2005 Proceedings of the Autumn Conference, Korea Statistical Society

웨이블렛(wavelet)을 이용한 경제시계열의 분해 및 예측

이 긍 희[1]

요 약

경제정책과 관련하여 경제시계열을 작성하는 중요한 목적중 하나는 순환변동을 파악할 수 있는 정보를 제공하는 것이다. 그런데 월별 또는 분기별로 작성되는 경제시계열은 계절변동 및 불규칙변동으로 인해 순환변동 등 기조적 변화를 잘못 파악하기 쉽다. 경제시계열의 기조적 변화를 파악하기 위해서는 원래의 경제시계열에서 계절변동, 불규칙변동을 분해 후 제거해서 분석해야 한다. 이 논문에서는 웨이블렛(wavelet)을 이용하여 시계열을 분해하고 이를 통해 경제시계열의 순환변동 등을 구하고 분해 요소들을 따로 예측한 후 결합된 예측을 시도한다.

주요용어 : 웨이블렛, 순환변동, 계절변동, 예측 실질 GDP

1. 머리말

경제시계열은 시간영역과 진동수영역에서 따로 분석되고 있다. 통상 시계열의 시간영역분석에서는 진동수영역의 변동을 모형화하기 어려우며 마찬가지로 진동수영역분석으로는 시간영역의 변동을 모형화하기 어렵다. 특히 국지적 주기변동과 같이 시간영역 및 진동수영역의 정보가 혼재되어 있는 경우 어느 영역 하나의 분석만으로 시계열의 특성을 파악하기 어렵다.

우리나라 경제시계열은 선진국에 비해 시장기능에 의해 이루어지는 내생적 조정이 미흡한 반면 빈번한 정책변화 등 외생적 변동요인이 많아 대체로 주기가 일정하지 않을 뿐만 아니라 특이항 및 구조변화 등 국지적 변동도 포함되어 있다. 이에 따라 시계열분석에서 널리 이용되고 있는 시간영역분석만으로 분석하고 예측하는 데에는 많은 어려움이 따른다. 경제변수간의 관계는 주기별로 특성이 다르게 나타나는데 통상 주기가 짧아질수록 이들 간의 관계성의 정도가 약해진다. 예를 들어, 소득과 소비의 관계는 중장기적으로는 일정하지만 분석주기가 단기화되면 그 관계가 미미해진다. 따라서 웨이블렛(wavelet) 분석과 같이 시간영역 및 진동수영역 정보를 동시에 활용하여 주기별로 시계열을 모형화할 수 있는 방법을 도입할 필요가 있다.

이 논문에서는 웨이블렛(wavelet)분석을 어떻게 경제시계열에 적용하여 경제시계열을 분해하고 예측하는 것을 정리하였다. 특히 경제시계열의 주요 정보인 순환변동, 계절변동을 찾는 방법을 정리하였으며 이를 바탕으로 미래를 예측하는 방법을 고려하였다.

본고의 구성은 다음과 같다. 제2장에서는 웨이블렛을 소개하였으며 제3장에서는 웨이블렛을 이용한 경제시계열의 분해방법과 예측방법을 소개하였다. 제4장에서는 이를 실질 GDP에 대해 적용하여 유용성을 점검하였다.

---

1) 한국방송통신대학교 정보통계학과 조교수

2. 웨이블렛의 개요

웨이블렛은 삼각함수와 같이 일정한 시간 t의 함수 $f(t)$를 근사할 수 있는 기초함수(basis)이다. 웨이블렛은 특정시간에 한정된 작은 파동을 의미하는데 대표적 웨이블렛으로는 Haar 웨이블렛, Daubechies 웨이블렛, Chui · Wang 웨이블렛 등이 있다(Chui 1992). 웨이블렛 함수는 평탄한 부분(저주파, 장기변동)을 설명하는 부웨이블렛(father wavelets, scaling function; ϕ)과 평탄하지 않은 부분(고주파, 단기변동)을 설명하는 모웨이블렛(mother wavelets; ψ)으로 구성되어 있다.

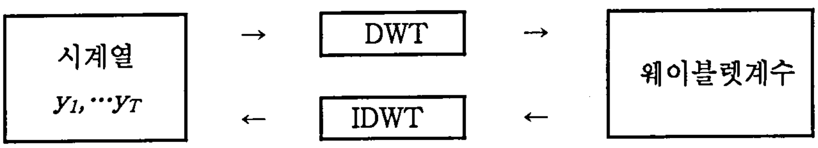
경제시계열 y_t가 함수 $f(t)$와 오차로 구성되어 있을 경우 (5)와 같은 웨이블렛급수추정량(wavelet series estimator)으로 $f(t)$를 추정할 수 있다. 이 때 $\phi(t)$는 시계열변동의 평탄한(smoothing, 추세순환변동)부문을, ψ_{jk}는 평탄하지 않은 부문(계절 · 불규칙변동)을 나타낸다.

$$\widehat{f}(t) = \sum_k \widehat{c_{j_0,k}}\,\phi_{j_0,k}(t) + \sum_j \sum_k \widehat{d_{j,k}}\,\psi_{jk}(t)$$

여기서 $\hat{c}_{j_0,k}$는 $\frac{1}{T}\sum_{i=1}^{T} y_i \phi_{j_0,k}(\frac{i}{T})$, $\hat{d}_{j,k}$는 $\frac{1}{T}\sum_{i=1}^{T} y_i \psi_{j,k}(\frac{i}{T})$이다.

웨이블렛계수는 직접 계산하기보다는 일정한 알고리즘 - 이산형웨이블렛변환(DWT, discrete wavelet transformation)으로 계산한다. 웨이블렛계수는 역이산형웨이블렛변환(IDWT, inverse discrete wavelet transformation)을 이용하여 원계열로 다시 복원할 수 있다.

<그림 1> 웨이블렛변환을 이용한 변환 및 복원



이산형웨이블렛변환은 2^n개의 자료에 적합한 변환이다. 따라서 시계열의 길이가 이와 다를 경우 여러 형태로 자료의 양끝에 대한 연장이 필요하다. 또 이산형웨이블렛변환으로 경제시계열에 대해 변환할 경우 어느 시점에서 시작하느냐에 따라서 그 결과가 달라지는 문제가 있다. 이를 개선하기 위한 변환으로는 maximal overlap DWT(MODWT)가 있는데 이는 앞서의 문제를 완화할 수 있다(Percival and Walden 2000). 또한 각 분해주기(scale)에 있는 계수의 수가 일정하여 시계열 분해시 유용하게 이용될 수 있다. MODWT는 DWT와 달리 직교성은 없으며 계산속도가 이산형웨이블렛변환보다는 느리다는 단점이 있다. 본 연구에서 다루는 분기별 경제시계열의 경우 그 자료의 수가 많지 않기 때문에 계산속도는 큰 문제가 되지 않는다.

3. 경제시계열의 분해와 예측

경제시계열을 분석할 때 장기적 추세는 물론 단기적 변동도 자세하게 살펴볼 필요가 있다. 그러나 이제까지 대부분의 시계열분석은 하나의 주기에 대한 분석에 초점을 맞추었기 때문에 여러 형태의 주기별 변동을 동시에 파악하기 어렵다는 한계를 가지고 있다.

웨이블렛계수는 시계열을 평탄한 부분(S_j)과 평탄하지 않은 부분(D_j)으로 구분할 수 있는데 주기수준 j가 커지면서 평탄한 정도가 증가한다.

$$S_{j-1}(t)=S_J(t)+D_J(t)+\cdot\cdot\cdot+D_j(t)$$

$$\text{여기에서}\quad S_J(t)=\sum_k c_{J,k}\phi_{J,k}(t),\quad D_j(t)=\sum_k d_{j,k}\psi_{j,k}(t)$$

이 때 $\{S_J, S_1, \cdot\cdot\cdot, S_j\}$를 각각 시계열의 다중주기(multiresolution) 근사, 다중주기분해로 정의된다. 시계열의 다중주기근사의 결과를 살펴보면 주기가 길어지면서 단기변동은 사라지고 추세변동이 나타나는 것을 알 수 있다.

웨이블렛 분석에서는 시계열을 기존 방법보다 세분된 주기변동으로 분해할 수 있다. Arino · Vidakovic(1995)은 시계열(Y)이 해당주기 변동(X_j)의 합(X_0 + X_2 + ⋯ + X_{n-1})으로 구성된 것으로 가정하고 3단계의 과정으로 시계열을 분해하였다. 먼저, 시계열을 이산형웨이블렛변환(DWT)을 이용하여 다음과 같이 웨이블렛 계수로 변환하였다.

$$d=(\hat{c}_{0,0}, \hat{d}_{0,0}, \hat{d}_{1,0}, \hat{d}_{1,1}, \hat{d}_{2,0}, \cdots \hat{d}_{n-1,2^{n-1}-1})$$

그리고 $d_0, d_1, \cdots, d_{n-1}$에 해당하는 주기 수준 웨이블렛계수 이외의 계수값을 "0"으로 대치한다.

$$d_n=(0, \widehat{d_{0,0}}, 0, \cdots 0)$$
$$d_{n-1}=(0,0, \widehat{d_{1,0}}, \widehat{d_{1,1}}, \cdots 0)$$
$$\vdots$$
$$d_{n-j}=(0,0,\cdots, \hat{d}_{j,0}, \cdots \hat{d}_{j,2^j-1}, \cdots, 0)$$
$$\vdots$$
$$d_1=(0,0,\cdots 0, \hat{d}_{n-1,0} \cdots \hat{d}_{n-1,2^{n-1}-1})$$

마지막으로 d_1, d_2, ⋯ d_{n-1}을 다시 역이산형웨이블렛변환(IDWT)을 이용하여 시계열로 복원한다. 이러한 과정을 통해 시계열 Y를 해당 주기요소시계열 X_0, X_1, · · ·, X_{n-1}로 분해할 수 있다.

이긍희(1998)는 웨이블렛으로 회사채유통수익률을 분해하고 분해된 주기별 자료를 AR모형으로 예측한 후 다시 결합하는 방법으로 예측모형을 설정하였다. 그 결과 분해 구성요소별로 예측하여 결합하는 것이 단순 ARIMA모형 등에 비해 우수한 것으로 나타났다. Aussem et al. (1997)은 웨이블렛 계수를 예측하고 이를 결합하는 방식으로 시계열을 예측하였다. 여기에서 f는 선형 또는 비선형 함수의 추정량이다.

$$\hat{c}_{N,t+p}=\hat{f}_0(c_{N,t}, c_{N,t-1}, \ldots, c_{N,t-r_0}),$$

$$\hat{d}_{j,t+p}=\hat{f}_j(d_{j,t}, d_{j,t-1}, \ldots, d_{j,t-r_j}),\quad j=1,2,\ldots,N$$

$$\hat{y}_{t+p}=\hat{c}_{N,k+p}+\hat{d}_{N,t+p}+\ldots+\hat{d}_{1,t+p}$$

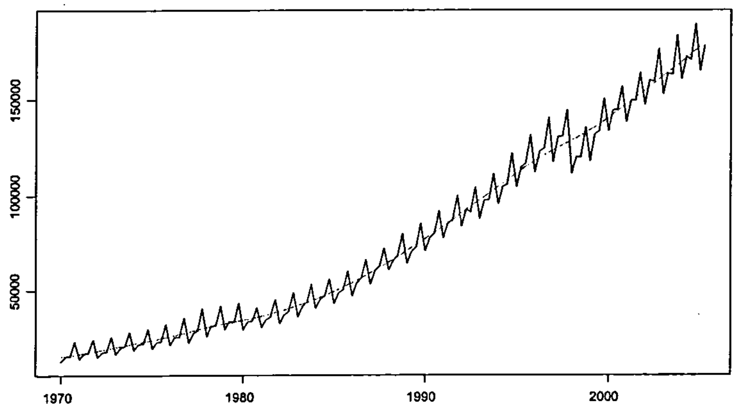
4. 실질 GDP의 분해

경제시계열(Y_t)은 다음과 같이 계절변동성분(S_t), 추세변동성분(T_t), 순환변동성분(C_t) 및 불규칙변동성분(I_t)의 합(가법형) 또는 곱(승법형)으로 구성되었다고 가정한다. 통상 계절변동에는 요일구성과 명절변동이 포함한다.

$$\text{가법형 모형 : } Y_t = T_t + C_t + S_t + I_t$$
$$\text{승법형 모형 : } Y_t = T_t \times C_t \times S_t \times I_t$$

<그림 2>는 실질 국내총생산(GDP)이다. GDP는 계절변동으로 인해 매년 1/4분기에는 작게, 4/4분기에는 크게 보이는 계절변동을 포함하고 있으며 추세성분이 강한 모습을 보이고 있다.

<그림 2> 실질 GDP 추이



웨이블렛변환인 MODWT의 시계열 양끝(boundary) 문제를 해결하기 위해서 GDP에 대해서 로그변환후 시간의 HP필터로 추세를 <그림 2>의 점선과 같이 추정한 후 이를 제거하여 추세제거계열 Y_t^B를 생성한다.

$$Y_t^B = C_t \times S_t \times I_t$$

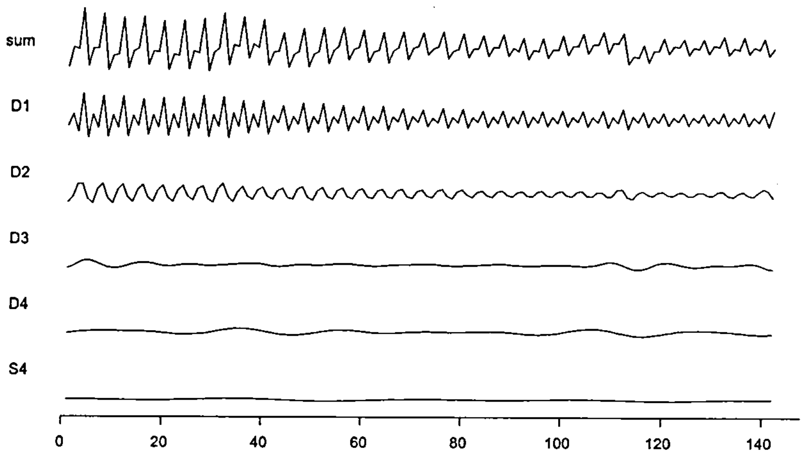
다음으로 MODWT라는 웨이블렛변환 방법으로 Y_t^B를 주기별로 분해하면 다음과 같이 정리된다.

$$\begin{aligned}\log(Y_t^B) &= \sum_k \widehat{c_{j_0,k}}\phi_{j_0,k}(t) + \sum_j\sum_k \widehat{d_{j,k}}\phi_{jk}(t) \\ &= S_{j_0}(t) + \sum_{j=1}^{j_0} D_j(t) = S_{j_0}(t) + \sum_{j=3}^{j_0} D_j(t) + D_2(t) + D_1(t)\end{aligned}$$

여기에서 Y_t^B는 순환변동(C_t)에 해당하는 $\sum_{j=3}^{j_0} S_{j_0}(t) + \sum_{j=0}^{S_0} D_j(t)$와 계절변동 및 불규칙 변동(S_t 와 I_t)이 포함된 $D_2(t)+D_1(t)$로 구분할 수 있다.

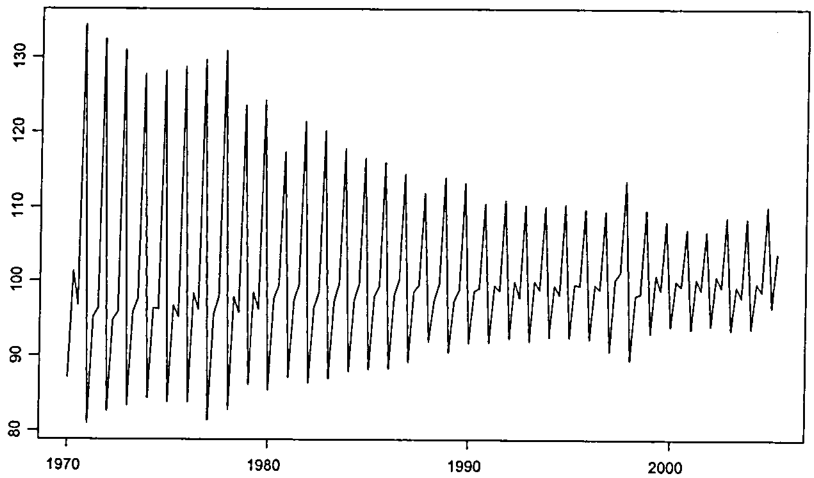
웨이블렛 변환 결과에 대한 다중주기 분석결과는 <그림 3>과 같다. 이를 보면 D1, D2에는 계절변동이 나머지에는 순환변동의 영향이 큰 것으로 볼 수 있다.

<그림 3> 추세를 제거한 실질 GDP의 다중주기분석

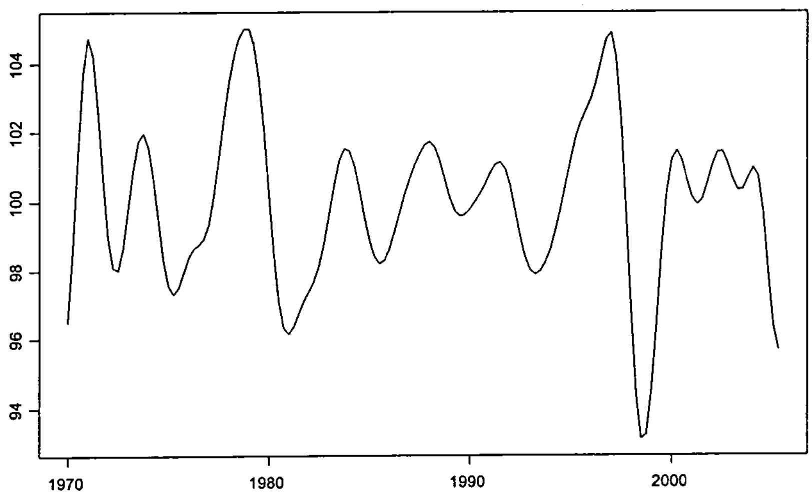


D1, D2를 합한 결과는 <그림 4>, 나머지를 합 결과는 <그림 5>와 같다. 이를 각각 계절변동 + 불규칙변동, 순환변동이라고 할 수 있다. 계절변동조정을 위해 순수 계절변동을 구하려면 계절변동 + 불규칙변동을 중심화 3분기 이동평균을 하면 된다. <그림 5>를 보면 순환변동의 결과과 기존의 우리나라 경기변동 결과와 대체로 잘 대응되는 것으로 보인다. 한편 예측을 위해서는 위의 구성요소들을 예측한 후 결합하여 최종예측치를 구할 수 있다.

<그림 4> 실질 GDP의 계절 및 불규칙변동



<그림 5> 실질 GDP의 순환변동



참고문헌

이긍희(1998), "소파동(wavelet)을 이용한 회사채유통수익률의 분해와 예측", 「경제분석」, 제4권 제3호, 한국은행 금융경제연구소.

Arino, M.A. · B. Vidakovic(1995), "On Wavelet Scalograms and their Applications in Economic Time Series," Discussion Paper 95-21, ISDS, Duke University.

Assem, A. and Murtagh(1997), "Combining neural netwaork forecasts on wavelet transformed time series, Conn. Sci. 9, pp 113-122.

Chui, C.K.(1992), *An Introduction to Wavelets.* Boston, MA: Academic Press.

Daubechies, I(1992)., *Ten Lectures on Wavelets*, Philadelphia: SIAM.

Lee, G.H.(1998). "Wavelets and Wavelet Estimation : A Review," *Journal of Economic Theory and Econometrics, 4(1).*

Lee, H.S.(2001), "Recent Advances in Wavelet Methods for Economic Time Series," Journal of Economic Theory and Econometrics", *Journal of Economic Theory and Econometrics,* Vol. 7 No. 1 43-65.

Percival D.B. and Walden A.T.(2000)., *Wavelet Methods for Time Series Analysis,* Cambridge University Press.