Time Series Analysis & Recurrent Neural Networks

lecturer: Daniel Durstewitz
tutors: Georgia Koppe, Leonard Bereska
WS2019/2020

Exercise 5

To be uploaded before the exercise group on November 27th, 2019

Task 1. Kullback-Leibler divergence of two normal distributions
The Kullback-Leibler divergence of two continuous distributions $\mathcal{P}$ and $\mathcal{Q}$ is

$$KL(P \parallel Q) = \int_{-\infty}^{\infty} p(x) log\left(\frac{p(x)}{q(x)}\right) dx.$$

Compute it analytically in the case where $\mathcal{P}$ and $\mathcal{Q}$ are two normal distributions $\mathcal{P} = \mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{Q} = \mathcal{N}(\mu_2, \sigma_2)$.

1. What is the specific value of $KL(P \parallel Q)$ when $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$?

2. What are the specific values of $KL(P \parallel Q)$ and $KL(Q \parallel P)$ when $\mu_1 = 0, \sigma_1 = 2, \mu_2 = 1, \sigma_2 = 1$?

Task 2. Kalman filter smoother
In file 'ex5.mat' (or alternatively, the corresponding '*.xls' files) you will find the variables 'z', 'x', 'A', 'B', 'C', 'u', 'Σ', and 'Γ', which specify a model as given in (script Ch7, eq. 7.53) with an additional control variable term, as described below. The variables represent a linear dynamical system, where the state of the system (in this case a drone) is contained in the multivariate vector variable 'z' of size $p = 2 \times T = 100$, where p specifies the number of latent states, and T the number of time points. The states correspond to the position of the drone (row 1 in variable 'z', referring for simplicity to height only) and velocity (row 2 in variable 'z'). In addition, the drone is controlled by external forces (i.e. the controller) contained in variable 'u'. The complete dynamical system evolves according to:

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{C}\mathrm{u}_t + \epsilon_t, \quad \text{with} \quad \epsilon_t \sim N(0, \Sigma)$$

where $\mathbf{A}$ is the transition matrix, $\mathbf{C}$ marks the influence of control variables u_t on the state, and Σ is the process noise covariance matrix (e.g. air currents). Assume the states are not directly observable, but only noisy and mixed measurements from the position and the velocity of the drone are available, contained in vector variable 'x' (with rows 1-2 corresponding to a mixture of the latent states 'z'). These measurements are modeled as:

$$\mathbf{x}_t = \mathbf{B}\mathbf{z}_t + \eta_t, \quad \text{with} \quad \eta_t \sim N(0, \Gamma)$$

where B is the mixture matrix and Γ the measurement noise covariance matrix.

1. Retrieve an estimate for the latent states 'z' purely from the observations 'x' by implementing the Kalman-filter recursions.
   (Clue: Use eq. (7.63) to implement the 'filter' going forward in time from $t = 1...T$. Make sure to include the extra control variable term by changing the top row of (7.63) to
   $\mu_t = \mathbf{A}\mu_{t-1} + \mathbf{C}\mathrm{u}_t + \mathbf{K}_t(\mathbf{x}_t - \mathbf{B}(\mathbf{A}\mu_{t-1} + \mathbf{C}\mathrm{u}_t))$. For $t = 1$, use μ_0 and $\mathbf{L}_0$ (also provided in the files) as initial conditions for the latent states and state covariance matrix.)

2. Plot the obtained predicted latent states against the true latent states (in variable 'z'). How well can you recover the true latent states? What could the drone be doing?

3. Examine the parameter values of **A** and **C**? Why are they chosen the way they are?

4. Now implement the Kalman-'smoother'. Clue: use eq. 7.70 to implement the Kalman smoother going backwards from $t = T$, to $t = 1$, in steps of 1. Note that $\boldsymbol{\mu}_T = \tilde{\boldsymbol{\mu}}_T$, and $\mathbf{V}_T = \tilde{\mathbf{V}}_T$.

5. Apply the Kalman-filter-smoother to obtain estimates for your latent state path. Plot the obtained predicted latent states against the true latent states (variable 'z'), and against the latent states obtained by only using the filter. Quantify these results by computing the mean squared error between true and recovered latent states (by both Kalman-filter, and Kalman-filter-smoother). How do they differ?