Course Instructor
Dr. Raymond C. Rumpf
Office: A-337
Phone: (915) 747-6958
E-Mail: rcrumpf@utep.edu

Computational Methods in EE

Topic 5a – Introduction to Curve Fitting & Linear Regression

EE 4386/5301 Computational Methods in EE

Outline

Computational Methods in EE

- Introduction
- Statistics of Data Sets
- Best Fit Methods
  - Linear Regression (ugly math)
  - Linear Least Squares (clean math)

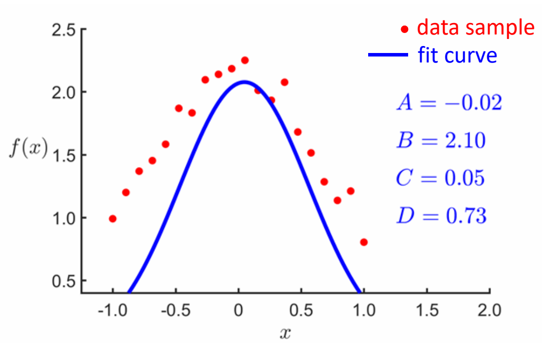
Computational Methods in EE

Introduction

What is Curve Fitting?

Computational Methods in EE

Curve fitting is simply fitting an analytical equation to a set of measured data.



$$f(x) = A + Be^{-\left(\frac{x-C}{D}\right)^2}$$

"Curve fitting" determines the values of A, B, C, and D so that $f(x)$ best represents the given data.

Why Fit Data to a Curve?

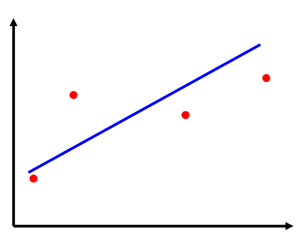
Computational Methods in EE

- Estimate data between discrete values (interpolation)
- Find a maximum or minimum.
- Deriving finite-difference approximations.
- Fit measured data to an analytical equation to extract meaningful parameters.
- Remove noise from a function.
- Observe and quantify general trends.

Two Categories of Curve Fitting

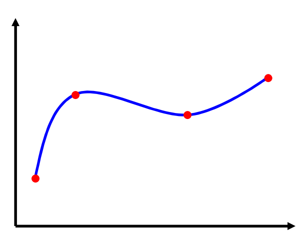
Computational Methods in EE

Best Fit – Measured data has noise so the curve does not attempt to intercept every point.



- Linear regression (ugly math)
- Linear least-squares (clean math)
- Nonlinear regression (moderate math)

Exact Fit – Data samples are assumed to be exact and the curve is forced to pass through each one.



- Fitting to polynomials

Computational Methods in EE

Statistics of Data Sets

Arithmetic Mean

Computational Methods in EE

If we had to come up with a single number that be represents an entire set of data, the arithmetic mean would probably be it.

$$f_{\text{avg}} = \frac{f_1 + f_2 + \cdots + f_M}{M} = \frac{1}{M} \sum_{m=1}^{M} f_m$$

Geometric Mean

Computational Methods in EE

The geometric mean is defined as

$$f_{\text{gm}} = \sqrt[M]{f_1 f_2 \cdots f_M}$$

The arithmetic mean tends to suppress the significance of outlying data samples. With the geometric mean, even a single small value among many large values can dominate the mean.

This is useful in optimizations where multiple parameters must be maximized at the same time and it is not acceptable to have any one of them low.

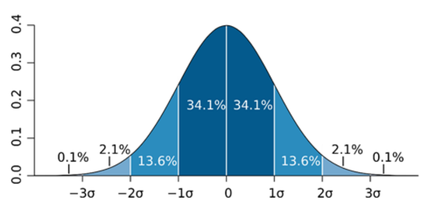
Variance & Standard Deviation

Computational Methods in EE

Standard Deviation σ_f

The standard deviation is a measure of the "spread" of the data about the mean. It is convenient because it shares the same units as the data.

$$\sigma_f = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(f_m - f_{\text{avg}}\right)^2}$$



Variance v_f

Variance is used more commonly in calculations, but carries the same information as the standard deviation.

$$v_f = \sigma_f^2 = \frac{1}{M}\sum_{m=1}^{M}\left(f_m - f_{\text{avg}}\right)^2$$

Coefficient of Variation

Computational Methods in EE

The coefficient of variation (CV) is the standard deviation normalized to the mean.

Think of it as "relative standard deviation."

$$\text{CV} = \frac{\sigma_f}{f_{\text{avg}}}$$

Computational Methods in EE

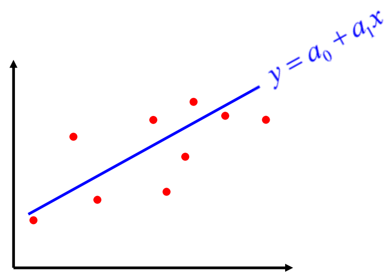
Linear Regression

(Best Fit, Ugly Math)

Goal of Linear Regression

Computational Methods in EE

The goal of linear regression is to fit a straight line to a set of measured data that has noise.



$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_M, y_M)$$
$$\downarrow$$
$$y = a_0 + a_1 x$$

Statement of Problem

Computational Methods in EE

Given a set of measured data points: (x_1,y_1), (x_2,y_2), …, (x_M,y_M), we write our equation of the line for each point.

$$\begin{aligned} y_1 &= a_0 + a_1 x_1 + e_1 \\ y_2 &= a_0 + a_1 x_2 + e_2 \\ &\vdots \\ y_M &= a_0 + a_1 x_M + e_M \end{aligned}$$

To make these equations correct, we must introduce an error term e called the *residual*.

We wish to determine values of a_0 and a_1 such that the residual terms e_m are as small as possible.

Criteria for "Best Fit"

Computational Methods in EE

We must define a single quantity that tells us how "good" the line fits our set of data.

Guess #1 – Sum of Residuals

$$E = \sum_{m=1}^{M} e_m$$

This does not work because negative and positive residuals can cancel and mislead the overall criteria to think there is no error.

Guess #2 – Sum of Magnitude of Residuals

$$E = \sum_{m=1}^{M} \left| e_m \right|$$

This does not work because is does not lead to a unique best fit.

Guess #3 – Sum of Squares of Residuals

$$E = \sum_{m=1}^{M} e_m^2$$

This works and leads to a unique solution.

Equation for Criterion

Computational Methods in EE

Our line equation for the mth sample is

$$y_m = a_0 + a_1 x_m + e_m$$

Solving this for the residual e_m gives

$$e_m = y_m - \left(a_0 + a_1 x_m \right)$$

This is the value of y of our line at point x_m.

This is our measured value of y.

From this, we can write our criterion as

$$E = \sum_{m=1}^{M} e_m^2 = \sum_{m=1}^{M} \left(y_{\text{measured},m} - y_{\text{line},m} \right)^2 = \sum_{m=1}^{M} \left(y_m - a_0 - a_1 x_m \right)^2$$

Least-Squares Fit

Computational Methods in EE

We wish to minimize the error criterion E.

We can identify minimums when the first-order derivative is zero.

$$\frac{\partial E}{\partial a_0} = 0 \quad \text{and} \quad \frac{\partial E}{\partial a_1} = 0$$

We seek the values of a_0 and a_1 that satisfy these equations.

This approach is solving the problem by least-squares (we are minimizing the squares of the residuals).

The Fun Math

Computational Methods in EE

Step 1 – Differential E with respect to each of the unknowns.

$$\begin{aligned}\frac{\partial E}{\partial a_0} &= \frac{\partial}{\partial a_0}\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)^2 \\ &= \sum_{m=1}^{M}\frac{\partial}{\partial a_0}\left(y_m - a_0 - a_1 x_m\right)^2 \\ &= \sum_{m=1}^{M}2\left(y_m - a_0 - a_1 x_m\right)(-1) \\ &= -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)\end{aligned}$$

$$\begin{aligned}\frac{\partial E}{\partial a_1} &= \frac{\partial}{\partial a_1}\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)^2 \\ &= \sum_{m=1}^{M}\frac{\partial}{\partial a_1}\left(y_m - a_0 - a_1 x_m\right)^2 \\ &= \sum_{m=1}^{M}2\left(y_m - a_0 - a_1 x_m\right)(-x_m) \\ &= -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)x_m\end{aligned}$$

The Fun Math

Computational Methods in EE

Step 2 – We set the derivatives to zero to find the minimum of E.

$$0 = \frac{\partial E}{\partial a_0} = -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right) = \sum_{m=1}^{M} y_m - \sum_{m=1}^{M} a_0 - \sum_{m=1}^{M} a_1 x_m = \sum_{m=1}^{M} y_m - Ma_0 - \sum_{m=1}^{M} a_1 x_m$$

$$0 = \frac{\partial E}{\partial a_1} = -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)x_m = \sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)x_m = \sum_{m=1}^{M} y_m x_m - a_0\sum_{m=1}^{M} x_m - \sum_{m=1}^{M} a_1 x_m^2$$

The Fun Math

Computational Methods in EE

Step 3 – We write these as two simultaneous equations. These are called the *normal equations*.

$$0 = \frac{\partial E}{\partial a_0} = -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right) = \sum_{m=1}^{M} y_m - \sum_{m=1}^{M} a_0 - \sum_{m=1}^{M} a_1 x_m = \sum_{m=1}^{M} y_m - Ma_0 - a_1\sum_{m=1}^{M} x_m$$

$$\downarrow$$

$$Ma_0 + a_1\sum_{m=1}^{M} x_m = \sum_{m=1}^{M} y_m$$

$$0 = \frac{\partial E}{\partial a_1} = -2\sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)x_m = \sum_{m=1}^{M}\left(y_m - a_0 - a_1 x_m\right)x_m = \sum_{m=1}^{M} y_m x_m - a_0\sum_{m=1}^{M} x_m - \sum_{m=1}^{M} a_1 x_m^2$$

$$\downarrow$$

$$a_0\sum_{m=1}^{M} x_m + a_1\sum_{m=1}^{M} x_m^2 = \sum_{m=1}^{M} y_m x_m$$

The Fun Math

Computational Methods in EE

Step 4 – The normal equations are solved simultaneously and the solution is

$$a_0 = y_{\text{avg}} - a_1 x_{\text{avg}}$$

$$a_1 = \frac{M\sum_{m=1}^{M} x_m y_m - \sum_{m=1}^{M} x_m \sum_{m=1}^{M} y_m}{M\sum_{m=1}^{M} x_m^2 - \left(\sum_{m=1}^{M} x_m\right)^2}$$

$$x_{\text{avg}} = \frac{1}{M}\sum_{m=1}^{M} x_m$$

$$y_{\text{avg}} = \frac{1}{M}\sum_{m=1}^{M} y_m$$

Yikes!
There has to be an easier way!

Computational Methods in EE

Linear Least-Squares

(Best Fit, Clean Math)

Statement of Problem

Computational Methods in EE

We wish to fit a set of M measured data points to a curve containing $N+1$ terms:

$$f = a_0 z_0 + a_1 z_1 + a_2 z_2 + \cdots + a_N z_N$$

$$f \equiv \text{measured value}$$
$$z_n \equiv \text{parameters from which } f \text{ is evaluated}$$
$$a_n \equiv \text{coefficients for the curve fit}$$

Formulation of Matrix Equation

Computational Methods in EE

We start by writing the function f for each of our M measurements. We also include the residual term.

$$\begin{aligned} f_1 &= a_0 z_{0,1} + a_1 z_{1,1} + \cdots + a_N z_{N,1} + e_1 \\ f_2 &= a_0 z_{0,2} + a_1 z_{1,2} + \cdots + a_N z_{N,2} + e_2 \\ &\vdots \\ f_M &= a_0 z_{0,M} + a_1 z_{1,M} + \cdots + a_N z_{N,M} + e_M \end{aligned}$$

This large set of equations is put into matrix form.

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{M-1} \\ f_M \end{bmatrix} = \begin{bmatrix} z_{0,1} & z_{1,1} & \cdots & z_{N,1} \\ z_{0,2} & z_{1,2} & \cdots & z_{N,2} \\ z_{0,3} & z_{1,3} & \vdots & z_{N,3} \\ \vdots & \vdots & & \vdots \\ z_{0,M-1} & z_{1,M-1} & \cdots & z_{N,M-1} \\ z_{0,M} & z_{1,M} & \cdots & z_{N,M} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{M-1} \\ e_M \end{bmatrix} \Rightarrow \begin{matrix} [f] = [Z][a] + [e] \\ \text{or} \\ \mathbf{f} = \mathbf{Z}\mathbf{a} + \mathbf{e} \end{matrix}$$

Formulation of Solution by Least-Squares (1 of 4)

Computational Methods in EE

Step 1 – Solve matrix equation for $\mathbf{e}$.

$$\mathbf{f} = \mathbf{Za} + \mathbf{e} \quad \rightarrow \quad \mathbf{e} = \mathbf{f} - \mathbf{Za}$$

Step 2 – Calculate the error criterion E from $\mathbf{e}$.

$$E = \sum_{m=1}^{M} e_m^2 = \begin{bmatrix} e_1 & e_2 & e_3 & \cdots & e_{M-1} & e_M \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{M-1} \\ e_M \end{bmatrix} = \mathbf{e}^T \mathbf{e}$$

Step 3 – We substitute our equation for $\mathbf{e}$ from Step 1 into our equation for E from Step 2.

$$E = \mathbf{e}^T \mathbf{e} = (\mathbf{f} - \mathbf{Za})^T (\mathbf{f} - \mathbf{Za})$$

Topic 5a -- Intro & Linear Regression 25

Formulation of Solution by Least-Squares (2 of 4)

Computational Methods in EE

Step 4 – Our new matrix equation is algebraically manipulated as follows in order to make it easier to find its first-order derivative.

$$E = (\mathbf{f} - \mathbf{Za})^T (\mathbf{f} - \mathbf{Za}) \qquad \text{original equation}$$

$$= (\mathbf{f}^T - \mathbf{a}^T \mathbf{Z}^T)(\mathbf{f} - \mathbf{Za}) \qquad \text{distribute the transpose}$$

$$= \mathbf{f}^T \mathbf{f} - \underbrace{\mathbf{f}^T \mathbf{Za} - \mathbf{a}^T \mathbf{Z}^T \mathbf{f}}_{\text{These are scalars and transposes of each other so they are equal.}} + \mathbf{a}^T \mathbf{Z}^T \mathbf{Za} \qquad \text{expand equation}$$

$$= \mathbf{f}^T \mathbf{f} - 2\mathbf{a}^T \mathbf{Z}^T \mathbf{f} + \mathbf{a}^T \mathbf{Z}^T \mathbf{Za} \qquad \text{combine terms}$$

Topic 5a -- Intro & Linear Regression 26

Formulation of Solution by Least-Squares (3 of 4)

Computational Methods in EE

Step 5 – We differential E with respect to $\mathbf{a}$.

We wish to determine $\mathbf{a}$ that minimizes E.
We can do this using the first-derivative rule.

$$E = \mathbf{f}^T\mathbf{f} - 2\mathbf{a}^T\mathbf{Z}^T\mathbf{f} + \mathbf{a}^T\mathbf{Z}^T\mathbf{Z}\mathbf{a}$$

$$\frac{\partial E}{\partial \mathbf{a}} = \frac{\partial}{\partial \mathbf{a}}\left(\mathbf{f}^T\mathbf{f} - 2\mathbf{a}^T\mathbf{Z}^T\mathbf{f} + \mathbf{a}^T\mathbf{Z}^T\mathbf{Z}\mathbf{a}\right) \qquad \text{substitute in expression for } E$$

$$= \cancel{\frac{\partial}{\partial \mathbf{a}}\mathbf{f}^T\mathbf{f}} - 2\frac{\partial}{\partial \mathbf{a}}\mathbf{a}^T\mathbf{Z}^T\mathbf{f} + \frac{\partial}{\partial \mathbf{a}}\mathbf{a}^T\mathbf{Z}^T\mathbf{Z}\mathbf{a} \qquad \mathbf{f} \text{ is not a function of } \mathbf{a}$$

$$= -2\mathbf{Z}^T\mathbf{f} + 2\mathbf{Z}^T\mathbf{Z}\mathbf{a} \qquad \text{finish differentiation}$$

Formulation of Solution by Least-Squares (4 of 4)

Computational Methods in EE

Step 6 – We find the value of $\mathbf{a}$ that makes the derivative equal to zero.

$$\frac{\partial E}{\partial \mathbf{a}} = -2\mathbf{Z}^T\mathbf{f} + 2\mathbf{Z}^T\mathbf{Z}\mathbf{a} = 0$$

$$-2\mathbf{Z}^T\mathbf{f} + 2\mathbf{Z}^T\mathbf{Z}\mathbf{a} = 0$$

$$2\mathbf{Z}^T\mathbf{Z}\mathbf{a} = 2\mathbf{Z}^T\mathbf{f}$$

$$\mathbf{Z}^T\mathbf{Z}\mathbf{a} = \mathbf{Z}^T\mathbf{f}$$

Observe that this is the original equation premultiplied by $\mathbf{Z}^T$.

$$\boxed{\mathbf{a} = \left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{f}}$$

DO NOT SIMPLIFY FURTHER!

Computational Methods in EE

If we were to simplify our least-squares equation, we would get

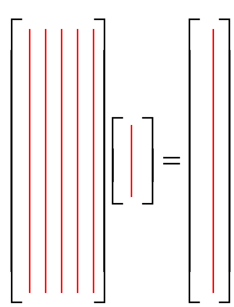
$$\begin{aligned}\mathbf{a} &= \left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{f}\\ &= \mathbf{Z}^{-1}\underbrace{\left(\mathbf{Z}^T\right)^{-1}\mathbf{Z}^T}_{\mathbf{I}}\mathbf{f}\\ &= \mathbf{Z}^{-1}\mathbf{f}\end{aligned}$$

This is just our original equation again ($\mathbf{f} = \mathbf{Za}$) without the least-squares approach incorporated.

Visualizing Least-Squares (1 of 3)

Computational Methods in EE

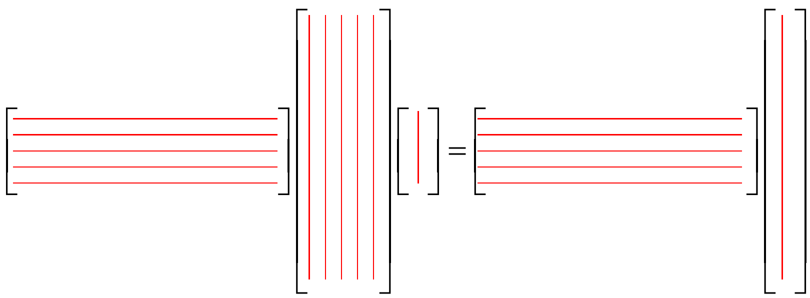
We are initially given a matrix equation with more equations than unknowns.



Visualizing Least-Squares (2 of 3)

Computational Methods in EE

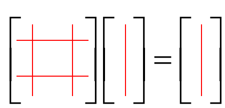
We premultiply by the transpose of **A**.



Visualizing Least-Squares (3 of 3)

Computational Methods in EE

The matrix equations reduces to the same number of equations as unknowns, which is solvable by many standard algorithms.



Least-Squares Algorithm

Computational Methods in EE

Step 1 – Construct matrices. $\mathbf{Z}$ is essentially just a matrix of the coordinates of the data points. $\mathbf{f}$ is a column vector of the measurements.

$$\mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{M-1} \\ f_M \end{bmatrix} \qquad \mathbf{Z} = \begin{bmatrix} z_{0,1} & z_{1,1} & \cdots & z_{N,1} \\ z_{0,2} & z_{1,2} & \cdots & z_{N,2} \\ z_{0,3} & z_{1,3} & \cdots & z_{N,3} \\ \vdots & \vdots & & \vdots \\ z_{0,M-1} & z_{1,M-1} & & z_{N,M-1} \\ z_{0,M} & z_{1,M} & & z_{N,M} \end{bmatrix}$$

Step 2 – Solve for the unknown coefficients $\mathbf{a}$.

$$\mathbf{a} = \left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{f}$$

Step 3 – Extract the coefficients from $\mathbf{a}$.

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix}$$

Least-Squares for Solving $\mathbf{Ax} = \mathbf{b}$

Computational Methods in EE

Suppose we wish to solve $\mathbf{Ax} = \mathbf{b}$, but we have more equations than we have unknowns.

We must solve this as a “best fit” because a perfect fit is impossible in the presence of noise.

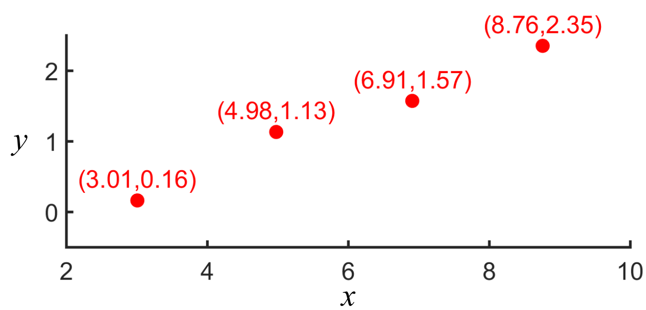
We apply least-squares by premultipling by $\mathbf{A}^T$.

$$\mathbf{Ax} = \mathbf{b} \quad \rightarrow \quad \mathbf{A}'\mathbf{x} = \mathbf{b}'$$
$$\mathbf{A}' = \mathbf{A}^T\mathbf{A}$$
$$\mathbf{b}' = \mathbf{A}^T\mathbf{b}$$

Example 1 (1 of 3)

Computational Methods in EE

Let's fit a line to the following set of points.



$$y = mx + b$$

Example 1 (2 of 3)

Computational Methods in EE

Step 1 – Build matrices

$$\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ y_3 &= mx_3 + b \\ y_4 &= mx_4 + b \end{aligned} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} \rightarrow \mathbf{f} = \begin{bmatrix} 0.16 \\ 1.13 \\ 1.57 \\ 2.35 \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} 3.01 & 1 \\ 4.98 & 1 \\ 6.91 & 1 \\ 8.76 & 1 \end{bmatrix}$$

With some practice, you will be able to write matrices directly from measured data.

Step 2 – Solve by least squares.

$$\mathbf{Z}^T\mathbf{Z}\mathbf{x} = \mathbf{Z}^T\mathbf{f} \rightarrow \begin{bmatrix} 3.01 & 4.98 & 6.91 & 8.76 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3.01 & 1 \\ 4.98 & 1 \\ 6.91 & 1 \\ 8.76 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 3.01 & 4.98 & 6.91 & 8.76 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.16 \\ 1.13 \\ 1.57 \\ 2.35 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 158.3462 & 23.6600 \\ 23.6600 & 4.0000 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 37.5437 \\ 5.2100 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0.3656 \\ -0.8602 \end{bmatrix}$$

Example 1 (3 of 3)

Computational
Methods in EE

Step 3 – Extract coefficients

$$\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0.3656 \\ -0.8602 \end{bmatrix} \quad \rightarrow \quad \begin{matrix} m = 0.3656 \\ b = -0.8602 \end{matrix}$$

Step 4 – Plot the result