

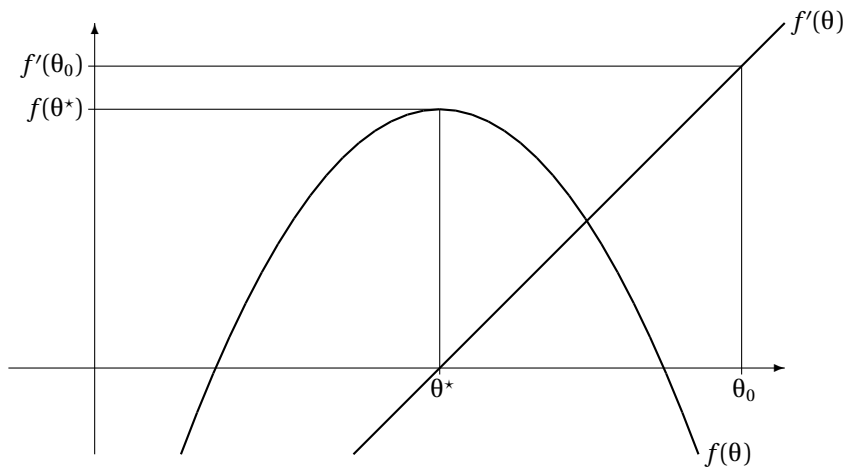
Fiche de travail n°6 Optimisation numérique

Je vous propose dans cette sixième fiche de travail de prendre la mesure de la complexité de la résolution numérique d'un problème de maximisation non quadratique.

1 L'algorithme de NEWTON dans le cas univarié

On part d'une fonction $y = f(\theta)$, d'une seule variable, que l'on suppose bien élevée – elle est définie partout, continue et dérivable. Elle admet un maximum global en θ^* .

FIG. 1 – Le cas quadratique

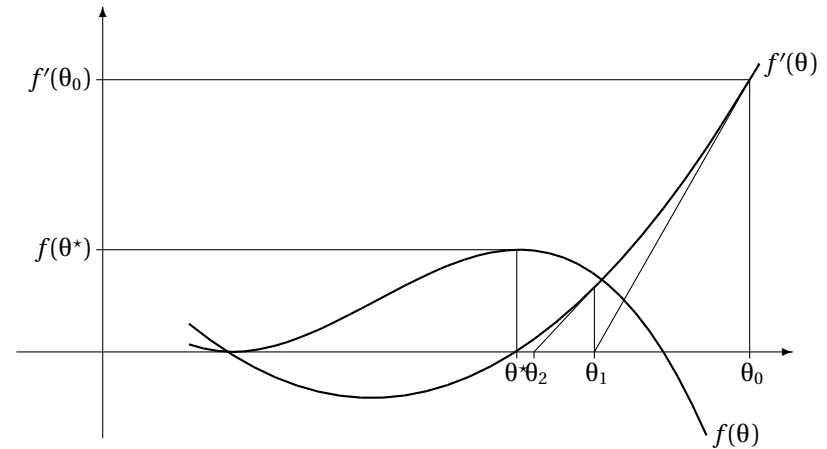


$$\theta^* = \theta_0 - \frac{f'(\theta_0)}{f''(\theta_0)}$$

Il est d'abord possible de rechercher, analytiquement, θ^* . La condition du premier ordre est :

$$f'(\theta) = 0$$

FIG. 2 – Un cas peu quadratique



$$\theta_1 = \theta_0 - \frac{f'(\theta_0)}{f''(\theta_0)} \quad \text{et} \quad \theta_2 = \theta_1 - \frac{f'(\theta_1)}{f''(\theta_1)}$$

Il faut donc résoudre cette équation pour trouver θ^* . Les ordinateurs sont beaucoup plus capables de calculer que d'effectuer des manipulations symboliques. Quand l'équation $f'(\theta) = 0$ n'est pas linéaire, il est plus efficace, en général, d'utiliser un algorithme de résolution numérique pour trouver θ^* .

L'algorithme de NEWTON repose sur une valeur initiale de θ , θ_0 , et sur l'approximation linéaire de $f'(\cdot)$ au voisinage de θ_0 :

$$f'(\theta) \approx f'(\theta_0) + (\theta - \theta_0)f''(\theta_0)$$

Pour θ^* , on trouve ainsi

$$\theta^* \approx \theta_0 - \frac{f'(\theta_0)}{f''(\theta_0)}$$

L'algorithme de NEWTON est itératif. Formellement, il s'exprime comme une suite définie de la manière suivante :

$$\theta_j = \theta_{j-1} - \frac{f'(\theta_{j-1})}{f''(\theta_{j-1})} \quad j = 1, 2, \dots \quad \text{et} \quad \theta_0 \text{ fixé}$$

La solution est atteinte pour un point fixe de cette suite.

Au cas quadratique, la solution est obtenue dès la première itération de l'algorithme. La fonction est de la forme $a\theta^2 + b\theta + c$; on a $f'(\theta) = 2a\theta + b$ et $f''(\theta) = 2a$. Aussi θ^* est-il égal à

$$\theta^* = \theta_0 - \frac{f'(\theta_0)}{f''(\theta_0)} = \theta_0 - \frac{2a\theta_0 + b}{2a} = -\frac{b}{2a}$$

Le cas quadratique est illustré à la figure 1. On en retire la conjecture que l'algorithme est d'autant plus efficace que la fonction est (localement) quadratique. Quand la fonction n'est pas quadratique, un plus grand nombre d'itérations est nécessaire. À la figure 2, j'ai cherché à illustrer ce cas là en portant les deux premières itérations de l'algorithme.

Je vous demande de trouver la solution du problème

$$\max_{\theta} (\theta - 2)(\theta - 6)$$

i) analytiquement ; ii) en utilisant l'algorithme de NEWTON et en prenant $\theta_0 = 1$.

Je vous demande aussi de calculer les deux premières itérations de l'algorithme de NEWTON, en prenant $\theta_0 = 1$, du problème suivant

$$\max_{\theta} 2\ln(\theta) - \theta$$

et de comparer les valeurs obtenues à la solution calculée analytiquement.

2 L'algorithme de NEWTON dans le cas multivarié

L'algorithme de NEWTON se généralise sans trop de difficultés dans le cas où la fonction dépend d'un vecteur de variables que l'on peut noter $\underline{\theta}$ de taille $K \times 1$. Au cas quadratique, la fonction serait :

$$f(\underline{\theta}) = \underline{\theta}' A \underline{\theta} + \underline{b}' \underline{\theta} + c$$

où A est une matrice (symétrique) de taille $K \times K$, \underline{b} est un vecteur de taille $K \times 1$ et c un nombre.

La suite de l'algorithme de NEWTON est maintenant définie comme suit :

$$\underline{\theta}_j = \underline{\theta}_{j-1} - [H(\underline{\theta}_{j-1})]^{-1} \underline{g}(\underline{\theta}_{j-1}) \quad j = 1, 2, \dots \quad \text{et } \underline{\theta}_0 \text{ fixé}$$

où $H(\underline{\theta})$ est la matrice hessienne de la fonction $f(\cdot)$ évaluée au point $\underline{\theta}$ et $\underline{g}(\underline{\theta})$ est le vecteur gradient de la fonction $f(\cdot)$ évaluée au point $\underline{\theta}$. Ces deux grandeurs, au cas quadratique, sont respectivement :

$$H(\underline{\theta}) = 2A \quad \text{et} \quad \underline{g}(\underline{\theta}) = 2A\underline{\theta} + \underline{b}$$

L'estimateur des MCO peut se mettre sous une forme qui s'interprète comme la première itération de l'algorithme de NEWTON en prenant comme point de départ $\underline{0}$:

$$\hat{\underline{a}} = \underline{0} + (X'X)^{-1} X'y$$

En effet, le critère des MCO (en prenant son opposé pour maximiser l'opposé de la grandeur à minimiser) s'écrit :

$$-\underline{a}' X' X \underline{a} + 2(X'y)' \underline{a} - y'y$$

Je vous demande, en prenant $\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, de trouver la solution du problème suivant

$$\max_{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}' \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} 4 \\ 6 \end{pmatrix}' \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

i) analytiquement ; ii) en utilisant l'algorithme de NEWTON et en prenant $\theta_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.