The Basic Iterative Method

15-859N
9/30/16

Goal: Solve $Au=b$ approximatly

<u>Simpler prob</u> solve $a \cdot x = 1$ approx without division?

Def $\bar{a} = 1 - a$ or $a = 1 - \bar{a}$

thus

$$\frac{1}{a} = \frac{1}{1-\bar{a}} = 1 + \bar{a} + \bar{a}^2 + \cdots \quad \text{for } |\bar{a}| < 1$$

or $0 < a < 2$.

Let $u^{(k)} = 1 + \bar{a} + \cdots + \bar{a}^{k-1}$

Recurrance $u^{(k+1)} = 1 + \bar{a}(u^{(k)})$

Suppose goal: $a \cdot x = b$

Set $v^{(k)} = b\,u^{(k)}$ or $u^{(k)} = v^{(k)}/b$

recurrence:

$$v^{(k+1)} = b\left(1 + \frac{\bar{a}}{b} v^{(k)}\right) = b + \bar{a} v^{(k)}$$

$$= b + (1-a) v^{(k)} = v^{(k)} - \left(b - a v^{(k)}\right)$$

$$0 < a < 2$$

<u>Error Analysis</u>

Set $\varepsilon^{(k)} = \frac{b}{a} - v^{(k)}$ $\quad \varepsilon^{(0)} = \frac{b}{a}$

$$\varepsilon^{(k)} = b(1 + \bar{a} + \bar{a}^2 + \cdots) - b(1 + \bar{a} \cdots \bar{a}^{k-1})$$

$$= b\bar{a}^k (1 + \bar{a} + \cdots) = \frac{b\bar{a}^k}{a} = \bar{a}^k \left(\frac{b}{a}\right)$$

$$= \bar{a}^k \varepsilon^{(0)}$$

converges iff $|\bar{a}| < 1$

Richardson's Method

view 1 $u^{(m+1)} = u^{(m)} + (b - Au^{(m)})$

residual error

view 2 $u^{(m+1)} = (I - A)u^{(m)} + b$

$$u^{(m+1)} = Gu^{(m)} + b \quad G = (I - A)$$

Good: 1) $O(w(A))$ work per step.
(parallel)

2) A fixed point is a solution.

Bad: 1) May not converge.

2) May do so very slowly.

The error for Richardson

Suppose $A\bar{u}=b$ & $G=(I-A)$

Def Error $\varepsilon^{(m)}=u^{(m)}-\bar{u}$

Claim $\varepsilon^{(m+1)}=G\varepsilon^{(m)}$

pf
$$
\begin{aligned}
G\varepsilon^{(m)} &= G(u^{(m)}-\bar{u}) = Gu^{(m)}-(I-A)\bar{u}\\
&= Gu^{(m)}-\bar{u}+A\bar{u}\\
&= Gu^{(m)}-\bar{u}+b\\
&= (Gu^{(m)}+b)-\bar{u}\\
&= u^{(m+1)}-\bar{u} = \varepsilon^{(m+1)}
\end{aligned}
$$

Note: To converge $\lim_{m\to\infty} G^m \varepsilon^{(0)} = 0$

to converge $\forall \varepsilon^{(0)}$ then $\lim_{m\to\infty} G^m = 0$

Note A sym $\Rightarrow$ G sym

thus $G = \sum \lambda_i (v_i v_i^T)$ $\quad \lambda_i \in \lambda(G)$

v_i eigenvector

and converge iff $\forall i \; |\lambda_i| < 1$

In general $\exists i$ s.t. $|\lambda_i| \geq 1$

Simple modification

to get convergence!

The Extrapolated Method

A sym

<u>Goal</u> Get Richardson to converge.

<u>Note</u> $Ax=b$ iff $\gamma Ax=\gamma b$ $\gamma\neq 0$

New recurrence

$$u^{(k+1)} = u^{(k)} - \gamma(b - Ax)$$

or $$u^{(k+1)} = (I-\gamma A)u^{(k)} + \gamma b$$

set $G_\gamma = (I-\gamma A)$

Goal: $\min\limits_{\gamma} \rho(G_\gamma) = \max\limits_{i} |\lambda_i(G_\gamma)|$

Set $M = \lambda_{max}(A)$ & $m = \lambda_{min}(A)$

Claim $\gamma = \frac{2}{M+m}$ is optimal

Note $Ax = \lambda_A x$ iff $G_\gamma x = (1-\gamma\lambda_A)x$

Simple fact: Let $Q(z)$ be a polynomial then $\lambda_i \in \lambda(G)$ iff $Q(\lambda_i) \in \lambda(Q(G))$

Here $Q(z) = 1-\gamma z$

$$\lambda_{max}(G) \le (1-\gamma\lambda_A) = 1-\gamma m = 1-\frac{2m}{M+n} = \frac{M-m}{M+m}$$

$$\lambda_{min}(G) \ge 1-\gamma M = 1-\frac{2M}{M+m} = \frac{m-M}{M+n}$$

$$= -\frac{M-m}{M+m}$$

thus $\sigma(G_\gamma) \le \frac{M-m}{M+m}$

<u>Def</u> condition number of $A \equiv \dfrac{M_A}{m_A} = k(A)$

$$\sigma(G_\gamma) = \frac{M-m}{M+m} = \frac{\frac{M}{m} - \frac{m}{m}}{\frac{M}{m} + \frac{m}{m}} = \frac{k(A)-1}{k(A)+1}$$

Set $k = k(A)$

$$\sigma = 1 - \frac{2}{k+1} \approx 1 - \frac{1}{k}$$

after k iteration $\mathrm{Err} \approx \left(1-\frac{1}{k}\right)^k = \frac{1}{e}$

We need $\approx k$ iteration / bit.

e.g. $A = L(P_n)$ $\quad k(A) \approx n^2$

Need $\approx n^2$ iteration / bit.

Polynomial Acceleration

Assume $Ax=b$ spd $A\bar{x}=b$

$$x^{(m+1)} = Gx^{(m)} + b \qquad G = (I-A)$$

$$\varepsilon^{(m)} = G^m \varepsilon^{(0)}$$

Idea: Use all previous $x^{(i)}$

---

ie Compute $x^{(0)}, \dots, x^{(n)}$

Pick $\alpha_i \in \mathbb{R}$ s.t $\sum \alpha_i = 1$

return $u^{(n)} = \sum \alpha_i x^{(i)}$

Question: How to pick the α_i?

A different view!

Def $\tilde{\varepsilon}^{(n)} = u^{(n)} - \bar{X}$

$$= \sum_{i=0}^{n-1} \alpha_i X^{(i)} - \bar{X} \quad \sum \alpha_i = 1$$

$$= \sum \alpha_i (X^{(i)} - \bar{X})$$

$$= \sum \alpha_i \varepsilon^{(i)} = \sum \alpha_i G^i \varepsilon^{(0)}$$

$$= \left(\sum \alpha_i G^i\right) \varepsilon^{(0)} = Q_n(G)\, \varepsilon^{(0)}$$

where $Q_n(z) = \alpha_0 + \cdots + \alpha_{n-1} z^{n-1}$

$$\sum \alpha_i = 1$$

Goal: For each n pick $Q_n(z)$

$$\min \sigma(Q_n(G))$$

Thus $\lambda(Q_n(G)) = Q_n(\lambda(G))$

Chebyshev Polys & Acceleration

Def Chebyshev Polynomials

$$T_0(w) = 1 \quad T_1(w) = w$$

$$T_{n+1}(w) = 2w\,T_n(w) - T_{n-1}(w)$$

eg $T_2(w) = 2w \cdot w - 1 = 2w^2 - 1$

$T_3(w) = 2w(2w^2 - 1) - w =$

$= 4w^3 - 2w - w = 4w^3 - 3w$

Note $T_n(1) = 1 \Rightarrow \sum \text{coef} = 1$

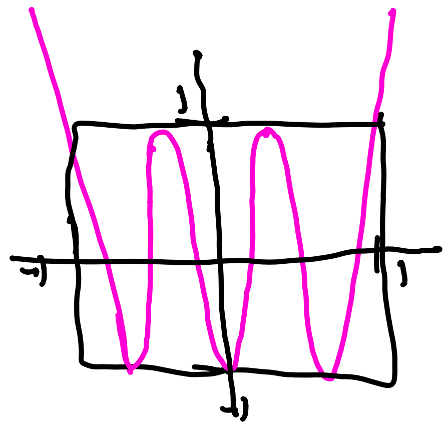
Claim: $T_n(-1) = (-1)^n$

Thm: $T_n(w)$

$$= \frac{1}{2}\left[\left(w + \sqrt{w^2 - 1}\right)^n + \left(w + \sqrt{w^2 - 1}\right)^{-n}\right]$$

HW.

$T_6(w)$



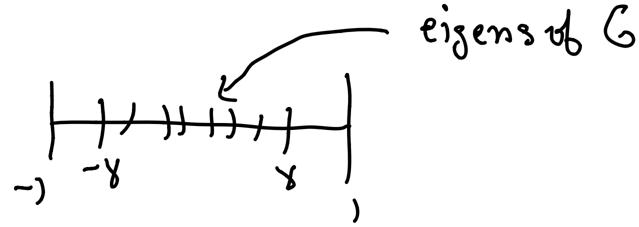
<u>**Thm**</u> $d > 1$ **Let** $H_n(w) = T_n(w) / T_n(d)$ then

1) $\displaystyle\max_{-1 \le w \le 1} |H_n(w)| = 1/T_n(d)$

2) $H_n(w)$ is min such poly

Chebyshev Plus Extrapolated Method

Let $\kappa = \kappa(A)$ & $\gamma = \frac{\kappa - 1}{\kappa + 1}$

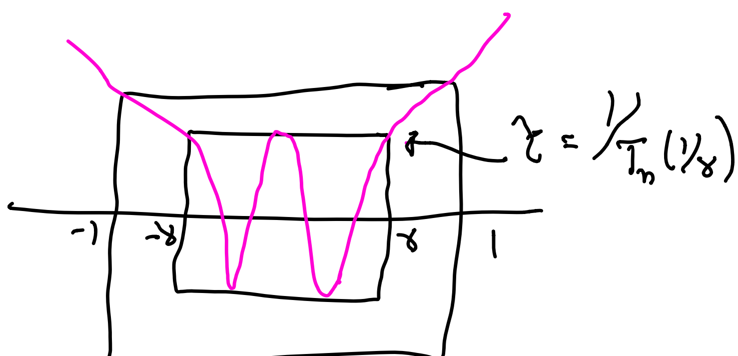


We will use $\tilde{T}_n = T_n(x/\gamma) / T_n(1/\gamma)$

for our poly excelleration!

Note $\tilde{T}_n(1) = 1$ thus $\sum \text{coef} = 1$

picture



Note: $|\lambda(G)| \leq \gamma$ thus $|\lambda(Q_n(G))| < \tau$

<u>Goal</u> bd from below $T_n(1/\gamma)$

$$1/\gamma = \frac{\kappa+1}{\kappa-1} = 1 + \frac{2}{\kappa-1} = 1 + 2\mu \quad \text{where } \mu = \frac{1}{\kappa-1}$$

<u>Note</u> $T_n(x) = \frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^n + \left(x+\sqrt{x^2-1}\right)^{-n}\right]$

$$\geq \frac{1}{2}\left(x+\sqrt{x^2-1}\right)^n \quad \text{for } x \geq 1$$

$$T_n(1+2\mu) \geq \frac{1}{2}\left(1+2\mu+\sqrt{(1+2\mu)^2-1}\right)^n$$

$$= \frac{1}{2}\left(1+2\mu+2\sqrt{\mu(\mu+1)}\right)^n$$

$$= \frac{1}{2}\left(\sqrt{\mu}+\sqrt{\mu+1}\right)^{2n} \quad (*)$$

<u>note</u>

$$\left(\sqrt{\mu}+\sqrt{\mu+1}\right)^2 = \left(\frac{1}{\sqrt{K-1}} + \frac{\sqrt{K}}{\sqrt{K-1}}\right)^n = \frac{\left(1+\sqrt{K}\right)^2}{K-1} = \frac{\sqrt{K}+1}{\sqrt{K}-1}$$

---

thus $(*) = \frac{1}{2}\left(\frac{\sqrt{K}+1}{\sqrt{K}-1}\right)^n$

Finally: $T_n(1/\gamma) \geq \frac{1}{2}\left(\frac{\sqrt{K}+1}{\sqrt{K}-1}\right)^n$

or: $\tau \leq 2\left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right)^n$

<u>Thm</u> Convergence Rate for Extrapolated Chebishev is $\Theta(1/\sqrt{\kappa})$

Back to P_n $\quad \kappa(L(P_n)) \approx n^2$

$\therefore$ Convergence for Cheb is $\Theta(1/n)$

n-iteration per bit.

Optimal for multi-term recurrence

since we are only doing "local" ops!

From Polynomial Recurance to Iterative Alg

Consider a poly defined by

$$Q_0(x) = 1$$

$$Q_1(x) = \alpha_1 x + \beta_1 \qquad \alpha_1 + \beta_1 = 1$$

$$Q_{n+1}(x) = \alpha_n x\, Q_n(x) + \beta_n Q_{n-1}(x) \qquad \alpha_n + \beta_n = 1$$

Iterative Alg

$$u^{(n+1)} = \alpha_n\left(G u^{(n)} + b\right) + \beta_n u^{(n-1)}$$

<u>to show</u> $\varepsilon^{(n)} = u^{(n)} - \bar{u}$ then $\varepsilon^{(n)} = Q_n(G)\,\varepsilon^{(0)}$

$$Q_{n+1}(G)\,\varepsilon^{(0)} = \left[\alpha_n G\, Q_n(G) + \beta_n Q_{n-1}(G)\right]\varepsilon^{(0)} \qquad \text{induct!}$$

$$= \alpha_n G \underbrace{Q_n(G)\,\varepsilon^{(0)}} + \beta_n Q_{n-1}(G)\,\varepsilon^{(0)}$$

$$= \alpha_n G\left(\varepsilon^{(n)}\right) + \beta_n\left(\varepsilon^{(n-1)}\right)$$

$$= \alpha_n G(u^{(n)} - \bar{u}) + \beta_n (u^{(n-1)} - \bar{u})$$

$$= \alpha_n (G u^{(n)} - \bar{u} + b) + \beta_n u^{(n-1)} - \beta_n \bar{u}$$

$$= \alpha_n (G u^{(n)} + b) + \beta_n u^{(n-1)} - \bar{u}$$

$$= u^{(n+1)} - \bar{u} = \varepsilon^{(n+1)}$$