# Exercise Sheet 6

## Exercise 21 Method of Least Squares/Regression

Determine a best fit line $y = a + bx$ (regression line) for the data set already considered in exercise 10, that is, for

| $x$ | 0 | 1 | 1 | 2 | 3 | 3 | 4 | 5 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y$ | 0 | 1 | 2 | 3 | 2 | 3 | 4 | 4 | 6 | 5 |

a) using the covariance and the variances/standard deviations
(see the lecture slides, section on correlation coefficients)

b) using the method of least squares/the system of normal equations!

Draw a diagram of the data points and the regression line!

## Exercise 22 Method of Least Squares/Regression

Determine a best fit parabola $y = a + bx + cx^2$ (regression parabola) for the data set $(x, y) = ((0, 0), (2, 1), (3, 2), (4, 4))$ with the method of least squares and draw this parabola!

## Exercise 23 Multilinear Regression

Determine a best fit plane $z = a + bx + cy$ for the following data set with the method of least squares: $(x, y, z) = ((0, 1, 0), (0, 4, 2), (2, 0, 1), (3, 1, 2), (2, 3, 3), (4, 4, 4))$.

## Exercise 24 Logistic Regression

The following table shows the number of American intercontinental ballistic missiles (ICBMs) in the years from 1960 to 1969:

| year, $x$ | 1960 | 1961 | 1962 | 1963 | 1964 | 1965 | 1966 | 1967 | 1968 | 1969 |
|---|---|---|---|---|---|---|---|---|---|---|
| number, $y$ | 18 | 63 | 294 | 424 | 834 | 854 | 904 | 1054 | 1054 | 1054 |

Find a best fit curve for this data set using logistic regression ($Y = 1060$)! Draw the original data and sketch the curve $y = \frac{1060}{1+e^{a+bx}}$!

**Additional Exercise** Exponential Regression

Radioactive substances decay according to the law $N(t) = N_0 e^{-\lambda t}$, where $t$ is the time, $\lambda$ a substance-specific decay parameter, $N_0$ the number of atoms of the substance at the beginning and $N(t)$ the number of remaining atoms at time point $t$. With the help of Geiger–Müller counter the following values $n$ were measured for the number of $\alpha$ particles that were emitted by a small amount of a radioactive substance up to different time points $t$:

| $t$ (in s) | 0 | 30 | 60 | 90 | 120 | 150 | 180 | 210 | 240 |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 0 | 306 | 552 | 655 | 768 | 863 | 901 | 919 | 956 |

Each counted $\alpha$ particle indicates that one atom of the radioactive substance decayed. Determine the half-life of the radioactive substance! What element is this substance?

Procedure: Find a best fit curve $n = n_0(1 - e^{a+bt})$!
(Hint: You have to find a transformation that reduces the problem to the problem of finding a best fit line (regression line); $n_0 = 1000$.) Although the value for $a$ may differ from zero with this approach, $-b$ may be seen as an approximation of the decay parameter $\lambda$, from which the half-life can easily be determined. The half-life of a substance is the time after which only half of the originally present atoms remain.