

# Acceleration I: Chebyshev Polynomials

Monday, March 9, 2020 8:38 PM

yellow

Recall the Richardson Iteration (Assume  $A$  is symmetric)

$$\begin{aligned}x^{(t)} &= x^{(t-1)} - (A x^{(t-1)} - b) \\&= (I - A)x^{(t-1)} + b \\&= \sum_{k=0}^t (I - A)^k b = p_t(A) \cdot b\end{aligned}$$

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

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

or  $\|I - A p_t(A)\|_F \leq \varepsilon.$

i.e.  $\|I - \lambda(A) p_t(\lambda)\| \leq \varepsilon$

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

i.e.  $\|I - \lambda p_t(\lambda)\| \leq \varepsilon \quad \text{if eigenvalues } \lambda \text{ of } A.$

$\|I - x p_t(x)\| \leq \varepsilon \quad \text{if } x \in [\lambda_{\min}, \lambda_{\max}]$ .

$$q_t(x) = 1 - x p(x) \quad q(0) = 1$$

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

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

This is the Richardson Iteration.

Q. Can we use lower degree?

We want a polynomial  $q_t(x)$  with  $q(0) = 1$  and  $\|q_t(x)\|$  as small as possible for  $x \in [-1, 1]$  (say, after normalizing  $[\lambda_{\min}, \lambda_{\max}] \rightarrow [-1, 1]$ ).

Ans. Chebyshev polynomials!

$t^{\text{th}}$  C.P.  $\xrightarrow{\text{is the degree}} T_t(\cos \theta) = \cos(t\theta)$

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

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

$$\therefore (L-1) \sin \alpha$$

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

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

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

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

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

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

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

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

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

so we get (\*) again.

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

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

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

Lemma,  $T_t(x) = \frac{1}{2} e^{t \cosh^{-1}(x)} + e^{-t \cosh^{-1}(x)}$

Pf.

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

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

$$\begin{aligned} & \geq \frac{1}{2} \left( x + \sqrt{x^2 - 1} \right)^t \\ x = 1 + r & \quad = \frac{1}{2} \left( (1+r) + \sqrt{2r+r^2} \right)^t \geq \frac{1}{2} (1+\sqrt{2r})^t. \end{aligned}$$

Thm.  $\exists q$  of degree  $t = O(\sqrt{k} \log \frac{1}{\epsilon})$ .

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

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

$$q(f(x)) = \frac{T_t(f(x))}{T_t(f(0))}$$

$$\text{s.t. } q(0) = 1.$$

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

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

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

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

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

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

---

Another, more general proof.

Thm.  $\forall s, d \exists p(x)$  of degree  $d$  s.t.  
 $\max_{x \in [-1, 1]} |p(x) - x^s| \leq 2e^{-d/2s}$

Cos.  $\left(1 - \frac{x}{x_{\max}}\right)^s$  can be approximated to within  $\varepsilon$   
 by  $p(x)$  of degree  $\sqrt{s \log \frac{1}{\varepsilon}}$

$\max$  by  $p(x)$  of degree  $\sqrt{\log \frac{1}{\epsilon}}$

Since we use  $d = O(k \log \frac{1}{\epsilon})$

$p(x)$  of degree  $O(\sqrt{k} \log \frac{1}{\epsilon})$  suffices.

Pf. (of Thm). Let  $Y_s = \sum_{i=1}^s Y_i$        $Y_i \sim \{-1, 1\}$  uniform.

$$\mathbb{E} T_{z_s} = \frac{1}{2} (\mathbb{E} T_{z_{s-1}+1} + \mathbb{E} T_{z_{s-1}-1})$$

$$\text{But we know } x T_z(x) = \frac{1}{2} (T_{z+1}(x) + \frac{1}{2} T_{z-1}(x))$$

$$\Rightarrow \mathbb{E} T_{z_s} = x \mathbb{E} T_{z_{s-1}}(x) = x^s.$$

$$\text{Let } p(x) = \mathbb{E} T_{z_s}(x) \cdot \mathbf{1}_{|z_s| \leq d}$$

$$\text{Then } \max_{x \in [-1, 1]} |p(x) - x^s| = \max_{x \in [-1, 1]} \left| \mathbb{E}_{z_s: |z_s| > d} T_{z_s}(x) \cdot \mathbf{1}_{|z_s| > d} \right|$$

$$= \max_{z_s} |\mathbb{E}_{z_s} T_{z_s}(x)| \cdot \mathbf{1}_{z_s > d}$$

$$= \max_{x \in [-1, 1]} |E_{z_0} T_{z_0}(x)| \cdot \frac{1}{|z_0| > d}$$

$$\leq P_R(|z_0| > d)$$

$$\leq 2e^{-\frac{d^2}{2s}}.$$

---