14-Apr-2025 Conjugate Gradient

Last lecture saw iterative lin equ solvers for $Ax=b$. Richardson & Chebyshev.

each iter has a matvec & a few vector ops.

After t steps, corresponds to a poly $p_t(A)$, $x_t = p_t(A)b$.

CG: For each b, finds best polynomial $p_t(A)$ of degree t. Will be diff poly for each b.
Still, one matvec per iteration.

Exact* solution in # steps $\leq$ # eigenvalues   * if ∞-precision arithmetic

---

Measures of accuracy $\|x_t - x\| / \|x\|$ $\|Ax_t - b\| / \|b\|$

We will use A-norm $\|v\|_A = \sqrt{v^T A v} = \|A^{1/2} v\|$ for $A \succcurlyeq 0$

Ex. To compute $R_{eff_{a,b}} = (\delta_a - \delta_b)^T L^+ (\delta_a - \delta_b)$ set $v = \delta_a - \delta_b$. $x = L^+ v$, $R_{eff} = x^T L x$
$= \|x\|_L$

Assume $\|x_t - x\|_L \leq \varepsilon \|x\|_L$. Estimate $R_{eff_{a,b}}$ by $\|x_t\|_L$

have $\|x_t\|_L \leq \|x\|_L + \|x - x_t\|_L \leq R_{eff} + \varepsilon \cdot R_{eff} = (1+\varepsilon) R_{eff}$. Similarly

$\|x\|_L \leq \|x_t\|_L + \|x - x_t\|_L \Rightarrow \|x_t\|_L \geq \|x\|_L - \|x - x_t\|_L \geq (1-\varepsilon) R_{eff}$

Chebyshev does this. Let p be a poly s.t. $\|p(A)A - I\| \le \varepsilon$

because $A^{1/2}$ commutes with $p(A)$

Then, $\|p(A)Ax - x\|_A = \|A^{1/2}p(A)b - A^{1/2}x\| = \|(p(A)A - I)A^{1/2}x\| \le \|p(A)A - I\| \cdot \|A^{1/2}x\|$

$$\le \varepsilon\|A^{1/2}x\| = \varepsilon\|x\|_A$$

"Krylov Space"

CG finds best polynomial means best solution in $\operatorname{span}\{b, Ab, A^2b, \dots, A^t b\} \triangleq S_t(A,b)$

CG finds $\arg\min_{x_t \in S_t} \|x_t - x\|_A$ using t matvecs

---

$\|x_t - x\|_A = x_t^T A x_t - 2x^T A x_t + x^T A x = x_t^T A x_t - 2b^T x_t + x^T A x$

We don't know $x^T A x$, but it is constant, so just minimize $x_t^T A x_t - 2b^T x_t$

Will use a special basis of S_t. For a basis $y_0, \dots, y_t$, $x_t = \sum_{i=0}^{t} c_i y_i$

$$x_t^T A x_t - 2b^T x_t = \sum_i c_i^2 y_i^T A y_i + \underbrace{\sum_{i \ne j} c_i c_j y_i^T A y_j}_{\text{will choose basis s.t. } y_i^T A y_j = 0 \text{ for } i \ne j} - 2\sum_i c_i b^T y_i$$

Gives $\sum_i \left(c_i^2 y_i^T A y_i - 2c_i b^T y_i\right)$ minimize by $c_i = \dfrac{b^T y_i}{y_i^T A y_i}$ $\quad x_t = \sum_i y_i \dfrac{b^T y_i}{y_i^T A y_i}$

To construct $y_0 \cdots y_t$, $y_i \in S_t(A,b)$ $\quad y_0 = b \quad y_1 = Ab + \alpha y_0 = Ay_0 + \alpha y_0$, choose α

want $y_0^T A y_1 = 0$ so $y_0^T A^2 y_0 + \alpha y_0^T A y_0$, so $\alpha = \dfrac{-y_0^T A^2 y_0}{y_0^T A y_0}$ $\quad y_1 = Ay_0 - y_0 \dfrac{y_0^T A^2 y_0}{y_0^T A y_0}$

$$y_{t+1} = Ay_t - \sum_{i \le t} y_i \frac{y_i^T A^2 y_t}{y_i^T A y_i}$$ (note: only matvec needed is Ay_t)

Check: for $j \le t$, $$y_j^T A y_{t+1} = y_j^T A^2 y_t - \sum_{i \le t} \overbrace{y_j^T A y_i}^{=0 \text{ for } i \neq j} \frac{y_i^T A^2 y_t}{y_i^T A y_i}$$

$$= y_j^T A^2 y_t - y_j^T A y_j \frac{y_j^T A^2 y_t}{y_j^T A y_j} = 0$$

Simplify expression for y_{t+1}:

Claim: for $i < t-1$, $y_i^T A^2 y_t = 0$. proof $Ay_i \in \text{span}(y_0, \dots y_{i+1})$ are A-orth to y_t as $i+1 < t$

So, $$y_{t+1} = Ay_t - y_t \frac{y_t^T A^2 y_t}{y_t A y_t} - y_{t-1} \frac{y_{t-1}^T A^2 y_{t-1}}{y_{t-1} A y_{t-1}}$$

Can compute y_{t+1} by Ay_t and a constant number of vector operations

Compute $x_{t+1} = x_t + c_{t+1} y_{t+1}$ $\qquad c_{t+1} = \dfrac{b^T y_{t+1}}{y_{t+1}^T A y_{t+1}}$

Lem: # iterations $\le$ # eigvals. (Ignores b) Let $\lambda_1 \ldots \lambda_k$ be distinct eigvals of A.

(e.g. hypercube)

Consider $q(x) = \dfrac{\prod_{i=1}^{k} (\lambda_i - x)}{\prod_{i=1}^{k} \lambda_i}$ $\quad q(\lambda_i) = 0$, $q(0) = 1$

So, $\exists$ poly that gives exact solve in k iterations

$\rightarrow$ exact solve in $\le n$ iterations.

If L has m non-zeros, time is $O(mn)$, space is $O(n)$

In contrast, writing inverse needs space n^2, in general.