Feuille de TP 1 – Algorithme de descente de gradient

1 Algorithme à pas fixe

1.1 En dimension 1

Exercice 1.1. On veut minimiser la fonction suivante :

$$f(x) = (x-1)(x-2)(x-3)(x-5).$$

1. Tracer cette fonction sur l'intervalle $[0, 6]$.
2. Écrire l'algorithme de descente du gradient à pas fixe en R. On arrêtera l'algorithme quand l'écart entre deux valeurs successives est inférieur à 0,001.
3. Si x_0 désigne le point de départ de l'algorithme et η le pas (fixe), tester avec les valeurs suivantes :
   - $x_0 = 5$, $\eta = 0,001$.
   - $x_0 = 5$, $\eta = 0,01$.
   - $x_0 = 5$, $\eta = 0,1$.
   - $x_0 = 5$, $\eta = 0,17$.
   - $x_0 = 5$, $\eta = 1$.
   - $x_0 = 0$, $\eta = 0,01$.

   On fera tourner un compteur calculant le nombre d'itérations avant l'arrêt de la boucle. Que remarque-t-on ?

1.2 En dimension 2

Exercice 1.2. On considère d'abord la fonction

$$f(x, y) = x^2 + 0,25xy + y^2.$$

Observez l'algorithme de descente de gradient quand le point initial est $(x_0 = -0,9, y_0 = 1)$ et quand il vaut $(x_0 = -1, y_0 = 1)$ avec un pas $\eta = 0,01$.

Exercice 1.3. La fonction de *Rosenbrock* est la suivante :

$$f(x, y) = \frac{1}{100}(1-x)^2 + (y - x^2)^2.$$

Elle est connue pour être difficile à minimiser. Son minimum vaut 0 et est atteint au point $(1, 1)$. Programmer l'algorithme de descente de gradient à partir du point initial est $(x_0 = -1, y_0 = 0)$. Qu'observe-t-on ?

1.3 En dimension supérieure

On considère la matrice

$$A = \begin{pmatrix} 3 & -1 & 0 & 0 & 0 \\ -1 & 12 & -1 & 0 & 0 \\ 0 & -1 & 24 & -1 & 0 \\ 0 & 0 & -1 & 48 & -1 \\ 0 & 0 & 0 & -1 & 96 \end{pmatrix}$$

et le vecteur

$$b = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}.$$

Sur $\mathbb{R}^5$, on veut minimiser la fonction

$$J(x) = \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle.$$

Exercice 1.4.

1. Vérifier que A est symétrique définie positive. On calculera ses valeurs propres via la fonction `eigen`. Donner également ρ_1 et ρ_5 la plus petite et la plus grande valeur propre.

J admet donc un unique minimum sur $\mathbb{R}^5$. L'algorithme du gradient à pas fixe défini par

$$x_{k+1} = x_k - \eta d_k, \quad d_k = Ax_k - b,$$

converge vers x pour toute initialisation x_0 si η est choisi tel que $0 < \eta < 2\rho_1/(\rho_5)^2$ (cf. cours). Pour ce problème, on peut montrer qu'il suffit de prendre $\rho < 2/\rho_5$ et que la vitesse maximale de convergence est obtenue pour le choix $\eta = 2/(\rho_1 + \rho_5)$.

2. L'implémenter sous R. On le testera avec $x_0 = (0, 0, 0, 0, 0)$ et $\eta = 1/(\rho_1 + \rho_5)$ et $\eta = 2/(\rho_1 + \rho_5)$. On arrêtera l'algorithme quand l'écart entre deux valeurs successives est inférieur à 10^{-5}.

2 Algorithme à pas variable

Exercice 2.1. On reprend la fonction de l'exercice 1.2.

1. Montrer que cette fonction se met sous la forme

$$f(x, y) = \frac{1}{2}\langle (x, y), M \begin{pmatrix} x \\ y \end{pmatrix}\rangle$$

où M est une matrice de dimension 2×2 à expliciter.

2. On considère l'algorithme de gradient qui suit :

$$x_{k+1} = x_k - \eta_k d_k, \quad d_k = Mx_k, \quad \eta_k = \frac{\langle d_k, d_k\rangle}{\langle d_k, Md_k\rangle}.$$

Implémenter cet algorithme en initialisant $x_0 = (-1, 1)$, et comparer avec les résultats obtenus dans l'exercice 1.2.

Exercice 2.2. On considère le même problème que dans l'exercice 1.4. Mais au lieu de prendre un pas η, on va l'optimiser en choisissant :

$$x_{k+1} = x_k - \eta_k d_k, \quad d_k = Ax_k - b, \quad \eta_k = \frac{\langle d_k, d_k\rangle}{\langle d_k, Ad_k\rangle}.$$

Implémenter cet algorithme et comparer avec les résultats obtenus dans l'exercice 1.4.

