First we introduce Cauthy criteria for real numbers and random variables.

Proposition 1 (Cauthy criterion for real numbers). *Let* $a_n = \sum_{j=0}^{n} a_j$ *be a sequence of real numbers. The Cauthy criterion states that* $\sum_{j=0}^{\infty} a_j$ *converges (i.e,,* $a_n \to a$ *as* $n \to \infty$*) if and only if for any* $M > N$*,*

$$\left| \sum_{j=0}^{M} a_j - \sum_{j=0}^{N} a_j \right| \longrightarrow 0 \ \ \textit{as} \ \ M, N \longrightarrow \infty$$

Proposition 2 (Cauthy criterion for random variables). *Let* $\{X_n\}$ *be a sequence of random variables. The Cauthy criterion states that* X_n *converges in mean square (i.e.,* $X_n \overset{m.s.}{\longrightarrow} X$*) if and only if for any* $M > N$*,*

$$E[(X_M - X_N)^2] \longrightarrow 0, \ \ \textit{as} \ \ M, N \longrightarrow \infty$$

We then establish the following propositions.

Proposition 3. *Absolute summability implies square summability. That is,*

$$\sum_{j=0}^{\infty} |\psi_j| < \infty \ \ \textit{implies} \ \ \sum_{j=0}^{\infty} \psi_j^2 < \infty$$

Proof. Suppose that $\{\psi_j\}_{j=0}^{\infty}$ is absolutely summable: $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Then there exist an $N < \infty$ such that

$$|\psi_j| < 1, \ \ \forall\, j \geq N,$$

which implies that

$$\psi_j^2 < |\psi_j|, \ \ \forall\, j \geq N.$$

Then

$$\begin{aligned} \sum_{j=0}^{\infty} \psi_j^2 &= \sum_{j=0}^{N-1} \psi_j^2 + \sum_{j=N}^{\infty} \psi_j^2 \\ &< \sum_{j=0}^{N-1} \psi_j^2 + \sum_{j=N}^{\infty} |\psi_j| \end{aligned}$$

Since N is finite, $\sum_{j=0}^{N-1}\psi_j^2$ is also finite. Moreover, $\sum_{j=N}^{\infty}|\psi_j|$ is finite since $\{\psi_j\}_{j=0}^{\infty}$ is absolutely summable. Hence,

$$\sum_{j=0}^{\infty}\psi_j^2 < \infty.$$

$\square$

Finally, we show the following proposition.

Proposition 4. *An MA(∞) process with square-summable MA coefficients converges (in mean square). That is, for some random variable y_t,*

$$\sum_{j=0}^{n}\psi_j e_{t-j} \xrightarrow{m.s.} y_t \text{ as } n \longrightarrow \infty.$$

Proof. According to the *Cauchy criterion*, $\sum_{j=0}^{\infty}\psi_j e_{t-j}$ converges if and only if for any integers $M > N$,

$$E\left[\left(\sum_{j=0}^{M}\psi_j e_{t-j} - \sum_{j=0}^{N}\psi_j e_{t-j}\right)^2\right] \longrightarrow 0 \quad \text{as} \quad M, N \longrightarrow \infty.$$

Note that

$$\begin{aligned}
E\left[\left(\sum_{j=0}^{M}\psi_j e_{t-j} - \sum_{j=0}^{N}\psi_j e_{t-j}\right)^2\right] &= E\left[\left(\sum_{j=N+1}^{M}\psi_j e_{t-j}\right)^2\right] \\
&= (\psi_{N+1}^2 + \psi_{N+2}^2 + \cdots + \psi_M^2)\sigma^2 \\
&= \left[\sum_{j=0}^{M}\psi_j^2 - \sum_{j=0}^{N}\psi_j^2\right]\sigma^2
\end{aligned}$$

Since $\sum_{j=0}^{\infty}\psi_j^2$ converges, then

$$\left[\sum_{j=0}^{M}\psi_j^2 - \sum_{j=0}^{N}\psi_j^2\right] = \left|\sum_{j=0}^{M}\psi_j^2 - \sum_{j=0}^{N}\psi_j^2\right| \longrightarrow 0 \quad \text{as} \quad M, N \longrightarrow \infty,$$

by Cauchy criterion. Hence,

$$E\left[\left(\sum_{j=0}^{M}\psi_j e_{t-j}-\sum_{j=0}^{N}\psi_j e_{t-j}\right)^2\right]=\left[\sum_{j=0}^{M}\psi_j^2-\sum_{j=0}^{N}\psi_j^2\right]\sigma^2\longrightarrow 0 \text{ as } M,N\longrightarrow\infty$$

which implies that

$$\sum_{j=0}^{n}\psi_j e_{t-j}\xrightarrow{m.s.} y_t \text{ as } n\longrightarrow\infty.$$

$\square$

Summary In sum, by Propositions 3 and 4, we establish that an MA(∞) process with absolutely summable MA coefficients converges.

Inverting Lag Polynomials Given a p-th degree lag polynomial

$$\beta(L) = 1 - \beta_1 L - \beta_2 L^2 - \cdots - \beta_p L^p$$

Let the inverse of $\beta(L)$ be defined as

$$\psi(L) = \beta(L)^{-1}.$$

That is,

$$\beta(L)\psi(L) = 1,$$

or

$$(1 - \beta_1 L - \beta_2 L^2 - \cdots - \beta_p L^p)(\psi_0 + \psi_1 L + \psi_2 L^2 + \cdots) = 1$$

The above equation yields

$$\begin{aligned}
&\text{constant:} & \psi_0 &= 1 \\
&L: & \psi_1 - \beta_1\psi_0 &= 0 \\
&L^2: & \psi_2 - \beta_1\psi_1 - \beta_2\psi_0 &= 0 \\
&L^3: & \psi_3 - \beta_1\psi_2 - \beta_2\psi_1 - \beta_3\psi_0 &= 0 \\
& & &\vdots \\
&L^{p-1}: & \psi_{p-1} - \beta_1\psi_{p-2} - \beta_2\psi_{p-3} - \cdots - \beta_{p-1}\psi_0 &= 0 \\
&L^p: & \psi_p - \beta_1\psi_{p-1} - \beta_2\psi_{p-2} - \cdots - \beta_{p-1}\psi_1 - \beta_p\psi_0 &= 0 \\
&L^{p+1}: & \psi_{p+1} - \beta_1\psi_p - \beta_2\psi_{p-1} - \cdots - \beta_{p-1}\psi_2 - \beta_p\psi_1 &= 0 \\
&L^{p+2}: & \psi_{p+2} - \beta_1\psi_{p+1} - \beta_2\psi_p - \cdots - \beta_{p-1}\psi_3 - \beta_p\psi_2 &= 0 \\
& & &\vdots
\end{aligned}$$

These equations can be solved successively for (ψ_0, ψ_2,...) as

$$\begin{aligned}
\psi_0 &= 1\\
\psi_1 &= \beta_1\\
\psi_2 &= \beta_2 + \beta_1^2\\
\psi_3 &= \beta_1[\beta_2 + \beta_1^2] + \beta_2\beta_1 + \beta_1\\
&\vdots
\end{aligned}$$

For $j \geq p$, ψ_j follows the p-th order homogeneous difference equation:

$$\psi_j - \beta_1\psi_{j-1} - \cdots - \beta_p\psi_{j-p} = 0. \tag{1}$$

Hence, once the first p coefficients (ψ_0,ψ_1,...,ψ_{p-1}) are obtained, we can use this p-th order homogeneous difference equation to calculate the rest of the coefficients (ψ_j, $j \geq p$) with the initial condition, (ψ_0,ψ_1,...,ψ_{p-1}).

Lemma 1. *Let $0 \leq \xi < 1$ and let n be a non-negative integer. Then there exist two real numbers, A and b, such that $\xi < b < 1$ and*

$$(j)^n\xi^j < Ab^j, \ \ \textit{for} \ \ j = 0, 1, 2, \ldots$$

Proof. Pick any b such that $\xi < b < 1$. Since $j\log\left(\frac{b}{\xi}\right)$ eventually gets larger than $n\log(j)$ as j increases, b^j eventually gets larger than $(j)^n\xi^j$ as j increases. Hence, there exists a large J, such that

$$(j)^n\xi^j < b^j, \ \ \text{for all} \ \ j \geq J. \tag{2}$$

Moreover, let

$$B_j = \frac{(j)^n\xi^j}{b^j},$$

and define

$$B = \max\{B_0, B_1, B_2, \ldots, B_{J-1}\}$$

Then, by construction, for $j = 0, 1, 2, \ldots, (J-1)$,

$$B \geq B_j = \frac{(j)^n \xi^j}{b^j},$$

or

$$(j)^n \xi^j \leq B b^j. \tag{3}$$

Now choose A such that $A > 1$ and $A > B$. Then by equation (2),

$$(j)^n \xi^j < b^j < A b^j, \quad \text{for} \quad j \geq J.$$

By equation (3),

$$(j)^n \xi^j \leq B b^j < A b^j, \quad \text{for} \quad j = 0, 1, 2, \ldots, (J-1).$$

That is,

$$(j)^n \xi^j < A b^j$$

for $j = 0, 1, 2, \ldots$. □

Theorem 1 (Absolutely summable inverses of lag polynomials)**.** *Consider a p-th degree lag polynimial*

$$\beta(L) = 1 - \beta_1 L - \beta_2 L^2 - \cdots - \beta_p L^p$$

and let

$$\psi(L) = \beta(L)^{-1}.$$

If all the roots of $\beta(z) = 0$ lie outside the unit circle, then the coefficient sequence $\{\psi_j\}$ of $\psi(L)$ is absolutely summable.

Proof. Recall from equation (1) that for $j \geq p$,

$$\psi_j - \beta_1 \psi_{j-1} - \cdots - \beta_p \psi_{j-p} = 0.$$

with the initial condition that $(\psi_0,\psi_1,\ldots,\psi_{p-1})$ are given. The associated polynomial equation is

$$\beta(z) = 1 - \beta_1 z - \beta_2 z^2 - \cdots - \beta_{p-1} z^{p-1} - \beta_p z^p = 0. \tag{4}$$

Let λ_k $(k = 1, 2, \ldots, K)$ be the distinct roots of equation (4), and r_k be the multiplicity of λ_k. Clearly,

$$\sum_{k=1}^{K} r_k = r_1 + r_2 + \cdots + r_K = p.$$

The polynomial equation can be rewritten as

$$1 - \beta_1 z - \beta_2 z^2 - \cdots - \beta_{p-1} z^{p-1} - \beta_p z^p = \left(1 - \frac{z}{\lambda_1}\right)^{r_1} \left(1 - \frac{z}{\lambda_2}\right)^{r_2} \cdots \left(1 - \frac{z}{\lambda_K}\right)^{r_K} = 0$$

Or

$$\prod_{k=1}^{K} (1 - \lambda_k^{-1} z)^{r_k} = 0.$$

It is well-known that the solution of the p-th order homogeneous difference equation is (check any textbook on mathematical economics)

$$\psi_j = \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} c_{kn} \cdot (j)^n \cdot \lambda_k^{-j},$$

where c_{kn} are constants that can be determined by the initial condition, $(\psi_0,\psi_1,\ldots,\psi_{p-1})$.

Hence,

$$\begin{aligned}
|\psi_j| &= \left| \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} c_{kn} \cdot (j)^n \cdot \lambda_k^{-j} \right| \\
&\leq \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} |c_{kn}| \cdot (j)^n \cdot |\lambda_k^{-1}|^j \\
&\leq c \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} (j)^n \cdot |\lambda_k^{-1}|^j,
\end{aligned}$$

where $c = \max\{|c_{kn}|\}$.

Since $|\lambda_k^{-1}| < 1$ by the condition that all roots of $\beta(z) = 0$ lie outside the unit circle, we can apply Lemma 1 with $\xi = |\lambda_k^{-1}|$ and claim that for some $A_k > 0$ and $|\lambda_k^{-1}| < b_k < 1$,

$$(j)^n |\lambda_k^{-1}|^j < A_k (b_k)^j, \quad \text{for all} \quad j, k$$

Set $\tilde{A} = \max\{A_k\}$ and $b = \max\{b_k\}$, so that

$$A_k (b_k)^j \leq \tilde{A} b^j, \quad \text{for all} \quad k.$$

That is,

$$|\psi_j| \leq c \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} (j)^n \cdot |\lambda_k^{-1}|^j < c \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} A_k (b_k)^j \leq c \sum_{k=1}^{K} \sum_{n=0}^{r_k - 1} \tilde{A} b^j = cp\tilde{A} b^j.$$

Define $A = cp\tilde{A}$, we then obtain

$$|\psi_j| < A b^j, \quad \text{for all} \quad j.$$

Hence,

$$\sum_{j=0}^{\infty} |\psi_j| < \sum_{j=0}^{\infty} A b^j = \frac{A}{1-b} < \infty,$$

which suggests that $\{\psi_j\}$ is absolutely summable.

$\square$