

## Gauss-Newton 방법에서의 유사 Hessian 행렬의 구축과 이를 이용한 파형역산

하태영<sup>1)</sup>

### Construction the pseudo-Hessian matrix in Gauss-Newton Method and Seismic Waveform Inversion

Taeyoung Ha<sup>1)</sup>

**요 약 :** 탄성과 역산에는 고전적인 Gauss-Newton 방법이 주로 사용된다. 이 방법은 Jacobian을 직접 계산하여 거대한 크기의 Hessian 행렬을 만드는 것을 필요로 한다. Hessian 행렬의 구성은 몇 가지의 요소들에 의해 결정되는데, 음원과 수신기의 위치, 영상화 구역(image zone), 음원 파형의 형태 등 다양한 형태의 모델링에 영향을 미치는 요소에 따라서 다른 모습으로 나타난다. 이 논문에서는 Gauss-Newton 방법에 나타나는 거대한 Hessian 행렬을 조절함으로써 Marmousi 탄성과 모델 자료를 역산하고자 한다. 또한 근사 Hessian 행렬의 대안으로 두 가지의 유사 Hessian 행렬들을 제시하고자 한다. 하나는 유한 폭을 갖는 Hessian 행렬이고 다른 하나는 자동안정함수(automatic gain function, AGC)를 이용한 Hessian 행렬이다. 작은 크기의 모델에 대한 수치결과로부터 몇 가지의 사실을 알 수 있다. 하나는 유한 폭을 갖는 Hessian 행렬을 이용하여 얻어진 한번 근사된 속도모델은 원래의 Hessian 행렬을 이용하여 얻은 결과와 매우 유사하다는 것이고, 둘째로 자동안정함수를 이용한 근사 Hessian 행렬의 안정성이 많이 개선된다는 것이다.

**주요어 :** Gauss-Newton 방법, 유사 Hessian 행렬, 파형역산

**Abstract :** Seismic waveform inversion can be solved by using the classical Gauss-Newton method, which needs to construct the huge Hessian by the directly computed Jacobian. The property of Hessian mainly depends upon a source and receiver aperture, a velocity model, an illumination zone and a frequency content of source wavelet. In this paper, we try to invert the Marmousi seismic data by controlling the huge Hessian appearing in the Gauss-Newton method. We make the two kinds of the approximate Hessian. One is the banded Hessian and the other is the approximate Hessian with automatic gain function. One is that the 1st updated velocity model from the banded Hessian is nearly the same of the result from the full approximate Hessian. The other is that the stability using the automatic gain function is more improved than that without automatic gain control.

**Keywords :** Gauss-Newton method, pseudo-Hessian matrix, waveform inversion

## 서 론

수년 동안 많은 응용수학자와 지구물리학자들은 파형역산을 다루어 왔다. 파형역산의 목적은 지진파탐사로부터 얻은 정보를 분석하여 보다 정확한 지질구조를 얻고자하는 것이다. 이 연구의 중요한 발전은 Lailly(1983)와 Tarantola(1984)에 의하여 제안된 이론에 바탕을 두는 것으로 data residual들을 역전 파하고 순방향으로 전파된 파형들을 서로 연관시키면서 반복적으로 더 나은 영상을 얻어가는 방법이다. 이 연구에 대한 시공간에서의 수치해석적인 결과는 Kolb *et al.*(1986), Gauthier *et al.*(1986) 등에 의해 제시되었고, 주파수공간에서의 결과는 Geller and Hara(1993), Pratt *et al.*(1998), Pratt(1999) 등에 의해 제시되었다. 그 중 주파수영역에서의 연구들 중에서 Pratt

*et al.*(1998)은 Jacobian 행렬을 이용한 탄성파역산을 다루었고 Shin *et al.*(2001b)은 파동방정식의 상호관계를 이용하여 거대한 Jacobian행렬을 효율적으로 구하는 방법을 제시하였다. 또한 Shin *et al.*(2001a)은 거대한 Hessian 행렬로 구조보정 영상( $J^T J$ )를 정규화 하는 대신에 단지 근사행렬의 대각행렬만을 이용하여 구조보정 영상의 정규화를 시도하였다.

이번 연구에서 우리는 두 가지 종류의 유사 Hessian 행렬을 제시하고 Marmousi 모형에 대하여 검증해 보았다. 두 가지의 행렬 중 하나는 대각선을 중심으로 일정한 폭을 갖는 행렬이고 다른 하나는 AGC를 이용하여 만들어진 행렬이다. 우리는 이들을 이용한 파형역산의 결과들을 제시하고 이 방법의 장점과 단점을 제시하고자 한다.

\*2004년 7월 8일 접수

1) 서울대학교 수리과학부

### 파형역산이론

유한차분법 또는 유한요소법을 이용하여 시간영역에서의 파동방정식을 행렬형태로 나타내면 다음과 같이 표현된다.

$$\mathbf{M} \frac{\partial^2 \mathbf{u}(t)}{\partial t^2} + \mathbf{K} \mathbf{u}(t) = \mathbf{f}(t), \tag{1}$$

여기서  $\mathbf{M}$ 은 질량(mass)행렬,  $\mathbf{K}$ 는 강성(stiffness)행렬, 그리고  $\mathbf{f}$ 는 송신원(source)항이다. 이를 주파수영역으로 푸리에변환을 하면 (2)식을 얻을 수 있다.

$$\mathbf{K} \tilde{\mathbf{u}}(\omega) - \omega^2 \mathbf{M} \tilde{\mathbf{u}}(\omega) = \mathbf{f}(\omega). \tag{2}$$

편의상 (2)식을 다음과 같이 쓰기로 한다.

$$\mathbf{S} \tilde{\mathbf{u}}(\omega) = \mathbf{f}(\omega) \text{ 또는 } \tilde{\mathbf{u}} = \mathbf{S}^{-1} \mathbf{f}(\omega), \tag{3}$$

여기서  $\mathbf{S}$ 는 complex impedance matrix라고 부르고  $\mathbf{K} - \omega^2 \mathbf{M}$ 으로 주어진다.

수신기가 놓여있는 격자점에서 얻은 측정값 벡터를  $\mathbf{d}$ 라고 가정하자. 모델 알고리즘으로 예측하고자 하는 모델변수를  $\mathbf{p}$ 라 하자. 이때, 초기모델변수를  $\mathbf{p}^0$ 라고 하면, (3)식을 이용하여  $\tilde{\mathbf{u}}(\omega)$ 를 구할 수 있다. 이때,  $n$ 개의 받는 위치에서의 나머지 오차  $\delta \mathbf{d}$ 라 하고  $i$ 번째 성분을  $\delta d_i$ 라 하면

$$\delta d_i = u_i - d_i, \quad i = (1, 2, \dots, n). \tag{4}$$

따라서 우리의 목적은 (4)식의 나머지 오차의  $l^2$ -크기를 최소화 하는  $\tilde{\mathbf{u}}$ 를 찾는 것이다. 그러므로 우리는 목적함수를

$$E(\mathbf{p}) = \frac{1}{2} \delta \mathbf{d}' \mathbf{d} \tag{5}$$

로 놓자. 이때 첨자 \*는 공액복소수(complex conjugate)를 나타낸다.

식 (5)의 최소값을 구하기 위하여 우리는 반복적 Newton 방법을 이용하려고 한다. 테일러급수를 이용하면,

$$E(\mathbf{p} + \delta \mathbf{p}) = E(\mathbf{p}) + \delta \mathbf{p}' \nabla_p E(\mathbf{p}) + \frac{1}{2} \delta \mathbf{p}' \mathbf{H} \delta \mathbf{p} + O(|\delta \mathbf{p}|^3) \tag{6}$$

로 표현된다. 식 (6)에서  $\mathbf{H}$ 는 Hessian 행렬이다. 우리는 목적함수  $E$ 를 최소화하는  $\delta \mathbf{p}$ 를 구하기를 원하기 때문에 (6)식을  $\delta \mathbf{p}$ 에 대하여 미분하면

$$0 = \nabla_p E + \mathbf{H} \delta \mathbf{p}.$$

이 된다. 그러므로 우리가 찾고자하는 해는 다음과 같이 주어진다.

$$\delta \mathbf{p} = -\mathbf{H}^{-1} \nabla_p E. \tag{7}$$

따라서 식 (7)으로부터,  $k$ 를 반복횟수라 할 때, 반복적 Newton

방법은

$$\mathbf{p}^{k+1} = \mathbf{p}^k - \mathbf{H}^{-1} \nabla_p E \tag{8}$$

로 표현된다. 식 (6)에서 Hessian 행렬의 각 항을 구해보면

$$H_{ij} = \frac{\partial^2 E}{\partial p_i \partial p_j} = Re \left\{ \left[ \frac{\partial u_1}{\partial p_i} \dots \frac{\partial u_n}{\partial p_i} \right] \begin{pmatrix} \frac{\partial u_1^*}{\partial p_j} \\ \vdots \\ \frac{\partial u_n^*}{\partial p_j} \end{pmatrix} + \left[ \frac{\partial^2 u_1}{\partial p_i \partial p_j} \dots \frac{\partial^2 u_n}{\partial p_i \partial p_j} \right] \begin{pmatrix} \delta d_1^* \\ \vdots \\ \delta d_n^* \end{pmatrix} \right\} \tag{9}$$

로 표현된다. 이 때, Jacobian을  $\mathbf{J} = \left[ \frac{\partial u_1}{\partial p_i} \dots \frac{\partial u_n}{\partial p_i} \right]$ 라 하면, 식 (9)을 이용하여 Hessian 행렬을 다음과 같이 나타낼 수 있으며

$$\mathbf{H} = Re \left\{ \mathbf{J}' \mathbf{J}^* \right\} + Re \left( \frac{\partial}{\partial p_1} \mathbf{J}' \right) \delta \mathbf{d}^* \left( \frac{\partial}{\partial p_2} \mathbf{J}' \right) \delta \mathbf{d}^* \dots \left( \frac{\partial}{\partial p_n} \mathbf{J}' \right) \delta \mathbf{d}^* \tag{10}$$

$$= \mathbf{H}_a + \mathbf{R}.$$

여기서  $\mathbf{H}_a = Re \left\{ \mathbf{J}' \mathbf{J}^* \right\}$ 로 두고 이것을 근사 Hessian이라고 부르고,  $\mathbf{R}$ 은 나머지 항이다.  $\mathbf{R}$ 항은 계산하기 어렵고 Tarantola (1987)에서 오차  $\delta \mathbf{d}$ 가 작거나 모델방정식이 유사선형이면  $\mathbf{R}$ 항의 값은 작다는 것을 보여주었다. 따라서 식 (8)에서  $\mathbf{R}$ 항을 무시하고, Hessian 행렬을 쓰는 대신 근사 Hessian  $\mathbf{H}_a$ 로 다시 표현하면,

$$\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} - \mathbf{H}_a^{-1} \nabla_p E \tag{11-1}$$

$$\text{즉, } \delta \mathbf{p} = -\mathbf{H}_a^{-1} \nabla_p E \tag{11-2}$$

이다.

### 파형역산

#### 파형역산의 문제점

파형역산에서 근사 Hessian 행렬  $\mathbf{H}_a$ 에 영향을 미치는 요소로는 속도모형, 송신원과 수신원의 위치, 영상화 영역, 송신파형의 주파수, 자동안정함수의 종류 등이며 이러한 요소들에 따라 행렬의 모양이 달라진다.

근사 Hessian 행렬을 사용하여 식 (11-2)를 풀어나가는데 크게 두 가지의 문제점이 발생한다. 첫 번째는 행렬  $\mathbf{H}_a$ 의 불안정성이다. 행렬  $\mathbf{H}_a$ 는 거의 특이치 행렬(singular matrix)에 가깝기 때문에 역행렬을 구하기가 매우 어렵다. 따라서 행렬의 안정화 방법이 필요하며 보통  $\mathbf{H}_a$ 에 감쇄항을 적용하여  $\mathbf{H}_a$ 의 불안정성을 개선한다. 따라서 식 (11-2)를 다시 쓰면

$$\delta \mathbf{p} = -(\mathbf{H}_a + \lambda \mathbf{I})^{-1} \nabla_p E \tag{12}$$

로 표현된다. 큰 값을 가진 감쇄항  $\lambda$ 를 사용하면 식 (12)의

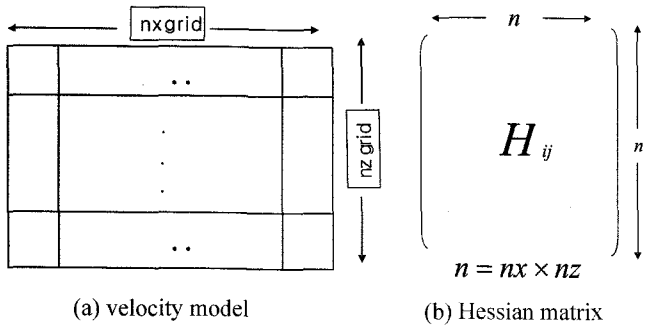


Fig. 1. The Hessian matrix constructed for velocity model whose size is  $n \times n$ .

$H + \lambda I$ 는 안정된 행렬이지만 gradient 방법과 비슷한 방법이 되고 작은 값을 가진 감쇄항  $\lambda$ 를 사용하면 행렬의 불안정성이 그대로 남아있어 역산이 잘 구해지지 않는다.

두 번째의 문제는  $H_a$ 를 구하는데 너무 많은 계산시간이 요구되어 계산하는 비용이 비싸다는 점이다. Fig. 1에서 보듯이 역산할 지층모델을 각각 수평 방향으로  $n_x$ 개, 수직방향으로  $n_z$ 개의 요소로 나눌 경우, 총 미지수의 개수는  $n_x \times n_z$ 개가 된다. 이 모델에 대한 파형역산을 할 경우 Hessian 행렬은  $n \times n (n = n_x \times n_z)$  행렬이 된다. 주파수 영역에서 계산된 Hessian 행렬의 한 행 또는 열을 계산하는데 필요한 연산횟수는 통상적인 single arrival Kirchoff 구조보정의 연산횟수와 동일하다.

예를 들어 100개의 주파수를 각각 수평 수직방향으로  $1000 \times 1000$ 의 요소로 나누었을 경우 Hessian 행렬을 계산하기 위한 연산횟수는  $1 \times 10^{20}$ 년의 연산이 요구된다. 현재의 고성능 슈퍼컴퓨터로도 2차원 속도모델에 대한 Hessian 행렬은 계산하기가 불가능하고, Hessian의 역산도 불가능하다. 대표적인 모델인 Marmousi 모형에 대한 파형역산에서 수평방향으로 150개, 수직방향으로 50개의 격자간격(60 m 간격)으로 나눈 경우에 대하여 주파수 영역에서 최대주파수를 5 Hz로 할 때, 필요한 메모리는 4 Giga byte, 디스크 용량은 30 Giga가 필요

하다. 또한 한번 반복실행을 하는데 걸리는 시간은 SNU의 IBM RS6000의 25개 CPU로 48시간 정도 걸린다. Fig. 2는 근사 Hessian  $H_a$ 를 이용하여 파형역산을 한 결과이다. 이 모델에 대하여 4 m 간격으로 역산을 하고자 한다면 1년 이상의 시간이 필요하다.

문제점의 해결을 위한 유사 Hessian 구축

탄성파역산에서 많은 시간이 요구되는 근사 Hessian 행렬을 구하는 것 대신 이를 대체할 다른 행렬들을 구축하였다. 이 행렬들을 만드는 목적은 시간을 단축시키거나 근사 Hessian 행렬보다 안정된 행렬을 얻고자 하는 것이다. 여기서 제시하는 방법은 두 가지다.

우선 첫 번째로 Jacobian 행렬이 다른 Jacobian 행렬의 한 부분과 자기상관(autocorrelation) 또는 상호상관(crosscorrelation)의 일정한 폭으로 제한한 행렬이다. 근사 Hessian 행렬의 i번째 열벡터는 Jacobian 행렬의 i번째 열벡터들끼리 의해 만들어지고 다른 Jacobian 열벡터들은 상호상관에 의해 구해진다. 그러므로 근사 Hessian 행렬의 모든 원소는 거의 모두 0이 아니다(Fig. 3). 그러므로 우리는 상호상관 폭을 변화하고자 하는 원소 주변의 Jacobian 열벡터로 제한하고자 하면 대각선을 중심으로 일정한 폭을 갖는 유사 Hessian 행렬은 상호상관하는

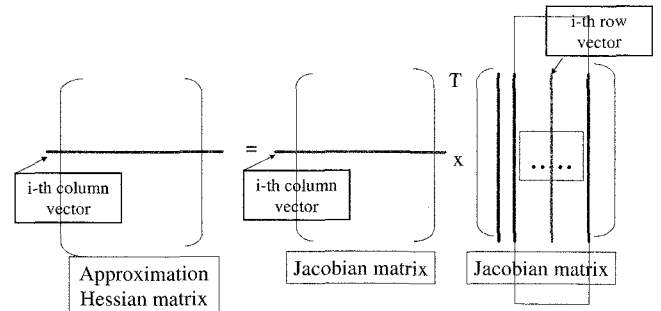


Fig. 3. Hessian matrix constructed by autocorrelation itself and crosscorrelating.

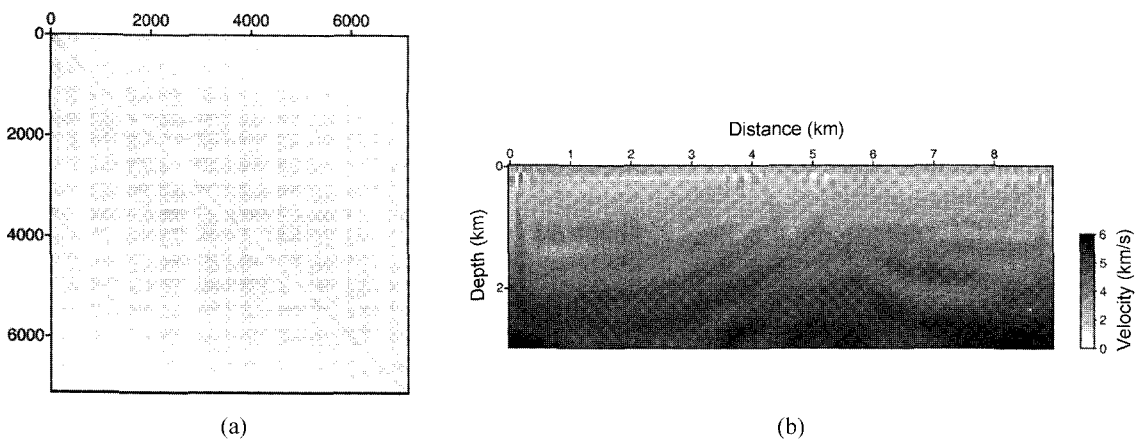


Fig. 2. Numerical examples for the Marmousi model: (a) approximate Hessian matrix and (b) velocity image inverted by using the approximate Hessian matrix using the iterative Newton method.

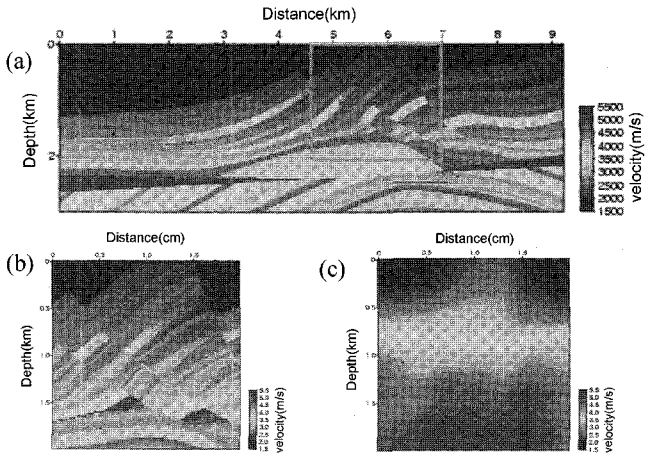


Fig. 4. (a) The Marmousi model, (b) selected area used to compute the Hessian matrix and (c) smoothed initial velocity model.

원소를 제외한 나머지 원소들은 모두 0인 값을 갖는다. 이러한 행렬을 이용하여 역산을 하면 행렬의 폭에 따라 계산시간이 많이 단축된다. 그러므로 우리는 근사 Hessian 행렬을 이용

한 것에 비하여 유사한 역산을 갖는 폭의 길이를 구하고자 시도하였다.

두 번째로 자동안정함수를 이용한 유사 Hessian 행렬이다. 자동안정함수는 양의 값을 갖는 증가함수이며 일반적으로  $t$  또는  $\rho$ 가 많이 쓰인다. 이 함수는 유사 Hessian 행렬의 불안정성을 개선하는데 상당한 역할을 한다는 것을 실험을 통해 알 수 있다. 따라서 몇 가지의 실험으로부터 가장 적당한 함수를 찾아보고자 한다.

## 결 과

### Marmousi 모형

수치해석결과는 모두 반복적 뉴턴방법을 한번 시행한 결과이다. 또한 모형은 Marmousi 모형의 일부분으로 실험하였다. Fig. 3으로부터 모형의 크기는  $2\text{ km} \times 2\text{ km}$ 이고 격자간격은  $20\text{ m}$ , 최대주파수는  $15\text{ Hz}$ , 그리고 구하고자 하는 최대시간은  $2\text{ 초}$ 까지이다. 그리고 속도모형은 원래모형을 부드럽게 수정한 속도모형으로 주었다(Fig. 4).

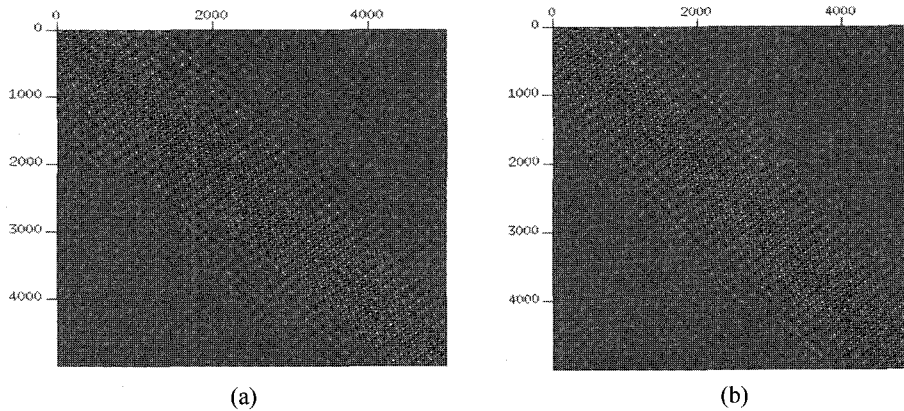


Fig. 5. The part of Hessian matrix whose size is  $5000 \times 5000$ : (a) approximate Hessian matrix and (b) Hessian matrix with fixed band width.

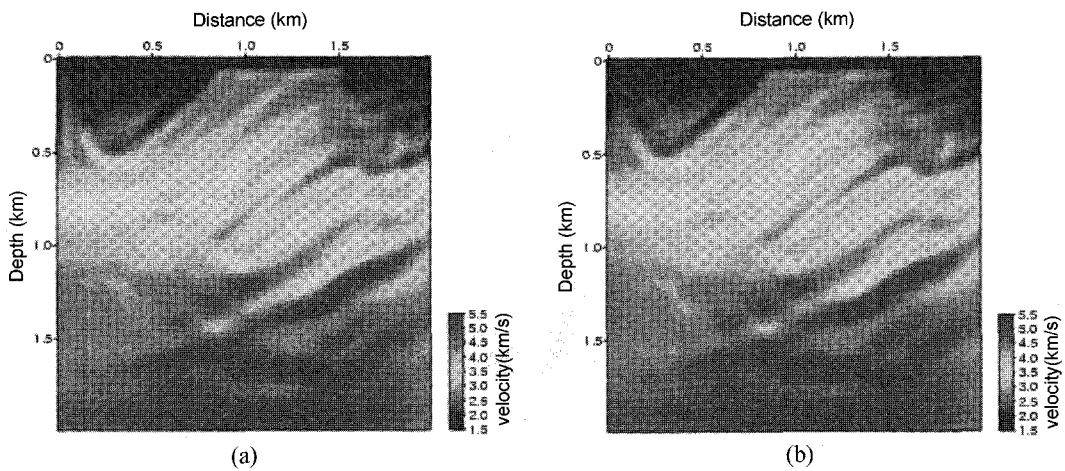
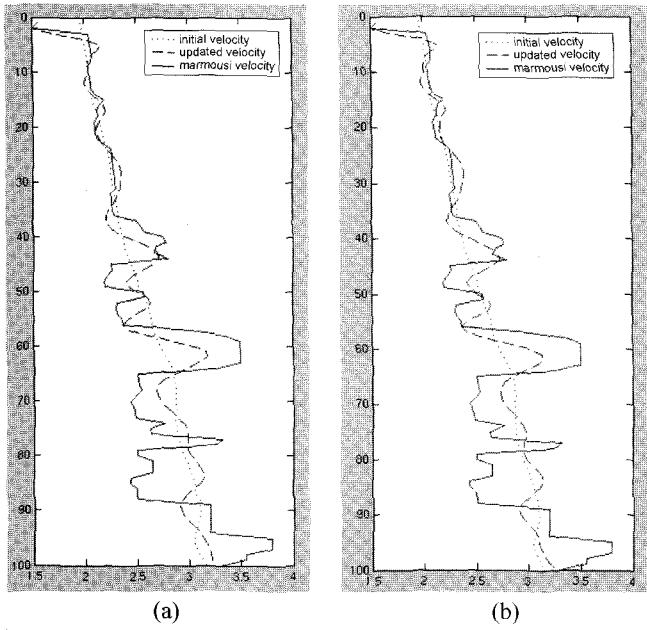


Fig. 6. The 1st updated velocity model using iterative Newton method with (a) approximate Hessian matrix and (b) Hessian matrix with fixed band width.



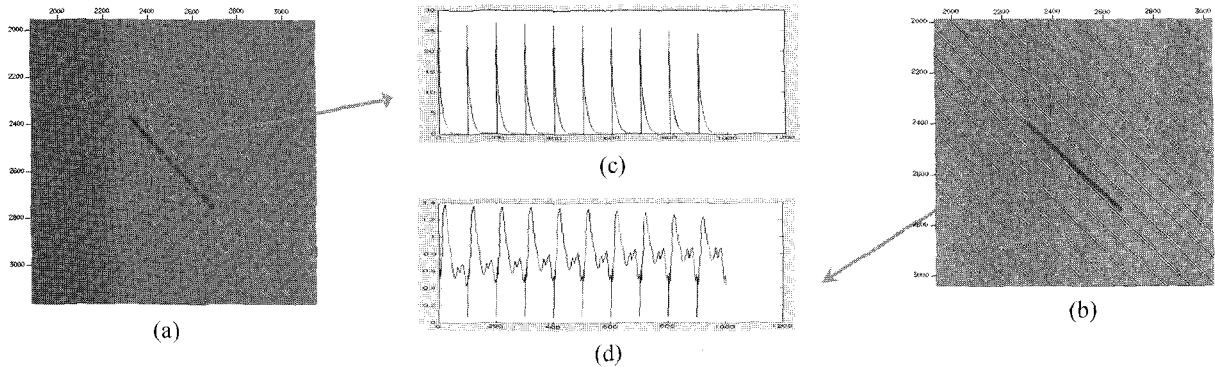
**Fig. 7.** Velocities at 1.0 km from the left corner of the model obtained (a) using the full approximate Hessian and (b) using the banded Hessian. True Marmousi model (dash), initial velocity model (dot) and updated velocity model (dotted dash).

**폭을 가진 Hessian 행렬을 이용한 역산**

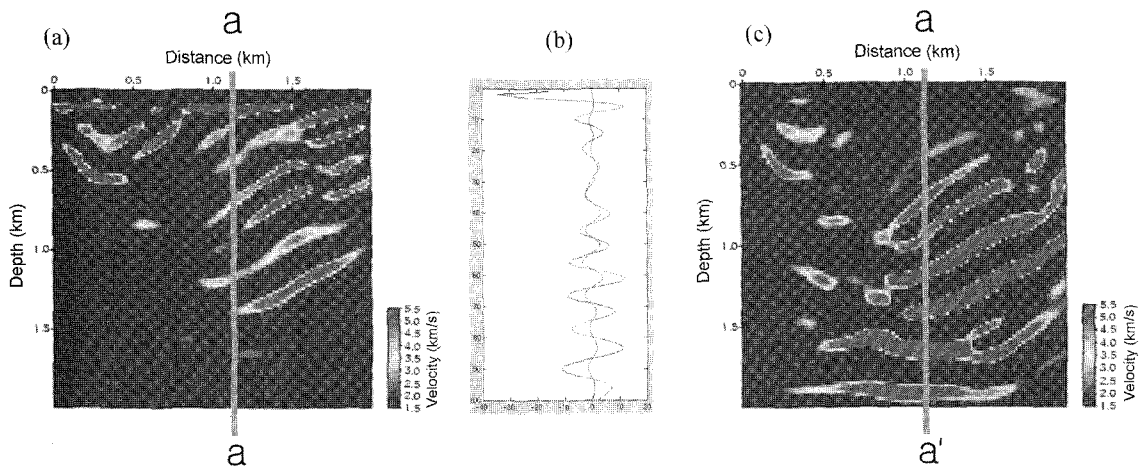
Fig. 4(b)에 대한 full 근사 Hessian 행렬과 일정한 폭을 갖는 Hessian 행렬은 Fig. 5와 같이 주어진다. Fig. 5(a)에서 보듯이 근사 Hessian 행렬의 거의 모든 원소들이 0이 아님을 알 수 있다. 이를 이용한 역산은 Fig. 6과 같이 주어진다. 우리는 이로부터 두 역산이 거의 비슷함을 알 수 있다. 또한 Fig. 7로부터 표면 1 km 지점에서의 파형을 비교하면 두 Hessian 행렬에 의해 구한 파형 역시 매우 유사함을 알 수 있다.

**자동안정함수(AGC)를 이용한 Hessian 행렬의 역산**

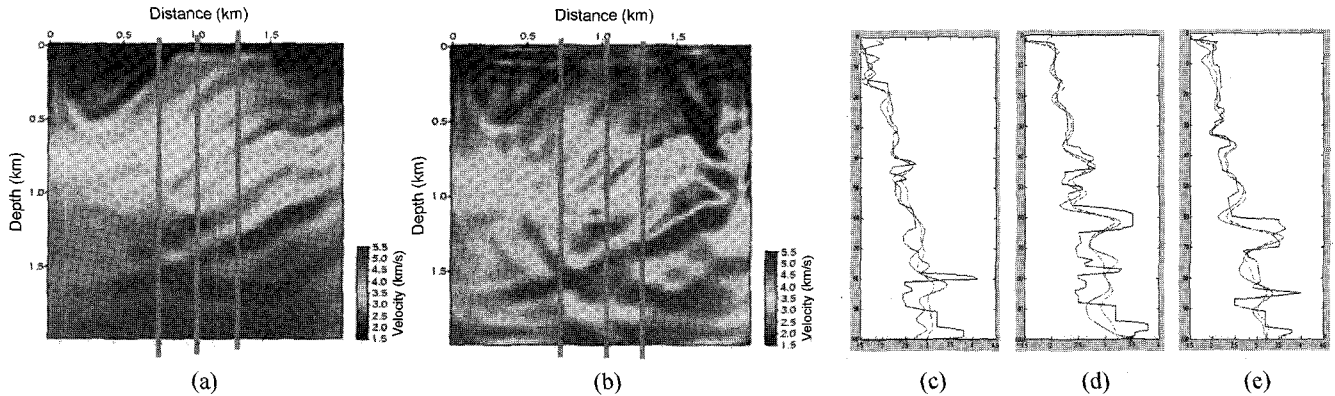
자동안정함수를 이용한 효과를 알아보기 위하여 우리는 시간영역에서 구한 Hessian 행렬에 대하여 실험을 해 보았다. 송신원 간격을 40 m, 수신원 간격을 40 m로 할 때, 자동안정함수를 사용하지 않은 Hessian 행렬 또는 Gradient 벡터와 자동안정함수를 사용한 Hessian 행렬 또는 Gradient 벡터를 비교해 보았다(Figs. 8, 9). 이 때, 자동안정함수로써 함수  $t$ 를 사용하였다. Fig. 8(d)로부터 Hessian 행렬의 대각원소의 최대값과 최소값을 비교할 때, 두 값의 비가 540에서 20으로 낮아진 것을 알 수 있다. 이 사실로부터 유사 Hessian 행렬의 불안정성이 많이 개선되었음을 알 수 있다. Fig. 9에서 gradient 값의 변화



**Fig. 8.** Comparison between some elements of the diagonal elements of the full approximate Hessians computed (a) without and (b) with automatic gain function, and the diagonal elements of the full approximation Hessian (c) with and (d) without automatic gain function.



**Fig. 9.** The gradient ( $J^T d$ ) computed (a) without (c) with automatic gain function respectively. (b) Gradient values along the cross-section a-a'.



**Fig. 10.** The 1<sup>st</sup> updated velocity model (a) without and (b) with automatic gain function. Green lines represent depth slices. (c), (d) and (e) are velocities at 0.6 km, 1 km and 1.4 km point from the left corner of the model, respectively. Black line indicates velocities for the Marmousi model, dot line denotes the velocity model without automatic gain function, and dotted dash line is the velocity model with automatic gain function.

를 나타내고 있다. Fig. 9(b)에서 보면 자동안정함수를 이용한 후엔 시간의 앞쪽의 파형은 상대적으로 작아지고 뒤쪽의 파형은 상당히 증폭되었음을 알 수 있다. 그리고 Fig. 10에서 자동안정함수를 사용한 Hessian 행렬을 이용한 역산과 그 단면이 주어져 있다.

## 결 론

우리는 Gauss-Newton 방법을 이용한 역산을 하는데 필요한 근사 Hessian 행렬의 성질을 살펴보고 이를 대체할 수 있는 행렬을 제시하였다. 일정한 폭을 갖는 유사 Hessian 행렬은 작은 크기의 행렬에 대해 근사 Hessian 행렬을 이용하여 얻은 영상과 매우 비슷함을 알 수 있다. 그러나 좀 더 큰 크기를 가진 모형에 대하여 비슷한 결과가 나올지에 대해서는 계속 연구해 봐야 할 것이다. 또한 어느 정도의 폭을 가져야 근사 Hessian 행렬과 유사한 역산 결과가 나올지에 대해서도 계속 연구해야 한다. 자동안정함수를 이용한 유사행렬도 역시 작은 모형에 대해 근사 Hessian 행렬이 가지고 있던 불안정성이 많이 제거됨을 알 수 있었다. 따라서 앞으로 전체 Marmousi 모델에 대해 실험을 계속하고 이를 3차원에 적용하고자 한다.

## 감사의 글

본 연구는 과학기술부 국가지정연구실 사업의 지원과 서울대학교 지구환경시스템공학부 신창수 교수님과 민동주 박사의 커다란 도움을 받았으며 이에 감사드립니다.

## 참고문헌

- Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional non-linear inversion of seismic waveforms, Numerical results, *Geophysics*, **51**, 1387-1403.
- Geller, R. J., and Hara, T., 1993, Two efficient algorithms for iterative linearized inversion of seismic waveform data, *Geophys. J. Int.*, **115**, 699-710.
- Kolb, P., 1983, Pre-stack inversion of a 1-D medium, *Proc. IEEE*, **74**, 498-508.
- Lailly, P., 1986, The seismic inverse problem as a sequence of before stack migration, in Bednar, J. B., Redner, R., Robinson, E., and Weglein, A., Eds., *Conference on Inverse Scattering, Theory and Application, Soc. Industr. Appl. Math.*
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain I-Theory and verification in a physical scale model, *Geophysics*, **64**, 888-901.
- Pratt, R. G., Shin, C. S., and Hicks, G. J., 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341-362.
- Shin, C., Jang, S. and Min, D. J., 2001a, Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophys. Prosp.*, **49**, 592-606.
- Shin, C. Yoon, K., Markurt, K. J. Park, Young, D. Lim, H. Chung, S., and Shin, S., 2001b, Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion, *Geophysics*, **66**, 1-8.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259-1266.
- Tarantola, A., 1987, *Inverse Problem Theory: Methods for data fitting and parameter estimation*, Elsevier, Amsterdam.