Review of "Waveform inversion via reduced order modeling", GEO22-0070

April 18, 2022

Reduced order modeling (ROM) refers to a collection of techniques for control or parameter identification of dynamical systems that approximate the dynamics by projection of the system state onto a collection of time samples, or snapshots. In effect, the collection of snapshots taken from the time history itself is used as an approximation to the full time history, and a reduced dynamics deduced that operates on this collection. The dimension (order) is smaller, perhaps much smaller, and control of the dynamics and or parameter estimation is computationally simpler that is the case with the original model. Observation of the snapshots is ordinarily a requirement.

Application of ROM to waveform inversion would seem to face an immediate obstacle: the system state is the seismic wavefield, unobservable except (perhaps) in a very few places in the earth's interior, and on or near the surface. The approach to inversion described in this work relies on two clever observations about the structure of wavefields that permit the construction of a Galerkin projection of wave dynamics onto the span of the snapshots, directly from surface data - *without access to the snapshots.* The idea has been described in a number of publications, and applied in various ways. The contribution of this paper is to show how the ROM construction can lead to velocity estimation, perhaps avoiding the notorious cycle-skipping obstacle of FWI. The paper describes in detail a velocity inversion algorihm based on ROM, and illustrates it with several examples, based on the well-known Tarantola Camembert and Marmousi 2D models.

The ROM construction requires a good deal of set-up, converting a familiar type of acoustic model for seismic data into a rather unfamiliar one, and relying (apparently) on a rather strict set of assumptions. I will first give an overview of the inversion approach, concentrating on the core constructions that make it quite different from other approaches to inversion. In effect, the result is a remarkable representation of the data in the form of a reduced-order model of the acoustic wave operator. The rest of the algorithm is involved but more conventional, estimating the velocity field by regularized least-squares fit to this data-derived reduced-order operator representation. After this overview of the methodology, I make some observations and pose some questions for the authors to consider.

1 Method

The acoustic pressure field is assumed to be a solution of the constant-density wave equation. Sources (right-hand sides) take the form form $\theta(\mathbf{x} - \mathbf{x}_s)\partial_t f(t)$, in which f has a non-negative, even Fourier transform, that is, is an autocorrelation (always achievable, assuming the the physical source is known). The spatial factor $\theta(\mathbf{x})$ is zero outside of a "small" ball centered at $\mathbf{x} = 0$, so that the sensors and sources have localized support near their nominal locations. The pressure field is scaled by $1/c(\mathbf{x})$, which has the effect of replacing the spatial operator $c(x)^2\nabla^2$ in the wave equation by the symmetric operator $\mathcal{A} = c(\mathbf{x})\nabla^2 c(\mathbf{x})$. The causal pressure field $p^{(s)}$ for the source at $\mathbf{x}_s$ is replaced by the symmetric-in-time field $w^{(s)}(\mathbf{x}, t) = (p^{(s)}(\mathbf{x}, t) + p^{(s)}(\mathbf{x}, -t))/c(\mathbf{x})$.

Data traces are defined by averaging the time-symmetrized pressure field with the same kernel θ used to define the source:

$$D^{(r,s)}(t) = \int dx \frac{\theta(\mathbf{x} - \mathbf{x}_r)}{c(\mathbf{x}_r)} w^{(s)}(\mathbf{x}, t)$$

It is assumed tht the sets $\{\mathbf{x}_s\}$ and $\{\mathbf{x}_r\}$ of source and receiver positions are the same, so that D is a square $m \times m$ matrix of traces. Note that these are not the traces of a causal pressure field resulting from localized sources, but can be computed from such "normal" traces, given the validity of the assumptions outlined above.

Finally, w is filtered by a local operator defined as a function of $\mathcal{A}$ (equation 11). This transformation (based on Duhamel's principle) converts w, which satisfies a wave equation with homogeneous initial values and a energy input through a right-hand side, into u, which satisfies a wave equation with zero right hand side and energy input through an initial condition:

$$\begin{aligned}
\partial_t^2 u^{(s)} + \mathcal{A}u^{(s)} &= 0, \\
u^{(s)}(\mathbf{x}, 0) &= u_0^{(s)}(\mathbf{x}), \\
\partial_t u^{(s)}(\mathbf{x}, 0) &= 0.
\end{aligned} \tag{1}$$

A consequence of this construction is a key relation between the data and the transformed field u (equation 16, more or less):

$$D^{(r,s)}(t) = \int dx\, u_0^{(r)}(\mathbf{x}) u^{(s)}(\mathbf{x}, t) = \langle u_0^{(r)}, u^{(s)}(t)\rangle$$

where the angle brackets on the right signify the L^2 (Euclidean) inner product. That is, averaging $w^{(s)}$ against $\theta(\mathbf{x} - \mathbf{x}_r)/c(\mathbf{x}_r)$ is equivalent to averaging $u^{(s)}$ against the $t = 0$ field $u_0^{(r)}$.

One of the fundamental ideas of this work is that you can represent the Green's function of problem 1 via the so-called functional calculus, as an operator cosine:

$$u^{(s)}(t, \mathbf{x}) = \cos(t\sqrt{\mathcal{A}}) u_0^{(s)}(\mathbf{x}) \tag{2}$$

This would be obvious if $\mathcal{A}$ were a number. It is also true for symmetric operators like $\mathcal{A}$ (officialliy, once they are made into self-adjoint operators by proper choice of domain), as a consequence of the spectral theory of such operators, as is explained in the manuscript.

After these preliminaries, the ROM is based on orthogonal projection of the vector of fields $u^{(s)}(\mathbf{x}, t), s = 0, ...m - 1$ onto the subspace spanned by the first n snapshots at time interval τ, i.e. $u^{(s)}(\mathbf{x}, j\tau), s = 0, ...m - 1, j = 0, ...n - 1$. Organize the fields into a block vector $u(\mathbf{x}, t) = (u^{(0)}(\mathbf{x}, t), ..., u^{(m-1)}(\mathbf{x}, t))$, and the snapshots into the nm block (row) vector $U(\mathbf{x})$, $U_{sm+j}(\mathbf{x}) = u^{(s)}(\mathbf{x}, j\tau)$. The block matrix projection (Galerkin approximation) may be written

$$u^{\text{gal}}(\mathbf{x}, t) = U(\mathbf{x})g(t)$$

in which $g(t) = (g_0(t), ..., g_{n-1}(t))$ is the $nm \times m$ matrix of Galerkin coefficients, chosen so that the residual in the wave operator applied to u^{gal} is orthogonal to the span of the elements of U. That is,

$$\langle U^T, (\partial_t^2 Ug + \mathcal{A}Ug)\rangle = \partial_t^2 \langle U^T, U\rangle \partial_t^2 g + \langle U^T, \mathcal{A}U\rangle g$$

$$= \mathbf{M}\partial_t^2 g + \tilde{\mathbf{S}}g = 0 \tag{3}$$

where the $nm \times nm$ Galerkin mass and stiffness matrices are

$$\mathbf{M} = \langle U^T, U\rangle = \int dx U^T(\mathbf{x})U(\mathbf{x}),$$

$$\tilde{\mathbf{S}} = \langle U^T, \mathcal{A}U\rangle$$

This is all very well, and the stiffness matrix $\tilde{\mathbf{S}}$ represents $\mathcal{A}$ and contains information about the velocity field c, though in a somewhat non-transparent way, since the snapshot vector U also depends on c. However the snapshots U are precisely the missing ingredient - they represent the acoustic field in the interior of the earth. They could be calculated if one knew c, by solving the problem 1 numerically, however of course c is precisely the unknown being sought.

This difficulty is overcome through two very clever observations. The first rests on the high-school identity

$$\cos a \cos b = \frac{1}{2}(\cos(a + b) + \cos(a - b))$$

From the Green's function representation 2, suppressing $\mathbf{x}$) and grouping the fields for all sources together into the m-vector u,

$$u_j = u(j\tau) = \cos(j\tau\sqrt{\mathcal{A}})u_0 \tag{4}$$

So the (i, j) $m \times m$ block of the $nm \times nm$ mass matrix $\mathbf{M}$ can be re-written (equation 31)

$$\mathbf{M}_{i,j} = \langle u_i, u_j\rangle = \langle \cos(i\tau\sqrt{\mathcal{A}})u_0, \cos(j\tau\sqrt{\mathcal{A}})u_0\rangle$$

$$= \langle u_0, \cos(i\tau\sqrt{\mathcal{A}}) \cos(j\tau\sqrt{\mathcal{A}})u_0\rangle$$

3

(since real functions of the symmetric operator $\mathcal{A}$ are symmetric and commute)

$$= \frac{1}{2} \langle u_0, [\cos((i+j)\tau\sqrt{\mathcal{A}}) + \cos(|i-j|\tau\sqrt{\mathcal{A}})u_0 \rangle$$

$$= \frac{1}{2}(D_{i+j} + D_{|i-j|}),$$

where we have written $D_j = D(j\tau)$.

In other words, *the mass matrix, of inner products of acoustic fields in the interior of the earth, can be calculated directly from the data.*

The same is true of the stiffness matrix $\tilde{\mathbf{S}}$, by a similar calculation (equation 49):

$$\tilde{\mathbf{S}}_{i,j} = -\frac{1}{2}((\partial_t^2 D)_{i+j} + (\partial_t^2 D)_{|i-j|}),$$

where the second time derivative traces $\partial_t^2 D$ are computed by any convenient numerical approximation.

The stiffness matrix contains information about c, but that is still hidden behind the interior fields U, in view of the defintiion. To tease it out, introduce block Gram-Schmidt orthogonalization of the vector $U = VR$. Here V has orthogonal m-dimensional time blocks, and R is an $nm \times nm$ dimensional block upper triangular block matrix. Then you can re-write the left-hand side of equation 3 as

$$U^T(U\partial_t^2 g + \mathcal{A}Ug) = R^T R\partial_t^2 g + R^T V^T \mathcal{A} V Rg = 0$$

Note in particular,

$$R^T R = \mathbf{M}. \tag{5}$$

Define the ROM fields $\tilde{u}^{\mathrm{ROM}} = Rg$. Then

$$\partial^2 \tilde{u}^{\mathrm{ROM}} + \mathcal{A}^{\mathrm{ROM}}\tilde{u}^{\mathrm{ROM}} = 0$$

where $\mathcal{A}^{\mathrm{ROM}} = V^T \mathcal{A} V$ is the "ROM of the wave operator", out of which will be extracted an estimate of c.

This still appears to be unsatisfactory, as the definition of $\mathcal{A}^{\mathrm{ROM}}$ involves the orthogonalized interior fields. Now the second very clever observation intervenes: R is block upper triangular, so equation 5 expresses the block Cholesky decomposition of $\mathbf{M}$. Thus R can be computed from $\mathbf{M}$ by a block version of Gram-Schmidt orthogonalization, for instance. But $\mathbf{M}$ is computable directly from the data, therefore the same is true of R. $\tilde{\mathbf{S}}$ is also computable directly from the data, and

$$\mathcal{A}^{\mathrm{ROM}} = V^T \mathcal{A} V = (R^T)^{-1}\tilde{\mathbf{S}}R^{-1},$$

whence $\mathcal{A}^{\mathrm{ROM}}$ *is computable directly from the data, entirely without knowledge of the interior fields.*

It now remains to extract the velocity, or an approximation, from $\mathcal{A}^{\mathrm{ROM}}$. The method used is basically straightforward nonlinear least squares with Tikhonov regularization. The

version of $\mathcal{A}^{\mathrm{ROM}}$ computed from the data as just described is treated as the data in this least squares problem. The prediction operator maps a trial velocity v to synthetic data, whence to a synthetic ROM operator $\mathcal{A}^{\mathrm{ROM}}(v)$. The Frobenius norm of the difference of these two operators is then minimized by updating v using a combination of layer-stripping and Gauss-Newton algorithms. Of course there are number of twists: for instance, the decay of the matrix entries of $\mathcal{A}^{\mathrm{ROM}}$ away from the diagonal (shown in Appendix B) is used to reduce the computational complexity of the algorithm.

2 Observations

I offer a few questions and comments.

2.1 Cycle-skipping mitigation - why?

The examples, based on Tarantola's Camembert and the Marmousi model, are impressive and certainly suggest that avoidance of cycle-skipping is somehow built into this approach to velocity inversion. However the mechanics of the ROM construction are sufficiently unfamiliar to this reviewer that he has no intuitive understanding of why this should be so. I believe the same will be true for most, or even all, readers of *Geophysics* who tackle this work. What is the mechanism by which cycle-skipping is avoided? I am not asking for proof - there are very few informative theoretical results about convexity or monomodality in the literature on wave inverse problems, and the results that do exist are mostly not so useful. I merely suggest that a plausibility argument in favor of cycle-skipping mitigation would make this paper considerably more valuable.

2.2 Data restrictions

From the outset, this paper presumes that sources and receivers occupy the same positions and are parametrized in the same way. Such assumptions are natural and more or less correct in some laboratory ultrasonics settings, in which the sources and receivers are the same transducers, and perhaps in some other settings, for example in ocean acoustics. However the same cannot be said of seismic data acquisition, where source and receiver positions often do not coincide and physical characteristics of sources and receivers (frequency response, radiation pattern,...) usually differ, often considerably. That is, Green's functions are reciprocal, but data often are not. Can this approach be modified to deal directly with the non-reciprocal data usually encountered in all forms of seismology? The algebraic rigidity of the ROM construction suggests that the answer is "no". If that is the case, then a form of data completion and direction-dependent signature modification would be required as preprocess. As a simple initial exercise without the complication of radiation pattern modification, the authors could consider what modifications would be necessary to use data in the original Marmousi streamer geometry. This is not a minor point - it bears on whether this approach can ever be practical for seismic inversion.

Note that, for all of its defects, FWI naturally accommodates more or less arbitrary data geometry and source/receiver characteristics. Of course, the latter may not be known with the necessary precision, but in fact these are nowadays often incorporated into the parameters to be determined by FWI. To become useful, the ROM construction will have to acquire some of this flexibility.

2.3 Linear independence of snapshots

A major assumption, that the snapshots are linearly independent, appears as a footnote on p. 10: "The space S is nm-dimensional because the snapshots $u_j(\mathbf{x})$ for $j = 0, .., n$ are assumed linearly independent." Linear independence is an open property, and seems reasonable to assume, but an assumption on which the entire construction rests does not belong in a footnote. Also, one must ask why this assumption is reasonable. A hint appears on p. 31, in a footnote to the statement "Because in the ROM construction we need the wavefront to progress downward at all times..."; the footnote is "If this does not hold, the snapshots that define the approximation space S may be linearly dependent." That is extremely interesting. First, what wavefronts? Does the remark refer to the wavefronts of geometric asymptotics, i.e. level sets of geodesic distance, emitted by the sources? In that case, it sounds like the so-called "DSR" assumption underlying one-way prestack migration. The Marmousi model produces many overturned wavefronts, yet the ROM-based inversion succeeds. Is it only some part of the geometric wavefronts that must "progress downwards"? Or does this statement not pertain to geometric wavefronts at all? This is a fundamental point, and the authors really need to clear it up.

2.4 Computational complexity

The authors do not directly address this topic at all. The algorithm involves a number of disparate subsidiary calculations - time domain finite difference simulation, QR factorization of large matrices, Gauss-Newton iteration, etc. The reader cannot be expeccted to estimate the computational cost of this algorithm or the factors on which it depends. The authors should add a serious attempt to assess the cost - maybe per Gauss-Newton itertion in each layer computation with reasonable off-diagoal truncation, and some rough estimate of overall cost and its critical parameters.

The straw-man algorithm, FWI, is now a commercial product, applied many times to 3D field data. What is a rough estimate of cost for a 3D ROM inversion?

2.5 Audience

As currently written, comprehending this paper requires some mathematical background, at the level of basic graduate courses in functional and numerical analysis. The vast majority of *Geophysics* readers have not acquired this background. Some attempt to compensate for this disparity would likely enlarge the pool of interested readers. For example, a graphical repre-

sentation of block Cholesky decomposition would give the reader with minimal background in numerical linear algebra a fighting chance.

2.6   Miscellaneous

The construction of the ROM propagator is really irrelevant, and could usefully be removed.

Can this construction be applied to transmission data? For instance, what about one of the other Camembert experiments from the Gauthier et al. paper, in which the receivers are on the opposite side (in addition, to keep the coincident source-receiver hypothesis valid)?

3   Conclusion

This paper presents truly interesting and original work, using concepts from numerical linear algebra and functional analysis to recast the velocity inversion problem in a form that appears to avoid cycle-skipping, at least in some cases. In its current form, it has two major drawbacks as a contribution to *Geophysics*: it presumes background that most readers do not have, and makes assumptions about data that are simply not true of almost all data collected in the field. If the authors address the questions and comments I have listed here, they will go a long ways towards making a revised version suitable for publication in *Geophysics*.