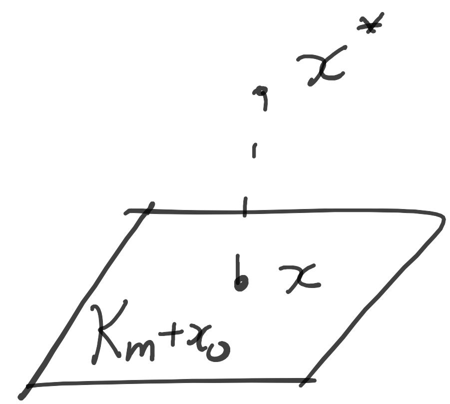
Thm. x_m is the sol of CG at step m,

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:

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

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

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

A-inner prod

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

$$\Rightarrow x_m = \underset{\text{s.t. } x = x_0 + v,\ v \in V_m}{\arg\min} \| x - x^* \|_A = \underset{\substack{\text{s.t. } x = x_0 + q(A) r_0, \\ q \in P_{m-1}}}{\arg\min} \| x - x^* \|_A$$

□.

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.

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

SD : K.   CG : $\sqrt{K}$.

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.

$$T_m = \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \ddots & \beta_3 & \\ & \ddots & \ddots & \beta_m \\ & & \beta_m & \alpha_m \end{pmatrix} = L_m U_m. \leftarrow \text{LU decomp.}$$

(a.k.a. Gauss elimination)

$$L_m = \begin{pmatrix} 1 & & & \\ \lambda_2 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda_m & 1 \end{pmatrix}, \quad U_m = \begin{pmatrix} \eta_1 & \beta_2 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & \beta_m \\ 0 & & & \eta_m \end{pmatrix}$$

bi-diagonal → no-fill-in.

$$\lambda_k \eta_{k-1} = \beta_k \quad \leftarrow T_{k,k-1}$$

$$\lambda_k \beta_k + \eta_k = \alpha_k \quad \leftarrow T_{k,k}$$

$$\beta_k = \beta_k \quad \leftarrow T_{k,k+1}$$

$$\Rightarrow \begin{cases} \lambda_k = \beta_k / \eta_{k-1} \\ \eta_k = \alpha_k - \underbrace{\beta_k^2 / \eta_{k-1}}_{\text{schur complement}} \end{cases}$$

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

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

$$P_{m-1} = [\, p_0 \cdots p_{m-1} \,]$$

Need to show

$$x_{m+1} = x_m + p_m \alpha_m, \qquad p_i^T A p_j = 0 \, , \; i \neq j$$

$$L_{m+1} = \left(\begin{array}{c|c} L_m & 0 \\ \hline \lambda_{m+1} e_m^T & 1 \end{array}\right) \qquad U_{m+1} = \left(\begin{array}{c|c} U_m & \beta_{m+1} e_m \\ \hline 0 & \eta_{m+1} \end{array}\right)$$

$$L_{m+1}^{-1} = \left(\begin{array}{c|c} L_m^{-1} & 0 \\ \hline -\lambda_{m+1} e_m^T L_m^{-1} & 1 \end{array}\right)$$

$$U_{m+1}^{-1} = \left(\begin{array}{c|c} U_m^{-1} & -U_m^{-1} e_m \beta_{m+1} \eta_{m+1}^{-1} \\ \hline & \eta_{m+1}^{-1} \end{array}\right)$$

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

$$P_m = \begin{bmatrix} V_m & v_{m+1} \end{bmatrix} U_{m+1}^{-1}$$

$$= \Big[\underbrace{V_m U_{m-1}^{-1}}_{P_{m-1}} , \; - \underbrace{V_m U_m^{-1} e_m}_{p_{m-1}} \beta_{m+1} \eta_{m+1}^{-1} + U_{m+1} \eta_{m+1}^{-1} \Big]$$

$$= \Big[P_{m-1} , \; \underbrace{p_{m-1} (*) + U_{m+1} (*)}_{p_m} \Big]$$

3 term recursion

$$w_m = L_{m+1}^{-1}\beta e_1 = \begin{bmatrix} L_m^{-1}\beta\tilde{e}_1 \\ -\lambda_{m+1} e_m^T L_m^{-1}\beta\tilde{e}_1 \end{bmatrix} = \begin{bmatrix} w_{m-1} \\ \hline * \end{bmatrix}$$

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

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

$$= (U_m^{-1})^T \underset{\underset{L_m U_m}{\parallel}}{T_m} U_m^{-1}$$

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

Lower triangular.

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

Why this is CG?

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

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

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

residual   conjugate direction.

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

A-orthogonality

x_{m+1} minimizing in $x_0 + K$

$\Big\}$ 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.

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

$V_m^T V_m = I$

T_m tri-diag.

Trade orthogonality w. tri-diagonality?

$\downarrow$

bi-orthogonality.

Lanczos biorthogonalization method

(for non-sym matrices).

$$AV = VD \qquad \text{diagonalizable.}$$

$$\underset{\substack{\| \\ W^T}}{V^{-1}} V = I \qquad W \text{ orthogonal to } V!$$

Keep track of both V, W

$\longrightarrow$ tri-diagonal.

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

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

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

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

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

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

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

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

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

$$\Rightarrow \quad \tilde{w}_{k+1}^T \tilde{v}_{k+1} = \beta_{k+1} \delta_{k+1} .$$ $\leftarrow$ distribute to v_{k+1}, w_{k+1}

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

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

$\alpha_{k+1} = w_{k+1}^T A v_{k+1}$, $\quad v_{k+1} = \tilde{v}_{k+1} / \delta_{k+1}$

$w_{k+1}^T = \tilde{w}_{k+1}^T / \beta_{k+1}$.

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)

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

$$x_m = x_0 + V_m y_m .$$

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

Summary

Solve \ Krylov	Arnoldi	Lanczos	Bi-Lanczos
square $Hy = \beta e_1$	FOM non-sym high mem, A	CG sym. pos. def. low mem, A	Bi-CG non-sym low mem, A, A^T
extended $\inf \lVert \tilde{H} y - \beta \tilde{e}_1 \rVert_2$	GMRES non-sym high mem, A	MINRES sym. indef. low mem, A	QMR non-sym. low mem. A, A^T

Preconditioner: makes huge diff in practice.

Need detailed knowledge of application & structure of eq.

$$Ax = b.$$

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

CG # 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

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

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

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

- 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.