Using the Iterative Ensemble Kalman Smoother for Seismic Waveform Inversion

Michael Gineste and Jo Eidsvik

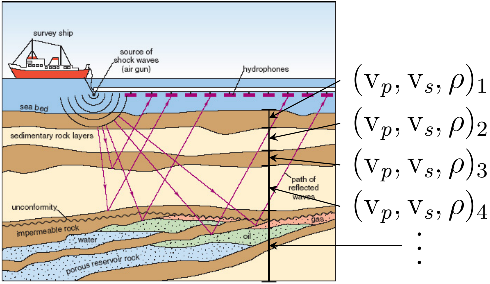
NTNU, Trondheim

June 2021

Outline

Seismic waveform inversion

- Seismic data $\mathbf{y}$ are wave reflections from subsurface model.
- Elastic parameters $\mathbf{x} = [\mathbf{v}_p, \mathbf{v}_s, \boldsymbol{\rho}]$.
- Information on elastic attributes are in waveform amplitude and phase.



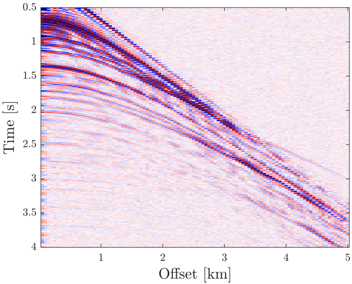
Marine seismic survey[1]

Seismic waveform data

- Data corrupted by noise $\rightarrow$ observation model $\mathbf{y} = h(\mathbf{x}) + \mathbf{e}, \mathbf{e} \sim \mathcal{N}(\mathbf{0},\ \mathbf{R})$.
- Nonlinear forward model $h(\mathbf{x})$: seismic wave propagation.
- $h(\mathbf{x})$ use much computation resources / time.
- Data dimension ~ 1000000

Reflectivity method

- Layered subsurface assumption (1.5D)
- Reflectivity method $h(\mathbf{x})$, a solution to elastic wave equation.



Seismic common midpoint gather

Probabilistic inversion

Bayes' rule

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x}) \ .$$

Observations model likelihood

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\tfrac{1}{2}\left\|\mathbf{y}^{\mathrm{o}} - h(\mathbf{x})\right\|_{\mathbf{R}}^{2}\right) \ .$$

Gaussian prior/forecast $\mathbf{x} \sim \mathcal{N}\left(\mathbf{x}^{\mathrm{f}}, \mathbf{P}_{\mathrm{f}}\right)$

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\tfrac{1}{2}\left(\left\|\mathbf{y}^{\mathrm{o}} - h(\mathbf{x})\right\|_{\mathbf{R}}^{2} + \left\|\mathbf{x} - \mathbf{x}^{\mathrm{f}}\right\|_{\mathbf{P}_{\mathrm{f}}}^{2}\right)\right) \ .$$

Nonlinear forward model, no closed form available.

No fluid dynamical model here. No time-lapse seismic data.

Outline

Ensemble representation

Ensemble

$$\mathbf{E}^{\mathrm{f}} = \left[\, \mathbf{x}^{\mathrm{f}}_{[1]} \; \mathbf{x}^{\mathrm{f}}_{[2]} \; \cdots \; \mathbf{x}^{\mathrm{f}}_{[n]} \,\right] \text{ member } \mathbf{x}^{\mathrm{f}}_{[i]} \sim p(\mathbf{x})\,.$$

Ensemble mean and covariance

$$\bar{\mathbf{x}}^{\mathrm{f}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}^{\mathrm{f}}_{[i]},\, \mathbf{P}_{\mathrm{f}} = \mathbf{X}_{\mathrm{f}}\mathbf{X}_{\mathrm{f}}^{\mathrm{T}} \text{ where } \mathbf{X}_{\mathrm{f}} = \left(\mathbf{E}^{\mathrm{f}} - \bar{\mathbf{x}}^{\mathrm{f}}\,\mathbf{1}^{\mathrm{T}}\right)/(n{-}1)^{1/2}.$$

Analysis state is linear combination in ensemble subspace

$$\mathbf{x}^{\mathrm{a}} \in \left\{\bar{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\,\mathbf{w} \,\middle|\, \mathbf{w} \in \mathbb{R}^{n}\right\}\,.$$

Square root version of EnKF. Update mean and anomaly matrix separately

$$\mathbf{x}^{\mathrm{a}} = \bar{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\,\mathbf{w}^{\mathrm{a}} \text{ and } \mathbf{X}_{\mathrm{a}} = \mathbf{X}_{\mathrm{f}}\,\mathbf{T}\,.$$

Iterative Ensemble Kalman Smoother

- Change of variable cost function

$$J(\mathbf{w}) = \frac{1}{2}\left\|\mathbf{y}^{\mathrm{o}} - h(\bar{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\,\mathbf{w})\right\|_{\mathbf{R}}^{2} + \frac{1}{2}\left\|\mathbf{w}\right\|^{2}.$$

- $\mathbf{w}^{\mathrm{a}} = \arg\min_{\mathbf{w}} J(\mathbf{w})$, transform matrix $\mathbf{T} = \mathbb{H}^{-1/2}\big|_{\mathbf{w}^{\mathrm{a}}}$.
- Gauss-Newton iteration $\mathbf{w}_{j+1} = \mathbf{w}_j - \mathbb{H}_j^{-1}\nabla J_j$.
  Sensitivities ∇J_j and $\mathbb{H}_j$ via iterate ensemble evaluation, centered at mean $\mathbf{x}_j = \mathbf{x}(\mathbf{w}_j)$.
- SVD of $\mathbf{R}^{-1/2}\mathbf{Y} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}$, where $\mathrm{diag}(\boldsymbol{\Sigma})_i = \lambda_i$ and
  $\Delta\tilde{\mathbf{y}} = \mathbf{R}^{-1/2}(\mathbf{y}^{\mathrm{o}} - \bar{\mathbf{y}}^{\mathrm{f}}) \rightarrow$

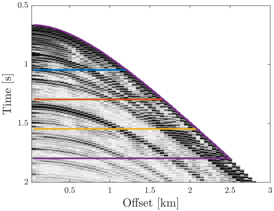
$$\Delta\mathbf{w}_{j+1} = \sum_{i=1}^{n} \mathbf{v}_i \left(\frac{-\mathbf{v}_i^{\mathrm{T}}\mathbf{w}_j}{1+\lambda_i^2} + \frac{\lambda_i(\mathbf{u}_i^{\mathrm{T}}\Delta\tilde{\mathbf{y}})}{1+\lambda_i^2} \right) = \Delta\mathbf{w}_{x,j} + \Delta\mathbf{w}_{y,j}\,.$$

Initial step analysis

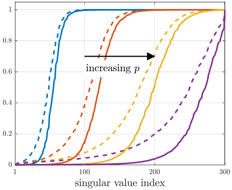
$$\Delta \mathbf{w}_1 = \sum_{i=1}^{n} \mathbf{v}_i \left(\frac{-\mathbf{v}_i^{\mathrm{T}} \mathbf{w}_0}{1+\lambda_i^2} + \frac{\lambda_i (\mathbf{u}_i^{\mathrm{T}} \Delta \tilde{\mathbf{y}})}{1+\lambda_i^2} \right) = \Delta \mathbf{w}_x + \Delta \mathbf{w}_y \; .$$

- Coefficients of vector components $\Delta \mathbf{w}_x$ and $\Delta \mathbf{w}_y$ are weighting of prior and likelihood projection coefficients.
- This weighting depends on the data size.

Data size and eigenvalues



Increasing data dimension p

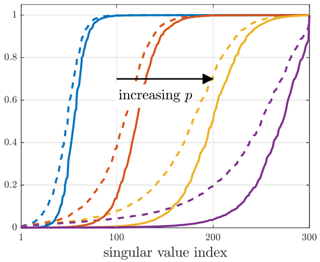


$(1+\lambda_i^2)^{-1}$ and $(1+\lambda_i^2)^{-1/2}$

Ensemble update look

Update mean and anomaly matrix

$$\mathbf{x}^{\mathrm{a}} = \bar{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\mathbf{w}^{\mathrm{a}} \text{ and } \mathbf{X}_{\mathrm{a}} = \mathbf{X}_{\mathrm{f}}\mathbf{T}\,.$$



Outline

Sequential estimation

Batch/sequential processing:

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{x}) \prod_{k=1}^{K} p(\mathbf{y}_k|\mathbf{x})$$

Sequential estimation:

$$p(\mathbf{x}) \rightarrow p(\mathbf{x}|\mathbf{y}_1) \rightarrow p(\mathbf{x}|\mathbf{y}_1, \mathbf{y}_2) \rightarrow \ldots \rightarrow p(\mathbf{x}|\mathbf{y}_1, \ldots, \mathbf{y}_K)$$

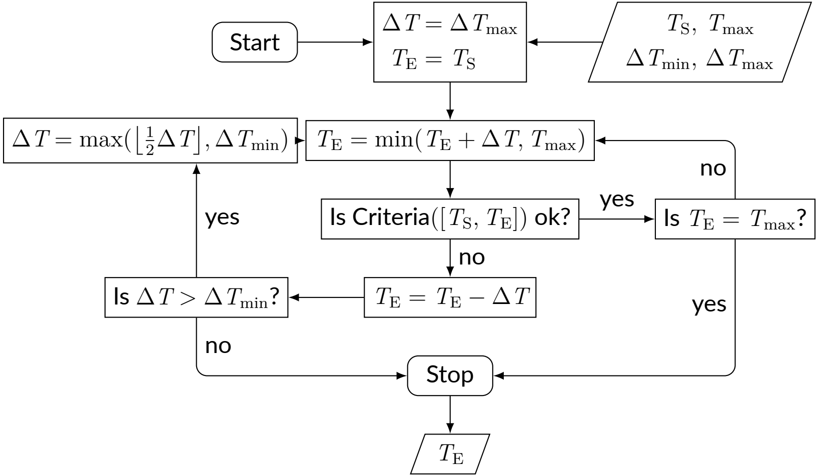
- Breaks down full depth dependency into intervals $\rightarrow$ inversion works top-down.

Sequential estimation

- Reduces nonlinearity $\rightarrow$ facilitates ensemble-linearization.
- Reduces amount of data in each conditioning step $\rightarrow$ prevents overfitting.
- Reduces the tendency to go into wrong posterior modes.
- Must balance the batch approach with time consuming reflectivity method runs.

Batch size selection

Flowchart



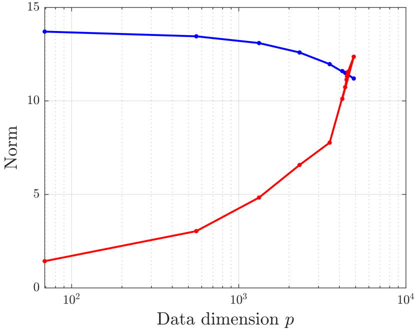
Batch size selection

Norm criteria

- Initially $\mathbf{w}_0 = \mathbf{0}$ so $\Delta\mathbf{w}_x = \mathbf{0}$.
- $$\widehat{\mathbf{a}}_i = \frac{1}{B}\sum_{b=1}^{B}\left|\frac{-\mathbf{v}_i^{\mathrm{T}}\mathbf{w}^b}{1+\lambda_i^2}\right|, \quad \mathbf{w}^b \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$
- Criteria in batch selection: $\widehat{\|\Delta\mathbf{w}_x\|}/\|\Delta\mathbf{w}_y\| \leq \beta$.

Batch size selection

Norm criteria visualization

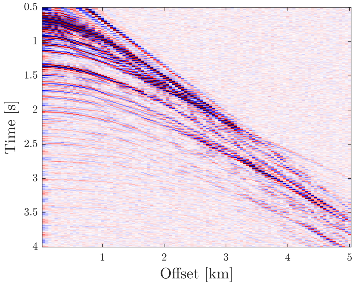


Prior $\|\Delta \mathbf{w}_x\|$ and likelihood $\|\Delta \mathbf{w}_y\|$ for $\beta = 1$.

Outline

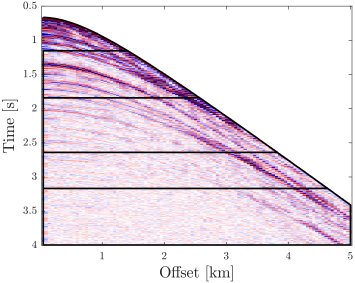
Results on a case

- Upscaled well is used as ground truth.
- Reflectivity method $h(\mathbf{x})$ and observation noise model used to generate data.
- Goal is to infer the truth, with uncertainties, from common mid point gather data.



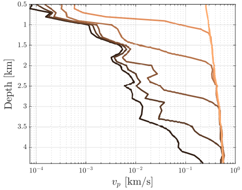
Seismic common midpoint gather

Results

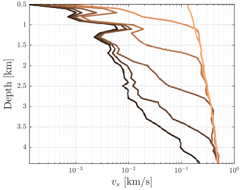


Partitioning windowed batches of seismic data.

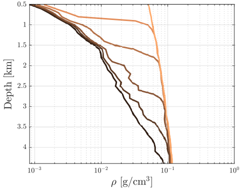
Results



Acoustic velocity



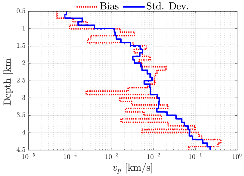
Shear velocity



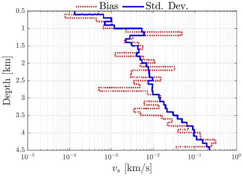
Density

Ensemble **standard deviation** over analysis cycles.
Order is from lightest (initial ensemble) to darkest (final analysis).

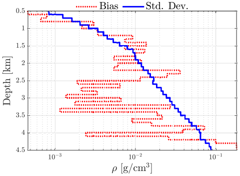
Results



Acoustic velocity



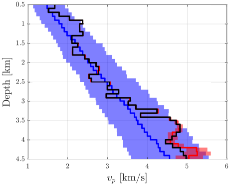
Shear velocity



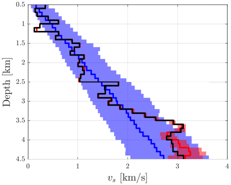
Density

Estimation bias and standard deviation.

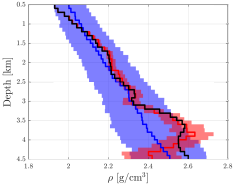
Results



Acoustic velocity



Shear velocity



Density

Prior and Posterior and Truth.

Outline

Summary

- Seismic waveforms are complicated data. Non-uniqueness (multimodal posterior) is an inherent problem with seismic inversion. Ensemble-based method does not (directly) support such solutions.
- Use of iterative scheme and norm criteria for adaptive data assimilation window gives less tendency of wrong mode.
- Many more areas to look into; colored noise, model error (layering), prior specification, etc.

Gineste, M. and J. Eidsvik (2021). "Batch seismic inversion using the iterative ensemble Kalman smoother". In: *Computational Geosciences*.

Gineste, M., J. Eidsvik, and Y. Zheng (2020). "Ensemble-based seismic inversion for a stratified medium". In: *GEOPHYSICS*.

Thank you for your attention!