Probabilistic Time Series Analysis: Lab 2

Tim Kunisky

September 12, 2018

Some Organizational Details

- **Office hours.** Unless otherwise announced on Piazza, my office hours will be in room 60FA-C15 (on this floor) from 6:10 to 7:10 on Wednesdays, immediately after this lab.

- **Coding assignments.** Today will be our first lab with a deliverable coding assignment. You will need to fill in some (small) pieces of code in an iPython notebook (available on the Github) and submit it on NYU Classes by the beginning of lab next week.

Time Series Basics Review

Let w_t be independent with $\mathbb{E}w_t = 0$ and $\mathbb{E}w_t^2 = \sigma^2$.

1. Let $x_t = w_{t-1} + 2w_t + w_{t+1}$.
   - What is the ACF of x_t?
   - Is x_t (weakly) stationary?
   - If not, can you transform it to be stationary?

2. Let $x_t = \delta + x_{t-1} + w_t$.
   - What is the ACF of x_t?
   - Is x_t (weakly) stationary?
   - If not, can you transform it to be stationary?

Parameter Redundancy in ARMA Models

Parameter redundancy is when a higher-order model is really representing the same process as a lower-order model.

1. Let w_t be white noise, $x_t = 0.5x_{t-1} - 0.5w_{t-1} + w_t$.

   Is x_t the same as w_t (if, say, $x_0 = w_0$)?

2. Let $w_t \sim \mathcal{N}(0, 25)$, $x_t = w_t + 0.2w_{t-1}$,
   and $v_t \sim \mathcal{N}(0, 1)$, $y_t = v_t + 5v_{t-1}$.

   Is x_t the same as y_t?

The Backshift Operator

- We define the *backshift operator* $\boldsymbol{B}$, acting on time series $\boldsymbol{x} = (x_t)$, to shift a series backwards one step:

$$(\boldsymbol{B}\boldsymbol{x})_t = x_{t-1}.$$

- If $\boldsymbol{x}$ were finite and written as a vector, what kind of object would $\boldsymbol{B}$ be? **A matrix.**
- ARMA models can be written as polynomial equations in $\boldsymbol{B}$:

$$\phi(\boldsymbol{B})\boldsymbol{x} = \theta(\boldsymbol{B})\boldsymbol{w}.$$

  Let's do this for $x_t = 0.5x_{t-1} - 0.5w_{t-1} + w_t$. What do you notice?
- **General principle: redundancy corresponds to cancellations in the ARMA polynomial equation.**

Causality and Invertibility

- AR, MA, and ARMA describe certain *linear processes*:

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \text{ or } \boldsymbol{x} = \boldsymbol{\psi} * \boldsymbol{w},$$

  with ψ_j a summable series.
- Is a linear process always *stationary*?
- A linear process is *causal* if $\psi_j = 0$ for $j < 0$.
- A linear process is *invertible* if:

$$w_t = \sum_{j=0}^{\infty} \pi_j x_{t-j}$$

  with π_j summable ($\boldsymbol{w}$ is a causal process of $\boldsymbol{x}$).
- **These can be tested for an ARMA in terms of ϕ and θ; see Shumway and Stoffer for details.**

Inference for AR Processes

General AR(p) model:

$$x_{t+1} = \phi_1 x_t + \cdots + \phi_p x_{t-p+1} + w_{t+1}$$

- Technique 1: *maximum likelihood* (discussed in lecture).
  - We write out $\log \mathbb{P}[\text{data} \mid \boldsymbol{\phi}]$ (not so bad since w is gaussian) and optimize over $\boldsymbol{\phi}$ using our favorite optimization technique.
  - The optimization problem is hard (very non-convex)! Need either an approximation or a good starting point (basically the same thing).

- Technique 2: *method of moments*.

Inference for AR Processes: Method of Moments

$$x_{t+1} = \phi_1 x_t + \cdots + \phi_p x_{t-p+1} + w_{t+1}$$

- A *moment* is the mean of a product of random variables.
- The *method of moments* is matching the moments the model should have with the empirical moments of the data.

First moments (means) are all zero, but...
Second moments ↔ autocovariances! Compute $c(k)$:

$$\begin{aligned} c(k) &= \mathbb{E}[x_{t-k+1} x_{t+1}] \\ &= \sum_{j=1}^{p} \phi_j \mathbb{E}[x_{t-k+1} x_{t-j+1}] + \mathbb{E}[x_{t-k+1} w_{t+1}] \\ &= \sum_{j=1}^{p} \phi_j c(|j-k|) \end{aligned}$$

Yule-Walker Equations

Dividing through by $c(0)$ to form the autocorrelations $\rho(k) = c(k)/c(0)$, we just derived

$$\begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \cdots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}$$

or just $\boldsymbol{\rho} = \boldsymbol{R}\boldsymbol{\phi}$. We can estimate $\hat{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{R}}$ from the data by taking empirical autocorrelations, and solve the linear system for $\boldsymbol{\phi}$.

How long does it take to solve this system?

Durbin-Levinson Algorithm

To speed this up, instead of solving for the order p AR, we can solve iteratively for order $1, 2, \ldots, p-1, p$.

It turns out that updating to go up by one order is easy (because, as you can check at home, the Yule-Walker equations are just regression of x_t onto $x_{t-1}, \ldots, x_{t-p-1}$, so we are effectively adding one variable to a linear regression).

- $\phi_{0,0} \leftarrow 0$.
- For $i = 1, \ldots, p$:
  - $\phi_{i,i} \leftarrow \frac{\rho(i) - \sum_{k=1}^{i-1} \phi_{i-1,k}\rho(i-k)}{1 - \sum_{k=1}^{i-1} \phi_{i-1,k}\rho(k)}$.
  - For $j = 1, \ldots, i-1$:
    - $\phi_{i,j} \leftarrow \phi_{i-1,j} - \phi_{i,i}\phi_{i-1,i-j}$.

How long does this algorithm take?

Coding Assignment

- Find `labs/lab2-student.ipynb` on the course Github.
- Try running everything, but you will only be graded on what you fill in for the missing bits of code marked with a TODO.
- **If you do not have Python/iPython and the numpy, scipy, and matplotlib libraries installed, and are not comfortable doing these things yourself, come see me now or in office hours today!**