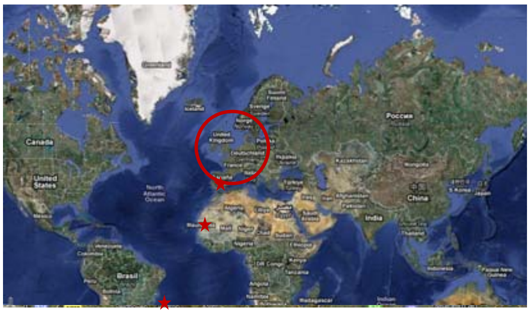
Bayesian Machine Learning for Decoding the Brain

Tom Heskes
Institute for Computing and Information Sciences
Radboud University Nijmegen, The Netherlands

June 10, IWANN 2011, Torremolinos

Radboud University Nijmegen, The Netherlands



Faculty of Science

Huygens building

Emerging Technologies

10 emerging technologies that will change your world

Technology Review 2004

- Universal Translation
- Synthetic Biology
- Nanowires
- Bayesian Machine Learning
- T-Rays (Terahertz Imaging System)
- Distributed Storage
- RNA interference Therapy
- Power Grid Control
- Microfluidic Optical Fibers
- Personal Genomics

Intel research director David Tennenhouse:

"These techniques are going to impact everything we do with computers - from user interfaces to sensor data processing to data mining."

Recent Developments

infer.net

Microsoft launches infer.net (December 2008), a framework for Bayesian inference in graphical models

Navia develops probabilistic computers to do Markov Chain Monte Carlo sampling "on the fly" (2009)

Lyric Semiconductor builds chips specifically geared to probability processing; "the most promising new technology" (2011 EE TIMES ACE Award)

Contents

- Theoretical challenges
  - Bayesian machine learning
  - Exact inference
  - Approximate inference
- Ongoing applications
  - Brain-computer interfacing
  - Bayesian source localization
  - Decoding fMRI

THEORY

Machine Learning

- Learning models out of data
- Bayesian:
  - Enumerate all reasonable models and assign a prior belief
  - Observing the data, evaluate how probable it was under each model
  - Multiply both terms and *normalize* the result

$$P(\Theta|X) = \frac{P(X|\Theta)P(\Theta)}{P(X)}$$

- Following Bayes rule is rational, consistent, coherent, optimal, …

The Challenge

- Compute

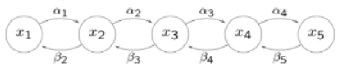
  P(whatever you're interested in | all observations)

- Exact Bayesian methods often require integrations or summations that are intractable / exponentially hard
- Approximate inference techniques
  - Markov Chain Monte Carlo (sampling)
  - Variational approximations
  - *Message passing algorithms*
- Models that are appropriate and (approximately) tractable

Belief Propagation

- Exact algorithm for probabilistic inference on chains and trees

forward messages



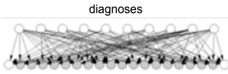
backward messages

- Converged after one forward and one backward pass
- For hidden Markov models (messages are tables) and Kalman filter/smoother (messages are Gaussian potentials)

Loopy Belief Propagation

Approximate inference in probabilistic graphical models with loops

diagnoses



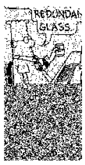
findings

- Bayesian networks for medical diagnosis

  Quick Medical Reference, Promedas

- Turbo codes for signal transmission

  "most exciting and potentially important development in coding theory in many years"



original with 7.5% bit flips

parity bits

Courtesy of David MacKay

Interpretation: fixed points of LBP correspond to local minima of a so-called Bethe free energy

(Yedidia, Freeman, Weiss: NIPS 2001)

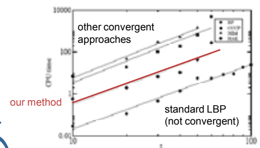
Stable Alternatives

Goal: robust learning and inference in loopy graphical models

Challenge: loopy belief propagation is often unstable

Solution: explicitly solve the implied constrained optimization problem with bound optimization

$$\cup + \cap \leq \cup + \cap$$



From: Pelizzola, J. Phys. A, 2005

Patented; applications to genetic linkage analysis and medical diagnosis

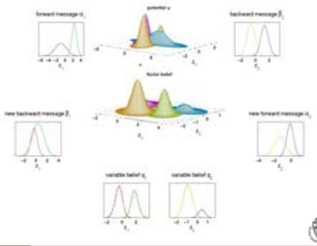
- Heskes: JAIR, 2006
- Albers, Heskes, and Kappen: Genetics, 2007
- Cseke and Heskes: UAI, 2009; JAIR, 2011

Expectation Propagation (Minka, 2002)

Goal: learning and inference in continuous/hybrid Bayesian networks such as switching and extended Kalman filters

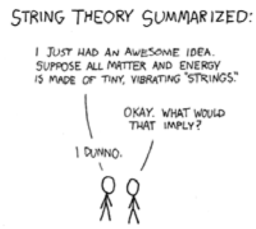
Challenge: exact belief propagation intractable

Solution: local projections to exponential family; only propagate expectations



- Zoeter and Heskes: Statistics and Computing, 2006
- Cseke and Heskes: AISTATS, 2010; JMLR, 2011

APPLICATIONS



Brain-Computer Interfacing

Goal: classifying mental states of the brain

Data: “single trial” EEG/MEG/fMRI

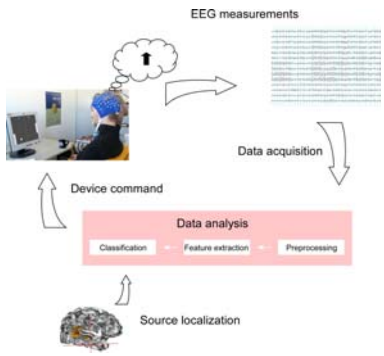
Applications: ALS/CVA patients, games, pilots, …

- How to build robust, adaptive methods for classification?
- How to incorporate neurobiological knowledge?

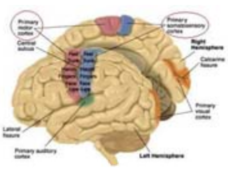
NWO/STW Cognition project in collaboration with the FC Donders Centre for Neurocognitive Imaging

Smartmix Braingain project

The BCI Loop



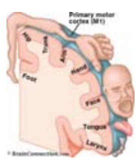
Imagined Movement



Basic assumption: movement *and imagery* elicit an event related desynchronization (ERD) in sensorimotor cortical areas in μ (8-12 Hz) and β (18-26 Hz) bands

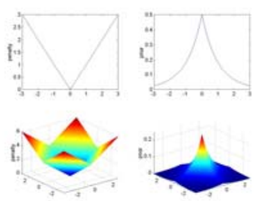
Typical experiment: imagined finger tapping

Data: EEG/MEG across many channels for several seconds for different trials, usually preprocessed to (log) power over frequency bands



L1/Lp Regularized Logistic Regression

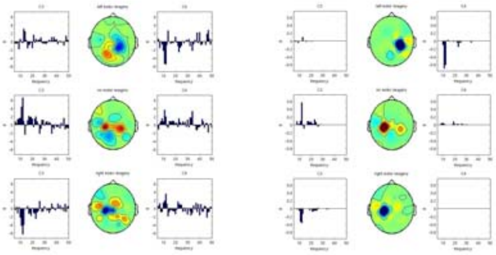
- Logistic regression: equivalent of linear regression, but then for classification
- L1 regularization: sparse solutions, many weights to zero
- L1/L2 regularization: L1 over groups of weights, L2 over weights within a group, sparsity in terms of groups
- Applications: grouping over channels, frequency bands, or subjects (transfer learning)



- Van Gerven, Hesse, Jensen, and Heskes: Neuroimage, 2009

Results on EEG Data

- Regularization leads to a slightly better classification performance
- Much sparser solutions, that are a lot easier to interpret



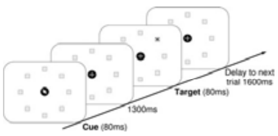
without regularization | with L1/L2 regularization over subjects

The Nijmegen MEG BCI System

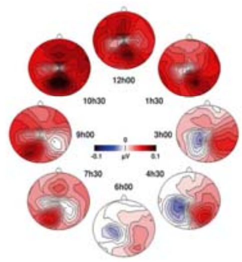
Christian Hesse, Robert Oostenveld, Marcel van Gerven

Covert Attention Modulates Alpha in 2D

Rihs, Michel, and Thut (Eur J Neurosci, 2007):
"Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization"



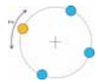
Task: count the number of times the target is on cue



A new paradigm for BCI?

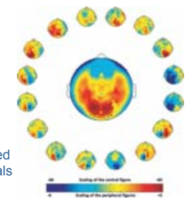
Continuous BCI Based on Covert Attention

Task: fixate in the middle, while covertly attending to a rotating target

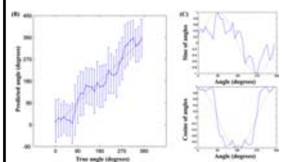


Results:

- Clear changes in alpha power for each angle
- Strong enough to predict the angle attended to in a single-trial setting using circular regression



averaged over trials



single trial performance

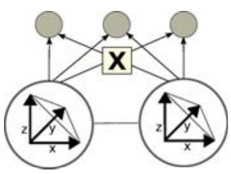
- Bahramisharif, Van Gerven, Heskes, Jensen: Eur J of Neurosci, 2010

Bayesian Source Localization

Assumption: Sensor readings Y are related to source currents S through

Y = **X S** + noise

with **X** a known lead field matrix.



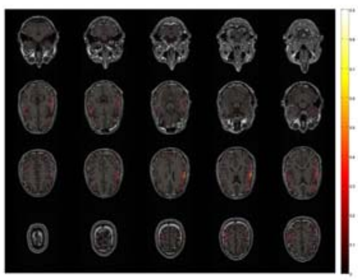
✓ Van Gerven, Cseke, Oostenveld, Heskes: NIPS, 2009

Goal: reconstruct the sources **S** from sensors **Y**

Method: Bayesian linear regression with a multivariate Laplace prior

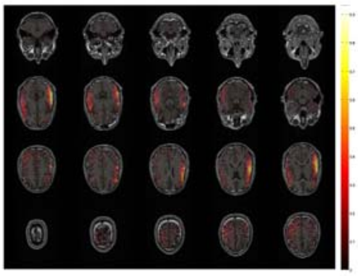
- sparse: just a few sources are relevant
- spatial smoothness: neighboring sources are likely to be both (ir)relevant

Without Spatial Smoothness



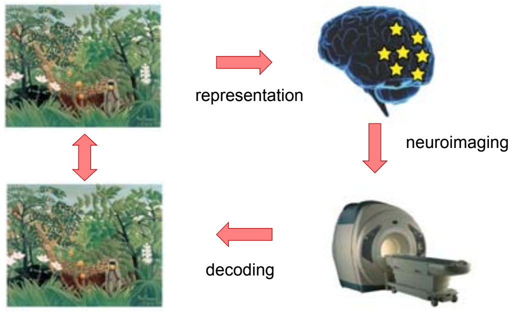
Not so clear

With Spatial Smoothness



Sources where you'd expect them to see

Reading the Brain



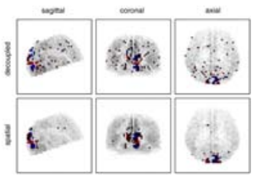
Bayesian Classification of fMRI

Goal: classify 6's versus 9's based on fMRI data

Assumptions:

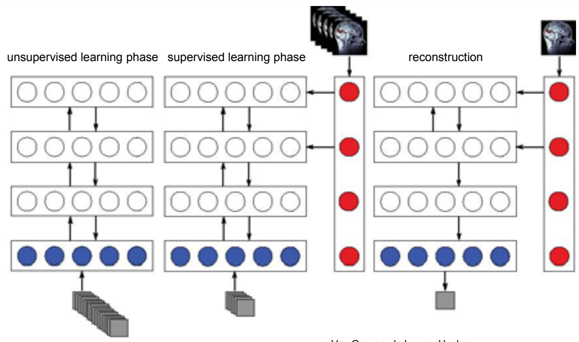
- A small number of relevant voxels (sparsity)
- If a voxel is relevant, then probably also its neighbor (spatial smoothness)

Solution: logistic regression from voxel activities to category labels with a multivariate sparsifying prior



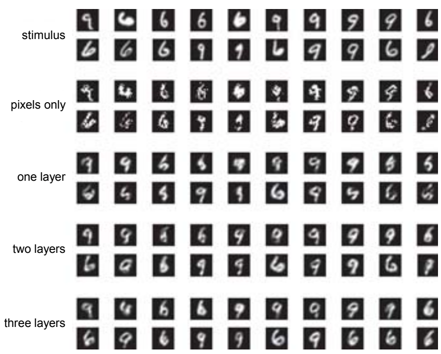
- Van Gerven, Cseke, de Lange, Heskes: Neuroimage, 2010

Reconstruction using Deep Boltzmann Machines



- Van Gerven, de Lange, Heskes: Neural Computation, 2010

Reconstructions



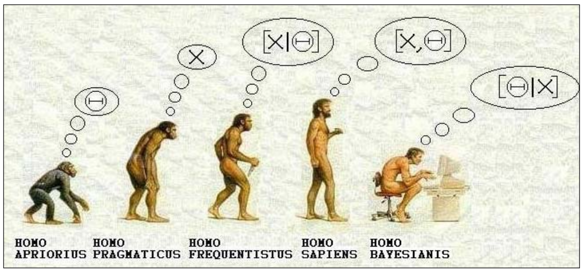
Summary

Bayesian machine learning has many applications, in neuroscience as well as in other fields

Richer, more realistic and more complex models

Better (understanding of) algorithms for approximate inference

Evolution…

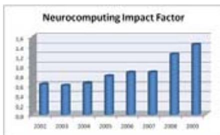


Thanks to:

- Ali Bahramisharif, Botond Cseke, Marcel van Gerven, Onno Zoeter, …
- Kees Albers, Christian Hesse, Ole Jensen, Bert Kappen, Floris de Lange, Robert Oostenveld, …

Neurocomputing

- Theory and applications of neural networks, from computational neuroscience to machine learning
- Impact factor steadily rising: 1.440 (2009)



- IWANN 2009 special issue soon to appear
- Proud to publish the best papers of this year's IWANN!