10.34, Numerical Methods Applied to Chemical Engineering
Professor William H. Green
Lecture #36: TA-Led Final Review.

BVP: Finite Differences or Method of Lines

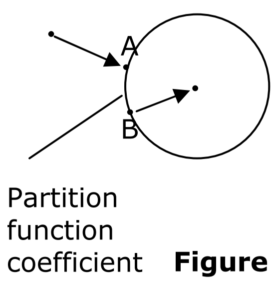
$\frac{\partial C}{\partial x}$ = Forward/Upwind/Central difference formulas

$\frac{\partial^2 C}{\partial x^2}$ = Central difference-like

Understand when to use the different formulas.

Boundary Condition (Flux) $D \left. \frac{\partial C}{\partial x} \right|_{\text{boundary}}$ =Reaction per surface area [moles/m^2·s]

[m^2/s] Internal Flux [(mol/m^3)/m]

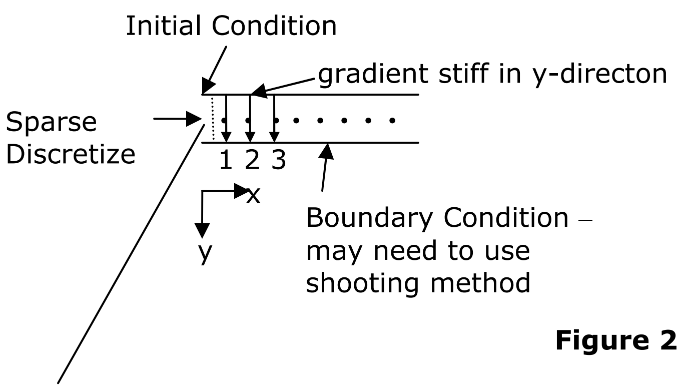


The flux is the same for these two arrows

can solve even if A and B are not known

Figure 1. The flux is the same for arrows at A and B.

Method of Lines



Solve a differential equation along line i = 2, …, N-1

$$\left. \frac{\partial C}{\partial x} \right|_2 = \frac{C_3 - C_1}{2x}$$

Figure 2. Example problem good for method of lines.

If this is the B.C.: $\left. \frac{\partial C}{\partial x} \right|_1 = \frac{C_2 - C_1}{\Delta x}$

Use this additional equation with rest to solve for C_1 D.A.E.

Models vs. Data

$\underline{y} = f(\underline{x},\underline{\theta})$

$y_1 = f(\underline{x}_1,\underline{\theta})$

$y_2 = f(\underline{x}_2,\underline{\theta})$

|

$y_n = f(\underline{x}_n,\underline{\theta})$

Assumption: 1) y distributed normally around $\hat{y}$

2) $\underline{x}$ are known exactly

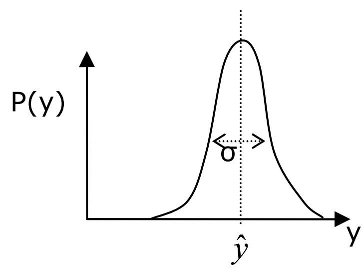


Figure 3. A normal distribution.

WANT:

1) Find the best $\underline{\theta}$
2) Is the model consistent?
3) Error bars on parameters $\underline{\theta}$

Assume model is exact

$\underline{x}_i \rightarrow y_i$ $\hat{y}_i = f(\underline{x}_i,\theta)$ ← data will be distributed around model

$\underline{x}_1 \rightarrow y_1$ $\underline{x}_2 \rightarrow y_2$ ... $\underline{x}_n \rightarrow y_n$

$P(y_i) \propto \exp\left[\frac{-(y_i - f(\underline{x}_i,\theta))^2}{2\sigma^2}\right]$

$P(\underline{y}) \propto \prod_{i=1}^{N} \exp\left[\frac{-(y_i - f(\underline{x}_i,\theta))^2}{2\sigma^2}\right] \propto \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{N}(y_i - f(\underline{x}_i,\theta))^2\right]$

FIT: Max $P(\underline{y})$ → Min $\sum_{i=1}^{N}(y_i - f(\underline{x}_i,\theta))^2$

$k = A \cdot \exp(^{-Ea}/_{RT})$

$\ln k = \ln A - {}^{Ea}/_{R}\, ({}^{1}/_{T})$ Linear in parameters ln k, ln A, ${}^{Ea}/_{R}$

$\underline{y} = \underline{\underline{x}}_n \cdot \underline{\theta} \rightarrow \underline{\theta} = [\underline{\underline{x}}^T \underline{\underline{x}}]^{-1} \underline{\underline{x}}^T \cdot \underline{y}$

$\underline{\underline{x}}_n$: n rows (measurements), m parameters

S.V.D.: $\underline{\underline{x}} = U_{mxm}\Sigma_{mxn}V_{nxn}$

$$\underline{\theta} = \sum_{i=1}^{N}\left(\frac{v_i \cdot \underline{y}}{\sigma_i}\right)v_i \qquad \text{Sample variance guess for } \sigma\text{:} \quad s^2 = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N - \dim(\theta)}$$

$\bar{y}$ is mean y, $f(\underline{x},\underline{\theta})$

If non-linear, use optimization methods.

For correctness, compare s to σ. Quantitatively, use χ^2 (chi squared)

$$\chi^2 = \sum_{i=1}^{N}\frac{\left(y_i - f(\underline{x}_i,\theta)\right)^2}{\sigma^2}$$

Transform to z

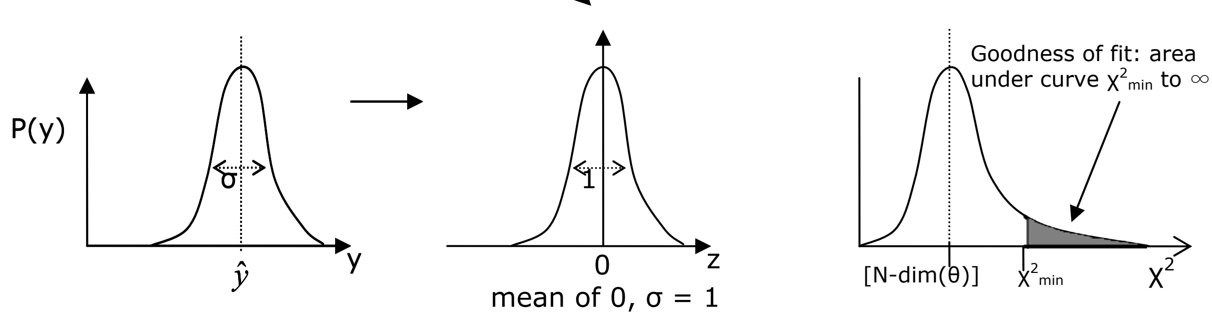


Figure 4. Usually we will accept a model with the integral greater than 5%, but we would like it higher. If 99% chance it is wrong, reject.

Error Bars – Difficult

If linear in parameters and σ is known, covariance(θ) = $\sigma^2[\underline{\underline{x}}^T\underline{\underline{x}}]^{-1}$ (diagonal mxm matrix)

$\theta_i = \theta_{min,i} \pm z_{2,5}\sigma[\underline{\underline{x}}^T\underline{\underline{x}}]_{i,i}^{-1/2}$ m = # parameters

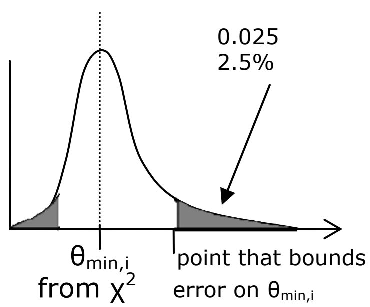


Figure 5. Chi-squared distribution.

Non-linear: $\sigma[\underline{\underline{x}}^T\underline{\underline{x}}]_{i,i}$ $\qquad x_{i,j} = \dfrac{\partial f(\underline{x}_i,\theta)}{\partial\theta_j}$ $\qquad$ Find $x_{i,j}$

In MATLAB, use nlinfit, nlparei

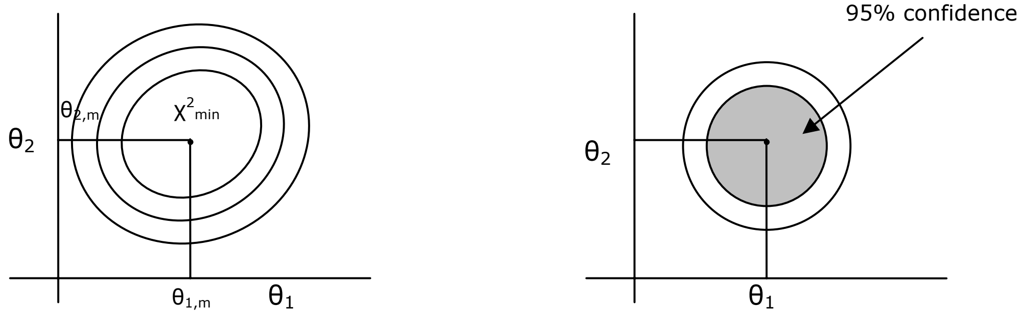
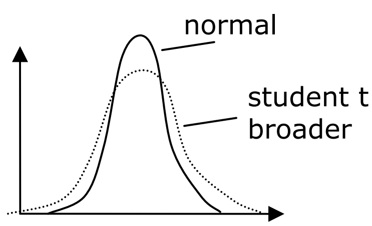


Figure 6. Location of chi-squared and 95% confidence interval in θ_1-θ_2 space.
$\Delta\chi^2 \equiv [\chi_1^2 - \chi_{min}^2]$ v = 2 additional degrees of freedom: let θ_1, θ_2 vary

If σ unknown, use student t distribution based on s.



Report T(χ,v), v being N-dim. θ

as N increases, student t approaches normal distribution

Figure 7. Comparison of normal and Student-t distributions.

$y_i = \theta$ (← you want to calculate θ)

σ is known, y_i is to be measured.

Average value of parameter: $\theta_m = (\Sigma y_i)/N$

$\underline{\underline{x}} = \left[\ \right]_N$ $\underline{\underline{x}}^T\underline{\underline{x}} = [\quad]\left[\ \right] = N$ $\sigma[\underline{\underline{x}}^T\underline{\underline{x}}]^{-1/2} \rightarrow \sigma/\sqrt{N}$

Global Optimization

Convex function – $\underline{\underline{H}} \geq 0$ (Hessian Matrix is positive definite)

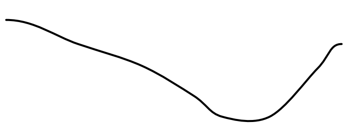


Figure 8. Example of a convex function.
Only 1 minimum

Non-convex:

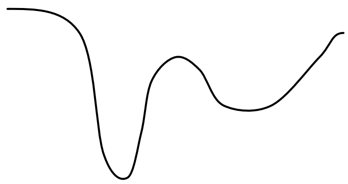
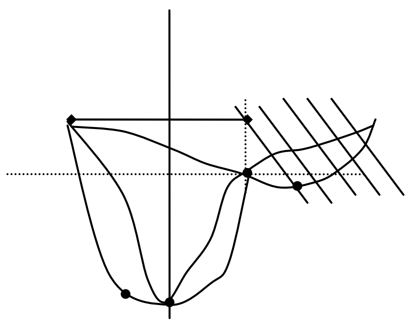


Figure 9. An example of a non-convex function.

Branch and bound

Professor Barton – Non convex function guarantees global minimum



Divide domain
Bound from above
Underestimate below

Find mimina.
Bound again…

Figure 10. An illustration of the branch and bound algorithm.
If new upper bound is lower than the lower bound, use other region; can stop considering that section.

Multistart:

Take a bunch of initial guesses and then run local minimization.

No guarantee.

100 points, 6 variables – 100^6 calculations.

Simulated annealing

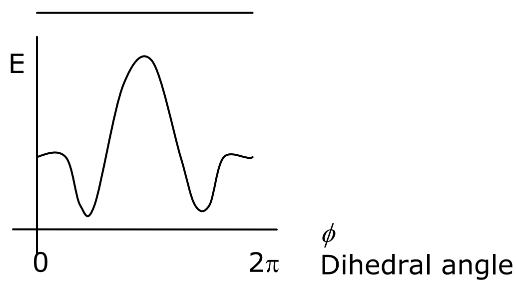


Figure 11. The energy varies with dihedral angle.
Start at high temperature, decrease T eventually can sample wells once the point is caught in a minimum.

Genetic Algorithms

Hybrid system: integer variables and continuous variables

Sample space by allowing function values to live, die, replicate, switch values, etc.

Monte Carlo: Metropolis Monte Carlo

Gillespie Kinetics Monte Carlo

Stochastics

Look at homework solutions to 10 and 11.

Additional Topics

Fourier Transforms and operator splitting may make a showing.