MDO (Multi-disciplinary design & optimisation)

→ Syllabus (Gradient opti, Non-gradient, Robust opti, Multi-objective opti, Multi-disciplinary opti, Calculation gradient, ...)

→ Grading → ~~Quiz 1~~, Quiz 2 (20%), Endsem (40%)

→ Intro. Internal - (40%)

Srinivasan - OR
Deepanjali - OR
Chaitanya - OR
Gokul - OR
Honhar - OR
Rajkumar - OR
Dharmik - OR
Pratik (Structure) - OK
Mital (Struct) -

Constraints / cost /
Mathematical
Design

Min: ① # of machines required to produce a product.

→ 100 ball bearing

② → time of machining

③ → Cost of "

f_1 f_2 f_3

$f_2 = \alpha \frac{1}{f_1}$

→ # of machine
→ cost of production of unit product
→ time of machining " " "

Constraints

Aircraft

① Objectives: ~~① Min. Drag~~
~~② Fuel consumption~~
③ Range
④ Payload
~~⑤ Max. altitude~~
⑥ Cost (R&D, production, maintenance, operations)
⑦ Size

ISRO
P

Military

PROFIT

L/D

Aircraft design problem

① Objective function.

② Constraints

③ Design variables.

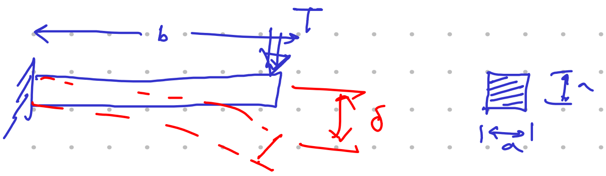
Objective functions <==> Constraints

These two are sometimes interchangeable.

1. Objective function

2. Constraints

3. Design variables



Constraint: $\Rightarrow g(T, \alpha, a, b) = \delta \leqslant \delta_{max}$

geometry)

Objective: $f(a,b) = b \cdot a^2 \cdot \underbrace{\rho_{mat.}}_{const.}$

a, b – Design variables.

28/01/2021

Objective functions are dependent on the perspective (the end client).
For e.g., I can optimise an aircraft for:
- manufacturer (low manufacturing cost)
- user (maximise safety and comfort)
- Airlines (minimise operating costs) etc.

We note that there are multiple disciplines.

But this course is about only the engineering aspect of the design and optimisation process.

So what are the various disciplines (engineering streams) required to design an aircraft?

→ 1

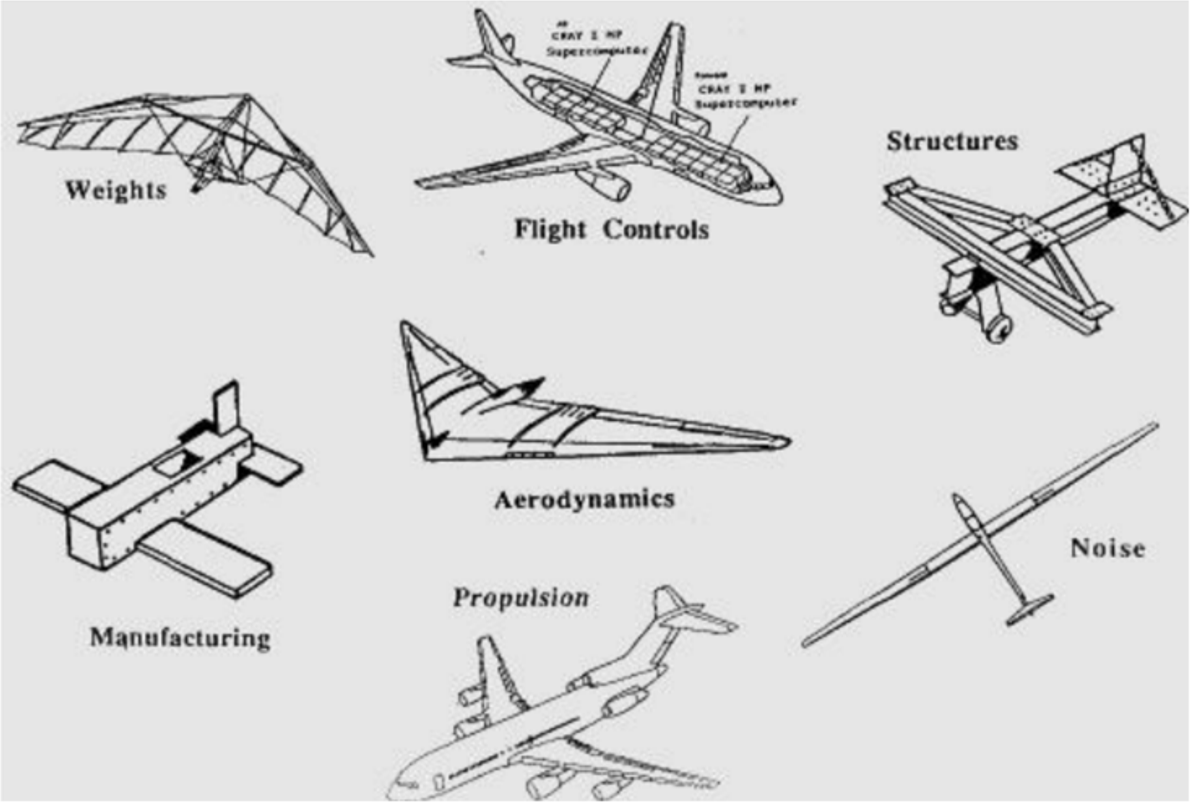
Aerospace

Aerodynamics / Controls / Structures / Avionics / Propulsion

AFM

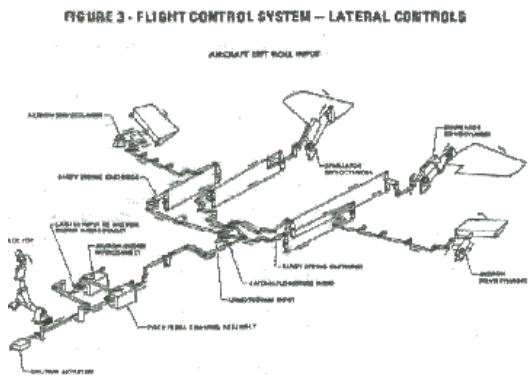
$\boxed{\text{DESIGN}}$

max. S_1 but $S_1 \geq S_1^0$ (DGCA)

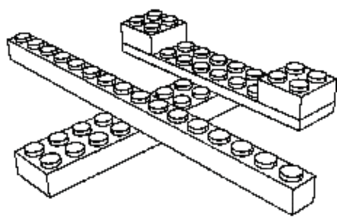


Aerodynamics group

Propulsion group

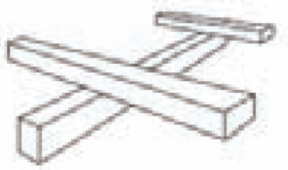


Flight control systems group



Manufacturing group

Stealth group



Structure group

MDO formalises the methodology of give and take between various disciplines.

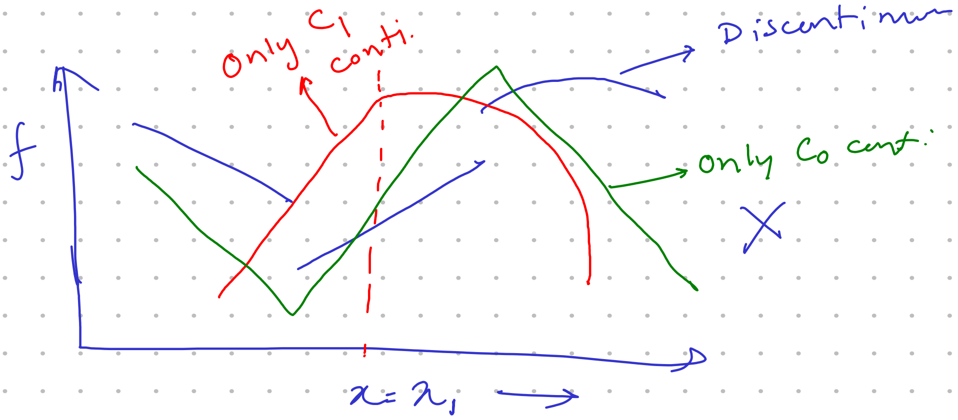
We begin our journey with 1D Unconstrained Optimisation problem.

$$\min_{\bar{x} \in \mathbb{R}^N} f(\bar{x}) \qquad (N=1)$$

What about $\bar{x} \in \mathbb{Z}^N$ (Integers) E.g. How many engines should I have?

$\bar{x} \in \mathbb{B}^N$ (Boolean) E.g. Should I use canard configuration?

$x \in \mathbb{R}^N \rightarrow$ Continuous opti.



C_0 continuous $\rightarrow$ f is continuous.

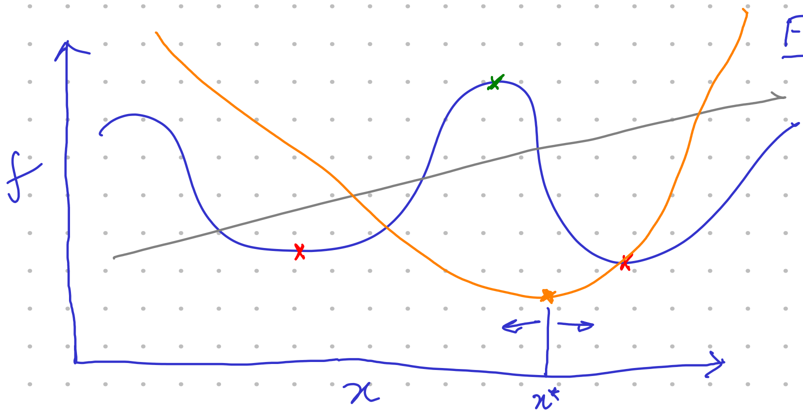
$\frac{df}{dx}$ is continuous — C_1 continuous

$\frac{d^2f}{dx^2}$ is " — C_2 conti

⋮

All derivatives are continuous — Analytic function.

$\underbrace{e^x, \sin(x), \cos(x)}$



For local optima

Necessary

$\frac{df}{dx} = 0$

↳ Stationary point.

Sufficient.

Minima $\frac{d^2f}{dx^2} > 0$, $\frac{df}{dx} = 0$

$f(x)$ Affine function (Linear function) $ax+b$ → Linear programming

$f(x)$ Quadratic function ax^2+bx+c

(1 stationary point) → Quadratic convex program

Any minima/maxima = Extrema is the GLOBAL extrema.

$f(x)$ | Cubic ax^3+bx^2+cx+d

General nonlinear function [Nonlinear optimisation]

Lets move from 1D to N-D.

Necessary condition for local optima :- $\frac{\partial f}{\partial x_i} = 0 \quad \forall i \in [1, n]$

$$\begin{bmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \\ \vdots \\ \partial/\partial x_n \end{bmatrix} f = 0 \implies \nabla f = 0$$

Sufficient condition " " " :- $\frac{\partial^2 f}{\partial x_i^2} > \frac{\partial^2 f}{\partial x_i \partial x_j}$ $i, j \in [1, n]$

1D → 2D

$\frac{d^2f}{dx^2} > 0$

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$ = Hessian matrix has to be positive definite.

Introduction to Optimum Design — Jasbir Arora

Optimization for Engg. Design — Kalyanmoy Deb (~2005)

Prof. Alonso (Stanford) — Notes from 2012.

Lec 3

Design a bucket of circular cross-section that can store at least 10 litres of water.

→ Design variables. h – height of bucket
r – radius " "

Obj function $f(h,r) = (\pi r^2 + 2\pi r h)$

Constraint $V(h,r) \geq 10$
$\pi r^2 h \geq 10$
$h_1 \leq h \leq h_2$
$r_1 \leq r \leq r_2$

Assumption
① Bucket is cylindrical.
② Cost of bucket ∝ surface area of bucket
③ All other obj. are ignored.

$V(h,r) = \pi r^2 h$

(r^3, h^3)
↑
(r, h) → (d, h)
→ $(h+2, r+1)$
$H = h+2, \; R = r+1$
$\Rightarrow h = H-2, \; r = R-1$

$(r, h) \Rightarrow (1, 2)$
$(d, h) \Rightarrow (2, 2)$
$(h+2, r+1) \Rightarrow (2+2, 1+1)$
$(4, 2)$

$f(h,r) = $ Same.

(x_1, x_2) ⇓ All linear combinations are also possible.

$X_1 = \alpha x_1 + \beta$
$X_2 = \gamma x_2 + \delta$

→ Matlab "fmincon".
Also read, the documentation of fmincon.
List names of of algos in fmincon.
List stopping for "fmincon".

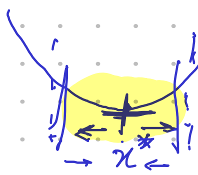
By Friday. 12pm.

Lec 4 :-

Positive definitive.

1D unconstrained $\quad \frac{df}{dx} = 0, \quad \frac{d^2f}{dx^2} > 0$

$$f(x^* + \Delta x) = f(x^*) + \left.\frac{df}{dx}\right|_{x=x^*} \cdot \Delta x + \frac{1}{2} \left.\frac{d^2f}{dx^2}\right|_{x=x^*} \cdot (\Delta x)^2 + HOT$$



$(\bar{x}^*)$

$f(x) \geq f(x^*)$

Local minima.

$$\Delta f = f(x^* + \Delta x) - f(x^*) = \underbrace{\frac{df}{dx} \cdot \Delta x}_{} + \underbrace{\frac{1}{2} \frac{d^2f}{dx^2} \cdot (\Delta x)^2}_{(\Delta x)^T H \Delta \bar{x}}$$

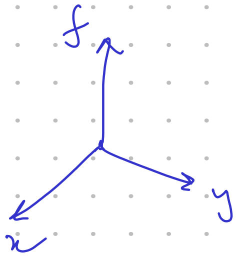
$\Delta f > 0$

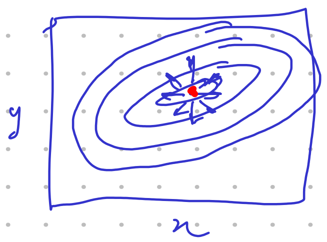
$\frac{df}{dx} = 0 \qquad \frac{d^2f}{dx^2} > 0$

$\Delta x = 10^{-10}$

$(\Delta x)^2 = 10^{-20}$

Quadratic forms





$$f(\bar{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} x_i x_j$$

$$f(x_1, x_2) = P_{11} x_1^2 + P_{12} x_1 x_2 + P_{21} x_2 x_1 + P_{22} x_2^2 = P_{11} x_1^2 + (P_{12} + P_{21}) x_1 x_2 + P_{22} x_2^2$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} - \alpha \\ P_{21} + \alpha & P_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

If $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then $f(x_1, x_2) = \underset{1 \times n}{\bar{x}^T} \underset{n \times n}{P} \underset{n \times 1}{\bar{x}}$

$$f(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underbrace{\begin{bmatrix} P_{11} & (P_{12} + P_{21})/2 \\ (P_{12} + P_{21})/2 & P_{22} \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \bar{x}^T A \bar{x}$$

Taylor series exp. n^{th} fun

$$f(\bar{x}^* + \Delta \bar{x}) = f(\bar{x}^*) + \nabla f(\bar{x}^*)^T \cdot \Delta \bar{x} + \frac{1}{2} \Delta \bar{x}^T H \Delta \bar{x} + HOT$$

$F(x) = \underline{x_1^2 + 2x_2^2 + 7x_3^2 + 2x_1x_2 + 4x_2x_3 + 14x_3x_1}$

$$P = \begin{bmatrix} 1 & 0.5 & 10 \\ 1.5 & 2 & 1 \\ 4 & 3 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 1 & 7 \\ 1 & 2 & 2 \\ 7 & 2 & 7 \end{bmatrix}$$

If matrix (A) is positive definite then $\Delta\bar{x}^T A \Delta\bar{x} > 0$

" semidefinite " $\Delta\bar{x}^T A \Delta\bar{x} \geq 0$

negative definite

" semi-definite

Indefinite

$\bar{x}^T A \bar{x}$

$f(x) = x_1^4 + x_2^4 + 4x_1^3 x_2^3$

If at $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, H is positive definite ?

|| Dharmik
Rajkumar Jha.
Mital Patel

$\rightarrow \quad \nabla f(x) = \begin{bmatrix} 4x_1^3 + 12x_1^2 x_2^3 \\ 4x_2^3 + 12x_1^3 x_2^2 \end{bmatrix} \Rightarrow \nabla^2 f(x) = \text{Hessian} = H(x) = \begin{bmatrix} 12x_1^2 + 24x_1x_2^3 & 36x_1^2x_2^2 \\ 36x_1^2x_2^2 & 12x_2^2 + 24x_1^3x_2 \end{bmatrix}$

At $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $H(x) = \begin{bmatrix} 36 & 36 \\ 36 & 36 \end{bmatrix} = 36 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

Characteristic eq^n:- $(1-\lambda)^2 - 1 = 0 \Rightarrow \lambda(\lambda - 2) = 0 \Rightarrow \lambda = 0, 2$

Positive semidefinite.

Q:- Find a quadratic approx. for the above function at (2,2). Write down the entire derivation.

Q:- Now, if $\nabla f(\bar{x}^*) = 0$ and $\text{eig}(\nabla^2 f(\bar{x}^*)) = 0$, then can $\bar{x}^*$ be a minima?

A: The lowest non-zero derivative must be even ordered for a stationary point and positive for a local minimum point.

$f(x)$ $\quad \left.\frac{df}{dx}\right|_{x^*} = 0$, $\quad \left.\frac{d^2f}{dx^2}\right|_{x^*} = 0$ $\quad \left.\frac{d^3f}{dx^3}\right|_{x^*} = 0$ $\quad \left.\frac{d^4f}{dx^4}\right|_{x^*} > 0$

→ Can I say x^* is not a minima.

→ I cannot comment.

Explicit (b) $y = z^2$, $z = x^3$, $\Rightarrow$ $y = x^6 \Rightarrow y = f(x)$

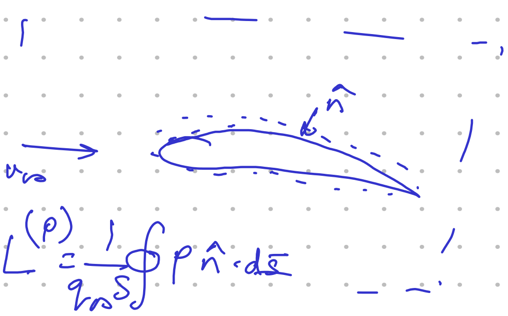
y ↑ Dependent, z ↑ intermediate variable, x ↑ independent

(a) $y = \sin(x) \Rightarrow$

Implicit

$$C_L(\bar{x}) \equiv C_L(p(\bar{x}))$$

$$\underline{GE(\bar{p}, \bar{x}) = 0}$$



$$C_L(p) = \frac{1}{q_\infty S} \oint p \, \hat{n} \cdot d\bar{s}$$

GE = Euler / N·S

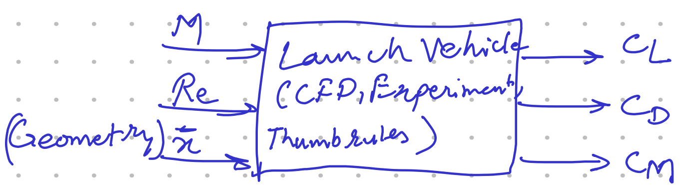
$$y = z^2, \quad —①$$

$$\frac{dz}{dx} + \sin(x) \cdot \cos(z) = 0 \quad —②$$

$x = x^* \xrightarrow{②} z^* \xrightarrow{①} y^*$

$y(x) =$

Black-Box → Don't know if this function is C^0, C^1, or C^2 continuous.



$$C_L(\vec{x})$$

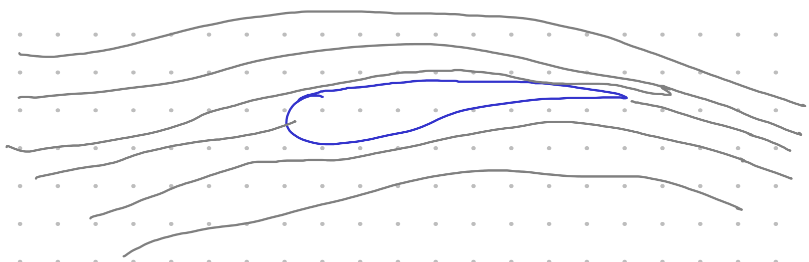
↑ independent

Lec 5: (04/02/21)

Recap: Given the functional form of an objective function f(x), we know the necessary and sufficient condition to find a minima.

Q: Do we always know f(x) explicitly? (Explicit, implicit & black-box functions)

Let us consider a standard Aerospace optimisation problem.

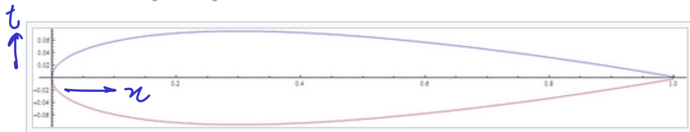


Design an airfoil that generates minimum drag (C_D) at
Cl=1.2, (C_L)
AOA=3 degrees, (α)
mach=0.6, (M)
11 km ISA conditions.

$$\min_{\bar{x} \in \mathbb{R}} C_D(\bar{x})$$
subject to
$$C_L(\bar{x}, \alpha, M) = 1.2$$

Equation for a symmetrical 4-digit NACA airfoil [edit]

The formula for the shape of a NACA 00xx foil, with "xx" being replaced by the percentage of thickness to chord, is[4]



Plot of a NACA 0015 foil generated from formula

$$y_t = 5t\left[0.2969\sqrt{x} - 0.1260x - 0.3516x^2 + 0.2843x^3 - 0.1015x^4\right],$$ [5][6]

where:

x is the position along the chord from 0 to 1.00 (0 to 100%),
y_t is the half thickness at a given value of x (centerline to surface),
t is the maximum thickness as a fraction of the chord (so t gives the last two digits in the NACA 4-digit denomination divided by 100).

Source: Wikipedia.

Lec 6 :- 05/02/21

Descent methods. (Gradient descent methods) [Only using 1^{st} order information i.e. only using $\nabla f(\bar{x}^{(k)})$]

Do not use $H(\bar{x}^{(k)})$

Problem statement $\min_{\bar{x} \in \mathbb{R}^N} f(\bar{x})$

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

$$\nabla f = \underbrace{\begin{bmatrix} 2x_2 + 2x_1 \\ 2x_1 \end{bmatrix}}_{2\times 1}$$

$N \times 1$

→ $\nabla f(\bar{x}) = \bar{c}$
→ $\bar{x}^* -$ minima for $f(\bar{x})$
→ $\underset{N\times 1}{\bar{d}}$ - direction vector.

1. Initialise $\bar{x}^{(0)}$ (randomly/heuristic/Prior experience)
2. Calculate $\nabla f(\bar{x}^{(0)})$
3. $\bar{d}^{(0)} = \underbrace{F(\nabla f(\bar{x}^{(0)}))}$
4. Calculate stepsize $\alpha^{(0)}$
5. $\bar{x}^{(1)} = \bar{x}^{(0)} + \alpha \bar{d}^{(0)}$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha^{(k)} \cdot \bar{d}^{(k)}$$

$$\|\nabla f(\bar{x}^{(k+1)})\|_2 \leq \underbrace{10^{-6}}_{\text{Tolerance}} \Rightarrow \text{Stoppage criteria}$$

③ Choice of descent direction $-\nabla f$

(A.) $\bar{d}^{(k)} = -\nabla f(\bar{x}^{(k)}) \Rightarrow$ Steepest descent method.

(B) $\bar{d}^{(k)} = -\begin{bmatrix} 0 \\ 0 \\ \frac{\partial f(\bar{x}^{(k)})}{\partial x_i} \\ 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow$ Coordinate " "

(C) $\bar{d}^{(k)} = \begin{bmatrix} 0 \\ 0 \\ -\frac{\partial f}{\partial x_i}(\bar{x}^{(k)}) \\ \vdots \\ 0 \end{bmatrix}$ where i is chosenly randomly. (Uniform/Normal) $\Rightarrow$ Stochastic " "

(D) Slight variation (A)

Descent direction

$\bar{d}^{(k)}$

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha^{(k)} \cdot \bar{d}^{(k)}$$

$$f(\bar{x}^{(k)} + \alpha^{(k)} \bar{d}^{(k)}) \approx \underbrace{f(\bar{x}^{(k)}) + \alpha^{(k)} \cdot \underbrace{\nabla f(\bar{x}^{(k)})^T}_{} \cdot \bar{d}^{(k)}}_{} \quad \text{— ①}$$

$$\bar{C}^{(k)^T} \cdot \bar{d}^{(k)}$$

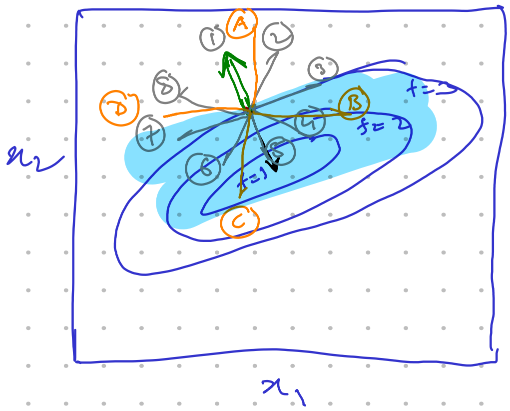
After every step, sol^n has to improve. — ②

i.e. $$f(\bar{x}^{(k+1)}) \leq f(\bar{x}^{(k)})$$

From ① & ②,

$$\bar{C}^{(k)^T} \cdot \bar{d}^{(k)} \leq 0$$

→ If $\bar{c}^{(k)^T} \cdot \bar{d}^{(k)} = 0 \implies$ Exact stepsize calculation



f $\qquad \nabla f(\bar{x}^{(k)})^T \cdot \bar{d}^{(k)} \leq 0$

$\bar{x}^{(k)}$

$\nabla f(\bar{x}^{(k)})$

$$\underbrace{\bar{\nabla f} \cdot \bar{d}}_{} = \underbrace{|\bar{\nabla f}| |\bar{d}|}_{} \cdot \underbrace{\cos(\angle_{\nabla f, d})}_{}$$

3, 7 → $\nabla f \cdot d = 0$

1, 2, 8 → $\nabla f \cdot d > 0$

4, 5, 6 → $\nabla f \cdot d < 0$

5 — steepest descent.

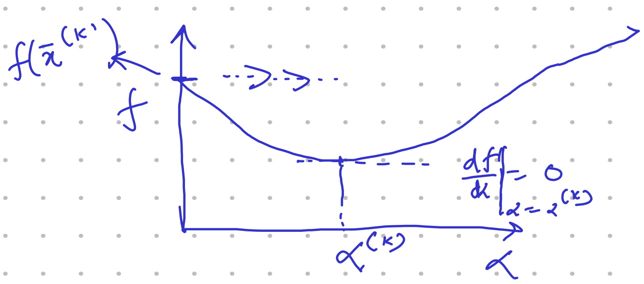
$$\bar{d}^{(k)} = -\nabla f(x^{(k)}) = -C^{(k)}$$

1. Steepest descent
2. Coordinate "
3. Stochastic "

Lec 7 :- **Stepsize calculation.**

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha^{(k)} \bar{d}^{(k)}$$

$$f\left(\bar{x}^{(k+1)}\right) = f\left(\bar{x}^{(k)} + \alpha^{(k)} \cdot \bar{d}^{(k)}\right) = f(\alpha)$$



for $\alpha = 0$, $f(x^{(k+1)}) = f(\bar{x}^{(k)})$

$$\left.\frac{df}{d\alpha}\right|_{\alpha=\alpha^{(k)}} = 0 \Rightarrow \text{Exact step-size}$$

Eg. $f(\bar{x}) = 0.1 x_1^2 + x_2^2 - 10$, $x^{(0)} = (5, 1)$

$\rightarrow$ $\bar{c}^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\bar{d}^{(0)} = -\bar{c}^{(0)} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$

$\alpha^{(0)} = ?$

$$f(\alpha) = f\left(\begin{bmatrix} 5 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ -2 \end{bmatrix}\right) = f(5-\alpha, 1-2\alpha)$$

$$= 0.1(5-\alpha)^2 + (1-2\alpha)^2 - 10$$

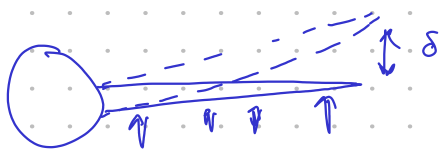
$\frac{df}{d\alpha} \leq 10^{-15}$

$\alpha = 0.64$

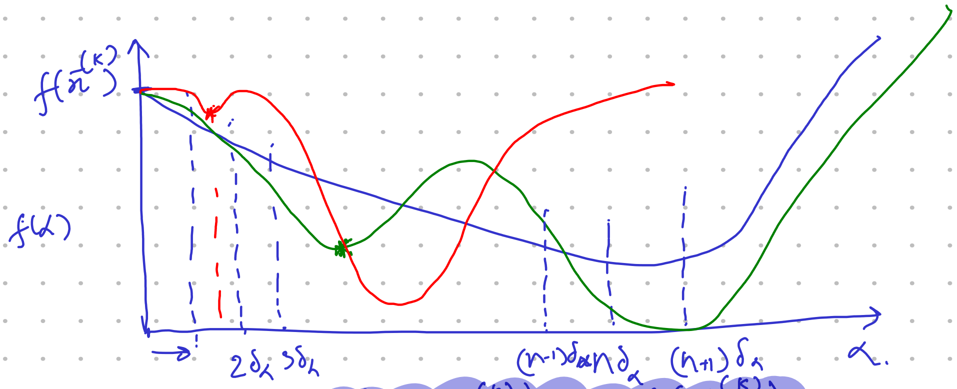
$x^{(1)} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + 0.64 \begin{bmatrix} -1 \\ -2 \end{bmatrix} =$

$f(\bar{x}) = \delta$

$\bar{x}$ - wing geometry



Equal interval search.



δ_α

$$f(\bar{x}^{(k)} + \delta_\alpha \cdot d^{(k)}) \leq f(\bar{x}^{(k)})$$

① $\Rightarrow$ $f(\bar{x}^{(k)} + (n+1)\delta_\alpha \cdot \bar{d}^{(k)}) < f(\bar{x}^{(k)} + n\delta_\alpha \bar{d}^{(k)})$

② $\Rightarrow$ $f(\bar{x}^{(k)} + (n+1)\delta_\alpha \bar{d}^{(k)}) \geq f(\bar{x}^{(k)} + n\delta_\alpha \bar{d}^{(k)})$

$$\boxed{\alpha^{(k)} = n\delta_\alpha}$$

Opti - 100 steps X 10 (stepsize) = 1000 (function evals.)

1 CFD solution

1 FEM sol^n.

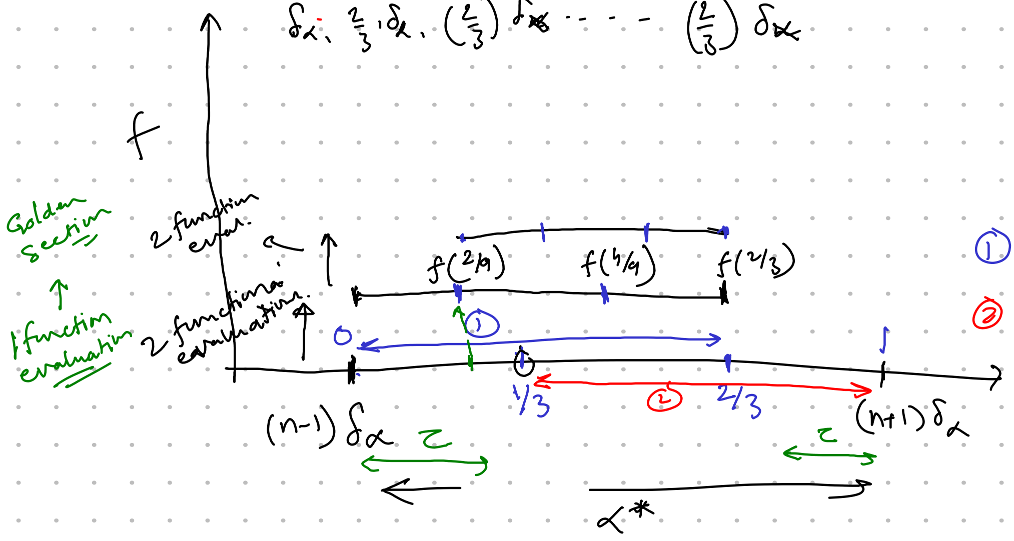
② Golden search method.

$\rightarrow$

③ Approximate step-size. (Armijo's)

Lec 8 :- 15/02/21

$\delta_\alpha,\ \frac{2}{3}\delta_\alpha,\ \left(\frac{2}{3}\right)^2\delta_\alpha \ \cdots\cdots \ \left(\frac{2}{3}\right)^{n-1}\delta_\alpha$



$f(0)$
$f(1/3)$
$f(2/3)$
$f(1)$

① $f(0) \geq f(1/3) \leq f(2/3)$

② $f(1/3) \geq f(2/3) \leq f(1)$

τ = Golden ratio

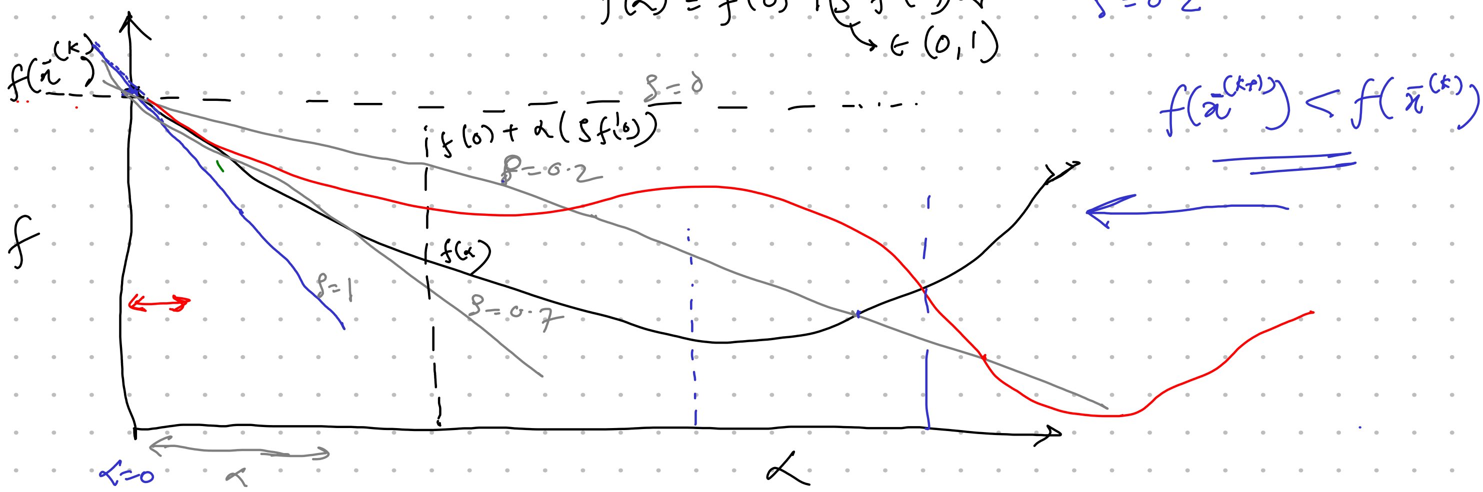
Accuracy Vs. Computational Cost

→ Golden section method.

Jasbir Arora. [Read]

→ Inexact line search. (Armijo's Rule)

$f(\alpha) = f(0) + \rho f'(0)\alpha$, $\rho \in (0,1)$ $\quad \rho = 0.2$



$f(\bar{x}^{(k+1)}) < f(\bar{x}^{(k)})$

If $f(\alpha) \leq f(0) + \alpha(\rho f'(0))$ then check if
$f(\eta\alpha) \leq f(0) + \eta\alpha(\rho f'(0))$, $\eta = 2, 3, 4, 5 \ldots$

If $f(\alpha) \geq f(0) + \alpha(\rho f'(0))$ then
$f(\alpha/\eta) \leq f(0) + \alpha/\eta\,(\rho f'(0))$, $\eta = 2, 3, 4 \ldots$

How to judge opti. Algos.

1. Accuracy (Convergence)
2. Time (wall clock time) $\equiv$ # of function evaluations
3. Global minima (Ensure)
4. Cost. $\to$ Parallel opti Algo.

<u>Steepest Descent</u> Vs. Coordinate descent.

$$\min.\ f(x_1, x_2, x_3) = x_1^2 + (x_2 - 10)^2 + e^{-(x_1^2 + \sin(x_2 x_3))}$$

10 iters. each SD & CD.

Let us assume that $\alpha^{(k)}$ calculation requires 10 function evaluation in both.

FD is used to calculate gradient.

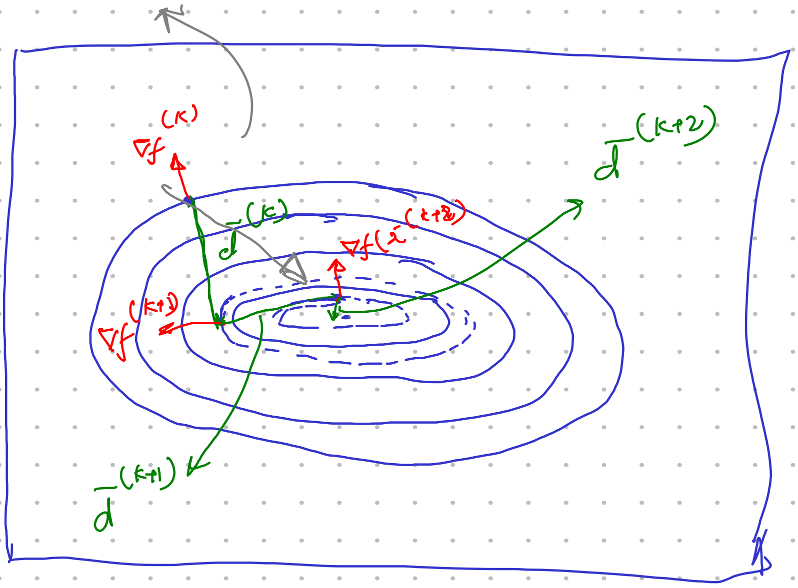
$\rightarrow$ k^{th} iter.

$$f(\bar{x}^{(k)}), \quad \nabla f = \begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \partial f/\partial x_3 \end{bmatrix} \quad \frac{f(\bar{x}^{(k)} + \Delta x_1) - f(\bar{x}^{(k)})}{\Delta x_1} \quad \alpha^{(k)}$$

SD: 1 + 3 + 10 = 14 × 10 = 140 func.

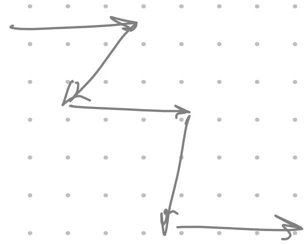
CD: 1 + 1 + 10 = 12 × 10 = 120 funct.

Conjugate Gradient



$$f(x_1, x_2) = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - 1$$

$$f(x_1, x_2) = x_1^2 + x_2^2$$



Exact step size calculation.

$$\bar{c}^{(k+1)^T} \cdot \bar{d}^{(k)} = 0$$

Gradient at $(k+1)^{th}$ iter. is orthogonal to descent direction at $(k)^{th}$ iter.

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha \bar{d}^{(k)}$$

$$f(\bar{x}^{(k+1)}) = f(\bar{x}^{(k)} + \alpha \bar{d}^{(k)})$$

$$\frac{df}{d\alpha} = 0 \Rightarrow \frac{df(\bar{x}^{(k+1)})}{d\alpha} = 0$$

$$\Downarrow$$

$$\nabla f(\bar{x}^{(k+1)}) \cdot \bar{d}^{(k)} = 0$$

$$\Downarrow$$

$$\boxed{\bar{c}^{(k+1)} \bar{d}^{(k)} = 0}$$

$$\min_{\bar{x} \in \mathbb{R}^n} f(\bar{x})$$

$\left.\begin{array}{l} n = 10 \\ n = 100 \end{array}\right\}$ SD. $\rightarrow$ Computational cost $(n=10)$
$\rightarrow$ " $(n=100)$

Algo 1

- $\rightarrow \nabla f$
- $\rightarrow$ Forward FD.
- $\rightarrow$ Ignore step-size calc.
- $\rightarrow$ # of function evaluation (# of design variables)

$n \rightarrow n$ functiona evaluations

Algo 2.

- $\rightarrow \nabla f, \nabla^2 f$
- $\rightarrow$ Forward FD.
- $\rightarrow$ Ignore step-size Cal.
- $\rightarrow$

n function evaluation (∇f) + $(\nabla^2 f)$

$$\frac{d^2 f}{dx^2} = \frac{d}{dx}\left[\frac{f(x+\Delta x) - f(x)}{\Delta x}\right] = \frac{f(x+2\Delta x) - f(x+\Delta x) - f(x+\Delta x) + f(x)}{(\Delta x)^2}$$

$$\frac{d^2f}{dx^2} = \frac{f(x+2\Delta x) - 2f(x+\Delta x) + f(x)}{(\Delta x)^2}$$

$$\frac{d^2f}{dx_i\,dx_j} =$$

$$\left(\frac{n(n+1)}{2}\right)$$

Lec 10 :-

① Discussion of stopping criteria

1. $\|f^{k+1} - f^{k}\|_2 < \epsilon$
2. $\|\bar{x}^{k+1} - \bar{x}^{k}\|_2 < \epsilon$
3. $\|\nabla f\|_2 < \epsilon \; (10^{-6}, 10^{-12})$
4. # function evaluations $\leq N$; # iterations.

② Slow convergence of SD.

Post optimality analysis.

Rate of convergence

$\bar{x}^{(0)}, \bar{x}^{(1)}, \bar{x}^{(2)}, \bar{x}^{(3)}, \ldots, \bar{x}^{(k)}, \bar{x}^{(k+1)}, \ldots, \bar{x}^{(N)}$

$\bar{x}^*$ $\qquad \bar{x}^* + \epsilon$

$x^* = 0$

Opti 1 → $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$

Opti 2 → $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$

Opti 3 → $1, \frac{1}{2^2}, \frac{1}{2^4}, \frac{1}{2^8}, \frac{1}{\cdot}, \ldots$

Rate of convergence.

$$\frac{\|\bar{x}^{(k+1)} - \bar{x}^*\|_2}{\|\bar{x}^{(k)} - \bar{x}^*\|_2^{\,r}} = \beta$$

$\bar{x}^* = 0 \qquad \beta = 1.2, \; r = 1$

$\bar{x}^{(0)} = 1 \qquad \|x^{(1)} - \bar{x}^*\| = \|x^{(0)} - x^*\| \times 1.2$

$r = 1, \; 0 < \beta < 1$ — Linear convergence.

— superlinear ..

$r = 2$ — Quadratic convergence

$r = 3$

$r = 4$

Two ways to estimate the rate of convergence :-

① Experimental. (Assumes that you know $\bar{x}^*$).
Usually, $\bar{x}^{(N)}$ is assumed to be the $\bar{x}^*$ and then the rate of convergence is found out.

② Asymptotic theoretical analysis under simplifying assumptions.

① Quad. - bad convergence const.

$x^* = 0$

$x^0 = 1$

$r = 2$

$\beta = 0.9$

$x^{(1)} = \beta \cdot (x^{(0)})^2 = 0.9 \times 1 = 0.9$

$x^{(2)} = \beta \cdot (x^{(1)})^2 = 0.9 \times (0.9)^2 = 0.729$

$x^{(3)} = (0.9) \cdot (0.729)^2 =$

⋮

② Quadratic - good convergence const.

$\beta = 0.1$

$x^{(1)} = 0.1 \times (1)^2 = 0.1$

$x^{(2)} = 0.1 \times (0.1)^2 = 0.001$

$x^{(3)} = 0.1 \times (10^{-3})^2 = 10^{-7}$

⋮

Quadratic

③ Good initial guess

$x^0 = 0.1$

$\beta = 0.1$

$x^{(1)} = 0.1 \times (0.1)^2 = 10^{-3}$

$x^{(2)} = 0.1 \times (10^{-3})^2 = 10^{-7}$

⋮

∇f info - 1^{st} order info. — 1 (linear conv.)

∇f, approx $\nabla^2 f$ - 1^{st} order info + history — $1 < r < 2$ (superlinear)

∇f, $\nabla^2 f$ info. - 1^{st} & 2^{nd} order info - 2 (Quadratic)

$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$

Newton interpolation

x_1, x_2, x_3

$f(x_1)\ f(x_2)\ f(x_3)$

$$\frac{df}{dx} = \frac{f(x+\Delta x) - f(x)}{\Delta x} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

$\Delta f(x_1) = f(x_2) - f(x_1)$

Δ

$x^{(0)}, x^{(1)}, x^{(2)}, \ldots \ldots x^{(n)} \to x^*$ — β, r

$$\left(x^{(k+1)} - x^{(n)}\right) - \beta_0 \left(x^{(k)} - x^{(n)}\right)^{r_0} = 0 \quad \forall\ k = 1, \ldots, n-2$$

(Least squares method)
fit

$$-2 \le x \le 2 \Rightarrow 0 \le y \le 1$$

$$y = \frac{x+2}{4}$$

$$f(x^{(k)} + \Delta x) = f(x^{(k)}) + c^T \Delta x + \frac{1}{2} \Delta x^T H \Delta x$$

$\Delta \bar{x} = \alpha \bar{d}$ GD
$= -\alpha \bar{c}$ SD
$= -H^{-1} \bar{c}$ (NS)

$$\frac{\partial f}{\partial \Delta \vec{x}} = 0 \Rightarrow \bar{C} + H \Delta \bar{x} = 0$$

$$H \Delta \bar{x} = -\bar{c}$$

$$\boxed{\Delta \bar{x} = -H^{-1} \bar{c}}$$

$\bar{d} = -\vec{\nabla f}$

$\vec{d_1}^T \cdot \bar{c} \le 0$

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \Delta \bar{x}^{(k)} = \bar{x}^{(k)} - H^{-1} \bar{c}$$

$$\Delta \bar{x}^T \bar{c} \le 0 \Rightarrow -\bar{c}^T (H^{-1})^T c \le 0$$

$$c^T \Delta x \le 0 \Rightarrow -c^T H^{-1} \bar{c} \le 0$$

$$\Delta \bar{x} = \bar{d}^{(k)} = -H^{-1} \bar{c} \Rightarrow \bar{c}^{(k)} = -H \bar{d}^{(k)}$$

$$\bar{c}^T \bar{d} \le 0$$

$-\bar{d}^T H \bar{d} \le 0 \Rightarrow H$ should be +tive definite.

In the neighbourhood of the minima, H is +tive definite.

→ Newton direction is a descent direction in the neighbourhood of minima.

Lec 12 :-

Q:- But what happens away from the optima?

→ We cannot guarantee that the Newton direction is a descent direction.

→ One option is instead of doing

$$x^{(k+1)} = x^{(k)} - \underbrace{H^{-1(k)} \cdot \bar{c}^{(k)}}_{\text{Newton step}}$$

we do,

$$\bar{x}^{(k+1)} = x^{(k)} - \underbrace{\alpha_k H^{-1(k)} \bar{c}^{(k)}}_{\text{Modified Newton step.}}$$

where α_k can be found from Armijo's rule.

If $\alpha_k = 1$, ⇒ Pure Newton's method. If $\alpha_k \ne 1$, modified Newton's method.

Second approach is to modify the Hessian and make it + tive definite.

Define $B^{(k)} = H^{(k)} + \Delta H^{(k)}$, and use

$$x^{(k+1)} = x^{(k)} - \alpha_k {B^{(k)}}^{-1} \bar{c}^{(k)}$$

$\rightarrow$ Quasi-Newton methods.
- Direct update — Approx. $H^{(k)} \approx B^{(k)}$
- Indirect update — Approx. $H^{-1} \approx A^{(k)}$

Q:- What is the fastest method of calculating $d^{(k)}$ using Newton eq^n

$$H^{(k)} \bar{d}^{(k)} = -\bar{c}^{(k)}$$

$$\Downarrow$$

$\rightarrow$

$$M \bar{x} = \vec{b}$$

$\rightarrow$ Gauss-Elemination

$\rightarrow$ Jacobi (point-line)

$\rightarrow$ PSOR

$$\bar{d}^{(k)} =$$

$$\frac{A-B}{\Delta x} = \frac{\frac{A^2-B^3}{A+B}}{\Delta x} \qquad \frac{A^2-B^3}{A+B}$$

$$\approx \frac{-\Delta A}{\Delta x}$$

Internal Analysis

$$\approx \frac{A^2-(A+\Delta A)^2}{2A+\Delta A}$$

$$= \frac{A^2-A^2-2A\Delta A-(\Delta A)^2}{2A+\Delta A}$$

$$\Downarrow$$

$$= -\frac{\Delta A\,(2A+\Delta A)}{2A+\Delta A}$$

$$= \frac{-\Delta A}{\Delta x}$$

Finite difference :-

→ 16 digits. (Matlab) → Excel, C, Julia, Python

$f(x) = \sin(x)$

$x \in \mathbb{R}$

$f(x) \in \mathbb{R}$.

$$\frac{df}{dx} \approx \frac{\sin(x_0+\Delta x) - \sin(x_0)}{\Delta x}$$

$$\frac{\sin(x_0+\Delta x) - \sin(x_0-\Delta x)}{2\Delta x}$$

Floating point Representation (IEEE-754)

16 digits

8 digits

$\pi = \underbrace{3.1415 \ldots\ldots}_{}\,|\ldots\ldots$ 2.0

$(\pm)a \cdot 10^b$

$1.01110\,1\ldots\ldots\ 1_{\,0}$ 53 = 64 bits (Double precision)

$\longleftarrow 64 \longrightarrow$

= 32 bits

double

float

Subtractive Cancellation error.

$\sin(x_0 + 10^{-64}) = 1.234\ldots|\ldots|$ 4

$\sin(x_0) =$

$0 \approx 10^{-16}$

Stability Analysis :-

$\bar{x}_0 \in \mathbb{R}^n$

# of functions ewations to calculate $\nabla f(x)$ using

F.D → $1+n \approx O(n)$ Time complexity.

$$\nabla f(x) = \begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix} \approx \begin{bmatrix} \frac{f(\bar{x}_0+\Delta x_1) - f(\bar{x}_0)}{\Delta x_1} \\ \frac{f(\bar{x}_0+\Delta x_2) - f(\bar{x}_0)}{\Delta x_2} \\ \vdots \end{bmatrix}$$

Central Difference → $2n \approx O(n)$

$$\frac{f(x_0+\Delta x) - f(x_0-\Delta x)}{2\Delta x}$$

This calculation assumes serial execution of the function evaluation. If parallel computing is available then time complexity is $O(1)$.

Space complexity for serial execution – $O(1)$

parallel " – $O(n)$

$\bar{x} \longrightarrow \boxed{\text{Black-box objective function.}} \longrightarrow f(\bar{x})$

Finite difference is the only choice.

→ Step-size calculation is done prior to starting the opti. Algo.

Space complexity :- $f(x) = L(\vec{x})$

10^6 grid points.

$\to \underbrace{\rho, P, \bar{v}}_{5} \longrightarrow 5 \times 10^6$ numbers.

$\to \underbrace{5 \times 10^6 \times 2.5}$ double variable

↓ 4 Bytes

Memory required to execute one function evaluation ← $\boxed{4 \times 5 \times 10^6 \times 2.5}$

(CFD or FEM solⁿ)

$\boxed{50\,MB}$

Serial execution :- $f(\bar{x}_0), f(\bar{x}_0 + \Delta x_1), f(\bar{x}_0 + \Delta x_2), \ldots, f(\bar{x}_0 + \Delta x_n)$

$O(1)$

Parallel " - $50\,MB + 50\,MB + \cdots \cdots + 50 = 50n\,MB$

$O(n)$

Assignment :- Space time

serial

$H(\bar{x}_0)$

parallel.

using F-D.