Gradient à pas optimal : détermination du pas à l'aide de la méthode de Newton

Considérons la fonction f définie pour tout $v = (v_1, v_2) \in \mathbb{R}^2$ par

$$f(v) = (v_1 - 4)^2 + 2(v_2 - 3)^2 + v_1 v_2$$

On rappelle que dans la méthode de gradient à pas optimal, le pas ρ_n change à chaque itération :

$$v_{n+1} = v_n - \rho_n \nabla f(v_n) \quad \text{où} \quad \rho_n = \arg\min_{\rho>0} f(v_n - \rho \nabla f(v_n))$$

Posons $g_n(\rho) = f(v_n - \rho \nabla f(v_n))$, définie de $\mathbb{R}$ dans $\mathbb{R}$. À chaque itération, pour déterminer le pas optimal ρ_n, nous cherchons la solution du problème (possiblement non linéaire) suivant :

$$g_n'(\rho) = 0$$

Montrons que, pour tout $n \geq 0$, la dérivée de g_n est

$$\forall \rho \in \mathbb{R}, \quad g_n'(\rho) = -\nabla f(v_n) \cdot \nabla f(v_n - \rho \nabla f(v_n))$$

Afin de déterminer une valeur approchée de la solution $g_n'(\rho) = 0$, nous allons considérer une variante de la méthode de Newton : dans la formule de Newton donnée par

$$\rho_{k+1} = \rho_k - \frac{g_n'(\rho_k)}{g_n''(\rho_k)},$$

pour ne pas avoir à calculer g_n'', on remplace $g_n''(\rho_k)$ par le taux d'accroissement :

$$\frac{g_n'(\rho_k + \delta) - g_n'(\rho_k)}{\delta},$$

avec $\delta = 10^{-8}$.

Ainsi, à chaque étape n de l'algorithme du gradient, on construit la suite

$$\begin{cases} \rho_0 \in \mathbb{R} \text{ donné} \\ \rho_{k+1} = \rho_k - \delta \frac{g_n'(\rho_k)}{g_n'(\rho_k + \delta) - g_n'(\rho_k)} \end{cases}$$

qui converge vers ρ_n, la solution du problème $\rho_n = \arg\min_{\rho>0} f(v_n - \rho \nabla f(v_n))$.