Four-dimensional Data Assimilation Problems in the infinite-dimensional setting

Paula M. Castro C.
Joint work with Juan Carlos De los Reyes

Research Center on Mathematical Modelling (ModeMat)
Escuela Politécnica Nacional

ALOP Autumn School on Bilevel Optimization
October 2020

Motivation

Data Assimilation in Numerical Weather Prediction

- Data assimilation methods aim at finding a good initial condition of the atmospheric system in order to get better weather forecasts;
- Weather observations come from many sources: ground stations, radiosondes or satellite images;

Source: World Meteorological Organization

- Reconstruction results depend strongly on the number of observations, which can be very limited in some cases.

4D-Var

Le Dimet and Talagrand (1986)

Finite dimensional problem

$$\begin{aligned}\min_u J(y,u) &= \frac{1}{2}\sum_{i=0}^{l}\left[H(y(t_i)) - z_o(t_i)\right]^T R_i^{-1}\left[H(y(t_i)) - z_o(t_i)\right] \\ &\quad + \frac{1}{2}\left[u - u_b\right]^T B^{-1}\left[u - u_b\right]\end{aligned}$$

subject to:

$$\begin{aligned} y(t_l) &= M(y(t_0)) && \text{(Dynamical system)} \\ y(t_0) &= u && \text{(Initial condition).}\end{aligned}$$

Features

- The dynamic problem incorporates given observations in a time window;
- The nonlinear dynamics may be taken into account;
- The operational use is still a computational challenge.

Infinite dimensional Semilinear DA problem

$$\min_u J(y,u) = \frac{1}{2}\sum_{k,i}[y(x_k,t_i) - z_0(x_k,t_i)]^2 + \frac{1}{2}\|u-u_b\|^2_{B^{-1}}$$

subject to:

$$\begin{aligned} \frac{\partial y}{\partial t} + Ay + g(y) &= 0 \quad \text{in } Q = \Omega\times]0,T[\\ y &= 0 \quad \text{on } \Sigma = \Gamma\times]0,T[\\ y(x,0) &= u \quad \text{in } \Omega \end{aligned}$$

Difficulties

- Higher regularity of the initial condition is required to get some sort of continuity of the state variable.
- Pointwise evaluations in the cost lead to right hand sides in $\mathcal{M}(Q)$ for the adjoint equation. Ill-posedness!

Infinite dimensional Semilinear DA problem

$$\min_u J(y,u) = \frac{1}{2}\int_0^T \sum_{k,i} w_k \sigma_i \rho_i(t)[y(x_k,t) - z_o(x_k,t)]^2\, dt + \frac{1}{2}\|u-u_b\|^2_{B^{-1}} + \frac{\vartheta}{2}\|\nabla(u-u_b)\|^2_{L^2(\Omega)}$$

subject to:

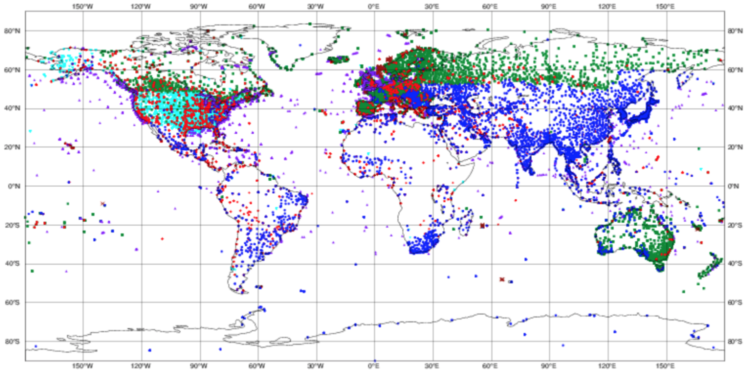
$$\begin{aligned} \frac{\partial y}{\partial t} + Ay + g(y) &= 0 && \text{in } Q = \Omega\times]0,T[\\ y &= 0 && \text{on } \Sigma = \Gamma\times]0,T[\\ y(x,0) &= u && \text{in } \Omega \end{aligned}$$

- w and σ are binary vectors, and $\rho_i(t)$ support functions
- ϑ is a regularization parameter, which in practice can be as small as required.

Analysis of the Variational Data Assimilation

- Well-posedness: With assumptions like the twice continuously differentiability of the nonlinear term, we get that the semilinear equation has a unique solution $y \in H^{2,1}(Q) \hookrightarrow L^2(0,T;C(\Omega))$
- Differentiability: We also prove that the contro-to-state mapping is Gâteaux differentiable
- Optimality system: We derive the optimality system and get the regularity of the adjoint state $p \in L^2(0,T;W_0^{1,r}(\Omega))$, with $r \in [1, \frac{m}{m-1}[$
  Casas-Clason-Kunisch (2013), Meyer-Susu (2017)

Observation stations



Problem

One has to make the best possible decision about where to locate the next observation stations to obtain better reconstructions of the initial condition.

Bilevel learning problem

Mixed integer-infinite dimensional optimization

ModeMat

$$\min_{w,\sigma \in \{0,1\}} \iint_Q \sum_{j=1}^{N} L(y_j, y_j^\dagger)\, dxdt + \beta \int_\Omega \sum_{j=1}^{N} l(u_j, u_j^\dagger)\, dx + \beta_w \sum_k w_k + \beta_\sigma \sum_i \sigma_i$$

subject to $(\forall j = 1, \dots, N)$:

$$\begin{cases} \min\limits_{u_j} \frac{1}{2} \int_0^T \sum\limits_{k,i} \left(w_k \sigma_i \rho_i(t) [y_j(x_k,t) - z_{oj}(x_k,t)]^2 \right) dt \\ \qquad + \frac{1}{2} \|u_j - u_{bj}\|_{B^{-1}} + \frac{\vartheta}{2} \|\nabla (u_j - u_{bj})\|^2_{L^2(\Omega)} \\ \text{subject to:} \\ \frac{\partial y_j}{\partial t} + A y_j + g(y_j) = 0 \quad \text{in } Q \\ y_j = 0 \quad \text{on } \Sigma \\ y_j(0) = u_j \quad \text{in } \Omega. \end{cases}$$

Bilevel learning problem

Mixed integer-infinite dimensional optimization

ModeMat

$$\min_{w,\sigma\in\{0,1\}} \iint_Q \sum_{j=1}^{N} L(y_j, y_j^\dagger)\, dxdt + \beta \int_\Omega \sum_{j=1}^{N} l(u_j, u_j^\dagger)\, dx + \beta_w \sum_k w_k + \beta_\sigma \sum_i \sigma_i$$

- $\beta_w > 0$ and $\beta_\sigma > 0$ are sparsity penalty terms for w and σ, respectively;
- For every $j = 1, \ldots, N$, $\left(u_j^\dagger, y_j^\dagger\right)$ represents an element of the training set, which is built from improved reconstructions of the initial condition and the observed state;
- L and l are loss functions that point us how close the state and initial condition are to the traing set (quality measure)
- We aim at learning the vector of placements w and time intervals σ at which the measurements should be carried out, such that the quality measure of the training set is minimized in average

Control of a singular system with measures

Replacing the lower level problems by their necessary optimality condition:

$$\min_{0\le w,\sigma\le 1} J(y,p,u,w) = \iint_Q \sum_{j=1}^{N} L(y_j, y_j^\dagger)\, dxdt + \beta \int_\Omega \sum_{j=1}^{N} l(u_j, u_j^\dagger)\, dx + \beta_w \sum_k w_k + \beta_\sigma \sum_i \sigma_i \quad (1)$$

subject to:

$$\begin{aligned}
\frac{\partial y_j}{\partial t} + Ay_j + g(y_j) &= 0 \\
y_j|_\Gamma &= 0 \\
y_j(0) &= u_j \\
-\frac{\partial p_j}{\partial t} + A^* p_j + g'(y_j) p_j &= \sum_{k,i} w_k \sigma_i \rho_i(t) \left[y_j(x,t) - z_{oj}(x,t) \right] \otimes \delta(x - x_k) \\
p_j|_\Gamma &= 0 \\
p_j(T) &= 0 \\
-\vartheta \Delta(u_j - u_b) + B^{-1}(u_j - u_b) &= -p_j(0) \\
u_j|_\Gamma &= 0.
\end{aligned}$$

Difficulty: No unique solution of the optimality system!

Analysis of Bilevel problem

- Existence theorem: We prove that the minimization problem (1) has at least one solution
- Adjoint system: We prove existence of Lagrange multipliers for (1), by analyzing an adapted penalized version of the problem
- Consistency of the penalizad problem: Convergence of the multipliers

Experiment 1

- Placement is allowed in any grid point
- Observations are taken in every time step
- Goal: observe how the solution structure changes with respect to β_w

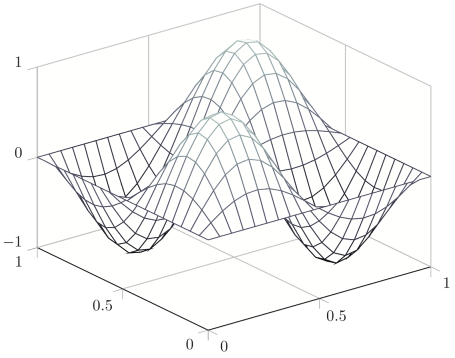


Figure 1. Initial condition $u^{\dagger}(x, y)$.

Experiment 1

- Placement is allowed in any grid point
- Observations are taken in every time step
- Goal: observe how the solution structure changes with respect to β_w