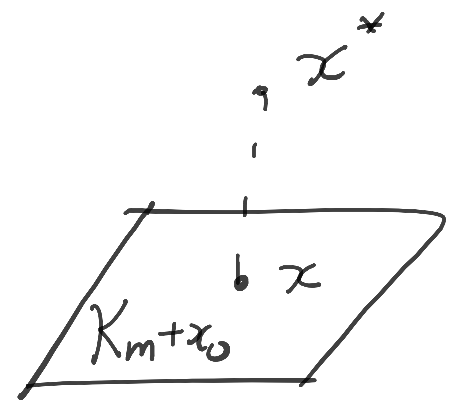
Thm. $x_m$ is the sol of CG at step $m$,

then

$$x_m = \arg\min \|x - x^*\|_A$$

s.t. $x = x_0 + q(A) r_0$,

$q \in \mathbb{P}_{m-1}$



Pf: CG:

$$K_m = \operatorname{span} V_m = \operatorname{span}[r_0, Ar_0, \cdots, A^{m-1}r_0]$$

$$V_m^T(b - Ax_m) = V_m^T V_{m+1}(\beta \tilde{e}_1 - \tilde{T}_m y_m)$$

$$= \beta e_1 - T_m y_m = 0$$

A-inner prod

$$\Rightarrow V_m^T A(x^* - x_m) = 0 \Rightarrow x_m - x^* \perp_A V_m$$

$$\Rightarrow x_m = \arg\min \| x - x^* \|_A = \arg\min \| x - x^* \|_A$$

s.t. $x = x_0 + v$, $v \in V_m$

s.t. $x = x_0 + q(A) r_0$, $q \in P_{m-1}$

$\square$.

Convergence of CG.

$\forall\ x \in x_0 + K_m$.

$$x_* - x = x_* - (x_0 + c_1 r_0 + \cdots + c_m A^{m-1} r_0)$$

$$= A^{-1}(r_0 - c_1 A r_0 - \cdots - c_m A^m r_0)$$

$$= A^{-1} P_m(A) r_0 .$$

$$P_m(z) = 1 - c_1 z - \cdots - c_m z^m . \quad P_m(0) = 1$$

$$\|x_* - x_{m+1}\|_A \le \left(\max_{\lambda_i} |P_m(\lambda_i)|\right) \underbrace{\|A^{-1}r_0\|_A}_{\substack{\| \\ \|x_0 - x_*\|_A}} .$$

$$\le \left(\min_{\substack{p \in \mathcal{P}_m \\ p(0)=1}} \max_{\lambda_1 \le t \le \lambda_n} |p(t)|\right) \|x_0 - x_*\|_A .$$

$$\downarrow$$

sol. Chebyshev poly.

$$\le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{m+1} \|x_0 - x_*\|_A$$

SD : $\kappa$.　　CG : $\sqrt{\kappa}$.

why this is CG? (same alg?)

CG: $A$ sym. pos. def.

$$A V_m = V_{m+1} \tilde{T}_m = V_m T_m + V_{m+1} e_m^T \cdot T_{m+1,m}$$

Tri-diagonal.

Arnoldi $\longrightarrow$ Lanczos.

FOM. $T_m y_m = \beta e_1 \longrightarrow$ CG.

$$T_m = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \ddots & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{m-1} \\ & & & \beta_{m-1} & \alpha_m \end{pmatrix} = L_m U_m . \leftarrow \text{LU decomp.}$$

(a.k.a. Gauss elimination)

$$L_m = \begin{pmatrix} 1 & & & 0 \\ \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ 0 & & \lambda_{m-1} & 1 \end{pmatrix}, \quad U_m = \begin{pmatrix} \eta_1 & \beta_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & \beta_{m-1} \\ 0 & & & \eta_m \end{pmatrix}$$

bi-diagonal

no-fill-in.

$$L_{m+1} = \left(\begin{array}{c|c} L_m & 0 \\ \hline 0 \cdots \lambda_m & 1 \end{array}\right) \qquad U_{m+1} = \left(\begin{array}{c|c} U_m & \begin{matrix} 0 \\ \beta_m \end{matrix} \\ \hline 0 & \eta_{m+1} \end{array}\right)$$

$$y_m = U_m^{-1} L_m^{-1} \beta e_1$$

$$x_m = x_0 + V_m y_m = x_0 + \underbrace{V_m U_m^{-1}}_{P_{m-1}} \underbrace{L_m^{-1} \beta e_1}_{\omega_{m-1}}$$

$$P_m = [\, p_0 \cdots p_m \,]$$

$$x_{m+1} = x_0 + P_m w_m$$

$$P_m = [V_m \quad v_{m+1}] U_{m+1}^{-1} = [\underbrace{V_m U_m^{-1}}_{P_{m-1}} \quad \underbrace{p_{m-1}\cdot(*) + v_{m+1}\cdot(*)}_{p_m}]$$

$$w_m = L_{m+1}^{-1} \beta e_1 = \begin{bmatrix} w_{m-1} \\ \hline * \end{bmatrix}$$

$$x_{m+1} = \underbrace{x_0 + P_{m-1} w_{m-1}}_{x_m} + p_m \alpha_m$$

$\leftarrow$ no need to keep track $P_m$!

---

prop. $P_m = [p_0, \cdots, p_{m-1}]$

A-conjugacy: $p_i^T A p_j = 0$. $i \neq j$.

Pf:
$$\underbrace{P_{m-1}^T A P_{m-1}}_{\text{sym}} = (U_m^{-1})^T V_m^T A V_m (U_m^{-1})$$

$$= (U_m^{-1})^T T_m U_m^{-1}$$

$$T_m = L_m U_m$$

$$= \underbrace{(U_m^{-1})^T L_m}$$

Lower triangular.

sym + lower triang $\rightarrow$ diagonal $\square$.

Why this is CG?

$$v_{m+1} = (*)\, r_m$$

$$p_m = (*)\, v_{m+1} + (*)\, p_{m-1}$$

$$= (*)\, \big( r_m + (*)\, p_{m-1} \big)$$

residual      conjugate direction.

$$x_{m+1} = x_m + (*)\, p_m$$

A-orthogonality
$x_{m+1}$ minimizing in $x_0 + K$ } uniquely $\Rightarrow$ CG

Lanczos + minimization.

Solve $y_m$ by

$$\min_{y_m} \| \tilde{T}_m y_m - \beta \tilde{e}_1 \|_2$$

MINimal RESidual method (MINRES)

[Paige, Saunders, 1975].

use of MINRES, sym , indefinte mat

Lanczos

CG can fail.

General unsym.

$$A V_m = V_{m+1} \tilde{H}_m \quad , \quad V_m^T V_m = I.$$

$H_m$ upper Hessenberg

sym mat.

$$A V_m = V_{m+1} \tilde{T}_m \qquad V_m^T V_m = I$$

$T_m$ tri-diag.

Trade orthogonality w. tri-diagonality?

$\downarrow$

bi-orthogonality.

Lanczos biorthogonalization method

(for non-sym matrices).

$AV = VD$ diagonalizable.

$$\underset{\substack{\| \\ W^T}}{V^{-1}} V = I$$

$W$ orthogonal to $V$!

Keep track of both $V, W$

$\longrightarrow$ tri-diagonal.

$$AV = VT. \qquad W = (V^{-1})^T$$

$$W^T A V = W^T V = T. \qquad V = [v_1, \cdots, v_n]$$

$$W = [w_1, \cdots, w_n]$$

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & & 0 \\ \delta_2 & \ddots & \ddots & \beta_n \\ & \ddots & \ddots & \\ 0 & & \delta_n & \alpha_n \end{pmatrix} \qquad W^T A = T W^T$$

$$A v_k = v_{k-1} \beta_k + v_k \alpha_k + v_{k+1} \delta_{k+1}$$

$$w_k^T A = w_{k-1}^T \delta_k + w_k^T \alpha_k + w_{k+1}^T \beta_{k+1}$$

$$\begin{cases} w_i^T A v_j = T_{ij} & \text{tridiag.} \\ w_i^T v_j = \delta_{ij} & \text{bi-orthogonality.} \end{cases}$$

$$\tilde{v}_{k+1} = A v_k - v_k \alpha_k - v_{k-1} \beta_k$$

$$\tilde{w}_{k+1}^T = w_k^T A - w_k^T \alpha_k - w_{k-1}^T \delta_k .$$

$$\Rightarrow \quad \tilde{W}_{k+1}^T \tilde{V}_{k+1} = \beta_{k+1} \delta_{k+1} .$$

Pick convention $\delta_{k+1} = \sqrt{|\tilde{w}_{k+1}^T v_{k+1}|}$

$$\beta_{k+1} = \frac{\tilde{W}_k^T \tilde{V}_{k+1}}{\delta_{k+1}} .$$

$$\alpha_{k+1} = W_{k+1}^T A \, V_{k+1} , \qquad V_{k+1} = \tilde{V}_{k+1} / \delta_{k+1}$$

$$W_{k+1}^T = \tilde{W}_{k+1}^T / \beta_{k+1} .$$

BiCG : Bi ortho Lanczos + FOM.

$$A V_m \approx V_m T_m$$

$$x_m = x_0 + V_m y_m. \qquad y_m = T_m^{-1}(\beta e_1)$$

LU factorization

$$T_m = L_m U_m$$

$$x_m = x_0 + \underbrace{V_m U_m^{-1}}_{P_m} L_m^{-1}(\beta e_1)$$

$P_m^T A P_m$ impossible to be diag in general.

BiCG solves an auxiliary system.

$$A^T z = b. \quad \Rightarrow z^T A = b^T$$

$$W_m^T A \approx T_m W_m^T$$

$$z_m^T = z_0^T + \tilde{\beta} e_1^T T_m^{-1} W_m^T$$

$$= z_0^T + \tilde{\beta} e_1^T U_m^{-1} \underbrace{L_m^{-1} W_m^T}_{\tilde{P}_m^T}$$

<u>Prop</u>. $\tilde{P}_m^T A P_m = I$.

Pf: $L_m^{-1} \underbrace{W_m^T A V_m}_{T_m = L_m U_m} U_m^{-1} = L_m^{-1} L_m U_m U_m^{-1} = I$. $\square$

# Quasi-minimal residual (QMR)

$$A V_m = V_m \tilde{T}_m \qquad (m+1) \times m$$

$$x_m = x_0 + V_m y_m .$$

$$y_m = \arg\min_y \| \beta e_1 - \tilde{T}_m y \|_2 .$$

Preconditioner: makes huge diff in practice.

Need detailed knowledge of application & structure of eq.

$$Ax = b.$$

$A$ pos. def. SD. $\#\text{iter} \sim \kappa(A)$

CG $\#\text{iter} \sim \sqrt{\kappa(A)}$

Assume $M$. sym. pos. def. $\begin{cases} M \approx A \\ M^{-1} \text{ easy to apply} \end{cases}$

Try to solve

$$\underbrace{M^{-1}A}_{\tilde{A}} x = \underbrace{M^{-1}b}_{\text{modified rhs. } \tilde{b}} .$$

$$\Rightarrow \tilde{A} x = \tilde{b} .$$ $\kappa(M^{-1}A) \ll \kappa(A)$

$M^{-1}A$ not sym (as a matrix).

But concept of sym. dep. on choice of inner prod.

e.g. inner prod $(x,y)$

operator $B$ sym.

$$(x, By) = (Bx, y), \quad \forall x, y.$$

$M$ pos. def. induces $M$-inner prod.

$$(x,y)_M = x^T M y.$$

$$(x, M^{-1}Ay)_M = x^T M M^{-1} A y = x^T A y.$$

$$\|$$

$$(M^{-1}Ax, y)_M = x^T A^T M^{-1} M y = x^T A y$$

$\Rightarrow$ $M^{-1}A$ sym. (in $M$-inner-prod)!

$\Rightarrow$ precond. SD/CG.

---

Preconditioned GMRES.

- Left preconditioning.

$$M^{-1}Ax = M^{-1}b.$$

- right preconditioning.

$$AM^{-1}u = b, \qquad x = M^{-1}u.$$

- split preconditioning

$$M_1^{-1} A M_2^{-1} u = M_1^{-1} b.$$

common scenario: $M \approx A$

$M = LU$, $M_1 = L$, $M_2 = U$.

Common choices of preconditioners.

- diagonal : $M = \mathrm{diag}(A_{11}^{-1}, \ldots, A_{NN}^{-1})$

- incomplete LU factorization. (ILU)

$$A \approx \hat{L}\hat{U}. \qquad M = \hat{U}^{-1}\hat{L}^{-1}.$$

- $A = A_0 + A_1$. $A_0^{-1}$ easy (e.g. $A_0 = (-\Delta - c)$ for $A = -\Delta + V$. $M = A_0^{-1}$).

dramatic reduction of # iterations.