Efficient Learning from Data

Michael Gutmann

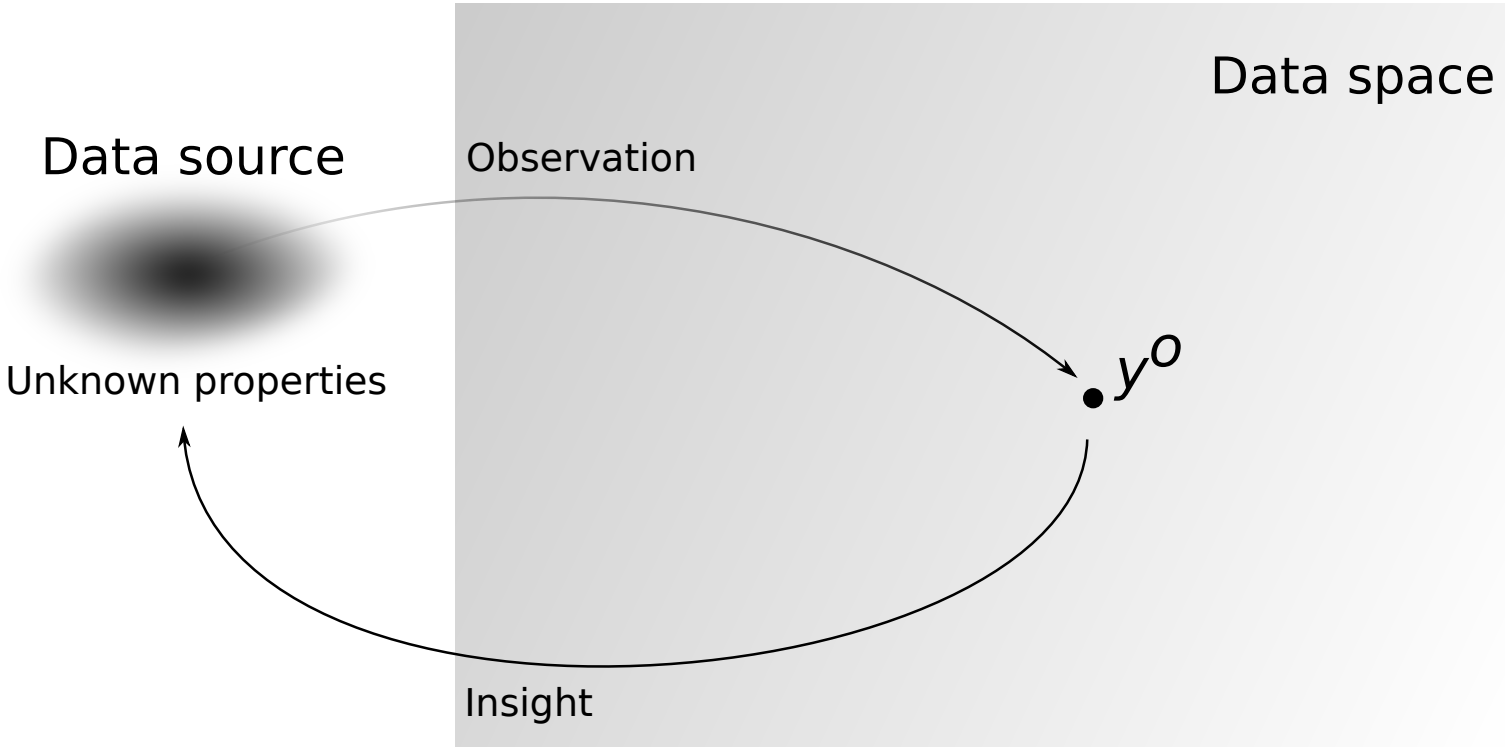
https://sites.google.com/site/michaelgutmann

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

October 2016

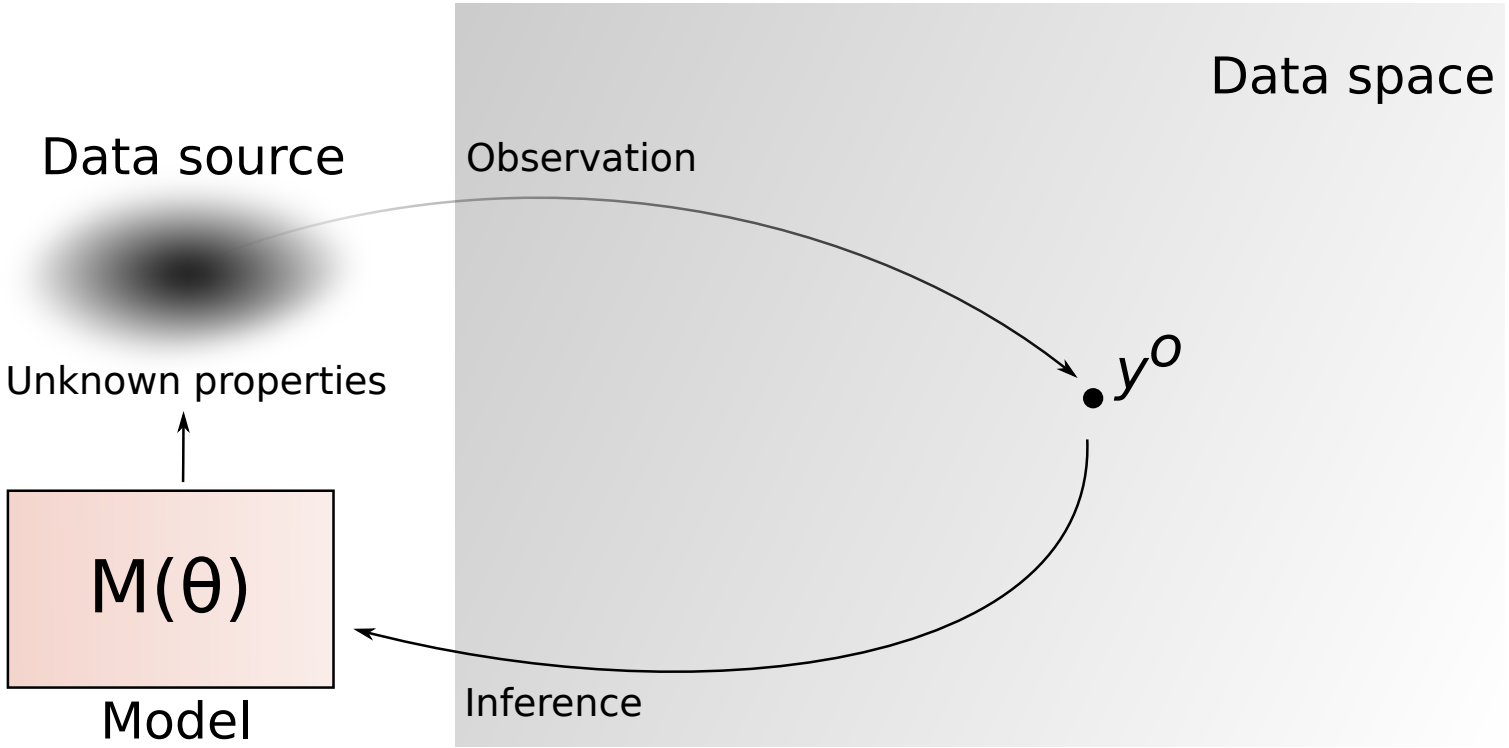
Learning from data

- Goal: Using observed data y^o, learn about their source
- Enables decision making, predictions, ...



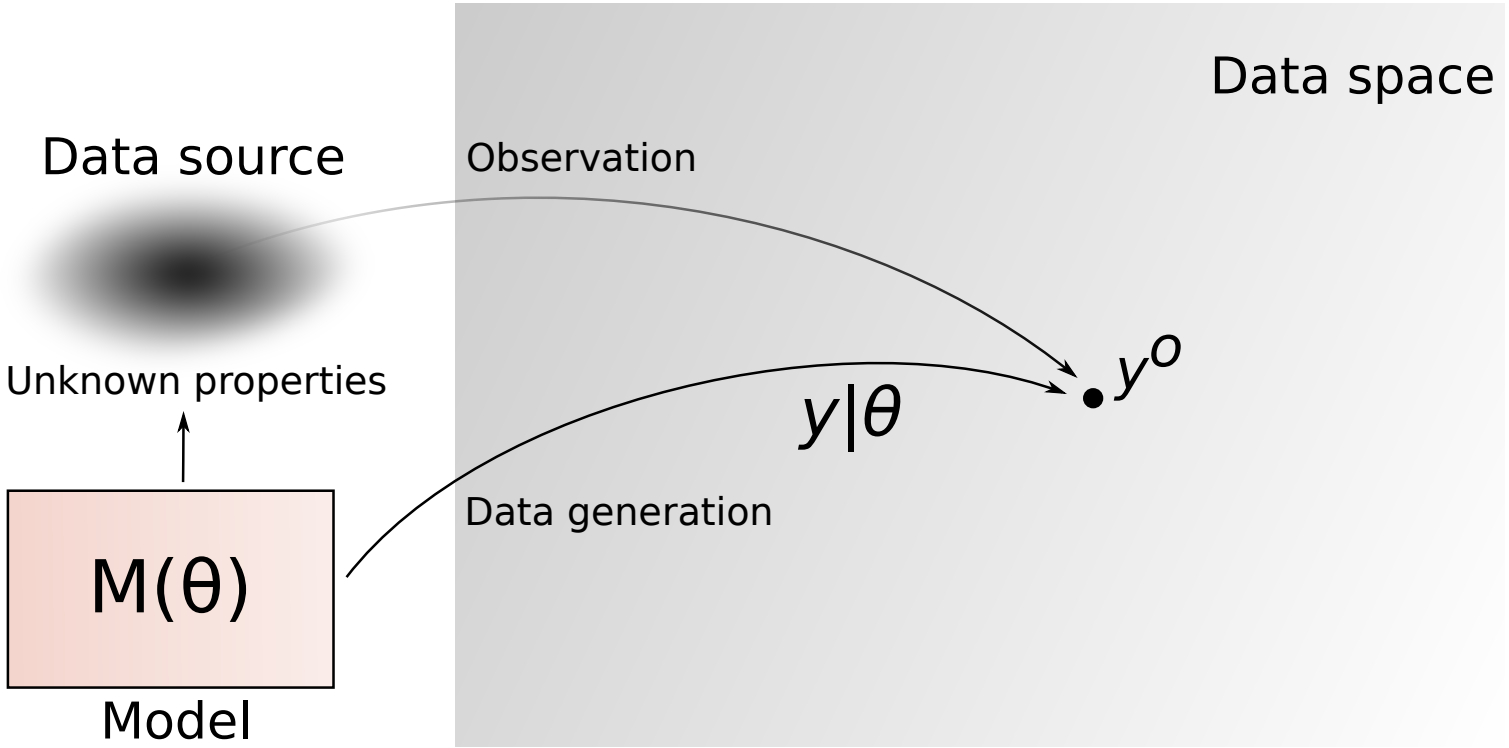
Statistical approach

- Set up a model with potential properties θ (parameters)
- See which θ are in line with the observed data y^o



The likelihood function

- Measures agreement between θ and the observed data y^o
- Probability to generate data like y^o if hypothesis θ holds



Computational cost

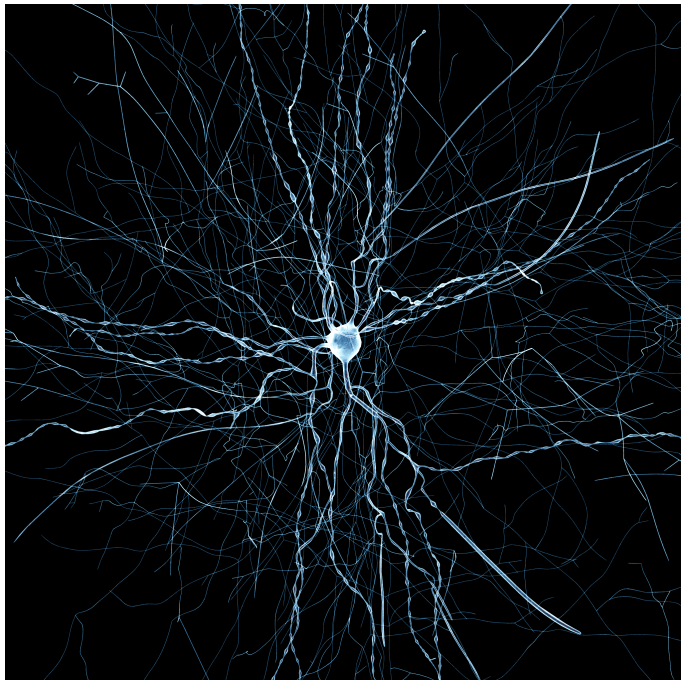
- Likelihood function often too expensive to compute: models / likelihood are "intractable"
- Exact inference impossible
- Efficient approximate solutions are needed: good trade-off between speed and accuracy

My research

- Develop methods for efficient approximate inference
  - Generative models
  - Unnormalized models
- Apply them to real problems
  - Unsupervised deep learning
  - Computational biology
    (visual neuroscience, infectious disease epidemiology)

Generative models

- Models which specify a mechanism for generating data
  - computer program that simulates some complex process
- Widely used
  - Evolutionary biology: Simulating evolution
  - Neuroscience: Simulating neural circuits
  - Health science: Simulating the spread of an infectious disease
  - Computer graphics: Simulating natural scenes
  - ...



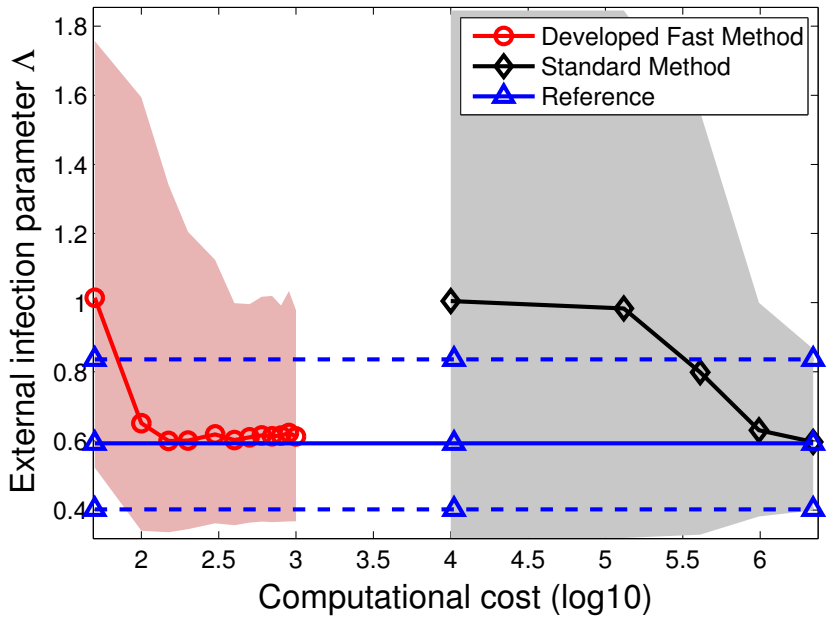
Simulated neural activity in rat brain
(Figure from `https://bbp.epfl.ch/nmc-portal`)

Efficient inference for generative models

- We used machine learning to accelerate the inference by factors of 1000 or more. (JMLR, 2016)
- Techniques used: Optimization, nonlinear regression, decision making under uncertainty (Bayesian optimization)

Open questions and projects:

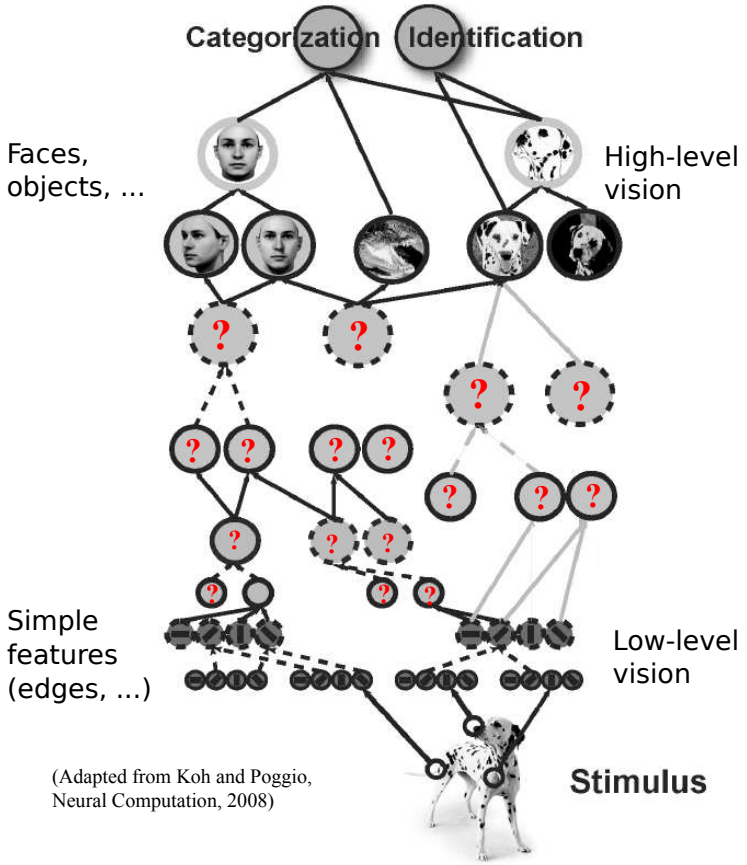
- Scalability: Can we use the approach to jointly infer 1000 variables?
- Automation: Can we find ways to fully automate the inference?
- ...



Unnormalized models

- Models specified in terms of energies rather than probabilities
  - normalizing scale factor (partition function) not computable

- Widely used
  - to model images: Markov random fields
  - to model text: neural probabilistic language models
  - to model networks: exponential random graphs
  - unsupervised deep learning
  - ...

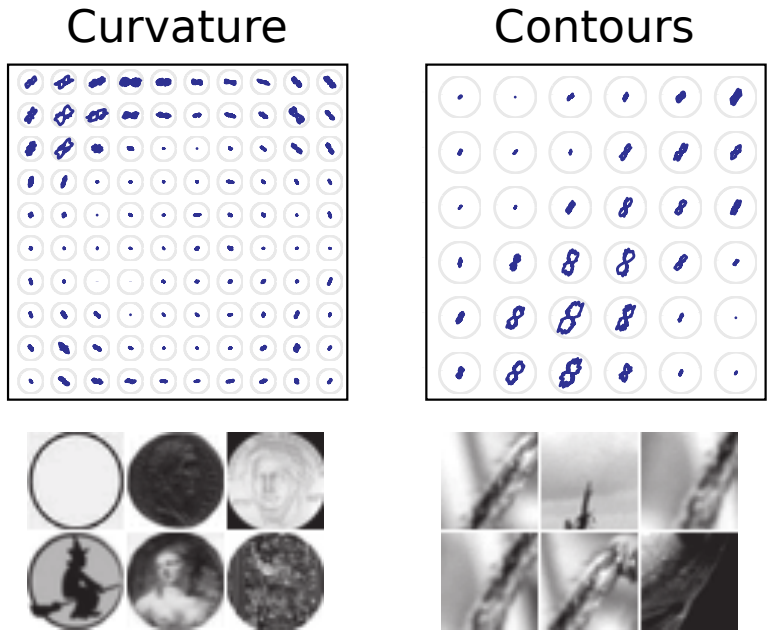


Efficient inference for unnormalized models

- We developed a large class of general inference methods for unnormalized models (AISTATS, 2010; UAI, 2011; JMLR, 2012)
- Techniques used: Optimization, nonlinear classification, math

Open questions and projects:

- Scalability: Method suited for unsupervised deep learning on larger images?
- Automation: How to automatically pick the best inference method?
- ...



Results on 32×32 "images"

(J Physiology-Paris, 2013)

Summary

- Statistical approach to learning from data
- Approximate inference: methods and applications
- Methods and applications related projects available