

# Time Series Analysis

## Basics

### Sample Autocorrelations

$$\gamma_0 = \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})$$
$$\gamma_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})$$
$$\rho_0 = 1$$
$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

## Autoregressive Models

### AR(1) Models

$$X_t - \phi X_{t-1} = a_t$$
$$(1 - \phi z) = 0$$

### AR(1) Properties

- Positive  $\phi$ 
  - Realizations appear to be wandering (aperiodic)
  - Autocorrelations are damped exponentials.
  - Spectral densities have peaks at zero
- Negative  $\phi$ 
  - Realizations appear to be oscillating
  - Autocorrelations are damped oscillating exponentials
  - Spectral densities have peaks at  $f = 0.5$

### AR(2) Models

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = a_t$$
$$(1 - \phi_1 z - \phi_2 z^2) = 0$$

### AR(2) Properties

- Two Real Roots - Both Pos
  - The realization will appear to be wandering
  - The autocorrelations will be exponentially damped
  - There will be a peak at 0
- Two Real Roots - Both Neg
  - The realization will appear to be oscillating
  - The autocorrelations will be damped oscillating exponentials
  - There will be a peak at 0.5
- Two Real Roots - One Each

- The realization will appear to be wandering and an oscillation will run on the realization
- The autocorrelations will be exponentially damped with a hint of oscillation
- There will be peaks at 0 and 0.5 in the spectral density
- One Complex
  - The realization will appear to have a pseudo-cyclic behavior with a cycle length of  $\frac{1}{f_0}$
  - The autocorrelations will be damped exponentials oscillating in a sinusoid envelope with a frequency of  $f_0$
  - There will be a peak at  $f_0$  (between 0 and 0.5)

$$f_0 = \frac{1}{2\pi} \cos^{-1} \left( \frac{\phi_1}{2\sqrt{-\phi_2}} \right)$$

### AR(p) Models

$$X_t - \beta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} = a_t$$
$$x_t - \phi_1 B X_t - \phi_2 B^2 X_t - \dots - \phi_p B^p X_t = a_t$$

### Key Concepts

- An AR(p) model is stationary if and only if all the roots of the characteristic equation are outside the unit circle.
- Any AR(p) characteristic equation can be numerically factored into 1st and 2nd order elements.
- These factors are interpreted as contributing AR(1) and AR(2) behaviors to the total behavior of the AR(p) model.

### Factor Contributions

AR(p) models reflect a contribution of AR(1) and AR(2) contributions. Roots that are close to the unit circle will be the dominate behavior.

- First order factors  $(1 - \phi_1 B)$ 
  - Associated with real roots
  - Contribute AR(1)-type behavior to the AR(p) model
  - Associated with a system frequency of 0 if  $\phi_1$  is positive or 0.5 if  $\phi_1$  is negative
- Second order factors  $(1 - \phi_1 B - \phi_2 B^2)$ 
  - Associated with complex roots
  - Contribute cyclic AR(2) behavior to the AR(p) model
  - Associated with a system frequency of  $f_0$

## Moving Average Models

### MA(1) Models

$$X_t = a_t - \theta a_{t-1}$$
$$(1 - \theta_1 z) = 0$$
$$\rho_0 = 0$$
$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$$
$$\rho_k = 0 |_{k>1}$$

### MA(2) Models

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$
$$(1 - \theta_1 z - \theta_2 z^2) = 0$$
$$\rho_0 = 0$$
$$\rho_1 = \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2}$$
$$\rho_k = 0 |_{k>2}$$

### MA(q) Models

$$X_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$
$$x_t = a_t - \theta_1 B a_t - \dots - \theta_q B^q X_t$$

### Key Concepts

- MA models are a finite GLP
- MA models are always stationary
- MA models are invertable iff all the roots are outside of the unit circle.

## ARMA(p,q) Models

$$X_t = \beta + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$
$$x_t - \phi_1 B X_t - \dots - \phi_p B^p X_t = a_t - \theta_1 B a_t - \dots - \theta_q B^q X_t$$

### Key Concepts

- Valid when the model is stationary and invertable
  - Stationary: roots of  $\phi(z)$  are outside the unit circle
  - Invertable: roots of  $\theta(z)$  are outside the unit circle
- $\phi(z)$  and  $\theta(z)$  have no common factors (check)

# ARIMA

## General Form

$$\phi(B)(1-B)^dX_t = \theta(B)a_t$$

## Properties

- The roots on the unit circle dominate the behavior of the realization
- The autocorrelations are defined to have a magnitude of 1 ( $\rho_k = 1$ )
- The variance of ARIMA is not well defined

# ARUMA

ARUMA is an generalization of ARIMA that includes a term or term(s) for seasonality.

$$\phi(B)(1-B)^d(1-B^s)X_t = \theta(B)a_t$$

## Monthly Seasonality

$$(1-B^4) = (1-B)(1+B)(1+B^2)$$

# General Linear Processes

## General Form

Use `psi.weights.wge` to calculate  $\psi$ s

$$X_t - \mu = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

## MA GLP

This is a finite GLP

$$X_t - \mu = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

$$\psi_0 = 1$$

## AR GLP

This is an infinite GLP

$$X_t - \mu = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

$$\psi_0 = 1$$

# Forecasting

## Notation

- $t_0$  - origin of the forecast
- $l$  - number of time units to forecast (lead time)
- $\hat{X}_{t_0}(l)$  - the forecast of  $X_{t_0+l}$  given data up to  $t_0$

## ARMA Forecasting

Use `fore.arma.wge()` for forecasting.

$$\hat{X}_{t_0}(l) = \sum_{i=1}^p \phi_i \hat{X}_{t_0}(l-i) - \sum_{j=1}^q \theta_j \hat{a}_{t_0+l-j} + \bar{x} \left[ 1 - \sum_{i=1}^p \phi_i \right]$$

## Facts

$$e_{t_0}(l) = X_{t_0+l} - \hat{X}_{t_0}(l)$$

$$var[e_{t_0}(l)] = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$$

$$CI : \hat{X}_{t_0}(l) \pm z_{1-\alpha/2} \sigma_a \left[ \sum_{k=0}^{l-1} \psi_k^2 \right]^{1/2}$$

## ARIMA Forecasting

Use `fore.aruma.wge()` for forecasting.

- Limits become unbounded as  $l$  increases
- A factor of  $(1-B)$  does not forecast a trend. An order of  $d > 1$  is required to forecast a trend.

## ARIMA with Seasonality Forecasting

The forecast for step  $l$  is same as the last  $s$  value. Use `fore.aruma.wge()` for forecasting.

- Limits become unbounded as  $l$  increases
- A factor of  $(1-B)$  does not forecast a trend. An order of  $d > 1$  is required to forecast a trend.
- $(1-B)(1-B^s) = a_t$  is called an airline model.

# Linear Forecasting

Use `fore.sigplusnoise.wge()` for forecasting.

- Fit an OLS to  $X_t$
- Fit an AR(p) to the residuals ( $Z_t$ )

$$\hat{X}_{t_0}(l) = b_0 + b_1 t + \hat{Z}_{t_0}(l)$$

$$CI : b_0 + b_1 t + \hat{Z}_{t_0}(l) \pm z_{1-\alpha/2} \hat{\sigma}_a \left[ \sum_{k=0}^{l-1} \psi_k^2 \right]^{1/2}$$

# Filtering

Filters transform time series.

$$Z_t \rightarrow H(B) \rightarrow X_t$$

$$X(t) = Z(t)H(B)$$

There are four basic types of filters.

- High pass - filters out low frequencies
- Low pass - filters out high frequencies
- Band pass - filters out frequencies outside the band
- Band stop - filters out frequencies inside the band

## Difference Filter

The first order difference is expressed by the following

$$X_t = Z_t - Z_{t-1}$$

$$H(B) = B^0 - B$$

This is a high pass filter.

## Moving Average Filter

A 5-point moving average filter can be expressed as

$$X_t = \frac{Z_{t+2} + Z_{t+1} + Z_t + Z_{t-1} + Z_{t-2}}{5}$$

$$H(B) = \frac{B^{-2} + B^{-1} + B^0 + B + B^2}{5}$$

This is a low pass filter.

## Band-Type Filter

High pass and low pass filters can be combined to produce band pass and band stop filters.