The Integrated ARMA model: ARIMA\((p, d, q)\)

- Some series are nonstationary, but their differences are stationary; e.g. the random walk.

- Recall: the first differences of \(x_t\) are
  \[
  x_t - x_{t-1} = (1 - B)x_t = \nabla x_t.
  \]

- The second differences are
  \[
  \nabla x_t - \nabla x_{t-1} = (1 - B)\nabla x_t = \nabla^2 x_t.
  \]

- If \(\nabla^d x_t\) is ARMA\((p, q)\), we say that \(x_t\) is ARIMA\((p, d, q)\).
Under-differencing

- Suppose that $x_t$ is ARIMA$(p, d, q)$, but we analyze $y_t = \nabla^{d'} x_t$ for some $d' < d$.

- In this case, $y_t$ satisfies

$$\nabla^{d-d'} \phi(B) y_t = \phi^*(B) y_t = \theta(B) w_t$$

where $\phi^*(z) = (1 - z)^{(d-d')} \phi(z)$ has $d - d'$ roots at $z = 1$.

- This looks like an ARMA$(p + d - d', q)$ model, but it is not causal.
Over-differencing

• Suppose that $x_t$ is ARIMA($p,d,q$), but we analyze $y_t = \nabla^{d'} x_t$ for some $d' > d$.

• In this case, $y_t$ satisfies

$$
\phi(B)y_t = \nabla^{d'-d} \theta(B) w_t = \theta^*(B)w_t
$$

where $\theta^*(z) = (1 - z)^{(d'-d)} \theta(z)$ has $d' - d$ roots at $z = 1$.

• This looks like an ARMA($p,q + d' - d$) model, but it is not invertible.
Simplest model with \( d > 0 \): \textbf{ARIMA}(0, 1, 1)

- Many nonstationary series are found to be fitted quite well as \textbf{ARIMA}(0, 1, 1).

- This model is connected with the \textit{exponentially weighted moving average} (EWMA) method of forecasting.

- If the model is written \( x_t - x_{t-1} = w_t - \lambda w_{t-1} \), the one-step forecast is

\[
\tilde{x}_{n+1} = (1 - \lambda) \sum_{j=0}^{\infty} \lambda^j x_{n-j},
\]

the exponentially weighted moving average.
• We can calculate the forecast recursively:

\[ x_{n+1} = x_n - \lambda w_n + w_{n+1}. \]

• We can find \( w_n \) from \( x_n, x_{n-1}, \ldots \), so the one-step forecast is the first part:

\[ \tilde{x}_{n+1} = x_n - \lambda w_n \]
• But $w_n$ is the previous forecast error, $x_n - \tilde{x}_n$, so

\[
\tilde{x}_{n+1} = x_n - \lambda(x_n - \tilde{x}_n) = (1 - \lambda)x_n + \lambda\tilde{x}_n.
\]

• In words,

the new forecast is a weighted average of the current forecast and the current value.

• Also

\[
\tilde{x}_{n+1} = \tilde{x}_n + (1 - \lambda)(x_n - \tilde{x}_n),
\]

so the new forecast is the current forecast plus a correction based on the current forecast error.
Strategy for Building ARIMA Models

1. First choose $d$:
   - ACF of an integrated series tends to die away slowly, so difference until it dies away quickly;
   - the IACF of a non-invertible series tends to die away slowly, which indicates over-differencing.
   - You may want to try more than one value of $d$.

2. Next choose $p$ and $q$, e.g. using MINIC.
3. Next estimate the model.

4. Finally check the model diagnostics:
   - Significance of highest order coefficients, \( \hat{\phi}_p \) (if \( p > 0 \)) and \( \hat{\theta}_q \) (if \( q > 0 \));
   - Non-significance in autocorrelation check of residuals;
   - Low value of AIC or SBC.

5. Repeat from step 2 until satisfactory.
   - Note: You may not find a completely satisfactory model, especially for a long data series.
Unit Root Tests

• Choice of $d$ can be formulated as a hypothesis test.

• E.g. in the AR(1) model $x_t = \phi x_{t-1} + w_t$, set:

  - $H_0 : \phi = 1$, $x_t$ is ARIMA$(0,1,0)$ (nonstationary, $d = 1$);
  - $H_A : |\phi| < 1$, $x_t$ is ARIMA$(1,0,0)$ (stationary, $d = 0$).

Test using proc arima’s stationarity keyword on the identify statement.

• E.g. the global temperature data: proc arima program and output.
• The statistics on the “Lags 0” rows in the panel “Augmented Dickey-Fuller Unit Root Tests” refer to the three models

- Zero Mean:

\[ x_t = \phi x_{t-1} + w_t; \]

- Single Mean:

\[ x_t - \mu = \phi (x_{t-1} - \mu) + w_t; \]

- Trend:

\[ x_t - \mu - \beta t = \phi [x_{t-1} - \mu - \beta (t-1)] + w_t. \]
• Note that under $H_0$, these models reduce to

$$x_t = x_{t-1} + w_t,$$
$$x_t = x_{t-1} + w_t,$$
$$x_t = x_{t-1} + \beta + w_t,$$

the first two being random walks with no drift, the latter being a random walk with drift.

• The statistics on the “Lags 1” rows refer to corresponding AR(2) models, which reduce to integrated AR(1) models under the null hypothesis.

• The “Tau” tests are generally preferred to the “Rho” tests.
• E.g. Case-Shiller housing data: proc arima program and output.