Calcul Scientifique

Cours 7: Introduction à l'optimisation

Alexis Lechervy

Sommaire

L'optimisation (mathématique)

Principe

L'optimisation est une branche des mathématiques qui cherchent à trouver les maximums et/ou les minimums d'une fonction sur un ensemble.
C'est un domaine qui joue un rôle important dans nombreux domaines notamment en informatique, en économie, dans l'industrie, en théorie des jeux, en recherche de stratégie...

Les origines

- Dès l'antiquité des problèmes d'optimisation sont apparu pour résoudre des problèmes d'optique. En trouve des traces chez Euclide (~ -300) et chez Héron d'Alexandrie (1er siècle) étudiant le principe de "plus court chemin" pour des rayons lumineux.
- Au 17ème les travaux de Newton sur la recherche d'extremum et ceux de Leibniz sur la dérivée posent un problème cadre formel.
- Au 18ème Euler et Lagrange ménent des travaux sur le calcul de variations et introduisent de nouvelles méthodes de résolutions.
- Le 19ème et 20ème vont permettre de grand avancées poussés par les besoins en économie, en logistique, en planification, en informatique...

Des exemples d'applications

Exemples

- Combien d'élément faut-il produire pour maximiser ses profits connaissant le coût de production, le prix de vente et les besoins du marché ? Comment faire des choix d'investissement ? Optimiser le rendement d'un appareil ?
- Comment répartir des éléments en fonction des préférences de chacun ?
- Comment construire un emploi du temps en minimisant les trous ou les horaires tardifs ?
- Trouvez le plus court chemin sur un réseau informatique ou pour un GPS.
- Optimiser des lignes ferroviaires. Faire du contrôle aérien.
- Définir la forme d'un objet. Exemple de la cannette qui minimise la quantité d'aluminium pour un volume donnée.
- Modéliser une réaction chimique.
- Faire de la reconnaissance de visage.
- C'est l'outil de base des méthodes "d'intelligence artificiel"...

$\Longrightarrow$ L'optimisation de ces systèmes permet de trouver une configuration idéale, d'obtenir un gain d'effort, de temps, d'argent, d'énergie, de matière première, de satisfaction...

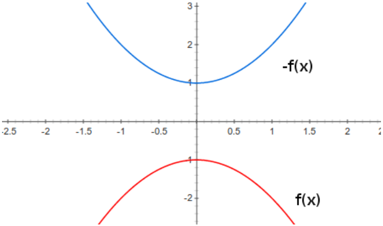
Problématique

Formulation

Un problème d'optimisation se formalise ainsi :
Soit une fonction $f : A \longrightarrow \mathbb{R}$, on cherche $x^\star$ tel que $\forall x \quad f(x^\star) \leq f(x)$.
Cela revient à minimiser la fonction f sur A. On note aussi $x^\star = \arg\min_x f(x)$.

Maximisation de fonction

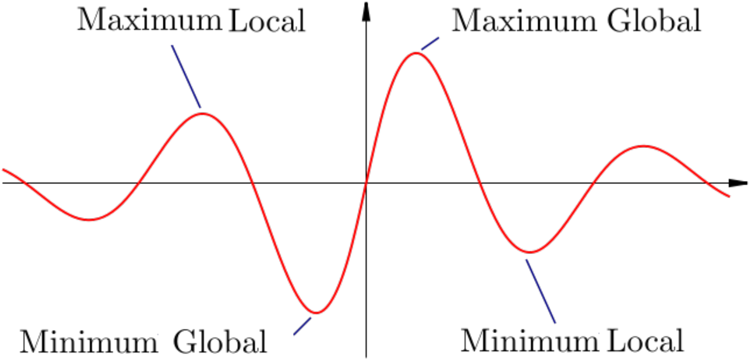
Chercher le x qui maximise une fonction revient à trouver le x qui minimise la fonction $-f(x)$.



En pratique : recherche d'extrema locaux

Optimisation en pratique

Il est généralement difficile de trouver les minimums globaux d'une fonction. En général, on se contentera de trouver un minimum local de la fonction.

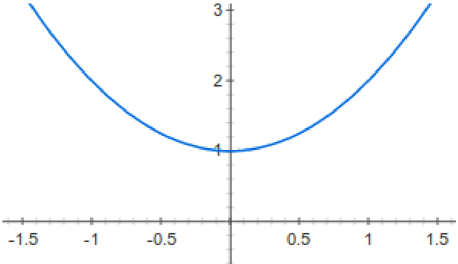


Sommaire

La méthode des trois points

Principe

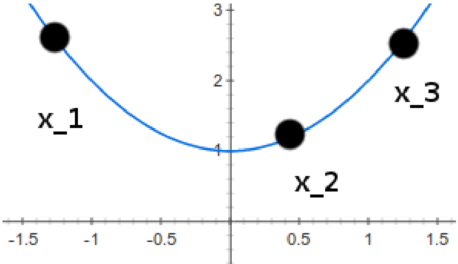
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

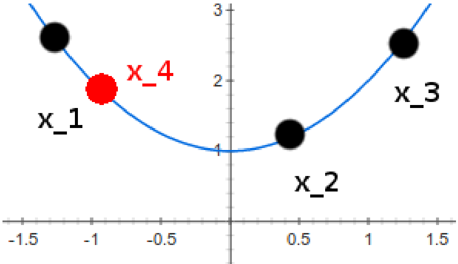
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

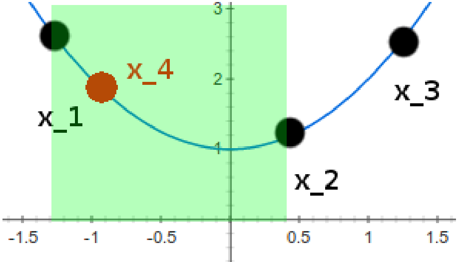
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

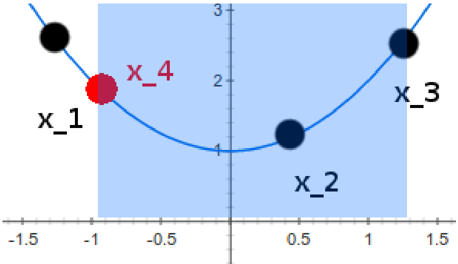
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

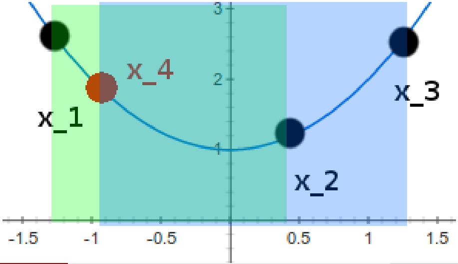
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

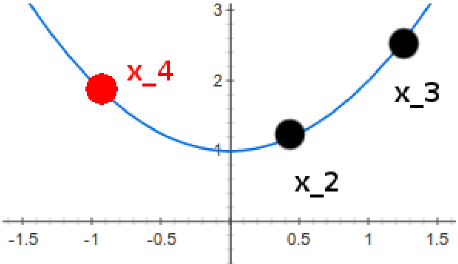
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

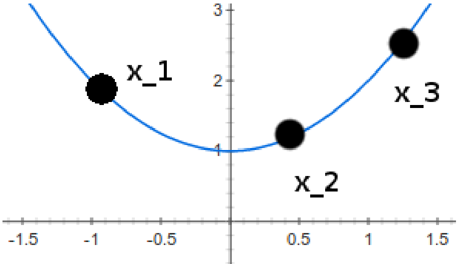
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

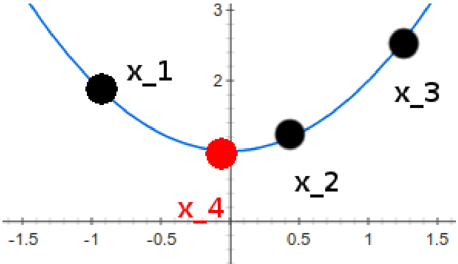
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

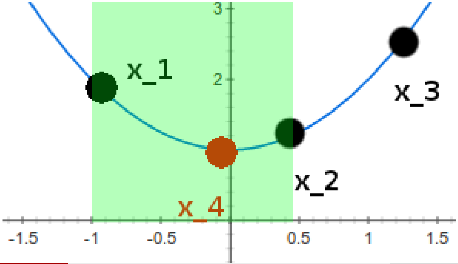
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

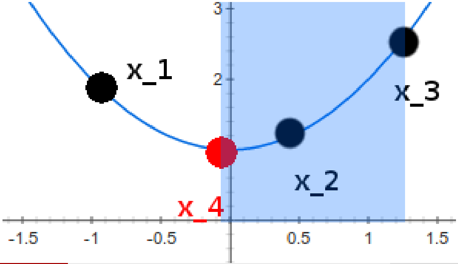
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

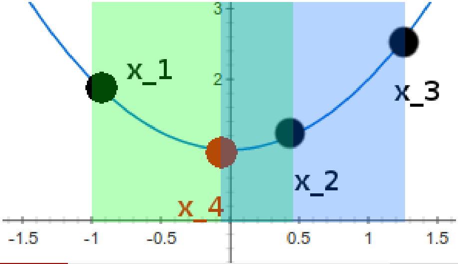
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

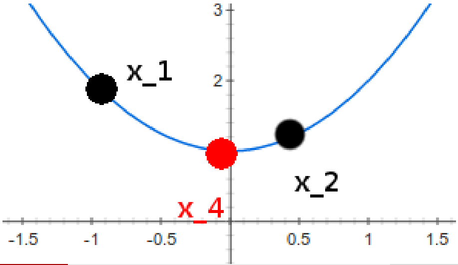
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

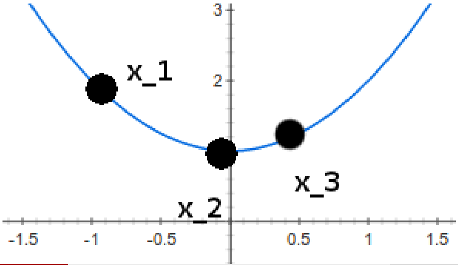
La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode des trois points

Principe

La méthode des trois points permet de trouver un minimum local de la fonction. Elle est assez similaire à une recherche par dichotomie (cf le cours sur la recherche des zéros d'une fonction). Elle nécessite de partir de trois points dont les points extrême encadrent la solution que l'on recherche.
On choisi trois points $x_1 < x_2 < x_3$ définissant deux ensembles disjoints $[x_1, x_2]$ et $[x_2, x_3]$, on prend un point x_4 dans un des deux ensembles (par exemple le premier). On regarde si $f(x_4) > f(x_2)$ si c'est le cas on continue avec les points $x_4 < x_2 < x_3$ sinon on prend les points $x_1 < x_4 < x_2$.



La méthode du nombre d'or

Principe

La méthode du nombre d'or est un cas particulier de la méthode précédente. Elle précise comment choisir le point x_4 pour avoir le moins d'étape en général. L'idée est de faire en sorte de garder toujours la même proportion entre le segment à une itération et la suivante quelque soit l'ensemble que l'on choisi.

Équation de calcul de x_4

$x_4 = x_1 + \dfrac{x_3 - x_1}{1+\varphi}$ avec le nombre d'or $\varphi = \frac{1+\sqrt{5}}{2}$.

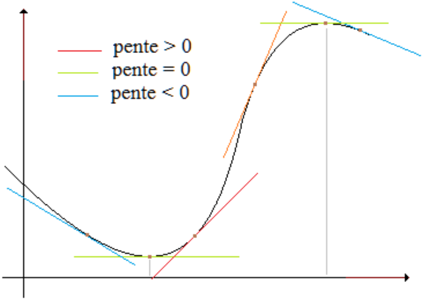
Implémentation scipy

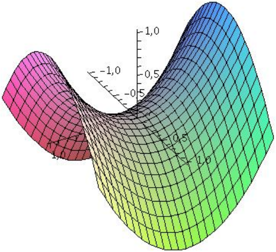
```
def f(x) :
    return x**2
scipy.optimize.minimize_scalar(f,method='golden')
```

Méthode par annulation du gradient

Constat

Les minimums ont leurs dérivées nulle (en 2D), leurs gradients nulles (dans le cas générale). En effet pour un minimum la tangente est horizontale, la courbe ne descend plus.





Attention

La dérivée est nulle pas uniquement au minimum. Par exemple la dérivée et le gradient sont nulle pour les maximums et le gradient est nulle pour les points selles.

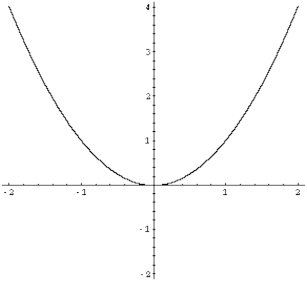
Solution analytique directe

Recherche d'une solution par les équations

Parfois on dispose de l'équation correspondant à notre fonction et on peut en déduire la dérivée. Il est alors parfois possible de calculer analytiquement les zéros de la dérivée. Les minimums de notre fonction se trouve nécessairement parmi ces valeurs.

Exemple

- La fonction : $f(x) = x^2$.
- La dérivée : $f'(x) = 2x$.
- Les zéros de la dérivée : $f'(x) = 0 \Longrightarrow x = 0$.
- $x = 0$ est bien le minimum de f.



$y = x^2$

Recherche des zéros de la dérivée

Principe

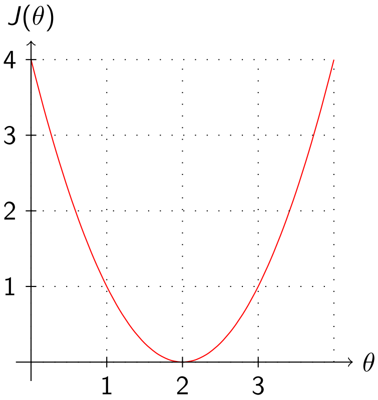
Si l'on dispose d'une méthode pour calculer la dérivée (2D) ou le gradient (en général), on peut rechercher ses zéros par les méthodes vu en cours. Par exemple avec les méthodes :

- par dichotomie,
- de la sécante,
- de Newton...

Descente de gradient en 2D

Idée

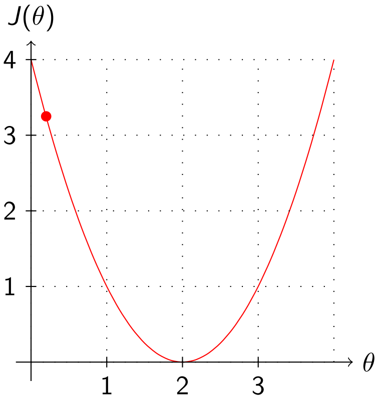
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Descente de gradient en 2D

Idée

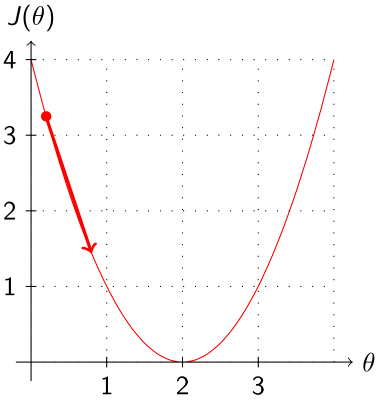
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Descente de gradient en 2D

Idée

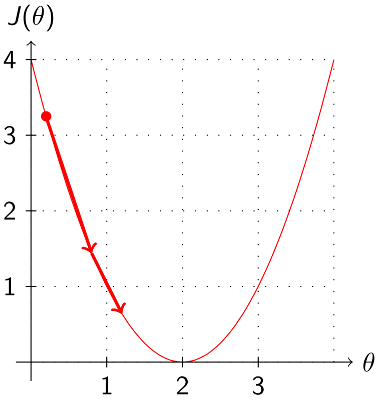
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Descente de gradient en 2D

Idée

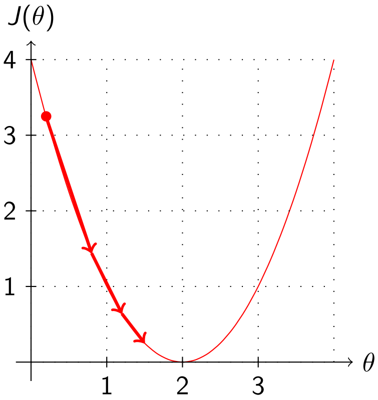
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Descente de gradient en 2D

Idée

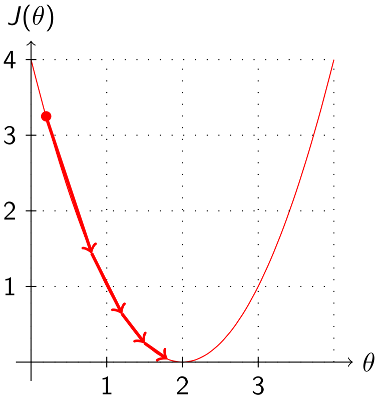
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Descente de gradient en 2D

Idée

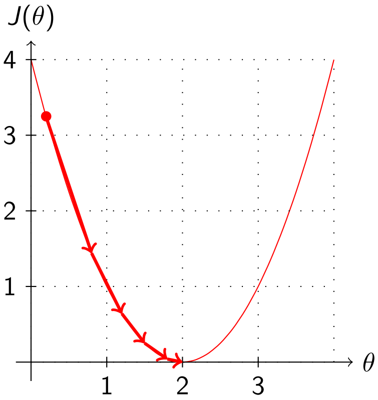
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



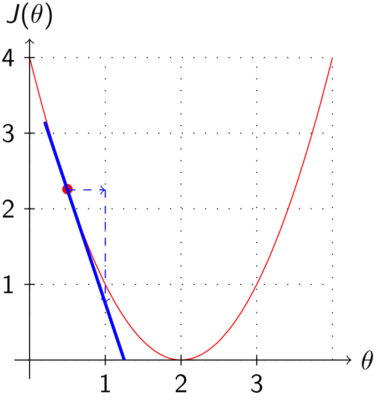
Descente de gradient en 2D

Idée

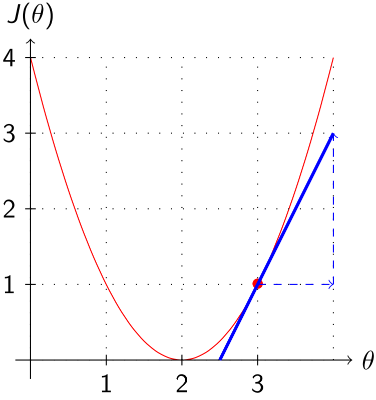
On peut atteindre le minimum en se dirigeant itérativement dans la direction de plus forte pente.



Direction de la pente.

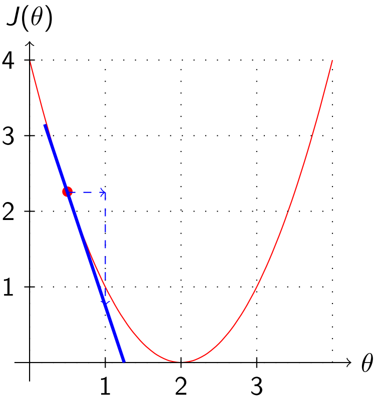


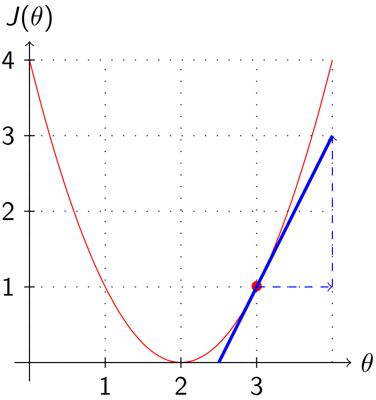
Valeur de la dérivée en 0.5 : -3
Direction de la pente : positive.



Valeur de la dérivée en 3 : 2
Direction de la pente : négative.

Direction de la pente.





Comment trouver la direction de la pente ?

Le signe opposé de la dérivée nous donnes la direction de la pente.

Résolution par descente de gradient en 2D

Problème d'optimisation à résoudre

$$\arg\min_{\theta} J(\theta).$$

Solution par descente de gradient

Initialiser avec un θ choisi aléatoirement.
Répéter jusqu'à convergence :

$$\theta := \theta - \alpha \frac{\partial}{\partial \theta} J(\theta).$$

α est une constante correspondant au taux d'apprentissage.
Lorsque l'on s'approche du minimum la dérivée tend vers 0, en conséquence les pas sont de plus en plus petit. Il n'est pas nécessaire d'avoir une valeur de α qui décroit.

Résolution par descente de gradient dans le cas général

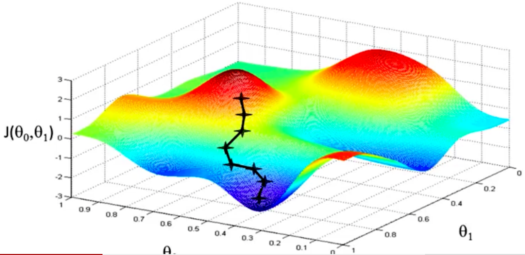
Problème d'optimisation à résoudre

$$\arg\min_{\theta} J(\theta).$$

Solution par descente de gradient

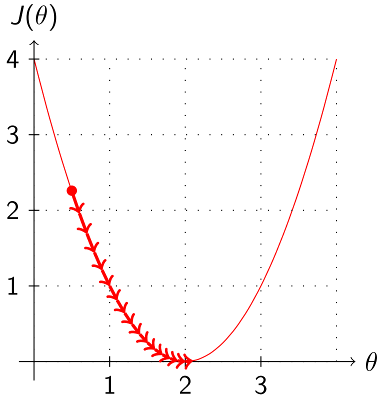
Initialiser avec un θ choisi aléatoirement.
Répéter jusqu'à convergence :

$$\theta := \theta - \alpha \nabla_{\theta} J(\theta).$$



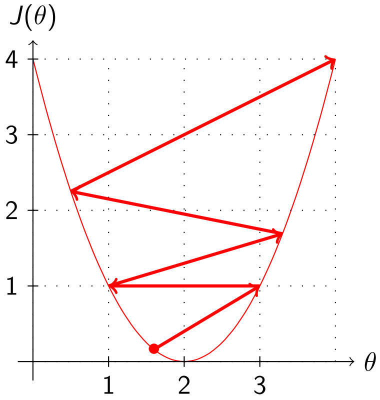
Bien choisir la valeur de α

Valeur de α trop petite



La convergence est très lente.

Valeur de α trop grande

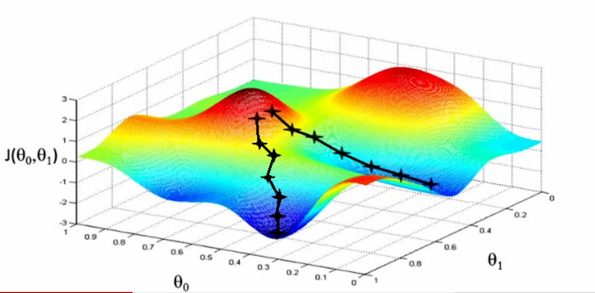


L'algorithme diverge.

Attention au point de départ

Remarques

- Si le minimum global est l'unique minimum locale, l'algorithme pourra le trouver.
- Si le point de départ est un maximum locale, le gradient sera nul et par conséquence l'algorithme ne bougera pas.
- S'il existe plusieurs minimum locaux, l'algorithme ira vers l'un d'entre eux. Des petites variations dans les conditions initiales peuvent conduire à des résultats différents.



Implémentation avec scipy

Implémentation avec scipy en utilisant la méthode de Newton

```
def f(x) :
    return x**2
def f2(x) :
    return 2*x
res = scipy.optimize.minimize(f,jac=f2,x0=1,method='BFGS').x
```

Méthode de Newton en estimant le gradient

```
def f(x) :
    return x**2
res = scipy.optimize.minimize(f,jac=False,x0=1,method='BFGS').x
```

Remarque

La méthode BFGS va choisir par elle même le taux d'apprentissage α.
Contrairement à ce qui a été présenté, le pas est variable dans la méthode 'BFGS'.
Il est choisi pour converger rapidement vers la solution.

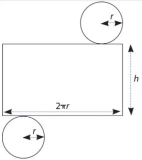
Exemple applicatif : dimension d'une canette

Construire une cannette à moindre coût

On cherche à construire une cannette de 33cl ($330cm^3$) dont la surface est la plus petite possible

Modélisation du problème

- Volume d'une canette : $\pi r^2 h = 330$.
- Aire d'une canette : $2\pi(hr + r^2)$.
- On a $h = \dfrac{330}{\pi r^2}$.
- Problème d'optimisation : $\arg\min_r 2\pi \left(\dfrac{330}{\pi r} + r^2 \right)$



Solution scipy

Code

```
import scipy as sc
import scipy.optimize
import numpy as np
r = sc.optimize.minimize(lambda x :(330/(np.pi*x)+x**2),
    x0=3,jac=False,method='BFGS').x[0]
h = 330/(np.pi*r**2)
```

Résultat

$r = 3.744938536896202$ et $h = 7.489876876504729$