Curve fitting

- On many occasions one has sets of ordered pairs of data $(x_1, y_1), \ldots, (x_n, y_n)$ 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

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

$\mathrm{SLOPE}(y_1, \ldots, y_n, x_1, \ldots, x_n) \rightarrow \alpha$

$\mathrm{INTERCEPT}(y_1, \ldots, y_n, x_1, \ldots, x_n) \rightarrow \beta$

- 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 $r \simeq 1$.

With VBA we can write a code which does the same job, see Lab-session 4 of Part II.

ii) LINEST - function

this function is more sophisticated than the previous ones

LINEST($y_1,\ldots,y_n, x_1,\ldots,x_n$, *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

- 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

	A	B	C	D	E	F	G
1	1	2					
2	2	2,3					
3	3	2,6					
4	4	3					
5	5	3,6					
6	6	8,5			1,280606	-0,63333	
7	7	9			0,135361	0,839895	
8	8	10,1			0,917952	1,22948	
9	9	11					
10	10	12					

slope (E6)

intercept (F6)

r-squared (E8)

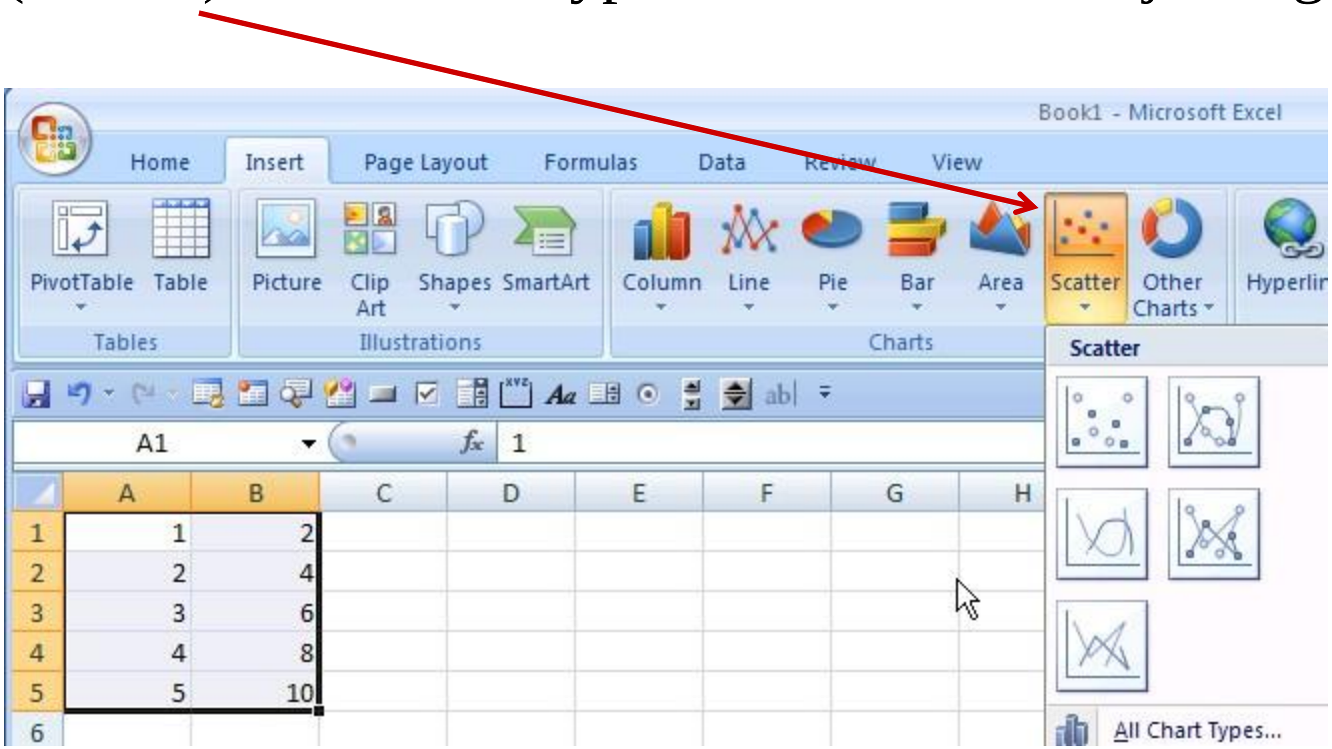
In the example we did =linest(B1:B10;A1:A10;true;true)

The value of r^2 is slightly away from 1, which shows that the points do not really fall into a line!

iii) adding a trendline

- First we need to have a set of points that we want to plot. Type the coordinates of the points that you want to plot. For example, the y-values in column B and the x-values in column A, as in the example before.

- Select the range containg the values you just entered and choose an XY-chart (Scatter) with the subtype which has no line joining the points

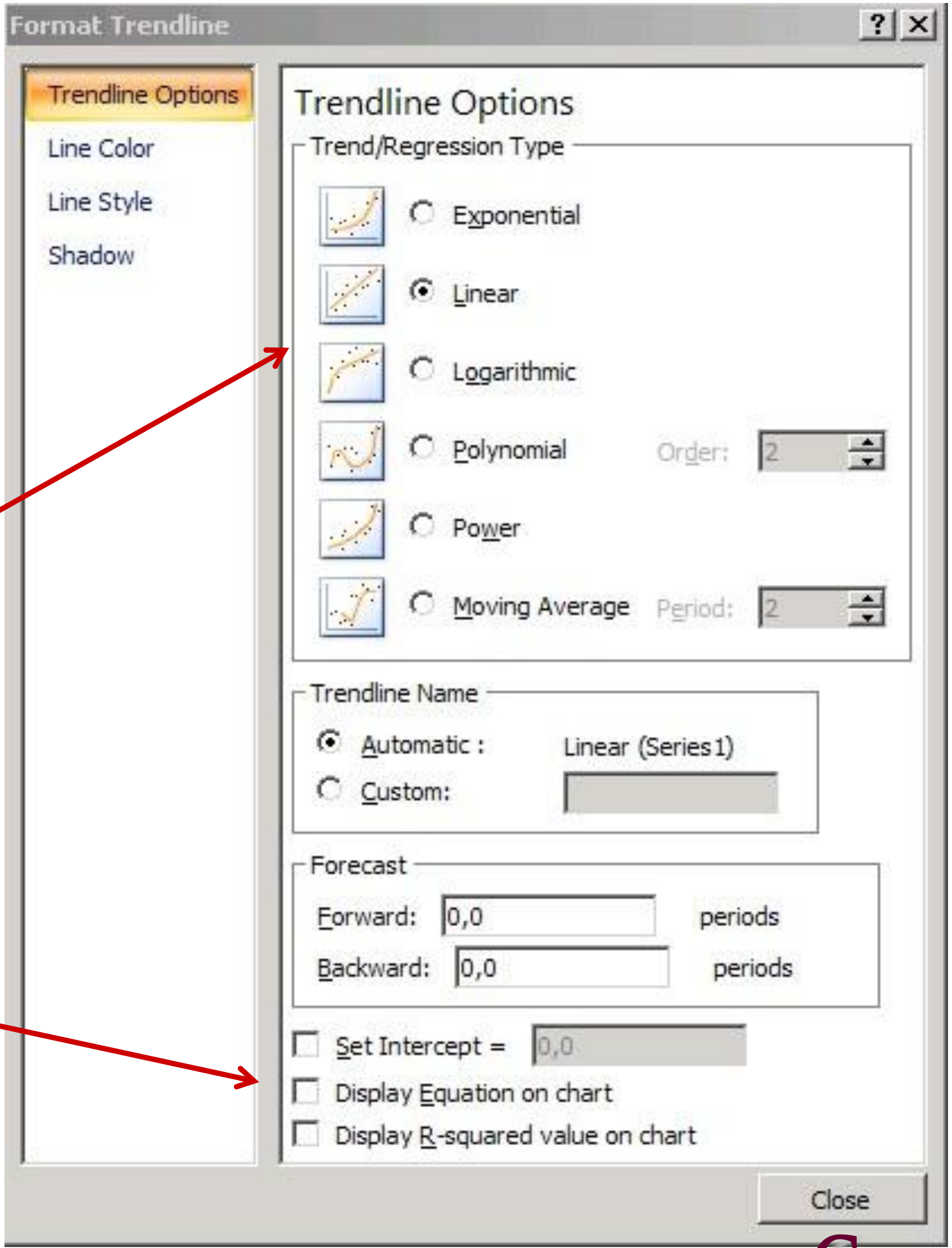


• right click on any of the plotted points

⇒ Add Trendline window opens

• select the type of correlation, e.g. Linear, polynomial, ...

• in Options decide if you want to add the computed equation or the r^2 value on the chart



Example:

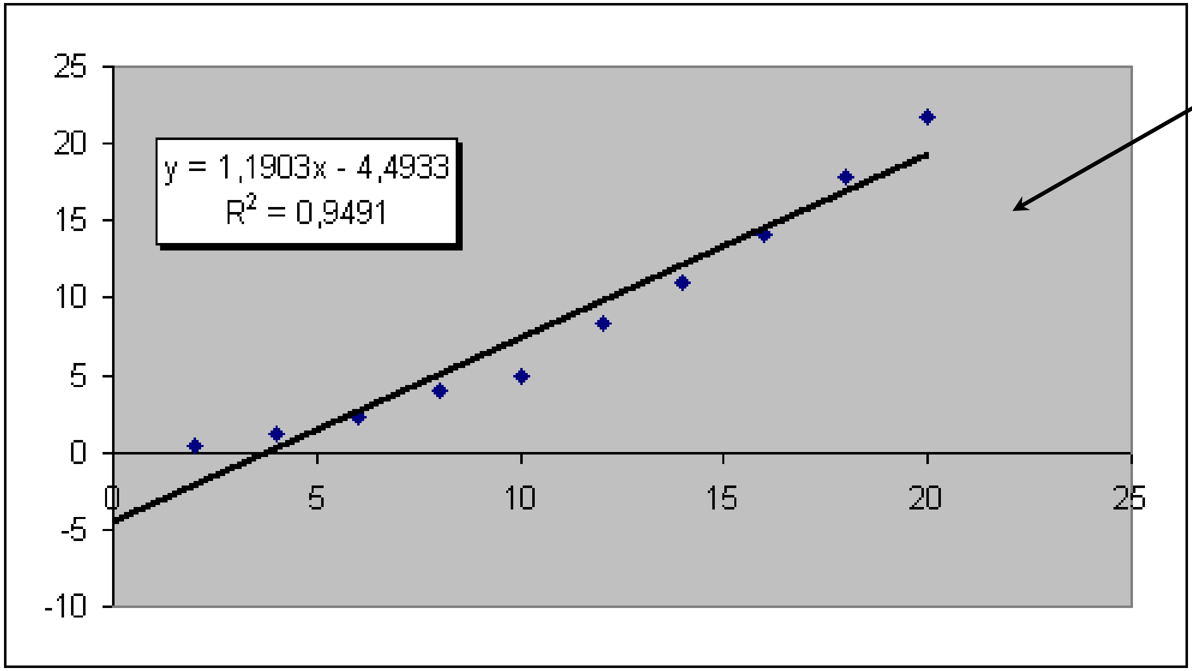
Consider the data:

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



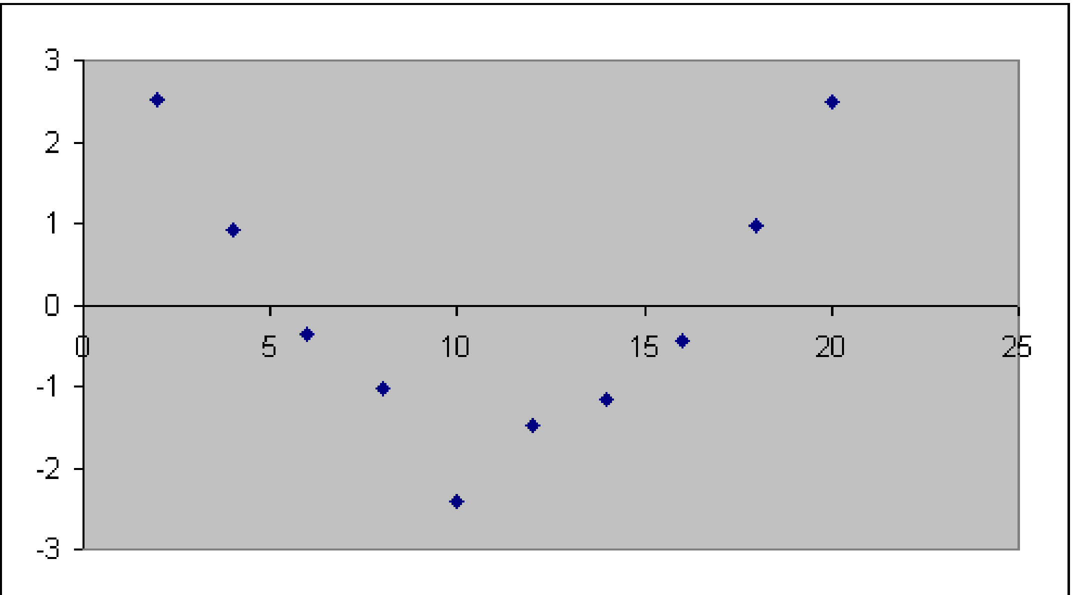
with trendline adding

looks more or less linear?

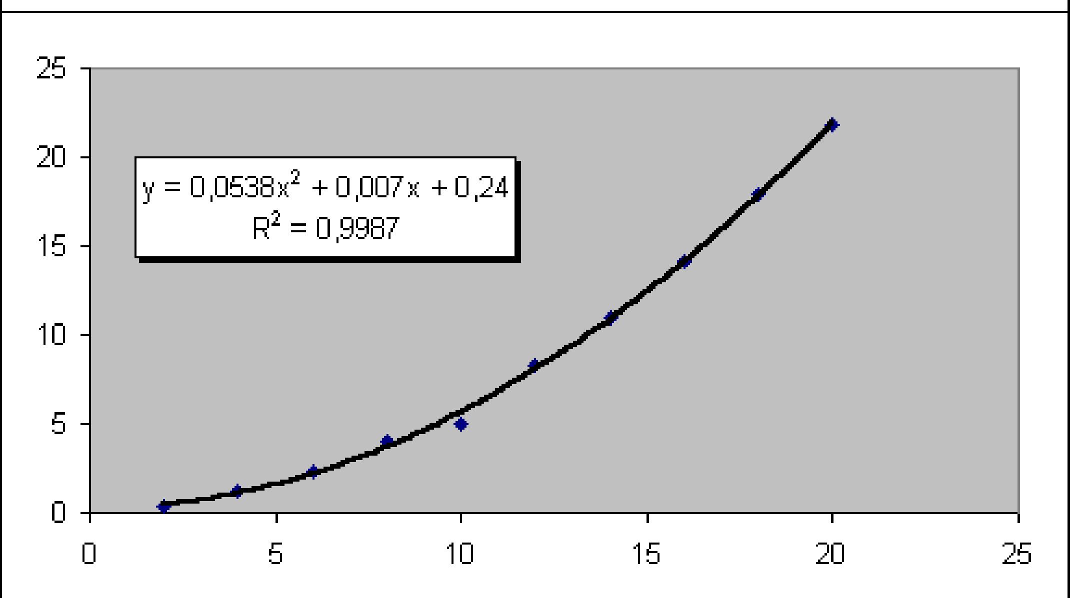
Compute the residuals, i.e. (the predicted values - the given ones):

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

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!

A simple VBA code that generates a set of points (x_i, $f(x_i)$)

Microsoft Visual Basic - Book1 - [Module1 (Code)]

(General)

```
Function f(x)
f = x ^ 2 + x - 3
End Function
Sub plott()
Range("a1").Select
For j = -5 To 5
ActiveCell.Offset(j + 5, 0).Value = j
ActiveCell.Offset(j + 5, 1).Value = f(j)
Next j
End Sub
```

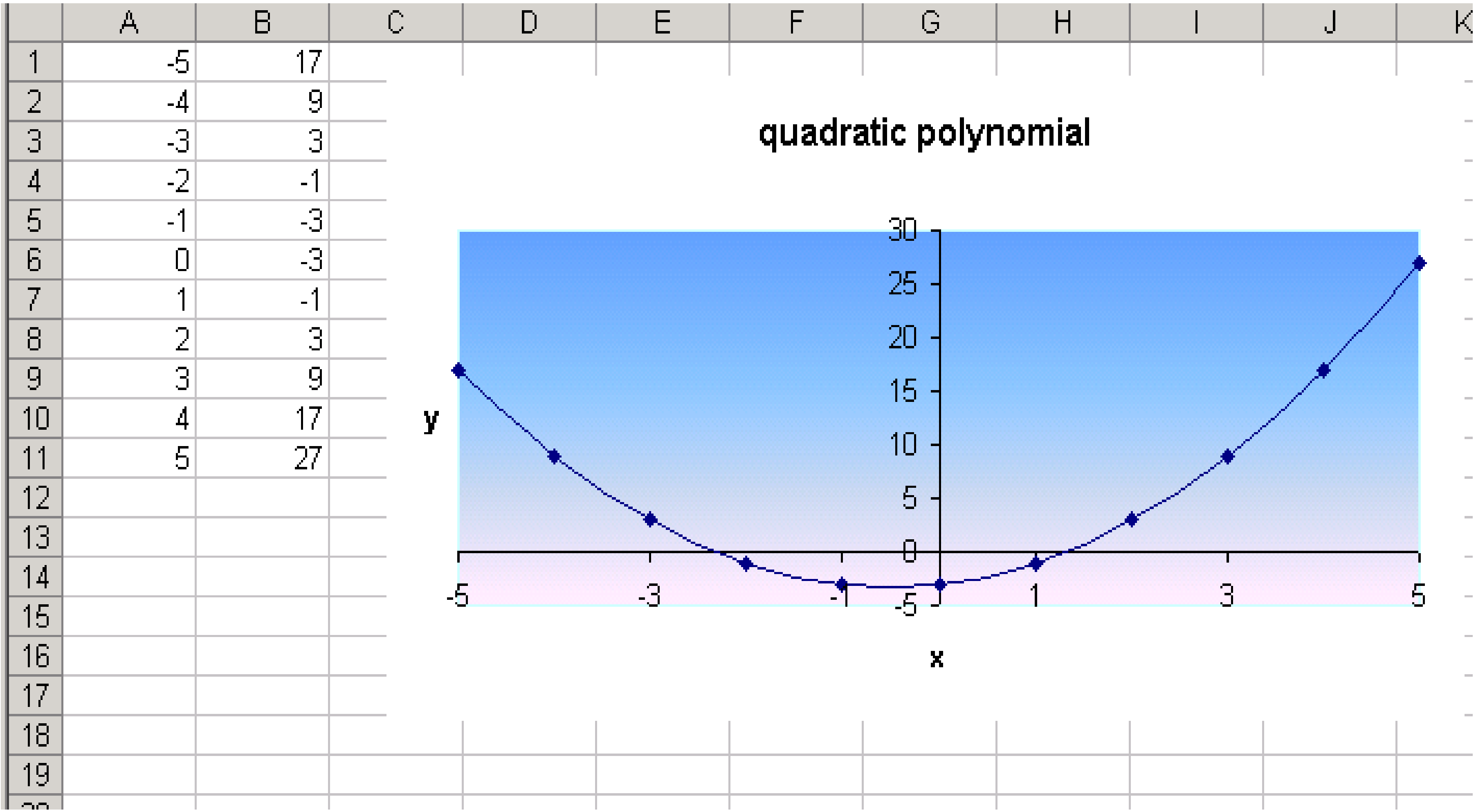
A1 is the active cell

	A	B
1	-5	17
2	-4	9
3	-3	3
4	-2	-1
5	-1	-3
6	0	-3
7	1	-1
8	2	3
9	3	9
10	4	17
11	5	27

The loop generates 11 pairs of points (xi, f(xi)) and writes them in columns 0 (A) and 1 (B)

When we run this code we obtain:

We can now plot the function as before:



	A	B
1	-5	17
2	-4	9
3	-3	3
4	-2	-1
5	-1	-3
6	0	-3
7	1	-1
8	2	3
9	3	9
10	4	17
11	5	27