Time series and Kalman Filters

Lecture 10

Xiaofei Shi

Learning objectives

- Learn how time series modeling extend from the regression model
- Learn how Dynamic Linear Models are related to time series framework

How do we model the data?

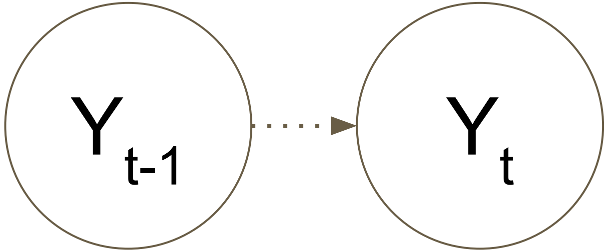
Why it is very hard to model time series data?

How do we model the data?

Why it is very hard to model time series data?

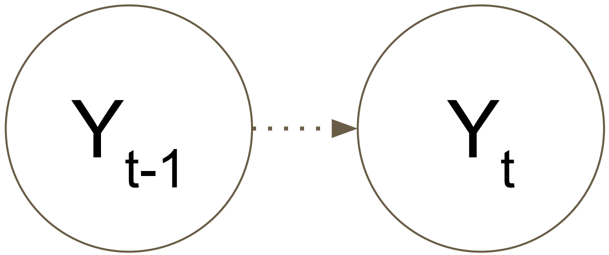
Traditional Time Series Setup

$$Y_t = Y_{t-1} + \epsilon_t$$



Traditional Time Series Setup

$$Y_t = Y_{t-1} + \epsilon_t$$



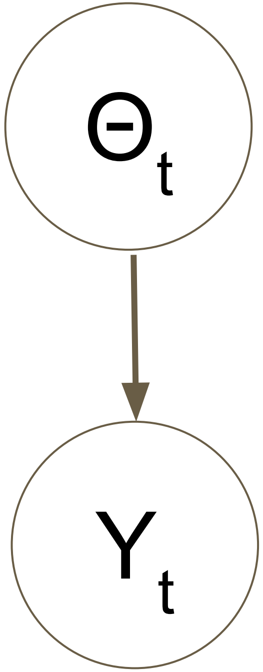
More complicated version:

- Autoregression Model: https://en.wikipedia.org/wiki/Autoregressive_model

- ARMA Model: https://en.wikipedia.org/wiki/Autoregressive–moving-average_model

Linear Regression Model - Data is a Noisy Measurement of the Process/State

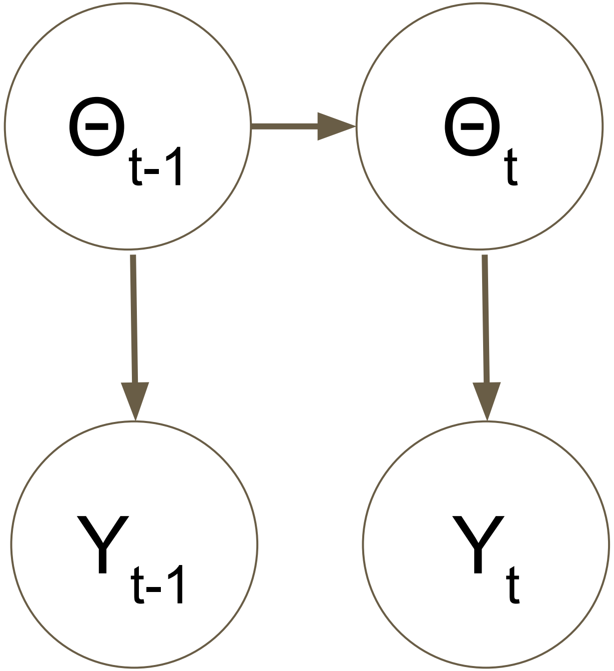
$$Y_t = \theta_t + \upsilon_t$$



To add temporal dependency, make the states depend on its previous one

$$Y_t = \theta_t + v_t$$

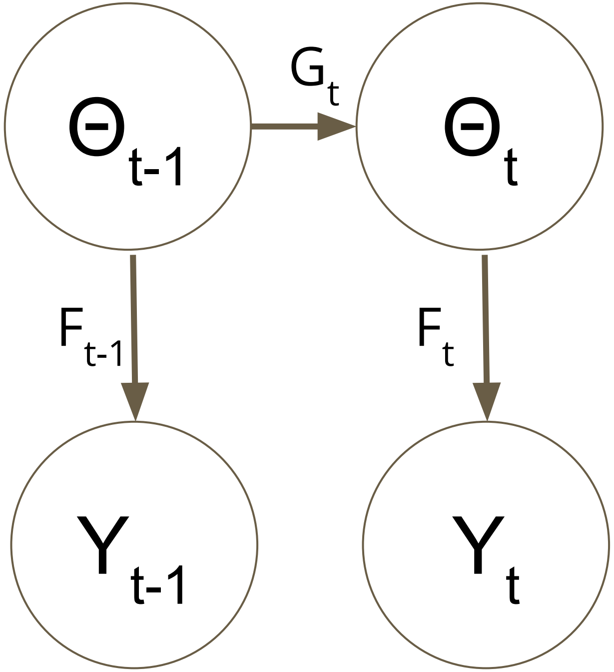
$$\theta_t = \theta_{t-1} + w_t$$



Dynamic Linear Model (State Space Models)

$$Y_t = F_t\theta_t + v_t$$

$$\theta_t = G_t\theta_{t-1} + w_t$$



Error terms can have different meanings

$$Y_t = F_t\theta_t + v_t \qquad v_t \sim N(0, V_t)$$

$$\theta_t = G_t\theta_{t-1} + w_t \qquad w_t \sim N(0, W_t)$$

- Random terms are all independent

Temporal models need to specify the “initial state”

$$Y_t = F_t\theta_t + v_t \qquad v_t \sim N(0, V_t)$$

$$\theta_t = G_t\theta_{t-1} + w_t \qquad w_t \sim N(0, W_t)$$

$$\theta_0 \sim N(m_0, C_0)$$

- Random terms are all independent

What are we interested in the model?

Bayesians always want the posterior!

$$P(\ ?\ | Y_1, \ldots, Y_{t-1})$$

Kalman filter is an algorithm to get the posterior expectation/covariance for "states" in the DLM

$$E(\theta_t | Y_1, \ldots, Y_{t-1})$$

$$Cov(\theta_t | Y_1, \ldots, Y_{t-1})$$

How to spot correlation across records?

If the correlation between Y_t and Y_(t-1) is not 0, they cannot be independent!

- What is the definition of the usual correlation?

Diagnostics via auto-correlation function

If the correlation between Y_t and Y_(t-1) is not 0, they cannot be independent!

$$\sum (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})$$

Diagnostics via auto-correlation function

If the correlation between Y_t and Y_(t-1) is not 0, they cannot be independent!

$$\sum_{t=k-1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})$$

Diagnostics via auto-correlation function

If the correlation between Y_t and Y_(t-1) is not 0, they cannot be independent!

$$\frac{1}{(n-k)} \sum_{t=k-1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})$$

Diagnostics via auto-correlation function

If the correlation between Y_t and Y_(t-1) is not 0, they cannot be independent!

$$\frac{1}{(n-k)\sigma^2}\sum_{t=k-1}^{n}(Y_t-\bar{Y})(Y_{t-k}-\bar{Y})$$

Reading the autocorrelation function

acf() in R!