Time Series and Sequence Learning

Validation, Order selection

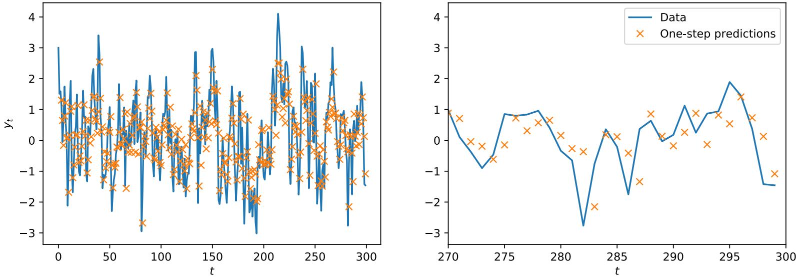
Fredrik Lindsten, Linköping University

ex) Toy model

We simulate an AR(3) model for $n = 300$ time steps,

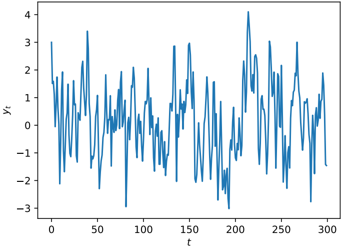
$$y_t = 0.9 y_{t-1} - 0.4 y_{t-2} + 0.2 y_{t-3} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, 1)$$

Estimating the model parameters with OLS gives:
$\widehat{\theta} = (0.84, -0.33, 0.16)$ and $\widehat{\sigma}_\varepsilon^2 = 0.95$.



Order selection

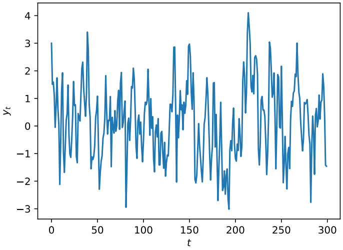
In practice we only observe the data.



How do we know which model order p to pick?!

Order selection

In practice we only observe the data.



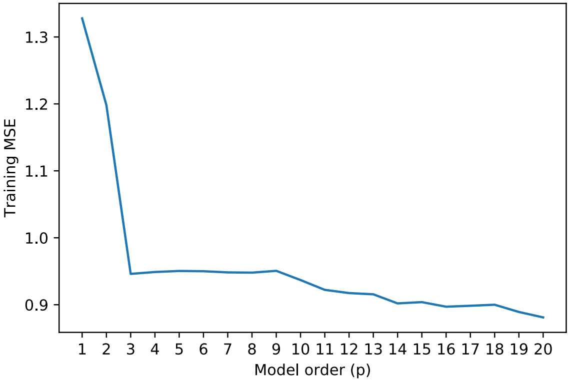
How do we know which model order p to pick?!

Two approaches:

1. Try to figure it out ***before*** fitting the model ("exploratory data analysis")
2. Estimate multiple models of different orders and perform model selection by validation!

ex) Toy model, cont'd

1. Look for the "bend" in training error plot



plot the MSE of training data 1st

Residual analysis

2. Look at the residuals!

The model assumption is

$$y_t = \theta^\mathsf{T} \phi_t + \varepsilon_t, \qquad \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2).$$

Residual analysis

2. Look at the residuals!

The model assumption is

$$y_t = \theta^\top \phi_t + \varepsilon_t, \qquad \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2).$$

Hence, if the model is accurate, we expect

$$y_t - \hat{\theta}^\top \phi_t \approx \varepsilon_t.$$

The residuals should be white Gaussian noise!

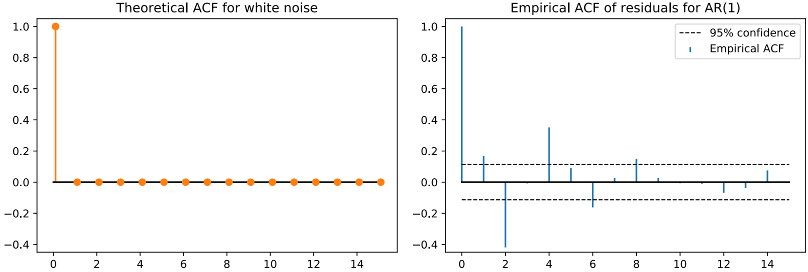
1. Auto-correlation
2. QQ-plots for marginal Gaussianity
3. ...

ex) Toy model, cont'd

X_t = phi * X_{t-1} + epsilon
phi(h) = Phi^h(big one) at lag t
\hat{ phi} (h) = sum_{t=1}^{n-h} (x_t-bar{x})(x_{t+h} -bar{x})/
sum_{t=1}^{n} (x_t -X)^2
when N is big, \hat{ phi} (h) is approximately normal
so \hat{ phi} (h) ~ N(phi(h),1/n) phi(h) is true autocorr
and var of \hat{ phi} (h) is appr 1/n, because
var (\hat{ phi} (h)) ~ 1/n, so sqrt(var) = 1/sqrt(n)
Reason:
Asymptotic Normality of Sample Autocorrelations
样本自相关的渐近正态性

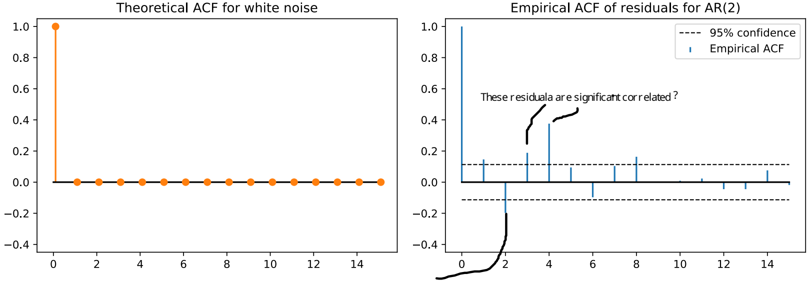
$\hat{\rho}(h) \sim N(0, \frac{1}{\sqrt{n}})$ $h > 0$, n large

Estimated model: AR(1)



ex) Toy model, cont'd

Estimated model: AR(2)

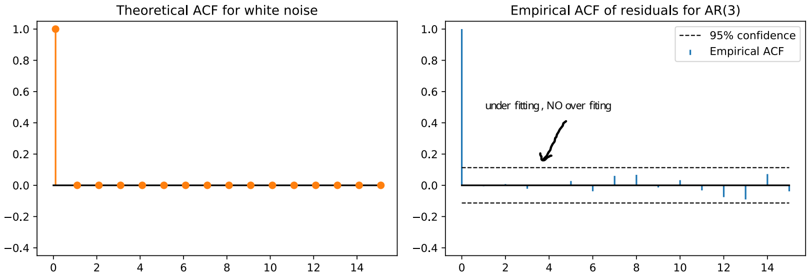


This could indicate that there is some degree of autocorrelation at those lags which is not explained by the AR(1) model.

?

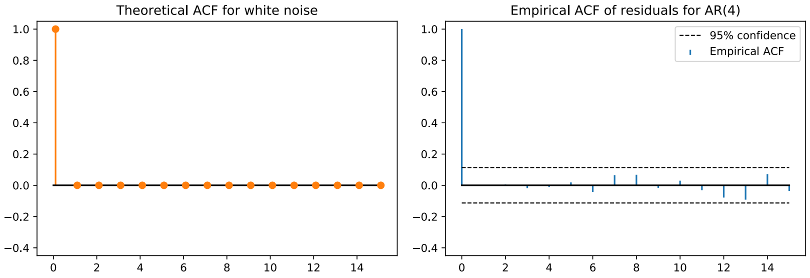
ex) Toy model, cont'd

Estimated model: AR(3)



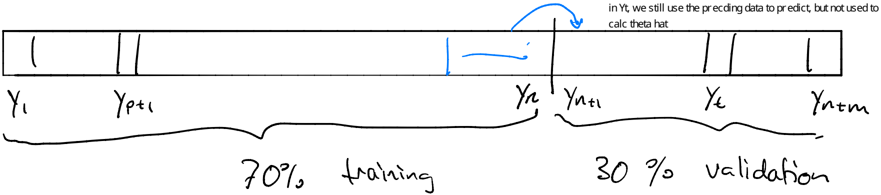
ex) Toy model, cont'd

Estimated model: AR(4)



Prediction error validation

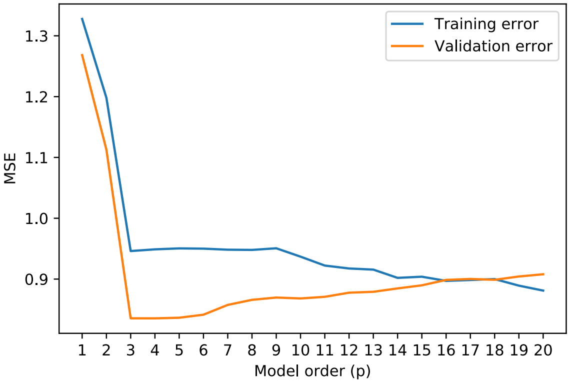
3. Evaluate on held-out validation data!



Validation mean-squared error, using one-step-ahead predictions:

$$\text{Val-MSE}(\widehat{\theta}) = \frac{1}{m} \sum_{t=n+1}^{n+m} (y_t - \widehat{\theta}^\mathsf{T} \phi_t)^2$$

ex) Toy model, cont'd



Testing the model