Numerical Analysis

Math 370 MWF 2:30-3:25
Spring 1997 127 Fowler

(03.10.97) Approximation Theory (Continued)

In approximation theory, like with interpolation, we have a set of m data points (x_k, y_k) but this time we do not know that these values accurately reflect coordinates that the unknown function $f(x)$ passes through.
So 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 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.

Let's assume that the polynomial we choose for $P(x)$ is linear so that $P(x) = ax + b$ then the problem of finding P becomes ...

How do we find the minimum of this function? What are the variables in this function?

Let's do it:

The solution is:

Therefore

$$a = \qquad\qquad b =$$

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

Polynomial Of Best Fit

What if $P(x)$ is not a line but a polynomial of degree n? Let's assume that

$$P(x) = \sum_{k=0}^{n} a_k x^k, \qquad \text{where } n < m - 1$$

Again, $E = \sum_{k=1}^{m} [P(x_k) - y_k]^2$ which we want to minimize

This leads to $n + 1$ **normal equations** in the $n + 1$ unknowns $a_0, a_1, \ldots a_n$

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.

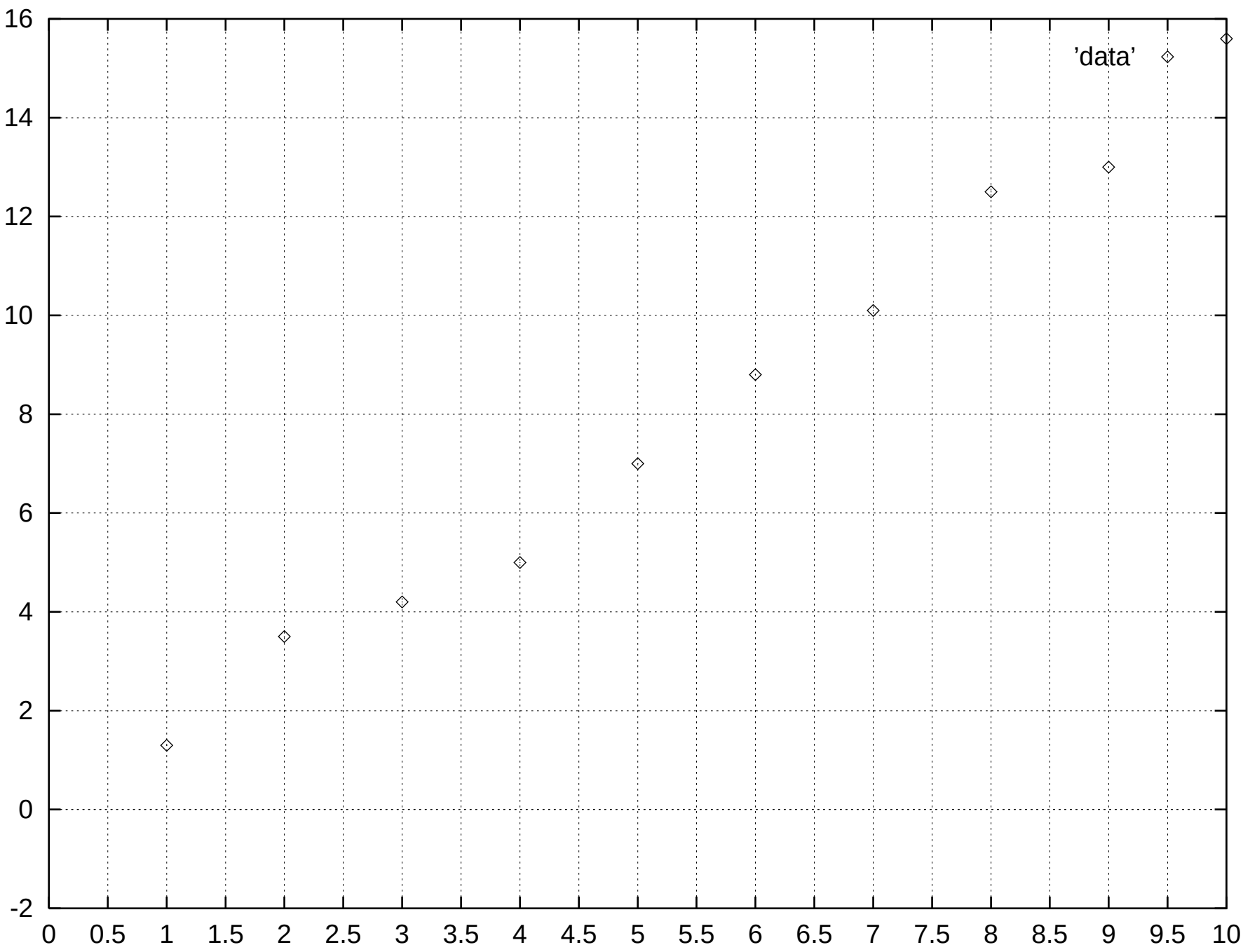
x_i	y_i	x_i^2	$x_i y_i$	$P(x_i)$	$\lvert y_i - P(x_i)\rvert$
1	1.3				
2	3.5				
3	4.2				
4	5.0				
5	7.0				
6	8.8				
7	10.1				
8	12.5				
9	13.0				
10	15.6				

Example

Let's try and compute the least squares polynomial of degree 2 for the following data:

i	1	2	3	4	5
x_i	0	0.25	0.5	0.75	1.00
y_i	1.0000	1.2840	1.6487	2.1170	2.7183

How many unknowns are you solving for? How many equations do you have?
Write down the equations:


'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