Statistics 434: Bullet Points for Day 7
AR(p) with $p \geq 2$

We now consider the autoregressive models with $p \geq 2$, beginning with a look at the most useful of these — the modest AR(2) model that was made famous by Yule's study of sun spot data.

We need to understand the conditions for stationarity of the AR(p) models, and this will lead us to further consideration of the lag operators. We'll then develop the so-called AR polynomial and make the surprising acquaintance of the unit circle of the complex plane. We will also introduce the Yule-Walker equations — partially historical, partially practical. Finally we'll take up the important *partial autocorrelation function*, which gives us a sloppy but credible way to try to "guess an appropriate value for p."

- The lag operator L and the AR(p) model
  1. AR(2) as the leading case
  2. The AR polynomial $\Phi(L)$
- The Wold Representation for AR(p)
  1. Reviewing the Wold Representation for AR(1)
  2. Doing an AR(2) example in some detail
  3. Criterion for stationarity and the unit circle in the complex plane
  4. How the Wold decomposition helps us "see" stationarity
  5. Interpretation of the "impulse function"
- Fitting the AR(p) model
  1. Derivation of the Yule-Walker equations (of historical interest)
  2. Gaussian Maximum likelihood (what S-Plus uses)
  3. The PACF — the main tool for diagnosing the order

The Parsimony Paradox

With an AR(p) model, the larger we take p to be, the better we can fit our sample data, so why not take a giant p and do a really good job? You have some experience from regression that tells you that this is probably not a good idea — but can you really explain it? Why would you intentionally accept a worse fit for the data at hand? What are you promised in return for accepting a less good fit, and how can you tell if you really "get paid" for taking such a noble action.

Don't feel bad if you don't have good answers. Neither do full-time statisticians. Most of them just trot out some anecdote, and claim that it is a lesson of experience one does well to have a preference for parsimonious models. One can argue little more persuasively than this, but it is hard to do substantially better. It's definitely true that "simpler is better" is one of our most powerful heuristics, but there is still a lot of unexplained mystery why this is so.