22MAT220

MATHEMATICS FOR COMPUTING-3

Dr. Sarada Jayan
9448455760
6364333760

Observation about gradient descent method

1. Number of steps required to converge depends on the starting point.
2. No way to predict the number of steps required.

If x is from R^n , can we get convergence in n steps?.

Yes, we have to use a special direction called Conjugate direction.

Krylov Subspaces

- The Krylov sequence was created by Russian mathematician and engineer Alexei Krylov.

Let $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$. The Krylov subspace $K_j(A,b)$ is defined as

$$K_j(A,b) = \text{Span}\{\bar{b}, \overline{Ab}, A^2b, A^3b, \dots, A^{j-1}b\}$$

$K_j(A,b)$ is Subspace of $\mathbb{R}^n$.

Eg:- $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\bar{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$Ab = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$; $A^2b = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$; $A^3b = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 14 \\ 13 \end{pmatrix}$; $A^4b = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{pmatrix} 14 \\ 13 \end{pmatrix} = \begin{pmatrix} 41 \\ 40 \end{pmatrix}$

$K_1(A,b) = \text{Span}\{\bar{b}\} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\} \longrightarrow K_1(A,b)$ is a Subspace of $\mathbb{R}^2$.

$K_2(A,b) = \text{Span}\{\bar{b}, A\bar{b}\} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\} \longrightarrow K_2(A,b) = \mathbb{R}^2$

$K_3(A,b) = \text{Span}\{\bar{b}, \overline{Ab}, A^2b\} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}\right\} = K_2(A,b) = \mathbb{R}^2$.

$K_4(A,b) = \text{Span}\{b, Ab, A^2b, A^3b\} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 14 \\ 13 \end{pmatrix}\right\} = K_2(A,b) = \mathbb{R}^2$.

$K_5(A,b) = \text{Span}\{b, Ab, A^2b, A^3b, A^4b\} = \text{Span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 14 \\ 13 \end{pmatrix}, \begin{pmatrix} 41 \\ 40 \end{pmatrix}\right\} = K_2(A,b) = \mathbb{R}^2$ //

Krylov Subspaces

eigenvalues of $A \to 3, 1$
eigenvectors of $A \to \begin{pmatrix}1\\1\end{pmatrix} \begin{pmatrix}1\\-1\end{pmatrix}$

Eg 2:- $A = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$, $\bar{b} = \begin{pmatrix}1\\1\end{pmatrix}$. Find all Krylov subspaces upto $j=10$.

$b = \begin{pmatrix}1\\1\end{pmatrix}$; $Ab = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix}$; $A^2b = \begin{pmatrix}9\\9\end{pmatrix}$; $A^3b = \begin{pmatrix}27\\27\end{pmatrix}$; $A^nb = 3^n\begin{pmatrix}1\\1\end{pmatrix}$

$K_1(A,b) = \text{Span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$; $K_2(A,b) = \text{Span}\{b, Ab\} = \text{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}3\\3\end{pmatrix}\right\} = \text{Span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$

$K_1(A,b) = K_2(A,b) = K_3(A,b) = \dots = K_{10}(A,b)$

Eg 3:- $A = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$, $\bar{b} = \begin{pmatrix}1\\-1\end{pmatrix}$. Find all Krylov subspaces.

$b = \begin{pmatrix}1\\-1\end{pmatrix}$; $A\bar{b} = \begin{pmatrix}1\\-1\end{pmatrix}$; $A^2b = \begin{pmatrix}1\\-1\end{pmatrix} = A^3b = \dots = A^nb$

$K_1(A,b) = \text{span}\left\{\begin{pmatrix}1\\-1\end{pmatrix}\right\} = K_2(A,b) = K_3(A,b)$

$K_j(A,b)$ is subspace of $\mathbb{R}^2$ & is same for all j.

Eg 4:- $A = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}$, $\bar{b} = \begin{pmatrix}1\\0\\1\end{pmatrix}$. Find $K_2(A,b)$ & $K_3(A,b)$

$b = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $Ab = \begin{pmatrix}2\\2\\2\end{pmatrix}$, $A^2b = \begin{pmatrix}6\\6\\6\end{pmatrix}$

$K_2(A,b) = \text{Span}\left\{\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}2\\2\\2\end{pmatrix}\right\} = \text{Span}\left\{\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$

$K_3(A,b) = \text{span}\{b, Ab, A^2b\} = \text{Span}\left\{\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}2\\2\\2\end{pmatrix}, \begin{pmatrix}6\\6\\6\end{pmatrix}\right\} = K_2(A,b)$

Krylov Matrix

Let $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$

$\{b, Ab, A^2b, \ldots\} \rightarrow$ Krylov Sequence.

$$\begin{bmatrix} | & | & | & & | \\ b & Ab & A^2b & \cdots & A^{j-1}b \\ | & | & | & & | \end{bmatrix}_{n\times j} = j^{th} \text{ order Krylov matrix}$$

Krylov Subspace $K_j(A,b)$ is the column space of the Krylov matrix of j^{th} order.

Eg:- $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

2^{nd} order K. matrix $= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

3^{rd} " " $= \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 4 \end{bmatrix}$

4^{th} " " $= \begin{bmatrix} 1 & 2 & 5 & 14 \\ 0 & 1 & 4 & 13 \end{bmatrix}$

Motivation for Krylov Subspaces

Consider $A = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 3 & 1 \\ -2 & 2 & 3 \end{bmatrix}$.

The characteristic eqn. of A, $|A - \lambda I| = 0$ is $\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$

By Cayley Hamilton Theorem (every sq. matrix satisfies it's own characteristic eqn.)

$$\therefore A^3 - 6A^2 + 11A - 6I = O_{3\times 3}$$

Premultiplying with A^{-1},

$$A^2 - 6A + 11I - 6A^{-1} = O_{3\times 3}$$

$$6A^{-1} = A^2 - 6A + 11I$$

$$A^{-1} = \frac{1}{6}\left[A^2 - 6A + 11I\right]$$

$\longrightarrow$ Cayley-Hamilton helps in finding inverse.

If $Ax = b$ is a linear system with A square matrix.

Soln. of $Ax = b$ is $x = A^{-1}b$

$$\bar{x} = \left(\frac{A^2 - 6A + 11I}{6}\right)b$$

$$= \frac{1}{6}\left[A^2 b - 6Ab + 11b\right]$$

$\longrightarrow$ $\bar{x}$ is a span of b, Ab, A^2b.

$\bar{x} \in K_3(A, b)$.

Motivation for Krylov Subspaces

1) Solns of $\underset{n\times n}{A}x = b$ are elts. of $K_n(A, b)$.

2) Calculation of Ab, A^2b, A^3b, are faster if A is sparse matrix.

[Do not calculate, $A^2, A^3, \ldots A^n$, Do Ab, $A^2b = A(Ab)$
$A^3b = A(A^2b)$].

Need for Krylov Subspaces

- Direct algorithms require on the order of m^3 operations for a matrix of m-dimension. This significantly limits the ability to work with larger matrices.

- Most matrices in engineering applications are sparsely populated, such as those produced from Finite Element Method (FEM) or Finite Difference Method (FDM). When direct methods are used on dense matrices, sparsity is often lost, as zero elements above/below a diagonal become non-zero.

- Iterative solutions that use Krylov Subspaces can reduce the order of operation to m, a drastic improvement over m^3. They also can maintain whatever sparsity exists within a matrix.

Krylov Subspace Methods

- The idea is that it is less costly to look for approximations to the solution space ***x*** or the eigenvalue λ in the Krylov subspace that minimize the residual than perform a direct method, such as QR factorization.

- However, the vectors $\boldsymbol{b, Ab, A^2b, \dots A^kb}$ can quickly become linearly dependant, forming a poor basis. Orthonormalization is usually required to form an orthogonal basis

$$\mathcal{K}_k = span\{b, Ab, A^2b, A^3b, \dots A^{k-1}b\} = span\{q_1, q_2, q_3, \dots q_k\}$$

Krylov Subspace Methods

- Based on projection onto expanding subspaces
- Used for solving linear systems and for solving eigenvalue problems

	Solving the system AX=b	**Solving eigenvalue problem AX=λX**
When Matrix is Symmetric ($A=A^T$) positive definite	Conjugate Gradient Method (CG)	Lanczos Method
When Matrix is not Symmetric ($A\neq A^T$)	Generalized Minimum Residuals (GMRES)	Arnoldi Iteration

What are conjugate directions?.

$Let\ A$ be a real symmetric $n \times n$ matrix with rank n.

3

The directions $d_0, d_1, \ldots., d_{\mathrm{n-1}}$ are A−Conjugate if, for all $i \neq j$, we have $d_i^T A d_j = 0$

$$d_i^T A d_j = d_j^T A d_i = 0, \forall i \neq j$$

A new type of orthogonality.

It is defined w.r.t. a symmetric matrix A

Examples of conjugate directions

Given a matrix, $A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. Verify whether the given directions are A-conjugate or not.

(a) $\bar{d}_1 = \begin{pmatrix} 2 \\ 3 \\ -6 \end{pmatrix}$; $\bar{d}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$$d_1^T A d_2 = \langle d_1, Ad_2 \rangle = \left\langle \begin{pmatrix} 2 \\ 3 \\ -6 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 0$$

$$d_2^T A d_1 = \langle d_2, Ad_1 \rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -10 \end{pmatrix} \right\rangle = 0$$

$\bar{d}_1$ & $\bar{d}_2$ are A conjugate.

If $d_1^T A d_2 = 0$
$(d_1^T A d_2)^T = 0$
$d_2^T (d_1^T A)^T = 0$
$d_2^T A^T d_1 = 0$
$d_2^T A d_1 = 0$

(b) $d_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $d_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$

$$d_1^T A d_2 = \langle d_1, Ad_2 \rangle = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 8 \\ -2 \\ 0 \end{pmatrix} \right\rangle = 4 \neq 0$$

$\therefore$ d_1 & d_2 are not A-conjugate.

For which all matrices A will orthogonal vectors be A-conjugate?

Soln: If A is identity matrix, I or $A = kI$.

Generating Conjugate directions for 3x3 symmetric matrix.
We can find 3 independent conjugate directions.
We will use this as paths along which we descend to minima point
(These directions are not unique)

$$Let\ A=\begin{pmatrix} x & y & z \\ y & p & r \\ z & r & q \end{pmatrix}$$

$Let\ \ d_1^T = [1\ 0\ 0\ ]$.

$$d_1^T A = [x, y, z], \text{ or } \mathrm{Ad}_1 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$d_2 \text{ such that } d_1^T A d_2 = 0 \Rightarrow d_2 \perp A d_1 \Rightarrow d_2 \perp \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \text{ one } d_2 = \begin{pmatrix} yz \\ -2xz \\ xy \end{pmatrix}$$

$$d_2^T = [yz\ \ -2xz\ xy]$$

d_3 shd be A-conjugate with d_1 & d_2.

$\Rightarrow \begin{cases} d_3 \perp Ad_1 \\ d_3 \perp Ad_2 \end{cases}$

If $\bar{x} \perp \bar{y}$
$\bar{x} \perp \bar{z}$
Then $\bar{x} = \bar{y} \times \bar{z}$

Find d_3 such that
$\mathrm{d}_1^T \mathrm{Ad}_3 = 0 \in \mathrm{R}$
$\mathrm{d}_2^T \mathrm{Ad}_3 = 0 \in \mathrm{R}$
or
$\mathrm{d}_3^T \mathrm{Ad}_1 = 0 \in \mathrm{R}$
$\mathrm{d}_3^T \mathrm{Ad}_2 = 0 \in \mathrm{R}$

$$\left.\begin{matrix} \mathrm{d}_1^T \mathrm{Ad}_3 = 0 \in \mathrm{R} \\ \mathrm{d}_2^T \mathrm{Ad}_3 = 0 \in \mathrm{R} \end{matrix}\right\} \Rightarrow d_3 \text{ is orthogonal to vectors } \mathrm{Ad}_1 \text{ and } \mathrm{Ad}_2$$

So, d_3 can be easily obtained by cross producting Ad_1 and Ad_2

How will we create A conjugate vectors

$$Let\ \text{A}=\begin{pmatrix} x & y & z \\ y & p & r \\ z & r & q \end{pmatrix}. \qquad \text{Choose } d_1=\begin{pmatrix}1\\0\\0\end{pmatrix},$$

$$d_1=\begin{pmatrix}1\\0\\0\end{pmatrix}, \Rightarrow d_1^T A = \begin{pmatrix}1 & 0 & 0\end{pmatrix}\begin{pmatrix} x & y & z \\ y & p & r \\ z & r & q \end{pmatrix} = [x \quad y \quad z]$$

$$d_2 : d_1^T A d_2 = 0 \Rightarrow \begin{pmatrix}x & y & z\end{pmatrix}\begin{pmatrix}d_{21}\\d_{22}\\d_{23}\end{pmatrix} = 0 \Rightarrow d_2 \perp \begin{pmatrix}x\\y\\z\end{pmatrix} \Rightarrow \text{d}_2 \text{ can be} \begin{pmatrix} yz \\ -2xz \\ xy \end{pmatrix}$$

$$Temp1 = d_1^T A = [x \quad y \quad z]$$

$$Let\ Temp2 = d_2^T A = [x_1 \quad y_1 \quad z_1]$$

We need d_3 such that $d_1^T A d_3 = d_2^T A d_3 = 0$

$Then$ $\text{d}_3 = cross$ product of Temp1 and temp 2

$\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}2\\3\\-6\end{pmatrix}$ are conj. directions //

$$A=\begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix} \qquad d_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

$$d_1^T A = \begin{pmatrix}1 & 0 & 0\end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

$= (3 \quad 0 \quad 1)$; or $Ad_1 = \begin{pmatrix}3\\0\\1\end{pmatrix}$

$d_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$; $Ad_2 = \begin{pmatrix}0\\4\\2\end{pmatrix}$

$$d_2^T A = \begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

$= (0 \quad 4 \quad 2)$

$$d_3 = \begin{pmatrix}3\\0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\4\\2\end{pmatrix} = \begin{pmatrix}-4\\-6\\12\end{pmatrix} \cong \begin{pmatrix}2\\3\\-6\end{pmatrix} //$$

$$Ad_1 \times Ad_2 = \begin{vmatrix} i & j & k \\ 3 & 0 & 1 \\ 0 & 4 & 2 \end{vmatrix} = i(-4) - j(6) + k(12)$$

Create A-conjugate directions for the given A

Qn) $A = \begin{bmatrix} 4 & 1 & -1 \\ 1 & 6 & 1 \\ -1 & 1 & 5 \end{bmatrix}$;

cross pdt.

$d_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$; $Ad_1 = \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix}$; $d_2 = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$; $Ad_2 = \begin{pmatrix} 0 \\ 9 \\ -7 \end{pmatrix}$; $d_3 = Ad_1 \times Ad_2$

$= \begin{pmatrix} 2 \\ 28 \\ 36 \end{pmatrix} \sim \begin{pmatrix} 1 \\ 14 \\ 18 \end{pmatrix}$

$\therefore d_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$ & $d_3 = \begin{pmatrix} 1 \\ 14 \\ 18 \end{pmatrix}$ are A conjugate dirns.

$d_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, d_3 = \begin{pmatrix} 13 \\ -24 \\ 28 \end{pmatrix}$ are also A conjugate directions.

Observation

Conjugate directions are independent w.r.t all others in the set

Conjugate gradient method (CG)

- invented by Hestenes and Stiefel around 1951
- the most widely used iterative method for solving $Ax = b$, with $A > 0$
- can be extended to non-quadratic unconstrained minimization

Conjugate Gradient Method

Algorithm:-

Step 1:- Input $\rightarrow A, b, x^{(0)}$

Step 2:- Compute $\bar{r}_0 = b - Ax^{(0)}$ $\left(\bar{r}_0 = -g(x^{(0)}) \right)$

Set $\bar{d}_0 = \bar{r}_0$

Step 3:- For $k = 0, 1, 2, \ldots$ until convergence

Calculate $\alpha_k = \dfrac{\langle r_k, r_k \rangle}{\langle d_k, Ad_k \rangle}$ $\left(\alpha_k = \dfrac{r_k^T r_k}{d_k^T A d_k} \right)$

Step 4:- $x^{(k+1)} = x^{(k)} + \alpha_k d_k$

Step 5:- $r_{k+1} = r_k - \alpha_k A d_k$ $\longrightarrow$

$$\left(\begin{aligned} r_{k+1} &= b - Ax^{(k+1)} \\ &= b - A\left(x^{(k)} + \alpha_k d_k\right) \\ &= \underbrace{b - Ax^{(k)}} - \alpha_k A d_k \\ &= r_k - \alpha_k A d_k \end{aligned} \right)$$

Step 6:- If $r_{k+1} = 0$, then stop

Else, calculate $\beta_k = \dfrac{\langle r_{k+1}, r_{k+1} \rangle}{\langle r_k, r_k \rangle}$

Step 7:- Find $d_{k+1} = r_{k+1} + \beta_k d_k$ (d_{k+1} which will be A conjugate with $d_1, d_2, \ldots d_k$)

Explanation: The conjugate gradient algorithm returns approximations

$$x_j \in x_0 + \underline{K_j(A, r_0)} \quad \text{for } j = 0, 1, 2 \cdots$$

such that

$$\|x - x_j\|_A = \min_{q \in P_{j-1}} \|(I - A q(A))(x - x_0)\|_A$$

where $\|x\|_A = \sqrt{x^T A x}$

First step: $x_0 \to x_1$ $\qquad \alpha_0 r_0 \in K_1\{A, r_0\}$

$$x_1 = x_0 + \underline{\alpha_0 r_0} \in x_0 + K_1$$

Here, $\alpha_0 = \dfrac{\langle r_0, r_0 \rangle}{d_0^T A d_0} = \dfrac{\|r_0\|_2^2}{\|d_0\|_A^2} \neq 0$

$$r_1 = r_0 - \alpha A_0 d_0 \in K_2(A, b)$$

If $r_1 = 0$ STOP

otherwise $\{r_0, r_1\}$ is an orthonormal basis of $\underline{K_2(A, b)}$.

Why solution of Ax=b is taken as $x=x0$ + an element of $K_n(A,r0)$?

An element of $K_n(A,r0)$ is solution of $Au = r0$ i.e., $u=A^{-1}(r0)$

So $x=x0$ + an element of $K_n(A,r0)$ is

$x=x0 + A^{-1}(r0)$

We can check if this x is a solution of $Ax=b$

$$
\begin{aligned}
Ax &= A\,(x0 + A^{-1}(r0)) \\
&= A\,x0 + AA^{-1}(r0)) \\
&= Ax0 + r0 \\
&= Ax0 + b - Ax0 \\
&= b
\end{aligned}
$$

Hence $x=x0$ + an element of $K_n(A,r0)$ is a solution of the system $Ax=b$.

HW

Solve the system $2x_1 - x_2 = 1$; $-x_1 + 2x_2 = 0$ using CG with $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$$r_0 = b - Ax_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad d_0 = r_0$$

$$\alpha_0 = \frac{\langle r_0, r_0 \rangle}{\langle d_0, Ad_0 \rangle} = \frac{1}{2}$$

$$x^{(1)} = x^{(0)} + \alpha_0 d_0 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$$

$$r_1 = r_0 - \alpha_0 A d_0 \quad (\text{or } r_1 = b - Ax^{(1)})$$

$$= \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \longrightarrow r_1 \neq 0$$

$$\beta_0 = \frac{\langle r_1, r_1 \rangle}{\langle r_0, r_0 \rangle} = 1/4$$

$$d_1 = r_1 + \beta_0 d_0 = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/2 \end{pmatrix}$$

$$\alpha_1 = \frac{\langle r_1, r_1 \rangle}{\langle d_1, Ad_1 \rangle} = \frac{2}{3}$$

$$x^{(2)} = x^{(1)} + \alpha_1 d_1 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} 1/4 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$$

$$r_2 = r_1 - \alpha_1 A d_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\|r_2\| = 0 \Longrightarrow \boxed{x^{(2)} \text{ is the exact soln. } x^* = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}}$$

Conjugate Gradient Method

Solve Ax=b using conjugate gradient method with initial point as $(0,0,0)^T$ $x^{(0)}$ if $A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$

$$r_0 = b - Ax^{(0)} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}; \quad d_0 = r_0 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

$$\alpha_0 = \frac{r_0' r_0}{d_0' A d_0} = \frac{5}{18} = 0.2778;$$

$$x^{(1)} = x^{(0)} + \alpha_0 d_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \frac{5}{18}\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5/6 \\ 0 \\ 5/18 \end{pmatrix} = \begin{pmatrix} 0.8333 \\ 0 \\ 0.2778 \end{pmatrix}$$

$$r_1 = r_0 - \alpha_0 A d_0 = \begin{pmatrix} 2/9 \\ -5/9 \\ -2/3 \end{pmatrix} = \begin{pmatrix} 0.2222 \\ -0.5556 \\ -0.6667 \end{pmatrix} \rightarrow \|r_1\| \neq 0.$$

$$\beta_0 = \frac{\langle r_1, r_1 \rangle}{\langle r_0, r_0 \rangle} = 0.0802$$

$$d_1 = r_1 + \beta_0 d_0 = \begin{pmatrix} 0.4630 \\ -0.5556 \\ -0.5864 \end{pmatrix}$$

$$\alpha_1 = \frac{r_1' r_1}{d_1' A d_1} = 0.2187$$

$$x^{(2)} = x^{(1)} + \alpha_1 d_1 = \begin{pmatrix} 0.9346 \\ -0.1215 \\ 0.1495 \end{pmatrix}$$

$$r_2 = r_1 - \alpha_1 A d_1 = \begin{pmatrix} 0.0467 \\ 0.1869 \\ -0.1402 \end{pmatrix} \rightarrow \|r_2\| \neq 0$$

$$\beta_1 = \frac{\langle r_2, r_2 \rangle}{\langle r_1, r_1 \rangle} = 0.0707$$

$$d_2 = r_2 + \beta_1 d_1 = \begin{pmatrix} 0.0795 \\ 0.1476 \\ -0.1817 \end{pmatrix}$$

$$\alpha_2 = \frac{\langle r_2, r_2 \rangle}{d_2' A d_2} = 0.8231$$

$$x^{(3)} = x^{(2)} + \alpha_2 d_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$r_3 = r_2 - \alpha_2 A d_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

```
A=[3,0,1;0,4,2;1,2,3];
b=[3;0;1];
x0=[0;0;0];
r0=b-A*x0;
d0=r0;
alpha0=(r0'*r0)/((A*d0)'*d0);
x1=x0+alpha0*d0;
r1=r0-alpha0*A*d0;
beta0=(r1'*r1)/(r0'*r0);
d1=r1+beta0*d0;
alpha1=(r1'*r1)/((A*d1)'*d1);
x2=x1+alpha1*d1;
r2=r1-alpha1*A*d1;
beta1=(r2'*r2)/(r1'*r1);
d2=r2+beta1*d1;
alpha2=(r2'*r2)/((A*d2)'*d2)
x3=x2+alpha2*d2
r3=r2-alpha2*A*d2
```

Conjugate Gradient Method

```
A=[3,0,1;0,4,2;1,2,3];
b=[3;0;1];
x=randi([-9, 9],length(b),1);
r = b - A * x;
d = r;
rsold = r' * r;
for i = 1:length(b)
Ad = A * d;
alpha = rsold / (d' * Ad);
x = x + alpha * d;
r = r - alpha * Ad;
rsnew = r' * r;
if sqrt(rsnew) < 1e-10
break;
end
d = r + (rsnew / rsold) * d;
rsold = rsnew;
end
x
residue=b-A*x
```

```
x = 3×1

    1.0000
   -0.0000
    0.0000
```

```
residue = 3×1
10^-15 ×

    0.4441
    0.5017
    0.1110
```

Conjugate Gradient Method

Solve the system AX=B, where $A = \begin{bmatrix} 5 & 1 & 2 & -1 \\ 1 & 9 & 1 & 3 \\ 2 & 1 & 4 & 0 \\ -1 & 3 & 0 & 6 \end{bmatrix}$ and $b = \begin{bmatrix} 7 \\ 14 \\ 7 \\ 8 \end{bmatrix}$ using conjugate gradient method with any initial vector

```
A=[5,1,2,-1;1,9,1,3;2,1,4,0;-1,3,0,6];
b=[7;14;7;8];
x=randi([-9, 9],length(b),1)
r = b - A * x;
d = r;
rsold = r' * r;
for i = 1:length(b)
Ad = A * d;
alpha = rsold / (d' * Ad);
x = x + alpha * d;
r = r - alpha * Ad;
rsnew = r' * r;
if sqrt(rsnew) < 1e-10
break;
end
d = r + (rsnew / rsold) * d;
rsold = rsnew;
end
x
residue=b-A*x
```

Solution of this is $(1,1,1,1)^T$

Conjugate Gradient Method

Solve the optimization problem:

$$f(x_1, x_2) = 2x_1^2 + 1x_2^2 + 2x_1x_2 + x_1 - x_2$$

$$= \frac{1}{2}\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \frac{1}{2}x^T A x - b^T x$$

```
A=[4,2;2,2];
b=[-1;1];
x=randi([-9, 9],length(b),1)
r = b - A * x;
d = r;
rsold = r' * r;
for i = 1:length(b)
Ad = A * d;
alpha = rsold / (d' * Ad);
x = x + alpha * d;
r = r - alpha * Ad;
rsnew = r' * r;
if sqrt(rsnew) < 1e-10
break;
end
d = r + (rsnew / rsold) * d;
rsold = rsnew;
end
x
residue=b-A*x
```

```
x = 2×1
   -1.0000
    1.5000

residue = 2×1
10^-14 ×
   -0.3553
    0.0888
```