Objectives

- ☐ I understand that a matrix is a map that sends vectors to vectors.
- ☐ Given data for two measurements, I can fit that data to *any* kind of function.
- ☐ Given data for any number of measurements, I can fit that data to *any* kind of function.

We first introduced matrices as a tool to find a solution to a system of linear equations. That is, we turned a set of linear equations like these

$$\begin{aligned} 2x - 3y + z &= 10 \\ 2y + 2z &= 4 \\ 2x \qquad - 3z &= 1 \end{aligned}$$

into an equation like this

$$\begin{pmatrix} 2 & -3 & 1 \\ 0 & 2 & 2 \\ 2 & 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 1 \end{pmatrix}$$

When written with matrices, we can think of the expression as

$$A\vec{x} = \vec{b}$$

In the above equation, notice that the matrix A sends $\vec{x}$ to $\vec{b}$. When we solve the equation (when we calculate $\vec{x} = A^{-1}\vec{b}$) we get the solution

$$\begin{pmatrix} \frac{26}{7} \\ -\frac{1}{7} \\ \frac{15}{7} \end{pmatrix}.$$

We know this is the solution to our set of equations above. That is, if we plug in these values into our original set of equations, we find that they evaluate to exactly the values we want.

$$2\left(\frac{26}{7}\right) - 3\left(-\frac{1}{7}\right) + \left(\frac{15}{7}\right) = 10$$

$$2\left(-\frac{1}{7}\right) + 2\left(\frac{15}{7}\right) = 4$$

$$2\left(\frac{26}{7}\right) \qquad\qquad - 3\left(\frac{15}{7}\right) = 1$$

Alternatively, we can also say that the following equation is true as well.

$$\begin{pmatrix} 2 & -3 & 1 \\ 0 & 2 & 2 \\ 2 & 0 & -3 \end{pmatrix} \begin{pmatrix} 26/7 \\ -1/7 \\ 15/7 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 1 \end{pmatrix}$$

In other words, the matrix

$$\begin{pmatrix} 2 & -3 & 1 \\ 0 & 2 & 2 \\ 2 & 0 & -3 \end{pmatrix}$$

sends the vector

$$\begin{pmatrix} 26/7 \\ -1/7 \\ 15/7 \end{pmatrix}$$

to the vector

$$\begin{pmatrix} 10 \\ 4 \\ 1 \end{pmatrix}.$$

Although we introduced matrices as a table of numbers, it's better to think of a matrix as a map between vectors. When we multiply a vector with a matrix[6], we get another vector.

9.1. **How Your Calculator Fits Curves to Data.** How does your calculator come up with these equations? How does it know which line or other function *best* fits the data?

It does this using matrices! Essentially, it tries to solve a problem very similar to the problem we solve with the system of linear equations. Unfortunately, it's not *exactly* the same because a few things go wrong. Let's see how this works and what goes wrong with an example.

To begin this process, you need two things:

(1) a data set, and
(2) a function to fit to the data set.

For our example, suppose you have the data below

[6]Remember that in order to multiply a matrix with a vector, we need the matrix's number of columns to be equal to the vector's number of rows.

x	**y**
1	3
2	20.9
3	43.3
4	68.6
5	96.3
6	125.9

and you suspect that the data should be fit to a function of the form below.

$$y = ax\ln(x) + b\ln(x) + cx + d$$

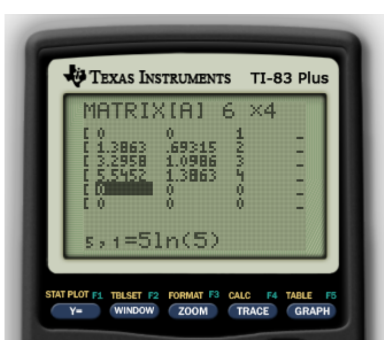
Notice that this function *does not* exist in your calculator. How do you do it? The trick is to change your perception. Rather than thinking of x and y as variables, we can think of the coefficients a, b, c and d as variables. If we use our data set, we can define a set of *linear* equations. We can plug in the values of x and y from our data and try to solve for the coefficients.

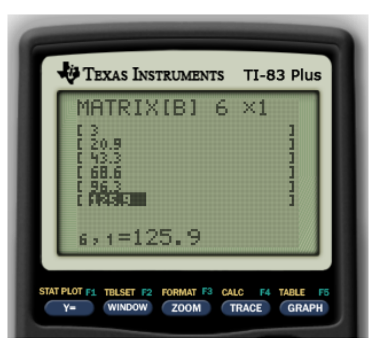
$$\begin{aligned}
a(1)\ln(1) + b\ln(1) + c(1) + d &= 3\\
a(2)\ln(2) + b\ln(2) + c(2) + d &= 20.9\\
a(3)\ln(3) + b\ln(3) + c(3) + d &= 43.3\\
a(4)\ln(4) + b\ln(4) + c(4) + d &= 68.6\\
a(5)\ln(5) + b\ln(5) + c(5) + d &= 96.3\\
a(6)\ln(6) + b\ln(6) + c(6) + d &= 125.9
\end{aligned}$$

This system of linear equations can be written as we have in the past.

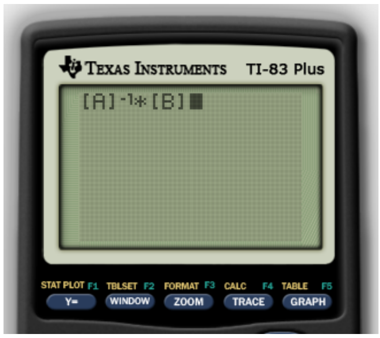
$$\begin{pmatrix} 0 & 0 & 1 & 1\\ 2\ln(2) & \ln(2) & 2 & 1\\ 3\ln(3) & \ln(3) & 3 & 1\\ 4\ln(4) & \ln(4) & 4 & 1\\ 5\ln(5) & \ln(5) & 5 & 1\\ 6\ln(6) & \ln(6) & 6 & 1 \end{pmatrix} \begin{pmatrix} a\\ b\\ c\\ d \end{pmatrix} = \begin{pmatrix} 3\\ 20.9\\ 43.3\\ 68.6\\ 96.3\\ 125.9 \end{pmatrix}$$

Suppose we attempt to solve this equation as we have other systems of linear equations. Remember that we access the `MATRIX` menu by pressing `2nd` and x^{-1}. Under the `EDIT` menu, change the entries of matrices `[A]` and `[B]`.





Next, we quit by pressing `2nd` and `MODE`. Then we type in the equation by returning to the `MATRIX` menu and selecting matrices `[A]` and `[B]` under the `NAMES` menu. If we enter our normal equation, we get an error message.



Why is there an error message? Essentially, it's because we have too many equations for the number of unknowns we have.

For example, suppose we have two equations with one variable:

$$2x = 4$$
$$3x = 6.1$$

What is the solution? It has no solution! The first equation implies that $x = 2$ but the second equation implies that $x = 61/30 \approx 2.03333$. Because x cannot have two different values, the system has no true solution.

The same problem is occurring with our data. we only have 4 variables to determine, but we have too many equations. When we graph our best-fit curves, the curve doesn't go through *all* points. It simply needs to follow the pattern of the data and minimize the data.

Because we cannot find a true solution, we can approximate our solution with a different equation:

$$\vec{x} = (A^T A)^{-1}(A^T \vec{b}) \tag{9.1}$$

This equation is a *projection*. When we learn about projects for vectors, we will return to this equation and explain why it is a projection. For now, we will simply use the equation and take for granted what it achieves.

What does the notation mean? A^T is called the *transpose* of A. A transpose reorganizes a matrix so that its columns are rows and vice versa.

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 2\ln(2) & \ln(2) & 2 & 1 \\ 3\ln(3) & \ln(3) & 3 & 1 \\ 4\ln(4) & \ln(4) & 4 & 1 \\ 5\ln(5) & \ln(5) & 5 & 1 \\ 6\ln(6) & \ln(6) & 6 & 1 \end{pmatrix} \qquad \text{means}$$

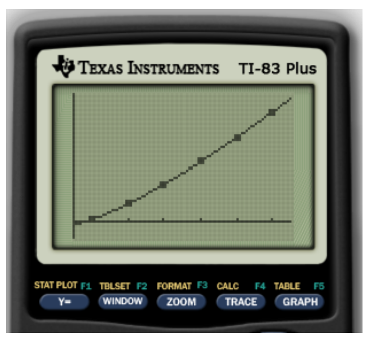
$$A^T = \begin{pmatrix} 0 & 2\ln(2) & 3\ln(3) & 4\ln(4) & 5\ln(5) & 6\ln(6) \\ 0 & \ln(2) & \ln(3) & \ln(4) & \ln(5) & \ln(6) \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Now, let us use our calculator to find the solution to this equation. You'll need to access the transpose by accessing the `MATRIX` menu. Under `MATH` select transpose, which is indicated by a little "T."

According to our calculation, our best-fit for the data with the given equation is

$$y = 9.898x\ln(x) + 2.726\ln(x) + 2.320x + 0.676$$

To check our work, let's see how this function fits with our data.



9.2. **Multi-Variable Best-Fit Curves.** Now that we know how to fit data to a function, we can consider multiple independent variables. Suppose we have the data set

x	**y**	**z**
1	1	3
2	1	9
1	2	4
2	2	16
3	1	27
3	2	27
1	3	5
2	3	23
3	3	39

and we wish to fit the data to the function

$$z = ay\ln(x) + bxy + cy + d.$$

Then we treat the coefficients a, b, c, and d as variables by plugging in data points. We get the following set of equations:

$$\begin{aligned}
a(1)\ln(1) + b(1)(1) + c(1) + d =& 3\\
a(1)\ln(2) + b(2)(1) + c(1) + d =& 9\\
a(2)\ln(1) + b(1)(2) + c(2) + d =& 4\\
a(2)\ln(2) + b(2)(2) + c(2) + d =& 16\\
a(1)\ln(3) + b(1)(1) + c(1) + d =& 27\\
a(2)\ln(3) + b(2)(3) + c(2) + d =& 27\\
a(3)\ln(1) + b(1)(3) + c(3) + d =& 5\\
a(3)\ln(2) + b(2)(3) + c(3) + d =& 23\\
a(3)\ln(3) + b(3)(3) + c(3) + d =& 39
\end{aligned}$$

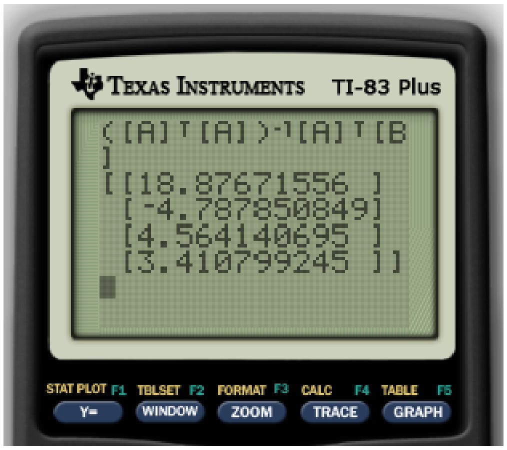
Notice that the only difference is plugging in more than one variable! We simply plug in the values according to the format $z = ay\ln(x) + bxy + cy + d$.

Then our matrix expression is

$$\begin{pmatrix} 0 & 1 & 1 & 1\\ \ln(2) & 2 & 1 & 1\\ 0 & 2 & 2 & 1\\ 2\ln(2) & 4 & 2 & 1\\ \ln(3) & 1 & 1 & 1\\ 2\ln(3) & 6 & 2 & 1\\ 0 & 3 & 3 & 1\\ 3\ln(2) & 6 & 3 & 1\\ 3\ln(3) & 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} a\\ b\\ c\\ d \end{pmatrix} = \begin{pmatrix} 3\\ 9\\ 4\\ 16\\ 27\\ 27\\ 5\\ 23\\ 39 \end{pmatrix}$$

Finally, we enter this information into our calculator and solve $\vec{x} = (A^TA)^{-1}(A^T\vec{b})$.

When we do, we get the solution



That means the equation that best fits the data is

$$z = 18.9y\ln(x) - 4.6xy + 4.6y + 3.4.$$

Unfortunately, our graphing calculator cannot graph functions in higher dimensions. So we cannot compare our graph with our data.

Summary of Ideas: Lecture 9

- Matrices are functions that take in a vector and produce a vector.
- We can find a best-fit curve by setting it up as a system of linear equations.
- Generally, we cannot solve this system. It often has more equations than unknown variables.
- To solve, we use an equation based on projections. It is

$$\vec{x} = (A^T A)^{-1}(A^T \vec{b})$$

- To fit data with more than one independent variable, we can use the *same* procedure.

PENN STATE UNIVERSITY, UNIVERSITY PARK, PA 16802