Adjoint problem ensemble algorithms for inverse modeling of advection-diffusion-reaction processes

A.V. Penenko, Z.S. Mukatova, A.B. Salimova

Institute of Computational Mathematics and Mathematical Geophysics SB RAS
Novosibirsk State University

EGU General Assembly 2019,
Vienna (Austria), 7–12 April 2019
Motivation

• The progress in nonlinear ill-posed operator equation solution and analysis methods (different regularization methods, SVD, convergence theory, etc.)

• The progress in the parallel computations technologies: the speedup is achieved through the intensive parallelization (ensemble algorithms, splitting, decomposition, etc.)

• Variety of applications for the inverse and data assimilation problems for advection-diffusion-reaction models. E.g.
  • Air quality studies (environmental applications)
  • Morphogen theory (developmental biology)

• Image-type measurement data in air quality applications (large volume of data with unknown value w.r.t. the considered inverse modelling task):
  • Time-series (in situ)
  • Vertical concentration profiles (aircraft sensing, lidar profiles, etc).
  • Satellite images (total column 2D images).
### Advection-diffusion-reaction model

The domain \( \Omega_T = \Omega \times [0,T] \)

\[
\frac{\partial \varphi_l}{\partial t} - \nabla \cdot \left( \text{diag} \left( \mu_l \right) \nabla \varphi_l - \mathbf{u} \varphi_l \right) + P_l(t, \varphi, y) \varphi_l = \Pi_l(t, \varphi, y) + f_l + r_l, \quad l = 1, \ldots, N_c
\]

Model scale: 0D,2D

\( \Omega \) rectangular in (0D,)1D,2D

### Advection-diffusion

\[
\frac{\partial \varphi_l}{\partial t} - \nabla \cdot \left( \text{diag} \left( \mu_l \right) \nabla \varphi_l - \mathbf{u} \varphi_l \right)
\]

\( \beta_l \varphi_l = \alpha_l^R, \quad (x, t) \in \Gamma_{\text{out}} \subset \partial \Omega \times (0,T], \)

\( \varphi_l = \alpha_l^D, \quad (x, t) \in \Gamma_{\text{in}} \subset \partial \Omega \times (0,T], \)

<table>
<thead>
<tr>
<th>Direct problem</th>
<th>Linear measurement operators, e.g.</th>
<th>Inverse problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varphi : {R \times Y \rightarrow \Phi } )</td>
<td>• Pointwise concentrations&lt;br&gt;• Total column 2D images&lt;br&gt;• Vertical profiles&lt;br&gt;Subspace ( \text{Span } U_{\text{mes}} )</td>
<td>( \mathbf{I} = \text{Pr} \varphi [\mathbf{r}^{(<em>)}, \mathbf{y}^{(</em>)}] + \delta \mathbf{I}, \quad U_{\text{mes}} )</td>
</tr>
<tr>
<td><strong>Given</strong></td>
<td><strong>To find (or)</strong></td>
<td><strong>Noise</strong></td>
</tr>
</tbody>
</table>
Adjoint problem

Lagrange type identity (sensitivity relation)
\[
\langle h, \delta \Phi \rangle_{\Phi} = \langle \delta r, \Psi[h] \rangle_{R} + \langle \delta y, K(t, \varphi^{(2)}, y^{(2)}, \varphi^{(1)}, y^{(1)}) \odot \Psi[h] \rangle_{Y}
\]

Sensitivity functions
\[
\varphi^{(m)} = \varphi[r^{(m)}, y^{(m)}]
\]

\[
K(t, \varphi^{(2)}, y^{(2)}, \varphi^{(1)}, y^{(1)}) \odot = \nabla_{y} \Pi(t, \varphi^{(2)}; y^{(2)}, y^{(1)}) \odot - \nabla_{y} P(t, \varphi^{(2)}; y^{(2)}, y^{(1)}) \odot \text{diag}(\varphi^{(1)}),
\]

Adjoint problem: Given \( h, \varphi^{(m)}, y^{(m)}, m = 1, 2 \), find \( \Psi \):

\[
- \frac{\partial \Psi_{l}}{\partial t} - u \cdot \nabla \Psi_{l} - \nabla \cdot \left( \text{diag} (\mu) \nabla \Psi_{l} \right) + \left( G(t, \varphi^{(2)}, y^{(2)}, \varphi^{(1)}, y^{(1)}) \Psi \right)_{l} = h_{l},
\]

\[
G(t, \varphi^{(2)}, y^{(2)}, \varphi^{(1)}, y^{(1)}) = \text{diag} \left( P(t, \varphi^{(2)}, y^{(2)}) \right) + \nabla \cdot \left( \text{diag} \left( \varphi^{(1)} \right) - \nabla \Pi(t, \varphi^{(2)}, \varphi^{(1)}; y^{(1)}) \right) \text{diag} \left( \varphi^{(1)} \right),
\]

\[
+ \text{adjoint problem boundary conditions}
\]

\[ \boxed{\text{TC: } \Psi(T) = 0} \]

Linear parametrizations
\[
\delta r = \sum_{m} \beta_{m} \delta r_{m} \quad \langle h, \delta \Phi \rangle_{\Phi} = \sum_{m} \beta_{m} \langle \delta r_{m}, \Psi[h] \rangle_{R}
\]
Gradient algorithms
(inverse source problem)

Given the cost function

\[ J(r) = \sum_{l \in L_{mes}} \| \varphi_l[r] - I_l \|_{L_2(\Omega_T)}^2 \rho_l. \]

if the parameters are smooth enough, then

\[ \varphi[r] \mapsto h = \left\{ \begin{array}{ll}
2(\varphi_l[r] - I_l), & l \in L_{mes} \\
0, & l \notin L_{mes}
\end{array} \right\}^{N_c}_{l=1} \]

\[ \nabla J(r) = \Psi[r, r, h], \]

E.g. Polak-Ribiere conjugate gradient algorithm implemented in GSL

\[ r^{(k+1)} := r^{(k)} - \alpha^{(k)} s^{(k)}, \quad \alpha^{(k)} = \arg \min_{\alpha > 0} J\left( r^{(k)} - \alpha s^{(k)} \right), \]

\[ s^{(k)} = \begin{cases} \beta^{(k)} s^{(k-1)}, & k > 1 \\ g^{(k)}, & k = 1 \end{cases}, \quad \beta^{(k)} = \frac{\langle g^{(k)}, g^{(k)} - g^{(k-1)} \rangle}{\langle g^{(k-1)}, g^{(k-1)} \rangle}, \quad g^{(k)} = -\nabla_r J(r^{(k)}). \]
Adjoint problem solution ensembles in inverse problem algorithms

Cost function based

• **Cost functional gradients with adjoint problem solution** (single element ensemble for the discrepancy)

• **Gradient computation with adjoint ensemble when adjoint is independent of direct solution** [Karchevsky, A., Eurasian journal of mathematical and computer applications, 2013, 1, 4-20]

• **Representer method** (optimality system decomposition, ensemble generated for discrepancies for each measurement data) [Bennett, A. F. Inverse Methods in Physical Oceanography (Cambridge Monographs on Mechanics) Cambridge University Press, 1992]

Sensitivity relation based

• **Coarse-fine mesh method** (Sequential solution refinement with sequential adjoint problems solving) [Hasanov, A.; DuChateau, P. & Pektas, B. An adjoint problem approach and coarse-fine mesh method for identification of the diffusion coefficient in a linear parabolic equation// Journal of Inverse and Ill-Posed Problems, 2006, 14, 1-29]

Sensitivity operator
(inverse source problem)

Image (model) to structure operator [Dimet et al, 2015]

Given $\Xi$ functions $U = \{u^{(\xi)}\}_{\xi \in \Xi} \subset \text{Span}U_{\text{meas}}$

$$H_U \left( \phi[\mathbf{r}^{(2)}] - \phi[\mathbf{r}^{(1)}] \right) = \sum_{\xi \in \Xi} \left\langle \phi[\mathbf{r}^{(2)}] - \phi[\mathbf{r}^{(1)}], u^{(\xi)} \right\rangle e^{(\xi)}$$

Sensitivity relation (Lagrange type identity)

$$\left\langle \phi[\mathbf{r}^{(2)}] - \phi[\mathbf{r}^{(1)}], u^{(\xi)} \right\rangle = \left\langle \Psi[\mathbf{r}^{(2)}, \mathbf{r}^{(1)}; u^{(\xi)}], \mathbf{r}^{(2)} - \mathbf{r}^{(1)} \right\rangle$$

Sensitivity operator $M_U[\mathbf{r}^{(2)}, \mathbf{r}^{(1)}]$: $R \rightarrow \mathbb{R}^\Xi$

$$z \mapsto \sum_{\xi \in \Xi} \left\langle \Psi[\mathbf{r}^{(2)}, \mathbf{r}^{(1)}; u^{(\xi)}], z \right\rangle e^{(\xi)}, \quad \text{Parallel w.r.t. } U$$

The inverse problem solution $\mathbf{r}^{(*)}$ for any $\mathbf{r}$ and $U$ satisfy

$$H_U \left( I - \Pr_{U_{\text{mes}}} \phi[\mathbf{r}] \right) = M_U[\mathbf{r}^{(*)}, \mathbf{r}] (\mathbf{r}^{(*)} - \mathbf{r}) + H_U \delta \mathbf{I}. \quad \text{Parametric family of quasi-linear operator equations}$$
Adjoint ensemble construction
(inverse source problem, 2D, $L_{mes}$ components are measured)

Defined by the adjoint problem sources sets

«a priori» approach – ensemble for the class of problems (smoothness)

- Fourier cos-basis $U_\Theta = \left\{ e_{\eta \theta_x \theta_y \theta_t} | \theta_x \in [0, \Theta_x], \theta_y \in [0, \Theta_y], \theta_t \in [0, \Theta_t], \eta \in L_{mes} \right\}$,

$$e_{\eta \theta_x \theta_y \theta_t} = \left\{ \frac{1}{\sqrt{\rho_\eta}} C(T, \theta_t, t^k) C(X, \theta_x, x_i) C(Y, \theta_y, y_j), l = \eta \right\}_{l=1}^{N_c}$$,

$$0, \ l \neq \eta$$

- Wavelets, curvlets, etc. [Dimet et al., 2015]

«a posteriori» approach – ensemble for the considered problem

- «Adaptive basis»: chose elements of $U_\Theta$ with maximal projections

$$\left\langle \Pr \varphi[r^{(0)}] - I, e_{\eta \theta_x \theta_y \theta_t} \right\rangle_{U_{mes}}$$


- «Informative basis»: Use left singular vectors of the operator $m_{U_\Theta} [r^{(0)}, r^{(0)}]$

**Inversion algorithm**

\[
H_U \left( I - \Pr_{U_{mes}} \varphi[q] \right) = m_U[q, q] \left( q^{(*)} - q \right) + \left( m_U[q^{(*)}, q] - m_U[q, q] \right) \left( q^{(*)} - q \right) + H_U \delta I,
\]

\[
m \leftarrow m_U[q, q] \quad N_{\text{unknowns}} = \begin{cases} 
|L_{src}| \cdot N_t \cdot N_x \cdot N_y, & \text{inverse source problem} \\
N_{\text{coeff}}, & \text{inverse coefficient problem}
\end{cases}
\]

**Newton-Kantorovich-type update**

\[
\delta q = \begin{cases} 
\left( m^T \left[ mm^T \right]_{\Sigma}^+ \right), & \Xi < N_{\text{unknowns}} \\
\left[ m^T m \right]_{\Sigma}^+ m^T, & \Xi > N_{\text{unknowns}}
\end{cases}
\]

\[
[C]_{\Sigma}^+ - \text{truncated SVD inversion parametrized by conditional number } \Sigma
\]

**Nonlinearity:** sequential increase of the conditional number

**Noise:** discrepancy principle

**Admissible solutions:** projection regularization

**Optional monotonicity:** monotonic decrease of the discrepancy

**Theoretical foundations:** [Issartel, J.-P., 2003], [Cheverda V.A., Kostin V.I., 1995], [Kaltenbacher et al, 2008], [Vainikko, Veretennikov, 1986]
Inverse source problem (0D)

\[ L_{mes} = \{CO_2, O_3\} \]

\[ r^{(0)} = 0 \]

\[ T = 10 \times 3600 \]

\[ N_t = 3000 \]

\[ N_c = 22 \]

Larger ensembles and better solutions (0D)

- Lorenz’63 model,
- Inverse coefficient problem (2 unknown + 1 fixed coefficient)
- Regular in time state function measurements ($N_{meas} = T \times 3 \times 10$)
- Monotonic discrepancy decrease

Reconstruction error dynamics WRT computation time

Data assimilation mode (2D)
(data assimilation problem = sequence of inverse problems)

«Inverse problem mode»
(1 assimilation window)

Source: \( NO \)
Measurements: \( O_3 \) concentration images (movies)
Initial guess: zero

\[
T = 0.5 \times 3600 \\
N_t = 100 \\
X = Y = 600 \\
N_x = 100 \\
\theta_x = \theta_y = 5 \\
\theta_t = 10
\]

Assimilation window boundary

«Data assimilation mode»
(2 assimilation windows)

Sources identification with direct and indirect *in situ* (5 sites) measurements

\[ T = 4 \times 24 \times 3600 \quad \Xi = 5 \times 10 \]

NO concentrations are measured

![NO concentration map](image)

O3 concentrations are measured

![O3 concentration map](image)

Exact stationary NO source function (city traffic)

![Projected sources](image)

Projection to orthogonal complement of sensitivity operator kernel

![Reconstructed sources](image)

Summary

• Given the adjoint model, the sensitivity operator allow reformulating the inverse problem stated as a PDE system to a parametric family of quasilinear operator equations.

• Nonlinear ill-posed operator equation methods can be applied to the analysis and solution of the considered inverse problems.

• To solve the operator equations, the Newton-Kantorovich-type inversion algorithm has been proposed using:
  • The sequential increase of the considered spectrum in TSVD
  • Discrepancy principle and the iterative regularization.

• Both ensemble size and its construction affects the efficiency of the inverse problem solution (accuracy, time, local convergence).

• The algorithm was tested numerically in inverse modeling (inverse and data assimilation, source and coefficient) problems for advection-diffusion-reaction-model.

Thank you for your attention!

The work has been supported by RSF project 17-71-10184.


Adjoint ensemble methods

- Adjoint problem
- Sensitivity function
- Sensitivity operator analysis
- Inverse Problem Statement (PDE)
- Lagrange-type Identity
- Family of quasi-linear ill-posed operator equations
- Inverse problem solution algorithm
- Data assimilation mode
- Total ensemble
- Finite ensemble
- # computation threads

ILL-POSED NONLINEAR OPERATOR EQUATION METHODS