12. EXERCISE SHEET *Bayesian inference and data assimilation*

Due date: Saturday, 3rd of July 2021, 23:59 PM (CEST)
Please refer to Assignment Submission Guideline on Moodle.

Exercise 1. Write a code to implement the SIR filter for Example 4.13 of the textbook, i.e. for the stochastically perturbed Lorenz-63 model,

$$\begin{aligned} x_1^{n+1} &= x_1^n + \delta t \sigma (x_2^n - x_1^n) + \sqrt{\delta t} \xi_1^n \\ x_2^{n+1} &= x_2^n + \delta t (x_1^n (\rho - x_3^n) - x_2^n) + \sqrt{\delta t} \xi_2^n \\ x_3^{n+1} &= x_3^n + \delta t (x_1^n x_2^n - \beta x_3^n) + \sqrt{\delta t} \xi_3^n . \end{aligned}$$

Use the same model to generate a reference trajectory $\{z_{ref}^n\}_{n\geq 0}$ from the initial condition $z_{ref}^0 = [-0.587 \quad -0.563 \quad 16.87]$. Use this reference trajectory to generate a total of $N_{obs} = 10\ 000$ observations of the x_1-component in time intervals of $\delta t = 0.001$ and $\delta t_{out} = 0.01$ and with measurement error variance $R = \frac{1}{15}$. The initial PDF is Gaussian with mean $z = z_{ref}^0$ and diagonal covariance matrix $P = \sigma^2 I$ with $\sigma = 0.1$. The sample size varies between $M = 20, 50$ and $M = 100$.

(i) Make three plots, one for each component. In each of the plots, plot the reference trajectory and the analysis means for all values of M. The x axis should vary over $t \in [0, 5]$.

(ii) Make a plot for the ensemble variance in the x-coordinate for $t \in [0, 5]$.

(iii) Make a plot for the effective sample size for $t \in [0, 5]$.

(iv) Calculate the RMSE for each M and each coordinate. Make one plot for each coordinate where you plot the ensemble size M (x-axis) against the RMSE for that coordinate.

Exercise 2. We have mentioned that the accuracy of importance sampling for Bayesian inference depends on the effective sample size

$$M_{\text{effective}} := \frac{M}{\rho}$$

with

$$\rho = \frac{\pi_Z^{\text{prior}}[\pi_Y(y_{\text{obs}}|\cdot)^2]}{\pi_Z^{\text{prior}}[\pi_Y(y_{\text{obs}}|\cdot)]^2}$$

(i) **Bonus:** Assume that $\tilde{w}_i = C\pi_Y(y_{\text{obs}}|z_i)$ where $z_i \sim \pi_Z^{\text{prior}}$ are independent samples and C is an unknown normalizing constant. We normalize the weights by defining

$$w_i = \frac{1}{\sum_{i=1}^M \tilde{w}_i} \tilde{w}_i$$

to have $\sum_{i=1}^{M} w_i = 1$. Show that

$$\hat{\rho} := M \sum_{i=1}^{M} w_i^2,$$

is an approximation to ρ. Hint: Use the law of large numbers.

(ii) Show that

$$\rho \geq \exp\left(D_{KL}(\pi_Z^{\text{post}} \| \pi_Z^{\text{prior}}) \right) .$$

(iii) Let μ and ν be two PDFs. We denote by $\mu^{\otimes n}$ the n-fold product of μ with itself i.e.

$$\mu^{\otimes n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \mu(x_i).$$

This measure samples an independent μ-distributed random variable for each coordinate. Show that $D_{KL}(\mu^{\otimes n} | \nu^{\otimes n}) = n D_{KL}(\mu | \nu)$.

(iv) What does (iii) tell you about the effective sample size as the dimension of the the space increases?

Exercise 3. Suppose you would like to obtain samples from the following bimodal distribution with density function

$$\pi(x) = \sum_{i=1}^{2} w^i N(\mu_2^i, \sigma_2^i).$$

with

	$i = 1$	$i = 2$
μ_2^i	-2	2
σ_2^i	0.5	0.5
w^i	0.5	0.5

i) Write pseudo-code using the method of importance sampling to generate an approximation to π, by using a Gaussian proposal density $\pi' = N(\mu = 0, \sigma^2)$ for both $\sigma^2 = 1$ and $\sigma^2 = 4$.
ii) Implement the above pseudo-code numerically. Produce a histogram of these samples and overlay the exact density π.
iii) Briefly discuss (1-2 sentences) the consequences of each choice of σ^2.