Forecasting for ARIMA(p, d, q) models

(given a specific choice of d)

$$\min_{c_1,\dots,c_m} \mathbb{E}\left[\left(y_{m+1} - \left(\sum_{j=1}^{m} c_{mj} y_{m+1-j}\right)\right)^2\right] =$$

$$\mathbb{E}[y_{m+1}^2] + \underset{\sim}{c_m}' A_m \underset{\sim}{c_m} - 2\underset{\sim}{b_m}' \underset{\sim}{c_m},$$

where $A_{mij} = \mathbb{E}[y_{m+1-i}\, y_{m+1-j}]$

$b_{mj} = \mathbb{E}[y_{m+1}\, y_{m+1-j}]$

previously these were autocovar. of a $\underline{\underline{\text{stationary}}}$ process

When $d>0$, y_t isn't stationary anymore, A_n, $\underset{\sim}{b_n}$ difficult to evaluate and work with

How we'll fix this... write

$$y_t = y_0 + \sum_{i=1}^{t} \nabla^d y_t$$] works for $d=1$

$$\min_{c_1,\dots,c_m} \mathbb{E}\left[\left(y_0 + \sum_{i=1}^{m+1} \nabla^d y_i - \sum_{j=1}^{m} c_{mj}\left(y_0 + \sum_{k=1}^{m+1-j} \nabla^d y_k\right)\right)^2\right]$$

* This depends on

$\mathbb{E}[y_0 \nabla^d y_i]$ for $i > 0$] it's not a big deal to assume these are zero

$\boxed{\mathbb{E}[\nabla^d y_i \nabla^d y_j]}$ autocovariances of a stationary process! :)

Tedious to work out by hand, very similar to ARMA setting, so we'll just use R for this

In general, it's useful to rewrite

$$y_t = \nabla^d y_t - \sum_{j=1}^{d} \binom{d}{j} (-1)^j y_{t-j}$$

we get this by rearranging $\nabla^d y_t = (1-B)^d y_t$ so that y_t is on the left hand side, everything else on right

Forecast one step ahead from $t=m$, $m \leq n$,

$$y_{m+1} = \underbrace{\nabla^d y_{m+1}}_{\text{stationary process}} - \boxed{\sum_{j=1}^{d} \binom{d}{j} \cdot (-1)^j y_{m+1-j}}$$

all the y_{m+1-j}'s for $j=1,\dots,d$ are known at time m

Forecast h steps ahead from $t=m$, $m\le n$, $h>1$

$$(*) \quad y_{m+h} = \underline{\nabla^d y_{m+h}} \sum_{j=1}^{d} \binom{d}{j} \cdot (-1)^j \, y_{m+h-j}$$

stationary process

not all of these were observed, can use the forecasts defined by $(*)$

variance of our forecasts will now have two components, variance from not knowing future increments and the variance from not knowing future values

forecasts will require one extra assumption
$y_1, \dots, y_d$ are independent of stationary increm. $\nabla^d y_t$

State Space Models!

$$y_t = a x_t + v_t$$ ← observation equation

$$x_t = \phi x_{t-1} + w_t$$ ← state equation

$v_t \overset{iid}{\sim} N(0, \sigma_v^2)$, $w_t \overset{iid}{\sim} N(0, \sigma_w^2)$, $x_1 = \mu$

$\sigma_v^2 = 0$, $a = 1$, then this is basically an AR(1) process for y_t, but with a fixed starting value $y_1 = \mu$.

y_t observed, x_t all unobserved

State Space Models!

$$y_t = a x_t + v_t$$ ← observation equation

$$x_t = \phi x_{t-1} + w_t$$ ← state equation

$$v_t \overset{iid}{\sim} N(0, \sigma_v^2), \quad w_t \overset{iid}{\sim} N(0, \sigma_w^2), \quad x_1 = \mu$$

First, assume that $a, \phi, \sigma_v^2, \sigma_w^2$, and μ are <u>known</u>

might be interested in:

example: $\mathbb{E}[x_2 \mid y_1]$, $\mathbb{V}[x_2 \mid y_1]$

* Predicting future x_t given $y_1, \ldots, y_{t-1}$.
* "Filtering" estimating future x_t given $y_1, \ldots, y_t$,
  example: $\mathbb{E}[x_2 \mid y_1, y_2]$, $\mathbb{V}[x_2 \mid y_1, y_2]$
* Smoothing estimating x_t given everything we saw, $y_1, \ldots, y_n$, example $\mathbb{E}[x_1 \mid y_1, y_2]$, $\mathbb{V}[x_1 \mid y_1, y_2]$

State Space Models!

$$y_t = a x_t + v_t \quad \leftarrow \text{observation equation}$$

$$x_t = \phi x_{t-1} + w_t \quad \leftarrow \text{state equation}$$

$$v_t \overset{iid}{\sim} N(0, \sigma_v^2), \quad w_t \overset{iid}{\sim} N(0, \sigma_w^2), \quad x_1 = \mu$$

Allows us to define a joint probability distribution for $\underset{\sim}{x}$ and $\underset{\sim}{y}$

$$\begin{pmatrix} \underset{\sim}{y} \\ \underset{\sim}{x} \end{pmatrix} \sim N\left(\begin{pmatrix} E[\underset{\sim}{y}] \\ E[\underset{\sim}{x}] \end{pmatrix}, \begin{pmatrix} V[\underset{\sim}{y}] & Cov[\underset{\sim}{y}, \underset{\sim}{x}] \\ Cov[\underset{\sim}{y}, \underset{\sim}{x}] & V[\underset{\sim}{x}] \end{pmatrix} \right)$$

Very Important Multivariate Normal Facts

Suppose $\begin{pmatrix} \underset{\sim}{u} \\ \underset{\sim}{v} \end{pmatrix} \sim N\left(\begin{pmatrix} \underset{\sim}{a} \\ \underset{\sim}{b} \end{pmatrix}, \begin{pmatrix} C & D \\ D' & E \end{pmatrix}\right)$

then $\underset{\sim}{u} \sim N(\underset{\sim}{a}, C)$, $\underset{\sim}{v} \sim N(\underset{\sim}{b}, E)$

also $\underset{\sim}{u} \mid \underset{\sim}{v} \sim N\left(\underset{\sim}{a} + D'E^{-1}(\underset{\sim}{v} - \underset{\sim}{b}), C - D'E^{-1}D\right)$

$\underset{\sim}{v} \mid \underset{\sim}{u} \sim N\left(\underset{\sim}{b} + DC^{-1}(\underset{\sim}{u} - \underset{\sim}{a}), E - DC^{-1}D'\right)$

State Space Models!

$$y_t = a x_t + v_t \quad \leftarrow \text{observation equation}$$

$$x_t = \phi x_{t-1} + w_t \quad \leftarrow \text{state equation}$$

$$v_t \overset{iid}{\sim} N(0, \sigma_v^2), \quad w_t \overset{iid}{\sim} N(0, \sigma_w^2), \quad x_1 = \mu$$

Allows us to define a joint probability distribution for $\underset{\sim}{x}$ and $\underset{\sim}{y}$

$$\begin{pmatrix} \underset{\sim}{y} \\ \underset{\sim}{x} \end{pmatrix} \sim N\left(\begin{pmatrix} a\mathbb{E}[\underset{\sim}{x}] \\ \mathbb{E}[\underset{\sim}{x}] \end{pmatrix}, \begin{pmatrix} a^2\mathbb{V}[\underset{\sim}{x}] + \sigma_v^2 I & a\mathbb{V}[\underset{\sim}{x}] \\ a\mathbb{V}[\underset{\sim}{x}] & \mathbb{V}[\underset{\sim}{x}] \end{pmatrix} \right)$$

Smoothing asks about the conditional distribution of $\underset{\sim}{x} \mid \underset{\sim}{y}$

State Space Models!

$$y_t = a x_t + v_t$$ ← observation equation

$$x_t = \phi x_{t-1} + w_t$$ ← state equation

$$v_t \overset{iid}{\sim} N(0, \sigma_v^2), \quad w_t \overset{iid}{\sim} N(0, \sigma_w^2), \quad x_1 = \mu$$

Smoothed $\underset{\sim}{x}$ distribution given $\underset{\sim}{y}$ has

$$\mathbb{E}[\underset{\sim}{x} \mid \underset{\sim}{y}] = \mathbb{E}[\underset{\sim}{x}] + a\mathbb{V}[\underset{\sim}{x}]\left(a^2\mathbb{V}[\underset{\sim}{x}] + \sigma_v^2 I\right)^{-1}\left(\underset{\sim}{y} - a\mathbb{E}[\underset{\sim}{x}]\right)$$

$$\mathbb{V}[\underset{\sim}{x} \mid \underset{\sim}{y}] = \mathbb{V}[\underset{\sim}{x}] - a^2\mathbb{V}[\underset{\sim}{x}]\left(a^2\mathbb{V}[\underset{\sim}{x}] + \sigma_v^2 I\right)^{-1}\mathbb{V}[\underset{\sim}{x}]$$