Vitesse de convergence du gradient conjugué

Théorème. *Soit $n \geq 1$, $|\alpha| > 1$. Il existe un unique polynôme T_n tel que pour tout θ, $T_n(\cos(\theta)) = \cos(n\theta)$.*
Ce polynôme est de degré n et vérifie de plus $T_n(\cosh(\theta)) = \cosh(n\theta)$.
Enfin, si $Q \in \mathbb{R}_n[X]$ vérifie $Q(\alpha) = T_n(\alpha)$ et $\max_{t\in[-1,1]} |Q(t)| < 1$, alors $Q = T_n$. Par conséquent,

$$\min_{deg(Q)\leq n, Q(\alpha)=1} \max_{t\in[-1,1]} |Q(t)| = \frac{1}{|T_k(\alpha)|}$$

Démonstration. Comme

$$\cos(n\theta) = Re((\cos(\theta) + i\sin(\theta))^n) = \sum_{0\leq k\leq n/2} \binom{n}{2k} \cos(\theta)^{n-2k}(\cos(\theta)^2 - 1)^k$$

Il existe donc bien un tel polynôme donné par :

$$T_n(X) = \sum_{0\leq k\leq n/2} \binom{n}{2k} X^{n-2k}(X^2 - 1)^k$$

Deux tels polynômes coïncident sur $[-1, 1]$, d'où l'unicité de T_n, et on constate que $\deg(T_n) = n$ (les termes sont de degrés n, avec des coefficient dominants positifs).

Comme $\cosh(z) = \cos(iz)$ et $T_n(\cos(z))$, $\cos(nz)$ sont des fonctions entières qui coïncident sur un segment, elles coïncident partout et $T_n(\cosh(\theta)) = \cosh(n\theta)$ (remarque : on peut donner des arguments bien plus élémentaires..)

Soit enfin Q comme dans l'énoncé, on note $x_k = \cos(\pi k/n)$. On a $1 = x_0 > x_1 > .. > x_n = -1$, et $T_n(x_k) = \cos(k\pi) = (-1)^k$. Alors pour tout $k \in \{0, .., n\}$, on a

$$Q(x_k) \begin{cases} < T_n(x_k) \text{ si } k \text{ est pair} \\ > T_n(x_k) \text{ si } k \text{ est impair} \end{cases}$$

Donc T_n et Q s'intersectent sur chaque $]x_{i+1}, x_i[$ et en α, d'où $Q = T_n$; absurde, il n'existe donc pas de tel polynôme. □

Corollaire. *Soit $A \in S_N^{++}(\mathbb{R})$ de valeurs propres $\lambda_1 \leq .. \leq \lambda_N$, $b \in \mathbb{R}^N$, $J(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$. Soit $x_0 \in \mathbb{R}^N$ et $(x_n)_n$ la suite construite à partir de x_0 dans la méthode du gradient conjugué appliqué à J, et $e_n = x_n - \bar{x}$.*
Soit $\bar{x} := A^{-1}b$, et $c = Cond(A) = \frac{\lambda_N}{\lambda_1}$. Alors

$$\|e_n\|_2 \leq 2\sqrt{c}\left(\frac{\sqrt{c}-1}{\sqrt{c}+1}\right)^n \|e_0\|_2$$

Démonstration. Un calcul rapide donne $J(x_n) - J(\bar{x}) = \frac{1}{2}\langle Ae_n, e_n\rangle = \frac{1}{2}\|\sqrt{A}e_n\|^2$. On note E_n cette quantité, on a $\frac{\lambda_1}{2}\|e_n\|^2 \leq E_n \leq \frac{\lambda_N}{2}\|e_n\|^2$.
Par définition de $x_n = \operatorname{argmin}\left(J(x), x \in x_0 + \mathbb{R}_{n-1}[A]Ae_0\right)$, on a :

$$\begin{aligned}
\frac{1}{2}\|\sqrt{A}e_n\|^2 &= \min_{e\in(I+A\mathbb{R}_{n-1}[A])e_0} \frac{1}{2}\|\sqrt{A}e\|^2 \\
&= \min_{\deg(Q)\leq n, Q(0)=1} \frac{1}{2}\|\sqrt{A}Q(A)e_0\|^2 \\
&\leq E_0 \min_{\substack{\deg(Q)\leq n\\ Q(0)=1}} \|Q(A)\|^2 \\
&\leq E_0 \min_{\substack{\deg(Q)\leq n\\ Q(0)=1}} \max_{[\lambda_1,\lambda_N]} |Q|^2 \\
&= E_0 \min_{\substack{\deg(Q)\leq n\\ Q(\alpha)=1}} \max_{[-1,1]} |Q|^2,
\end{aligned}$$

où $\alpha = (c+1)/(c-1)$. Précisons cette étape : on fait un changement de variable affine sur les polynôme, qui envoie λ_1 sur 1, λ_N sur -1, et donc 0 sur α. Grâce au théorème précédent, on a donc :

$$\|e_n\| \leq \sqrt{\frac{2}{\lambda_1}E_n} \leq \sqrt{c}T_n(\alpha)^{-1}\|e_0\|.$$

On conclut en minorant $T_n(\alpha)$.
En effet, si on note θ tel que $\cosh(\theta) = \alpha$, alors

$$T_n(\cosh(\theta)) = \cosh(n\theta) \geq \frac{1}{2}e^{n\theta} = \frac{1}{2}\left(\alpha + \sqrt{\alpha^2-1}\right)^n$$

En remplaçant $\alpha = (c+1)/(c-1)$ dans cette expression, on obtient $T_n(\alpha) \geq \frac{1}{2}\left(\frac{\sqrt{c}+1}{\sqrt{c}-1}\right)^n$, d'où le résultat.

$\square$

Référence : Serre