Méthode des gradients conjugués

Initialisation : choisir un point de départ x_0 et une direction initiale $d_0 = -\nabla f(x_0)$.

Pour k variant de 0 à n, faire:

- Choisir s_k minimisant la fonction $f(x_k + sd_k)$ par rapport à s.
-
$$x_{k+1} \longleftarrow x_k + s_k d_k$$
-
$$b_k \longleftarrow \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}$$
-
$$d_{k+1} \longleftarrow -\nabla f(x_{k+1}) + b_k d_k.$$

Cette méthode à deux avantage : elle necessite le stockage de trés peu d'informations et sa vitesse de convergence est trés supérieure à celle de la méthode du gradient. Le fait qu'elle converge au plus n itérations est bien un atout important.

Exercise 1 *On considère le problème d'optimisation sans contraintes :*

$$\text{minimiser } f(x_1, x_2) = 3x_1^2 - 2x_1x_2 + x_2^2 - 8x_2.$$

Partant de $x^0 = (0,0)$, déterminons la solution optimale du problème en utilisant la méthode du gradient conjugué. Le gradient de f est :

$$\nabla f(x) = \nabla f(x_1, x_2) = \begin{pmatrix} 6x_1 - 2x_2 \\ -2x_1 + 2x_2 - 8 \end{pmatrix}.$$

- *Pour l'instant $k = 0$. Ainsi,*

$$d_0 = -\nabla f(x^0) = -\begin{pmatrix} 0 \\ -8 \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \end{pmatrix}.$$

Ainsi,

$$x_0 + sd_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 8s \end{pmatrix}.$$

- *Cherchons* s_0 *minimisant la fonction* $g(s) = f(x_0 + sd_0)$ *par rapport à* s. *Pour cela, il faut résoudre l'équation*

$$g'(s) = 0 \iff d_0^t \nabla f(x_0 + sd_0) = 0 \iff (0,\ 8) \begin{pmatrix} -16s \\ 16s - 8 \end{pmatrix} = 0 \iff 16s - 8 = 0.$$

On trouve

$$s_0 = \frac{1}{2}.$$

- *Puisque* $x_1 = x_0 + s_0 d_0$, *alors*

$$x_1 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

- *Puisque*

$$\nabla f(x_0) = \begin{pmatrix} 0 \\ -8 \end{pmatrix} \ et\ \nabla f(x_1) = \begin{pmatrix} -8 \\ 0 \end{pmatrix}$$

alors

$$b_0 = \frac{\|\nabla f(x_1)\|^2}{\|\nabla f(x_0)\|^2} = \frac{64}{64} = 1.$$

En conséquence,

$$d_1 = -\nabla f(x_1) + b_0 d_0 = \begin{pmatrix} 8 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 8 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \end{pmatrix}.$$

- *Maintenant,* $k = 1$. *Ainsi,*

$$x_1 + sd_1 = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + s \begin{pmatrix} 8 \\ 8 \end{pmatrix} = \begin{pmatrix} 8s \\ 4 + 8s \end{pmatrix}$$

et

$$\nabla f(x_1 + sd_1) = \begin{pmatrix} 6(8s) - 2(4 + 8s) \\ -2(8s) + 2(4 + 8s) - 8 \end{pmatrix} = \begin{pmatrix} 48s - 8 - 16s \\ -16s + 8 + 16s - 8 \end{pmatrix} = \begin{pmatrix} 32s - 8 \\ 0 \end{pmatrix}.$$

- *Cherchons* s_1 *minimisant la fonction* $g(s) = f(x_1 + sd_1)$ *par rapport à* s. *Pour cela, il faut résoudre l'équation*

$$g'(s) = 0 \iff d_1^t \nabla f(x_1 + sd_1) = 0 \iff (8,\ 8) \begin{pmatrix} 32s - 8 \\ 0 \end{pmatrix} = 0 \iff 32s - 8 = 0.$$

On trouve

$$s_1 = \frac{1}{4}.$$

– *Puisque* $x_2 = x_1 + s_1 d_1$, *alors*

$$x_2 = \begin{pmatrix} 2 \\ 6 \end{pmatrix}.$$

– *On vérifie bien que*

$$\nabla f(x_2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Le point x_2 *est un minimum global car* f *est convexe.*