The Korean Communications
in Statistics Vol. 12 No. 3, 2005
pp. 797-805

A Cointegration Test Based on Weighted Symmetric Estimator[1)]

Bu-Il Son[2)] and Key-Il Shin[3)]

Abstract

Multivariate unit root tests for the VAR(p) model have been commonly used in time series analysis. Several unit root tests were developed and recently Shin(2004) suggested a cointegration test based on weighted symmetric estimator. In this paper, we suggest a multivariate unit root test statistic based on the weighted symmetric estimator. Using a small simulation study, we compare the powers of the new test statistic with the statistics suggested in Shin(2004) and Fuller(1996).

Keywords : Unit Root Test, Unconditional MLE, Vector Autoregressive Model.

1. 서론

다음과 같은 k-차원의 VAR(p) 모형을 고려하자.

$$Z_t - \mu + A_1(Z_{t-1} - \mu) + \cdots + A_p(Z_{t-p} - \mu) = e_t \tag{1}$$

여기서 $Z_0 = Z_{-1} = \cdots = Z_{1-p} = \mu, \{e_t : t = 1, 2, \cdots\}$는 독립이고 평균이 "0", 분산이 Σ_{ee}인 다변량 정규분포를 따른다고 가정하자. 또한 (1)식은 다음에 정의된 특성 방정식의 가정을 만족한다고 가정하자.

$$\Phi(m) = |Im^p + A_1 m^{p-1} + \cdots + A_p| = 0 \tag{2}$$

즉 (2)식에서 g 개 근의 절대값은 "1"이고 나머지 근의 절대값은 "1"보다 작다고 가정하자.

---

1) This work was supported by Hankuk University of Foreign Studies Research Fund of 2005.
2) Graduate student, Department of Statistics, Hankuk University of Foreign Studies, Mohyun, Cheoingu, Yongin, Kyunggi, Korea, 449-791.
E-mail : sbona@naver.com
3) Corresponding Author, Professor, Department of Statistics, Hankuk University of Foreign Studies, Mohyun Cheoingu, Yongin, Kyunggi, Korea, 449-791.
E-mail : keyshin@hufs.ac.kr

이제 (1)식을 다음과 같이 표시하자.

$$Y_t = \Phi_1 Y_{t-1} + \sum_{t=2}^{p} \Phi_i \Delta Y_{t-i+1} + e_t \tag{3}$$

여기서 $Y_t = Z_t - \mu$, $\Delta Y_t = Y_t - Y_{t-1}$, $\Phi_1 =-\sum_{j=1}^{p} A_j$, $\Phi_p = A_p$ 그리고 $\Phi_{p-i} = \sum_{j=p-i}^{p} A_j$ $j = 2, \cdots, p-1$ 이다. 모형 (3)은 Fuller(1996)의 정리 2.4.2에서 알 수 있듯이 다음을 만족하는 Q 행렬이 존재한다.

$$Q^{-1}\Phi_1 Q = \Lambda \tag{4}$$

여기서 $\Lambda = block\{I, \Lambda_{22}\}$이고 Λ_{22}는 정칙행렬이다. 행렬 Q^{-1}에 관한 연구는 공적분 벡터에 관한 연구와 연결되며 많은 연구가 진행되고 있다. 이에 관한 자세한 내용은 Hamilton(1994)와 Fuller(1996)을 참고하기 바란다. 따라서 본 논문에서는 정준형태의 모형인 다음의 $k-$차원 VAR(p) 모형을 고려하였다.

$$Y_t = \Phi_1 Y_{t-1} + \sum_{t=2}^{p} \Phi_i \Delta Y_{t-i+1} + e_t \tag{5}$$

여기서 $Y_t' = (Y_{1t}', Y_{2t}')$라 하고 $Y_{1,t}$는 Y_t의 첫 번째 g 차원 벡터이고 $Y_{2,t}$는 $k-g$차원 벡터라 하자. 그러면 오차의 분산을 Σ_{ee}라 할 때 왼편 상단의 행렬은 I_g이 된다. 또한 $\Phi_1 = blockI_g$, $\Phi_{1,22}$ 그리고 $\Phi_{1,22} - I_{k-g}$ 행렬은 정칙행렬이다.

이제 일반적인 관심은 Φ_1의 I_g에서 몇 개의 단위근이 존재하느냐에 있다. Phillips와 Durlauf(1986)은 VAR(1) 모형에서 $H_0 : \Phi_1 = I_k$의 검정을 위한 검정 통계량을 제안하였다. Fountis와 Dickey(1989)는 비정상 VAR(p) 모형에서 하나의 단위근이 존재하는 경우에 관하여 연구하였다. 최근 Shin(2004)은 가중대칭추정량(weighted symmetric estimator)의 벡터형태를 이용한 단위근 검정법을 제안하였으며 Fuller(1996)의 단위근 검정법과 검정력을 비교하였다. 본 논문에서는 벡터형태의 가중대칭추정량을 이용한 단위근 검정법을 제안하였다. 특히 본 논문에서 제안한 통계량은 가설 “ H_0 : 1개의 단위근이 존재한다 대 H_1 : 정상 시계열이다”을 검정하기 위한 통계량이다. 2절에서는 VAR(p) 모형에서 Shin(2004), Fuller(1996) 그리고 본 논문에서 제안한 검정 통계량을 설명하였다. 3절에서는 VAR(1) 모형을 이용하여 단변량 단위근 검정통계량과 제안한 통계량과의 관계를 설명하였으며 4절에서 각 통계량의 검정력을 비교하였다. 5절에 결론이 있다.

2. 검정 통계량

(5)에서 정의된 정준형태의 VAR(p) 모형을 고려하자. 즉

$$Y_t = \Phi_1 Y_{t-1} + \sum_{t=2}^{p} \Phi_i \Delta Y_{t-i+1} + e_t$$

을 고려하자. 여기서 $\Phi_1 = block\{I_g, \Phi_{1,22}\}$ 그리고 $\Phi_{1,22} - I_{k-g}$ 행렬은 정칙행렬이다. Fuller(1996)는 단위근 검정을 위하여 다음의 최소제곱추정량을 사용하였다.

$$\hat{\Phi} = (\sum_{t=p+1}^{n} L_{t-1}' L_{t-1})^{-1} \sum_{t=p+1}^{n} L_{t-1}' Y_t' \tag{6}$$

그리고

$$\hat{\Sigma}_{ee} = (n-p-pk)^{-1} \sum_{t=p+1}^{n} \hat{e_t} \hat{e_t}' \tag{7}$$

이다. 여기서 $L_{t-1} = (Y_{t-1}', \Delta Y_{t-1}', \cdots, \Delta Y_{t-p+1}')'$, $\Phi' = (\Phi_1, \Phi_2, \cdots, \Phi_p)'$ 이고 $\hat{e_t}' = Y_t' - L_{t-1}\hat{\Phi}'$이다. 이제 $\hat{S}_{hh} = (\hat{\Phi}_1 - I) V_{11}^{-1} (\hat{\Phi}_1 - I)'$라 하고 V_{11}은 다음 행렬의 역행렬에서 Y_{t-1}에 해당되는 부분을 의미한다고 하자.

$$\sum_{t=p+1}^{n} (Y_{t-1}', \Delta Y_{t-1}', \cdots, \Delta Y_{t-p+1}')' (Y_{t-1}', \Delta Y_{t-1}', \cdots, \Delta Y_{t-p+1}')$$

그리고 $\hat{\lambda_i}$, $i = 1, 2, \cdots, k$를 다음 방정식의 근이라 하자.

$$|\hat{S}_{hh} - \lambda \hat{\Sigma}_{ee}| = 0 \tag{8}$$

Fuller(1996)는 $\lambda_i^* = (1 + d_f^{-1} \hat{\lambda_i})^{-1} \hat{\lambda_i}$을 검정통계량으로 제안하였다. 여기서 d_f는 자유도를 의미한다. 최근 Shin(2004)은 (6)식을 변형한 다음의 식을 이용할 것을 제안하였다.

$$\hat{\Phi}_{adj} = ((n+p)n^{-1} \sum_{t=2p+1}^{n} L_{t-1}' L_{t-1})^{-1} \sum_{t=p+1}^{n} L_{t-1}' Y_t' \tag{9}$$

그리고

$$\hat{\Sigma}_{ee,adj} = (n-p-pk)^{-1} \sum_{t=p+1}^{n} \hat{e}_{t,adj} \hat{e}_{t,adj}' \tag{10}$$

여기서 L_{t-1}, Φ'은 최소제곱추정량에서와 같은 식으로 정의되며 $\hat{e}_{t,adj}' = Y_t' - L_{t-1}\hat{\Phi}_{adj}'$이다. 이제 $\hat{S}_{hh,adj} = (\hat{\Phi}_{1,adj} - I) V_{11,adj}^{-1} (\hat{\Phi}_{1,adj} - I)'$라 하고 $V_{11,adj}$은 최소제곱추정량에서와 같이 다음 행렬의 역행렬에서 Y_{t-1}에 해당되는 부분을 의미한다고 하자.

$$(n+p)n^{-1} \sum_{t=p+1}^{n} (Y_{t-1}', \Delta Y_{t-1}', \cdots, \Delta Y_{t-p+1}')' (Y_{t-1}', \Delta Y_{t-1}', \cdots, \Delta Y_{t-p+1}')$$

이제 최소제곱추정량에서 사용한 방법을 그대로 적용하여 얻은 통계량을 $W_{adj,i}$, $i = 1, \cdots, k$라 할 때 Shin(2004)은 (9)를 기본으로 한 추정량이 Fuller(1996)보다 검정력이 우수함을 보였다. 본 논문에서는 특히 가설 " H_0 : 적어도 1개의 단위근이 존재 한다 대 H_1 : 정상 시계열이다"을 검정하는 통계량에 관심이 있다. 이 경우는 서론에서도 설명하였듯이 Fountis 와 Dickey(1989)에서처럼 추정된 회귀계수를 사용할 수 있을 것이다. 따라서 본 논문에서는 위의 가설을 검정하기위한 검정통계량으로 회귀계수의 추정량 $\hat{\Phi}_{1,adj}$ 를 제안하였다.

3. VAR(1) 모형인 경우

이 절에서는 VAR(1) 모형을 이용하여 본 논문에서 제안한 통계량의 특성을 간단히 살펴보도록 하자. 먼저 정준형태의 VAR(1) 모형은 다음과 같다.

$$Y_t = \Phi_1 Y_{t-1} + e_t \tag{11}$$

여기서 $\Phi_1 = block\{I_g, \Phi_{1,22}\}$이고 $\Phi_{1,22} - I_{k-g}$ 행렬은 정칙행렬이다. 단위근 검정을 위한 최소제곱추정량을 $\hat{\Phi}_{1,ols}$라 하면 최소제곱추정량은 다음과 같다.

$$\hat{\Phi}_{1,ols} = (\sum_{t=2}^{n} Y_{t-1} Y_{t-1}')^{-1} \sum_{t=2}^{n} Y_{t-1} Y_t' \tag{12}$$

이제 단변량 AR(1) 모형에서 얻어진 최소제곱추정량을 $\hat{\phi}_{1,ols}$라 하자. 그러면 $\hat{\phi}_{1,ols}$는 다음과 같이 얻어진다.

$$\hat{\phi}_{1,ols} = [\sum_{t=2}^{n} y_t y_{t-1}] / [\sum_{t=2}^{n-1} y_t^2 + y_1^2]$$

다음으로 Fuller(1996)에 자세히 설명되어 있는 단변량 가중대칭추정량(weighted symmetric estimator: WSE)은 위의 식에서 y_1^2 대신에 $n^{-1}\sum_{t=1}^{n} y_t^2$이 사용된 것을 알 수 있다. 즉 가중대칭추정량을 $\hat{\phi}_{1,wse}$라 하면 $\hat{\phi}_{1,wse}$은 다음과 같이 정의된다.

$$\hat{\phi}_{1,wse} = [\sum_{t=2}^{n} y_t y_{t-1}]/[\sum_{t=2}^{n-1} y_t^2 + n^{-1}\sum_{t=1}^{n} y_t^2]$$

Pantula 등(1994)은 비조건부 MLE(unconditional MLE)에 기초한 검정통계량과 가중대칭추정량에 기초한 검정통계량의 검정력이 매우 유사함을 보였으며 Shin(2002)과 Shin(2004)은 근사적으로 두 추정량이 같은 검정력을 갖는 이유를 설명하였다. 물론 Pantula 등(1994)의 검정력 비교 결과를 살펴보면 가중대칭추정량과 비조건부 MLE가 매우 우수한 검정력을 보여주고 있음을 알 수 있다. 따라서 본 논문에서처럼 OLS 추정량을 기초로 하는 대신 WSE에 기초한 통계량을 고려할 수 있을 것이다. 2절에서 설명하였듯이 Shin(2004)이 제안한 통계량은 Fuller(1996)에서와 같이 고유값을 이용하여 단위근을 검정하였다. 이에 반하여 본 논문에서 제안한 통계량은 회귀계수 추정량에 기초를 둔 다음의 통계량을 이용하였다. 즉

$$\Phi_{1,wse} = (\frac{n+1}{n}\sum_{t=3}^{n} Y_t Y_{t-1}')^{-1}\sum_{t=2}^{n} Y_{t-1} Y_t' \tag{13}$$

을 구하여 이중에서 $h_{11} = [\Phi_{1,wse}]_{11}$을 사용하여 검정하는 것이다. 물론 이 검정통계량은 가설 "H_0 : 1개의 단위근이 존재 한다 대 H_1 : 정상 시계열이다"을 검정하기 위한 검정통계량이다. 다음절에서는 VAR(1) 모형에서 Fuller(1996)가 제안한 λ^*, Shin(2004)가 제안한 W_{adj} 그리고 본 논문에서 제안한 h_{11}의 검정력을 비교하였다.

4. 모의실험

이절에는 본 논문에서 제안한 검정통계량과 Shin(2004), Fuller(1996)에서 제안한 통계량의 검정력을 비교하였다. 먼저 검정통계량의 5% 임계값을 구하기 위하여 각 통계량의 분포를 구하였다. 여기서 λ^*는 Fuller(1996)과 거의 비슷한 값을 보이고 있으며 W_{adj}는 Shin(2004)와 같은 값이다. 임계값을 위한 표는 <표 1>에 나와 있으며 경험적 검정력은 <표 2>와 <표 3>에 나와 있다.

4.1 임계값

평균이 알려진 자료를 분석하는 경우는 거의 없기 때문에 평균이 "0"인 모형은 고려하지 않았다. 또한 Shin(2004)에서처럼 다음의 2차원 AR(1)모형만을 고려하였다.

$$Y_t - \mu = \Phi_1 (Y_{t-1} - \mu) + e_t$$

여기서 $Y_0 = e_0$, $\mu = 0$ 그리고 $\Phi_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$을 사용하였다. 또한 오차항 $e_t : t = 1, 2, \cdots$ 는 정규분포, 평균 "0", 분산 I_2 를 이용하여 생성하였다. FORTRAN의 RANNOR 함수가 사용되었으며 자료 수 $n = 50, 100, 250$에서 50,000번의 반복을 이용하여 분포를 얻었다.

4.2 경험적 검정력

검정력 비교를 위하여 다음의 모형이 사용되었다.

$$Y_t - \mu = \Phi_1 (Y_{t-1} - \mu) + e_t$$

여기서 $Y_0 = e_0$, $\mu = 0$이고, 오차항 $e_t : t = 1, 2, \cdots$ 은 $E(e_t) = 0$, $Var(e_t) = \Sigma$이다. 특히 검정력 비교를 위하여 $\Sigma = \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}$, $\Phi_1 = \begin{pmatrix} \alpha_1 & 0 \\ \gamma & \alpha_2 \end{pmatrix}$에서 $(\alpha_1, \alpha_2) = 0.8, 0.85, 0.9$ 그리고 0.95를 사용하고 $\beta = 0, 1$, 그리고 $\gamma = 0, 0.2$가 사용되었다. 여기서 사용된 숫자는 Shin(2004)에서 사용된 숫자로 λ^*, W_{adj} 와의 검정력을 비교하기 위하여 같은 값을 사용하였다. 또한 반복수는 $n = 100$ 만을 사용하여 비교하였다. <표 2>와 <표 3>에 결과가 나와 있으며 다음의 결론을 얻을 수 있다. 먼저 <표 2>를 살펴보면 본 논문에서 제안한 검정통계량, h_{11}이 가장 우수한 것을 알 수 있다. 특히 λ^*와 비교할 경우 검정력이 매우 우수한 것을 확인 할 수 있다. 특히 $\beta = 1$에 비하여 $\beta = 0$인 경우 h_{11} 이 더 좋아지는 것을 알 수 있다. 다음으로 <표 3>을 살펴보면 전체적으로 h_{11} 의 검정력이 우수한 것을 확인 할 수 있다. 그러나 α_2이 0.9 이상인 경우에는 다른 검정 통계량에 비하여 나쁜 검정력을 보이고 있다. 특히 $\alpha_1 = 0.95, \alpha_2 = 0.95$인 경우 0.05보다 작은 검정력을 보이고 있다. 이러한 현상은 <표 2>에서 알 수 있듯이, Fuller(1996)에서 제안된 λ^*가 갖고 있는 특징이기도 한다. 이런 점에서 W_{adj} 를 이용한 검정법이 상대적인 장점을 갖고 있다고 말 할 수도 있을 것이다. 그러나 h_{11}이 갖고 있는 문제점을 보완한다면 좋은 검정력을 갖는 통계량을 만들 수 있으리라 생각된다.

<표 1> 하나의 단위근이 있을 경우의 W_{adj}, λ^*와 h_{11}의 Percentiles

Length of data		Percentiles							
		0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
$n=50$	λ^*	0.00	0.02	0.07	0.25	5.83	7.11	8.34	9.86
	W_{adj}	0.00	0.00	0.02	0.07	4.61	5.83	6.98	8.48
	h_{11}	-17.55	-14.25	-11.73	-9.08	1.04	1.71	2.28	2.93
$n=100$	λ^*	0.00	0.02	0.07	0.25	6.24	7.68	9.08	10.82
	W_{adj}	0.00	0.00	0.02	0.07	4.80	6.15	7.47	9.02
	h_{11}	-18.02	-14.60	-11.89	-9.13	1.02	1.65	2.18	2.80
$n=250$	λ^*	0.00	0.02	0.07	0.26	6.51	8.08	9.65	11.65
	W_{adj}	0.00	0.01	0.02	0.07	4.95	6.37	7.77	9.61
	h_{11}	-18.22	-14.73	-11.96	-9.18	1.02	1.63	2.12	2.73

<표 2> H_0 : 한개의 단위근 존재 H_1 : 정상 시계열 ($\gamma=0$)

			α_1							
			$\beta=0$				$\beta=1$			
			0.8	0.85	0.9	0.95	0.8	0.85	0.9	0.95
α_2	0.8	λ^*	70.69	49.93	26.65	9.65	70.76	49.02	25.58	1.91
		W_{adj}	93.80	82.62	55.75	24.06	93.68	80.76	53.22	24.29
		h_{11}	94.73	84.17	58.15	27.18	94.70	82.85	58.65	32.13
	0.85	λ^*		35.35	18.06	7.15		34.84	18.29	8.95
		W_{adj}		71.75	48.20	21.13		71.83	47.34	22.46
		h_{11}		74.42	51.70	24.41		74.49	54.42	25.62
	0.9	λ^*			9.43	3.70			9.60	4.84
		W_{adj}			32.43	14.72			33.13	16.56
		h_{11}			35.58	16.76			36.17	16.96
	0.95	λ^*				1.67				1.61
		W_{adj}				7.14				6.98
		h_{11}				7.75				7.65

<표 3> H_0 : 한개의 단위근 존재 H_1 : 정상 시계열 ($\gamma = 0.2$)

			α_1							
			$\beta = 0$				$\beta = 1$			
			0.8	0.85	0.9	0.95	0.8	0.85	0.9	0.95
α_2	0.8	λ^*	55.06	40.35	24.09	11.26	55.65	43.62	25.73	10.07
		W_{adj}	82.26	69.02	47.95	23.99	82.32	73.02	53.12	24.27
		h_{11}	86.79	76.54	57.99	32.29	86.72	78.19	58.84	28.18
	0.85	λ^*		30.39	19.67	10.45		30.39	20.54	9.19
		W_{adj}		57.88	41.75	22.35		57.88	43.78	22.50
		h_{11}		66.18	46.42	22.60		66.18	48.93	25.87
	0.9	λ^*			14.99	9.53			15.36	8.98
		W_{adj}			32.54	19.50			32.91	19.70
		h_{11}			27.72	10.36			28.21	12.80
	0.95	λ^*				8.26				8.31
		W_{adj}				14.62				14.60
		h_{11}				2.93				2.97

5. 결 론

단위근 검정, 공적분 검정 그리고 오차수정모형등은 최근의 시계열 분석에서 보편적으로 사용되고 있다. 그러나 아직도 단위근 검정의 검정력은 매우 낮아 검정 결과에 의문을 낳는 경우도 있다. 이에 본 논문에서는 검정력을 높일 수 있는 검정 통계량을 제안하였다. 이 검정 통계량이 가설 "H_0 : 1개의 단위근이 존재 한다 대 H_1 : 정상 시계열이다"를 검정하기 위한 통계량으로 만들어졌지만 전체적으로 매우 우수한 검정력을 보여주고 있다. 그러나 제안된 검정 통계량은 기존의 검정 통계량에 비해 전 모수공간에서 검정력이 우수한 것은 아니며 또한 λ^*가 가지고 있는 문제점 또한 갖고 있다. 그러나 제안된 통계량은 검정력이 높은 검정 통계량을 만들 수 있는 가능성을 보여 주고 있다. 즉 단변량에서처럼 분산을 고려한 통계량을 이용한다면 검정력은 높아질 수 있으리라 생각된다.

참고문헌

[1] Fountis, N. G., and Dickey, D. A. (1989). Testing for a unit root nonstationary in multivariate autoregressive time series, The Annals of Statistics, 17, 419-248.

[2] Fuller, W. A. (1996). Introduction to statistical time series, New York, Wiley

[3] Hamilton, J. D. (1994). Time Series Analysis, Princeton University Press.

[4] Pantula, S. G., Gonzales-Farias, G., and Fuller, W. A. (1994). A comparison of unit root criteria, Journal of Business and Economic Statistics, Vol. 13, 449-459.
[5] Phillips, P. C. B., and Durlauf, S. N. (1986), Multiple time series regression with integrated process, Review of Economic Studies, Vol. 53, 473-495.
[6] Shin, K.-I. (2002), An alternative unit root test statistic based on least squares estimator, The Korean Communications in Statistics Vol. 9, No. 3, 639-647.
[7] Shin, K.-I. (2004). A multivariate unit root test based on the modified weighted symmetric estimator for VAR(p), Journal of Applied Statistics, Vol. 31, No. 5, 587-596.

[2005년 8월 접수, 2005년 11월 채택]