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

\[
\begin{array}{c|cccccccc}
  x & 0 & 1 & 1 & 2 & 3 & 4 & 5 & 6 \\
  y & 0 & 1 & 2 & 3 & 2 & 3 & 4 & 6 & 5 \\
\end{array}
\]

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:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>number, (y)</td>
<td>18</td>
<td>63</td>
<td>294</td>
<td>424</td>
<td>834</td>
<td>854</td>
<td>904</td>
<td>1054</td>
<td>1054</td>
<td>1054</td>
</tr>
</tbody>
</table>

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

Radioactive substances decay according to the law $N(t) = N_0e^{-\lambda t}$, 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 $t$:

<table>
<thead>
<tr>
<th>$t$ (in s)</th>
<th>0</th>
<th>30</th>
<th>60</th>
<th>90</th>
<th>120</th>
<th>150</th>
<th>180</th>
<th>210</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>0</td>
<td>306</td>
<td>552</td>
<td>655</td>
<td>768</td>
<td>863</td>
<td>901</td>
<td>919</td>
<td>956</td>
</tr>
</tbody>
</table>

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.