Lanczos Iterations & Conjugate Gradients

Recap

$$K_n(A,b) = \text{span}\{b, Ab, \dots, A^{n-1}b\}$$

$$\begin{bmatrix} | & | & & | \\ b & Ab & \dots & A^{n-1}b \\ | & | & & | \end{bmatrix} = Q_n R, \qquad AQ_n = Q_{n+1}\tilde{H}_n$$

$$Q_n^* A Q_n = \underline{H_n}$$

Approx.
$\Rightarrow$ solve $Ax = \lambda x$ and $Ax = b$
using Q_n and H_n.

Lanczos Iteration

$A = A^T$ (real symmetric)

$$H_n = Q_n^T A Q \iff H_n^T = (Q_n^T A Q_n)^T$$
$$= Q_n^T A^T Q_n$$
$$= Q_n^T A Q_n$$
$$= H_n$$

$$\begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} = T_n \; (= H_n)$$

T

$$A\begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \left[\begin{array}{cccc|c} | & | & & | & | \\ q_1 & q_2 & \cdots & q_n & q_{n+1} \\ | & | & & | & | \end{array}\right]\begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \beta_2 & \\ & \beta_2 & \alpha_3 & \ddots \\ & & \ddots & \beta_{n-1} \\ & & & \alpha_n \\ \hline & & & \beta_n \end{bmatrix}$$

$$A q_k = \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}$$

$$q_{k+1} = \left(-A q_k + \alpha_k q_k + \beta_{k-1} q_{k-1}\right)\frac{1}{\beta_k}$$

$$q_{k+1} = \frac{1}{\beta_k}\left[(\alpha_k I - A) q_k + \beta_{k-1} q_{k-1}\right]$$

3-term recurrence $\Rightarrow$ no <u>explicit</u> orthogonalization $q_1, \cdots, q_{k-2}$

$\Rightarrow$ Use $T_n w = \theta w$ to approx. $Ax = \lambda x$.

$\Rightarrow$ Convergence related to theory of orthogonal polynomials and their roots.

$\Rightarrow$ Loss of orthogonality is a problem in floating-point.
-> reorthogonalization
-> Implicit restarts

The method of Conjugate Gradients

"SPD"

$$\underset{\text{symmetric}}{A}\, x_* = b \qquad \underset{\text{positive def.}}{x^T A x > 0} \quad x \neq 0$$

CG is an analogue of GMRES for SPD matrices: approx. $x_* = A^{-1} b$ by solving an optimization problem over $K_n(A, b)$.

$$x^T A y = \langle x, y \rangle_A$$

"A-inner product"

$$\|x\|_A = \sqrt{\langle x, y \rangle_A} = \sqrt{x^T A y}$$

"A-norm"

$\Rightarrow$ CG minimizes $e_n = x_* - x_n$ over

$x_n \in K_n(A,b)$ in the norm $\|\cdot\|_A$.

CG Iteration

$$x_0 = 0,\; r_0 = b,\; \overbrace{p_0 = r_0}^{\text{search directions}}$$

for $n = 1, 2, 3, \ldots$

$$\alpha_n = (r_{n-1}^T r_{n-1}) / (\overbrace{p_{n-1}^T A p_{n-1}}^{\|p_n\|_A^2}) \qquad \text{step size}$$

$$x_n = x_{n-1} + \alpha_n p_{n-1} \qquad \text{approx soln}$$

$$r_n = r_{n-1} - \alpha_n A p_{n-1} \qquad \text{residual}$$

$$\beta_n = (r_n^T r_n) / (r_{n-1}^T r_{n-1})$$

$$p_n = r_n + \beta_n p_{n-1}$$

$\Rightarrow$ Only 1 mat-vec / iteration

$\Rightarrow$ No <u>explicit</u> orthogonalization

<u>Thm 1</u> If $r_{n-1} \neq 0$ (not yet converged)

$$\begin{aligned} K_n &= \text{span}\{b, Ab, \ldots, A^{n-1} b\} \\ &= \text{span}\{x_1, x_2, \ldots, x_n\} \end{aligned}$$

$$= \operatorname{span}\{\overset{b}{r_0}, r_1, \dots, r_{n-1}\}$$

$$= \operatorname{span}\{\overset{b}{p_0}, p_1, \dots, p_{n-1}\}$$

$\underset{\langle\,,\,\rangle}{r_n^T r_j} = 0$ for $j<n$, and $\underset{\langle\,,\,\rangle_A}{p_n^T A p_j} = 0$ $\quad j<n$

<u>Thm 2</u> $r_{n-1} \neq 0$, x_n is the unique minimizer of $\|e_n\|_A$ and $e_n = 0$ for some $n \le m$

$$e_n = x_* - x_n.$$