Cours n°00: Optimisation en dimension 1: rappels, premières intuitions et difficultés

Nicolas DUPIN

http://nicolasdupin2000.wixsite.com/research

version du 15 mars 2022

Issu et étendu d'enseignements donnés à l'ENSTA Paris, Polytech' Paris-Saclay, et à l'université Paris-Saclay

Optimisation continue en dimension 1, bien connu ?

- Etudier une fonction de la variable réelle, les variations, tracé du graphe, qqch de bien connu pour vous normalement.
- Y a t'il des applications à l'optimisation de fonctions de la variable réelle ?
- Permet d'appréhender/rappeler des premières propriétés de l'optimisation, si on a une bonne vision de la dimension 1, est ce que tout se généralise en dimension supérieure ?
- On regarde ici la généralité des méthodes pour résoudre un problème d'optimisation en dimension 1, pas que les cas qui marchent bien.
- Permet d'appréhender les premières propriétés et difficultés des problèmes d'optimisation.
- Aspects algorithmiques : Si on a les propriétés d'existence, comment avoir des procédés constructifs permettant une implémentation efficace ?

Plan

Commençons par un exemple

Formalisations et généralisations

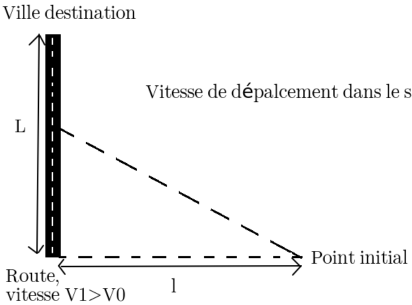
D'autres exemples

Conclusions

Commençons par un exemple en dimension 1

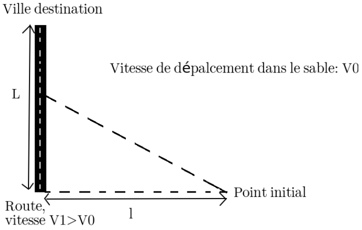
Un véhicule tout terrain veut rejoindre une ville, sa vitesse de déplacement est v_0 dans le désert ou $v_1 > v_0$ sur la route, situées à une distance de l.

(ou vous êtes à la plage, quel chemin le plus court en temps pour aller acheter une glace au bord de la route)



Question : quelle trajectoire suivre pour minimiser le temps de parcours ?

Propriété de Bellman



"Le plus court chemin entre deux points est la ligne droite", vrai à vitesse constante.

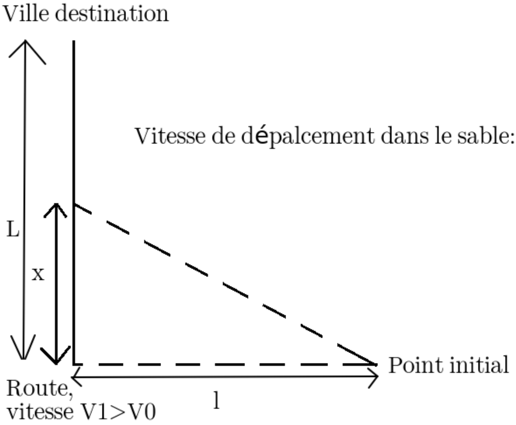
Une fois qu'un point de la route est atteint par un chemin optimal, le fin du chemin optimal se termine sur la route. Preuve : par l'absurde, sinon, on construirait un chemin strictement améliorant l'optimum avec la ligne droite.

Avant d'atteindre la route, le chemin optimal est en ligne droite dans le sable. Preuve : par l'absurde, sinon, on construirait un chemin strictement améliorant l'optimum avec la ligne droite.

$\Longrightarrow$ le chemin optimal est une ligne brisée, la seule variable à optimiser est le point d'entrée sur la route

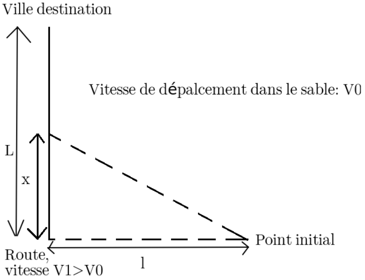
Trajectoire optimale en dimension 1, modélisation

Question : quel angle viser pour minimiser le temps de parcours ?



Modélisation : $x \in [0, L]$ point où la trajectoire rencontre la route

Optimisation, ne pas confondre variables et paramètres



Variable(s) d'un problème d'optimisation : le "levier" pour prendre une décision.

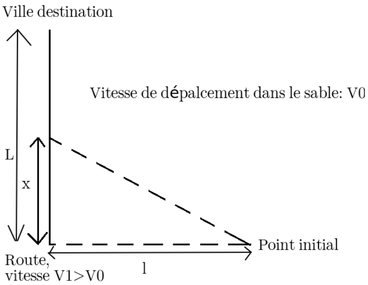
Ici, une seule variable $x \in [0, L]$, modélisant le point d'entrée sur la route.

Ici, la décision optimale se calcule en fonction de données d'entrées, nommés des *paramètres* : v_0, v_1, l, L. Le calcul de la décision optimale x^* se ferait idéalement en exprimant x^* en fonction des 4 paramètres, ie en exhibant un fonction F telle que $x^* = F(v_0, v_1, l, L)$.

Pour une valeur donnée des paramètres, on calculerait la solution optimale en calculant $F(v_0, v_1, l, L)$

Si les paramètres varient dans leur domaine réalisable, attention à ne pas nommer les paramètres "variables" qui définissent les décisions possibles, pour optimiser la fonction objectif en fonction des valeurs des paramètres d'entrée.

Mise en équations



Longueur du trajet : $\sqrt{l^2 + x^2} + (L - x)$

Durée du trajet : $T(x) = \frac{1}{v_0}\sqrt{l^2 + x^2} + \frac{1}{v_1}(L - x)$

On veut donc minimiser $T(x)$ pour $x \in [0, L]$.

Question : Comment vous-y prendriez vous à ce stade ?

Résolution

$$T(x) = \frac{1}{v_0}\sqrt{l^2 + x^2} + \frac{1}{v_1}(L - x)$$

$$T'(x) = \frac{1}{2v_0}\frac{2x}{\sqrt{l^2 + x^2}} - \frac{1}{v_1}$$

$$T'(x) = 0 \iff \frac{x}{\sqrt{l^2 + x^2}} = \frac{v_0}{v_1} = \alpha < 1$$
$$T'(x) = 0 \iff \frac{x^2}{l^2 + x^2} = \alpha^2$$
$$T'(x) = 0 \iff x^2 = \alpha^2(l^2 + x^2)$$
$$T'(x) = 0 \iff (1 - \alpha^2)x^2 = \alpha^2 l^2$$

Avec $x \geqslant 0$, il vient

$$T'(x) = 0 \iff x = \frac{\alpha}{\sqrt{1 - \alpha^2}} l$$
$$T'(x) > 0 \iff x > \frac{\alpha}{\sqrt{1 - \alpha^2}} l$$
$$T'(x) < 0 \iff x < \frac{\alpha}{\sqrt{1 - \alpha^2}} l$$

Tableau de variation

x	0		β		$+\infty$
$T'(x)$		$-$	0	$+$	
$T(x)$	$T(0)$	$\searrow$	$T(\beta)$	$\nearrow$	

On prouve que T admet un minimum sur $[0, L]$ en $\beta = \dfrac{\alpha}{\sqrt{1-\alpha^2}} l$ si $\beta < L$, sinon, le minimum est atteint en L par monotonie

$$x^* = F(v_0, v_1, l, L) \text{ avec } F(v_0, v_1, l, L) = \min\left(L, \frac{\frac{v_0}{v_1}}{\sqrt{1 - \frac{v_0^2}{v_1^2}}} l\right)$$

Calcul en temps $O(1)$, on peut utiliser la solution analytique pour coder cette simple fonction de meilleure réponse

Plan

Ordre total dans $\mathbb{R}$

Soit $a, b \in \mathbb{R}$.

On a soit $a \leqslant b$ soit $b \leqslant a$.

On a soit $a < b$, soit $b < a$, soit $a = b$.

Propriétés d'ordre total, fondamentales pour faire de l'optimisation, on sait toujours comparer des réels

Optimisation, définition de base

Soit X l'ensemble des décisions possibles d'un problème d'optimisation. On dit aussi que X est l'ensemble des *solutions* (ou *solutions réalisables*) du problème d'optimisation.

On définit une fonction de score f, dite *fonction objectif*, pour pouvoir évaluer les décisions possibles, que l'on cherchera à maximiser ou minimiser.

Si on cherche à minimiser f , x_1 est préféré à x_2, si $f(x_1) < f(x_2)$. L'ordre total sur $\mathbb{R}$ permet de toujours comparer des solutions deux à deux

Quitte à changer f en $-f$, on peut supposer que l'on minimise une fonction objectif pour la suite.

Exemple : pour le meilleur itinéraire entre 2 points, on peut minimiser le temps, la distance, le coût du trajet (essence+péages) ou l'impact carbone, ou maximiser le profit (avec du car-sharing). Dans ce cours, un objectif unique à choisir, ou alors composer une unique fonction de score à partir de ces trois critères (extension optimisation multi-objectif en M2)

Plus petit élément

Soit $A \subset \mathbb{R}$. On dit que A admet un plus petit élément si :

$$\exists a \in A, \ \ \forall a' \in A, \ \ a \leqslant a' \tag{1}$$

Un tel a est dit plus petit élément de A, et noté $a = \min A$,

Le plus petit élément est la meilleure formalisation pour un problème d'optimisation, on chercherait à calculer $\text{OPT} = \min\{f(x)|x \in X\}$. on cherche la valeur optimale OPT et une solution associée x telle que $\text{OPT} = f(x)$, i.e. $x \in$argmin$\{f(x) \mid x \in X\}$, i.e. trouver un élément de $\{x \in X \mid f(x) =$OPT$\}$.

Problème : si $A \subset \mathbb{R}$ sans hypothèse supplémentaire, il n'existe pas forcément de plus petit élément. ex : $A =]0, 1]$.

Si A est fini ou si A est un ensemble discret minoré, A admet un plus petit élément.

$\Longrightarrow$ Formaliser un problème d'optimisation comme $\text{OPT} = \min\{f(x) \mid x \in X\}$ correspond à ce qu'on veut en pratique, mais cela n'est pas toujours défini.

Remarque : en optimisation, on cherchera à fournir une solution à valeur optimale, et pas à obtenir l'ensemble des solutions optimales (plus combinatoire).

Borne inférieure

Soit $A \subset \mathbb{R}$. On dit que A admet une borne inférieure a notée $a =$inf A, si l'ensemble des minorants de A admet un plus grand élément :

$$a = \inf A \iff a = \max\{x \in \mathbb{R} \mid \forall a' \in A,\ x \leqslant a'\} \quad (2)$$

$$a = \inf A \iff \forall x \in \mathbb{R},\ \left(\forall a' \in A,\ \ x \leqslant a' \Longrightarrow x \leqslant a\right) \quad (3)$$

$$a = \inf A \iff \forall \varepsilon > 0,\ \exists a' \in A,\ a \leqslant a' \leqslant a + \varepsilon \text{ et } \forall a' \in A, a \leqslant a' \quad (4)$$

Si A est une partie de $\mathbb{R}$ non vide et minorée , A admet une borne inférieure. (hypothèses minimales pour faire de l'optimisation en pratique). Si $A =]0, 1]$, $\inf A = 0$

Si A admet un plus petit élément, on a $\min A = \inf A$, c'est le cas où $\inf A \in A$. Les cas où $\min A$ n'existe pas correspond à $\inf A \notin A$.

$\Longrightarrow$ Formaliser un problème d'optimisation comme $\mathrm{OPT} = \inf\{f(x) \mid x \in X\}$ est toujours défini sous les hypothèses minimalistes, mais ne fournit pas toujours une solution réalisable du problème.

$\Longrightarrow$ Comment peut on assurer l'existence de $\mathrm{OPT} = \min\{f(x) \mid x \in X\}$?

Intérieur et ouverts

Soit $A \subset \mathbb{R}$.

On dit que a est un *point intérieur* de A si $a \in A$ et s'il existe $\varepsilon > 0$ tel que $]a-\varepsilon, a+\varepsilon[\subset A$.

L'*intérieur* de A, noté $\mathring{A}$ est l'ensemble des points intérieurs de A. On a toujours $\mathring{A} \subset A$.

A est *ouvert* de $\mathbb{R}$ si et seulement si $\mathring{A} = A$. Exemples : $]0,1[$, $]a,b[$, $]a,+\infty[$, $]-\infty,b[$, $\mathbb{R}$

Une intersection d'un ensemble fini d'ouverts est un ouvert.

Une union quelconque d'ouverts est un ouvert.

Adhérence, fermeture et fermés

Soit $A \subset \mathbb{R}$.

On dit que a est un *point d'adhérence* de A si on peut toujours trouver des points de A aussi proches que possibles de a :
$\forall \varepsilon > 0,\ \exists a' \in A,\ a \leqslant a' \leqslant a + \varepsilon$

La *fermeture* de A, noté $\overline{A}$ est l'ensemble des points d'adhérence de A. On a toujours $A \subset \overline{A}$.

A est *fermé* de $\mathbb{R}$ si et seulement si $\overline{A} = A$. Exemples : $[0,1]$, $[a,b]$, $[a,+\infty[$, a $]-\infty,b]$, $\mathbb{R}$

Un ensemble compact $A \subset \mathbb{R}$ se définit comme un ensemble fermé et borné de $\mathbb{R}$.

Une union d'un ensemble fini de fermés est un fermé.

Une intersection quelconque de fermés est un fermé.

Définitions de limite d'une suite

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle, aussi noté $u \in \mathcal{F}(\mathbb{N},\mathbb{R})$. On dit que la suite u est convergente vers $l \in \mathbb{R}$, noté $\lim\limits_{n\to+\infty} u_n = l$ si et seulement si :

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |u_n - l| \leqslant \varepsilon$$

Dans les autres cas, u est dite divergente. Deux cas particuliers intéressants :

$$\lim_{n\to+\infty} u_n = +\infty \iff \forall M > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ u_n \geqslant M$$
$$\lim_{n\to+\infty} u_n = -\infty \iff \forall M > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ u_n \leqslant -M$$

Caractérisation séquentielle des fermés et des compacts

f est point d'adhérence de $F \subset \mathbb{R}$ si et seulement si il existe $(u_n) \in \mathcal{F}(\mathbb{N}, F)$ une suite d'éléments de F avec $\lim_{n \to +\infty} u_n = f$.

K est un compact de $\mathbb{R}$ si et seulement si de toute suite d'éléments de K, on peut en extraire une sous -suite convergente. (Théorème de Bolzano-Weierstrass)

O est un ouvert de $\mathbb{R}$ si et seulement si $\mathbb{R} \setminus O$ est un fermé, la caractérisation séquentielle d'un fermé s'applique au complémentaire $\mathbb{R} \setminus O$

Définitions de limite d'une fonction en un point

Soit $A \subset \mathbb{R}$. Soit $f : A \longrightarrow \mathbb{R}$ une application, soit $a \in \mathbb{R} \cap \overline{A}$.

$$\lim_{x \to a} f(x) = l \in \mathbb{R} \iff \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a+\alpha] \implies |f(x) - l| \leqslant \varepsilon$$
$$\lim_{x \to a} f(x) = +\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a+\alpha] \implies f(x) \geqslant M$$
$$\lim_{x \to a} f(x) = -\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a+\alpha] \implies f(x) \leqslant -M$$

$$\lim_{x \to a+} f(x) = l \in \mathbb{R} \iff \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A \cap]a, a+\alpha] \implies |f(x) - l| \leqslant \varepsilon$$
$$\lim_{x \to a+} f(x) = +\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap]a, a+\alpha] \implies f(x) \geqslant M$$
$$\lim_{x \to a+} f(x) = -\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap]a, a+\alpha] \implies f(x) \leqslant -M$$
$$\lim_{x \to a-} f(x) = l \in \mathbb{R} \iff \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a[\implies |f(x) - l| \leqslant \varepsilon$$
$$\lim_{x \to a-} f(x) = +\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a[\implies f(x) \geqslant M$$
$$\lim_{x \to a-} f(x) = -\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [a-\alpha, a[\implies f(x) \leqslant -M$$

Définitions de limite d'une fonction à l'infini

Soit $A \subset \mathbb{R}$ non minoré ou non majoré. Soit $f : A \longrightarrow \mathbb{R}$ une application.

$$\lim_{x \to +\infty} f(x) = l \in \mathbb{R} \iff \forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [\alpha, +\infty[\Longrightarrow\ |f(x) - l| \leqslant \varepsilon$$
$$\lim_{x \to +\infty} f(x) = +\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [\alpha, +\infty[\Longrightarrow\ f(x) \geqslant M$$
$$\lim_{x \to +\infty} f(x) = -\infty \iff \forall M > 0,\ \exists \alpha > 0,\ \forall x \in A \cap [\alpha, +\infty[\Longrightarrow\ f(x) \leqslant -M$$

$$\lim_{x \to -\infty} f(x) = l \in \mathbb{R} \iff \forall \varepsilon > 0,\ \exists \alpha < 0,\ \forall x \in A \cap]-\infty, \alpha] \Longrightarrow\ |f(x) - l| \leqslant \varepsilon$$
$$\lim_{x \to -\infty} f(x) = +\infty \iff \forall M > 0,\ \exists \alpha < 0,\ \forall x \in A \cap]-\infty, \alpha] \Longrightarrow\ f(x) \geqslant M$$
$$\lim_{x \to -\infty} f(x) = -\infty \iff \forall M > 0,\ \exists \alpha < 0,\ \forall x \in A \cap]-\infty, \alpha] \Longrightarrow\ f(x) \leqslant -M$$

Caractérisation séquentielle des limites

Soit $A \subset \mathbb{R}$. Soit $f : A \longrightarrow \mathbb{R}$ une application, soit $a \in \overline{A}$. Soit $l \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$

$\lim_{x \to a} f(x) = l$ si et seulement si pour toute suite réelle $(u_n)_{n \in \mathbb{N}}$ à valeur dans A telle que $\lim_{n \to +\infty} u_n = a$, on a $\lim_{n \to +\infty} f(u_n) = l$

Unification des définitions par rapport aux cas précédents, mais pas forcément plus facile à utiliser en pratique

Rappels : fonctions continues, dérivables, $\mathcal{C}^\infty$

Soit $f : A \to \mathbb{R}$.

f est continue en $a \in A$ si on a $\lim\limits_{x \to a} f(x) = l$

f est continue sur A si pour tout $a \in A$, f est continue en a.

f est dérivable en $a \in A$ si on a $\lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a} = l \in \mathbb{R}$, on note alors $f'(a) = l$

f est dérivable sur A si pour tout $a \in A$, f est dérivable en a. $x \in A \mapsto f'(x)$ est dite dérivée de f.

f est dérivable en $a \in A$ implique que f est continue en $a \in A$

f est de classe $\mathcal{C}^1$ sur A si f dérivable sur A et f' est continue sur A.

f est de classe $\mathcal{C}^2$ sur A si f dérivable sur A et f' est $\mathcal{C}^1$ sur A.

Extremum local/global

Soit $f : A \to \mathbb{R}$. On dit que

x_0 est un minimum local si

$$\exists \varepsilon > 0 \, \forall x \in]x_0 - \varepsilon, x_0 + \varepsilon[\cap A, \; f(x_0) \leqslant f(x)$$

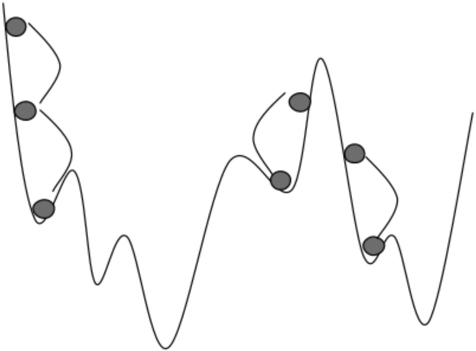
x_0 est un minimum global si

$$\forall x \in A, \; f(x_0) \leqslant f(x)$$

Définitions similaires de maximums locaux/globaux en changeant le sens des inégalités.

Un *extremum* local/global se définit comme un minimum ou un maximum local/global.

Illustration



Fonctions continues et extremum globaux

Theorem

Soit $A \subset \mathbb{R}$ un compact. Soit $f : A \to \mathbb{R}$ une fonction continue. Alors f est bornée et atteint ses bornes, il existe a_0, a_1 tels que pour tout $x \in A$, $m = f(a_0) \leqslant f(x) \leqslant f(a_1) = M$, ie. $f(A) \subset [m, M]$.

Justifie l'existence d'un plus petit élément, et d'une solution optimale aux problèmes de minimisation ou de maximisation de f sur A.

Question subsidiaire : a t'on un mode de calcul constructif de ces $m = f(a_0)$ et $M = f(a_1)$ ou est-ce qu'on ce sont juste des résultats d'existence ? Pour cela, on regarde la démonstration.

Pour la démonstration, on prouve uniquement l'existence du minimum global, la preuve est équivalente pour le maximum local.

Soit $m = \inf\{f(x) \mid x \in A\}$, bien définie une fois que l'on a prouvé par l'absurde que f est bornée sur A en utilisant la continuité.

Soit $n \in \mathbb{N}$. Par définition de la borne inférieure, il existe $a_n \in A$ tel que $|f(a_n) - m| \leqslant 1/(n+1)$ (ou $1/2^n$)

Avec le thm de Bolzano-Weierstrass au compact A, on extrait une sous suite convergente de (a_n), $(a_{\phi(n)})$ avec $\phi(n) \geqslant n$ et $\lim\limits_{n \to +\infty} a_{\phi(n)} = a_0$

A partir de $|f(a_{\phi(n)}) - m| \leqslant 1/(\phi(n)+1) \leqslant 1/(n+1)$, le passage à la limite donne $|f(a_0) - m| = 0$ avec la continuité de f.

Donc $m = f(a_0)$ et la borne inférieure est un plus petit élément !

$\Longrightarrow$ Si la preuve est constructive, elle repose sur des résultats d'existence et on n'a pas de procédé algorithmique issu de cette preuve.

Fonction continue et minimum global, autre résultat

Theorem

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue telle que $\lim\limits_{-\infty} f = \lim\limits_{+\infty} f = +\infty$. Alors f est minorée et atteint un minimum, il existe $a_0 \in \mathbb{R}$ tq pour tout x, $f(x) \geqslant f(a_0)$.

Preuve : on a $M_+ > 0$ tel que pour $x \geqslant M_+$, $f(x) \geqslant f(0) + 1$ et $M_- < 0$ tel que pour $x \leqslant M_-$, $f(x) \geqslant f(0) + 1$ en utilisant $\lim_{-\infty} f = \lim_{+\infty} f = +\infty$.

Sur le compact $[M_-, M_+]$, on a a_0 tq $f(a_0) = \min\limits_{x \in [M_-, M_+]} f(x)$.

Si $x \notin [M_-, M_+]$, $f(a_0) \leqslant f(0) < f(0) + 1 \leqslant f(x)$.

Dans tous les cas, $f(a_0) \leqslant f(x)$ pour tout $x \in \mathbb{R}$.

Autres résultats similaires

Theorem

Soit $f : [a, +\infty[\to \mathbb{R}$ *une fonction continue telle que* $\lim_{+\infty} f = +\infty$. *Alors* f *est minorée et atteint un minimum, il existe* $a_0 \in [a, +\infty[$ *tq pour tout* x, $f(x) \geqslant f(a_0)$.

Theorem

Soit $f : \mathbb{R} \to \mathbb{R}$ *une fonction continue telle qu'il existe un* x_0 *avec* $\lim_{-\infty} f = l_- > f(x_0)$ *et* $\lim_{+\infty} f = l_+ > f(x_0)$. *Alors* f *est minorée et admet un minimum.*

Et variantes hybrides ...

$\Longrightarrow$ Peut on avoir des schémas constructifs pour calculer des minimums globaux ?

Minimum local et dérivée première

Si $f : A \to \mathbb{R}$ est dérivable, et si un point intérieur $x \in \mathring{A}$ est un extremum local, cela implique $f'(x_0) = 0$.

Si l'existence d'un minimum est prouvée par les théorèmes d'existence et qu'on peut résoudre l'équation $f'(x) = 0$ et qu'il existe un nombre fini de solutions, on va chercher le minimum global parmi les solutions de $f'(x) = 0$ et les extremités $(A \diagdown \mathring{A})$.

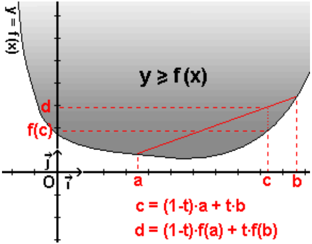
Minimum local et dérivée seconde

Si $f : A \to \mathbb{R}$ est de classe $\mathcal{C}^2$, les dérivées successives donnent une condition suffisante pour qu'un point intérieur $x \in \mathring{A}$ soit un extremum local

Si $f'(x_0) = 0$ et $f''(x_0) > 0$ alors x est un minimum local

Si $f'(x_0) = 0$ et $f''(x_0) < 0$ alors x est un maximum local

Cas particulier d'une fonction convexe



Rappel : Une fonction est convexe si et seulement si toute corde est au dessus de sa courbe : $\forall x, x',\ \forall \lambda \in [0;1]$,

$$f(\lambda x + (1-\lambda)x') \leqslant \lambda f(x) + (1-\lambda) f(x')$$

f est strictement convexe si les inégalités précédentes sont toujours strictes.

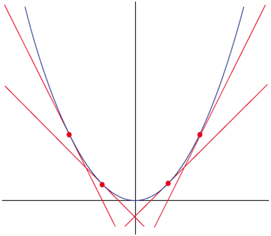
Si f est strictement convexe et définie sur un intervalle, un minimum local est minimum global (au plus unicité d'un minimum local)

Si f est convexe et définie sur un intervalle, l'ensemble des minimas globaux est soit vide, soit un intervalle.

$\Longrightarrow$ l'existence d'une solution (alors unique) à $f'(x) = 0$ implique d'avoir un minimum global

Caractérisation d'une fonction convexe par la dérivée

Soit $f : A \to \mathbb{R}$, définie sur un intervalle A



Si f est dérivable, f est convexe ssi f' est croissante

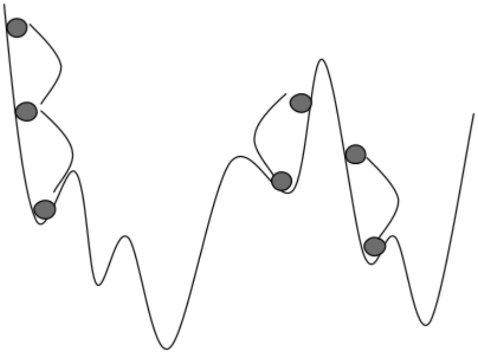
Si f est de classe $\mathcal{C}^2$, f est convexe ssi f'' est positive, i.e. pour tout $x \in A$, $f''(x) \geqslant 0$

Géométriquement, cela revient à dire qu'une fonction f est convexe si et seulement si sa courbe est au dessus de toutes ses tangentes

Si pour tout $x \in A$, $f''(x) > 0$, f est strictement convexe, mais cette condition n'est pas nécessaire.

Si pour tout $x \in A$, $f''(x) \geqslant 0$, et l'ensemble des zéros de f'' est isolé (par exemple fini, dénombrable ou discret), f est bien strictement convexe.

De manière générale



$\Longrightarrow$ De multiples extremums locaux ne sont pas extremums globaux en général

Contre-exemples pour illustrer l'importance d'hypothèses

Trouver des contre-exemples de fonctions réelles f définie sur un ensemble K :

- minorée, mais pourtant n'admettant pas de minimum local
- dérivable sur $[0,1]$, ayant un minimum local $f(x_0)$ en x_0, mais pourtant $f'(x_0) \neq 0$
- dérivable sur $[0,1]$ ayant une infinité de minimums locaux
- une fonction strictement convexe admettant plusieurs minimums locaux
- une fonction strictement convexe de $\mathbb{R}$ dans $\mathbb{R}$ admettant aucun minimum local.

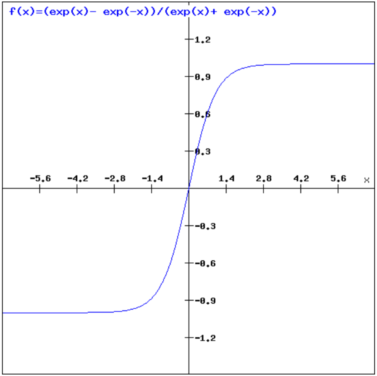
Fonction minorée, mais pourtant n'admettant pas de minimum local (1)

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$x \longmapsto \text{th}x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Minorée par 0, $\lim_{-\infty} f = 0$

pas de minimum local dans $\mathbb{R}$



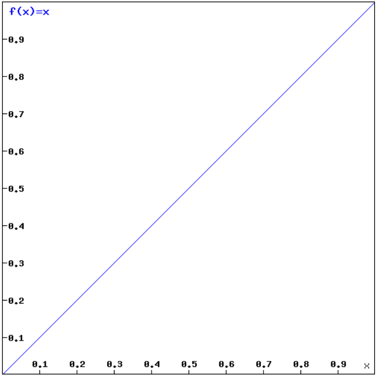
Fonction minorée, mais pourtant n'admettant pas de minimum local (2)

$f :]0,1] \longrightarrow \mathbb{R}$

$x \longmapsto x$

Minorée par 0, $\lim_0 f = 0$

pas de minimum local dans $]0,1]$



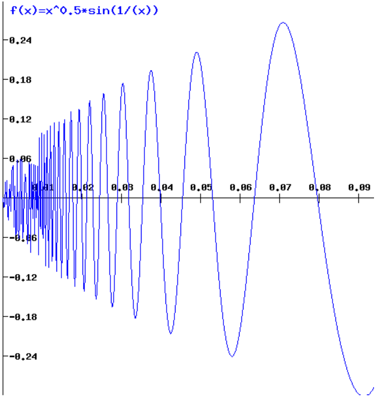
Fonction dérivable sur $[0,1]$ ayant une infinité de minimums locaux

$f : [0,1] \longrightarrow \mathbb{R}$

$x \longmapsto \sqrt{x} \sin \frac{1}{x}$

$f(0) = 0$ et $f'(0) = 0$, prolongement par continuité

Dans ce cas : trouver le minimum global en énumérant tous les minimums locaux ?



Fonction ayant un minimum local $f(x_0)$ en x_0, mais pourtant $f'(x_0) \neq 0$

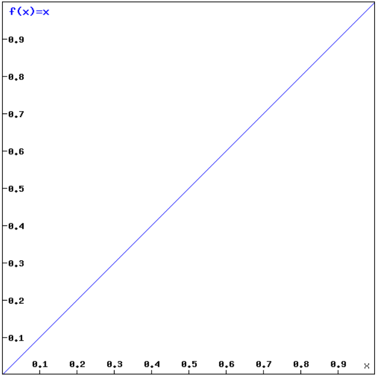
$f : [0,1] \longrightarrow \mathbb{R}$

$x \longmapsto x$

Minimum local en 0

$f'(0) = 1$

Rq : c'est même un minimum global



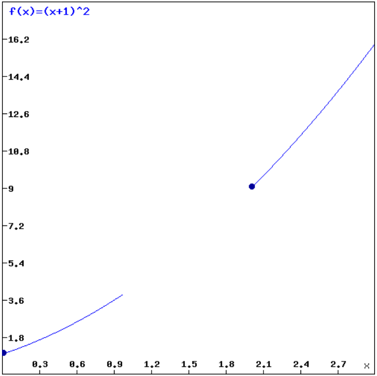
une fonction strictement convexe admettant plusieurs minimums locaux

$f : [0,1] \cup [2,3] \longrightarrow \mathbb{R}$

$x \longmapsto (x+1)^2$

Minimums locaux en 0 et 2

$\Longrightarrow$ Attention à la convexité de l'ensemble de définition ! !



Une fonction strictement convexe admettant aucun minimum local

Fonction strictement convexe admettant aucun minimum local :

$f : \mathbb{R} \longrightarrow \mathbb{R}$

$x \longmapsto e^x$

On peut même avoir une fonction strictement convexe et non minorée :

$f : \mathbb{R} \longrightarrow \mathbb{R}$

$x \longmapsto e^x - x$

$\Longrightarrow$ Attention, une fonction strictement convexe ne ressemble pas forcément à une parabole ! !

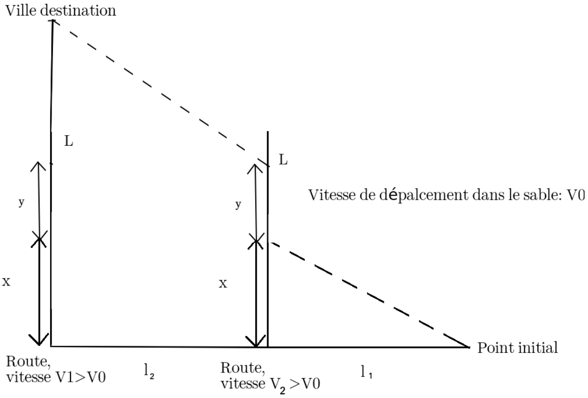
Optimisation continue en dimension 1, bilan partiel

- En dimension 1, une unique décision à optimiser.
- Extremums global/local
- Pour une fonction dérivable, extremums locaux et dérivées sont liés.
- En dimension 1, on a la notion de monotonie (tableaux de variation), contrairement en dimension supérieure.
- Sur l'exemple, si on sait calculer la dérivée et les solutions de l'équation $f'(x) = 0$, on peut obtenir les solutions de l'optimisation de manière analytique.
- Même en dimension 1, on a pu trouver des exemples mettant en défaut la méthodologie précédente.
- Question : sur des problèmes applicatifs avec des fonctions et structures usuelles, peut on rencontrer ces mêmes difficultés ?

Plan

Variante de l'exemple préliminaire

On rajoute une seconde route intermédiaire :



Modélisation : $x \in [0, L]$ point où la trajectoire rencontre la première route

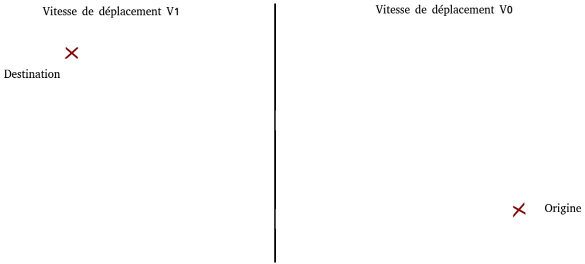
Modélisation : $y \in [0, L]$, tq y est la distance parcourue sur la deuxième route

$\Longrightarrow$ Optimiser une fonction à valeur réelle avec deux variables réelles

Un nouvel exemple en dimension 1

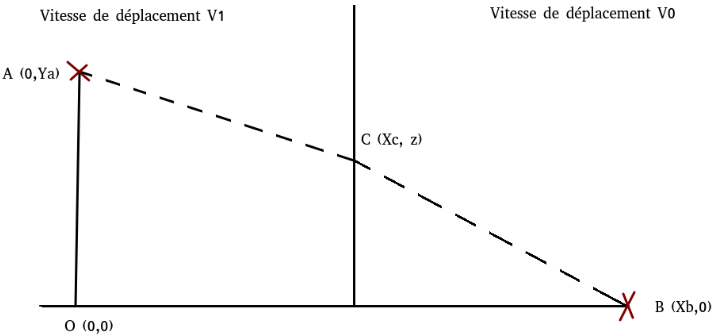
Vous êtes à la plage, vous sortez de l'eau, quel chemin prendre pour revenir le plus vite possible à votre serviette située sur une portion de sable sec où on marche plus lentement que sur le sable mouillé.

(ou pour les plus belliqueux, trajectoire optimale pour un char d'assault sur deux terrains différents où le char progresse à deux vitesses différentes)



Question : quelle trajectoire suivre pour minimiser le temps de parcours ?

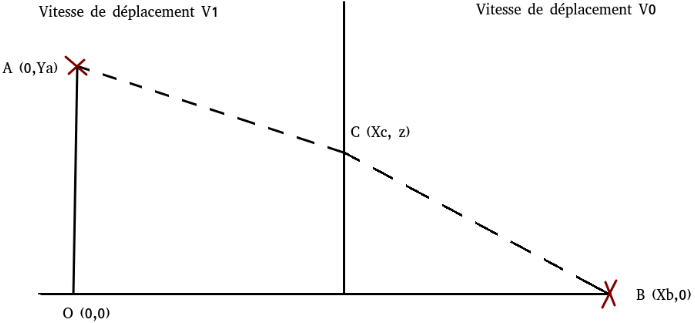
Propriété de Bellman



De même que sur l'exemple précédent, si on considère dans le chemin optimal deux points d'un même milieu, le chemin optimal est en ligne droite entre ces points. Preuve : par l'absurde.

$\Longrightarrow$ le chemin optimal est une ligne brisée, la seule variable à optimiser est le point de changement des milieux

Trajectoire optimale en dimension 1, modélisation



On place le repère de sorte à ce que $x_A = 0$ et $y_B = 0$. x_C est connu, définissant la séparation. La seule variable à optimiser est $z = y_C$, à optimiser en fonction des paramètres connus y_A, x_B et x_C. On peut considérer que $y_A, x_B, x_C \geqslant 0$.

$t_{AC} = \frac{AC}{v_0} = \frac{1}{v_0}\sqrt{x_C^2 + (z - y_A)^2}$
$t_{BC} = \frac{BC}{v_1} = \frac{1}{v_1}\sqrt{(x_C - x_B)^2 + z^2}$

On cherche donc à minimiser pour $z \in \mathbb{R}$:

$$f(z) = \frac{1}{v_0}\sqrt{x_C^2 + (z - y_A)^2} + \frac{1}{v_1}\sqrt{(x_C - x_B)^2 + z^2}$$

On cherche à minimiser pour $z \in \mathbb{R}$ la fonction de classe $\mathcal{C}^\infty$:

$$f(z) = \frac{1}{v_0}\sqrt{x_C^2 + (z - y_A)^2} + \frac{1}{v_1}\sqrt{(x_C - x_B)^2 + z^2}$$

On a $\lim_{z \to +\infty} f(z) = \lim_{z \to -\infty} f(z) = +\infty$, donc f admet bien un minimum global.

N.B : généralise pour qu'une fonction continue définie sur un compact est bornée et atteint ses bornes

On peut aussi remarquer que le minimum est nécessairement atteint sur $[0, y_A]$ avec des arguments géométriques simples.

$$f(z) = \frac{1}{v_0}\sqrt{x_C^2 + (z - y_A)^2} + \frac{1}{v_1}\sqrt{(x_C - x_B)^2 + z^2}$$

$$f'(z) = \frac{2(z - y_A)}{2v_0\sqrt{x_C^2 + (z - y_A)^2}} + \frac{2z}{2v_1\sqrt{(x_C - x_B)^2 + z^2}}$$

$$f'(z) = \frac{z - y_A}{v_0\sqrt{x_C^2 + (z - y_A)^2}} + \frac{z}{v_1\sqrt{(x_C - x_B)^2 + z^2}}$$

Ici, il n'est pas évident de déterminer le signe de f' comme sur le premier exemple. Cela se ramène à étudier le signe d'un polynôme de degré 4, on peut utiliser la méthode de Ferrari pour trouver les racines (bon courage avec les paramètres ...) et il faut ensuite identifier si les racines sont réelles ou complexes (bon courage encore ...). Ici, on trouverait (prouvé par la suite) que le polynôme en question a une unique racine réelle ...

Dans la suite, on effectue une dérivation supplémentaire pour déterminer le signe de f'.

$$f'(z) = \frac{z - y_A}{v_0\sqrt{x_C^2 + (z - y_A)^2}} + \frac{z}{v_1\sqrt{(x_C - x_B)^2 + z^2}}$$

$$f''(z) = \frac{v_0\sqrt{x_C^2 + (z - y_A)^2} - (z - y_A)v_0\frac{z - y_A}{\sqrt{x_C^2 + (z - y_A)^2}}}{v_0^2(x_C^2 + (z - y_A)^2)}$$
$$+ \frac{v_1\sqrt{(x_C - x_B)^2 + z^2} - zv_1\frac{z}{\sqrt{(x_C - x_B)^2 + z^2}}}{v_1(x_C - x_B)^2 + z^2}$$

$$f''(z) =$$
$$\frac{x_C^2 + \cancel{(z - y_A)^2} - \cancel{(z - y_A)^2}}{v_0(x_C^2 + (z - y_A)^2)\sqrt{x_C^2 + (z - y_A)^2}} + \frac{(x_C - x_B)^2 + \cancel{z^2} - \cancel{z^2}}{(v_1^2(x_C - x_B)^2 + z^2)\sqrt{(x_C - x_B)^2 + z^2}}$$

On a donc $f''(z) > 0$ et f est strictement convexe sur $\mathbb{R}$.

Cela prouve l'unicité du minimum global, $f(z_0)$ où z_0 est défini comme l'unique solution réelle de $f'(z_0) = 0$. Pour $z < z_0$, on a $f'(z_0) < 0$ et pour $z > z_0$, on a $f'(z_0) > 0$.

Remarque : de manière générale, on détermine le signe de f' à partir du signe de f'' en utilisant un tableau de variation. Ici, on prouve que f' est strictement croissante, négative puis positive, un seule racine de $f'(x) = 0$.

$$f'(0) = \frac{-y_A}{v_0\sqrt{x_C^2 + y_A^2}} < 0$$

$$f'(y_A) = \frac{y_A}{v_1\sqrt{(x_C - x_B)^2 + y_A^2}} > 0$$

Cela prouve que $0 < z_0 < y_A$

On définit les suites $(\alpha_n), (\beta_n)$ récursivement par :
$\alpha_0 = 0, \beta_0 = y_A$
$\alpha_{n+1}, \beta_{n+1} = \alpha_n, \frac{1}{2}(\beta_n + \alpha_n)$ si $f'(\frac{1}{2}(\beta_n + \alpha_n)) \geqslant 0$
$\alpha_{n+1}, \beta_{n+1} = \frac{1}{2}(\beta_n + \alpha_n), \beta_n,$ si $f'(\frac{1}{2}(\beta_n + \alpha_n)) < 0$

Le suites $(\alpha_n), (\beta_n)$ sont adjacentes et convergentes, $\lim\limits_{n \to +\infty} \alpha_n = \lim\limits_{n \to +\infty} \beta_n = z_0$ où z_0 est l'unique solution réelle de $f'(z_0) = 0$.

On a $\beta_n - \alpha_n = \dfrac{y_A}{2^n}$ et $\alpha_n < z_0 \leqslant \beta_n$, ce qui permet de calculer z_0 avec une précision décimale définie

N.B : vous avez déjà vu ce schéma pour prouver le théorème des valeurs intermédiaires. L'intérêt n'est pas que théorique mais aussi pratique, la vitesse de convergence des suites permet de calculer en peu d'itérations un arrondi de z_0 à la précision voulue.

Algorithme dichotomique général

Algorithme : Résolution dichotomique de $f'(x) = 0$

Entrée :
- f' continue
- $a < b$ tels que $f'(a)f'(b) < 0$
- une tolérance numérique $\varepsilon > 0$

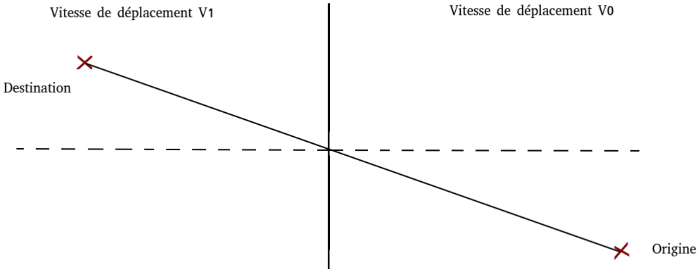
Sortie : une approximation numérique à ε près d'une racine de f'

$\alpha := a, \beta := b$
  tant que $\beta - \alpha > 2\varepsilon$
    $c := \frac{\alpha+\beta}{2}$
    si signe($f'(c)$) =signe($f'(\alpha)$)
      $\alpha := c$
    sinon $\beta := c$
  fin tant que
return $\frac{\alpha+\beta}{2}$

Solution approchée vs analytique

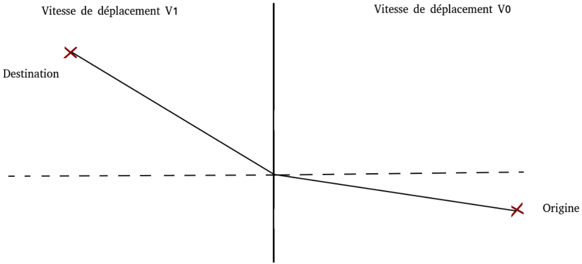
- Dans le premier exemple, on avait une solution analytique, il suffit d'implémenter la fonction de réponse optimale en fonction des paramètres.
- Avec l'algorithme dichotomique approché, on effectue quelques étapes d'algorithme (convergence logarithmique rapide) pour avoir une valeur numérique, en utilisant les valeurs définie des paramètres.
- Est-ce important d'avoir uniquement une valeur approchée ? à quelle précision est on limité pour l'application pratique pour implémenter/mettre en pratique une solution optimale ?
- Avoir une solution analytique est un cas "chanceux", la résolution analytique d'équation $F(x) = 0$ n'est pas générique.
- Sur le dernier exemple, peut-on avoir une solution analytique ? Ne reconnaissez vous pas un tel problème ?

Intuition graphique (1)



Si $v_0 = v_1$, l'optimum est une ligne droite

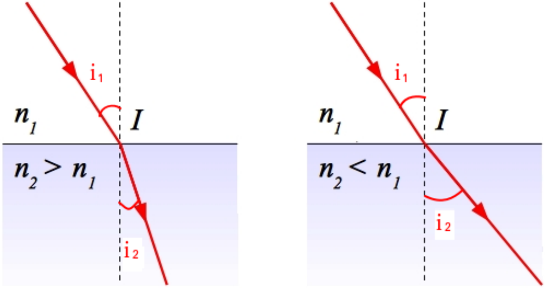
Intuition graphique (2)



Si $v_0 < v_1$, l'optimum traverse moins le secteur où $v = v_0$ et fait une distance plus grande sur le secteur $v = v_1$

$\Longrightarrow$ ça en vous rappelle rien ?

Lois de Snell-Descartes

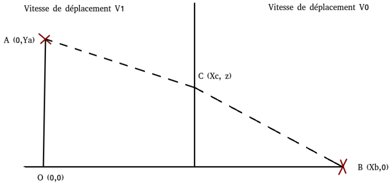


Lois de Snell-Descartes : $n_1 \sin i_1 = n_2 \sin i_2$

Principe de Fermat en Optique Géométrique : la lumière suit le chemin le plus court en temps.

n est le rapport entre la vitesse (célérité) de la lumière dans le vide divisée par celle de la lumière dans le milieu en question.

On aurait alors la relation : $\dfrac{1}{v_1} \sin i_1 = \dfrac{1}{v_0} \sin i_0$ sur notre problème.



Utilisant $\frac{1}{v_1}\sin i_1 = \frac{1}{v_0}\sin i_0$ pour définir une telle solution exprimée avec la variable z, cela donne :

$$\sin i_1 = \frac{y_A - z}{AC} = \frac{y_A - z}{\sqrt{x_c^2 + (y_A - z)^2}}$$

$$\sin i_0 = \frac{z}{\sqrt{z^2 + (x_B - x_C)^2}}$$

$\frac{1}{v_1}\sin i_1 = \frac{1}{v_0}\sin i_0$ donne la même équation que $f'(z) = 0$! ! ! !

Comme il a été prouvé que la solution est unique, cela démontre bien que la solution optimale vérifie $\frac{1}{v_1}\sin i_1 = \frac{1}{v_0}\sin i_0$

Remarques sur ce dernier résultat

On a prouvé rigoureusement que la solution optimale vérifie la loi de Snell-Descartes, mais c'est décevant, preuve "tirée du chapeau"

Au prochain cours, de nouveaux outils permettront de retrouver de manière directe et constructive ce résultat.

Souvenirs de physique : l'application des lois de Descartes n'impliquait pas de trouver les racines d'un polynôme de degré 4. Pourquoi ?

Dans les exercices de physique, la direction incidente était donnée, la lumière suit alors le chemin le plus court en temps, ce qui explique les lois de réfraction.

Ici, on donnait les points origine et destination. Au final, la trajectoire suit la même loi. En utilisant les lois de Descartes pour trouver de le meilleur chemin (ici la direction incidente), le problème se résout analytiquement avec un polynôme de degré 4, quelque soit la méthode

Plan

Commençons par un exemple

Formalisations et généralisations

D'autres exemples

Conclusions

Bilan, optimisation continue en dimension 1

- Optimisation de la variable réelle : dimension 1, une unique décision à optimiser. On a naturellement de l'optimisation d'une fonction à valeurs réelle avec plusieurs variables réelles. En dimension 1, on a de l'application en optique géométrique.

- Extremums global/local. Pour une fonction dérivable, extremums locaux et dérivées sont liés. Résoudre des équations $f'(x) = 0$ est une base pour déterminer des minimum/maximums locaux. Savoir résoudre ces équations et éventuellement déterminer le signe de f' permet de calculer analytiquement des optimums.

- Des arguments de convexité, localement ou globalement peuvent aider à prouver que l'on a un minimum unique.

- Problème majeur : La résolution d'équations $f'(x) = 0$ ne peut pas toujours se faire de manière analytique, même en dimension 1. En dimension 1, si f' est continue, et qu'a a a, b tels que $f'(a)f'(b) < 0$, une recherche dichotomique calcule efficacement avec une bonne précisions numérique de telles racines.

- En dimension n quelconque, comment ces résultats s'étendent ?