Computational inference in dynamical system
– a PhD course at ACFR

Thomas Schön
Division of Automatic Control, Linköping University
E-mail: schon@isy.liu.se

Overview

Part 1 - Getting started

1.1 Brief introduction (explaining what we will do during these 8 hours and provide some basic model assumptions).

1.2 Derive the general expressions for computing the filtering and various smoothing densities for the states in nonlinear dynamical systems.

Part 2 - EM and MCMC explained via linear system identification

2.1 Derive the expectation maximisation (EM) for computing maximum likelihood (ML) estimates.

2.2 Identification of linear, Gaussian state-space (LGSS) models using EM.

2.3 Markov chain Monte Carlo (MCMC) methods for computing Bayesian estimates. Introducing the Metropolis Hastings sampler and the Gibbs sampler in general and illustrate their use on identification of LGSS models.

Part 3 - Nonlinear state inference using sequential Monte Carlo (SMC)

3.1 Particle filtering (introduced via importance sampling). Theory and applications.

3.2 Particle smoothing (PS).

Part 4 - Nonlinear system identification

4.1 Computing ML estimates using EM and PS. Use Wiener system identification as one example.

4.2 Computing Bayesian estimates using particle Markov chain Monte Carlo (PMCMC) methods.

Course home page: `www.rt.isy.liu.se/~schon/Courses/CILDS_USYD2012/index.html`

Abstract

The overall aim in this course is to provide an introduction to the theory and application of computational methods (some of them only a couple of years old) for inference in dynamical systems. More specifically, the computational methods we are referring to are sequential Monte Carlo (SMC) methods (including particle filters and particle smoothers) for nonlinear state inference problems and expectation maximisation (EM) and Markov chain Monte Carlo (MCMC) methods for nonlinear system identification.

Dealing with the nonlinear system identification problem will require nonstandard combinations of these methods. We will work almost exclusively with state-space models, linear models to introduce the methods and nonlinear models to illustrate the capabilities of the methods. It is our firm belief that even if you aim for solving nonlinear problems, you should always make sure that the method under study is capable of solving basic linear problems first. If that cannot be done, the method does not stand a chance in solving the nonlinear problem either. Furthermore, a good understanding of linear models is important in order to be able to understand nonlinear models. Our aim throughout this course is to introduce the methods by answering simple questions relating to linear models and then (most importantly) show that the methods are capable of tackling nonlinear problems as well.

After a brief introduction we will derive general expressions for computing filtering and various smoothing densities for the states in nonlinear dynamical models. The basic strategies employed in both maximum likelihood (ML) and Bayesian system identification are then reviewed. The EM algorithm is then derived and we will show how it can be used to compute ML estimates in linear models. We then turn our attention to the MCMC methods and show how these methods can be used to solve the linear system identification problem. This involves some interesting developments requiring the use of the matrix-Normal and the inverse-Wishart distributions and the so called simulation smoothers.

The linear models have so far served the purpose of testing grounds for introducing the EM and the MCMC methods. However, it is now time to leave the linear models behind and turn our attention to nonlinear models instead. The SMC methods (focusing on the particle filter) will be introduced and the basic theory is provided. We will also show how the particle filter has been used to solve some nontrivial nonlinear filtering problems we have been working on together with various companies. The particle smoother is briefly introduced.

Finally, we will show how the methods introduced above can be used to solve various problems in nonlinear system identification. We start by showing how to compute ML estimates using EM (involving particle smoothers and nonlinear optimisation) and we will illustrate how this can be used to solve various problems, including some Wiener identification problems. Finally, the recent (and exciting) development referred to as the particle MCMC (PMCMC) methods will be introduced. Using PMCMC we are capable of solving nonlinear Bayesian system identification problems by a nontrivial combination of the MCMC methods and the SMC methods.