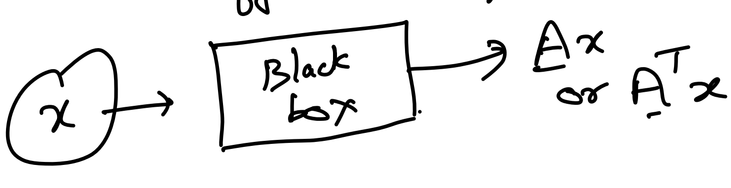
(1) Direct methods require $O(m^3)$ work

→ You get the answer in finite number of steps!

(2) Iterative method require infinite number of iterations to get the exact answer but for practical purposes we can terminate the iterations using certain tolerance criteria. This reduces computational cost specifically when dealing with sparse matrices arising for discretization of partial different equations!



Krylov Subspace methods:-

Given a matrix $\underset{\sim}{A}$ and a vector $\underset{\sim}{b}$, the associated Krylov sequence is the set of vectors $\underset{\sim}{b}, \underset{\sim}{A}\underset{\sim}{b}, \underset{\sim}{A}^2\underset{\sim}{b}, \underset{\sim}{A}^3\underset{\sim}{b}, \ldots$

Heart of most iterative methods rely on projection into Krylov subspace!

Arnoldi Iteration:-

In phase I, we saw the reduction of $\underset{\sim}{A} \in \mathbb{R}^{m\times m}$ to upper Hessenberg form by choosing reflectors.

$$\boxed{\underset{\sim}{A} = \underset{\sim}{Q}\,\underset{\sim}{H}\,\underset{\sim}{Q}^T}$$

$$\Rightarrow \boxed{\underset{\sim}{A}\,\underset{\sim}{Q} = \underset{\sim}{Q}\,\underset{\sim}{H}} \quad - (1)$$

Now we will consider "n" columns of

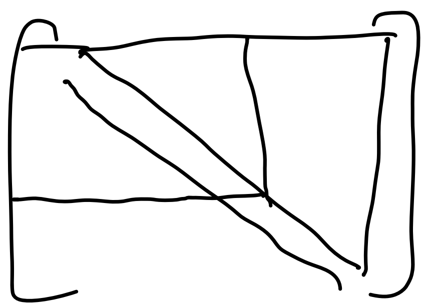
$$AQ = QH$$

Let $\hat{Q}_n$ be the $m \times n$ matrix whose columns are the first 'n' columns of Q:

$$\hat{Q}_n = \underbrace{\left[q_1 \,\middle|\, q_2 \,\middle|\, \cdots \,\middle|\, q_n \right]}$$

Let $\tilde{H}_n$ be the $(n+1) \times n$ upper left section of H is also Hessenberg matrix

$$\tilde{H}_n = \begin{bmatrix} h_{11} & \cdots & \cdots & h_{1n} \\ h_{21} & h_{22} & & \vdots \\ & \ddots & \ddots & \vdots \\ & & h_{n,n-1} & h_{nn} \\ & & & h_{n+1,n} \end{bmatrix}$$



$$\boxed{\underbrace{AQ = QH}}$$ $\hat{Q}_n$; $\tilde{H}_n$

Then we have

$$\boxed{A\hat{Q}_n = \hat{Q}_{n+1}\tilde{H}_n} \quad (2)$$

The n^{th} column of Equation (2) can be written as follows:-

$$\boxed{A q_n = h_{1n} q_1 + h_{2n} q_2 + \cdots h_{n+1,n} q_{n+1}} \quad — (3)$$

q_{n+1} satisfies $(n+1)$-term recurrence relation involves itself and previous vectors $q_1, q_2, \ldots q_n$.

Arnoldi iteration is essentially iterative procedure which implements (1).

Algorithm:-

$$b = \text{arbitrary}$$

$$q_1 = \frac{b}{\|b\|}$$

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

$$\boxed{v = A q_1}$$

for $j = 1$ to n

$$h_{jn} = \underset{\sim}{q}_j^T \underset{\sim}{v}$$

$$\boxed{\begin{aligned}\underset{\sim}{A}\underset{\sim}{q}_1 &= h_{11}\underset{\sim}{q}_1 \\ &+ h_{21}\underset{\sim}{q}_2\end{aligned}} \to (*)$$

$$\underset{\sim}{v} = \underset{\sim}{v} - h_{jn}\underset{\sim}{q}_j$$

$$h_{n+1,n} = \|v\|$$

$$\underset{\sim}{q}_{n+1} = \frac{v}{h_{n+1,n}}$$

$$K_n = \langle \underset{\sim}{q}_1, \underset{\sim}{q}_2, \ldots \underset{\sim}{q}_n \rangle = \langle \underset{\sim}{b}, \underset{\sim}{A}\underset{\sim}{b}, \underset{\sim}{A}^2\underset{\sim}{b}, \ldots \underset{\sim}{A}^{n-1}\underset{\sim}{b} \rangle$$

Intuitive:-
proof

$$\underset{\sim}{q}_1 = \frac{\underset{\sim}{b}}{\|\underset{\sim}{b}\|}$$

$$\underset{\sim}{p} = \beta_1 \underset{\sim}{q}_1 + \beta_2 \underset{\sim}{q}_2$$

$$= \beta_1 \frac{\underset{\sim}{b}}{\|\underset{\sim}{b}\|} + \frac{\beta_2}{h_{21}}\left(\underset{\sim}{A}\underset{\sim}{q}_1 - h_{11}\underset{\sim}{q}_1\right)$$

$$= \beta_1 \frac{\underset{\sim}{b}}{\|\underset{\sim}{b}\|} + \frac{\beta_2}{h_{21}}\left(\underset{\sim}{A}\frac{\underset{\sim}{b}}{\|\underset{\sim}{b}\|} - h_{11}\frac{\underset{\sim}{b}}{\|\underset{\sim}{b}\|}\right)$$

$$= \underset{\sim}{b}\underbrace{\left(\frac{\beta_1}{\|\underset{\sim}{b}\|} - \frac{\beta_2 h_{11}}{h_{21}\|\underset{\sim}{b}\|}\right)} + \underset{\sim}{A}\underset{\sim}{b}\left(\frac{\beta_2}{h_{21}\|\underset{\sim}{b}\|}\right)$$

$$= \underset{\sim}{b}\,\alpha_1 + \underset{\sim}{A}\underset{\sim}{b}\,\alpha_2 = \alpha_1 \underset{\sim}{b} + \alpha_2 \underset{\sim}{A}\underset{\sim}{b}$$

$\underline{\underline{n=3}}$

$$\bar{P} = \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_3$$
$$= \alpha_1 b + \alpha_2 A b + \alpha_3 A^2 b$$

$$K_n = \hat{Q}_n \hat{R}_n$$

$$K_n = \begin{pmatrix} | & | & | & & | \\ b & Ab & A^2 b & \cdots & A^{n-1} b \\ | & | & | & & | \end{pmatrix}$$

Projection onto Krylov Subspace :-

$$x \rightarrow \boxed{\begin{matrix}\text{Black box} \\ Ax,\ A^T x\end{matrix}} \rightarrow Ax$$
$$A \rightarrow$$

$$A Q_n = Q_{n+1} \tilde{H}_n$$

$$\underset{n \times m}{\hat{Q}_n^T}\ \underset{m \times (n+1)}{\hat{Q}_{n+1}} = \begin{bmatrix} 1 & \cdots & & & 0 \\ & 1 & & & 0 \\ & & 1 & & 0 \\ 0 & & & 1 & 0 \\ & & & 1 & 0 \end{bmatrix}_{m \times (n+1)}$$

$$\tilde{H}_n = \begin{bmatrix} h_{11} & \cdots & \cdots & h_{1n} \\ h_{21} & h_{22} & & \vdots \\ & \ddots & \ddots & h_{nn} \\ & & h_{n\,n-1} & \\ & & & h_{n+1\,n} \end{bmatrix}$$

$$\boxed{\hat{Q}_n^T \hat{Q}_{n+1} \tilde{H}_n = H_n}$$

H_n is actually $n \times n$ Hessenberg matrix obtained by removing the last row of $\tilde{H}_n$

$$H_n = \hat{Q}_n^T \hat{Q}_{n+1} \tilde{H}_n$$

$$A \hat{Q}_n = \hat{Q}_{n+1} \tilde{H}_n$$

$$\hat{Q}_n^T A \hat{Q}_n = \hat{Q}_n^T \hat{Q}_{n+1} \tilde{H}_n = H_n$$

$$\boxed{\underset{\sim}{H}_n = \hat{\underset{\sim}{Q}}_n^T \underset{\sim}{A} \hat{\underset{\sim}{Q}}_n}$$ ✓

This matrix $\underset{\sim}{H}_n$ can be interpreted as the representation of the orthogonal projection of matrix $\underset{\sim}{A}$ onto the subspace spanned by the basis $\{\underset{\sim}{q}_1, \underset{\sim}{q}_2, \ldots \underset{\sim}{q}_n\}$ in terms of the same basis vectors $\{\underset{\sim}{q}_1, \underset{\sim}{q}_2, \ldots \underset{\sim}{q}_n\}$

$$\underset{\sim}{A}_p = \underset{\sim}{P}\,\underset{\sim}{A} \qquad \underset{\sim}{P} = \hat{Q}_n \hat{Q}_n^T$$

$$\underset{\substack{\boxed{m\times m}}}{\underset{\sim}{A}_p} = \underset{\boxed{m\times n}}{\hat{Q}_n}\,\underset{\boxed{n\times m}}{\hat{Q}_n^T}\,\underset{\boxed{m\times m}}{\underset{\sim}{A}} \quad \text{—} (*) \qquad \hat{Q}_n = [\underset{\sim}{q}_1 | \underset{\sim}{q}_2 \ldots | \underset{\sim}{q}_n]$$

We have to express $\underset{\sim}{A}p$ in the basis of $K_n = \left[\underset{\sim}{q}_1 \middle| \underset{\sim}{q}_2 \middle| \cdots \middle| \underset{\sim}{q}_n \right]$

i.e $\quad \tilde{A}_{p_{ij}} = \underset{\sim}{q}_i^T \underset{\sim}{A}p \underset{\sim}{q}_j$

$$\Rightarrow \underset{\sim}{\tilde{A}}p = \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{A}p \underset{\sim}{\hat{Q}}_n$$

$$n \times n \qquad n \times m \quad m \times m \quad m \times n$$

$$\Rightarrow \underset{\sim}{\tilde{A}}p = \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{\hat{Q}}_n \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{A} \underset{\sim}{\hat{Q}}_n$$

$$\boxed{\underset{\sim}{\tilde{A}}p = \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{A} \underset{\sim}{\hat{Q}}_n}$$

$$\boxed{\underset{\sim}{H}_n = \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{A} \underset{\sim}{\hat{Q}}_n}$$

Since $\underset{\sim}{H}_n$ is a projection of $\underset{\sim}{A}$

onto K_n, there is a possibility that eigenvalues of $\underset{\sim}{H_n}$ is related to $\underset{\sim}{A}$ in a useful fashion!

$$\{\theta_j\} = \{\text{eigenvalues of } H_n\}$$

are called Arnoldi eigenvalue estimates (at step n) or Ritz values (with respect to K_n)

__Thm:-__ The matrices $\hat{Q}_n$ generated by the Arnoldi iteration are reduced QR factors of the Krylov matrix

$$\underset{\sim}{K}_n = \underset{\sim}{\hat{Q}}_n \underset{-}{\hat{R}}_n$$

The Hessenberg matrices $\underset{\sim}{H}_n$ are the corresponding projections

$$\underset{\sim}{H}_n = \underset{\sim}{\hat{Q}}_n^T \underset{\sim}{A} \underset{\sim}{\hat{Q}}_n$$ and successive

iterates are related by the formula $\underset{-}{A} \hat{Q}_n = \underset{\sim}{\hat{Q}}_{n+1} \underset{-}{\tilde{H}}_n$

How do we compute eigenvalues of $\underset{-}{A}$ using Arnoldi iteration?

Project $A^{m \times m} \longrightarrow K_n$, you get projected matrix $\underset{\sim}{H}_n = \underset{\sim}{\hat{Q}}_n^T \underset{-}{A} \underset{-}{\hat{Q}}_n$

you can find eigenvalues of $\underset{\sim}{H}_n$ by QR algorithm

<u>Arnoldi iteration as a polynomial approximation</u> :-

If $\underset{\sim}{x} \in K_n$, $K_n = \langle \underset{\sim}{b}, \underset{\sim}{A}\underset{\sim}{b}, \ldots A^{n-1}\underset{\sim}{b} \rangle$

$$\underset{\sim}{x} = c_0\underset{\sim}{b} + c_1\underset{\sim}{A}\underset{\sim}{b} + c_2\underset{\sim}{A}^2\underset{\sim}{b} + \ldots + c_{n-1}\underset{\sim}{A}^{n-1}\underset{\sim}{b} \quad -(4)$$

Introduce $q(z) = c_0 + c_1 z + c_2 z^2 + \ldots c_{n-1} z^{n-1}$, then

eqn (4) can be written as

$$\boxed{\underset{\sim}{x} = q(\underset{\sim}{A})\underset{\sim}{b}} \quad -(5)$$

Define $P^n = \{$monic polynomials of degree $n\}$

→ monic polynomials are those polynomials whose coefficient of Z^n is 1 (coefficient of term of degree n is 1)

<u>Arnoldi approximation problem</u> :-

Find $p^n \in P^n$ such that

$\boxed{\|p^n(\underset{\sim}{A})\underline{b}\|}$ is minimum.

<u>Thm</u> :- Let $\underline{A} \in \mathbb{R}^{m\times m}$, $\underline{b} \in \mathbb{R}^m$ and P^n denotes the space of monic polynomial of degree n. If $K_n = \langle \underline{b}, \underline{A}\underline{b}, \ldots A^{n-1}\underline{b} \rangle$ is of full rank, the solution of

$\boxed{\min_{p^n \in P^n} \|p^n(\underline{A})\underline{b}\|}$ is the characteristic polynomial of $\underset{\sim}{H_n}$

$p_{H^n}(z) = \det(z\underset{\sim}{I} - \underset{\sim}{H_n})$ where

$\underset{\sim}{H_n} = \underset{\sim}{\hat{Q}_n^T} \underset{\sim}{A} \underset{\sim}{\hat{Q}_n}$ with

$\hat{Q}_n = (\underset{\sim}{q_1}, \underset{\sim}{q_2} \ldots q_n)$ constructed

by Arnoldi iteration i.e

$$\boxed{p_{H_n} = \arg \min_{p^n \in P^n} \| p^n(A) \underset{\sim}{b} \|}$$

* The goal of Arnoldi iteration is to solve a polynomial approximation problem.

* If one's aim is to compute a monic polynomial p^n with the

property $\|p^n(\underline{A})\underline{b}\|$ is small, we need to pick p^n such a way that it has zeros close to eigenvalues of $\underline{A}$.

GMRES

$\longrightarrow$ $$A x = \underline{b} \quad (1)$$ $A \in \mathbb{R}^{m \times m}$, $\underline{x} \in \mathbb{R}^m$, $\underline{b} \in \mathbb{R}^m$

$\longrightarrow$ A is nonsingular and our goal is solve eqn (1) and let x^* be solution which is exact $\boxed{\underline{x}^* = \underline{A}^{-1}\underline{b}}$

$\longrightarrow$ Key idea in GMRES:-

At every step n, we approximate

x^* by a vector $\underline{x}_n \in K_n$ the minimizes the norm of the residual $\underline{r}_n = \underline{b} - \underline{A}\,\underline{x}_n$

In other words we want to solve the problem $\min_{z \in K_n} \|\underline{A}\,\underline{z} - \underline{b}\|_2$

Why Krylov subspace for GMRES :-

→ A matrix satisfies its characteristic polynomial i.e if $A \in \mathbb{R}^{m \times m}$. $\det(A - \lambda I) = p(\lambda)$

if $p(\lambda) = \lambda^m + c_{m-1}\lambda^{m-1} + \dots c_1\lambda + c_0$

Then $p(A) = 0$

i.e $\boxed{A^m + c_{m-1} A^{m-1} + \dots + c_1 A + c_0 I_n = 0}$

$c_0 \neq 0$

as A is nonsingular

and $(\det A) \neq 0$

Multiply with b

we get

$A^m b + c_{m-1} A^{m-1} b + \dots + c_1 A b + c_0 b = 0$

Multiply with A^{-1} we get

$A^{m-1} b + c_{m-1} A^{m-2} b + \dots c_1 b = -c_0 A^{-1} b$

$x^* = A^{-1} b = -\frac{1}{c_0} \left[c_1 b + c_2 A b + c_3 A^2 b + \dots + A^{m-1} b \right]$

This Krylov subspace

$K_m = \langle b, Ab, A^2 b, \dots A^{m-1} b \rangle$

is a good subspace to search for solution of $\underset{\sim}{A}^{-1}\underset{\sim}{b}$ or in otherwords approximate $\underset{\sim}{x}^{*}$ to lie in the Krylov subspace of dimension $n < m$ and GMRES exactly does this!

→ GMRES tries to solve the problem $\min_{z \in K_n} \|\underset{\sim}{A}\underset{\sim}{z} - \underset{\sim}{b}\|_2$ since $\underset{\sim}{z} \in K_n$, let trial $z = \hat{Q}_n y$ where columnspace $\underset{\sim}{\hat{Q}}_n$ span K_n

$$\hat{Q}_n = [\underset{\sim}{q}_1 | \underset{\sim}{q}_2 \ldots | \underset{\sim}{q}_n]$$

→ The above minimization problem

can be posed as

$$\min_{y \in \mathbb{R}^n} \| A \hat{Q}_n y - b \|_2$$

Now from Arnoldi iteration

$$\boxed{A \hat{Q}_n = \hat{Q}_{n+1} \tilde{H}_n}$$

$$\min_{y \in \mathbb{R}^n} \| A \hat{Q}_n y - b \|_2 = \min_{y \in \mathbb{R}^n} \| \hat{Q}_{n+1} \tilde{H}_n y - b \|_2$$

$$= \min_{y \in \mathbb{R}^n} \| \hat{Q}_{n+1}^T (\hat{Q}_{n+1} \tilde{H}_n y - b) \|_2$$

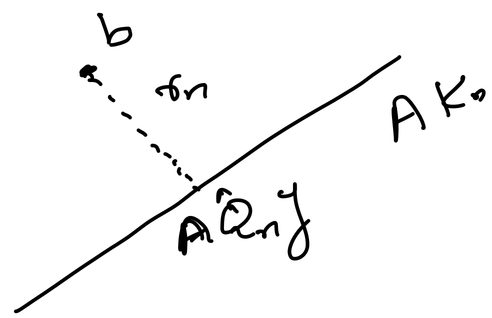
$$= \min_{y \in \mathbb{R}^n} \| \underbrace{\hat{Q}_{n+1}^T \hat{Q}_{n+1}}_{n+1 \times m \quad m \times n+1} \tilde{H}_n y - \hat{Q}_{n+1}^T b \|_2$$

$$= \min_{y \in \mathbb{R}^n} \| \tilde{H}_n y - \hat{Q}_{n+1}^T b \|_2$$

$$= \boxed{\min_{y \in \mathbb{R}^n} \| \tilde{H}_n y - \|b\| e_1 \|_2}$$

This $(n+1) \times n$ least squares problem

At each step n of GMRES we solve the least square problem of $(n+1) \times n$ and $x_n = Q_n y$



Algo:-

$$q_1 = \frac{b}{\|b\|}$$

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

< step n of Arnoldi iteration >

Find y to minimize $\|\tilde{H}_n y - \|b\| e_1\|$

$$x_n = Q_n y$$

Polynomial approximation of GMRES

$P_n = \{$ polynomials p of degree $\leq n$ with $p(0) = 1\}$

$x_n \in K_n$

$$x_n = c_0 \underline{b} + c_1 \underline{A}\,\underline{b} + c_2 \underline{A}^2 \underline{b} + \dots c_{n-1} \underline{A}^{n-1} \underline{b}$$

$$q_n(z) = c_0 + c_1 z + \dots c_{n-1} z^{n-1}$$

$$\underline{x}_n = \underbrace{q_n(\underline{A})}\,\underline{b}$$

$$\underline{r}_n = \underline{b} - \underline{A}\,\underline{x}_n$$

$$= \underline{b} - \underline{A}\, q_n(A) b$$

$$\underline{r}_n = [\underline{I} - \underline{A}\, q_n(\underline{A})]\,\underline{b}$$

If p_n is the polynomial

$$\boxed{p_n(z) = 1 - z q(z)}$$

GMRES approximation problem:

Find $p_n \in P_n$ such that $\| p_n(\underline{A}) b \|$ is minimum

Convergence of GMRES

What properties of A determine the size of $\|r_n\|$?

→ GMRES converges monotonically

$\|r_{n+1}\| \leq \|r_n\|$ $\quad K_n \subseteq K_{n+1}$

→ At most m steps are required for the process to converge

$\|r_m\| = 0$

→ From a practical view point I need $n \ll m$ for it to be useful method

$$\underbrace{\|r_n\| = \|p_n(A)b\|} \leq \underbrace{\|p_n(A)\|}_{\text{is minimal}}\|b\|$$

$$\frac{\|r_n\|}{\|b\|} \le \min_{p_n \in P^n} \|p_n(A)\|$$

Given $A \in \mathbb{R}^{m \times m}$ and step n of GMRES iteration how small $\|p_n(A)\| b$?

Let us assume $\underset{\sim}{A}$ is diagonalizable satisfying $\underset{\sim}{A} = \underline{V}\underline{\Lambda}\underline{V}^{-1}$ for some non singular matrix $\underset{\sim}{V}$ and diagonal matrix Λ. Then

$$\|p_n(A)\| \le \|V\|\|p(\Lambda)\|\|V^{-1}\|$$

$$= \underline{K}(V)\|p(\Lambda)\|$$

$$\underbrace{\|p_n(A)\| \le K(V) \max_{\lambda \in \sigma(A)} |p_n(\lambda)|}$$

Thm:- At step n of GMRES iteration, the residual $\underline{r}_n$ satisfies $\frac{\|\underline{r}_n\|}{\|\underline{b}\|} \le \min_{p_n \in P^n} \|p_n(A)\|$

$$\le \underbrace{\underline{K(V)} \left[\min_{p_n \in P^n} \left(\max_{\lambda \in \underline{\sigma(A)}} |p_n(\lambda)| \right) \right]}$$

$AA^T = A^T A$

$$\|\underline{r}_n\| \le \left(\frac{(K(A))^2 - 1}{(K(A))^2} \right)^{n/2}$$ for $\underline{A} \in \mathbb{R}^{m \times m}$ being symmetric

Conjugate Gradient

$\underline{A} \in \mathbb{R}^{m \times m}$ (Symmetric and positive definite matrices)

$\rightarrow \underline{A}\underline{x} = \underline{b}$

$\rightarrow$ $K_n = \langle \underline{b}, \underline{A}\underline{b}, \ldots \underline{A}^{n-1}\underline{b} \rangle$

$\rightarrow$ $\|\cdot\|_A$ defined by $\|\underline{x}\|_A = \sqrt{\underline{x}^T \underline{A}\, \underline{x}}$

is a norm on $\mathbb{R}^m$ and this is called A-norm.

$\rightarrow$ Conjugate Gradient method is a set of recurrence formulae which generate sequence of iterates $\{x_n \in K_n\}$ with the property that at step n, $\boxed{\|\underline{e}_n\|_A}$ is minimized where $\underbrace{\underline{e}_n = x^* - x_n}$

Algo for CG iterations

$x_0 = 0, \; r_0 = b, \; p_0 = r_0$

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

$$\alpha_n = \frac{r_{n-1}^T r_{n-1}}{p_{n-1}^T A \, p_{n-1}} \quad \text{(step length)}$$

$$x_n = x_{n-1} + \alpha_n p_{n-1} \quad \text{(update of approximate solution)}$$

$$r_n = r_{n-1} - \alpha_n A \, p_{n-1} \quad \text{(residual)}$$

$$\beta_n = \frac{r_n^T r_n}{r_{n-1}^T r_{n-1}} \quad \text{(improvement of search direction)}$$

$$p_n = r_n + \beta_n p_{n-1} \quad \text{(search direction)}$$

→ A is dense and unstructured $\approx O(m^2)$ ✓

→ A is sparse $\approx O(m)$ ✓

Thm (1):-

Let $A \in \mathbb{R}^{m\times m}$ be a symmetric and a positive-definite matrix. Let $Ax=b$ be solved by applying CG iteration. As long as the iteration is not converged (i.e $r_{n-1} \neq 0$), the algo proceeds satisfying the following identities of subspaces :→

$$K_n = \langle x_1, x_2, x_3, \dots x_n \rangle$$

$$= \langle p_0, p_1, \dots p_{n-1} \rangle$$

$$= \langle r_0, r_1, \dots r_{n-1} \rangle$$

$$= \underbrace{\langle b, Ab, \dots A^{n-1} b \rangle}$$

K_4 r_3

r_2

r_0 r_1 Ax_4

$-b$

$r_4 \perp r_0$ =?

$r_4 \perp r_1$ =?

$r_4 \perp r_i$

r_4

Moreover the residuals are orthogonal $\boxed{r_n^T r_j = 0}$ $(j<n)$

and search directions are A-conjugate

$$\boxed{p_n^T A p_j = 0} \quad (j<n)$$

Thm 2:-

Let CG iteration be applied to a symmetric positive definite matrix $\underline{A} \in \mathbb{R}^{m \times m}$ $\underline{A}\underline{x} = \underline{b}$. If the iteration is not fully converged, then $\underline{x}_n$ is the unique point in K_n that minimizes $\|\underline{e}_n\|_A$. The convergence is monotonic

$$\|e_n\|_A \leq \|e_{n-1}\|_A$$

Pf:- We know that $\underset{\sim}{x}_n \in K_n$. We need to show that it is the unique point which minimizes $\|e\|_A$.

Consider an arbitrary point $\underset{\sim}{x}$

$\underset{\sim}{x} \in K_n$ and let $\underset{\sim}{x} = \underset{\sim}{x}_n + \Delta \underset{\sim}{x}$ where $\Delta \underset{\sim}{x} \neq \underset{\sim}{0}$

Note $\Delta x \in K_n$.

$\underset{\sim}{e} = \underset{\sim}{x}^* - \underset{\sim}{x}$ where x^* is the exact solution for $A\underset{\sim}{x} = \underset{\sim}{b}$

$$\underset{\sim}{e} = \underset{\sim}{x}^* - (\underset{\sim}{x}_n + \Delta \underset{\sim}{x})$$

$$= \underset{\sim}{x}^* - \underset{\sim}{x}_n - \Delta \underset{\sim}{x} = \underset{\sim}{e}_n - \Delta \underset{\sim}{x}$$

$$\|\underset{\sim}{e}\|_A^2 = (\underset{\sim}{e}_n - \Delta \underset{\sim}{x})^T A (\underset{\sim}{e}_n - \Delta \underset{\sim}{x})$$

$$= \underset{\sim}{e}_n^T A \underset{\sim}{e}_n + \Delta \underset{\sim}{x}^T A \Delta \underset{\sim}{x} - 2 \underset{\sim}{e}_n^T A \Delta \underset{\sim}{x}$$

Let us consider this term $\underset{\sim}{e}_n^T A \Delta \underset{\sim}{x}$

$$\underset{\sim}{e}_n^T A \Delta \underset{\sim}{x} = (\underset{\sim}{x}^* - \underset{\sim}{x}_n)^T A \Delta \underset{\sim}{x}$$

$$= \underset{\sim}{x}^{*T} A \Delta \underset{\sim}{x} - \underset{\sim}{x}_n^T A \Delta \underset{\sim}{x} \quad — (1)$$

$$A x^* = b \Rightarrow (x^*)^T A = b^T \quad - (2)$$

$$r_n = b - A x_n$$

$$\Rightarrow A x_n = b - r_n$$

$$\Rightarrow x_n^T A = b^T - r_n^T \quad - (3)$$

using eqn (2) and (3) in (1)

$$e_n^T A \Delta x = b^T \Delta x - (b^T - r_n^T) \Delta x$$

$$= r_n^T \Delta x$$

Δx is a vector in K_n and we know CG generates sequence of iterates such that r_n is perpendicular Krylov subspace K_n, $r_n^T \Delta x = 0$

$$\|e\|_A^2 = e_n^T A e_n + (\Delta x)^T A (\Delta x) \quad - (4)$$

Since A is positive definite, $e_n^T A e_n \geq 0$ and $(\Delta x)^T A (\Delta x) > 0$

$\|e\|_A^2$ will be minimum if and only if $(\Delta x)^T A \Delta x = 0 \Rightarrow \Delta x = 0$

i.e $x = x_n$

CG as an optimization Algo :-

$$x_n = x_{n-1} + \alpha_n p_{n-1}$$

$$\|e_n\|_A^2 = e_n^T A e_n$$

$$= (x^* - x_n)^T A (x^* - x_n)$$

$$= x_n^T A x_n - 2 x_n^T A x^* + x^{*T} A x^*$$

$$= \underbrace{x_n^T A x_n - 2 x_n^T b} + x^{*T} A x^*$$

$$= 2\phi(x_n) + \text{Constant}$$

$$\underbrace{\phi(x) = \frac{1}{2} x^T A x - x^T b}$$

CG iteration can be interpreted as an iterative process for minimizing quadratic function $\phi(\underset{\sim}{x})$ of $\underset{\sim}{x} \in \mathbb{R}^m$

<u>CG</u> <u>as</u> a <u>polynomial</u> <u>app</u>roximation

$P_n = \{$ polynomial p of degree $\leq n$ with $p(0) = 1\}$

Find $p_n \in P_n$ such that $\|p_n(A)e_0\|_A =$ minimum where $e_0 = x^* - x_0 = x^*$

<u>Thm</u>:- If the CG iteration has already not converged before step n then the solution of $\min_{p_n} \|p_n(A)e_0\|$

is unique and the iterate $\underset{\sim}{x}_n$ has error $\underset{\sim}{e}_n = \underset{\sim}{x}^* - \underset{\sim}{x}_n = p_n(A)\, e_0$ for this minimum polynomial p_n and then

$$\frac{\|e_n\|_A}{\|e_0\|_A} = \min_{p \in P_n} \frac{\|p(A)\, e_0\|_A}{\|e_0\|_A}$$

$$\leq \underbrace{\min_{p \in P_n} \max_{\lambda \in \sigma(A)} |p(\lambda)|}$$

$$\frac{\|e_n\|_A}{\|e_0\|_A} \leq 2 \left(\frac{\sqrt{K} - 1}{\sqrt{K} + 1} \right)^n$$

$K \to \infty$
convergence
$O(\sqrt{K})$