Thm. CG algorithm:

$$\underset{\sim}{x^{(0)} = 0} \quad r^{(0)} = p^{(0)} = b \implies$$

i) $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$

$$\alpha_k = \frac{p^{(k)T} r^{(k)}}{p^{(k)T} A p^{(k)}}$$

ii) $r^{(k+1)} = b - A x^{(k+1)}$

iii) $p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)}$

$$\beta_k = \frac{- r^{(k+1)T} A p^{(k)}}{p^{(k)T} A p^{(k)}}$$

$\implies$ CG converges to the true soln x^* in at most "m" steps

(a) $r^{(k)T} r^{(i)} = 0, \quad \forall i < k$

(b) $p^{(k)T} A p^{(i)} = 0, \quad \forall i < k$

$$\mathcal{K}^{(k)} = \mathrm{Span}\{x^{(1)}, \ldots, x^{(k)}\}$$

$$= \mathrm{Span}\{p^{(0)}, \ldots, p^{(k-1)}\}$$

$$= \mathrm{Span}\{r^{(0)}, \ldots, r^{(k-1)}\}$$

Krylov Space

$= \operatorname{span}\{b, Ab, \dots, A^k b\}$

(c) $e^{(k)} = x^* - x^{(k)}$

$$\|e^{(k)}\|_A = \inf_{u \in K^{(k)}} \|x^* - u\|_A$$

Rmk. $\|e^{(k+1)}\|_A \le \|e^{(k)}\|_A$

pf $K^{(k)} \subseteq K^{(k+1)}$

Note

$a > b > 0$

$$\frac{1}{a+b} < \frac{1}{a} < \frac{1}{b}$$

Rmk. Rate of convergence of CG

$$\|e^{(n)}\|_A \le \frac{2}{(⊛)^n + (⊛)^{-n}} \|e^{(0)}\|_A$$

$$\left(⊛ = \frac{\sqrt{k}+1}{\sqrt{k}-1} \; (>1), \quad k = \operatorname{cond}(A) \right)$$

$$\le \frac{2}{⊛^n} \|e^{(0)}\|_A$$

i) $\circledast \sim 1 \Rightarrow$ slow convergence

(if k is too large)

ii) $\circledast \gg 1 \Rightarrow$ fast convergence

(if k is not too large)

Alg. Basic CG

$x = x_0 \quad (=0)$

$r = b - Ax \quad (= b)$

$p = r \quad (= b)$

do while $\|r\| >$ large

$y = Ap$

$\alpha = \dfrac{p^T r}{p^T y}$

$x = x + \alpha p$

$r = r - \alpha y \quad = (b - Ax) - \alpha Ap$

$= b - A(x + \alpha p)$

$= b - Ax$

Calculate $\|r\|$

$$\beta = -\frac{r^T y}{p^T y}$$

$$p = r + \beta p$$

end do

3 Preconditioning for CG

$$Ax = b$$

Goal: Find M (invertible) s.t

i) $M \sim A$,

ii) M^{-1} is easy.

iii) $M^{-1}Ax = M^{-1}b$

vi) We hope: $k(M^{-1}A) \ll k(A)$

Rmk: M is called a <u>preconditioner</u>

Ex i) $M = A$ $\boxed{M^{-1}Ax = M^{-1}b}$

$\Downarrow$ $\Updownarrow$

$k(M^{-1}A) = k(I) = 1$, $Ix = A^{-1}b$

$\longrightarrow$ useless!

ii) $M = I$, $M^{-1} = I$ (easy!)

$\boxed{M^{-1}Ax = M^{-1}b}$

$\Longleftrightarrow \boxed{IAx = Ib}$

(useless!)

iii) $M \sim A$ & M^{-1} : easy

What we want

$\rightarrow$ $M = \mathrm{diag}(A)$ if A is diagonally dominant.

$M^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & & 0 \\ & \ddots & \\ & & \end{bmatrix}$

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$$

$$\begin{bmatrix} \ddots & \\ 0 & \frac{1}{a_{mm}} \end{bmatrix}$$

(Jacobi preconditioner)

This works well for A: diagonally dominant

Rmk. Not all $M = \mathrm{diag}(A)$ works all the times

Ex Consider A : $\begin{cases} a_{ii} = D, & \forall i. \\ a_{ij} = 1, & \forall i \neq j \end{cases}$

For example $\underset{\sim 3\times 3}{A} = \begin{bmatrix} D & 1 & 1 \\ 1 & D & 1 \\ 1 & 1 & D \end{bmatrix}$

$$\rightarrow M^{-1} = \begin{bmatrix} \frac{1}{D} & & \\ & \frac{1}{D} & \\ & & \frac{1}{D} \end{bmatrix}$$

$$\rightarrow \quad M^{-1}A = \begin{bmatrix} 1 & \frac{1}{D} & \frac{1}{D} \\ \frac{1}{D} & 1 & \frac{1}{D} \\ \frac{1}{D} & \frac{1}{D} & 1 \end{bmatrix} = \frac{1}{D} A$$

$\rightarrow$ $\text{cond}(M^{-1}A) = \text{cond}(\frac{1}{D}A) = \text{cond}(A)$!

$\rightarrow$ Not good!

$\rightarrow$ If $a_{ii} \neq$ const, then

$M = \text{diag}(A)$ would work!

Q $M^{-1}A$ is NOT spd even if A is spd!

A Find C s.t $M = CC^T$
$\uparrow$ (invertible)

$\rightarrow$ $M^{-1} = C^{-T}C^{-1}$

$\rightarrow$ $M^{-1}Ax = \cancel{C^{-T}}C^{-1}Ax$

$\|$

$$M^{-1}b = \cancel{C}^{T} C^{-1} b$$

$\rightarrow$ $\boxed{C^{-1}AX = C^{-1}b}$

$\rightarrow$ $\underbrace{(C^{-1}AC^{-T})}_{=\tilde{A}} \underbrace{(C^{T}x)}_{\tilde{x}} = \underbrace{C^{-1}b}_{=\tilde{b}}$

$\rightarrow$ $\boxed{\tilde{A}\tilde{x} = \tilde{b}}$, when $\tilde{A}$ is spd

pf i) $\tilde{A}$ is symm

$$\tilde{A}^{T} = (C^{-1}AC^{-T})^{T}$$

$$= C^{-1}AC^{-T} = \tilde{A}$$

ii) $\tilde{A}$ is pd. :

$$u^{T}\tilde{A}u = (u^{T}C^{-1})A\underbrace{(C^{-T}u)}_{v}$$

$$= \underbrace{v^{T}Av}\ \underline{> 0}, \text{ plus}$$

$\left(v \neq 0, \text{ since } C^{-T} \text{ is} \right)$

invertible

∴ $\tilde{A}$: spd

$\longrightarrow$ PCG

(preconditioned — CG)

Rmk. "Exit condition"

i) $\|r^{(k)}\| >$ large $\Leftarrow$ one good way

or

ii) $\alpha_k \sim 0 \Rightarrow$ Converged!

$\boxed{4}$ GMRES (Generalized Minimal RESidual Method)

$A \in \mathbb{R}^{m \times m}$ $(n < m)$

i) Search space $= K^{(n)}$

ii) works for non-spd.

iii) Create $K^{(n)} \subset \mathbb{R}^m$, n: fixed

find x_n s.t. $\| b - Ax_n \| = \inf_{u \in K^{(n)}} \| b - Au \|$

§4.1. Krylov subspace

Def. $\mathcal{K}_n(A, v)$: Krylov subspace of A w.r.t. v

$\Big(\Big($ def

$$\operatorname{span}\{v, Av, A^2 v, \ldots, A^{n-1} v\}$$

$(\subset \mathbb{R}^m)$

Def. K_n : $m \times n$ krylov matrix

((def

$$\begin{bmatrix} v & Av & \cdots & A^{n-1}v \end{bmatrix}$$

$$\equiv \begin{bmatrix} k_1 & k_2 & \cdots & k_n \end{bmatrix}$$

<u>Rmk</u>. $\mathcal{K}_n(A,v) = \mathrm{span}\{k_1, \ldots, k_n\}$

$= \mathrm{Span}\{v, Av, \ldots A^{n-1}v\}$.

<u>Rmk</u>. Goal:

$$\min \|b - Ax_n\|_2 \quad \text{where } x_n \in \mathcal{K}_n(A,v)$$

$\to \exists\, c \in \mathbb{R}^n$ s.t

$$x_n = c_1 k_1 + \cdots + c_n k_n, \quad c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$= K_n c$$

$$\to \min_{x_n \in \mathcal{K}_n(A,v)} \|b - Ax_n\| \iff \min_{c \in \mathbb{R}^n} \|b - \overbrace{\underbrace{AK_n}_{m\times n} \underbrace{c}_{n\times 1}}\|$$

$\rightarrow$ $\underbrace{\underset{\sim}{A}}_{m\times m}\underbrace{\underset{\sim}{K_n}}_{m\times n}$: $m \times n$ $(m > n)$

over-determined system!

$\rightarrow$ $\hat{Q}\hat{R} = AK_n$ (reduced QR)

$\rightarrow$ $P_{AK_n} = \hat{Q}\hat{Q}^T$

$\rightarrow$ $AK_n c = \underbrace{\hat{Q}\hat{Q}^T}_{P_{AK_n}} b$

$\rightarrow$ $\underbrace{\hat{Q}^T\hat{Q}}_{= I}\hat{R}c = \underbrace{\hat{Q}^T\hat{Q}}_{= I}\hat{Q}^T b$

$\rightarrow$ $\boxed{\hat{R}c = \hat{Q}^T b}$

Def. $A: m \times n$

$\rightarrow$ A minimal poly of A is "p" uniquely defined as a monic poly with the least possible degree ($\alpha_n = 1$) satisfying $p(A)v = \underset{\sim}{0}$, $\forall v \in \mathbb{R}^m$

i.e, $p(A)v = \left(\sum_{i=0}^{n} \alpha_i A^i\right) v$

$$= (\alpha_0 I + \alpha_1 A + \cdots + A^n) v$$

$$= \underset{\sim}{0}$$

Rmk. Why $\mathcal{K}_n(A, b)$ is a good choice?

(A) $0 = p(A) b = (\alpha_0 I + \cdots + A^n) b$

$\rightarrow$ $$I = -\frac{1}{\alpha_0} \left(\sum_{i=1}^{n} \alpha_i A\right) b$$

$$\rightarrow \quad \underbrace{A^{-1} b}_{\parallel} = -\frac{1}{\alpha_0} \left(\sum_{i=0}^{n-1} \alpha_{i+1} A^i \right) b$$

$$\rightarrow \quad x = A^{-1} b = -\frac{1}{\alpha_0} \sum_{i=0}^{n-1} \alpha_{i+1} A^i b$$