Econ 582
Univariate Stationary Time Series

Eric Zivot

April 1, 2013
Time Series Concepts

A stochastic process \( \{Y_t\}_{t=1}^{\infty} \) is a sequence of random variables indexed by time \( t \):

\[
\{\ldots, Y_1, Y_2, \ldots, Y_t, Y_{t+1}, \ldots\}
\]

A realization of a stochastic process is the sequence of observed data \( \{y_t\}_{t=1}^{\infty} \):

\[
\{\ldots, Y_1 = y_1, Y_2 = y_2, \ldots, Y_t = y_t, Y_{t+1} = y_{t+1}, \ldots\}
\]
We are interested in the conditions under which we can treat the stochastic process like a random sample, as the sample size goes to infinity. Under such conditions, at any point in time $t_0$, the *ensemble average*

$$\frac{1}{N} \sum_{k=1}^{N} Y^{(k)}_{t_0}$$

will converge to the sample *time average*

$$\frac{1}{T} \sum_{t=1}^{T} Y_t$$

as $N$ and $T$ go to infinity. If this result occurs then the stochastic process is called *ergodic*. 
Definition 1  *Strict stationarity*

A stochastic process $\{Y_t\}_{t=1}^{\infty}$ is *strictly stationary* if, for any given finite integer $r$ and for any set of subscripts $t_1, t_2, \ldots, t_r$ the joint distribution of

$$(Y_t, Y_{t_1}, Y_{t_2}, \ldots, Y_{t_r})$$

depends only on $t_1 - t, t_2 - t, \ldots, t_r - t$ but not on $t$. 

**Stationary Stochastic Processes**

Definition 1  *Strict stationarity*
Remarks

1. For example, the distribution of $(Y_1, Y_5)$ is the same as the distribution of $(Y_{12}, Y_{16})$.

2. For a strictly stationary process, $Y_t$ has the same mean, variance, moments etc. (if they exist) for all $t$.

3. Any function/transformation $g(\cdot)$ of a strictly stationary process, $\{g(Y_t)\}$ is also strictly stationary.
Example 1  iid sequence

If \( \{Y_t\} \) is an iid sequence, then it is strictly stationary.

Let \( \{Y_t\} \) be an iid sequence and let \( X \sim N(0, 1) \) independent of \( \{Y_t\} \). Let \( Z_t = Y_t + X \). Then the sequence \( \{Z_t\} \) is strictly stationary.

Since \( \{Z_t\} \) is strictly stationary, \( \{Z_t^2\} \) is also strictly stationary.
**Definition 2**  *Covariance (Weak) stationarity*

A stochastic process \( \{Y_t\}_{t=1}^{\infty} \) is *covariance stationary* (weakly stationary) if

1. \( E[Y_t] = \mu \) does not depend on \( t \)

2. \( \text{cov}(Y_t, Y_{t-j}) = \gamma_j \) exists, is finite, and depends only on \( j \) but not on \( t \) for \( j = 0, 1, 2, \ldots \)

Remark:

A strictly stationary process is covariance stationary if the mean and variance exist and the covariances are finite.
For a weakly stationary process \( \{Y_t\}_{t=1}^\infty \) define the following moments:

\[
\begin{align*}
\gamma_j &= \text{cov}(Y_t, Y_{t-j}) = j^{th} \text{ order autocovariance} \\
\gamma_0 &= \text{var}(Y_t) = \text{variance} \\
\rho_j &= \frac{\gamma_j}{\gamma_0} = j^{th} \text{ order autocorrelation}
\end{align*}
\]

Autocorrelation Function (ACF)

plot \( \rho_j \) vs. \( j \)

**Remark**

A weakly stationary process is uniquely determined by its mean, variance and autocovariances.
**Example:** Independent White Noise, \( IWN(0, \sigma^2) \)

\[
Y_t = \varepsilon_t, \quad \varepsilon_t \sim \text{iid} \ (0, \sigma^2) \\
E[Y_t] = 0, \quad \text{var}(Y_t) = \sigma^2, \quad \gamma_j = 0, \ j \neq 0
\]

**Example:** Gaussian White Noise, \( GWN(0, \sigma^2) \)

\[
Y_t = \varepsilon_t, \quad \varepsilon_t \sim \text{iid} \ N(0, \sigma^2)
\]

**Example:** White Noise, \( WN(0, \sigma^2) \)

\[
Y_t = \varepsilon_t \\
E[\varepsilon_t] = 0, \quad \text{var}(\varepsilon_t) = \sigma^2, \quad \text{cov}(\varepsilon_t, \varepsilon_{t-j}) = 0
\]
Nonstationary Processes

**Example:** Deterministically trending process

\[ Y_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \]

\[ E[Y_t] = \beta_0 + \beta_1 t \quad \text{depends on } t \]

Note: A simple detrending transformation yield a stationary process:

\[ X_t = Y_t - \beta_0 - \beta_1 t = \varepsilon_t \]
Nonstationary Processes

Example: Random Walk

\[ Y_t = Y_{t-1} + \varepsilon_t, \varepsilon_t \sim WN(0, \sigma^2), \ Y_0 \text{ is fixed} \]

\[ = Y_0 + \sum_{j=1}^{t} \varepsilon_j \Rightarrow \text{var}(Y_t) = \sigma^2 t \text{ depends on } t \]

Note: A simple detrending transformation yield a stationary process:

\[ \Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t \]
**Definition 3  **\textit{Ergodicity}

Loosely speaking, a stochastic process \( \{Y_t\}_{t=1}^{\infty} \) is \textit{ergodic} if any two collections of random variables partitioned far apart in the sequence are almost independently distributed. The formal definition of ergodicity is highly technical and involves measure theory.

Result:

Let \( \{Y_t\} \) be a covariance stationary and ergodic process with mean \( E[Y_t] = \mu \) and autocovariances \( \gamma_j = \text{cov}(Y_t, Y_{t-j}) \). Then

\[
\sum_{j=0}^{\infty} |\gamma_j| < \infty
\]

and \( \text{cov}(Y_t, Y_{t-j}) = 0 \) for \( j \) large.
Theorem 1  Ergodic Theorem

Let \( \{Y_t\} \) be stationary and ergodic with \( E[Y_t] = \mu \). Then

\[
\bar{Y} = \frac{1}{T} \sum_{t=1}^{T} Y_t \xrightarrow{p} E[Y_t] = \mu
\]

Remarks

1. The ergodic theorem says that for a stationary and ergodic sequence \( \{Y_t\} \) the time average converges to the ensemble average as the sample size gets large. That is, the ergodic theorem is a LLN for stochastic processes.

2. The ergodic theorem is a substantial generalization of Kolmogorov’s LLN because it allows for serial dependence in the time series.
3. Any transformation $g(\cdot)$ of a stationary and ergodic process $\{Y_t\}$ is also stationary and ergodic. That is, $\{g(Y_t)\}$ is stationary and ergodic. Therefore, if $E[g(Y_t)]$ exists then the ergodic theorem gives

$$\bar{g} = \frac{1}{T} \sum_{t=1}^{T} g(Y_t) \xrightarrow{p} E[g(Y_t)]$$

This is a very useful result. For example, we may use it to prove that the sample autocovariances

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^{T} (Y_t - \bar{Y})(Y_{t-j} - \bar{Y})$$

converge in probability to the population autocovariances $\gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)] = \text{cov}(Y_t, Y_{t-j})$. 
Example 2  *Stationary but not ergodic process* (White, 1984)

Let \( \{Y_t\} \) be an iid sequence with \( E[Y_t] = \mu \) and let \( X \sim N(0, 1) \) independent of \( \{Y_t\} \). Let \( Z_t = Y_t + X \). Note that \( E[Z_t] = \mu \).

Claim: \( Z_t \) is stationary but not ergodic.
Wold’s Decomposition Theorem

Any covariance stationary time series \( \{Y_t\} \) can be represented in the form

\[
Y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots
\]

\[
= \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},
\]

\[
\psi_0 = 1, \quad \sum_{j=0}^{\infty} \psi_j^2 < \infty, \quad \varepsilon_t \sim WN(0, \sigma^2)
\]

Lag operator notation

\[
Y_t = \mu + \psi(L) \varepsilon_t,
\]

\[
\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j
\]
Properties:

\[ E[Y_t] = \mu \]
\[ \gamma_0 = \text{var}(Y_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 < \infty \]
\[ \gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)] \]
\[ = E[(\varepsilon_t + \psi_1 \varepsilon_{t-1} + \cdots + \psi_j \varepsilon_{t-j} + \psi_{j+1} \varepsilon_{t-j-1} + \cdots) \times (\varepsilon_{t-j} + \psi_1 \varepsilon_{t-j-1} + \cdots)] \]
\[ = \sigma^2 (\psi_j + \psi_{j+1} \psi_1 + \cdots) \]
\[ = \sigma^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+j} \]
Autoregressive moving average models (ARMA) Models (Box-Jenkins 1976)

**Idea:** Approximate Wold form of stationary time series by parsimonious parametric models

ARMA(p,q) model:

\[ Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \cdots + \phi_p(Y_{t-p} - \mu) + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \]
\[ \epsilon_t \sim WN(0, \sigma^2) \]

Lag operator notation:

\[ \phi(L)(Y_t - \mu) = \theta(L)\epsilon_t \]
\[ \phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p \]
\[ \theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q \]
ARMA(0,1) Process (MA(1) Process)

\[ Y_t = \mu + \varepsilon_t + \theta \varepsilon_{t-1} = \mu + \theta(L)\varepsilon_t \]

\[ \theta(L) = 1 + \theta L, \; \varepsilon_t \sim WN(0, \sigma^2) \]

Moments:

\[
\begin{align*}
E[Y_t] &= \mu \\
\text{var}(Y_t) &= \gamma_0 = E[(Y_t - \mu)^2] \\
&= E[(\varepsilon_t + \theta \varepsilon_{t-1})^2] \\
&= \sigma^2(1 + \theta^2) \\
\gamma_1 &= E[(Y_t - \mu)(Y_{t-1} - \mu)] \\
&= E[(\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-1} + \theta \varepsilon_{t-2})] \\
&= \sigma^2 \theta \\
\rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{\theta}{1 + \theta^2}
\end{align*}
\]

Note: Sign of \( \rho_1 \) depends on the sign of \( \theta \).
Result: The MA(1) process only has memory for one period. That is, $\gamma_j = 0$ for $j > 1$:

$$
\gamma_2 = E[(Y_t - \mu)(Y_{t-2} - \mu)] = E[(\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-2} + \theta \varepsilon_{t-3})] = 0
$$

$$
\rho_2 = 0
$$

$$
\gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)] = 0 \text{ for } j > 1
$$

$$
\rho_j = 0 \text{ for } j > 1
$$

Result: The MA(1) process is covariance stationary and ergodic. Note,

$$
\sum_{j=0}^{\infty} |\gamma_j| = \sigma^2(1 + \theta^2 + |\theta|) < \infty
$$
**Remark:** There is an identification problem for

\[-0.5 < \rho_1 < 0.5\]

The values \(\theta\) and \(\theta^{-1}\) produce the same value of \(\rho_1\). For example, \(\theta = 0.5\) and \(\theta^{-1} = 2\) both produce \(\rho_1 = 0.4\).

**Invertibility Condition:** The MA(1) is invertible if \(|\theta| < 1\). This is important for MLE of the parameters.

**Note:** More on the invertibility condition later on.
ARMA(0,2) (MA(2) Process)

\[ Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} = \mu + \theta(L)\varepsilon_t \]

\[ \theta(L) = 1 + \theta_1 L + \theta_2 L^2, \quad \varepsilon_t \sim WN(0, \sigma^2) \]

Moments:

\[ E[Y_t] = \mu \]

\[ \text{var}(Y_t) = \gamma_0 = \sigma^2(1 + \theta_1^2 + \theta_2^2) \]

\[ \text{cov}(Y_t, Y_{t-1}) = \gamma_1 \]

\[ = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3})] \]

\[ = \sigma^2(\theta_1 + \theta_1 \theta_2) \]

\[ \text{cov}(Y_t, Y_{t-2}) = \gamma_2 \]

\[ = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2})(\varepsilon_{t-2} + \theta_1 \varepsilon_{t-3} + \theta_2 \varepsilon_{t-4})] \]

\[ = \sigma^2 \theta_2 \]

\[ \text{cov}(Y_t, Y_{t-j}) = \gamma_j = 0 \text{ for } j > 2 \]
ARMA(1,0) Model (AR(1) Process) without Mean

\[ Y_t = \phi Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \]

Solution by recursive substitution:

\[
Y_t = \phi^{t+1}Y_{-1} + \phi^t\varepsilon_0 + \cdots + \phi\varepsilon_{t-1} + \varepsilon_t \\
= \phi^{t+1}Y_{-1} + \sum_{i=0}^{t} \phi^i\varepsilon_{t-i} \\
= \phi^{t+1}Y_{-1} + \sum_{i=0}^{t} \psi_i\varepsilon_{t-i}, \quad \psi_i = \phi^i
\]

Alternatively, solving forward \( j \) periods from time \( t \):

\[
Y_{t+j} = \phi^{j+1}Y_{t-1} + \phi^j\varepsilon_t + \cdots + \phi\varepsilon_{t+j-1} + \varepsilon_{t+j} \\
= \phi^{j+1}Y_{t-1} + \sum_{i=0}^{j} \psi_i\varepsilon_{t+j-i}
\]
Dynamic Multiplier:

\[
\frac{dY_j}{d\varepsilon_0} = \frac{dY_{t+j}}{d\varepsilon_t} = \phi^j = \psi_j
\]

Impulse Response Function (IRF)

Plot \(\psi_j\) vs. \(j\)

Cumulative impact (up to horizon \(j\))

\[
\sum_{i=0}^{j} \psi_j
\]

Long-run cumulative impact

\[
\sum_{i=0}^{\infty} \psi_j = \psi(1)
\]

\[= \psi(L) \text{ evaluated at } L = 1\]
Stability and Stationarity Conditions

If $|\phi| < 1$ then

$$\lim_{j \to \infty} \phi^j = \lim_{j \to \infty} \psi_j = 0$$

and the stationary solution (Wold form) for the AR(1) becomes.

$$Y_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

This is a stable (non-explosive) solution. Note that

$$\psi(1) = \sum_{j=0}^{\infty} \phi^j = \frac{1}{1-\phi} < \infty$$
If $\phi = 1$ then the AR(1) process becomes the *random walk process*

\[
Y_t = Y_{t-1} + \varepsilon_t
\]

\[
= Y_0 + \sum_{j=0}^{t} \varepsilon_j, \quad \psi_j = 1, \quad \psi(1) = \infty
\]

which is not stationary or stable.
AR(1) in Lag Operator Notation

\[(1 - \phi L)Y_t = \varepsilon_t\]

If \(|\phi| < 1\) then

\[
(1 - \phi L)^{-1} = \sum_{j=0}^{\infty} \phi^j L^j = 1 + \phi L + \phi^2 L^2 + \cdots
\]

such that

\[
(1 - \phi L)^{-1}(1 - \phi L) = 1
\]
Trick to find Wold form:

\[ Y_t = (1 - \phi L)^{-1}(1 - \phi L)Y_t = (1 - \phi L)^{-1}\varepsilon_t \]

\[ = \sum_{j=0}^{\infty} \phi^j L^j \varepsilon_t \]

\[ = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \]

\[ = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \; \psi_j = \phi^j \]
Moments of Stationary AR(1) with Mean

Mean adjusted form:

\[ Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad |\phi| < 1 \]

Regression form:

\[ Y_t = c + \phi Y_{t-1} + \varepsilon_t, \quad c = \mu(1 - \phi) \]

Trick for calculating moments: use stationarity properties

\[
E[Y_t] = E[Y_{t-j}] \text{ for all } j \\
cov(Y_t, Y_{t-j}) = \text{cov}(Y_{t-k}, Y_{t-k-j}) \text{ for all } k, j
\]
Mean of AR(1)

\[ E[Y_t] = c + \phi E[Y_{t-1}] + E[\varepsilon_t] \]
\[ = c + \phi E[Y_t] \]
\[ \Rightarrow E[Y_t] = \frac{c}{1 - \phi} = \mu \]
Variance of AR(1)

\[ \gamma_0 = \text{var}(Y_t) = E[(Y_t - \mu)^2] = E[(\phi(Y_{t-1} - \mu) + \varepsilon_t)^2] \]
\[ = \phi^2 E[(Y_{t-1} - \mu)^2] + 2E[(Y_{t-1} - \mu)\varepsilon_t] + E[\varepsilon_t^2] \]
\[ = \phi^2 \gamma_0 + \sigma^2 \text{ (by stationarity)} \]
\[ \Rightarrow \gamma_0 = \frac{\sigma^2}{1 - \phi^2} \]

Note: From the Wold representation

\[ \gamma_0 = \text{var}\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right) = \sigma^2 \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma^2}{1 - \phi^2} \]
Autocovariances and Autocorrelations

Trick: multiply $Y_t - \mu$ by $Y_{t-j} - \mu$ and take expectations

\[
\gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)] = E[\phi(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + E[\varepsilon_t(Y_{t-j} - \mu)] = \phi \gamma_{j-1} \text{ (by stationarity)}
\]

\[
\Rightarrow \gamma_j = \phi^j \gamma_0 = \phi^j \frac{\sigma^2}{1 - \phi^2}
\]

Autocorrelations:

\[
\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\phi^j \gamma_0}{\gamma_0} = \phi^j = \psi_j
\]

Note: for the AR(1), $\rho_j = \psi_j$. However, this is not true for general ARMA processes.
The half-life is a measure of the speed of mean reversion.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>half-life</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>68.97</td>
</tr>
<tr>
<td>0.9</td>
<td>6.58</td>
</tr>
<tr>
<td>0.75</td>
<td>2.41</td>
</tr>
<tr>
<td>0.5</td>
<td>1.00</td>
</tr>
<tr>
<td>0.25</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 1: Half lives for AR(1)

Half-Life of AR(1): lag at which IRF decreases by one half

$\psi_j = \phi^j = 0.5$

$\Rightarrow j \ln \phi = \ln(0.5)$

$\Rightarrow j = \frac{\ln(0.5)}{\ln \phi}$
Application: Half-Life of Real Exchange Rates

The real exchange rate is defined as

\[ z_t = s_t - p_t + p_t^* \]
\[ s_t = \log \text{ nominal exchange rate} \]
\[ p_t = \log \text{ of domestic price level} \]
\[ p_t^* = \log \text{ of foreign price level} \]

Purchasing power parity (PPP) suggests that \( z_t \) should be stationary.
ARMA$(p, 0)$ Model (AR(p) Process)

Mean-adjusted form:

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \cdots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t$$

$$\varepsilon_t \sim W N(0, \sigma^2)$$

$$E[Y_t] = \mu$$

Lag operator notation:

$$\phi(L)(Y_t - \mu) = \varepsilon_t$$

$$\phi(L) = 1 - \phi_1L - \cdots \phi_pL^p$$
Unobserved Components representation

\[ Y_t = \mu + X_t \]
\[ \phi(L)X_t = \varepsilon_t \]

Regression Model formulation

\[ Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \]
\[ \phi(L)Y_t = c + \varepsilon_t, \quad c = \mu \phi(1) \]
\[ \phi(1) = 1 - \phi_1 - \cdots - \phi_p \]
Stability and Stationarity Conditions

Trick: Write AR(p) as a vector AR(1) (VAR(1))

\[
\begin{bmatrix}
X_t \\
X_{t-1} \\
X_{t-2} \\
\vdots \\
X_{t-p+1}
\end{bmatrix}
= \begin{bmatrix}
\phi_1 & \phi_2 & \cdots & \cdots & \phi_p \\
1 & 0 & \cdots & \cdots & 0 \\
0 & 1 & \ddots & \cdots & 0 \\
\vdots & \vdots & \ddots & \cdots & \vdots \\
0 & 0 & \cdots & 1 & 0
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
X_{t-2} \\
X_{t-3} \\
\vdots \\
X_{t-p}
\end{bmatrix}
+ \begin{bmatrix}
\varepsilon_t \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

or

\[
\xi_t = F \xi_{t-1} + v_t
\]

(\(p \times 1\)) \((p \times p)(p \times 1)\) \((p \times 1)\)

Use insights from AR(1) to study behavior of VAR(1):

\[
\xi_{t+j} = F^{j+1} \xi_{t-1} + F^j v_t + \cdots + F v_{t+j-1} + v_t
\]

\[
F^j = F \times F \times \cdots \times F \ (j \ \text{times})
\]
Intuition: Stability and stationarity requires

$$\lim_{j \to \infty} F^j = 0$$

Initial value has no impact on eventual level of series.
Example: AR(2)

\[ X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \]

or

\[
\begin{bmatrix}
X_t \\
X_{t-1}
\end{bmatrix} =
\begin{bmatrix}
\phi_1 & \phi_2 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
X_{t-2}
\end{bmatrix}
+ \begin{bmatrix}
\varepsilon_t \\
0
\end{bmatrix}
\]

Iterating \( j \) periods out gives

\[
\begin{bmatrix}
X_{t+j} \\
X_{t+j-1}
\end{bmatrix} =
\begin{bmatrix}
\phi_1 & \phi_2 \\
1 & 0
\end{bmatrix}^{j+1}
\begin{bmatrix}
X_{t-1} \\
X_{t-2}
\end{bmatrix}
+ \begin{bmatrix}
\phi_1 & \phi_2 \\
1 & 0
\end{bmatrix}^j
\begin{bmatrix}
\varepsilon_t \\
0
\end{bmatrix}
+ \cdots +
\begin{bmatrix}
\phi_1 & \phi_2 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{t+j-1} \\
0
\end{bmatrix}
+ \begin{bmatrix}
\varepsilon_{t+j} \\
0
\end{bmatrix}
\]
First row gives $X_{t+j}$

$$X_{t+j} = [f_{11}^{(j+1)}X_{t-1} + f_{12}^{(j+1)}X_{t-2}] + f_{11}^{(j)} \varepsilon_t$$

$$+ \cdots + f_{11} \varepsilon_{t+j-1} + \varepsilon_{t+j}$$

$$f_{11}^{(j)} = (1, 1) \text{ element of } \mathbf{F}^j$$

Note:

$$\mathbf{F}^2 = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \phi_1^2 + \phi_2 & \phi_1 \phi_2 \\ \phi_1 & \phi_2 \end{bmatrix}$$
Result: The ARMA$(p, 0)$ model is covariance stationary and has Wold representation

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \; \psi_0 = 1$$

with $\psi_j = (1, 1)$ element of $F^j$ provided all of the eigen-values of $F$ have modulus less than 1.
Finding Eigenvalues

λ is an eigenvalue of \( F \) and \( x \) is an eigenvector if

\[
Fx = \lambda x \Rightarrow (F - \lambda I_p)x = 0
\]

\[
\Rightarrow F - \lambda I_p \text{ is singular} \Rightarrow \det(F - \lambda I_p) = 0
\]

Example: AR(2)

\[
\det(F - \lambda I_2) = \det\left(\begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)
\]

\[
= \det\left(\begin{pmatrix} \phi_1 - \lambda & \phi_2 \\ 1 & -\lambda \end{pmatrix}\right)
\]

\[
= \lambda^2 - \phi_1 \lambda - \phi_2
\]
The eigenvalues of $F$ solve the “reverse” characteristic equation
\[ \lambda^2 - \phi_1 \lambda - \phi_2 = 0 \]
Using the quadratic equation, the roots satisfy
\[ \lambda_i = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}, \quad i = 1, 2 \]
These root may be real or complex. Complex roots induce periodic behavior in $y_t$. Recall, if $\lambda_i$ is complex then
\[ \lambda_i = a + bi \]
\[ a = R \cos(\theta), \quad b = R \sin(\theta) \]
\[ R = \sqrt{a^2 + b^2} = \text{modulus} \]
To see why $|\lambda_i| < 1$ implies $\lim_{j \to \infty} F^j = 0$ consider the AR(2) with real-valued eigenvalues. By the spectral decomposition theorem

$$F = T \Lambda T^{-1}, \quad T^{-1} = T'$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix},$$

$$T^{-1} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}$$

Then

$$F^j = (T \Lambda T^{-1}) \times \cdots \times (T \Lambda T^{-1})$$

$$= T \Lambda^j T^{-1}$$

and

$$\lim_{j \to \infty} F^j = T \lim_{j \to \infty} \Lambda^j T^{-1} = 0$$

provided $|\lambda_1| < 1$ and $|\lambda_2| < 1$. 
Note:

\[
F^j = T \Lambda^j T^{-1}
\]

\[
= \begin{bmatrix}
t_{11} & t_{12} \\
t_{21} & t_{22}
\end{bmatrix}
\begin{bmatrix}
\lambda_1^j & 0 \\
0 & \lambda_2^j
\end{bmatrix}
\begin{bmatrix}
t_{11} & t_{12} \\
t_{21} & t_{22}
\end{bmatrix}
\]

so that

\[
f_{11}^{(j)} = (t_{11}t^{11})\lambda_1^j + (t_{12}t^{22})\lambda_2^j
\]

\[
= c_1\lambda_1^j + c_2\lambda_2^j = \psi_j
\]

where

\[
c_1 + c_2 = 1
\]

Then,

\[
\lim_{j \to \infty} \psi_j = \lim_{j \to \infty} (c_1\lambda_1^j + c_2\lambda_2^j) = 0
\]
Examples of AR(2) Processes

\[ Y_t = 0.6Y_{t-1} + 0.2Y_{t-2} + \varepsilon_t \]

\[ \phi_1 + \phi_2 = 0.8 < 1 \]

\[ F = \begin{bmatrix} 0.6 & 0.2 \\ 1 & 0 \end{bmatrix} \]

The eigenvalues are found using

\[ \lambda_i = \frac{\phi \pm \sqrt{\phi_1^2 + 4\phi_2}}{2} \]

\[ \lambda_1 = \frac{0.6 + \sqrt{(0.6)^2 + 4(0.2)}}{2} = 0.84 \]

\[ \lambda_2 = \frac{0.6 - \sqrt{(0.6)^2 + 4(0.2)}}{2} = -0.24 \]

\[ \psi_j = c_1(0.84)^j + c_2(-0.24)^j \]
\[ Y_t = 0.5Y_{t-1} - 0.8Y_{t-2} + \varepsilon_t \]

\[ \phi_1 + \phi_2 = -0.3 < 1 \]

\[ F = \begin{bmatrix} 0.5 & -0.8 \\ 1 & 0 \end{bmatrix} \]

Note:

\[ \phi_1^2 + 4\phi_2 = (0.5)^2 - 4(0.8) = -2.95 \]

\[ \Rightarrow \text{complex eigenvalues} \]

Then

\[ \lambda_i = a \pm bi, \ i = \sqrt{-1} \]

\[ a = \frac{\phi_1}{2}, \ b = \frac{\sqrt{-(\phi_1^2 + 4\phi_2)}}{2} \]
Here

\[ a = \frac{0.5}{2} = 0.25, \quad b = \frac{\sqrt{2.95}}{2} = 0.86 \]

\[ \lambda_i = 0.25 \pm 0.86i \]

modulus \quad \begin{align*} R &= \sqrt{a^2 + b^2} = \sqrt{(0.25)^2 + (0.86)^2} = 0.895 \end{align*} \\

Polar co-ordinate representation:

\[ \lambda_i = a + bi \text{ s.t. } a = R \cos(\theta), \quad b = R \sin(\theta) \]

\[ = R \cos(\theta) + R \sin(\theta)i = R e^{i\theta} \]

Frequency \( \theta \) satisfies

\[ \cos(\theta) = \frac{a}{R} \implies \theta = \cos^{-1}\left(\frac{a}{R}\right) \]

period \quad \begin{align*} &\frac{2\pi}{\theta} \end{align*}
Here,

\[ R = 0.895 \]
\[ \theta = \cos^{-1} \left( \frac{0.25}{0.985} \right) = 1.29 \]
\[ \text{period} = \frac{2\pi}{1.29} = 4.9 \]

Note: the period is the length of time required for the process to repeat a full cycle.

Note: The IRF has the form

\[ \psi_j = c_1 \lambda_1^j + c_2 \lambda_2^j \]
\[ \propto R^j \left[ \cos(\theta j) + \sin(\theta j) \right] \]
Stationarity Conditions on Lag Polynomial $\phi(L)$

Consider the AR(2) model in lag operator notation

$$(1 - \phi_1 L - \phi_2 L^2)X_t = \phi(L)X_t = \varepsilon_t$$

For any variable $z$, consider the characteristic equation

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 = 0$$

By the fundamental theorem of algebra

$$1 - \phi_1 z - \phi_2 z^2 = (1 - \lambda_1 z)(1 - \lambda_2 z)$$

so that

$$z_1 = \frac{1}{\lambda_1}, \quad z_2 = \frac{1}{\lambda_2}$$

are the roots of the characteristic equation. The values $\lambda_1$ and $\lambda_2$ are the eigenvalues of $F$. 
Note: If $\phi_1 + \phi_2 = 1$ then $\phi(z = 1) = 1 - (\phi_1 + \phi_2) = 0$ and $z = 1$ is a root of $\phi(z) = 0$. 
Result: The inverses of the roots of the characteristic equation

\[ \phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots \phi_p z^p = 0 \]

are the eigenvalues of the companion matrix \( F \). Therefore, the AR(p) model is stable and stationary provided the roots of \( \phi(z) = 0 \) have modulus greater than unity (roots lie outside the complex unit circle).

Remarks:

1. The reverse characteristic equation for the AR(p) is

\[ z^p - \phi_1 z^{p-1} - \phi_2 z^{p-2} - \cdots - \phi_{p-1} z - \phi_p = 0 \]

This is the same polynomial equation used to find the eigenvalues of \( F \).
2. If the AR(p) is stationary, then

$$\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p = (1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_p L)$$

where $|\lambda_i| < 1$. Suppose, all $\lambda_i$ are all real. Then

$$(1 - \lambda_i L)^{-1} = \sum_{j=0}^{\infty} \lambda_i^j L^j$$

$$\phi(L)^{-1} = (1 - \lambda_1 L)^{-1} \cdots (1 - \lambda_p L)^{-1}$$

$$= \left( \sum_{j=0}^{\infty} \lambda_1^j L^j \right) \left( \sum_{j=0}^{\infty} \lambda_2^j L^j \right) \cdots \left( \sum_{j=0}^{\infty} \lambda_p^j L^j \right)$$
The Wold solution for $X_t$ may be found using

$$X_t = \phi(L)^{-1}\varepsilon_t$$

$$= \left( \sum_{j=0}^{\infty} \lambda_1^j L^j \right) \left( \sum_{j=0}^{\infty} \lambda_2^j L^j \right) \ldots \left( \sum_{j=0}^{\infty} \lambda_p^j L^j \right) \varepsilon_t$$
3. A simple algorithm exists to determine the Wold form. To illustrate, consider the AR(2) model. By definition

\[
\phi(L)^{-1} = (1 - \phi_1 L - \phi_2 L^2)^{-1} = \psi(L),
\]

\[
\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j
\]

\[
\Rightarrow 1 = \phi(L)\psi(L)
\]

\[
\Rightarrow 1 = (1 - \phi_1 L - \phi_2 L^2)
\times (1 + \psi_1 L + \psi_2 L^2 + \cdots)
\]

Collecting coefficients of powers of \( L \) gives

\[
1 = 1 + (\psi_1 - \phi_1)L + (\psi_2 - \phi_1\psi_1 - \phi_2)L^2 + \cdots
\]
Since all coefficients on powers of $L$ must be equal to zero, it follows that

\[
\begin{align*}
\psi_1 &= \phi_1 \\
\psi_2 &= \phi_1 \psi_1 + \phi_2 \\
\psi_3 &= \phi_1 \psi_2 + \phi_2 \psi_1 \\
&\vdots \\
\psi_j &= \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2}
\end{align*}
\]
Moments of Stationary AR(p) Model

\[ Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \cdots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t \]
\[ \varepsilon_t \sim WN(0, \sigma^2) \]

or

\[ Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \]
\[ c = \mu(1 - \pi) \]
\[ \pi = \phi_1 + \phi_2 + \cdots + \phi_p \]

Note: if \( \pi = 1 \) then \( \phi(1) = 1 - \pi = 0 \) and \( z = 1 \) is a root of \( \phi(z) = 0 \). In this case we say that the AR(p) process has a unit root and the process is nonstationary.
Straightforward algebra shows that

\[
E[Y_t] = \mu \\
\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma^2 \\
\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p} \\
\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}
\]

The recursive equations for \(\rho_j\) are called the Yule-Walker equations.

Result: \((\gamma_0, \gamma_1, \ldots, \gamma_{p-1})\) is determined from the first \(p\) elements of the first column of the \((p^2 \times p^2)\) matrix

\[
\sigma^2[I_{p^2} - (F \otimes F)]^{-1}
\]

where \(F\) is the state space companion matrix for the AR(p) model.