CONVERGENCE THEORY

- **Background: Best uniform approximation;**
- **Chebyshev polynomials;**
- **Analysis of the CG algorithm;**
- **Analysis in the non-Hermitian case (short)**

Background: Best uniform approximation

We seek a function ϕ (e.g. polynomial) which deviates as little as possible from f in the sense of the $\|.\|_\infty$ -norm, i.e., we seek the

$$\min_{\phi} \max_{t \in [a,b]} |f(t) - \phi(t)| = \min_{\phi} \|f - \phi\|_\infty$$

where ϕ is in a finite dimensional space (e.g., space of polynomials of degree $\leq n$)

- Solution is the "best uniform approximation to f"
- Important case: ϕ is a polynomial of degree $\leq n$
- In this case ϕ belongs to $\mathbb{P}_n$

The Min-Max Problem:

$$\rho_n(f) = \min_{p \in \mathbb{P}_n} \max_{x \in [a,b]} |f(t) - p(t)|$$

➤ If f is continuous, best approximation to f on $[a, b]$ by polynomials of degree $\leq n$ exists and is unique

➤ ... and $\lim_{n\to\infty} \rho_n(f) = 0$ (Weierstrass theorem).

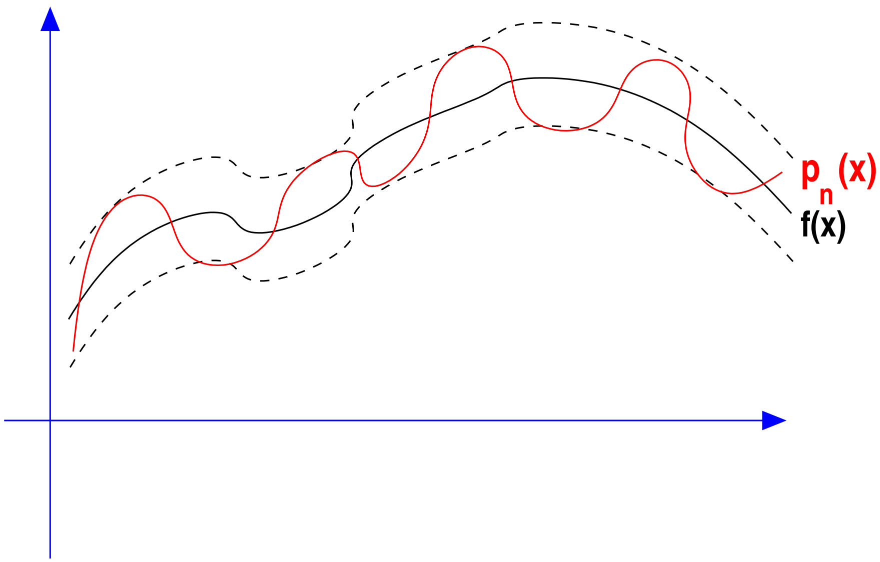
Question: How to find the best polynomial?

Answer: Chebyshev's equi-oscillation theorem.

Chebyshev equi-oscillation theorem: p_n is the best uniform approximation to f in $[a, b]$ **if and only if** there are $n + 2$ points $t_0 < t_1 < \ldots < t_{n+1}$ in $[a, b]$ such that

$$f(t_j) - p_n(t_j) = c(-1)^j \|f - p_n\|_\infty \quad \text{with} \quad c = \pm 1$$

[p_n 'equi-oscillates' $n + 2$ times around f]



Application: Chebyshev polynomials

Question: Among all **monic** polynomials of degree $n+1$ which one minimizes the infinity norm? Problem:

$$\text{Minimize } \|t^{n+1} - a_n t^n - a_{n-1} t^{n-1} - \cdots - a_0\|_\infty$$

Reformulation: Find the best uniform approximation to t^{n+1} by polynomials p of degree $\leq n$.

➤ $t^{n+1} - p(t)$ should be a polynomial of degree $n+1$ which equi-oscillates $n+2$ times.

- Define Chebyshev polynomials:

$$C_k(t) = \cos(k \cos^{-1} t) \text{ for } k = 0, 1, ..., \text{ and } t \in [-1, 1]$$

- Observation: C_k is a polynomial of degree k, because:
- the C_k's satisfy the three-term recurrence :

$$C_{k+1}(t) = 2xC_k(t) - C_{k-1}(t)$$

with $C_0(t) = 1$, $C_1(t) = t$.

✍1 Show the above recurrence relation

✍2 Compute $C_2, C_3, \dots, C_8$

✍3 Show that for $|t| > 1$ we have

$$C_k(t) = \text{ch}(k\ \text{ch}^{-1}(t))$$

- C_k Equi-Oscillates $k+1$ times around zero.
- Normalize C_{n+1} so that leading coefficient is 1

The minimum of $\|t^{n+1} - p(t)\|_\infty$ over $p \in \mathbb{P}_n$ is achieved when $t^{n+1} - p(t) = \frac{1}{2^n} C_{n+1}(t)$.

- Another important result:

Let $[\alpha, \beta]$ be a non-empty interval in $\mathbb{R}$ and let γ be any real scalar outside the interval $[\alpha, \beta]$. Then the minimum

$$\min_{p \in \mathbb{P}_k, p(\gamma)=1} \; \max_{t \in [\alpha,\beta]} |p(t)|$$

is reached by the polynomial: $\hat{C}_k(t) \equiv \dfrac{C_k\left(1 + 2\frac{\alpha - t}{\beta - \alpha}\right)}{C_k\left(1 + 2\frac{\alpha - \gamma}{\beta - \alpha}\right)}$.

Convergence Theory for CG

- Approximation of the form $x = x_0 + p_{m-1}(A) r_0$. with $x_0 =$ initial guess, $r_0 = b - A x_0$;
- Recall property: *x_m minimizes $\|x - x_*\|_A$ over $x_0 + K_m$*
- $\boxed{\text{Consequence:}}$ Standard result

Let $x_m = m$-th CG iterate, $x_* =$ exact solution and

$$\eta = \frac{\lambda_{min}}{\lambda_{max} - \lambda_{min}}$$

Then: $$\|x_* - x_m\|_A \le \frac{\|x_* - x_0\|_A}{C_m(1 + 2\eta)}$$

where $C_m =$ Chebyshev polynomial of degree m.

- Alternative expression. From $C_k = ch(kch^{-1}(t))$:

$$C_m(t) = \frac{1}{2}\left[\left(t+\sqrt{t^2-1}\right)^m + \left(t+\sqrt{t^2-1}\right)^{-m}\right] \geq \frac{1}{2}\left(t+\sqrt{t^2-1}\right)^m .$$

Then:

$$C_m(1+2\eta) \geq \frac{1}{2}\left(1+2\eta+\sqrt{(1+2\eta)^2-1}\right)^m \geq \frac{1}{2}\left(1+2\eta+2\sqrt{\eta(\eta+1)}\right)^m .$$

- Next notice that:

$$1+2\eta+2\sqrt{\eta(\eta+1)} = \left(\sqrt{\eta}+\sqrt{\eta+1}\right)^2 = \frac{(\sqrt{\lambda_{min}}+\sqrt{\lambda_{max}})^2}{\lambda_{max}-\lambda_{min}}$$

$$= \frac{\sqrt{\lambda_{max}} + \sqrt{\lambda_{min}}}{\sqrt{\lambda_{max}} - \sqrt{\lambda_{min}}}$$
$$= \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1}$$

where $\kappa = \kappa_2(A) = \lambda_{max}/\lambda_{min}$.

➤ Substituting this in previous result yields

$$\|x_* - x_m\|_A \leq 2 \left[\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right]^m \|x_* - x_0\|_A.$$

➤ Compare with steepest descent!

Theory for Nonhermitian case

- Much more difficult!
- No convincing results on *'global convergence'* for most algorithms: FOM, GMRES(k), BiCG (to be seen) etc..
- Can get a general a-priori – a-posteriori error bound

Convergence results for nonsymmetric case

- Methods based on minimum residual better understood.

- If $(A + A^T)$ is positive definite ($(Ax, x) > 0 \ \forall x \neq 0$), all minimum residual-type methods (ORTHOMIN, ORTHODIR, GCR, GMRES,...), $+$ their restarted and truncated versions, converge.

- Convergence results based on comparison with one-dim. MR [Eisenstat, Elman, Schultz 1982] $\longrightarrow$ not sharp.

MR-type methods: if $A = X\Lambda X^{-1}$, Λ diagonal, then

$$\|b - Ax_m\|_2 \leq \text{Cond}_2(X) \min_{p \in \mathcal{P}_{m-1}, p(0)=1} \ \max_{\lambda \in \Lambda(A)} |p(\lambda)|$$

($\mathcal{P}_{m-1} \equiv$ set of polynomials of degree $\leq m - 1$, $\Lambda(A) \equiv$ spectrum of A)