Math 343 Lab 7: Line and Curve Fitting

Objective

In this lab, we explore another use of linear algebra in statistics. Specifically, we discuss the notion of least squares as a way to fit lines and curves to data.

Hooke's Law

It is well known that the displacement of a spring is proportional to the force acting upon it, that is, $F = kx$. The proportionality constant k is called Hooke's spring constant. Consider a laboratory experiment where different loads are placed on a spring and the displacement is measured and recorded in the table below:

x (cm)	F (dyne)
1.04	3.11
2.03	6.01
2.95	9.07
3.92	11.99
5.06	15.02
6.00	17.91
7.07	21.12

To find the spring constant k, we simply need to solve the following linear system

$$\begin{pmatrix} 1.04 \\ 2.03 \\ 2.95 \\ 3.92 \\ 5.06 \\ 6.00 \\ 7.07 \end{pmatrix} \begin{pmatrix} k \end{pmatrix} = \begin{pmatrix} 3.11 \\ 6.01 \\ 9.07 \\ 11.99 \\ 15.02 \\ 17.91 \\ 21.12 \end{pmatrix}.$$

However, there is no solution to this system because it is overdetermined. Instead, we seek the "best" k that fits the data.

Least Squares

In class, you will soon discuss the least squares problem, which finds the "best" solution to the (usually overdetermined) linear system

$$Ax = b.$$

The main result goes as follows:

Theorem: Consider the linear system $Ax = b$, where A is an $m \times n$ matrix of rank n. Then $A^T A$ is nonsingular, and the least squares solution of the system is $\widehat{x} = (A^T A)^{-1} A^T b$.

In other words, even though an overdetermined linear system doesn't generally have a solution, there is a vector $\widehat{x}$ that minimizes the error between Ax and b. This is called the least squares solution.

As an example, lets revisit our spring problem. Although it doesn't have a solution, we can find the least squares solution by computing the following in Matlab:

```
>> A = [1.04;2.03;2.95;3.92;5.06;6.00;7.07];

>> b = [3.11;6.01;9.07;11.99;15.02;17.91;21.12];

>> k = inv(A'*A)*A'*b
```

Hence, we find the spring constant to be $k = 2.9957$. We plot the data against the best fit as follows:

```
>> x0 = linspace(0,8,100);

>> y0 = k*x0;

>> plot(A,b,'*',x0,y0)
```

See Figure 1 to see how well the line fits the data.

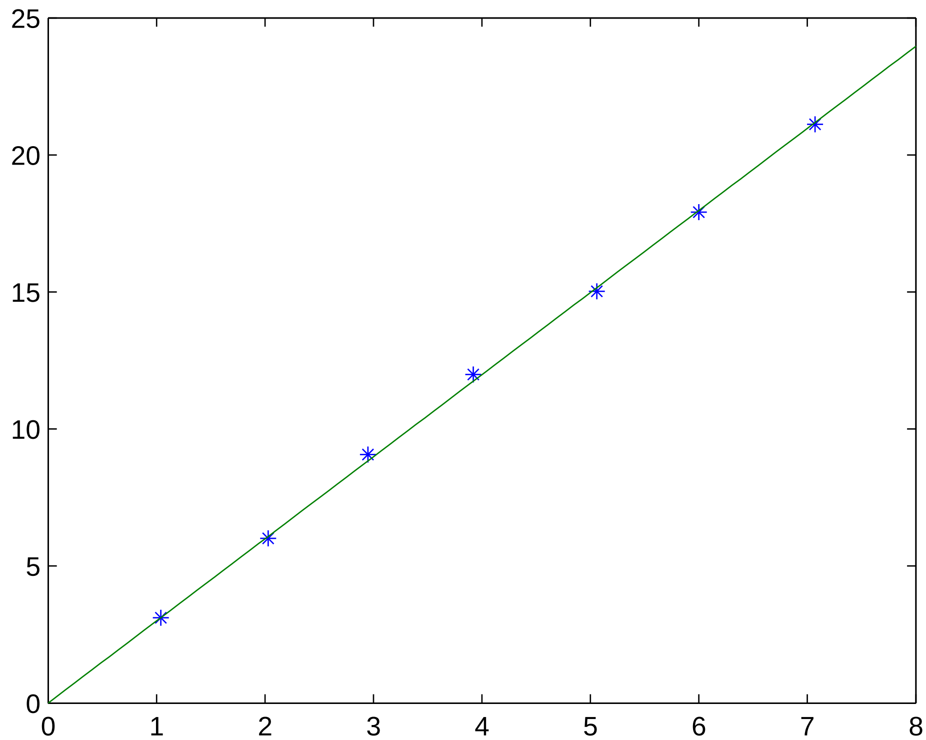

Figure 1: The graph of the spring data together with its linear fit

General Line Fitting

Suppose that we wish to fit a general line, that is $y = mx + b$, to the data set $\{(x_k, y_k)\}_{k=1}^n$. Assume that the line does not cross through the origin, as in the previous example. Then we seek both a slope and a y-intercept. In this case, we set up the following linear system $Ax = b$, or more precisely

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}.$$

Note that A has rank 2 as long as not all of the x_k values are the same. Hence, the least squares solution will fit the best line for this data.

Fitting data to a circle

Recall that the equation of a circle, with radius r centered at (c_1, c_2), is given by

$$(x - c_1)^2 + (y - c_2)^2 = r^2. \qquad (1)$$

Suppose we are given a set of data points closely forming a circle $\{(x_i, y_i)\}_{i=1}^{n}$. The "best" fit is found via least squares by expanding (1) to get

$$2c_1 x + 2c_2 y + c_3 = x^2 + y^2,$$

where $c_3 = r^2 - c_1^2 - c_2^2$. Then we can write the linear system $Ax = b$ as

$$\begin{pmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots \\ 2x_n & 2y_n & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ \vdots \\ x_n^2 + y_n^2 \end{pmatrix},$$

where the matrix A and the vector b are obtained by the given data and the unknown x contains the information about the center and radius of the circle and is obtained by finding the least squares solution.

Example

In this section, we fit the following points to a circle:

$$(134, 76), (104, 146), (34, 176), (-36, 146),$$
$$(-66, 76), (-36, 5), (34, -24), (104, 5), (134, 76)$$

We enter them into Matlab as the 9×2 matrix:

```
>> P = [134 76; 104 146; 34 176; -36 146; -66 76; -36 5;
        34 -24; 104 5; 134 76]
```

Then we can separate the x and y coordinates by the commands `P(:,1)` and `P(:,2)`, respectively. Hence, we compute A and b by entering the following:

```
>> A = [2*P(:,1) 2*P(:,2) ones(9,1)]

>> b = [P(:,1).^2 + P(:,2).^2];
```

Hence, we get the least squares solution

```
>> x = inv(A'*A)*A'*b
```

Then we find c_1, c_2, and r by:

```
>> c1 = x(1)

>> c2 = x(2)

>> c3 = x(3)

>> r = sqrt(c1^2 + c2^2 + c3)
```

We plot this by executing

```
>> theta = linspace(0,2*pi,200);

>> plot(r*cos(theta)+c1,r*sin(theta)+c2),'-',P(:,1),P(:,2),'*')
```

Assignment

Problem 1. *Download the file lab7a.txt from the following link:*

`http://www.math.byu.edu/~jeffh/teaching/m343/labs/lab7a.txt`

You can load this datafile by typing

`load lab7a.txt`

Note that the data is available in the matrix `lab7a`*. This consists of two columns corresponding to the x and y values of a given data set. Use least squares to find the slope and y-intercept that best fits the data. Then plot the data points and the line on the same graph. Finish off the problem with a discussion of what you've learned.*

Problem 2. *Download the file lab7b.txt from the following link:*

`http://www.math.byu.edu/~jeffh/teaching/m343/labs/lab7b.txt`

You can load this datafile into Matlab by typing

```
load lab7b.txt
```

Note that the data is available in the matrix `lab7b`*. This consists of two columns corresponding to the x and y values of a given data set. Use least squares to find the center and radius of the circle that best fits the data. Then plot the data points and the circle on the same graph. Finish off the problem with a discussion of what you've learned.*