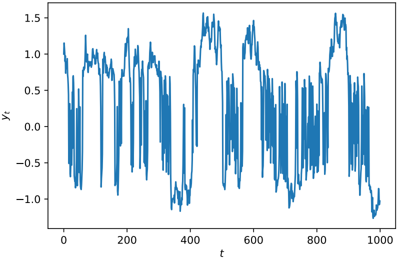
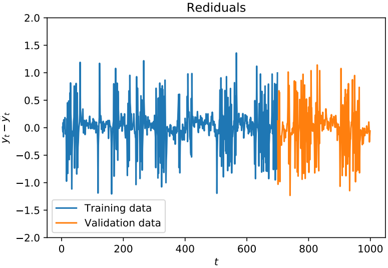
Time Series and Sequence Learning

Discussion seminar for Lecture 3

Johan Alenlöv, Linköping University

2022-09-05

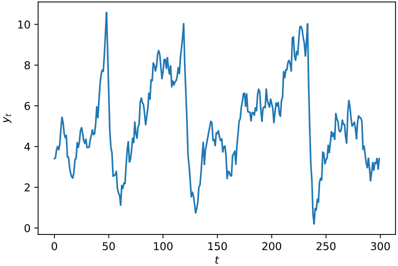




Left: Data. *Right:* Residuals for a fitted AR(3) model.

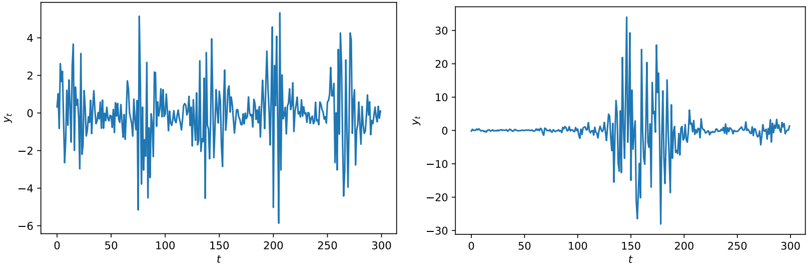
Discussion questions:

1. How can we see from the **residuals** that the data is not well explained by the model?
2. Is it possible to see that that data is not well explained by a *linear process* by **just looking at the data**? How?



Discussion questions:

1. Does the data above **appear** to be generated by a linear AR(p) model?
2. If $\{y_t\}_{t\geq 1}$ is a **stationary** process generated by an AR(p) model with coefficients $(a_1, \dots, a_p)$, then what can be said about the **negated process** $\{-y_t\}_{t\geq 1}$



Left: Non-stationary (cyclic variance).
Right: Stationary but nonlinear (stochastic variance).

When we are asked the question *"Is this time series stationary?"*, the meaning is often *"Could this time series have been generated by a stationary linear process?"*

In practice it is hard to distinguish between non-stationarity and nonlinearity!

Nonlinear auto-regressive model, NAR(p):

$$y_t = f_{\theta}(y_{t-1}, y_{t-2}, \ldots, y_{t-p}) + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_{\varepsilon}^2),$$

Using an 2-layer NN to model f_{θ}:

$$\begin{aligned} \mathrm{h}_t &= \sigma(W^{(1)} H_{t-1} + b^{(1)}) \\ y_t &= W^{(2)} \mathrm{h}_t + b^{(2)} + \varepsilon_t. \end{aligned}$$

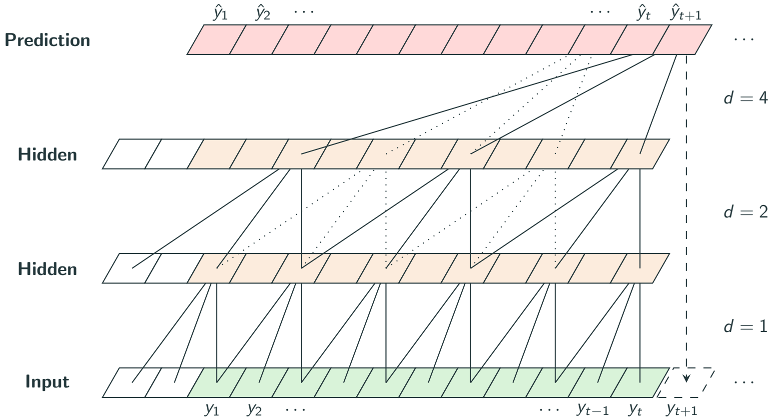
Discussion questions:

1. Contrary to a linear AR(p) model it is customary to include "bias terms" $b^{(j)}$ in the MLP-NAR. Why? Why don't we include a bias/intercept in the linear AR model?
2. How is the squared loss function

$$L(\theta) = \frac{1}{n-p} \sum_{t=p+1}^{n} \left(y_t - f_\theta(y_{t-1}, y_{t-2}, \ldots, y_{t-p})\right)^2$$

related to the **data likelihood**?

lecture3{b,c} - TCN



Discussion questions:

1. Which statements are correct?
   a) NAR is a special case of TCN.
   b) TCN is a special case of NAR.
   c) TCN and NAR are equivalent model classes.
2. How is the receptive field of a TCN related to the order p of a NAR(p) model?
3. Will the receptive field always increase exponentially with depth for a TCN with dilated convolutions?