Bayesian Methods for Multimedia Signal Processing

[Extended Abstract] *

A. Taylan Cemgil
University of Cambridge
Signal Processing and Communications Lab., Department of Engineering
Trumpington Street, CB2 1PZ Cambridge UK
atc27@cam.ac.uk

ABSTRACT

In the last years, there have been a significant growth of multimedia information processing applications that employ ideas from statistical machine learning and probabilistic modeling. In this paradigm, multimedia data (music, audio, video, images, text, ...) are viewed as realizations from highly structured stochastic processes. Once a model is constructed, several interesting problems such as transcription, coding, classification, restoration, tracking, source separation or resynthesis etc. can be formulated as Bayesian inference problems. In this context, graphical models provide a "language" to construct models for quantification of prior knowledge. Unknown parameters in this specification are estimated by probabilistic inference. Often, however, the problem size poses an important challenge and in order to render the approach feasible, specialized inference methods need to be tailored to improve the computational speed and efficiency.

The scope of the proposed tutorial is as follows: First, we will review the fundamentals of probabilistic models, with some focus on music, video and text data. Then, we will discuss the numerical techniques for inference in these models. In particular, we will review exact inference, approximate stochastic inference techniques such as Markov Chain Monte Carlo, Sequential Monte Carlo and deterministic (variational) inference techniques. Our ultimate aim is to provide a basic understanding of probabilistic modeling for multimedia processing, associated computational techniques and a roadmap such that information retrieval researchers new to the Bayesian approach can orient themselves in the relevant literature and understand the current state of the art.

Categories and Subject Descriptors

H.3.1. [**Information Storage and Retrieval**]: Content Analysis and Indexing; G.3 [**Probability and Statistics**]

General Terms

Algorithms

*Slides for this tutorial can be downloaded from `http://www-sigproc.eng.cam.ac.uk/∼atc27/acm-tutorial`

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.
ACM 978-1-59593-701-8/07/0009.

Keywords

Graphical Models, Bayesian Networks, Factor Graphs, Markov Chain Monte Carlo, Sequential Monte Carlo, Variational Bayes, Multimedia Signal Processing

1. DESCRIPTION OF THE TUTORIAL

The Bayes' theorem is a fundamental result in probability theory, which has surprisingly broad practical applications. It is a statement about the relation between conditional and marginal probability distributions of random variables λ and $\mathcal{D}$:

$$\begin{aligned} p(\lambda|\mathcal{D}) &= \frac{p(\mathcal{D}|\lambda)p(\lambda)}{p(\mathcal{D})} \qquad (1)\\ \text{Posterior} &= \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}} \end{aligned}$$

If we interpret λ as an unknown parameter and $\mathcal{D}$ as observed data, this can be simply paraphrased as "What you know about a parameter $\boldsymbol{\lambda}$ after the data $\mathcal{D}$ arrive is what you knew before about $\boldsymbol{\lambda}$ and what the data $\mathcal{D}$ told you" [7]. Hence, the Bayes' theorem tells how to revise our beliefs in light of new evidence – this process is known as *Bayesian inference.*

In the Bayesian paradigm, multimedia data (music, audio, video, images, text, ...) is viewed as realizations from highly structured probabilistic models. Here $\mathcal{D}$ denotes the raw data or some features derived from raw data using deterministic transformations such as feature extractors. The variable λ denotes hidden structure or the quantities we wish to infer. The first component $p(\mathcal{D}|\lambda)$ describes the observation model, referred in equation (1) as the likelihood. The second component is the prior $p(\lambda)$. Once a model is constructed – that is when $p(\mathcal{D}|\lambda)$ and $p(\lambda)$ are specified – several interesting queries about data can be formulated as Bayesian inference problems. This seemingly rather abstract approach has two potential benefits: first, the model specification is separated in a clean and transparent manner from computational details. Second, the computational requirements of various methods can be assessed easily and various alternative computational methods can be employed to balance quality/computation time tradeoff.

Constructing useful models and inference algorithms is facilitated through the use of *graphical models* [5, 1]. In this respect, graphical models can be viewed as a formal language/notation for specification of probability distributions and the associated inference algorithms. This powerful language provides graph based algorithms for derivations and computation as well as pedagogical insight and motivation for model construction. Historically, they were introduced in probabilistic expert systems [8] as a visual guide for representing expert knowledge. Today, they are extensively used as an almost standard tool in machine learning, statistics and signal processing. Moreover, software tools for automated code generation start to emerge, making the design - implement - test cycle for deploying applications shorter.

Often, however, realistic signal models for multimedia signals form a tough class to solve: the problem size may pose significant challenges. In particular, these models require sophisticated Bayesian model uncertainty management (e.g. to infer dimensionality or hyper-parameters) and high dimensional parameter inference, for which there are no known analytical solutions. Therefore, numerical integration and optimisation techniques have to be applied and novel methods need to be developed to improve solution quality and computational efficiency.

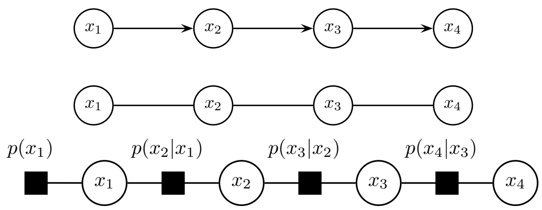

Figure 1: Graphical model denoting the distribution $p(x_1)p(x_2|x_1)p(x_3|x_2)p(x_4|x_3)$. Top to bottom, a directed graphical model (Bayesian network), an undirected model (a Markov Random field) and a factor graph. Although all these models are equivalent they are complimentary for model construction and understanding inference algorithms.

The scope of the proposed tutorial is as follows: Using the graphical model formalism, we will review the fundamentals of probabilistic models for music, video and text processing and introduce popular model structures such as hidden Markov models, Kalman filters, independent component analysis, latent semantic indexing, latent Dirichlet allocation, e.t.c. In the context of individual models, we will also introduce exact inference and approximate inference methods. In particular, we will describe sampling based methods for Bayesian inference, namely Markov Chain Monte Carlo (MCMC) [4, 6] and Sequential Monte Carlo (particle filters) [3, 2]. We will also introduce computationally cheaper optimisation and approximation methods. In particular, we will incorporate variational methods, based on tractable analytical approximations to otherwise intractable posterior distributions, which recently had good success in the machine learning community, see [9] for a general review. Our treatment will aim at contrasting Monte Carlo methods with potentially faster variational methods. This will form the basis for developing hybrid algorithms.

Our treatment will be aimed at an introductory level so no prior knowledge of advanced statistical concepts is required. Our ultimate aim is to provide a basic understanding of probabilistic modeling for multimedia processing, associated computational techniques and a roadmap such that information retrieval researchers new to the Bayesian approach can orient themselves in the relevant literature and understand the current state of the art.

2. REFERENCES

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[2] A. Doucet, N. de Freitas, and N. J. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer Verlag, 2001.

[3] A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.

[4] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov Chain Monte Carlo in Practice*. CRC Press, London, 1996.

[5] M. I. Jordan, editor. *Learning in graphical models*. MIT Press, 1999.

[6] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer, 2004.

[7] D. J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[8] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, San Mateo, CA., 1988.

[9] M. Wainwright and M. I. Jordan. Graphical models, exponential families, and variational inference. Technical Report 649, Department of Statistics, UC Berkeley, September 2003.