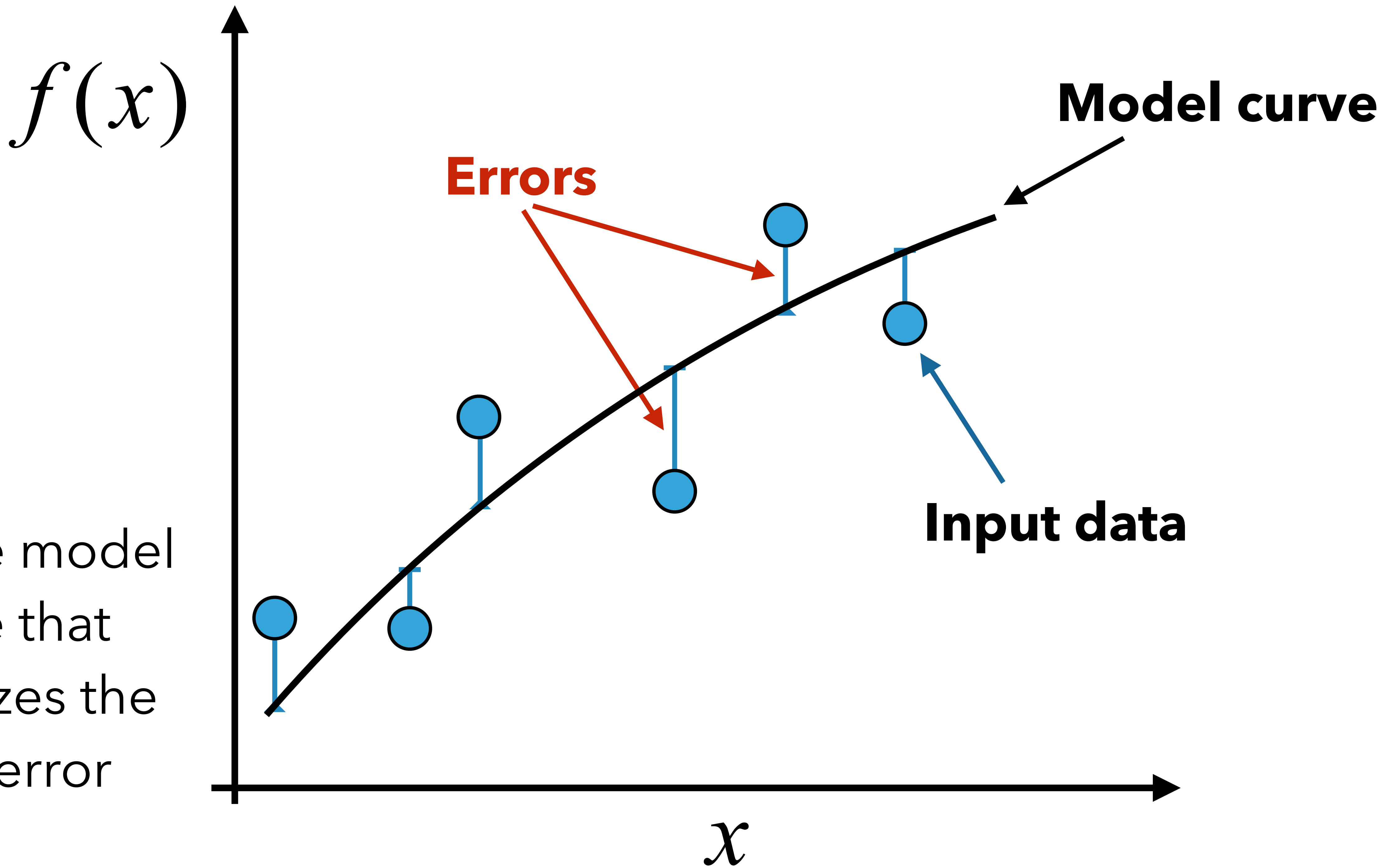
Exam 2 Review

CH EN 2450
Numerical Methods
Fall 2022

Prof. Tony Saad

Department of Chemical Engineering
University of Utah

Basic Idea of Regression



Find the model curve that minimizes the total error

regression to high-order polynomials

$$\text{model} : f(x) = a_m x^m + \ldots + a_2 x^2 + a_1 x + a_0$$

$$\begin{bmatrix} N & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{m+1} \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{m+2} \\ \vdots & & & & \\ \sum x_i^m & \sum x_i^{m+1} & \sum x_i^{m+2} & & \sum x_i^{m+m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^m y_i \end{bmatrix}$$

$m+1$ Equations, $m+1$ Unknowns

Generalized Linear Regression

$$y_{\text{model}} : f(x) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \ldots + \alpha_M \phi_M(x)$$

$$\mathbf{A} = \begin{array}{cc}
\begin{array}{cccc} \text{First basis function} & \text{Second basis function} & & \text{mth basis function} \end{array} & \\
\left[\begin{array}{cccc}
\phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_m(x_1) \\
\phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_m(x_2) \\
\vdots & \vdots & \ldots & \vdots \\
\phi_1(x_N) & \phi_2(x_N) & \ldots & \phi_m(x_N)
\end{array}\right] &
\begin{array}{l} \text{First observation} \\ \text{Second observation} \\ \\ \text{Nth observation} \end{array}
\end{array}$$

$$\mathbf{A}^T\mathbf{A}\boldsymbol{\alpha} = \mathbf{A}^T\mathbf{y}$$

All you need to know is the matrix A!

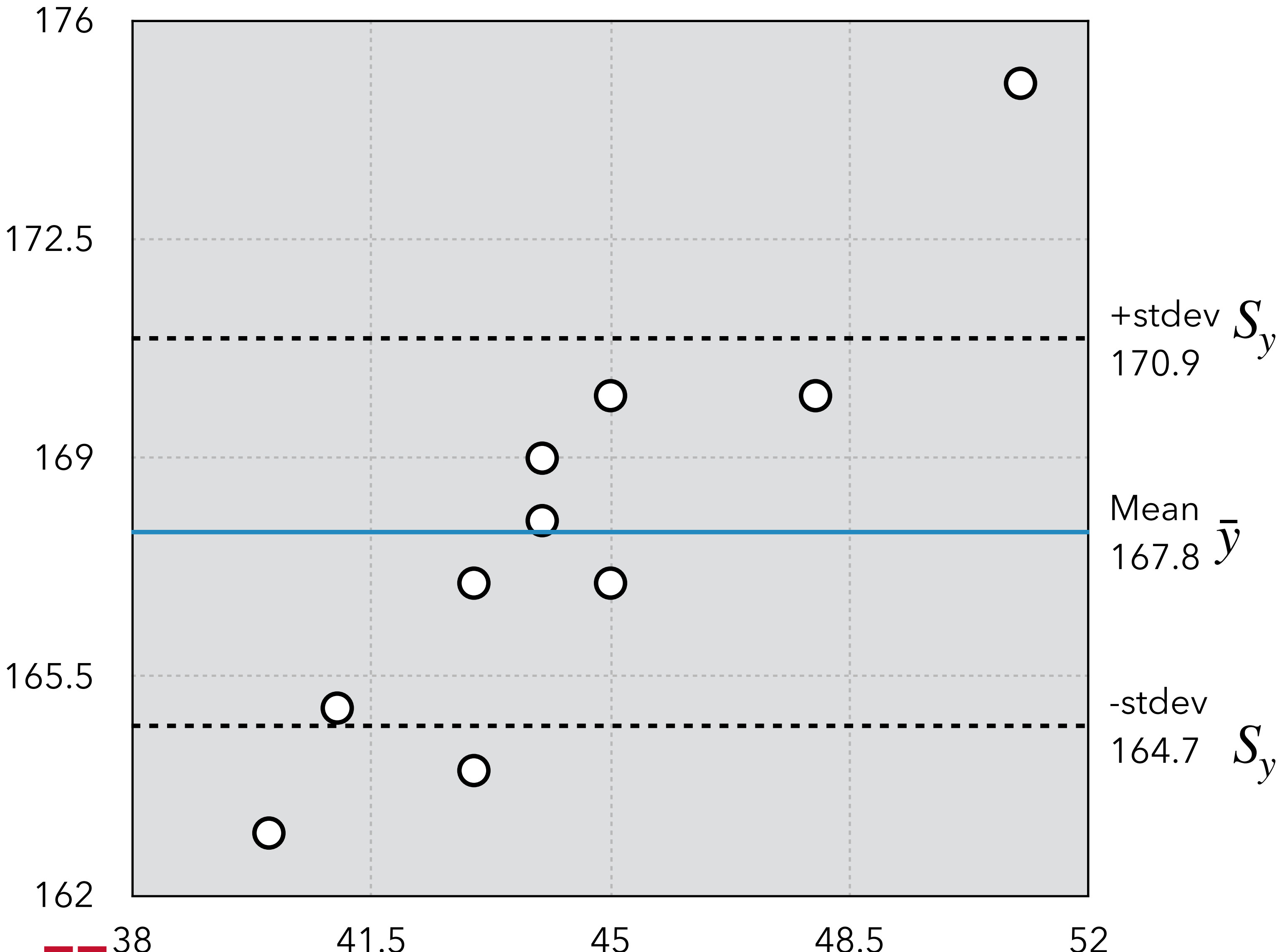
These are called the Normal Equations

Usage

- Given N data pairs in the form $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ and a linear regression model of the form
  $y_{\text{model}} : f(x) = \alpha_1\phi_1(x) + \alpha_2\phi_2(x) + \ldots + \alpha_M\phi_M(x)$
- Build the regression matrix $\mathbf{A}$
  - Column number corresponds to which basis function goes there (first column, first basis function etc…)
  - Row number corresponds to which observation the basis function is applied to (first row corresponds to first observation x_1)
- Build the $\mathbf{y}$ vector which contains the values of the observed dependent variable
- Solve the system: $\mathbf{A}^T\mathbf{A}\boldsymbol{\alpha} = \mathbf{A}^T\mathbf{y}$ - the solution vector, $\boldsymbol{\alpha}$, contains the coefficients of the regression model

Standard Deviation

Standard Deviation: Measures spread of data around the mean value



Mean Value: $\bar{y} = \dfrac{\sum y_i}{n}$

$$S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$$

For a normal distribution,

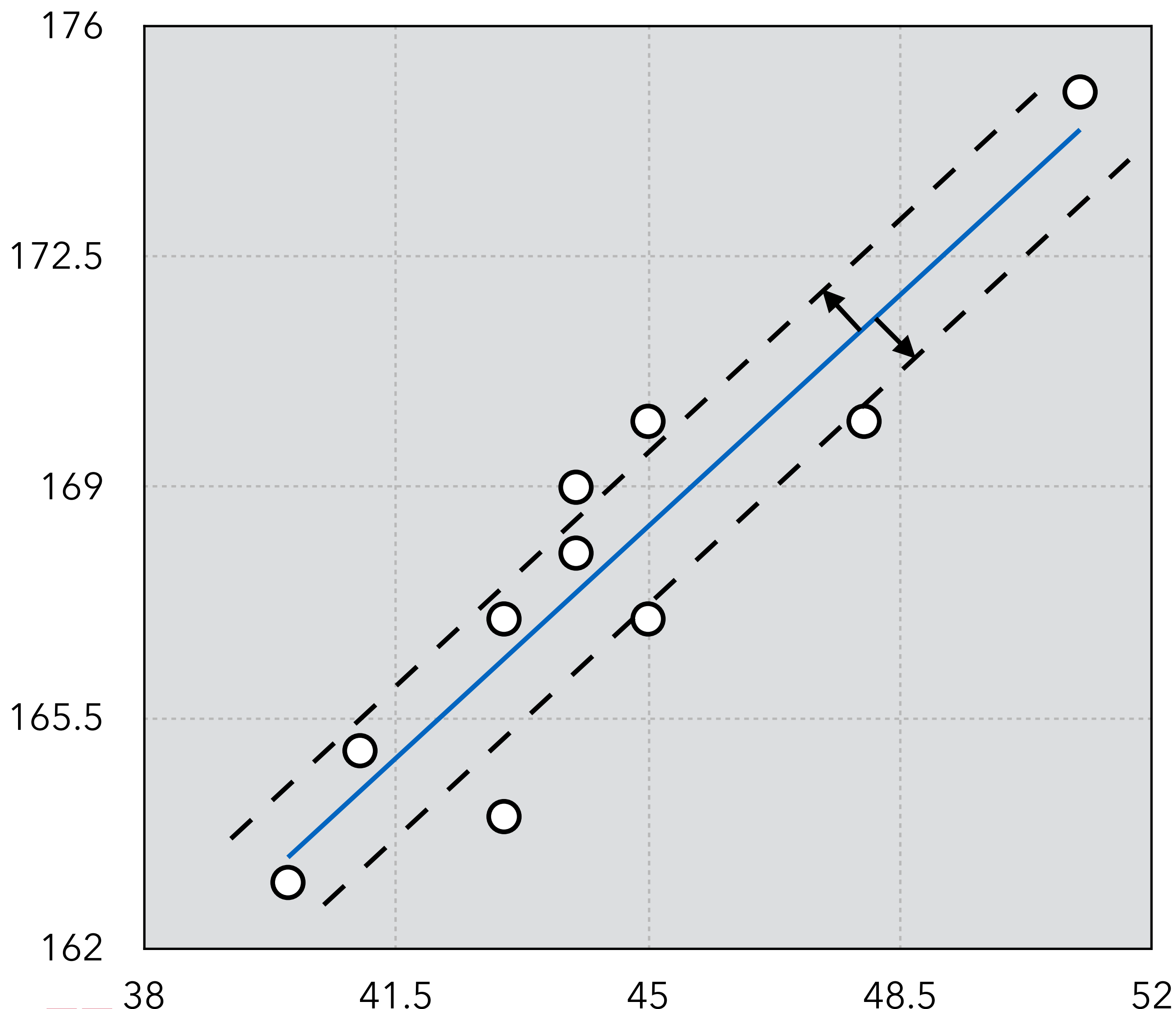
- 68% of the data fall within a standard deviation of the mean
- 95% of the data fall within two standard deviations of the mean

Note that in some cases you will see n used in the denominator for the standard deviation. This has to do with sample vs population stats and is beyond the scope of this course. The standard deviation is a property of a population while it is a statistic for a sample.

Standard Error

Standard Error: Measures spread of data around the regression line

$$S_{y/x} = \sqrt{\frac{\sum (y_i - f_i)^2}{n-2}}$$



For a normal distribution,

- 68% of the data fall within a standard error of the model fit
- 95% of the data fall within two standard errors of the model fit

Standard Error represents the average distance that the observed values fall from the **regression** line. It tells you how wrong the **regression** model is on average using the units of the response variable.

You can think of the standard error as the standard deviation with respect to the regression line

$S_t = \sum (y_i - \bar{y})^2$ Magnitude of error **prior** to regression, with respect to the **mean**

$S_r = \sum (y_i - f_i)^2$ Magnitude of error **after** to regression, with respect to the **regression line**

This is called the R-Squared value
of Coefficient of Determination

$$R^2 = \frac{S_t - S_r}{S_t} = 1 - \frac{S_r}{S_t}$$

- $R^2 \to 0$: Bad fit. No improvement over using a mean value to describe the data
- $R^2 \to 1$: Perfect fit! Significant improvement over just using a mean value to describe the data

What Makes a Regression Model Linear?

- A regression model is linear if
  - The parameters show up linearly OR
  - the least-squares equations can be written as a system of linear equations
- There are two ways to show the latter:
  - Write the least-squares equations (a lot of algebra)
  - Write down part the normal equations (much easier)

y_{model}	**Linear**	**Nonlinear**
$y_{\text{model}} = ax + b$	X	
$y_{\text{model}} = ax + \frac{b}{x}$	X	
$y_{\text{model}} = \sin(a)x + b$	X	
$y_{\text{model}} = \sin(ax) + b$		X
$y_{\text{model}} = (ax)^2 + b$	X	
$y_{\text{model}} = x^a + b$		X
$y_{\text{model}} = a^x + b$		X
$y_{\text{model}} = e^{\alpha}x + \beta x^2$	X	
$y_{\text{model}} = \beta e^{\alpha x}$		X

Nonlinear Least-Squares Regression

Assume that your model, $f(x)$, has n parameters a_k that we want to determine via regression.

$$S = \sum (y_i - f_i)^2$$

$$\frac{\partial S}{\partial a_k} = 0, \quad \forall k$$

$$y_{\text{model}} : f(x) = \cos(a_0 x) + e^{-a_1 x^2} + \sin(a_2 x^2)$$

Same approach as before, but now the parameters of $f(x)$ may enter *nonlinearly*!

Example

$$S = \sum (y_i - f_i)^2 \qquad \frac{\partial S}{\partial a_k} = 0$$

$$f(x) = \alpha e^{\beta x}$$

$$S = \sum (y_i - \alpha e^{\beta x_i})^2$$

$$\frac{\partial S}{\partial \alpha} = 0 \qquad -2\sum e^{\beta x_i}\left(y_i - \alpha e^{\beta x_i}\right) = 0$$

$$\frac{\partial S}{\partial \beta} = 0 \qquad -2\sum \alpha x_i e^{\beta x_i}\left(y_i - \alpha e^{\beta x_i}\right) = 0$$

We will learn how to solve this soon!

Two NONLINEAR equations with two unknowns

$$\sum y_i e^{\beta x_i} - \alpha \sum y_i e^{\beta x_i} = 0$$

$$\alpha \sum x_i y_i e^{\beta x_i} - \alpha^2 \sum x_i y_i e^{\beta x_i} = 0$$

Linear Regression of a Nonlinear Models

Sometimes, it is possible to linearize nonlinear models via appropriate transformations

$$y_{\text{model}} = \alpha e^{\beta f(x)} \quad \text{e.g.} \quad k(T) = A e^{-E\frac{1}{RT}}$$

$$\ln(y_{\text{model}}) = \ln(\alpha) + \beta f(x)$$

$$\underbrace{\ln(y_{\text{model}})}_{Y_{\text{model}}} = \underbrace{\ln(\alpha)}_{A_0} + \underbrace{\beta}_{A_1} \underbrace{f(x)}_{X}$$

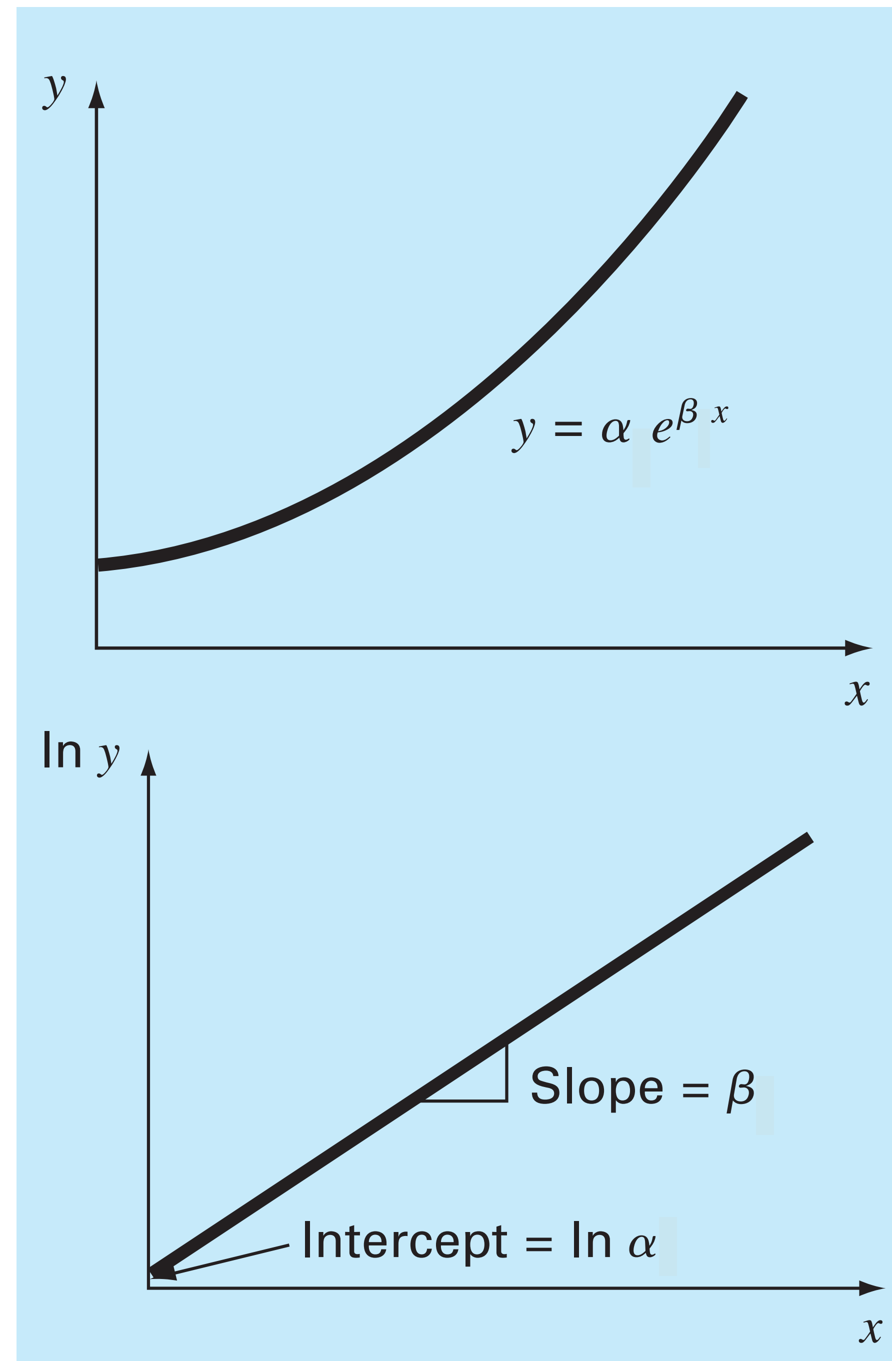
$$Y_{\text{model}} = A_0 + A_1 X$$

$$Y_{\text{model}} = \ln(y_{\text{model}})$$

$$A_0 = \ln(\alpha)$$

$$A_1 = \beta$$

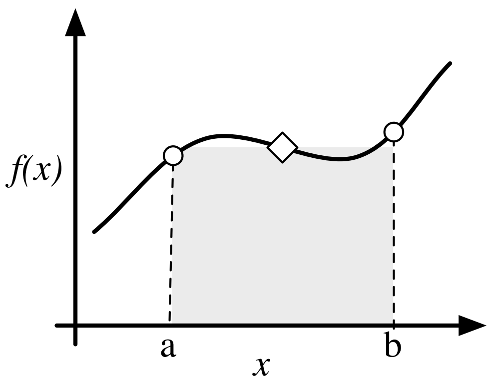
$$X = f(x)$$



Midpoint Rule

<u>Concept</u>: Approximate $f(x)$ as a *constant* on the interval $[a,b]$.

$$\int_a^b f(x)\mathrm{d}x \approx (b-a)\, f\left(\frac{b+a}{2}\right)$$

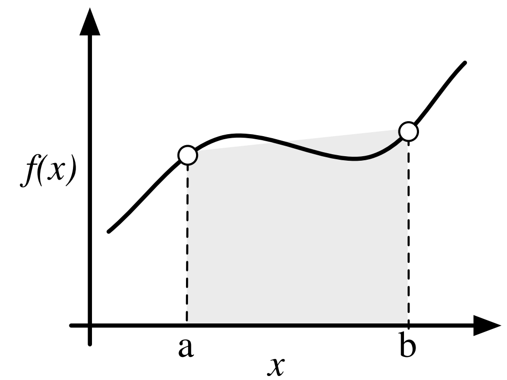


- Requires function value at the midpoint (can be a problem for tabular/discrete data).

Trapezoid Rule

<u>Concept</u>: Approximate $f(x)$ as a *linear* function on the interval $[a,b]$.

$$\int_a^b f(x)\mathrm{d}x \approx \frac{(b-a)}{2}\left[f(a)+f(b)\right]$$

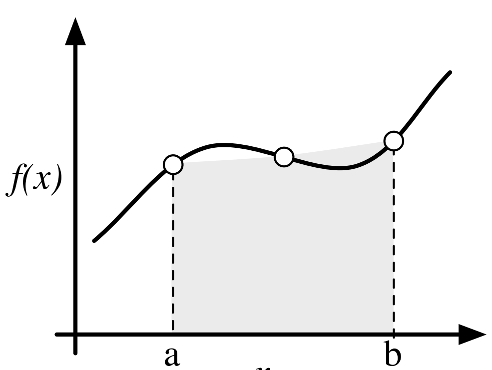


- Convenient form for tabular (discrete) data.
- Doesn't require equally spaced data.
- $\Delta x = b-a$

Simpson's 1/3 Rule

<u>Concept</u>: Approximate $f(x)$ as a *quadratic* on the interval $[a,b]$.

$$\int_a^b f(x)\mathrm{d}x \approx \frac{\Delta x}{3}\left[f(a) + 4f\left(\tfrac{a+b}{2}\right) + f(b)\right]$$

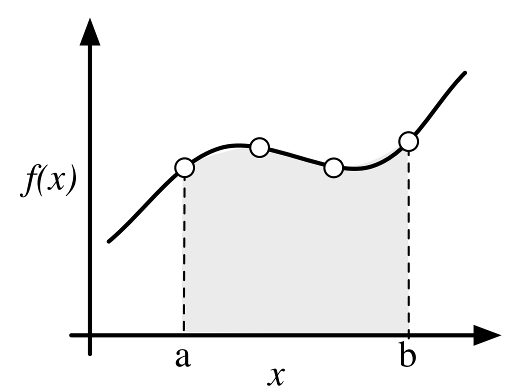


- Requires three equally spaced points on interval $[a,b]$.
- $\Delta x = (b-a)/2$

Simpson's 3/8 Rule

<u>Concept</u>: Approximate $f(x)$ as a *cubic* on the interval $[a,b]$.

$$\int_a^b f(x)\mathrm{d}x \approx \frac{3\Delta x}{8}\left(f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)\right)$$



- Requires four equally spaced points on interval $[a,b]$.
- $\Delta x = (b-a)/3$
- $x_i = a + i\Delta x$

Error Bounds

$$\int_a^b f(x)\mathrm{d}x \quad |f''(x)| \le K \quad |f''''(x)| \le M$$ on [a,b]

Midpoint

$$E_M \le K\frac{(b-a)^3}{24n^2}$$

Trapezoid

$$E_T \le K\frac{(b-a)^3}{12n^2}$$

Simpson's 1/3

$$E_S \le M\frac{(b-a)^5}{18n^4}$$

n = Number of segments

Note, these are only UPPER bounds. The actual error may be much smaller depending on the function.

First Derivatives

formula	order		Constant Δx ?
$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right\|_i \approx \frac{f_{i+1}-f_i}{\Delta x}$	Δx	Forward difference	no
$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right\|_i \approx \frac{f_i-f_{i-1}}{\Delta x}$	Δx	backward difference	no
$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right\|_i \approx \frac{f_{i+1}-f_{i-1}}{2\Delta x}$	Δx^2	central difference	yes
$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right\|_i \approx \frac{-3f_i+4f_{i+1}-f_{i+2}}{2\Delta x}$	Δx^2	Forward difference	yes
$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right\|_i \approx \frac{3f_i-4f_{i-1}+f_{i-2}}{2\Delta x}$	Δx^2	backward difference	yes

Some Approximations for Second Derivatives

All assume constant Δx.

$\frac{\partial^2 f}{\partial x^2}$	Order	Comments
$\left.\frac{\partial^2 f}{\partial x^2}\right\|_i \approx \frac{f_i - 2f_{i+1} + f_{i+2}}{\Delta x^2}$	Δx	Forward difference
$\left.\frac{\partial^2 f}{\partial x^2}\right\|_i \approx \frac{f_i - 2f_{i-1} + f_{i-2}}{\Delta x^2}$	Δx	Backward difference
$\left.\frac{\partial^2 f}{\partial x^2}\right\|_i \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2}$	Δx^2	Central difference
$\left.\frac{\partial^2 f}{\partial x^2}\right\|_i \approx \frac{2f_i - 5f_{i+1} + 4f_{i+2} - f_{i+3}}{\Delta x^2}$	Δx^2	Forward difference
$\left.\frac{\partial^2 f}{\partial x^2}\right\|_i \approx \frac{-f_{i-3} + 4f_{i-2} - 5f_{i-1} + 2f_i}{\Delta x^2}$	Δx^2	Backward difference