Numerical Analysis

Math 370 Fall 2002 — MWF 9:30am - 10:25pm

 — Fowler 127

Worksheet 20

SUMMARY Introduction to Approximation Theory
READING Recktenwald, Sec 9.1, pp 455–468

Approximation Theory

In approximation theory we have a set of m data points (x_k, y_k) for which we do not know what the actual function $f(x)$ which reflects the relationship between the input variable x and the output y.

Suppose we define the **deviation** as $\delta_k = P(x_k) - y_k$ and find a function $P(x)$ such that the total deviations between the function $P(x)$ and the data points (x_k, y_k) is minimised. There is more than one way to do this.

We define a function E which represents the total deviation we are trying to minimize and we want to find P which minimizes E, where E can have different forms.
Some Ways To Formulate E Are:

$$E_1 = \sum_{k=1}^{m} |P(x_k) - y_k|$$

OR

$$E_\infty = \max_{1 \le k \le m} |P(x_k) - y_k|$$

OR

$$E_2 = \sum_{k=1}^{m} [P(x_k) - y_k]^2$$

From statistics we know that if the data are **normally distributed** then the square error (E_2) is the best form of the error to use to measure how well $P(x)$ is approximating the unknown function $f(x)$ represented by the data y_k.

Linear Fit

If we assume that the polynomial we choose for $P(x)$ is linear so that $P(x) = ax + b$ then the problem of finding P becomes a minimization problem. If we consider E is a function of the parameters a and b what it the problem we have to solve?

Therefore

$$a = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2} \qquad b = \frac{\overline{x^2} \cdot \overline{y} - \overline{xy} \cdot \overline{x}}{\overline{x^2} - \overline{x}^2}$$

The line $P(x) = ax+b$ is known as the "least squares" line, or "line of best fit" or "regression line"

Example

Consider the following data. We shall compute the **line of best fit** for the data and sketch it on the graph paper on the next page. You may try drawing what looks like a line of best fit by "eye" in one ink color and seeing how that compares with the computed regression line in a different ink color.

x_i	y_i	x_i^2	$x_i y_i$	$P(x_i)$	$\lvert y_i - P(x_i)\rvert$
1	1.3				
2	3.5				
3	4.2				
4	5.0				
5	7.0				
6	8.8				
7	10.1				
8	12.5				
9	13.0				
10	15.6				

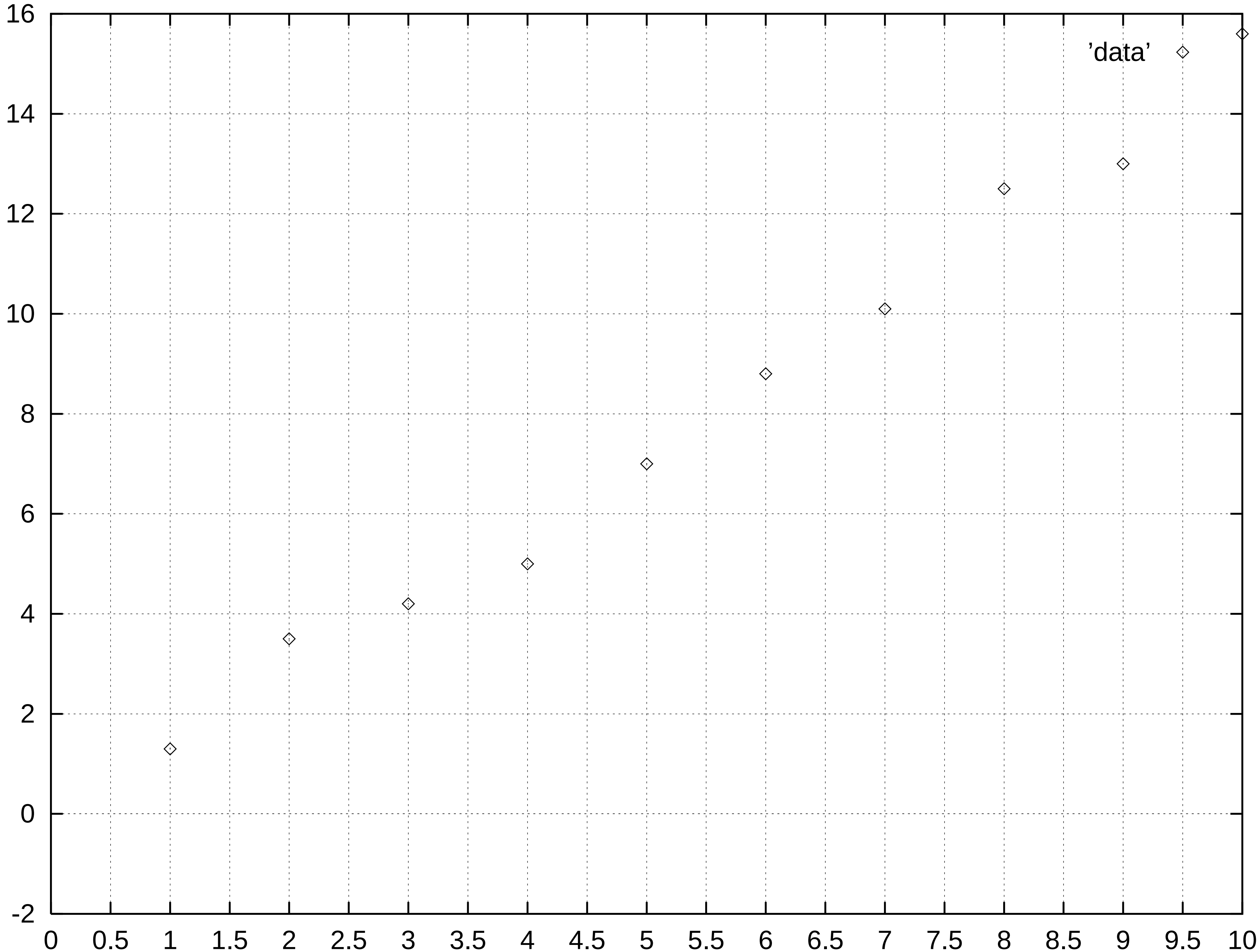
Using MATLAB to compute the line of best fit

Given a vector of inputs in **x** and **y** MATLAB will compute the slope and intercept of the line of best fit using the `linefit` command.

GROUPWORK

Use `linefit` to find the line of best fit for the above data.

Use the `plot` and `linspace` commands to plot the original data and the line of best fit on the same graph.



'data'
16
14
12
10
8
6
4
2
0
-2
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
8.5
9
9.5
10

Fitting Data to Nonlinear Functions: OR Making Nonlinear Relationships Appear Linear

Isn't there some way we could transform the equation $y = be^{ax}$ and $y = bx^a$ so that a linear relationship would appear? Then we could use our previously defined normal equations.

Think about introducing some new variables Y and X such that there is a linear relationship between Y and X even though y and x are non-linearly related.

A Harder One

How could you pick Y and X so that you could solve the normal equations and fit data to $y = \alpha x e^{\beta x}$?

For an example of what this would look like, try the command `xexpdemo(100)`, which is found in the "datafit" directory of the NMM.