Algorithme du gradient conjugué

Fabien Navarro

1 Directions conjuguées

Définition 1.

Soit A une matrice symétrique $n \times n$, définie positive. On dit que deux vecteurs x et y de $\mathbb{R}^n$ sont $A-$conjugués (ou conjugués par rapport à A) s'il vérifient

$$x^T A y = 0. \tag{1}$$

La matrice A étant définie positive, la forme bilinéaire $a(x, y) = x^T A y$ définie un produit scalaire et la relation (1) traduit l'orthogonalité des vecteurs x et y pour ce produit scalaire.

Définition 2.

Soit $\{d_0, d_1, \ldots, d_n\}$ une famille de vecteur $A-$ conjugués. On appelle alors méthode de directions conjuguées la méthode

$$\begin{cases} x_0 \in \mathbb{R}^n \text{ , donné} \\ x_{k+1} = x_k - \rho_k d_k. \end{cases} \tag{2}$$

2 Algorithme de la méthode du gradient conjugué

L'idée de la méthode est de construire itérativement des directions $d_0, \ldots, d_k$ mutuellement conjuguées. A chaque étape k la direction d_k est obtenue comme combinaison linéaire du gradient en x_k et de la direction précédente d_{k-1}, les coefficients étant choisis de telle manière que d_k soit conjuguée avec toutes les directions précédentes. L'algorithme peut s'écrire sous la forme de l'algorithme 1.

Théorème 1.

Supposons que f est une fonctionnelle quadratique sur $\mathbb{R}^n$, alors :
Pour toute initialisation x_0, l'algorithme converge en au plus n itérations.

3 Exercices

Exercice 1.

On considère la fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par :

$$f(x) = \frac{1}{2} x^T A x - x^T b,$$

Algorithm 1: Gradient conjugué

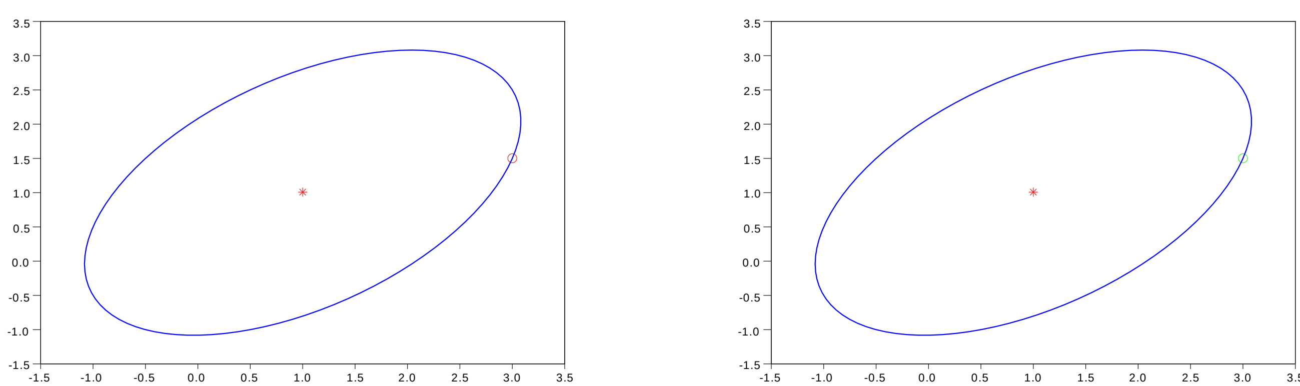
input : x_0, A, b, ε, $n_{\max}$;
output: $x^\star$;
Initialisation;
$g_0 \longleftarrow Ax_0 - b$;
$d_0 \longleftarrow -g_0$;
for $k = 0 : n_{\max}$ **do**
  $\rho_k \longleftarrow -\dfrac{g_k^T d_k}{d_k^T A d_k}$;
  $x_{k+1} \longleftarrow x_k + \rho_k d_k$;
  $g_{k+1} \longleftarrow g_k + \rho_k A d_k$;
  $\beta_k \longleftarrow \dfrac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$;
  $d_{k+1} \longleftarrow -g_{k+1} + \beta_k d_k$;
  if $\|g_{k+1}\|^2 < \varepsilon$ **then**
    Stop;
  end
end

où A et b sont repectivement une matrice symétrique définie positive d'ordre n et un vecteur de $\mathbb{R}^n$.

On testera la méthode sur l'exemple suivant :

$$A = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

1) Écrire une fonction `gradientconj` sur le modèle de l'alorithme 1.
2) Résoudre le problème (2) pour $n = 10$, $\varepsilon = 1e\text{-}7$ et $x_0 = (0, \dots, 0)^T$.
3) Comparer vos résultats avec ceux obtenus avec les méthodes de gradient à pas fixe et à pas optimal (en prenant par exemple $n = 2, 10, 100, \dots$).



(a) gradient fixe (b) gradient optimal

FIGURE 1 – Comparaison des méthodes de gradient dans le cas $n = 2$.

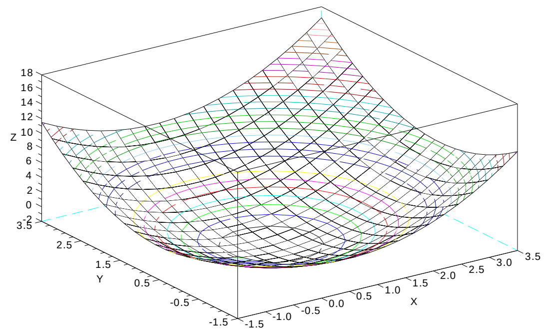


FIGURE 2 – graph de la fonction à minimiser ($n = 2$).

Exercice 2.

On considère la fonction :

$$\begin{aligned} g : \mathbb{R}^3 &\to \mathbb{R} \\ (x, y, z) &\mapsto x^2 + (x-y)^2 + 3(y-z)^2 - 2x + 6y - 6z \end{aligned}$$

1) Calculer ∇g et $\nabla^2 g$ (vous pouvez utiliser `derivative`).
2) Écrire le problème de minimisation sous la forme d'un système linéaire de type $Ax = b$.
3) Vérifier que la matrice A est définie positive.
4) Résoudre le problème $\min_{\mathbb{R}^3} g(x)$ en utilisant l'algorithme du gradient conjugué, avec $(x_0, y_0, z_0) = (0, 0, 0)^T$.