Problems on *statistics and data analysis (MVComp2)*

lecturer: Björn Malte Schäfer
head tutor: Alessio Spurio Mancini
summer term 2016

Problem sheet 9

To be handed in during the exercise group on 04.July.2016 (independence day)

1. ***python exercise*** *(10 points)*

Set up a Metropolis-Hastings-sampler to explore the likelihood of a fit of the model

$$y(x) = \exp(-\alpha x) \tag{I}$$

to data $y_i = y(x_i) + \Delta y$ in the interval $x \in [-1 \ldots +1]$ at 21 equidistant points with $\alpha = 1$. Δy follows from a Gaussian noise process,

$$p(\Delta y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\Delta y^2}{2\sigma^2}\right), \tag{II}$$

with $\sigma = 0.1$. Show that the likelihood $\mathcal{L}(\alpha)$ is non-Gaussian by providing a plot, but becomes more Gaussian if

(a) the range of x-values, or

(b) the noise amplitude σ, or

(c) the value of α

is decreased. Why is this the case? Finally, please compare your MCMC-sampled likelihood $\mathcal{L}(\alpha)$ with a direct evaluation of $\exp(-\chi^2/2)$ on a grid in α.

2. ***errors in a nonlinear fit*** *(10 points)*

Please compute the Fisher-matrix $F_{\alpha\alpha}$ in the setup of the first exercise.

3. ***bivariate Cauchy-distribution*** *(10 points)*

The bivariate, uncorrelated Cauchy-distribution $p(x, y)$ is given by

$$p(x, y) = \frac{1}{(1 + x^2 + y^2)^{3/2}}. \tag{III}$$

(a) Is it possible to quantify the curvature matrix C,

$$C = \begin{pmatrix} \partial^2_{xx} \ln p & \partial^2_{xy} \ln p \\ \partial^2_{xy} \ln p & \partial^x_{yy} \ln p \end{pmatrix} \tag{IV}$$

and set it into relation with suitably defined confidence intervals?

(b) Please construct a quantification of statistical independence between x and y.

4. *volume and surface of a sphere in* n *dimensions* *(10 points)*

One can use the Gaussian distribution to derive in a straightforward way the volume and the surface of a sphere in n dimensions: The volume V_n is given by

$$V_n(R) = \prod_{i=1}^{n} \int_{\sum_{i=1}^{n} x_i^2 \leq R^2} \mathrm{d}x_i = C_n R^n, \tag{V}$$

which needs to be proportional to R^n on dimensional arguments, with a constant of proportionality C_n. The relation between surface $S_n(R)$ and volume $V_n(R)$ follows from

$$S_{n-1}(R) = \frac{\mathrm{d}}{\mathrm{d}R} V_n(R) = n C_n R^{n-1}, \tag{VI}$$

such that there must be this relation between the volume elements

$$\prod_{i=1}^{n} \mathrm{d}x_i = r^{n-1}\, \mathrm{d}r\, \mathrm{d}\Omega_{n-1}, \tag{VII}$$

in cartesian and in spherical coordinates. Please show that

$$V_n(R) = \int_0^R \mathrm{d}r\, S_{n-1}(r) \tag{VIII}$$

and from that the relation eqn. (VII) as well as the relation

$$\int \mathrm{d}\Omega_{n-1} = n C_n. \tag{IX}$$

Then, starting from integrating an n-dimensional Gaussian,

$$\prod_{i=1}^{n} \int \mathrm{d}x_i\, \exp\left(-\frac{x_i^2}{2}\right) = \int \mathrm{d}r\, r^{n-1} \int \mathrm{d}\Omega_{n-1}\, \exp\left(-\frac{r^2}{2}\right), \tag{X}$$

in cartesian as well as in spherical coordinates with $r^2 = \sum_{i=1}^{n} x_i^2$, please show that

$$C_n = \frac{\pi^{n/2}}{\Gamma(1 + n/2)}. \tag{XI}$$

and continue to demonstrate the final result

$$V_n(R) = \frac{\pi^{n/2}}{\Gamma(1 + n/2)} R^n \quad \text{and} \quad S_{n-1}(R) = \frac{2\pi^{n/2}}{\Gamma(n/2)} R^{n-1}. \tag{XII}$$

Do you recover familiar results for $n = 1, 2, 3$? Both $V_n(R)$ and $S_{n-1}(R)$ vanish for $n \to \infty$ at fixed R. Please convince yourself of this by plotting $V_n(R)$ and $S_{n-1}(R)$ for $R = 1$ and $n = 1 \ldots 10$. Can this be true?