DIMENSIONALITY REDUCTION AND UNCERTAINTY QUANTIFICATION FOR SEISMIC INVERSION

Tristan van Leeuwen
Mathematical Institute, Utrecht University, Utrecht, the Netherlands.

Key words full waveform inversion, uncertainty quantification, co-variance matrix, dimensionality reduction

In seismic waveform inversion, the goal is to estimate subsurface parameters, $\mathbf{m}$, from a collection of data $\{\mathbf{d}_i\}_{i=1}^k$ from various sources, indexed by i. Here, a single datum $\mathbf{d}_i \in \mathbb{R}^n$ represents a collection of n measurements, for example recorded by various stations or geophones. These recordings are modeled as noisy versions of the true signal, i.e.,

$$\mathbf{d}_i = F_i(\mathbf{m}) + \mathbf{n}_i,$$

where $F_i(\mathbf{m})$ is the (deterministic) forward operator, parametrized by $\mathbf{m}$ and $\mathbf{n}_i$ is a (stochastic) noise term. If these noise terms are Gaussian with mean zero and known covariance C (i.e., $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, C)$) we arrive at the well-known non-linear least-squares formulation. The problem is that the covariance of the noise is typically not known a-priori. If there is enough data, i.e. $k > n$, we can formulate an extended problem that allows for estimation of both C and $\mathbf{m}$

$$\min_{\mathbf{m},C} \log\left(|C|\right) + \frac{1}{n}\sum_{i=1}^{k} \|F_i(\mathbf{m}) - \mathbf{d}_i\|_{C^{-1}}^2, \tag{1}$$

where $\|\mathbf{r}\|_A^2 = \mathbf{r}^T A \mathbf{r}$ is a weighted norm. The minimization over C has a closed-form solution

$$\widehat{C} = \frac{1}{n}\sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^T,$$

where $\mathbf{r}_i = F_i(\mathbf{m}) - \mathbf{d}_i$, allowing one to design an efficient alternating optimization algorithm for small-scale problems. Such an approach is based on a singular-value-decomposition (SVD) of the residual matrix $R = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k] = USV^T$, allowing one to express the covariance matrix as $C = n^{-1}RR^T = US^2U^T$. This, in turn, allows us to compute its inverse and determinant needed for the evaluation of the objective in (1) and its derivatives and apply a standard gradient-based method to solve for $\mathbf{m}$.

For large-scale problems, however, it is not feasible to compute a full SVD of the residual matrix. Instead, we propose to first reduce the dimensionality of the residual matrix by compressing it with a random (Gaussian) matrix $W \in \mathbb{R}^{k\times k'}$ with $k' \ll k$ to get $\widetilde{R} = RW$. Note that we can perform such a compression in practice at the cost of only k' applications of the forward model if the principle of superposition holds. If we normalize the random matrix we can approximate the covariance matrix as $\widehat{C} \approx n^{-1}\widetilde{R}\widetilde{R}^T$. This is an unbiased approximation, i.e., $\mathbb{E}(\widetilde{R}\widetilde{R}^T) = nC$, where $\mathbb{E}$ denotes expectation over W. The compressed residual matrix is much smaller, allowing us to compute a (truncated) SVD and proceed with the algorithm as stated above. Preliminary results are shown in figure 1.

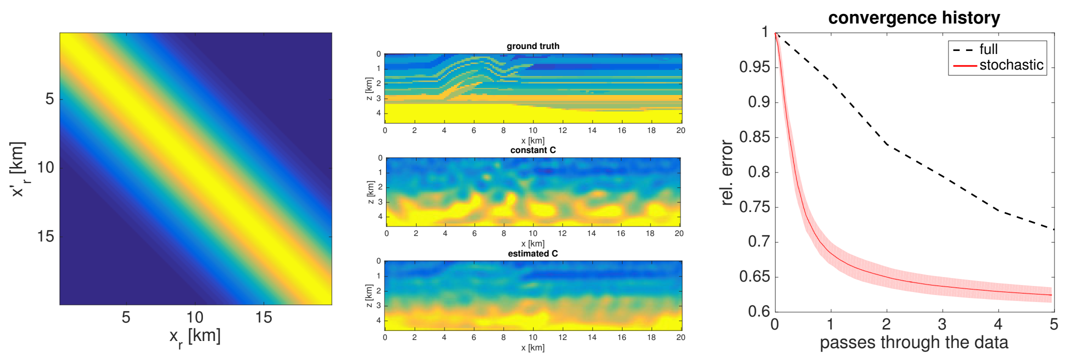

Figure 1. The Gaussian noise is generated according to the covariance matrix (left), inversion of the data for with constant covariance leads to a very poor reconstruction, while estimating the covariance on-the-fly gives a reasonable result (middle). The convergence plot (right) shows that the stochastic approach ($k' = 10$) converges much faster than the naive approach ($k = 200$), leading to an order of magnitude in computational savings.