DS-GA 3001.001 Special Topics in Data Science: Modeling Time Series
Homework 2

Yves Greatti - yg390

Problem 1. LDS model, 10p
Consider a special case of LDS with $\mathbf{C} = \mathbf{I}$ and $\mathbf{R} = \sigma^2\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. Show that in the limit where there is no observation noise the best estimate for latent $\mathbf{z}_i$ is to simply use the observation $\mathbf{x}_i$: formally, in the limit when $\sigma^2 \to 0$ the posterior for $\mathbf{z}_i$ has mean $\mathbf{x}_i$ and vanishing variance.
Given the LDS model specifications for the posterior $\mathbf{z}_i$:

$$\begin{aligned}\mu_{i|i} &= \mu_{i|i-1} + \boldsymbol{K}_i(\mathbf{x}_i - \boldsymbol{C}\mu_{i|i-1}) \\ \Sigma_{i|i} &= \Sigma_{i|i-1} - \boldsymbol{K}_i\boldsymbol{C}\Sigma_{i|i-1}\end{aligned}$$

where $\boldsymbol{K}_i = \Sigma_{i|i-1}\boldsymbol{C}^T(\boldsymbol{C}\Sigma_{i|i-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}$. If $\mathbf{C} = \mathbf{I}$ and $\mathbf{R} = \sigma^2\mathbf{I}$, then the Kalman gain becomes $\boldsymbol{K}_i = \frac{\Sigma_{i|i-1}}{\Sigma_{i|i-1}+\sigma^2\mathbf{I}}$, taking the limit $\sigma^2 \to 0$ then $\boldsymbol{K}_i \to \mathbf{I}$, thus in the limit the posterior $\mathbf{z}_i$ has for parameters:

$$\begin{aligned}\mu_{i|i} &= \mu_{i|i-1} + (\mathbf{x}_i - \mu_{i|i-1}) = \mathbf{x}_i \\ \Sigma_{i|i} &= \Sigma_{i|i-1} - \Sigma_{i|i-1} = \mathbf{0}\end{aligned}$$

Problem 2. LDS prediction, 10p
Given the standard parametrization of the LDS model, and the Kalman filtering estimates $\mu_{i|i}$ and $\Sigma_{i|i}$, obtained for a dataset $\mathbf{x}_{1:t}$ write down the expressions for predicting the following 2 observations in the sequence $\mathbf{x}_{t+1}$ and $\mathbf{x}_{t+2}$.
Given the prior with parameters $\mu_{i|i}$ and $\Sigma_{i|i}$, we compute the prior for $\mathbf{z}_{t+1} \approx \mathcal{N}(\mu_{i+1|i}, \Sigma_{i+1|i})$:

$$\begin{aligned}\mu_{i+1|i} &= \boldsymbol{A}\mu_{i|i} \\ \Sigma_{i+1|i} &= \boldsymbol{A}\Sigma_{i|i}\boldsymbol{A}^t + \boldsymbol{Q}\end{aligned}$$

We obtain the prediction $\mathbf{x}_{t+1} \approx \mathcal{N}(\boldsymbol{C}\mu_{i+1|i}, \boldsymbol{C}\Sigma_{i+1|i}\boldsymbol{C}^t + \boldsymbol{R})$, we then incorporate the evidence to obtain the posterior parameters:

$$\begin{aligned}\mu_{i+1|i+1} &= \mu_{i+1|i} + \boldsymbol{K}_{i+1}(\mathbf{x}_{t+1} - \boldsymbol{C}\mu_{i+1|i}) \\ \Sigma_{i+1|i+1} &= \Sigma_{i+1|i} - \boldsymbol{K}_{i+1}\boldsymbol{C}\Sigma_{i+1|i}\end{aligned}$$

where the Kalman gain matrix $\boldsymbol{K}_{i+1} = \Sigma_{i+1|i}\boldsymbol{C}^t(\boldsymbol{C}\Sigma_{i+1|i}\boldsymbol{C}^t + \boldsymbol{R})^{-1}$ which we use to update the posterior $\mathbf{z}_{t+1} \approx \mathcal{N}(\mu_{i+1|i+1}, \Sigma_{i+1|i+1})$. With the updated $\mathbf{z}_{t+1}$ we can compute the next prior $\mathbf{z}_{t+2} \approx \mathcal{N}(\mu_{i+2|i+1}, \Sigma_{i+2|i+1})$ where

$$\begin{aligned}\mu_{i+2|i+1} &= \boldsymbol{A}\mu_{i+1|i+1} \\ \Sigma_{i+2|i+1} &= \boldsymbol{A}\Sigma_{i+1|i+1}\boldsymbol{A}^t + \boldsymbol{Q}\end{aligned}$$

which is used to obtain the prediction $\mathbf{x}_{t+2} \approx \mathcal{N}(\boldsymbol{C}\mu_{i+2|i+1}, \boldsymbol{C}\Sigma_{i+2|i+1}\boldsymbol{C}^t + \boldsymbol{R})$.

Problem 3. LDS inference with missing observations, 10p
Consider a variation of the original LDS graphical model with one single missing value $\mathbf{x}_j$. Everything else is as in the original; the only difference is that the graphical model loses the downward observation arrow and the corresponding $\mathbf{x}_j$.) How do the Kalman filtering/smoothing updates change?
During filtering the current state $\mathbf{z}_{j+1}$ updated as a function of $\mathbf{z}_j$ and the most recent observation $\mathbf{x}_{j-1}$ is:

$$\begin{aligned}\mu_{j+1|j+1} &= \mu_{j+1|j} - \boldsymbol{K}_{j+1}\boldsymbol{C}\mu_{j+1|j} \\ \Sigma_{j+1|j+1} &= \Sigma_{j+1|j} - \boldsymbol{K}_{j+1}\boldsymbol{C}\Sigma_{j+1|j}\end{aligned}$$

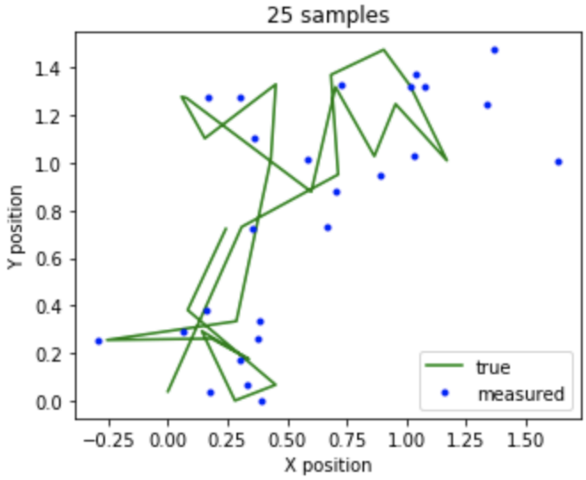
where $\boldsymbol{K}_{j+1} = \Sigma_{j+1|i}\boldsymbol{C}^t(\boldsymbol{C}\Sigma_{j+1|j}\boldsymbol{C}^t + \boldsymbol{R})^{-1}$.

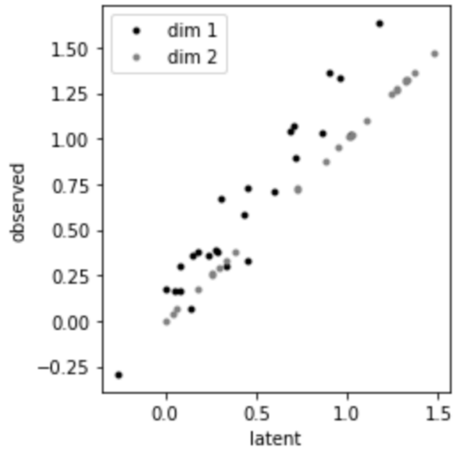
Problem 4. LDS filtering, 10p

Given the model parameters: $A = \begin{bmatrix} 0.65 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$, $C = \begin{bmatrix} 1.1 & 0.2 \\ 0.5 & 0.95 \end{bmatrix}$, $Q = 0.1\mathbf{I}_2$, $R = 0.01\mathbf{I}_2$ with initial condition parameters $\mu_0 = [0\ 0]^t$, $\Sigma_0 = 0.5\mathbf{I}$. Generate 25 samples from the model and plot them. Use these observation for inference (filtering and smoothing). How much does the Kalman gain $\mathbf{K}_i$, and $\mathbf{F}_i$ vary across timepoints? Try playing around with the parameters. Does the result change? Discuss.

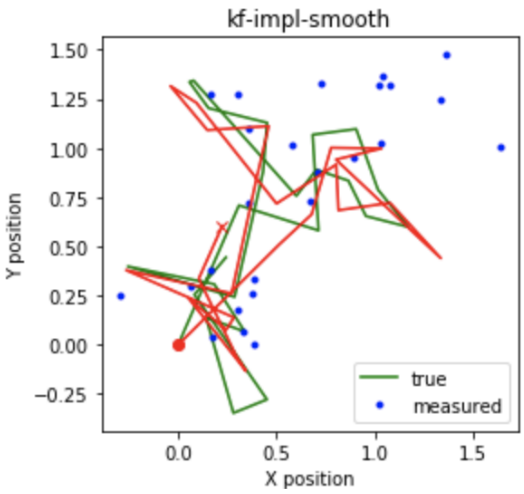
Note: you can use code from the lab as starting point, if you want. Not need to submit the code, only your figures.

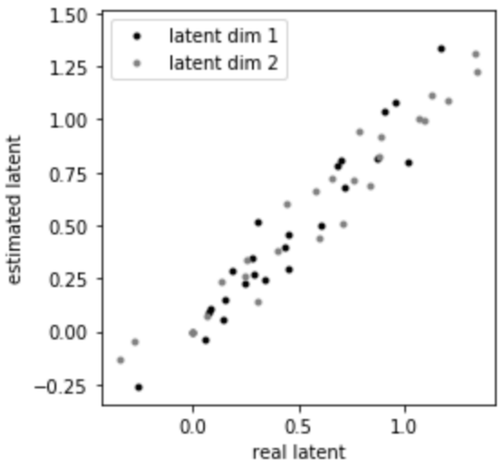
With the given parameters, using the code from the lab, the sampling of 25 observations gives us the plots below:



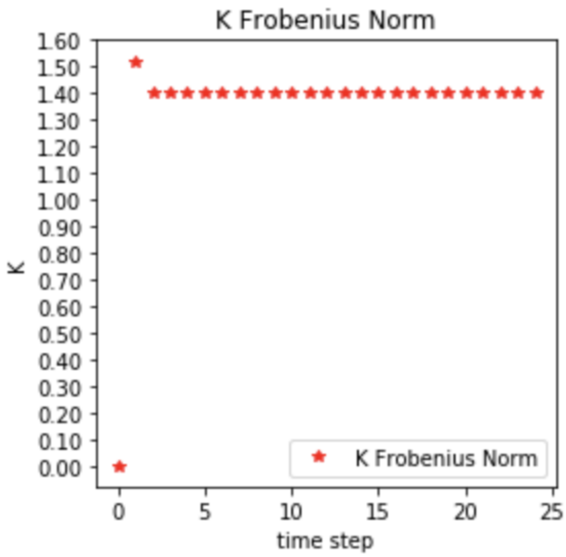


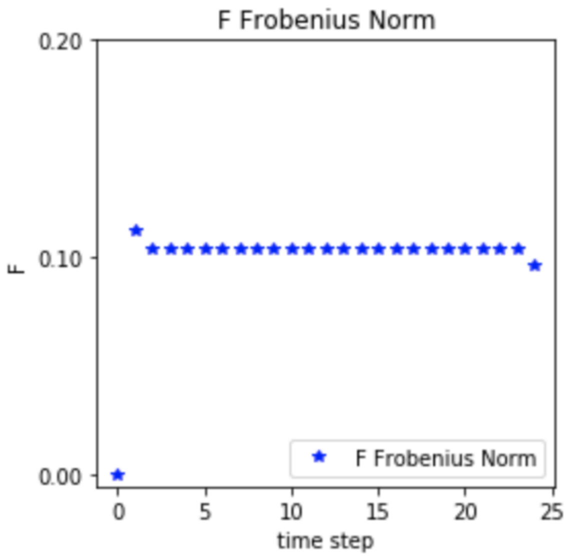
After inference (filtering and smoothing) we obtain the estimates in red:



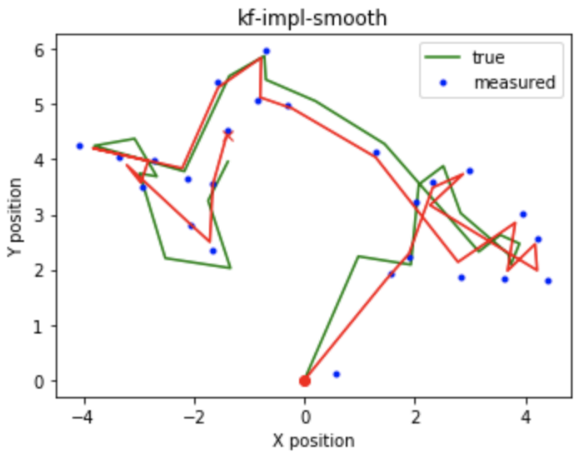


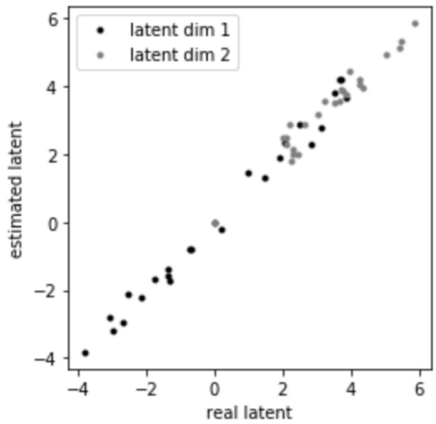
At each time step, we compute the frobenius norm: $\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = \sqrt{\operatorname{trace}(A^*A)} = \sqrt{\sum_{i=1}^{\min\{m,n\}}\sigma_i^2(A)}$ of the Kalman gain and F matrices and we plot them. On the plots below, we can see that once established since the system is linear, the Kalman matrix does not change.





Using now a smoother dynamic system, as described below, we vary different parameters and observe the results listed below:





As we reduce the observation noise the Kalman Gain and the F "Gain" decrease.

```
22]: df
```

22]:

	K-Norm-Mean	K-Norm-Var	F-Norm-Mean	F-Norm-Var
0	1.24374	0.06445	0.11394	0.00054
1	1.24374	0.06445	0.11394	0.00054
2	1.34438	0.07531	0.01331	0.00001
3	1.35630	0.07665	0.00135	0.00000

As we increase the temporal dynamics (transition matrix) the Kalman Gain and F increase.

```
4]: df
```

4]:

	K-Norm-Mean	K-Norm-Var	F-Norm-Mean	F-Norm-Var
0	1.24374	0.06445	0.11394	0.00054
1	1.24374	0.06445	0.11394	0.00054
2	1.34532	0.07541	0.16918	0.05159
3	1.35751	0.07678	0.52721	6.33260

As we increase the observation dynamics the Kalman Gain and F decrease.

```
[26]: df
```

[26]:

	K-Norm-Mean	K-Norm-Var	F-Norm-Mean	F-Norm-Var
0	1.24374	0.06445	0.11394	0.00054
1	1.24374	0.06445	0.11394	0.00054
2	0.13563	0.00077	0.00135	0.00000
3	0.01358	0.00001	0.00001	0.00000

Problem 5 Particle filtering, 10p

Consider the usual LDS model, but where inference is done using particle filtering instead of the traditional Kalman filter. Write down pseudocode for the particle updates. Given the generated samples, $\{\mathbf{z}_i^{(k)}\}_{k=1:K,i=1:t}$, how would you go about computing the quantities $\mu_{i|i}$, $\Sigma_{i|i}$ and $\mathbb{E}[\mathbf{z}_i\mathbf{z}_{i+1}^t]$?

Hint: Use the general form from the lecture, and plug in the expressions for the different probabilities of the LDS model. The mean and variance can be written as expectations and approximated accordingly.