Introduction to Time Series Analysis. Lecture 15.

Last lecture: Maximum likelihood estimation

1. Integrated ARMA models
2. Diagnostics
3. Model selection
4. Seasonal ARMA models

Integrated ARMA Models: ARIMA(p,d,q)

For $p, d, q \geq 0$, we say that a time series $\{X_t\}$ is an **ARIMA (p,d,q) process** if $Y_t = \nabla^d X_t = (1 - B)^d X_t$ is ARMA(p,q). We can write

$$\phi(B)(1 - B)^d X_t = \theta(B) W_t.$$

Recall the random walk: $X_t = X_{t-1} + W_t$.
X_t is not stationary, but $Y_t = (1 - B)X_t = W_t$ is a stationary process.
In this case, it is white, so $\{X_t\}$ is an ARIMA(0,1,0).

Also, if X_t contains a trend component plus a stationary process, its first difference is stationary.

ARIMA models example

Suppose $\{X_t\}$ is an ARIMA(0,1,1): $X_t = X_{t-1} + W_t - \theta_1 W_{t-1}$.
If $|\theta_1| < 1$, we can show

$$X_t = \sum_{j=1}^{\infty} (1-\theta_1)\theta_1^{j-1} X_{t-j} + W_t,$$

and so

$$\begin{aligned} \tilde{X}_{n+1} &= \sum_{j=1}^{\infty} (1-\theta_1)\theta_1^{j-1} X_{n+1-j} \\ &= (1-\theta_1)X_n + \sum_{j=2}^{\infty} (1-\theta_1)\theta_1^{j-1} X_{n+1-j} \\ &= (1-\theta_1)X_n + \theta_1 \tilde{X}_n. \end{aligned}$$

Exponentially weighted moving average.

Building ARIMA models

1. Plot the time series.

   Look for trends, seasonal components, step changes, outliers.

2. Nonlinearly transform data, if necessary

3. Identify preliminary values of d, p, and q.

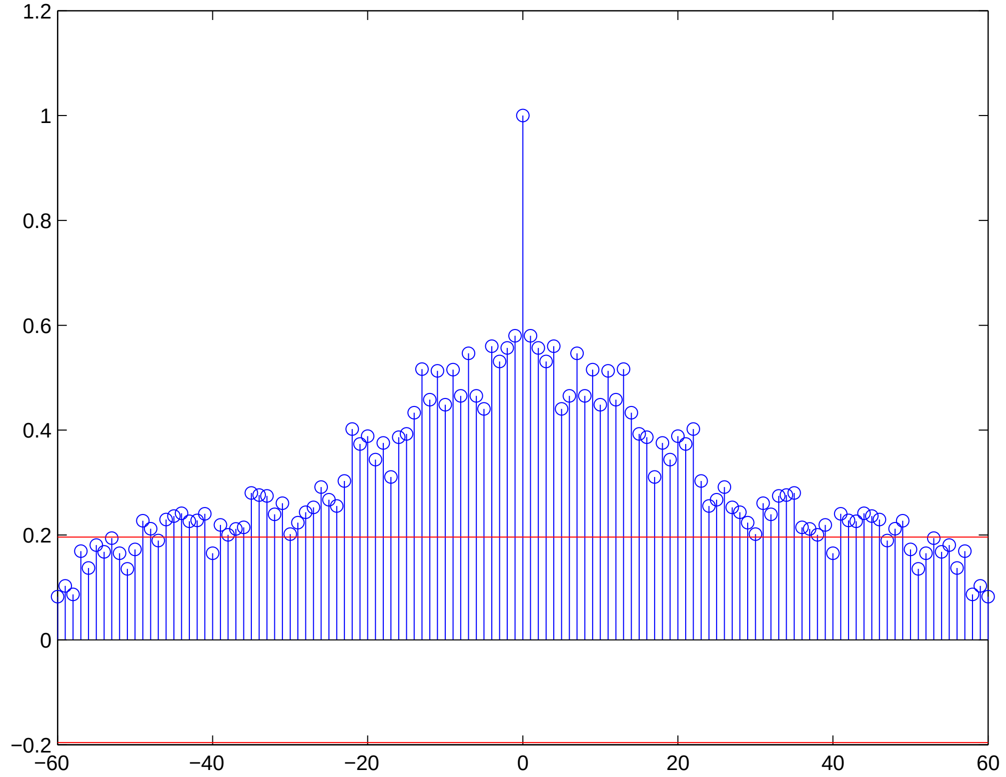
4. Estimate parameters.

5. Use diagnostics to confirm residuals are white/iid/normal.

6. Model selection.

Identifying preliminary values of d: Sample ACF

Trends lead to slowly decaying sample ACF:



Identifying preliminary values of d, p, and q

For identifying preliminary values of d, a time plot can also help.

Too little differencing: not stationary.
Too much differencing: extra dependence introduced.

For identifying p, q, look at sample ACF, PACF of $(1-B)^d X_t$:

Model:	**ACF:**	**PACF:**
AR(p)	decays	zero for $h > p$
MA(q)	zero for $h > q$	decays
ARMA(p,q)	decays	decays

Diagnostics

How do we check that a model fits well?

The residuals (innovations, $x_t - x_t^{t-1}$) should be white.
Consider the *standardized innovations*,

$$e_t = \frac{x_t - \hat{x}_t^{t-1}}{\sqrt{\hat{P}_t^{t-1}}}.$$

This should behave like a mean-zero, unit variance, iid sequence.

- Check a time plot
- Turning point test
- Difference sign test
- Rank test
- Q-Q plot, histogram, to assess normality

Model Selection

We have used the data x to estimate parameters of several models. They all fit well (the innovations are white). We need to choose a single model to retain for forecasting. How do we do it?

If we had access to independent data y from the same process, we could compare the likelihood on the new data, $L_y(\hat{\phi}, \hat{\theta}, \hat{\sigma}_w^2)$.

We could obtain y by leaving out some of the data from our model-building, and reserving it for model selection. This is called *cross-validation*. It suffers from the drawback that we are not using all of the data for parameter estimation.

Model Selection: AIC

We can approximate the likelihood defined using independent data: asymptotically

$$-\ln L_y(\hat{\phi}, \hat{\theta}, \hat{\sigma}_w^2) \approx -\ln L_x(\hat{\phi}, \hat{\theta}, \hat{\sigma}_w^2) + \frac{(p+q+1)n}{n-p-q-2}.$$

AIC_c: corrected Akaike information criterion.

Notice that:

- More parameters incur a bigger penalty.
- Minimizing the criterion over all values of $p, q, \hat{\phi}, \hat{\theta}, \hat{\sigma}_w^2$ corresponds to choosing the optimal $\hat{\phi}, \hat{\theta}, \hat{\sigma}_w^2$ for each p, q, and then comparing the penalized likelihoods.

There are also other criteria: BIC.

Pure seasonal ARMA Models

For $P, Q \geq 0$ and $s > 0$, we say that a time series $\{X_t\}$ is an **ARMA(P,Q)$_s$ process** if $\Phi(B^s)X_t = \Theta(B^s)W_t$, where

$$\Phi(B^s) = 1 - \sum_{j=1}^{P} \Phi_j B^{js},$$

$$\Theta(B^s) = 1 + \sum_{j=1}^{Q} \Theta_j B^{js}.$$

It is **causal** iff the roots of $\Phi(z^s)$ are outside the unit circle.
It is **invertible** iff the roots of $\Theta(z^s)$ are outside the unit circle.

Pure seasonal ARMA Models

Example: $P = 0, Q = 1, s = 12$. $X_t = W_t + \Theta_1 W_{t-12}$.

$$\gamma(0) = (1 + \Theta_1^2)\sigma_w^2,$$

$$\gamma(12) = \Theta_1 \sigma_w^2,$$

$$\gamma(h) = 0 \qquad \text{for } h = 1, 2, \ldots, 11, 13, 14, \ldots.$$

Example: $P = 1, Q = 0, s = 12$. $X_t = \Phi_1 X_{t-12} + W_t$.

$$\gamma(0) = \frac{\sigma_w^2}{1 - \Phi_1^2},$$

$$\gamma(12i) = \frac{\sigma_w^2 \Phi_1^i}{1 - \Phi_1^2},$$

$$\gamma(h) = 0 \qquad \text{for other } h.$$