4. EXERCISE SHEET *Bayesian inference and data assimilation*, SS2018

Hand in on 14.05.2018

Exercise 1. Consider a transition matrix $\mathrm{P} \in \mathbb{R}^{M\times M}$ which we assume to be also symmetric, *i.e.*, $\mathrm{P}^T = \mathrm{P}$. Eigenvectors $v \in \mathbb{R}^M$ with eigenvalue $\lambda = 1$ are called *stationary* under the transition matrix P.

(i) Show that all eigenvalues λ of P are real and satisfy $\lambda \in [-1, 1]$.

(ii) A transition matrix P is called *geometrically ergodic* if it possesses a unique stationary eigenvector and all other eigenvalues λ satisfy $|\lambda| < 1$. Show that the entries v_i^* of a stationary eigenvector v^* of a geometrically ergodic transition matrix satisfy $v_i^* \geq 0$ and can be normalised to $\sum_{i=1}^M v_i^* = 1$. (Of course, if v^* is an eigenvector so is $-v^*$.) Show that other eigenvectors v satisfy $\sum_{i=1}^M v_i = 0$.

Exercise 2.

(i) Verify that the Kullback-Leibler divergence of two univariate Gaussians $X_i \sim \mathrm{N}(\bar{x}_i, \sigma_i^2)$, $i = 1, 2$, is given by

$$\begin{aligned} d_{\mathrm{KL}}(F_{X_1}, F_{X_2}) &= \int_{\mathbb{R}} \ln \frac{\pi_{X_1}(x)}{\pi_{X_2}(x)} \pi_{X_1}(x)\mathrm{d}x \\ &= \frac{1}{2}\left(\sigma_2^{-2}\sigma_1^2 + \sigma_2^{-2}(\bar{x}_2 - \bar{x}_1)^2 - 1 - 2\log\frac{\sigma_1}{\sigma_2}\right). \end{aligned}$$

(ii) Verify that the Wasserstein distance between π_{X_1} and π_{X_2} is given by

$$W(\pi_{X_1}, \pi_{X_2})^2 = (\bar{x}_1 - \bar{x}_2)^2 + (\sigma_1 - \sigma_2)^2.$$

Exercise 3. Consider the linear, discrete–time Brownian dynamics model,

$$x^{n+1} = x^n - \Delta\tau x^n + \sqrt{2\Delta\tau}\xi^n, \qquad \xi^n \sim \mathrm{N}(0,1), \tag{1}$$

with $\Delta\tau = 0.01$ and initial value $x^0 = 0$. This model has the Gaussian distribution $\mathrm{N}(0, 1)$ as the invariant distribution in the limit $\Delta\tau \to 0$. Use the model in order to generate a total of $M_{\max} = 13000$ samples x_i from the Gaussian distribution by storing every tenth iterate of (**??**). Compute empirical means and variances and compare those with the exact values (mean zero, variance one). Repeat the experiment one hundred time to assess the distribution of estimation errors in the mean and variance.