02417: Time Series Analysis
Lasse Engbo Christiansen
Fall 2016

DTU Compute
Technical University of Denmark

Assignment 4: Multivariate Processes

The GPS chip in smartphones has some flavor of a Kalman filter onboard to reduce noise and potentially also to socalled fuse accelerations and gravitational measurements into the estimated position. To mimic what goes on inside you'll get data logged from a GPS with a little added noise. The noise that was added has a standard deviation of about 1.5m. To reduce numerical problems in your implementations the data has furthermore been transformed by:

$$mlat = (lat - 55.73) \cdot 1000 \qquad mlong = (long - 12.44) \cdot 1000$$

The data can be considered milli degrees around a point. You are expected to present the data on the transformed scale as it also eases the readability of plots. Sampling was done every second and 1000 samples are in `A4_gps_log.csv`.
You are expected to both use a marima approach and your own implementation of the Kalman filter. The last 50 observations should be left out when fitting models and only be used to compare with predictions.

Question 4.1: Presenting the data Present the data. Both the temporal and spatial aspects should be presented. Do comment on what you see.

Question 4.2: ACF, PACF, CCF and PCCF Calculate the crosscorrelation function and partial cross correlation function (Including ACF and PACF) for the data. If relevant also for differenced versions of the data.

Comment on the structures you find.

Question 4.3: MARIMA model Find a suitable MARIMA model for the bivariate system consisting of mlat and mlong. Reduce the model and validate it. Then present the final model.

Question 4.4: Predicting with MARIMA model Predict 50 steps ahead (From observation 950). Plot the last part of the data and the predictions. Include a 95% prediction interval (Assuming the normal distribution). Include a table with the 1, 10, 25 and 50 step predictions.

Compare with the 50 observations that where left out and comment on the results.

Question 4.5: Kalman filter formulation We choose the state vector $[x_t, v_{x,t}, y_t, v_{y,t}]^T$, where x_t represent the position and $v_{x,t}$ the velocity per sample in the x direction and similar for the y direction. The model for the x direction is given by:

$$\begin{aligned} x_t &= x_{t-1} + v_{x,t-1} + \varepsilon_{x,t} \\ v_{x,t} &= v_{x,t-1} + \varepsilon_{v_x,t} \end{aligned}$$

where $\varepsilon_{x,t} \sim N(\mu = 0, \sigma_x^2 = 0.0003)$ and $\varepsilon_{v_x,t} \sim N(\mu = 0, \sigma_v^2 = 10^{-5})$. Again it is similar for y. Finally, the system is observed by:

$$\begin{aligned} mlong_t &= x_t + \varepsilon_{1,t} \\ mlat_t &= y_t + \varepsilon_{2,t} \end{aligned}$$

where both $\varepsilon_{1,t}$ and $\varepsilon_{2,t} \sim N(\mu = 0, \sigma_o^2 = 10^{-4})$.

From the supplied information you are supposed to formulate the system on state space form.

Question 4.6: Kalman filter implementation Implement you own version of the Kalman filter you formulated in the previous question in R, Matlab, Python or a similar language. You should provide a listing of your code showing your implementation in the report. (Not just as an appendix).

Question 4.7: Using your Kalman filter implementation Use the implementation to reconstruct (estimate) the position of the gps logger including a 95% confidence interval. Present the reconstructions.

Predict the position 50 steps ahead and present the predictions as for the MARIMA model.

Comment on the results.

Question 4.8: Bonus: Optimize noise parameters The provided values for the variances are not the optimal values. If you have time for an extra challenge then find the maximum likelihood estimates of the three variances (σ_x^2, σ_v^2, and σ_o^2).

Question 4.9: Comparison Comment on the performance of the two modelling approaches.

Suggest alternative formulations of the Kalman filter that may perform better.