Conjugate gradient Method :- To minimize quadratic function

- $x_1, x_2, x_3, \text{---}$

$x_{i+1} = x_i + \alpha_i d_i$

α_i — optimal length

$\min f(x_i + \alpha_i d_i)$

$f(x_1) > f(x_2) > f(x_3) > \text{---}$

$d_i \rightarrow$ direction

Conjugate direction :- Let Q be $n \times n$ positive definite symmetric matrix. The non-zero vectors $d_1, d_2 \in R^n$ are said to be Q-conjugate if

column vectors $\leftarrow d_1, d_2$

$$\underline{\underline{d_1^t Q d_2}} = 0.$$

I-conjugate d_1, d_2 then d_1 & d_2 orthogonal vectors.

Example :-

$$Q = \begin{bmatrix} 4 & -2 \\ -2 & 32 \end{bmatrix}$$

$d_1 = (15, -1)$

$d_2 = (1, 1)$

symmetric

positive definite $\quad 4 > 0$

$4 \times 32 - 4 > 0$

- d_1 and d_2 are Q-conjugate direction

$$d_1^t Q d_2 = [15, -1] \begin{bmatrix} 4 & -2 \\ -2 & 32 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = [15, -1] \begin{bmatrix} 2 \\ 30 \end{bmatrix} = [0].$$

Theorem :- Let Q be a positive definite symmetric matrix.

$$\text{quadratic function} \leftarrow \min f(x) = \frac{1}{2} X^t \overset{n\times n \text{ matrix}}{Q} X + \underset{\text{Row matrix}}{b^t} X + \underset{\text{constant number}}{C}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}$$

$X = \begin{pmatrix} x_2 \\ | \\ x_n \end{pmatrix}$ matrix

Let $d_1, d_2, -, d_n$ be the complete set of Q- conjugate directions in $\mathbb{R}^n$. Then By conjugate gradient method, we find $\min f(x)$ in at most n iterations

↳ same as the no. of variables in the functions.

Example :-

$$f(x,y) = 2x^2 + 16y^2 - 2xy - x - 6y - 5$$

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \quad f(X) = \frac{1}{2} X^t Q X + b^t X + C$$

$$f(x,y) = \frac{1}{2}[x, y]\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [k_1 \; k_2]\begin{bmatrix} x \\ y \end{bmatrix} + c$$

$$f(x,y) = \frac{1}{2}\left[x, y\right]\begin{bmatrix} ax + by \\ bx + dy \end{bmatrix} + k_1 x + k_2 y + c$$

$$f(x,y) = \frac{1}{2}\left(ax^2 + bxy + bxy + dy^2\right) + k_1 x + k_2 y + c$$

$$f(x,y) = \frac{a}{2}x^2 + \underline{bxy} + \frac{d}{2}y^2 + \underline{k_1 x} + \underline{k_2 y} + c$$

$$f(X) = \frac{1}{2} X^t Q X + b^t X + C, \quad Q = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

$f(x) = \frac{1}{2} x^t Q x + b^t x + c$, $Q = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$

$b^t = [k_1, k_2]$

$\frac{a}{2} \rightarrow$ coef of x^2

$\frac{d}{2}$ - coef of y^2

$f(x,y) = 2x^2 + 16y^2 - 2xy - x - 6y - 5$

$Q = \begin{bmatrix} 4 & -2 \\ -2 & 32 \end{bmatrix}$ $\quad b^t = [-1, -6]$

$f(x,y) = \frac{1}{2} x^t \begin{bmatrix} 4 & -2 \\ -2 & 32 \end{bmatrix} x + [-1, -6] x - 5$

- Symmetric matrix $\quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$

$A = A^t \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \Rightarrow \boxed{b = c}$

Conjugate Gradient Algorithm/Method

Min $f(x)$, initial point, stopping condition → $|\nabla f(x_k)| < \varepsilon$ OR $|x_k - x_{k-1}| < \varepsilon$ OR $|\nabla f(x_k) - \nabla f(x_{k-1})| < \varepsilon$

① Let the initial point be x_1

② first search direction $d_1 = -\nabla f(x_1)$

③ $x_2 = x_1 + \alpha_1 d_1$, where α_1 is the optimal length and it is determined by minimizing the function $f(x_1 + \alpha_1 d_1)$

④ set $i = 2$ $\quad$ find $\nabla f_i = \nabla f(x_i)$ and

$|\nabla f_i|^2$ ...

(4) set $i=2$ ...

$$\text{set } d_i = -\nabla f_i + \frac{|\nabla f_i|^2}{|\nabla f_{i-1}|^2} d_{i-1}$$

[d_2 will be Q-conjugate to d_1
d_3 will be Q-conjugate to both d_1 and d_2]

(5) Compute α_i by minimizing the function $f(X_i + \alpha_i d_i)$

(6) set $X_{i+1} = X_i + \alpha_i d_i$

(7) If X_{i+1} satisfies optimality condition, the stop, Otherwise set $i = i+1$ & go to step (4).

Question :- Minimize $f(x,y) = x - y + 2x^2 + 2xy + y^2$, starting from the point $X_1 = (0,0)$. $|\nabla f(X_k)| = 0$ is stopping condition

- Suppose the question is

$$\text{Max } g(x,y) = -x + y - 2x^2 - 2xy - y^2$$

- $\text{Min } -g(x,y) = \boxed{\text{Min } f(x,y) = x - y + 2x^2 + 2xy + y^2}$

minima of $f(x,y)$
$\shortparallel$
maxima of $g(x,y)$

Solution :-

$X_1 = (0,0)$ $\qquad$ $f(x) = x - y + 2x^2 + 2xy + y^2$

$$\nabla f(x) = \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \end{pmatrix} = \begin{pmatrix} 1 + 4x + 2y \\ -1 + 2x + 2y \end{pmatrix}$$

$1 - 4 + 3 = 0$

$-1 - 2 + 3 = 0$

$$\nabla f(x_1) = \nabla f(0,0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad d_1 = -\nabla f(0,0) = (-1,1)$$

To find optimal length α_1

$$\min f(x_1 + \alpha_1 d_1) = \min f\big((0,0) + \alpha_1(-1,1)\big)$$
$$= \min f(-\alpha_1, \alpha_1)$$

$$f(-\alpha_1, \alpha_1) = -2\alpha_1 + \alpha_1^2$$

$$\frac{d\, f(-\alpha_1,\alpha_1)}{d\alpha_1} = -2 + 2\alpha_1 = 0 \Rightarrow \boxed{\alpha_1 = 1}$$

is minima

$$\frac{d^2 f(-\alpha_1,\alpha_1)}{d\alpha_1^2} = 2 > 0$$

$$x_2 = x_1 + \alpha_1 d_1 = (0,0) + 1(-1,1)$$

$$x_2 = (-1,1)$$

$|\nabla f(x_2)| \neq |(0,0)| = 0$, x_2 is not optimal solution.

iteration 2

$$d_2 = -\nabla f(x_2) + \frac{|\nabla f(x_2)|^2}{|\nabla f(x_1)|^2} d_1$$

$$\nabla f(0,0) = \nabla f(x_1) = (1,-1) \qquad |\nabla f(x_1)|^2 = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

$$\nabla f(-1,1) = \nabla f(x_2) = (-1,-1) \qquad |\nabla f(x_2)|^2 = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}$$

$$d_2 = -(-1,-1) + \frac{\sqrt{2}}{\sqrt{2}}(-1,1) = (1,1) + (-1,1)$$

$$d_2 = -(-1,-1) + \frac{\sqrt{2}}{\sqrt{2}}(-1,1) = (1,1)+(-1,1)$$

$$d_2 = (0,2)$$

- To find optimal length α_2

$$\min f(x_2 + \alpha_2 d_2) = \min f((-1,1) + \alpha_2 (0,2))$$

$$\min f(-1, 1+2\alpha_2)$$

$$f(x,y) = x - y + 2x^2 + 2xy + y^2$$

$$f(-1, 1+2\alpha_2) = -1 - 1 - 2\alpha_2 + 2 + 2(-1)(1+2\alpha_2) + (1+2\alpha_2)^2$$

$$\frac{df(-1, 1+2\alpha_2)}{d\alpha_2} = -2 - 4 + 2(1+2\alpha_2)\cdot 2$$

$$= -6 + 4 + 8\alpha_2$$

$$\frac{df(-1,1+2\alpha_2)}{d\alpha_2} = -2 + 8\alpha_2 = 0 \qquad \alpha_2 = \frac{2}{8} = \frac{1}{4}$$

$\downarrow$

minima

$$\frac{d^2 f(-1, 1+2\alpha_2)}{d\alpha_2^2} = 8 > 0$$

- $$X_3 = X_2 + \alpha_2 d_2$$

$$= (-1,1) + \frac{1}{4}(0,2) = (-1, 3/2)$$

$\hookrightarrow$ is the optimal solution.

$$|\nabla f(X_3)| = |(0,0)| = 0$$

* $d_1 = (-1,1)$ $\quad d_2 = (0,2)$

$$f(x,y) = x - y + 2x^2 + 2xy + y^2$$

$$f(x,y) = \frac{1}{2} X^t Q X + b^t X + c$$

$$f(x,y) = \frac{1}{2}[x,y]\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [k_1, k_2]\begin{bmatrix} x \\ y \end{bmatrix} + c$$

$$f(x,y) = \frac{1}{2}[x,y]\begin{bmatrix} ax + by \\ bx + dy \end{bmatrix} + k_1 x + k_2 y + c$$

$$f(x,y) = \frac{1}{2}\left(ax^2 + 2bxy + dy^2\right) + k_1 x + k_2 y + c$$

$$f(x,y) = a/2\, x^2 + bxy + d/2\, y^2 + \underline{k_1} x + k_2 y + c$$

$$f(x,y) = \underline{x} - y + 2x^2 + 2xy + y^2$$

$b^t = [k_1, k_2] = [1, -1]$

$Q = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$ $\quad \frac{a}{2} = 2 \Rightarrow a = 4, \quad b = 2, \quad \frac{d}{2} = 1 \Rightarrow d = 2$ $\quad c = 0$

To verify that $d_1 = (-1,1)$ $d_2 = (0,2)$ are Q-conjugate directions

- $d_1^t Q d_2 = 0$ $\quad [-1, 1]\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = [-1, 1]\begin{bmatrix} 4 \\ 4 \end{bmatrix}$

$= -4 + 4 = 0.$

. System of linear equations $\quad$ matrix of coefficients

- System of linear equations

$$3x + 4y + 5z = 11$$
$$x + y + 2z = 12$$
$$5x + y + z = 2$$

$$[A:B] = \begin{bmatrix} 3 & 4 & 5 & 11 \\ 1 & 1 & 2 & 12 \\ 5 & 1 & 1 & 2 \end{bmatrix}$$

$$AX = B$$

A → matrix of coefficients

X → matrix of variables

B → matrix of constants

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 1 & 2 \\ 5 & 1 & 1 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad B = \begin{bmatrix} 11 \\ 12 \\ 2 \end{bmatrix}$$

$$AX = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 1 & 2 \\ 5 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 4y + 5z \\ x + y + 2z \\ 5x + y + z \end{bmatrix} = \begin{bmatrix} 11 \\ 12 \\ 2 \end{bmatrix}$$

$AX = B$ system of linear equations.

$AX = B$ →

- Unique solution
  - ↓ A will an invertible matrix
  - $A^{-1}B = X$
- No solution
- Infinitely many solutions.

$$\boxed{A^{-1} = \frac{Adj(A)}{\det(A)}}$$

$$AX = B$$

$\text{Rank}[A]_{n \times n}$

$$AX = B$$

$\text{Rank}[A]_{n\times n}$

$\text{Rank}[A:B]_{n\times(n+1)}$

$\text{Rank}[A] = \text{Rank}[A:B]$ → Solution exists

$\text{Rank}[A] \neq \text{Rank}[A:B]$ → No solution

Solution exists:

- $\text{Rank}[A] =$ No. of variables → unique solution
- $\text{Rank}(A) \neq$ No. of variables → infinitely many solutions.

* Solving the system of linear Equations $\underline{AX = B}$, when A is symmetric matrix, is same as solving optimization problem

$$\underline{\text{Min}\left(\tfrac{1}{2} X^t A X - B^t X\right)}$$

Example:

$$x + y = 3$$
$$x - y = 1$$

$$\boxed{\begin{aligned} x &= 2 \\ y &= 1 \end{aligned}}$$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \end{bmatrix} \quad B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$f(x,y) = \frac{1}{2}[x,y]\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} - [3,1]\begin{bmatrix}x\\ y\end{bmatrix} = \frac{1}{2}[x,y]\begin{bmatrix}x+y\\ x-y\end{bmatrix} - 3x - y$$

$$f(x,y) = \frac{1}{2}\left(x^2 + xy + xy - y^2\right) - 3x - y$$

$$f(x,y) = \frac{x^2}{2} + xy - \frac{y^2}{2} - 3x - y$$

$$\text{Min } f(x,y) = \text{Min}\left(\frac{x^2}{2} + xy - \frac{y^2}{2} - 3x - y\right)$$

$$\nabla f(x,y) = \begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix} = \begin{pmatrix}x+y-3\\ x-y-1\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$$

$$x + y = 3$$
$$x - y = 1$$