Plan

- Problèmes quadratiques
- Méthode de Newton pure
- Méthodes de descente avec recherche linéaire
- Méthodes de région de confiance
- Méthodes quasi-Newton

Problèmes quadratiques

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^T Q x + b^T x + c$$

- Q symétrique définie positive
- Minimum global unique
- Comme $\nabla f(x) = Qx + b$, nous avons

$$x^* = -Q^{-1} b$$

- Algorithme direct: résoudre $Qx = -b$

Equivalent à la résolution d'un système d'équations linéaires

Algorithme : Résolution directe

Objectif

Trouver le minimum global de $\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}x^T Q x + b^T x$

Input

- La matrice Q symétrique définie positive.
- Le vecteur b.

Output

La solution $x^* \in \mathbb{R}^n$.

Algorithme : Résolution directe

Résolution

1. Calculer la factorisation de Cholesky $Q = LL^T$.
2. Calculer y^* solution du système triangulaire inférieur $Ly = -b$.
3. Calculer x^* solution du système triangulaire supérieur $L^T x = y^*$.

Méthode des gradients conjugués

- Méthode itérative
- Appropriée pour les problèmes de grande taille

Directions conjuguées

Soit une matrice $Q \in \mathbb{R}^{n \times n}$, définie positive.
Les vecteurs non nuls de $\mathbb{R}^n$ $d_1, \ldots, d_k$ sont dits Q-conjugués si

$$d_i^T Q d_j = 0 \quad \forall i, j \text{ tels que } i \neq j.$$

Si $Q = I$, alors conjugués $\Longleftrightarrow$ orthogonaux

Directions conjuguées

Indépendance des directions conjuguées *Soit $d_1, \ldots, d_k$ un ensemble de directions non nulles et Q-conjuguées.*
Alors les vecteurs $d_1, \ldots, d_k$ sont linéairement indépendants.

(p. 231)

Corollaire immédiat : il y a toujours au maximum n directions Q-conjuguées

Directions conjuguées

Idée de la méthode :

- Considérons n directions conjuguées $d_1, \ldots, d_n$
- A chaque itération $k = 1, \ldots, n$

$$x_{k+1} = x_k + \alpha_k d_k$$

avec

$$\alpha_k = \operatorname{argmin}_\alpha f(x_k + \alpha d_k).$$

Directions conjuguées

Lemme : *Soient $d_1, \ldots, d_n$ un ensemble de directions Q-conjuguées dans $\mathbb{R}^n$. Soient $x_1, \ldots, x_{n+1}$ les itérés générés par une méthode de directions conjuguées. Alors, pour tout $k = 1, \ldots, n$,*

- *Le pas α_k est défini par $\alpha_k = -\frac{d_k^T(Qx_k+b)}{d_k^T Q d_k} = -\frac{d_k^T \nabla f(x_k)}{d_k^T Q d_k}$*
- *$\nabla f(x_k)$ est orthogonal à $d_1, \ldots, d_{k-1}$, c'est-à-dire*

$$\nabla f(x_k)^T d_i = 0, \quad i = 1, \ldots, k-1.$$

- $\nabla f(x_{n+1}) = 0$
- *Pour tout k tel que $\nabla f(x_k) = 0$, alors*

$$\nabla f(x_i) = 0 \quad i = k, \ldots, n+1.$$

(p. 231)

Directions conjuguées

Méthode des directions conjuguées *Soient* $d_1,\ldots,d_\ell$, $\ell \leq n$, *un ensemble de directions* Q*-conjuguées, soit* $x_1 \in \mathbb{R}^n$ *et soit*

$$M_\ell = x_1 + \langle d_1, \ldots, d_\ell \rangle = \{x | x = x_1 + \sum_{k=1}^{\ell} \lambda_k d_k\}$$

le sous-espace affine engendré par les directions $d_1,\ldots,d_\ell$*. Alors, le minimum global du problème*

$$\min_{x \in M_\ell} f(x) = \frac{1}{2} x^T Q x + b^T x$$

est...

(suite...)

Directions conjuguées

Méthode des directions conjuguées (suite)

$$x_{\ell+1} = x_1 + \sum_{k=1}^{\ell} \alpha_k d_k$$

avec

$$\alpha_k = \min_{\alpha} f(x_k + \alpha d_k) = -\frac{d_k^T (Q x_k + b)}{d_k^T Q d_k}$$

(p. 233)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Directions conjuguées

Convergence de la méthode des directions conjuguées *Soit $Q \in \mathbb{R}^{n\times n}$ définie positive. Soient $d_1, \ldots, d_n$, un ensemble de directions Q-conjuguées. Soit $x_1 \in \mathbb{R}^n$ arbitraire. L'algorithme basé sur la récurrence $x_{k+1} = x_k + \alpha_k d_k$ avec*

$$\alpha_k = -\frac{d_k^T(Qx_k + b)}{d_k^T Q d_k}$$

identifie le minimum global du problème

$$\min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}x^T Q x + b^T x$$

en au plus n itérations.

(corollaire)

Générer les directions conjuguées

- Idée : généralisation de Gram-Schmidt
- Soient ℓ vecteurs $\xi_1, \ldots \xi_\ell$, linéairement indépendants
- Générer ℓ vecteurs $d_1, \ldots, d_\ell$, Q-conjugués, tels que, pour tout i $=1, \ldots, \ell$,

$$\langle \xi_1, \ldots \xi_i \rangle = \langle d_1, \ldots, d_i \rangle.$$

- On choisit $d_1 = \xi_1$ et

$$d_i = \xi_i + \sum_{k=1}^{i-1} \alpha_k^i d_k.$$

- Calcul de α_k^i (p. 235)

Gradients conjugués

$$\xi_i = -\nabla f(x_i) = -Qx_i - b$$

On peut appliquer Gram-Schmidt, car les $\nabla f(x_i)$, $i = 1, \ldots, n$ sont orthogonaux, et donc linéairement indépendants.

Gradients conjugués

Lemme : *Considérons la méthode de directions conjuguées où chaque direction d_i est générée par la méthode de Gram-Schmidt appliquée aux directions $-\nabla f(x_1), \ldots, -\nabla f(x_i)$, c'est-à-dire*

$$d_i = -\nabla f(x_i) + \sum_{k=1}^{i-1} \frac{d_k^T Q \nabla f(x_i)}{d_k^T Q d_k} d_k.$$

Alors,

$$\langle \nabla f(x_1), \ldots, \nabla f(x_i) \rangle = \langle d_1, \ldots, d_i \rangle$$

et

$$\nabla f(x_i)^T \nabla f(x_k) = 0 \quad k = 1, \ldots, i-1.$$

(p. 236)

Gradients conjugués

Lemme : *Considérons la méthode de directions conjuguées où chaque direction d_i est générée par la méthode de Gram-Schmidt appliquée aux directions $-\nabla f(x_1), \ldots, -\nabla f(x_i)$.*
Si $\nabla f(x_i) \neq 0$, alors

$$d_i = -\nabla f(x_i) + \beta_i d_{i-1}$$

avec

$$\beta_i = \frac{\nabla f(x_i)^T \nabla f(x_i)}{\nabla f(x_{i-1})^T \nabla f(x_{i-1})}$$

(p. 237)

Algorithme : Gradients conjugués

Objectif

Trouver le minimum global de $\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^T Q x + b^T x$.

Input

- Une première approximation x_1 de la solution.
- La matrice Q symétrique définie positive.
- Le vecteur b.

Output

La solution $x^* \in \mathbb{R}^n$.

Algorithme : Gradients conjugués

Initialisation

$k = 1$, $d_1 = -Qx_1 - b$.

Itérations

- Calculer le pas
$$\alpha_k = -\frac{d_k^T(Qx_k + b)}{d_k^T Q d_k}$$
- Calculer l'itéré suivant $x_{k+1} = x_k + \alpha_k d_k$

Algorithme : Gradients conjugués

Itérations (suite)

- Calculer

$$\beta_{k+1} = \frac{\nabla f(x_{k+1})^T \nabla f(x_{k+1})}{\nabla f(x_k)^T \nabla f(x_k)} = \frac{(Qx_{k+1}+b)^T(Qx_{k+1}+b)}{(Qx_k+b)^T(Qx_k+b)}$$

- Calculer la nouvelle direction

$$d_{k+1} = -Qx_{k+1} - b + \beta_{k+1} d_k$$

- $k = k + 1$

Critère d'arrêt

Si $\|\nabla f(x_k)\| = 0$ ou $k = n + 1$, alors $x^* = x_k$.

Gradients conjugués

Exemple :

$$Q = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} \qquad b = \begin{pmatrix} -4 \\ -7 \\ -9 \\ -10 \end{pmatrix}$$

k	x_k	$\nabla f(x_k)$	d_k	α_k	β_k
1	+5.00000e+00	+1.60000e+01	-1.60000e+01	+1.20766e-01	
	+5.00000e+00	+2.80000e+01	-2.80000e+01		
	+5.00000e+00	+3.60000e+01	-3.60000e+01		
	+5.00000e+00	+4.00000e+01	-4.00000e+01		
2	+3.06775e+00	+1.50810e+00	-1.52579e+00	+1.02953e+00	+1.10547e-03
	+1.61856e+00	+9.48454e-01	-9.79407e-01		
	+6.52430e-01	-2.29750e-01	+1.89953e-01		
	+1.69367e-01	-1.06038e+00	+1.01616e+00		
3	+1.49690e+00	+1.70656e-01	-1.97676e-01	+2.37172e+00	+1.77089e-02
	+6.10224e-01	-1.55585e-01	+1.38241e-01		
	+8.47993e-01	-9.20500e-02	+9.54138e-02		
	+1.21554e+00	+1.23492e-01	-1.05497e-01		
4	+1.02806e+00	+5.77796e-03	-8.27569e-03	+3.39118e+00	+1.26355e-02
	+9.38093e-01	-1.65085e-02	+1.82552e-02		
	+1.07429e+00	+2.31118e-02	-2.19062e-02		
	+9.65332e-01	-1.15559e-02	+1.02229e-02		
5	+1.00000e+00	-1.66356e-12			
	+1.00000e+00	-3.12639e-12			
	+1.00000e+00	-4.21174e-12			
	+1.00000e+00	-4.78906e-12			