5) Curve fitting:

- On many occasions one has sets of ordered pairs of data $(x_1,...,x_n, y_1,...,y_n)$ which are related by a concrete function Y(X) e.g. some experimental data with a theoretical prediction

➤ suppose Y(X) is a linear function

$$Y = \alpha X + \beta$$

- Excel offers various ways to determine α and β

  i) SLOPE, INTERCEPT - functions
  based on the method of least square

$$\min = \sum_{i=1}^{n} [y_i - (\beta + \alpha x_i)]^2$$

SLOPE$(y_1,...,y_n,x_1,...,x_n,) \rightarrow \alpha$

INTERCEPT$(y_1,...,y_n,x_1,...,x_n,) \rightarrow \beta$

- How does Excel compute this? (see other courses for derivation)

· mean values: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

· slope: $\alpha = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) / \sum_{i=1}^{n} (x_i - \bar{x})^2$

· intercept: $\beta = \bar{y} - \alpha\bar{x}$

· regression coefficient:

$$r = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) / \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}$$

A good linear correlation between the x_i and y_i -values is $r \cong 1$.

With VBA we can write a code which does the same job, see Lab-session 5 of Part II.

ii) LINEST - function

this function is more sophisticated than the previous one

LINEST(y_1,...,y_n,x_1, ...,x_n,*constant*,*statistics*)

- if *constant* = TRUE or omitted the intercept is computed otherwise it is zero
- if *statistics* = TRUE the function returns regression statistic values with the output:

slope	**intercept**
standard error in the slope	**standard error in the intercept**
r-squared	**standard error in the y estimation**

- we restrict ourselves here to

51

- notice that LINEST is an array function, such that you have to prepare for an output bigger than one cell:
  · select a range for the output, e.g. 2×3 cells
  · type the function, e.g. =LINEST(.....)
  · complete with [Ctrl] + [Shift] + [Enter]

iii) adding a trendline

- this option also works for nonlinear, logarithmic, exponential ... correlations between the x- and y-values
  · choose an XY-chart with the subtype which has no line
  · right click any of the plotted points
  ⇒ Add Trendline windows opens
  · select the type of correlation, e.g. Linear, polynomial, ...
  · in Options decide if you want to add the computed equation the r-squared value etc on the chart

52

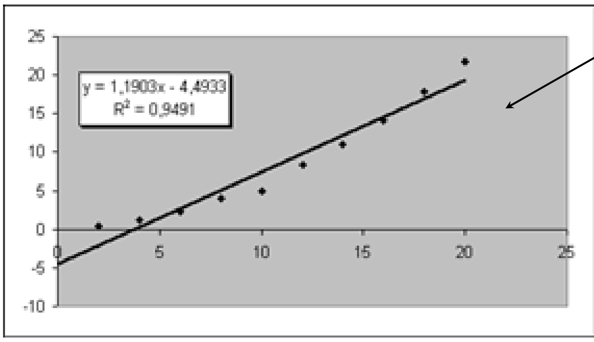
Example:
Consider the data:

2	0,4
4	1,2
6	2,3
8	4
10	5
12	8,3
14	11
16	14,1
18	17,9
20	21,8

assume linear correlation:

slope → 1.1903

intercept → -4,4933



with trendline adding

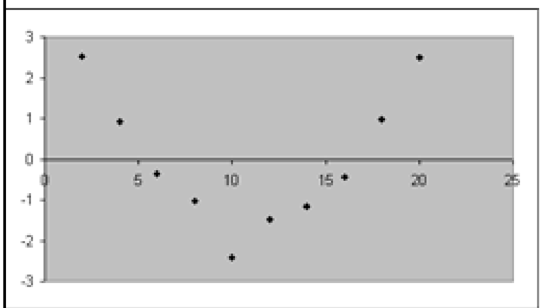
looks more or less linear?

53

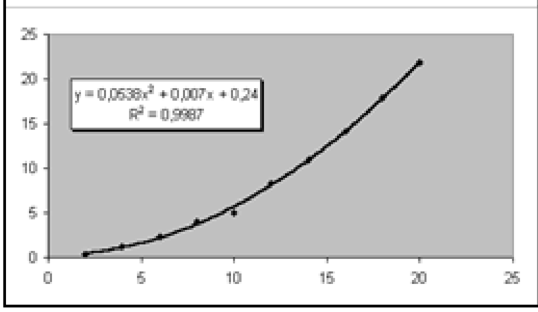
Compute the residuals, i.e. (the predicted values - the given ones):

$(1.1903\ x_i - 4.4933) - y_i \rightarrow$

2,512727
0,932121
-0,34848
-1,02909
-2,4097
-1,4903
-1,17091
-0,45152
0,967879
2,487273



not random!



quadratic fit is better!

54