Lecture Notes on Autoregression

Econ 624

March 26, 2021

1These notes are for instructional purposes only and are not to be distributed outside of the classroom.
Consider the process

\[ y_t = \mu + \alpha y_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}, \]

where we assume that

\[ |\alpha| < 1, \]
\[ \{\epsilon_t\} \equiv i.i.d. (0, \sigma^2_{\epsilon}) \text{ with } 0 < \sigma^2_{\epsilon} < \infty. \]

To obtain the moving average representation of the AR(1) process, we shall consider here an approach due to Kasparis (2016). To proceed, note that, by using the lag operator notation, we can rewrite this process as

\[ \alpha(L)y_t = (1 - \alpha L)y_t = \mu + \epsilon_t, \quad t \in \mathbb{Z}, \]

where

\[ \alpha(L) = 1 - \alpha L. \]
Hence, a moving-average representation can be obtained if we can invert the lag operator $\alpha(L)$ to obtain

\[
y_t = \alpha(L)^{-1}(\mu + \epsilon_t)
= (1 - \alpha L)^{-1}(\mu + \epsilon_t), \ t \in \mathbb{Z}.
\]
**Definition:** Let \( V \) be a real (or complex) linear space (vector space). A function \( \| \cdot \| : V \to \mathbb{R} \) with the properties

- **P1**
  \[ \| \varphi \| \geq 0 \text{ (positivity)} \]

- **P2**
  \[ \| \varphi \| = 0 \text{ if and only if } \varphi = 0 \text{ (definiteness)} \]

- **P3**
  \[ \| \alpha \varphi \| = |\alpha| \| \varphi \| \text{ (homogeneity)} \]

- **P4**
  \[ \| \varphi + \psi \| \leq \| \varphi \| + \| \psi \| \text{ (triangle inequality)} \]

for all \( \varphi, \psi \in V \) and for all \( \alpha \in \mathbb{R} \) (or \( \mathbb{C} \)) is called a norm on \( V \). A linear space \( V \) equipped with a norm is called a normed space.
**Definition:** A linear operator $A : \mathcal{V} \rightarrow \mathcal{W}$ from a normed space $\mathcal{V}$ into a normed space $\mathcal{W}$ is called bounded if there exists a positive constant $C$ such that

$$\|A\varphi\| \leq C \|\varphi\|$$

for all $\varphi \in \mathcal{V}$. Each number $C$ for which the inequality holds is called a bound for the operator $A$.

**Remark:** With the aid of linearity of the operator $A$, it is easy to see that there exists a positive constant $C$ such that

$$\|A\varphi\| \leq C \|\varphi\|$$

if and only if

$$\frac{\|A\varphi\|}{\|\varphi\|} = \left\| A \frac{\varphi}{\|\varphi\|} \right\| \quad \text{(by homogeneity)}$$

$$\leq C$$

for all $\varphi \in \mathcal{V}$.
Remark (con’t): Moreover, define

\[ \phi^* = \frac{\phi}{\|\phi\|} \]

so that

\[ \|\phi^*\| = \left\| \frac{\phi}{\|\phi\|} \right\| \] (again by homogeneity)

\[ = \frac{\|\phi\|}{\|\phi\|} \]

\[ = 1 \]

Note that \( \phi^* \in \mathbb{V} \) since a linear space is closed under scalar multiplication.
Hence, $A$ is bounded if and only if there exists a positive constant $C$ such that

$$\|A\varphi^*\| \leq C$$

for all $\varphi^* \in \mathbb{V}$ such that $\|\varphi^*\| = 1$. It, thus, follows that $A$ is bounded if and only if

$$\|A\| := \sup_{\|\varphi\| = 1} \|A\varphi\| = \sup_{\|\varphi\| \leq 1} \|A\varphi\| < \infty$$

where the number $\|A\|$ is the smallest bound for $A$ and is called the norm of $A$. Thus, a linear operator is bounded if and only if it maps bounded sets in $\mathbb{V}$, i.e., $\{\varphi : \|\varphi\| \leq 1\}$ into bounded sets in $\mathbb{W}$. 
Definition (Cauchy Sequence): A sequence \( \{ \varphi_n \} \) of elements in a normed space \( \mathbb{V} \) is called a Cauchy sequence if for every \( \epsilon > 0 \), there exists an integer \( N(\epsilon) \) such that

\[
\| \varphi_n - \varphi_m \| < \epsilon
\]

for all \( n, m \geq N(\epsilon) \); that is, if

\[
\lim_{n, m \to \infty} \| \varphi_n - \varphi_m \| = 0.
\]

Definition (Completeness and Banach Space): A subset \( \mathbb{U} \) of a normed space is called complete if every Cauchy sequence of elements in \( \mathbb{U} \) converges to an element in \( \mathbb{U} \). A normed space \( \mathbb{U} \) is called a Banach space if it is complete.
Theorem 1: Let $\mathbb{B}$ be a Banach space. If $T : \mathbb{B} \rightarrow \mathbb{B}$ is bounded linear operator and $\|I - T\| < 1$, where $I : \mathbb{B} \rightarrow \mathbb{B}$ denotes the identity operator. Then, $T$ has a bounded inverse operator on $\mathbb{B}$ which is given by the Neumann series

$$T^{-1} = \sum_{k=0}^{\infty} (I - T)^k$$

and which satisfies

$$\|T^{-1}\| \leq \frac{1}{1 - \|I - T\|}.$$

The iterated operators $(I - T)^k$ are defined by $(I - T)^0 := I$ and $(I - T)^k := (I - T)(I - T)^{k-1}$ for $k \in \mathbb{N}$. 
First-Order Autoregressive Process (or AR(1) Process)

- Consider the space $\mathbb{V}$ of sequences $Y(\omega) = \{Y_t(\omega)\}_{t \in \mathbb{Z}}$ on some probability space $(\Omega, \mathcal{F}, P)$ that satisfies

$$\sup_t E|Y_t| < \infty.$$ 

We can take $\mathbb{V}$ to be a normed space, equipped with the (pseudo or semi) norm

$$\|Y\| = \sup_{t \in \mathbb{Z}} E|Y_t|.$$ 

- **Remark:** Note that any covariance stationary sequence belongs to the space $\mathbb{V}$, but the setup here is more general since covariance stationarity requires a higher (second) moment condition as well as homogeneity of the first two moments with respect to $t$.

- **Lemma 1 (Kasparis, 2016):** The normed space $\mathbb{V}$ is complete and therefore is a Banach space.
Some Calculations: Since
\[
\sup_{t \in \mathbb{Z}} E \left| Y_t \right| = \sup_{t \in \mathbb{Z}} E \left| Y_{t-1} \right| = \sup_{t \in \mathbb{Z}} E \left| LY_t \right|,
\]
we have
\[
\| L \| = \sup_{\| Y \| = 1} \| LY \| = \sup_{\{ Y_t \} \in \mathcal{V} : \sup_{t \in \mathbb{Z}} E \left| Y_t \right| = 1} \left( \sup_{t \in \mathbb{Z}} E \left| Y_{t-1} \right| \right) = 1.
\]
Some Calculations (con't): Similarly, we have

\[
\sup_{t \in \mathbb{Z}} E \left| Y_t \right| = \sup_{t \in \mathbb{Z}} E \left| Y_{t+1} \right| = \sup_{t \in \mathbb{Z}} E \left| L^{-1} Y_t \right|, 
\]

so that

\[
\left\| L^{-1} \right\| = \sup_{\|Y\| = 1} \left\| L^{-1} Y \right\| 
\]

\[
= \sup_{\{Y_t\} \in \mathcal{V} : \sup_{t \in \mathbb{Z}} E |Y_t| = 1} \left( \sup_{t \in \mathbb{Z}} E |Y_{t+1}| \right) 
\]

\[
= 1. 
\]

where \( L^{-1} \) is the inverse of \( L \).
Some Calculations (con’t):

Now, given the assumption $|\alpha| < 1$, we have

\[
\|1 - \alpha (L)\| = \|1 - (1 - \alpha L)\| \\
= \|\alpha L\| \\
= |\alpha| \|L\| \\
< 1.
\]
Some Calculations (con’t): We can, thus, apply Theorem 1 to the AR (1) case by taking

\[ T = \alpha (L) = 1 - \alpha L \]

and note that

\[ \alpha (L)^{-1} = \sum_{j=0}^{\infty} (1 - [1 - \alpha L])^j \]

\[ = \sum_{j=0}^{\infty} \alpha^j L^j. \]
Some Calculations (con’t): It then follows that

\[
y_t = \alpha (L)^{-1} (\mu + \epsilon_t)
\]

\[
= \sum_{j=0}^{\infty} \alpha^j L^j (\mu + \epsilon_t)
\]

\[
= \mu \sum_{j=0}^{\infty} \alpha^j + \sum_{j=0}^{\infty} \alpha^j \epsilon_{t-j}
\]

\[
= \frac{\mu}{1 - \alpha} + \sum_{j=0}^{\infty} \alpha^j \epsilon_{t-j}
\]

where the last equality is obtained by applying the summation formula for geometric series

\[
\sum_{j=0}^{\infty} \alpha^j = \frac{1}{1 - \alpha}
\]

given that \(|\alpha| < 1\).
Remarks:

(i) The representation

\[ y_t = \frac{\mu}{1 - \alpha} + \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j} \]  

is often referred to as the stationary solution of the AR(1) process

\[ y_t = \mu + \alpha y_{t-1} + \varepsilon_t, \ t \in \mathbb{Z}. \]

Moreover, this AR(1) process is said to be causal because it only depends on past innovations.

(ii) From expression (1), we see that \{y_t\} is strictly stationary and ergodic since it is a measurable transformation of \{\varepsilon_t\}, which is an i.i.d. sequence and, thus, strictly stationary and ergodic.
Remarks (con’t):

(iii) Note, in addition, that in this case

\[
\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\alpha^j| = \sum_{j=0}^{\infty} |\alpha|^j = \frac{1}{1 - |\alpha|} < \infty
\]

so that the moving average coefficients are absolutely summable.
Moments: We can exploit the stationarity of the AR(1) process to calculate its mean, variance, and autocovariances.

1. Mean:

\[ E[y_t] = E[\mu] + E[\alpha y_{t-1}] + E[\varepsilon_t] \]
\[ = \mu + \alpha E[y_{t-1}] \quad (\text{since } E[\varepsilon_t] = 0 \text{ by assumption}) \]

This implies that
\[ E[y_t] = \alpha E[y_{t-1}] = \mu \]
or
\[ (1 - \alpha) E[y_t] = \mu \quad (\text{since by stationarity } E[y_t] = E[y_{t-1}]) \]

from which we deduce that
\[ E[y_t] = \frac{\mu}{1 - \alpha} \]

which is well-defined given that \(|\alpha| < 1\).
Moments (con’t):

2. Variance:

\[ \text{Var} \left( y_t \right) = \text{Var} \left( \mu + \alpha y_{t-1} + \varepsilon_t \right) \]
\[ = \alpha^2 \text{Var} \left( y_{t-1} \right) + \text{Var} \left( \varepsilon_t \right) + 2 \text{Cov} \left( y_{t-1}, \varepsilon_t \right) \]
\[ = \alpha^2 \text{Var} \left( y_{t-1} \right) + \sigma^2_{\varepsilon}. \]

It follows that

\[ \sigma^2_{\varepsilon} = \text{Var} \left( y_t \right) - \alpha^2 \text{Var} \left( y_{t-1} \right) \]
\[ = (1 - \alpha^2) \text{Var} \left( y_t \right) \quad \text{(by stationarity)} \]

or

\[ \text{Var} \left( y_t \right) = \frac{\sigma^2_{\varepsilon}}{1 - \alpha^2} < \infty \quad \text{(given that } 0 < \sigma^2_{\varepsilon} < \infty \text{ and } |\alpha| < 1) \]
Moments (con’t):

3. Autocovariances: For positive integer \( h \), note that

\[
\text{Cov} \left( y_t, y_{t-h} \right) = \text{Cov} \left( \mu + \alpha y_{t-1} + \varepsilon_t, y_{t-h} \right) = \alpha \text{Cov} \left( y_{t-1}, y_{t-h} \right) + \text{Cov} \left( \varepsilon_t, y_{t-h} \right) = \alpha \text{Cov} \left( y_{t-1}, y_{t-h} \right)
\]

Now, for \( h = 1 \), we have

\[
\text{Cov} \left( y_t, y_{t-1} \right) = \alpha \text{Cov} \left( y_{t-1}, y_{t-1} \right) = \alpha \text{Var} \left( y_t \right) \text{ (by stationarity)} = \frac{\alpha \sigma^2_\varepsilon}{1 - \alpha^2} = \gamma_y (1).
\]
3. **Autocovariances (con’t):** Iterating backwards, we also see that for \( h = 2 \)

\[
\text{Cov}(y_t, y_{t-2}) = \alpha \text{Cov}(y_{t-1}, y_{t-2}) \\
= \alpha \text{Cov}(y_t, y_{t-1}) \quad \text{(by stationarity)} \\
= \frac{\alpha^2 \sigma^2}{1 - \alpha^2} \\
= \gamma_y(2).
\]

It is, thus, clear that for any positive integer \( h \), we would have

\[
\text{Cov}(y_t, y_{t-h}) = \frac{\alpha^h \sigma^2}{1 - \alpha^2} \\
= \gamma_y(h).
\]
3. Autocovariances (con’t): In addition, for negative integer $h$, we have

\[ \text{Cov} \left( y_t, y_{t-h} \right) = \text{Cov} \left( y_t, y_{t+|h|} \right) \]
\[ = \text{Cov} \left( y_{t-|h|}, y_t \right) \quad \text{(by stationarity)} \]
\[ = \text{Cov} \left( y_t, y_{t-|h|} \right) \]
\[ = \alpha^{|h|} \sigma^2 \]
\[ = \frac{\alpha^{|h|} \sigma^2}{1 - \alpha^2} \]

Hence, in general, for any integer $h$, we have

\[ \gamma_y (h) = \frac{\alpha^{|h|} \sigma^2}{1 - \alpha^2}. \]
Consider the process

\[ y_t = \mu + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}, \]

where \( \{\varepsilon_t\} \equiv i.i.d. (0, \sigma^2_{\varepsilon}) \) with \( 0 < \sigma^2_{\varepsilon} < \infty \). Again, using the lag operator notation, we can rewrite this process as

\[ \alpha(L) y_t = \mu + \varepsilon_t, \quad t \in \mathbb{Z}, \]

where

\[ \alpha(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p \]
Now, factor this lag operator polynomial as follows:

\[ \alpha(L) = \left(1 - \frac{1}{\rho_1}L\right) \left(1 - \frac{1}{\rho_2}L\right) \times \cdots \times \left(1 - \frac{1}{\rho_p}L\right) \]

\[ = \alpha_1(L) \alpha_2(L) \times \cdots \times \alpha_p(L) \]

where \( \{\rho_i : i = 1, \ldots, p\} \) are the roots of the polynomial equation \( \alpha(z) = 0, \ z \in \mathbb{C}, \) and where

\[ \alpha_i(L) = 1 - \frac{1}{\rho_i}L \text{ for } i = 1, \ldots, p. \]

We assume that \( |\rho_i| > 1 \) for every \( i \in \{1, \ldots, p\} \).

**Lemma 2:** Let \( V \) be a normed space and suppose that the operators \( T_i : V \to V, \) with \( i \in \{1, \ldots, p\}, \) commute. Define \( T \) as \( T_1 T_2 \cdots T_p. \) Then, \( T \) is invertible if and only if each \( T_i \) is invertible.
Next, let

\[ T_i = \alpha_i (L) = 1 - \frac{1}{\rho_i} \]

and note that

\[ \|1 - T_i\| = \left\|1 - \left(1 - \frac{1}{\rho_i} L\right)\right\| = \left\|\frac{1}{\rho_i} L\right\| = \left|\frac{1}{\rho_i}\right| \|L\| < 1 \]

Hence, by Theorem 1 given earlier,

\[ T_i^{-1} = \alpha_i (L)^{-1} = \sum_{j=0}^{\infty} \left(1 - \left[1 - \frac{1}{\rho_i} L\right]\right)^j \]

\[ = \sum_{j=0}^{\infty} \left(\frac{1}{\rho_i}\right)^j L^j \text{ for every } i \in \{1, \ldots, p\} . \]
Hence, a moving-average representation for the $AR(p)$ process can be obtained by inverting the lag operator $\alpha(L)$ to obtain

$$y_t = \alpha(L)^{-1}(\mu + \varepsilon_t)$$

$$= \left(1 - \frac{1}{\rho_1}L\right)^{-1}\left(1 - \frac{1}{\rho_2}L\right)^{-1} \times \cdots \times \left(1 - \frac{1}{\rho_p}L\right)^{-1}(\mu + \varepsilon_t)$$

To give a more explicit form of the moving average representation, we look first at the explicit case where $p = 2$. 
Example (AR(2) process): In this case,

\[ \alpha(L) = \left(1 - \frac{1}{\rho_1} L \right) \left(1 - \frac{1}{\rho_2} L \right) \]

so that

\[ \alpha(L)^{-1} = \left(1 - \frac{1}{\rho_1} L \right)^{-1} \left(1 - \frac{1}{\rho_2} L \right)^{-1} \]

\[ = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left( \left(1 - \frac{1}{\rho_1} L \right) \right)^{j_1} \left( \left(1 - \frac{1}{\rho_2} L \right) \right)^{j_2} \]

\[ = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left( \frac{1}{\rho_1} \right)^{j_1} \left( \frac{1}{\rho_2} \right)^{j_2} L^{j_1+j_2} \]
Example (cont'): AR(2) Process

Let \( j = j_1 + j_2 \) and \( k = j_1 \). By rearranging the sums, we can further write

\[
\alpha (L)^{-1} = \sum_{j=0}^{\infty} \sum_{k=0}^{j} \left( \frac{1}{\rho_1} \right)^k \left( \frac{1}{\rho_2} \right)^{j-k} L^j \\
= \sum_{j=0}^{\infty} \psi_j L^j
\]

where

\[
\psi_j = \sum_{k=0}^{j} \left( \frac{1}{\rho_1} \right)^k \left( \frac{1}{\rho_2} \right)^{j-k}.
\]
AR(2) Process (cont): Making use of the representations of $\alpha (L)^{-1}$ given previously, we have

$$y_t = \alpha (L)^{-1} (\mu + \epsilon_t)$$

$$= \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left( \frac{1}{\rho_1} \right)^{j_1} \left( \frac{1}{\rho_2} \right)^{j_2} L^{j_1+j_2} \mu$$

$$+ \sum_{j=0}^{\infty} \sum_{k=0}^{j} \left( \frac{1}{\rho_1} \right)^k \left( \frac{1}{\rho_2} \right)^{j-k} L^j \epsilon_t$$

$$= \mu \left[ \frac{1}{1 - (1/\rho_1)} \right] \left[ \frac{1}{1 - (1/\rho_2)} \right] + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

$$= \mu \left( \prod_{i=1}^{2} \frac{1}{1 - (1/\rho_i)} \right) + \psi (L) \epsilon_t$$

where $\psi (L) = \sum_{j=0}^{\infty} \psi_j L^j$. 
**AR (2) Process (con’t):** Moreover, note that

\[
\mu \left( \prod_{i=1}^{2} \frac{1}{1 - (1/\rho_i)} \right) = \left( 1 - \frac{1}{\rho_1} L \right)^{-1} \left( 1 - \frac{1}{\rho_2} L \right)^{-1} \mu \\
= \left[ \left( 1 - \frac{1}{\rho_1} L \right) \left( 1 - \frac{1}{\rho_2} L \right) \right]^{-1} \mu \\
= (1 - \alpha_1 L - \alpha_2 L^2)^{-1} \mu \\
= \frac{\mu}{1 - \alpha_1 - \alpha_2}.
\]
**AR (2) Process (con’t):** Finally, note that the moving-average coefficients in this case are absolutely summable since

\[
\sum_{j=0}^{\infty} \left| \psi_j \right| \leq \sum_{j=0}^{\infty} \sum_{k=0}^{j} \left| \frac{1}{\rho_1} \right|^k \left| \frac{1}{\rho_2} \right|^{j-k}
\]

\[
\leq \sum_{k=0}^{\infty} \left| \frac{1}{\rho_1} \right|^k \sum_{j=0}^{\infty} \left| \frac{1}{\rho_2} \right|^j
\]

\[
= \left( \frac{1}{1 - \left| 1/\rho_1 \right|} \right) \left( \frac{1}{1 - \left| 1/\rho_2 \right|} \right)
\]

\[
< \infty
\]
Returning to the more general $AR(p)$ process, note that, in a similar manner, we have

$$y_t = \alpha (L)^{-1} (\mu + \varepsilon_t)$$

$$= \left(1 - \frac{1}{\rho_1} L\right)^{-1} \left(1 - \frac{1}{\rho_2} L\right)^{-1} \times \cdots \times \left(1 - \frac{1}{\rho_p} L\right)^{-1} (\mu + \varepsilon_t)$$

$$= \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_p=0}^{\infty} \left\{ \left(\frac{1}{\rho_1}\right)^{j_1} \left(\frac{1}{\rho_2}\right)^{j_2} \times \cdots \times \left(\frac{1}{\rho_p}\right)^{j_p} L^{j_1 + \cdots + j_p} (\mu + \varepsilon_t) \right\}$$
By letting \( j = j_1 + \cdots + j_p \) and rearranging the sums, we further obtain

\[
y_t = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_p=0}^{\infty} \left\{ \left( \frac{1}{\rho_1} \right)^{j_1} \left( \frac{1}{\rho_2} \right)^{j_2} \times \cdots \times \left( \frac{1}{\rho_p} \right)^{j_p} \times L^{j_1 + \cdots + j_p} (\mu + \varepsilon_t) \right\}
\]

\[
= \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_p=0}^{\infty} \left( \frac{1}{\rho_1} \right)^{j_1} \left( \frac{1}{\rho_2} \right)^{j_2} \times \cdots \times \left( \frac{1}{\rho_p} \right)^{j_p} L^{j_1 + \cdots + j_p} \mu
\]

\[
+ \sum_{j=0}^{\infty} \sum_{k_1=0}^{j} \sum_{k_2=0}^{j-k_1} \cdots \sum_{k_{p-1}=0}^{j-(k_1 + \cdots + k_{p-2})} \left\{ \left( \frac{1}{\rho_1} \right)^{k_1} \left( \frac{1}{\rho_2} \right)^{k_2} \times \cdots \times \left( \frac{1}{\rho_p} \right)^{j-(k_1 + \cdots + k_{p-1})} L^j \varepsilon_t \right\}.
\]
p-th Order Autoregressive Process (or AR(p) Process)

- It follows by argument similar to that given previously for the $AR(2)$ case that

$$y_t = \alpha (L)^{-1} (\mu + \varepsilon_t)$$

$$= \left(1 - \frac{1}{\rho_1} L\right)^{-1} \left(1 - \frac{1}{\rho_2} L\right)^{-1} \times \cdots \times \left(1 - \frac{1}{\rho_p} L\right)^{-1} (\mu + \varepsilon_t)$$

$$= \mu \left[\frac{1}{1 - (1/\rho_1)}\right] \left[\frac{1}{1 - (1/\rho_2)}\right] \times \cdots \times \left[\frac{1}{1 - (1/\rho_p)}\right]$$

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

$$= \mu \left(\prod_{i=1}^{p} \frac{1}{1 - (1/\rho_i)}\right) + \psi (L) \varepsilon_t$$
p-th Order Autoregressive Process (or AR(p) Process)

where

$$\psi_j = \sum_{k_1=0}^{j} \sum_{k_2=0}^{j-k_1} \cdots \sum_{k_{p-1}=0}^{j-(k_1+\cdots+k_{p-2})} \left\{ \left( \frac{1}{\rho_1} \right)^{k_1} \left( \frac{1}{\rho_2} \right)^{k_2} \times \cdots \times \left( \frac{1}{\rho_p} \right)^{j-(k_1+\cdots+k_{p-1})} \right\}.$$
In addition, note that

\[
\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} \left( \sum_{k_1=0}^{j-k_1} \sum_{k_2=0}^{j-(k_1+\cdots+k_{p-2})} \cdots \sum_{k_{p-1}=0}^{j-(k_1+\cdots+k_{p-2})} \left\{ \left| \frac{1}{\rho_1} \right|^{k_1} \left| \frac{1}{\rho_2} \right|^{k_2} \times \cdots \times \left| \frac{1}{\rho_p} \right|^{j-(k_1+\cdots+k_{p-1})} \right\} \right) \sum_{k_1=0}^{\infty} \left| \frac{1}{\rho_1} \right|^{k_1} \sum_{k_2=0}^{\infty} \left| \frac{1}{\rho_2} \right|^{k_2} \times \cdots \times \sum_{j=0}^{\infty} \left| \frac{1}{\rho_p} \right|^{j}.
\]
p-th Order Autoregressive Process (or AR(p) Process)

so that, applying the summation formula for a convergent geometric series, we obtain

\[
\sum_{j=0}^{\infty} |\psi_j| \\
\leq \sum_{k_1=0}^{\infty} \left|\frac{1}{\rho_1}\right|^{k_1} \sum_{k_2=0}^{\infty} \left|\frac{1}{\rho_2}\right|^{k_2} \times \cdots \times \sum_{j=0}^{\infty} \left|\frac{1}{\rho_p}\right|^j \\
= \left(\frac{1}{1-|1/\rho_1|}\right) \left(\frac{1}{1-|1/\rho_2|}\right) \times \cdots \times \left(\frac{1}{1-|1/\rho_p|}\right) \\
< \infty
\]

which shows the absolute summability of the moving average coefficients of the AR(p) process.
Remark: It follows from the moving average representation

\[ y_t = \mu \left( \prod_{i=1}^{p} \frac{1}{1 - (1/\rho_i)} \right) + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \]

that \( \{y_t\} \) is strictly stationary and ergodic since it is measurable transformation of \( \{\epsilon_t\} \), which is an i.i.d. sequence and, thus, strictly stationary and ergodic.
p-th Order Autoregressive Process (or AR(p) Process)

- **Moments:** Exploiting the strictly stationary property, we can obtain the following expressions for the mean, variance, and autocovariances of an AR(p) process.

1. **Mean:**

   \[
   E[y_t] = E[\mu] + E[\alpha_1 y_{t-1}] + \cdots + E[\alpha_p y_{t-p}] + E[\varepsilon_t]
   \]

   \[
   = \mu + \alpha_1 E[y_{t-1}] + \cdots + \alpha_p E[y_{t-p}]
   \]

   (since \( E[\varepsilon_t] = 0 \) by assumption)

   This implies that

   \[
   \mu = E[y_t] - \alpha_1 E[y_{t-1}] - \cdots - \alpha_p E[y_{t-p}]
   \]

   \[
   = (1 - \alpha_1 - \cdots - \alpha_p) E[y_t]
   \]

   (since by stationarity \( E[y_t] = E[y_{t-h}] \) for all \( h \in \mathbb{Z} \))
Mean: From this, we deduce that

\[
\begin{align*}
E[y_t] & = \frac{\mu}{1 - \alpha_1 - \cdots - \alpha_p} \\
& = \mu \left( \prod_{i=1}^{p} \frac{1}{1 - (1/\rho_i)} \right)
\end{align*}
\]

which is well-defined, since by assumption \( z = 1 \) is not a solution of the polynomial equation \( \alpha(z) = 1 - \alpha_1 z - \cdots - \alpha_p z^p = 0 \).
2. Variance:

\[
\text{Var} \left( y_t \right) = E \left[ \left( y_t - E \left[ y_t \right] \right)^2 \right]
\]

\[
= E \left[ \left( y_t - \frac{\mu}{1 - \alpha_1 \cdots - \alpha_p} \right)^2 \right]
\]

\[
= E \left[ \left( \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right)^2 \right]
\]

\[
= \sum_{j=0}^{\infty} \psi_j^2 E \left[ \varepsilon_{t-j}^2 \right]
\]

\[
= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2 < \infty
\]

where the last equality follows from the fact that absolute summability implies square summability and that by assumption \( \sigma_\varepsilon^2 < \infty \).
2. **Autocovariances:** Let \( h > 0 \)

\[
\gamma(h) = \text{Cov}(y_t, y_{t-h})
\]
\[
= E \left[ (y_t - E[y_t]) (y_{t-h} - E[y_{t-h}]) \right]
\]
\[
= E \left[ \left( y_t - \frac{\mu}{1 - \alpha_1 - \cdots - \alpha_p} \right) \left( y_{t-h} - \frac{\mu}{1 - \alpha_1 - \cdots - \alpha_p} \right) \right]
\]
\[
= E \left[ \left( \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) \left( \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-h-j} \right) \right]
\]
\[
= E \left[ \left( \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) \left( \sum_{k=h}^{\infty} \psi_{k-h} \varepsilon_{t-k} \right) \right] \quad \text{(by setting } k = j + h)\]
\[
= \sum_{j=h}^{\infty} \psi_j \psi_{j-h} E \left[ \varepsilon_{t-j}^2 \right] \quad \text{(given the assumption of independence)}
\]
2. **Autocovariances (con’t):** The assumption that \( \{\varepsilon_t\} \) is identically distributed then implies that

\[
\gamma(h) = \sigma^2_{\varepsilon} \sum_{j=h}^{\infty} \psi_j \psi_{j-h} \quad \text{(i.e., } E[\varepsilon_{t-j}^2] = \sigma^2_{\varepsilon} \text{ for all } j \in \mathbb{Z}_+ \cup \{0\})
\]

Moreover, since covariance stationarity implies that \( \gamma(h) = \gamma(-h) \), we also have for \( h < 0 \)

\[
\gamma(h) = \gamma(-h) = \gamma(|h|) = \sigma^2_{\varepsilon} \sum_{j=|h|}^{\infty} \psi_j \psi_{j-|h|}
\]
2. **Autocovariances (con’t):** It follows from these calculations that for all $h \in \mathbb{Z}$,

$$
\gamma(h) = \text{Cov}(y_t, y_{t-h}) = \sigma^2 \sum_{j=\lfloor |h| \rfloor}^{\infty} \psi_j \psi_{j-|h|}.
$$