Time Series

Don van den Bergh

March 7, 2018

Decomposing Time Series

Time series consist of:

- Trend M
- Seasonal S
- Random R

Decomposing Time Series

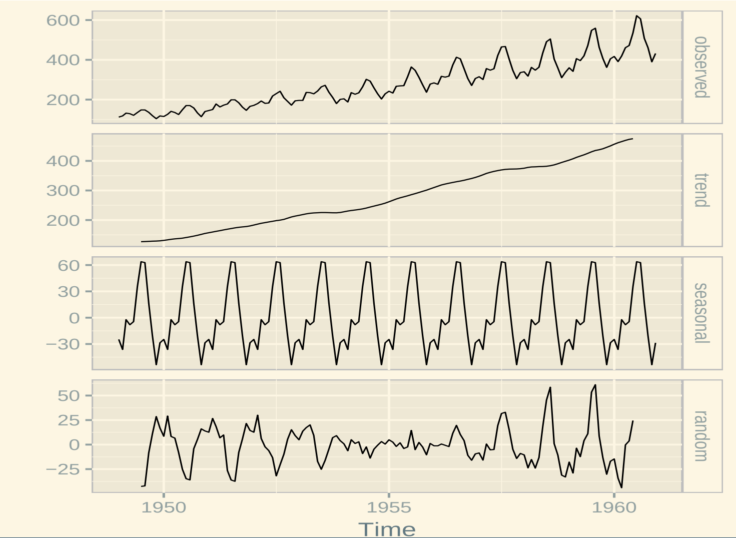
Time series consist of:

- Trend M
- Seasonal S
- Random R

Models are additive or multiplicative:

$$X_t = M_t + S_t + R_t$$
$$X_t = M_t S_t R_t$$

Decomposing Time Series

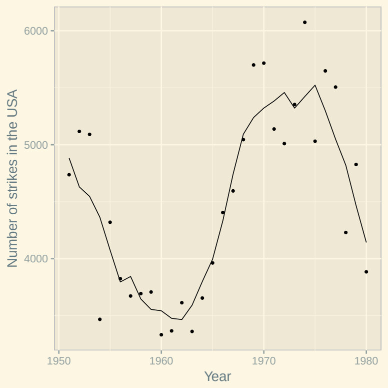


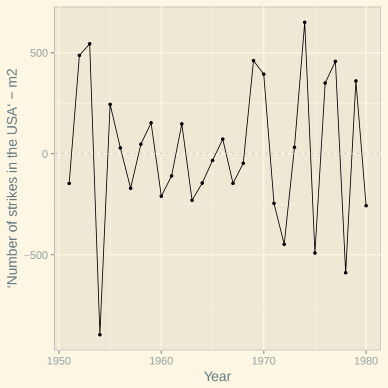
Method 1: Trend Estimation

$$\hat{m}_t^q = (2q-1)^{-1} \sum_{j=-q}^{q} X_{t-j}$$

$$\hat{Y}_t = X_t - \hat{m}_t^q$$

Method 1: Trend Estimation $q = 2$





Method 2: Trend Elimination by Differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$
$$BX_t = X_{t-1}$$
$$B^j X_t = X_{t-j}$$

Method 2: Trend Elimination by Differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$
$$BX_t = X_{t-1}$$
$$B^j X_t = X_{t-j}$$
$$\nabla^2 X_t = \nabla(\nabla X_t) = X_t - 2X_{t-1} - X_{t-2}$$

Method 2: Trend Elimination by Differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$
$$BX_t = X_{t-1}$$
$$B^j X_t = X_{t-j}$$
$$\nabla^2 X_t = \nabla(\nabla X_t) = X_t - 2X_{t-1} - X_{t-2}$$

Imagine:

$$m_t = c_0 + c_1 t$$

Method 2: Trend Elimination by Differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$
$$BX_t = X_{t-1}$$
$$B^j X_t = X_{t-j}$$
$$\nabla^2 X_t = \nabla(\nabla X_t) = X_t - 2X_{t-1} - X_{t-2}$$

Imagine:

$$m_t = c_0 + c_1 t$$
$$\nabla m_t = m_t - m_{t-1} = c_0 + c_1 t - (c_0 + c_1(t-1))$$
$$\nabla m_t = c_1$$

Method 2: Trend Elimination by Differencing

This generalizes:

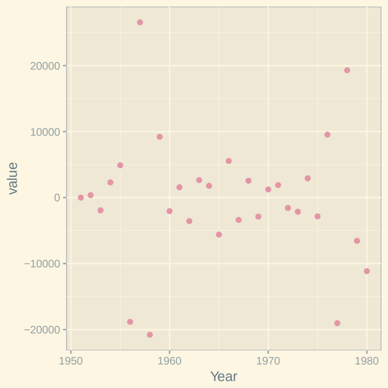
$$X_t = m_t + Y_t \qquad Y_t \text{ is stationary with } \mu_Y(t) = 0$$

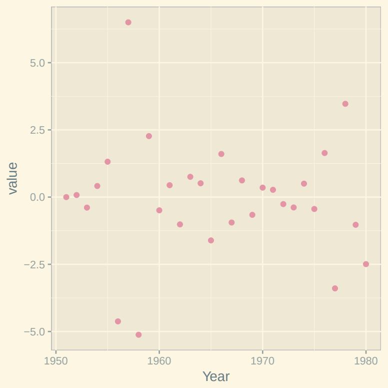
$$m_t = \sum_{j=0}^{k} c_j t^j$$

$$\nabla^k X_t = k! c_k + \nabla^k Y_t$$

Now, $\nabla^k X_t$ is stationary with mean $k! c_k$.

Method 2: Trend Elimination by Differencing $q = 2$





Summarizing the preprocessing step

- Visualize the time series
  - Detect Trend and Seasonal effects
- Transform data such that residuals are stationary
  - Estimate and substract M_t and S_t
  - Differencing
  - Transformations (log, $\sqrt{}$)
- Fit time series model to the stationary residuals

Stationarity

Some notation:

$$\mu_X(t) = E[X_t]$$
$$\gamma_X(s,t) = \text{Cov}(X_s, X_t) = E[(X_s - \mu_s)(X_t - \mu_t)]$$

Stationarity

Strict:

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})$$

Weak:

- $\mu_X(t)$ is independent of t
- $\gamma_X(t+h, t)$ is independent of t for each h.

Stationarity

Strict:

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})$$

Weak:

- $\mu_X(t)$ is independent of t
- $\gamma_X(t+h, t)$ is independent of t for each h.

Strict usually implies weak but not the other way around.

Stationarity: Random Walk

$S_t = \sum_{i=1}^{t} X_i$ where X_i i.i.d., $\mu_X = 0$ and $\gamma_X = \sigma^2$
$E[S_t] = 0$ and $E[S_t^2] = t\sigma^2$. Hence:

Stationarity: Random Walk

$S_t = \sum_{i=1}^{t} X_i$ where X_i i.i.d., $\mu_X = 0$ and $\gamma_X = \sigma^2$
$E[S_t] = 0$ and $E[S_t^2] = t\sigma^2$. Hence:

$$\begin{aligned} \gamma_S(t+h, t) &= \text{Cov}\left(S_{t+h}, S_t\right) \\ &= t\sigma^2 - 0^2 \end{aligned}$$

Stationarity: Random Walk

$S_t = \sum_{i=1}^{t} X_i$ where X_i i.i.d., $\mu_X = 0$ and $\gamma_X = \sigma^2$
$E[S_t] = 0$ and $E[S_t^2] = t\sigma^2$. Hence:

$$\begin{aligned} \gamma_S(t+h, t) &= \text{Cov}\left(S_{t+h}, S_t\right) \\ &= t\sigma^2 - 0^2 \end{aligned}$$

We conclude: Random walks violate stationarity.

Linear process

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

where $W_t \sim \mathcal{N}\left(0, \sigma^2\right)$, and $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Linear process

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

where $W_t \sim \mathcal{N}\left(0,\, \sigma^2\right)$, and $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Wold's decomposition

Every second-order stationary process is, or can be transformed to, a linear process.

MA(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$,

$$\psi_j = \begin{cases} 1, & \text{if } j = 0 \\ \theta, & \text{if } j = 1 \\ 0, & \text{otherwise} \end{cases}$$

we obtain:

$$X_t = W_t + \theta W_{t-1}$$

Properties of the MA(1)

MA(1):

$$\text{Var}(X_t) = \gamma_X(t,t) = \sigma^2(1+\theta^2)$$

$$\text{Cov}(X_t) = \begin{cases} \gamma_X(t,t+h) = \theta\sigma^2, \text{ if } h = 1 \\ \gamma_X(t,t+h) = 0, \text{ if } h > 1 \end{cases}$$

$$\text{Cor}_{X_t}(t,t+h) = \frac{\gamma_X(t,t+h)}{\gamma_X(t,t)} = \begin{cases} \frac{\theta}{1+\theta^2}, \text{ if } h = 1 \\ 0, \text{ if } h > 1 \end{cases}$$

AR(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$,

$$\psi_j = \begin{cases} \rho^j, & \text{if } j \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Then if $|\rho| < 1$ we obtain:

$$X_t = \rho X_{t-1} + W_t$$

Properties of the AR(1)

AR(1):

$$\text{Var}(X_t) = \gamma_X(t,t) = \frac{\sigma^2}{(1-\rho^2)}$$

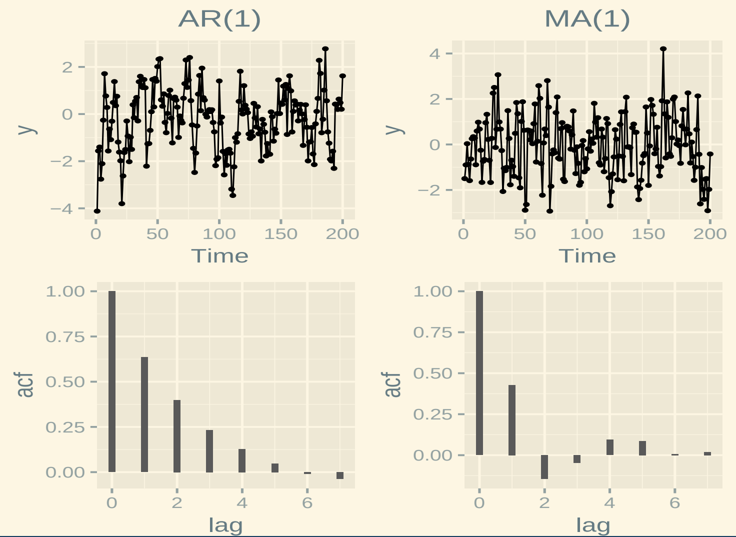
$$\text{Cov}(X_t) = \gamma_X(t,t+h) = \frac{\rho^{|h|}\sigma^2}{(1-\rho^2)}$$

$$\text{Cor}_{X_t}(t,t+h) = \frac{\gamma_X(t,t+h)}{\gamma_X(t,t)} = \rho^{|h|}$$

Comparison of MA(1) and AR(1)

- The difference is in the correlation and covariance function
- Useful for estimation
- Useful for model comparison

Comparison of MA(1) and AR(1): ACF



Extensions: AR(p)

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t$$

where $W_t \sim \mathcal{N}(0, \sigma^2)$

Extensions: AR(p)

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t$$

where $W_t \sim \mathcal{N}(0, \sigma^2)$

Equivalently,

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$
$$\phi(B) X_t = W_t$$

Remember: $BX_t = X_{t-1}$ and $B^j X_t = X_{t-j}$

Extensions: AR(p)

Conditions on $\phi(B)$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

- Unique solution iff no roots in the unit circle
- Does not depend on future iff roots outside of unit circle

Extensions: ARMA(p, q)

ARMA(1, 1)

$$X_t - \phi X_{t-1} = W_t + \theta W_{t-1}$$

where $\phi + \theta \neq 0$

Extensions: ARMA(p, q)

ARMA(1, 1)

$$X_t - \phi X_{t-1} = W_t + \theta W_{t-1}$$

where $\phi + \theta \neq 0$

ARMA(p, q)

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$
$$\phi(B) X_t = \theta(B) W_t$$

For identifiability, $\phi(B)$ and $\theta(B)$ have no common factors.

Extensions: predictors

- VAR model
- VMA model
- VARMA model

R: `stats::arima` and `forecasting`