Ciarlet

Méthode du gradient conjugué.

Théorème: Soit $A \in S_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$. On veut résoudre $Ax = b$. Soit $J : x \longmapsto \frac{1}{2} \langle Ax ; x \rangle - \langle b ; x \rangle$. On cherche alors à minimiser J. Soit $u_0 \in \mathbb{R}^n$. À chaque itération, k, on pose $G_k = \mathrm{Vect}(\nabla J(u_\ell))_{\ell \in [\![0, k]\!]}$, et on définit le vecteur suivant, u_{k+1}, comme $J(u_{k+1}) = \inf J_{|u_k + G_k}$. On s'arrête dès que $\nabla J(u_k) = 0$.

La méthode s'arrête en au plus n étapes et converge vers $A^{-1}b$.

Démonstration: ① Mq la famille $(\nabla J(u_\ell))_{\ell \leq k}$ est orthogonale.

u_{k+1} minimise $J_{|u_k + G_k}$. Donc par les extremas liés, $\forall w \in G_k$, $\langle \nabla J(u_{k+1}) ; w \rangle = 0$

Donc en particulier, $\forall i \leq k$, $\langle \nabla J(u_{k+1}) ; \nabla J(u_i) \rangle = 0$, i.e. $\nabla J(u_{k+1}) \in G_k^\perp$.

On en déduit que la famille $(\nabla J(u_\ell))_{\ell \leq k}$, est libre si tous les $\nabla J(u_\ell)$ sont non nuls. Donc $\dim G_k = k+1 \leq n$. L'algorithme se finit donc en au plus n étapes puisque si $\nabla J(u_\ell) \neq 0$ pour $\ell \leq n-1$, $\nabla J(u_n) = 0$ (sinon, on aurait $(n+1)$ vecteurs indépendants de $\mathbb{R}^n$).

② Il faut maintenant voir que chaque minimisation est facile à faire

Supposons que $(u_0 \cdots u_k)$ soient construits et $\nabla J(u_k) \neq 0$. On définit alors, pour $0 \leq \ell \leq k-1$

$u_{\ell+1} - u_\ell = \Delta_\ell$. $u_{\ell+1} \in G_\ell$ et $u_\ell \in G_{\ell-1} \subset G_\ell$. Donc $\Delta_\ell \in G_\ell$, i.e.

$$\exists (\delta_i^\ell)_{\substack{\ell \leq k-1 \\ i \leq \ell}} \text{ tels que } \Delta_\ell = \sum_{i=0}^{\ell} \delta_i^\ell \nabla J(u_i).$$

Propriété: Les (Δ_ℓ) forment une famille orthogonale pour le p.s associé à A.

Preuve: $\forall x \in \mathbb{R}^n$, $\nabla J(x) = Ax - b$.

Donc $\nabla J(u_{\ell+1}) = \nabla J(u_\ell + \Delta_\ell) = Au_\ell - b + A\Delta_\ell = \nabla J(u_\ell) + A\Delta_\ell$ si $\ell \leq k-1$

- D'une part $0 = \langle \nabla J(u_{\ell+1}) ; \nabla J(u_\ell) \rangle = \|\nabla J(u_\ell)\|^2 + \langle A\Delta_\ell ; \nabla J(u_\ell) \rangle$

  Donc $\Delta_\ell \neq 0$ (car $\nabla J(u_\ell) \neq 0$)
- D'autre part $0 = \langle \nabla J(u_{\ell+1}) ; \nabla J(u_i) \rangle = \underbrace{\langle \nabla J(u_\ell) ; \nabla J(u_i) \rangle}_{=0} + \langle A\Delta_\ell ; \nabla J(u_i) \rangle$

  $= \langle A\Delta_\ell ; \nabla J(u_i) \rangle$ si $i < \ell \leq k-1$

  Or, $\Delta_i \in G_i = \mathrm{Vect}(\nabla J(u_\ell))_{\ell \leq i}$

Donc $\forall i < \ell$, $(A\Delta_\ell \mid \Delta_i) = 0 = <\Delta_\ell ; \Delta_i>_A$.

Une conséquence immédiate est la liberté des $(\Delta_\ell)_{\ell \leq k-1}$ dans G_{k-1}.

Donc $(\Delta_0 \cdots \Delta_{k-1})$ est une base de G_{k-1} et comme les $(\Delta_\ell)_{\ell \leq k-1}$ sont échelonnés en les $\nabla J(u_\ell)$, il s'ensuit que $\delta_\ell^\ell \neq 0$ pour tout ℓ.

$$\Delta_\ell = \delta_\ell^\ell \Big[\sum_{k=0}^{\ell-1} \underbrace{\frac{\delta_i^\ell}{\delta_\ell^\ell}}_{=\lambda_i^\ell} \nabla J(u_k) + \nabla J(u_\ell) \Big] \stackrel{def}{=} \delta_\ell^\ell \, d_\ell = \delta_\ell^\ell \Big[\sum_{i=0}^{\ell-1} \lambda_i^\ell \nabla J(u_i) + \nabla J(u_\ell) \Big]$$

③ Calcul de $u_{k+1} = u_k + \delta_k^k d_k = \inf J_{|u_k+G_k}$

(On calcule séparément δ_k^k et d_k (i.e les λ_i^k pour $i \leq k-1$)

- On obtient δ_k^k comme annulateur de la dérivée de $t \mapsto J(u_k + t d_k)$

Si on connaît d_k, il vient que $\delta_k^\ell = \dfrac{<d_k ; \nabla J(u_k)>}{<d_k ; A d_k>}$.

- Si $i < k$, $0 = <A d_k ; \Delta_i>$ par la propriété.

$$= <d_k ; A\Delta_i> = <d_k ; \nabla J(u_{i+1}) - \nabla J(u_i)>$$

Donc $0 = <\sum_{i=0}^{k-1} \lambda_i^k \nabla J(u_i) + \nabla J(u_k) ; \nabla J(u_{i+1}) - \nabla J(u_i)>$ si $i \leq k-1$

Comme les $(\nabla J(u_i))_{i \leq k-1}$ forment une BON de l'espace, on a :

$$\begin{cases} \to \text{Si } i = k-1 & \|\nabla J(u_k)\|^2 - \lambda_{k-1}^k \|\nabla J(u_{k-1})\|^2 = 0 \\ \to \text{Sinon} & \lambda_{i+1}^k \|\nabla J(u_{i+1})\|^2 - \lambda_i^k \|\nabla J(u_i)\|^2 = 0. \end{cases}$$

Donc $\lambda_i^k = \dfrac{\|\nabla J(u_k)\|^2}{\|\nabla J(u_i)\|^2}$ et en réinjectant dans l'expression de d_k,

il vient que $d_k = \nabla J(u_k) + \dfrac{\|\nabla J(u_k)\|^2}{\|\nabla J(u_{k-1})\|^2} d_{k-1}$.

④ Algorithme complet : $d_0 = \nabla J(u_0)$. Si $d_0 = 0$, c'est bon.

Sinon $u_1 = u_0 + r_0 d_0$ où $r_0 = \dfrac{<d_0 \mid d_0>}{<d_0 ; A d_0>}$ (cf premier point de ③).

Sinon $d_1 = \nabla J(u_1) + \dfrac{\|\nabla J(u_1)\|^2}{\|\nabla J(u_0)\|^2} d_0$ et on itère.