

# IFT 3515

## Fonctions à une variable

Fabian Bastin  
DIRO  
Université de Montréal

Concentrons-nous dans un premier temps sur des fonctions univariées.

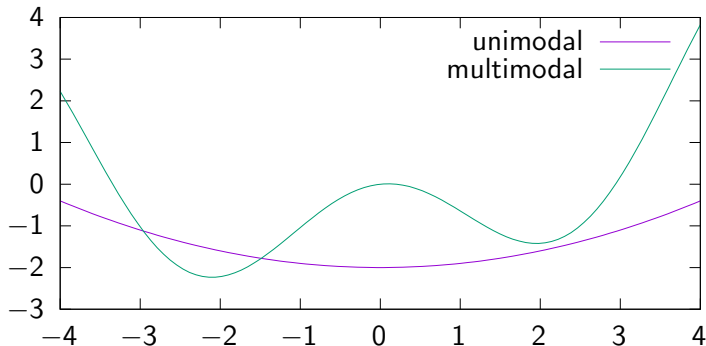
Le problème que nous cherchons à résoudre est

$$\begin{aligned} \min_x f(x) \\ \text{t.q. } x \in \mathcal{X} \subseteq \mathcal{R}. \end{aligned}$$

Même pour ce problème simple, nous allons devoir étudier les caractéristiques de la fonction pour développer des méthodes appropriées.

# Unimodalité

Une fonction est dite unimodale si elle possède un seul mode, c'est-à-dire un et un seul point où elle atteint un maximum local, dans le cas d'un problème de maximisation, ou un et un seul point où elle atteint un minimum local, dans le cas d'un problème de minimisation.



## Définition (Fonction unimodale – version maximisation)

*Une fonction  $f(x)$  est une fonction unimodale si pour une certaine valeur  $m$ , elle est monotonement croissante pour  $x \leq m$  et monotonement décroissante pour  $x \geq m$ . Dans ce cas, la valeur maximale de  $f(x)$  est  $f(m)$  et il n'y a pas d'autres maximums locaux.*

## Définition (Fonction unimodale – version minimisation)

*Une fonction  $f(x)$  est une fonction unimodale si pour une certaine valeur  $m$ , elle est monotonement décroissante pour  $x \leq m$  et monotonement croissante pour  $x \geq m$ . Dans ce cas, la valeur minimale de  $f(x)$  est  $f(m)$  et il n'y a pas d'autres minimums locaux.*

Une fonction unimodale n'est pas nécessairement continue, et encore moins dérivable.

# Principe général

Soit  $f : [a, b] \rightarrow \mathcal{R}$  unimodale. Considérons le problème

$$\min_{x \in [a, b]} f(x),$$

et soit  $x^* \in \arg \min \{f(x) \mid x \in [a, b]\}$ .

Soient  $x^G$  et  $x^D$  tels que  $a < x^G < x^D < b$ . Alors

$$f(x^G) < f(x^D) \Rightarrow x^* \in [a, x^D]$$

$$f(x^G) > f(x^D) \Rightarrow x^* \in [x^G, b]$$

$$f(x^G) = f(x^D) \Rightarrow x^* \in [x^G, x^D].$$

# Méthode de Fibonacci

**Hypothèse:**  $f(x)$  est unimodale.

L'approche consiste

- à choisir un certain nombre de points selon une stratégie basée sur les nombres de Fibonacci
- à évaluer séquentiellement la valeur de la fonction à ces points
- identifier un plus petit intervalle contenant le minimum local en se basant sur la propriété d'unimodalité de la fonction objectif

Étant données les valeurs de la fonction en deux points de l'intervalle, l'unimodalité permet d'identifier une partie de l'intervalle où le minimum ne peut se retrouver

# Principes

- Choisir deux points  $x_1$  et  $x_2$  symétriques et à la même distance de chaque extrémité de l'intervalle  $[a, b]$
- Choisir le prochain point symétriquement par rapport au point déjà dans l'intervalle résultant.

# Stratégie optimale de sélection des points d'évaluation

Notation:

- $d_0 = d_1 = b - a$ : longueur de l'intervalle initial
- $d_k$ : longueur de l'intervalle après avoir utilisé  $k$  points d'évaluation

La suite  $F_k$  des nombres de Fibonacci récursivement comme:

$$F_0 = 0, \quad F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2}, \quad n \geq 2.$$



## Fibonacci et nombre d'or

Nous avons

$$\frac{F_{n+1}}{F_n} = \frac{F_n + F_{n-1}}{F_n} = 1 + \frac{F_{n-1}}{F_n}$$

En supposant que

$$\lim_{n \rightarrow \infty} \frac{F_{n+1}}{F_n} = \tau$$

avec  $\tau \in \mathbb{R}$ , nous obtenons en prenant la limite de part et d'autre de la première équation

$$\tau = 1 + \frac{1}{\tau}$$

donnant

$$\tau^2 - \tau - 1 = 0.$$

Cette dernière équation a comme racines

$$\frac{1 \pm \sqrt{5}}{2}$$

## Fibonacci et nombre d'or

Comme  $F_j > 0$  pour tout  $j \geq 1$ , nous retenons la racine positive pour obtenir

$$\lim_{n \rightarrow \infty} \frac{F_{n+1}}{F_n} = \tau = \frac{1 + \sqrt{5}}{2}$$

Il est aussi possible de montrer (formule de Binet) que

$$F_n = \frac{1}{\sqrt{5}}(\tau^n - (1 - \tau)^n)$$

de sorte que

$$\lim_{n \rightarrow \infty} \frac{F_{n+1}}{F_n} = \frac{\tau^{n+1}}{\tau^n} = \tau$$

comme  $(1 - \tau)^n \rightarrow 0$  quand  $n \rightarrow \infty$ . Il n'est plus nécessaire ici de supposer en partant que la limite existe.

Preuve de la formule de Binet: par récurrence (exercice!).

## Procédure de Fibonacci

Soit  $N$ , fixé, le nombre de fois ou l'on évaluera la fonction.

1. Initialisation:  $x_{\min} = a$ ,  $x_{\max} = b$ ,  $d_1 = x_{\max} - x_{\min}$ .
2. À chaque itération  $k$ ,  $x^G$  et  $x^D$  sont choisis symétriques à une distance  $\frac{F_{N-k-1}}{F_{N-k+1}} d_k$  de chacune des extrémités de l'intervalle  $[a, b]$ :

$$\begin{aligned}x^G &= x_{\min} + \frac{F_{N-k-1}}{F_{N-k+1}} d_k, \\x^D &= x_{\max} - \frac{F_{N-k-1}}{F_{N-k+1}} d_k = x_{\max} - \frac{F_{N-k+1} - F_{N-k}}{F_{N-k+1}} d_k \\&= x_{\max} - d_k + \frac{F_{N-k}}{F_{N-k+1}} d_k = x_{\min} + \frac{F_{N-k}}{F_{N-k+1}} d_k\end{aligned}$$

## Procédure

- 2 Une partie de l'intervalle est éliminée en se basant sur l'unimodalité de la fonction. Il en résulte un intervalle de longueur  $d_{k+1} = \frac{F_{N-k}}{F_{N-k+1}} d_k$ . Plus précisément,

$$f(x^G) < f(x^D) \Rightarrow x_{\max} = x^D$$

$$f(x^G) > f(x^D) \Rightarrow x_{\min} = x^G$$

$k := k + 1$ . Dans le cas particulier où  $f(x^G) = f(x^D)$ , on posera  $x_{\min} = x^G$  et  $x_{\max} = x^D$  et  $k := k + 2$ .

## Calcul de $x^G$ et $x^D$

Calcul de  $x_{k+1}^G$  et  $x_{k+1}^D$ . Supposons  $f(x_k^G) < f(x_k^D)$  (le cas  $f(x_k^G) > f(x_k^D)$  se traite de façon similaire).

$$\begin{aligned}x_{k+1}^D &= x_{\min,k+1} + \frac{F_{N-(k+1)}}{F_{N-(k+1)+1}}(x_{\max,k+1} - x_{\min,k+1}) \\&= x_{\min,k} + \frac{F_{N-k-1}}{F_{N-k}}(x_k^D - x_{\min,k}) \\&= x_{\min,k} + \frac{F_{N-k-1}}{F_{N-k}} \frac{F_{N-k}}{F_{N-k+1}}(x_{\max,k} - x_{\min,k}) \\&= x_{\min,k} + \frac{F_{N-k-1}}{F_{N-k+1}}(x_{\max,k} - x_{\min,k}) \\&= x_k^G.\end{aligned}$$

Par conséquent, seul  $f(x_{k+1}^G)$  doit être évaluée à l'itération  $k + 1$ .

## Note

L'algorithme place le dernier point  $N$  au centre de l'intervalle, superposé au point s'y trouvant déjà. En effet, puisqu'en utilisant cette stratégie de sélection des points d'évaluation, nous avons

$$d_k = \frac{F_{N-k+1}}{F_N} d_1, \quad k = 2, 3, \dots, N$$

il s'ensuit que

$$\begin{aligned} d_{N-1} &= \frac{F_{N-(N-1)+1}}{F_N} d_1 = \frac{F_2}{F_N} d_1 = \frac{2}{F_N} d_1 \\ d_N &= \frac{F_{N-N+1}}{F_N} d_1 = \frac{F_1}{F_N} d_1 = \frac{1}{F_N} d_1 \end{aligned}$$

et donc

$$d_N = \frac{d_{N-1}}{2}$$

Pour remédier à cette situation, le dernier point est plutôt placé à une distance  $\epsilon$  (à gauche ou à droite) de celui s'y trouvant déjà.

# Convergence

Il est possible de démontrer que la longueur finale d'intervalle obtenue,

$$d_N = \frac{1}{F_N} d_1$$

est la plus petite longueur d'intervalle qu'il est possible d'obtenir en utilisant une stratégie basée sur  $N$  points d'évaluations.

Ainsi, lorsque le nombre de points d'évaluation  $N$  devient très grand pour tendre vers l'infini, la suite des valeurs

$$\{d_k\} \rightarrow 0$$

plus rapidement qu'en utilisant toute autre stratégie

## Méthode de la section dorée

Nom issu du nombre d'or  $\tau = \frac{1+\sqrt{5}}{2} \approx 1.618 \dots$

La méthode de la section dorée utilise la même stratégie que la méthode de Fibonacci pour sélectionner les points d'évaluation, mais le nombre de points d'évaluation n'est pas fixé au départ.

Pour spécifier les deux premiers points, nous procédons comme dans la méthode de Fibonacci en les prenant symétriques à une distance  $\frac{F_{N-1}}{F_N} d_1$  de chaque extrémité en considérant que  $N \rightarrow \infty$ .

Par conséquent les deux premiers points sont choisis symétriquement à une distance  $\tau$  de chaque extrémité:

$$x_G = x_{\max} - \frac{1}{\tau} d_1$$
$$x_D = x_{\min} + \frac{1}{\tau} d_1$$



## Fonctions continues

Faisons à présent l'hypothèse suivante:  $f : [a, b] \rightarrow \mathcal{R}$ ,  $f \in C^0$  (i.e.  $f$  est continue sur  $[a, b]$ ).

(Wikipedia)

### Théorème (Théorème des bornes)

*Soient  $f$  une fonction définie sur un intervalle  $[a, b]$  de réels et à valeurs réelles. Si  $f$  est continue, alors la fonction  $f$  est bornée et atteint ses bornes, autrement dit il existe deux réels  $c$  et  $d$  de l'intervalle  $[a, b]$  tels que pour tout  $x$  de  $[a, b]$ , les inégalités suivantes soient vérifiées:*

$$f(c) \leq f(x) \leq f(d).$$

Nous cherchons  $c$ .

## Méthode de bisection (ou de bipartition)

Méthode pour identifier le 0 d'une fonction  $g(x)$  sur un intervalle  $[a, b]$ .

Supposons  $f$  continûment différentiable. Si  $g(x) = f'(x)$ , alors la méthode de bisection peut être utilisée pour identifier un point où la dérivée d'une fonction s'annule.

**Hypothèse:** sur l'intervalle  $[a, b]$ , la fonction  $g$  est continue et telle que  $g(a)g(b) < 0$ .

Ainsi, la fonction change de signe entre  $a$  et  $b$ , et par continuité, passe par 0. La méthode génère une suite d'intervalles de longueur décroissante jouissant de la même propriété.

Notons  $\overline{X} = \{x \in [a, b] \mid g(x) = 0\}$ . Nous cherchons à trouver un point  $\bar{x} \in \overline{X}$ .

# Algorithme

Principe de la méthode: à chaque itération, réduire la longueur de l'intervalle contenant  $\bar{x}$  en la divisant en deux. Soit  $L_0 = b - a$ .

Tant que  $b - a > \delta$ , répéter

**Étape 0** Soit  $a$  et  $b$  tels que  $g(a)g(b) < 0$ . Soit  $\delta$  le niveau de tolérance sur la longueur de l'intervalle contenant la racine à la fin de l'algorithme.

**Étape 1** Soit  $c = \frac{a+b}{2}$ .

- Si  $g(c) = 0$ , alors  $\bar{x} := c$ . Poser  $a := c$ ,  $b := c$ .
- Si  $g(a)g(c) > 0$ , poser  $a := c$ . Sinon, poser  $b := c$ .

## Convergence

La longueur de l'intervalle à l'itération  $k$  est dès lors

$$L_k = \frac{L_0}{2^k}$$

Le nombre  $n$  d'itérations requises pour atteindre une longueur inférieure ou égale à  $\delta$  doit satisfaire

$$\frac{L_0}{2^n} \leq \delta$$

et donc

$$n \geq \log_2 \frac{L_0}{\delta}$$

Il suffit donc de prendre

$$n = \left\lceil \log_2 \frac{L_0}{\delta} \right\rceil$$

## Vitesse de convergence

La suite des longueurs des intervalles converge vers 0.

La convergence est linéaire avec un rapport de convergence de  $\frac{1}{2}$ :

$$\lim_{k \rightarrow \infty} \frac{\frac{L_0}{2^{k+1}} - 0}{\frac{L_0}{2^k} - 0} = \frac{1}{2}$$

## Application à l'optimisation

On ne cherche pas un zéro d'une fonction en optimisation, par contre on peut appliquer l'approche sur la dérivée de la fonction.

Soit  $f : \mathcal{R} \rightarrow \mathcal{R}$ , dérivable en  $c$ :

$$f'(c) = \frac{df(c)}{dx} := \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}$$

Si  $f'(c) \neq 0$ , alors  $c$  n'est ni un minimum local ni un maximum local de  $f$ .

De plus, si  $f'(c) > 0$ , il existe un intervalle  $(c - \delta, c + \delta)$ , avec  $\delta > 0$ , tel que

$$f(x) < f(c) \quad \forall x \in (c - \delta, c)$$

$$f(c) < f(x) \quad \forall x \in (c, c + \delta)$$

## Preuve

Puisque  $f'(c) = \lim_{x \rightarrow c, x \neq c} \frac{f(x) - f(c)}{x - c} > 0$ , alors  $\exists \delta > 0$  tel que

$$\frac{f(x) - f(c)}{x - c} > 0, \quad \forall x \in (c - \delta, c + \delta), x \neq c.$$

Par conséquent, les signes de  $f(x) - f(c)$  et  $(x - c)$  doivent être identiques, d'où le résultat.



Similairement, si  $f'(c) < 0$ , il existe un intervalle  $(c - \delta, c + \delta)$ , avec  $\delta > 0$ , tel que

$$f(x) > f(c) \quad \forall x \in (c - \delta, c)$$

$$f(c) > f(x) \quad \forall x \in (c, c + \delta)$$

## Corollaire

Soit  $c \in \mathcal{R}$ .

1. si  $f'(c) > 0$ , il existe un intervalle  $(c - \delta, c + \delta)$  avec  $\delta > 0$  t.q.

$$f(x) < f(c), \quad \forall x \in (c - \delta, c)$$

$$f(x) > f(c), \quad \forall x \in (c, c + \delta)$$

et la fonction  $f$  est croissante au point  $c$ .

2. si  $f'(c) < 0$ , il existe un intervalle  $(c - \delta, c + \delta)$  avec  $\delta > 0$  t.q.

$$f(x) > f(c), \quad \forall x \in (c - \delta, c)$$

$$f(x) < f(c), \quad \forall x \in (c, c + \delta)$$

et la fonction  $f$  est décroissante au point  $c$ .

3. si  $c$  est un minimum ou un maximum local de  $f$ , alors  $f'(c) = 0$ .



## Condition nécessaire, non suffisante

Si  $f : \mathcal{R} \rightarrow \mathcal{R}$ , et  $f'(c) = 0$ ,  $c$  n'est pas nécessairement un minimum ou un maximum local.

Prenons par exemple  $f(x) = x^3$ .  $f'(0) = 0$ , mais 0 n'est ni minimum, ni maximum.

### Définition (Point stationnaire)

*$c$  est un point stationnaire d'une fonction  $f$  si  $f$  est dérivable en  $c$  et  $f'(c) = 0$ .*

Nous devons étudier sous quelles conditions il est possible de garantir qu'un point stationnaire est un extremum de la fonction.

## Condition suffisante

### Théorème

Soit  $f : \mathcal{R} \rightarrow \mathcal{R}$  deux fois dérivable en  $c$  et telle que  $f'(c) = 0$ .  
Supposons également que la dérivée  $f'(x) \forall x \in \mathcal{B}(c, \epsilon)$  existe.

- i) Si  $f''(c) < 0$  alors  $c$  est un maximum local de  $f$ .
- ii) Si  $f''(c) > 0$  alors  $c$  est un minimum local de  $f$ .

Note: ce n'est pas une condition nécessaire. Considérons par exemple  $f(x) = x^4$ . 0 est minimum local (et global) de  $f$  et  $f''(x) = 12x^2$ , d'où  $f''(0) = 0$ .

Justification intuitive de ii): si  $f''(c) > 0$ , alors  $f'$  est une fonction croissante au point  $c$  et par conséquent  $c$  est un minimum local.

# Optimisation sous contrainte

Soit  $f : S \rightarrow \mathcal{R}$ , avec  $S \subset \mathcal{R}$ .

Si  $x^*$  est à l'intérieur de  $S$ , les critères sur la dérivée s'appliquent, mais ce n'est plus le cas si  $x^*$  est sur la frontière de  $S$ .

Considérons par exemple  $f : [-1, 1] \rightarrow \mathcal{R}$ ,  $f(x) = x^3$ . Le minimum est atteint en  $x^* = -1$ , mais  $f'(-1) \neq 0$ .

Nous devons donc réviser les conditions d'optimalité!

# Convergence

On voudrait garantir que la suite d'itérés produits par un algorithme converge vers une solution optimale du problème, et ce le plus rapidement possible.

## Définition (Ordre de convergence)

Soit  $\{r_k, k = 1, 2, \dots\}$  une suite de vecteurs dans  $R^n$  convergente vers  $r^*$ . L'ordre de convergence est défini comme

$$\sup \left\{ p \mid 0 \leq \lim_{k \rightarrow \infty} \frac{\|r_{k+1} - r^*\|}{\|r_k - r^*\|^p} < \infty \right\}$$

## Ordre de convergence

Plus la valeur de  $p$  est grande, plus la convergence est rapide à la limite puisque la distance à  $r^*$  décroît plus rapidement. En effet, si

$$\lim_{k \rightarrow \infty} \frac{\|r_{k+1} - r^*\|}{\|r_k - r^*\|^p} = \beta$$

alors asymptotiquement

$$\|r_{k+1} - r^*\| = \beta \|r_k - r^*\|^p.$$

# Convergence linéaire

## Définition (Convergence linéaire)

Soit  $\{r_k\}$  une suite de vecteurs convergeant vers  $r^*$  telle que

$$\lim_{k \rightarrow \infty} \frac{\|r_{k+1} - r^*\|}{\|r_k - r^*\|} = \beta < 1$$

Alors la suite  $\{r_k\}$  converge linéairement avec un rapport de convergence  $\beta$ .

Notes:

- la convergence linéaire est aussi appelée *convergence géométrique*
- le cas  $\beta = 0$  est appelé convergence super-linéaire

## Exemples

Considérons la suite

$$\left\{ \frac{1}{k} \right\} \rightarrow 0$$
$$\lim_{k \rightarrow \infty} \frac{\frac{1}{k+1}}{\frac{1}{k}} = 1$$

Convergence d'ordre 1, mais pas linéaire.

$$\left\{ \frac{1}{k^k} \right\} \rightarrow 0$$
$$\lim_{k \rightarrow \infty} \frac{\frac{1}{(k+1)^{k+1}}}{\frac{1}{k^k}} = \lim_{k \rightarrow \infty} \left( \frac{k}{k+1} \right)^k \frac{1}{k+1} = 0$$

Convergence super-linéaire.