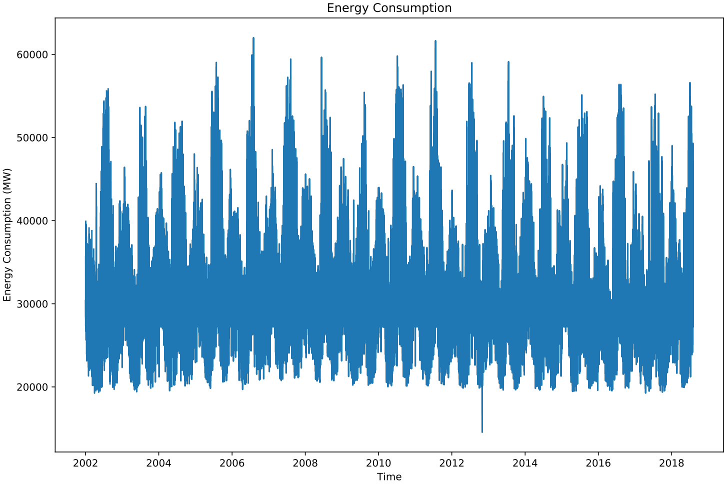
Time Series and Sequence Learning

Discussion seminar for Lecture 4

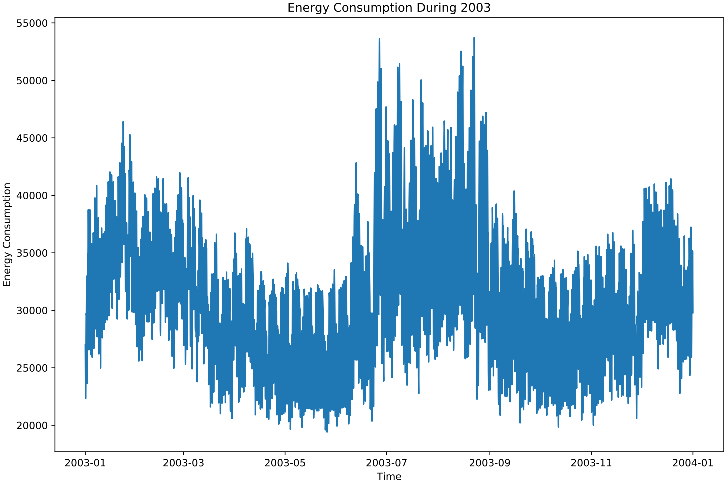
Johan Alenlöv, Linköping University

2022-09-06

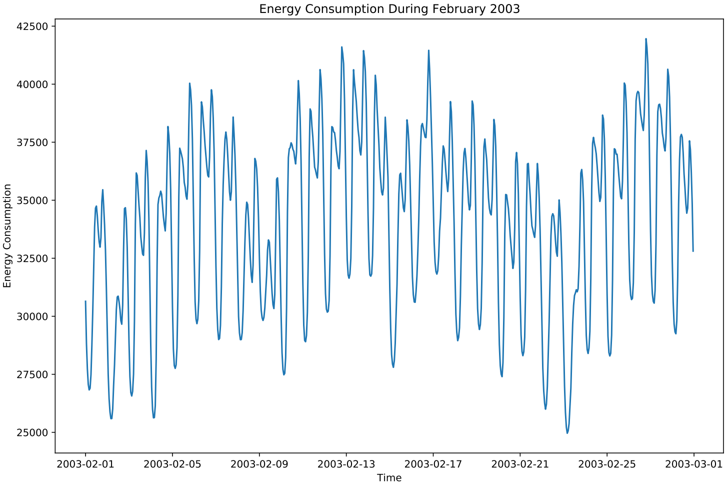
lecture4a - Classical Decomposition



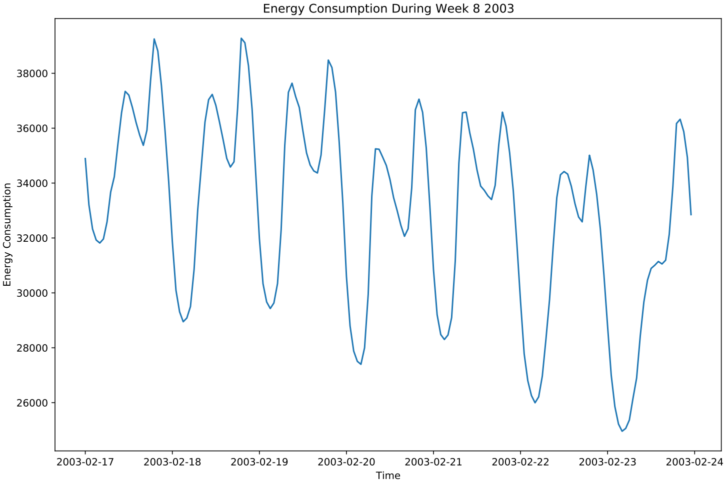
lecture4a - Classical Decomposition



lecture4a - Classical Decomposition



lecture4a - Classical Decomposition



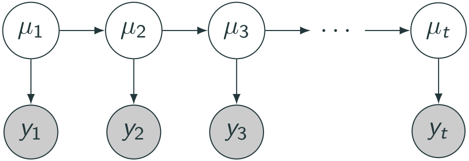
exp(sinx)

Discussion questions:

1. What components can you find in the data? Can you explain them?
2. Would this be an additive or multiplicative model?

The Local Level Model Two stochastic processes $y_1, y_2, y_3, \ldots$ and $\mu_1, \mu_2, \mu_3, \ldots$

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t, && \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2) \\
\mu_{t+1} &= \mu_t + \eta_t, && \eta_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\eta^2) \\
& && \mu_1 \sim \mathcal{N}(a_1, P_1)
\end{aligned}
$$



Discussion questions:

1. In the simple random walk, if $\mu_0 = b$, what is $\mathbb{E}[\mu_t]$ and $\mathrm{Var}[\mu_t]$? What does this imply?
2. Which of the following problems relate to **filtering**, **prediction** and **smoothing**? Also think about what could the state and observations be.
   a) Find the trajectory an airplane based on noisy observations. (Does it change of we do it realtime or if we collect all the data first?)
   b) Make a statement about the weather tomorrow.
   c) Calculate the current position of a robot based on sensors.

Kalman Filter for local-level model

For each iteration $t = 1, 2, 3, \ldots$ repeat the following steps:

- Measurement updates
  1. Forecasting error: $v_t = y_t - \hat{\mu}_{t\,|\,t-1}$
  2. Forecasting variance: $F_t = P_{t\,|\,t-1} + \sigma_\varepsilon^2$
  3. Kalman gain: $K_t = P_{t\,|\,t-1}/F_t$
  4. Filter mean: $\hat{\mu}_{t\,|\,t} = \hat{\mu}_{t\,|\,t-1} + K_t v_t$
  5. Filter variance: $P_{t\,|\,t} = P_{t\,|\,t-1}(1 - K_t)$
- Time updates
  6. Predictor mean: $\hat{\mu}_{t+1\,|\,t} = \hat{\mu}_{t\,|\,t}$
  7. Predictor variance: $P_{t+1\,|\,t} = P_{t\,|\,t} + \sigma_\eta^2$

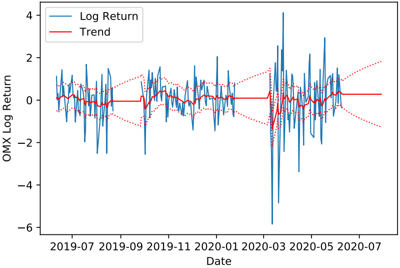
Initialized using $\hat{\mu}_{1\,|\,0} = a_1$ and $P_{1\,|\,0} = P_1$.

Discussion questions:

1. If we change the state equation to $\mu_t = a \cdot \mu_{t-1} + \eta_t$, how would that change the Kalman filter?
2. We are able to calculate the log-likelihood as

$$\ell(\theta) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(\log F_t(\theta) + \frac{(y_t - \hat{\mu}_{t\,|\,t-1}(\theta))^2}{F_t(\theta)}\right),$$

how would you try to maximize this? Is direct calculations of derivatives feasible?



Discussion questions:

1. For us there is a difference to forecasting and missing data, but the Kalman filter does not see any difference. Can you explain this?
2. As for the Kalman filter what would happen if we changed the state equation to $\mu_t = a \cdot \mu_{t-1} + \eta_t$?