Applied Time Series Analysis
FS 2014 – Week 13

Marcel Dettling

Institute for Data Analysis and Process Design

Zurich University of Applied Sciences

marcel.dettling@zhaw.ch

http://stat.ethz.ch/~dettling

ETH Zürich, May 19, 2014

Applied Time Series Analysis
FS 2014 – Week 13

State Space Models

Basic idea: There is a stochastic process/time series X_t which we cannot directly observe, but only under the addition of some measurement noise.

Thus: We observe the time series $Y_t = X_t + V_t$, with iid measurement errors $V_t \sim N(0, \sigma_V^2)$

Example: X_t = # of fish in a lake
Y_t = # estimated number of fish from a sample

Other:
- Dynamic linear modeling
- Regression with time-varying coefficients

Applied Time Series Analysis
FS 2014 – Week 13

State Space Formulation

State space models are always built on two different equations, one of which aims for the process, and the other for the measurement noise:

State Equation: $X_t = G_t X_{t-1} + W_t$, where $W_t \sim N(0, w_t)$

Observation Equation: $Y_t = F_t X_t + V_t$, where $V_t \sim N(0, v_t)$

All matrices in this model, i.e. G_t, F_t, w_t, v_t can be time-varying. However, often they are time-constant, if anything, then F_t is adapting over time.

Note: such models are usually estimated with the Kalman filter.

AR(1) with Measurement Noise

We assume that the true underlying process is an AR(1), i.e.

$$X_t = \alpha_1 X_{t-1} + W_t,$$

where

$W_t \sim N(0, \sigma_W^2)$ are i.i.d. innovations, „process noise".

In practice, we only observe y_t, as realizations of the process

$Y_t = X_t + V_t$, with $V_t \sim N(0, \sigma_V^2)$, i.i.d.

and additionally, the V_t are independent of X_s, W_s for all s,t, thus they are independent „observation white noise".

Applied Time Series Analysis
FS 2014 – Week 13

More Terminology

We call

$$X_t = \alpha_1 X_{t-1} + W_t$$

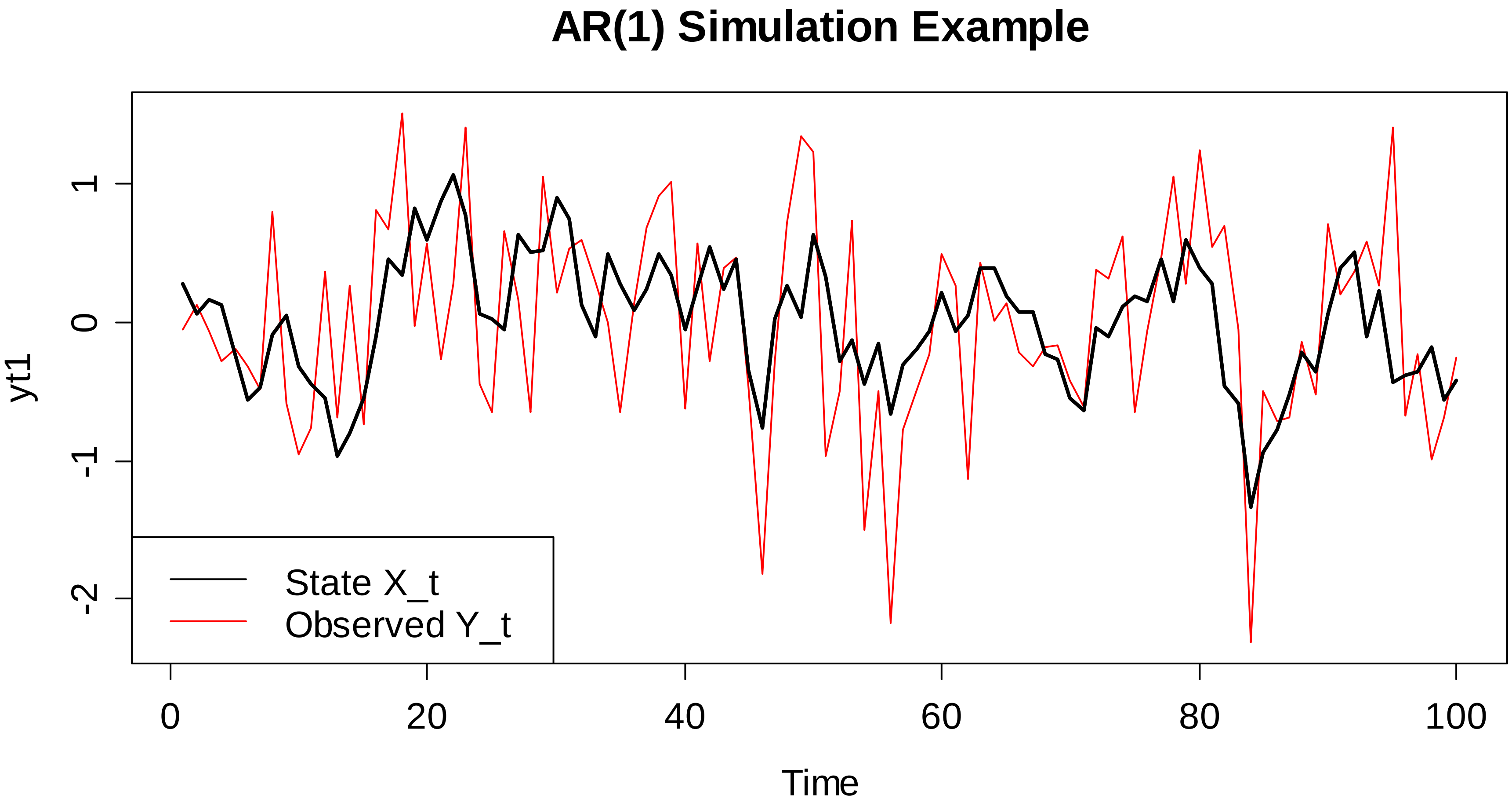
the „state equation“, and

$$Y_t = X_t + V_t$$

the „observation equation“.

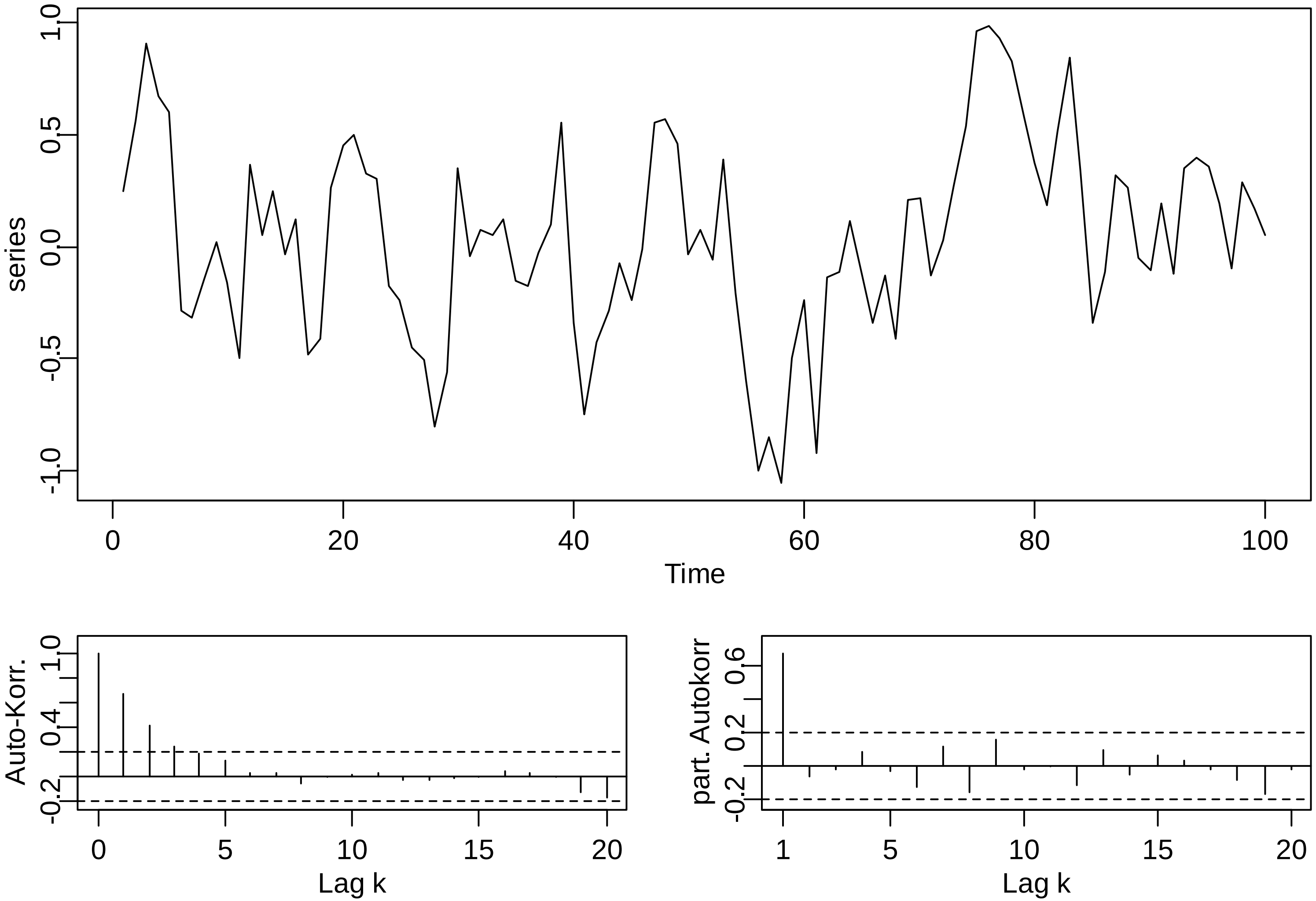
On top of that, we remember once again that the „process noise“ W_t is an innovation that affects all future values X_{t+k} and thus also Y_{t+k}, whereas V_t only influences the current observation Y_t, but no future ones.

AR(1)-Example with α=0.7



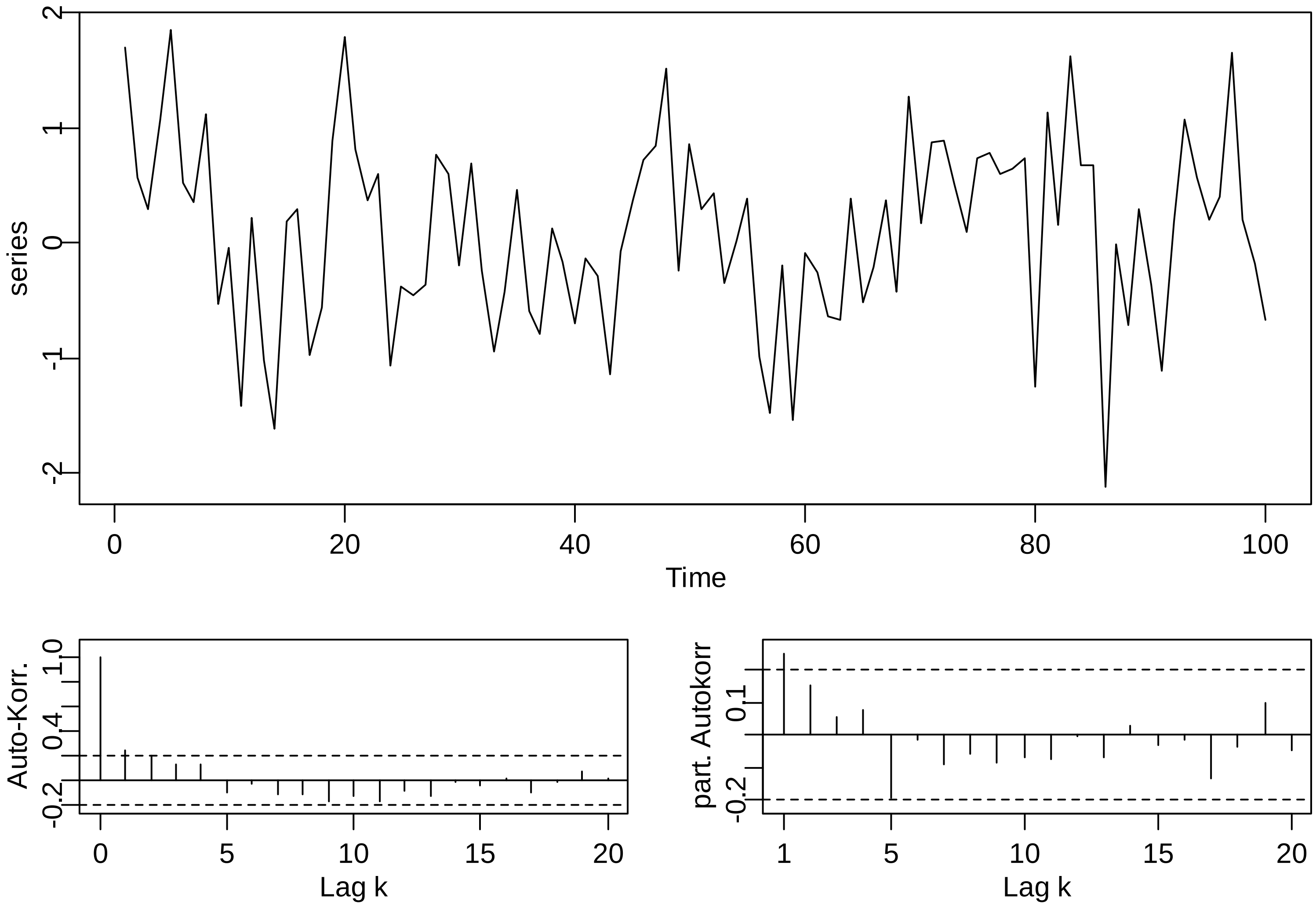
Applied Time Series Analysis
FS 2014 – Week 13

ACF/PACF of X_t



Applied Time Series Analysis
FS 2014 – Week 13

ACF/PACF of Y_t



Applied Time Series Analysis
FS 2014 – Week 13

What is the goal?

The goal of State Space Modeling/Kalman Filtering is:

To uncover the „de-noised“ process X_t from the observed process Y_t.

- The algorithm of Kalman Filtering works with non-stationary time series, too.
- The algorithm is based on a maximum-likelihood-principle where one assume normal distortions.
- There are extensions to multi-dimensional state space models. **See blackboard** for an example how the state space formulation of an AR(2) is set up .

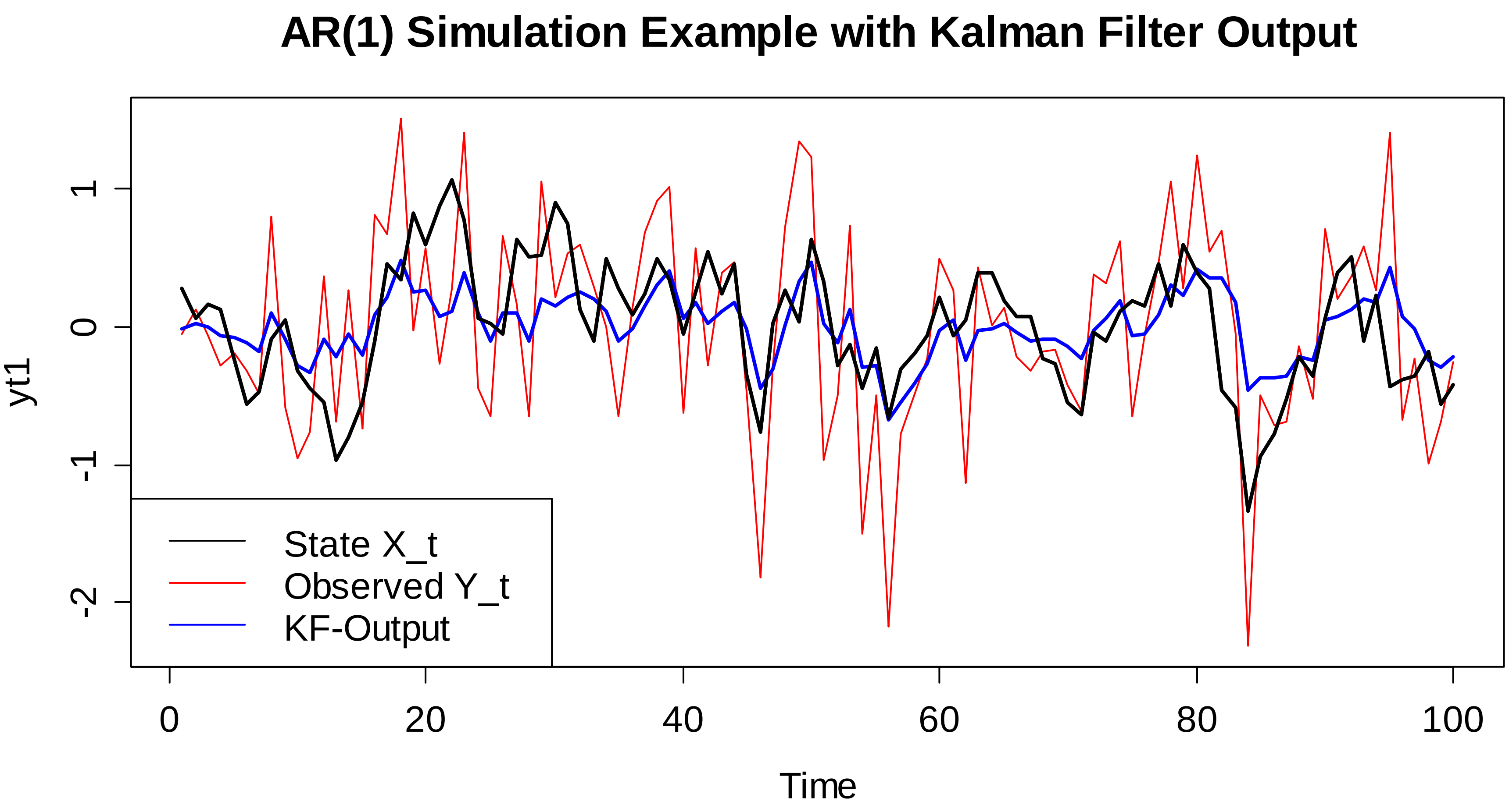
State Space and Kalman Filtering in R

```
## Load the package for Kalman filtering
library(sspir)

## State Space Formulation
ssf <- SS(y = as.matrix(obs),
         Fmat = function(tt,x,phi) { return(matrix(1)) },
         Gmat = function(tt,x,phi) { return(matrix(0.7)) },
         Vmat = function(tt,x,phi) { return(matrix(0.5)) },
         Wmat = function(tt,x,phi) { return(matrix(0.1)) },
         m0 = matrix(0), C0 = matrix(0.1))

## Kalman Filtering
fit  <- kfilter(ssf)
```

Kalman Filter Solution



State Space Formulation of an AR(2)

→ **see blackboard...**

Applied Time Series Analysis
FS 2014 – Week 13

Dynamic Linear Models

In particular: *regression models with time-varying coefficients*

Example: the sales of a housing company depend on the general level of sales in that area at time t, and on the pricing policy at time t.

$$S_t = L_t + \beta_t P_t + V_t$$

This is a regression model with price as the predictor, and the general sales level as the intercept. They are time-varying:

$$L_t = L_{t-1} + \Delta L_t \qquad \beta_t = \beta_{t-1} + \Delta\beta_t$$

Here, $V_t, \Delta L_t, \Delta\beta_t$ are random elements, noise & perturbations

Applied Time Series Analysis
FS 2014 – Week 13

Simulation Example

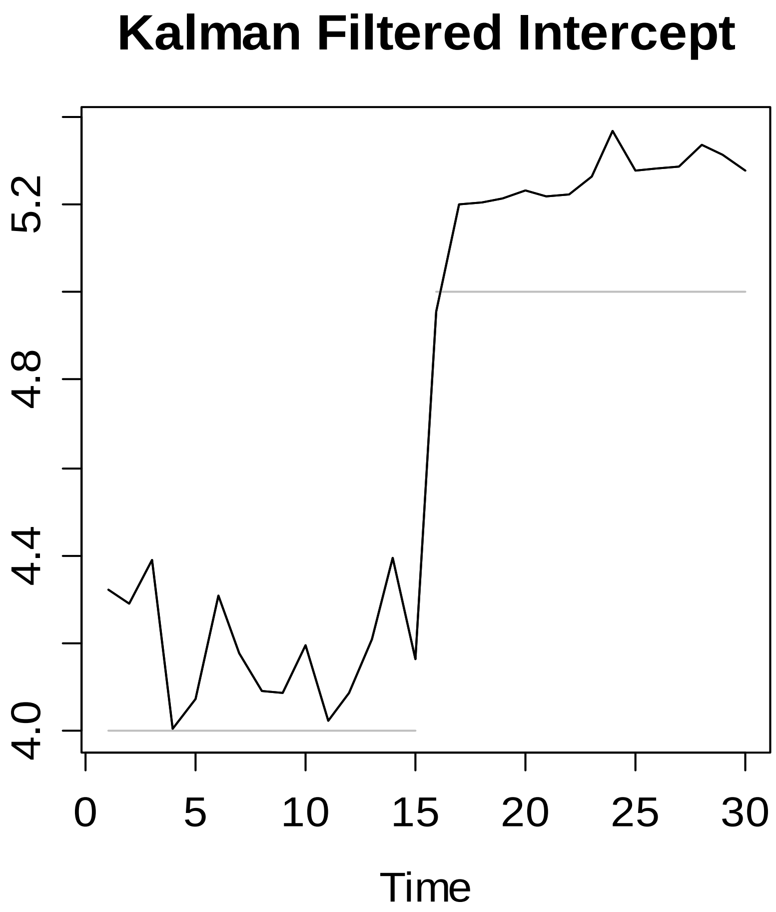
→ **see blackboard...**

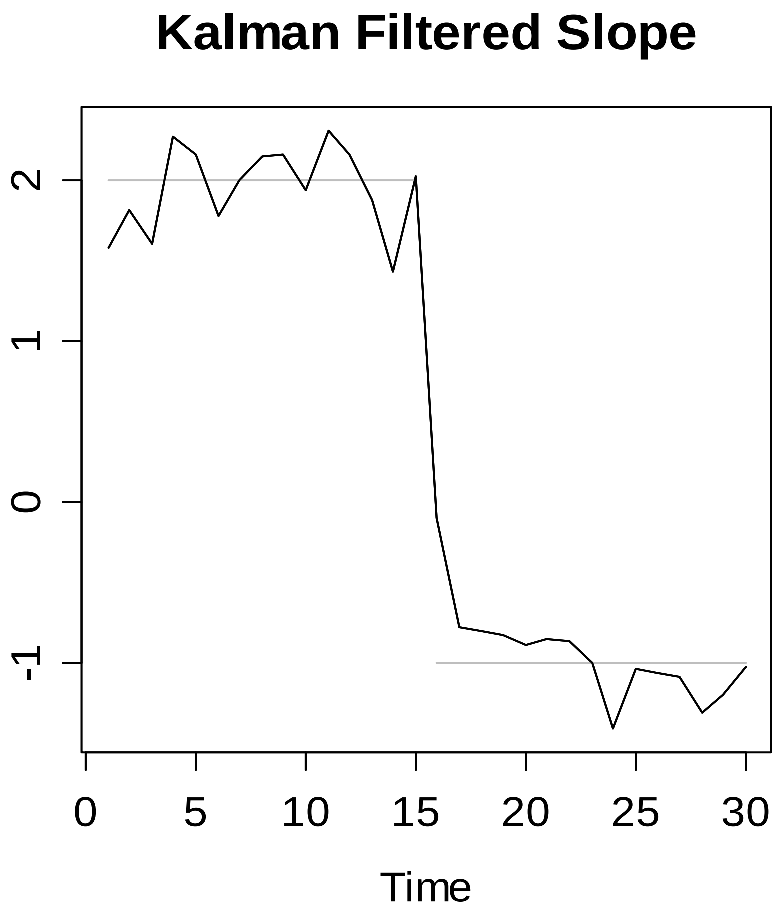
Kalman Filtering for Regression

```
### State Space Formulation
ssf <- SS(y=y.mat, x=x.mat,
          Fmat=function(tt,x,phi)  return(matrix(c(x[tt,1],x[tt,2]),2,1)),
          Gmat=function(tt,x,phi) return(diag(2)),
          Wmat=function(tt,x,phi) return(0.1*diag(2)),
          Vmat=function(tt,x,phi) return(matrix(1)),
          m0=matrix(c(5,3),1,2),C0=10*diag(2))

## Kalman-Filtering
fit <- kfilter(ssf)
plot(fit$m[,1], type="l", xlab="Time", ylab="Intercept")
plot(fit$m[,2], type="l", xlab="Time", ylab="Slope")
```

Kalman Filter Solution





Summary of Kalman Filtering

Summary:

1) The Kalman Filter is a recursive algorithm

2) It relies on an update idea, i.e. we update the forecast $\hat{X}_{t+1,t}$ with the difference $(y_{t+1} - \hat{Y}_{t+1,t})$.

3) The weight of the update is determined by the relation between the process variance σ_W^2 and the measurement noise σ_V^2.

4) This relies on the knowledge of G, F, σ_W^2, σ_V^2. In R we have procedures where everything is estimated simultaneously.

Additional Remarks

1) For the recursive approach of Kalman filtering, initial values are necessary. Their choice is not crucial, their influence cancels out rapidly.

2) The procedures yield forecast and filter intervals:
$\hat{X}_{t+1,t} \pm 1.96 \cdot \sqrt{R_{t+1,t}}$ and $\hat{X}_{t+1,t+1} \pm 1.96 \cdot \sqrt{R_{t+1,t+1}}$

3) State space models are a very rich class. Every ARIMA(p,d,q) can be written in state space form, and the Kalman filter can be used for estimating the coefficients.