Time	Group	Submission in Moodle; Mails with subject: [SMD2023]
Th. 12:05–13:05	A	tristan.gradetzke@udo.edu and samuel.haefs@udo.edu
Fr. 09:00–10:00	B	lucas.witthaus@udo.edu and david.venker@udo.edu

Exercise 5 *Regularized Least Squares* **5 p.**

A colleague has measured a distribution. This distribution becomes part of a Monte Carlo simulation. To be able to work better with this distribution, you look for a suitable parameterization. You know that the distribution can be well described by a sixth-degree polynomial. However, the measurement is very noisy and your colleague was also only able to take eight pairs of values (x, y).

(a) Fit a sixth degree polynomial to the data in the file `ex_a.csv` using the least squares method. State the resulting coefficients and plot the fitted polynomial and data.

(b) Fit a sixth degree polynomial to the data in the file `ex_a.csv` using the least squares method and additionally use the regularization via the second derivative ($\mathit{\Gamma} = \sqrt{\lambda} CA$). For the regularization strength use $\lambda \in (0.1, 0.3, 0.7, 3, 10)$. State the resulting coefficients and plot the fitted polynomial and the data.

Your colleague makes the effort to produce 50 new measurements of the spectrum.

(c) Fit a sixth degree polynomial to the mean values of the data from the file `ex_c.csv` using the least squares method. Weight the calculated means with the uncertainty of the mean. Use these weights when fitting. Plot the fitted polynomial and the averaged data.

Exercise 6 *γ-Astronomy* **5 p.**

In a typical measurement in γ-astronomy, the telescope is pointed at a position (*on*-position) where a γ-ray source is suspected. In the subsequent measurement, N_{on} events are recorded over a period of time t_{on}. The measured events N_{on} contain both background and signal photons. To determine how many background photons are present, measurements are also taken at another position without a known source (*off*-position). In this measurement, N_{off} photons are measured in a time t_{off}.

In order to decide whether there is a source at the *on* position, a likelihood ratio test will be used to test whether a significant excess of photons over the background expectation was measured for the *on* position (not yet here, not until chapter *Testing*).

The aim of this task is to prepare the correct likelihood function for the likelihood ratio test.

Use these expressions to complete the task:

- $\alpha = \frac{t_{\text{on}}}{t_{\text{off}}}$: Quotient of the different measuring times
- $b = \langle N_{\text{off}} \rangle$: Unknown expected value for the number of background photons during the measurement time t_{off}
- s: Unknown expected value for the number of signal photons during the measurement time t_{on} from the γ-ray source. Not to be confused with the whole expectation for the *on* position.

(a) What is the expected value $\langle N_{\text{on}} \rangle$ expressed by s, b and α?

(b) Which probability distributions do N_{on} and N_{off} follow?
Hint: The counted events arrive at the detector independently of each other.

(c) What is the likelihood function $\mathcal{L}(b, s)$ for the parameters b and s?

(d) Which values for $\hat{b}$ and $\hat{s}$ maximise $\mathcal{L}$?
Hint: Use the negative log-likelihood function, the calculation becomes easier.

(e) Calculate the covariance matrix of $\hat{b}$ and $\hat{s}$. How is the covariance matrix related to the likelihood? Is this type of error calculation accurate?