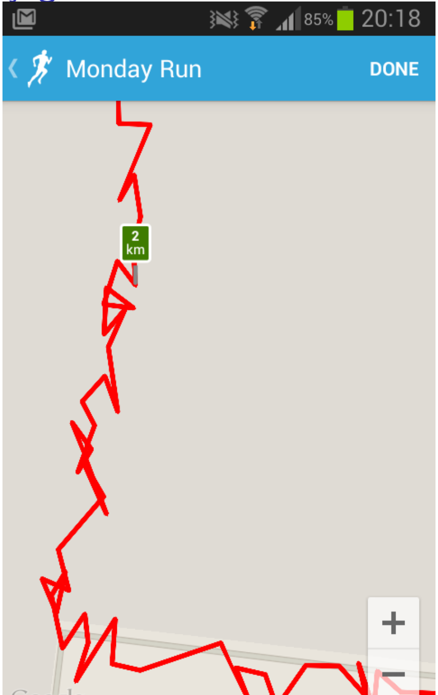
Time Series Analysis

Lasse Engbo Christiansen

Department of Applied Mathematics and Computer Science
Technical University of Denmark

December 1, 2017

Where did he actually go?



Outline of the lecture

State space models, 2nd part:

- ARMA-models on state space form, Sec. 10.4
- Example: Random walk with measurement noise
- The Kalman filter when some observations are missing
- ML-estimates in state space models, Sec. 10.6
- Time-varying systems
- Example AR(1) through measurement noise

Cursory material:

- Signal extraction, Sec. 10.4.1
- Time series with missing observations, Sec. 10.5

The linear stochastic state space model

$$\begin{aligned} \text{System equation:} \quad & \boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{B}\boldsymbol{u}_{t-1} + \boldsymbol{e}_{1,t} \\ \text{Observation equation:} \quad & \boldsymbol{Y}_t = \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{e}_{2,t} \end{aligned}$$

- $\boldsymbol{X}$: State vector
- $\boldsymbol{Y}$: Observation vector
- $\boldsymbol{u}$: Input vector
- $\boldsymbol{e}_1$: System noise
- $\boldsymbol{e}_2$: Observation noise

- $\dim(X_t) = m$ is called the order of the system
- $\{\boldsymbol{e}_{1,t}\}$ and $\{\boldsymbol{e}_{2,t}\}$ mutually independent white noise
- $V[\boldsymbol{e}_1] = \boldsymbol{\Sigma}_1$, $V[\boldsymbol{e}_2] = \boldsymbol{\Sigma}_2$
- $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_2$ are **known** matrices

- The state vector contains all information available for future evaluation; the state vector is a *Markov process*.

The ARMA(p, q) model as a state space model

$$Y_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

State space form:

$$\begin{aligned} \boldsymbol{X}_t &= \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{G}\boldsymbol{\varepsilon}_t \\ \boldsymbol{Y}_t &= \boldsymbol{C}\boldsymbol{X}_t \end{aligned}$$

Or:

$$\boldsymbol{X}_t = \begin{bmatrix} -\phi_1 & 1 & 0 & \cdots & 0 \\ -\phi_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_{d-1} & 0 & 0 & 0 & 1 \\ -\phi_d & 0 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{X}_{t-1} + \begin{pmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{d-1} \end{pmatrix} \boldsymbol{\varepsilon}_t$$

$C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$

where $d = max(p, q+1)$ and any extra parameter is fixed to zero.

For multivariate processes, just plug in matrices, and use I in stead of 1.

Random walk with measurement noise

Consider the state space model

$$
\begin{aligned}
X_t &= X_{t-1} + \eta_t \\
Y_t &= X_t + \varepsilon_t
\end{aligned}
$$

where $\{\eta_t\}$ and $\{\varepsilon_t\}$ are white noise processes with $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

- $\{X_t\}$ is a random walk, that is not directly observed.
- The observations, $\{Y_t\}$, are influenced by measurement noise.
- What is the ARIMA structure of the Y process?
- Hint:

$$\nabla Y_t = \nabla X_t + \nabla \varepsilon_t = \eta_t + \varepsilon_t - \varepsilon_{t-1}$$

Random walk with measurement noise II

$$\nabla Y_t = \eta_t + \varepsilon_t - \varepsilon_{t-1}$$

ACF for ∇Y_t:

$$\rho(k) = \begin{cases} 1 & k = 0 \\ -\sigma_\varepsilon^2/(\sigma_\eta^2 + 2\sigma_\varepsilon^2) & k = 1 \\ 0 & k > 1 \end{cases}$$

This is the ACF of an $MA(1)$ process; thus, Y is $IMA(1,1)$
Alternative formulation:

$$\nabla Y_t = \xi_t + \theta_1 \xi_{t-1}\,, \quad \theta_1 < 0\,,$$

where ξ is white noise with variance σ_ξ^2.

Random walk with measurement noise III

Parameter relations in the two formulations, found by equaling the ACF expressions:

$$
\begin{aligned}
(1+\theta_1^2)\sigma_\xi^2 =& \sigma_\eta^2 + 2\sigma_\varepsilon^2 \\
\theta_1\sigma_\xi^2 =& -\sigma_\varepsilon^2
\end{aligned}
$$

- The ARMA process coefficients for the MA-parts are covariance parameters in the State Space formulation.
- The ARMA representation may be used to derive estimates for Σ_1, Σ_2.

The linear stochastic state space model

$$\text{System equation:} \quad \boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{B}\boldsymbol{u}_{t-1} + \boldsymbol{e}_{1,t}$$
$$\text{Observation equation:} \quad \boldsymbol{Y}_t = \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{e}_{2,t}$$

- $\boldsymbol{X}$: State vector
- $\boldsymbol{Y}$: Observation vector
- $\boldsymbol{u}$: Input vector
- $\boldsymbol{e}_1$: System noise
- $\boldsymbol{e}_2$: Observation noise

- $\dim(X_t) = m$ is called the order of the system
- $\{\boldsymbol{e}_{1,t}\}$ and $\{\boldsymbol{e}_{2,t}\}$ mutually independent white noise
- $V[\boldsymbol{e}_1] = \boldsymbol{\Sigma}_1$, $V[\boldsymbol{e}_2] = \boldsymbol{\Sigma}_2$
- $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_2$ are **known** matrices

The Kalman filter

Initialization

$$\widehat{\boldsymbol{X}}_{1|0} = E\left[\boldsymbol{X}_1\right] = \boldsymbol{\mu}_0$$
$$\boldsymbol{\Sigma}^{xx}_{1|0} = V\left[\boldsymbol{X}_1\right] = \boldsymbol{V}_0$$
$$\boldsymbol{\Sigma}^{yy}_{1|0} = \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{1|0}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2$$

For: $t = 1, 2, 3, \ldots$

Reconstruction:

$$\boxed{\begin{aligned} \boldsymbol{K}_t &= \boldsymbol{\Sigma}^{xx}_{t|t-1}\boldsymbol{C}^T\left(\boldsymbol{\Sigma}^{yy}_{t|t-1}\right)^{-1} \\ \widehat{\boldsymbol{X}}_{t|t} &= \widehat{\boldsymbol{X}}_{t|t-1} + \boldsymbol{K}_t\left(\boldsymbol{Y}_t - \boldsymbol{C}\widehat{\boldsymbol{X}}_{t|t-1}\right) \\ \boldsymbol{\Sigma}^{xx}_{t|t} &= \boldsymbol{\Sigma}^{xx}_{t|t-1} - \boldsymbol{K}_t\boldsymbol{\Sigma}^{yy}_{t|t-1}\boldsymbol{K}_t^T \end{aligned}}$$

Prediction:

$$\boxed{\begin{aligned} \widehat{\boldsymbol{X}}_{t+1|t} &= \boldsymbol{A}\widehat{\boldsymbol{X}}_{t|t} + \boldsymbol{B}\boldsymbol{u}_t \\ \boldsymbol{\Sigma}^{xx}_{t+1|t} &= \boldsymbol{A}\boldsymbol{\Sigma}^{xx}_{t|t}\boldsymbol{A}^T + \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}^{yy}_{t+1|t} &= \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{t+1|t}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2 \end{aligned}}$$

- What happens if the observation Y_t is missing for some t?

Estimation in ARMA(p, q)-models using the KF

- Using the Kalman filter we can get the mean and variance of the one-step predictions of the observations:

$$\widehat{\boldsymbol{Y}}_{t+1|t} = \boldsymbol{C}\widehat{\boldsymbol{X}}_{t+1|t}$$
$$\boldsymbol{\Sigma}^{yy}_{t+1|t} = \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{t+1|t}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2$$

- The Kalman filter can handle missing observations
- An ARMA(p, q)-model can be written as a state space model

- This gives us a way of calculating ML-estimates in the ARMA(p, q)-model even when some observations are missing.

The ARMA(p, q) model as a state space model

$$Y_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

State space form:

$$\begin{aligned} \boldsymbol{X}_t &= \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t \\ \boldsymbol{Y}_t &= \boldsymbol{C}\boldsymbol{X}_t \end{aligned}$$

For $p \geq q$:

$$\boldsymbol{X}_t = \begin{bmatrix} -\phi_1 & 1 & 0 & \cdots & 0 \\ -\phi_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_{d-1} & 0 & 0 & 0 & 1 \\ -\phi_d & 0 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{X}_{t-1} + \begin{pmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{d-1} \end{pmatrix} \boldsymbol{\varepsilon}_t$$

$C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$
where $d = max(p, q+1)$ and any extra parameter is fixed to zero.

ML-estimates in state space models

$$\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{G}\boldsymbol{e}_{1,t}$$
$$\boldsymbol{Y}_t = \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{e}_{2,t}$$

- $\{\boldsymbol{e}_{1,t}\}$ and $\{\boldsymbol{e}_{2,t}\}$ are mutually uncorrelated normally distributed white noise
- $V(\boldsymbol{e}_{1,t}) = \Sigma_1$ and $V(\boldsymbol{e}_{2,t}) = \Sigma_2$
- For ARMA(p, q)-models we have $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{G}$ as stated on the previous slide. Furthermore, $\boldsymbol{e}_{1,t} = \varepsilon_t$, $\boldsymbol{\Sigma}_1 = \sigma_\varepsilon^2$, and $\boldsymbol{\Sigma}_2 = 0$

Maximum Likelihood Estimates

- Let $\mathcal{Y}_{N^*}$ contain the available observations and let $\boldsymbol{\theta}$ contain the parameters of the model
- The likelihood function is the density of the random vector corresponding to the observations and given the set of parameters:

$$L\left(\boldsymbol{\theta}; \mathcal{Y}_{N^*}\right) = f\left(\mathcal{Y}_{N^*} | \boldsymbol{\theta}\right)$$

- The ML-estimates is found by selecting $\boldsymbol{\theta}$ so that the density function is as large as possible at the actual observations
- The random variables $\boldsymbol{Y}_{N^*} | \mathcal{Y}_{N^*-1}$ and $\mathcal{Y}_{N^*-1}$ are independent:

$$\begin{aligned} L\left(\boldsymbol{\theta}; \mathcal{Y}_{N^*}\right) = f\left(\mathcal{Y}_{N^*} | \boldsymbol{\theta}\right) &= f\left(\boldsymbol{Y}_{N^*} | \mathcal{Y}_{N^*-1}, \boldsymbol{\theta}\right) f\left(\mathcal{Y}_{N^*-1} | \boldsymbol{\theta}\right) \\ &= f\left(\boldsymbol{Y}_{N^*} | \mathcal{Y}_{N^*-1}, \boldsymbol{\theta}\right) f\left(\boldsymbol{Y}_{N^*-1} | \mathcal{Y}_{N^*-2}, \boldsymbol{\theta}\right) \cdots f\left(\boldsymbol{Y}_1 | \boldsymbol{\theta}\right) \end{aligned}$$

- The conditional densities can be found using the Kalman filter

MLE / KF – Prediction

- Assume that at time t we have:

$$\widehat{\boldsymbol{X}}_{t|t} = E\left[\boldsymbol{X}_t|\mathcal{Y}_t\right] \quad \text{and} \quad \boldsymbol{\Sigma}^{xx}_{t|t} = V\left[\boldsymbol{X}_t|\mathcal{Y}_t\right]$$

- Using the model we obtain predictions for time $t+1$:

$$\begin{aligned}
\widehat{\boldsymbol{X}}_{t+1|t} &= \boldsymbol{A}\widehat{\boldsymbol{X}}_{t|t} \\
\boldsymbol{\Sigma}^{xx}_{t+1|t} &= \boldsymbol{A}\boldsymbol{\Sigma}^{xx}_{t|t}\boldsymbol{A}^T + \boldsymbol{G}\boldsymbol{\Sigma}_1\boldsymbol{G}^T \\
\widehat{\boldsymbol{Y}}_{t+1|t} &= \boldsymbol{C}\widehat{\boldsymbol{X}}_{t+1|t} \\
\boldsymbol{\Sigma}^{yy}_{t+1|t} &= \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{t+1|t}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2
\end{aligned}$$

- Due to the normality of the white noise process $f(\boldsymbol{Y}_{t+1}|\mathcal{Y}_t, \boldsymbol{\theta})$ is then the (multivariate) normal density (see Chapter 2) with mean $\widehat{\boldsymbol{Y}}_{t+1|t}$ and variance-covariance $\boldsymbol{\Sigma}^{yy}_{t+1|t}$ $(= \boldsymbol{R}_{t+1})$

MLE / KF – Reconstruction

At time $t+1$ there are two possibilities for the reconstruction part:

The observation $\boldsymbol{Y}_{t+1}$ is available:

We update the state estimate using the reconstruction step of the Kalman Filter:

$$\begin{aligned}
\boldsymbol{K}_{t+1} &= \boldsymbol{\Sigma}^{xx}_{t+1|t}\boldsymbol{C}^T\left(\boldsymbol{\Sigma}^{yy}_{t+1|t}\right)^{-1} \\
\widehat{\boldsymbol{X}}_{t+1|t+1} &= \widehat{\boldsymbol{X}}_{t+1|t} + \boldsymbol{K}_{t+1}\left(\boldsymbol{Y}_{t+1} - \widehat{\boldsymbol{Y}}_{t+1|t}\right) \\
\boldsymbol{\Sigma}^{xx}_{t+1|t+1} &= \boldsymbol{\Sigma}^{xx}_{t+1|t} - \boldsymbol{K}_{t+1}\boldsymbol{\Sigma}^{yy}_{t+1|t}\boldsymbol{K}^T_{t+1}
\end{aligned}$$

The observation $\boldsymbol{Y}_{t+1}$ is missing:

We get no new information and we use:

$$\begin{aligned}
\widehat{\boldsymbol{X}}_{t+1|t+1} &= \widehat{\boldsymbol{X}}_{t+1|t} \\
\boldsymbol{\Sigma}^{xx}_{t+1|t+1} &= \boldsymbol{\Sigma}^{xx}_{t+1|t}
\end{aligned}$$

MLE / KF – The likelihood function

- Using the prediction errors and variances

$$\widetilde{\boldsymbol{Y}}_i = \boldsymbol{Y}_i - \widehat{\boldsymbol{Y}}_{i|i-1}$$
$$\boldsymbol{R}_i = \boldsymbol{\Sigma}^{yy}_{i|i-1}$$

- The likelihood function can be expressed as

$$L\left(\boldsymbol{\theta}; \mathcal{Y}_{N^*}\right) = \prod_{i=1}^{N^*} \left[(2\pi)^m \det \boldsymbol{R}_i\right]^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\widetilde{\boldsymbol{Y}}_i^T \boldsymbol{R}_i^{-1} \widetilde{\boldsymbol{Y}}_i\right]$$

- In practice optimization is based on

$$\log L\left(\boldsymbol{\theta}; \mathcal{Y}_{N^*}\right) = -\frac{1}{2} \sum_{i=1}^{N} \left(\log \det \boldsymbol{R}_i + \widetilde{\boldsymbol{Y}}_i^T \boldsymbol{R}_i^{-1} \widetilde{\boldsymbol{Y}}_i\right) + c$$

- The variance of the estimates can be approximated by the 2nd order derivatives of the log-likelihood.

MLE / KF IV – Initialization

- The only outstanding issue is "prediction" of $\boldsymbol{Y}_1$, i.e. calculation of $\widehat{\boldsymbol{Y}}_{1|0}$
- This can be done by setting $\widehat{\boldsymbol{X}}_{0|0} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{0|0}^{xx} = \alpha \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix and α is a 'large' constant (we don't know what it is)
- Alternatively, we can fix the initial state $\widehat{\boldsymbol{X}}_{0|0}$ and set $\boldsymbol{\Sigma}_{0|0}^{xx} = \mathbf{0}$, whereby $\boldsymbol{\Sigma}_{1|0}^{xx} = \boldsymbol{G}\boldsymbol{\Sigma}_1\boldsymbol{G}^T$
- Or combinations thereof - recommended
- The important part is that the (un-)certainty of $\widehat{\boldsymbol{X}}_{0|0}$ is reflected in $\boldsymbol{\Sigma}_{0|0}^{xx}$.

Autocovariance functions with missing data

Define the observation indicator a as

$$a(t) = \begin{cases} 1 & \text{if } Y_t \text{ is observed;} \\ 0 & \text{if } Y_t \text{ is missing.} \end{cases}$$

Define similarly

$$C_a(k) = \frac{1}{N} \sum_{t=1}^{N-|k|} a(t)a(t+|k|)\,.$$

- a indicates which data points are present
- For large N, $C_a(k)$ measures the fraction of pairs of data k time steps away from each other that are observed.

Autocovariance functions with missing data II

The indicator, $a(t)$ is used to define and estimate of the mean of $\{Y_t\}$

$$\overline{\mu}_y = \frac{\sum_{t=1}^{N} a(t)Y_t}{\sum_{t=1}^{N} a(t)}$$

- $\overline{\mu}_y$ is the mean of the observed Y_t's.

And defining:

$$C_a^{\square}(k) = \frac{1}{N} \sum_{t=1}^{N-|k|} a(t)a(t+|k|)(Y_t - \overline{\mu}_y)(Y_{t+|k|} - \overline{\mu}_y)\,.$$

- Note: $C_a^{\square}(k)$ leave pairs out if one of the observations is missing.
- The sample autocovariance is estimated by:

$$C_{YY}(k) = \frac{C_a^{\square}(k)}{C_a(k)}$$

- These estimates are available with the `acf` function in R, by using
```
> acf(...,na.action=na.omit)
```

Time-varying systems

$$\begin{aligned} \text{System equation:} \quad \boldsymbol{X}_t &= \boldsymbol{A}_t \boldsymbol{X}_{t-1} + \boldsymbol{B}_t \boldsymbol{u}_{t-1} + \boldsymbol{e}_{1,t} \\ \text{Observation equation:} \quad \boldsymbol{Y}_t &= \boldsymbol{C}_t \boldsymbol{X}_t + \boldsymbol{e}_{2,t} \end{aligned}$$

- $\boldsymbol{X}$: State vector
- $\boldsymbol{Y}$: Observation vector
- $\boldsymbol{u}$: Input vector
- $\boldsymbol{e}_1$: System noise
- $\boldsymbol{e}_2$: Observation noise

- $\dim(X_t) = m$ is called the order of the system
- $\{\boldsymbol{e}_{1,t}\}$ and $\{\boldsymbol{e}_{2,t}\}$ mutually independent white noise
- $V[\boldsymbol{e}_{1,t}] = \boldsymbol{\Sigma}_{1,t}$, $V[\boldsymbol{e}_{2,t}] = \boldsymbol{\Sigma}_{2,t}$
- $\boldsymbol{A}_t$, $\boldsymbol{B}_t$, $\boldsymbol{C}_t$, $\boldsymbol{\Sigma}_{1,t}$, and $\boldsymbol{\Sigma}_{2,t}$ are **known** matrices at any point in time

The Kalman filter for time varying systems

Initialization

$$\widehat{\boldsymbol{X}}_{1|0} = E\left[\boldsymbol{X}_1\right] = \boldsymbol{\mu}_0$$
$$\boldsymbol{\Sigma}^{xx}_{1|0} = V\left[\boldsymbol{X}_1\right] = \boldsymbol{V}_0 \Rightarrow$$
$$\boldsymbol{\Sigma}^{yy}_{1|0} = \boldsymbol{C}_1 \boldsymbol{\Sigma}^{xx}_{1|0} \boldsymbol{C}_1^T + \boldsymbol{\Sigma}_{2,1}$$

For: $t = 1, 2, 3, \ldots$

Reconstruction:

$$\boxed{\begin{aligned} \boldsymbol{K}_t &= \boldsymbol{\Sigma}^{xx}_{t|t-1} \boldsymbol{C}_t^T \left(\boldsymbol{\Sigma}^{yy}_{t|t-1}\right)^{-1} \\ \widehat{\boldsymbol{X}}_{t|t} &= \widehat{\boldsymbol{X}}_{t|t-1} + \boldsymbol{K}_t \left(\boldsymbol{Y}_t - \boldsymbol{C}_t \widehat{\boldsymbol{X}}_{t|t-1}\right) \\ \boldsymbol{\Sigma}^{xx}_{t|t} &= \boldsymbol{\Sigma}^{xx}_{t|t-1} - \boldsymbol{K}_t \boldsymbol{\Sigma}^{yy}_{t|t-1} \boldsymbol{K}_t^T \end{aligned}}$$

Prediction:

$$\boxed{\begin{aligned} \widehat{\boldsymbol{X}}_{t+1|t} &= \boldsymbol{A}_{t+1} \widehat{\boldsymbol{X}}_{t|t} + \boldsymbol{B}_{t+1} \boldsymbol{u}_t \\ \boldsymbol{\Sigma}^{xx}_{t+1|t} &= \boldsymbol{A}_{t+1} \boldsymbol{\Sigma}^{xx}_{t|t} \boldsymbol{A}_{t+1}^T + \boldsymbol{\Sigma}_{1,t+1} \\ \boldsymbol{\Sigma}^{yy}_{t+1|t} &= \boldsymbol{C}_{t+1} \boldsymbol{\Sigma}^{xx}_{t+1|t} \boldsymbol{C}_{t+1}^T + \boldsymbol{\Sigma}_{2,t+1} \end{aligned}}$$

Example: An AR(1) and obs. noise

The heat transfer from a certain body to its surroundings is dominated by conduction. The temperature of the body is given by

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{1}{R}(T_{surr} - T)$$

Let the surrounding temperature be constantly 0. Then

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{1}{R}(T) = aT$$

The solution to the differential equation is

$$T = T_0 e^{at}$$

A discretization of this yields

$$T_{t+1} = e^{a\Delta t} \cdot T_t$$

We know in reality, such a process is influenced by noise:

$$T_{t+1} = -\phi \cdot T_t + e_t, \quad e_t \sim \mathcal{N}(0, \sigma_e^2)$$

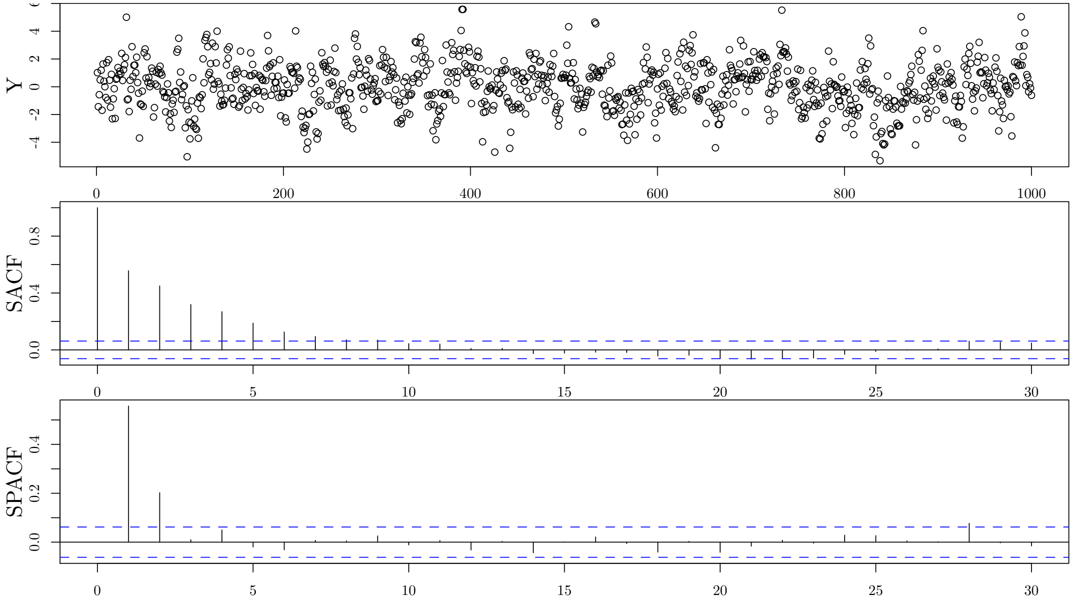
We measure the temperature at each time step:

$$Y_t = T_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$$

Example: An AR(1) and obs. noise – Simulation

```r
a <- 0.8
N <- 1000
X <- numeric(N)
X[1] <- 1
for (i in 2:N){
    X[i] <- a * X[i-1] + rnorm(1, sd=0.8)
}
Y <- X + rnorm(N, sd=1.0)
```

Example: An AR(1) and obs. noise – What process is this?



Highlights

- ARMA models on State space form
- Kalman filter
  - Handling missing values
  - Prediction
  - Maximum likelihood estimation of parameters
  - Comparing models using likelihood