

LINEAR ALGEBRA

STRANG, SECTION 4.3

1. THE ASSIGNMENT

- Read section 4.3 of Strang (pages 218-226).
- Read the following.
- Complete exercises 1 - 11 of section 4.3 in Strang.
- Prepare the items below for presentation.

2. LEAST SQUARES APPROXIMATIONS

(It is probably best to read this after you read the section in Strang.)

2.1. Some Perspective. Scientific problems often come down to something as simple as this: make a bunch of observations, and then try to fit those observations with some sort of model for greater understanding.

But data found in scientific problems is often noisy, or infected with error in some way. This leads researchers to gather *more* data so that chance variations and small errors might get smoothed out. How might we fit a curve to a lot of data? Lots of data points means that we likely have too many points to have a curve of our specified model type actually hit all of those points. For example, fitting a line to five points is already problematic: any two points gives us a line, and there is no reason to believe that the other three points will all sit on that line.

If we set the problem up as a system of equations, things go like this: We have a bunch of data set up as input-output pairs $\{(a_i, y_i)\}$; we are looking for a function f which has a specified type (linear, quadratic, exponential, ...) that passes through those points.

- Each data point leads us to an equation $f(a_i) = y_i$.
- The modelling function f has some parameters in it, and we want to find the best value of those parameters so that the curve “fits” the data well. These parameters are the unknowns in our equations.

This is generally a challenging problem. The method of least squares is a technique for solving it when the resulting equations make a linear system.

2.2. Some History. Gauss discovered the technique described in this section in the late 1790's. In 1801 he used it to help astronomers calculate the orbit of the newly discovered asteroid Ceres, and thus find it after it re-emerged from behind the sun.

See how the pattern fits? Several weeks worth of data about the position of Ceres was known, but it surely had measurement errors in it. Since the time of Kepler (Newton), we have known that the motion of the asteroid must be an ellipse. This is a simple equation with only a few parameters (the coefficients of the equation defining the ellipse). So, the question confronting Gauss was this: find the ellipse which best fits the data.

But plugging all the data into the correct model shape (a conic!) leads to a rather large system of linear equations where the unknowns are the coefficients we seek.

2.3. So, what is really happening here? In the end, we get a system of the form $Ax = y$. Here A is an $m \times n$ matrix and y is an n -vector, where m is the number of equations and n is the number of parameters we must find. Typically, m is much larger than n , so the matrix A is tall and skinny.

So the system likely has no solution. Instead, we will find the orthogonal projection \hat{y} of y onto the column space $\text{col}(A)$ of A , and then solve $Ax = \hat{y}$. That's the secret. Since we have already mastered projections, this is no big deal.

3. SAGE INSTRUCTIONS

I have made a Sage worksheet file with some basic commands that you might find useful in investigating matrices. The file is called `section4_3.sagews`.

4. QUESTIONS FOR SECTION 4.3

The best thing you can do to understand this is work some examples. Do Strang 1 - 11 from section 4.3. We will present these:

Exercise 127. Exercise 8 from section 4.3 of Strang.

Exercise 128. Exercise 9 from section 4.3 of Strang.

Exercise 129. Exercise 10 from section 4.3 of Strang.

Exercise 130. Exercise 11 from section 4.3 of Strang.