

디컨볼루션 기법을 이용한 BSR 신호특성 연구

Analysis of signal characteristics of the BSR using
deconvolution

한국해양연구원

한국지질자원연구원

제 출 문

한국지질자원연구원 원장 귀하

본 보고서를 “딤킨볼루션을 이뵙한 BSR 신호특성 연구”의 보고서로 제출합니다.

2006. 11.

주관연구기관명 : 한국해양연구원

주관연구책임자 : 김 한 준

연 구 원 : 주 형 태

요 약 문

I. 제목:

딤킨볼루션 기법을 이용한 BSR 신호특성 연구

II. 연구개발의 목적 및 필요성

1. 연구개발의 필요성

● 기술적 측면

- 탄성과 탐사를 이용해서 메탄 하이드레이트가 분포하는 층을 찾기 위해 가장 보편적으로 확인하는 것은 BSR (Bottom Simulating Reflector)의 존재이다. 하이드레이트 BSR은 높은 속도의 하이드레이트 안정지역과 그 아래 자유가스 (free gas)가 존재할 수 있는 낮은 속도의 퇴적층의 경계면에서 발생하는 음향 임피던스의 차이 때문에 발생한다. 따라서 가스 하이드레이트 BSR은 진폭이 크며 해저면과 비교했을 때 극성이 반전된 반사신호이다. 탄성과 탐사자료에서 BSR의 반사계수와 극성을 정확히 아는 것은 매우 중요하다. 이 목적을 위해 지금까지 스파이킹 디컨볼루션이나 예측 디컨볼루션 (predictive deconvolution)을 많이 사용하고 있지만 이 방법을 성공적으로 적용하기 위해서는 탄성과 자료의 주파수 특성이 무색이고 부분적으로 파형요소 (source wavelet)가 최소 위상이어야 한다. 하지만 대개의 경우 지층경계면을 정의하는 반사계수는 무색 스펙트럼을 가지지 못하고 파형요소를 정확히 아는 것 역시 매우 힘이 든다. Wiggins (1978)가 제시한 최소엔트로피 디컨볼루션 (MED: Minimum Entropy Deconvolution)은 새로운 방법의 디컨볼루션 기술이다. MED는 탄성과 자료와 부합하는 최소한의 개수의 큰 스파이크로 반사계수를 구하는 방법으로 앞의 방법들에서 요구하는 제약조건을 요구하지 않는 장점이 있다. 이러한 차이에도 불구하고 이들 두 가지 접근방법은 선형연산으로 구현된다. 다만, 두 방법 다

Toeplitz 행렬을 역산하지만 MED 필터는 반복적으로 Toeplitz 행렬을 역산하여 구한다. 선형연산을 채용하는 경우에는 주파수 대역이 제한된 (band-limited) 탄성과 자료를 정확히 처리하기에 곤란하다.

- 대역이 제한된 자료의 디컨볼루션을 위해서 두 가지의 방법이 있다. 첫 번째가 선형프로 그래밍 (LP: linear programming) 방법 (Levy and Fullagar, 1981)으로 탄성과 자료와 부합하는 최소 절대 norm (minimum absolute norm)을 이용하여 반사계수를 구하는 것이다. 두 번째 방법은 탄성과 자료의 스펙트럼을 복소 AR (autoregressive) 모델로 계산하는 것으로 (Lines and Clayton, 1977) 탄성과 자료의 실제 주파수대역으로 외삽하여 반사계수의 저주파 및 고주파 성분을 구한다. 하지만 LP 와 AR 방법은 파형요소 (wavelet)를 미리 알아야 하는 제약을 가지며 연산이 상대적으로 불안정하다.
- Sacchi et al. (1994)는 새로운 방법의 MED를 소개하였다. 이 방법은 MED의 장점을 살리면서 반사계수 스펙트럼에서 손실된 주파수 대역에 관한 norm을 극대화하는 비선형 연산을 수행함으로써 선형연산이 제대로 처리하지 못하는 주파수 대역이 제한된 자료를 처리할 수 있다. 이 방법은 주파수 영역에서 반사계수의 스펙트럼을 확장하기 위해 주파수 영역에서의 제한조건을 사용하므로 FMED (MED with frequency domain constraints)라 불리운다.
- 따라서 FMED를 이용해서 탄성과자료를 처리하는 것이 BSR의 탐지와 반사특성을 이해하는 데에 매우 유용할 것이다.

2. 목적

- FMED를 이용해서 탄성과자료로부터 BSR의 반사신호를 분리하고 특성을 해석
- 탄성과자료로부터 파형요소를 추출

III. 연구개발의 내용 및 범위

- FMED 방법 구현
- FMED를 이용하여 한국지질자원연구원이 획득한 탄성과자료를 처리

- BSR 반사신호의 특성 분석
- 탄성과자료로부터 과형요소를 추출하고 특성을 분석

IV. 연구개발결과

탄성과 탐사자료로부터 가스 하이드레이트의 존재를 지시할 수 있는 BSR의 특성을 분석하기 위해 FMED 방법을 구현하여 한국지질자원연구원에서 획득한 탄성과 프로파일 자료에 적용하였다.

탄성과 프로파일은 해저면에서 약 210- 230 ms의 아래에 해저면의 반사신호와 평행한 BSR이 잘 발달되어 있음을 보여 준다. 탄성과 프로파일에서 BSR은 해저면의 반사신호와 반대인 음의 극성을 가지고 있는 것으로 보이지만 여러 개의 신호가 중첩되어 뚜렷이 구분되지 않는 경우도 많다.

FMED를 이용하여 프로파일의 자료를 처리한 결과는 과형요소의 영향을 완전히 제거함으로써 해저면과 그 아래에서 반사신호들을 완전히 분리하여 각각의 반사계수의 크기와 극성을 확인할 수 있음을 보여 주었다. 특히 BSR의 반사신호는 매우 잘 분리할 수 있었다. 조사지역의 BSR은 진폭의 변화가 매우 크며 지역적으로 이중 BSR (double BSR)의 특징을 보여주기도 한다. BSR 아래에 상당한 진폭의 반사면이 존재하므로 BSR 하부에서 자유가스의 존재와 관련된 구조를 지시할 가능성이 있다. 하지만 BSR 아래의 반사신호는 그 극성이 일관되지 못하므로 이에 대해 타당한 해석이 필요하다.

탄성과 프로파일의 자료로부터 구한 과형요소는 반향신호가 매우 효과적으로 억제되어 분해능이 우수한 것으로 판단된다. 과형요소의 중심 주파수는 50 Hz로서 가스 하이드레이트 탐사에 적합하다.

이 연구에서는 제한된 탄성과 프로파일을 처리하였지만 이 연구에서 사용한 방법을 차후 시험시추를 계획하고 있는 지역에서 얻은 탄성과자료에 적용함으로써 우리나라 해역에서 가스 하이드레이트 탐사와 개발의 효율을 증대시킬 수 있을 것으로 기대한다.

V. 연구개발결과의 활용계획

1. 한국지질자원연구원에서 획득한 탄성파자료를 이용하여 가스 하이드레이트 유망 구조를 해석하고 시추를 계획하는 작업에 이 연구의 결과를 적용할 수 있다.
2. 가스 하이드레이트 안정영역 아래에 가스가 존재할 경우 발생하는 탄성과 감쇠를 구명하기 위해서는 파형요소를 아는 것이 필수적이다. 이 연구에서 구한 파형 요소는 탄성과 역산과 가스의 존재에 대한 자료 분석 등에 매우 중요하게 사용될 수 있을 것을 기대한다.
3. 탄성파 자료에서 보이는 가스 하이드레이트 부존 구조에서 BSR과 그 아래 반사면의 성질을 해석할 수 있다. 이런 과정을 통하여 유망한 구조를 찾고 시추의 성공율을 높일 수 있을 것으로 기대된다.

SUMMARY

I. Title:

**Tectonic evolution and crustal deformation
of the southwestern Korean margin**

II. Necessity and Objectives of the Study

1. Necessity

- Technological aspects

- Identification of the BSR (Bottom Simulating Reflector) is one of the most widely used criteria to evaluate the existence of methane hydrates in sediments. The gas hydrate BSR results from a sharp acoustic impedance contrast between the high-velocity gas hydrate stability zone above and low-velocity sediments below that can contain free gas. Therefore, the gas hydrate BSR in the seismic profile is high in amplitude and shows reversed polarity with respect to the seafloor reflection. These features point to the importance of the correct knowledge of reflection coefficient and polarity of the BSR in seismic data. While spiking deconvolution and predictive deconvolution are conventional approaches to this goal, their implementation requires crucial assumptions that the spectrum of reflectivity is white and the source wavelet is minimum-phase. However, reflection coefficients that define the interfaces of the subsurface show a non-white spectrum and the source wavelet is not minimum-phase in most real cases. The minimum entropy deconvolution (MED) introduced by Wiggins (1978)

offers a different approach to seismic deconvolution; MED seeks the smallest number of large spikes that are consistent with the data. MED does not require the assumptions for spiking and predictive deconvolution methods. Despite these differences, both methods constitute a linear approach to seismic deconvolution. The spiking and predictive filters are obtained by inverting the Toeplitz matrix; the MED filter is calculated in an iterative procedure in which the Toeplitz matrix is inverted at each step. Because they are linear operators, none of them can handle band-limited data properly.

- Two main procedures to deconvolve band-limited data have been developed. The first method is the linear programming (LP) approach (Levy and Fullagar, 1981). This method attempts to find the reflectivity series with the minimum absolute norm that remain constant with the data (Levy and Fullagar, 1981). The second approach (Lines and Clayton, 1977) fits a complex autoregressive (AR) model to the data spectrum and attempts to extrapolate the missing low and high frequencies. Nevertheless, these approaches require that the source wavelet be available and their computation is relatively unstable.
- Sacchi et al. (1994) introduced a new MED method. This method not only retains the advantages of MED but handles band-limited data by performing a nonlinear procedure to maximize the norm for the missing frequencies, which is not achievable using linear procedures. This method, derived from constraints to extrapolate the frequency spectrum of reflectivity, is defined as FMED (MED with frequency domain constraints).
- Therefore, FMED is expected to provide an efficient and useful tool with which to identify the BSR and investigate its reflection characteristics.

2. Objectives

- Separation of the BSR and its interpretation using FMED
- Extraction of the source wavelet

III. General Scope of the Study

- Implementation of FMED
- Processing of seismic data obtained by KIGAM to recover the sparse reflectivity series
- Analysis of the BSR signal
- Extraction of the source wavelet
- Analysis of BSR

IV. Results of the Study

The BSR provides indications of the presence of gas hydrates on seismic data. In this study, FMED was implemented and applied to seismic data obtained by KIGAM (Korea Institute of Geoscience & Mineral Resources) for the purpose of analyzing the quantitative characteristics of the BSR.

Seismic profiles document well-imaged BSRs 210 - 220 ms below the sea floor that parallel the sea floor. The BSRs, although estimated to have negatively polarity, are superposed in many cases with other signals and thus are difficult to resolve.

The result of deconvolution of seismic data using FMED demonstrates that the effect of the source wavelet was completely removed from the data. Consequently, amplitude and polarity of the reflections from the seafloor and subsurface interfaces was accurately resolved. In particular, this method was proved to be suited for resolving the BSR. The BSR in the survey area shows large variations in amplitude. Locally, occurrence of the

double BSR is recognized.

The existence of the enhanced reflection of significant amplitude immediately below the BSR possibly indicates structure associated with free gas. However, the polarity of the enhanced reflection is not consistent, pointing to further proper interpretations.

The source wavelet estimated from seismic data is estimated to provide high resolution by efficiently suppressing reverberations. The main frequency of the source wavelet is 50 Hz, adequate for detection of gas hydrates.

Although a limited amount of seismic data was processed in this study, the application of the implemented method is expected to be applicable to the seismic data from the target area and benefit the exploration and development of gas hydrates.

V. Suggestions for Applications

1. The implemented method in this study can significantly contribute to the correct interpretation of prospective gas hydrate structure and planning of drilling based on seismic data obtained by KIGAM.
2. It is imperative to know the source wavelet to investigate seismic attenuation when free gas is present below the gas hydrate stability zone. The estimated source wavelets can be very useful for seismic inversion and analysis of data to identify free gas.
3. The use of the method in this study can help properly address the characteristics of BSR and ensuing reflectors, thereby improving the discovery of prospective gas hydrate structure and the success rate of drilling.

CONTENTS

Summary

Contents

List of Figures and Tables

Chapter 1. Introduction

Chapter 2. Contents and Results of the Study

1. FMED

1-1. Seismic Model

1-2. Entropy Norm

1-3. Maximization of $V(y)$

1-4. Computation of FMED

2. Processing of seismic data

1-1. Seismic data and results of processing

1-2. Extracted source wavelets

3. Conclusions

Chapter 3. Plans for application

Chapter 4. References

Appendix 1. FMED Algorithm

Appendix 2. Statistical Extraction of the Wavelet

목 차

요약문	
목차	
그림 및 표목차	
제 1 장 서 론	
제 2 장 연구개발 수행내용 및 결과	
제 1 절 FMED	
1. 탄성과 모델	
2. 엔트로피 norm	
3. $V(y)$ 의 최대화	
4. FMED 연산	
제 2 절 탄성과 자료 처리	
1. 탄성과 자료 및 처리 결과	
2. 추출된 파형요소	
제 3 절 결론	
제 3 장 연구개발결과의 활용계획	
제 4 장 참고문헌	
부록 1. FMED 연산	
부록 2. 통계학적 방법을 이용한 파형요소의 추출	

그림 및 표 목차

- Fig. 2-1. MCS profile 05gh_1002_s003.
- Fig. 2-2. MCS profile 05gh_1036_s013.
- Fig. 2-3. MCS profile 05gh_1059_s018.
- Fig. 2-4. MCS profile 05gh_1043_s037.
- Fig. 2-5. MCS profile 05gh_1043_s048.
- Fig. 2-6. Trace CDP 461 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.
- Fig. 2-7. Trace CDP 3281 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.
- Fig. 2-8. Trace CDP 3721 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.
- Fig. 2-9. Trace CDP 13081 on Profile 05gh_1043_s037 before (upper) and after (lower) FMED.
- Fig. 2-10. Trace CDP 14301 on Profile 05gh_1043_s037 before (upper) and after (lower) FMED.
- Fig. 2-11. Trace CDP 1521 on Profile 05gh_1043_s048 before (upper) and after (lower) FMED.
- Fig. 2-12. Part of Profile 05gh_1043_s037 before (left) and after (right) deconvolution.
- Fig. 2-13. Deconvolved traces in Fig. 2-12.
- Fig. 2-14. Part of Profile 05gh_1043_s048 before (left) and after (right) deconvolution.
- Fig. 2-15. Estimated source wavelet from Trace CDP 401 on Profile 05gh_1002_s003.
- Fig. 2-16. Estimated source wavelet from Trace CDP 15101 on Profile 05gh_1043_s037.

List of Figures and Tables

Fig. 2-1. MCS profile 05gh_1002_s003.

Fig. 2-2. MCS profile 05gh_1036_s013.

Fig. 2-3. MCS profile 05gh_1059_s018.

Fig. 2-4. MCS profile 05gh_1043_s037.

Fig. 2-5. MCS profile 05gh_1043_s048.

Fig. 2-6. Trace CDP 461 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.

Fig. 2-7. Trace CDP 3281 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.

Fig. 2-8. Trace CDP 3721 on Profile 05gh_1002_s003 before (upper) and after (lower) FMED.

Fig. 2-9. Trace CDP 13081 on Profile 05gh_1043_s037 before (upper) and after (lower) FMED.

Fig. 2-10. Trace CDP 14301 on Profile 05gh_1043_s037 before (upper) and after (lower) FMED.

Fig. 2-11. Trace CDP 1521 on Profile 05gh_1043_s048 before (upper) and after (lower) FMED.

Fig. 2-12. Part of Profile 05gh_1043_s037 before (left) and after (right) deconvolution.

Fig. 2-13. Deconvolved traces in Fig. 2-12.

Fig. 2-14. Part of Profile 05gh_1043_s048 before (left) and after (right) deconvolution.

Fig. 2-15. Estimated source wavelet from Trace CDP 401 on Profile 05gh_1002_s003.

Fig. 2-16. Estimated source wavelet from Trace CDP 15101 on Profile 05gh_1043_s037.

제 1 장 서 론

디컨볼루션(deconvolution)은 탄성과 자료로부터 반사계수를 구하는 연산으로 주변 매질과 음향임피던스의 차이가 큰 석유와 가스를 찾는 데에 필수적이다. 지금까지 스파이킹 디컨볼루션이나 예측 디컨볼루션 (Robinson and Treitel, 1980) 을 많이 사용하고 있지만 이 방법을 성공적으로 적용하기 위해서는 탄성과 자료의 주파수 특성이 무색이고 부분적으로 파형요소 (source wavelet) 가 최소위상이어야 한다. 하지만 대개의 경우 지층경계면을 정의하는 반사계수는 무색 스펙트럼을 가지지 못하고 파형요소를 정확히 아는 것 역시 매우 힘이 든다. Wiggins (1978)가 제시한 최소엔트로피 디컨볼루션 (MED: Minimum Entropy Deconvolution)은 새로운 방법의 디컨볼루션 기술이다. MED는 탄성과 자료와 부합하는 최소한의 개수의 큰 스파이크로 반사계수를 구하는 방법으로 앞의 방법들에서 요구하는 제약조건을 요구하지 않는 장점이 있다. 이러한 차이에도 불구하고 이들 두 가지 접근방법은 선형연산으로 구현된다. 다만, 두 방법 다 Toeplitz 행렬을 역산하지만 MED 필터는 반복적으로 Toeplitz 행렬을 역산하여 구한다. 선형연산을 채용하는 경우에는 주파수 대역이 제한된 (band-limited) 탄성과 자료를 정확히 처리하기에 곤란하다.

대역이 제한된 자료의 디컨볼루션을 위해서 두 가지의 방법이 있다. 첫 번째가 선형프로그래밍 (LP: linear programming) 방법 (Levy and Fullagar, 1981) 으로 탄성과 자료와 부합하는 최소 절대 norm (minimum absolute norm)을 이용하여 반사계수를 구하는 것이다. 두 번째 방법은 탄성과 자료의 스펙트럼을 복소 AR (autoregressive) 모델로 계산하는 것으로 (Lines and Clayton, 1977) 탄성과 자료의 실제 주파수대역으로 외삽하여 반사계수의 저주파 및 고주파 성분을 구한다. 하지만 LP 와 AR 방법은 파형요소 (wavelet)를 미리 알아야 하는 제약을 가지며 연산이 상대적으로 불안정하다.

Sacchi et al. (1994)는 새로운 방법의 MED를 소개하였다. 이 방법은 MED의 장점을 살리면서 반사계수 스펙트럼에서 손실된 주파수 대역에 관한 norm을 극대화하는 비선형 연산을 수행함으로써 선형연산이 제대로 처리하지 못하는 주파수 대역이 제한된 자료를 처리할 수 있다. 이 방법은 주파수 영역에서 반사계수의 스펙트럼을 확장하기 위해 주파수 영역에서의 제한조건을 사용하므로 FMED (MED with frequency domain constraints)라 불리운다. 이러한 방법을 이용하여 탄성과자료로부터 지하의 경계면에 해당하는 반사계수의 크기와 극성을 구한다면 매질간 음향임

피턴스의 대비가 큰 가스 하이드레이트층과 극성이 음인 가스하이드레이트 층의 기저, 그리고 그 아래 존재하는 자유가스의 존재를 파악하는 데에 매우 유용할 것이다.

이 보고서에서는 한국지질자원연구원이 가스수화물을 탐사하기 위해 동해에서 획득한 탄성파자료를 FMED를 이용하여 처리한 결과를 기술하였다.

제 2 장 연구개발 수행 내용 및 결과

제 1 절. FMED

1. 탄성과 모델

탄성파가 수직입사하는 경우의 모델은 두 가지 기본 성분인 반사계수 $y(t)$ 와 파형요소 $w(t)$ 의 컨볼루션으로 표현된다. 잡음이 없는 탄성과 트레이스 $x(t)$ 는

$$x_t = w_t * y_t \quad (1)$$

로 쓸 수 있다. 여기에서 $*$ 는 컨볼루션을 나타낸다. 디컨볼루션의 목적은 x_t 로부터 y_t 를 복원하는 것이다. 만약 디컨볼루션 필터 f_t 를 구한다면

$$y_t = x_t * f_t \quad (2)$$

로서 y_t 를 구할 수 있다. 만약 x_t 가 주파수 대역이 제한된 신호라면 y_t 는 부분적으로 구해진다.

통상, 탄성과 신호에는 잡음(n_t)이 포함되므로 탄성과기록은

$$x_t = w_t * y_t + n_t \quad (3)$$

가 된다. 우리가 원하는 디컨볼루션 필터 f_t 는 $f_t * w_t = \delta_t$ 의 성질을 가진다. 하지만 실제로 f_t 의 추정치 fe_t 를 구할 수 있으므로 디컨볼루션의 결과는 $fe_t * w_t = a_t$ 가 된다. 여기에서 a_t 는 averaging function으로 정의되며 이상적인 경우 델타함수 δ_t 처럼 된다 (Oldenburg, 1981). fe_t 를 이용하여 디컨볼루션을 수행하면

$$\begin{aligned}
y_{e_t} &= f_{e_t} * x_t = f_{e_t} * (w_t * y_t + n_t) \\
&= a_t * y_t + f_{e_t} * n_t \\
&= a_t * y_t + (y_t - \delta_t * y_t) + f_{e_t} * n_t \\
&= y_t + (a_t - \delta_t) * y_t + f_{e_t} * n_t
\end{aligned} \tag{4}$$

를 얻는다. 여기에서 y_{e_t} 는 대역이 제한된 (band-limited) 반사계수 그리고 y_t 는 전 대역 (full-band) 반사계수이다. (4)식은 디컨볼루션 필터가 a_t 를 δ_t 에 최대한 근사화 시키면서 동시에 잡음수준을 최대한 작게 억제해야 함을 지시한다. 따라서 분해능과 통계학적 안정성 사이를 조절할 수 있는 균형(trade-off)을 생각해야 한다.

아무리 탄성과 자료가 양호하고 처리기술이 뛰어나도 지하매질의 특성상 탄성과 자료는 특정의 주파수 대역만 가지게 된다. 즉, 과형요소를 제거한다 해도 반사계수 스펙트럼의 일부분만 얻게 된다. 다시 말해, $y_{e_t} = y_t * a_t$ 는 주파수 대역이 제한된다. 디컨볼루션의 결과로서 과형요소를 제거하였다고 가정한다. 그러므로 a_t 는 특정한 주파수 대역 $[\omega_L, \omega_H]$ 에서 진폭이 일정하며 영위상 (zero-phase)을 가진다.

대역이 제한된 반사계수로부터 전대역의 반사계수를 추정하는 것은 해가 유일하지 않은 선형 역산 문제이다. 잡음을 무시하면 (4)식의 푸리에 변환은

$$Y_{e_t}(\omega) = Y(\omega) + A(\omega) \cdot Y(\omega) - Y(\omega) = A(\omega) \cdot Y(\omega) \tag{5}$$

이 된다. 여기에서, $Y(\omega)$ 는 $A(\omega)$ 가 사라지는 대역에서 어떤 값을 가져도 되지만 $Y_{e_t}(\omega)$ 의 대역은 $A(\omega)$ 가 영이 아닌 대역과 일치한다. 따라서 $Y_{e_t}(\omega)$ 를 아무리 좋게 추정해도 $A(\omega)$ 가 영이 아닌 대역 바깥쪽에 있는 $Y(\omega)$ 를 추정하는 데에는 부족하다. 따라서, (5)식을 만족하는 $y(t)$ 는 무수히 많은 수가 존재할 수 있다. 다른 말로, $Y_{e_t}(\omega)$ 는 null space에 속하는 $Y(\omega)$ 의 부분에 대해서는 아무런 정보도 제공하지 못한다.

2. 엔트로피 norm

앞에서 이야기한 문제에 대한 모든 가능한 해중에서 반사계수를 합리적으로 구할 수 있는 방법을 찾고자 한다. 보통, 절약성(parsimony)은 타당한 모델이 가져야 할 특성이다. “최소 구조(minimum structure)” 혹은 “단순한 해(simple solution)”는 절약성이 있는 모델을 정의하는 데에 사용된다. Wiggins (1977)의 최초 연구에서 최소 엔트로피는 최대 질서 (maximum order)와 같은 개념으로 사용되었다. 이 용어는 MED와 스파이킹 혹은 예측 디컨볼루션 사이의 주된 차이를 지시한다. 스파이킹 혹은 예측 디컨볼루션이 자료의 스펙트럼을 무색대역으로 만들려고 (whitening) 하는 반면, MED는 몇 개의 스파이크들로 이루어진 해를 찾는다. Wiggins의 엔트로피는 인수분석에서 유래된 것으로 다음 형태의 여러 가지 norm에서 특정한 하나로 간주 된다:

$$V(\mathbf{y}) = \frac{1}{NF(N)} \sum_{i=1}^N q_i \cdot F(q_i) \quad (6)$$

여기에서 vector \mathbf{y} 는 길이 N인 반사계수 급수를 나타내며 q_i 는

$$q_i = \frac{y_i^2}{\sum_k y_k^2 / N} \quad (7)$$

에 의해 주어지는, 진폭이 정규화된 측정치이다.

(6)식에서, $F(q_i)$ 는 q_i 의 단조증가하는 함수이며 엔트로피 함수로 불리운다 (De Vries and Berkhout, 1984).

$$F(1) \leq F(2) \leq \dots \leq F(N)$$

이므로 $F(1)/F(N) \leq 1$ 이다. 만약 $F(q_i)$ 를 정의하면 $F(1)/F(N) \leq V(\mathbf{y}) \leq 1$ 의 부등식을 얻을 수 있다. (6)식에서 정규화하는 인수는 어떤 엔트로피 함수에 대해서라

도 상한이 같게 만든다. 가장 간단한 경우로서 어떤 급수가 하나의 스파이크와 나머지는 전부 영으로 이루어진다면 norm은 상한 $V(\mathbf{y}) = 1$ 이 될 것이다. 모든 샘플이 전부 같다면 $V(\mathbf{y})$ 는 하한이 될 것이다.

원래의 MED norm은 $F(q_i) = q_i$ 일 때 얻어진다.

3. $V(\mathbf{y})$ 의 최대화

(2)식에서 표현된 문제의 trivial solution은

$$\begin{aligned} f_e(t) &= x^{-1}(t), \\ y_e(t) &= \delta(t) \end{aligned} \tag{8}$$

이다. 실제로 유한한 길이의 $f(t)$ 를 고려해야 하므로

$$y_n = \sum_{l=1}^{LF} f_l \cdot x_{n-l}$$

의 방식을 따른다. 연산자 f_k 를 설계하기 위한 기준은 다음과 같다:

$$\begin{aligned} \frac{\partial V(\mathbf{y})}{\partial f_k} &= 0, \quad k = 1, 2, \dots, LF, \\ \frac{\partial V(\mathbf{y})}{\partial f_k} &= \frac{\partial}{\partial f_k} \frac{1}{N F(N)} \sum_i q_i \cdot F(q_i) \\ &= \frac{1}{N F(N)} \sum_i \left\{ F(q_i) + q_i \frac{\partial F(q_i)}{\partial q_i} \right\} \frac{\partial q_i}{\partial f_k} \end{aligned} \tag{9}$$

$G(q_i) = \left\{ F(q_i) + q_i \frac{\partial F(q_i)}{\partial q_i} \right\}$ 라고 하면 (2)식으로부터, $\partial y_n / \partial f_k = x_{n-k}$ 이므로 (9)식

은 다음과 같이 전개할 수 있다.

$$0 = \sum_i G(q_i) \left(\frac{2y_i}{\frac{1}{N} \sum_j y_j^2} \frac{\partial y_i}{\partial f_k} - \frac{y_j^2 \sum_j 2y_i \frac{\partial y_i}{\partial f_k}}{\frac{1}{N} \sum_j y_j^2} \right)$$

$$\downarrow \sum_i G(q_i) y_i \frac{\partial y_i}{\partial f_k} = \sum_i G(q_i) \frac{y_i^2}{\sum_j y_j^2} \sum_j y_j \frac{\partial y_j}{\partial f_k}$$

(7)식으로부터 $\frac{y_i^2}{\sum_j y_j^2} = \frac{q_i}{N}$ 이므로

윗 식은

$$\sum_i G(q_i) y_i \frac{\partial y_i}{\partial f_k} = \sum_j y_j \frac{\partial y_j}{\partial f_k} \sum_i G(q_i) \frac{q_i}{N} = \sum_i y_i \frac{\partial y_i}{\partial f_k} \sum_j G(q_j) \frac{q_j}{N}$$

$$\downarrow \sum_i \frac{G(q_i) y_i}{\sum_j G(q_j) \frac{q_j}{N}} \frac{\partial y_i}{\partial f_k} = \sum_n y_n \frac{\partial y_n}{\partial f_k} \text{ 이 되며}$$

윗식에서 좌변의

$$\frac{G(q_i) y_i}{\sum_j G(q_j) \frac{q_j}{N}} \text{ 을 } b_i \text{ 라 한다.} \quad (9-a)$$

또한 윗식 좌변의 $\frac{\partial y_i}{\partial f_k}$ 는 x_{i-k} 이고 우변에서 y_n 은 $\sum_l f_l x_{n-l}$ 그리고 $\frac{\partial y_n}{\partial f_k}$ 는 x_{n-k}

이므로 윗식은

$$\sum_i b_i x_{i-k} = \sum_l f_l \sum_n x_{n-l} x_{n-k}$$

$$\downarrow \sum_l f_l \sum_n x_{n-l} x_{n-k} = \sum_i b_i x_{i-k} \quad (10)$$

(10)식은 잘 알려진 Wiener 혹은 shaping 필터 (Robinson and Treitel, 1980)를 설계하기 위해 사용되는 시스템으로 행렬로는

$$\mathbf{f} \cdot \mathbf{R} = \mathbf{g}(\mathbf{f}) \quad (11)$$

로 쓸 수 있다. 여기에서 \mathbf{R} 은 자료의 Toeplitz 행렬이며 벡터 $\mathbf{g}(\mathbf{f})$ 는 \mathbf{b} 와 \mathbf{x} 의 상호상관이다. 위의 시스템은 다음 방식으로 반복적으로 풀어야 한다.

$$\mathbf{f}^{(n)} = \mathbf{R}^{-1} \cdot \mathbf{g}(\mathbf{f}^{(n-1)}) \quad (12)$$

여기에서 위첨자 n 은 반복횟수를 나타낸다. 각 반복단계에서 이 시스템은 Levinson 연산으로 푼다. 시스템의 초기값은 $\mathbf{f}^{(0)} = (0, 0, \dots, 1, \dots, 0, 0, 0)$ 이다. 각 반복단계에서 이 시스템은 반사계수를 구현한다. 적당한 반복계산 후에 이 시스템은 최대값을 가지게 되며 반사계수들을 추정하게 된다.

4. FMED 연산

주파수 영역에서, 엔트로피의 최대화는 다음의 제약조건에 종속된다:

$$Y(\omega) = Y_e(\omega), \quad \omega \in [\omega_L, \omega_H] \quad (13)$$

실제적인 목적상, (5)식을 다음과 같이 수치적인 푸리에변환으로 표현한다:

$$Y_e(\omega) = A_k \cdot Y_k \quad (14)$$

여기에서, 아랫첨자 k 는 주파수이다. 중간대역의 제한조건으로써 $V(\mathbf{y})$ 를 최대화시키는 것은 다음과 같이 쓸 수 있다:

$$\text{Maximize } V(\mathbf{y}), \quad (15)$$

subjected to

$$Y_k = \sum_{n=0}^{N-1} \{y\}_n e^{-\frac{i2\pi kn}{N}}, k = k_L, \dots, k_H \quad (16)$$

여기에서 $\{y\}$ 는 y 의 추정치이며 K_L 과 K_H 는 ω_L 과 ω_H 에 해당하는 샘플들이다. 연산과정에서 중간대역 $[\omega_L, \omega_H]$ 은 바뀌지 말아야 한다.

위의 최대화 문제의 해는 다음 시스템의 식들을 풀어서 구한다:

$$\frac{\partial V(y)}{\partial Y_k^*} + \sum_{l=k_L}^{l=k_H} \lambda_l \frac{\partial}{\partial Y_k^*} \left(y_l - \sum_{n=0}^{N-1} \{y\}_n e^{-\frac{i2\pi knl}{N}} \right) = 0 \quad (17)$$

$$Y_k - \sum_{n=0}^{N-1} \{y\}_n e^{-\frac{i2\pi kn}{N}} = 0, k = k_L, \dots, k_H \quad (18)$$

여기에서 $*$ 는 conjugate이며 λ_l 은 Lagrange multipliers 이다. 도함수를 구하고 제약 조건에 Y_K 를 대입하고 (17)식에 multipliers를 대입하면 다음의 결과를 얻을 수 있다:

$$Y_k = \frac{\sum_{n=0}^{N-1} G(q_n) y_n e^{-\frac{i2\pi kn}{N}}}{\sum_j G(q_j) q_j / N}, \quad k \notin [k_L, k_H]$$

$$Y_k = \{Y\}_{k'}, \quad k \in [k_L, k_H] \quad (19)$$

(11)과 (13)식으로부터

$$Y_k = B_k, \quad k \notin [k_L, k_H]$$

$$Y_k = Y_{e_{k'}}, \quad k \in [k_L, k_H] \quad (20)$$

임을 알 수 있다. 여기에서 B_k 는 b_t 의 푸리에 변환이다. b_t 는 y_t 의 비선형함수이므로 FMED는 다음과 같이 실행된다:

- (1) $y_t = \{y\}_t$ 로 놓음으로써 연산을 시작한다 (단, $\{y\}$ 는 y 의 추정치).
- (2) b_t 와 B_k 의 요소들을 계산한다
- (3) 복원해야할 저주파와 고주파성분을 B_k 로 치환한다.
- (4) 역푸리에변환으로부터 반사계수의 추정치를 계산한다. 수렴성을 측정하기 위해

norm $V(y)$ 를 계산하며 (2)단계에서 반복한다.

제 2 절 탄성과 자료 처리

1. 탄성과 자료 및 처리 결과

한국지질자원연구원이 제공한 탄성과 프로파일은 5 개로서 Fig. 2-1에서 Fig. 2-5에 나와 있다. 프로파일에서 각 트레이스의 길이는 7 초로서 1 초당 1023 개의 샘플이므로 트레이스 전체는 7169개의 샘플로 이루어져 있다. 전체적으로 프로파일은 해저면에 평행한 반사면이 잘 발달되어 있다. 특히, 관심의 대상인 가스 하이드레이트 BSR로 해석할 수 있는, 극성이 반전된 뚜렷한 반사면은 해저면 아래 220 - 230 ms에 분포한다.

FMED를 이용한 처리결과를 검증하기 위해 프로파일에서 몇 개의 트레이스를 추출하여 디컨볼브한 결과를 Fig. 2-6에서 Fig. 2-10 에 나타내었다.

Fig. 2-6의 위 패널은 프로파일 05gh_1002_s003의 CDP 461에 해당하는 트레이스로서 1초부터 2초까지를 그린 것이다. 처음 나타나는 신호가 해저면의 반사신호로서 잇달아 여러 개의 반사신호가 중첩되어 있다. 프로파일에서 BSR로 추정되는 반사신호는 1.3초와 1.4초사이에 존재하는데 (Fig. 2-1) 트레이스를 살펴보면 그 크기가 해저면 반사신호보다 더 큰 것을 알 수 있지만 반사계수의 크기와 극성을 추정하기가 곤란하다. FMED의 처리 결과는 Fig. 2-6의 아래 패널로서 의미있는 반사신호들을 잘 분리한 것으로 판단된다. FMED의 결과는 파형요소를 효과적으로 델타함수로 전환시킴으로써 반사계수만을 추출한 것을 보여 준다. 디컨볼브된 트레이스에서 해저면 반사계수는 0.14초 (실제로는 1.14초)에서 나타나며 그 바로 아래에 음의 반사신호가 존재한다. 프로파일에서 천부의 퇴적물들은 저탁류(turbidite)의 퇴적상을 보여주는데 표층 바로 아래의 지층이 표층보다 밀도가 더 낮은 영향으로 볼 수 있다. 이런 음의 반사계수를 제외하면 표층과 BSR 사이에서 대부분의 반사계수들은 양의 값을 보이며 지층들의 밀도가 아래로 갈수록 증가함을 알 수 있다. 프로파일에서 BSR로 추정되는 신호의 반사계수는 해저면에서 217 ms 아래에 있으며 (트레이스에서는 360 ms) 반사계수는 해저면의 반사신호보다 더 크게 나타난다.

이 반사신호 다음에 큰 진폭의 양의 반사계수가 트레이스에서 390 ms 에 존재한다. 가스 하이드레이트층이 존재할 경우 흔히 그 아래에 자유가스 (free gas)가 존재한다고 보고되어 있는데 만약 여기에서 BSR이 가스 하이드레이트 BSR이라면 그 아래의 양의 큰 반사계수는 BSR과 그 사이에 자유가스가 있음을 지시한다고 볼 수 있다.

Fig. 2-7은 앞의 예와 동일한 프로파일 05gh_1002_s003의 CDP 3281번을 처리한 것을 보여 준다. 프로파일에서는 BSR로 추정되는 반사신호가 해저면에서 약 220 ms 아래에 있을 것으로 보이지만 명확하게 분리되어 나타나지 않으며 (Fig. 2-1) 트레이스에서도 BSR을 구분하기가 힘들다 (Fig. 2-7). 디컨볼브된 결과는 분리하기 힘든 반사계수들을 잘 보여 준다. 해저면 반사계수와 BSR로 추정되는 반사계수는 트레이스에서 각각 380 ms와 600 ms에서 발생하는데 역시 그 간격이 220 ms임을 알 수 있다. 이 경우에는 BSR의 반사계수가 앞의 예보다 더 작은 값을 갖는다. 프로파일에서 해저면에서 약 400 ms 아래에서도 음의 반사계수를 갖는 것으로 추정되는 반사면이 존재하는데 디컨볼브된 트레이스에서도 780 ms에서 그것을 확인할 수 있다. 앞의 예와는 달리 BSR 아래에서 보이는 반사신호는 큰 진폭이 아니며 반사계수의 부호도 음으로 나타난다.

Fig. 2-8은 같은 프로파일상의 CDP 3721의 처리결과를 그린 것이다. 앞의 예와 달리 트레이스는 상대적으로 많은 반사계수로 이루어져 있다. 이것은 연산과정에서 threshold를 낮춘 것과 원래의 자료에 반사계수가 더 많이 존재하기 때문일 것이다. 프로파일에서 해저면이 함몰되어 있으므로 상대적으로 BSR은 해저면에서 더 짧은 거리인 180 ms아래에서 발생하는 것으로 추정된다. 처리결과에서도 그 범위에서 (트레이스에서는 570 ms) 극성이 음인 반사계수를 확인할 수 있다. BSR 아래에는 많은 수의 음의 반사신호가 존재한다.

프로파일 05gh_1043_s037의 경우 BSR로 추정되는 반사면이 잘 발달하여 있으며 CDP 1400을 중심으로 층서를 자르는 양상을 보인다 (Fig. 2-4). 하지만 BSR은 다른 반사신호와 겹쳐서 나타나므로 진폭과 크기를 구하기 힘들다. Fig. 2-9는 CDP 13081 트레이스와 처리결과이다. 해저면 반사와 극성이 반대인 반사는 해저면에서 200 ms 아래에서부터 여러 개가 나타나고 있다. Fig. 2-10은 동일한 프로파일에서 CDP 14301에 대한 처리예이다. 여기에서는 해저면에서 약 240 ms 아래에서부터 여러 개의 음의 반사계수가 나타난다. BSR과 관련된 반사신호가 여러 개의 음의 반

사계수들로서 나타날 때 전체 트레이스를 동시에 그려보면 BSR을 구분할 수 있을 것으로 판단된다.

프로파일 05gh_1043_s048 에서도 해저면에서 BSR로 해석할 수 있는 반사면이 약 220 ms 아래에 존재한다 (Fig. 2-5). Fig. 2-11은 프로파일 의 CDP 1501을 처리한 것이다. BSR 반사계수는 해저면의 반사계수보다 훨씬 작으며 바로 아래에 극성이 음인 반사계수가 따라서 존재한다.

Fig. 2-12의 왼쪽은 05gh_1043_s037의 CDP 11241 에서 CDP 15501까지의 부분이다. 이 그림에서 BSR이 해저면에서 약 220 ms 아래에 해저면과 평행하게 발달해 있으며 주위의 퇴적층서를 자르는 것을 볼 수 있지만 BSR이 다른 반사신호와 겹쳐 있어서 분리하기가 곤란하다. 이 프로파일을 FMED로써 처리하여 구한 결과는 Fig. 2-12의 오른쪽 그림이다. Fig. 2-13은 반사계수의 파형을 보여주기 위해 100 개 CDP 마다 처리결과를 그린 것으로 채워지지 않은 파형은 양의 극성을, 그리고 채워진 파형은 음의 극성을 갖는다. 전체적으로 BSR은 연속성이 매우 좋은 음의 극성을 갖는 반사계수로서 표현이 되며 그 진폭도 해저면 반사보다 더 크게 나타나고 있다. 해저면과 평행하면서 반대의 극성을 갖는 반사면의 성질을 갖는 BSR은 여러 트레이스에서 잇다른 두 개의 반사신호로 나타난다. 만약 위에 있는 BSR이 가스 하이드레이트의 안정영역의 기저에 해당한다면 자유가스가 BSR 바로 아래에서부터 존재할 경우 두 개의 반사신호 사이에 있을 가능성도 있지만 이 부분에 대해서는 보다 정확한 지질해석이 요구된다. Fig. 2-12에서 프로파일의 오른쪽 부분에 가스가 해저면부근까지 올라온 형태를 볼 수 있는데 해저면 바로 아래에 음의 값을 갖는 반사면이 존재하는 것은 가스의 상부로 해석할 수 있다.

Fig. 2-14는 프로파일 05gh_1043_s048의 일부에 해당하는 CDP 1501 에서 CDP 1981 까지를 처리한 결과이다. 역시 앞의 예와 마찬가지로 해저면 아래 약 220 ms 에서 보이는 BSR이 매우 잘 분해되었으며 그 아래에 대체로 음의 부호가 우세한 반사면이 존재한다.

2. 추출된 파형요소

탄성과 자료로부터 추출한 파형요소는 Fig. 2-15와 Fig. 2-16이다. 파형요소는

음의 압력을 나타내는 trough와 양의 압력을 나타내는 peak, 그리고 다시 음의 압력을 나타내는 trough로 구성되며 잔향(reverberation)이 매우 효과적으로 억제되어 있음을 보여준다. 이러한 특징은 탐사에 사용된 에어건이, 복잡한 지층구조를 대상으로 하는 탄성과 탐사에 필수적인 분해능이 높은 파형요소를 만들었음을 제시한다. 전체적으로 파형요소의 중심주파수는 50 Hz로 추정된다.

제 3 절. 결 론

탄성과 트레이스로부터 지하의 반사계수에 해당하는 sparse한 스파이크의 급수를 추출하는 FMED방법을 구현하여 가스 하이드레이트 탐사를 위해 얻은 자료에 적용하여 그 결과를 검증하고 BSR의 신호특성을 해석하였다.

이 연구에서 적용한 방법을 이용하여 탄성과 자료로부터 해저면과 그 아래 경계면에 해당하는 반사층의 반사계수의 크기와 극성을 구할 수 있었다. BSR로 추정되는 반사신호는 대부분 해저면으로부터 210 - 220 ms 아래에 해저면과 평행하게 발달해 있으며 극성이 음인 특징을 잘 보여 준다. 이 BSR은 진폭이 매우 커서 많은 경우에 해저면 반사신호의 진폭보다 더 크다. 조사지역의 BSR은 진폭의 변화가 매우 크며 지역적으로 이중 BSR (double BSR)의 특징을 보여주기도 한다. BSR 아래에 상당한 진폭의 반사면이 존재하므로 BSR 하부에서 자유가스의 존재와 관련된 구조를 지시할 가능성이 있다. 하지만 BSR 아래의 반사신호는 그 극성이 일관되지 못하므로 이에 대해 타당한 해석이 필요하다.

탄성과 프로파일의 자료로부터 구한 파형요소는 반향신호가 매우 효과적으로 억제되어 분해능이 우수한 것으로 판단된다. 파형요소의 중심 주파수는 50 Hz로서 가스 하이드레이트 탐사에 적합하다.

이 연구에서는 제한된 탄성과 프로파일을 처리하였지만 이 연구에서 사용한 방법을 차후 시험시추를 계획하고 있는 지역에서 얻은 탄성과자료에 적용함으로써 우리나라 해역에서 가스 하이드레이트 탐사와 개발의 효율을 증대시킬 수 있을 것으로 기대한다.

제 3 장 연구개발결과의 활용계획

1. 탄성과 자료로부터 지하의 반사계수를 추출하는 것은 특히 가스 하이드레이트의 기저를 표현하므로 반사계수가 크고 극성이 음인 BSR의 탐지와 그 신호특성에 매우 유리하다. 이 연구에서 구현한 FMED는 한국지질자원연구원에서 획득한 탄성과자료에 적용했을 때 해저면과 그 아래에서 발생한 반사신호를 분리하고 진폭과 극성을 계산함으로써 가스 하이드레이트의 탐사에 유용하게 활용될 수 있다.
2. 탄성과 자료로부터 지하의 음향임피던스를 복원하는 것은 가스 수화물의 안정영역과 그 아래 존재하는 가스층으로 이루어진 지질구조의 구명에 매우 유용하다. 이 연구에서 구한 반사계수들은 탄성과자료로부터 지하의 음향임피던스를 복원하는 데에 바로 이용될 수 있다.
3. 가스 하이드레이트 안정영역 아래에 가스가 존재할 경우 탄성과의 감쇠가 심하게 발생하므로 이것을 이용한 가스 하이드레이트 및 자유가스의 탐사가 중요하다. 이를 위해서는 파형요소를 아는 것이 필수적이다. 이 연구에서 구한 파형요소는 탄성과 역산과 가스의 존재에 대한 자료 분석 등에 매우 중요하게 사용될 수 있을 것을 기대한다.
4. 탄성과 자료에서 보이는 가스 하이드레이트 부존 구조에서 BSR과 그 아래 반사면의 성질을 해석할 수 있다. 이런 과정을 통하여 유망한 구조를 찾고 시추의 성공율을 높일 수 있을 것으로 기대된다.

제 4 장 참고문헌

Levy, S., and Fullagar, P.K., 1981, Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution: *Geophysics*, v.46, p.1235-1243.

Lines, L.R., and Clayton, R.W., 1977, A new approach to vibroseis deconvolution: *Geophysics*, v.25, p.417-433.

Robinson, E.A., and Treitel, S., 1980, *Geophysical signal analysis*: Prentice-Hall Inc.

Sacchi, M.D., Veils, D.R., and Cominguez, A.H., 1994, Minimum entropy deconvolution with frequency-domain constraints: *Geophysics*, v.59, p.938-945.

Wiggins, R.A., 1978, Minimum entropy deconvolution: *Geopl.*, v.16, p.21-35.

부록 1. FMED 연산

여기에서는 1장에서 다룬 (20)식의 유도과정을 정리하였다.

주파수 대역이 제약조건으로 사용되는 알고 있는 대역은 $B_\omega = k_l, \dots, k_H$ 이다.

$$\sum_{l=B_\omega} \equiv \sum_{l=k_l}^{k_H}$$

Case 1 : 알려진 주파수 대역의 바깥영역 ($k \notin B_\omega$) 에서의 연산

이 경우 $l \in B, k \notin B$ 이므로 (17) 식에서 좌변의 둘째 항이 다음과 같이 영이 되므로

$$\frac{\partial}{\partial Y_k^*} \left(y_l - \sum_{n=0}^{N-1} \{y\}_n e^{-i2\pi knl} \right) = 0$$

(17) 식은

$$\frac{\partial V(y)}{\partial Y_k^*} = 0 \tag{a-1}$$

가 된다. (a-1)식을 풀기 위해서 chain rule을 적용하면

$$\begin{aligned} \frac{\partial V(y)}{\partial Y_k^*} &= \frac{\partial}{\partial Y_k^*} \left(\sum_i \frac{1}{NF(N)} F(q_i) q_i \right) \\ &= \frac{1}{N F(N)} \sum_i \left(F(q_i) + q_i \frac{\partial F(q_i)}{\partial q_i} \right) \frac{\partial q_i}{\partial Y_k^*} \\ &= \frac{1}{N F(N)} \sum_i G(q_i) \frac{\partial q_i}{\partial Y_k^*} \left(\because F(q_i) + q_i \frac{\partial F(q_i)}{\partial q_i} = G(q_i) \text{ from (9)} \right) \end{aligned} \tag{a-2}$$

가 된다.

(a-2)식을 풀기 위해 먼저, (7)식으로부터 다음 관계를 얻을 수 있다:

$$\frac{\partial q_i}{\partial Y_k^*} = \frac{2y_i}{\frac{1}{N} \sum_l y_l^2} \frac{\partial y_i}{\partial Y_k^*} - \frac{y_i^2 \sum_l 2y_l \frac{\partial y_l}{\partial Y_k^*}}{\frac{1}{N} \left(\sum_l y_l^2 \right)^2} \tag{a-3}$$

(a-3)과 (a-2)식을 결합하고 (9-a)식에 나온 b_i 의 정의를 이용하면 다음을 얻는다.

$$\frac{1}{N} \sum_i G(q_i) \left(\frac{2y_i}{\frac{1}{N} \sum_l y_l^2} \frac{\partial y_i}{\partial Y_k^*} - \frac{y_i^2 \sum_l 2y_l \frac{\partial y_l}{\partial Y_k^*}}{\frac{1}{N} \sum_l y_l^2 \sum_l y_l^2} \right) = 0$$

윗 식에서 없애도 성립하는 항 (예를 들어, $1/N$)을 소거해서 간단히 쓰면

$$\sum_i G(q_i) y_i \frac{\partial y_i}{\partial Y_k^*} = \sum_i G(q_i) \frac{y_i^2 \sum_l y_l \frac{\partial y_l}{\partial Y_k^*}}{\sum_l y_l^2}$$

$$\downarrow \sum_i G(q_i) y_i \frac{\partial y_i}{\partial Y_k^*} = \sum_i G(q_i) \frac{\sum_l y_l \frac{\partial y_l}{\partial Y_k^*}}{\frac{N}{q_i}} \left(\because \sum_l y_l^2 = \frac{N}{q_i} y_i^2 \leftarrow \text{eq. (7)} \right)$$

$$\downarrow \sum_i G(q_i) y_i \frac{\partial y_i}{\partial Y_k^*} = \sum_i G(q_i) q_i \frac{1}{N} \sum_l y_l \frac{\partial y_l}{\partial Y_k^*}$$

(9-a)식에서 $G(q_i) y_i = b_i \frac{1}{N} \sum_j G(q_j) q_j$ 인데 이것을 윗 식에 대입하면

$$\downarrow \sum_i b_i \cdot \frac{1}{N} \sum_j G(q_j) q_j \cdot \frac{\partial y_i}{\partial Y_k^*} = \sum_i G(q_i) q_i \cdot \frac{1}{N} \sum_l y_l \frac{\partial y_l}{\partial Y_k^*}$$

$$\downarrow \sum_i b_i \frac{\partial y_i}{\partial Y_k^*} \sum_j G(q_j) q_j = \sum_i y_i \frac{\partial y_i}{\partial Y_k^*} \sum_j G(q_j) q_j$$

$$\begin{aligned} \because \sum_i f(y_i) \sum_l y_l \frac{\partial y_l}{\partial Y_k^*} &= \sum_i f(y_i) \left(y_1 \frac{\partial y_1}{\partial Y_k^*} + \dots + y_N \frac{\partial y_N}{\partial Y_k^*} \right) \\ &= f(y_1) \left(y_1 \frac{\partial y_1}{\partial Y_k^*} + \dots + y_N \frac{\partial y_N}{\partial Y_k^*} \right) + \dots + f(y_N) \left(y_1 \frac{\partial y_1}{\partial Y_k^*} + \dots + y_N \frac{\partial y_N}{\partial Y_k^*} \right) \\ &= y_1 \frac{\partial y_1}{\partial Y_k^*} \{f(y_1) + \dots + f(y_N)\} + \dots + y_N \frac{\partial y_N}{\partial Y_k^*} \{f(y_1) + \dots + f(y_N)\} \\ &= \sum_l y_l \frac{\partial y_l}{\partial Y_k^*} \sum_i f(y_i) = \sum_i y_i \frac{\partial y_i}{\partial Y_k^*} \sum_l f(y_l) \end{aligned}$$

결국 윗 식은

$$\sum_i b_i \frac{\partial y_i}{\partial Y_k^*} = \sum_i y_i \frac{\partial y_i}{\partial Y_k^*} \quad (\text{a-4})$$

이 된다.

$\frac{\partial y_i}{\partial Y_k^*}$ 를 계산해야 하는데 푸리에 변환의 정의에 따라

$$Y_k^* = \sum_n y_n e^{\frac{j2\pi kn}{N}} \rightarrow y_i = \frac{1}{N} \sum_l Y_l^* e^{\frac{-j2\pi il}{N}}$$

이므로

$$\frac{\partial y_i}{\partial Y_k^*} = \frac{1}{N} e^{\frac{-j2\pi il}{N}} \quad (l = k) \quad (\text{a-5})$$

가 된다. 최종적으로 (a-5)식을 (a-4)식에 대입하면

$$\frac{1}{N} \sum_i b_i e^{\frac{-j2\pi ik}{N}} = \frac{1}{N} \sum_i y_i e^{\frac{-j2\pi ik}{N}}$$

$$\therefore B_k = Y_k$$

가 된다. 즉, Y_k ($k \notin B$) 의 DFT는 B_k ($=b_i$ 의 DFT)를 이용해서 추정한다.

Case 2 : 알려진 주파수 대역 ($k \in B_0$) 에서의 연산

이 경우에는 $Y_k = Y_{e_k}$ (단, Y_{e_k} 는 Y_k 의 추정치로서, 이미 알고 있는 제약조건)

이므로 trivial 하다.

(1)과 (2)의 경우를 다 합하면 (20)식에 해당하는

$$Y_k = B_k, \quad k \notin [k_L, k_H]$$

$$Y_k = Y_{e_k}, \quad k \in [k_L, k_H] \quad (\text{a-6})$$

를 얻는다.

부록 2. 탄성파자료의 파형요소 추출

I. 서론

신호는 사람과 사람, 기계와 기계 그리고 사람과 기계들 사이의 통신에 사용된다. 쉽게 관찰할 수 없는 구조나 상태의 세부사항을 밝히기 위해서 주변 환경을 조사하거나 에너지와 정보를 이용하는데 사용되어져 왔다. 신호 처리는 신호와 그 기기에 포함되어 있는 정보의 표현, 변환 그리고 조작에 관계되는 내용이다. 예를 들어 여러 신호로 결합되어 있는 중복신호를 분리하고자 할 때, 어떤 신호 모델의 성분 또는 매개 변수의 성능을 향상시키고자 할 때 신호처리가 사용되곤 한다. 수십 년에 걸쳐 신호처리는 음성통신, 데이터 통신, 생의학 공학, 음향심리학, 수중음파 탐지기, 레이더, 지진학, 석유탐사, 각종기기장비, 로봇 그리고 가정용 전자제품들과 같은 다양한 분야에서 중요한 역할을 수행하여 왔다.

정교한 신호처리 알고리즘과 하드웨어는 산업용 또는 고도의 기술을 필요로 하는 특수 군사용 시스템에서, 그에 비해 낮은 기술과 부피가 큰 가정용 전기·전자 제품에 이르기까지 넓은 시스템 영역에서 유용하게 사용되어왔다. 비록 우리들이 텔레비전, 고급 오디오등과 같은 가정용 오락기의 시스템 성능을 일상적으로 당연한 것처럼 평범하게 여겨왔지만 이들 시스템은 언제나 최첨단 기술의 신호처리에 크게 의지하여 왔다. 또 하나의 예로서 음성합성이 있다. 이 기술은 자동 소리응답 시스템이나 학습보조도구와 장난감과 같은 일반 소비자용 제품에 빠르게 진입하고 있다. 그리고 이 기술의 이동은 연구학문 수준에서 군사용, 산업용, 그리고 일반 소비자용 상품화 수준으로 급속도로 빠르게 진행되고 있다.

종래의 많은 신호처리에서는 또 다른 신호를 얻기 위해서 하나의 신호를 처리하기도 하였다. 신호처리의 과제 중에서 또 하나의 중요한 부분이 신호번역(Signal Interpretation)이다. 여기서 신호처리의 목적은 출력신호를 얻는 것이 아니고 입력신호의 특성을 얻고자 하는 것이다. 예를 들어 음성인식 시스템의 경우, 그 목적은 입력신호를 번역하거나 그 입력신호로부터 정보를 추출하는 것이다.

신호처리의 중요한 분야 중에 하나가 지진 신호등 탄성과 신호에 대한 처리 문제이다. 특히 최근 동남아에서 발생한 자연재해로 인한 참변은 지진으로 인해 바다 속 해면이 두 쪽으로 갈라진 것이며, 이로 인한 해일이 대륙을 덮치면서 일

어난 것으로 알려졌다. 자연재해는 인간이 피할 수 없는 재난이지만, 만약 미리 지진을 예상했다면 많은 인명 피해는 줄일 수 있었을 것이다. 전 세계적으로 지각 구조에 관한 연구는 지진활동이 활발한 곳뿐만 아니라 상대적으로 지진활동이 적은 국내에서도 꾸준히 다양한 방법으로 수행되고 있다. 이는 지진의 진앙, 진원깊이를 더욱 정확히 결정하여 활성 정량화하며, 지진위험도를 합리적으로 평가하는데 기본적인 자료가 되기 때문이다.

대부분 관측소에 기록된 지진파에는 지진원, 전파경로 및 관측소 하부의 지각 구조의 효과가 모두 포함되어 있으며, 이로부터 디컨볼루션 통해 수신 지역의 구조의 효과를 제외한 복잡한 지진원과 전파경로를 효과적으로 제거 할 수 있다. 따라서 디컨볼루션을 통하여 관측소 하부의 지각과 맨틀구조에 대한 정보를 얻을 수 있다.

또한 이와 비슷한 분야로 탄성과 탐사가 있다. 이는 지하자원 탐사와 기반암까지의 깊이 결정, 지하매질의 복잡하고 기본적인 물성정보까지 얻는 데 사용된다. 탄성과 신호는 지하매질을 주행하다가 지하 경계면을 만나면 그 일부가 굴절하거나 반사되어 되돌아 온다. 이렇게 기록된 신호를 분석하여 지하구조를 조사하는 것이 탄성과 탐사의 기본 원리이다. 여기서 되돌아오는 출력신호는 지하구조를 추정할 때 쓰이며 그 추정방법에는 크게 ML(Maximum Likelihood)기법, 디컨볼루션(Deconvolution)기법, FML(Fast Maximum Likelihood)기법, 등이 있다. 여기서 관측되는 출력 신호는 반사계수와 웨이블릿이 컨볼루션되어진 형태이다.

이 보고서에서는 시간에 따라 컨볼브된 출력신호를 가지고 디컨볼루션 기법을 사용하여 반사계수가 알려졌다고 가정했을 때, 파형요소를 추정하는 방법을 제시한다. 또한, 제시된 추정방법의 검증을 위해서 몇 가지 경우의 모의실험을 하고자 한다.

2. 기초이론

A. 컨볼루션(Convolution)

두 개의 시간영역의 함수가 컨볼브된 형태는 서론에서도 언급했듯이, 여러 과

학 분야, 심지어 우리의 일상생활 깊은 곳에서도 다양하게 활용되고 있다. 여기서 컨볼루션에 대해 자세히 알아보기로 하자.

먼저, 두 신호 r_t 와 w_t 의 컨볼루션은

$$y_t = r_t * w_t$$

와 같이 표시하며 다음과 같이 정의 된다.

$$y_t = \sum_{s=-\infty}^{\infty} r_t w_{t-s}$$

컨볼루션은 시불변 선형 시스템인 경우에 나타난다. 이는

$$w_s \rightarrow w_0$$

이고 시간 s 에서의 임펄스 응답은 시간 0 에서의 임펄스 응답과 형태는 같 으며 또한,

$$w_t \rightarrow w_{t-s}$$

인 형태와도 같다. 즉 s 만큼 시간 지연만 됐었을 뿐, 시간에 따라 시스템은 변하지 않는다. 그러므로 아래와 같이

$$y_t = \sum_{s=-\infty}^{\infty} r_t w_{t-s} = r_t * w_t$$

이 된다.

다음은 컨볼루션의 성질에 대해 간단히 살펴보자. 첫 번째 성질인 교환성은

$$r_t * w_t = w_t * r_t$$

으로 입력 r_t 과 단위 임펄스 응답 w_t 을 갖는 시불변 선형 시스템의 출력은 입력 w_t 과 임펄스 응답 r_t 을 갖는 시불변 선형 시스템의 출력과 동일하다. 두 번째 성질 결합성은

$$r_t * (w_t * w_t') = (r_t * w_t) * w_t'$$

으로 하나의 입력 r_t 와 두 개의 단위 임펄스 응답 (w_t, w_t') 의 컨볼루션으로 얻어진 출력이 다시 컨볼루션되어 나타난 시불변 선형 시스템의 출력은 입력 r_t 와 한 개의 단위 임펄스 응답 (w_t) 이 컨볼루션 되어 나타난 출력이 다른 한 개의 단위 임펄스 응답 (w_t') 과 또 다시 컨볼루션 되어 나타나는 시불변 선형 시스템의 출력과 동일하다. 세 번째 성질 분배성은

$$r_t * (w_t + w_t') = r_t * w_t + r_t * w_t'$$

으로 입력 r_t 과 두 개의 단위 임펄스의 합 (w_t, w_t') 과 컨볼루션된 시불변 선형 시스템의 출력은 입력 r_t 이 각각의 단위 임펄스 (w_t, w_t') 과 컨볼루션 되어 합하여 지는 시불변 선형 시스템의 출력과 동일하다.

위의 컨볼루션 성질은 연속시간, 이산시간인 경우 모두에 해당되며, 시불변 선형 시스템에서만 성립되는 것이다. 보다 자세한 내용은 Openheim & Willsky(1997)를 참고하기 바란다.

대부분 우리가 얻은 출력신호는 이렇게 두 개 또는 여러 개의 신호가 컨볼루션된 것들이다. 이러한 출력신호를 받아 우리는 노이즈를 제거하고 두 개 혹은 여러 개의 신호로 각각 분리해야한다. 이런 방법 중 하나가 다음에 언급할 디컨볼루션이다.

B. 디컨볼루션(Deconvolution)

컨볼루션된 두 개 혹은 그 이상의 신호들을 분리해 내기 위해서는 디컨볼루션 기법이 사용된다. 여기서 디컨볼루션에 대해 좀 더 자세히 알아보자.

디컨볼루션은 유일하게 얻어진 관측치 y_t 에서 두 개의 신호 r_t 와 w_t 을 분리하여 추정하는 것이다. 가장 간단한 방법으로는 w_t , 즉 과형요소를 직접 측정하여 관측치 y_t 에서 제거하는 것이다. 이렇게 하면 반사계수인 r_t 을 구할 수 가 있다. 이를 결정론적 디컨볼루션 (deterministic deconvolution) 이라 한다. 그러나 대부분의 경우 과형요소를 직접 관찰하기는 힘들기 때문에, 관찰치 y_t 만으로 r_t 와 w_t 을 추정해야 한다. 그래서 이를 해결하기 위해 blind deconvolution 방법을 사용 한다.

입력 신호 r_t 와 시스템 모수 w_t 가 모두 시계열이며, r_t 는 평균이 0이고 분산이 1로 서로 독립인 확률변수로 가정한다. 또한 w_t 는 $\sum w_t^2 < \infty$ 을 만족하며 , y_t 은 w_t 와 r_t 가 컨볼루션 되어 관측되는 값이라 가정하자.

$$y_t = r_t * w_t = \sum_{s=-\infty}^{\infty} r_t w_{t-s} \quad t=1, 2, \dots, n$$

위와 같이 가정했을 때, Brockwell & Davis(1991)의 논문과 Azencott & Dacunha-Castelle(1984) 논문에서는 r_t 가 최소위상이고 w_t 의 정상성을 가정할 때, 선형 예측의 전형적인 방법을 사용하여 디컨볼루션 추정 방법을 제시하였다. 그 후 Li(1992)의 논문에서 r_t 의 최소위상과 정상성을 가정하지 않고, w_t 의 분포 정보를 이용하지 않는 일반적인 경우의 디컨볼루션 방법이 제안되었다.

탄성과 자료의 디컨볼루션을 위해 개발된 방법으로는 Wiener(1949)가 제시한 Wiener filtering 또는 Mendel(1983)이 제시한 Minimum variance deconvolution 등이 있다. 이는 정상성을 띠는 가우시안 가정을 전제로 하며, 계산이 용이하다는 장점이 있으나 실제 데이터가 정상성을 띠는 가우시안을 따르지 않기에 현실성이 없는 방법으로 여겨졌다. 이외에도 다른 많은 디컨볼루션 방법이 제시되었다. 그 방법으로는 Kaaresen & Torfinn Taxt(1998)이 제시한 "Block Component Method"와 Wiggins(1978)의 "Minimum Entropy Deconvolution"이 있다. 또한, Minimum Entropy criterion을 이용해 반사계수를 추정하는 방법이 Mauricio et al.(1994)에 의해 제시 되었다.

C. 파형요소 (Wavelet)

관측치 y_t 에서 파형요소의 추정에 대해 알아보자. 파형요소는 전파경로에 따라 일정한 값을 가지며 반사된 신호는 늦게 들어온 신호에 중첩되어 나타난다. 신호가 반사된 위치는 탄성파로 예를 들자면, 매질이 변하는 위치이다. 이로 인해 우리는 지층의 구조를 알아낼 수 있다.

파형요소의 추정 방법으로는 Lazear(1993)이 제시한 "Mixed-phase wavelet estimation using fourth-order cumulants"이 있다. 여기서는 관측치인 트레이스(trace)는 컨볼루션된 모델로써 노이즈가 가우시안을 따르나, 반사계수는 가우시안을 따르지 않고 정상성을 띠며 독립적으로 확률과정을 따른다고 가정하고 있다. 이런 가정 하에 파형요소의 4차 적률은 관측치의 4차 cumulant의 척도화된 버전으로 표현할 수 있다. 그리고 이것의 정합은 최소제곱오차를 감지할 수 있다. 그러나 이는 일반적인 비선형적 문제를 가지고 있으며, 신뢰성있는 추정값을 얻기 위해서는 수백 개의 트레이스를 필요로 한다. 이 문제점을 보완한 Velis & Ulrych(1995)이 제시한 "Improved wavelet estimation using fourth-order cumulants based on very fast simulated annealing"이 있다. 여기서는 수치적인 신뢰성을 제공하며, 일반적인 것을 포괄하는 알고리즘을 사용한다. 또한, 다차원

의 매끄러운 트레이스 cumulant를 지원하고 정합을 향상 시킨다. 또한 이 연구에서는 큰 데이터 사용의 문제를 해결하였으며, 이를 250개의 샘플을 사용하여 신뢰할 수 있는 추정값을 얻을 수 있음을 모의 실험을 통해 설명하고 있다.

그리고 Rietsch(1997)가 제시한 “Euclid and the art of wavelet estimation, part I: Basic algorithm for noise-free data”에서는 지진파 관측치의 자기상관과 교차상관의 행렬의 고유값이 0인 개수는 파형요소의 샘플수와 동일하고, 고유벡터는 반사계수와 관련이 있다는 것을 이용해, 지진파 파형요소를 다항식에서 찾을 수 있다고 제시한다.

또한 Lian et al.(2002)이 제시한 “Using high-order cumulants to extrapolate spatially variant seismic wavelet”이 있다. 이는 어떠한 공간에서 특정한 지진 파형요소를 추정하는 방법이다. 이들은 지진파 자료를 가지고 third -and fourth-order cumulants를 사용하여 파형요소를 추정하는 방법을 제시하였다. 이는 지진파의 파형요소의 고차원 적률과 지진파 데이터의 고차원 cumulants를 정합시킴으로써 다차원 필터를 고안한 것이다.

3. 추정방법

3장에서는 여러 개의 반사계수가 1개의 공통 파형요소과 컨볼루션되어 여러 개의 출력신호가 얻어졌을 때 파형요소를 추정하는 방법을 제시한다. 즉 n 개의 파형요소 $\{r_{it}, t=0, 1, \dots, T, i=1, \dots, n\}$ 과 1개의 공통 반사계수 $\{w_t, t=0, 1, \dots, T\}$ 의 컨볼루션으로 n 개의 트레이스 $\{y_{it}=r_{it} * w_t, t=0, 1, \dots, T, i=1, \dots, n\}$ 이 관측되었을 때 이 관측치로부터 반사계수를 추정하는 문제를 논하고자 한다

3.1 적률법

파형요소의 추정 방법을 제시하기 전에 다음과 같은 가정이 필요하다.

- 1) r_{it} 은 모든 i, t 에 대해 평균 0 분산 1 이고 서로 독립이다.

2) 파형요소는 연속이며, 진동한다.

다음 식은 반사계수와 파형요소의 컨볼루션을 나타낸 것이다.

$$y_{it} = r * w_t, \quad i=1, 2, \dots, n \quad t=0, 1, \dots, T$$

$$= \sum_{s=0}^t r_{i, t-s} w_s$$

여기에서 *은 컨볼루션의 연산기호이며, r_{it} 는 반사계수를 나타내며, w_t 는 파형요소를 나타낸다.

위 식에서 파형요소를 추정하기 위해 먼저 파형요소의 제곱값을 추정해야 한다. 파형요소의 제곱값의 추정은 적률법(Method of moment)을 사용하며, 이 추정 방법의 개요에 대해 설명하고자 한다. n 개의 확률변수 X, \dots, X_n 가 서로 독립이고 동일한 분포를 따르며(*iid*), 모수가 Θ 인 확률밀도함수 $p(x|\theta)$ 을 따른다고 가정하자. 이때 X 가 이산형 확률변수이면

$$\mu_k(\theta) = E[X^k | \theta] = \sum x^k p(x|\theta), \quad k=1, 2, \dots$$

이 되고, 이것을 k 차 모적률(k -th population moment)이라 한다. 이것의 샘플버전인

$$m_k = \sum_{i=1}^n X_i^k / n, \quad k=1, 2, \dots$$

을 k 차 표본적률(k -th sample moment)이라고 한다. 적률법이라 함은 방정식 $\mu_k(\theta) = m_k$ 에 의해 원하는 모수를 추정하는 것이다.

다음 예는 위의 추정방법을 이용한 가장 대표적인 예이다. n 개의 확률변수 X, \dots, X_n 가 *iid* $p(x|\theta)$ 을 따르고 $\theta = E(X)$ 가 X 의 모평균일 때, 이 경우 1차 표본적률을 이용하면, θ 의 추정치는

$$\hat{\theta} = m_1 = \sum_{i=1}^n X_i / n = \bar{X}$$

으로 표본평균 $\hat{\theta} = \bar{X}$ 을 구할 수 있다. 즉 표본 평균의 추정으로 1차 표본적률을 이용하는 경우에 해당된다. 표본분산 추정시에는 1차 표본적률과 2차 표본적률을 동시에 이용하여 추정값을 얻어낼 수 있다. 즉 θ 가 X 의 모분산일 경우 아래와 같

이 표본분산을 구하기 위해 2차 표본적률을 이용하면,

$$\widehat{\Theta} = m_2 - m_1^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / n = s^2$$

으로 표본 분산 $\widehat{\Theta} = s^2$ 를 구할 수 있다.

3.2 제곱 파형요소(w_t^2)의 추정

컨볼루션의 가정 하에서 r_{it} 는 평균 0 분산 1인 서로 독립인 확률변수 이므로 위의 적률법(Method of moment)을 이용하면,

$$\text{var}(y_{it}) = \text{var}\left(\sum_{s=0}^t r_{i,t-s} w_s\right) = \sum_{s=0}^t w_s^2, \quad i=1, 2, \dots, n$$

이므로 이에 대응하는 샘플버전은,

$$\frac{1}{n} \sum_{i=1}^n y_{it}^2 = \sum_{s=0}^t \widehat{w}_s^2, \quad t=0, 1, 2, \dots, T$$

이 된다. 이를 \widehat{w}_t^2 에 대해 축차적으로 풀면, 다음과 같이 $w_0^2, w_1^2, \dots, w_T^2$ 의 추정값을 구할 수 있다.

$$t=0 \text{ 일 때, } \widehat{w}_0^2 = \frac{1}{n} \sum_{i=1}^n y_{i0}^2,$$

$$t=1 \text{ 일 때, } \widehat{w}_1^2 = \frac{1}{n} \sum_{i=1}^n y_{i1}^2 - \widehat{w}_0^2$$

⋮

$$t=T \text{ 일 때, } \widehat{w}_T^2 = \frac{1}{n} \sum_{i=1}^n y_{iT}^2 - \sum_{s=0}^{T-1} \widehat{w}_s^2$$

이 된다. 이를 일반화 시키면, 다음과 같이 축차적으로 파형요소 제곱 추정값 식을 얻을 수 있다.

$$\widehat{w}_t^2 = \frac{1}{n} \sum_{i=1}^n y_{it}^2 - \sum_{s=0}^{t-1} \widehat{w}_s^2, \quad i=1, 2, \dots, n, \quad t=0, 1, \dots, T$$

3.3 파형요소(w_t)의 추정

앞에서 언급한 적률법(Method of moment)을 사용하여 w_t^2 의 추정값을 구했다고 가정하자. 이를 이용하여 \widehat{w}_t 를 추정하는 법을 고안하고자 한다.

\widehat{w}_t^2 을 구하는 추정식은

$$\widehat{w}_t^2 = \frac{1}{n} \sum_{i=1}^n y_{it}^2 - \sum_{s=0}^{t-1} \widehat{w}_s^2$$

이며, $\frac{1}{n} \sum_{i=1}^n y_{it}^2$ 이 $\sum_{s=0}^{t-1} \widehat{w}_s^2$ 의 값 보다 작은 경우, \widehat{w}_t^2 은 음수값을 갖게 된다.

여기서 음수값을 갖는 추정값 \widehat{w}_t^2 에 대해서 생각해보자. 간단한 예로 $t=0$ 일 때를 들어보자. $t=0$ 이면 $\widehat{w}_0^2 = \frac{1}{n} \sum_{i=1}^n y_{i0}^2$ 이 되며, 이는 \widehat{w}_0^2 이 y_{i0} 의 분산인 셈이고 항상 양의 값을 갖는다. 그러나 $t=1$ 인 경우를 생각하면,

$$\widehat{w}_1^2 = \frac{1}{n} \sum_{i=1}^n y_{i1}^2 - \sum_{s=0}^{t-1} \widehat{w}_s^2$$

이고 만일 $\frac{1}{n} \sum_{i=1}^n y_{i1}^2 < \widehat{w}_0^2 = \frac{1}{n} \sum_{i=1}^n y_{i0}^2$ 이면 분산을 추정함에 있어 y_{i1} 의 분산에 해

당되는 \widehat{w}_1^2 이 음수값을 갖게 된다. 이와 같이 3.2에서 추정한 \widehat{w}_t^2 은

음수가 될 가능성이 존재한다. 그러나 분산은 항상 0보다 크거나 같은 값을 갖는다. 또한 우리는 보통 \widehat{w}_t^2 에서 \widehat{w}_t 을 추정하기 위해서는 \widehat{w}_t^2 의 제곱근을 사용한다. 그러나 위의 경우는 \widehat{w}_t^2 의 값이 음수이므로 이것의 제곱근을 사용한다면, \widehat{w}_t 은 허수를 갖게 될 것이다. 그러므로 추정된 \widehat{w}_t^2 의 값이 음수값일 경우를 고려해 줘야 한다.

본 연구에서는 이 문제를 다음과 같이 해결한다. 즉 \widehat{w}_t^2 의 제곱근을 구하기 위해서는 \widehat{w}_t^2 의 값이 $\widehat{w}_t^2 \geq 0$ 조건을 만족해야 한다. 그러므로 음의 값을 갖는 \widehat{w}_t^2 에 대해서 0이상의 값을 갖도록 인위적인 조작이 필요하다. 만약, 불가피하게 음수인 값을 양수인 값으로 대체를 해야 한다면, 0 이상의 값 중에서 0으로 추정하는 것이 가장 근사할 것이다. 그래서 다음과 같이 \widehat{w}_t^2 의 추정값 중 음수인 것에 대해서는 0으로 대체하였고, 새로 대체되어진 변수를 \widehat{w}_t^{2+} 로 표기하였다.

$$\widehat{w}_t^{2+} = \begin{cases} \widehat{w}_t^2 & \text{if } \widehat{w}_t^2 \geq 0 \\ 0 & \text{if } \widehat{w}_t^2 < 0 \end{cases}$$

다음은 \widehat{w}_t^{2+} 에서 \widehat{w}_t 을 구하는 과정이다. w_t^{2+} 의 추정값에서 \widehat{w}_t 의 값을 구하기 위해서는 일반적으로 제곱근을 많이 사용한다. 본 논문에서도 제곱근을 사용하여 w_t 의 추정값을 구하고자 한다. 그러나 어떤 제곱값의 제곱근도 \pm 인 어떤 수로 나오게 된다. 여기서 부호 선택의 문제가 생긴다. 두 가지의 부호 중 어떤 부호를 택일해야 할지에 대해 생각해보자. 본 연구에 서는 부호 선택의 문제를 다음과 같이 제시해 보았다. 우선, \widehat{w}_t^{2+} 의 제곱근 값의 절대값을 구한다. 여기서 이 값을 다음과 같이 \widehat{w}_t^+ 이라고 표기하기로 하였다.

$$\widehat{w}_t^+ = \sqrt{\alpha \widehat{w}_t^{2+}}$$

위와 같이 부호를 고려하지 않고 일괄적으로 w_t^{2+} 의 추정값의 제곱근은 방향을 가지고 있지 않는 크기로만 지정해준 후, 부호를 결정해주시기로 한다. 부호만 결정되면, 과형요소의 추정치가 결정되는 것이다.

다음은 \widehat{w}_t^+ 의 부호를 결정해주는 과정이다. 최종적으로 결정된 과형요소의 추정치는 \widehat{w}_t 으로 표장기로 하였다. 먼저, \widehat{w}_t^+ 의 부호 결정 과정은 $w_t^2 \approx 0$ 일 경우 w_t 역시 $w_t \approx 0$ 이 되며, 앞에서 가정한 것과 같이 과형요소는 연속인 진동함수라는 것에 근거를 두고 있다.

부호 결정을 시작하기 전에 처음 초기값이 양수인지 음수인지를 결정해줘야 한다. 본 논문에서는 초기값을 양수로 지정하고 있다. 부호 결정의 방법의 핵심 아이디어는 $t = t_0$ 에서 w_t 의 값이 0에 가까우면, t_0 을 기점으로 전후의 w_t 의 추정값의 부호가 반대가 된다는 것이다. 본 논문의 표기에 맞춰 \widehat{w}_t 의 추정방법 설명에 앞서 \widehat{w}_t^+ 의 값이 어떤 시점 t 에서 어떤 작은 양의 값 ε 보다 같거나 작아지는 것을 ε 조건이라 하자. 첫 번째 시점 $t = t_i$ 에서의 \widehat{w}_t^+ 의 값이 ε 조건을 만족하면, $\widehat{w}_{t_{i+1}}^+$ 의 부호를 $\widehat{w}_{t_{i-1}}^+$ 의 부호와 반대로 바꾸어 준다. 그리고 두 번째 \widehat{w}_t^+ 의 값이 ε 의 조건에 해당 되지 않는 경우에는 $\widehat{w}_{t_{i+1}}^+$ 의 부호는 \widehat{w}_t^+ 의 부호를 따라간다.

또한 세 번째 ε 조건에 해당하는 \widehat{w}_t^+ 의 값은 부호의 변화없이 \widehat{w}_t^+ 의 값이 \widehat{w}_t^+ 의 값으로 대체되며, ε 조건에 만족하는 \widehat{w}_t^+ , \widehat{w}_t^- 의 값의 부호는 참조되지 않는다. 즉, ε 보다 작은 값이 연속적으로 관측된 경우에는 한번 관측된 것으로 간주하게 된다.

예로서, 다음과 같이 추정값 \widehat{w}_t^+ 가 얻어졌다고 가정해보고, 파형요소 추정값인 \widehat{w}_t^- 의 부호를 결정해 보자.

t	0	1	2	3	4	5	6	7	8	9	10	11	...
\widehat{w}_t^+	0	0.13	0	0.13	0	5×10^{-4}	0.58	0.95	0.53	1.4×10^{-7}	0.47	0.71	...
\widehat{w}_t^-	0	0.13	0	-0.13	0	5×10^{-4}	0.58	0.95	0.53	1.4×10^{-7}	0.47	-0.71	...

전반부 시작값을 양수라 지정하고, ε 을 0.001 라 하자. 처음 $t = 0$ 일 때, $\widehat{w}_0^+ = 0$ 이므로 0.001보다 작다. 그러므로 $t = 1$ 일 때, \widehat{w}_1^+ 의 값은 \widehat{w}_1^+ 의 값과 반대 부호를 취해야 하나, \widehat{w}_1^+ 에 대한 정보는 없다. 그러므로 앞에서 전반부 시작값을 지정해야 하는 것이며, 본 연구에서는 이것을 양수로 지정하였다. 그러므로 $\widehat{w}_1^- = 0.13$ 이 된다. \widehat{w}_2^- 의 부호는 \widehat{w}_1^- 의 값이 0.001보다 크므로 \widehat{w}_1^- 과 같은 양의 방향을 유지하여, $\widehat{w}_2^- = 0$ 이 된다. $t = 2$ 일 때, $\widehat{w}_2^- = 0$ 으로, 0.001보다 작다. 그러므로 $t = 3$ 일 때의 값 $\widehat{w}_3^- = 0.13$ 은 \widehat{w}_1^- 의 방향과 반대인 음으로 바꾸어 $\widehat{w}_3^- = -0.13$ 으로 만들어 준다. \widehat{w}_4^- 의 값은 \widehat{w}_3^- 의 값이 0.001보다 크므로 \widehat{w}_3^- 의 부호인 음수를 유지한다. 다음 $t=4$, $t=5$ 의 특별한 경우를 주의하여 살펴보자. 이 경우의 \widehat{w}_t^- 값이 ε 값보다 작은 값이 연속으로 나오는 경우이다. \widehat{w}_4^- 의 값이 0.001보다 작으므로 \widehat{w}_5^- 의 값은 \widehat{w}_3^- 의 부호와 반대인 양수가 되어야 한다. 그러나 \widehat{w}_5^- 의 값 역시 0.001보다 작은 값을 가지므로 \widehat{w}_5^- 의 값에 부호변화를 시키지 않는다. 그러므로 $\widehat{w}_4^- = 0$ 이고 $\widehat{w}_5^- = 5 \times 10^{-4}$ 가 된다. 이렇게 연속적으로 ε 보다 작은 값이 관측될 때, 다음 관측값의 부호결정에 주의하자. \widehat{w}_5^- 이 0.001보다 작은 값을 가지므로 \widehat{w}_6^- 의 부호는 \widehat{w}_4^- 의 부호를 참조하여 결정되나, \widehat{w}_4^- 이 ε 보다 작은 값을 가졌으므로, 부호결정에 참조되지 않는다. 그러므로 \widehat{w}_6^- 의 부호는 \widehat{w}_3^- 의 부호에 의해 결정되어진다. 따라서

\widehat{w}_3 의 값은 다음과 같이 -0.58 이 아닌 0.58 이 되는 것이다. 결국, 연속적으로 ε 값보다 작은 값이 관측될 경우 한번 관측된 결로 간주하게 되는 것이다. $t=7, 8, 9$ 까지는 $\widehat{w}_6, \widehat{w}_7, \widehat{w}_8$ 의 값들이 0.001 보다 작거나 같은 경우가 없으므로 그에 해당되는 $\widehat{w}_7, \widehat{w}_8, \widehat{w}_9$ 의 부호는 바뀌지 않으며, 각각 $0.95, 0.53, 1.4 \times 10^{-7}$ 의 값을 갖는다. 그러나 $t=9$ 일 때 $\widehat{w}_9^+ = 1.4 \times 10^{-7}$ 로 0.001 보다 작은 값을 갖는다. 그러므로 \widehat{w}_9 의 값은 \widehat{w}_8 의 값의 부호와 반대로 -0.47 이 된다. $t=11$ 일 때는 \widehat{w}_{10} 의 값이 0.47 로 0.001 보다 크므로 \widehat{w}_{10} 의 부호를 따라 -0.71 이 된다.

4. 모의 실험

이 장에서는 3장에서 제시된 파형요소 추정방법의 통계적 성질을 모의실험을 통해 연구한다. 반사계수의 분포가 정규분포, t-분포, 이중지수분포일 때를 가정하고, T_0 를 파형요소이 소멸되는 최초의 시점이라 할 때, T_0 를 알고 있을 경우와 모르고 있을 경우를 나누어 n 의 개수와 T 의 횟수를 몇 가지로 바꿔가며 100번의 반복 실험을 해보았다.

4.1 실험의 개요

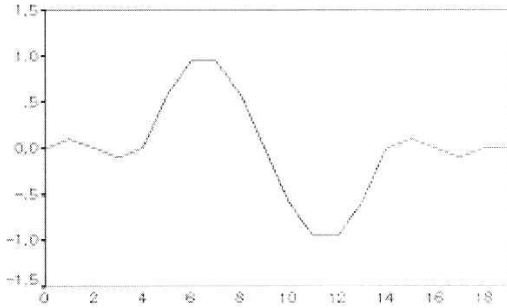
다음과 같이 반사계수와 파형요소이 컨볼루션된 자료 y_{it} 가 얻어지는 경우를 생각한다.

$$y_{it} = r * w_t \quad i=1, 2, \dots, n \quad t=0, 1, \dots, T$$

여기서 r_{it} 는 평균이 0인 iid오차이다. 또 파형요소으로는

$$w_t = \begin{cases} 0.1 \sin\left(\frac{2\pi}{4} t\right), & t=0, 1, 2, 3, 4 \\ \sin\left(\frac{2\pi}{10}(t-4)\right), & t=5, \dots, 14 \\ 0.1 \sin\left(\frac{2\pi}{4}(4-14)\right), & t=15, 16, 17, 18 \end{cases}$$

을 고려한다. 이 파형요소의 그래프가 아래 그림에 나타나있다. 이는 탄성과 실험에서 나타나는 파형요소의 전형적인 한 형태이다.



[파형요소 (w_t)]

다음의 조합에 대해 모의실험을 하고자 한다. :

T_0 : 알려짐(known), 모름(unknown) ;

r_{it} 의 분포 : 표준정규분포($\mathcal{N}(0, 1)$), 자유도 3인 t -분포($t(3)$),
이중지수분포(Double exponential) ;

$T = 50, 100$;

$n = 20, 50$.

여기서 T_0 는 w_t 가 이 시점 이후에 소멸되는 최초의 t 값 즉 $w_t = 0, t \geq T_0$ 이다. 위 그림에서 $T_0 = 19$ 이다. 위의 조합에 대해서 파형요소 추정을 100번 반복했을 때의 추정치 $\{\widehat{w}_t, t=0, 1, \dots, T\}$ 100개의 평균값과 100개의 추정값 중 무작위로 한 개를 골라 실선으로 그렸으며, 참값 w_t 은 점선으로 그렸다.

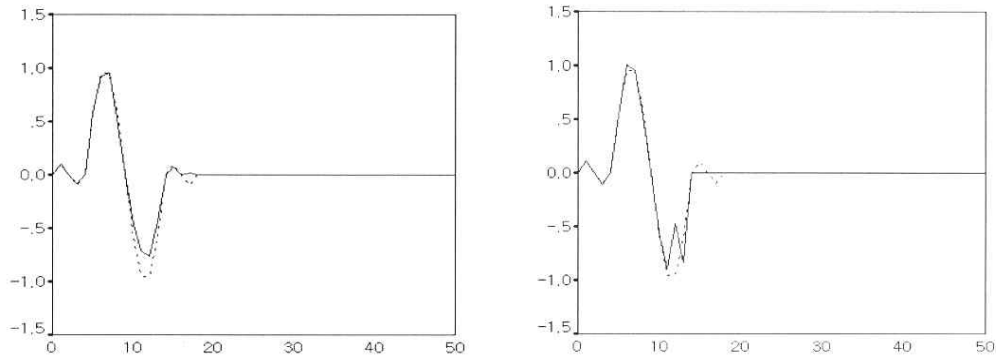
4.2 T_0 가 알려졌을 때

먼저 $T_0 = 19$ 를 알고 있을 때의 추정을 논한다. 이 경우 $t \geq 19$ 일 때 $w_t = 0$ 임을 추정시 이용하게 된다.

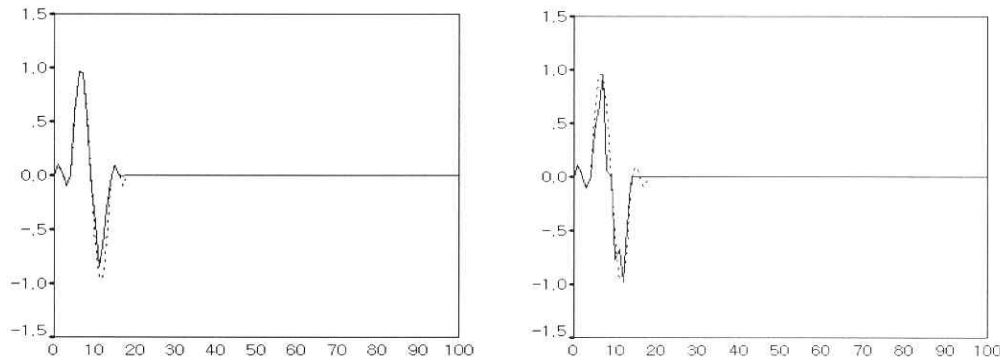
1) 반사계수의 분포가 표준정규분포인 경우($n=20$)

이 경우 $\{\widehat{w}_t, t=0, 1, \dots, T\}$ 의 100개 평균값과 100개 중 랜덤하게 선택된 어떤 하나 추정값이 $T=50, 100$ 에 대해 각각 [그림1]-[그림2]에 도시되어 있다. 아래 결과에서 [그림1]-[그림2]를 비교해 보면, $T=50$ 일 때와 $T=100$ 일 때 모두 참값에 근사하는 정도가 비슷하다. $T=50, 100$ 일 때의 추정값의 평균이 나타난 두 그림은 모두 $T=10$ 일 때 까지는 참값과 아주 근사하게 움직이고 있으나, $T=10$ 이후부터는 몇몇 시점에서 참값을 잘 맞추지 못하고 있다. 그리고 $T=17$ 쯤에 와서는 거의

추정값이 0으로 사라지고 있다. 이는 오른쪽에 개별 추정치의 그림을 보면 좀 더 명확히 나타난다. 개별 추정치가 $T=10$ 이후부터 불안정해지기 시작한다. 급기야 $T=15$ 에서부터는 추정값이 0으로 사라지고 있다.



[그림 1] $N(0, 1)$ 반사계수, ($n=20$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

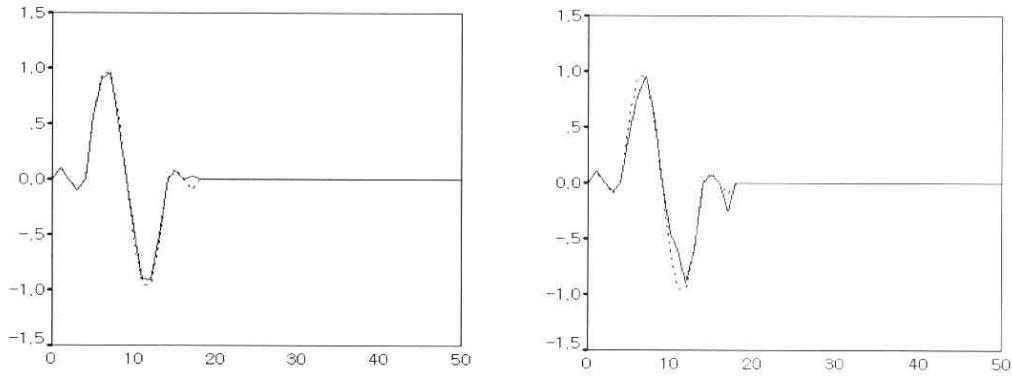


[그림 2] $N(0, 1)$ 반사계수, ($n=20$: $T=100$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

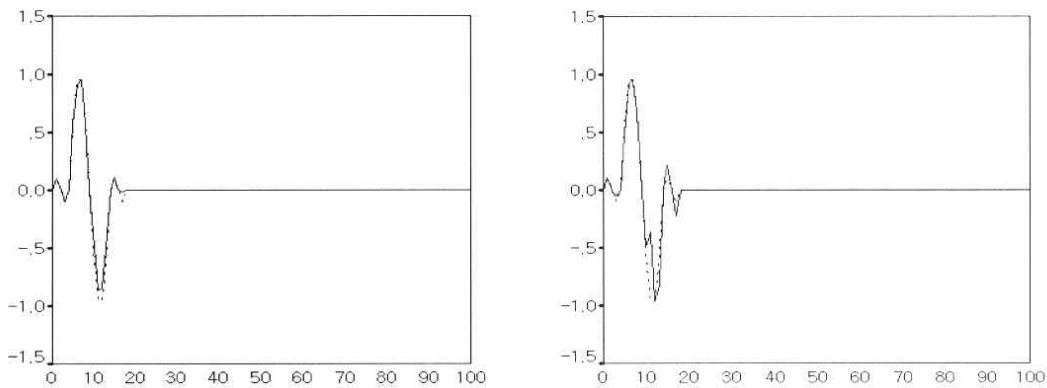
2) 반사계수의 분포가 표준정규분포인 경우($n=50$)

이 경우 추정 결과가 [그림3]-[그림4]에 나타나 있다. 여기서 $n=50$ 일 때도 $n=20$ 일 때와 마찬가지로 $T=17$ 쯤에 와서는 추정값이 참값에 잘 근사되지 않고 있다. 그러나 $T=17$ 이후의 값이 [그림1]-[그림2]에서 처럼 0으로 근사되지는 않으며, 부호의 방향이 다를 뿐 참값에 근사한 크기를 유지하고 있다. 이는 개별 추정치의 그림에 더 명확히 나타난다. 개별 추정치의 그림을 보면 전반적으로 참값에 근사

하며, 부호의 방향도 잘 맞고 있다. 또한 $T=10$ 이후부터 불안정한 모습을 보여준 [그림1]-[그림2]와는 달리 보다 안정된 모습을 보이고 있다. $T=15$ 이후의 추정값 역시 0으로 사라지지 않고 어느 정도 참값의 부호 방향을 맞춰가며 유지되고 있다.



[그림 3] $N(0, 1)$ 반사계수, ($n=50$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)



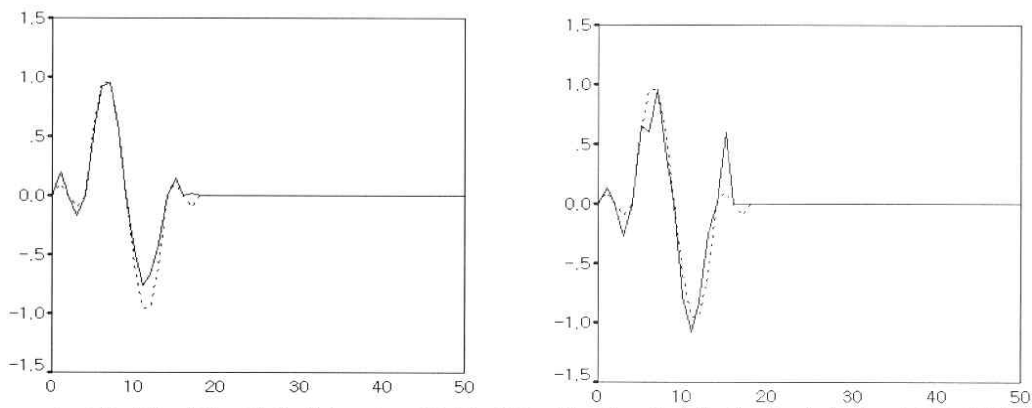
[그림 4] $N(0, 1)$ 반사계수, ($n=50$: $T=100$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

[그림1]-[그림2]와 [그림3]-[그림4]를 비교해 보면 위의 그림과 같이 $n=20$ 인 경우보다 $n=50$ 인 경우에 파형요소의 참값이 더 잘 추정되며, $T=17, 18$ 부분의 부호 방향도 $n=20$ 일 때 보다 $n=50$ 일 때가 더 잘 맞는 것을 알 수 있다. 또 $n=20$ 인 경우와 $n=50$ 인 경우의 개별 추정치 그림을 비교해보면, $n=50$ 인 경우의 그림이 $n=20$ 인 경우의 그림보다 값이 많이 안정되었음을 볼 수 있으며, $T=17, 18$ 부분에서도

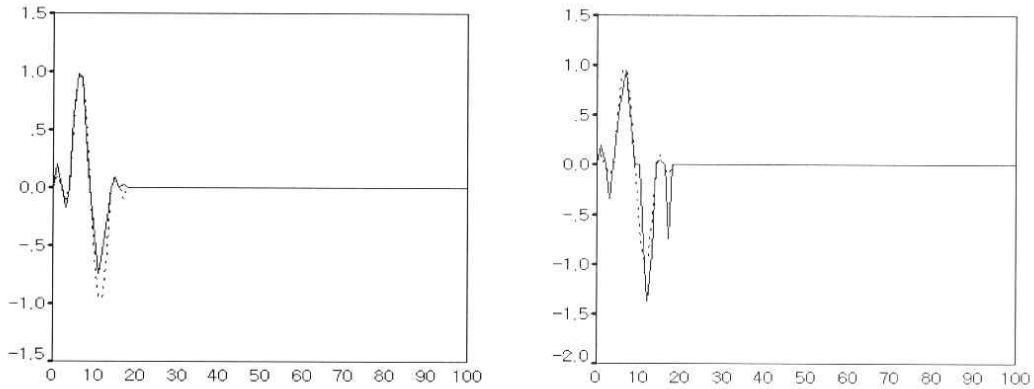
이는 출력신호 즉 관측값의 정보가 많을수록 파형요소의 추정값의 정확도가 높아질 수 있다는 것을 알려준다. 또한 평균은 파형요소의 참값과 비슷하게 매끄러운 굴곡으로 부호방향도 잘 맞추고 있다. 그러나 개별 추정치는 두 경우 모두 $T=15$ 전부터 불안정해지는 것을 볼 수 있으며, $n=20$ 인 경우는 $T=15$ 에서부터 파형요소의 참값에 근사하지 못하고 있다.

3) 반사계수의 분포가 $t(3)$ 인 경우($n=20$)

이 경우 추정결과가 [그림5]-[그림6]에 나타나 있다. 이를 보면 t -분포일 경우 파형요소의 추정값은 피크점에서 잘 맞지 않음을 알 수 있다. 그리고 개별 추정치 역시 매끄럽지 못하며, 불안정함을 볼 수 있다. 특히 개별 추정치 그림의 $T=50$ 인 경우는 $T=15$ 에서 $T=17$ 까지의 추정값이 참값보다 너무 크게 추정되었으나, 부호의 방향은 잘 맞음을 볼 수 있다. 그러나 $T=17$ 이후부터 추정값이 0으로 사라지고 있다. 이에 반해 $T=100$ 인 경우의 개별 추정치는 $T=16$ 까지는 잘 맞으나 $T=17$ 부터는 값이 많이 떨어짐을 볼 수 있다. 그럼에도 불구하고 $T=100$ 일 때의 $T=17$ 부분의 평균값이 0근처로 근사되는 것을 보면, 개별 추정치들이 $T=17$ 부터 상당히 들쭉날쭉하게 0사이를 왔다 갔다 하는 것으로 보여진다.



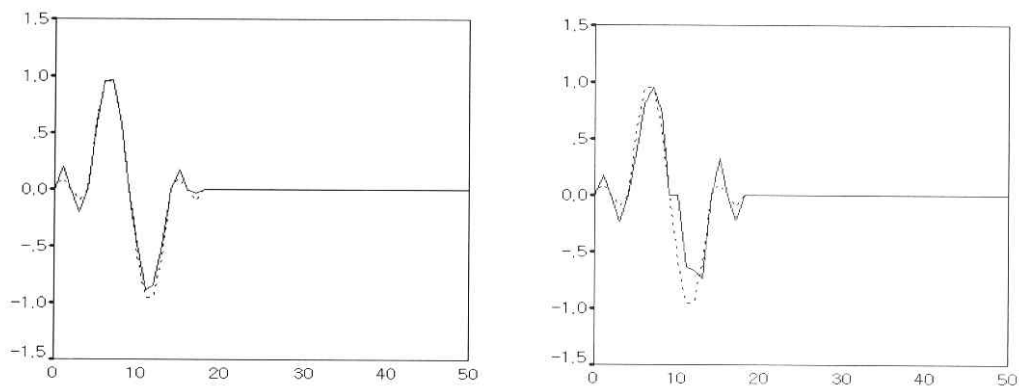
[그림 5] $t(3)$ 반사계수, ($n=20$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)



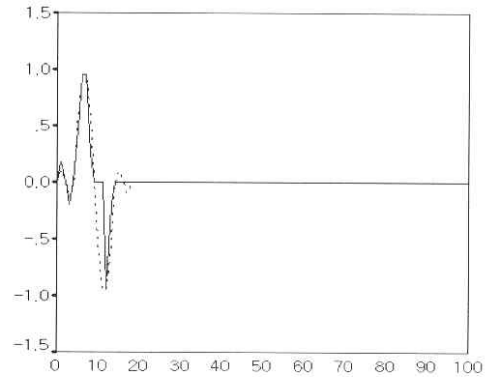
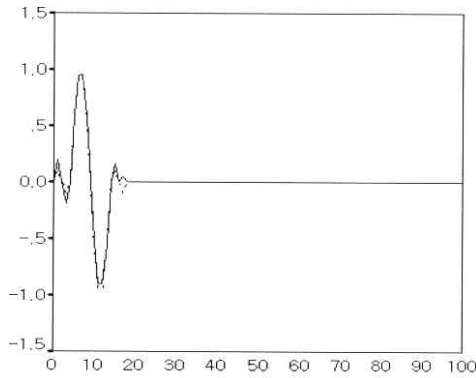
[그림 6] $t(3)$ 반사계수, ($n=20$: $T=100$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

4) 반사계수의 분포가 $t(3)$ 인 경우($n=50$)

이 경우 추정 결과가 [그림7]-[그림8]에 나타나있다. 이를 보면 $T=50$ 일 때의 평균값의 그림을 보면 파형요소의 추정값이 참값에 잘 근사되며, 부호 방향 또한 잘 맞는 것을 볼 수 있다. 그러나 여전히 피크점에서는 근사값이 참값보다 큰 경우가 종종 보인다. 그러나 $T=100$ 일 때의 추정값의 평균값은 $T=17$ 이후부터의 방향을 잘 맞추지 못하고 있다. [그림7]-[그림8]의 개별 추정치 그림을 보면 $T=50$ 인 경우 방향을 잘 맞추고 있으나 참값의 근사가 잘 이루어지지 않고 있다. $T=100$ 인 경우의 참값 근사는 $T=50$ 일 경우보다 전반적으로 잘 근사되고 있으나, $T=15$ 이후의 추정값이 0으로 사라지고 있다.



[그림 7] $t(3)$ 반사계수, ($n=50$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

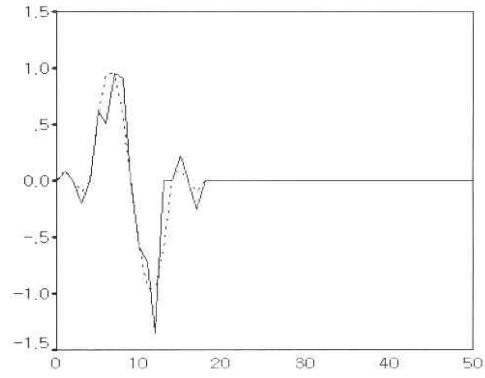
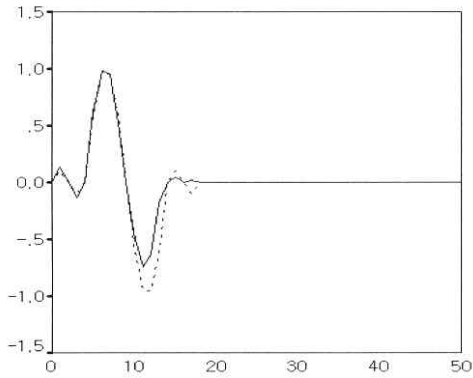


[그림 8] $t(3)$ 반사계수, ($n=50$: $T=100$) 일 때 과형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

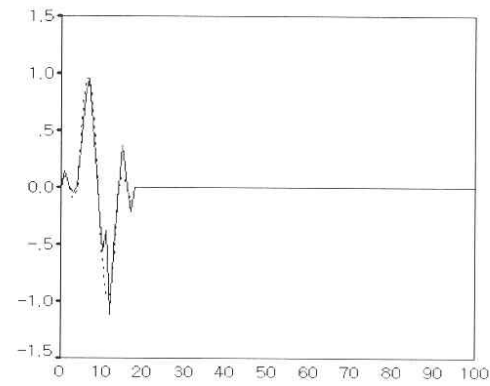
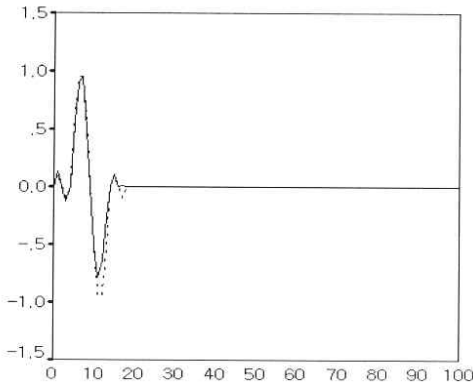
[그림5] - [그림6]와 [그림7] - [그림8]를 비교해 보면 t -분포 역시 정규분포 일 때와 마찬가지로 $n=20$ 인 경우보다 $n=50$ 인 경우에 더 과형요소의 참값에 잘 맞는다. 과형요소 추정값의 평균 그림을 보면 $T=17, 18$ 부분의 부호 방향도 $n=20$ 일 때 보다 더 잘 맞는 것으로 나왔다. 또한, 정규분포인 경우와 마찬가지로 전반적으로 매끄럽게 잘 맞는다. 그러나 개별 추정치는 전반적으로 정규분포인 경우보다 추정값이 매끄럽지 않으며, 불안정하다.

5) 반사계수의 분포가 이중지수분포인 경우($n=20$)

이 경우 추정 결과가 [그림9]-[그림10]에 나타나있다. 이중지수분포의 경우 역시 위의 두 개의 분포와 마찬가지로 과형요소 추정값의 평균은 참값에 잘 근사되나, $T=17$ 부터는 0으로 근사되고 있다. 그러나 개별 추정치의 그림을 보면 위의 두 분포보다 상당히 불안정함을 알 수 있다. 또한 $T=100$ 일 때 보다 $T=50$ 일 때가 더 심하게 불안정함을 볼 수 있다. 이는 추정값이 피크값 근처로 갈수록 불안정한 정도가 커지는 것으로 보여진다. 그에 반해 평균값이 매끄럽고 안정정인 것은 개별 추정치가 참값의 근처에서 불안정한 형태를 보이고 있다고 보여진다.



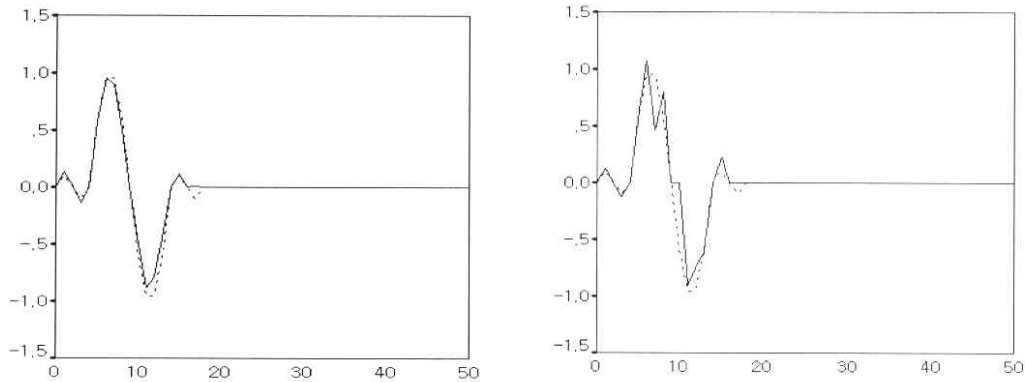
[그림 9] 이중지수 반사계수, ($n=20$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)



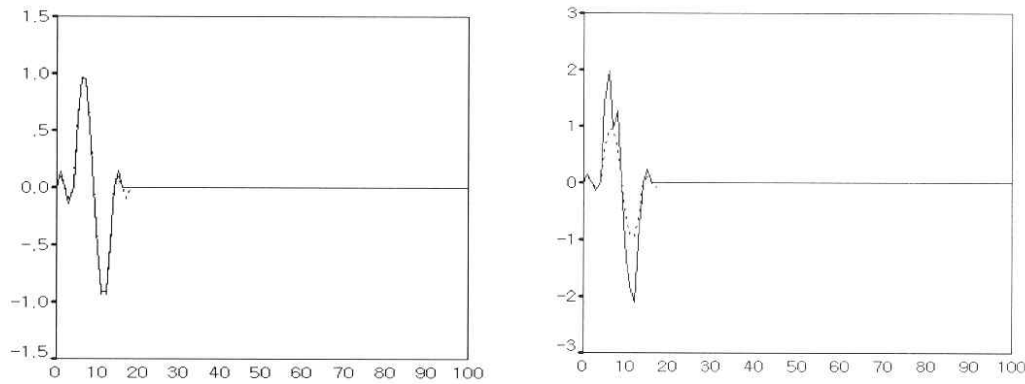
[그림 10] 이중지수 반사계수, ($n=20$: $T=100$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

6) 반사계수의 분포가 이중지수분포인 경우($n=50$)

이 경우 추정 결과가 [그림11]-[그림12]에 나타나있다. 여기서도 위 경우와 마찬가지로 $T=16$ 까지는 파형요소의 추정값이 참값에 잘 근사하고 있다. 특히 $T=100$ 인 경우는 추정값들의 평균은 참값과 거의 일치하는듯 보여 진다. 그러나 여전히 $T=17$ 부터는 참값을 맞추지 못하고 있다. 그리고 개별 추정치의 그림을 보면 더 불안정해짐이 심해짐을 알 수 있다. 또한 피크값이 잘 근사되고 있지 않으며 $T=17$ 부터 0으로 근사되고 있다.



[그림 11] 이중지수 반사계수, ($n=50$: $T=50$) 일 때 과형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)



[그림 12] 이중지수 반사계수, ($n=50$: $T=50$) 일 때 과형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

[그림9]-[그림10]와 [그림11]-[그림12]를 비교해 보면 이중지수분포인 경우도 $n=20$ 인 경우보다 $n=50$ 인 경우에 더 과형요소의 참값에 잘 근사한다. 그러나 정규분포와 t -분포와는 달리 $T=17, 18$ 부분부터는 전혀 참값을 맞추지 못하고 과형요소의 추정값이 사라진다. 관측값을 살펴보면, t -분포보다 더 불안정한 모양을 하고 있음을 볼 수 있다.

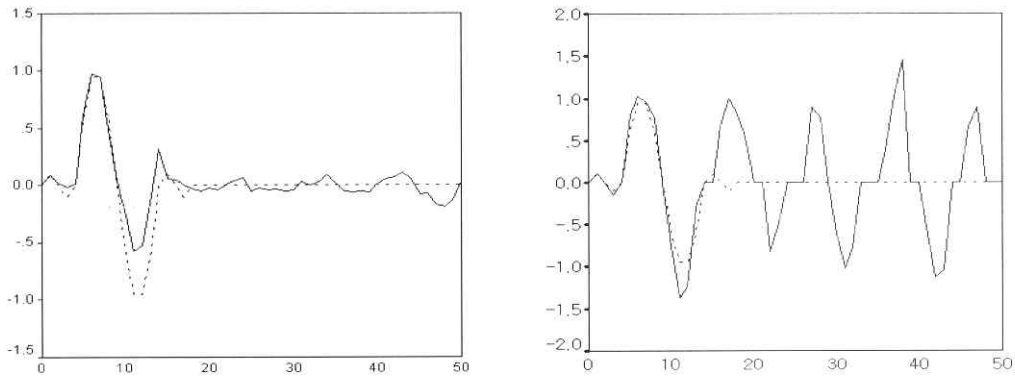
4.3 T_0 를 모를 때

여기서는 T_0 를 모를 때의 추정을 논한다. 각 분포에 대해 $n=50$ 으로 통제하고

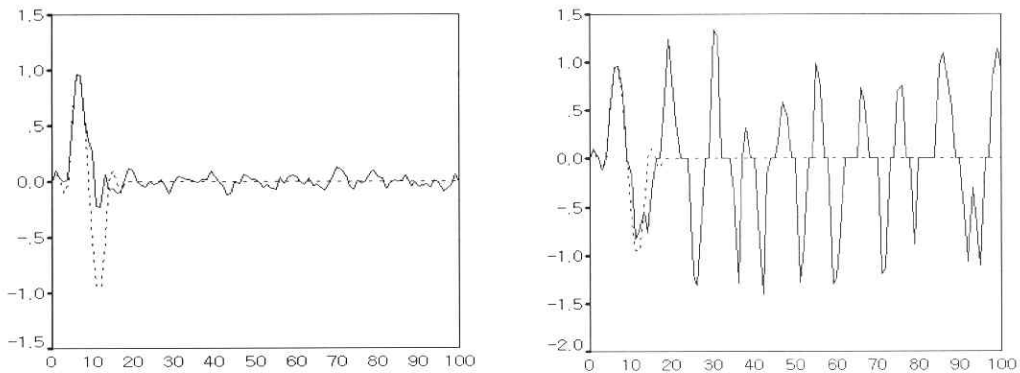
T=50, 100 일 때를 비교한다.

1) 반사계수의 분포가 정규분포인 경우

이 경우 $\{\widehat{w}_t, t=0,1,\dots,T\}$ 의 100개 평균값과 100개 중 랜덤하게 선택된 어떤 하나 추정값이 T=50, 100에 대해 각각 [그림13]-[그림14]에 도시되어 있다. 아래 [그림13]-[그림14]의 과형요소 추정값의 평균을 보면, T=50, 100 인 경우 모두 참값의 방향을 잘 맞추고 있다. 그러나 추정값의 근사는 T=10인 부분부터 잘 맞추지 못하고 있다. 또한 T=50 일 때가 T=100 일 때 보다 참값에 더 잘 근사되고 있다. T가 알려지지 않을 때는 T가 작을수록 과형요소를 추정하는데 더 유리하다는 것을 알 수 있다. 개별 추정치를 보면 T=16이후부터 높은 값과 낮은 값을 오가며 심하게 들쭉날쭉하다. 그러나 이 개별 추정치들의 평균 그림을 보면 과형요소의 추정값들이 0근처로 모이고 있음을 볼 수 있다.



[그림 13] N(0, 1) 반사계수, (n=50: T=50) 일 때 과형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

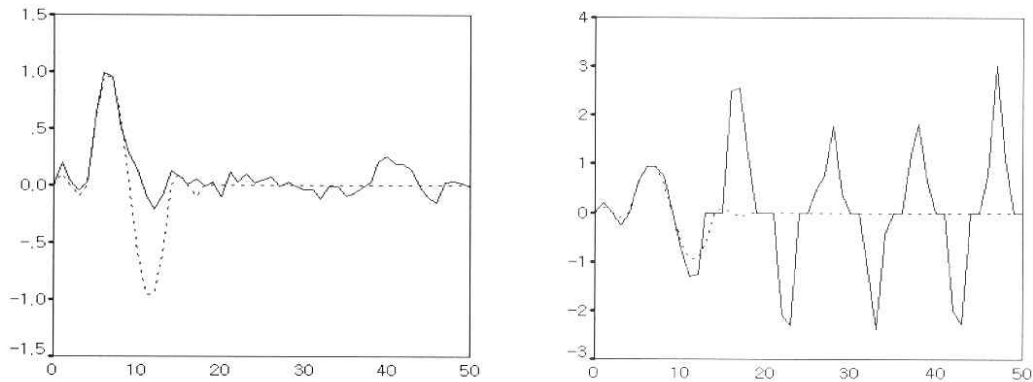


[그림 14] N(0, 1) 반사계수, (n=50: T=100) 일 때 과형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

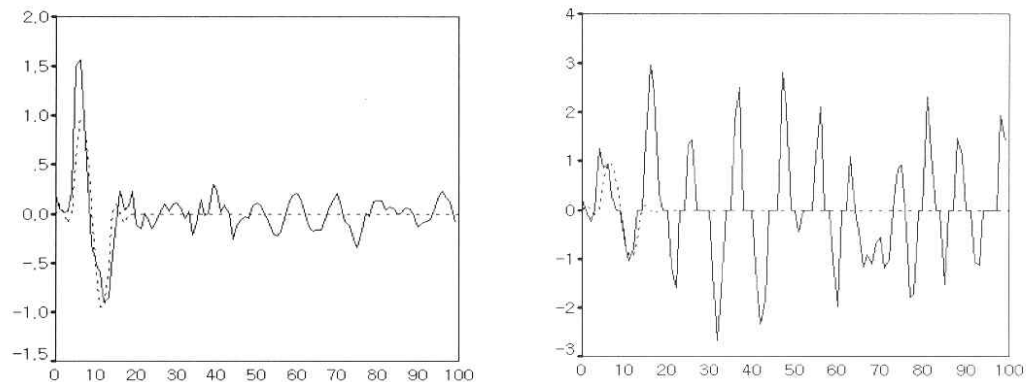
쪽)과 개별 추정치(오른쪽)

2) 반사계수의 분포가 t-분포인 경우

이 경우 추정 결과가 [그림15]-[그림16]에 나타나있다. t-분포의 경우도 마찬가지로 정규분포와 비교해 피크값에서 근사가 잘 안되고 있음을 볼 수 있다. 그리고 $T=17$ 이후의 개별 추정치의 근사값도 불안정하며, 평균을 보면 정규분포일 때보다 더 심하게 0을 중심으로 들쭉날쭉하다. 평균값을 보면 $T=50$ 일 때 보다 $T=100$ 일 때가 참값의 방향에 더 잘 근사하는 것으로 나왔으나, 추정값의 근사율은 $T=50, 100$ 인 두 경우 모두 비슷하게 나타난다. 그러나 $T=100$ 인 경우가 개별 추정치를 보았을 때, 가장 큰 값과 가장 작은 값의 차이가 상당히 크며, 아주 불안정함을 알 수 있다.



[그림 15] $t(3)$ 반사계수, ($n=50$: $T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)

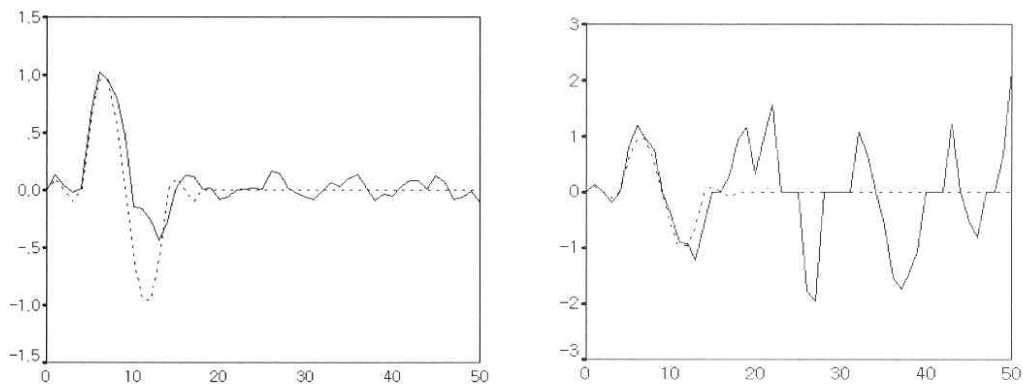


[그림 16] $t(3)$ 반사계수, ($n=50$: $T=100$) 일 때 파형요소 추정값의 평균(왼쪽)

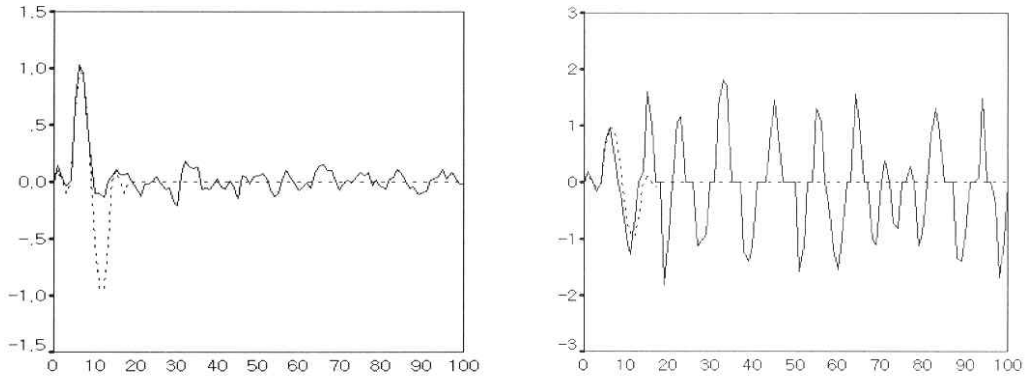
과 개별 추정치(오른쪽)

3) 반사계수의 분포가 이중지수분포인 경우

이 경우의 결과는 아래 [그림18]-[그림19]에 나타나있다. 아래와 같이 이중지수 분포일 때의 평균 그림에서 양의 값의 피크점에서 잘 맞고 있으나 음의 값의 피크점은 거의 맞추지 못하고 있음을 볼 수 있다. 반면, 개별 추정치에서는 양의 값 음의 값 각각의 피크점에 근사하게 맞고 있다. 이는 음의 값 피크점에서 개별 추정치들이 불안정하게 퍼져있음을 알 수 있다. 또한 $T=50, 100$ 이 모두 참값의 방향을 잘 유지하고 있다. 그러나 $T=50$ 인 경우와 $T=100$ 인 경우보다 참값에 더 잘 근사되고 있는 것으로 보인다.



[그림 18] 이중지수 반사계수, ($n=50: T=50$) 일 때 파형요소 추정값의 평균(왼쪽)과 개별 추정치(오른쪽)



[그림 19] 이중지수 반사계수, ($n=50$: $T=100$) 일 때 파형요소 추정값의 평균 (왼쪽)과 개별 추정치(오른쪽)

5. 결론

4장에서 다뤄진 모의 실험을 요약해보면 다음과 같다. 반사계수와 파형요소가 컨볼루션되어 얻어진 관측값으로 반사계수가 주어졌다는 가정하에 파형요소를 추정해 보았다. 위 실험은 반사계수가 정규분포, t-분포, 이중지수 분포인 경우를 각각 가정하였으며, 시간에 따라 얻어지는 파형요소의 소멸시점(T_0)를 알고 있는 경우와 모르고 있는 경우를 가정하였다. 또한, 얻어진 관측값의 개수(n)를 다르게 하면서 실험을 해보았다.

첫 번째 T_0 가 알려졌다고 가정한 경우는 (n, T)의 조합(20, 50), (20, 100), (50, 50), (50, 100)에 대해 실험해 보았다. 실험결과 T 의 시점 50, 100과는 무관하게 오로지 관측값의 개수, n 에 따라 추정값의 근사율의 차이를 보였다. 본 실험의 경우 세 가지 분포에서 모두 $n=20$ 일 때 보다 $n=50$ 일 때가 참값에 더 잘 근사하는 추정값을 얻었다. 이는 샘플 수가 많은 관측값은 파형요소의 추정이 더 유리함을 보여 주고 있다. 또한 T_0 가 알려진 경우에는 T 에 따른 관측값의 개수가 추정값의 근사정도에 영향을 미치지 않음을 알 수 있다.

두 번째 실험에서는 T_0 를 모르고 있다는 가정 하에 (n, T)의 조합 (50, 50), (50, 100)에 대해 실험해 보았다. 위의 첫 번째 실험 결과에서 n 이 클 수록 더 좋

은 추정값을 얻을 수 있다는 사실을 알았으므로, n 의 개수를 50으로 통제하고 T 에 따른 추정값의 근사율 차이가 있는지를 알아보았다. 본 실험의 경우 세 가지 분포 모두 $T=100$ 인 경우보다 $T=50$ 인 경우에 참값에 더 잘 근사하는 추정값을 얻을 수 있었다. 이는 T 에 대한 정보가 많다고 해서 과형요소의 추정에 좋은 영향을 주는 것은 아니라 T 로 인한 관측값의 증가는 오히려 과형요소의 불확실성만 증가시키게 되기 때문인 것 같다. 과형요소의 추정에 있어서 관측값의 데이터는 횡적인 크기보다 종적인 크기가 추정의 근사율을 좋게 하는데 더 많은 영향을 미침을 알 수 있다.

주 의

1. 이 보고서는 산업자원부에서 시행한 국가지질조사 및 자원탐사 사업의 연구보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 산업자원부에서 시행한 국가지질조사 및 자원탐사의 연구결과임을 밝혀야 합니다.
3. 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니 됩니다.