

Math 2270 - Computer Lab 3 : Least Squares Approximation

Dylan Zwick

Fall 2012

In this lab we're going to use Maple to calculate and plot the line that best fits a number of points.

1 Introduction

We'll start with a refresher. First, you should load the following packages into Maple:

with(LinearAlgebra) :

with(ArrayTools) :

Consider the three points $(0, 1)$, $(1, 2.1)$, $(2, 3.9)$. If these points lie on a line $y = mx + b$ then the matrix equation

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} 1 \\ 2.1 \\ 3.9 \end{pmatrix}$$

is satisfied. These three points do not, however, lie on a line, and so there is no vector $\begin{pmatrix} b \\ m \end{pmatrix}$ satisfying the above equation. So, we want to

find the next best thing. That is to say, we want to find the projection of $\mathbf{b} = \begin{pmatrix} 1 \\ 2.1 \\ 3.9 \end{pmatrix}$ onto the column space of $\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$.

To do this we want to find the vector $\begin{pmatrix} b \\ m \end{pmatrix}$ that satisfies

$$A^T A \begin{pmatrix} b \\ m \end{pmatrix} = A^T \begin{pmatrix} 1 \\ 2.1 \\ 3.9 \end{pmatrix}.$$

As $A^T A$ is invertible this will be:

$$\begin{pmatrix} b \\ m \end{pmatrix} = (A^T A)^{-1} A^T \begin{pmatrix} 1 \\ 2.1 \\ 3.9 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 5.3 \\ 8.7 \end{pmatrix}.$$

So, the best fit line is $y = \frac{5.3}{6} + \frac{8.7}{6}x$.

We can plot these points and this best fit line in Maple with the following commands:

```
Points := [[0, 1], [1, 2.1], [2, 3.9]];
y := 5.3/6 + 8.7/6*x;
plot([Points, y], x = -1..3, color = [red, blue], style = [point, line]);
```

The Maple code for doing this is attached to the end of this handout.

2 Best Fit Line

We're going to do this again, but with a larger data set. Get the file *DataSet.txt* from the course webpage, and save it to your home directory (or somewhere else that you can find them).

In Maple enter

```

datadir := "pathtothedatafiles";
DataSet := readdata(cat(datadir," DataSet"),2);

```

Take a look at this set of points by plotting them:

```

plot(DataSet, x = 0..10, color = red, style = point);

```

The sequence of points looks approximately linear, and we want to find the line that best fits them. Right now, the data set is something called a "list of lists". We can make it a matrix by entering the command:

```

M := Matrix(DataSet);

```

M is not the matrix we usually call A . Column 1 of M is column 2 of A , and column 2 of A is what we usually call \mathbf{b} . We can pick out a column of M by multiplying it by a vector. If we multiply $M \cdot \langle 1, 0 \rangle$ we get the first column of M , and $M \cdot \langle 0, 1 \rangle$ gives us the second column. So, to get A and \mathbf{b} we want to do the following:

```

b := M .< 0, 1 >;
A :=< Vector(column, Size(M)(1), 1)|M .< 1, 0 >>;

```

We want to use A and \mathbf{b} to find the slope and y -intercept of the best fit line. Then we want to plot the points of this dataset along with the best fit line.

For this computer lab, you should download the data set from the course webpage, set up A and \mathbf{b} using the instructions above, and then use them to find the best fit line. Finally, you should plot this best fit line, along with the points of the dataset. You should hand in the Maple code used to do this.

```
# Dylan Zwick
# Lab 3 Maple for Handout
```

```
with(LinearAlgebra) :
with(ArrayTools) :
A := Matrix([[1, 0], [1, 1], [1, 2]]);
```

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad (1)$$

```
b := <1, 2.1, 3.9>;
```

$$\begin{bmatrix} 1 \\ 2.1 \\ 3.9 \end{bmatrix} \quad (2)$$

```
T := Transpose(A);
```

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad (3)$$

```
p := (T.A)-1.T.b;
```

$$\begin{bmatrix} 0.88333333333333304 \\ 1.44999999999999995 \end{bmatrix} \quad (4)$$

```
Points := [[0, 1], [1, 2.1], [2, 3.9]];
```

$$[[0, 1], [1, 2.1], [2, 3.9]] \quad (5)$$

```
y := p(1) + p(2)x;
```

$$0.88333333333333304 + 1.44999999999999995 x \quad (6)$$

```
plot([Points, y], x=-1..3, color=[red, blue], style=[point, line]);
```

