# Numerical Methods, Spring 2023
# Assignment 10 [State Space Models - 2]
# Total: 160, Deadline: 19 May

## I. SUGGESTED READING

- [1], Chapter 18 State space models

## II. EXERCISES

1. (25) Consider a familiar LDS, but this time with non-stationary variance of the noise:

$$x_t = x_{t-1} + u_{t-1} + \epsilon_{t-1}, \quad z_t = x_t + \eta_t \tag{1}$$

with $\epsilon_t \sim \mathcal{N}(0, \sigma_t)$ and $P(\eta_t) \sim \mathcal{N}(0, \rho_t)$. Using all data available in `10_1.pickle` [pickled dict of $u_t$, $\sigma_t$ and $\rho_t$], compute the smoothed trajectory of the hidden state $x_t$. Plot it as a function of time together with the error bar. Compute the prediction for the location of the object at $T+1$ (that is for the time instant, following the latest available observation).

2. (35) Consider 2D motion with random acceleration. Let $z_{1t}$ and $z_{2t}$ be the horizontal and vertical locations of the object, and $\dot{z}_{1t}$ and $\dot{z}_{2t}$ be the corresponding velocity. We can represent this as a state vector $z_t \in R^4$ as follows:

$$z_t = (z_{1t}, z_{2t}, \dot{z}_{1t}, \dot{z}_{2t})^T.$$

Let us assume that the object would-be-moving at constant velocity, but its motion is perturbed by random Gaussian noise (e.g., due to the wind). Additionally, we can only observe the noised location of the object $y_t \in R^2$, but not its velocity. We can model the system dynamics as follows:

$$z_t = A_t z_{t-1} + \epsilon_t, \quad y_t = C_t z_t + \delta_t,$$

with

$$A_t = \begin{pmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad C_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

and

$$\epsilon_t \sim \mathcal{N}(0, \sigma \hat{1}), \quad \delta_t \sim \mathcal{N}(0, \rho \hat{1}).$$

For the data in `10_2.pickle` [pickled dict of $y_t$, $\Delta$, $\sigma$ and $\rho$], compute the filtered trajectory $z_t$. Draw the filtered trajectory on the plane together with the observations and compute the prediction for the location of the object at $T+1$ (that is for the time instant, following the latest available observation).

3. (50) Consider the model in Eq. (1) with stationary but unknown $\sigma$ and $\rho$ and constant velocity, $u \equiv 1$. You are given a series of $T$ observations $z_1, ..., z_T$ in `10_3.pickle` [pickled dict of $z_t$]. Your task is to estimate the parameters $\sigma$ and $\rho$ using the EM (expectation–maximization) algorithm.

   - Derive expression for the log-likelihood $\ln P(x, z, \sigma, \rho)$ where $x$ and $z$ are the hidden state and observed time series.

   Pick some reasonable estimates for $\sigma$ and $\rho$ (lets denote them $\sigma_0$ and $\rho_0$ correspondingly). Iterate the following two steps ($i = 0, 1, ...$), until reasonable convergence:

   - Compute expected value of $\ln P(x, z, \sigma, \rho)$ – that is, average over $x$ with the distribution, given by $x \sim P(x, z, \sigma_i, \rho_i)$. The result is a function $C_i(\sigma, \rho)$ of a very simple analytical form, only the coefficients are to be computed numerically.

- Maximize the function $C_i(\sigma, \rho)$ over its arguments and set $(\sigma_{i+1}, \rho_{i+1}) = \mathrm{argmax}_{\sigma,\rho} C_i(\sigma, \rho)$.

What are the estimated values of $\sigma$ and $\rho$?

4. (50) Consider the model in Eq. (1) with stationary but unknown $\sigma$ and $\rho$ and constant velocity, $u \equiv 1$. You are given a series of $T$ observations $z_1, ..., z_T$ in `10_4.pickle` [pickled dict of $z_t$]. Your task is to estimate the parameters $\sigma$ and $\rho$ using maximization of the log-likelihood function.

   - The likelihood of given values of $\sigma$ and $\rho$ is proportional to $P(\sigma, \rho) \propto \int P(x, z, \sigma, \rho) dx$, where the integral can be evaluated explicitly. Using `JAX`, write fast code to compute $\ln P(\sigma, \rho)$ and its gradients up to the 2nd order.
   - Perform numerical maximization to find the optimize values for $\sigma$ and $\rho$.

   What are the estimated values of $\sigma$ and $\rho$?

5. (75*) Consider the local linear trend model:

$$y_t = a_t + \epsilon_t^y, \quad \epsilon_t^y \sim \mathcal{N}(0, \sigma_y)$$

$$a_t = a_{t-1} + b_{t-1} + \epsilon_t^a, \quad \epsilon_t^a \sim \mathcal{N}(0, \sigma_a)$$

$$b_t = b_{t-1} + \epsilon_t^b, \quad \epsilon_t^b \sim \mathcal{N}(0, \sigma_b)$$

with unknown parameters $\sigma_y$, $\sigma_b$. For the data in `10_5.pickle` [pickled numpy array with observed time series of $y_t$ and $\sigma_a$], compute, using the likelihood maximization (in spirit of Exercise 4 above), the most probable values of $\sigma_y$, $\sigma_b$. With these values, compute the smoothed trajectory of the local slope coefficient $b_t^*$ and plot it as a function of time. Compute the prediction for $y_{T+1}$ (that is for the time instant, following the latest available observation).

## REFERENCES

[1] Kevin P Murphy. *Machine learning: a probabilistic perspective.* MIT press, 2012.