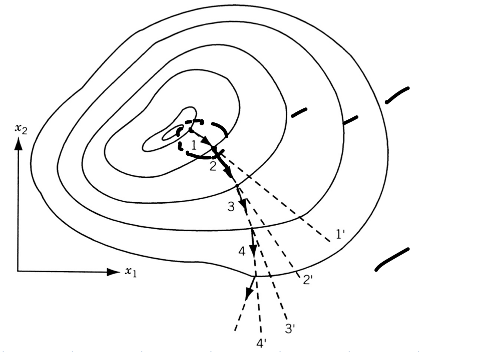
01-oct-2021 # Indirect Search Methods. (Descent)

Gradient of an objective function.

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{Bmatrix} \qquad \nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix}$$



If we move along the gradient direction, the value of the function increases at the fastest rate

However, the problem is ~~that~~ it is a local property

Steepest Ascent direction : ∇f

Steepest Descent " : $-\nabla f$

· Arbitrary point $\underset{\sim}{X}$ in $\mathbb{R}^n$

f be the objective function value at $\underset{\sim}{X}$

Consider a neighbouring point $\underset{\sim}{X} + d\underset{\sim}{X}$

$\underset{\sim}{X} \qquad \underset{\sim}{X}+d\underset{\sim}{X}$

$$d\underset{\sim}{X} = \begin{Bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{Bmatrix} \qquad d\underset{\sim}{X}^T d\underset{\sim}{X} = (ds)^2$$

$$\boxed{d\underset{\sim}{X} = ds\, \underset{\sim}{u}}$$ unit vector along $d\underset{\sim}{X}$

$f + df$ represents the value of the objective fn. @ $\underset{\sim}{X} + d\underset{\sim}{X}$

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \nabla f^T d\underset{\sim}{X}$$

(02)

$$df = \nabla f^T d\underset{\sim}{x}$$

$$= ds \nabla f^T \underset{\sim}{u}$$

$$\therefore \boxed{\frac{df}{ds}} = \nabla f^T \underset{\sim}{u} = \|\nabla f\| \, \|\underset{\sim}{u}\| \cos\theta$$

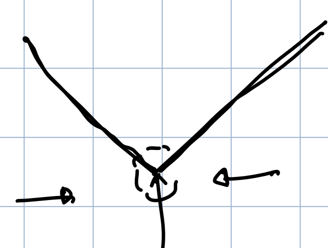
The maximum variation of f w.r.t s takes place when

$\theta = 0 \Rightarrow \underset{\sim}{u}$ is in the same direction as ∇f

$$\boxed{\left.\frac{df}{ds}\right|_{max} = \|\nabla f\|}$$

$$\left.\frac{\partial f}{\partial x_i}\right|_{x_m} \simeq \frac{f(\underset{\sim}{x}_m + \Delta x_i \underset{\sim}{u}_i) - f(\underset{\sim}{x}_m)}{\Delta x_i} \quad \text{F.D.}$$

$$\left.\frac{\partial f}{\partial x_i}\right|_{x_m} \simeq \frac{f(\underset{\sim}{x}_m + \Delta x_i \underset{\sim}{u}_i) - f(\underset{\sim}{x}_m - \Delta x_i \underset{\sim}{u}_i)}{2 \Delta x_i}$$



Rate of change of a function along a direction.

$$\underset{\sim}{x}_i \rightarrow \underset{\sim}{x}_\cdot = \underset{\sim}{x}_i + \lambda \underset{\sim}{s}_i$$

$$\frac{df}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial \lambda}$$

$$\frac{\partial x_j}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left(x_{ij} + \lambda s_{ij}\right) = s_{ij}$$

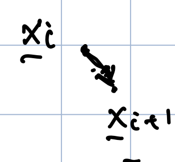
x_{ij} is the j^{th} component of $\underset{\sim}{X_i}$ & s_{ij} is the j^{th} component of $\underset{\sim}{S_i}$

$$\frac{df}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} s_{ij} = \nabla f^T \underset{\sim}{S_i}$$

If λ^* minimizes f in the direction of $\underset{\sim}{S_i}$

$$\left.\frac{df}{d\lambda}\right|_{\lambda=\lambda^*} = 0 \Rightarrow \nabla f^T\Big|_{\lambda^*} \underset{\sim}{S_i} = 0$$

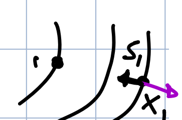
$$\underline{\underset{\sim}{X_i} + \lambda^* S_i}$$



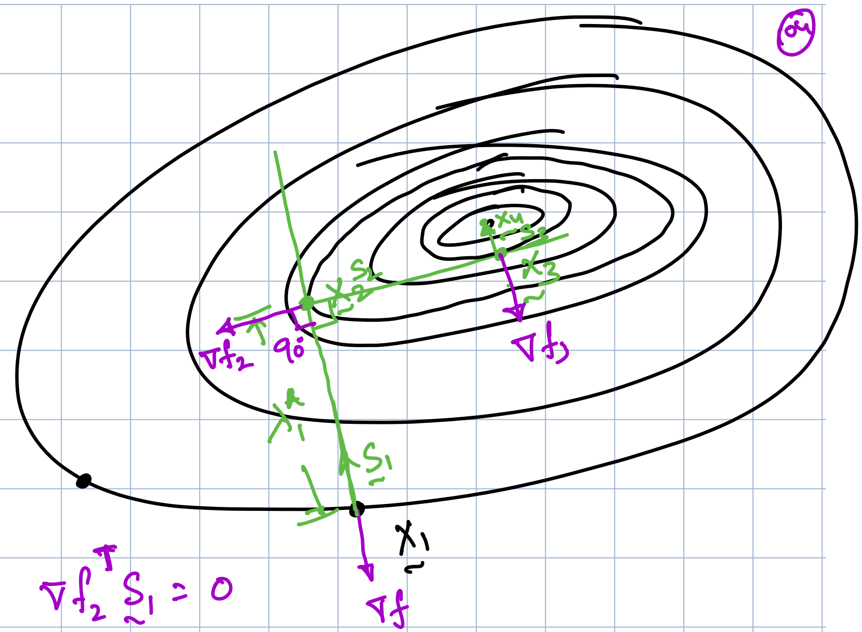
- **Steepest Descent Method:** (Cauchy)

1. Start with initial guess $\underset{\sim}{X_1}$, $i=1$
2. Find the search direction S_i

$$S_i = -\nabla f_i = -\nabla f(\underset{\sim}{X_i})$$



3. $$\underset{\sim}{X_{i+1}} = \underset{\sim}{X_i} + \lambda_i^* \underset{\sim}{S_i} = \underset{\sim}{X_i} - \lambda_i^* \nabla f_i$$
4. Check $\underset{\sim}{X_{i+1}}$ for optimality. If $\underset{\sim}{X_{i+1}}$ is not the optimum, $i = i+1$, go to step 2



$\therefore \quad \nabla f_2^T S_1 = 0$

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

from $X_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$

$i = 1$ $\quad \nabla f = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_2 + 2x_1 \end{Bmatrix}$ ✓

$\nabla f\big|_1 = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} \qquad S_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$

$X_2 = X_1 + \lambda_1 S_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + \lambda_1 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -\lambda_1 \\ \lambda_1 \end{Bmatrix}$ ✓

find λ_1^*

$\frac{df}{d\lambda_1} = 0 \Rightarrow f = -\lambda_1 - \lambda_1 + \cancel{2\lambda_1^2} - \cancel{2\lambda_1^2} + \lambda_1^2$

$\frac{df}{d\lambda} = -2 + 2\lambda_1 = 0 \Rightarrow \lambda_1^* = 1$

(05)

$$\therefore \ \underset{\sim}{X}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

Check for optimality of $\underset{\sim}{X}_2$.

$$\nabla f_2 = \begin{Bmatrix} 1+4(-1)+2(1) \\ -1+2(-1)+2(1) \end{Bmatrix} = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}.$$

$\therefore \ \underset{\sim}{X}_2$ is not optimum.

$$\underset{\sim}{S}_2 = -\nabla f_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

$$\therefore \ \underset{\sim}{X}_3 = \underset{\sim}{X}_2 + \lambda_2^* \underset{\sim}{S}_2$$

$$f(\underset{\sim}{X}_2 + \lambda_2 \underset{\sim}{S}_2) = f(-1+\lambda_2, 1+\lambda_2)$$

$$= 5\lambda_2^2 - 2\lambda_2 - 1$$

$$\therefore \ \frac{df}{d\lambda_2} = 10\lambda_2 - 2 = 0 \Rightarrow \lambda_2^* = \frac{1}{5}$$

$$\therefore \ \underset{\sim}{X}_3 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{5}\begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix}$$

$$\nabla f_3 = \begin{Bmatrix} 0.2 \\ -0.2 \end{Bmatrix} \neq \underset{\sim}{0} \quad \therefore \ \underset{\sim}{X}_3 \text{ is not optimum.}$$

<u>iteration 3</u>

$$\underset{\sim}{S}_3 = -\nabla f_3 = \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix}$$

$$f(\underset{\sim}{X}_3 + \lambda_3 \underset{\sim}{S}_3) = f(-0.8 - 0.2\lambda_3, 1.2 + 0.2\lambda_3)$$

$$\frac{df}{d\lambda_3} = 0 \Rightarrow \lambda_3^* = 1$$

106

$$\underset{\sim}{X_4} = \underset{\sim}{X_3} + \lambda_3^* \underset{\sim}{S_3}$$

$$= \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix} + 1 \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1.4 \end{Bmatrix} \checkmark$$

$$\underset{\sim}{X^*} = \begin{Bmatrix} -1 \\ 1.5 \end{Bmatrix}$$

$$\nabla f_4 = \begin{Bmatrix} -0.2 \\ -0.2 \end{Bmatrix} \neq 0$$

$$S_4 = \begin{Bmatrix} 0.2 \\ 0.2 \end{Bmatrix}$$

i=4

$$\underset{\sim}{X_5} = \underset{\sim}{X_4} + \lambda_4 \underset{\sim}{S_4}$$

$$\underset{\sim}{X_5} = \begin{Bmatrix} -1 \\ 1.4 \end{Bmatrix} + \lambda_4 \begin{Bmatrix} 0.2 \\ 0.2 \end{Bmatrix}$$

$$f(-1 + 0.2\lambda_4 ,\ 1.4 + 0.2\lambda_4)$$

$$\frac{df}{d\lambda_4} = 0 \ ? \quad \Rightarrow \lambda_4^* = 0.375$$

$$\underset{\sim}{X_5} = \begin{Bmatrix} -0.25 \\ 2.15 \end{Bmatrix} \checkmark$$