Bayesian Experimental Design for Large Scale Signal Acquisition Optimization

Matthias Seeger

Laboratory for Probabilistic Machine Learning
Ecole Polytechnique Fédérale de Lausanne
http://lapmal.epfl.ch/

9/12/2013
Magnetic Resonance Imaging

- Extremely versatile
- Noninvasive, no ionizing radiation
- Very expensive
- **Long scan times**: Major limiting factor
Motivation

Magnetic Resonance Imaging

- Faster scans by undersampled reconstruction
- Which fast designs give best images?
Magnetic Resonance Imaging

- Faster scans by undersampled reconstruction
- Which fast designs give best images?
Image Reconstruction

Ideal Image $u$

Measurement

Data $y$

Design

$y \approx Xu$
scan time $\propto \# \text{ phase encodes}

\[ y \approx Xu \]

Reconstruction
Bayesian Experimental Design

- Posterior: **Uncertainty** in reconstruction
- Experimental design: Find poorly determined directions
- Sequential search with interjacent partial measurements
Bayesian Experimental Design

- Posterior: **Uncertainty** in reconstruction
- Experimental design: Find poorly determined directions
- Sequential search with interjacent partial measurements
Score design extension $X_*$ by information gain:

$$I(X_*) = I(y_*, u|y) = H[P(u|y)] - H[P(u|y_*, y)]$$

Most work: Combinatorial aspects
- Assume: $I(X_*)$ tractable to compute
- Assume: $I(X_*)$ cheap to compute (many $X_*$)
- Simple greedy forward works well in practice...
Score design extension $X_*$ by information gain:

$$I(X_*) = I(y_*, u|y) = H[P(u|y)] - H[P(u|y_*, y)]$$

Most work: Combinatorial aspects
- Assume: $I(X_*)$ tractable to compute
- Assume: $I(X_*)$ cheap to compute (many $X_*$)
- Simple greedy forward works well in practice . . .
Score design extension $X_*$ by information gain:

$$I(X_*) = I(y_*, u|y) = H[P(u|y)] - H[P(u|y_*, y)]$$

Most work: Combinatorial aspects

- Assume: $I(X_*)$ tractable to compute
- Assume: $I(X_*)$ cheap to compute (many $X_*$)
- Simple greedy forward works well in practice . . .

So is it . . .?
$I(X_*) = H[P(u|y)] - H[P(u|y_*, y)]$

- Assume: $I(X_*)$ tractable to compute.
  - Only if $P(u|y)$ Gaussian . . .
Whatever images are ... they are not Gaussian!
\[ I(X_*) = H[P(u|y)] - H[P(u|y_*, y)] \]

- Assume: \( I(X_*) \) tractable to compute?
  - No: Needs approximate inference
- Assume: \( I(X_*) \) cheap to compute (many \( X_* \)).
Global covariances
Scores $I(X_\star)$ need full $\text{Cov}_P[u|y]$

Massive scale
$\mathbb{R}^{131072}$ (just one slice).

Many times
Posterior after each design extension
\[ I(X_*) = H[P(u|y)] - H[P(u|y_*, y)] \]

- Assume: \( I(X_*) \) tractable to compute?
  No: Needs approximate inference

- Assume: \( I(X_*) \) cheap to compute (many \( X_* \))?
  No: Needs new algorithms and high performance computing
Approximate inference for non-Gaussian models
Computations driven by Gaussian inference
**Variational Bayesian Inference**

\[ P(u|y) = \frac{P(y|u) \times P(u)}{P(y)} \]

**Variational Inference Approximation**

- Write intractable integration as optimization
- Relax to tractable optimization problem
Variational Bayesian Inference

\[ P(u|y) = \frac{P(y|u) \times P(u)}{P(y)} \]

- Variational Relaxation: Bound the master function

\[ - \log P(y) = - \log \int P(u, y) \, du \leq \frac{1}{2} \min_{u_*} \min_{\gamma} \phi(u_*, \gamma) \]

- Approximate posterior \( P(u|y) \) by Gaussian
- Integration \( \Rightarrow \) Convex optimization
Variational Bayesian Inference

\[ P(u|y) = \frac{P(y|u) \times P(u)}{P(y)} \]

- Variational Relaxation: Bound the master function

\[-\log P(y) = -\log \int P(u, y) \, du \leq \frac{1}{2} \min_{u^*} \min_{\gamma} \phi(u^*, \gamma)\]

- Approximate posterior \( P(u|y) \) by Gaussian
- Integration \( \Rightarrow \) Convex optimization
No Inference Without . . .

Gaussian Variances

Linear Systems

Peanuts $O(q)$
Double loop algorithm

- **Inner loop optimization:**
  Standard MAP Estimation

- **Outer loop update:**
  Gaussian Variances

\[
\min_{\gamma} \phi = \min_{\gamma} \min_{z} \min_{u^*} \phi \\
\overset{!}{=} \min_{z} \min_{u^*} \min_{\gamma} \phi
\]
Double loop algorithm

- **Inner loop optimization:** Standard MAP Estimation
- **Outer loop update:** Gaussian Variances

\[
\min_{\gamma} \phi = \min_{\gamma} \min_{z} \min_{u^*} \phi \\
\approx \min_{z} \min_{u^*} \min_{\gamma} \phi
\]
Monte Carlo Gaussian variances: Perturb&MAP

- Inner loop: Fast first-order MAP solvers
- Warmstarting variational optimization: Small changes after each \((x_*, y_*)\)
Monte Carlo Gaussian variances: Perturb&MAP, Papandreou, Yuille, NIPS 2010

Inner loop: Fast first-order MAP solvers

Warmstarting variational optimization:
Small changes after each \( (X_*, y_*) \)
Optimizing Cartesian MRI

Bayes Optim.  VD Random  Low Pass

Seeger *et al.*, MRM 63(1), 2010  Lustig, Donoho, Pauli, MRM 58(6), 2007  Common MRI practice
Experimental Results: Test Set Errors

- Sagittal short TE
- Sagittal long TE
- Axial short TE
- Axial long TE

Graphs show the $L_2$ reconstruction error for different number of phase encodes and TE settings. Three methods are compared: Low Pass, VD Random, and Bayes Optim.
Large Scale Bayesian Inference

- **Advanced Bayesian experimental design**
  - Signal acquisition optimization

- **Computer Vision**
  - Hierarchically structured image priors
  - Learning Image Models (fields of experts, ...)
  - Bayesian dictionary learning
  - Intelligent user interfaces (Bayesian active learning)

- **Advanced variational inference**
  - Speeding up expectation propagation

- **Generic framework**
  - You can do MAP estimation efficiently?
  - You can do variational Bayesian inference!

Ko, Seeger, ICML 2012

Seeger, Nickisch, AISTATS 2011
Large Scale Bayesian Inference

- Advanced Bayesian experimental design
  - Signal acquisition optimization

- Computer Vision
  - Hierarchically structured image priors
  - Learning Image Models (fields of experts, ...)
  - Bayesian dictionary learning
  - Intelligent user interfaces (Bayesian active learning)

- Advanced variational inference
  - Speeding up expectation propagation

- Generic framework
  - You can do MAP estimation efficiently?
  - You can do variational Bayesian inference!

Ko, Seeger, ICML 2012
Seeger, Nickisch, AISTATS 2011
glm-ie: Toolbox by Hannes Nickisch

mloss.org/software/view/269/

- Generalized sparse linear models
- MAP reconstruction and variational Bayesian inference (double loop algorithm for super-Gaussian bounding)
- Matlab 7.x, GNU Octave 3.2.x

- Hannes Nickisch (now Philips Research, Hamburg)
- Rolf Pohmann, Bernhard Schölkopf (MPI Tübingen)
- Young Jun Ko
- Emtiyaz Khan