

TP 1 - Gradient, Levenberg-Marquardt, Nelder-Mead

Jules Kozolinsky

1 Gradient et règle d'Armijo pour la recherche linéaire approchée

1.1 Fonction de Rosenbrock

Lignes de niveaux de la fonction de Rosenbrock

On remarque sur le graphique 5 que les lignes de niveaux ont une forme de crevasse. Il peut donc sembler difficile de trouver le minimum global.

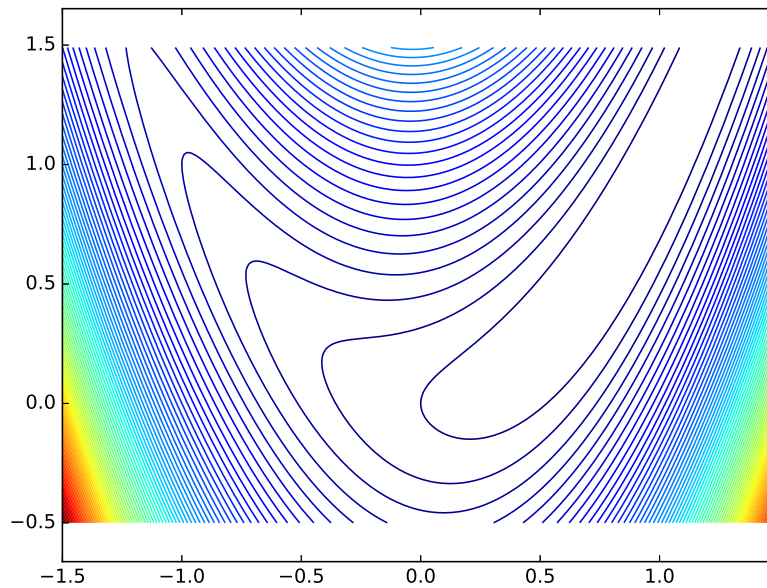


FIGURE 1 – Lignes de niveaux de la fonction de Rosenbrock

Calcul de $\nabla J(\mathbf{u})$

On a $J(\mathbf{u}) = (1 - u_1)^2 + p(u_2 - u_1^2)^2 = 1 - 2u_1 + u_1^2 + pu_2^2 - 2pu_1^2u_2 + pu_1^4$.

D'où :

- $\frac{\partial J}{\partial u_1} = -2 + 2u_1 - 4pu_1u_2 + 4pu_1^3 = 2(u_1 - 1) + 4pu_1(u_1^2 - u_2)$
- $\frac{\partial J}{\partial u_2} = 2p(u_2 - u_1^2)$

On remarque que $\nabla J(\mathbf{u}^*) = \mathbf{0}$.

Étude numérique de la suite $(u^k)_k$

On conjecture une convergence de la suite $(u^k)_k$ vers le minimum global u^* qui s'effectue très lentement (les itérés $(u^k)_k$ sont en effet très proches). On remarque que les itérés sont orthogonaux aux lignes de niveaux, ce qui s'explique par $u^{k+1} - u^k = \rho \frac{\nabla J(u^k)}{\|\nabla J(u^k)\|}$.

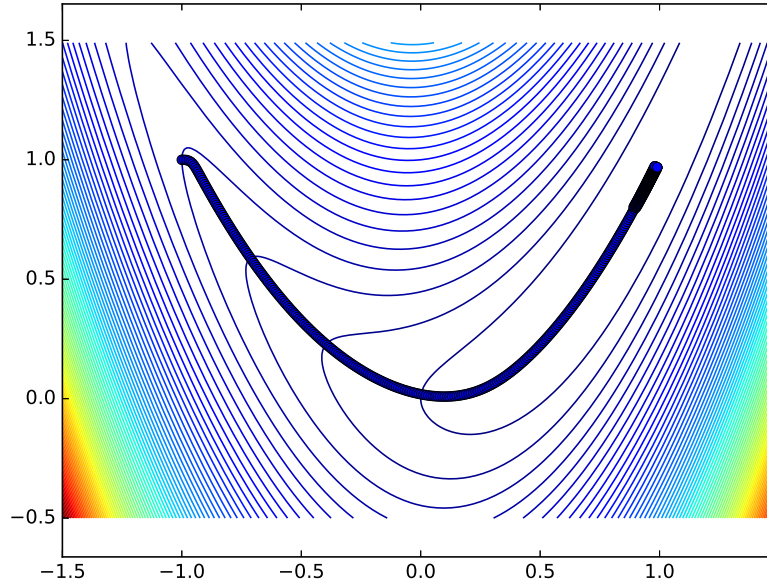


FIGURE 2 – Itérés $(u^k)_{k \leq 500}$ sur les lignes de niveaux de J

Critère de convergence de la suite $(u^k)_k$

Soit n , trouver k tel que $\frac{\|\nabla J(u^k)\|}{\|\nabla J(u^0)\|} \leq 10^{-n}$.

Pour $n = 4$, l'algorithme ne termine pas en un temps raisonnable.

Pour $n = 1$, on trouve $k = 229$.

1.2 Règle d'Armijo

On observe une fois de plus la convergence de la suite u mais cette fois-ci de manière bien plus rapide (surtout lors des premières itérations), ce qui est confirmé par le faible nombre d'itérations nécessaires pour atteindre le critère de convergence avec $n = 4$ (en reprenant les notations précédentes). On trouve $k = 771$.

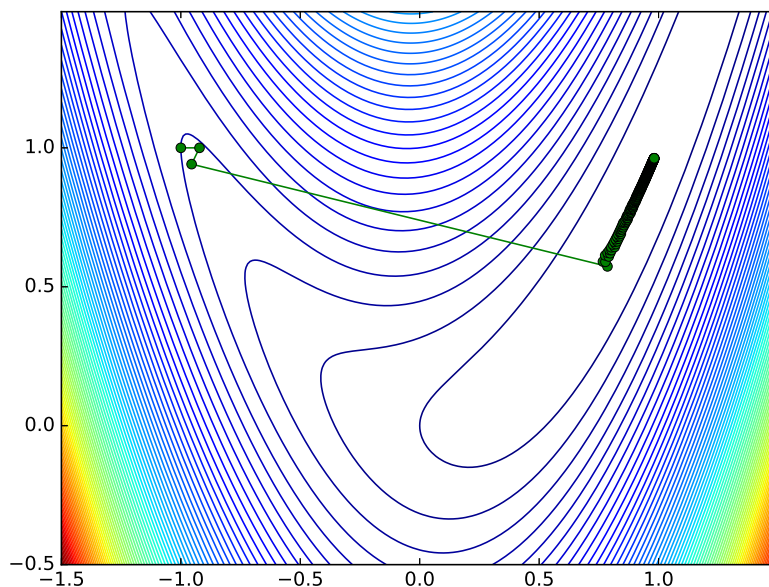


FIGURE 3 – Itérés $(u^k)_{k \leq 500}$ de la règle d'Armijo sur les lignes de niveaux de J

2 Régression non linéaire, algorithme de Levenberg-Marquardt

Calcul de S^k

Pour calculer S^k , on a besoin de calculer $\nabla \varphi_i(\mathbf{u})$ et donc $\frac{\partial f}{\partial \mathbf{u}}$.
On a :

- $\frac{\partial f}{\partial \alpha_j} = \exp\left(-\frac{1}{2} \frac{(x - x_j^0)^2}{\sigma_j^2}\right)$
- $\frac{\partial f}{\partial \sigma_j} = \alpha_j \frac{(x - x_j^0)^2}{\sigma_j^3} \exp\left(-\frac{1}{2} \frac{(x - x_j^0)^2}{\sigma_j^2}\right)$
- $\frac{\partial f}{\partial x_j^0} = \alpha_j \frac{x - x_j^0}{\sigma_j^2} \exp\left(-\frac{1}{2} \frac{(x - x_j^0)^2}{\sigma_j^2}\right)$

Pour la règle d'Armijo, on considère la direction de recherche \mathbf{d}^k pour atteindre la condition d'arrêt :
 $J(\mathbf{u}^k - \rho^k \mathbf{d}_k) > J(\mathbf{u}^k) + m\rho^k \langle \nabla J(\mathbf{u}^k), \mathbf{d}_k \rangle$

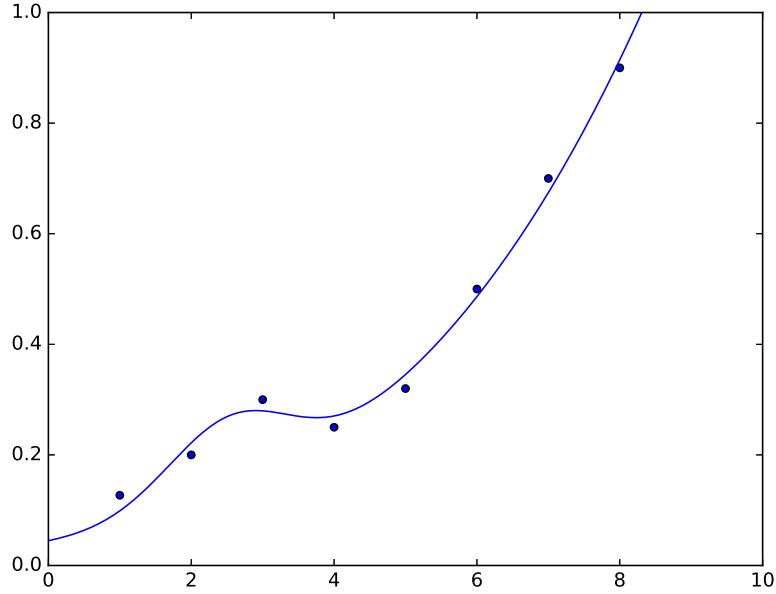


FIGURE 4 – Points de données $\{(x_i, y_i)\}_i$ et la fonction d'approximation correspondante

L'algorithme converge en 31 itérations avec $\frac{\|\nabla J(\mathbf{u}^k)\|}{\|\nabla J(\mathbf{u}^0)\|} \leq 10^{-4}$. On remarque de plus que la fonction passe près des points.

3 Méthode sans gradient. Algorithme de Nelder-Mead

Le nombre d'évaluation nécessaire pour atteindre une précision de $\frac{\|\nabla J(\mathbf{u}^k)\|}{\|\nabla J(\mathbf{u}^0)\|} \leq 10^{-4e}$ est de 46 itérations. On peut d'ailleurs remarquer que la condition d'arrêt semble absurde car elle requiert de calculer le gradient de la fonction J alors que tout l'intérêt de cette méthode est de s'en passer, on pourrait utiliser le test classique $\|u^{k+1} - u^k\| \leq 10^{-4}$.

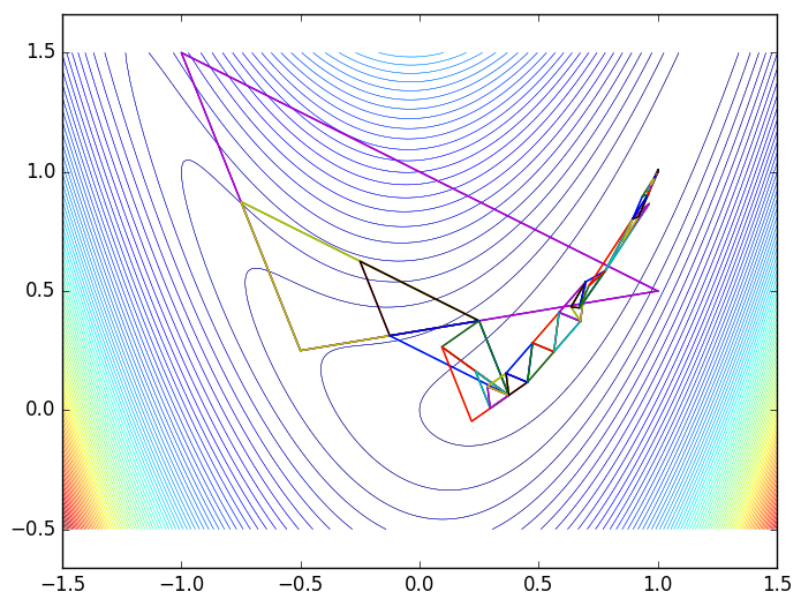


FIGURE 5 – Points de données $\{(x_i, y_i)\}_i$ et la fonction d'approximation correspondante