Deep Learning for Smoothing in Dynamical Systems

UPPSALA UNIVERSITET

Project Outline

- Implement state-of-the-art deep learning methods for solving the smoothing problem.
- Three methods have been examined for different state space models and compared to the ground truth given by the Kalman smoother.

PDF – Probability Density Function
CNN - Convolutional Neural Network
RNN - Recurrent Neural Network
BBVI - Black Box Variational Inference

State Space Model

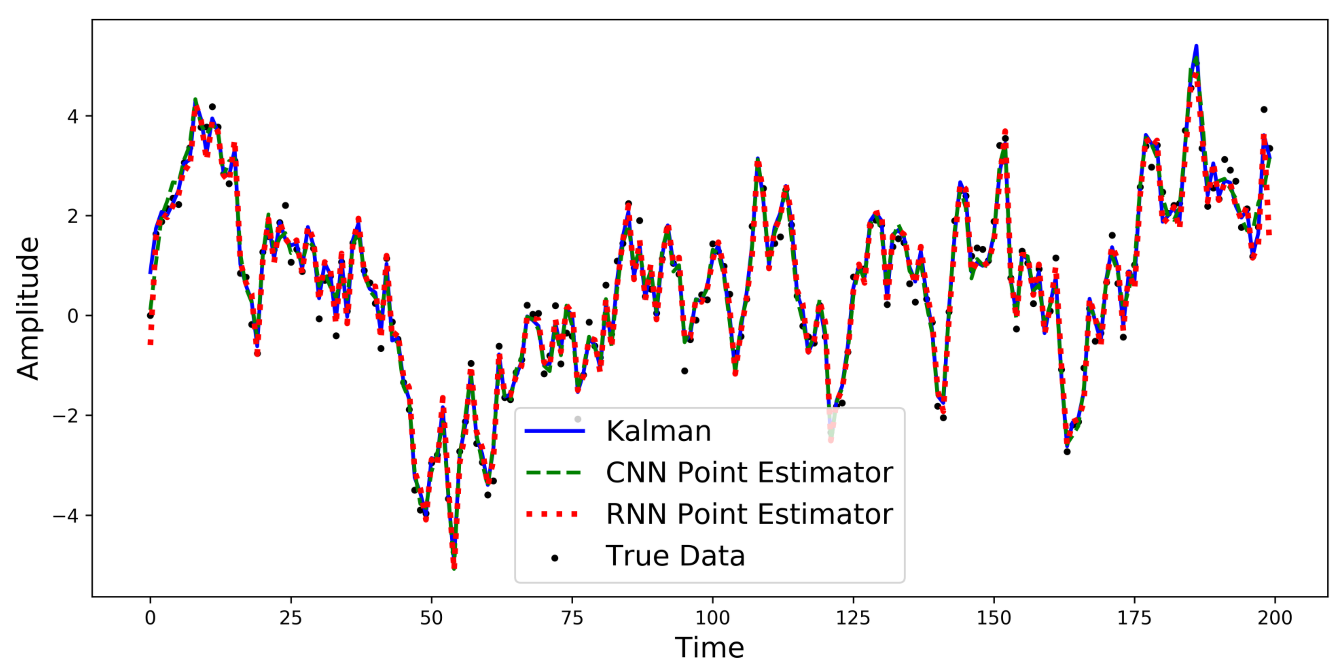
A state space model can be written as

$$z_{t+1} = f(z_t) + {v_1}_t$$
$$x_t = g(z_t) + {v_2}_t,$$

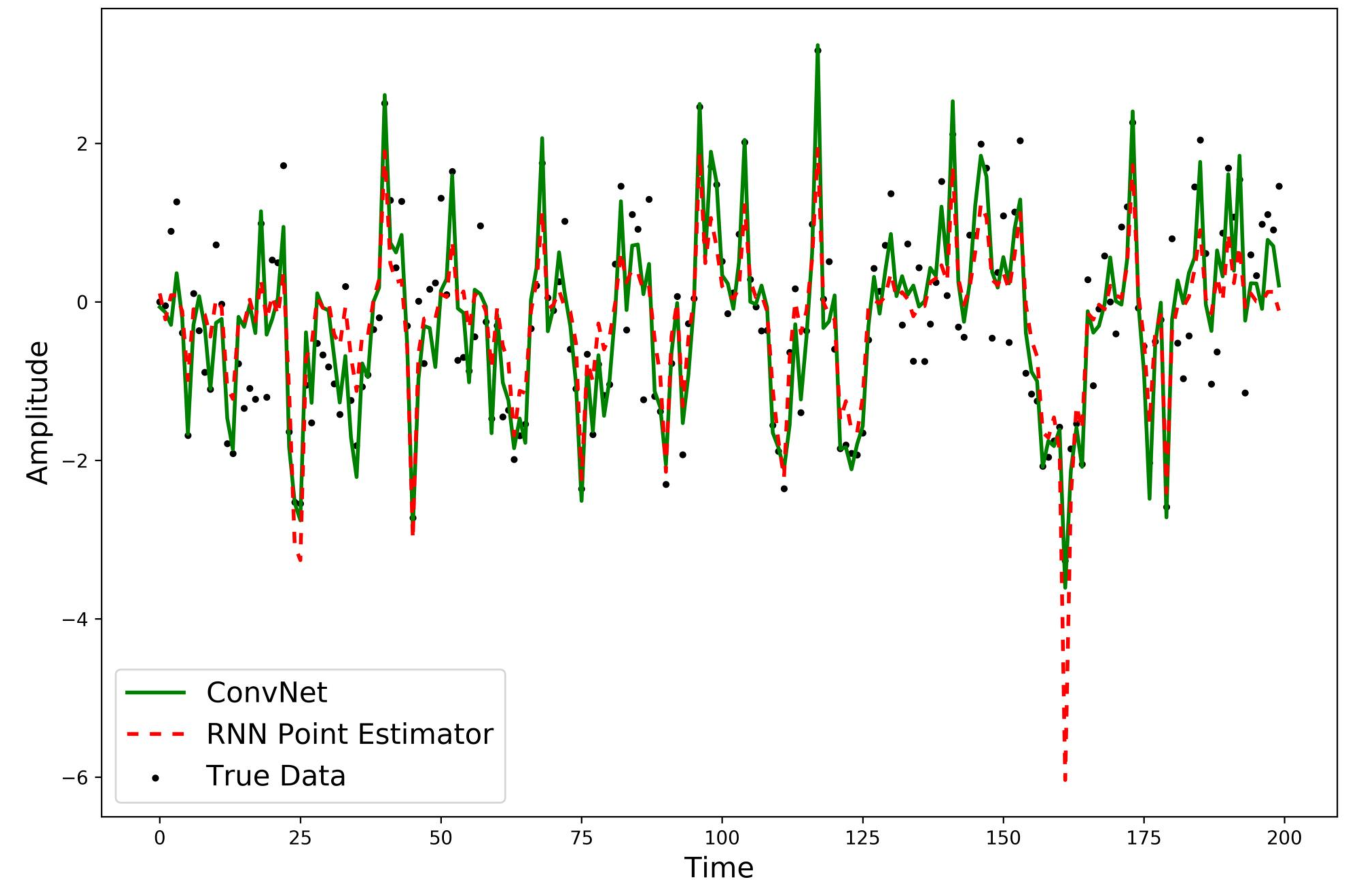
z - hidden states, x - measurements. Finding the posterior PDF $p(\boldsymbol{z}|\boldsymbol{x})$ is the goal of the smoothing problem. Here, $\boldsymbol{z} = (z_1, z_2, \dots, z_T)$ and $\boldsymbol{x} = (x_1, x_2, \dots, x_T)$.

Point Estimators

A CNN point estimator and a RNN point estimator were examined by training on generated data and minimising loss functions related to mean squared errors.



Results for linear state space model with Gaussian white noise.



Results for non-linear state space model with Gaussian white noise.

BBVI Using a RNN

BBVI is a method for hidden variable models that turns the smoothing problem into an optimisation problem. An approximate posterior $q_\theta(\boldsymbol{z}|\boldsymbol{x}) \approx p(\boldsymbol{z}|\boldsymbol{x})$ is found by maximising the objective function

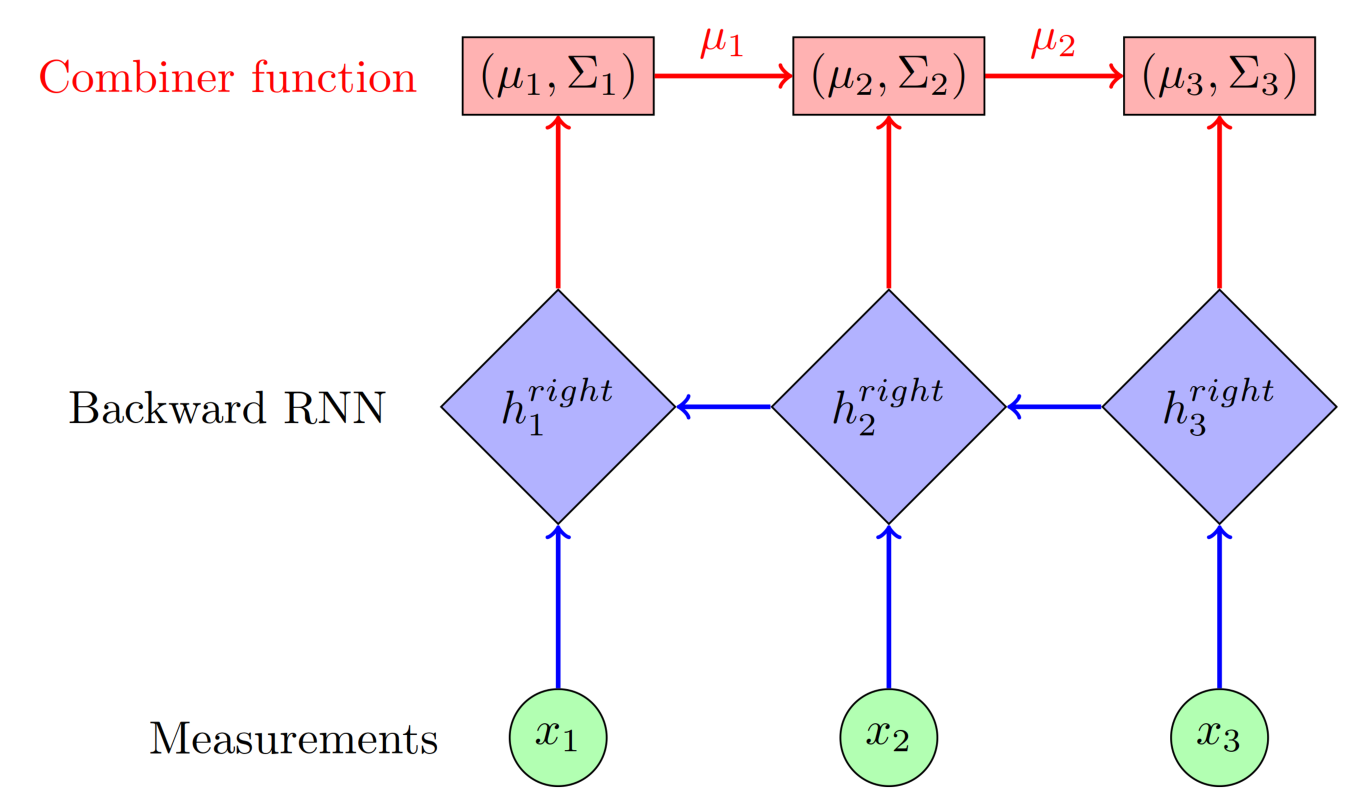
$$\mathcal{L} = \mathbb{E}_{q_\theta(\boldsymbol{z})}[\log p(\boldsymbol{x}, \boldsymbol{z}) - \log q_\theta(\boldsymbol{z}|\boldsymbol{x})].$$

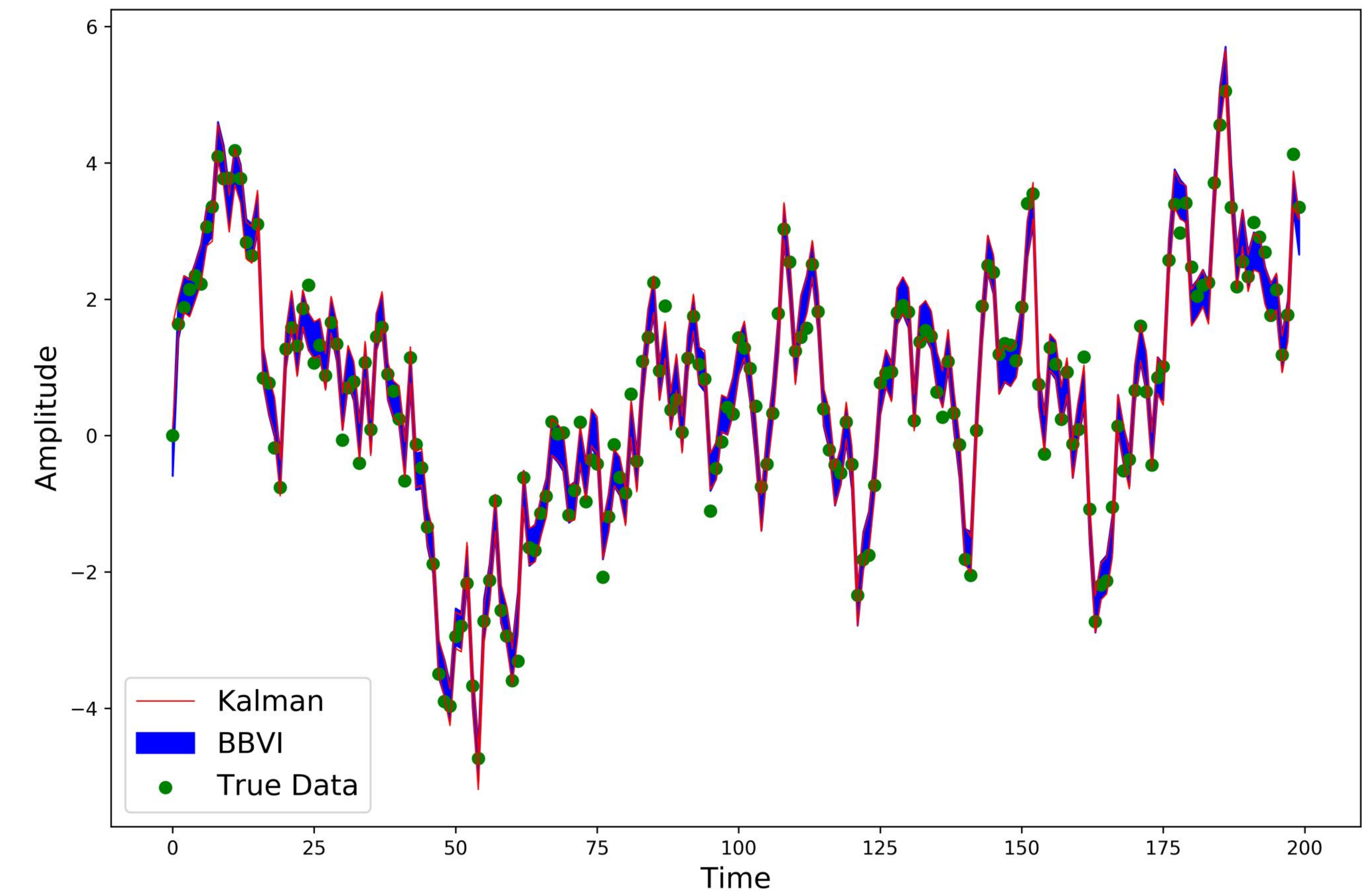
$q_\theta(\boldsymbol{z}|\boldsymbol{x})$ is factorised as

$$q_\theta(\boldsymbol{z}|\boldsymbol{x}) = q_\theta(z_1|x_1, \dots, x_T) \prod_{t=2}^{T} q_\theta(z_t|z_{t-1}, x_t, \dots, x_T),$$

$$q_\theta(z_t|z_{t-1}, x_t, \dots, x_T) \sim \mathcal{N}\big(\mu(z_{t-1}, x_t, \dots, x_T), \sigma^2(z_{t-1}, x_t, \dots, x_T)\big).$$

This was implemented with the RNN illustrated in the top right corner.





Full posterior PDF for linear state space model with Gaussian white noise.

Conclusions

- Both point estimators using the CNN and the RNN get close to recovering the expected value of the posterior PDF obtained by the Kalman smoother.
- Our BBVI using a RNN recovers the full posterior PDF obtained by the Kalman smoother.
- All of the methods are more flexible than the Kalman smoother.

Hannes Larsson
Hannes.Larsson.9471@student.uu.se
Songchen Li
lisongchenmichael@outlook.com

Supervisor
Fredrik Lindsten
fredrik.lindsten@it.uu.se

Course Coordinator
Maya Neytcheva
maya.neytcheva@it.uu.se

Project in Scientific Computing, 2017