APPLICATIONS OF MARKOV CHAIN MONTE CARLO AND POLYNOMIAL CHAOS EXPANSION BASED TECHNIQUES FOR STATE AND PARAMETER ESTIMATION

An Undergraduate Research Scholars Thesis

by

SHUILIAN XIE

Submitted to Honors and Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Research Advisor: Dr. Krishna R. Narayanan

May 2014

Major: Electrical Engineering

TABLE OF CONTENTS

Page

ABSTRACT 1

DEDICATION 3

ACKNOWLEDGMENTS 4

NOMENCLATURE 5

I INTRODUCTION 6

II STATE ESTIMATION USING MARKOV CHAIN MONTE CARLO METHODS 9

Problem Statement 9

Sequential Monte Carlo Approach for Estimation of States-SIS 10

Sequential Monte Carlo Approach for Estimation of States-SIR 11

Results of MCMC method in estimation of states 12

Conclusion 16

III PARAMETER ESTIMATION USING MARKOV CHAIN MONTE CARLO METHOD 17

Problem Statement 17

Kernel Smoothing Algorithm 17

Results of MCMC Method in Estimation of Parameter 19

Conclusion 21

IV PROBABILITY DENSITY FUNCTION ESTIMATION USING POLYNOMIAL-CHAOS EXPANSION 22

Pre-knowledge of Polynomial-Chaos Expansion 22

Problem Statement 22

Page

Results of Polynomial Chaos in Estimation PDF 23

Conclusion 26

V FUTURE WORKS 27

REFERENCES 28

ABSTRACT

APPLICATIONS OF MARKOV CHAIN MONTE CARLO AND POLYNOMIAL CHAOS EXPANSION BASED TECHNIQUES FOR STATE AND PARAMETER ESTIMATION .
(May 2014)

SHUILIAN XIE
Department of Electrical and Computer Engineering
Texas A&M University

Research Advisor: Dr. Krishna R. Narayanan
Department of Electrical and Computer Engineering

In this research thesis, we implement Markov Chain Monte Carlo techniques and polynomial-chaos expansion based techniques for states and parameters estimation in hidden Markov models (HMM). Our goal is to estimate the probability density function (PDF) of the states and parameters given noisy observations of the output of the hidden Markov model. We consider three problems, namely, (i) determining the PDF of the states in a non-linear HMM using sequential MCMC techniques, (ii) determining the parameters of discretized linear, ordinary differential equations (ODE) given noisy observations of the solutions and (iii) Determining the PDF of the solution of a linear ordinary differential equation when the parameters of the ODE are random variables. While these problems naturally arise in several areas in engineering, this thesis is motivated by potential applications in bio-mechanics. One of the interesting research questions that is being considered by some researchers is whether the formation of clots can be predicted by observing the mechanical properties of arteries, such as their stiffness. In order for this approach to be successful, it is critical to estimate the stiffness of arteries based on noisy measurements of their mechanical response. The parameters of these models can then be used to differentiate diseased arteries from healthy ones or, the parameters can be used to predict the probability of formation of plaques. From experimental data, we would like to infer the posterior density of the states and parameters

(such as stiffness), and classify it as being healthy or diseased. If it is accomplished, this will improve the state-of-the art in modeling mechanical properties of arteries, which could lead to better prediction, and diagnosis of coronary artery disease.

DEDICATION

I dedicate my undergraduate research work to my family and many friends. A special feeling of gratitude to my loving parents who support me to study at Texas A&M University. My sisters Lulu and Jinping have never left my side.

I also dedicate this research work to my friends, without whose help, I could not have been motivated and positive. I will always appreciate all they have done, especially Xintong Xia and Shan Wang for helping me develop my strong heart and optimistic altitude.

ACKNOWLEDGMENTS

First and foremost, I would like to take this opportunity to express my deepest gratitude to my research advisor, Dr. Krishna R. Narayanan, for providing me a great opportunity to learn about statistical signal processing. Ever since the first days when he taught me Digital Communication, Krishna has been an outstanding mentor and role-model. His enthusiasm for research and teaching is inexhaustible. His brilliant insight, patient guidance throughout my senior year helped to lit my route of future study. Working with him was indeed a pleasant and rewarding experience. Without his continuous support, my undergraduate research would have never become possible.

I would also like to thank Dr. Arun R. Srinivasa, a professor in Department of Mechanical Engineering at Texas A&M University. Dr. Srinivasa provided a bunch of profound insights and helpful resource to our project.

Furthermore, I'm very grateful to all the graduate students under Dr. Krishna Narayanan's advisory, especially Avinash Vem, who offered great help for my research project.

Last, I want to thank the Department of Electrical and Engineering, Texas A&M University for providing me a platform to pursue my research interest and giving me award to honor my research achievements as an undergraduate. Additionally, I would like to appreciate Dr. Narayanan's funding support, which helped ease my financial burdens.

NOMENCLATURE

MCMC	Markov Chain Monte Carlo
N_s	Particle Number
PCE	Polynomial Chaos Expansion
PDF	Probability Density Function
SIS	Sequential Importance Sampling
SIR	Sequential Importance Resampling
SMC	Sequential Markov Chain

CHAPTER I

INTRODUCTION

We consider three closely-related problems in statistical signal processing in this thesis. These problems pertain to inferring the posterior distribution of the states or parameters of a discrete-time hidden Markov model given noisy observations of the output of such a model. More specifically, we consider the following discrete-time hidden Markov model

$$x_k = f_k\left(x_{k-1}, \theta_k\right) + v_{k-1} \tag{I.1}$$

$$z_k = g_k(x_k) + n_k \tag{I.2}$$

where x_k and z_k denote the state and observation at time instant k, respectively. θ_k is a vector of parameters and v_k and n_k denote i.i.d noise sequences.

The following three problems are considered. (i) Estimating the probability density function (PDF) of x_k given observations z_ks, (ii) Estimating the PDF of the parameters θ_k given observations z_ks and (iii) Estimating the PDF of z_ks given the PDF of θ_ks. Two main techniques are used to accomplish these tasks. We use Markov chain monte carlo techniques to accomplish tasks (i) and (ii) and we use polynomial-chaos based expansion techniques to accomplish task (iii).

While hidden Markov models and the aforementioned estimation problems naturally occur in several engineering applications, the study in this thesis is mainly motivated by applications in bio-mechanics. The broader context within which this study was undertaken is described below. Currently, there is interest within the bio-mechanics research community to answer the question of whether the mechanical properties of arteries can be used to predict the formation of arterial plaques. An important first step in addressing this question is to find a mathematical model that explains the mechanical behavior of the artery. In particular, if we can conduct an experiment (on dead tissue) where we apply a combination of forces and torques and measure the expansion of the artery, can we then fit a mathematical model that will explain the response of the artery? Current modeling methods in biomechanics largely assume that the expansion of the artery is a deterministic function of the input force and try to find models, typically the shape of the artery is given by the solution to a differential equation. It is true that such deterministic models have been very successful in modeling man-made materials. However, using such deterministic models to obtain biomechanics models has failed because there is no simple one to one relationship between the parameters of the model and measured values. In addition, there is huge variation in the response from one tissue sample to the other. Our approach is to model the unknown parameters (e.g. elasticity) as random variables and obtain stochastic models for the arterial response.

More specifically, we will assume that the shape of the artery is the solution to a (possibly non-linear) differential equation whose parameters are unknown. Further, from experiments, we can observe the shape of the artery either entirely or partially in the presence of some measurement noise. When this model is appropriately discretized, it can be seen that the resulting model falls in the framework of a hidden Markov model as given in equations (I.1) and (I.2).

Performing the estimation task described above is not an easy because the underlying models are often non-linear. The objective of this project is to explore two powerful ideas in statistical signal processing to carry out these estimation tasks. The first one is the idea of Markov Chain Monte Carlo (MCMC) methods [1],[2]. The second one is the idea of using polynomial-chaos based model fitting [3]. We believe that these will be powerful, effective and feasible ways to perform estimation tasks in the presence of non-linearities and/or unknown parameters.

An important consequence of being able to perform these estimation tasks well is that the results of estimation can be used for diagnostic purposes. For example, if one can obtain the distribution of the unknown parameters from experimental data, this can be used to classify the artery as being healthy or diseased or the likelihood of the artery developing into a diseased artery can be estimated. Even though these classification problems are not addressed in this thesis, the estimation step can be seen to be crucial for the classification problem.

The rest of the thesis is organized as follows. In Chapter II, we discuss sequential monte carlo techniques, in particular, the particle filtering technique for estimating the states of a HMM. We discuss the degeneracy problem associated with naive particle filtering techniques and we consider improved sampling techniques based on resampling. The algorithms considered in this chapter are based on those in [5]. In Chapter III, we consider the problem of estimating unknown parameters of a linear ordinary differential equation by observing a noisy version of the output of the differential equation. We discuss why traditional sequential monte carlo techniques are not well-suited for this problem. We implement a kernel-smoothing based sequential monte carlo technique based on [6] for the estimating the parameters. We discuss the limitations of such a scheme for the problem that we studied. Finally in Chapter IV, we consider the use of polynomial-chaos based expansion techniques for estimating the PDF of the output of a linear ODE, when the parameters in the ODE are random variables. We implement this scheme to estimate the PDF of the output of a first-order ODE. Chapter V discusses some future work that can be performed to continue research along the direction of research considered in this thesis.

CHAPTER II

STATE ESTIMATION USING MARKOV CHAIN MONTE CARLO METHODS

Problem Statement

Consider a hidden Markov state-space model given by

$$x_k = f_k(x_{k-1}, \theta) + v_k \tag{II.1}$$

$$z_k = g_k(x_k) + n_k \tag{II.2}$$

where (II.1) and (II.2) give the state x_k and the observation z_k at time instant k, respectively. Note that v_k and n_k denote i.i.d noise sequences. We wish to determine the posterior pdf of x_k given the observations z_0^T, where z_0^T denotes the vector $\{z_k, i = 0, .., T\}$ and T is the maximum time for which observations are available. This is a special case of (I.1) and (I.2) with θ being fixed.

Is is well known that the Kalman filter is optimal for determining the posterior pdf under the following conditions:

- v_k and n_k are drawn from Gaussian distribution of known parameters.
- $f_k(x_{k-1}, \theta)$ is known and is a linear function of x_{k-1}.
- $g_k(x_k)$ is a known linear function of x_k.

However, when the noise sequences are not Gaussian, nor f and g are linear, the Kalman filter is not an optimal solution for this tracking. In this case, sequential monte carlo approaches have been very successful for the estimation of states x_ks. [5]

Sequential Monte Carlo Approach for Estimation of States-SIS

The Sequential Importance Sampling (SIS) Particle Filter is a Monte Carlo method that is used for states and parameters estimation. In this approach, we use $\{x_k^i\}$ and a corresponding set of weights w_k^i to characterize the posterior density pdf $p(x|z)$. The key idea of this approach is to represent the pdf by these random samples x_k^i with associate weights w_k^i, under the noisy measurements z_ks. In the SIS algorithm, the random sample $x_{0:k}^i$ are drawn from (II.1), which shows the relationship between the previous state and current state. The next step is to assign the particle a weight. The weights are updated according to

$$w_k^i \propto \frac{p\left(z_k|x_k^i\right) p\left(x_k^i|x_{k-1}^i\right)}{q\left(x_k^i|x_{0:k-1}^i, z_{1:k}\right)}. \tag{II.3}$$

where $q\left(x_k^i|x_{0:k-1}^i, z_{1:k}\right)$ is called the proposal density function, and we can choose $q\left(x_k^i|x_{0:k-1}^i, z_{1:k}\right)$ to be anything that is easy to sample from. To simplify our problem, we define

$$q\left(x_k^i|x_{0:k-1}^i, z_{1:k}\right) \triangleq p\left(x_k^i|x_{k-1}^i\right). \tag{II.4}$$

so that it follows that

$$w_k^i \propto w_{k-1}^i p\left(z_k|x_k^i\right). \tag{II.5}$$

The weights w_k^i are normalized such that $\sum_i w_k^i = 1$.

Once we get the random measure $\left[\{x_k^i, w_k^i\}_{i=1}^{N_s}\right]$, we can calculate the posterior filtered density $p\left(x_k|z_{1:k}\right)$ as

$$p\left(x_k|z_{1:k}\right) \approx \sum_{i=1}^{N_s} w_k^i \delta\left(x_k - x_k^i\right) \tag{II.6}$$

It can be shown that as $N_s \to \infty$, the approximation approaches the true posterior density.[5]

A pseudocode for the algorithm is presented below:

Algorithm 1: Sequential Important Sampling (SIS) Particle Filter [5]

$\left[\{x_k^i, w_k^i\}_{i=1}^{N_s}\right] = SIS\left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}\right]$

- FOR $i = 1 : N_s$
  - Draw $x_k^i \sim q\left(x_k | x_{k-1}^i, z_k\right)$
  - Assign the particle a weight, $w_k^i \propto w_{k-1}^i \frac{p\left(z_k|x_k^i\right)p\left(z_k|x_{k-1}^i\right)}{q\left(x_k^i|x_{k-1}^i, z_k\right)}$

    where $q\left(x_k^i | x_{k-1}^i, z_k\right)$ is called the proposal density function
- END FOR

Sequential Monte Carlo Approach for Estimation of States-SIR

After the implementation of SIS, we find that there are some negligible weights whose contribution to $p\left(x_k|z_{1:k}\right)$ is almost zeros. When this happens, small weights take a large computational effort to update; however they do not contribute substantially to the overall pdf. This is called degeneracy problem. In order to solve this problem, it is common to use resampling algorithm. The basic idea of resampling is to eliminate particles that have small weights and to concentrate on particles with large weights. In the algorithm, we firstly construct a CDF of the weights. To determine whether the weight is small or large, we utilized a vector called u_j, shown in the resampling algorithm below. If u_j is less than the value of CDF, we regard the corresponding weight large and then assign a new weight as $\frac{1}{N_s}$. Otherwise, we can say that the weight is small enough to eliminate. Therefore, we could see that resampling involves generating a new set weights w_k^i as $\frac{1}{N_s}$

A pseudo code for the algorithm is presented:

Algorithm 2: Resampling Algorithm[5]

$\left[\left\{x_k^{j*}, w_k^{j*}, i^j\right\}_{i=1}^{N_s}\right] = RESAMPLE\left[\left\{x_{k-1}^i, w_{k-1}^i\right\}_{i=1}^{N_s}\right]$

- Initialize the CDF: $c_1 = 0$
- FOR $i = 2 : N_s$
  - Construct CDF: $c_i = c_{i-1} + w_k^i$
- END FOR
- Start at the bottom of the CDF: $i = 1$

- Draw a starting point: $u_1 \sim U\left[0, \frac{1}{N_s}\right]$
- FOR $j = 1 : N_s$
  - Move along the CDF: $u_j = u1 + \frac{1}{N_s}(j-1)$
  - WHILE $u_j \geq c_i$
  - $* \; i = i + 1$
  - END WHILE
  - Assign sample: $x_k^{j*} = x_k^i$
  - Assign weight: $w_k^j = \frac{1}{N_s}$
  - Assign parent: $i^j = i$
- END FOR

Sequential Importance Resampling (SIR) Algorithm is a combination of SIS and Resampling algorithm, which means that we firstly obtain the random measure $\left[\{x_k^i, w_k^i\}_{i=1}^{N_s}\right]$, and then implement resampling algorithm to generate a new set of the random measure $\left[\left\{x_k^{j*}, w_k^i\right\}_{i=1}^{N_s}\right]$. After that, we get the posterior filtered density $p\left(x_k|z_{1:k}\right)$, which is showed in the previous section.

Results of MCMC method in estimation of states

Example 1 We consider the estimation of x_k by the SIS algorithm for the following example:

$$\begin{aligned} x_k &= \frac{x_{k-1}}{2} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + v_{k-1} \\ z_k &= \frac{x_k^2}{20} + n_k \end{aligned} \tag{II.7}$$

where v_k and n_k are zero mean Gaussian random variables with variance 10 and 1, respectively. We consider 1,000 particles and time up to 50 units. The value of x_0 is drawn uniformly between -25 and 25. Fig II.1 presents the tracking of the states x_k as time, which we call the estimation of posterior density function of x_k. Red dots mean there are higher possibility for x_k to fall into corresponding small interval. Fig II.2 shows the true value of x_k versus time k. We could mainly see that as time becomes larger, the estimation is more accurate. The posterior density function shows the effectiveness of Sequential Importance Sampling algorithm. Besides, calculating the Root Mean Squared Error (RMSE) can also represent the performance of sequential monte carlo filter. Where

$$RMSE = \sqrt{\sum_{k=1}^{T} \left(x_k - x_k^i\right)^2 w_k^i} \tag{II.8}$$

From above equation, we could compute the RMSE of SIS is about 6.83.

Due to degeneracy problem of SIS, we implement another algorithm, Sequential Importance Resampling. Fig II.3 presents the tracking of the states x_k as time, which we call the estimation of posterior density function of x_k. Fig II.4 shows the true value of x_k versus time k. we would see the SIR method also works well. Also, we compute the RMSE of SIR, which is about 5.69.

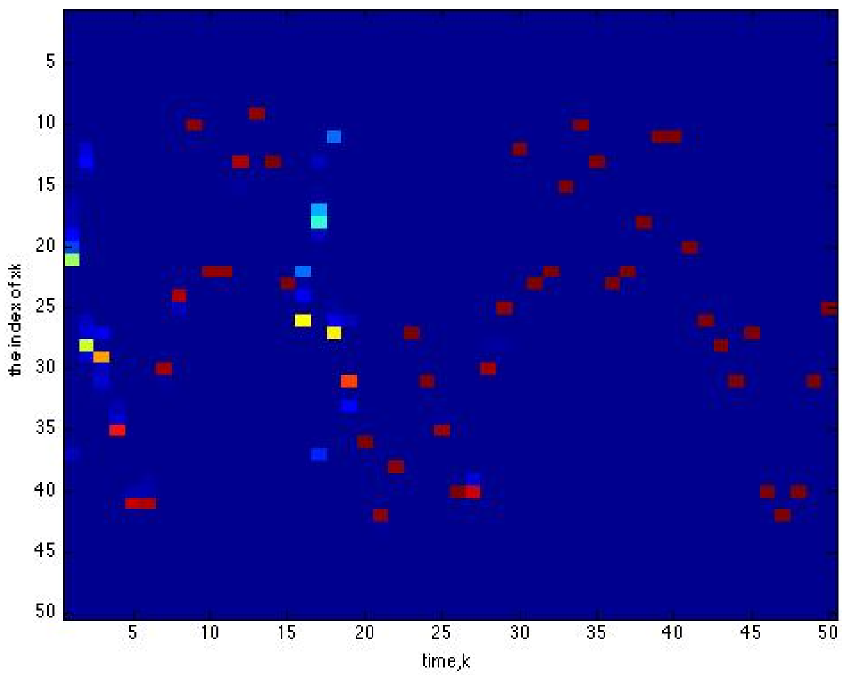

Fig. II.1.: The posterior density function of x_k by SIS

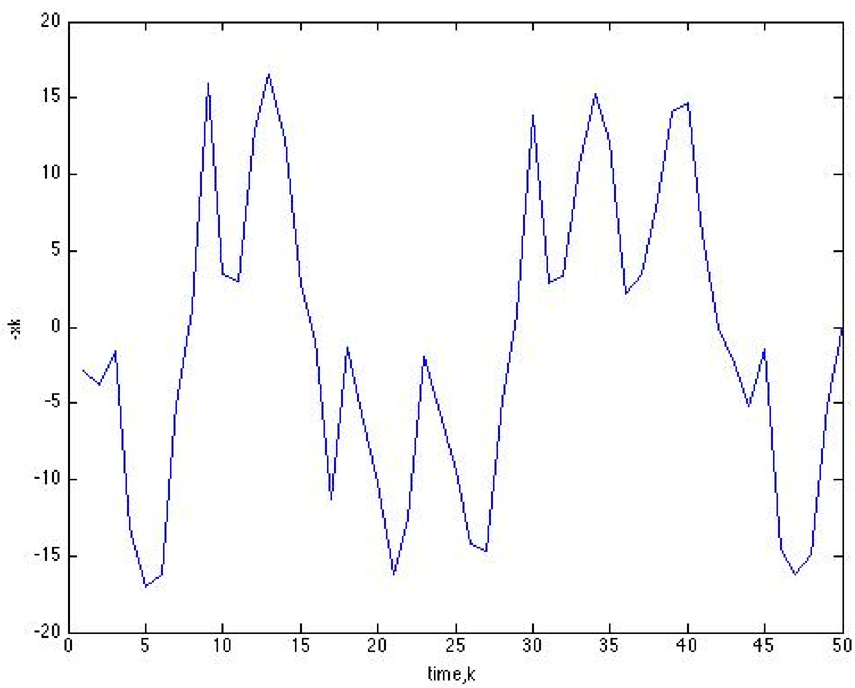

Fig. II.2.: The plot of the true value of x_k vs k

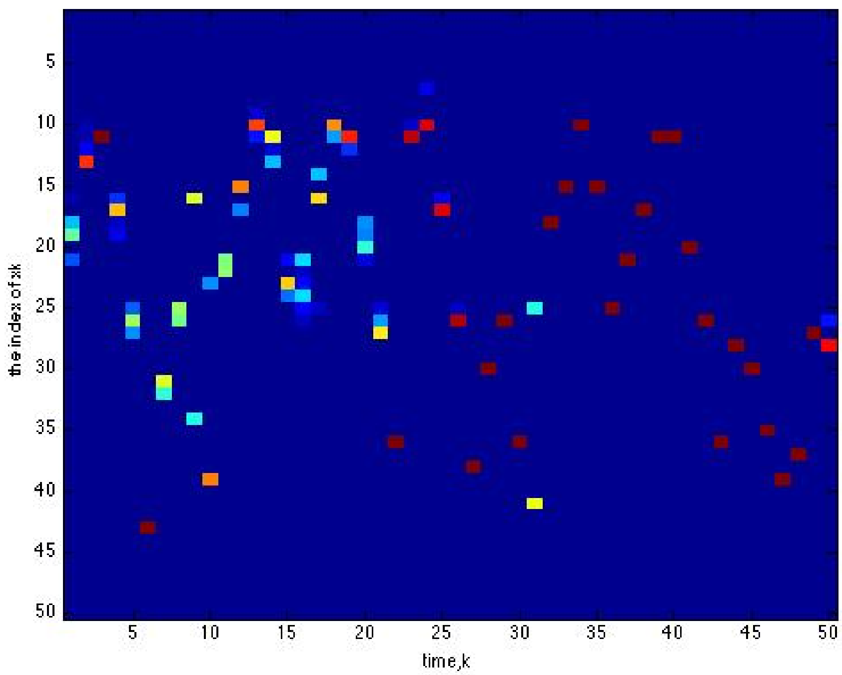

Fig. II.3.: The posterior density function of x_k by SIR

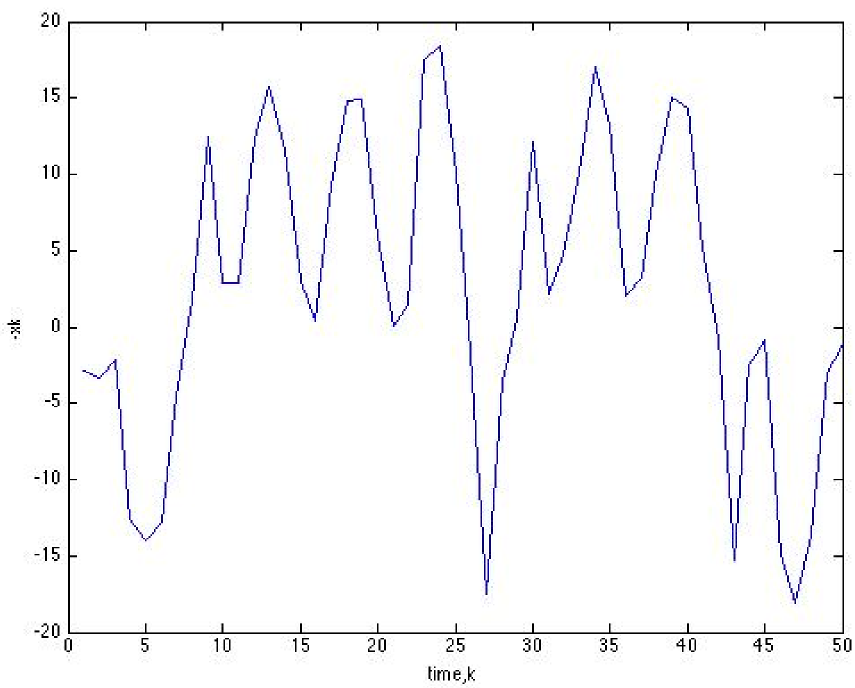

Fig. II.4.: The plot of the true value of x_k vs k

Conclusion

In this chapter we showed that Markov chain monte carlo methods can be very effective for state estimation in hidden Markov models. Our simulation results shows that after an initial period, the particle filtering algorithm is able to track the states well. Resampling is an effective technique to deal with the degeneracy problem. By concentrating the updating effort on large weights, the resampling technique is able to decrease the estimation error.

CHAPTER III

PARAMETER ESTIMATION USING MARKOV CHAIN MONTE CARLO METHOD

Problem Statement

The previous chapter deals with the state estimation when the parameter θ is being fixed. In this chapter, we mainly focus on the estimation of parameter θ. Again, we assume a Hidden Markov model given by

$$y_k = f_k(y_{k-1}, \theta) + v_k \tag{III.1}$$

$$z_k = g_k(y_k) + n_k \tag{III.2}$$

where (III.1) and (III.2) give the state y_k and the observation z_k at time instant k, respectively. Note that v_k and n_k denote i.i.d noise sequences. In this problem, we are going to deal with the estimation of posterior density function of θ.

Kernel Smoothing Algorithm

The combination of Kernel Smoothing Algorithm and particle has been shown to work well in some cases. Suppose we have $\{y_k^i, \theta_k^i\}$, and associated weights $\{w_k^i\}$ that together represent a monte carlo importance sample. We can see that the weights w_k^i are able to represent the probability density function of θ.

The basic idea is to regard θ as time-varying with small random perturbations[6]. One way is to add an independent, zero-mean normal increment to the parameter at each time. That is,

$$\theta_{k+1} = \theta_k + \zeta_{k+1} \tag{III.3}$$

$$\zeta_{k+1} \sim N(0, W_{k+1})$$

where W_{k+1} is independent with given states and observations.

When updating θ_k^i, we computer the new sample as $N\left(m_k^{(k)}, h^2 V_k\right)$, where $m_k^{(k)}$ is called the locations of θ, shown as the algorithm below, and h is chozen to make the new sample more concentrated about to their locations. In terms of corresponding weights w_k^i, we have the same method shown as SIS algorithm, that is

$$w_k^i \propto w_{k-1}^i p\left(z_k | y_k^i\right). \tag{III.4}$$

The weights w_k^i are normalized such that $\sum_i w_k^i = 1$.

Algorithm 3: Kernel Smoothing Algorithm[6]

- Sample an auxiliary integer variable from the set $\{1...N_s\}$ with probabilities of w_k^i, call the sampled index k

- Sample a new parameter vector θ_{k+1}^i from the k^{th} normal component of the kernel density, namely $\theta_{k+1}^{(k)} \sim N\left(m_k^{(k)}, h^2 V_k\right)$ where $m_k^{(k)} = a\theta_k^{(k)} + (1-a)\overline{\theta_k}$, $a = \frac{3\delta-1}{2\delta}$, $h = \sqrt{1-a^2}$, δ is called discount factor, typically around $0.95 - 0.99$. V_k is the variance of θ_k^i

- Sample a value of current state vector from the system equation $p\left(x_{k+1}^i | x_k^{(k)}, \theta_{k+1}^{(k)}\right)$

- Evaluate the corresponding weight $w_k^i \propto w_{k-1}^i \frac{p\left(z_k|x_k^i\right)p\left(z_k|x_{k-1}^i\right)}{q\left(x_k^i|x_{k-1}^i, z_k\right)}$

Results of MCMC Method in Estimation of Parameter

Example 2 In this example, We have the following model:

$$y_{k+1} = y_k + y_k \theta \Delta \tag{III.5}$$
$$z_k = y_k + n_k$$

In our experiment, to generate the observations, we set the value of θ as -0.2. During the estimation step, we set the number of particles N_s as 10,000 and time T as 200. Fig III.1 has four subfigures, the first one shows the distribution of θ when $k = 1$, and from the figure, we can see that the values of θ are around a fixed value as -0.2. The rest of subfigures also have shown the estimated θ is a true distribution. Fig III.2 is a histogram showing the distribution of θ when $k = T$. From these two figures, we can see that during the implementation of Kernel Smoothing Algorithm, θs are approaching to a certain value, which shows the effectiveness of this method to estimate parameters.

Discussion of this example: When we implemented this algorithm, we found that when Δ is very small, say 0.01, the estimated θ will concentrate to a random number between -1 and 0, instead of a fixed number. We believe the reason for this is as follows: when Δ is very small, the truly estimated value of θ will not affect the first few steps of evolution of θ, especially when we utilize resampling to make estimated θ more concentrated to "wrong" θ, which is a random value between -1 and 0. In addition, the joint estimation does not produce results consistently. This is because the correct deduction of next state from the current state should be

$$y_{k+1}\left(\theta^i_{k+1}\right) = y_k\left(\theta^i_k\right) + y_k\left(\theta^i_k\right)\Delta\theta^i_{k+1} \tag{III.6}$$

while in our example, the computation for a new sample of y_k is:

$$y_{k+1}\left(\theta^i_{k+1}\right) = y_k\left(\theta^i_{k+1}\right) + y_k\left(\theta^i_{k+1}\right)\Delta\theta^i_{k+1} \tag{III.7}$$

Eq III.6 and Eq III.7 are not the same. Therefore, choosing the value of Δ has a significant effect on the results.

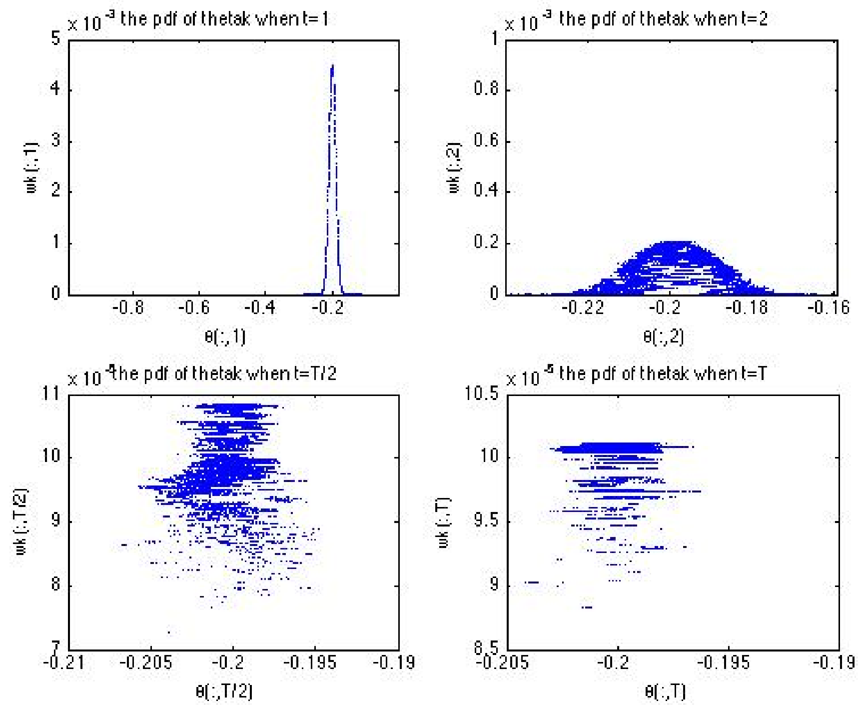

Fig. III.1.: The estimated θ by Kernel Smoothing Algorithm

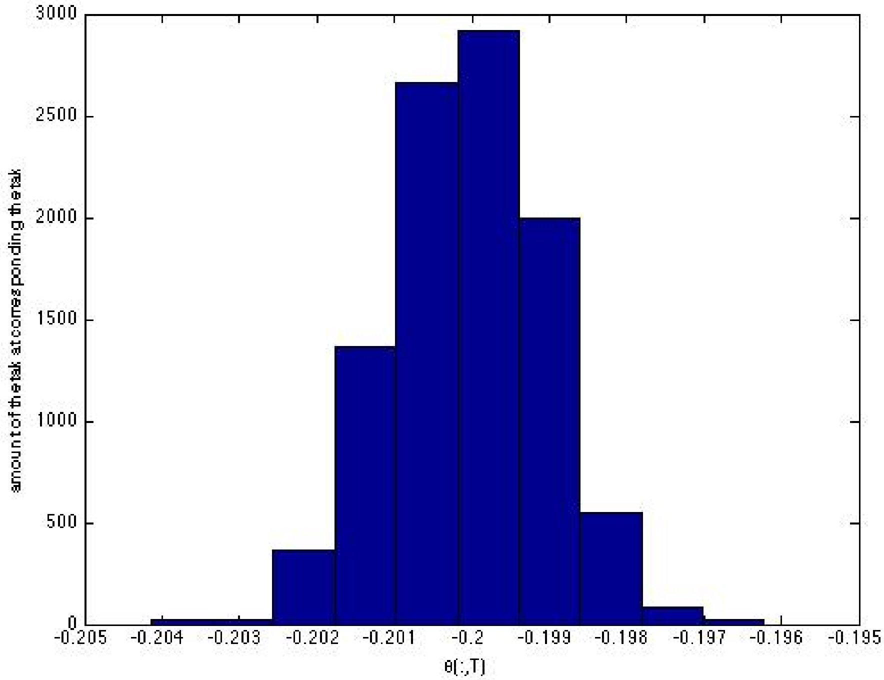

Fig. III.2.: The histogram of θ when $k = T$

Conclusion

Markov chain monte carlo techniques along with kernel smoothing can be used for parameter estimation in first-order linear ordinary differential equations. However, the performance of this algorithm is highly sensitive to the time step resulting from the discretization of the differential equation. When the time step is very small, the performance of the algorithm is very poor. A qualitative explanation for this was given in this chapter.

CHAPTER IV

PROBABILITY DENSITY FUNCTION ESTIMATION USING POLYNOMIAL-CHAOS EXPANSION

Pre-knowledge of Polynomial-Chaos Expansion

In this chapter, we will use the polynomial chaos expansion to find the pdf of random processes that satisfy stochastic ODEs. A PC expansion (PCE) is a way of representing a random variable as a function of another random variable with a given distribution, and of representing that function as a polynomial expansion[7], with the following format:

$$X(t) \approx \sum_{j=0}^{p} x_j(t)\,\psi_j(\Xi) \tag{IV.1}$$

where ψ_j is a polynomial of order j and they satisfy the orthogonality condition that for all $j \neq k$, $\langle \psi_j, \psi_j \rangle = 0$. Ξ is called the germ and it is a random variable. Usually we assume that Ξ is a scalar. In PC theory, x_j is called the mode strength and ψ_j is mode function. Note that the total number of expansion terms is $P+1$. Given f and there is a unique expansion in which the mode strengths are given by

$$x_j = \frac{\langle f, \psi_j \rangle}{\langle \psi_j, \psi_j \rangle} \tag{IV.2}$$

Problem Statement

In our problem, we consider the ordinary differential equation

$$\frac{dy(t)}{dt} = -ky, \qquad y(0) = y_0 \tag{IV.3}$$

where the decay rate coefficient k is considered to be a random variable $k(\theta)$ with certain distribution, whose probability function is $f(k)$. we compute y_j in a differential equation by polynomial chaos expansion so that we would know the pdf of $y(t)$.

By applying the polynomial chaos expansion to the solution y and random input k

$$y(t) = \sum_{i=0}^{P} y_i(t)\,\Phi_i, \qquad k = \sum_{i=0}^{P} k_i \Phi_i \tag{IV.4}$$

and substituting the expansion into the differential equation, we obtain

$$\sum_{i=0}^{P} \frac{dy_i(t)}{dt}\Phi_i = -\sum_{i=0}^{P}\sum_{j=0}^{P} \Phi_i \Phi_j k_i y_j(t) \tag{IV.5}$$

By taking $\langle ., \Phi_l \rangle$ and utilizing the orthogonality condition, we obtain the following set of equations:

$$\frac{y_l(t)}{dt} = -\frac{1}{\langle \Phi_l^2 \rangle} \sum_{i=0}^{P}\sum_{j=0}^{P} \langle \Phi_i \Phi_j, \Phi_l \rangle\, k_i y_j(t) \tag{IV.6}$$

Now, we have converted the problem of estimating the pdf of $y(t)$ in to one of the estimating the coefficients $y_l(t)$ which all satisfy a set of differential equations given in (IV.6). Note that any standard ODE solver can be employed here to solve these coefficients.

Results of Polynomial Chaos in Estimation PDF

We consider the ordinary differential equation

$$\frac{dy(t)}{dt} = -ky(t), \qquad y(0) = 1 \tag{IV.7}$$

where k is assumed to be a uniform random variable with $\Phi_1 = 1$ and $\Phi_i s = 0$ for $i \neq 1$. We choose P=4. By applying the polynomial chaos expansion to the solution y and random input k

$$y(t) = \sum_{i=0}^{4} y_i(t)\,\Phi_i, \qquad k = \Phi_1 \tag{IV.8}$$

and substituting the expansion into the differential equation, we obtain

$$\sum_{i=0}^{4} \frac{dy_i(t)}{dt} \Phi_i = -\sum_{j=0}^{4} \Phi_1 \Phi_j y_j(t) \tag{IV.9}$$

By taking $\langle ., \Phi_l \rangle$ and utilizing the orthogonality condition, we obtain the following set of equations:

$$\frac{y_l(t)}{dt} = -\frac{1}{\langle \Phi_l^2 \rangle} \sum_{j=0}^{4} \langle \Phi_1 \Phi_j, \Phi_l \rangle y_j(t), \quad l = 0, 1, 2, 3, 4 \tag{IV.10}$$

We choose the polynomials Φ_i uniform distribution. It is easy to get $y_i(t)$ by ODE solver and the solutions $y_l(t)$ for $l = 0, 1, 2, 3, 4$ are showed Fig IV.1.

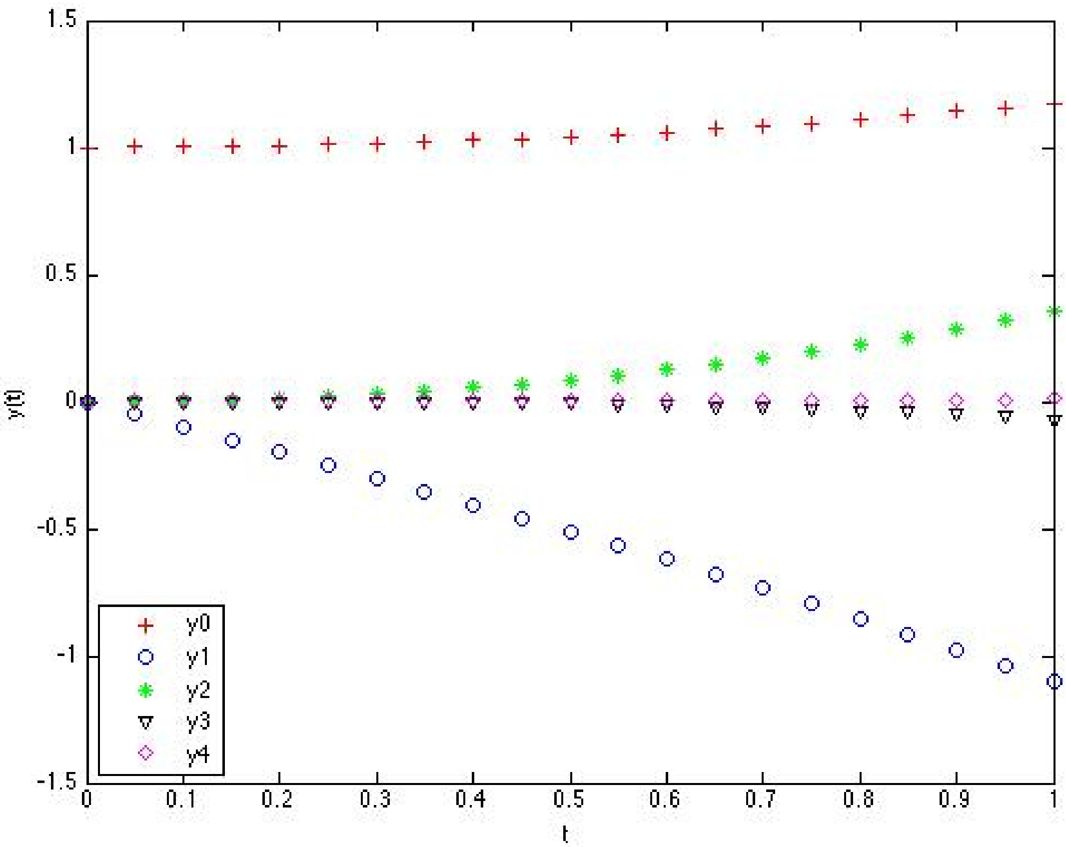

Fig. IV.1.: $y_i(t)$ in the range of $t \in (0, 1)$

Noting that the value of P in this example, we set as 4. To prove its correctness, we implement a program, which computes the error measures for the mean when P is 1,2,3 and 4.

$$\varepsilon(t) = \left| \frac{\overline{y}(t) - \overline{y}_{exact}(t)}{\overline{y}_{exact}(t)} \right| \tag{IV.11}$$

where

$$\overline{y}(t) = y_0(t), \quad \overline{y}_{exact}(t) = \frac{e^t - e^{-t}}{2t} \tag{IV.12}$$

From Fig IV.2 , we would see that at t=1, when P=4, the error of mean of $y_i(t)$ is small enough. There is no need to increase P to a larger number, which will lead to the complexity of computation and longer running time consumption.

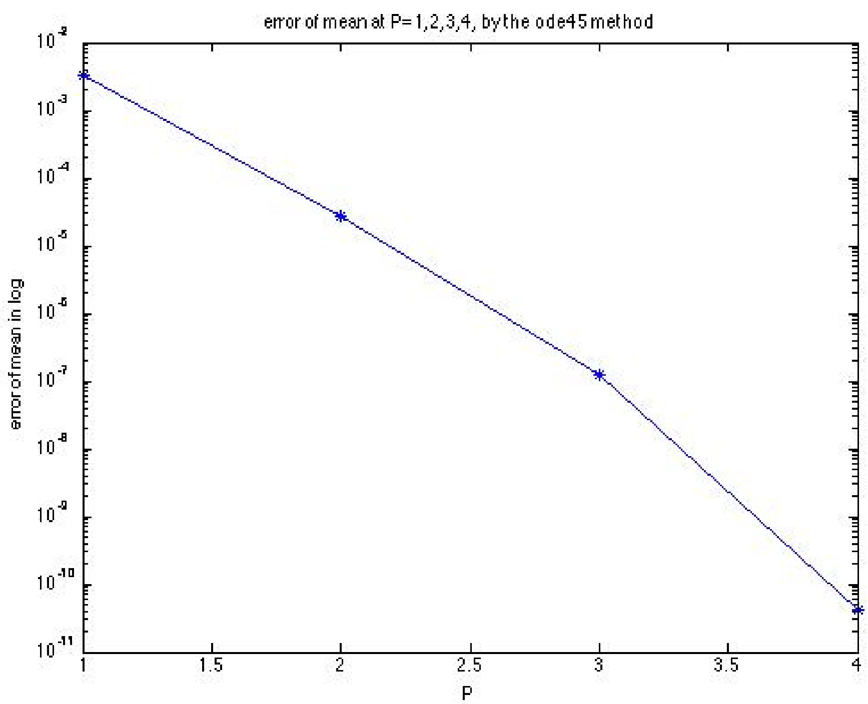

Fig. IV.2.: The error mean of of $y(1)$ when P=1,2,3,4

Conclusion

In this chapter, we considered the use of polynomial-chaos based expansion techniques for estimating the PDF of the output of a linear ODE, when the parameters in the ODE are random variables. Polynomial-Chaos expansion is a powerful tool to estimate the PDF of the solution of stochastic differential equations. In addition, by calculating the error of mean square root of $y(t)$(at a certain time), we can find that only a small number of terms need to be retained in the expansion to obtain good estimates of the PDF.

CHAPTER V

FUTURE WORKS

State and parameter estimation in bio-mechanics is a vast research topic and the research presented in this thesis represents only a first step in the estimation of states and parameters for certain problems. The following are important problems that need to be addressed in the future.

- In Chapter III, only a linear ODE is considered. The response of the artery to forces is typically given by the solution to a non-linear differential equation and hence non-linear HMM have to be considered.

- Even for the ODE considered in Chapter III, the performance of the kernel-smoothing algorithm is not very robust. Certain inconsistencies in the sequential monte carlo approach were pointed out. One easy way to fix this problem is to use a non-sequential version where an initial population is chosen for θs and fixed. However, such an approach would not be viable with θ changed with k. This model is really what is of interest since for the purpose of diagnosis, one is interested in determining changes in the elasticity in the artery as a function of length of the artery. Hence, there is a need to design robust estimation techniques that work for a variety of models of change of θ_k with k.

- Even though Chapter IV shows that the PDF of the output of the ODE can be found, our interest is in using polynomial chaos expansion based methods for parameter estimation. We still need to develop a Bayesian inference technique based on PC for determining the posterior density of the parameters of the ODE/HMM. Then, the advantages and disadvantages of MCMC and polynomial chaos methods for parameter estimation problems should be compared.

REFERENCES

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

[2] L. R. Rabiner, B. H. Juang. An Introduction to Hidden Markov Models. *IEEE ASSP* Jan 1986. Web. 3 Oct. 2013

[3] A. O'Hagan(2013). Polynomial Chaos: A Tutorial and Critique from a Statistician's Perspective. *Submitted to SIAM/ASA Journal of Uncertainty Quantification.*

[4] Moon, Todd K, and Wynn C. Stirling. *Mathematical Methods and Algorithms for Signal Processing.* Upper Saddle River, NJ: Prentice Hall, 2000. Print.

[5] M.S. Arulampalam, S. Maskell, A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, *IEEE Transactions on Signal Processing*, Vol.50, No. 2, Feb 2002.

[6] J. Liu and M. West, Combined parameter and state estimation in simulation-based filtering, in *Sequential Monte Carlo Methods* , A. Doucet, J.F.G de Freitas, and N.J. Gordon, Eds New York: Springer-Verlag, 2001

[7] D. Xiu and G.E. Karniadakis, The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations, *SIAM J. Sci. Comput.*, 24(2), 619-644 (2002)