Homogeneous Green's functions as a dictionary for wavefield reconstruction

Paolo Terenghi, PGS*

Summary

Techniques for wavefield interpolation, extrapolation and regularization often rely on general-purpose dictionaries originating in the discipline of signal or image processing. However, the linearity of the Helmholtz equation suggests that a dictionary composed of homogeneous Green's functions may be specifically suitable for the reconstruction of the acoustic wavefields relevant to seismic exploration. Through such dictionary, data can be linearly mapped into a dual domain interpreted as an equivalent source distribution. The theoretical argument behind this mapping operation is constructed from first principles in this document. The scheme is then applied to the interpolation of a set of simple synthetic data.

Introduction

The quest for higher imaging resolution and the move towards fully 3-dimensional data-processing workflows justify the current interest for interpolation, extrapolation and regularization techniques. The ultimate goal is to produce 3-dimensional images and reservoir models which are free of imprint from the acquisition devices employed in the field.

In the discipline of compressive sensing it is common to formulate the reconstruction of a desired quantity, based on coarsely and non-uniformly distributed field measurements, as a linear problem $\mathbf{d} = L\mathbf{m}$, where $\mathbf{d}$ and $\mathbf{m}$ represent the data and model vectors, respectively, and L the kernel matrix which implements a linear relationship between them.

The columns of L contain the individual analysing functions acting as individual elements or "words" in the dictionary. Within the workings of the matrix-vector product $L\mathbf{m}$, each coefficient in $\mathbf{m}$ always multiplies the same function and can therefore be interpreted as the coefficient relative to that specific element in the dictionary.

The model vector can be seen as a representation of the data in a dual domain. If the problem is suitably parameterised, the dictionary describes a "complete" transform, so that any signal of interest can be mapped into its dual form, and vice-versa. In case of aliasing, the information in $\mathbf{d}$ is insufficient to characterise the model $\mathbf{m}$ uniquely and two or more analysing functions may represent the measurements equally well.

The approach of compressive sensing in this case is to privilege the model containing the least information. Such preference is implemented by specific sparse solvers (e.g. Scales and Gersztenkorn, 1988; Mallat and Zhang, 1993; van den Berg and Friedlander, 2008; Becker et Al., 2009), which effectively introduce additional information through the assumption that the model is small in some sense. The extent to which such assumption is satisfied determines the scheme's potential to solve the problem beyond classical aliasing rules. The ideal dictionary is therefore one which can represent the measured data using the least amount of information (e.g. a small number of coefficients).

Since the onset of compressive sensing, many such dictionaries have been proposed (e.g. Mallat, 1998; Candes and Donoho, 2000), many of them originating in the context of image processing and signal compression. In many cases the basis functions in those dictionaries are devoid of any physical interpretation as they are specifically engineered to mimic certain macroscopic features of the signal to be analysed.

This work proposes a less generic dictionary with close ties to the acoustic wave equation and therefore specific to the reconstruction of seismic wavefields. It will become apparent that such dictionary implies utilising concepts and ideas akin to seismic imaging rather than to signal or image processing.

Theory

Let G_0 be the Green's function for the homogeneous unbounded medium characterized by a constant velocity c_0. G_0 can be seen as the wavefield, measured at $\mathbf{r}$, produced by an isotropic impulsive point-source located at the origin

$$[\nabla^2 + \omega^2/c_0^2]\, G_0(\omega, \mathbf{r}) = -\delta(\mathbf{r})$$

In the same homogeneous medium, the wavefield P_0 produced by a multiplicity of point-sources satisfies

$$[\nabla^2 + \omega^2/c_0^2]\, P_0(\omega, \mathbf{r}) = \rho_0(c_0, \mathbf{r}) A(\omega)$$

where A is the signature (wavelet) and ρ_0 is the function describing the location and magnitude of each point-source. By the linearity of the wave operator, the following relationship holds between the fundamental wavefield G_0 and the composite wavefield P_0

$$P_0(\omega, \mathbf{r}) = \int \mathrm{d}r'\; \rho_0(c_0, \mathbf{r}')\, G_0(\omega, \mathbf{r} - \mathbf{r}')$$

Although the support for ρ_0 is theoretically infinite (all space), in exploration seismology wavefields are triggered by man-made sources positioned at known locations, whose effects are causal and presumably vanish at great distance. Contributions for large $\mathbf{r}$ can then be ignored, so the problem can be discretized and written in the desired form $\mathbf{d} = L\mathbf{m}$.

DOI http://dx.doi.org/10.1190/segam2014-1212.1

However, the concepts so far laid out refer to an ideal homogeneous unbounded medium with scarce similarity with the scenarios in exploration seismology. Nevertheless, using a simple perturbative approach, it is straightforward to show how the linearity principle applies to media with arbitrary variations in space $c(\mathbf{r})$ and to wavefields satisfying a wave equation of a more general kind, such as

$$[\nabla^2 + \omega^2/c(\mathbf{r})^2]\, P(\omega, \mathbf{r}, \mathbf{r_s}) = -\delta(\mathbf{r} - \mathbf{r}_s) A(\omega)$$

The previously introduced single velocity parameter c_0 now describes a conventional reference medium. A scattering potential α can be established as a squared slowness perturbation with respect to the reference medium (Weglein et al., 2003):

$$1/c^2(\mathbf{r}) = 1/c_0^2\ (1 - \alpha(\mathbf{r}))$$

For the common case of marine seismic exploration c_0 can be conveniently chosen to agree with sea-water properties, so that α is zero at least in the vicinity of the receivers. The wave equation can then be rewritten in a form suggesting that an equivalent source distribution exists which would cause, in a homogeneous medium, the same wavefield as a point-source would in a heterogeneous medium:

$$[\nabla^2 + \omega^2/c_0^2]\, P(\omega, \mathbf{r}) = -\delta(\mathbf{r}) A(\omega) + \omega^2/c_0^2\, \alpha(\mathbf{r})\, P(\omega, \mathbf{r})$$

The distribution can be chosen to include or exclude the portion of the wavefield which radiates directly from the physical source, through the background medium, to the receivers. Either way, it is possible to rewrite the wave equation in the form

$$[\nabla^2 + \omega^2/c_0^2]\, P(\omega, \mathbf{r}) = A(\omega) \rho(c_0, \mathbf{r})$$

It should be noted that the Helmholtz operator on the LHS is identical to the homogeneous case. It is therefore possible to conclude that the wave equation admits solutions of the type

$$P(\omega, \mathbf{r}) = \int \mathrm{d}r'\ \rho(c_0, \mathbf{r}')\, G_0(\omega, \mathbf{r} - \mathbf{r}')$$

where G_0 is the homogeneous Green's function introduced earlier. The wavefield propagating in a medium with lateral variations is thus seen as a weighted superposition of the Green's functions related to the reference medium.

Again, the current expression complies with the general definition of a linear compressive sensing problem $\mathbf{d} = L\mathbf{m}$, where the analysing functions (the columns of L) identify with individual Green's functions and ρ is the dual space (the $\mathbf{m}$ vector). However, the potential of a reconstruction scheme based on these concepts depends on whether the dictionary produces a sparse representation of the wavefield.

The *method of images* applied to Green's functions (Morse and Feshbach, 1956, ch 7) suggests the equivalence between the reflection produced by a plane boundary (in an otherwise homogeneous medium) and the effect of an additional isolated point source at a "mirrored" location across the boundary. By analogy, since G_0 is a solution to the wave equation for the homogeneous reference medium, it is reasonable to infer that ρ will be sparse only for reflections where the overburden behaves approximately as the reference medium (e.g. the water bottom reflection).

Furthermore, that very same argument justifies the definition of an augmented problem

$$P(\omega, \mathbf{r}) = \int \mathrm{d}c \int \mathrm{d}r'\ \rho(c, \mathbf{r}')\, G_c(\omega, \mathbf{r} - \mathbf{r}')$$

where P receives contributions from Green's functions at different velocities. Note that the additional velocity scan increases the algorithm's complexity, but does not infringe its linearity, so the expression still complies with the framework $\mathbf{d} = L\mathbf{m}$. It is worth emphasizing that the proposed methodology does not require any knowledge of the medium's velocities. On the contrary, the estimated ρ can be analysed to determine best-fitting replacement velocities for reflected events.

Related geophysical algorithms

Wavefield reconstruction applications based on the scattering potential have been known for several years. The term itself refers to a generic quantity which can be said to hold a linear relationship with the observed data. Where the application uses data from all sources and all receivers (e.g. Stolt, 2002, Kaplan , 2010, ch. 4; Kutsha et al., 2010), the scattering potential naturally relates to the seismic image and thus directly to the earth's structures and properties.

The algorithms derived under this setup are very powerful, but also very costly, as they require all physical experiments to enter a single numerical optimization.

In other cases, the scheme only involves data acquired during a single physical experiment (Trad, 2003; Kaplan, 2010, ch. 6). In this domain, the extent to which the potential can be focused is limited.

The methodology presented here belongs to this second category, where the scattering potential is simply seen as a convenient equivalent representation of the data, devoid of any attempts to resolve medium properties.

Numerical example

The proposed methodology is tested using a synthetic wavefield composed of hyperbolic events representing the seismic responses of 3 plane reflectors, as summarized in

DOI http://dx.doi.org/10.1190/segam2014-1212.1

Table 1. The wavefield is uniformly sampled with a trace spacing of 12.5m in the offset interval between -500m and 2500m (Figure 2b) and conditioned to contain temporal frequencies up to approximately 60Hz (Figure 3b). The fully sampled wavefield is then uniformly under-sampled by a factor of 10 (one trace is kept every 10) to form the input dataset (Figure 2a), where temporal frequencies above 12.5Hz are spatially aliased (Figure 3a).
The proposed scheme is parameterised to scan the model space for the offset range between -1000 and +4000m, for all depths compatible with the trace's temporal duration, and for velocities between 1300 and 2150m/s. The equivalent source distribution (ρ as defined above) represented in Figure 1 is obtained using a weighted conjugate gradient solver configured according to the Iteratively Reweighted Least Squares scheme (IRLS, Scales and Gersztenkorn, 1988).
It can be observed that each of the hyperbolic reflections maps to a small number of coefficients located in a constrained portion of the model space. The horizontal and vertical locations of the equivalent source responsible for each of the three reflections are particularly well-resolved (Figure 1b) whereas the replacement velocity can be estimated within a confidence range of approximately 100m/s (Figure 1a).
Application of the adjoint transformation produces the reconstructed wavefield shown in Figure 2c. The interpolation error (difference between the reconstructed wavefield and the original fully sampled data is shown in Figure 1d.

Table 1: Depth, dip angle and velocity associated with the three reflections in the test data

Reflector	depth (m)	dip (degrees)	velocity (m/s)
1	300	10	1500
2	600	20	1750
3	800	-30	2000

Conclusions

The linearity of the Helmholtz equation suggests that a dictionary composed of homogeneous Green's functions may be specifically suitable for the reconstruction of seismic wavefields through the standard approach of compressive sensing. Such dictionary maps the data to a dual space which can be described as a spatial distribution of point-sources capable of generating a wavefield identical to the one being processed. As a proof of concept, the proposed scheme was applied successfully to the reconstruction of simple synthetic wavefield after it had been decimated to a tenth of its initial size.

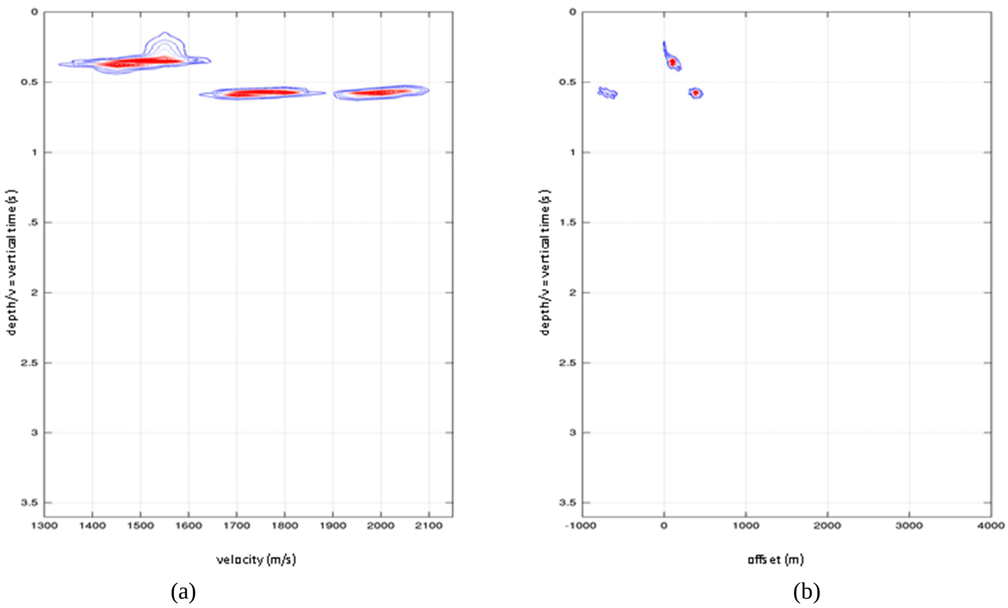
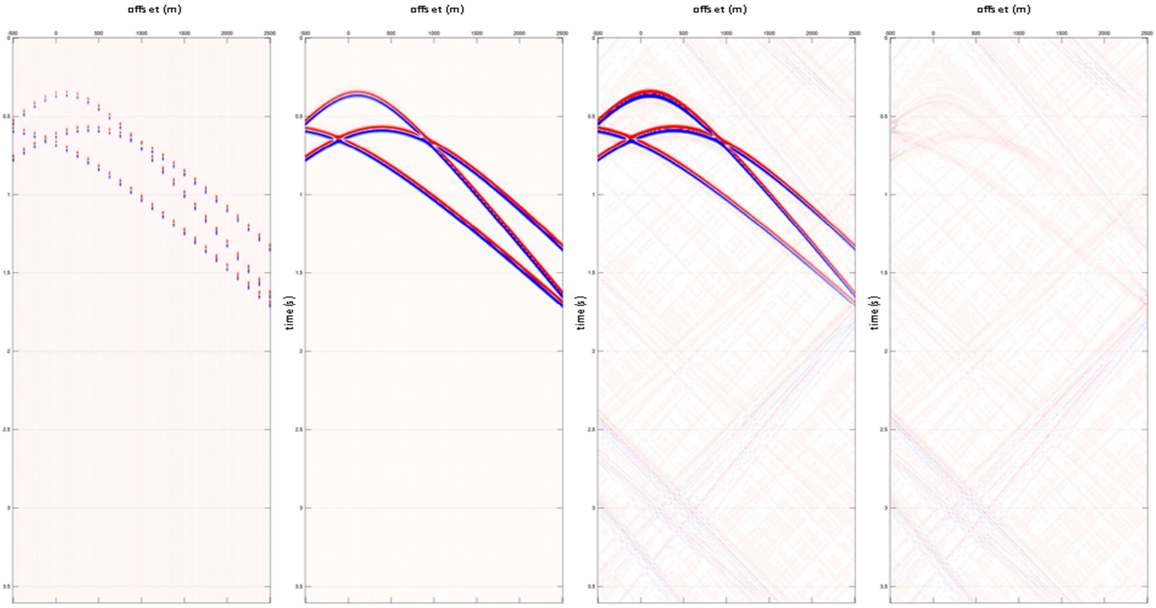


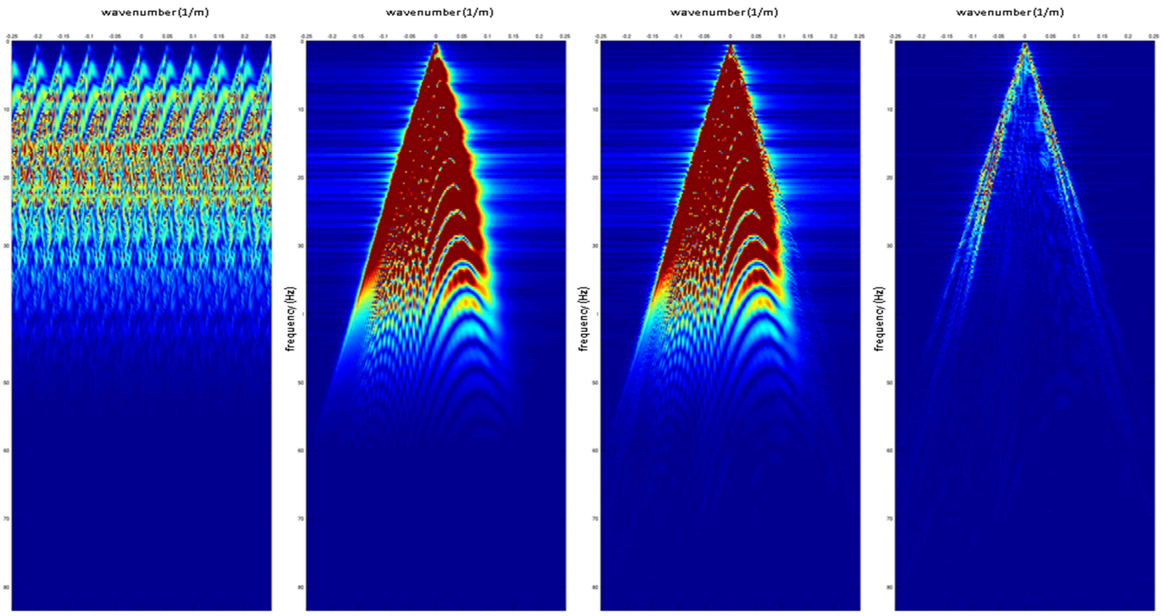
Figure 1: 2-dimensional representations of the equivalent source distribution (ESD, ρ in the main text) describing the input data and the recovered wavefield. (a) magnitude of the ESD stacked along the offset axis; (b) magnitude of the ESD stacked along the velocity axis. For ease of comparison with Figure 2, depth is represented as vertical travel time.

DOI http://dx.doi.org/10.1190/segam2014-1212.1



(a) (b) (c) (d)

Figure 2: Numerical example: (a) input data decimated to a trace spacing of $125m$; (b) fully sampled wavefield with a trace spacing of $12.5m$; (c) reconstructed wavefield, with a trace spacing of $12.5m$; (d) difference between (b) and (c) (residual).



(a) (b) (c) (d)

Figure 3: F-k spectra of quantities displayed in Figure 2.

DOI http://dx.doi.org/10.1190/segam2014-1212.1

http://dx.doi.org/10.1190/segam2014-1212.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Becker, S., J. Bobin, and E. J. Candes, 2009, NESTA: A fast and accurate first-order method for sparse recovery: Applied and computational mathematics: Caltech Pasadena.

Candes, E. J., and D. L. Donoho, 2000, Curvelets — A surprisingly effective nonadaptive representation for objects with edges: Vanderbilt University Press.

Kaplan, S. T., 2010, Regularized wave equation migration for imaging and data reconstruction: Ph.D. thesis, University of Alberta.

Kutsha, H., D. J. Verschuur, and A. J. Berkhout, 2010, High-resolution double focal transformation and its application to data reconstruction: 80th Annual International Meeting, SEG, Expanded Abstracts, 3589–3593.

Mallat, S. G., 1998, A wavelet tour of signal processing: Elsevier.

Mallat, S. G., and Z. Zhang, 1993, Matching pursuit with time-frequency dictionaries: IEEE Transactions on Signal Processing, **41**, no. 12, 3397–3415, http://dx.doi.org/10.1109/78.258082.

Morse, P. M., and H. Feshbach, 1953, Methods of theoretical physics: Feshbach Publishing.

Scales, J. A., and A. Gersztenkorn, 1988, Robust methods in inverse theory: Inverse Problems, **4**, no. 4, 1071–1091, http://dx.doi.org/10.1088/0266-5611/4/4/010.

Stolt, R. H., 2002, Seismic data mapping and reconstruction: Geophysics, **67**, 890–908, http://dx.doi.org/10.1190/1.1484532.

Trad, D., 2003, Interpolation and multiple attenuation with migration operators: Geophysics, **68**, 2043–2054, http://dx.doi.org/10.1190/1.1635058.

van den Berg, E., and M. P. Friedlander, 2009, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, no. 2, 890–912, http://dx.doi.org/10.1137/080714488.

Weglein, A. B., F. V. Araújo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coats, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse Problems, **19**, no. 6, R27–R83, http://dx.doi.org/10.1088/0266-5611/19/6/R01.