infinie est essentielle:

- souvent la variable d'optimisation est une fonction (comme en contrôle optimal), donc il faut travailler en dimension infinie,
- même en dimension finie, si la dimension est grande, le point de vue et les méthodes de la dimension infinie sont très utiles,
- les espaces de Hilbert sont les plus simples en dim. infinie.

1. Transport optimal
2. Problème d'affectation
3. Consommation des ménages
4. Apprentissage machine
5. Calcul des variations

Transport optimal

Optimisation de la livraison à N clients à partir de M entrepôts.

- stock s_i pour l'entrepôt $1 \leq i \leq M$
- commande r_j du client $1 \leq j \leq N$
- coût c_{ij} de transport unitaire entre l'entrepôt i et le client j
- quantité v_{ij} de marchandise allant de i vers le client j

On minimise le coût du transport

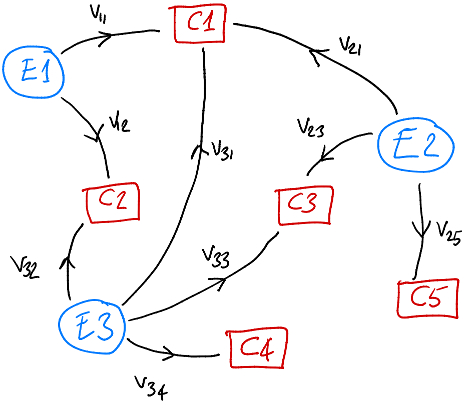
$$\min_{(v_{ij}) \in \mathbb{R}^{N \times M}} \left(\sum_{i=1}^{M} \sum_{j=1}^{N} c_{ij} v_{ij} \right)$$

sous les contraintes de limites des stocks et de satisfaction des commandes des clients

$$v_{ij} \geq 0, \ \sum_{j=1}^{N} v_{ij} \leq s_i, \ \sum_{i=1}^{M} v_{ij} = r_j \ \text{ pour } 1 \leq i \leq M, \ 1 \leq j \leq N.$$

Il s'agit d'un exemple de programmation linéaire.

Transport optimal



Entrepôts E i

Clients C j

Problème d'affectation

Soit N candidats, $1 \leq i \leq N$, et N postes, $1 \leq j \leq N$. Soit $a_{ij} = 1$ si le candidat i et le poste j sont en adéquation, sinon $a_{ij} = 0$. On optimise le nombre d'affectations "satisfaisantes". Si $\mathcal{S}_N$ est l'ensemble des permutations de $\{1, ..., N\}$, on maximise

$$\max_{\sigma \in \mathcal{S}_N} \sum_{i=1}^{N} a_{i\sigma(i)}.$$

Si on introduit les variables de décision v_{ij}, on obtient

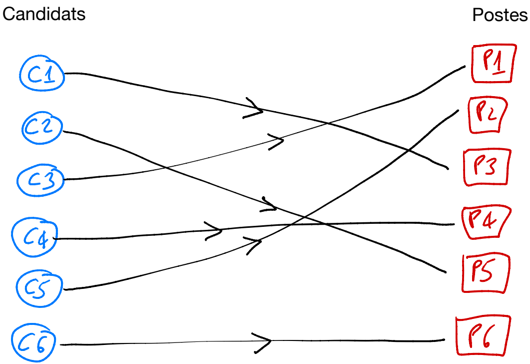
$$\max_{(v_{ij}) \in \mathbb{R}^{N \times N}} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} v_{ij} \right)$$

sous les contraintes

$$v_{ij} = 0 \text{ ou } 1, \ \sum_{j=1}^{N} v_{ij} \leq 1, \ \sum_{i=1}^{M} v_{ij} \leq 1 \ \text{ pour } 1 \leq i, j \leq N.$$

Il s'agit d'un exemple d'optimisation combinatoire ou en variables entières.

Problème d'affectation



Consommation des ménages

- Soit n marchandises dont les prix forment un vecteur $p \in \mathbb{R}_+^n$.
- Soit $x \in \mathbb{R}_+^n$ le vecteur des quantités achetées par un ménage.
- Soit $u(x)$ fonction croissante et concave de $\mathbb{R}_+^n$ dans $\mathbb{R}$ qui mesure la "satisfaction" que le ménage tire de sa consommation.
- Soit $b > 0$ le budget dont dispose le ménage.

La consommation du ménage est le vecteur x qui maximise

$$\max_{x \in \mathbb{R}_+^n, \, x \cdot p \leq b} u(x)$$

$u(x)$ est appelée fonction d'utilité.

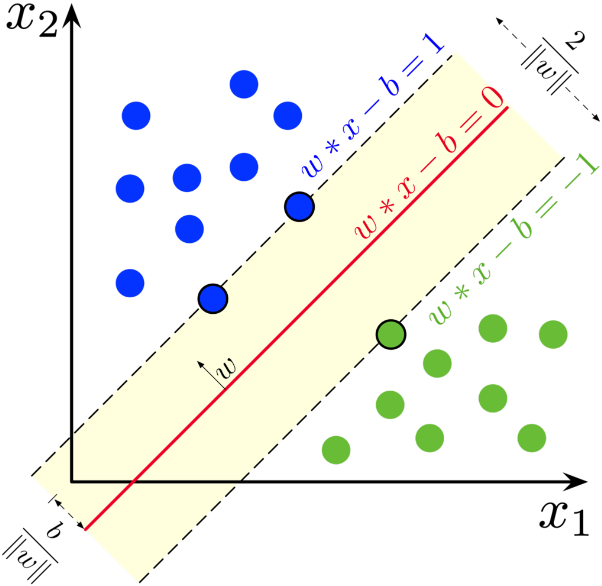
Apprentissage machine (ou SVM)

- Soit des données $(x_i)_{1\leq i\leq n}$ avec $x_i \in \mathbb{R}^d$ qui sont classées par un label $y_i = -1$ ou $+1$.
- Soit la fonction de prédiction, $h_{w,b}(x) = w\cdot x - b$, où $w \in \mathbb{R}^d$ et $b \in \mathbb{R}$ sont des paramètres à optimiser pour que
$$\operatorname{sgn}\Big(h_{w,b}(x_i)\Big) \approx y_i$$
- Soit la fonction de perte qui approche le "bon signe"
$$P(h,y) = \log(1+\exp(-hy)) \quad (h,y) \in \mathbb{R}\times\{-1,+1\}$$

Pour "apprendre", on minimise
$$\min_{w\in\mathbb{R}^d, b\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} P(h_{w,b}(x_i), y_i).$$

On peut "régulariser" le vecteur w et, pour $\ell > 0$, on minimise
$$\min_{w\in\mathbb{R}^d, b\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} P(h_{w,b}(x_i), y_i) + \frac{\ell}{2}\|w\|_2^2.$$

SVM (séparateur à vaste marge)



Calcul des variations

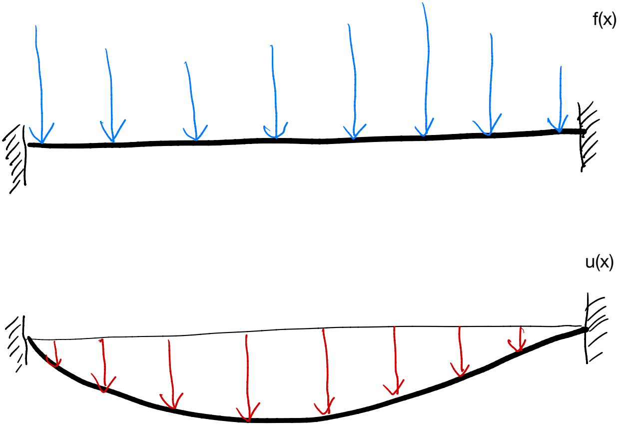
- Soit une poutre horizontale $(0, L)$.
- Soit une force verticale $f(x)$ qui agit sur la poutre.
- Soit $u(x)$ le déplacement vertical.
- Soit $\mu > 0$ le coefficient de rigidité de la poutre.

La position d'équilibre de la poutre est obtenue par minimisation de l'énergie (somme des énergies de déformation et potentielle)

$$\min_{u \in C^1(0,L)} \frac{1}{2} \int_0^L \mu |u'(x)|^2 dx - \int_0^L f(x)\, u(x)\, dx,$$

Si la poutre est encastrée à une ou deux extrémités on rajoute les contraintes $u(0) = 0$ et/ou $u(L) = 0$.
Beaucoup de problèmes de physique ou de mécanique sont issus de minimisation d'énergie.

Poutre horizontale



But du cours en optimisation

1. Existence (et parfois unicité) de solutions de problèmes d'optimisation:
   facile en dimension finie, délicat en dimension infinie (mais pas le but de ce cours).
2. Caractérisation des solutions: conditions d'optimalité:
   très important en pratique !
3. Algorithmes numériques d'optimisation:
   tous basés sur des dérivées (dans ce cours).

Optimisation dans la "vraie vie"

1. Pour des raison pédagogiques, ce cours se limite à des problèmes d'optimisation simples.
2. Dans les applications en ingénierie, les problèmes d'optimisation peuvent être plus complexes:
   1. modélisation de la fonction objectif et des contraintes,
   2. modèles stochastiques ou avec incertitudes,
   3. modèles avec équations aux dérivées partielles.
3. Pour aller plus loin en 3ème année:
   1. MAP 557 (recherche opérationnelle),
   2. MAP 562 (conception optimale de structures),
   3. MAP 545 (optimisation et apprentissage profond).

Qu'est ce que le contrôle ?

Pour une fonction $f : [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}^d$ on considère un système dynamique dont l'inconnue est $x : [0,T] \to \mathbb{R}^d$

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad \forall t \in [0,T], \\ x(0) = x_0 \in \mathbb{R}^d, \end{cases}$$

où $u : [0,T] \to \mathbb{R}^k$ est un "contrôle" à notre disposition.

Peut-on choisir le contrôle $u(t)$ pour atteindre une cible ?

$$x(T) \in C$$

Ou bien suivre au mieux une trajectoire ?

$$x(t) \approx x_{traj}(t) \quad \forall t \in [0,T]$$

Est-ce possible même si $k < d$?

Quelques exemples en contrôle

1. Contrôlabilité d'un tram
2. Aspirateur robot
3. Contrôle optimal
4. Traversée d'un canal

Contrôlabilité d'un tram

Soit $X(t) : [0, T] \to \mathbb{R}$ la position d'un tram sur une ligne.
Le tram est controlé par son accélération $u(t) : [0, T] \to \mathbb{R}$

$$m\ddot{X}(t) = u(t), \quad \forall t \in [0, T].$$

On introduit la vitesse $V(t) := \dot{X}(t)$ pour réécrire

$$\dot{x}(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{=:A} x(t) + \underbrace{\begin{pmatrix} 0 \\ 1/m \end{pmatrix}}_{=:B} u(t), \quad \text{avec } x(t) := \begin{pmatrix} X(t) \\ V(t) \end{pmatrix}.$$

Il y a $d = 2$ degrés de liberté et $k = 1$ variable de contrôle. Pour $T > 0$ et $x(0) = 0$, existe-t-il un contrôle u qui permette d'atteindre la prochaine gare (et de s'arrêter !) $x(T) = (X_c, 0)$?

(On peut aussi vouloir limiter l'accélération $u(t) \in [u_{\min}, u_{\max}]$.)

Aspirateur robot

Le robot est décrit par $(X, Y, \theta) : [0, T] \to \mathbb{R}^3$ $(d = 3)$ où (X, Y) est la position dans le plan et θ l'angle des roues par rapport à l'axe des X.

Le contrôle est dans le volant qui oriente les roues: $u : [0, T] \to \mathbb{R}$ $(k = 1)$. Pour une vitesse v constante, la dynamique est:

$$\begin{cases} \dot{X}(t) = v\cos(\theta(t)), \\ \dot{Y}(t) = v\sin(\theta(t)), \\ \dot{\theta}(t) = u(t), \qquad\qquad \leftarrow \text{action sur le système} \end{cases}$$

On peut envisager plusieurs problèmes:

1. atteindre une cible prescrite,
2. le faire en temps minimum,
3. le faire en action (du contrôle) minimum,
4. suivre une trajectoire.

Il s'agit d'un exemple de contrôle d'un système non-linéaire.

On mélange contrôle et optimisation !
Exemple du système linéaire-quadratique. EDO linéaire:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + f(t) & \forall t \in [0,T], \\ x(0) = x_0, \end{cases}$$

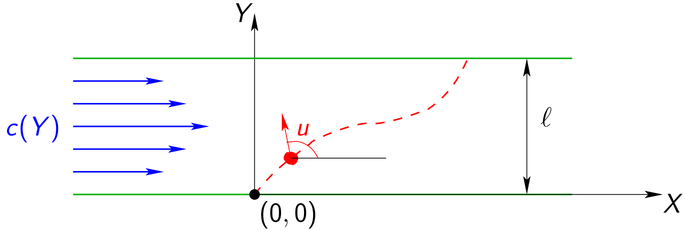
avec $x(t) \in \mathbb{R}^d$, $u(t) \in \mathbb{R}^k$, A et B matrices $d \times d$ et $d \times k$.
Soit une cible z_T et une trajectoire $z(t)$ que l'on veut approcher de manière optimale. Le problème est donc de résoudre

$$\min_{u(t) \in \mathbb{R}^k, \ t \in [0,T]} J(u),$$

avec un critère quadratique, défini pour 3 matrices $R, Q, D \geq 0$

$$J(u) = \int_0^T Ru \cdot u \, dt + \int_0^T Q(x-z) \cdot (x-z) \, dt + D\,(x(T) - z_T) \cdot (x(T) - z_T)$$

Traversée d'un canal



Une barque doit traverser un canal horizontal de largeur ℓ.
La barque a une vitesse (relative) constante v, et le courant a une vitesse $c(Y) > v$.
Le contrôle est l'angle $u \in [0, 2\pi]$ de la vitesse de la barque.
L'état de la barque est décrit par le couple $x := (X, Y) \in \mathbb{R}^2$:

$$\dot{x}(t) = f(x(t), u(t)) = \begin{pmatrix} v\cos(u(t)) + c(Y(t)) \\ v\sin(u(t)) \end{pmatrix}, \quad \forall t \in [0, T],$$

avec $x(0) = (0, 0)$.

Traversée d'un canal (suite)

On veut atteindre la berge opposée (contrainte de cible $Y(T) = \ell$) et aussi:

1. soit minimiser le déport latéral

$$J_1(u) = X(T),$$

   ce qui donne le problème de minimisation sous contraintes

$$\min_{u(t)\in[0,2\pi],\ t\in[0,T],\ Y(T)=\ell} J_1(u).$$

2. soit minimiser le temps de traversée

$$J_2(u) = \int_0^T dt = T,$$

   ce qui donne le problème de minimisation sous contraintes

$$\min_{u(t)\in[0,2\pi],\ t\in[0,T],\ Y(T)=\ell} J_2(u).$$

Dans ces deux problèmes, le temps T n'est pas fixé et est une inconnue supplémentaire.

But du cours en contrôle

1. Contrôlabilité et existence de contrôles:
   critère de Kalman dans le cas linéaire, contrôlabilité locale dans le cas non-linéaire.
2. Caractérisation des contrôles optimaux:
   conditions d'optimalité, notion d'état adjoint.
3. Principe du minimum de Pontryaguine:
   méthode pratique pour trouver un contrôle optimal.

Pour aller plus loin en 3ème année: MAP 561 (Automatique).

Soit V un espace de Hilbert réel. Soit $K \subset V$ un sous-ensemble non-vide. Soit $J : V \to \mathbb{R}$. Un problème d'optimisation s'écrit:

$$\inf_{v \in K} J(v) \quad \text{ou bien} \quad \min_{v \in K} J(v)$$

Notation. On utilise inf lorsqu'on ne sait pas a priori si la valeur minimum est atteinte et min dans le cas contraire.

S'il existe $v^* \in K$ tel que $J(v^*) \leq J(v)$ pour tout $v \in K$, alors

$$J(v^*) = \min_{v \in K} J(v).$$

Définition. L'infimum est le plus grand des minorants de J sur K:

$$\inf_{v \in K} J(v) = \max \left\{ C \in \mathbb{R} \cup \{-\infty\} \text{ tel que } C \leq J(v) \text{ pour tout } v \in K \right\}.$$

Suite minimisante

Définition. Une suite minimisante du critère (ou objectif) J sur l'ensemble K est une suite $(u^n)_{n\in\mathbb{N}}$ telle que

$$u^n \in K \ \forall\, n \quad \text{et} \quad \lim_{n\to+\infty} J(u^n) = \inf_{v\in K} J(v).$$

Par définition de l'infimum de J sur K **il existe toujours des suites minimisantes !**

Définition

Définition. On dit que u est un minimum global de J sur K si

$$u \in K \quad \text{et} \quad J(v) \geq J(u) \quad \forall v \in K \ .$$

On dit que u est un minimum local de J sur K si

$$u \in K \quad \text{et} \quad \exists \delta > 0 \ , \ \forall v \in K \ , \ \|v - u\| < \delta \Longrightarrow J(v) \geq J(u) \ .$$

Remarque. **Il y a une différence entre la théorie et la pratique numérique !** En théorie on considère un minimum global, alors qu'en pratique on calcule un minimum local.

Cas facile: optimisation en dimension finie $V = \mathbb{R}^N$.

Théorème. Soit K un ensemble fermé non vide de $\mathbb{R}^N$, et J une fonction continue sur K à valeurs dans $\mathbb{R}$ vérifiant la propriété, dite "infinie à l'infini",

$$\forall (u^n)_{n\geq 0} \text{ suite dans } K\ ,\ \lim_{n\to+\infty} \|u^n\| = +\infty \Longrightarrow \lim_{n\to+\infty} J(u^n) = +\infty\ .$$

Alors il existe au moins un point de minimum de J sur K. De plus, on peut extraire de toute suite minimisante de J sur K une sous-suite convergeant vers un point de minimum sur K.

(Idée: les fermés bornés sont compacts en dimension finie.)

Remarque. La propriété "infinie à l'infini" est évidente pour des ensembles bornés K.

Preuve. Soit $(u^n)_{n\in\mathbb{N}}$ une suite minimisante. Comme K est non vide, $J(u^n) \le C < +\infty$ pour tout n.
A cause de la propriété "infinie à l'infini", la suite est bornée

$$\|u^n\| \le C < +\infty \qquad \text{pour tout } n \in \mathbb{N}.$$

Il existe donc une sous-suite $(u^{n_k})_{k\in\mathbb{N}}$ qui converge vers une limite u. Comme K est fermé, u appartient à K. Comme J est continue, $J(u^{n_k})$ converge vers $J(u)$ qui est donc le minimum.

Autrement dit, si on note B_C la boule fermée de rayon C, on a

$$\inf_{v\in K} J(v) = \inf_{v\in K\cap B_C} J(v)$$

et la fonction continue J atteint son minimum sur le fermé borné $K \cap B_C \subset \mathbb{R}^N$.

Nettement plus délicat ! En général le théorème précédent est faux...

Problème: une fonction continue sur un fermé borné n'atteint pas toujours son minimum !

Il existe donc des problèmes d'optimisation "raisonnables" en dimension infinie qui n'admettent pas de solution.

Donnons deux contre-exemples (un théorique et un plus concret).

Mécanisme de non-existence: oscillation ou fuite à l'infini de la suite minimisante.

Premier contre-exemple

Soit l'espace de Hilbert des suites de carré sommable dans $\mathbb{R}$

$$\ell_2(\mathbb{R}) = \left\{ x = (x_i)_{i\geq 1} \text{ tel que } \sum_{i=1}^{+\infty} x_i^2 < +\infty \right\},$$

muni du produit scalaire $\langle x, y\rangle = \sum_{i=1}^{+\infty} x_i y_i$. Soit le problème

$$\inf_{x\in\ell_2(\mathbb{R})} J(x) = \left(\|x\|^2 - 1\right)^2 + \sum_{i=1}^{+\infty} \frac{x_i^2}{i}.$$

Lemme. Il n'existe pas de point de minimum $x^* \in \ell_2(\mathbb{R})$ tel que $J(x^*) \leq J(x)$ pour tout $x \in \ell_2(\mathbb{R})$.

Preuve. On va montrer que

$$\inf_{x\in\ell_2(\mathbb{R})} J(x) = 0.$$

Comme $J(x) > 0$ pour tout $x \in \ell_2(\mathbb{R})$, cela prouve le Lemme.

Premier contre-exemple (fin)

$$J(x) = \left(\|x\|^2 - 1\right)^2 + \sum_{i=1}^{+\infty} \frac{x_i^2}{i}.$$

On construit une suite minimisante x^n définie par $x_i^n = \delta_{in}$ pour tout $i \geq 1$

$$x^n = \underbrace{(0, ..., 0, 1, 0, 0, ...)}_{n}$$

On a $\|x^n\| = 1$,

$$J(x^n) = \frac{1}{n} \longrightarrow 0 \quad \text{quand } n \to +\infty,$$

et comme $J(x) \geq 0$ pour tout $x \in \ell_2(\mathbb{R})$, on en déduit

$$\inf_{x \in \ell_2(\mathbb{R})} J(x) = 0.$$

Remarque: on voit que la suite minimisante x^n "part à l'infini" et ne converge pas (bien qu'elle soit bornée).

Second contre-exemple

Problème modèle en science des matériaux pour expliquer le changement de phase solide/solide et les alliages à mémoire de forme.

Soit l'espace V des fonctions continues sur $[0,1]$ et dérivables par morceaux. On considère le problème

$$\inf_{v\in V} J(v) = \int_0^1 \Big((|v'(x)| - 1)^2 + v(x)^2 \Big) dx \,.$$

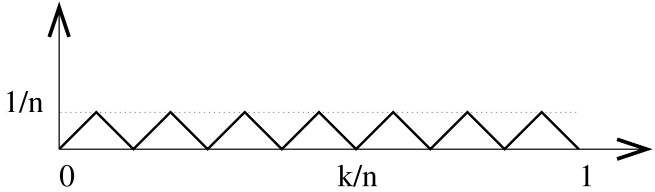
Lemme. Il n'existe pas de solution à ce problème de minimisation.

Remarque: difficulté indépendante du choix de l'espace. On peut aussi prendre l'espace de Hilbert $V = H^1(0,1)$.

Preuve. Il n'existe aucun $v \in V$ tel que $J(v) = 0$, et pourtant

$$\inf_{v \in H^1(0,1)} J(v) = 0,$$

car si on définit la suite u^n telle que $(u^n)' = \pm 1$



on a $J(u^n) = \int_0^1 u^n(x)^2 dx = \frac{1}{4n} \to 0$.

Remarque. On voit que la suite minimisante u^n oscille de plus en plus (elle converge mais pas sa dérivée).

Non-existence de solutions en dimension infinie

Remarque. Dans la pratique numérique, la non-existence de solutions se manifeste par la variation de la solution numérique selon les paramètres de calcul !

- Pour calculer une solution numérique, on approche l'espace V de dimension infinie par un espace V_N de dimension finie N (discrétisation).
- On appelle u_N la solution approchée dans V_N.
- Si on augmente N, la suite u_N ne converge pas !

Définition de la convexité

Pour obtenir des résultats d'existence en dimension infinie on rajoute une hypothèse de convexité.

Définition. Un ensemble $K \subset V$ est dit **convexe** si, pour tout $x, y \in K$ et tout réel $\theta \in [0,1]$, l'élément $(\theta x + (1-\theta)y)$ appartient à K.

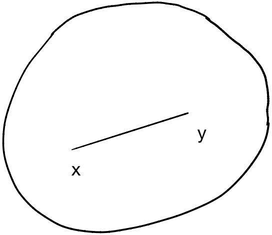
Définition. On dit qu'une fonction J définie sur un ensemble convexe non vide $K \in V$ et à valeurs dans $\mathbb{R}$ est **convexe** sur K si et seulement si

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) \quad \forall\, u, v \in K \;,\; \forall\, \theta \in [0,1] \;.$$

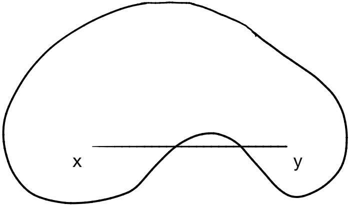
De plus, J est dite **strictement convexe** si l'inégalité est stricte lorsque $u \neq v$ et $\theta \in]0,1[$.

Interprétation de la convexité: la fonction J est sous sa corde.

Ensemble convexe K

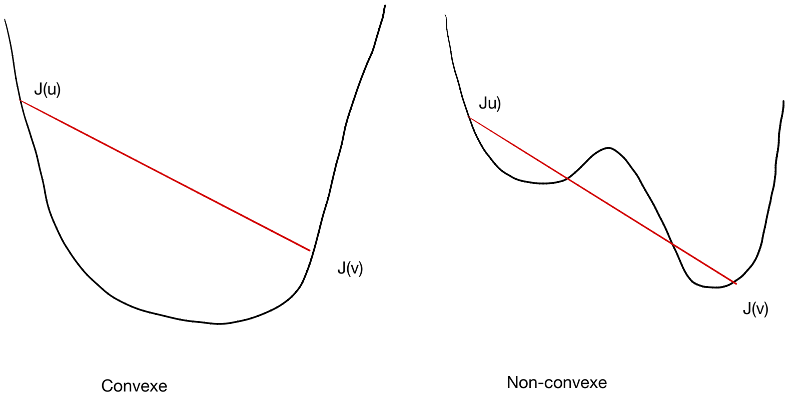


Convexe



Non-convexe

Fonction convexe J



Continuité des fonctions convexes

Lemme. Une fonction convexe J de V dans $\mathbb{R}$, localement majorée, est continue.

Preuve. Sans perte de généralité, on prouve la continuité en 0 et on suppose $J(0) = 0$. Soit $v \neq 0$, $\theta = \|v\| \leq 1$ et M la borne de J sur la boule unité. Par convexité

$$J(v) = J(\theta \frac{v}{\|v\|} + (1-\theta)0) \leq \theta J(\frac{v}{\|v\|}) + (1-\theta)J(0) \leq M\|v\|.$$

Par ailleurs, toujours par convexité

$$0 = J(0) \leq \frac{1}{1+\|v\|} J(v) + \frac{\|v\|}{1+\|v\|} J(\frac{-v}{\|v\|}) \leq \frac{1}{1+\|v\|}(J(v) + M\|v\|),$$

d'où l'on déduit la continuité

$$|J(v)| \leq M\|v\|.$$

Proposition. Si J est une fonction convexe sur un convexe K, alors tout minimum local est un minimum global.

Preuve. Soit u un minimum local de J sur K

$$\exists\, \delta > 0\ ,\ \forall\, w \in K\ ,\ \|w - u\| < \delta \implies J(w) \geq J(u)\ .$$

Soit $v \in K$. Pour $0 < \theta \ll 1$, $w_\theta = \theta v + (1-\theta)u \in K$ vérifie $\|w_\theta - u\| < \delta$. Donc

$$J(u) \leq J(w_\theta) \leq \theta J(v) + (1-\theta)J(u),$$

et ainsi $J(u) \leq J(v)$, i.e. u est un minimum global sur K.

Unicité dans le cas strictement convexe

Proposition. Si J est strictement convexe sur un convexe K, alors il existe au plus un point de minimum dans K.

Preuve. Si u_1 et u_2 sont deux minima et si $\theta \in]0,1[$, alors $w = \theta u_1 + (1-\theta)u_2$ est un minimum puisque $w \in K$ et que

$$\inf_{v \in K} J(v) \leq J(w) \leq \theta J(u_1) + (1-\theta)J(u_2) = \inf_{v \in K} J(v) \, .$$

Si J est strictement convexe, alors nécessairement $u_1 = u_2$.

Exemple et exercice

Exemple. La fonction $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$, définie sur $\mathbb{R}^n$, est convexe si la matrice A est positive. Elle est strictement convexe si A est définie positive.

Exercice.
(i) La somme de deux fonctions convexes est convexe.
(ii) Le maximum de deux fonctions convexe est convexe.

Définition. Une fonction $J(v)$ est concave si $-J(v)$ est convexe.

Forte convexité

Définition. On dit qu'une fonction J définie sur un ensemble convexe K est fortement convexe si et seulement si il existe $\alpha > 0$ tel que, pour tout $u, v \in K$ et pour tout $0 \leq \theta \leq 1$,

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) - \frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2$$

On dit aussi dans ce cas que J est α-convexe.

Exercice. Vérifier que $J(u)$ est fortement convexe, si et seulement si $J(u) - \frac{\alpha}{2}\|u\|^2$ est convexe.

Minoration des fonctions convexes

Proposition. Si J est convexe continue sur un convexe fermé non vide $K \subset V$, alors il existe une forme linéaire continue $L \in V'$ et une constante $\delta \in \mathbb{R}$ telles que

$$J(v) \geq L(v) + \delta \quad \forall\, v \in K \ .$$

Si de plus J est fortement convexe sur K, alors il existe deux constantes $\gamma > 0$ et $\eta \in \mathbb{R}$ telles que

$$J(v) \geq \gamma \|v\|^2 + \eta \quad \forall\, v \in K \ .$$

Remarque. Ce résultat nous sera utile pour majorer les suites minimisantes (voir la prochaine leçon).

Preuve. On définit l'épigraphe de J comme l'ensemble

$$Epi(J) = \{(\lambda, v) \in \mathbb{R} \times K \; , \; \lambda \geq J(v)\}.$$

Si J est convexe, alors $Epi(J)$ est un convexe dans $\mathbb{R} \times V$. Soit $v_0 \in K$ et $\lambda_0 < J(v_0)$. Puisque $(\lambda_0, v_0) \notin Epi(J)$, on déduit du théorème de séparation d'un point et d'un convexe l'existence de $\alpha, \beta \in \mathbb{R}$ et d'une forme linéaire continue $L \in V'$ tels que

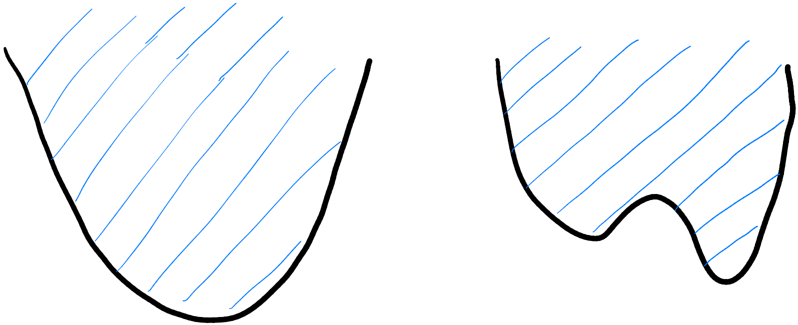
$$\beta\lambda + L(v) > \alpha > \beta\lambda_0 + L(v_0) \quad \forall\, (\lambda, v) \in Epi(J)\ .$$

Nécessairement $\beta \geq 0$ quand $\lambda \to +\infty$. De plus, $\beta \neq 0$ quand $v = v_0$. On choisit alors $\lambda = J(v)$, ce qui donne

$$J(v) + L(v)/\beta > \alpha/\beta \quad \forall v \in K.$$

Donc J est minorée par un hyperplan affine.

Epigraphe d'une fonction



Epigraphe d'une fonction J

Minoration des fonctions convexes (fin)

Finalement, si J est α-convexe, pour $\theta = 1/2$ on a

$$\frac{J(v)}{2}+\frac{J(v_0)}{2} \geq J\left(\frac{v+v_0}{2}\right)+\frac{\alpha}{8}\|v-v_0\|^2 \geq \frac{L(v)+L(v_0)}{2}+\frac{\alpha}{8}\|v-v_0\|^2+\delta \, .$$

Comme L est linéaire continue, $|L(v)| \leq \|L\|_{V'}\|v\|$, on a

$$J(v) \geq \frac{\alpha}{4}\|v\|^2 - \left(\|L\|_{V'} + \frac{\alpha\|v_0\|}{2}\right)\|v\| + C_1 \geq \frac{\alpha}{8}\|v\|^2 + \eta \, ,$$

par Cauchy-Schwarz, pour $\eta \in \mathbb{R}$ bien choisi.

Conclusion

Lors de la prochaine leçon:

1. on montrera l'existence de point de minimum pour les fonctions fortement convexes en dimension infinie,
2. on commencera l'étude des conditions d'optimalité.