Computational Fluid Dynamics — Prof. Suman Chakraborty (1)

Lecture 1: Introduction to CFD

CFD: Computational Fluid Dynamics

Not just limited to Fluid Dynamics, but to any general transport phenomena (like heat transfer, mass transfer)

CFD applications:

- Aerospace — Interior: ventilation system, combustion engine
  Exterior: flow over A/C.
- automobile
- biomedical
- chemical — (mixing of chemicals ≠ bubbles at the interface, separation & mixing, injection of streams).
- Electronics — efficient cooling strategies.
- Energy -
- Fluid structure interaction (two way coupling & dynamic coupling - fluid interacts with structure & structure interacts with fluid)
- Marine (eg: flow past ships & boats)
- Materials processing (~~cast~~ grain growth etc; CFD need not be deterministic approach; can be stochastic like Monte Carlo as well, but welding etc, mold filling etc
  ↳ helps choose the correct filling process to avoid casting defects etc.
  find out if the distribution of impurities via their convection across the material is okay for the material itself)
- Micro fluidics — microscale fluid flows (micron or sub-micron scales)
  (eg: mix two streams — good mixing via pulsating flows, droplet dynamics — micro reactor studies,
  Fluid structure interactions at small scales, flap placed in micro reaching flow — need to optimise the flap movement — flap acting like a mixer & a pump)

- sports: (racing cars, golf balls, running motion etc.)
- Turbomachines: (flow over blade passages

- Is CFD inevitable?

  Numerical vs Analytical vs Experimental.

  - Experimental investigation:
    - full scale (no substitute for this: seeing is believing)
      - expensive & often impossible
      - measurement errors.
    - on a scaled model
      - simplified (could upscale / downscale in such a manner that the flow physics doesn't changes & eg: micro capillary scaled up: at small scales, surface tension forces are significant) (need to maintane all similarities - kinematic, dynamic etc....
      - difficult to extrapolate results
      - measurement errors.
- Theoretical calculation:
  - analytical solutions:
    - if exists, gives us exact answers but exists only for a few cases.

{ CFD cannot stand on its own without experimental or analytical solutions. B/c we need to benchmark our solution }

    - Sometimes complex

  - numerical solutions:
    - exists for almost any problems.
    - Continuous nature of the problem is compromised, but at its expense we get answers to complex problems

Modeling vs Experimentation.

- Advantages of modeling:
  - cheaper
  - more complete information (all details of all variables can be obtained)
  - can handle any degree of complexity as long as.

- Disadvantages of modeling:
  - deals with a mathematical description not a reality
    Numerical solution is as good as the input fed to the problem
  - mathematical description can be inadequate
    (Governing eq^{ns} may / may not capture the correct physics).
  - multiple solutions can exist
    (non-linear problems may have multiple different solutions depending on the IC).

In conclusion: no real substitute for experimentation, but experimentation is limited by many restrictions & cannot handle multiple trials).

Usual plan of action: try analytical solution → do numerical simulation → create good experimental design & validate results from numerical simulation.

→ Cross validate the goodness of numerical solution.

$\boxed{\text{Lecture 2: Classification of PDEs.}}$

$$\frac{\partial}{\partial t}(\rho\phi) + \nabla\cdot(\rho\vec{V}\phi) = \nabla\cdot(\Gamma\nabla\phi) + S$$

— 2nd order PDE — from $\nabla\cdot(\Gamma\nabla\phi)$.

Consider a 2nd order PDE of the form:

$$A\phi_{xx} + B\phi_{xy} + C\phi_{yy} + \underbrace{D\phi_x + E\phi_y + F\phi + G}_{H} = 0 \quad —① $$

$\phi_x \to \frac{\partial\phi}{\partial x}$

$\phi_y \to \frac{\partial\phi}{\partial y}$

$\phi_{xx} \to \frac{\partial^2\phi}{\partial x^2}$

$\phi_{xy} \to \frac{\partial^2\phi}{\partial x\partial y}$

$\phi_{yy} \to \frac{\partial^2\phi}{\partial y^2}$

- A, B, C need not be consts. They can be fns / lower order partial derivatives / independent variables.
- $\phi = \phi(x,y)$.
- One classification —
  1. Linear : $A, B, C \to$ fns of x, y. remaining terms ~~D, E, F, G fns of~~ linear fns of ϕ, ϕ_x, ϕ_y.
  2. ~~Non linear~~ Quasi linear: $A, B, C \to$ fns of $x, y, \phi, \phi_x, \phi_y$.

- Characteristics of the PDE:

Highest order derivatives in a PDE may be continuous or discontinuous in the domain of consideration. There may be lines across which these highest order derivative could ~~be~~ have ~~are~~ discontinuities. Such lines are called characteristics lines of the PDE.

Why they're important? B/C in our numerical method, we have to account for such discontinuities apriori.

Objective: to get characteristics of the PDE

~~Goto~~ $\phi_x = \phi_x(x,y) \Rightarrow d\phi_x = \frac{\partial\phi_x}{\partial x}dx + \frac{\partial\phi_x}{\partial y}dy.$

$$d\phi_x = \phi_{xx}dx + \phi_{xy}dy \quad —② $$

$|||^{rly}$ $\phi_y = \phi_y(x,y) \Rightarrow d\phi_y = \phi_{yx}dx + \phi_{yy}dy.$

$$\because \phi_{xy} = \phi_{yx}, \quad d\phi_y = \phi_{xy}dx + \phi_{yy}dy. \quad —③ $$

From ①, ②, ③:

$$\begin{bmatrix} A & B & C \\ dx & dy & 0 \\ 0 & dx & dy \end{bmatrix} \begin{bmatrix} \phi_{xx} \\ \phi_{xy} \\ \phi_{yy} \end{bmatrix} = \begin{bmatrix} -H \\ d\phi_x \\ d\phi_y \end{bmatrix}$$

We are interested in the case where the solution of this matrix expression → $\begin{bmatrix} \phi_{xx} \\ \phi_{xy} \\ \phi_{yy} \end{bmatrix}$ doesn't exist.

Analogy with algebraic eq^{ns}:

$$\left.\begin{aligned} x+y &= 2 \\ 2x+2y &= 5 \end{aligned}\right\} \text{ sol}^n \text{ doesn't exist.}$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

↳ $\Delta = 0$.

We want to find locus of pts across which we have discontinuity in ϕ_{xx}, ϕ_{yy}, & ϕ_{xy}. That is possible when $\det\left(\begin{bmatrix} A & B & C \\ dx & dy & 0 \\ 0 & dx & dy \end{bmatrix}\right) = 0$

For ϕ_{xx}, ϕ_{yy}, ϕ_{xy} to be discontinuous,

$\Delta = 0$.

$$\begin{vmatrix} A & B & C \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = 0 \Rightarrow$$

{only coeffs of highest order derivatives matters}

$$A\,(dy)^2 - B\,(dx\,dy) + C\,((dx)^2) = 0$$

$$A\left(\frac{dy}{dx}\right)^2 - B\left(\frac{dy}{dx}\right) + C = 0$$

$$\frac{dy}{dx} = \frac{B \pm \sqrt{B^2 - 4AC}}{2A}.$$

Number of real characteristics existing will depend on whether $B^2 - 4AC \gtrless 0$ or < 0.

If $B^2 - 4AC = 0$ → only one real characteristic. (Parabolic).

$B^2 - 4AC < 0$ → no real characteristic (elliptic PDE)

$B^2 - 4AC > 0$ → 2 real characteristics (hyperbolic PDE)

Lecture 3 : Examples of PDEs

Ex 1 : $\nabla^2\phi = 0$ (Laplace eq^n - most commonly encountered simple PDE)

$\nabla^2 T = 0$ (T: Temperature, uniform heat conductivity, steady state no heat source).

2D example:

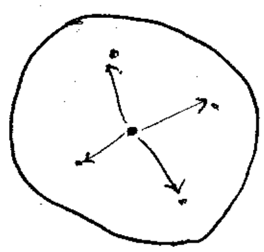
$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$$

$A = 1,\ B = 0,\ C = 1$

$B^2 - 4AC = -4 \longrightarrow$ elliptic equation

Say, we have a domain with uniform temperature throughout. Put a heat source at a pt $\rightarrow$ acts as a disturbance that propagates in all directions in the domain at infinite speed.

i.e. Thermal disturbance propagates in all directions with infinite speed.



This disturbance tries to nullify the temperature differences at different points in the domain to make the temp distribution homogeneous everywhere.

- So while numerically formulating the problem, a point on the domain under our consideration will be influenced by all the other points in our domain.

- B/C has to be specified at the boundary of the domain. The B/C can be discontinuous. For eg: half the boundary may be in contact with steam & the other half in contact with ice, so discontinuities are possible on the boundary. But since the disturbance travel at infinite speeds (message propagation is fast) means that there will not be discontinuities within the domain.

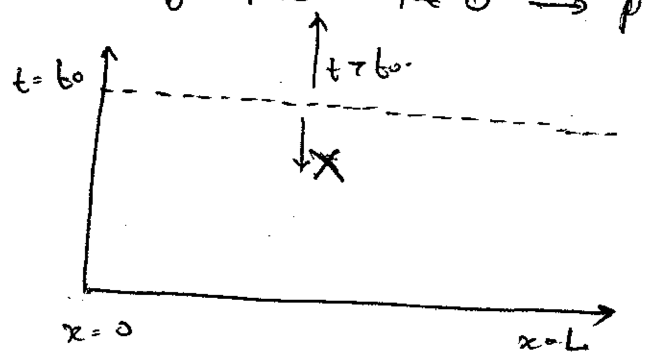
- This type of problem $\rightarrow$ BVP.

Ex-2 $\frac{\partial \phi}{\partial t} = \alpha \frac{\partial^2 \phi}{\partial x^2}$

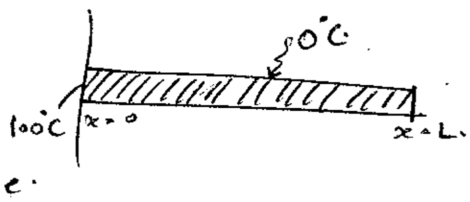
Eg of 1D unsteady heat conduction.

$A = \alpha$.
$B = 0$
$C = 0$

$B^2 - 4AC = 0 \rightarrow$ parabolic : one characteristic.



Say we have a 1D rod at uniform temp. At $t = t_0$, create a sudden disturbance in T at one end of the rod



Abrupt disturbance at $t = t_0$ will propagate in a direction forward in time.

The disturbance at $t = t_0$ will influence what will happen for $t > t_0$. It cannot influence back what has already happened sometime back.

~~advect~~ $\rightarrow$ time marching problems.

- will have only one type of discontuity. That discontinuty at reference time at which the distorbance is imposed.

- At time $t > t_0$, The abrupt distorbane originated at $t = t_0$ may have made its presence known Throughout some part of the domain. That part is called domain of influence.

- The total region is the doman where the presence of distorbance can potentially make its presence known is called domain of distorbance.

$t > t_0 \rightarrow$ domain of influence
~~for ellip~~ $t \le t_0 \rightarrow$ domain of disturbance
In elliptic case $\rightarrow$ entire domain is the domain of influence.

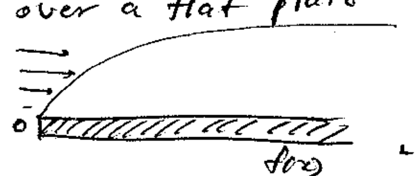
Initial-boundary value problem:
~~IBVP~~ since $t = t_0$ disturbance introduced.

As $t \to t_0$, steady state reached $\Rightarrow \frac{\partial}{\partial t} = 0$ or it becomes elliptic

So it has some elliptic nature build into it.

So more correct way of saying would be that this eq^n is parabolic in time and elliptic in space.

. Another eg: B/L over a flat plate

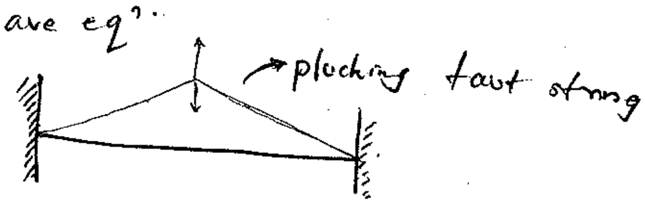


- Space marching problems — disturbance at $x = 0$. whatever happens before $x=0$ is not influenced by the disturbance at $x=0$.

Its possible b/c of high Re. High Re $\Rightarrow$ high inertial forces. Inertial force are predominantly uni-directional compared to viscous forces which spread out in all directions. High Re $\Rightarrow$ disturbances predominantly carried uni-directionally.

(~~that~~ until ~~&~~ unless ~~and~~ B/L separation occurs).

Ex-3: $\frac{\partial^2\phi}{\partial t^2} = c^2 \frac{\partial^2\phi}{\partial x^2}$ wave eq^n.



Use
$x = X$
$t = Y$ } to formulate in the form $A\phi_{XX} + B\phi_{XY} + C\phi_{YY} = -H$.

$A = +c^2$
$B = 0$
$C = -1$

$B^2 - 4AC = 4c^2$

$\frac{dY}{dX} = \frac{\pm 2c}{2c^2} = \pm\frac{1}{c}$

$\Rightarrow$

$\frac{dt}{dx} = \pm\frac{1}{c}$

$\therefore \frac{dx}{dt} = \pm c$

Integrating,

$x = \pm ct + c_1$ → Hyperbolic eq^{ns}.

Forget c_1 (It just shifts the sol^n by a const. amount everywhere).

main characteristics; $x - ct = \xi$

$x + ct = \eta$.

Effect is combined spatio-temporal effect.

It's possible to write the entire eq^n in terms of characteristic variables ξ and η.

$\boxed{\text{Lecture 4: Examples of partial differential equations (contd).}}$

$$\frac{\partial^2\phi}{\partial t^2} = c^2\frac{\partial^2\phi}{\partial x^2} \quad : \left.\begin{array}{l}\xi = x - ct\\ \eta = x + ct\end{array}\right\} \text{two characteristic variables.}$$

$\phi(x,t) \to \phi(\xi,\eta)$. Sol^n can be written in terms of characteristic variables

$$\frac{\partial\phi}{\partial t} = \frac{\partial\phi}{\partial\xi}\overset{-c}{\cancel{\frac{\partial\xi}{\partial t}}} + \frac{\partial\phi}{\partial\eta}\overset{c}{\cancel{\frac{\partial\eta}{\partial t}}}$$

$$\frac{\partial^2\phi}{\partial t^2} = \frac{\partial}{\partial\xi}\left[-c\frac{\partial\phi}{\partial\xi} + c\frac{\partial\phi}{\partial\eta}\right]\overset{-c}{\cancel{\frac{\partial\xi}{\partial t}}} + \frac{\partial}{\partial\eta}\left[-c\frac{\partial\phi}{\partial\xi} + c\frac{\partial\phi}{\partial\eta}\right]\overset{c}{\cancel{\frac{\partial\eta}{\partial t}}}$$

$$= c^2\frac{\partial^2\phi}{\partial\xi^2} + c^2\frac{\partial^2\phi}{\partial\eta^2} - 2c^2\frac{\partial^2\phi}{\partial\xi\partial\eta} \quad \cdots ①$$

$$\frac{\partial\phi}{\partial x} = \frac{\partial\phi}{\partial\xi}\overset{1}{\cancel{\frac{\partial\xi}{\partial x}}} + \frac{\partial\phi}{\partial\eta}\overset{1}{\cancel{\frac{\partial\eta}{\partial x}}}$$

$$\frac{\partial^2\phi}{\partial x^2} = \frac{\partial}{\partial\xi}\left(\frac{\partial\phi}{\partial\xi} + \frac{\partial\phi}{\partial\eta}\right)\frac{\partial\xi}{\partial x} + \frac{\partial}{\partial\eta}\left(\frac{\partial\phi}{\partial\xi} + \frac{\partial\phi}{\partial\eta}\right)\frac{\partial\eta}{\partial x}$$

$$= \frac{\partial^2\phi}{\partial\xi^2} + \frac{\partial^2\phi}{\partial\eta^2} + 2\frac{\partial^2\phi}{\partial\xi\partial\eta} \quad \cdots ②$$

Apply ① and ② in original wave eq^n.

gives, $\quad 4\dfrac{\partial^2\phi}{\partial\xi\partial\eta} = 0 \quad$ or $\quad \dfrac{\partial^2\phi}{\partial\xi\partial\eta} = 0 \to \dfrac{\partial}{\partial\xi}\left(\dfrac{\partial\phi}{\partial\eta}\right) = 0.$

$\Rightarrow \frac{\partial \phi}{\partial \eta} = f_1(\eta)$.

Integrating. $\phi = F(\eta) + G(\xi)$.

Conclusion: general sol^n can be written in terms of characteristic variables.

IC $\rightarrow$ At $t=0$, $\phi = f(x)$.

$t=0$, $\frac{\partial \phi}{\partial t} = g(x)$.

$F(x) + G(x) = f(x)$

$cF'(x) - cG'(x) = g(x)$

$\Rightarrow F(x) - G(x) = \frac{1}{c}\int_0^x g(\tau)\, d\tau$.

$F(x) = \frac{1}{2}f(x) + \frac{1}{2c}\int_0^x g(\xi)\, d\xi + C_{11}$

$G(x) = \frac{1}{2}f(x) - \frac{1}{2c}\int_0^x g(\tau)\, d\tau$.

$\phi = F(\eta) + G(\xi)$.

$= F(x+ct) + G(x-ct)$

$= \frac{1}{2}\left[f(x+ct) + f(x-ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\tau)\, d\tau$

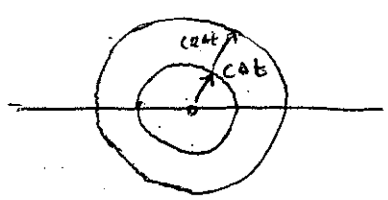
- Say a disturbance propagates in a fluid medium.

disturbance speed = Sonic speed.

c = sonic speed.

v = speed of the source of disturbance.

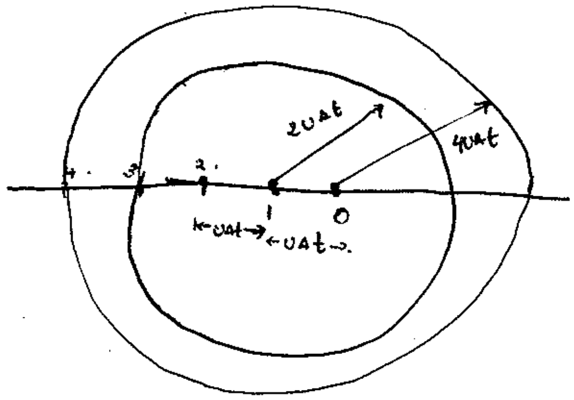
Ex-1 $M_a = 0$ (Source of disturbance doesn't move). $\{M_a = v/c.\}$



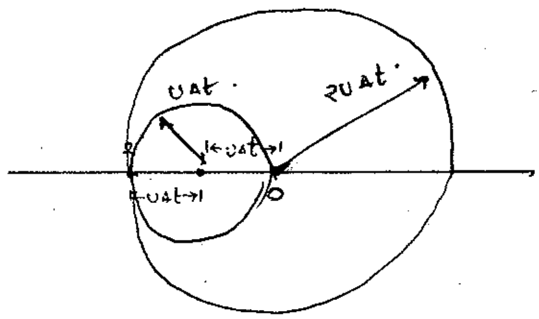
$\leftarrow$ Take Δt, $2\Delta t$, $3\Delta t$ etc....

Ex-2

$Ma = \frac{1}{2} \rightarrow \frac{U}{c} = \frac{1}{2}$

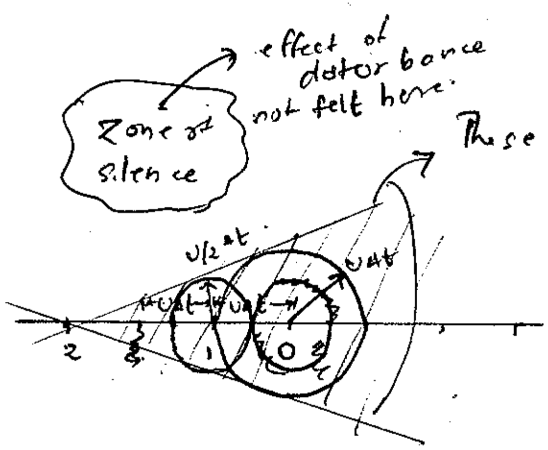


Ex-3 $Ma = 1 \rightarrow U = c$



Disturbance wave doesn't propagate more than where the source is located.

Ex-4 $Ma > 1$ $Ma = 2 \rightarrow U = 2c$

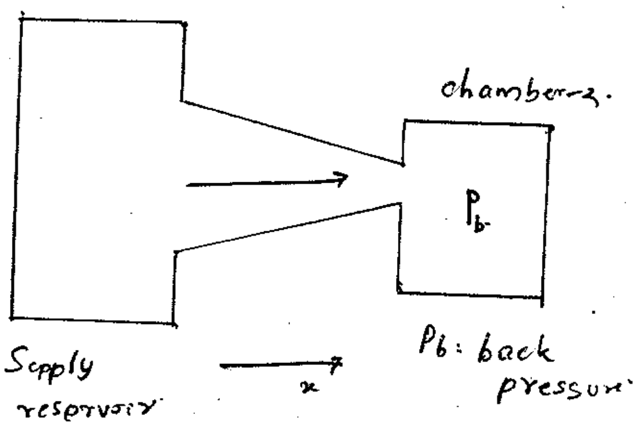


These two straight lines are characteristic line (weak discontinuity)

Mach cone - imaginary cone within which the effects of disturbance is felt.

Highly ~~incompressible~~ flows — disturbance can propagate only with finite speed (unlike in incompressible case where speed is infinite).
As such the effect of disturbance gets accumulated

Say we have a converging nozzle.



To increase flow rate, $P_b \downarrow$.
$\Rightarrow \dot{m} \uparrow$

P_b regulation ~~is~~ essentially creates ~~involves creating~~ ~~Creating~~ a disturbance in chamber-2 which is propagated upstream & gives message to supply reservoir to respond to that & send more mass flow.

But what actually happens is That the mass flow rate can be increased upto $Ma = 1$ by decreasing P_b. Beyond that limit it cannot be increased.

Explanation:

$\vec{V}_{DF}$ = velocity of disturbance relative to flow

$= \vec{V}_D - \vec{V}_F$

$\vec{V}_F = +c$

$\vec{V}_{DF}' = -c$

$\Rightarrow \vec{V}_D = 0$

In hyperbolic cases, when source of disturbance moves faster than the disturbance itself, it leads to discontinuities in the flow medium which has to be taken into account while designing the numerical simulation.

Lecture 5: Nature of the Characteristics of partial differential equation

General 2nd order pde form:

$$\sum_i \sum_j A_{ij} \frac{\partial^2 \phi}{\partial x_i \partial x_j} + \underbrace{B}_{\text{lower order terms}} = 0.$$

Nature of Characteristics depends on eigen value of A.
(coeff matrix)

To get eigs, use $\det |A - \lambda I| = 0 \longrightarrow \lambda_i$'s.

- If any λ is zero $\longrightarrow$ parabolic
- If none is zero and all λs are of the same sign $\longrightarrow$ elliptic
- If none is zero and al but one λ is opposite sign $\longrightarrow$ hyperbolic.

Ex $(1 - M_\infty^2) \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \longrightarrow$ eq^n relating velocity potential for isentropic inviscid compressible flow over ~~sle~~ bodies with slender shapes.

$$A_{11} \frac{\partial^2 \phi}{\partial x_1^2} + \cancel{A_{22} \frac{\partial^2 \phi}{\partial x_2^2}}$$

$$A_{12} \frac{\partial^2 \phi}{\partial x_1 \partial x_2} + A_{21} \frac{\partial^2 \phi}{\partial x_2 \partial x_1} + A_{22} \frac{\partial^2 \phi}{\partial x_2^2} + B = 0$$

$x_1 = x$ $\quad A_{11} = 1 - M_\infty^2$

$x_2 = y$ $\quad A_{12} = A_{21} = B = 0.$

$A_{22} = 1.$

$$\begin{vmatrix} (1 - M_\infty^2) - \lambda & 0 \\ 0 & 1 - \lambda \end{vmatrix} = 0$$

$$(1 - \lambda - M_\infty^2)(1 - \lambda) = 0.$$

$\lambda = 1,$

$\lambda = 1 - M_\infty^2$

If $M_\infty = 1 \longrightarrow$ parabolic

$M_\infty < 1 \longrightarrow$ elliptic

$M_\infty > 1 \longrightarrow$ hyperbolic

So the same stream eq^n depending on M_∞ can ~~be~~ be parabolic, hyperbolic or elliptic.

Ex. • Heat transfer
• Unsteady
• 1D
• low $\frac{k}{\rho C_p} \to 0$
• $U = U_\infty$

$$\frac{\partial}{\partial t}(\rho T) + \nabla(\rho \vec{V} T) = \nabla\left(\frac{k}{C\rho}\nabla T\right) + S$$

Using the conditions given,

$$\nabla(\rho \vec{V} T) = \frac{d}{dx}(\cdots)$$

$$\nabla\left(\frac{k}{C\rho}\nabla T\right) \to 0$$

$$\Rightarrow \frac{\partial T}{\partial t} + U_\infty \frac{\partial T}{\partial x} = S$$

$$T = T(x,t)$$

$$dT = \frac{\partial T}{\partial x} dx + \frac{\partial T}{\partial t} dt$$

$$\begin{bmatrix} 1 & U_\infty \\ dt & dx \end{bmatrix} \begin{bmatrix} \frac{\partial T}{\partial t} \\ \frac{\partial T}{\partial x} \end{bmatrix} = \begin{bmatrix} S \\ dT \end{bmatrix}$$

$$|\ | = 0 \longrightarrow dx - U_\infty dt = 0 \Rightarrow \frac{dx}{dt} = U_\infty \leftarrow \text{characteristics}$$

What is the nature of characteristics here? Not parabolic!

Converting to 2nd order form

$$\frac{\partial T}{\partial t} + U_\infty \frac{\partial T}{\partial x} = 0$$

$$\frac{\partial^2 T}{\partial t \partial x} + U_\infty \frac{\partial^2 T}{\partial x^2} = 0 \cdots ①$$

$$\frac{\partial T}{\partial t} + U_\infty \frac{\partial T}{\partial x} = 0$$

$$\frac{\partial^2 T}{\partial t^2} + U_\infty \frac{\partial^2 T}{\partial x \partial t} = 0 \cdots ②$$

$$① \times U_\infty - ② \rightarrow$$

$$U_\infty^2 \frac{\partial^2 T}{\partial x^2} - \frac{\partial^2 T}{\partial t^2} = 0$$

$$\frac{\partial^2 T}{\partial t^2} = U_\infty^2 \frac{\partial^2 T}{\partial x^2} \longrightarrow \text{hyperbolic!}$$

HW: $$\left.\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= 4v \end{aligned}\right\}$$ Find the nature of pde.

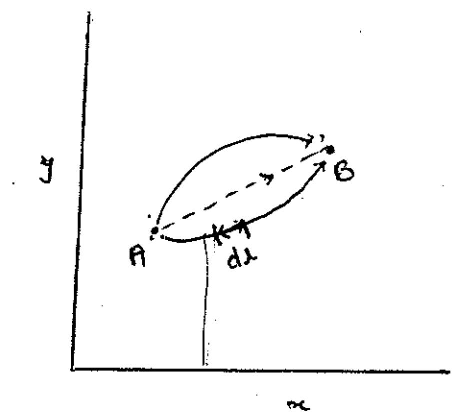
HW: Make a chart for the features of parabolic, hyperbolic, elliptic eq^n
Nature of the characteristics, how many characteristics, zone of influence, zone of disturbance, speed of propagation of disturbance.

Lecture 6: Euler-Lagrangian Equation

- Error minimization – key principle with which many numerical methods are founded.
  - B/c a good approximate sol^n is the one which incurs least error.
  - includes many considerations. one such is variations.

- Calculus of variations in brief:

  Say we have two pts. Objective: find the path with least distance b/w them.



$$dl = \sqrt{dx^2 + dy^2}$$

$$= \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \cdot dx$$

$$y' \equiv \frac{dy}{dx}, \quad y'' \equiv \frac{d^2y}{dx^2}$$

$$\therefore \; l = \int dl = \int \sqrt{1 + y'^2}\, dx$$

Problem statement becomes:

Find the path AB that minimizes l

$\Rightarrow$ minimizes $I = \int \sqrt{1+y'^2}\, dx$

$\sqrt{1+y'^2} \rightarrow F(x, y, y')$

x: independent variable; y, y': themselves fns of x

So F is a fn of functions

$\therefore$ F is a functional

$$I = \int F(x, y, y')\, dx$$

To minimize I, take δI.

δI – arbitrarily small virtual change in I.

δI will only involve changes in dependent variables, not on independent variable (x here).

Why? We are trying to find what the δy, $\delta y'$ etc. should be to minimize I, so that ~~our~~ our answer fall on the desired path i.e. the straight line.

$$\delta I = \int \left[\frac{\partial F}{\partial y} \delta y + \frac{\partial F}{\partial y'} \delta y' \right] dx$$

y is fixed at A and B. $\therefore$ $\delta y = 0$ at A, B.

Simplify 2nd term ~~by~~ using integration by parts:

$$\delta I = \int \left[\frac{\partial F}{\partial y}\right] \delta y + \left[\frac{\partial F}{\partial y'} \delta y\right]_A^B - \int \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \delta y \, dx$$

Minimize $I = \int F(x, y, y')\, dx$ subject to constraint $y_{@A}$, $y_{@B}$ specified.

$\therefore$ 2nd term $\left[\frac{\partial F}{\partial y'} \delta y\right]_A^B = 0$ $\quad (\delta y_{@A} = \delta y_{@B} = 0)$.

$$\therefore \delta I = \int \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] \delta y \, dx$$

For min I, $\delta I = 0$ for any arbitrary δy.

This is possible when integrand $= 0$.

$$\Rightarrow \frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 \leftarrow \text{Euler-Lagrange equation.}$$

for minimization of dist b/w two pts in this example.

$$I = \int \sqrt{1+y'^2}\, dx \Rightarrow F(x, y, y') = \sqrt{1+y'^2}$$

$$\frac{\partial F}{\partial y} = 0, \quad \frac{\partial F}{\partial y'} = \frac{2y'}{2\sqrt{1+y'^2}} = \frac{y'}{\sqrt{1+y'^2}}$$

$$\therefore \text{E-L eq}^n: \quad 0 - \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0$$

$$\text{or } \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 0. \Rightarrow \frac{y'}{\sqrt{1+y'^2}} = \text{Const}$$

$\Rightarrow y' = \text{Const} = C$
or $\frac{dy}{dx} = C$ $\rightarrow$ path is a straight line.

Lecture 7: Approximate Solutions of Differential Equations

(Prob) Show that an alternative form of the Euler-Lagrange eq^n is given by $\frac{\partial F}{\partial x} - \frac{d}{dx}\left(y' \frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial x} = 0$, where $F(x, y, y')$. ...(1)

(Ans) $dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{\partial F}{\partial y'} dy'$

$$\Rightarrow \frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} + \frac{\partial F}{\partial y'}\frac{dy'}{dx}$$

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{dy'}{dx}\left(\frac{\partial F}{\partial y'}\right)$$

$$\therefore \frac{dy'}{dx}\left(\frac{\partial F}{\partial y'}\right) = \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) - y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right).$$

Sub into main expression $\frac{dF}{dx}$

$$\Rightarrow \frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\cancel{\frac{dy}{dx}}^{\,y'} + \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) - y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \quad \ldots (2)$$

$$= \frac{\partial F}{\partial x} + \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) + y'\left[\cancel{\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)}\right]$$

0 (since this is same as original E-L eq^n).

$$\Rightarrow \frac{dF}{dx} - \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial x} = 0$$

(Prob-2) Hence show that the closed curve that minimizes the perimeter for a given area is a circle.

$$P = \int dl = \int\sqrt{dx^2 + dy^2} = \int\sqrt{1+y'^2}\,dx, \quad y' = \frac{dy}{dx}.$$

$$A = A^* = \int y\,dx.$$

Obj: minimize P s.t. A is a const.

Do this via Lagrange multiplier.

Introduce $F = P + \lambda A$, where λ is Lagrange multiplier.

$$I = \int (P + \lambda A) = \int \Big[\underbrace{\sqrt{1+y'^2} + \lambda y}_{F(x,y,y')}\Big] dx$$

Use alternate form of E-L eq^n.

$$\frac{d}{dx}\left[F - y'\frac{\partial F}{\partial y'}\right] - \cancel{\frac{\partial F}{\partial x}}^{0} = 0$$

$$\Rightarrow F - y'\frac{\partial F}{\partial y'} = \text{const } c$$

$$\sqrt{1+y'^2} + \lambda y - \frac{(y')^2}{\sqrt{1+(y')^2}} = c$$

$$\Rightarrow 1 + \cancel{y'^2} + \lambda y \sqrt{1+y'^2} - \cancel{(y')^2} = c\sqrt{1+(y')^2}$$

$$1 = (c - \lambda y)\sqrt{1+y'^2}$$

Use $y' = \tan\theta = \frac{dy}{dx}$.

$$\cos\theta = c - \lambda y.$$

$$\frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta}$$

$$\frac{d}{d\theta}(\cos\theta) = \frac{d}{d\theta}(c - \lambda y)$$

$$-\sin\theta = -\lambda\frac{dy}{d\theta}.$$

So, $\frac{dy}{d\theta} = \frac{1}{\lambda}\sin\theta$

So $$\frac{dy}{dx} = \frac{\frac{1}{\lambda}\sin\theta}{\left(\frac{dx}{d\theta}\right)}.$$

$$\tan\theta = \frac{\frac{1}{\lambda}\sin\theta}{\left(\frac{dx}{d\theta}\right)}$$

$$\therefore \frac{dx}{d\theta} = \frac{1}{\lambda}\cos\theta.$$

$$\therefore x = \frac{\sin\theta}{\lambda} + c_1$$

$$\sin\theta = \lambda(x - c_1)$$

$$\cos\theta = c - \lambda y$$

$$\sin^2\theta + \cos^2\theta = 1$$

$$\Rightarrow [\lambda(x - c_1)]^2 + [c - \lambda y]^2 = 1.$$

$$\lambda^2(x-c_1)^2 + \lambda^2(y-c_2)^2 = 1.$$

This is of the form

$$(x-a)^2 + (y-b)^2 = r^2$$ $\hookrightarrow$ eq^n of circle

Thus circle minimizes perimeter for const area shapes!

Ex

Functionals involving higher order derivatives:

Say, $I = \int F(x, y, y', y'')\, dx$

$$\delta I = \int\left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' + \frac{\partial F}{\partial y''}\delta y''\right] dx$$

{to denote first order variation (usually omitted)}

$$\frac{\partial F}{\partial y'}\delta y' \to \left[\frac{\partial F}{\partial y'}\delta y\right]_A^B - \int \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\delta y\, dx$$

$$\frac{\partial F}{\partial y''}\delta y'' \to \left[\left[\frac{\partial F}{\partial y''}\delta y'\right]_A^B - \int \frac{d}{dx}\left(\frac{\partial F}{\partial y''}\right)\delta y'\, dx\right]$$

$$\to -\left[\frac{d}{dx}\left(\frac{\partial F}{\partial y''}\right)\delta y\right]_A^B + \int \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right)\delta y\, dx$$

$$\cancel{\left[\frac{\partial F}{\partial y''}\delta y'\right]_A^B - \int}$$

Boundary terms.

$$\left[\frac{\partial F}{\partial y'}\delta y\right]_A^B, \quad \left[\frac{\partial F}{\partial y''}\delta y'\right]_A^B, \quad -\left[\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\delta y\right]_A^B$$

Boundary terms $= 0$ if $\delta y, \delta y' = 0$.

So assume y, y' are specified at A & B :

$\delta y, \delta y' = 0$.

$$\Rightarrow \delta I = \int\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right)\right]\delta y\, dx = 0$$

$$\cancel{\delta I = \int\left[\frac{\partial F}{\partial y} + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right)\right]\delta y\, dx = 0}$$

$$\Rightarrow \frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right) = 0$$ to minimize I.

Fore more higher order eq^n, we have the form:

$$\left\{\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right) - \frac{d^3}{dx^3}\left(\frac{\partial F}{\partial y'''}\right) + \frac{d^4}{dx^4}\left(\frac{\partial F}{\partial y''''}\right) - \cdots = 0\right\}$$

↳ Euler – Poisson form.

Approximate solutions of differential equations through variational formulation:

Eg: 1-D steady state heat conduction with constant heat source:

$$\frac{\partial}{\partial t}(\rho T) + \nabla(\rho \vec{V} T) = \nabla\left(\frac{k}{C_p}\nabla T\right) + \frac{S}{C_p}$$

Also assume: const. thermal properties k, C_p etc const.

Steady state $\Rightarrow$ $\frac{\partial}{\partial t}(\cdot\cdot) = 0$

Conduction problem $\Rightarrow$ no flow velocity involved.

$$\therefore \nabla(\rho \vec{V} T) = 0$$

1D Problem $\Rightarrow$ $\nabla(\,) = \frac{d}{dx}(\,)$

Final form: $\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S = 0.$ $\longrightarrow$ "D" form (Differential form) $\hookrightarrow$ Strong form.

To make it a variational form, multiply with a variational parameter & integrate over the domain.

$$\int\left[\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S\right] v\, dx = 0$$

v $\hookrightarrow$ carries the meaning of δT.

Use integration by parts:

$$\left[v k \frac{dT}{dx}\right]_1^2 - \int k \frac{dv}{dx}\frac{dT}{dx}\, dx + \int S v\, dx = 0$$? doubt with the expression!

Boundary conditions possible:

$\rightarrow$ T specified.

$\Rightarrow$ $\delta T = 0$ or $v = 0$.

$\rightarrow$ $\frac{dT}{dx}$ specified

} either of the two.

Ex: T specified at both boundaries.

$$\Rightarrow \boxed{\int k \frac{dT}{dx}\frac{dv}{dx}\, dx = \int S v\, dx} \longrightarrow \text{Weak form.}$$

"V" form.

Lecture 8: Variational formulation

In the weak form, it requires continuity only upto the first order derivative, while in the strong form, continuity upto the second order derivative is required. Hence why the name - weak form

Generic form of "V" form:

$$\underbrace{a(T, v)}_{\text{bilinear operator}} = \underbrace{l(v)}_{\text{linear operator}}$$

why this form?

Make some observations:

obs-1

$$a(\alpha_1 T + \alpha_2 v, \beta_1 T + \beta_2 v) = \int k \frac{d}{dx}(\alpha_1 T + \alpha_2 v) \frac{d}{dx}(\beta_1 T + \beta_2 v)\, dx$$

$$= \alpha_1 \beta_1 \int k \frac{dT}{dx} \frac{dT}{dx}\, dx + \alpha_1 \beta_2 \int k \frac{dT}{dx}\frac{dv}{dx}\, dx + \alpha_2 \beta_1 \int k \frac{dv}{dx}\frac{dT}{dx}\, dx + \alpha_2 \beta_2 \int k \frac{dv}{dx}\frac{dv}{dx}\, dx$$

$$= \alpha_1 \beta_1\, a(T,T) + \alpha_1 \beta_2\, a(T,v) + \alpha_2 \beta_1\, a(v,T) + \alpha_2 \beta_2\, a(v,v).$$

$\Rightarrow$ $a(\,,\,)$ is bilinear (i.e. linear in each slot)

obs-2

$$l(\alpha T + \beta v) = \cancel{\int s(\alpha T + \beta v)\, dx}$$

$= \alpha l(T) + \beta l(v)$, if such a property is satisfied, l is a linear operator

obs-3. $a(T, v) = a(v, T)$ $\Rightarrow$ a is symmetric (self-adjoint)

obs-4 $a(v,v) = \int k\left(\frac{dv}{dx}\right)^2 dx$ is a +ve definite operator

integral over the domain is +ve >0.

a is a scalar product on V.

Say we have a function $g(\varepsilon) = \frac{1}{2} a(T+\varepsilon v, T+\varepsilon v) - \ell(T+\varepsilon v)$

M problem : Minimize g at $\varepsilon = 0 \longrightarrow \frac{1}{2} a(T,T) - \ell(T)$

Assumptions:

$\longrightarrow$ a is bilinear & ℓ is linear

$$g(\varepsilon) = \frac{1}{2} a(T,T) + \frac{\varepsilon}{2} a(T,v) + \frac{\varepsilon}{2} a(v,T) + \frac{\varepsilon^2}{2} a(v,v) - \ell(T) - \varepsilon \ell(v)$$

To minimize 'g', $g'(\varepsilon)\Big|_{\varepsilon=0} = 0$

$$g'(\varepsilon)\Big|_{\varepsilon=0} = \frac{1}{2} a(T,v) + \frac{1}{2} a(v,T) + \cancel{\varepsilon a(v,v)}^{\,0} - \ell(v) = 0.$$

$$\frac{1}{2}\left(a(T,v) + a(v,T)\right) = \ell(v)$$

From obs-8, $\Rightarrow$ $a(T,v) = \ell(v)$ which is same as 'V' form.

i.e. a is symmetric $\Rightarrow a(T,v) = a(v,T)$

$$\Rightarrow a(T,v) = \ell(v) \rightarrow \text{"V" form.}$$

Conclusion: 'M' form $\Rightarrow$ 'V' form, provided a is bilinear, ℓ is linear & a is symmetric.

Question is: starting from 'V' form, is it possible to reach 'M' form?

Consider a as symmetric.

$$g(\varepsilon) = \frac{1}{2} a(T,T) - \ell(T) + \varepsilon\left[a(T,v) - \ell(v)\right] + \frac{\varepsilon^2}{2} a(v,v)$$

If 'V' form is true, $\varepsilon\left[a(T,v) - \ell(v)\right] = 0$

Then $g(\varepsilon) \geq \frac{1}{2} a(T,T) - \ell(T)$; provided $a(v,v)$ is +ve.

It is true when a is +ve definite ~~pe~~

$$\therefore\ g(\varepsilon) = \frac{1}{2} a(T,T) - \ell(T) \text{ is the minimum of } g(\varepsilon) \text{ at } \varepsilon = 0.$$

$\hookrightarrow$ 'M' form

for V to M, we require additional constraint i.e. a is +ve definite.

$$\int k \frac{dT}{dx}\frac{dv}{dx}\,dx = \int S v\,dx$$

sub δT in place of v,

$$\delta \underbrace{\int \left[\frac{1}{2} k \left(\frac{dT}{dx}\right)^2 dx - \int ST\right] dx}_{\pi} = 0.$$

Essentially, we are minimizing Π.

Q) How to get 'D'-form from 'V' form?

'V' form: $\displaystyle\int k \underset{1}{\underline{\underline{\frac{dT}{dx}}}}\ \underset{2}{\underline{\underline{\frac{d(\delta T)}{dx}}}}\,dx = \int S\,\delta T\,dx$

Integrate by parts:

$$k \frac{dT}{dx} v \Big]_1^2 - \int \frac{d}{dx}\left(k\frac{dT}{dx}\right) v\,dx = \int_1^2 S v\,dx$$

$$\int \left[\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S\right] v\,dx = 0$$

$$\Rightarrow \frac{d}{dx}\left(k\frac{dT}{dx}\right) + S = 0 \longrightarrow \text{'D' form.}$$

So far we took BC as T is specified. We can do the same with $\frac{dT}{dx}$ specified ~~as~~ a instead.

- Boundary conditions in the variational formulation:

T specified: variable for which variation appears in the boundary terms.
B/c v is variation of T. (primary variable)

$k\frac{dT}{dx}$ specified $\longrightarrow$ Coeff. of variation in the boundary term (secondary variable)

Specifying the primary variable at boundary is called as essential BC.

Specifying the secondary variable at boundary is called as natural BC. Natural BC that forms automatically appears in the eq^n (-ve of heat flux).

Lecture-9: Example of Variational formulation and introduction to Weighted Residual Method.

Eg: $\frac{d^2}{dx^2}\left[a(x)\frac{d^2y}{dx^2}\right] + b(x) = 0.$

Obj: cast in variational formulation.

$$\int \left[\frac{d^2}{dx^2}\left[a(x)\frac{d^2y}{dx^2}\right] + b(x)\right] v\,dx = 0.$$

Integrate by parts:

$$\left. v\frac{d}{dx}\left[a(x)\frac{d^2y}{dx^2}\right]\right|_{x=0}^{x=L} - \int \underset{1}{\underline{\frac{dv}{dx}}}\,\underset{2}{\underline{\frac{d}{dx}\left[a(x)\frac{d^2y}{dx^2}\right]}}\,dx + \int b(x)\,v\,dx = 0.$$

$$\underline{\left. v\frac{d}{dx}\left[a(x)\frac{d^2y}{dx^2}\right]\right|_{x=0}^{x=L}} - \underline{\left.\frac{dv}{dx}\left(a(x)\frac{d^2y}{dx^2}\right)\right|_{x=0}^{x=L}} + \underbrace{\int_{x=0}^{x=L}\frac{d^2v}{dx^2}\,a(x)\frac{d^2y}{dx^2}\,dx}_{a(y,v)} + \underbrace{\int_{x=0}^{x=L} b(x)\,v\,dx}_{l(v)} = 0.$$

Primary variable $\to y$. (E.B.C)

Secondary variable $\to \frac{d}{dx}\left[a(x)\frac{d^2y}{dx^2}\right]$

(N.B.C)

Primary variable: $\frac{dy}{dx}$ (E.B.C)

Secondary variable: $a(x)\frac{d^2y}{dx^2}$. (N.B.C).

Boundary terms need not be zero after applying B/Cs.

The terms that we get after BC application can be dubbed with $l(v)$ term.

Specifying B/Cs.

Let $y=0$ at $x=0$.

$\frac{d}{dx}\left[a(x)\frac{d^2y}{dx^2}\right] = c_1$ at $x=L$.

$$\frac{dy}{dx} = 0 \quad \text{at } x=0$$

$$a(x)\frac{d^2y}{dx^2} = c_2 \text{ at } x=L.$$

$$\Rightarrow \quad V_L c_1 - \frac{dv}{dx}\Big|_L c_2 + \int_{x=0}^{x=L} a_{(x)} \frac{d^2v}{dx^2}\frac{d^2y}{dx^2}\,dx + \int_{x=0}^{x=L} b(x)\, v\, dx = 0.$$

Write this in the form $A(y,v) = L(v)$

$$A(y,v) = \int_{x=0}^{x=L} a(x) \frac{d^2y}{dx^2}\frac{d^2v}{dx^2}\,dx$$

$$L(v) = -\int_{x=0}^{x=L} b(x)\, v\, dx - V_L c_1 + \frac{dv}{dx}\Big|_L c_2$$

$A(y,v) = L(v)$ is the required variational formulation.

Approximate solutions of differential equations:

Weighted residual approach:

: gets a clue from the variational form.

$\int (\ ) v = 0$ form. v is an arbitrarily small variation.

Say we wish to solve $\frac{d^2y}{dx^2} = 0$.

Call linear operator $\mathcal{L}(y) = 0 \longrightarrow \mathcal{L} = \frac{d^2}{dx^2}$.

v till now is an abstract variational parameter. We try to make it non-abstract by looking at the possibilities of functions we can use in place of v.

In the problem $\frac{d^2y}{dx^2} = 0$, to convert it into an algebraic eq^n (b/c algebraic are easier to solve), we replace y with an approximate y_{approx} polynomial & solve for it.

But $\frac{d^2}{dx^2}(y_{approx}) \neq 0$, in general.

Then, $L(y) - L(y_{approx}) = R$ $R \rightarrow$ residual

Our objective is to minimize R in an integral sense over the domain.

$$\int_\Omega R w \, d\Omega = 0$$

Try to minimize the error or the residual in a weighted integral sense.

→ y^* (y_{approx})
[trial function]

→ w
[weighting function]

Note: these functions need not be as rigorous as the variation formulations. Restrictions like $a(y,v)$ symmetry & positive definiteness not needed.

Lecture 10: Weighted Residual Method.

Say, governing differential eq^n $L(y) = 0$.

Substituting y_{approx}, $L(y_{approx}) \neq 0$.

$$L(y_{approx}) = R.$$

$$\int_\Omega R w \, d\Omega = 0.$$

In 1D problem, say $y = f(x)$,

$\int R w \, dx = 0$. try to make sure y_{approx} is appropriate to minimise the R.

Trial function → y_{approx},

- polynomial — most convenient form.
- Should satisfy the essential BC (key requirement)
- should be continuous
- derivatives of trial function must be square integrable.

$\int \left(\frac{dy_{approx}}{dx}\right)^2 dx < \infty$ — shows integral is not unbounded.

+ H^1 fn: 1^{st} derivative is square integrable.

Requirements for weighting function: $\rightarrow w$

- should satisfy homogeneous part of the EBC.

{Say if $y = 5$ is EBC, then $w = 0$ is the homogeneous part}

why? B/c if y is specified, then variation is $y = 0$. w has similar meaning as that of variation in y. $\therefore$ $w = 0$ at Boundary

- Should be continuous.

Some specific examples:

Prob: 1D, steady state heat transfer with uniform Thermal conductivity k, Source S.

Governing DE: $$k\frac{d^2T}{dx^2} + S = 0.$$

$$\frac{d^2T}{dx^2} + \frac{S}{k} = 0. \quad \frac{S}{k} \rightarrow 100\text{, say}$$

BCs: $x = 0$, $T = 0$
$x = 10$, $T = 0$.

$\longrightarrow$

$\frac{d^2y}{dx^2} + 100 = 0$	At $x = 0$, $y = 0$ $x = 10$, $y = 0$

Obj: find approx. sol^n.

Ex-1 Least squared method

$y \rightarrow y_{approx}$

$$\frac{d^2 y_{approx}}{dx^2} + 100 = R$$

Interested in minimizing R^2:

$\int R^2 dx \longrightarrow$ minimized.

$\equiv$ sum of square of errors.

{$\int R\,dx$ — not correct as some errors can be +ve & -ve & sum to 0, can be misleading

. $\int |R|\,dx$ — not used b/c of tedious algebra.}

$y_{approx} \rightarrow 2^{nd}$ order polynomial with 1 parameter.

General form : $ax(10-x) = y_{approx}$

$\rightarrow$ polynomial, 2^{nd} order

Find out a $\therefore$ $\int R^2 dx$ is minimized.

$y_{approx} = ax(10-x)$

$= 10ax - ax^2$

$\dfrac{dy_{approx}}{dx} = 10a - 2ax$

$\dfrac{d^2 y_{approx}}{dx^2} = \underline{\underline{-2a}}$

$R = y - y_{approx}$

$= -2a + 100.$

$\dfrac{\partial}{\partial a}\left(\int R^2 dx\right) = 0.$

$\Rightarrow \int 2R \dfrac{\partial R}{\partial a}\, dx = 0$

$\Rightarrow \int R \dfrac{\partial R}{\partial a}\, dx = 0.$

So $\dfrac{\partial R}{\partial a}$ is the weighting function.

$\dfrac{\partial R}{\partial a} = \dfrac{\partial}{\partial a}(-2a+100) = \underline{\underline{-2}}$

$\therefore \int_0^{10} (-2a+100)(-2)\, dx = 0$

$\Rightarrow -2a+100 = 0$

or $a = \underline{\underline{50}}$

Lecture 11: Point Collocation method, Galerkin's method & The 'M' form.

Ex-2. Point Collocation method:

$w = \delta(x - x_i)$ δ: Dirac-Delta function.

Idea: you try to satisfy the value of the function at chosen points x_i

$$\int_x R\,w\,dx = 0.$$

Keep trial function same: $ax(10-x)$

$R_{x=x_i} = 0$ Consider only 1 collocation point, say $x=5$.

$R_{x=5} = 0$

$-2a + 100 = 0 \Rightarrow a = 50$

Ex-3 Galerkin's method:

It considers the weighting function as the trial function.

$w = x(10-x)$ Putting a doesn't matter as its just a const & it will go away in the expression $\int Rw\,dx = 0$.

$$\int R w\,dx = 0.$$

$$\int_0^{10} (-2a+100)\, x(10-x)\,dx = 0$$

$\rightarrow -2a+100 = 0$

or $\underline{\underline{a = 50}}$

Going through routes of the 'M' form:

$$\int_0^{100} \left(\frac{d^2y}{dx^2} + 100\right) v\,dx = 0$$

$$\left[v\frac{dy}{dx}\right]_0^{100} - \int_0^{100} \frac{dv}{dx}\frac{dy}{dx}\,dx + \int 100\,v\,dx = 0$$

PV: y $y=0$ at $x=0$ (given)

SV: $\frac{dy}{dx}$

$$\Rightarrow \int_0^{100} \frac{dv}{dx}\frac{dy}{dx}\,dx = \int_0^{100} 100\,v\,dx$$

$a(y,v) = l(v)$

$a(y,y) > 0$. here $\Rightarrow$ 'M' form exists here!

M form:

Minimize $\Pi = \frac{1}{2} a(y, y) - l(y)$.

$$= \frac{1}{2}\int_0^{10} \left(\frac{dy}{dx}\right)^2 dx - \int_0^{10} 100\, y\, dx$$

Use $y_{approx} \rightarrow$ in place of y & minimize Π, & can see have reduced requirement for continuity.

Here we don't require any weighting fn. Only trial function is necessary. This convenience comes at the cost of additional requirement of positive definitness of $a(y, y)$.

$$\frac{\partial \Pi}{\partial a} = 0 \qquad \text{Use trial fn } y_{approx} = ax(10-x).$$

— Rayleigh – Ritz method.

$$\frac{dy_{approx}}{dx} = 10a - 2ax$$

$$\Pi = \frac{1}{2}\int_0^{10} \left(\frac{dy_{approx}}{dx}\right)^2 dx - \int_0^{10} 100\, y_{approx}\, dx$$

$$= \frac{1}{2}\int_0^{10} (10-2ax)^2 dx - \int_0^{10} 100\,(10a - 2ax)\, dx$$

$$= \cancel{\frac{\partial}{\partial a}\left(500 + \frac{2a^2}{3}\times 1000 - 100 x\right.} - \cancel{10000a} + \cancel{\left.4000a\right)} = 0$$

$$= \cancel{\frac{4a}{3}\times 1000}\ \cancel{\times 100}$$

$$\frac{\partial \Pi}{\partial a} = 0 \Rightarrow a = 100 \int_0$$

$$\frac{dy}{dx} = -100x + C_1$$

$$y = -\frac{100}{2}x^2 + C_1 x + C_2.$$

$$C_2 = 0 \longrightarrow 0 = -100 \times \frac{1000}{2} + 10C_1 \longrightarrow C_1 = 500.$$

$$y = 50(10x - x^2)$$

$$= 50x(10-x)$$ exact!

How to reduce calculations associated with higher order polynomials?
Divide the domain into smaller & smaller subdomain & use lower order polynomials in those subdomains. B/c even though across the entire domain our fn may be quite complex, in smaller subdomains, such functions may be simpler. Now the solution will be fitted b/w discrete points rather than over the entire domain → basic idea behind discretization.

here we lose the continuous nature of the domain.

<u>Discretization:</u>

- Divide the domain into a number of discrete subdomains.
  (element, control volume, ...)
- Each subdomain is represented by a discrete set of points.
  (grid points, nodes, ...)
- Objective is to convert the governing DE into a system of algebraic equations valid at each of these discrete points.

Lecture 12: Finite Element Method (FEM) of discretization

Discretization principles:

→ Divide the domain into a number of discrete subdomains; each subdomain being ~~characterized~~ represented by a number of discrete points.

→ Derive algebraic equations from the governing diff. eq^{ns}; valid at these discrete points.

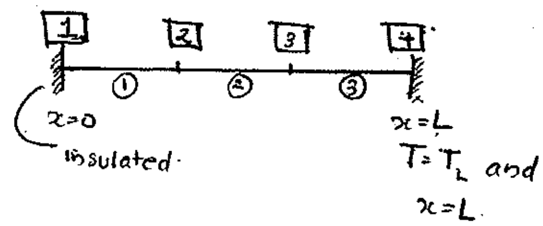
→ Solve the system of algebraic equations to obtain values of the dependent variables at the discrete points.

→ Broad steps in overall analysis:

- → Pre-processing: set-up geometry, discretized eq^{ns}, input data (property data), initial cond, BCs.
- → Solution: algebraic eq^{ns}
- → Post-processing: Graphical representation of the obtained results.

Finite Element Method (FEM)

Ex: $\frac{d}{dx}\left(k\frac{dT}{dS}\right) + S = 0$



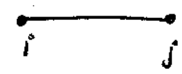
□1, □2, □3, □4 → nodes

①, ②, ③ → elements.

Prepare node-element connectivity chart:

Element	node i	node j
1	1 (,)	2 (,)
2	2 (,)	3. (,)
3	3 (,)	4 (,)

Consider any isolated element



Writing an algebraic eq^n corresponding to the governing diff eq^n.

$$\int_{x=0}^{x=L}\left[\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S\right]w\,dx = 0$$

Integrate by parts:

$$w\,k\frac{dT}{dx}\Big]_0^L - \int_0^L \frac{dw}{dx}k\frac{dT}{dx}\,dx + \int_0^L Sw\,dx = 0.$$

For ~~sa~~ the isolated element, This would become.

$$wk\frac{dT}{dx}\Big]_{x_i}^{x_j} - \int_{x_i}^{x_j} k\frac{dw}{dx}\frac{dT}{dx}\,dx + \int_{x_i}^{x_j} Sw\,dx = 0.$$

If it ~~are~~ not a two-noded element, we'd require a higher order polynomial to approximate ~~T~~ T.

i ——— j $\quad T = a_0 + a_1 x$ (trial function)

We choose trial function for each element, not for the whole domain.

At $x = x_i$, $T = T_i$

$x = x_j$: $T = T_j$

$$\left.\begin{aligned} T_i &= a_0 + a_1 x_i \\ T_j &= a_0 + a_1 x_j \end{aligned}\right] \rightarrow \quad a_1 = \frac{T_j - T_i}{x_j - x_i}$$

$$a_0 = T_i - a_1 x_i = T_i - \left(\frac{T_j - T_i}{x_j - x_i}\right)x_i$$

We finally get piecewise continuous function for the whole domain

17

$$a_0 = T_i - \left(\frac{T_j - T_i}{x_j - x_i}\right) x_i = \frac{T_i x_j - T_j x_i}{x_j - x_i} \qquad \Bigg| \qquad a_1 = \frac{T_j - T_i}{x_j - x_i}$$

$$T = \left(\frac{T_i x_j - T_j x_i}{x_j - x_i}\right) + \left(\frac{T_j - T_i}{x_j - x_i}\right) x$$

$$= \underbrace{\frac{(x_j - x)}{(x_j - x_i)}}_{N_i} T_i + \underbrace{\left(\frac{x - x_i}{x_j - x_i}\right)}_{N_j} T_j$$

$$T = N_i T_i + N_j T_j$$

Interpolation functions / shape functions.

Property of shape functions:

$N_i = 1$ at node i, $= 0$ at node j
$N_j = 0$ at node j, $= 1$ at node i.

Writing in matrix form

$$T = \underbrace{[N_i \quad N_j]}_{[N]} \underbrace{\begin{bmatrix} T_i \\ T_j \end{bmatrix}}_{[T]}$$

$$w = [N]\hat{[w]} \text{ in Galerkin form.} \qquad \hat{[w]} = \begin{bmatrix} w_i \\ w_j \end{bmatrix}$$

$$\Rightarrow w^T = [w]^T [N]^T$$

$$\therefore \; wk\frac{dT}{dx}\Big]_{x_i}^{x_j} - \int_{x_i}^{x_j} k \frac{dw}{dx}\frac{dT}{dx}\,dx + \int_{x_i}^{x_j} Sw\,dx = 0.$$

$$\rightarrow \; [w]^T [N]^T q''\Big]_i^j - [w]^T \int_{x_i}^{x_j} \left[\frac{dN^T}{dx}\right] k \left[\frac{dN}{dx}\right] [T]\,dx \; \cancel{+ \int_{x_i}^{x_j} S [N]^T [T]} + [w]^T \int S [N]^T dx = 0.$$

Lecture 13 : Finite Element Method of Discretization (contd)

Since $[w]^T$ is arbitrary,

$$\underbrace{\left[-[N]^T q''\right]_i^j}_{\text{Term-1}} - \underbrace{\left[\int_i^j \left[\frac{dN}{dx}\right]^T k \left[\frac{dN}{dx}\right]\right]}_{\text{Term-2}} [T] + \int_i^j S\,[N]^T dx = 0$$

$$\text{Term-1} = \left[-\begin{bmatrix} N_i \\ N_j \end{bmatrix} q''\right]_i^j$$

$$= -\begin{bmatrix} 0 \\ q''_j \end{bmatrix} + \begin{bmatrix} q''_i \\ 0 \end{bmatrix} = \begin{bmatrix} +q''_i \\ -q''_j \end{bmatrix}$$

$$\text{Term-2} = \int_i^j \begin{bmatrix} \frac{dN_i}{dx} \\ \frac{dN_j}{dx} \end{bmatrix}^T k \begin{bmatrix} \frac{dN_i}{dx} & \frac{dN_j}{dx} \end{bmatrix} dx$$

$$\frac{dN_i}{dx} = \frac{-1}{\underbrace{x_j - x_i}_{le}} = \frac{-1}{le} \qquad \frac{dN_j}{dx} = \frac{1}{\underbrace{x_j - x_i}_{le}} = \frac{+1}{le}$$

$$\Rightarrow \text{Term-2} = \int_i^j \frac{k}{le^2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} dx$$

$$= \int_i^j \frac{k}{le^2} \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix} dx$$

$$= \frac{k}{le^2} (x_j - x_i)$$

$$= \underline{\underline{\frac{k}{le} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}}}$$

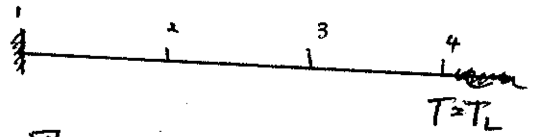
$$\text{Term-3} : \int_i^j S[N]^T dx = S\int_i^j \begin{bmatrix} \frac{x_j - x}{le} \\ \frac{x - x_i}{le} \end{bmatrix} dx = \underline{\begin{bmatrix} \frac{S\,le}{2} \\ \frac{S\,le}{2} \end{bmatrix}}$$

→ entire effect of the element is manifested by the behavior at nodes.
Total is shared equally b/w nodes.

Assembling all terms together,

$$(\text{Term-2})\cdot[T] = (\text{Term-1}) + (\text{Term-3})$$

$$\Rightarrow \frac{k}{l_e}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}T_i\\ T_j\end{bmatrix} = \begin{bmatrix}q_i''\\ -q_j''\end{bmatrix} + S l_e\begin{bmatrix}1/2\\ 1/2\end{bmatrix}.$$



This is of the form:

$$\underbrace{[\bar{k}]}_{\text{as if like the stiffness of the system}}\;\underbrace{[T]}_{\text{as if behaves like displacement}} = \underbrace{[F]}_{\text{Acts like a force term}} \longrightarrow \text{Similar to spring mass system.}$$

Stiffness matrix.

for 2 nodes we get 2x2 form. ||^ly for 4 nodes we get 4x4 form.

This part of coeff matrix activated for 1st element.

$$\begin{array}{c|c} & \begin{matrix}1 & \quad 2 & \quad 3 & \quad 4\end{matrix}\\ \begin{matrix}1\\2\\3\\4\end{matrix} & \begin{bmatrix}1 & -1 & & \\ -1 & 1+1 & -1 & \\ & -1 & 1+1 & -1\\ & & -1 & 1\end{bmatrix}\end{array}\begin{bmatrix}T_1\\ T_2\\ T_3\\ T_4\end{bmatrix} = \begin{bmatrix}q_1''\\ -q_2''+q_2''\\ -q_3''+q_3''\\ -q_4''\end{bmatrix} + S l_e\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}+\frac{1}{2}\\ \frac{1}{2}+\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$$

This part for the third element

The part activated for 2nd element.

Important assumptions: ① Thermal conductivity k as constant.
② All elements have same length. It need not be the case in reality.

Final form:

$$\begin{bmatrix}1 & -1 & 0 & 0\\ -1 & 2 & -1 & 0\\ 0 & -1 & 2 & -1\\ 0 & 0 & -1 & 1\end{bmatrix}\begin{bmatrix}T_1\\ T_2\\ T_3\\ T_4\end{bmatrix} = \begin{bmatrix}q_1''\\ \cancel{-q_2''+q_2''}\ 0\\ 0\\ q_4''\end{bmatrix} + S l_e\begin{bmatrix}1/2\\ 1\\ 1\\ 1/2\end{bmatrix}.$$

Complete the problem by imposing BCs.

At $x=0$, $q_1'' = 0$ (heat flux is zero as insulated)

At $x=L$, T_4 is specified.

Even if T_4 is specified, computer may try to find T_4 with the first BC. In that case, we require matching b/w specified T_4 & computed T_4.

$$k_{41} T_1 + k_{42} T_2 + k_{43} T_3 + k_{44} T_4 = R.$$

Let $T_{4,\text{specified}} = T_4^*$.

Use small computational trick.

Replace R with $L\, T_4^*$ &

replace k_{44} with $k_{44} + L$, where L is a large number.

Thus $k_{41} T_1 + k_{42} T_2 + k_{43} T_3 + (k_{44}+L) T_4 = L\, T_4^*$.

$$T_4 = \frac{-k_{41}}{k_{44}+L} T_1 - \frac{k_{42}}{k_{44}+L} T_2 - \frac{k_{43}}{k_{44}+L} T_3 + \frac{L\, T_4^*}{k_{44}+L}$$

If L is very large,

$$\frac{-k_{41}}{k_{44}+L} \to 0, \quad \frac{-k_{42}}{k_{44}+L} \to 0, \quad \frac{-k_{43}}{k_{44}+L} \to 0 \quad \&$$

$$\frac{L}{k_{44}+L} \to 1.$$

Thus we numerically get,

$$T_4 = T_4^*.$$

Lecture 14: Finite Difference Method (FDM) of discretization.

(4/10)

(Prob-1) Consider the DE: $\frac{d^2 u}{dx^2} + u + x = 0.$

with BC $u(0) = u(1) = 0.$

Solve the above eq^n using (1) Least square
(2) Point collocation
(3) Galerkin
(4) Rayleigh Ritz method.

Choose trial function $\rightarrow u = a \sin \pi x$

(Prob-2) Consider a heat conduction problem with the following governing DE:

$$\frac{d}{dx}\left(A k \frac{dT}{dx}\right) + Q = 0, \quad A = 10m^2,\ k = 5\ J/Kms,$$
$$Q = 100\ J/sm.$$

Domain $2cm \leq x \leq 8cm.$

BCs: $T(x = 2cm) = 0^\circ C$

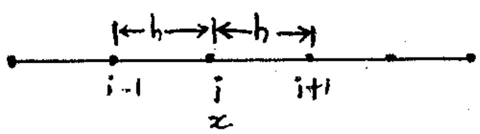
$q''(x = 8cm) = 15\ J/m^2 s.$

Obtain temperature distribution in the domain using FEM with three linear elements, and compare with the analytical solution.

- If solving a structural mechanics problem, 'M' form is essentially a statement of minimization of potential energy of that system, which governs the stability of the system at equilibrium.

- In FDM, we deal with the 'D'-form directly.

Express derivatives in terms of suitable algebraic differences by using Taylor series expansion.



Consider a 1D domain.
We represent the domain with a collection of discrete grid points.
(Don't have the concept of discrete elements here).

$$f(x+h) = f(x) + h f'(x) + \frac{h^2}{2!} f''(x) + \cdots \quad — ①$$

||ly $$f(x-h) = f(x) - h f'(x) + \frac{h^2}{2!} f''(x) \cdots \quad — ②$$

We are interested in an algebraic expression for $f'(x)$.

From ①

Then, $$f'(x) = \frac{f(x+h) - f(x)}{h} \mp \frac{h}{2!} f''(x) \cdots$$

here continuous derivative is represented as discrete algebraic quantities.

Error incurred $-\frac{h}{2!} f''(x) \cdots \rightarrow$ truncation error (TE)

$\sim O(h)$ dictated by the leading order term.

From ②;

$$f'(x) = \frac{f(x) - f(x-h)}{h} + \frac{h}{2!} f''(x) + \cdots \rightarrow \text{truncation error (TE)} \sim O(h)$$

$-② + ①$ gives:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{3!} f'''(x) + \cdots \rightarrow TE \sim O(h^2)$$

$f'(x) = \frac{f(x+h) - f(x)}{h} + O(h) \rightarrow$ Forward difference

$f'(x) = \frac{f(x) - f(x-h)}{h} + O(h) \rightarrow$ Backward difference

$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) \rightarrow$ Central difference

① + ② gives

$$f(x+h) + f(x-h) = 2f(x) + \frac{2 \cdot h^2}{2!} f''(x) + \frac{2h^4}{4!} f''''(x) + \cdots$$

$$f''(x) = \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} - \frac{h^4}{12} f''''(x) \cdots$$ → Central difference for 2nd order PDE.

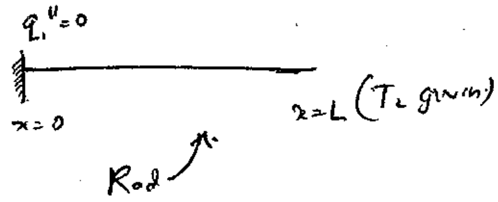
$$f''(x) = \frac{f'(x+h) - f'(x)}{h} \quad \text{forward-difference}$$

$$= \frac{\overbrace{\frac{f(x+h)-f(x)}{h}}^{\to BD} - \overbrace{\frac{f(x)-f(x-h)}{h}}^{\to BD}}{h}$$

$$= \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}$$

Ex Consider 1D, steady-state heat conduction problems.

$$\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S = 0$$



Assumption: k, S both constants

$$k\,\frac{T_{i+1} + T_{i-1} - 2T_i}{h^2} + S = 0$$

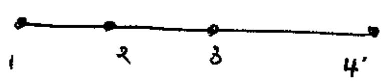
$$T_{i+1} + T_{i-1} - 2T_i + \frac{Sh^2}{k} = 0.$$

↳ algebraic eq^n.

[What determines the whether we should choose large h or small h?

It depends upon the temperature gradients in the domain. It may so happen that at some regions, the gradients may be steep. In such regions, we use finer value of h. At other regions, where the gradient is less, we may opt for a larger h. It all depends on the physics of the problem & our understanding about it]

Consider 4 grid pts:



The above eq^n is valid for internal grid pts not at the boundary. (B/c we don't have T_{i-1} @ left boundary & T_{i+1} at right boundary).

Grid 2 →

$$T_3 - T_1 - 2T_2 + \frac{Sh^2}{k} = 0$$

Grid 3 →

$$T_4 - T_2 - 2T_3 + \frac{Sh^2}{k} = 0$$

Grid 4 →

BC: $T_4 = T_L$ (given)

Grid 1 →

BC: $q'' = 0$

$$\Rightarrow k\frac{dT}{dx}\bigg|_1 = 0$$

So. $\frac{T_2 - T_1}{h} = 0$ FD formula

$\Rightarrow$ $\underline{T_1 = T_2}$

(Assign the value at the boundary with the interior value)

Lecture 15: Well Posed Boundary Value Problems.

- Well posed BVP problems requirements:
  - → Existence of solution
  - → Uniqueness of sol^n.
  - → A small perturbation in BC shouldn't lead to large changes in the solution.

    (This is important b/c such a perturbation may be unwillingly introduced through round-off errors; B/C of it, it may lead to large change in sol^n → oversensitive BC).

- Possible types of BCs: (2nd order problems)

1. Dirichlet BC:
   Value of the dependent variable is specified → EBC.

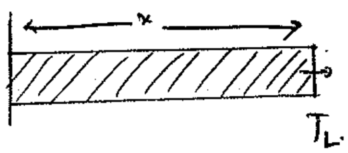
2. Neumann BC:

   Value of the gradient of the dependent variable is specified - NBC.

3. Mixed BC:

   Value of the dependent variable is expressed as a function of the grad.

   Eg: Convective heat transfer BC.



$T_\infty < T_L$.

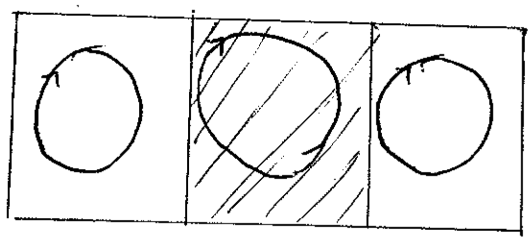
$$-k\frac{dT}{dx}\bigg|_{x=L} = h\,(T_L - T_\infty)$$

The expression can be used even in unsteady case. (replace $\frac{dT}{dx}$ with $\frac{\partial T}{\partial x}$)

This BC says, Whatever is the heat flux coming via conduction at $x=L$, the same is the heat flux leaving $x=L$ via convection. That remains true even if unsteady

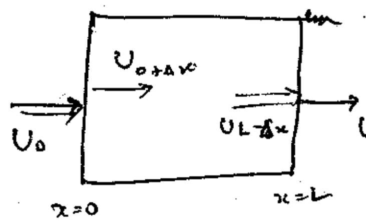
Here temperature at $x=L$ is expressed as a function of gradient of the temperature $(-k\frac{dT}{dx})$. Hence its a mixed BC.

4. Periodic BC:



Say, in the domain, the physics of the problem is such that the solution is periodically repeated.

So, solve 1 part of the domain & extrapolate or extend the solution to other parts of the domain.



$U_L = U_{\Delta x}$

$U_0 = U_{L-\Delta x}$

periodicity = $L - \Delta x$.

Here again we see boundary terms being represented in terms of interior terms, not the opposite.

Q) Is any condition specified at the boundary, a boundary condition?

Ans) Consider a simple 1D, steady-state, heat conduction problem, $S = 0$, $k = \text{const.}$

→ Neumann BC at both the boundaries.

$$\frac{d^2T}{dx^2} = 0$$

$$\frac{dT}{dx} = C_1$$

$$T = C_1 x + C_2.$$

Say, $k = 1\,W/mK$.

$x = 0$, $\quad -k\dfrac{dT}{dx} = 1$

$\Rightarrow \dfrac{dT}{dx} = -1 \Rightarrow C_1 = -1$

$|||^{ly}$ at $x = L$,

$-k\dfrac{dT}{dx} = 1 \Rightarrow \dfrac{dT}{dx} = -1 \Rightarrow C_1 = -1$

Basically
$T = -x + C_2$.

Cannot determine C_2.

⇒ Violates requirement for uniqueness of sol^n?

Plotting in T–x plane gives all sol^{ns} to be parallel straight lines; no 1 sol^n.

→ not legitimate BCs.

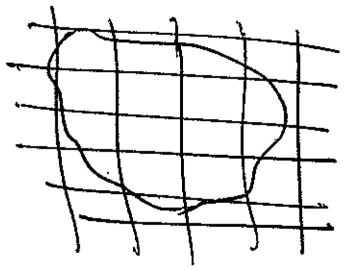
⇒ not well posed BV problem

Lecture 16: Finite Volume Method (FVM) of Discretization.

FDM → simple

Issues: ① Complex geometry.

One has to tediously ~~create~~ ~~to~~ handle the boundary while using cartesian grid.



Taylor series expansion:

$$f(x+h) = f(x) + h f'(x) + \frac{h^2}{2} f''(x) + \dots$$

truncate from here

② But what if $f''(x)$ is large & enough to be non-negligible? (B/c h cannot be tending to zero, it is still finite).

Ex: $f(x) = e^x \rightarrow f''(x)$ = non-negligible

⇒ significant errors while truncating.

→ limitation of Taylor series based method.

Q) What do we expect from the discretization?

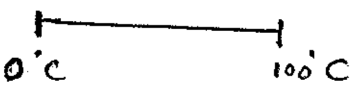
(Ans) → ① Conservativeness → Discretized versions of conservation eq^ns should exhibit that conservative nature.

→ ② Boundedness

→ ③ Transportiveness

- Finite Difference ~~FD~~ discretization may not satisfy conservativeness b/c we haven't explicitly enforced that condition while expanding out the function in Taylor series form and ~~ext~~ truncation. Conservativeness is not inbuilt while coming up with FDM.

• Boundedness: Say we have a rod:



We are interested in the temperature distribution in the rod.
We expect the values in the rod to lie b/w 0 & 100°C. The discretization should also ensure the physical nature of boundedness in the problem.

This boundedness is also not ensured while coming up with FDM.

• Transportiveness: If there is a predominant directionality to the flow involved in the problem, the ~~transport~~ properties should also have a predominant transport direction based on the flow direction.
(For high Re flow, for eg, enthalpy should predominantly be transported downstream).

FEM — relatively more complicated when compared to FDM.
- Strong mathematical basis in error minimization.
- not intuitive physically: all the V-formulation & M-formulation needs to have some physical meaning, which can be difficult to come up with.

For fluid flow, mass flow, conservation is important, while for structural mechanics, minimization of potential energy is important.

- Can handle complex geometries.

Finite Volume Method (FVM):

Step-1: Divide the domain into a number of finite size sub-domains (control Volumes)
Each sub-domain is represented by a finite no. of gridpts.

Step-2: Integrate the governing differential equation over each subdomain

Step-3: Consider a profile assumption for the dependent variable for evaluating the above integrals to express the result in terms of algebraic quantities at the grid points.

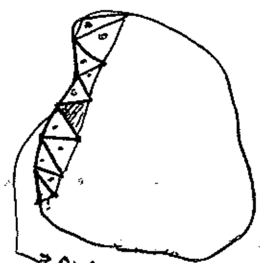
Lecture 17: Illustrative Examples of Finite Volume Method

Ex Steady state convection diffusion with $S=0$.

General transport equation:

$$\cancel{\frac{\partial}{\partial t}(\rho\phi)}^{0} + \nabla\cdot(\rho\vec{V}\phi) = \nabla\cdot(\Gamma\nabla\phi) + \cancel{S}^{0}$$

$$\int_{CV} \nabla\cdot\overbrace{\underbrace{(\rho\vec{V}\phi - \Gamma\nabla\phi)}_{g.d.e.}}^{\vec{J}}\, d\forall = 0$$



$\rho\vec{V}\phi \rightarrow$ advection flux
$\Gamma\nabla\phi \rightarrow$ diffusion flux.

Using divergence theorem,

$$\int_{C.S} \vec{J}\cdot\hat{n}\, ds = 0.$$

Grid pts located at the center of each CV.

Q) Why the name 'finite' volume method?

Ans) While deriving the transport equations we considered an infinitesimally small ~~control~~ element. ~~the~~ ~~A~~ Now we're integrating back to apply it to a finite volume. Hence the name finite volume method

Key step → step-2.

Similarly:

It can be thought of having wt is

$$\int_{CV} \nabla \cdot \left(\rho \vec{V} \phi - \Gamma \nabla \phi \right) \underset{\downarrow \atop r}{w} \, dV = 1 .$$

Galerkin method with wt.

- Requirement of conservation when applied to each control-volume will satisfy the conservation across the whole domain.

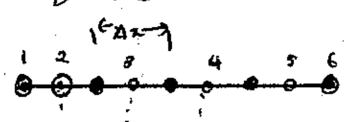
  ∴ unlike FDM, FVM takes into account conservation requirement implicitly.

- Profile assumption method is used just for step-3. Afterwards as post analysis, we no longer require profile assumption. Hence we have more flexibility in choosing profile assumption as compared to FEM.

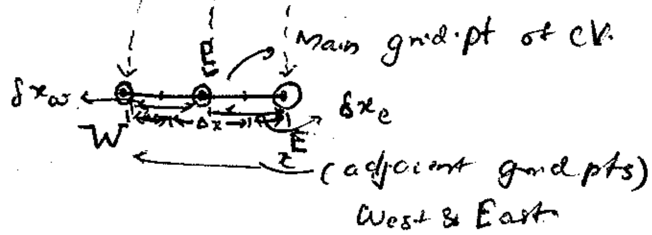
<u>Illustration:</u> 1D steady state heat conduction, with S const.

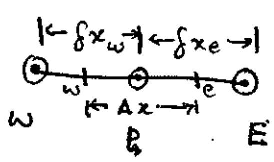
$$\frac{d}{dx}\left(k \frac{dT}{dx} \right) + S = 0$$

Divide 1-D domain into CVs.



In addition to centroids of CVs, we also consider grid pts at the boundary (just so we can impose BCs at those pts).

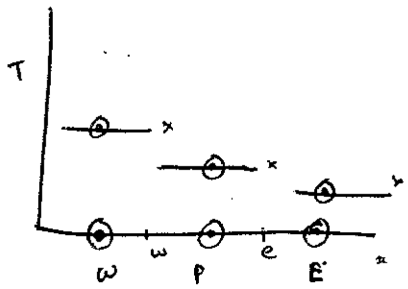




$$\int_w^e \frac{d}{dx}\left(k \frac{dT}{dx} \right) dx + \int_w^e S \, dx = 0 .$$

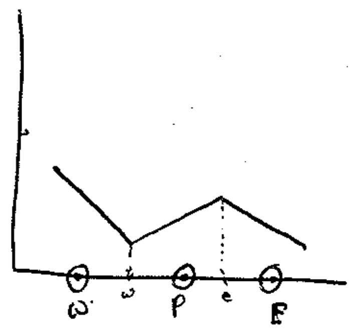
$$\left. k \frac{dT}{dx} \right]_e - \left. k \frac{dT}{dx} \right]_w + S \, \Delta x = 0 .$$

· Choosing a profile assumption:



Can consider piecewise continuous functions for each CV.
But not here. B/c we need $\frac{dT}{dx}$ here. Since T = const for each CV. This is not a valid profile assumption. (Here discontinuity is not a problem).

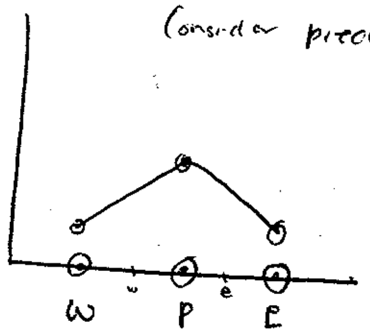
Piecewise linear non-const profile:



Problems with this case:
$k\frac{dT}{dx}$ is not continuous, i.e. physically it means heat flux is not continuous, which is incorrect.

⇒ not an acceptable profile.

Consider piecewise linear profile b/w the grid pts



This will work b/c $\frac{dT}{dx}$ is not evaluated @P, but rather at the faces of each control volume w, e.
⇒ valid profile assumption

Profile assumption: Piecewise linear T b/w grid pts.

$$\Rightarrow \quad k\frac{dT}{dx}\Big]_e - k\frac{dT}{dx}\Big]_w + S\Delta x = 0$$

$$\rightarrow \quad k_e \frac{T_E - T_P}{\delta x_e} - k_w \frac{T_P - T_W}{\delta x_w} + S\Delta x = 0$$

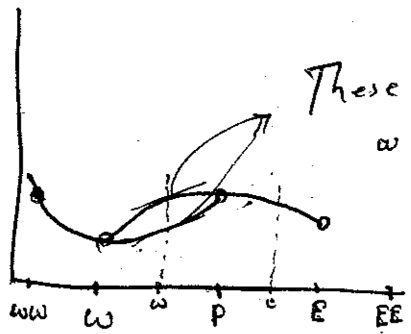
$$\Rightarrow -\left(\frac{k_e}{\delta x_e} + \frac{k_w}{\delta x_w}\right) T_P \cancel{-\left(\frac{k_e}{\delta x_e} + \frac{k_w}{\delta x_w}\right) T_P} + \frac{k_e}{\delta x_e} T_E + \frac{k_w}{\delta x_w} T_W + S\Delta x = 0 .$$

$$\underset{(i)}{a_P T_P} = \underset{(i+1)}{a_E T_E} + \underset{(i-1)}{a_W T_W} + b$$

where $a_E = \frac{k_e}{\delta x_e}$ $\quad a_W = \frac{k_w}{\delta x_w}$, $\quad a_P = a_E + a_W$, $\quad b = S\Delta x$

$\dfrac{k}{\delta x} \longrightarrow$ conductance (physical meaning)

Would it have been better to consider a higher order interpolation function?



These slopes need not necessarily be equal. It happens only when w lies exactly in b/w W and P. From Rolle's Thm., a chord b/w two pts has a slope that will be attained by the curve b/w those two pts at some pt b/w them, given the curve is differentiable at all points. From MVT, this will lie at the midpoint of the curve. ~~It is only happens when w lies~~

So higher order interpolate functions need not necessarily given more accurate results in FVM* (counter-intuitive)

Lecture-18: Illustrative Examples of Finite Volume Method (Contd)

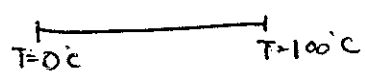
1-D steady state heat conduction equation:

$$\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S = 0$$

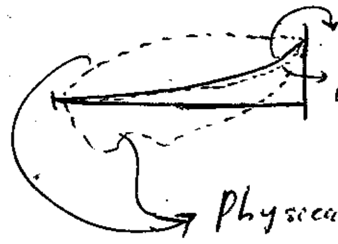
If $k, S \rightarrow fn(T) \longrightarrow$ non linear eq^n $\rightarrow$ solved via iterative process

- Requirements of discretization:
  (1) Physical consistency
  (2) Overall balance

Eg:



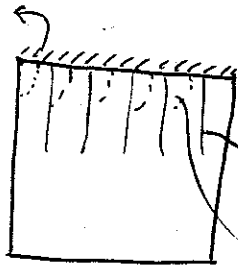
Assume $S=0$
$k =$ const.



Say, exact sol^n.
Good approx. sol^n.

Physically inconsistent solutions (for eg. Boundedness invalidated)

insulated



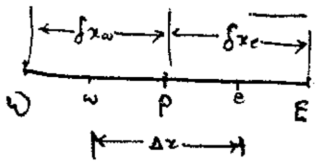
Iso thermal lines in 2D domain

These isotherms are physically consistent (normal to insulated surface)

These isotherms are physically inconsistent.

Variable S : $\longrightarrow$ let $S(T)$

Ex: S is a linear fn of T.
i.e. $S = a + bT$.



$$k\frac{dT}{dx}\Big]_e - k\frac{dT}{dx}\Big]_w + \int_w^e (a+bT)\,dx = 0.$$

→ piecewise const T profile within each CV.

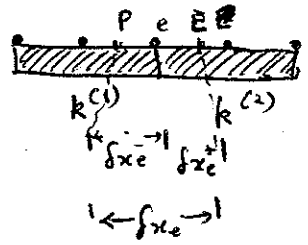
Profile assumption: piecewise linear T b/w grid pts.

$$\Rightarrow \quad k_e\frac{(T_E - T_P)}{\delta x_e} - k_w\frac{(T_P - T_W)}{\delta x_w} + (a + bT_P)\Delta x = 0.$$

$$a_P T_P = a_E T_E + a_W T_W + b.$$

where $a_E = \dfrac{k_e}{\delta x_e}$, $a_W = \dfrac{k_w}{\delta x_w}$, $a_E + a_W - b\Delta x = a_P$, $b = a\Delta x$.

Composite material with position dependent k:



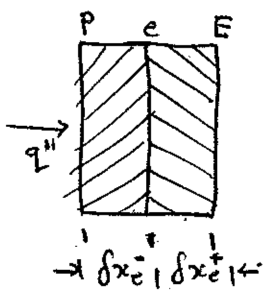
→ 4 CVs

→ Requirement: Thermal conductivity at the interface to find.

Since interface shared by both materials, an equivalent thermal conductivity needs to be described.

If k_P & k_E is known, intuitive: $k_e = \dfrac{k_E + k_P}{2}$ → linear interpolation

Physical assessment ~~of~~ of this k_e needs to be performed. → AN formulation.



$$q''_{left} = \left(\frac{T_e - T_P}{\delta x_e^-}\right)(-k_P)$$

$$= \frac{T_e - T_P}{\left(-\dfrac{\delta x_e^-}{k_P}\right)}$$

$$q''_{left} = q''_{right}$$

$$q''_{right} = \frac{T_E - T_e}{\delta x_e^+}(-k_P)$$

$$= \frac{T_E - T_e}{\left(-\dfrac{\delta x_e^+}{k_E}\right)}$$

$$\frac{T_P - T_e}{\dfrac{\delta x_e^-}{k_P}} = \frac{T_e - T_E}{\dfrac{\delta x_e^+}{k_E}} = \frac{T_P - T_E}{\dfrac{\delta x_e}{k_e}} = \frac{T_P - T_E}{\dfrac{\delta x_e^-}{k_P} + \dfrac{\delta x_e^+}{k_E}}.$$

→ where k_e is the equivalent thermal conductivity.

If $\delta x_e^- = \delta x_e^+$,

$$\frac{2}{k_e} = \frac{1}{k_P} + \frac{1}{k_E}$$

→ H.M. formulation

$$\text{or } k_e = \frac{2 k_P k_E}{k_P + k_E} = \frac{2}{\frac{1}{k_P} + \frac{1}{k_E}}$$

- Limiting cases:

$k_P >> k_E$

Then AM → $k_e \approx \frac{k_P}{2}$

HM → $k_e \approx 2k_E$

For interfacial conductivity variation, HM formulation is physically much more appealing. why?
Say $k_E = 0$. That is equivalent to highly insulated material-E. In that case k_e should also be 0 at the interface. This is reflected in HM, while not at all in AM formulation.

Lecture 19: Basic rules of finite Volume Discretization.

4 Basic rules (of 1D steady state diffusion type problem):

(1) Physical consistency of fluxes at Control Volume faces.

→ Profile should be chosen in such a way that there is no discontinuity of flux at control volume faces.

(2) All coefficients in the discretized equation must be of the same sign.

Eg; say, $b = 0$

$$10\,T_P = 15\,T_E - 5\,T_W$$

Say $T_E = 10$

$T_W = 100$

$$\therefore T_P = \frac{15 \times 10 - 5 \times 100}{10}$$

$$= -35$$

T_P is not bounded b/w T_E & T_W.

⇒ Physically inconsistent.

This inconsistency has originated b/c of the -ve sign in the discretization eq^n ($10\,T_P = 15\,T_E - 5\,T_W$)

If, $10 T_p = 5 T_E + 5 T_W$

$$T_P = \frac{5\times 10 + 5\times 100}{10}$$

$= 55$

Here $T_E \le T_P \le T_W$

Hence consistent.

(Same sign coefficients)

By sign convention, we will consider that sign to be +ve.

(3) If the source term is linearized as: $S = S_c + S_P T_P$,

Then S_P must be $-$ve $\longrightarrow$ Extension of consideration-#2.

(4) If a linear governing DE is discretized, its discretized version should satisfy the following requirement:

If T is a sol^n, then $T+c$ is also a sol^n.

We are interested to see

$$a_P (T_P + c) \stackrel{?}{=} a_E (T_E + c) + a_W (T_W + c) \cdots (1)$$

We already know

$$a_P T_P = a_E T_E + a_W T_W \cdots (2)$$

(1) $-$ (2)

$$\Rightarrow \quad a_P\, c = (a_E + a_W)\, c$$

$$\text{or} \quad a_P = a_E + a_W$$

Thus linearity is satisfied.

Q) What if the source term is non-linear?

Method to linearize a non-linear source term $\longrightarrow$

Source term linearization

Ex-1 Say, $S = 3 + 4T$ $\qquad S = S_c + S_P T_P$

$S_c = ?$ $\quad S_P = ?$

From the form, it may appear that $S_c = 3$, $S_P = 4$.

Just now we've seen that S_P should be taken $-$ve. So the above form is not valid.

Use iterative process. Initially take:

$$S_c = 3 + 4T_p^*, \quad S_p = 0.$$

If the iteration has a tendency to diverge fast, considering appropriate initial value for S_p may be required.

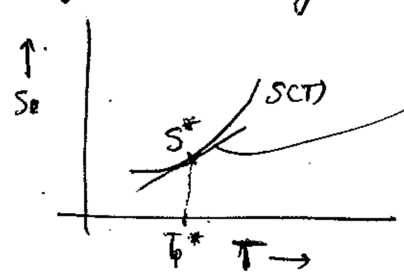
Ex-2 $\quad S = 3 - 4T$

Here $S_c = 3$, $S_p = -4$ is valid & the best choice.

To slow down the convergence, we may choose $S_c = 3 + 4T_p^*$, $S_p = -8$.

Ex-3. $\quad S = 3 - 4T^3$.

Say we have any arbitrary $S(T)$.



Best linear function for this → tangent at T_p^*.

(why? Same first order derivative for tangent as that of the non-linear curve).

$$\frac{S - S^*}{T - T_p^*} = \left.\frac{dS}{dT}\right|_{T_p^*}$$

$$S - S^* = \left.\frac{dS}{dT}\right|_{T_p^*}(T - T_p^*)$$

$$S^* = 3 - 4T_p^{*3}$$

$$\left.\frac{dS}{dT}\right|_{T_p^*} = -12T_p^{*2}$$

$$S - S^* = S - (3 - 4T_p^{*3})$$

$$= -12T_p^{*2}(T - T_p^*)$$

$$\rightarrow \quad S = 3 + 8T_p^{*3} - 12T_p^{*2}T.$$

$$\therefore \quad S_c = (3 + 8T_p^{*3})$$

$$S_p = (-12T_p^{*2}) \longrightarrow S_p \text{ -ve satisfied as } T_p^{*2} \geq 0$$

Ex-4

$$S = 3 + 4T^3$$

$$S^* = 3 + 4T_P^{*3} + \left.\frac{dS}{dT}\right|_{T_P^*} = 12T_P^{*2}(T - T_P^*)$$

$$S - [3 + 4T_P^{*3}] = 12T_P^{*2}(T - T_P^*)$$

$$S = 3 - 8T_P^{*3} + 12T_P^{*2}T$$

$$\Rightarrow S_C = 3 - 8T_P^{*3}$$

$S_P = 12T_P^{*2}$ → Here +ve S_P cannot work!

What to do then?

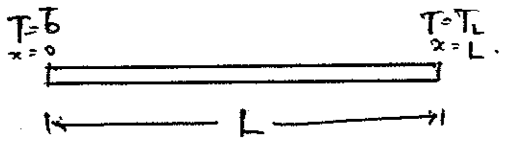
Dump the entirety to S_C.

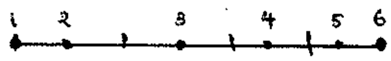
$$S_C = 3 + 4T_P^{*3}, \quad S_P = 0$$

This linearization is mathematically correct, but will give physically inconsistent solution

Lecture 20: Implementation of boundary conditions in FVM.

Ex-1





$$a_P T_P = a_E T_E + a_W T_W + b$$

For grid point-2,

$$a_2 T_2 = a_3 T_3 + a_1 T_1 + b$$

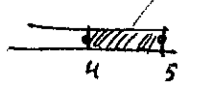
But at point-1, $T_1 = T_0$ (given)

(Use penalty approach in FEM for further analysis)

Ex-2.

To implement BC, consider CV at the boundary.

Take ½ of a CV. (Smaller length can capture sharper gradients accurately)



G. de: $\frac{d}{dx}\left(k\frac{dT}{dx}\right) + S = 0$.

Integrating gde over the half CV →

$$\int_4^5 \frac{d}{dx}\left(k\frac{dT}{dx}\right)dx + \int_4^5 S\,dx = 0$$

Total length of $\frac{1}{2}$ CV = $\frac{\Delta x}{2}$.

$$\Rightarrow \underbrace{k\frac{dT}{dx}\Big]_5}_{\text{as } -q_L''} - \underbrace{k\frac{dT}{dx}\Big]_4}_{} + S\frac{\Delta x}{2} = 0.$$

Use profile assumption of piecewise linear nonconstant T.

$$\Rightarrow k\frac{T_5 - T_4}{\left(\frac{\Delta x}{2}\right)}$$

$$\Rightarrow -q_L'' - \frac{2k}{\Delta x}(T_5 - T_4) + S\frac{\Delta x}{2} = 0.$$

$$\therefore \frac{2k}{\Delta x}(T_5 - T_4) = -q_L'' + \frac{S\Delta x}{2}$$

$$T_5 = T_4 + \frac{S\Delta x^2}{4k} - \frac{q_L''\Delta x}{2k}$$

→ Eq^n of the form: $a_5 T_5 = \cancel{a_6 T_6} + a_4 T_4 + b$

where $a_5 = 1$
$a_4 = 1$,
$b = \left(\frac{S\Delta x^2}{4k} - \frac{q_L''\Delta x}{2k}\right)$.

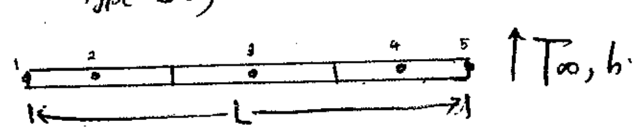
(Expression written as boundary term as a function of interior terms)

FD

$$q_L'' = -k\left(\frac{T_5 - T_4}{\frac{\Delta x}{2}}\right)$$

$$T_5 = T_4 - \frac{q_L''\Delta x}{2k}$$

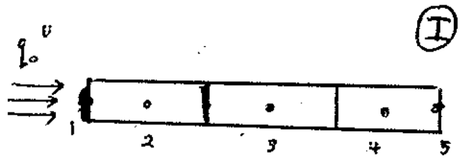
Ex-3 (Mixed type BC)

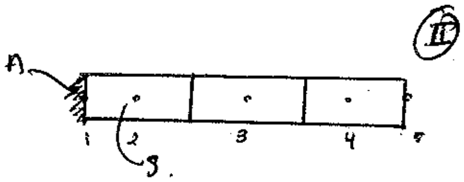


$$\underbrace{-k\frac{dT}{dx}\Big|_{x=L}}_{q_L''} = h\left(T_{x=L} - T_\infty\right)$$

$$\underbrace{-q_L''}_{h(T_5 - T_\infty)} - \frac{2k}{\Delta x}(T_5 - T_4) + \frac{S\Delta x}{2} = 0$$

$h(T_5 - T_\infty)$ → $(\,)T_5 = (\,)T_4 + (\,)$ → very much analogous to Neumann BC case.





$$S_{extra} = \frac{(q_0'' A_{face})}{V_{cv}}$$

$(q_0'' A_{face})$ → total rate of heat transfer

S_{extra} → Volumetric heat generation (Heat source term)

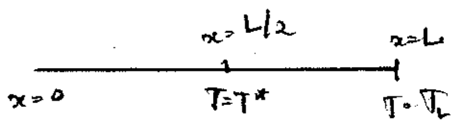
I → Heat flux entering through left boundary.

II → Source at 2. Requirement: equivalent heat flux passing through face b/w 2 & 3 same as in I. Boundary-1 insulated.

Case-I is usually treated equivalently as case-II by dumping all the flux through corresponding heat source term $(S)_{extra}$. Already present source term unaffected. Why do so?

More rapid convergence to solution.
In case I, heat flux has to penetrate through 1 & then through face 2-3. While in case II, with the introduction of S_{extra}, flux through 2-3 face is handled in 1 shot.

- Difference observed is marginal in most cases.

• 

Can we specify the B/C at an internal Grid pt & still expect the problem to be well-posed?

$$\frac{d}{dx}\left(k\frac{dT}{dx}\right) = 0 \quad (\text{Take } S = 0 \text{ for this eg}).$$
$k = \text{const}$

$$\Rightarrow \frac{dT}{dx} = C_1$$

$$T = C_1 x + C_2$$

$x = L/2,\ T = T^*$ →
$x = L,\ T = T_L$ → } can obtain both C_1 & C_2.

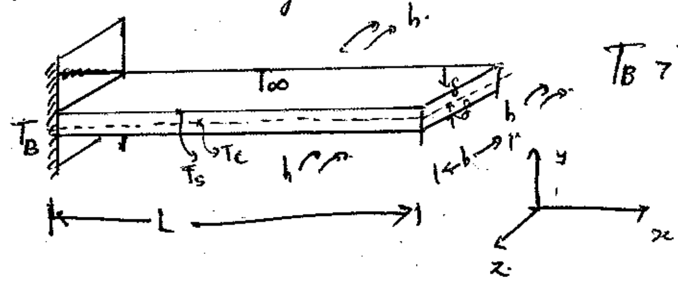
↳ perfectly valid BC. Only thing is that it isn't specified at the physical boundary.

Cannot do the same for IVP. why?

B/C time is a one way coordinate system. B/c whatever is happening now cannot influence what has already happened a while back.

Lecture 21: Implementation of boundary conditions in FVM (contd)

Ex 1-D steady state heat conduction in a fin.



$T_B > T_\infty$

Thermal resistance in x is most important. Out of y & z, since $b \gg 2\delta$, thermal resistance along z direction is the next significant one.

$\frac{\delta}{kA_1}$, $\frac{b}{kA_2}$ | Either the consideration $b \gg 2\delta$, or $2\delta \gg b$ makes this problem 2D.

$$k\frac{T_c - T_s}{\delta} \sim h(T_s - T_\infty)$$

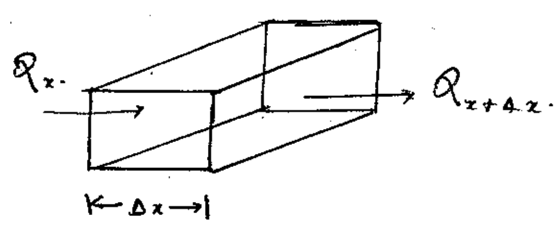
$$\frac{T_c - T_s}{T_s - T_\infty} \sim \frac{h\delta}{k}$$ → Bi_δ → assume this to be small.

If $\frac{T_c - T_s}{T_s - T_\infty} \ll 1 \Rightarrow T_c \sim T_s$ ⇒ no necessity to analyse the temp difference b/w centerline & outside surface

⇒ no necessity to analyse heat transfer in y direction

⇒ 1-D problem.

Take a section out of the fin:



$$\dot{Q}_x - \Delta Q_{conv} = \dot{Q}_{x+\Delta x}$$

$$\cancel{\dot{Q}_x + \dot{Q}_{x+\Delta x}} \quad \dot{Q}_x = q_x'' A$$

$$\dot{Q}_{x+\Delta x} = q''_{x+\Delta x} A$$

$$\dot{Q}_x - \dot{Q}_{x+\Delta x} = (q''_x - q''_{x+\Delta x})A$$

$$q''_{x+\Delta x} = q''_x + \Delta x \frac{dq''_x}{dx} + \dots$$

$$\Rightarrow \dot{Q}_x - \dot{Q}_{x+\Delta x} = -\left[\frac{dq''_x}{dx}\Delta x + \dots\right]A.$$

* Top $= 2hb\,\Delta x\,(T-T_\infty)$
  Front $= 2h(2\delta)\,\Delta x\,(T-T_\infty)$

$\Delta Q_{conv} = P h (T - T_\infty)\Delta x$
$P =$ perimeter $= 2(b + 2\delta)$.

~~P = perimeter~~

$\dot{Q}_x - \dot{Q}_{x+\Delta x} = \Delta Q_{conv}$

$\therefore\; -\dfrac{dq_x''}{dx}\Delta x + \cdots = \dfrac{Ph\,(T-T_\infty)\,\Delta x}{A}$

Take limit as $\Delta x \to 0$.

$\Rightarrow -\dfrac{dq_x''}{dx} A = Ph\,(T-T_\infty)$

and $q_x'' = -k\dfrac{dT}{dx}$

$+\dfrac{d}{dx}\left(kA\dfrac{dT}{dx}\right) = Ph\,(T-T_\infty)$.

$\dfrac{d}{dx}\left(kA\dfrac{dT}{dx}\right) - Ph\,(T-T_\infty) = 0$.

↳ Governing DE.

* Discretize using FVM:

The above eq^n is of the form

$\dfrac{d}{dx}\left(k\dfrac{dT}{dx}\right) + S = 0$

with $S = -\dfrac{Ph}{A}(T-T_\infty)$

↑ implicitly linear source term.

$S = S_c + S_p T$

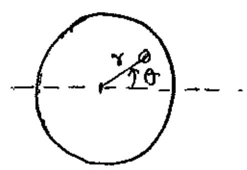
$S_c = \dfrac{Ph\,T_\infty}{A}$, $S_p = -\dfrac{Ph}{A}$ (−ve).

⇒ Its a well posed problem.

(H/W) (1) At $x=0$, $T = T_B$
(2) At $x = L$, $\dfrac{dT}{dx} = 0$.

Non-dim entire equations: $\bar{x} = \dfrac{x}{L}$.
Find Temp distribution with FDM, FEM & FVM.

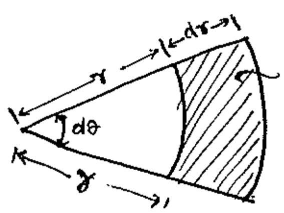
Ex 1D steady state heat conduction in cylindrical coordinates.



Consider z large & no variation along z.
Axially symmetric ⇒ no variation along θ.

Governing DE:

$\dfrac{1}{r}\dfrac{d}{dr}\left(rk\dfrac{dT}{dr}\right) + S = 0$.



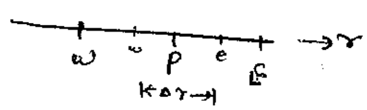
$dA = r\,d\theta\,dr$
$dV = r\,d\theta\,dr\,dz$

$dV = (r\,dr)(d\theta)(dz)$

form integral eq^n using this in ↓

$\displaystyle\int \frac{1}{r}\frac{d}{dr}\left(rk\frac{dT}{dr}\right) r\,dr + \int S\,r\,dr = 0$

↳ Multiplying with $r\,dr$ remove singularity induced by $\dfrac{1}{r}$.

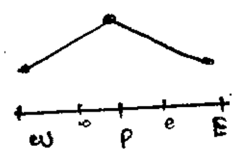


↳ length of the CV.

$\left. rk\dfrac{dT}{dr}\right]_{r+\Delta r} - \left. rk\dfrac{dT}{dr}\right]_{r} + \displaystyle\int_w^e (S_c + S_p T)\, r\,dr = 0$

(Assume, $S = S_c + S_p T$. Constant temp over CV)

Make piecewise linear profile ~~ass~~ assumption b/w the grid pts.



$$r_e k_e \frac{T_E - T_P}{\delta r_e} - r_w k_w \frac{T_P - T_W}{\delta r_w}$$

$$+ \left(S_c + S_p T_P\right) \frac{r_e^2 - r_w^2}{2} = 0$$

Organise the eq^n in the form:

$$a_P T_P = a_E T_E + a_W T_W + b$$

where $\cancel{a_P = \frac{r_e k_e}{\delta r_e} + \frac{r_w k_w}{\delta r_w}}$

$$a_E = \frac{r_e k_e}{\delta r_e} \qquad a_W = \frac{r_w k_w}{\delta r_w}$$

$$a_P = a_E + a_W - S_P \left(\frac{r_e^2 - r_w^2}{2}\right)$$

$$b = S_c \left(\frac{r_e^2 - r_w^2}{2}\right)$$

- If it were spherical coord system, eq^n would have been

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{dT}{dr}\right) + S = 0$$

and while forming the integral eq^n, we'd have to multiply with $r^2 dr$ instead of $r dr$ ~~...~~ before integrating to remove singularities induced by $\frac{1}{r^2}$ term.

$$\boxed{\text{Lecture 22: 1D- Unsteady state Diffusion Problem.}}$$

FVM for:
1D unsteady state diffusion problem: (diffusion; so no fluid flow terms)

$$\frac{\partial}{\partial t}(\rho C_p T) + \cancel{\nabla\cdot(\rho C_p T \vec{V})} = \nabla\cdot(k \nabla T) + S \qquad \left[\nabla\cdot(k\nabla T) = \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)\right]$$

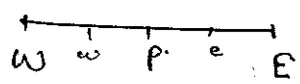
$$\Rightarrow \overbrace{\frac{\partial}{\partial t}(\rho C_p T)}^{\text{unsteady term}} = \overbrace{\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)}^{\text{conduction term}} + S.$$

Our obj: how to take care of the new term $\frac{\partial}{\partial t}(\rho C_p T)$?

Integrate eq^n over the domain.
Here domain consists of both t & x.
So our elemental domain consist of both dt & dx. {Take $S = 0$}

$$\therefore \int_w^e \int_t^{t+\Delta t} \frac{\partial}{\partial t}(\rho C_p T)\, dt\, dx = \int_t^{t+\Delta t}\int_w^e \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) dx\, dt.$$

Treat both terms separately.



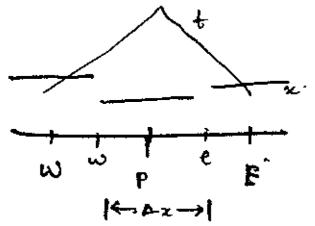
$$\int_w^e \int_t^{t+\Delta t} \frac{\partial}{\partial t}(\rho C_p T)\, dt\, dx$$

$$= \int_w^e \left[(\rho C_p T)^{t+\Delta t} - (\rho C_p T)^t\right] dx$$

Make profile assumption for T variation in space;

Simplest assumption: piecewise constant profile

(why? Here no need to take piecewise linear terms about grid pts b/c there are no derivative terms involved).



Assume ρC_p a constant.

$$\therefore \int_w^e \rho C_p T\Big]^{t+\Delta t} - \rho C_p T\Big]^{t}\, dx$$

$$= \rho C_p \left(T_P^{\,t+\Delta t} - T_P^{\,t}\right) \Delta x$$

Term-2:

$$\int_t^{t+\Delta t}\int_w^e \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x}\right) dx\, dt$$

$$= \int_t^{t+\Delta t} \left(k \frac{dT}{dx}\Big]^e - k\frac{dT}{dx}\Big]^w \right) dt$$

Profile assumption for t:

piecewise linear t.

$$\Rightarrow = \int_t^{t+\Delta t} \left\{ k_e \left(\frac{T_E - T_P}{\delta x_e}\right) - k_w \frac{(T_P - T_W)}{\delta x_w} \right\} dt$$

$$\int_t^{t+\Delta t} T\,dt = ??$$

Make profile assumption:

$$= \left[(1-f)\, T_P^{\,t} + f\, T^{\,t+\Delta t}\right] \Delta t$$

$$\Rightarrow = \left[\frac{k_e}{\delta x_e}\left\{(1-f)\, T_E^{\,t} + f\, T_E^{\,t+\Delta t}\right\} - \frac{k_e}{\delta x_e}\left\{(1-f)\, T_P^{\,t} + f\, T_P^{\,t+\Delta t}\right\} - \frac{k_w}{\delta x_w}\left\{(1-f)\, T_P^{\,t} + f\, T_P^{\,t+\Delta t}\right\} \right.$$

$$\left. + \frac{k_w}{\delta x_e}\left\{(1-f)\, T_W^{\,t} + f\, T_W^{\,t+\Delta t}\right\}\right] \Delta t.$$

For notational convenience, write

$$T^t = T^0$$

$$T^{t+\Delta t} = T^1 = T.$$

Term① = Term②.

$$\Rightarrow a_P T_P = a_E T_E + a_W T_W + \underbrace{a_P^0 T_P^0}_{\text{new term appears!}} + b.$$

where,

$$a_E = \frac{k_e f}{\delta x_e}, \quad a_W = \frac{k_w f}{\delta x_e}.$$

$$a_P^0 = -\frac{(1-f)}{f}(a_E + a_W) + \rho C_p \frac{\Delta x}{\Delta t}$$

$$b = \frac{k_e}{\delta x_e}(1-f)\, T_E^0 + \frac{k_w}{\delta x_w}(1-f)\, T_W^0$$

$$a_P = \rho C_p \frac{\Delta x}{\Delta t} + \frac{k_e f}{\delta x_e} + \frac{k_w f}{\delta x_w}$$

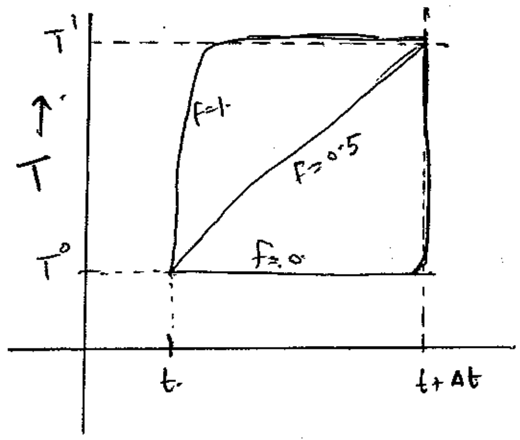
$$= a_E + a_W + \rho C_p \frac{\Delta x}{\Delta t}$$

Earlier there were two neighbors E and W (spatial neighbors). Now we have a temporal neighbor P^0.

We only have one neighbour for time b/c whatever has happened at time $t+\Delta t$ is influenced by the events at time t and not by events at time $t+2\Delta t$.

That's why 2 space neighbors & 1 time neighbor.

Choices of f



- Remember $0 \le f \le 1$

$T^t = T^0$

$T^{t+\Delta t} = T' = T$

$f=0 \rightarrow$ fully explicit scheme

$f=1 \rightarrow$ implicit scheme

$f=1/2 \rightarrow$ Crank-Nicholson scheme.

$f=0$:

$$a_E = 0, \quad a_W = 0, \quad a_P = \rho C_p \frac{\Delta x}{\Delta t},$$

$$a_P^0 = \rho C_p \frac{\Delta x}{\Delta t} - \frac{k_e}{\delta x_e} - \frac{k_w}{\delta x_w}$$

$$b = \frac{k_e}{\delta x_e} T_E^0 + \frac{k_w}{\delta x_w} T_W^0$$

$$a_P T_P = a_P^0 T_P^0 + b$$

$$T_P = \frac{a_P^0 T_P^0}{a_P} + \frac{b}{a_P}$$

$T_P = f(T_E^0, T_P^0, T_W^0)$ expressed in an explicit form.

$f=1$:

$$a_E = \frac{k_e}{\delta x_e}, \quad a_W = \frac{k_w}{\delta x_w}, \quad a_P = a_E + a_W + \rho C_p \frac{\Delta x}{\Delta t}$$

$$a_P^0 = \rho C_p \frac{\Delta x}{\Delta t}, \quad b = 0.$$

> **Lecture 23: 1-D Unsteady State Diffusion Problem (contd.)**

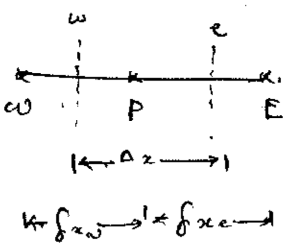
We should have $a_P^0 \geq 0$.

$f=0$.

$$a_P^0 = \rho C_p \frac{\Delta x}{\Delta t} - \frac{k_e}{\delta x_e} - \frac{k_w}{\delta x_w}$$

Let $k_e = k_w = k$ (for algebraic simplicity)

$\delta x_e = \delta x_w = \delta x = \Delta x$.



$$a_P^0 = \rho C_p \frac{\Delta x}{\Delta t} - \frac{2k}{\Delta x}.$$

Conditions for $a_P^0 \geq 0$:

$$\rightarrow \rho C_p \frac{\Delta x}{\Delta t} \geq \frac{2k}{\Delta x}$$

$$\Rightarrow \alpha \frac{\Delta t}{\Delta x^2} \leq \frac{1}{2} \qquad \alpha \text{- Thermal diffusivity} = \frac{k}{\rho C_p}.$$

$$\frac{\alpha t_c}{L_c^2} = F_o \rightarrow \text{Fourier number}$$

t_c → characteristic time

L_c → characteristic length.

Stability criterion for the explicit scheme.

Round off errors can propagate & amplify with calculations. If such a thing happens & it is inherent to the scheme itself, then such a scheme is an unstable scheme.

Key requirement of stable scheme:

→ physically consistent: coeffs are of same sign. If temp is increased at a pt, then that change will cause the temperature at neighbouring pts to increase as well & not decrease.

B/c of the condition $\frac{\alpha \Delta t}{\Delta x^2} \leq \frac{1}{2}$, change in grid spacing also affects the size of a time step to be chosen.

$\frac{\alpha t_c}{L^2} = \frac{t_1}{t_c}$, $t_1 \to \Delta t$

$t_c \to \frac{\Delta x^2}{\alpha}$ → characteristic thermal time over which a thermal disturbance propagates by thermal diffusion in a medium.

→ conditionally stable scheme: the scheme works as long as Δt, Δx constraint is followed.

$f=1$ →

all coeffs same sign

⇒ unconditionally stable.

$f=0.5$ →

$$\rho C_p \frac{\Delta x}{\Delta t} - \frac{2k}{\Delta x} \times \frac{1}{2} \geq 0$$

$$\Rightarrow \frac{\alpha \Delta t}{\Delta x^2} \leq 1.$$

Lecture 24: Consequences of Discretization of Unsteady State Problems.

Consequences of time-discretization in Finite Difference:

Errors associated with any Taylor series based discretization:

→ Consistency: - characteristic of a numerical scheme.

Consistent if in the limit of grid size & time step size → 0, the algebraic eqns mimic the same behavior as that of its parent diff. eqn.

This happens when error is nullified at such refined scales. ⇒ nullification of truncation error as grid size & time step size tends to zero in the limit.

⇒ discretized eqn tends to behave same as g.d.e.

→ Stability: Just like consistency talks about truncation errors, stability talks of round off errors. Round off errors in a numerical scheme is similar to physical perturbation. How strongly these perturbations propagate/amplify in the presence of numerical calculation determines stability.

Stable ⇒ no amplification of numerical perturbations due to propagation of round-off errors.

→ Convergence: ~~As~~ In the limit of grid size and time step size tends to 0, numerical soln → exact solution.

Lax equivalence theorem: for linear problems,

Consistency + stability ⇒ convergence.

Consistency + stability need not ensure convergence for a non-linear problems. Such problems can have multiple solutions. To test for convergence in non linear problems, the following is done. At a finite grid size & time step size evaluate the soln. Then take finer grid & smaller time step & evaluate soln. As finer & finer grid & time step is used, if the soln is found to be grid independent & time-step independent, the non-linear problem has convergence.

Finite difference schemes on 1-D unsteady state diffusion problems:

G.de: $\rho C_p \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2}$

(assume constant properties).

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \quad \left(\alpha = \frac{k}{\rho C_p}\right).$$

Use FTCS.

FT → Forward Time

CS → Central Space.

Corresponding Taylor series:

$x-\Delta x$ x $x+\Delta x$

$i-1$ (W) i (P) $i+1$ (E)

Superscript:

n → current time (t)

$n-1$ → previous time ($t-\Delta t$)

$n+1$ → next time. ($t+\Delta t$)

$$T_i^{n+1} = T_i^n + \left.\frac{\partial T}{\partial t}\right]_i \Delta t + \left.\frac{\partial^2 T}{\partial t^2}\right]_i \frac{\Delta t^2}{2} + O(\Delta t^3)$$

$T_i^{n+1} \to T_x(t+\Delta t)$

$$\left.\frac{\partial T}{\partial t}\right]_i = \frac{T_i^{n+1} - T_i^n}{\Delta t} - \left.\frac{\partial^2 T}{\partial t^2}\right|_i \frac{\Delta t}{2} + O(\Delta t^2)$$

Truncated representation of forward time derivative:

$$\left.\frac{\partial T}{\partial t}\right|_i = \frac{T_i^{n+1} - T_i^n}{\Delta t}.$$

$$T_{i+1}^n = T_i^n + \left.\frac{\partial T}{\partial x}\right|_i^n \Delta x + \left.\frac{\partial^2 T}{\partial x^2}\right|_i^n \frac{\Delta x^2}{2!} + \left.\frac{\partial^3 T}{\partial x^3}\right|_i^n \frac{\Delta x^3}{6} + \left.\frac{\partial^4 T}{\partial x^4}\right|_i^n \frac{\Delta x^4}{24} + O(\Delta x^5)$$

~~$O(\Delta x^4)$~~

$$T_{i-1}^{n} = T_i^n - \frac{\partial T}{\partial x}\Big|_i \Delta x + \frac{\partial^2 T}{\partial x^2}\Big|_i \frac{\Delta x^2}{2} - \frac{\partial^3 T}{\partial x^3}\frac{\Delta x^3}{6} + \frac{\partial^4 T}{\partial x^4}\frac{\Delta x^4}{24} + O(\Delta x^5)$$

$$T_{i+1}^n + T_{i-1}^n = 2T_i^n + \frac{\partial^2 T}{\partial x^2}\Big|_i \Delta x^2 + \frac{\partial^4 T}{\partial x^4}\Big|_i \frac{\Delta x^4}{12} + O(\Delta x^6).$$

$$\frac{T_{i+1}^n + T_{i-1}^n - 2T_i^n}{\Delta x^2} - \frac{\partial^4 T}{\partial x^4}\frac{\Delta x^2}{12} + O(\Delta x^4) = \frac{\partial^2 T}{\partial x^2}.$$

Sub into g.d.e.

$$\frac{T_i^{n+1} - T_i^n}{\Delta t} - \underbrace{\boxed{\frac{\partial^2 T}{\partial t^2}\Big|_i \frac{\Delta t}{2}}}_{\text{Term-1}} + O(\Delta t^2)$$

$$= \alpha\frac{T_{i+1}^n + T_{i-1}^n - 2T_i}{\Delta x^2} + \underbrace{\boxed{\alpha\frac{\partial^4 T}{\partial x^4}\frac{\Delta x^2}{12}}}_{\text{Term-2}} + O(\Delta x^4)$$

$$\text{g.d.e} \longrightarrow \frac{\partial^2}{\partial x^2} \Rightarrow \frac{\partial^3 T}{\partial t\,\partial x^2} = \alpha\frac{\partial^4 T}{\partial x^4}.$$

$$\text{g.d.e} \longrightarrow \frac{\partial}{\partial t} \Rightarrow \frac{\partial^2 T}{\partial t^2} = \alpha\frac{\partial^3 T}{\partial t\,\partial x^2}$$

$$\frac{\partial^2 T}{\partial t^2} = \alpha\frac{\partial^3 T}{\partial t\,\partial x^2} = \alpha^2\frac{\partial^4 T}{\partial x^4}$$

Term ① - term ②.

$$\Rightarrow \alpha\frac{\partial^4 T}{\partial x^4}\left[\frac{\alpha\Delta t}{2} - \frac{\Delta x^2}{12}\right]$$

$\hookrightarrow O(\Delta t), O(\Delta x^2)$

First order in time,

Second order in space.

(32).

$$\alpha\frac{\partial^4 T}{\partial x^4}\left[\alpha\frac{\Delta t}{2} - \frac{\Delta x^2}{12}\right] = 0.$$

when $\alpha\frac{\Delta t}{2} - \frac{\Delta x^2}{12} = 0$

$$\Rightarrow \frac{\alpha\Delta t}{\Delta x^2} = \frac{1}{6}.$$

Thus $O(\Delta t^2), O(\Delta x^4)$ are order of errors in time & space respectively.

Is it consistent? Yes! as $\Delta x \to 0$, $\Delta t \to 0$

error $\to 0$.

Lecture 25: FTCS scheme

FTCS: $\frac{\partial T}{\partial t} = \alpha\frac{\partial^2 T}{\partial x^2}$.

$$\frac{T_i^{n+1} - T_i^n}{\Delta t} = \alpha\frac{T_{i+1}^n + T_{i-1}^n - 2T_i^n}{\Delta x^2}$$

Checking for stability: (whether numerical perturbations amplify while propagating or not).

$$T_i^{n+1} - T_i^n = r\left[T_{i+1}^n + T_{i-1}^n - 2T_i^n\right]$$

$$\Rightarrow T_i^{n+1} = (1-2r)\,T_i^n + rT_{i+1}^n + rT_{i-1}^n. \quad ①$$

→ analogy with FV. Discretization.

→ Explicit scheme. T_i^{n+1} described in terms of previous time step.

Say, we get an approximate soln. That soln must also satisfy above eq?

i.e.

$$T_i^{*n+1} = (1-2r)T_i^{*n} + rT_{i+1}^{*n} + rT_{i-1}^{*n} \quad ②.$$

① − ② ⇒

$$(T_i - T_i^*)^{n+1} = (1-2r)(T_i - T_i^*)^n +$$

$$r(T_{i+1} - T_{i+1}^*)^n + r(T_{i-1} - T_{i-1}^*)^n$$

Say $\epsilon = T - T^*$
(error)

Then

$$\mathcal{E}_i^{n+1} = (1-2r)\mathcal{E}_i^n + r\,\mathcal{E}_{i+1}^n + r\,\mathcal{E}_{i-1}^n$$

errors satisfy the same eq^n as discretization variable.

$\mathcal{E} = \mathcal{E}(x,t)$

Write this in terms of a Fourier series.

$$\epsilon(x,t) = \sum A Z_j^{(n)} e^{ikx} \quad j = \sqrt{-1}$$

What is the particular form of Z?

A convenient form is e^{at}, where $a = fn(j)$

why e^{at}? exponential fn helps evaluate the growth/decay behavior easily.

If $e^{at+\Delta t} < e^{at} \longrightarrow$ decay;
$e^{at+\Delta t} > e^{at} \rightarrow$ growth.

to assess whether there is exponential growth/decay.

w.r.t space, there is a periodicity.

Over length of each CV, repeatability observed.

∴ corresponding to wave number (k) (number of waves over a time period), corresponding wave length = cell length.

$e^{at} e^{jkx} \quad j = \sqrt{-1}$

$$e^{a(t+\Delta t)} \cdot e^{jkx}$$

$$= (1-2r)\, e^{at} e^{jkx} + r\, e^{at} e^{jk(x+\Delta x)}$$

$$+ r\, e^{at} e^{jk(x-\Delta x)}$$

$$A: \text{Amplification factor} = \frac{e^{a(t+\Delta t)}}{e^{at}}$$

$$A = (1-2r) + r\left(e^{jk\Delta x} + e^{jk(-\Delta x)}\right)$$

Check for stability. $|A| < 1$

If regardless of r, $|A| < 1$, then unconditionally stable.

If regardless of r, $|A| > 1$, then unconditionally unstable.

If for some values of r, $|A| < 1$, then conditionally stable.

$$A = (1-2r) + r\left(e^{j\theta} + e^{-j\theta}\right)$$

$$= (1-2r) + r\left(\cos\theta + j\cancel{\sin\theta} + \cos\theta - j\cancel{\sin\theta}\right)$$

$$= (1-2r) + 2r\cos\theta.$$

$$= 1 - 2r(1-\cos\theta)$$

$$= 1 - 4r\sin^2\left(\tfrac{\theta}{2}\right).$$

For stability. $|A| \le 1$.

$$\Rightarrow -1 \le A \le 1.$$

$$\boxed{-1 \le 1 - 4r\sin^2\left(\tfrac{\theta}{2}\right)} \le 1$$

LHS $\hookrightarrow$ $4r\sin^2\left(\frac{\theta}{2}\right) \le 2$

$$r \le \frac{1}{2\sin^2(\theta/2)}$$

↳ conservative upper limit ⇒ $r \le \frac{1}{2}$

Right hand limit:

$$1 - 4r\sin^2\left(\frac{\theta}{2}\right) \le +1$$

$$\Rightarrow 4r\sin^2\left(\frac{\theta}{2}\right) \ge 0 \rightarrow \text{always true!}$$

$$r = \frac{\alpha\,\Delta t}{\Delta x^2} = F_0 \le \frac{1}{2}$$

$$\boxed{\text{Lecture 26: CTCS scheme (Leap Frog Scheme) and Dufort-Frankel Scheme}}$$

CTCS Scheme (Leap frog scheme).

$$\frac{\partial T}{\partial t} = \alpha\frac{\partial^2 T}{\partial x^2}$$

$$\frac{T_i^{n+1} - T_i^{n-1}}{2\Delta t} = \alpha\,\frac{T_{i+1}^n + T_{i-1}^n - 2T_i^n}{\Delta x^2}$$

$$T_i^{n+1} - T_i^{n-1} = 2r\left[T_{i+1}^n + T_{i-1}^n - 2T_i^n\right] \quad -(1)$$

Approx sol^n:

$$T_i^{*\,n+1} - T_i^{*\,n-1} = 2r\left[T_{i+1}^{*\,n} + T_{i-1}^{*\,n} - 2T_i^{*\,n}\right] \quad -(2)$$

(1) − (2) ⟶

$$\varepsilon_i^{n+1} - \varepsilon_i^{n-1} = 2r\left[\varepsilon_{i+1}^n + \varepsilon_{i-1}^n - 2\varepsilon_i^n\right]$$

Sol of the form $e^{at}e^{jkx}$

$$e^{a(t+\Delta t)}e^{jkx} - e^{a(t-\Delta t)}e^{jkx} = 2r\left[e^{at}e^{jk(x+\Delta x)} + e^{at}e^{jk(x-\Delta x)} - 2e^{at}e^{jkx}\right]$$

Divide both sides by $e^{at}e^{jkx}$:

$$e^{a\Delta t} - e^{-a\Delta t} = 2r\left[e^{jk\Delta x} + e^{-jk\Delta x} - 2\right], \quad e^{a\Delta t} \rightarrow A$$

$$A - \frac{1}{A} = 2r\left[e^{j\theta} + e^{-j\theta} - 2\right]$$

$$A - \frac{1}{A} = 2r\left[2\cos\theta - 2\right]$$

$$A - \frac{1}{A} = 4r(\cos\theta - 1) = -8r\sin^2(\theta/2)$$

$$A^2 + 8r\sin^2(\theta/2)\,A - 1 = 0.$$

Quadratic eq^n in A.

Product of the roots has a magnitude of 1.

$$A = \frac{-8r\sin^2(\theta/2) \pm \sqrt{64r^2\sin^4(\theta/2) + 4}}{2}$$

$$= -4r\sin^2(\theta/2) \pm \sqrt{16r^2\sin^4(\theta/2) + 1}$$

Greater magnitude of A ⟶

$$\left|-4r\sin^2(\theta/2) - \sqrt{16r^2\sin^4(\theta/2)+1}\right| \rightarrow > 1$$

⇒ unconditionally unstable!

The greed for higher accuracy (CT ~ $O(\varepsilon^2)$, FT ~ $O(\varepsilon)$) led to us using CTCS over FTCS. But CTCS is unconditionally unstable while FTCS is conditionally stable.

Modification of this scheme will provide accuracy as well as stability. Such a modification → Dufort-Frankel Scheme.

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2}$$

$$\frac{T_i^{n+1} - T_i^{n-1}}{2\Delta t} = \alpha \left[\frac{T_{i+1}^n + T_{i-1}^n - 2T_i^n}{\Delta x^2} \right].$$

Make adhoc change:

$$T_i^n = \frac{T_i^{n-1} + T_i^{n+1}}{2}$$

(temperature at a grid pt at an instant = average temperature at that grid pt at previous time and the next time).

This adhoc change :: CTCS ⟶ Dufort-Frankel Scheme.

$$T_i^{n+1} - T_i^{n-1} = 2r\left[T_{i+1}^n + T_{i-1}^n - T_i^{n-1} - T_i^{n+1} \right]$$

$$\varepsilon_i^{n+1} - \varepsilon_i^{n-1} = 2r\left[\varepsilon_{i+1}^n + \varepsilon_{i-1}^n - \varepsilon_i^{n-1} - \varepsilon_i^{n+1} \right]$$

$$e^{a(t+\Delta t)} e^{jkx} - e^{a(t-\Delta t)} e^{jkx}$$

$$= 2r\left[e^{at} e^{jk(x+\Delta x)} + e^{at} e^{jk(x-\Delta x)} - e^{a(t-\Delta t)} e^{jkx} - e^{a(t+\Delta t)} e^{jkx} \right]$$

Divide by $e^{at} e^{jkx}$,

$$A - A^{-1} = 2r\left[e^{jk\Delta x} + e^{-jk\Delta x} - A - A^{-1} \right]$$

$$A_t = \frac{1}{A}$$

$$A(1+2r) - \frac{1}{A}(1-2r) = 2r(2\cos\theta).$$

$$(1+2r)A^2 - 4r\cos\theta\, A - (1-2r) = 0.$$

$$A = \frac{4r\cos\theta \pm \sqrt{16r^2\cos^2\theta + 4(1-4r^2)}}{2(1+2r)}.$$

$$= \frac{2r\cos\theta \pm \sqrt{4r^2\cos^2\theta + 1 - 4r^2}}{(1+2r)}$$

$$= \frac{2r\cos\theta \pm \sqrt{1-4r^2\sin^2\theta}}{(1+2r)}$$

Case 1: $2r < 1$.

Then $4r^2 < 1$.

$\therefore 4r^2\sin^2\theta < 1.$

$$A = \frac{2r\cos\theta \pm \sqrt{1-4r^2\sin^2\theta}}{(1+2r)}$$

Take (+) for more conservative estimate,

$$A = \left(2r\cos\theta + \sqrt{1-4r^2\sin^2\theta}\right) / (1+2r)$$

$$\leq \frac{1+2r\cos\theta}{1+2r} \leq 1.$$

Case 2: $2r > 1$

$$A = \frac{2r\cos\theta \pm j\sqrt{4r^2\sin^2\theta - 1}}{1+2r}$$

$$|A|^2 = \frac{4r^2\cos^2\theta + 4r^2\sin^2\theta - 1}{(1+2r)^2}$$

$$= \frac{4r^2-1}{(1+2r)^2} = \frac{(2r+1)(2r-1)}{(1+2r)\cancel{(1+2r)}} < 1.$$

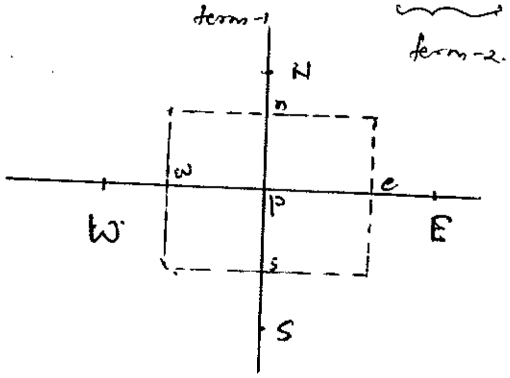
This scheme → Conditionally stable, but not consistent (H/W: prove)

↳ i.e. truncation error as $\Delta t \to 0$ & $\Delta x \to 0$ doesn't nullify itself.

$$\boxed{\text{Lecture 27: FV Discretization of 2D unsteady State Diffusion Type Problems:}}$$

Eg: heat conduction:

$$\underbrace{\rho C_p \frac{\partial T}{\partial t}}_{\text{term-1}} = \underbrace{\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)}_{\text{term-2}} + \underbrace{\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right)}_{\text{term-3}} + \underbrace{S}_{\text{term-4}}$$



Term-1 → $\displaystyle\int_s^n\int_w^e\int_t^{t+\Delta t} \rho C_p \frac{\partial T}{\partial t}\, dt\, dx\, dy$

$= \rho C_p (T_P - T_P^\circ)\Delta x \Delta y$ → can use piece wise const temperature profile (no derivative terms involved). Then $T = T_P$.

Term-2:

$$\int_t^{t+\Delta t}\int_s^n\int_w^e \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) dx\, dy\, dt$$

$\displaystyle\int_w^e \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) dx.$ Take piecewise-linear profile.

$$= k_e\frac{T_E - T_P}{\delta x_e} - k_w\frac{T_P - T_W}{\delta x_w}$$

Final integration form:

= Consider a fully implicit scheme.

$$\Rightarrow \left(k_e\frac{T_E - T_P}{\delta x_e} - k_w\frac{T_P - T_W}{\delta x_w}\right)\Delta t\,\Delta y.$$

Term 3:

$$\int_t^{t+\Delta t}\int_w^e\int_s^n \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) dy\, dx\, dt$$

$$\to \left(k_n\frac{T_N - T_P}{\delta y_n} - k_s\frac{T_P - T_S}{\delta y_s}\right)\Delta t\,\Delta x.$$

Term 4:

$$\int_t^{t+\Delta t}\int_s^n\int_w^e (S_c + S_p T_P)\, dx\, dy\, dt$$

$$= (S_c + S_p T_P)\,\Delta x\,\Delta y\,\Delta t$$

Assemble terms:

Divide all terms by Δt:

Then we get eqn of the form:

$$a_P T_P = a_E T_E + a_W T_W + a_S T_S + a_N T_N + a_P^\circ T_P^\circ + b$$

$$a_E = \frac{k_e}{\delta x_e}\Delta y,\quad a_W = \frac{k_w}{\delta x_w}\Delta y,$$

$$a_S = \frac{k_s}{\delta y_s}\Delta x,\quad a_N = \frac{k_n}{\delta y_n}\Delta x,\quad a_P^\circ = \frac{\rho C_p \Delta x \Delta y}{\Delta t}.$$

$a_P = a_E + a_W + a_S + a_N + \rho C_p \frac{\Delta x \Delta y}{\Delta t}$ $- S_p \Delta x \Delta y$

$b = S_c \Delta x \Delta y$.

of the form:

$$a_P T_P = \sum a_{nb} T_{nb} + b \quad (nb \to \text{neighbour})$$

The coeffs a_E, a_W, a_N, a_S etc physically represent thermal conductance. why?

Take $a_E = \frac{k_e \Delta y}{\delta x_e}$. Assuming unit length normal to the x-y plane, $(\Delta y \times 1)$ = area of face. So a_E is of the form $a_E = \frac{k_e A}{L}$, which is the formula for conductance, $\frac{L}{k_e A} \to$ resistance.

By setting Δt to very large number, this unsteady problem can be converted into a steady state problem as terms containg $(\Delta t)^{-1}$ becomes negligibly small.

(H/w): repeat same exercise for a fully explicit scheme.

- Solving system of algebraic eq^{ns}:-

① $x + y = 2$

$2x + 3y = 5$

$E_1 \times 2 - E_2$

$\Rightarrow -y = -1$

$\Rightarrow \underline{y = 1}$

E_1 : $x = 2 - y = 2 - 1 = \underline{1}$

using method of elimination.

$\underline{(x, y) = (1, 1)}$

② $x + y = 2$

$2x + 2y = 4$

~~③~~

lin. early dependent eq^{ns}.

No. of independent eqs. < no. of unknowns.

~~④~~ ~~⑤~~ ~~$x+y=0$~~ ~~$2x+2y=0$~~

$y = 2 - x$.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2-x \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

③ $x + y = 2$

$2x + 2y = 5$

$E_1 \times 2 \longrightarrow 2x + 2y = 4$

From $E_2 \to 2x + 2y = 5$

Equate them both $\Rightarrow 4 = 5$ ✗

$\longrightarrow$ no sol^n.

(parallel straight lines — no intersection pts).

④ $x + y = 0$

$2x + 3y = 0$

} If RHS = 0 $\to$ homogeneous system of eq^{n}s.

$x = 0, y = 0 \to$ trivial sol^{ns}.

no non-trivial sol^{ns}.

⑤ $x + y = 0$

$2x + 2y = 0$

trivial sol^n $x = 0, y = 0$

∞ non-trivial sol^{ns}.

$y = -x$.

what if number of eq^ns are large?

Use matrix forms.

$$① \rightarrow \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}}_{\text{Coeff matrix}} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 3 & 5 \end{array}\right] \rightarrow \text{Augmented matrix.}$$

Lecture 28: Solutions to linear algebraic equations (contd).

Identify rank of coeff matrix & augmented matrix. Why rank? B/c it gives an idea about the linear dependence/independence of equations in the system.

Rank of a matrix is r, if

(i) It has atleast one non-zero minor of order r.

(ii) all minors of order $> r$ vanishes (=0).

- For eg:1,

~~Take~~ $R_c = 2 \rightarrow$ coeff matrix

$R_A = 2 \rightarrow$ aug. matrix.

Observation: $R_c = R_A = 2 = n$ (no. of eq^ns = no. of unknowns).

- For eg: 2,

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 4 \end{array}\right]$$

$R_c = 1$

$R_A = 1$

Observation: $R_A = R_c = 1 < n$

- For eg: 3,

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 5 \end{array}\right]$$

$R_c = 1$

$R_A = 2$

Obs: $R_c \neq R_A$

For a system of non-homogeneous eq^ns:

→ $R_c = R_A = n \Rightarrow$ unique sol^n.

→ $R_c = R_A < n \Rightarrow$ infinitely large no. of solutions,

→ $R_c \neq R_A \Rightarrow$ no solution.

For homogeneous eq^ns, since RHS is all 0, there is no need for augmented matrix.

For system of homogeneous eq^n:

$\Delta \neq 0 \longrightarrow$ only trivial sol^n

$\Delta = 0 \longrightarrow$ infinitely large no. of solutions

It's important to see the nature of sol^n for system of algebraic eq^ns. Why? For eg, ~~B/c~~ for well-posed problems, uniqueness is an important criteria. But if the system of ~~n~~ non-homogeneous eq^ns has infinitely many solutions, the problem itself is ill defined & needs modifications.

Solution techniques for systems of linear algebraic equations:

→ Elimination

→ Iteration

→ Gradient search method.

Elimination method:

- $x_1 + x_2 = 2.$ (E_1)
  $2x_1 + 3x_2 = 5.$ (E_2)
  $2E_1 - E_2 \rightarrow$ Effort is to eliminate x_1

- $x_1 + x_2 + x_3 = 3$ (E_1)
  $2x_1 + 2x_2 + 3x_3 = 7$ (E_2)
  $3x_1 + x_2 + 2x_3 = 6.$ (E_3)

$2 \times E_1 - E_2 \Rightarrow x_3 = 1.$

$E_3 - 3E_1 \Rightarrow -2x_2 - x_3 = -3$

$-2x_2 - 8(1) = -3.$

$x_2 = 1$ .....

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 2 & 3 & 7 \\ 3 & 1 & 2 & 6 \end{array}\right] \rightarrow \left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 0 & 1 \\ \end{array}\right.$$

Reorder:

$x_1 + x_2 + x_3 = 3.$

$3x_1 + x_2 + 2x_3 = 6$

$2x_1 + 2x_2 + 3x_3 = 7$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -1 & -3 \\ 0 & 0 & 1 & 1 \end{array}\right]$$ Converted into upper triangular form.

Now, $x_3 = 1$

$x_2 = \frac{3 - x_3}{2} = 1$

$\Rightarrow x_1 = 3 - x_2 - x_3 = 1$

We can break this method into two parts:

part-I: forward elimination.

$\Rightarrow$ Convert C to upper Δ form.

part-II: Backwards substitution.

- $x_1 + x_2 + x_3 = 3.$
  $2x_1 + x_2 + 3x_3 = 6.$
  $3x_1 + 4x_2 + 2x_3 = 9.$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 1 & 3 & 6 \\ 3 & 4 & 2 & 9 \end{array}\right]$$

$E_2 - 2E_1 \rightarrow -x_2 + x_3 = 0$

$E_3 - 3E_1 \rightarrow x_2 - x_3 = 0$

linear dependency involved!

$(E_3 = 5E_1 - E_2).$

$\boxed{\text{Lecture 29: Elimination Methods.}}$ →

Ex $10x_1 + x_2 + x_3 = 12$. — E_1

$x_1 + 10x_2 + x_3 = 12$. — E_2

$x_1 + x_2 + 10x_3 = 12$. — E_3

Step-1 : Row 1 as the pivotal row

$E_2 \rightarrow E_2 - \frac{1}{10} \times E_1$

$E_3 \rightarrow E_3 - \frac{1}{10} E_1$

$$\left[\begin{array}{ccc|c} 10 & 1 & 1 & 12 \\ 1 & 10 & 1 & 12 \\ 1 & 1 & 10 & 12 \end{array}\right]$$

$$\hookrightarrow \left[\begin{array}{ccc|c} 10 & 1 & 1 & 12 \\ 0 & 9.9 & 0.9 & 10.8 \\ 0 & 0.9 & 9.9 & 10.8 \end{array}\right]$$

Step-2 :

$E_3 \rightarrow E_3 - \frac{0.9}{9.9} E_2$

$$\Rightarrow \left[\begin{array}{ccc|c} 10 & 1 & 1 & 12 \\ 0 & 9.9 & 0.9 & 10.8 \\ 0 & 0 & 9.818 & 9.818 \end{array}\right]$$

2 steps required for 3 equations.

So in general, for n equations, it would require n-1 steps.

→ Upto this forward elimination.

Next: Backward substitution:

From E_3; $x_3 = \frac{9.818}{9.818} = 1$.

From E_2, $9.9x_2 = 10.8 - \ldots$

& finally get x_1 from E_1

Generalization of Gaussian Elimination:

$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1$ — E_1

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$. — E_2

$\vdots$

$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$ — E_n.

Row-1 pivotal

$E_2 \rightarrow E_2 - \frac{a_{21}}{a_{11}} E_1$

$E_3 \rightarrow E_3 - \frac{a_{31}}{a_{11}} E_1$

$\vdots$

$E_i \rightarrow E_i - \frac{a_{i1}}{a_{11}} E_1$

for step-1:

$a_{ij} \rightarrow a_{ij} - \frac{a_{i1}}{a_{11}} \times a_{1j}$

for step 2:

$a_{ij} \rightarrow a_{ij} - \frac{a_{i2}}{a_{22}} \times a_{2j}$

Generalize for step no. k. (k for pivot)

$a_{ij} \rightarrow a_{ij} - \frac{a_{ik}}{a_{kk}} \times a_{kj}$

Lecture 30: Gaussian Elimination and LU Decomposition methods.

Formalize forward elimination:

for $k=1$ to $n-1$

for $j=k+1$ to $n+1$ ∵ augmented matrix considered.

for $i=k+1$ to n

a_{ik} is already ready if 3rd loop comes in 2nd loop's position. So

```
for k=1 to n-1
    for i = k+1 to n
        R = a_ik/a_kk;
        for j = k+1 to n+1
            a_ij = a_ij - R * a_kj;
        end
    end
end
```

Formalize backward substitution:

$$x_n = \frac{b_n}{a_{nn}} \rightarrow a_{nn}x_n = b_n.$$

$$a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1}.$$

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}}$$

$$a_{n-2,n-2}x_{n-2} + a_{n-2,n-1}x_{n-1} + a_{n-2,n}x_n = b_{n-2}.$$

$$x_{n-2} = \frac{1}{a_{n-2,n-2}}\left[b_{n-2} - a_{n-2,n-1}x_{n-1} - a_{n-2,n}x_n\right]$$

$$x_{n-i} = \frac{1}{a_{n-i,n-i}}\left[b_{n-i} - \left\{\sum_{j=0}^{j=i-1} a_{n-i,n-j}x_{n-j}\right\}\right]$$

- Formal algo for backward substitution:

$x_n = \frac{b_n}{a_{nn}}$;

```
for i=1 to n-1
    sum=0
    for j = 0 to i-1
        sum = sum + a_{n-i,n-j} x_{n-j}
    end
    x_{n-i} = (b_{n-i} - sum)/a_{n-i,n-i};
end
```

Assessment of number of computations:-

Forward elimination $\rightarrow O(n^3) \rightarrow$ (n for each loop).

Backward substitution $\rightarrow O(n^2)$

Rate determining step: forward elimination

∴ computational complexity of the algo $\rightarrow O(n^3)$.

LU decomposition evolved to reduce the complexity from $O(n^3)$.

L-U decomposition technique:

Factorize $A = LU$ → U: Upper triangular matrix; L: Lower triangular matrix.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ \vdots & & & \ddots \; 0 \\ l_{n1} & l_{n2} & l_{n3} \cdots & l_{nn} \end{bmatrix} \times$$

$$\begin{bmatrix} 1 & u_{12} & u_{13} & \cdots & u_{1n} \\ & 1 & u_{23} & u_{24} \cdots & u_{2n} \\ & & 1 & & \ddots \\ & 0 & & 1 & \ddots \\ & & & & 1 \end{bmatrix}$$

Crout's method:

$\cancel{a_{11} = l_{11}}$ $\quad l_{11} = a_{11}$

$\cancel{a_{21} = l_{21}}$ $\quad l_{21} = a_{21}$

$\quad l_{31} = a_{31}$

$\vdots$

$\cancel{a_{n1} = l_{n1}}$ $\quad l_{n1} = a_{n1}$

$Ax = b$

$L\boxed{Ux} = b$

$LZ = b$ ↳ triangular matrix.

$Z = L^{-1} b \to O(n^2)$ (Forward elim).

$UX = Z$

$X = U^{-1} Z \to O(n^2)$ (Back sub.)

Even though complexity may seem $O(n^2)$, Calculation required to factorize has complexity $O(n^3)$. Thus it is no better than Gaussian elimination.

Then why use L-U decomposition method at all?

> Lecture 31: Illustrative example of elimination method.

$Ax = b$

A ↓ Fixed, b ↓ Variable.

→ Special case: when A is symmetric and +ve definite $\Rightarrow U = L^T$ (Cholesky's L-U factorization).

Here number of calculations become half (But doesn't become $O(n^3) \to O(n^2)$; remains $O(n^3)$ itself)

Symmetric: $A = A^T$

+ve definite: $v^T A v > 0$.

Ex. $\varepsilon x_1 + x_2 = 1 \quad (E_1)$

$x_1 + x_2 = 2 \quad (E_2)$ $\quad \varepsilon \to$ a small no.

Solve by Gaussian elimination

Forward elimination:

$(E_2) - (E_1) \times \frac{1}{\varepsilon} \Rightarrow \left(1 - \frac{1}{\varepsilon}\right) x_2 = \left(2 - \frac{1}{\varepsilon}\right)$

$$x_2 = \frac{2 - \frac{1}{\varepsilon}}{1 - \frac{1}{\varepsilon}} \approx 1$$

$$\left.x_1 = \frac{1 - x_2}{\varepsilon} = 0.\right\} \times$$

∵ These sol^ns doesn't satisfy the equations, They aren't the correct solutions.

Test whether reordering eq^n help or not:

Reorder the eq^ns:

$x_1 + x_2 = 2 \;—\; (E_1)$

$\varepsilon x_1 + x_2 = 1 \;—\; (E_2)$.

Forward elimination:

$(E_2) - \varepsilon(E_1) \Rightarrow (1 - \varepsilon) x_2 = 1 - 2\varepsilon$

Backward substitution: $x_2 = \frac{1 - 2\varepsilon}{1 - \varepsilon} \approx 1 \to x_1 = 2 - x_2 = 1 \} ✓$

Reordering works!

Origin of the problem: Pivotal coefficient/diagonal entry being small

Reordering the equation makes diagonal entries not small.

What is the issue with diagonals being small?

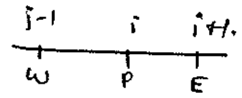
- During forward elimination, division by diagonal element is required. Division by small number ϵ makes the resulting number very large. This blowing off can overweigh any other coeff-number in the equation making them insignificant & making it difficult to extract out the difference b/w coeff numbers. → leading to errors.

Reorder equations to reassign pivotal rows
→ pivotization.

- Gaussian elimination need not work well in cases where there are issues in diagonal dominance.

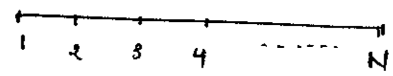
Lecture 32: Tri-diagonal Matrix Algorithm (TDMA).

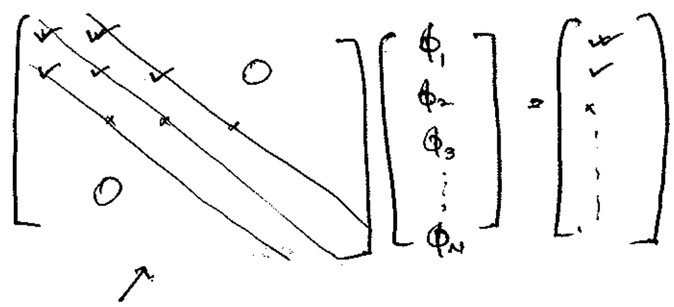
- has complexity: $O(n)$
- also known as Thomas algorithm.



$$a_P \phi_P = a_E \phi_E + a_W \phi_W + b$$
$$= \sum a_{nb} \phi_{nb} + b.$$

$$a_i \phi_i = b_i \phi_{i+1} + c_i \phi_{i-1} + d_i \quad \text{---} (*)$$





Tridiagonal matrix.

Instead of using 2 indices for storage of an element, use 1 index for storing entries related to three diagonals together. → 2D array storage system to linear storage system.

~~ME 497~~

$$a_1\phi_1 = b_1\phi_2 + d_1 \quad (c_1 = 0)$$

$$\phi_1 = \frac{b_1\phi_2 + d_1}{a_1} - f_1(\phi_2)$$

$$a_2\phi_2 = b_2\phi_3 + c_2 \overset{f(\phi_2)}{\cancel{\phi_1}} + d_2.$$

$$\phi_2 = f_2(\phi_3).$$

$$\vdots$$

$$\phi_n = f_n(\cancel{\phi_{n+1}})$$ const (∵ there is no ϕ_{n+1}, its just a const)

In general,

$$\phi_i = P_i\phi_{i+1} + Q_i \quad \text{— (3*)}$$ → in the form of a linear function.

Immediate step before:

$$\phi_{i-1} = P_{i-1}\phi_i + Q_{i-1} \quad (**)$$

Sub this recurrence formula (**) in (*):

$$a_i\phi_i = b_i\phi_{i+1} + c_i\left[P_{i-1}\phi_i + Q_{i-1}\right] + d_i$$

$$\Rightarrow (a_i - c_iP_{i-1})\phi_i = b_i\phi_{i+1} + d_i + c_iQ_{i-1}.$$

$$\phi_i = \frac{b_i}{a_i - c_iP_{i-1}}\phi_{i+1} + \frac{d_i}{a_i - c_iP_{i-1}} + \frac{c_i}{a_i - c_iP_{i-1}}Q_{i-1} \quad \text{— (4*)}$$

Compare (3*) and (4*) :-

$$\Rightarrow P_i = \frac{b_i}{a_i - c_iP_{i-1}}$$

$$Q_i = \frac{d_i + c_iQ_{i-1}}{a_i - c_iP_{i-1}}.$$

where $c_1 = 0$ (∵ there is not ϕ_0) &

$b_N = 0$ (∵ there is no ϕ_{n+1})

$$P_1 = \frac{b_1}{a_1}, \quad Q_1 = \frac{d_1}{a_1}.$$

$\phi_n = Q_N$ (∵ $\phi_n = P_n\phi_{n+1} + Q_n$

∵ there is no ϕ_{n+1},

$\phi_n = Q_n$).

Above ↑ stuff → like forward elimination.

$\phi_i = P_iQ_{i+1} + Q_i$ → use to backward sub.

Summary of TDMA:

- Input a_i, b_i, c_i, d_i

$$P_1 = \frac{b_1}{a_1}, \quad Q_1 = \frac{d_1}{a_1}$$

for $i = 2, N$ (forward elimination)

$$P_i = \frac{b_i}{a_i - c_iP_{i-1}}; \quad Q_i = \frac{d_i + c_iQ_{i-1}}{a_i - c_iP_{i-1}}$$

end

$\phi_n = Q_N$.

for $i = n-1, 1, i--$ (Backward substitution)

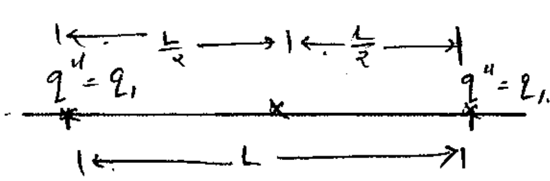
$$\phi_i = P_i\phi_{i+1} + Q_i$$

end

O(N) complexity!

Will TDMA always work?

Take an example:- $S = 0$.



Find steady state T distribution.

Steps:

→ FV discretization / FD discretization.
→ sol^n of eq^n by TDMA

Governing equation:

$$\frac{d}{dx}\left(k\frac{dT}{dx}\right) = 0.$$

If k const $\longrightarrow \frac{d^2T}{dx^2} = 0.$

FD discretization: Applied to grid-point (2)

$$\frac{T_3 + T_1 - 2T_2}{(L/2)^2} = 0$$

$$\Rightarrow T_2 = \frac{T_1 + T_3}{2} \longrightarrow 2T_2 = T_1 + T_3.$$

For B/c at ①,

$$q'' = -k\frac{dT}{dx}\Big|_1 = k\frac{(T_1 - T_2)}{(L/2)}$$

$$T_1 = T_2 + \frac{q''L}{2k}.$$

$$\rightarrow T_1 = T_2 + a$$

BC at ③,

$T_3 = T_2 - a$ in a similar way.

{ Physically obvious as heat is transferred from higher to lower temperature. $T_3 < T_2$ }

$a_1 T_1 = b_1 T_2 + d_1$

$a_1 = 1,\ b_1 = 1,\ d_1 = \alpha.$

$a_2 T_2 = b_2 T_3 + c_2 T_1 + d_2.$

$a_2 = 2,\ b_2 = 1,\ c_2 = 1,\ d_2 = 0.$

$a_3 T_3 = c_3 T_1 + d_3.$

$a_3 = 1,\ c_3 = 1,\ d_3 = -\alpha$

$$P_1 = \frac{b_1}{a_1} = 1, \quad Q_1 = \frac{d_1}{a_1} = \alpha.$$

$$P_2 = \frac{b_2}{a_2 - c_2 P_1} = \frac{1}{2 - 1\times 1} = \underline{\underline{1}} \qquad Q_2 = \frac{d_2 + c_2 Q_1}{a_2 - c_2 P_1} = \frac{0 + 1\times\alpha}{1} = \underline{\underline{\alpha}}$$

$$P_3 = \frac{b_3}{a_3 - c_3 P_2} = \frac{0}{1 - 1\times 1} = \frac{0}{0} \rightarrow \text{indeterminate.}$$

TDMA breaks down in this case.

why? Coeff matrix is singular (det = 0).

In $Ax = b$

$x = A^{-1}b,$

inverse of A A^{-1}

$$= \frac{\text{adj}(A)}{\det(A)}.$$

If $\det(A) = 0 \rightarrow A^{-1}$ has a problem.

ill posed BVP.

Why? 2 flux conditions doesn't give additional information regarding the system.

~~Ill posed~~ If ill posedness & physical unreality of the problem isn't detected, the mathematics of the problem will naturally reveal it.

Lecture 33: Elimination Methods: Error Analysis

Prob Given for L-U factorization (A=LU) by Crout's method, following steps are to be executed.

$$l_{i1} = a_{i1} \quad \text{for } i \in [1, n]$$

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} U_{kj} \quad \text{where}$$

$$U_{1j} = \frac{a_{1j}}{l_{11}} \quad \text{for } j \in [2, n]$$

$$U_{ij} = \frac{1}{l_{ii}} \left[a_{ij} - \sum_{k=1}^{i-1} l_{ik} U_{kj}\right] \text{ for } j \in [i+1, n]$$

Estimate the operational count. If $A_{N \times N}$.

Numerical error → combined effect of errors intrinsic to method and errors due to the machine on which the algo is performed.

— Errors intrinsic to elimination method:-
{Terminology}

* norm of a vector:

say, we have a vector x,

$$x = \{x_1, x_2, \dots, x_n\}$$

element of $x \to x_i$

$$\|x\|_p = \left[\sum_i |x_i|^p\right]^{1/p}$$

Ex $x = \{1, -2, 3, -4\}$

$$\|x\|_1 = |1| + |-2| + |3| + |-4|$$
$$= \underline{10}$$

$$\|x\|_2 = (1^2 + 2^2 + 3^2 + 4^2)^{1/2} = \sqrt{30}$$

length of a vector.

$$\|x\|_\infty = \max |x_i|$$
$$= \underline{4}$$

Norm of a matrix:

$\|A\|$?

Introduce a vector & find $\|Ax\|$.

$$\|A\| \longrightarrow \frac{\|Ax\|}{\|x\|} \quad \text{s.t. } \|x\| = 1.$$

Prob: $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Find $\|A\|_2$.

$$Ax = b \;; \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{s.t. } \|x\|_2 = 1$$

$$Ax = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 \end{bmatrix}$$

$$\|Ax\|_2 = \sqrt{(x_1 + x_2)^2 + x_1^2}$$

subject to the constraint:

$$\|x\|_2 = 1 \Rightarrow \sqrt{x_1^2 + x_2^2} = 1 \quad \text{— ①}$$

$$\|Ax\|_2 = \sqrt{2x_1^2 + x_2^2 + 2x_1x_2}$$

Using ①, $\|Ax\|_2$ can be written in terms of either x_1 or x_2 → unambiguously.

B/c of that we say, $\|A\|_2 = \max \frac{\|Ax\|_2}{\|x\|_2}$

Effectively find out

$$\max\left(2x_1^2 + x_2^2 + 2x_1x_2\right), \text{ given}$$

$$x_2 = \sqrt{1 - x_1^2}$$

$$y = 2x_1^2 + 1 - x_1^2 + 2x_1\sqrt{1-x_1^2}.$$

$$y = x_1^2 + 1 + 2x_1\sqrt{1-x_1^2}.$$

for max y, $\dfrac{dy}{dx} = 0$.

$$\rightarrow \quad 2x_1 + 2\left[\sqrt{1-x_1^2} + \frac{x_1(-2x_1)}{2\sqrt{1-x_1^2}}\right] = 0.$$

$$\Rightarrow \quad x_1\sqrt{1-x_1^2} + \left[1 - x_1^2 - x_1^2\right] = 0.$$

$$\Rightarrow \quad x_1\sqrt{1-x_1^2} = 2x_1^2 - 1$$

$$x_1^2(1-x_1^2) = 4x_1^4 - 4x_1^2 + 1.$$

$$\Rightarrow \quad 5x_1^4 - 5x_1^2 + 1 = 0.$$

$$x_1^2 = \frac{5 \pm \sqrt{5}}{2\times 5} = \frac{1}{2} \pm \frac{1}{2\sqrt{5}}$$

$$x_1^2 = \frac{1}{2} \pm \frac{\sqrt{5}}{10}$$

Find condition for maxima, out of the two possible roots.

↳ That gives $x_1^2 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow$ elements of Ax.

Lecture 34: Elimination method: Error Analysis (Contd).

$$y = 2x_1^2 + 1 - x_1^2 + 2x_1\sqrt{1-x_1^2}.$$

$$y = x_1^2 + 1 + 2x_1\sqrt{1-x_1^2}.$$

For max y, $\dfrac{dy}{dx_1} = 0$

$$\Rightarrow \cancel{2}x_1 + \cancel{2}\sqrt{1-x_1^2} + \cancel{2}x_1 \times \frac{1(-\cancel{2}x_1)}{\cancel{2}\sqrt{1-x_1^2}} = 0$$

$$\Rightarrow x_1\sqrt{1-x_1^2} + 1 - x_1^2 - x_1^2 = 0$$

$$\Rightarrow x_1\sqrt{1-x_1^2} = 2x_1^2 - 1.$$

$$\cancel{x_1^2(1-x_1^2) = 4x_1}$$

Norms with special meaning for matrices: 1-norm & ∞-norm.

- 1-norm $\|A\|_1 \longrightarrow$ Column sum norm.

$$\max_j \left\{ \sum_i |A_{ij}| \right\}$$

- ∞-norm $\|A\|_\infty \longrightarrow$ Maximum row sum norm.

$$\max_i \sum_j |A_{ij}|.$$

$$A = \begin{bmatrix} 1 & 4 & 6 \\ -5 & -2 & 1 \\ -8 & -1 & 3 \end{bmatrix}$$

$\|A\|_1 \longrightarrow \max(14, 7, 10) = 14$

$\|A\|_\infty \longrightarrow \max(11, 8, 12) = 12$

Some important properties of matrix norms:

1. $\|kA\| = |k|\,\|A\|$
2. $\|A+B\| \le \|A\| + \|B\|$
3. $\|AB\| \le \|A\|\cdot\|B\| \rightarrow \|AB\| = \max \dfrac{\|ABx\|}{\|x\|}$

$$= \max \frac{\|ABx\|}{\|Bx\|} \cdot \max \frac{\|Bx\|}{\|x\|}$$

$$= \max \frac{\|Ay\|}{\|y\|} \cdot \max \|Bx\| / \|x\|$$

$$\|AB\| = \max \frac{\|Ay\|}{\|y\|} \max \frac{\|Bx\|}{\|x\|} \le \|A\| \cdot \|B\|$$

Error analysis of elimination methods:

Consider

$Ax = b$

Let x_{approx} be the approximate numerical sol^n

$$A\underbrace{(x - x_{approx})}_{e\ (\text{error})} = \underbrace{b - Ax_{approx}}_{r\ (\text{residue})}$$

In error analysis, we look for an upper bound of error without actually knowing the exact solution. It involves estimation, not exact quantification.

Say we have $e_1 \sim 10^{-10}$ and $e_2 \sim 10^{-5}$, in which case error is more / less? We cannot really tell as it depends on the x itself.

What we require is relative error, not absolute error.

Relative error indicator: $\frac{\|e\|}{\|x\|}$

$\|Ax\| \le \|A\| \cdot \|x\|$

$\|b\| \le \|A\| \cdot \|x\|$

$\|x\| \ge \frac{\|b\|}{\|A\|}$

$x = A^{-1}b$

$\|x\| = \|A^{-1}b\| \le \|A^{-1}\| \cdot \|b\|$

If you have $Ae = r$,

$\|e\| \ge \frac{\|r\|}{\|A\|}$ and

$\|e\| \le \|A^{-1}\| \cdot \|r\|$

Bounds:

$$\frac{\|e\|_{min}}{\|x\|_{max}} \le \frac{\|e\|}{\|x\|} \le \frac{\|e\|_{max}}{\|x\|_{min}}$$

$$\frac{\|r\|}{\|A\| \cdot \|A^{-1}\| \cdot \|b\|} \le \frac{\|e\|}{\|x\|} \le \frac{\|A^{-1}\| \|r\| \|A\|}{\|b\|}$$

Focus on upper bound (for conservative approach):

$$\frac{\|e\|}{\|x\|} \le \boxed{\|A^{-1}\| \|A\|} \cdot \frac{\|r\|}{\|b\|}$$

Conclusion: Even with small residual, the relative error may be large if $\|A^{-1}\| \cdot \|A\|$ is large. Therefore the largeness of $\|A^{-1}\| \|A\|$ determines the condition for accuracy of the system of equations one is solving.

Condition number: $C(A) = \|A\| \cdot \|A^{-1}\|$

Large $C(A) \Rightarrow$ even a small $\frac{\|r\|}{\|b\|}$ can lead to large $\frac{\|e\|}{\|x\|}$.

$\|AA^{-1}\| \le \|A\| \cdot \|A^{-1}\|$.

$1 \le \|A\| \cdot \|A^{-1}\|$.

$\Rightarrow C(A) \ge 1$

Closer to 1, is better.

$C(A) \rightarrow$ very critical parameter for estimating error.

Lecture 35. Iteration Methods.

Example problem:

Given, $A = \begin{bmatrix} 2 & 1 \\ 2 & 1.01 \end{bmatrix}$.

Find $C(A)$.

$C(A) = \|A\| \cdot \|A^{-1}\|$

$$A^{-1} = \begin{bmatrix} 1.01 & -1 \\ -2 & 2 \end{bmatrix} \times \frac{1}{(2.02-2)}$$

$$= \frac{1}{0.02}\begin{bmatrix} 1.01 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 50.05 & -50 \\ -100 & 100 \end{bmatrix}$$

$\|A\|_\infty = 3.01$

$\|A^{-1}\|_\infty = 200$

$C(A) = \|A\|_\infty \|A^{-1}\|_\infty$

$= 602 \rightarrow$ quite large

$\Rightarrow$ ill conditioned system.

Source of this largeness: ~~large~~ smallness of the determinant

~~Concl~~ Inferences: Smaller the determinant, greater the chances of ~~ge~~ larger Condition number & ill-condition of the system.

Iteration methods:

Basic philosophy: start with an initial guess for solution & iterate on it till you get a final solution that converges.

Say, we have equations:

$$5x_1 + x_2 = 6 \quad \text{---} (E_1)$$

$$x_1 + 5x_2 = 6 \quad \text{---} (E_2)$$

$$x_1 = \frac{6 - x_2}{5} \qquad x_2 = \frac{6 - x_1}{5}$$

Make an initial guess:

$$x_1^{(0)} = 0 \;, \quad x_2^{(0)} = 0.$$

Now try to update on this initial guess:

Make an iterative formula out of the given system:

$$x_1^{(k+1)} = \frac{6 - x_2^{(k)}}{5} \;, \qquad x_2^{(k+1)} = \frac{6 - x_1^{(k)}}{5}$$

$\Rightarrow$ If we write iteration formula in this manner, it is called Jacobi's method / Jacobi's iteration scheme.

S-1 $\quad x_1^{(1)} = \frac{6-0}{5} = \frac{6}{5} \qquad x_2^{(1)} = \frac{6-0}{5} = \frac{6}{5}$

S-2 $\quad x_1^{(2)} = \frac{6 - x_2^{(1)}}{5} = \frac{6-(6/5)}{5} = \frac{24}{25}$

$\quad x_2^{(2)} = \frac{6 - x_1^{(1)}}{5} = \frac{6-(6/5)}{5} = \frac{24}{25}$.

S-3

S-4

S-5

⋮

$\hookrightarrow x_1 \longrightarrow 1 \;, \quad x_2 \longrightarrow 1$ convergence

Convergence $\Rightarrow$ result b/w the current & previous steps doesn't differ substantially. (within some tolerance)

To update iterations with a faster rate, use

$$x_1^{(k+1)} = \frac{6 - x_2^{k}}{5}$$

$$x_2^{(k+1)} = \frac{6 - x_1^{(k+1)}}{5}$$

(instead of x_1^k, we use more updated version).

If you do that, then this becomes Gauss~~sedel~~-Siedel method.

$$x_1^{(1)} = \frac{6 - x_2^{(0)}}{5} = 6/5.$$

$$x_2^{(1)} = \frac{6 - x_1^{(1)}}{5} = \frac{6 - (6/5)}{5} = \frac{24}{25}$$

$$x_1^{(2)} = \frac{6 - x_2^{(1)}}{5} = \frac{6 - 24/25}{5} = \frac{126}{125}$$

$$x_2^{(2)} = \frac{6 - x_1^{(2)}}{5} = \frac{6 - 126/125}{5} = \frac{624}{625}$$

Here within two steps, ~~they~~ sol^n is converged very fast.

- Where is the guarantee that the scheme will converge or not?

Lecture 36: Generalized Analysis of Iteration Method

(41)

Generalized analysis of the iterative methods:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n = b_n$$

$$[A][x] = [b]$$

$$x_1^{(k+1)} = \frac{b_1 - (a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \cdots + a_{1n}x_n^{(k)})}{a_{11}}$$

Jacobi's method:

$$x_2^{(k+1)} = \frac{b_2 - (a_{21}x_1^{(k)} + a_{23}x_3^{(k)} + \cdots + a_{2n}x_n^{(k)})}{a_{12}}$$

Gauss-Sredel:

$$x_2^{(k+1)} = \frac{b_2 - (a_{21}x_1^{(k+1)} + a_{23}x_3^{(k)} + \cdots + a_{2n}x_n^{(k)})}{a_{22}}$$

$$[A][x] = [b]$$
$$\downarrow$$
$$\underbrace{[L]}_{\text{lower }\Delta\text{lar}} + \underbrace{[D]}_{\text{diagonal}} + \underbrace{[U]}_{\text{upper }\Delta\text{lar}}.$$

{Don't confuse with LU decomposition}

$$L = \begin{bmatrix} a_{21} & & & & 0 \\ a_{31} & a_{32} & & & \\ a_{41} & a_{42} & a_{43} & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & & a_{nn-1} \end{bmatrix} \qquad D = \begin{bmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & a_{33} & \\ 0 & & & \ddots \\ & & & & a_{nn} \end{bmatrix}$$

$$U = \begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ & a_{23} & a_{24} \cdots & a_{2n} \\ & & a_{34}\ a_{35} \cdots & a_{3n} \\ 0 & & \ddots & \vdots \\ & & & a_{n-1,n} \end{bmatrix}$$

Jacobi: $x_2^{(k+1)} = \dfrac{b_2 - (\overset{L}{a_{21} x_1^{(k)}} + \overset{U}{a_{23} x_3^{(k)} + \cdots + a_{2n} x_n^{(k)}})}{a_{22}}$

Gauss-Siedel:

$$x_2^{(k+1)} = \frac{b_2 - (\overset{L}{a_{21} x_1^{(k+1)}} + \overset{U}{a_{23} x_2^{(k)} + \cdots + a_{2n} x_n^{(k)}})}{a_{22}}$$

In place of $Ax = b$,

$$(L+D+U)x = b.$$

Jacobi: $Dx^{k+1} + (L+U)x^k = b.$

$$x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b.$$

$$x^{(k+1)} = Mx^{(k)} + C,$$

where $M = -D^{-1}(L+U)$

$C = D^{-1}b$

Gauss-Siedel:

$$DX^{(k+1)} + LX^{(k+1)} + UX^{(k)} = b.$$

$$(L+D)X^{(k+1)} = -UX^{(k)} + b$$

$$X^{(k+1)} = -(L+D)^{-1}UX^{(k)} + (L+D)^{-1}b.$$

$$X^{(k+1)} = MX^{(k)} + C$$

$$M = -(L+D)^{-1}U$$

$$C = (L+D)^{-1}b$$

$$X^{(k+1)} = MX^{(k)} + C$$

Say, X^* is actual sol^n

$$X^* = MX^* + C.$$

$$X^{(k+1)} - X^* = M(X^k - X^*)$$

error in the $(k+1)^{th}$ step: $e^{(k+1)}$.

$$e^{(k+1)} = Me^{(k)}$$

$$e^{(1)} = Me^{(0)}$$

$$e^{(2)} = Me^{(1)} = M^2 e^{(0)}$$

$\vdots$

$$e^{(k)} = M^k e^{(0)}$$

$\dfrac{\|e^k\|}{\|e^0\|} < 1 \rightarrow$ requirement for convergence.

$\Rightarrow \|M^k\| < 1$

Not easy for raw computation of M^k.
Use eigenvalues & eigenvectors of M for ease of representation.

- Let λ_i & v_i be correspondingly the eigenvalues & eigenvectors of M.

$$e^0 = a_1 v_1 + a_2 v_2 + a_3 v_3 + \cdots + a_n v_n.$$

(eigenvalues are arranged in a way that $|\lambda_1| > |\lambda_2| > |\lambda_3| \cdots > |\lambda_n|$)

$$Me^0 = a_1 Mv_1 + a_2 Mv_2 + a_3 Mv_3 + \cdots + a_n Mv_n$$

$$Mv_1 = \lambda_1 v_1;\ Mv_2 = \lambda_2 v_2 \cdots;\ Mv_n = \lambda_n v_n.$$

$$M^2 e^0 = a_1 \lambda_1 M v_1 + a_2 \lambda_2 M v_2 + \cdots + a_n \lambda_n M v_n$$
$$= a_1 \lambda_1^2 v_1 + a_2 \lambda_2^2 v_2 + \cdots + a_n \lambda_n^2 v_n$$
$$\vdots$$
$$e^k = M^k e^0 = a_1 \lambda^k v_1 + a_2 \lambda_2^k v_2 + \cdots + a_n \lambda_n^k v_n.$$

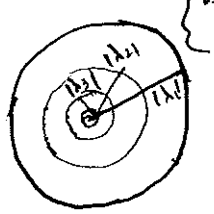
Compare leading order term of $M^k e^0$ with that of e^0.

$$\Rightarrow |\lambda_1|^k < 1 \longrightarrow \frac{\|e^k\|}{\|e^0\|}$$
$$= \frac{\| a_1 \lambda^k v_1 + \text{LOT} \|}{\| a_1 \lambda v_1 + \text{LOT} \|}$$
$$\sim |\lambda|^k.$$

$|\lambda_{max}| \rightarrow$ spectral radius of convergence.

{ In certain physical problems, eigenvalues represents the natural frequencies of the system. Something to do with frequency → "spectral" radius }

For convergence



{ Given $|\lambda_1| > |\lambda_2| > |\lambda_3| \cdots$
$|\lambda_1| = |\lambda_{max}|$ }

& radius of the biggest circle should be less than 1.

- Sufficient condition for convergence:

(If that condition is satisfied, you'll definitely satisfy convergence. But you may also have convergence without satisfying that condition).

- Rate of convergence:

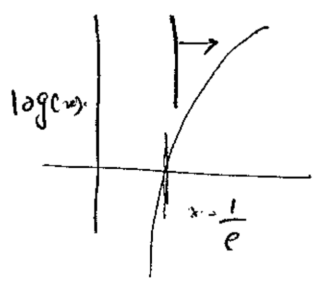
Requirement: no. of iterations to converge.

Say we require 'm' decimal accuracy.

Then, $\dfrac{\|e^k\|}{\|e^0\|} < 10^{-m}$.

$$\rho^k < 10^{-m}$$
$$k \log_{10} \rho < -m$$
$$m < k \log_{10}\left(\tfrac{1}{\rho}\right)$$
$$k > \frac{m}{\boxed{\log_{10}\left(\frac{1}{\rho}\right)}} \longrightarrow R = \text{rate of convergence.}$$

Lecture 37: Further discussion on Iterative Methods.



- Smaller spectral radius ⇒ better rate of convergence.

$$Mv = \lambda v$$
$$\|Mv\| \leq \|M\| \cdot \|v\|$$
$$\downarrow$$
$$|\lambda| \|v\| \leq \|M\| \cdot \|v\|$$
$$\Rightarrow |\lambda| \leq \|M\| \Rightarrow \rho \leq \|M\|.$$

spectral radius upper bound.

Which norm to choose?

An estimate of $\rho \rightarrow \max\left[\|M\|_1, \|M\|_\infty\right]$

• Sufficient condition for convergence:

$\max(\|M\|_1, \|M\|_\infty) \leq 1$

Jacobi's method:

$$M = -D^{-1}(L+U)$$

$$\underbrace{\begin{bmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & a_{33} \ddots & \\ 0 & & & a_{nn} \end{bmatrix}}_{D} \quad \underbrace{\begin{bmatrix} a_{21} & & & 0 \\ a_{31} & a_{32} & & \\ a_{41} & a_{42} & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{n\,n-1} \end{bmatrix}}_{L}$$

$$\underbrace{\begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ & a_{23} & a_{24} \cdots & a_{2n} \\ 0 & & a_{34} \cdots & a_{3n} \\ & & \ddots & \\ & & & a_{n-1\,n} \end{bmatrix}}_{U}$$

$$\|M\|_R = \|M\|_\infty \rightarrow \frac{\sum_{i \neq j} |a_{ij}|}{|a_{ii}|} \rightarrow$$

should be ≤ 1 to satisfy sufficient condition for convergence.

$$\downarrow$$

$$\frac{\sum |a_{nb}|}{|a_p|} \leq 1.$$

for 1D system $\longrightarrow$ TDM.
2D " $\longrightarrow$ Pentadiagonal system.
3D " $\longrightarrow$ 7-diag system.

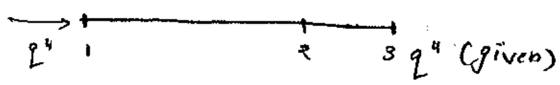
But matrix is generally sparse (sparse matrices).

Scarborough Criteria ~~for~~ of sufficient condition for convergence in Gauss-Siedel method.

$$\frac{\sum |a_{nb}|}{|a_p|} \leq 1 \text{ for all eq}^{ns}$$
$$< 1 \text{ for at least one eq}^{n}.$$

Lecture 38: Illustrative Examples of Iterative methods

Ex 1-D steady state heat conduction in a rod with uniform k, $S = 0$



gde: $\dfrac{d^2T}{dx^2} = 0$

CD: $\dfrac{T_3 + T_1 - 2T_2}{\Delta x^2} = 0$

$$\Rightarrow T_2 = \frac{T_1 + T_3}{2}$$

For gp 1, $q'' = -k\dfrac{(T_2 - T_1)}{\Delta x}$

$$T_1 = T_2 + \left(\frac{q'' \Delta x}{k}\right) \rightarrow c$$

For gp 3, $q'' = -k\dfrac{(T_3 - T_2)}{\Delta x}$

$$T_3 = T_2 - \left(\frac{q'' \Delta x}{k}\right) \rightarrow c$$

$Eq\,1 \longrightarrow T_1 = T_2 + T_C \rightarrow \frac{\sum |a_{nb}|}{|a_p|} = 1$

$Eq\,2 \longrightarrow 2T_2 = T_1 + T_3 \rightarrow \frac{\sum |a_{nb}|}{|a_p|} = \frac{2}{2} = 1$

$Eq\,3 \longrightarrow T_3 = T_2 - T_0$

$\rightarrow \frac{\sum |a_{nb}|}{|a_p|} = \frac{1}{1} = 1$

Ill-posed problems $\because$ we need atleast temp specified at a boundary.

In all cases $\frac{\sum |a_{nb}|}{|a_p|} \leq 1$. In none of the cases is it < 1. Can't satisfy Scarborough's criteria.

So specify at gp1 $T_{given} = T^*$ instead of q''

Then for gp2,

$2T_2 = T_1 + T_3$.

$\Rightarrow 2T_2 = T_1^* + T_3$.

{ Eqⁿ no more valid as q'' at gp1 not known }

$\because T_1^*$ is already known, mathematically it is no longer a neighbour to T_2. Only T_3 is the neighbour.

$\therefore$ for modified eqⁿ ②,

$$\frac{\sum |a_{nb}|}{|a_p|} = \frac{1}{2} \ne 1 < 1$$

With redefinition of the problem, Scarborough criterion is satisfied.

Ex Consider the system,

$$2x_1 + 3x_2 + 10x_3 = 10$$
$$5x_1 - 2x_2 + 2x_3 = 5$$
$$x_1 + 10x_2 + 5x_3 = 6.$$

Q: Is it possible to follow an iterative method (say Jacobi iteration) for the above system with guaranteed convergence?

$\rightarrow$ If yes, what is the estimated no. of iterations to achieve 4 decimal accuracy?

- $\sum_{i \ne j} \frac{|a_{ij}|}{|a_{ii}|} \leq 1 \longrightarrow$ representative of the diagonal dominance.

Diagonal terms are dominating over the sum of the off diagonal terms.

$R_1 \longrightarrow \frac{|3|+|10|}{|2|} \nleq 1$ doesn't satisfy sufficient condition.

How to somehow satisfy the suff cond?

Find an eqⁿ from the set of eqⁿs where the first term's coeff is largest. Here eq-2 has met that criteria. So swap eq-1 with eq-2. [eq-1* = eq-2].

Now $R_1^* \longrightarrow \frac{|-2|+|+2|}{|5|} = \frac{4}{5} \leq 1$.

sufficient condition is satisfied!

IIIly make eq-3 $\rightarrow$ eq-2* & eq-1 $\rightarrow$ eq-3*

Thus reordering equations can help satisfy the suff. cond.

eq-1* $\frac{|2|+|2|}{|5|} = \frac{4}{5}$

eq-2* $\frac{|5|+|1|}{|10|} = \frac{6}{10}$

eq-3* $\frac{|2|+|3|}{|10|} = \frac{5}{10}$

Row sum. norm = 0.8 as max = 0.8

$\rho = \max(\|M\|_R, \|M\|_c)$

Column sum norm $\|M\|_c$

$$\begin{matrix} 5 & -2 & 2 \\ 1 & 10 & 5 \\ 2 & 3 & 10 \end{matrix}$$

col-1 : 3/5

col-2 : 5/10

col-3 : 7/10

max: 0.7

$\|M\|_c = 0.7$

$\rho = \max(0.8, 0.7) = 0.8$

$\therefore \; k = \log_{10}\left(\frac{1}{\rho}\right) = \log_{10}(1.25)$

Ex. For a linear system of size N, the i^{th} eigenvalue of the Jacobi iteration matrix $[M = -D^{-1}(L+U)]$ is given by $\lambda_i = \cos\left(\frac{i\pi}{N+1}\right)$ where $i = 1, 2, \dots, N$. It is also known that for sufficiently large values of N, $\cos\frac{\pi}{N+1} \approx 1 - \frac{\pi^2}{2(N+1)^2}$

$$\approx \exp\left(\frac{-\pi^2}{2N^2}\right).$$

If the size of the coefficient matrix changes from size $10^3 \times 10^3$ to $10^2 \times 10^2$, to what proportion would the total no. of iterations expected to achieve a desired ~~total~~ level of accuracy will decrease?

for $|\lambda_{max}|$, when $i = 1$,

$$\cos\left(\frac{\pi}{N+1}\right) \approx \exp\left(\frac{-\pi^2}{2N^2}\right)$$

$k \longrightarrow \frac{N}{R}$

$\frac{k_2}{k_1} \approx \frac{R_1}{R_2}$ (keeping same level of accuracy, but diff coeff matrix)

$$\frac{k_2}{k_1} = \frac{-\log_{10}(|\lambda_1|_{max}) \rightarrow N = 10^3}{-\log_{10}(|\lambda_1|_{max}) \rightarrow N = 10^2}$$

$\approx 10^{-2}$

$k_2 = 10^{-2} k_1$

$\boxed{\text{Lecture 39: Gradient Search Based Methods}}$

f: function

Grad $f \rightarrow$ represents maximum rate of change of f.

Say, $f = \frac{1}{2}x^T A x - b^T x + c$,   Obj: Solve $Ax = b$.

c arbitrary const.

Two restrictions:

- A symmetric
- A +ve definite.

Next obj, find cond. for min f = ?

min $f \Rightarrow \nabla f = 0$.

$$f = \frac{1}{2}[x_1 \; x_2 \dots x_n] \begin{bmatrix} a_{11} & a_{12} & a_{13} \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

$$- [b_1\ b_2\ b_3 \cdots b_n] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + c$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$f = \frac{1}{2} [x_1\ x_2 \cdots x_n] \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix}$$

$$- [b_1\ b_2 \cdots b_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + c$$

$$f = \frac{1}{2} x_1 (a_{11}x_1 + \cdots + a_{1n}x_n) + \frac{1}{2} x_2 (a_{21}x_1 + \cdots + a_{2n}x_n)$$
$$+ \cdots + \frac{1}{2} x_n (a_{n1}x_1 + \cdots + a_{nn}x_n)$$
$$- (b_1x_1 + b_2x_2 + \cdots + b_nx_n) + c.$$

$$\frac{\partial f}{\partial x_1} = \frac{1}{2}(a_{11}x_1 + \cdots + a_{1n}x_n) + \frac{1}{2}x_1(a_{11}) +$$
$$\frac{1}{2} a_{21} x_2 + \frac{1}{2} a_{31} x_3 + \cdots + \frac{1}{2} a_{n1} x_n$$
$$- b_1$$

$$= \frac{1}{2}(a_{11}x_1 + \cdots + a_{1n}x_n) +$$
$$\frac{1}{2}(a_{11}x_1 + \cdots + a_{n1}x_n) - b_1$$

$\because$ A is symmetric, $a_{ij} = a_{ji}$

$$\Rightarrow \frac{\partial f}{\partial x_1} = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) - b_1$$
$$= [a_{11}\ a_{12} \cdots a_{1n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - b_1.$$

$|||^{ly}$

$$\frac{\partial f}{\partial x_2} = [a_{21}\ a_{22} \cdots a_{2n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - b_2.$$

$\vdots$

$$\Rightarrow \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= \underline{\underline{Ax - b}}$$

$\nabla f = 0$

$\Rightarrow Ax - b = 0$

$\Rightarrow \underline{\underline{Ax = b.}}$

Getting a sol^n $Ax = b$ is as good as extremizing $f = \frac{1}{2}x^T A x - b^T x + c$.

We're actually doing minimization of f & not maximization

How to show it? Why minimization?

$f_{(x)} = \frac{1}{2} x^T A x - b^T x + c$

$f(x+e)$

if $f(x) < f(x+e)$ $\forall e$, then $f(x)$ is minimum.

$$f(x+e) = \frac{1}{2}(x+e)^T A (x+e) - b^T(x+e) + c$$

$$= \frac{1}{2}(x^T A x + e^T A x + x^T A e + e^T A e) - b^T x - b^T e + c$$

$= \cancel{\left[\frac{1}{2}(x^TAx) - b^Tx + c\right]} + \cancel{e^TAx}$

- $x^TAe = (x^TAe)^T$ (transpose of scalar is scalar itself)

$= (Ae)^T(x^T)^T$

$= e^TA^Tx$ (A symmetric; $A^T = A$).

$= e^TAx$

- $b^Te = (b^Te)^T$ (again, scalar).

$= e^T(b^T)^T$

$= e^Tb$.

- $f(x+e) = f(x) + e^TAx - e^Tb + \frac{1}{2}e^TAe$

$= f(x) + e^T(Ax-b) + \frac{1}{2}e^TAe$

B/c $Ax = b$, $e^T(Ax-b) = 0$.

$= f(x) + e^T(Ax-b) + \frac{1}{2}e^TAe$

≥ 0 for arbitrary e (∵ A is positive definite).

$\Rightarrow f(x+e) \geq f(x)$.

$\Rightarrow f(x)$ is a minimum.

— Many gradient search methods!

1. Steepest descent method.

Ex Solve $Ax = b$, where $A = \begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}$, $b = \begin{bmatrix}1\\1\end{bmatrix}$

Ans $f = \frac{1}{2}x^TAx - b^Tx + c$

$$f = \frac{1}{2}[x_1\ x_2]\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} - [1\ 1]\begin{bmatrix}x_1\\ x_2\end{bmatrix} + c$$

$$= \frac{1}{2}[x_1\ x_2]\begin{bmatrix}x_1\\ 2x_2\end{bmatrix} - [x_1 + x_2] + c$$

$$= \frac{1}{2}x_1^2 + x_2^2 - x_1 - x_2 + c$$

What does $f = 0$ represent?

$$\frac{1}{2}(x_1^2 - 2x_1 + 1^2) + \left(x_2^2 - x_2 + \left(\frac{1}{2}\right)^2\right) + c - \frac{3}{4} = 0$$

$$= \frac{(x_1-1)^2}{2} + \frac{(x_2-\frac{1}{2})^2}{1} = \frac{3}{4} - c$$

If f satisfies $(0,0)$,

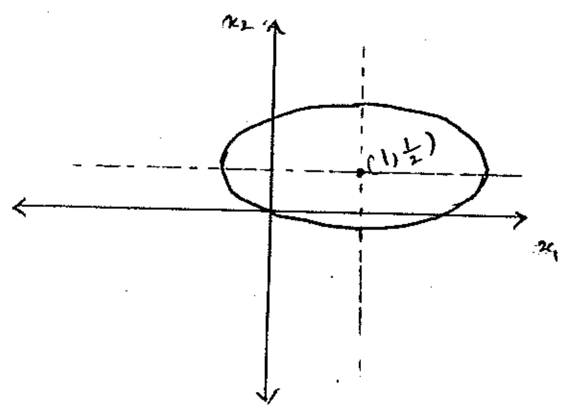
then $\frac{1}{2} + \frac{1}{4} = \frac{3}{4} - c$

$\Rightarrow c = 0$

{∵ choice of c is arbitrary, we can choose our curve to pass through a pt s.t. $c = 0$.}

$$\frac{(x_1-1)^2}{2} + \frac{(x_2-\frac{1}{2})^2}{1} = \frac{3}{4}$$

$$\frac{(x_1-1)^2}{(\sqrt{3/2})^2} + \frac{(x_2-\frac{1}{2})^2}{(\sqrt{3}/2)^2} = 1.$$



Say we start with $(x_1, x_2) = (0,0)$ & try to reach the actual solution $(1, \frac{1}{2})$. We move along the gradient direction.

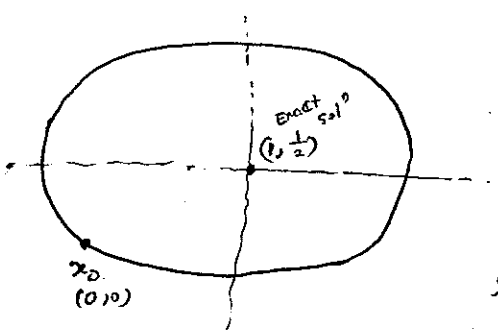
$$\nabla f = Ax - b.$$
$$= -r.$$

After moving some distance, we again stop & move in the gradient direction. The objective then becomes to find how much to travel in each segment.

Lecture 40: Steepest descent method (contd.)

~~for~~ Steepest descent method:

- move along direction of maximum rate of change.



$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\nabla f = Ax - b$$
$$= -r.$$
$$x^1 = x^0 + \alpha_0 r_0.$$
$$r_0 = \nabla f|_{x^0}$$

$\alpha_0 \to$ how much to move in that direction.

$r_0 \to$ tells direction in which you're moving

$$f(x^1) = \frac{1}{2}(x^0 + \alpha_0 r_0)^T A (x^0 + \alpha_0 r_0) - b^T(x^0 + \alpha_0 r_0) + c$$

For f to be min, $\frac{\partial f}{\partial \alpha_0} = 0.$

$$\frac{1}{2} r_0^T A (x^0 + \alpha_0 r_0) + \frac{1}{2}(x^0 + \alpha_0 r_0)^T A r_0 - b^T r_0 = 0.$$

$$\alpha_0 r_0^T A r_0 + \frac{1}{2} r_0^T A x^0 + \frac{1}{2} x^{0T} A r_0 - b^T r_0 = 0.$$

$$x^{0T} A r_0 = (x^{0T} A r_0)^T$$
$$= (A r_0)^T (x^{0T})^T$$
$$= r_0^T A^T x^0$$
$$= r_0^T A x^0 \qquad (A^T = A)$$

$$b^T r_0 = (b^T r_0)^T = r_0^T b.$$

$$\alpha_0 r_0^T A r_0 + r_0^T \underbrace{(Ax^0 - b)}_{-r_0} = 0$$

$$\Rightarrow \alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0} \quad \text{—— (1)}$$

$$x^1 = x^0 + \alpha_0 r_0.$$
$$\alpha_1 = \frac{r_1^T r_1}{r_1^T A r_1}.$$
$$x^2 = x^1 + \alpha_1 r_1$$

Relation b/w directions of r_0 & r_1:

$$r_0^T r_1 = r_0^T [b - A x_1]$$
$$= r_0^T [b - A(x^0 + \alpha_0 r_0)].$$
$$= r_0^T [\underbrace{(b - Ax^0)}_{+r_0} - \alpha_0 A r_0]$$
$$= r_0^T r_0 - \alpha_0 r_0^T A r_0.$$
$$= 0 \longrightarrow \{\text{Using (1)}\}$$

→ r_0 and r_1 are orthogonal to each other.

⇒ we'll be moving mutually perpendicular directions till we reach the solution.

$$r_0 = b - Ax^0 = \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}$$

$$r_0^T r_0 = \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = 2$$

$$r_0^T A r_0 = \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = 3.$$

$$\alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0} = \underline{\frac{2}{3}}$$

$$x^1 = x^0 + \alpha_0 r_0$$

$$= \begin{bmatrix}0\\0\end{bmatrix} + \frac{2}{3}\begin{bmatrix}1\\1\end{bmatrix} = \underline{\begin{bmatrix}2/3\\2/3\end{bmatrix}}$$

$$r_1 = b - Ax^1 = \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}2/3\\2/3\end{bmatrix}$$

$\downarrow$

$\alpha_1 \to \cancel{x} \; x_2 \to \cdots\cdots$ until $r_k = 0$.

Substantial no. of steps required for the naive method.

Improvement $\to$ Conjugate Gradient method.

Lecture 41: Conjugate Gradient Method

$P_0 = r_0$ $\quad x_c (b, \frac{1}{2})$

$P_1 = k(x_c - x^1)$

Initial direction same as steepest descent method

Suppose we want to reach x_c from x^1 in 1 step. Then we'll have to move according to.

$$AP_1 = k(\underbrace{Ax_c}_{b} - Ax^1)$$

($\because x_c$ is correct sol^n).

$$= k(r_1)$$

$$r_0^T r_1 = 0$$

Multiply LHS & RHS by r_0^T

$\rightarrow r_0^T A P_1 = k\, r_0^T r_1 = 0$

$\Rightarrow P_0^T A P_1 = 0$

P_0 is 'A orthogonal' to P_1.

Form,

$P_1 = r_1 - \beta_1 P_0$

$\hookrightarrow$ make new direction from old direction (Gram-Schmidt Conjecture).

$$P_0^T A(r_1 - \beta_1 P_0) = 0.$$

$$\Rightarrow \beta_1 = \frac{P_0^T A r_1}{P_0^T A P_0}$$

We get direction $\overline{P_1}$ from this.

Next to find: how much to go?

$$x^2 = x^1 + \alpha_1 P_1$$

In steepest descent, we used $x^2 = x^1 + \alpha_1 r_1$

(we moved along r_1. Now we move along different direction P_1 with the hope of reaching the target in 1 shot).

α_1 should be such that f should be a min.

$$\Rightarrow f(x^2) = \frac{1}{2}(x^1 + \alpha_1 P_1)^T A (x^1 + \alpha_1 P_1) - b^T(x^1 + \alpha_1 P_1) + c$$

For f to be minimum

$$\frac{\partial f}{\partial \alpha_1} = 0 \longrightarrow$$