Bayesian Inference and Experimental Design for Implicit Models

Michael Gutmann

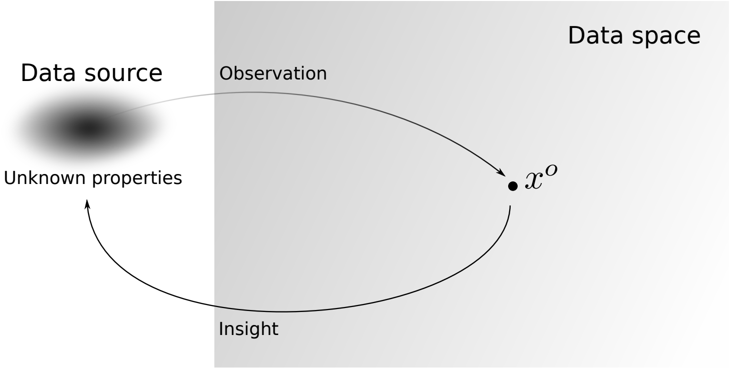
http://homepages.inf.ed.ac.uk/mgutmann

Institute for Adaptive and Neural Computation
School of Informatics, University of Edinburgh

7 December 2018

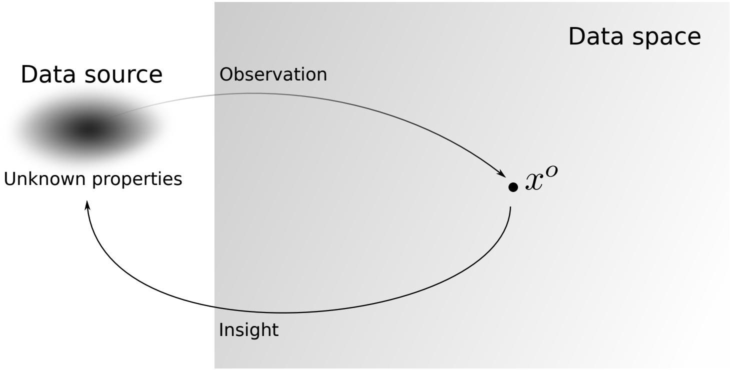
Learning from data

- Goal: Use data $\boldsymbol{x}^o$ to learn about properties of its source
- Enables predictions, decision making under uncertainty, ...



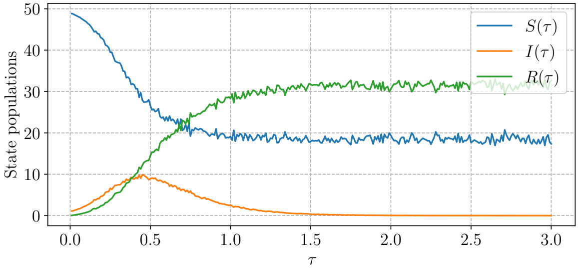
Two fundamental tasks

- Inference task : Given $\boldsymbol{x}^o$, what can we robustly say about the properties of the source?
- Experimental design task : How to obtain a $\boldsymbol{x}^o$ that is maximally useful for learning about the properties?



Example: stochastic SIR model

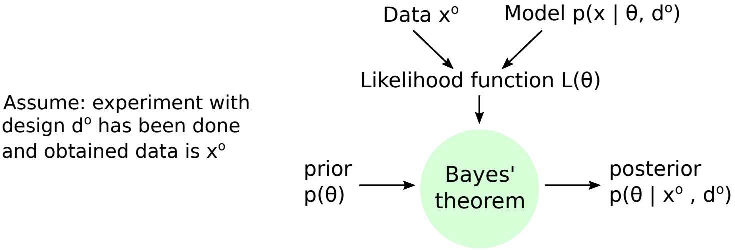
- Stochastic model describing the population of susceptibles $S(\tau)$, infected $I(\tau)$ and recovered $R(\tau)$ as a function of time.
- Parameters $\boldsymbol{\theta}$: rate of infection β and the rate of recovery γ.
- Inference task : determine plausible values of β and γ given some measurements of the population sizes.
- Exp design task : find the optimal times at which to perform the measurements to most accurately estimate β and γ.



(Figure by Steven Kleinegesse)

Bayesian approach

- Learning $\equiv$ Bayesian inference



- Exp design $\equiv$ utility optimisation problem
  e.g. maximise mutual information (MI) between $\boldsymbol{x}$ and $\boldsymbol{\theta}$

$$U(\boldsymbol{d}) = \mathbb{E}_{\boldsymbol{x}|\boldsymbol{d}}\left[\mathrm{KL}\left(p(\boldsymbol{\theta}|\boldsymbol{d},\boldsymbol{x}) \,||\, p(\boldsymbol{\theta})\right)\right]$$

Expected information gain for an experiment with design $\boldsymbol{d}$

Principled but very difficult

- Difficulty essentially due to high-dimensional integrals
- One reason for the integrals: unobserved variables $\boldsymbol{z}$ which makes the likelihood function intractable

$$
\begin{aligned}
L(\boldsymbol{\theta}) &\propto p(\boldsymbol{x}^o \,|\, \boldsymbol{\theta}, \mathbf{d}) \\
&\propto \int p(\boldsymbol{x}^o, \boldsymbol{z} \,|\, \boldsymbol{\theta}, \mathbf{d}) \mathrm{d}\boldsymbol{z}
\end{aligned}
$$

- Makes both Bayesian inference and experimental design very difficult

Research interests

1. Methods development:
   - efficient inference when the likelihood is intractable (e.g. because of unobserved variables)
   - efficient experimental design
2. Applications in biomedicine (in collaboration with domain experts)

Main tools: likelihood-free inference and modern machine learning techniques (Bayesian optimisation, neural networks, ratio estimation)

Example research: efficient inference

- We developed a fast inference method for models where the likelihood is intractable but sampling is possible (implicit models).
- Example: infer bacterial transmission dynamics in child care centres.
- Roughly equal results using 1000 times fewer simulations.

4.5 days with 200 cores

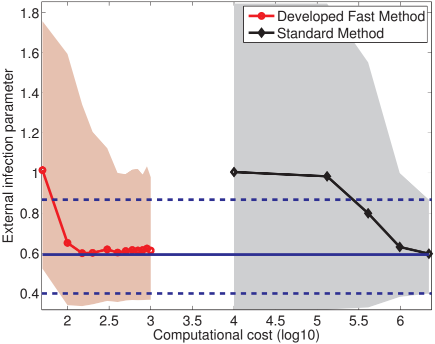
↓

90 minutes with seven cores

Posterior means: solid lines,

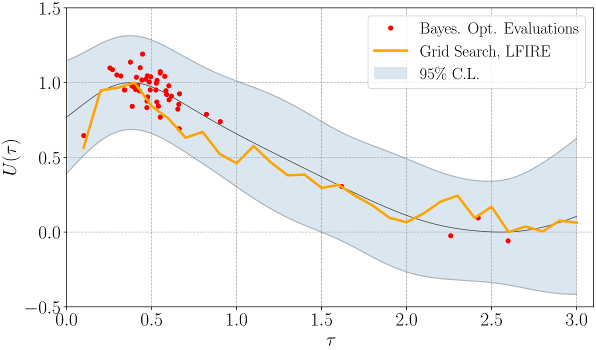
credibility intervals: shaded areas or dashed lines.

(Gutmann and Corander, JMLR, 2016)



Example research: exp design (Kleinegesse and Gutmann, arXiv:1810.09912)

- SIR model: find optimal measurement time by maximisation of mutual information
- Technical difficulties: (1) approximation of the posterior and the mutual information (2) maximisation
- For (1), we use likelihood-free inference by ratio estimation (LFIRE, arXiv:1611.10242). For (2), we use Bayesian optimisation



Research interests

1. Methods development:
   - efficient inference when the likelihood is intractable (e.g. because of unobserved variables)
   - efficient experimental design
2. Applications in biomedicine (in collaboration with domain experts)

Main tools: modern machine learning techniques (Bayesian optimisation, neural networks, ratio estimation)