Blending Mathematical Models and Data: Algorithms, Analysis and Applications

Andrew Stuart

D Blömker (Augsburg), M. Dashti (Sussex), D Kelly (NYU),
KJH Law (Oak Ridge), D Sanz-Alonso (Warwick),
J Voss (Leeds), KC Zygalakis (Southampton)

EMS 25^{th} Anniversary, Institut Henri Poincaré, Paris,
October 22^{nd} 2015

Funded by EPSRC, ERC and ONR

Enabling Quantification of EQUIP Uncertainty for Inverse Problems

THE UNIVERSITY OF WARWICK

Outline

Table of Contents

Current State of Play

Data Everywhere

In every realm of human experience, growing in volume.

Current State of Play

Data Everywhere

In every realm of human experience, growing in volume.

Statisticians Deal With Data

Systematic development of mathematical statistics over more than a century. But range of models arguably somewhat limited.

Current State of Play

Data Everywhere

In every realm of human experience, growing in volume.

Statisticians Deal With Data

Systematic development of mathematical statistics over more than a century. But range of models arguably somewhat limited.

What Is New For Mathematics?

Interfacing existing models with data, and developing new models from data, provide intellectual challenges that will define much of the development of mathematics in the next century.

Current State of Play

Data Everywhere

In every realm of human experience, growing in volume.

Statisticians Deal With Data

Systematic development of mathematical statistics over more than a century. But range of models arguably somewhat limited.

What Is New For Mathematics?

Interfacing existing models with data, and developing new models from data, provide intellectual challenges that will define much of the development of mathematics in the next century.

Where Are We Now?

Mathematics in relation to data is in the same state that it was in relation to analysis at the time of Fourier.

Current State of Play

Data Everywhere

In every realm of human experience, growing in volume.

Statisticians Deal With Data

Systematic development of mathematical statistics over more than a century. But range of models arguably somewhat limited.

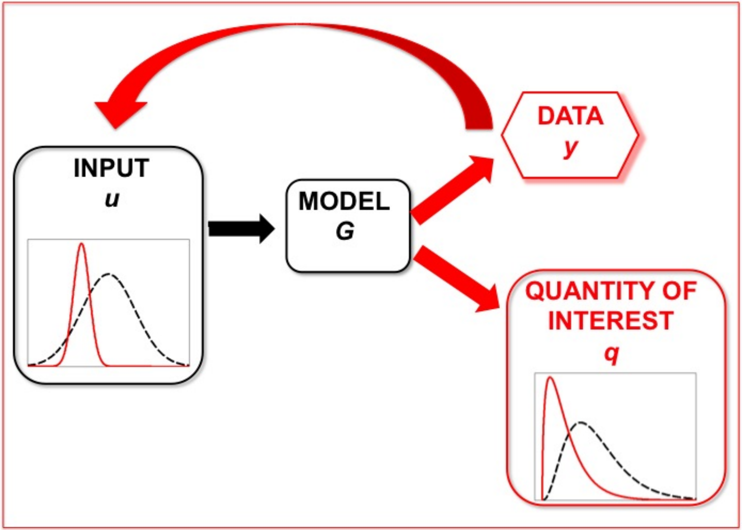
What Is New For Mathematics?

Interfacing existing models with data, and developing new models from data, provide intellectual challenges that will define much of the development of mathematics in the next century.

Where Are We Now?

Mathematics in relation to data is in the same state that it was in relation to analysis at the time of Fourier. Interesting times!

Model and Data



Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics). Strong belief in model.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics). Strong belief in model.
- Biology (Epidemics, Neuroscience). Models are (good) caricatures.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics). Strong belief in model.
- Biology (Epidemics, Neuroscience). Models are (good) caricatures.
- Social Sciences (Crowds, Economics, Voting). Models are (often weak) caracitures.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics). Strong belief in model.
- Biology (Epidemics, Neuroscience). Models are (good) caricatures.
- Social Sciences (Crowds, Economics, Voting). Models are (often weak) caracitures.
- Commerce (Amazon, Google, Netflix). Models barely exist.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics). Strong belief in model.
- Biology (Epidemics, Neuroscience). Models are (good) caricatures.
- Social Sciences (Crowds, Economics, Voting). Models are (often weak) caracitures.
- Commerce (Amazon, Google, Netflix). Models barely exist.

All Have Vast Amounts of Data!

Issues

- How to weight quality of model versus quality of data?

Issues

- How to weight quality of model versus quality of data?
- Pencil/paper model or machine-learnt model?

Issues

- How to weight quality of model versus quality of data?
- Pencil/paper model or machine-learnt model?
- Structure of data: time-ordered/all at once; how/low frequency; spatial structure; model-data link function?

Issues

- How to weight quality of model versus quality of data?
- Pencil/paper model or machine-learnt model?
- Structure of data: time-ordered/all at once; how/low frequency; spatial structure; model-data link function?
- How to deal with uncertainty in model and in data?

Issues

- How to weight quality of model versus quality of data?
- Pencil/paper model or machine-learnt model?
- Structure of data: time-ordered/all at once; how/low frequency; spatial structure; model-data link function?
- How to deal with uncertainty in model and in data?
- What extra do we learn from the data?

This Talk

- Based on physical models of high quality. NWP – Numerical Weather Prediction.

This Talk

- Based on physical models of high quality. NWP – Numerical Weather Prediction.
- Pencil-paper model: Navier-Stokes equation.

This Talk

- Based on physical models of high quality. NWP – Numerical Weather Prediction.
- Pencil-paper model: Navier-Stokes equation.
- Data: both time-ordered and all-at-once.

This Talk

- Based on physical models of high quality. NWP – Numerical Weather Prediction.
- Pencil-paper model: Navier-Stokes equation.
- Data: both time-ordered and all-at-once.
- Deterministic (used now) versus probabilistic (future).

This Talk

- Based on physical models of high quality. NWP – Numerical Weather Prediction.
- Pencil-paper model: Navier-Stokes equation.
- Data: both time-ordered and all-at-once.
- Deterministic (used now) versus probabilistic (future).

Challenges in Optimization, Dynamical Systems, Control and High Dimensional Probability

Table of Contents

Numerical Weather Prediction (NWP)

Signal dynamics on separable Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$. For $\Psi : \mathcal{H} \to \mathcal{H}$ and v_j velocity, temperature etc. at time j:

Signal Process

$$v_{j+1} = \Psi(v_j).$$

Numerical Weather Prediction (NWP)

Signal dynamics on separable Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$. For $\Psi : \mathcal{H} \to \mathcal{H}$ and v_j velocity, temperature etc. at time j:

Signal Process

$$v_{j+1} = \Psi(v_j).$$

For $P : \mathcal{H} \to \mathbb{R}^K$, observations y_j describe noisy satellite data, weather ballon data, aircraft data etc.:

Observation Process

$$y_{j+1} = Pv_{j+1} + \epsilon\eta_{j+1}.$$

Numerical Weather Prediction (NWP)

Signal dynamics on separable Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle, | \cdot |)$. For $\Psi : \mathcal{H} \to \mathcal{H}$ and v_j velocity, temperature etc. at time j:

Signal Process

$$v_{j+1} = \Psi(v_j).$$

For $P : \mathcal{H} \to \mathbb{R}^K$, observations y_j describe noisy satellite data, weather ballon data, aircraft data etc.:

Observation Process

$$y_{j+1} = Pv_{j+1} + \epsilon\eta_{j+1}.$$

Data Assimilation

Blend model and observations to learn $v_0 = u$. *Limit* $\epsilon \to 0$?

Challenges

In numerical weather prediction:

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.
- Map Ψ is chaotic.

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.
- Map Ψ is chaotic.
- Data is noisy.

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.
- Map Ψ is chaotic.
- Data is noisy.
- Model in imperfect (not discussed here):

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.
- Map Ψ is chaotic.
- Data is noisy.
- Model in imperfect (not discussed here):
- the data $\{y_j\}$ is not produced by the model Ψ.

Challenges

In numerical weather prediction:

- $\dim(\mathcal{H}) = \mathcal{O}(10^9)$.
- $J = \dim(\mathrm{Range}(P)) = \mathcal{O}(10^6)$.
- Map Ψ is chaotic.
- Data is noisy.
- Model in imperfect (not discussed here):
- the data $\{y_j\}$ is not produced by the model Ψ.

High Dimensional Probability!

Navier-Stokes Equation (all theorems apply)

Incompressible NSE on $\Omega_T = \mathbb{T}^2 \times (0, \infty)$

$$
\begin{aligned}
\partial_t v - \nu \triangle v + v \cdot \nabla v + \nabla p = f \quad & \text{in}\, \Omega_T, \\
\nabla \cdot v = 0 \quad & \text{in}\, \Omega_T, \\
v\big|_{t=0} = u \quad & \text{in}\, \mathbb{T}^2.
\end{aligned}
$$

Navier-Stokes Equation (all theorems apply)

Incompressible NSE on $\Omega_T = \mathbb{T}^2 \times (0, \infty)$

$$\begin{aligned} \partial_t v - \nu \triangle v + v \cdot \nabla v + \nabla p = f \quad & \text{in } \Omega_T, \\ \nabla \cdot v = 0 \quad & \text{in } \Omega_T, \\ v\big|_{t=0} = u \quad & \text{in } \mathbb{T}^2. \end{aligned}$$

Discrete-Time Semigroup

- $\mathcal{H} = L^2_{\mathrm{div}}(\mathbb{T}^2; \mathbb{R}^2)$.
- $\Psi : \mathcal{H} \to \mathcal{H}$ defined by:
- $v\big|_{t=h} := \Psi(u)$. Then:
- $v_{j+1} = \Psi(v_j)$.

Observations

- Eulerian:
- $Pv := \{v(x_k)\}_{k=1}^K$.
- (Divergence-Free) Fourier:
- $Pv := \{\int_{\mathbb{T}^2} v(x) e_k(x)\, dx\}_{k=1}^K$.

4DVAR (Talagrand and Courtier [8] Q. J. R. Met. Soc 1987)

(Covariance) Weighted Hilbert Spaces

$$A > 0, |\cdot|_A = |A^{-\frac{1}{2}} \cdot|.$$

4DVAR (Talagrand and Courtier [8] Q. J. R. Met. Soc 1987)

(Covariance) Weighted Hilbert Spaces

$$A > 0, |\cdot|_A = |A^{-\frac{1}{2}} \cdot|.$$

Objective Function

$$J(u) = |u - m_{\mathrm{pr}}|_C^2 + \epsilon^{-2} \sum_{j=1}^{J} |y_j - P\Psi^{(j)}(u)|_\Gamma^2.$$

4DVAR (Talagrand and Courtier [8] Q. J. R. Met. Soc 1987)

(Covariance) Weighted Hilbert Spaces

$$A > 0, |\cdot|_A = |A^{-\frac{1}{2}} \cdot|.$$

Objective Function

$$J(u) = |u - m_{\mathrm{pr}}|_C^2 + \epsilon^{-2} \sum_{j=1}^{J} |y_j - P\Psi^{(j)}(u)|_\Gamma^2.$$

Best Fit to Data and Model

$$m_{\mathrm{post}} = \mathrm{argmin}_{u \in \mathcal{H}} J(u).$$

3DVAR (Lorenc [5] Q. J. R. Met. Soc 1986)

Cycled 3DVAR Filter.

$$m_{j+1} = \mathrm{argmin}_{m \in \mathcal{H}} \{ |m - \Psi(m_j)|_C^2 + \epsilon^{-2} |y_{j+1} - Pm|_\Gamma^2 \}.$$

3DVAR (Lorenc [5] Q. J. R. Met. Soc 1986)

Cycled 3DVAR Filter.

$$m_{j+1} = \mathrm{argmin}_{m\in\mathcal{H}}\{|m-\Psi(m_j)|_C^2 + \epsilon^{-2}|y_{j+1} - Pm|_\Gamma^2\}.$$

Solve Variational Equations (with $C = \epsilon^2(\eta^{-2}\Gamma P + Q)$, Q=I-P)

$$m_{j+1} = (I-K)\Psi(m_j) + Ky_{j+1}, \quad K = (1+\eta^2)^{-1}P,$$

3DVAR (Lorenc [5] Q. J. R. Met. Soc 1986)

Cycled 3DVAR Filter.

$$m_{j+1} = \mathrm{argmin}_{m \in \mathcal{H}} \{ |m - \Psi(m_j)|_C^2 + \epsilon^{-2} |y_{j+1} - Pm|_\Gamma^2 \}.$$

Solve Variational Equations (with $C = \epsilon^2(\eta^{-2}\Gamma P + Q)$, Q=I-P)

$$m_{j+1} = (I - K)\Psi(m_j) + K y_{j+1}, \quad K = (1 + \eta^2)^{-1} P,$$

Variance Inflation (from NWP) $\eta \ll 1$

$$m_{j+1} = Q\Psi(m_j) + P y_{j+1}, \quad \eta = 0. \quad \textit{Synchronization Filter}.$$

The EnKF (Evensen [2] Journal of Geophysical Research 1994.)

Ensemble Kalman Filter. For $n = 1, \ldots, N$:

$$v_{j+1}^{(n)} = \mathrm{argmin}_{v \in \mathcal{H}} \{ |v - \Psi(v_j^{(n)})|_{C_j}^2 + \epsilon^{-2} |y_{j+1} - Pv|_{\Gamma}^2 \}.$$

The EnKF (Evensen [2] Journal of Geophysical Research 1994.)

Ensemble Kalman Filter. For $n = 1, \ldots, N$:

$$v_{j+1}^{(n)} = \mathrm{argmin}_{v \in \mathcal{H}} \{ |v - \Psi(v_j^{(n)})|_{C_j}^2 + \epsilon^{-2} |y_{j+1} - Pv|_\Gamma^2 \}.$$

Empirical Covariance

$$\bar{v}_j = \frac{1}{N} \sum_{n=1}^{N} v_j^{(n)}, \quad C_j = \frac{1}{N} \sum_{n=1}^{N} (v_j^{(n)} - \bar{v}_j) \otimes (v_j^{(n)} - \bar{v}_j).$$

The EnKF (Evensen [2] Journal of Geophysical Research 1994.)

Ensemble Kalman Filter. For $n = 1, \ldots, N$:

$$v_{j+1}^{(n)} = \mathrm{argmin}_{v \in \mathcal{H}} \{ |v - \Psi(v_j^{(n)})|_{C_j}^2 + \epsilon^{-2} |y_{j+1} - Pv|_\Gamma^2 \}.$$

Empirical Covariance

$$\bar{v}_j = \frac{1}{N} \sum_{n=1}^{N} v_j^{(n)}, \quad C_j = \frac{1}{N} \sum_{n=1}^{N} \big(v_j^{(n)} - \bar{v}_j\big) \otimes \big(v_j^{(n)} - \bar{v}_j\big).$$

Interacting Particles (Coupled Through K_j)

$$v_{j+1}^{(n)} = (I - K_j)\Psi(v_j^{(n)}) + K_j y_{j+1}$$

Table of Contents

Smoothing

Prior

$$\mathbb{P}(u) : u \sim \mu_0 = N(m_{\mathrm{pr}}, C).$$

Smoothing

Prior

$$\mathbb{P}(u) : u \sim \mu_0 = N(m_{\mathrm{pr}}, C).$$

Likelihood

$$\mathbb{P}(y_j|u) = N\big(P\Psi^{(j)}(u), \epsilon^{-2}\Gamma\big).$$

Smoothing

Prior

$$\mathbb{P}(u) : u \sim \mu_0 = N(m_{\mathrm{pr}}, C).$$

Likelihood

$$\mathbb{P}(y_j|u) = N\big(P\Psi^{(j)}(u), \epsilon^{-2}\Gamma\big).$$

Posterior

$$\mathbb{P}(u) : u \sim \mu^y.$$

Smoothing

Prior

$$\mathbb{P}(u) : u \sim \mu_0 = N(m_{\rm pr}, C).$$

Likelihood

$$\mathbb{P}(y_j|u) = N\big(P\Psi^{(j)}(u), \epsilon^{-2}\Gamma\big).$$

Posterior

$$\mathbb{P}(u) : u \sim \mu^y.$$

Bayes' Formula (to be justified) $\mathbb{P}(u|y) \propto \mathbb{P}(y|u)\mathbb{P}(u)$

$$\mu^y(du) \propto \exp\Big(-\epsilon^{-2}\sum_{j=1}^{J}|y_j - P\Psi^{(j)}(u)|_\Gamma^2\Big)\mu_0(du).$$

Smoothing is Well-Posed (S [7] Acta Numerica 2010.)

Theorem

Assume that prior covariance C is chosen so that $u \in \mathcal{H}$ μ_0 a.s. Then posterior $\mu^y \ll \mu_0$ and

$$\mu^y(du) \propto \exp\Big(-\epsilon^{-2}\sum_{j=1}^{J}|y_j - P\Psi^{(j)}(u)|_\Gamma^2\Big)\mu_0(du).$$

Furthermore $y \mapsto \mu^y$ is locally Lipschitz in Hellinger metric.

Smoothing is Well-Posed (S [7] Acta Numerica 2010.)

Theorem

Assume that prior covariance C is chosen so that $u \in \mathcal{H}$ μ_0 a.s. Then posterior $\mu^y \ll \mu_0$ and

$$\mu^y(du) \propto \exp\Big(-\epsilon^{-2}\sum_{j=1}^{J}|y_j - P\Psi^{(j)}(u)|_\Gamma^2\Big)\mu_0(du).$$

Furthermore $y \mapsto \mu^y$ is locally Lipschitz in Hellinger metric.

Hellinger Metric

The metric d_{Hell} is a natural one because, for $f \in L^2_{\mu_0}(\mathcal{H}; S)$,

$$\|\mathbb{E}^{\mu^y} f - \mathbb{E}^{\mu^{y'}} f\|_S \le K\, d_{Hell}(\mu^y, \mu^{y'}).$$

4DVAR is MAP Estimation (Dashti et al [1] Inverse Problems 2013.)

4DVAR

Recall the functional $J : E := \mathrm{Dom}(C^{-\frac{1}{2}}) \to \mathbb{R}^+$ (4DVAR):

$$J(u) := |u - m_{\mathrm{pr}}|_C^2 + \epsilon^{-2} \sum_{j=1}^{J} |y_j - P\Psi^{(j)}(u)|_\Gamma^2.$$

4DVAR is MAP Estimation (Dashti et al [1] Inverse Problems 2013.)

4DVAR

Recall the functional $J : E := \mathrm{Dom}(C^{-\frac{1}{2}}) \to \mathbb{R}^+$ (4DVAR):

$$J(u) := |u - m_{\mathrm{pr}}|_C^2 + \epsilon^{-2} \sum_{j=1}^{J} |y_j - P\Psi^{(j)}(u)|_\Gamma^2.$$

Theorem

Let $B^\delta(z) := \{v \in \mathcal{H} : |v - z| < \delta\}$. The probability measure μ^y and functional J are related by

$$\lim_{\delta \to 0} \frac{\mu^y\big(B^\delta(z)\big)}{\mu^y\big(B^\delta(0)\big)} = \exp\big(J(0) - J(z)\big).$$

4DVAR is MAP Estimation (Dashti et al [1] Inverse Problems 2013.)

4DVAR

Recall the functional $J : E := \mathrm{Dom}(C^{-\frac{1}{2}}) \to \mathbb{R}^+$ (4DVAR):

$$J(u) := |u - m_{\mathrm{pr}}|_C^2 + \epsilon^{-2} \sum_{j=1}^{J} |y_j - P\Psi^{(j)}(u)|_\Gamma^2.$$

Theorem

Let $B^\delta(z) := \{v \in \mathcal{H} : |v - z| < \delta\}$. The probability measure μ^y and functional J are related by

$$\lim_{\delta \to 0} \frac{\mu^y\big(B^\delta(z)\big)}{\mu^y\big(B^\delta(0)\big)} = \exp\big(J(0) - J(z)\big).$$

Thus maximum a posterior (MAP) estimators are minimizers of J.

$\epsilon \to 0$ Asymptotics (Dashti et al [1] Inverse Problems 2013.)

Bayesian Posterior Consistency

Fix $u^\dagger \in \mathcal{H}$, draw $\eta \sim N(0, \Gamma)$, and let J_ϵ be the 4DVAR objective functional corresponding to the data y_ϵ defined by:

$$y_\epsilon = \mathcal{G}(u^\dagger) + \epsilon\eta.$$

$\epsilon \to 0$ Asymptotics (Dashti et al [1] Inverse Problems 2013.)

Bayesian Posterior Consistency

Fix $u^\dagger \in \mathcal{H}$, draw $\eta \sim N(0, \Gamma)$, and let J_ϵ be the 4DVAR objective functional corresponding to the data y_ϵ defined by:

$$y_\epsilon = \mathcal{G}(u^\dagger) + \epsilon\eta.$$

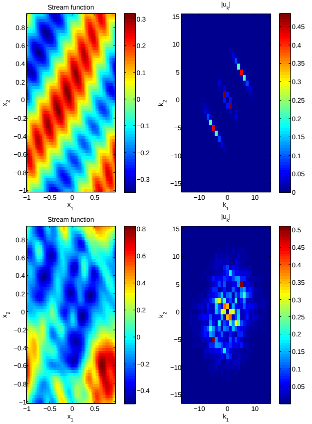
Theorem

There is a $u_\epsilon \in E$:

$$J(u_\epsilon) := \inf\{J_\epsilon(u) : u \in E\}.$$

Furthermore, $\eta-$a.s. there is subsequence $\{u_{\epsilon'}\}$ converging weakly to u_0 in E where $\mathcal{G}(u_0) = \mathcal{G}(u^\dagger)$.

Example: Navier-Stokes Inversion for Initial Condition



- Incompressible NSE on $\Omega_T = \mathbb{T}^2 \times (0,\infty)$:

$$\begin{aligned} \partial_t v - \nu \triangle v + v\cdot\nabla v + \nabla p = f &\quad \text{in } \Omega_T, \\ \nabla\cdot v = 0 &\quad \text{in } \Omega_T, \\ v\big|_{t=0} = u &\quad \text{in } \mathbb{T}^2. \end{aligned}$$

- $y_{k,\ell} = v(x_k, t_\ell) + \eta_{k,\ell}, \quad \eta_{k,\ell} \sim N(0, \sigma^2 I_{2\times 2})$.
- $y = \mathcal{G}(u) + \eta, \quad \eta \sim N(0, \sigma^2 I)$.
- $C_0 = (-\triangle_{\text{stokes}})^{-2}$; $\Phi = \frac{1}{10^3\sigma^2}|y - \mathcal{G}(u)|^2$.

Example: MAP estimator $u^\star$; Truth $u^\dagger$

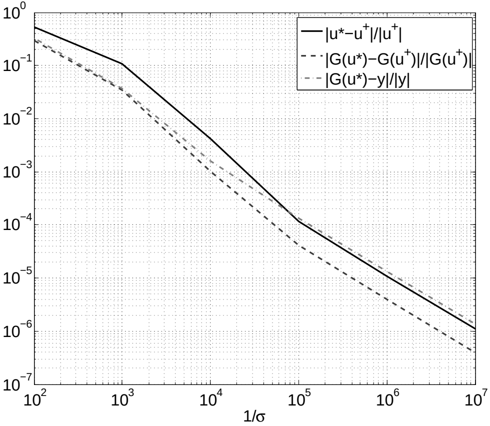


Table of Contents

Filtering

Filter

$$\mu_j(A) = \mathbb{P}\big(v_j \in A | \{y_\ell\}_{\ell=1}^{j}\big).$$

Filtering

Filter

$$\mu_j(A) = \mathbb{P}\big(v_j \in A | \{y_\ell\}_{\ell=1}^{j}\big).$$

Pushforward Under Dynamics

$$\widehat{\mu}_{j+1} = \Psi \star \mu_j.$$

Filtering

Filter

$$\mu_j(A) = \mathbb{P}\big(v_j \in A | \{y_\ell\}_{\ell=1}^{j}\big).$$

Pushforward Under Dynamics

$$\widehat{\mu}_{j+1} = \Psi\star\mu_j.$$

Observations via Bayes' Theorem

$$\mu_{j+1}(A) = \frac{\int_A \exp\big(-\epsilon^{-2}|y_{j+1} - Pv|_\Gamma^2\big)\widehat{\mu}_{j+1}(dv)}{\int_{\mathcal{H}} \exp\big(-\epsilon^{-2}|y_{j+1} - Pv|_\Gamma^2\big)\widehat{\mu}_{j+1}(dv)}.$$

Filtering is Well-Posed

From the pushforward property

$$\mu_j = \Psi^{(j)}{}_{\star}\mu^{(y_1,\cdots,y_j)}$$

and the well-posedness of the smoothing problem:

Theorem

Assume that prior covariance C is chosen so that $u \in \mathcal{H}$ μ_0 a.s. Then $(y_1, \cdots, y_j) \mapsto \mu_j$ is locally Lipschitz in Hellinger metric.

Detectability

Nonlinear Detectability (Recall $Q = I - P$)

The ball $B(R_0)$ in $\mathcal{H}$ is absorbing for the dynamics and there is a squeezing property: $\alpha(R_0) \in (0,1)$ such that, for all $u, v \in B(R_0)$,

$$\|Q(\Psi(u) - \Psi(v))\|^2 \leq \alpha(\mathbf{R_0}) \|u - v\|^2.$$

Detectability

Nonlinear Detectability (Recall $Q = I - P$)

The ball $B(R_0)$ in $\mathcal{H}$ is absorbing for the dynamics and there is a squeezing property: $\alpha(R_0) \in (0,1)$ such that, for all $u, v \in B(R_0)$,

$$\|Q(\Psi(u) - \Psi(v))\|^2 \leq \alpha(\mathbf{R_0})\|u - v\|^2.$$

Foais and Prodi '67, Mattingly et al '00s, Olsen, Titi et al '00s.

Consistency of the Filter (Sanz-Alonso and S, 2014, [6] SIAM JUQ.)

Theorem

Assume nonlinear detectability and Fourier observations. Then there is a constant $c > 0$ independent of the noise strength ϵ such that

$$\limsup_{j\to\infty} \mathbb{E}|v_j - \mathbb{E}^{\mu_j} v_j|^2 \leq c\epsilon^2$$

.

Consistency of the Filter (Sanz-Alonso and S, 2014, [6] SIAM JUQ.)

Theorem

Assume nonlinear detectability and Fourier observations. Then there is a constant $c > 0$ independent of the noise strength ϵ such that

$$\limsup_{j\to\infty} \mathbb{E}|v_j - \mathbb{E}^{\mu_j} v_j|^2 \leq c\epsilon^2$$

.

Idea of proof:

- Let m_j come from synchronization filter (3DVAR variant).
- Prove

$$\limsup_{j\to\infty} \mathbb{E}|v_j - m_j|^2 \leq c\epsilon^2.$$

- Use the L^2 optimality of the filtering distribution.

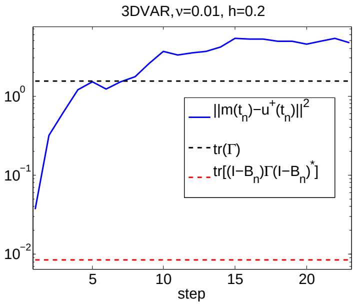
Idea of proof (sketch, Ψ globally Lipschitz):

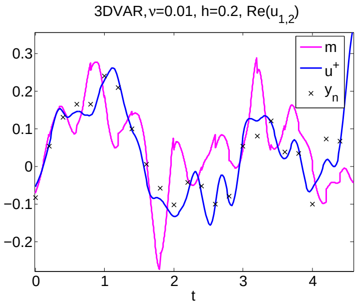
$$\begin{aligned} m_{j+1} &= Q\Psi(m_j) \qquad && + \overbrace{P\Psi(v_j) + \epsilon\eta_{j+1}}^{\mathbf{y_{j+1}}}, \\ v_{j+1} &= Q\Psi(v_j) \qquad && + P\Psi(v_j). \end{aligned}$$

Subtract and use independence plus contractivity of $Q\Psi$:

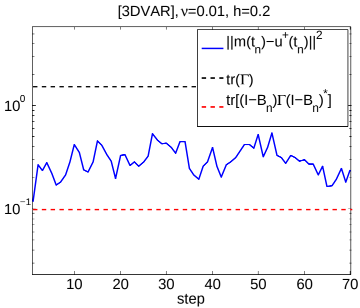
$$\begin{aligned} \mathbb{E}\|v_{j+1} - m_{j+1}\|^2 &= \mathbb{E}\|Q\left(\Psi(v_j) - \Psi(m_j)\right) - \epsilon\eta_{j+1}\|^2 \\ &\le \mathbb{E}\|Q\left(\Psi(v_j) - \Psi(m_j)\right)\|^2 + \epsilon^2\mathbb{E}\|\eta_{j+1}\|^2 \\ &\le \alpha\mathbb{E}\|v_j - m_j\|^2 + \epsilon^2\mathbb{E}\|\eta_{j+1}\|^2. \end{aligned}$$

Inaccurate: η too large. (NSE torus) Law and S [4], Monthly Weather Review, 2012





Accurate: smaller η. (NSE torus) Law and S [4], Monthly Weather Review, 2012



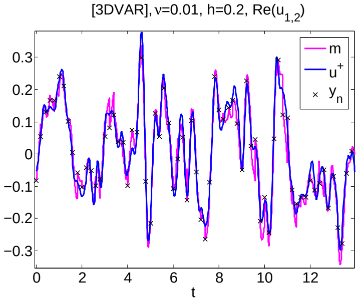


Table of Contents

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- The talk illustrates the first of these challenges in the context of numerical weather prediction.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- The talk illustrates the first of these challenges in the context of numerical weather prediction.
- Particular challenges are the dimension of the state space – $\mathcal{O}(10^9)$ – the dimension of the data – $\mathcal{O}(10^6)$ – and the chaotic nature of the models.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- The talk illustrates the first of these challenges in the context of numerical weather prediction.
- Particular challenges are the dimension of the state space – $\mathcal{O}(10^9)$ – the dimension of the data – $\mathcal{O}(10^6)$ – and the chaotic nature of the models.
- The ideal probabilistic formulation of the problem is currently beyond computational reach.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- The talk illustrates the first of these challenges in the context of numerical weather prediction.
- Particular challenges are the dimension of the state space – $\mathcal{O}(10^9)$ – the dimension of the data – $\mathcal{O}(10^6)$ – and the chaotic nature of the models.
- The ideal probabilistic formulation of the problem is currently beyond computational reach.
- However it provides a benchmark against which to test algorithms like 4DVAR, 3DVAR and EnKF. (Not discussed here).

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- The talk illustrates the first of these challenges in the context of numerical weather prediction.
- Particular challenges are the dimension of the state space – $\mathcal{O}(10^9)$ – the dimension of the data – $\mathcal{O}(10^6)$ – and the chaotic nature of the models.
- The ideal probabilistic formulation of the problem is currently beyond computational reach.
- However it provides a benchmark against which to test algorithms like 4DVAR, 3DVAR and EnKF. (Not discussed here).
- It also provides a theoretical framework in which to discuss ability of algorithms to reconstruct the "truth" in the data in small noise or large data volume limits. (Discussed in detail here for 3DVAR and 4DVAR; see Kelly 2014 [3] for EnKF).

References I

M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss.
MAP estimators and posterior consistency in bayesian nonparametric inverse problems.
Inverse Problems, 29:095017, 2013.

G. Evensen.
Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics.
Journal of Geophysical Research, pages 143–162, 1994.

D.T.B. Kelly, K.J.H. Law, and A.M. Stuart.
Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time.
Nonlinearity, 27:2579–2603, 2014.

K.J.H. Law and A.M. Stuart.
Evaluating data assimilation algorithms.
Monthly Weather Review, 140:3757–3782, 2012.

References II

A.C. Lorenc.
Analysis methods for numerical weather prediction.
Quarterly Journal of the Royal Meteorological Society, 112(474):1177–1194, 1986.

D. Sanz-Alonso and A.M. Stuart.
Long-time asymptotics of the filtering distribution for partially observed chaotic deterministic dynamical systems.
To appear: SIAM JUQ, arxiv.org/abs/1411.6510, 2014.

A. M. Stuart.
Inverse problems: a Bayesian perspective.
Acta Numer., 19:451–559, 2010.

References III

O. Talagrand and P. Courtier.
Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory.
Quarterly Journal of the Royal Meteorological Society, 113(478):1311–1328, 1987.

References IV

MATLAB files and book chapters freely available:

http://tiny.cc/damat

http://arxiv.org/abs/1506.07825