Blending Mathematical Models and Data: Algorithms, Analysis and Applications

Andrew Stuart

Data-Rich Phenomena and PDEs,
Cambridge, December 16^{th} 2015
Funded by DARPA, 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.

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!

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics).
  Strong belief in model.
  Pencil-paper models – centuries old.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics).
  Strong belief in model.
  Pencil-paper models – centuries old.
- Biology (Epidemics, Neuroscience, Healthcare).
  Models are (good) cartoons.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics).
  Strong belief in model.
  Pencil-paper models – centuries old.
- Biology (Epidemics, Neuroscience, Healthcare).
  Models are (good) cartoons.
- Social Sciences (Crowds, Economics, Voting).
  Models are (often weak) cartoons.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics).
  Strong belief in model.
  Pencil-paper models – centuries old.
- Biology (Epidemics, Neuroscience, Healthcare).
  Models are (good) cartoons.
- Social Sciences (Crowds, Economics, Voting).
  Models are (often weak) cartoons.
- Commerce (Amazon, Google, Netflix).
  Models in their infancy.
  Models are being machine learnt.
  Modeling principles are emerging.

Prototypical Problem Areas

- Physics (Newton's Laws, Quantum Mechanics).
  Strong belief in model.
  Pencil-paper models – centuries old.
- Biology (Epidemics, Neuroscience, Healthcare).
  Models are (good) cartoons.
- Social Sciences (Crowds, Economics, Voting).
  Models are (often weak) cartoons.
- Commerce (Amazon, Google, Netflix).
  Models in their infancy.
  Models are being machine learnt.
  Modeling principles are emerging.

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?
- Linking pencil/paper and machine-learning?

Issues

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

Issues

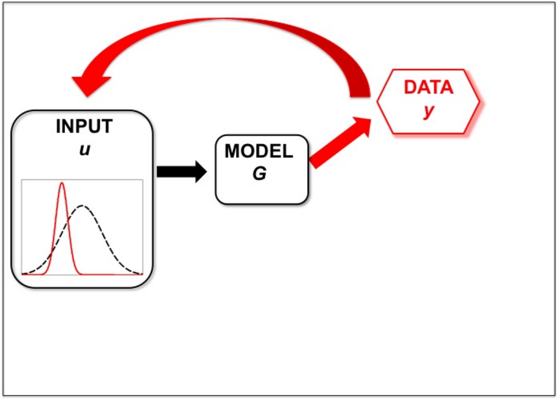
- How to weight quality of model versus quality of data?
- Pencil/paper model or machine-learnt model?
- Linking pencil/paper and machine-learning?
- Structure of data: time-ordered/all at once; how/low frequency; spatial structure; model-data link function?
- How to deal with uncertainty in model and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

Issues

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

Table of Contents

Bayes Theorem (Picture)



Model

$G : X \to \mathbb{R}^J$, X separable Banach space.

Bayes Theorem (Mathematics)

Prior

Probabilistic information about u before data is collected: $\mathbb{P}(du)$.

Likelihood

Since $y = \mathcal{G}(u) + \eta$, if $\eta \sim N(0, \Gamma)$, then $y|u \sim N(\mathcal{G}(u), \Gamma)$. The model-data misfit Φ is the negative log-likelihood:

$$\mathbb{P}(y|u) \propto \exp\big(-\Phi(u;y)\big), \quad \phi(u;y) = \frac{1}{2}\Big|\Gamma^{-1/2}\big(y - \mathcal{G}(u)\big)\Big|^2.$$

Posterior

Probabilistic information about u after data is collected:

$$\mathbb{P}(du|y) \propto \exp\big(-\Phi(u;y)\big)\mathbb{P}(du).$$

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

- Approximation by a Dirac: MAP/Tikhonov [2].

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

- Approximation by a Dirac: MAP/Tikhonov [2].
- Approximation by a Gaussian: Variational/ML [4].

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

- Approximation by a Dirac: MAP/Tikhonov [2].
- Approximation by a Gaussian: Variational/ML [4].
- Approximation by sampling: MCMC [1].

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

- Approximation by a Dirac: MAP/Tikhonov [2].
- Approximation by a Gaussian: Variational/ML [4].
- Approximation by sampling: MCMC [1].
- Approximation by sequential sampling: SMC [3].

Algorithms For Bayesian Inversion

We wish to get information about the structure of the posterior probability on input, given data. Possibilities:

- Approximation by a Dirac: MAP/Tikhonov [2].
- Approximation by a Gaussian: Variational/ML [4].
- Approximation by sampling: MCMC [1].
- Approximation by sequential sampling: SMC [3].
- Approximation by high dimensional integration: Harmonic analysis, sparse integration [5].

Table of Contents

Example 1: Piecewise Constant Reconstruction

Forward Problem

Let $K \in \mathcal{L}(X, \mathbb{R}^J)$ for some Banach space X. Given $\sigma \in X$

$$y = K\sigma.$$

Let $\eta \in \mathbb{R}^J$ be a realization of an observational noise.

Inverse Problem

Impose prior information that σ is piecwise constant:

$$\sigma = F(u) := \sigma^+ \chi_{\{u \geq 0\}}(x) + \sigma^- \chi_{\{u<0\}}(x).$$

Given $y \in \mathbb{R}^J$, find u such that

$$y = KF(u) + \eta.$$

Example 2: Electrical Impedance Tomography

M. Dunlop and A.M. Stuart
Bayesian formulation of EIT.
arXiv:1509.03136
Inverse Problems and Imaging, Submitted, 2015.

- Apply currents I_ℓ on e_ℓ, $\ell = 1, \ldots, L$.
- Induces voltages Θ_ℓ on e_ℓ, $\ell = 1, \ldots, L$.
- We have an Ohm's law $\Theta = R(\sigma)I$.
- Find conductivity σ from measurements (I, Θ).

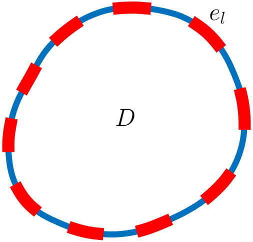


Table of Contents

Summary

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.
- These challenges require new ideas linking computer science, mathematics and statistics.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- These challenges require new ideas linking computer science, mathematics and statistics.
- The driving applications are central and will help define different problem classes.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- These challenges require new ideas linking computer science, mathematics and statistics.
- The driving applications are central and will help define different problem classes.
- Methods from data science may link with pencil/paper first-principles models.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- These challenges require new ideas linking computer science, mathematics and statistics.
- The driving applications are central and will help define different problem classes.
- Methods from data science may link with pencil/paper first-principles models.
- The UK has world-leading expertise in many of the constituent areas of the mathematical sciences.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- These challenges require new ideas linking computer science, mathematics and statistics.
- The driving applications are central and will help define different problem classes.
- Methods from data science may link with pencil/paper first-principles models.
- The UK has world-leading expertise in many of the constituent areas of the mathematical sciences.
- An essential barrier, which the ATI can help overcome, is to bring together this expertise to effectively identify significant intellectual challenges which underpin different problem classes.

Summary

- Blending models with data, and learning models from data, will be major challenges for the mathemtical sciences in the coming century.
- These challenges require new ideas linking computer science, mathematics and statistics.
- The driving applications are central and will help define different problem classes.
- Methods from data science may link with pencil/paper first-principles models.
- The UK has world-leading expertise in many of the constituent areas of the mathematical sciences.
- An essential barrier, which the ATI can help overcome, is to bring together this expertise to effectively identify significant intellectual challenges which underpin different problem classes.
- In doing so define data science as a discipline.

References I

S.L. Cotter, G.O. Roberts, A.M. Stuart, and D. White.
MCMC methods for functions: modifying old algorithms to make them faster.
Statistical Science, 28(3):424–446, 2013.

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.

M. Dashti and A.M. Stuart.
The Bayesian approach to inverse problems.
arXiv preprint arXiv:1302.6989, 2013.

F.J. Pinski, G. Simpson, A.M. Stuart, and H. Weber.
Algorithms for Kullback-Leibler approximation of probability measures in infinite dimensions.
arXiv preprint arXiv:1408.1920, 2014.

References II

C. Schwab and A.M. Stuart.
Sparse deterministic approximation of bayesian inverse problems.
Inverse Problems, 28(4):045003, 2012.

K.J.H. Law, A.M. Stuart and K.C.Zygalakis,
Data Assimilation: A Mathematical Introduction,
Springer, 2015.

MATLAB files and book chapters freely available:
`http://tiny.cc/damat`
`http://arxiv.org/abs/1506.07825`