METHODS FOR APPROXIMATING HESSIAN TIMES VECTOR OPERATION IN FULL WAVEFIELD INVERSION

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Patent Application 61/564,669, filed November 29, 2011 entitled METHODS FOR APPROXIMATING HESSIAN TIMES VECTOR OPERATION IN FULL WAVEFIELD INVERSION, the entirety of which is incorporated by reference herein.

FIELD OF THE INVENTION

[0002] This invention relates generally to the field of geophysical prospecting and, more particularly, to seismic data processing. Specifically, the invention is a method for faster estimation of a quantity known as *Hessian times vector* which arises in certain methods for numerical solving of partial differential equations, for example iterative inversion of seismic data to infer elastic properties of a medium, involving forward modeling of synthetic data by solving the wave equation in a model medium.

BACKGROUND OF THE INVENTION

[0003] Full wavefield inversion (FWI) in exploration seismic processing relies on the calculation of the gradient of an objective function with respect to the subsurface model parameters [9]. The gradient of the objective function is used to calculate an update to the model. An objective function E is usually given as an L_2 norm as

[0004] $$E = \frac{1}{2} \int\int\int \left| p(\mathbf{r_g}, \mathbf{r_s}; t) - p_b(\mathbf{r_g}, \mathbf{r_s}; t) \right|^2 \, dt \, dS_g \, dS_s, \tag{1}$$

where p and p_b are the measured pressure, i.e. seismic amplitude, and the modeled pressure in the background subsurface model at the receiver location $\mathbf{r_g}$ for a shot located at $\mathbf{r_s}$. In iterative inversion processes, the *background* medium is typically the medium resulting from the previous inversion cycle. In non-iterative inversion processes or migrations, the background medium is typically derived using conventional seismic processing techniques such as migration velocity analysis. The objective function is integrated over all time t, and the surfaces S_g and S_s that are defined by the spread of the receivers and the shots. We define $K_d(\mathbf{r}) = K(\mathbf{r}) - K_b(\mathbf{r})$ and $\rho_d(\mathbf{r}) = \rho(\mathbf{r}) - \rho_b(\mathbf{r})$, where $K(\mathbf{r})$ and $\rho(\mathbf{r})$ are the true bulk modulus and density, and $K_b(\mathbf{r})$ and $\rho_b(\mathbf{r})$ are the bulk modulus and the density of the background model at the subsurface location $\mathbf{r}$. (Bulk modulus is used as an example here,

but any of the 21 elastic constants might be used instead.) We also define the difference between the measured and the modeled pressure to be $p_d(\mathbf{r_g},\mathbf{r_s};t) = p(\mathbf{r_g},\mathbf{r_s};t) - p_b(\mathbf{r_g},\mathbf{r_s};t)$.

[0005] The measured pressure p, satisfies the wave equation

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla p\right) - \frac{\rho}{K}\ddot{p} = -q(t)\delta(\mathbf{r}-\mathbf{r_s}), \qquad (2)$$

where $q(t)$ is the source signature. It can be shown that the gradient of the objective function E with respect to the bulk modulus $K_b(\mathbf{r})$, for example, is given as

$$\begin{aligned}\frac{\partial E}{\partial K_b(\mathbf{r})} &= -\iiint p_d \frac{\partial p_b}{\partial K_b(\mathbf{r})}\,dt\,dS_g\,dS_s \\ &= -\frac{\rho_b(\mathbf{r})dV}{K_b^2(\mathbf{r})}\iint \dot{p}_b(\mathbf{r},\mathbf{r_s};t)\int \frac{\rho_b(\mathbf{r_g})}{\rho_b(\mathbf{r})} g_b(\mathbf{r},\mathbf{r_g};-t) * \dot{p}_d(\mathbf{r_g},\mathbf{r_s};t)\,dS_g\,dt\,dS_s,\end{aligned} \qquad (3)$$

where g_b is the Green's function in the background medium, and dV is an infinitesimal volume around $\mathbf{r}$ [9, 3]. The equations for the gradients of other subsurface medium parameters in general elastic cases can be found in Refs. [10, 6, 1]. One can then perform full wavefield inversion by minimizing the value of the objective function E in an iterative manner, using gradient equations for the medium parameters such as that in Eq. 3.

[0006] The convergence rate of full wavefield inversion can be improved when information on the Hessian of the objective function E is employed in the inversion process [7, 5]. The Hessian is a matrix of second partial derivatives of a function. The Hessian (with respect to the physical property bulk modulus) of the objective function E in Eq. 1 is given as

$$\frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')} = \iiint \left[\frac{\partial p_d}{\partial K_b(\mathbf{r})}\frac{\partial p_d}{\partial K_b(\mathbf{r}')} + p_d\frac{\partial^2 p_d}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')}\right] dt\,dS_s\,dS_g. \qquad (4)$$

The second term in the right-hand side is the term responsible for multiple scattering, and is often ignored due to difficulty in evaluation [11]. By dropping this second term, one obtains the equation for the Gauss-Newton Hessian,

$$\frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')} \approx \iiint \frac{\partial p_d}{\partial K_b(\mathbf{r})}\frac{\partial p_d}{\partial K_b(\mathbf{r}')}\,dt\,dS_s\,dS_g. \qquad (5)$$

[0007] Once the Hessian matrix is computed, the medium parameter update required for the minimization of E can be obtained by multiplying the inverse of the Hessian matrix and the gradient using the Newton's method [5]:

$$K_d(\mathbf{r}) = -\left(\frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')}\right)^{-1} \cdot \frac{\partial E}{\partial K_b(\mathbf{r})}$$

Direct computation of the inverse of the Hessian matrix, however, often requires prohibitively large memory space in full wavefield inversion, and so the inverse of the Hessian is computed iteratively using the conjugate gradient (CG) method. This iterative scheme is often referred to as Newton-CG method, which may be used on either the full Hessian from equation 4 or the Gauss-Newton Hessian from equation 5. An example of this Newton-CG method can be found in Algorithm 7.1 of Ref. [5], which is reproduced below. For notational convenience, we use H for the Hessian matrix, and ∇E for the gradient vector.

1. Define tolerance $\delta = \min(0.5, \sqrt{\| \nabla E \|}) \| \nabla E \|$
2. Set $z_0 = 0,\ r_0 = \nabla E,\ \tilde{K}_0 = -\nabla E$
3. For $j = 0, 1, 2, \ldots$
   a. If $\tilde{K}_j^T H \tilde{K}_j \le 0$
      i. If $j = 0$
         1. Return $K_d = -\nabla E$
      ii. Else
         1. Return $K_d = z_j$
   b. Set $l_j = r_j^T r_j / \tilde{K}_j^T H \tilde{K}_j$
   c. Set $z_{j+1} = z_j + l_j d_j$
   d. Set $r_{j+1} = r_j + l_j H \tilde{K}_j$
   e. If $\| r_{j+1} \| < \delta$
      i. Return $K_d = z_{j+1}$
   f. Set $\beta_{j+1} = r_{j+1}^T r_{j+1} / r_j^T r_j$
   g. Set $\tilde{K}_{j+1} = -r_{j+1} + \beta_{j+1} \tilde{K}_j$
4. Perform line search using K_d as the search direction for the medium parameter update, starting with the step size of 1 if possible.

The algorithm above requires repeated evaluation of the Hessian times vector $H\tilde{K}_j$ in a loop, where $\tilde{K}_j$ for $j = 0$ is set as the negative of the gradient. The value of $H\tilde{K}_j$ is then used to update z_j, which eventually becomes the inverse of the Hessian times gradient K_d in 3.a.ii.1 or 3.e.i.

[0008] The Newton-CG method requires evaluation of the Hessian matrix times (i.e. multiplied by) a medium perturbation vector $\tilde{K}$, which may be referred to hereafter in this document as the (Gauss-Newton) "Hessian times vector,"

$$\begin{aligned}
&\int \frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV \\
&\quad = \int \left[\iiint \frac{\partial p_d}{\partial K_b(\mathbf{r})}\frac{\partial p_d}{\partial K_b(\mathbf{r}')}\, dtdS_s dS_g \right] \tilde{K}(\mathbf{r}')dV' \\
&\quad = \iiint \frac{\partial p_d}{\partial K_b(\mathbf{r})}\left[\int \frac{\partial p_d}{\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV' \right] dtdS_s dS_g,
\end{aligned} \tag{6}$$

where the Hessian has been approximated as the Gauss-Newton Hessian. For example, in the estimation of parameters of a physical property model by iterative inversion of geophysical data, the medium perturbation vector will typically initially be the gradient (in model parameter space) of the objective function. As the iterations progress, this vector will gradually diverge from being the gradient, and lose physical meaning as such. Equation (6) may be evaluated by performing two forward-wave-propagation and two reverse-wave-propagation computations. The present invention is a method that allows this equation to be evaluated by performing only one forward-wave-propagation (one forward solve of a partial differential equation, such as the wave equation) and one reverse-wave-propagation computation (i.e., one gradient computation of the objective function), resulting in valuable saving of computer time.

SUMMARY OF THE INVENTION

[0009] In one embodiment, the invention is a method for determining a discrete physical properties model of a subsurface region, which may be referred to as the "model," by iteratively inverting measured geophysical data acquired from the subsurface region, comprising using a Hessian matrix of an objective function, then times a vector, called "Hessian times vector," to determine an update for the model, wherein the Hessian times vector is approximated, using a computer, with a single forward-wave simulated propagation

and a single computation of gradient of the objective function, in a modified subsurface model.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The present invention and its advantages will be better understood by referring to the following detailed description and the attached drawings in which:

Fig. 1 is a graph comparing the convergence rate of full waveform inversion using the present inventive method as compared to the traditional method for estimating the Hessian times vector, in the test example using the models from Figs. 3A and 3B;

Fig.2 is a flowchart showing basic steps in one embodiment of the present inventive method; and

Fig. 3A shows a "true" subsurface velocity model for generating seismic data in a synthetic test example of the present inventive method, and Fig. 3B shows the initial velocity model used in iterative full wavefield inversion employing the method of the present invention.

[0011] The invention will be described in connection with example embodiments. However, to the extent that the following detailed description is specific to a particular embodiment or a particular use of the invention, this is intended to be illustrative only, and is not to be construed as limiting the scope of the invention. On the contrary, it is intended to cover all alternatives, modifications and equivalents that may be included within the scope of the invention, as defined by the appended claims. Persons skilled in the technical field will readily recognize that all practical applications of the present inventive method are performed using a computer programmed according to the disclosure herein.

DETAILED DESCRIPTION OF EXAMPLE EMBODIMENTS

[0012] Let

$$\begin{aligned}\int \frac{\partial p_d(\mathbf{r_g},\mathbf{r_s};t)}{\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV' &= -\int \frac{\partial p_b(\mathbf{r_g},\mathbf{r_s};t)}{\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV' \\ &= -\int \frac{\rho dV}{K_b^2(\mathbf{r}')}\Big[g_b(\mathbf{r_g},\mathbf{r}';t)*\ddot{p}_b(\mathbf{r}',\mathbf{r_s};t)\Big]\tilde{K}(\mathbf{r}')dV' \qquad (7)\\ &\equiv -\tilde{p}_d(\mathbf{r_g},\mathbf{r_s};t).\end{aligned}$$

It may be noted that Eq. 7 is the equation for the Born scattered field $\tilde{p}_d$, where $\tilde{K}$ behaves as a scatterer distribution in the background medium with bulk modulus K_b. The Gauss-Newton Hessian times vector can then be seen as

$$\int \frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV' = -\iiint \tilde{p}_d \frac{\partial p_d}{\partial K_b(\mathbf{r})} dt\, dS_s dS_g = \iiint \tilde{p}_d \frac{\partial p_b}{\partial K_b(\mathbf{r})} dt\, dS_s dS_g. \quad (8)$$

Therefore, computation of the Gauss-Newton Hessian times vector is equivalent to gradient computation using artificial residual $\tilde{p}_d$. As mentioned above, the artificial residual $\tilde{p}_d$ is computed using the Born approximation in the background model with scatterers $\tilde{K}$. If the wave equation is linear so that the Born approximation is exact, and if $\tilde{K} = K - K_b = K_d$, then the artificial residual $\tilde{p}_d = p_d$. Therefore, the negative of Hessian times vector above should be equal to the gradient in this case.

[0013] It may be noted that the Born scattered field $\tilde{p}_d$ in Eq. 7 can be approximated as

$$\tilde{p}_{d,F} = \frac{1}{\varepsilon}(\tilde{p}_F - p_b), \quad (9)$$

where $\tilde{p}_F$ is the solution to the wave equation

$$\nabla^2 \tilde{p}_F(\mathbf{r}) - \frac{\rho}{K_b(\mathbf{r}) + \varepsilon\tilde{K}(\mathbf{r})}\ddot{\tilde{p}}_F(\mathbf{r}) = -q(t)\delta(\mathbf{r} - \mathbf{r}_s). \quad (10)$$

If ε in Eq. 10 is sufficiently small, $\tilde{p}_F$ should be approximately equal to the summation of p_b and the Born scattered field due to $\grave{o}\tilde{K}$. Since the amplitude of Born scattering is linear with respect to the medium property perturbation of the scatterers, $\tilde{p}_{d,F}$ in Eq. 9 should be approximate equal to $\tilde{p}_d$ in Eq. 7.

[0014] The Hessian times vector operation can be approximated as

$$\int \frac{\partial^2 E}{\partial K_b(\mathbf{r})\partial K_b(\mathbf{r}')}\tilde{K}(\mathbf{r}')dV' = \iiint \tilde{p}_d \frac{\partial p_b}{\partial K_b(\mathbf{r})} dt dS_s dS_g \approx \iiint \tilde{p}_{d,F} \frac{\partial p_b}{\partial K_b(\mathbf{r})} dt dS_s dS_g. \quad (11)$$

The equation above shows that one can compute the Hessian times vector by (1) creating a new subsurface model $K_b + \varepsilon\tilde{K}$, (2) computing the received field $\tilde{p}_F$ in the new subsurface model, (3) computing the gradient in the background model K_b by treating $\tilde{p}_F$ to be a field measurement, and (4) scaling the gradient by $1/\varepsilon$. Since these operations uses only forward wave propagation and gradient computation, this method eliminates the need to implement the Hessian times vector operators, i.e. eliminates the need to implement a computer program that computes the Hessian times vector. Instead, one can reuse already existing forward propagation and gradient computation routines to obtain the Hessian times vector. Of course,

this advantage disappears if one has already implemented the Hessian times vector operation. Furthermore, the operation disclosed herein requires roughly 3.5 wave propagations compared to 4 wave propagations in Eq. 6, and so is computationally more efficient.

[0015] One can further improve the convergence to minima by increasing the value of ε in Eq. 10. When the value of ε is increased, the scattered field $\tilde{p}_{d,F}$ in Eq. 9 departs from the Born scattered field, and it includes nonlinear effects such as multiple scattering, travel time change, and nonlinear amplitude scaling with respect to $\tilde{K}$. One special case of a large ε value is when $\varepsilon = 1$, and $\tilde{K} = K_d$. In this case, the wave equation 10 is the wave equation for the true subsurface medium, and so $\tilde{p}_F = p$. The Hessian times vector in Eq. 11 is then exactly equal to the gradient, which is never achieved in Eq. 6 even when $\tilde{K} = K_d$ due to the neglected higher order terms in Eq. 6.

[0016] For practical purposes, we may define ε to be

$$\varepsilon = \alpha \frac{\| K_b(\mathbf{r}) \|_{\max}}{\| \tilde{K}(\mathbf{r}) \|_{\max}}, \qquad (12)$$

where $\| K_b(\mathbf{r}) \|_{\max}$ and $\| \tilde{K}(\mathbf{r}) \|_{\max}$ are the maximum absolute values of K_b and $\tilde{K}$ in space, respectively. The parameter α then represents approximate fractional change of $\grave{o}\tilde{K}(\mathbf{r})$ in Eq. 10 with respect to $K_b(\mathbf{r})$. Thus, α represents the ratio of the magnitude of the vector to be added to the model to the magnitude of the model vector, where the model's medium parameters are the components of the model vector in model space. One can then choose the value of α to control the behavior of the Hessian times vector in Eq. 11. When the value of α is fairly small, on the order of 0.01, the Hessian operator in Eq. 11 mimics the behavior of Gauss-Newton Hessian in Eq. 5. When the value of α is relatively large and reaches on the order of 0.1, on the other hand, the Hessain operator in Eq. 11 starts to include the effect of nonlinearity and multiple scattering, and so behaves similar to that in Eq. 4.

[0017] While the equation above was derived specifically for the bulk modulus K, this method can be applied for the computation of the Hessian times vector of any general elastic parameters such as density ρ or any of the 21 elastic stiffness constants C_{ij}. Let $m_b(\mathbf{r})$ and $\tilde{m}(\mathbf{r})$ be the background medium properties and the multiplication vector of any of these elastic parameters. Then the Hessian times vector of any of these properties can be computed as

$$\begin{aligned}
&\int \frac{\partial^2 E}{\partial m_b(\mathbf{r})\partial m_b(\mathbf{r}')}\tilde{m}(\mathbf{r}')dV' \\
&\quad = \int\left[\iiint \frac{\partial p_d}{\partial m_b(\mathbf{r})}\frac{\partial p_d}{\partial m_b(\mathbf{r}')}dtdS_s dS_g\right]\tilde{m}(\mathbf{r}')dV' \\
&\quad = \iiint \frac{\partial p_d}{\partial m_b(\mathbf{r})}\left[\int \frac{\partial p_d}{\partial m_b(\mathbf{r}')}\tilde{m}(\mathbf{r}')dV'\right]dtdS_s dS_g \\
&\quad = -\iiint \tilde{p}_d \frac{\partial p_d}{\partial m_b(\mathbf{r})}dtdS_s dS_g = \iiint \tilde{p}_d \frac{\partial p_b}{\partial m_b(\mathbf{r})}dtdS_s dS_g,
\end{aligned} \tag{13}$$

where $\tilde{p}_d$ now can be computed using the same method as that in Eqs. 9 and 10. Note this derivation is generally applicable to any type of objective functions such as L_2 objective function given in Eq. 1 or cross-correlation objective function in Ref. 8. Note also that the method presented here is a special case of PDE (Partial Differential Equation) constrained optimization problems, and so the method is in general applicable to any PDE constrained optimization problems where Hessian times vector needs to be computed. For example, this method can also be applied to problems relating to electro-magnetic wave propagation.

[0018] Below is the procedure of the present invention as applied to, for example, determining the model update in a method for physical property parameter estimation by data inversion using the gradient of an objective function, with the steps as shown in the flowchart of Fig. 2:

Step **21** Form a multi-component vector whose components are related to the values of a selected physical property at discrete cells in a model of the subsurface according to a current subsurface model **20** of that property (for the present example, the vector would typically be the gradient of the objective function, but in a general application of the inventive method, this would be whatever vector is later to be multiplied by the Hessian);

Step **22** Scale the vector by a small scalar constant (ε), i.e. by multiplying by ε;

Step **23** Add the scaled vector to the current subsurface model;

Step **24** Forward propagate the wavefields in the new model to compute a synthetic seismic dataset;

Step **25** Compute the gradient of the objective function measuring misfit (typically a selected norm of the data residual) between forward modeled data using the current subsurface model and "measured" data, where the synthetic seismic dataset from step **24** is treated as the measured data for purposes of this step;

Step **26** Divide the gradient from step **25** by the scalar constant ε from step **22**,

Step **27** Use this scaled gradient as an estimate of the Hessian times vector in optimization routines that require evaluation of that Hessian times vector.

[0019] Then, steps **21** to **27** may be repeated iteratively, using a Newton-CG algorithm for example, to obtain the inverse of the Hessian times the gradient of the objective function, which gives the model (medium parameter) update. Here, the objective function measures misfit between the actual measured field data and the data modeled from the current model.

[0020] The sum in step **23** is performed, cell-by-cell, in the discrete subsurface property model. For example, for application to model parameter estimation by geophysical data inversion, the "vector" will typically start out being the multi-dimensional gradient of the objective function with respect to each model parameter. Thus, the gradient of the objective function is a vector with as many components as there are cells in the model, times the number of subsurface parameters, such as bulk modulus or density. For each cell, the corresponding component of the vector, after being scaled, is added to the model value, sometimes called *parameter*, for that cell.

[0021] The method of steps **21-27** may be used in the iterative solving of any partial differential equation involving the Hessian times vector operation.

Example

[0022] For an illustrative example, we use the Marmousi II model [4] shown in Fig. 3A. This is the "true" model that was used to generate the synthetic data that was assumed to be the measured data for purposes of this example. The initial model for the full wavefield inversion is given in Fig. 3B. The inversion was performed by the encoded-simultaneous-source method described in Ref. [2], which discusses the Marmousi II model and includes the two drawings reproduced here. The inversion result using the Newton's method as a function of computation time is shown in Fig. 1. The solid line is the result using Gauss-Newton Hessian, the dashed line is the result using the present invention where $\alpha = 0.01$ in equation (12), and the dotted line is the result using the present invention where $\alpha = 0.1$. In other words, the solid line uses the Gauss-Newton Hessian times vector in equation (8), with the inverse evaluated by the C-G method *without* benefit of the present invention, whereas the two broken lines use the Hessian times vector in equation (11) with the inverse evaluated by the C-G method *with* benefit of the present invention Convergence of full waveform

inversion is shown as a function of computation time. Both computation time and the model fit are normalized to 1. Model fit is defined to be an RMS velocity error between the inverted and the true model, as defined in Ref. [2]. The dots in the lines mark the time for 1 iteration of the Newton method.

[0023] One can see from Fig. 1 that the present invention yields roughly 40 % speed up in the convergence compared to the result using the Gauss-Newton Hessian when α = 0.01. One can further improve the convergence by increasing the value of α to 0.1 so that more nonlinear effects are included in the Hessian times vector approximation. However, if the value of alpha were to be further increased, the new model in Step **23** above might be non-physical, with the result that one may not even be able compute the forward wavefield in Step **24**. For example, the velocity of the new summation model might become negative.

[0024] The foregoing application is directed to particular embodiments of the present invention for the purpose of illustrating it. It will be apparent, however, to one skilled in the art, that many modifications and variations to the embodiments described herein are possible. All such modifications and variations are intended to be within the scope of the present invention, as defined in the appended claims.

References

[1] A. Baumstein, J. E. Anderson, D. Hinkley, and J. R. Krebs, "Scaling of the objective function gradient for full wavefield inversion," 79th SEG Annual International Meeting, Expanded Abstract, (2009).

[2] J. R. Krebs, J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, "Fast full-wavefield seismic inversion using encoded sources," *Geophysics*, 74:WCC177-188, (2009).

[3] S. Lee, J. R. Krebs, J. E. Anderson, A. Baumstein, and D. Hinkley, "Methods for subsurface parameter estimation in full wavefield inversion and reverse-time migration," 80th SEG Annual International Meeting, Expanded Abstract, (2010).

[4] G. S. Martin, R. Wiley, K. J. Marfurt, "Marmousi2: An elastic upgrade for Marmousi," *The Leading Edge* **25**, 156-166 (2006).

[5] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, Chap. 7.1, New York, 2nd edition (2006).

[6] R. E. Plessix, "A review of the adjoin-state method for computing the gradient of a functional with geophysical applications," *Geophys. J. Int.*, 167:495-503, (2006).

[7] R. G. Pratt, C. Shin, and G. J. Hicks, "Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion," *Geophys. J. Int.*, 133:341-362, (1998).

[8] P. Routh, J. Krebs, S. Lazaratos, A. Baumstein, S. Lee, Y. H. Cha, I. Chikichev, N. Downey, D. Hinkley, and J. Anderson, "Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data," 81st SEG Annual International Meeting, Expanded Abstract, (2011).

[9] A. Tarantola, "Inversion of seismic reflection data in the acoustic approximation," *Geophysics*, 49:1259-1266, (1984).

[10] A. Tarantola, "A strategy for nonlinear elastic inversion of seismic reflection data," *Geophysics*, 51:1893-1903, (1986).

[11] A. Tarantola, "Inverse Problem Theory and Methods for Model Parameter Estimation," *SIAM*, (2005).

CLAIMS

1. A method for determining a discrete physical properties model of a subsurface region, referred to herein as the model or the subsurface model, by iteratively inverting measured geophysical data acquired from the subsurface region, comprising:

approximating a Hessian matrix of an objective function, then times a vector, called Hessian times vector, using a computer, with a single forward-wave simulated propagation and a single computation of gradient of the objective function, in a modified subsurface model, thereby requiring only three forward-wave or reverse-wave propagations;

wherein said approximation is based on an approximate equation for a Born scattered pressure field $\tilde{p}_{d,F}$ where the Hessian times vector is approximated by a gradient computation using $\tilde{p}_{d,F}$ as an artificial residual;

then computing a direction in model parameter space for an update to a current model by multiplying inverse of the Hessian matrix times a gradient of the objective function, wherein the inverse of the Hessian matrix is computed iteratively using a conjugate gradient method in which said approximate Hessian times vector is used to evaluate the Hessian matrix times a model perturbation vector;

performing a line search to determine magnitude of the model update using said computed direction; and

adding the model update to the current model to form an updated model, and using the updated model for geophysical prospecting.

2. The method of claim 1, wherein the modified subsurface model is a linear combination of a current subsurface model and the vector.

3. The method of claim 2, wherein the linear combination = the current subsurface model + {(a scalar constant) x (the vector)}, wherein $0 <$ the scalar constant ≤ 1.

4. The method of claim 3, wherein the scalar constant is chosen according to whether the Hessian matrix will be computed exactly, which means taking multiple scattering into account, or will be approximated by the Gauss-Newton Hessian matrix.

5. The method of claim 4, wherein if the Gauss-Newton approximation is to be used, the scalar constant is chosen such that{(a scalar constant) x (the vector)} is approximately 1 % in magnitude of the current subsurface model; but if the Hessian matrix is to be computed exactly, then the scalar constant is chosen so that the {(a scalar constant) x (the vector)} is approximately 10 % in magnitude of the current subsurface model.

6. The method of claim 3, wherein the vector in a first iteration of the method is gradient of the objective function with respect to parameters of the model, wherein the objective function measures misfit between the measured geophysical data and corresponding synthetic geophysical data simulated using the model, said misfit being a selected norm of a data residual, wherein the data residual is the difference between the measured geophysical data and corresponding synthetic geophysical data simulated using the model; provided that for said single computation of gradient of the objective function, the measured geophysical data in the data residual is replaced by synthetic data simulated using the modified subsurface model, and the data residual is divided by the scalar constant.

7. The method of claim 1, where the measured geophysical data being inverted is a full wavefield of seismic data.

8. The method of claim 7, where the physical properties model is a model of at least one of a group consisting of 21 elastic constants including bulk modulus; density; or any combination of two or more of them.

9. A method for determining a discrete physical properties model of a subsurface region by iteratively inverting measured geophysical data acquired from the subsurface region, comprising using a Hessian matrix operating on an objective function, then times a vector, called Hessian times vector, to determine an update for an initial model, wherein the Hessian times vector is approximated, using a computer, with a single forward wave simulated propagation and a single gradient of the objective function computation in a modified subsurface model thereby requiring only three forward-wave or reverse-wave propagations, comprising:

forming a scaled vector by multiplying the vector by a number ε, where $0 < \varepsilon \leq 1$, and adding the scaled vector to the initial model;

simulating a first synthetic data set using the model with the scaled vector added to it;

forming the objective function to measure misfit between a second synthetic data set, simulated using the initial model, and said first synthetic data set;

computing a gradient of the formed objective function with respect to the physical property parameters, dividing the gradient by ε, and using that as an estimate of the Hessian times vector;

computing a direction in model parameter space for an update to the initial model by multiplying inverse of the Hessian matrix times a gradient of the objective function, wherein the inverse of the Hessian matrix is computed iteratively using a conjugate gradient method in which said estimate of the Hessian times vector is used to evaluate the Hessian matrix times a model perturbation vector;

performing a line search to determine magnitude of the model update using said computed direction; and

adding the model update to the initial model to form an updated model, and using the updated model for geophysical prospecting.

10. The method of claim 9, wherein the vector in a first iteration of the iterative inversion is gradient in model space of the objective function, wherein the objective function measures misfit between the measured geophysical data and corresponding model-simulated geophysical data.

11. A computer program product, comprising a non-transitory computer usable medium having a computer readable program code embodied therein, said computer readable program code adapted to be executed to implement a method for determining a discrete physical properties model of a subsurface region, referred to herein as the model or the subsurface model, by iteratively inverting measured geophysical data acquired from the subsurface region, said method comprising:

using a Hessian matrix of an objective function, then times a vector, called "Hessian times vector," to determine an update for the model, wherein the Hessian times vector is approximated with a single forward-wave simulated propagation and a single computation of gradient of the objective function, in a modified subsurface model, thereby requiring only three forward-wave or reverse-wave propagations;

wherein said approximation is based on an approximate equation for a Born scattered pressure field $\tilde{p}_{d,F}$ where the Hessian times vector is approximated by a gradient

computation using $\tilde{p}_{d,F}$ as an artificial residual;

then computing a direction in model parameter space for an update to a current model by multiplying inverse of the Hessian matrix times a gradient of the objective function, wherein the inverse of the Hessian matrix is computed iteratively using a conjugate gradient method in which said approximate Hessian times vector is used to evaluate the Hessian matrix times a model perturbation vector;

performing a line search to determine magnitude of the model update using said computed direction; and

adding the model update to the current model to form an updated model.

12. The method of claim 11, wherein the vector in a first iteration of the iterative inversion is gradient in model space of the objective function, wherein the objective function measures misfit between the measured geophysical data and corresponding model-simulated geophysical data.

13. The method of claim 1, wherein $\tilde{p}_{d,F}$ is approximated as

$$\tilde{p}_{d,F} = \frac{1}{\varepsilon}(\tilde{p}_F - p_b),$$

where ε is a number smaller than a pre-determined value, p_b is a pressure modeled using a current subsurface model, and $\tilde{p}_F$ is the solution to a wave equation

$$\nabla^2 \tilde{p}_F(\mathbf{r}) - \frac{\rho}{K_b(\mathbf{r}) + \varepsilon\tilde{K}(\mathbf{r})} \ddot{\tilde{p}}_F(\mathbf{r}) = -q(t)\delta(\mathbf{r} - \mathbf{r}_s).$$

14. The method of claim 13, wherein said approximation further includes taking $\tilde{p}_{d,F}$ to be equal to $\tilde{p}_d$, where $\tilde{p}_d$ is difference between measured pressure and the modeled pressure p_b.