Acceleration I: Chebychev Polynomials

Monday, March 9, 2020 8:38 PM

yullearn

Recall the Richardson Iteration (Assume A is symmetric)

$$x^{(t)} = x^{(t-1)} - (Ax^{(t-1)} - b)$$

$$= (I - A)x^{(t-1)} + b$$

$$= \sum_{k=0}^{t} (I-A)^k b = p_t(A) \cdot b$$

We want $\|p_t(A).b - x^*\| \leq \varepsilon \|x^*\|$

i.e. $\|(p_t(A).A - I)x^*\| \leq \varepsilon \|x^*\|$

or $\|I - Ap_t(A)\|_{op} \leq \varepsilon$.

i.e. $\|1 - \lambda(Ap_t(A))\| \leq \varepsilon$

or $\|1 - \lambda(A)\, p_t(\lambda(A))\| \leq \varepsilon$

i.e. $\|1 - \lambda p_t(\lambda)\| \leq \varepsilon \quad \forall$ eigenvalues λ of A.

$\|1 - xp_t(x)\| \leq \varepsilon \quad \forall x \in [\lambda_{min}, \lambda_{max}]$.

$q(x) = 1 - x\,p(x) \qquad q(0) = 1$

$q(x) = \left(1 - \frac{x}{\lambda_{max}(A)}\right)^t$ satisfies $\|q(x)\| \le \left(1 - \frac{\lambda_{min}}{\lambda_{max}}\right)^t$

So $t = O\left(\frac{\lambda_{max}}{\lambda_{min}} \log \frac{1}{\varepsilon}\right)$ suffices.

$= O\left(\kappa \log \frac{1}{\varepsilon}\right)$.

This is the Richardson Iteration.

Q. Can we use lower degree?

We want a polynomial $q(x)$ with $q(0) = 1$

and $\|q(x)\|$ as small as possible for $x \in [-1, 1]$

(say, after normalizing $[\lambda_{min}, \lambda_{max}] \longrightarrow [-1, 1]$).

Ans. Chebyshev polynomials!

t^{th} C.P. $\to T_t(\cos\theta) = \cos(t\theta)$

is the degree t poly s.t.

$\left(T_t(x) = \cos(t \cos^{-1}(x)) \text{ for } x \in [-1, 1]\right)$

$\cos(\theta + \alpha) = \cos\theta \cos\alpha - \sin\theta \sin\alpha$

$\ldots (t-1)\theta)\sin\theta$

$$\cos(\theta+\alpha) = \cos\theta\cos\alpha - \sin\theta\sin\alpha$$

$$\therefore \quad \cos(t\theta) = \cos((t-1)\theta)\cos\theta - \sin((t-1)\theta)\sin\theta$$

$$\cos((t-2)\theta) = \cos((t-1)\theta)\cos\theta + \sin((t-1)\theta)\sin\theta$$

$$\Rightarrow \quad \cos(t\theta) = 2\cos((t-1)\theta)\cos\theta - \cos((t-2)\theta)$$

$$T_t(x) = 2x\,T_{t-1}(x) - T_{t-2}(x) \quad (*)$$

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1 \ldots$$

Note that $T_t(\cosh\theta) = \cosh(t\cosh^{-1}(\theta))$ also holds

since $\cosh\theta = \frac{e^{\theta}+e^{-\theta}}{2}$ $\quad \sinh\theta = \frac{e^{\theta}-e^{-\theta}}{2}$

satisfies $\cosh(\theta+\alpha) = \cosh(\theta)\cosh(\alpha) + \sinh(\theta)\sinh(\alpha)$

So we get $(*)$ again.

For $x \geq 1$, $\quad T_t(x) = \cosh(t\cosh^{-1}(x))$.

$x \leq -1 \quad T_t(x) = (-1)^t\cosh(t\cosh^{-1}(-x))$

Lemma. $\quad T_t(1+\gamma) \geq \frac{1}{2}\left(1+\sqrt{2\gamma}\right)^t \quad \gamma > 0.$

Lemma. $T_t(1+\gamma) \geq \frac{1}{2}($

Pf. $$T_t(x) = \frac{1}{2}\left(e^{t\cosh^{-1}(x)} + e^{-t\cosh^{-1}(x)}\right)$$

$\cosh^{-1}(x) = \ln(x+\sqrt{x^2-1})$ for $x \geq 1$.

$$\geq \frac{1}{2}\left(x+\sqrt{x^2-1}\right)^t$$

$x = 1+\gamma$

$$= \frac{1}{2}\left(1+\gamma+\sqrt{2\gamma+\gamma^2}\right)^t \geq \frac{1}{2}\left(1+\sqrt{2\gamma}\right)^t.$$

Thm. $\exists\, q$ of degree $t = O(\sqrt{\kappa}\log\frac{1}{\varepsilon})$.

Pf. shift: $$f(x) = \frac{\lambda_{max}+\lambda_{min}-2x}{\lambda_{max}-\lambda_{min}}$$

$$f(x) = \begin{cases} -1 & x = \lambda_{max} \\ 1 & x = \lambda_{min} \\ \frac{\lambda_{max}+\lambda_{min}}{\lambda_{max}-\lambda_{min}} & x = 0. \end{cases}$$

$$q(f(x)) = \frac{T_t(f(x))}{T_t(f(0))} \qquad \text{s.t. } q(0) = 1.$$

$$q(f(x)) = \frac{T_t(f(x))}{T_t(f(0))} \quad \text{s.t. } q(0) = 1.$$

$\forall x$: $f(x) \in [-1, 1]$ $\quad |T_t(f(x))| \le 1$. $\quad (\cos\theta)$.

$$T_t(f(0)) = T_t\left(1 + \frac{2}{\frac{\lambda_{max}}{\lambda_{min}} - 1}\right) = T_t\left(1 + \frac{2}{K-1}\right)$$

$$\geq \frac{1}{2}\left(1 + \sqrt{\frac{2}{K-1}}\right)^t$$

$$\therefore |q(x)| \le \frac{2}{\left(1+\sqrt{\frac{2}{K-1}}\right)^t} \quad \forall x \in [-1, 1]$$

i.e. $t = O\left(\sqrt{K-1} \cdot \log \frac{1}{\varepsilon}\right)$ suffices.