A SIMPLE PROOF FOR THE INVERTIBILITY OF THE LAG POLYNOMIAL OPERATOR

By

Ioannis Kasparis

September 25, 2016

RESEARCH INSTITUTE FOR ECONOMETRICS

DISCUSSION PAPER NO. 3-16

Research Institute for Econometrics
מכון מחקר לאקונומטריקה

DEPARTMENT OF ECONOMICS
BAR-ILAN UNIVERSITY
RAMAT-GAN 5290002, ISRAEL

http://econ.biu.ac.il/en/node/2473

A SIMPLE PROOF FOR THE INVERTIBILITY OF THE LAG POLYNOMIAL OPERATOR

Ioannis Kasparis
University of Cyprus

April 2016

We provide a proof for the invertibility of the finite lag polynomial operator in the context of stochastic difference equations, for the case where the polynomial roots lie inside/outside the complex unit circle. We establish invertibility and provide a characterisation for the inverse, using an elementary result from functional analysis.

1 Motivation and Results

Time series models like ARMA processes are widely used in econometrics and statistics. These type of models are defined through Finite Lag Polynomial (FLP) operators. For instance, an AR(p) process $\{X_t\}_{t\in\mathbb{Z}}$, is:

$$\begin{aligned} &\phi(L)\{X_t\} = \{\epsilon_t\}, \quad \epsilon_t \sim \text{w.n.}(0,\sigma^2), \\ &\phi(L) = I - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p, \end{aligned} \tag{1}$$

where w.n.$(0,\sigma^2)$ denotes white noise sequence with mean zero and variance σ^2, while I and L are the identity and the lag operators respectively. The fact that the FLP operator $\phi(L)$ is invertible, when the polynomial roots lie inside/outside the unit root circle, is often stated in time series courses and in time series textbooks. If the polynomial $\phi(z)$, satisfies

$$\phi(z) \neq 0, \ \ |z| = 1,$$

then

$$\phi^{-1}(L) = \sum_{k=-\infty}^{\infty} \psi_k L^k, \ \ \sum_{k=-\infty}^{\infty} |\psi_k| < \infty,$$

and

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k \epsilon_{t-k}. \tag{2}$$

This is an important result for two reasons. First it provides a sufficient condition for stationarity, as it implies that the AR(p) process, X_t, is a linear process and therefore covariance stationary[1]. Secondly, it establishes that X_t is amenable to the well developed asymptotic theory for linear processes e.g. Phillips and Solo (1992), Peligrad and Utev (1997).

Our motivation is a pedagogical one. Although some version of the aforementioned result is stated in almost any econometric textbook, a rigorous proof is rarely provided. Deistler (1975) provides a proof by establishing an isomorphism between rings. The key idea in Deistler (1975) is to transform an operator problem into an equivalent algebraic problem. Although the Deistler's approach is straightforward, rings are not used often in econometrics. We avoid the use of rings and prove the result directly using operator theory. Operator theory has been employed in several recent econometric papers see for example Darolles, Florens, and Renault (2002), Linton and Mammen (2005), Carrasco, Florens, and Renault (2006) and Vanhems (2006) *intrer alia*.

In his recent book, Bosq (2000, Theorems 3.1 and 5.1) provides a proof, for the case where the polynomial roots lie outside the unit circle. Bosq relies on operator theory as well. In particular, he is utilising results for the spectral radius. Our approach is based on more elementary concepts, that can be found in any introductory functional analysis book e.g. Rynne and Youngson (2000). The technical level of our proof is about the same as that of the classical work of Brockwell and Davis (1991). Therefore, our results should be accessible to the reader of the aforementioned book.

To prove the invertibilty of the FLP operator, we employ a well known theorem in operator theory. The theorem states that if an operator is sufficiently close to the identity operator, with respect the operator norm, then is invertible. The operator norm, provides a notion of distance between two operators. The norm of a linear operator T, on some normed space, $\mathbf{V}$ say, is defined as $\|T\| = \sup_{\|x\|\leq 1} \|Tx\|$, with x in $\mathbf{V}$. Moreover the operator is bounded, if $\|T\| < \infty$. A formal statement of the aforementioned theorem is given below:

Theorem 1. *Let $\mathbf{B}$ be a Banach space. If $T : \mathbf{B} \to \mathbf{B}$ is a bounded linear operator and $\|I - T\| < 1$, then T is invertible with inverse:*

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

Note that Theorem 1 does not only provide a sufficient condition for the invertibility of the operator, but also provides a characterisation for the inverse, when the condition is satisfied. In particular Theorem 1 postulates that the inverse can be approximated by Neumann series.

In order to exploit Theorem 1, we need to define our time series process on some appropriate Banach space. We will consider the space $\mathbf{X}$, of sequences $X = \{X_t\}_{t\in\mathbb{Z}}$ on some probability space $(\Omega,\ \mathcal{F},\ P)$, that satisfy $\sup_t \mathbf{E}\,|X_t| < \infty$. Hence $\mathbf{X}$ is a normed space, equipped with the norm $\|X\|_\infty = \sup_t \mathbf{E}\,|X_t|$. Any covariance stationary sequence belongs to the space $\mathbf{X}$. The following lemma ensures, that the particular space is a Banach space[2].

Lemma 1. *The normed space $\mathbf{X}$ is complete and therefore is a Banach space.*

Next, we shall determine the lag operator norm. In view of the fact that $\sup_t \mathbf{E}\,|X_t| = \sup_t \mathbf{E}\,|X_{t-1}|$ we have

$$\|L\| = \sup_{\|X\|\le 1} \|LX\| = \sup_{\{X_t\}\in\mathbf{X}:\sup_t \mathbf{E}|X_t|\le 1} \left(\sup_t \mathbf{E}\,|X_{t-1}|\right) = 1$$

Using the same arguments as above it can be easily seen that $\|L^{-1}\| = 1$ as well, where L^{-1} is the inverse of L. Now it is straight forward to apply Theorem 1 to first order lag polynomials. Consider $\phi(L) = I - \phi_1 L$ with $|\phi_1| \neq 1$. For $|\phi_1| < 1$ we have

$$\|I - \phi(L)\| = \|\phi_1 L\| = |\phi_1|\,\|L\| = |\phi_1| < 1$$

Therefore, by virtue of Theorem 1

$$\phi(L)^{-1} = \sum_{k=0}^{\infty} \phi_1^k L^k. \tag{3}$$

For $|\phi_1| > 1$ notice that

$$\phi(L) = -\phi_1 L(I - \frac{1}{\phi_1}L^{-1}) =: -\phi_1 L \phi^+(L).$$

In addition,

$$\left\|I - \phi^+(L)\right\| = \left|\frac{1}{\phi_1}\right| \left\|L^{-1}\right\| = \left|\frac{1}{\phi_1}\right| < 1.$$

Hence, by Theorem 1

$$\phi(L)^{-1} = -\frac{1}{\phi_1} L^{-1} \sum_{k=0}^{\infty} \phi_1^{-k} L^{-k} = -\sum_{k=1}^{\infty} \phi_1^{-k} L^{-k} \tag{4}$$

Next, consider the higher order lag polynomial $\phi(L) = I - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$. Write

$$\phi(L) = \left(I - \frac{1}{\rho_1}L\right)\left(I - \frac{1}{\rho_2}L\right)...\left(I - \frac{1}{\rho_p}L\right),$$

where $\{\rho_i, i = 1, ..., p\}$ are the roots of the polynomial $\phi(z)$, $z \in \mathbb{C}$. The following result enables us to apply the partial results of (3) and (4) to higher order lag polynomials.

Lemma 2. *Let* $\mathbf{V}$ *be a normed space and suppose that the operators* $T_i : \mathbf{V} \to \mathbf{V}$, *with* $i = \{1, 2, ..., p\}$, *commute. Define* T as $T = T_1 T_2 ... T_p$. *Then* T *is invertible, if and only if each* T_i *is invertible.*

It is obvious from Lemma 2 that $\phi(L)^{-1}$ is determined by a product of terms

Theorem 2. *Consider the lag operator $\phi(L) = I - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$ on $\mathbf{X}_\gamma$. Suppose that the roots, $\{\rho_i, i = 1, ..., r, ..., p\}$, of the polynomial $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$, $z \in \mathbb{C}$ satisfy $|\rho_i| < 1$ for $i \leq r$ and $|\rho_i| > 1$ for $i > r$.* Then *the inverse of $\phi(L)$ exists and is of the form:*

$$\psi(L) = \sum_{k=-\infty}^{\infty} \psi_k L^k,$$

with

$$\begin{aligned}\psi_k &= (-1)^r \sum_{l=\max(r,-k)}^{\infty} \sum_{k_{p-1}=0}^{k+l} ... \sum_{k_{r+1}=0}^{k_{r+2}} \sum_{k_{r-1}=r-1}^{l-1} ... \sum_{k_1=1}^{k_2-1} \left(\frac{1}{\rho_1}\right)^{-k_1} ... \\ &\quad ... \left(\frac{1}{\rho_r}\right)^{-(l-k_{r-1})} \left(\frac{1}{\rho_{r+1}}\right)^{k_{r+1}} ... \left(\frac{1}{\rho_p}\right)^{k+l-k_{p-1}}\end{aligned}$$

and $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$.

Remark: Suppose that the processes X_t and Y_t satisfy the difference equation:

$$\phi(L) X_t = Y_t, \text{ with } Y_t \text{ in } \mathbf{X}.$$

(a) If $\sup_t \mathbf{E}|Y_t| < \infty$ and $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$ then, $\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k}$ is well defined *a.s.* (see for example Brockwell and Davis (1991)). If $\sup_t \mathbf{E}|Y_t| < \infty$ and Y_t satisfies the difference equation shown above, then $\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k}$ is well defined in $\mathbf{L}_1$ sense by virtue of Theorem 2, as $\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k}$ belongs to $\mathbf{X}$.

(b) Under the stronger requirement $\sup_t \mathbf{E}|Y_t|^2 < \infty$, $\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k}$ is well defined in $\mathbf{L}_2$ sense (cf. Fuller (1976), Brockwell and Davis (1991)). If $\sup_t \mathbf{E}|Y_t|^2 < \infty$ and Y_t satisfies the difference equation shown above, then $\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k}$ is well defined in $\mathbf{L}_2$ sense by virtue of Theorem 2[2].

(c) If the polynomial roots lie outside the unit circle, then Y_t is causal for X_t i.e.

$$X_t = \sum_{k=0}^{\infty} \psi_k Y_{t-k}.$$

2 Proofs

Proof of Theorem 1. Theorem 4.40 in Rynne and Youngson (2000). ■

Proof of Lemma 1. Denote by $\|.\|_{\mathbf{L}_1}$ the $\mathbf{L}_1$-norm and consider a Cauchy sequence $\{X^n\}_{n \in \mathbb{N}}$ in $\mathbf{X}$. By the definition of $\mathbf{X}$, X^n is a double indexed sequence, i.e. for each n, $X^n = \{X_t^n\}_{t \in \mathbb{Z}}$. By the completeness of $\mathbf{L}_1$ measurable spaces (e.g. Brockwell and Davis, 1991), $\|X_t^n - X_t\|_{\mathbf{L}_1} \overset{n \to \infty}{\to} 0$, for some X_t in $\mathbf{L}_1(\Omega, \mathcal{F}, P)$. Also note that due to the Cauchy property, for any $\epsilon > 0$ and some $N_\epsilon \in \mathbb{N}$, we have $\|X^n - X^m\|_\infty < \epsilon$, for all $n,\ m \geq N_\epsilon$. Define $X = \{X_t\}_{t\in\mathbb{Z}}$. Thus, for $n \geq N_\epsilon$ we have,

$$\|X_t^n - X_t\|_{\mathbf{L}_1} = \lim_{m\to\infty} \|X_t^n - X_t^m\|_{\mathbf{L}_1} \leq \lim\sup_{m\to\infty} \|X^n - X^m\|_\infty \leq \epsilon,$$

which implies $\|X^n - X\|_\infty \overset{n\to\infty}{\to} 0$. Moreover X is in $\mathbf{X}$ because, $\|X\|_\infty \leq \|X^n - X\|_\infty + \|X^n\|_\infty < \infty$ and the result follows. ■

Proof of Lemma 2. The proof is trivial and therefore omitted. ■

Proof of Theorem 2. Consider the operator $\phi_i(L) = I - \frac{1}{\rho_i} L$ on $\mathbf{X}$. By Theorem 1, (3) and (4) the inverse of $\phi_i(L)$ exists and is given by

$$\phi_i(L)^{-1} = \begin{cases} -\sum_{k=1}^{\infty} \left(\frac{1}{\rho_i}\right)^{-k} L^{-k} \text{ for } 1 \leq i \leq r, \\ \sum_{k=0}^{\infty} \left(\frac{1}{\rho_i}\right)^{k} L^{k} \text{ for } r < i \leq p. \end{cases}$$

Now, because $\phi_i(L)$'s commute, Lemma 2 implies that $\phi(L)$ is invertible with inverse

$$\phi(L)^{-1} = \phi_1(L)^{-1} ... \phi_p(L)^{-1}.$$

Next, we obtain an expression for the inverse in terms of the polynomial roots. Consider

$$\begin{aligned} \hat{\phi}(L)^{-1} &:= \phi_1(L)^{-1} ... \phi_r(L)^{-1} \text{ and} \\ \tilde{\phi}(L)^{-1} &:= \phi_{r+1}(L)^{-1} ... \phi_p(L)^{-1} \end{aligned}$$

It can be easily checked that

$$\hat{\phi}(L)^{-1} = (-1)^r \sum_{k=r}^{\infty} \hat{\psi}_k L^{-k}$$

with

$$\hat{\psi}_k = \sum_{k_{r-1}=r-1}^{k-1} \sum_{k_{r-2}=r-2}^{k_{r-1}-1} ... \sum_{k_1=1}^{k_2-1} \left(\frac{1}{\rho_1}\right)^{-k_1} ... \left(\frac{1}{\rho_i}\right)^{-(k_{r-1}-k_{r-2})} \left(\frac{1}{\rho_r}\right)^{-(k-k_{r-1})} \tag{5}$$

and

$$\begin{aligned} \sum_{k=r}^{\infty} \left|\hat{\psi}_k\right| &= \sum_{k=r}^{\infty} \left| \sum_{k_{r-1}=r-1}^{k-1} \sum_{k_{r-2}=r-2}^{k_{r-1}-1} ... \sum_{k_1=1}^{k_2-1} \left(\frac{1}{\rho_1}\right)^{-k_1} ... \left(\frac{1}{\rho_i}\right)^{-(k_{r-1}-k_{r-2})} \left(\frac{1}{\rho_r}\right)^{-(k-k_{r-1})} \right| \\ &\leq \sum_{k=1}^{\infty} \left|\frac{1}{\rho_1}\right|^{-k} ... \sum_{k=1}^{\infty} \left|\frac{1}{\rho_r}\right|^{-k} < \infty. \end{aligned} \tag{6}$$

Moreover,

$$\tilde{\phi}(L)^{-1} = \sum_{k=0}^{\infty} \tilde{\psi}_k L^k,$$

with

$$\tilde{\psi}_k = \sum_{k_{p-1}=0}^{k} \sum_{k_{p-2}=0}^{k_{p-1}} \dots \sum_{k_{r+1}=0}^{k_{r+2}} \left(\frac{1}{\rho_{r+1}}\right)^{k_{r+1}} \dots \left(\frac{1}{\rho_{p-1}}\right)^{k_{p-1}-k_{p-2}} \left(\frac{1}{\rho_p}\right)^{k-k_{p-1}} \tag{7}$$

and

$$\begin{aligned} \sum_{k=0}^{\infty} \left|\tilde{\psi}_k\right| &= \sum_{k=0}^{\infty} \left| \sum_{k_{p-1}=0}^{k} \sum_{k_{p-2}=0}^{k_{p-1}} \dots \sum_{k_{r+1}=0}^{k_{r+2}} \left(\frac{1}{\rho_{r+1}}\right)^{k_{r+1}} \dots \left(\frac{1}{\rho_{p-1}}\right)^{k_{p-1}-k_{p-2}} \left(\frac{1}{\rho_p}\right)^{k-k_{p-1}} \right| \\ &\leq \sum_{k=0}^{\infty} \left|\frac{1}{\rho_{r+1}}\right|^k \dots \sum_{k=0}^{\infty} \left|\frac{1}{\rho_p}\right|^k < \infty. \end{aligned} \tag{8}$$

Therefore,

$$\phi(L)^{-1} = \tilde{\phi}(L)^{-1}\hat{\phi}(L)^{-1} = \sum_{k=-\infty}^{\infty} \sum_{l=\max(r,-k)}^{\infty} \tilde{\psi}_{k+l}\hat{\psi}_l L^k.$$

In view of (5) and (7) we have,

$$\begin{aligned} \phi(L)^{-1} &= \sum_{k=-\infty}^{\infty} \psi_k L^k, \text{ with} \\ \psi_k &= \sum_{l=\max(r,-k)}^{\infty} \tilde{\psi}_{k+l}\hat{\psi}_l \\ &= \sum_{l=\max(r,-k)}^{\infty} \sum_{k_{p-1}=0}^{k+l} \dots \sum_{k_{r+1}=0}^{k_{r+2}} \sum_{k_{r-1}=r-1}^{l-1} \dots \sum_{k_1=1}^{k_2-1} \left(\frac{1}{\rho_1}\right)^{-k_1} \dots \\ &\quad \dots \left(\frac{1}{\rho_r}\right)^{-(l-k_{r-1})} \left(\frac{1}{\rho_{r+1}}\right)^{k_{r+1}} \dots \left(\frac{1}{\rho_p}\right)^{k+l-k_{p-1}} \end{aligned}$$

Finally, we show that the sequence ψ_k is summable. Notice that

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \sum_{l=\max(r,-k)}^{\infty} \left|\tilde{\psi}_{k+l}\hat{\psi}_k\right| &= \sum_{k=-r}^{\infty} \sum_{l=r}^{\infty} \left|\tilde{\psi}_{k+l}\hat{\psi}_l\right| + \sum_{k=-\infty}^{-r-1} \sum_{l=-k}^{\infty} \left|\tilde{\psi}_{k+l}\hat{\psi}_l\right| \\ &= \sum_{k=0}^{\infty} \sum_{l=r}^{k} \left|\tilde{\psi}_k\hat{\psi}_l\right| + \sum_{k=0}^{\infty} \sum_{l=k+r+1}^{\infty} \left|\tilde{\psi}_k\hat{\psi}_l\right| \\ &\leq 2\sum_{k=0}^{\infty} \left|\tilde{\psi}_k\right| \sum_{l=r}^{\infty} \left|\hat{\psi}_l\right| < \infty, \end{aligned}$$

by (6) and (8). ■

NOTES

1. Brockwell and Davis (1991), Proposition 3.1.2.

2. Actually the results we provide hold, when $\mathbf{X}$ is equipped with the norm $\|X\|_{\infty} = (\sup_t E\,|X_t|^{\nu})^{1/\nu}$, $\nu \geq 1$.

REFERENCES

Bosq, D. (2000). Linear processes in function spaces. Springer Lecture Notes in Statistics 149.

Brockwell, P.J. and R. A. Davis (1991). Time series: theory and methods. 2nd ed., Springer-Verlag.

Carrasco, M., J.P. Florens, and E. Renault (2006). Linear inverse problems in structural econometrics: estimation based on spectral decomposition and regularization, Handbook of Econometrics, J.J. Heckman and E.E. Leamer (eds.), vol. 6, North Holland.

Darroles S., M., J.P. Florens, and E. Renault (2002). Nonparametric instrumental regression. Working paper, University of Montreal.

Deistler M. (1975). z-Transform and identification of linear econometrics models with autocorrelated errors. *Metrika*, 22, 13-25.

Fuller, W.A. (1976). Introduction to statistical time series. New-York: Wiley.

Linton O., and E. Mammen (2005). Estimating Semiparametric ARCH(∞) models by kernel smoothing methods. *Econometrica*, 73, 771-836.

Peligrad M. and S. Utev (1997). Central limit theorem for stationary linear processes. *The Annals of Probability*, 25, 443-456.

Phillips P.C.B. and V. Solo (1992). Asymptotics for linear processes. *The Annals of Statistics*, 20, 971-1001.

Rynne, B.P. and M. A.Youngson (2000). Linear functional analysis. Springer-Verlag.

Vanhems A. (2006). Nonparametric study of solutions of differential equations. *Econometric Theory*, 22, 127-157.