

EFFICIENT LINE SEARCH METHODS FOR MULTI-PARAMETER FULL WAVEFIELD INVERSION

CROSS-REFERENCE TO RELATED APPLICATION

- 5 **[0001]** This application claims the benefit of U.S. Provisional Patent Application 61/990,860, filed May 9, 2014, entitled EFFICIENT LINE SEARCH METHODS FOR MULTI-PARAMETER FULL WAVEFIELD INVERSION, the entirety of which is incorporated by reference herein.

FIELD OF THE INVENTION

- 10 **[0002]** This disclosure relates generally to the field of geophysical prospecting for hydrocarbons and, more particularly, to seismic data processing. Specifically the disclosure relates to a method for performing efficient line searches in multi-parameter full wavefield inversion (“FWI”) of seismic data to infer a subsurface physical property model. Such a model may be useful in exploration for, or production of, hydrocarbons.

15 BACKGROUND OF THE INVENTION

- [0003]** Full wavefield inversion is a nonlinear inversion technique that recovers the earth model by minimizing the mismatch between the simulated and the observed seismic wavefields. Due to the high computational cost associated with FWI, conventional implementations utilize local optimization techniques to estimate optimal model parameters.
- 20 A widely used local optimization technique is the gradient-based first-order method, (e.g., steepest descent or nonlinear conjugate gradient), which utilizes only the gradient information of the objective function to define a search direction. Although a gradient-only first-order method is relatively efficient – it requires computing only the gradient of the objective function – its convergence is generally slow. The convergence of FWI can be improved
- 25 significantly by using a second-order method. This improved convergence is achieved because second-order methods utilize both the gradient and curvature information of the objective function to determine an optimal search direction in model parameter space. (The search direction unit vector $\mathbf{s}$ is related to the model update process by $\mathbf{m}_{\text{updated}} = \mathbf{m} + \alpha \mathbf{s}$, where α (a scalar) is the step size.)
- 30 **[0004]** The major difference between first and second order methods is that second-order methods precondition the gradient with the inverse Hessian (e.g., Gauss-Newton/Newton method), or with the inverse of a projected Hessian (e.g., subspace method). The Hessian is a

matrix of second-order partial derivatives of the objective function with respect to the model parameters. In general, second-order methods are attractive not only because of their relative fast convergence rate, but also because of the capability to balance the gradients of different parameter classes and provide meaningful updates for parameter classes with different data sensitivities (e.g., velocity, anisotropy, attenuation, etc.) in the context of multi-parameter inversion. In second-order methods, optimum scaling of parameter classes using the Hessian is crucial in multi-parameter inversion, if such parameter classes are to be simultaneously inverted. However, because it is very expensive to compute the inverse of the Hessian, this is a major obstacle for wide adoption of second-order methods in practice. Another disadvantage of second-order methods is that if the objective function is not quadratic or convex (e.g., where the initial model is far from the true model), the Hessian or its approximation may not accurately predict the shape of the objective function. Hence, the gradients for different parameter classes may not be properly scaled, thereby resulting in suboptimal search directions.

15

SUMMARY OF THE INVENTION

[0005] In one embodiment, the invention is a method for iteratively inverting seismic data to simultaneously infer a model for at least two physical properties of the subsurface, said method comprising:

- (a) for each physical property, computing a gradient of an objective function with respect to parameters of the physical property, said objective function measuring misfit between all or part of the seismic data and corresponding model-simulated seismic data;
- (b) for each physical property, computing a search direction in model space from the gradient;
- (c) alternating line searches between or among the search directions for the at least two physical properties to determine optimal step sizes along each of the search directions; and
- (d) using the optimal step sizes to update the model.

25

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] The present invention and its advantages will be better understood by referring to the following detailed description and the attached drawings in which:

30

Fig. 1 illustrates the optimal search direction in the alternating one-pass line search embodiment of the present invention;

Fig. 2 illustrates a sub-optimal search direction that may result from the alternating one-pass line search embodiment of the present invention;

- 5 Fig. 3 illustrates the optimal search direction using the alternating two-pass line search embodiment of the present invention;

Figs. 4A-4F show, for a test example of the present inventive method, the initial model and the final inverted model for two parameters (velocity and anisotropy), and compare the data after migration using the initial and final models;

- 10 Fig. 5 is a flow chart showing basic steps in the alternating one-pass line search of the present inventive method;

Fig. 6 is a flow chart showing basic steps in the alternating two-pass line search of the present inventive method; and

- 15 Fig. 7 is a flow chart showing basic steps in an extended alternating two-pass line search disclosed herein, in which a cascaded inversion approach is combined with the alternating two-pass line search.

[0007] Due to patent rule restrictions on the use of color in drawings, some of Figs. 4A-4F are black-and-white reproductions of drawings originally generated in color.

- [0008] The invention will be described in connection with example embodiments.
20 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.

25 **DETAILED DESCRIPTION OF EXAMPLE EMBODIMENTS**

- [0009] The present invention bridges the gap between first-order and second-order optimization methods by using what may be called an alternating one/two pass line search method, which requires no explicit information from the Hessian matrix, but approximates the second-order information through successive line searches. It will be shown that the
30 present inventive method can properly scale the gradients for parameter classes with different data sensitivity and can provide meaningful updates for multiple parameter classes simultaneously. In practice, the present inventive method can be more robust than the

Hessian-based second-order method, because it does not assume that the objective function is quadratic, and it can also be significantly cheaper if each line search can be efficiently implemented.

[0010] Although, for simplicity, the theory is described using two parameter classes, the invention is applicable to simultaneous inversion of any number of parameter classes, and extension of the method to more than two parameter classes is straightforward. For the case where two parameter classes are to be inverted, the model can be expressed as a vector containing two different sub-models, i.e., $\mathbf{m} = (\mathbf{m}_1 \ \mathbf{m}_2)^T$, where $\mathbf{m}_1$ and $\mathbf{m}_2$ are the first and second model parameter class, respectively, where T stands for the transpose. The search direction at a current iteration $\mathbf{s}$ is a concatenation of the search directions for both these parameter classes, and can be written as follows:

$$\mathbf{s} = \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix}, \quad (1)$$

where $\mathbf{s}_1$ and $\mathbf{s}_2$ are the search directions of the first and second parameter class, respectively. The first-order based approach typically updates model $\mathbf{m}$ along direction $\mathbf{s}$. One major issue of this approach is that if the model parameter classes are physical quantities with very different units and sensitivities to the FWI objective function (e.g., velocity, anisotropy, attenuation, etc.), the resulting search direction derived from the FWI gradient will typically have quite different magnitudes, i.e. the magnitude of $\mathbf{s}_1$ may differ substantially from that of $\mathbf{s}_2$. This often leads FWI to choose a convergence path along which it predominantly updates parameter classes that are more sensitive to the objective function, while keeping parameter classes that are less sensitive to the objective function barely updated. Consequently, FWI may converge to a suboptimal solution. Below are described two alternative embodiments of the present invention's alternating line-search-based approach, which address this issue, with the goal of providing optimal updates for all parameter classes simultaneously without biasing towards certain parameter classes due to different units or data sensitivity.

Method I: Alternating one-pass line search

[0011] To start, two basis vectors are defined, each of which contains the search direction for a particular model parameter class:

$$\tilde{\mathbf{s}}_1 = \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{0} \end{pmatrix}, \quad \tilde{\mathbf{s}}_2 = \begin{pmatrix} \mathbf{0} \\ \mathbf{s}_2 \end{pmatrix}, \quad (2)$$

where $\mathbf{0}$ denotes a vector containing zeros. Then, the original search direction can be written as the sum of the above two basis vectors, which are orthogonal to each other (i.e., $\tilde{\mathbf{s}}_1 \cdot \tilde{\mathbf{s}}_2 = 0$):

$$\mathbf{s} = \tilde{\mathbf{s}}_1 + \tilde{\mathbf{s}}_2. \quad (3)$$

[0012] In this first method, the two model parameter classes are updated in an alternating fashion, with basic steps shown in Fig. 5. The current model **52** is used to simulate predicted data, and that is combined with the measured data **51** to compute an objective function, and a gradient of the objective function with respect to the first model parameter is then computed in step **53** to yield a search direction $\tilde{\mathbf{s}}_1$ in step **54**. In step **55**, a line search is performed along direction $\tilde{\mathbf{s}}_1$ and the model is updated. Since $\tilde{\mathbf{s}}_1$ is nonzero only for the first parameter class, only parameter class $\mathbf{m}_1$ will be updated. After the model has been updated, a line search is then performed along direction $\tilde{\mathbf{s}}_2$ and the model is updated again (step **56**). In this case, since $\tilde{\mathbf{s}}_2$ is nonzero only for the second parameter class, only parameter $\mathbf{m}_2$ will be updated, as shown in Fig. 1. In this approach, one FWI iteration refers to the process of updating both parameter classes in this alternating manner, and at step **57** the process is repeated for another iteration unless it has converged according to some predetermined convergence criterion or has reached another stopping point. Because $\mathbf{m}_1$ and $\mathbf{m}_2$ are updated in an independent manner, the scaling-imbalance in the search directions (or gradients) described above is overcome. All parameter classes then receive significant updates in an alternating fashion. This method may be called the alternating one-pass line search. One potential drawback of this method is that the order of the successive line searches can influence the convergence path, and if the objective function has a relatively complicated shape, this may result in a zigzagged convergence path (Fig. 2).

[0013] In more detail, a line search involves model-simulating seismic data for various different "step" sizes along the search direction in the highly multi-dimensional model parameter space. The step size that minimizes the misfit between the model-simulated data and the measured data is selected. In the Method I described above, after the first physical property is updated in the model, then that updated model is used to simulate data for the line search to update the second physical property. Even though that second line search updates

only the second physical property, the update to the first physical property from the first line search will affect the model simulated data for the second line search, and therefore will affect the resulting update to the second physical property.

[0014] To further explain Fig. 1, the solution is represented by the star in the center of the objective function contours, which are ellipses instead of circles because model parameter $\mathbf{m}_1$ is more sensitive than $\mathbf{m}_2$ to the objective function. Figure 1 shows that in this situation the solution can be reached in a single cycle of the iteration process by making two independent steps using the alternating one-pass line search embodiment of the present inventive method. Figure 1 also shows that the traditional line search $\mathbf{s}$ is thrown off of the right direction by the different sensitivities, and will require more iterations to reach the solution. Figure 2 shows where the alternating one-pass line search might be problematic, i.e., when $\mathbf{m}_1$ and $\mathbf{m}_2$ are coupled, as shown by the objective function contour, which is tilted, meaning strong correlation between these two parameter classes. Under such circumstances, the alternating one-pass line search might require additional iterations to find an optimal search direction. Figure 3 shows how the alternating two-pass line search embodiment of the present invention, explained next, can solve this problem.

Method II: Alternating two-pass line search

[0015] Some of the above-identified drawbacks of Method I can be mitigated by a modification to the Method I. As indicated in Fig. 6, the observed data **61** and the current subsurface model **62** are used in computing the gradient of the objective function for each parameter class (step **63**), resulting in a search direction for each parameter class (**64**). In step **65**, a first-pass independent line search is performed for each parameter class using the search direction computed in step **64** for such parameter class. The computed step sizes are saved, but no model update is performed yet. In step **66**, this modified approach is similar in one of its aspects to the so-called subspace method (Kennett, et al., 1998) in that instead of updating all model classes in an alternating fashion, it first recombines the two basis vectors $\tilde{\mathbf{s}}_1$ and $\tilde{\mathbf{s}}_2$ using two scaling factors α and β , i.e. step sizes, to form a new search direction as follows:

$$\mathbf{s}_{\text{new}} = \alpha \tilde{\mathbf{s}}_1 + \beta \tilde{\mathbf{s}}_2. \quad (4)$$

Once this new combined search direction is obtained, a second-pass line search (utilizing the new search direction) is performed to update both parameter classes simultaneously (step **67**). A main difference between this approach and the well-known subspace approach (Kennett, et

al., 1988) is how the scaling factors α and β are estimated. In the subspace approach, the scaling factors are determined by inverting the following projected Hessian matrix:

$$\begin{pmatrix} \tilde{\mathbf{s}}_1^T \mathbf{H} \tilde{\mathbf{s}}_1 & \tilde{\mathbf{s}}_1^T \mathbf{H} \tilde{\mathbf{s}}_2 \\ \tilde{\mathbf{s}}_2^T \mathbf{H} \tilde{\mathbf{s}}_1 & \tilde{\mathbf{s}}_2^T \mathbf{H} \tilde{\mathbf{s}}_2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = - \begin{pmatrix} \mathbf{g}^T \tilde{\mathbf{s}}_1 \\ \mathbf{g}^T \tilde{\mathbf{s}}_2 \end{pmatrix}, \quad (5)$$

where $\mathbf{H}$ is the Hessian matrix. Instead of using the Hessian, preferred embodiments of

5 Method II estimate the optimal scaling factors through line searches (step 65). In this two-parameter example of the Method II of the present disclosure, two independent line searches are first performed as follows:

- (i) First, determine the optimal step size α such that it minimizes the objective function along the search direction defined by $\tilde{\mathbf{s}}_1$, but do not actually update the model;
- 10 (ii) then determine the optimal step size β such that it minimizes the objective function along the search direction defined by $\tilde{\mathbf{s}}_2$ also without actually updating the model.

[0016] Note that the fact that the method does not update the model parameters with the estimated step sizes represents one of the major differences between this two-pass line search method and the previously described one-pass line search method. Once the scaling factors are determined, a new search direction is formed using equation 4. A second pass line search is then performed using the new search direction $\mathbf{s}_{\text{new}}$ to update both parameter classes, i.e. we try to find a step length λ such that the objective function is minimized along direction $\mathbf{s}_{\text{new}}$. This approach may be called the alternating two-pass line search method: the first pass line search determines the relative scaling among different search components, and the estimated scaling effectively rotates the original search direction to obtain a new search direction; then the second pass line search updates the model parameters using the new search direction. Since the scaling factors are determined by independent line searches, the scaling of the original search direction (or gradient) can then be corrected using the estimated scaling factors. The relative magnitudes of each component of the combined search direction are mainly determined by how sensitive each parameter class is along the basis vector, i.e., the sensitivity of $\mathbf{m}_1$ along $\mathbf{s}_1$, and the sensitivity of $\mathbf{m}_2$ along $\mathbf{s}_2$.

Practical considerations and extensions

Determination of scaling by exhaustive line search

[0017] In the alternating two-pass line search method disclosed herein, one objective is that the first pass line search finds proper scaling factors (step sizes) for each parameter class. In order to properly do so, an exhaustive line search may be performed to find the inflection points when performing the first pass line search. Taking inversion of two parameter classes
 5 for example, that means iteratively scanning the objective function values by perturbing the concatenated model $\mathbf{m}$ along search direction $\tilde{\mathbf{s}}_1$ for a range of perturbation strengths (effectively, this means scanning $\mathbf{m}_1$ along $\mathbf{s}_1$, because the second component of $\tilde{\mathbf{s}}_1$ is zero). The line search is terminated when the objective function value is bigger than the previous objective function value. The returned α is used as the scaling factor to weight the search
 10 direction for the first parameter class. An algorithm is described next that illustrates an exhaustive line search:

do while

$$\alpha_i \leftarrow \alpha_i + \Delta\alpha$$

compute $J(\mathbf{m} + \alpha_i \tilde{\mathbf{s}}_1)$

15 if $J(\mathbf{m} + \alpha_i \tilde{\mathbf{s}}_1) > J(\mathbf{m} + \alpha_{i-1} \tilde{\mathbf{s}}_1)$, stop and return $\alpha = \alpha_i$

$$i \leftarrow i + 1$$

end

where $\Delta\alpha$ is a user-supplied increment value, and J is the objective function value. Similarly, the optimal scaling factor β for the second parameter class can be found by
 20 perturbing $\mathbf{m}$ along search direction $\tilde{\mathbf{s}}_2$.

Extended methodology--incorporating cascaded decomposition of data sets

[0018] The present inventive method can also be extended by incorporating the cascaded inversion method described in a companion patent application, "A Method for Estimating Multiple Subsurface Parameters by Cascaded Inversion of Wavefield Components,"
 25 inventors Ayeni et al., which is incorporated herein by reference in all jurisdictions that allow it. This extension allows readily conditioning the estimated gradients (and hence search directions) using parts of the data most sensitive to individual parameter classes. The extended embodiment of the present inventive method is summarized in the flow chart of Fig. 7. The extended approach disclosed herein differs from that described in Ayeni et al. at least
 30 because, in the present invention, all parameter classes are simultaneously updated in every

iteration whereas in the inversion taught by Ayeni et al., only one parameter class is updated at every iteration.

Cost considerations for multiple line searches

[0019] Compared to the conventional gradient-based first-order approaches, the alternating one/two pass approach requires performing many more line searches, which might be costly. To mitigate that issue, only a small subset(s) of the source shots may be used, which may preferably be randomly selected from the entire survey, to perform the line searches. These randomly selected shots can be used for the first pass and/or second pass line searches.

10 *EXAMPLES*

[0020] In a test example, the alternating two-pass line search method disclosed herein was applied to anisotropic VTI (vertical-transverse isotropic) inversion of a 3D field data example. In this example, two parameter classes were simultaneously inverted: Thompson's anisotropy parameter η and Normal Moveout velocity V_{NMO} . During the inversion, the other Thompson anisotropy parameter δ is set equal to zero, and it is fixed throughout the inversion. Figures 4A-4F compare the models and migrated gathers derived from the initial and inverted models. Figures 4A and 4D show Kirchhoff depth migration gathers derived from the initial (4A) and final (4D) models. Figures 4B and 4E show the initial and final η models respectively, while Figs. 4C and 4F show the initial and inverted V_{NMO} models. Note that the migrated gathers migrated using the FWI inverted models of V_{NMO} and η derived from the alternating two-pass line search method (the migrated gathers shown in Fig. 4D) are flat relative to the gathers shown in Fig. 4A that were migrated using the starting models – indicating accuracy of the derived models.

[0021] The foregoing description 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 by the appended claims. As will be apparent to those who work in the technical field, all practical applications of the present inventive method are performed using a computer, programmed in accordance with the disclosures herein.

References

Kennett, B. L. N., M. S. Sambridge, and P. R. Williamson, "Subspace methods for large inverse problems with multiple parameter classes," *Geophysical Journal* **94**, 237-247 (1988).

CLAIMS

1. A computer-implemented method for iteratively inverting seismic data to simultaneously infer a model for at least two physical properties of the subsurface, said method comprising:
 - (a) for each physical property, computing, with a processor, a gradient of an objective function with respect to parameters of the physical property, said objective function measuring misfit between all or part of the seismic data and corresponding model-simulated seismic data;
 - (b) for each physical property, computing, with the processor, a search direction in model space from the gradient;
 - (c) alternating line searches between or among the search directions for the at least two physical properties to determine, with the processor, optimal step sizes along each of the search directions;
 - (d) using the optimal step sizes to update the model with the processor; and
 - (e) generating, with the processor, an image of the subsurface displaying the at least two physical properties according to the updated model;wherein (c)-(d) comprise:
 - (i) performing a first line search for a first physical property, using an initial model, and saving the optimal step size from the first line search but not updating the model yet;
 - (ii) performing a second line search for a second physical property, using the initial model, and saving the optimal step size from the second line search but not updating the model yet;
 - (iii) performing a line search using the initial model and saving the optimal step size as in (ii) for any additional physical properties being modeled; and
 - (iv) using the saved step sizes from (i)-(iii) to compute a new search direction, then performing a second-pass line search along the new search direction, determining a second-pass optimal step size and using it to generate an updated model for a current iteration of the inversion.
2. The computer-implemented method of claim 1, wherein in (a) one or more parts of the seismic data are selected and used for computing the gradient for each physical property, said one or more parts being determined according to a selected criterion based on sensitivity of the physical property to the data.
3. The computer-implemented method of claim 2, further comprising, in a next iteration of the iterative inversion, expanding the selected part or parts of the seismic data for at least one of the at least two physical properties to include more parts of the seismic data.

4. The computer-implemented method of claim 1, wherein the first and second and any additional line searches in (iii) are exhaustive line searches, performed to find inflection points in the objective function as a function of the step size.
5. The computer-implemented method of claim 1, wherein some source shots in the seismic data are discarded and not used in the method in order to reduce computing cost and time.
6. The computer-implemented method of claim 1, wherein the data inversion is full wavefield inversion.
7. The computer-implemented method of claim 1, wherein the at least two physical properties comprise at least two of P-wave velocity, S-wave velocity, normal move-out velocity, one or more anisotropy parameters, and attenuation.
8. The computer-implemented method of claim 1, wherein the search direction is computed without preconditioning the gradient with a Hessian of the objective function.