Statistical inverse problem of partial differential equation: an example with stationary 1D heat conduction problem

Chettapong Janya-anurak

Lehrstuhl für Interaktive Echtzeitsysteme
Institute for Anthropomatics
Karlsruhe Institute of Technology (KIT), Germany
chettapong.janya-anurak@ies.uni-karlsruhe.de

Technical Report IES-20xx-xx

Abstract:

Local behaviour in a continuous system with spatially or temporally variable parameters is often described in terms of partial differential equations (PDEs). Given a system of PDEs, an inverse problem is to reconstruct parameters in every point given a limited number of observations or conditions. There exists a plethora of solution methods for various inverse problems, nevertheless, this is still an active field of research. In particular, non-linear systems, such as heat transfer equation, pose the biggest challenge. In this report we present a novel method based on Bayesian statistics. The parameter fields are represented in terms of some basis functions with unknown coefficients, treated as random variables. Their posterior probability distribution is then computed using Markov Chain Monte-Carlo approach. Finally, the field is reconstructed using the values that maximize likelihood. We illustrate the method with the example of the one-dimensional heat transfer equation, and discuss various choices of the basis functions and the accuracy issues.

1 Introduction

Many industrial processes e.g. thermal process or fluid process possess spatio-temporal interrelation. The spatio-temporal behavior of system is essential for analysis and control the processes. These systems are normally mathematically modeled in term of partial differential equations (PDEs). Due to their infinite-dimensionality the PDE models are not used directly for implementation. Practically these infinite-dimensional models are approximated into computable finite-dimensional model. Han-Xiong Li and Chenkun Qi arrange and classify various finite-dimensional approximation methods in [Li11]. Modeling with preferably few finite numbers of states that keep model computation efficient and still have good a performance is an active field of research which is known as model reduction. The overview of model reduction methods can be found in [AD01].

In fact the model cannot interpret reality exactly because there are generally some unknown uncertainties in model due to incomplete knowledge. In order to keep the model close to the reality, the measurement data are used to estimate unknown parameters. Such problems are known as parameter estimation or identification problems, which can be classified as inverse problems from mathematical point of view. A number of theories and algorithm have been developed to solve these inverse problems. Recently many methods using statistical inference to solve inverse problems are proposed especially Bayesian inference method e.g. [KS05]. The statistical inverse problem has many advantages comparing to the classic deterministic solving methods [Jin06] and therefore is an interesting active field of research.

Despite of many advantages the statistical inverse problem is normally required tremendous computational efforts. Therefore the combination of model reduction and statistical inference method is the new aspect in field of modeling research as presented in e.g. [Lie09] and [Jin06]. The residuum of the discretization or the approximation always has an effect in solving inverse problem using statistical approach [KS07]. A Key to avoid such error is to choose an appropriate approximation method, which is an aim of our study.

The organization of this paper is as follows. In section 2, the background about the partial differential equation and the spatial discretization for deterministic as well as for stochastic model is presented. Section 3 is devoted to the inverse problem and their solving methods. Section 4 shows the numerical experiments of Bayesian inverse problem of PDE by means of 1D heat conduction example. Conclusions and some future works are finally discussed in section 5.

2 Partial differential equation and spatial discretization methods

Many physical phenomena can be described with the relation between some continuously varying quantities and their rates of change in some independent variables. This relation can be mathematically formulated in form of differential equation. If there is only single independent variable in the differential equation, the equation will be called *ordinary differential equation* (ODE). On the other hand the equation involving an unknown function of several independent variables and their partial derivatives with respect to those variables is defined as *partial differential equation* (PDE). In engineering the independent variables are normally time and space coordinate. Therefore many engineering systems can be modeled mathematically in general form

$$\mathcal{D}\left(x(\boldsymbol{r},t)|\boldsymbol{\eta}\right) = s(\boldsymbol{r},t) \tag{2.1}$$

where $x(\boldsymbol{r},t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ and $s(\boldsymbol{r},t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ denotes the system state and inhomogeneous term at time t and at spatial coordinate $\boldsymbol{r} = [x,y,z]^T \in \Omega$. The inhomogeneous term, the system state and its derivatives, respect to time or space coordinate, are related by means of some operator $\mathcal{D}$. The dynamic behavior and distributed properties of the system depend on the parameter of the operator, which are collected in the process parameter vector $\boldsymbol{\eta}$.

Additional conditions are generally required in order that the PDE has unique solution. In engineering view, the additional condition can be classified in 2 types relating to the independent variable, namely the initial condition relating to time and the boundary condition relating to space. The initial condition can be formulated in general form

$$\mathcal{D}_t\left(x(\boldsymbol{r},t=0)\right) = h(\boldsymbol{r}) \tag{2.2}$$

where $\mathcal{D}_t$ denotes a differential operator in respect to time coordinate. The function $h(\boldsymbol{r})$ describe the relation of the state $x(\boldsymbol{r},t)$ for all spatial coordinate $\boldsymbol{r} \in [x,y,z]^T$ at the initial time $(t = 0)$. Similarly the boundary condition is formulated in general form

$$\mathcal{D}_r\left(x(\boldsymbol{r} \in \partial\Omega, t)\right) = b(t) \tag{2.3}$$

where $\mathcal{D}_r$ denotes a differential operator in respect to spatial coordinate. The function $b(t)$ describe the relation of the state $x(\boldsymbol{r},t)$ at boundary spatial coordinate $\boldsymbol{r} \in \partial\Omega$ for all time t.

All three equations (2.1) , (2.2) and (2.3) form an *initial-boundary value problem (IBVP)*, which are used as distributed parameter system models of physical phenomena in engineering and science.

2.1 Deterministic spatial discretization

The initial-boundary value problem is normally only in simple case analytically solvable. Therefore this IBVP is practically approximated from infinite-dimensional problem to finite dimensional problem. This approximation is usually applied in spatial domain and can be accomplished by transform the PDE and boundary conditions into system of ODEs. The overview of important spatial discretization methods can be founded in [Li11].

As showed in figure 2.1, the spatial discretization methods can be classified in 2 groups.

- Finite difference method (FDM): this method is an approximation of the differential operator in PDE by difference operator as

  $$\frac{\partial x}{\partial r} \approx \frac{\Delta x}{\Delta r}$$

  This approximation term can be forward, backward or central difference derived often from Taylor expansion.

- Weighted residual method (WRM): It is known that a continuous function can be represented as infinite series. Based on this principle, the solution of IBVP can be expanded by a set of spatial basis function $\{\varphi_i(\boldsymbol{r})\}_{i=1}^{\infty}$ as follows

  $$x(\boldsymbol{r},t) = \sum_{i=1}^{\infty} a_i(t)\varphi_i(\boldsymbol{r})$$

  From this infinite series the solution of IBVP can be approximated in finite dimension by truncating the infinite series at N

  $$x(\boldsymbol{r},t) = \sum_{i=1}^{\infty} a_i(t)\varphi_i(\boldsymbol{r}) \approx \sum_{i=1}^{N} a_i(t)\varphi_i(\boldsymbol{r}) = \hat{x}(\boldsymbol{r},t)$$

  This method is often used and very efficient. Based on types of spatial basis function either local or global, WRM can be classified in 2 groups namely the finite element method and spectral method. In both methods many approaches of proper spatial basis functions can be used to expand the solution into finite series. The selection of the basis functions is a critical point and always has impact to the performance of systems of ODEs.

It is noted that the FDM is an approximation of the derivatives operators in PDE, while the WRM approximate the solution of PDE.

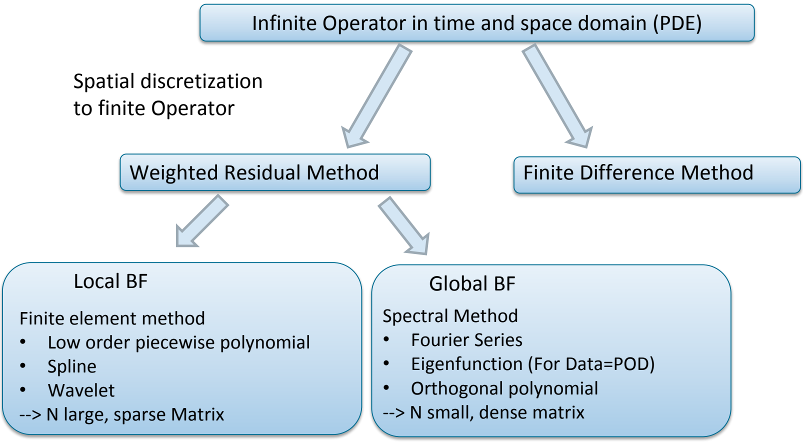


Figure 2.1: The discretization methods to approximate the PDE into system of ODEs

2.2 Discretization of random field

For some modeling aspect the uncertainty of some variables need to be considered. The state (or solution) of PDE can be modeled as random variable. In this case it will lead the PDE (2.1) to a *stochastic partial differential equation SPDE*, which represents the dynamic behavior, the distributed properties and the uncertainty of the underlying physical system.

The term stochastic process is well-known in many engineering discipline e.g in control engineering, signal processing etc. The stochastic process can be treated as functions of one deterministic argument (in most cases regarded as time) whose values are random variables. In case of several deterministic arguments (multidimensional vectors e.g. 2-dimensional space) the stochastic process can be called random field. A random field can be seen as a generalization of stochastic process. According to [SD00] the random field can be defined as:

Definition 1 (Definition of random field) Given a probability space $(\Theta, \mathcal{F}, \mathcal{P})$, a random field $H(\boldsymbol{r}, \omega)$ is defined as a curve in $\mathcal{L}^2(\Theta, \mathcal{F}, \mathcal{P})$, that is a collection of random variables indexed by a continuous parameter $\boldsymbol{r} \in \Omega$, where Ω is an open set of $\mathbb{R}^d$.

The random field can be considered as a continuous function as in deterministic case and therefore can be expanded by means of spatial basis function $\{\varphi_i(\boldsymbol{r})\}_{i=1}^{\infty}$:

$$H(\boldsymbol{r},\omega) = \sum_{i=1}^{\infty} \theta_i(\omega)\varphi_i(\boldsymbol{r})$$

This set of basis function can be either local or global as by deterministic case. Accordingly the discretized field can be expressed as a finite summation of the series:

$$\hat{H}(\boldsymbol{r},\omega) = \sum_{i=1}^{N} \theta_i(\omega)\varphi_i(\boldsymbol{r}) \tag{2.4}$$

where $\theta_i(\omega)$ denotes the random variables, which can be expressed as *weighted integrals* of $H(\cdot)$ over the domain of the system:

$$\theta_i(\omega) = \int_{\Omega} H(\boldsymbol{r},\omega)w_i(\boldsymbol{r})d\boldsymbol{r}$$

Various methods to random field discretization are proposed in [SD00]. All methods have the same approach as mentioned above, by truncating the obtained series after a finite number of terms. The difference between each method rely on different basis function and respectively weight function.

3 Inverse problem

In engineering and science mathematical models are mostly used to describe physical phenomena. The challenge of the modeling is to relate physicals parameters characterizing a **model** $\boldsymbol{\theta}$, to conceivable instrumental observations making up some set of **data** $\boldsymbol{y}$. Under assumption of physical laws the relation of $\boldsymbol{\theta}$ and $\boldsymbol{y}$ can be determined as a mathematical operator $\boldsymbol{F}(\cdot)$ as follows:

$$\boldsymbol{F}(\boldsymbol{\theta}) = \boldsymbol{y} \tag{3.1}$$

A forward problem is to find observable data $\boldsymbol{y}$ that would correspond to a given model $\boldsymbol{\theta}$. On the other hand the inverse problem is the problem of finding model $\boldsymbol{\theta}$ given observed data $\boldsymbol{y}$. The forward problem can normally be solved by mathematical manner because the problem is commonly well-posed. The term well-posed problem stems from a definition given by Jacques Hadamard [Had23]. He stated that the mathematical model of physical phenomena should have three properties, which are existence, uniqueness and continuous dependence of the solution on the data, in some reasonable topology.

By this well-posedness definition, an inverse problem is inherently ill-posed because the limited observation of data often inevitably implies non-uniqueness. In order to solve an inverse problem, some additional assumptions are involved to reformulate the problem. Two preferred approach are deterministic optimization and statistical inference, which be discussed in following subsections.

3.1 Deterministic inverse problem

In the deterministic setting, the inverse problem is formulated as an estimation problem of parameter $\boldsymbol{\theta}$. It is the minimization problem of an objective function $\mathcal{J}(\boldsymbol{\theta})$. The objective function is usually a normed difference between the observed data Y and the predicted output $\hat{Y} = F(\boldsymbol{\theta})$. Usually the quadratic of Euclidean norm is used, the mathematic formulation of inverse problem is to find optimal parameter $\boldsymbol{\theta}^*$, which

$$\boldsymbol{\theta}^* = \text{argmin}\mathcal{J}(\boldsymbol{\theta}) = \text{argmin}\left(\frac{1}{2}\left\|F(\boldsymbol{\theta}) - Y\right\|^2\right)$$

In the absence of measurement error, the optimal parameters $\boldsymbol{\theta}^*$ are the parameters, which make the objective function equal zero $\mathcal{J}(\boldsymbol{\theta}^*) = 0$. There can be not only one parameter, which produce the data Y. The solution of this problem is still not unique, therefore this problem is still ill-posed.

In order to make this problem well-posed, the term $\beta R(\boldsymbol{\theta})$ is added in the objective function. The $R(\cdot) : \mathbb{R}^{N_\theta} \to \mathbb{R}$ is a nonnegative functional, called *regularization functional*. $\beta \in \mathbb{R}_+$ is a regularization parameter. The regularized inverse problem is reformulated as:

$$\boldsymbol{\theta}^* = \text{argmin}\mathcal{J}(\boldsymbol{\theta}) = \text{argmin}\left(\frac{1}{2}\left\|F(\boldsymbol{\theta}) - Y\right\|^2 + \beta R(\boldsymbol{\theta})\right)$$

The existence and uniqueness of the solution of this problem depend on the choice of functional $R(\cdot)$. For example the Tikhonov regularization set the regularization functional as:

$$R(\boldsymbol{\theta}) = \left\|\boldsymbol{\theta} - \boldsymbol{\theta_0}\right\|^2$$

where $\boldsymbol{\theta_0}$ is some arbitrary parameter. This inverse problem with regularization is a constrained nonlinear optimization. There are no guarantees about the convexity of either the solution space or the objective function [Lie09]. The extremal problem can be solved by means of gradient method. The solution of this extremal problem is a single point in parameter space. There is no further information about the uncertainty of the solution.

3.2 Bayesian statistic formulation of inverse problem

By statistical formulation, the inverse problem is reformulated as a problem of statistical inference by means of Bayesian statistics [KS05]. The unknown parameters are modeled as random variables. The randomness of the unknown parameters describes the degree of information concerning their realizations in Bayesian statistics. The degree of information concerning these values is coded in the form of probability distributions. The solution of statistical inverse problem is the posterior distribution. This solution as a probability distribution is the main difference between the statistical approach and deterministic approach, which gives only the optimal single point solution. By statistical inverse problem the single point solution can be also estimated from the probability distribution.

The ill-posedness of the inverse problem is handled by restate the problem as a well-posed extension in a large space of probability distributions. Moreover it is allowed to add the prior knowledge that is often hidden in the deterministic regularization view. This *prior information* of unknown parameters $\boldsymbol{\theta}$ can be coded into a probability density. This probability distribution called *a priori* distribution $\pi(\boldsymbol{\theta})$ represent our *a priori* knowledge of the parameters.

With joint probability density of unknown parameters $\boldsymbol{\theta}$ and measurement y, which is denoted by $\pi(\boldsymbol{\theta}, y)$, there is a relation :

$$\pi(\boldsymbol{\theta}, y) = \pi(y|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|y)\pi(y)$$

where the conditional probability $\pi(y|\boldsymbol{\theta})$, called *likelihood function*, expresses the likelihood of different measurement outcomes by given parameters $\boldsymbol{\theta}$. The conditional probability $\pi(\boldsymbol{\theta}|y)$, called *posterior distribution* of $\boldsymbol{\theta}$, expresses the information of $\boldsymbol{\theta}$ after realized observation y.

In the Bayesian framework, the inverse problem is expressed in following way: Given the data y, find the conditional probability distribution $\pi(\boldsymbol{\theta}|y)$ of the variable $\boldsymbol{\theta}$. The statistical formulation of inverse problem can be concluded in the following theorem [KS05]:

Theorem 1 (Bayes' theorem of inverse problems) *Assume that the random variable $\Theta \in \mathbb{R}^n$ has a known prior probability density $\pi(\boldsymbol{\theta})$ and the data consist of the* *observed value y of an observable random variable $Y \in \mathbb{R}^k$ such that $\pi(y) > 0$. Then the posterior probability distribution of θ, given the data y is*

$$\pi_{post}(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{\theta})\pi(\boldsymbol{y}|\boldsymbol{\theta})}{\pi(\boldsymbol{y})} \tag{3.2}$$

The marginal probability density in the equation (3.2)

$$\pi(\boldsymbol{y}) = \int_{\mathbb{R}^n} \pi(\boldsymbol{\theta}, y) d\boldsymbol{\theta} = \int_{\mathbb{R}^n} \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

acts only as a norming constant and therefore is normally neglect.

In summary, solving the inverse problem from the Bayes's point of view can be divided into three subtasks.

1. Determine an appropriate prior probability density $\pi_{pr}(\boldsymbol{\theta})$ relied on the available prior information of unknown $\boldsymbol{\theta}$
2. Construct an appropriate likelihood function $\pi(y|\boldsymbol{\theta})$ that describes the interrelation between the observation and the unknown.
3. Compute the posterior probability density $\pi_{post}(\boldsymbol{\theta})$

3.3 Inverse problem for partial differential equation

In this work, we present the inverse problem of PDEs. It means, that from the IBVP (2.1), (2.3) and (2.2) discussed in the last section, the solution of PDEs accordingly the distribution and the dynamic of state $x(\boldsymbol{r}, t)$ is normally to be founded and therefore can be seen as the data $\boldsymbol{y}$ according to the equation (3.1). The mathematical operator in this case is also the differential operator $\mathcal{D}$. The source term $s(\boldsymbol{r}, t)$, the process parameters $\boldsymbol{\eta}$, the initial condition and the boundary condition are needed in order to find the solution of PDEs and hence can be seen as a model $\boldsymbol{\theta}$.

The computation of state $x(\boldsymbol{r}, t)$ from the IBVP (2.1), (2.3) and (2.2) is forward problem. But in case the state $x(\boldsymbol{r}, t)$ is known or observable and the others are unknown. Typical inverse problems of PDEs is normally the problem of:

- Source location : find $s(\boldsymbol{r}, t)$
- Parameter estimation : find $\boldsymbol{\eta}$

- Boundary reconstruction : find $b(t)$
- Initial reconstruction : find $h(\boldsymbol{r})$

In the statistical inverse problem the unknown variable is casted as a random variable. In case that the distribution (e.g. source distribution $s(\boldsymbol{r}, t)$) is sought, the sought distribution has to be approximated by discretization, because it is impossible to find the total distribution in infinite space. The distribution can be discretized by means of a set of basis function as showed in section 2. In this case the coefficients of the basis function is the unknown to be found and therefore are modeled as random variables.

As discussed in the last subsection there are three tasks to be worked on to solve the Bayesian inverse problem. Firstly the unknown parameter prior density are commonly modeled by Gaussian density for solve such inverse problem of PDE, which can be formulated in following form.

$$\pi_{pr}(\boldsymbol{\theta}) = exp\left(-\frac{1}{2}\beta(\boldsymbol{\theta} - \boldsymbol{\theta_0})^T S^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta_0})\right)$$

where $S \in \mathbb{R}^{n\times n}$ denotes covariance matrix of random variables $\boldsymbol{\theta}$ and $\beta \in \mathbb{R}_+$ is some arbitrary constant. Under the assumption of additive white noise in the measurement and giving the forward model $F(\boldsymbol{\theta})$ the likelihood function is:

$$\pi(y|\boldsymbol{\theta}) = exp\left(-\frac{1}{2\sigma^2}\left\|F(\boldsymbol{\theta}) - Y\right\|^2\right)$$

where σ denotes the standard deviation of the measurement noise. More information about both formulation can be found in [CO10],[Jin06] and [KS05].

Finally we have to find the method to compute the posterior distribution π_{post}. For large dimension, the integration over space $\mathbb{R}^n$ cannot be done with numerical quadrature. One effective class of technique to compute the integration in large dimensional space is known as *Markov Chain Monte Carlo (MCMC)* simulation. The MCMC method use a set of point from the given distribution, a *sample*, to approximate the Monte Carlo integration. The sample ensembles are generated by using the Markov chain random walk algorithm, for example *Metropolis-Hasting algorithm* or *Gibbs sampling*. There are also some modifications of this algorithm for better performance and computation such as *Delay rejection (DR)* or *Adaptive Metropolis-Hasting (AM)* [HLMS06]. The MCMC realizes the computation of the posterior distribution and its estimation of expected values.

4 Numerical experiment of Bayesian invese problem of PDE by means of the 1D heat conduction example

In this work, the estimation problems of heat source in one dimensional stationary heat conduction serve to illustrate the use of Bayesian statistic to solve the PDEs inverse problem. One dimensional stationary heat conduction, for example heat conduction in rod, can be described with PDE as follow.

$$\frac{\partial}{\partial x}\left(Ak\frac{\partial T(x)}{\partial x}\right) = -Q(x)$$

where $T(x)$ is the temperature distribution in considered domain. $Q(x)$ is the heat source distribution function. A is the surface area of the rod. k is the heat conductivity of the material. Normally in heat transfer literature, this problem is formulated to find the temperature distribution $T(x)$ by given boundary condition and the other fixed parameters. This formulation is the forward problem as equation (3.1). In contrast, our heat conduction inverse problem example can be formulated as follow. By giving temperature T at measure point d_i, the heat source distribution $Q(x)$ is to be estimated under the conditions that the boundary condition and parameters Ak are known (see figure 4.1).

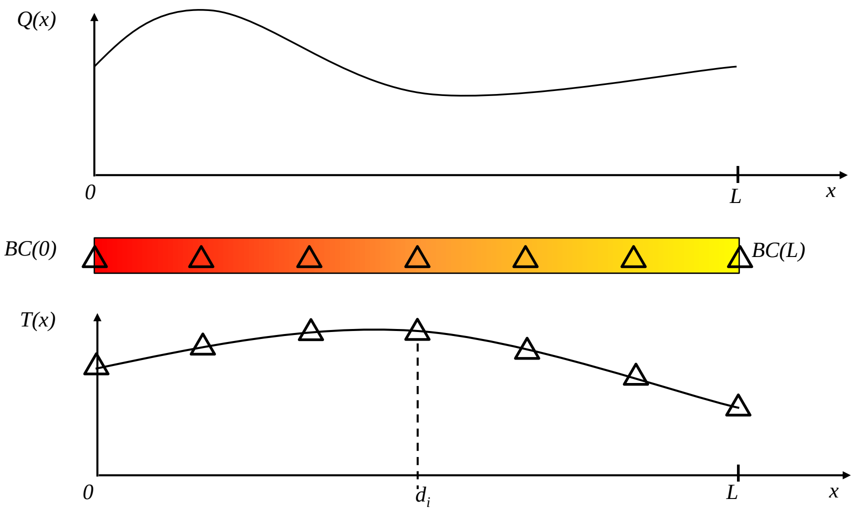


Figure 4.1: the inverse problem in 1D stationary heat conduction as an example

For simplification the parameters A and k are supposed to be constant. In our work the boundary conditions are set to be Neumann boundary condition on the left end and Dirichlet boundary condition on the right end as following equation (4.1).

$$\begin{aligned} \frac{\partial T}{\partial x}(0) &= b_1(const.) \\ T(L) &= b_2(const.) \end{aligned} \tag{4.1}$$

Because the estimation total heat source distribution $Q(x)$ in the whole infinite space directly is impossible, it has to be discretized in finite space with some spatial basis functions in order to keep the unknowns in finite number. In this work we present three type of spatial basis function which we go in detail in the following subsection.

- linear local basis function
- quadratic local basis function
- global basis function

4.1 Linear local basis function

Suppose we have the temperature measurement data $T(d_i)$ at sensor position $d_i = \{0, 0.1, ..., 5.9, 6\}$ as shown in the figure 4.2. The ground truth of heat source distribution function $Q(x)$, which is afterwards estimated , is piecewise linear. In this example it is discretized in 4 elements with 5 nodes by linear local basis function. The approximation of heat source distribution for example in first element is:

$$\hat{H}_I(\xi) = \theta_1 \left(1 - \frac{\xi}{l_I}\right) + \theta_2 \frac{\xi}{l_I}$$

where ξ is the local coordinate in the element and l_i the length of the i-th element. Transform the local coordinate ξ to the global coordinate x, the heat source distribution for total domain is:

$$\hat{Q}(x) = \sum_{i=I}^{IV} \hat{H}_i(x) = \hat{H}_I(x) + \hat{H}_{II}(x) + \hat{H}_{III}(x) + \hat{H}_{IV}(x)$$

Instead of estimation source distribution of total domain, we need to estimate only 5 variables $\theta_1, ..., \theta_5$. By using the Bayesian statistical method as shown in section

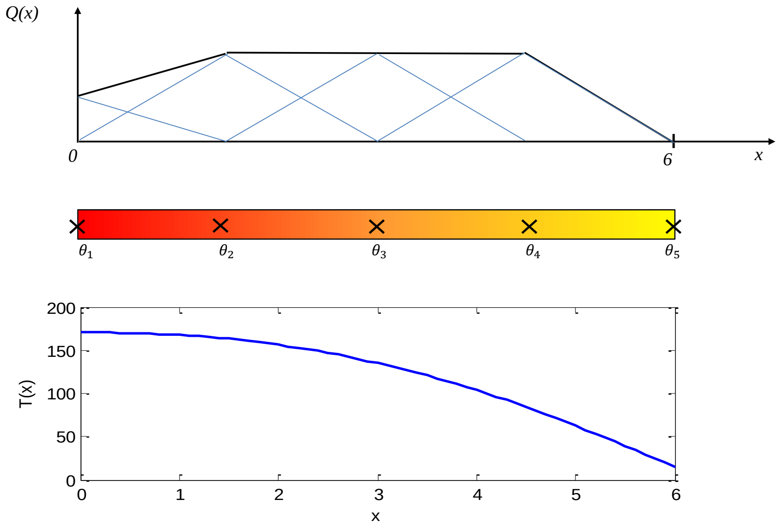
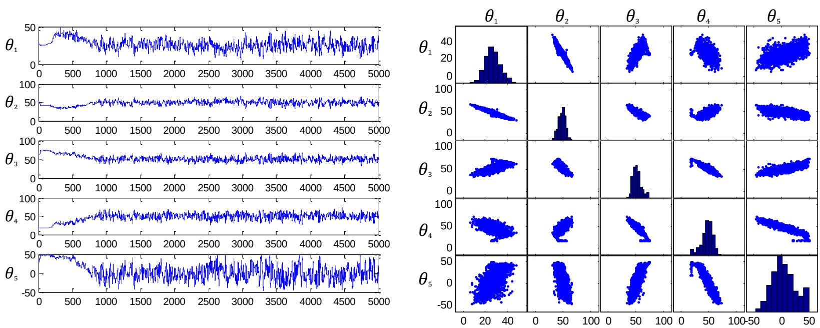


Figure 4.2: shows the example in 4.1. The ground truth of heat source distribution function $Q(x)$, the position of the variables θ and the temperature data



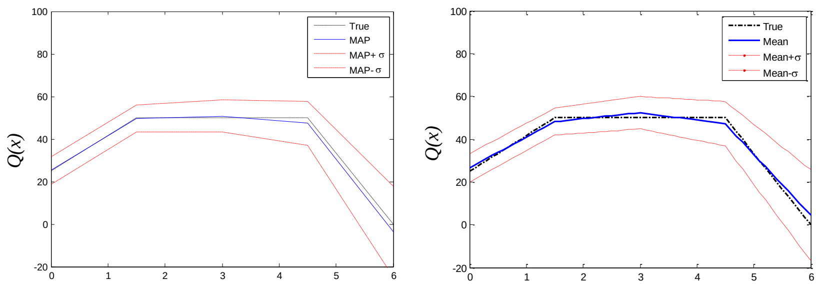
(a) chain of the parameter

(b) posterior distribution and scatter plot

Figure 4.3: The result distribution of unknown parameter $[\theta_1, ..., \theta_5]$ from MCMC simulation in the example 4.1

3.2 these 5 unknowns are casted as a random variable. The MCMC simulation results the chain of parameters as shown in figure 4.3a. The posterior distribution and the scattering plot, which present the probability and the dependency of each parameters, are shown in figure 4.3b.

The estimated heat source distribution $\hat{Q}(x)$ in total domain can be reconstructed by using the expected values from the posterior distribution as a coefficient of basis function by using equation (2.4). The expected value of the parameter can be estimated with *Maximum a Posterior (MAP)* or *conditional mean (CM)* from the posterior distribution. The figure 4.4 shows the plot of the reconstructed heat source distribution, which are satisfyingly accurate. Only at the right boundary one can see that there is a large deviation. The scattering plot also shows that the variable θ_5 is quite independent from the other variables. Because of the Dirichlet boundary condition at the right boundary, the variable θ_5 has a little effect to the solution of the forward model. This shows how the Bayesian statistic handle the ill-posed problem.



(a) reconstructed source distribution with MAP (b) reconstructed source distribution with CM

Figure 4.4: the reconstructed source distribution using linear local basis function in the example 4.1

4.2 Quadratic local basis function

In some case the linear local basis function cannot reproduce original function accurately. Compare to the quadratic local basis function the linear local basis function need more nodes to keep good accuracy. The figure 4.5 show that at the same number of variables the quadratic local basis function can approximate the original function better than the local basis function. The quadratic local basis function need 3 coefficient in 1 element as shown in equation (4.2).

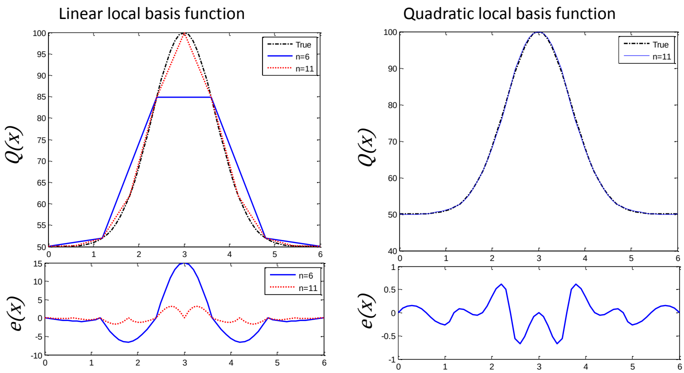


Figure 4.5: comparing the reconstruction of $Q(x)$ using linear basis function and quadratic basis function

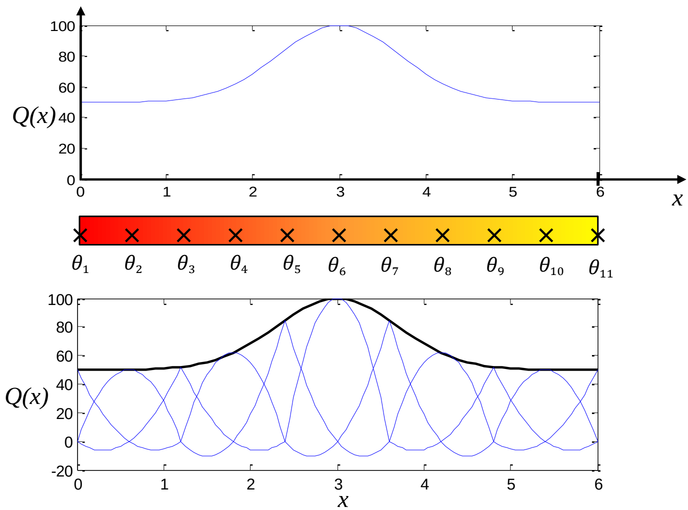


Figure 4.6: shows the example 4.2. The ground truth of heat source distribution function $Q(x)$, the position of the variables θ and the temperature data

$$\hat{H}_I(\xi) = \theta_1 \left(1 - \frac{\xi}{l}\right)\left(1 - 2\frac{\xi}{l}\right) + \theta_2 4\frac{\xi}{l}\left(1 - \frac{\xi}{l}\right) + \theta_3 \frac{\xi}{l}\left(2\frac{x}{l} - 1\right) \quad (4.2)$$

The approximation of heat source distribution is the summation of all local element after transformation the local coordinate to the global coordinate system.

$$\hat{Q}(x) = \sum_{i=I}^{N} \hat{H}_i(x)$$

The figure 4.6 shows the ground truth of the heat source distribution and its discretization with the quadratic local basis function. In this example the ground truth of heat source distribution is determined as follow

$$Q(x) = 50(1 + e^{(x-3)^2})$$

The MCMC simulation results the posterior distributions of all unknown coefficients (11 variables in this example). The reconstructed heat source distribution with the MAP- and CM-estimators are shown in figure 4.7

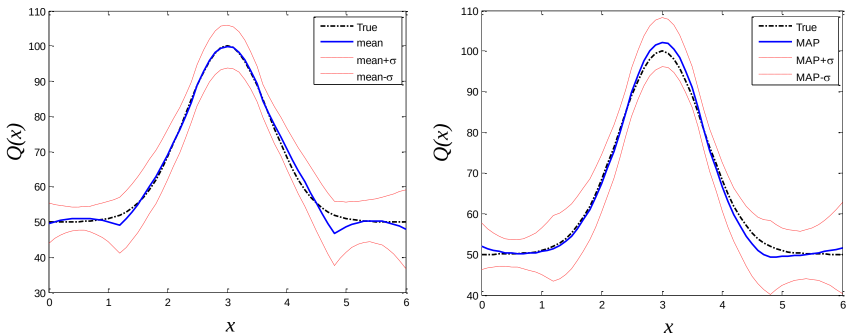


Figure 4.7: the reconstructed source distribution using quadratic local basis function in the example 4.2

4.3 Global basis function

The quadratic local basis function can give a satisfactory result but it need a lot of coefficient to reconstruct the original function. Sometimes it can lead to the curse of dimensionality. As we discussed in the section 2.1, the global basis function can be also used to reconstruct the original distribution. There are a plenty of

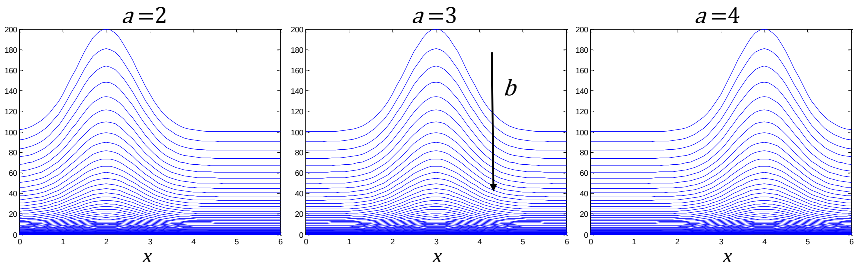


Figure 4.8: set of the assumed possible heat source distribution functions

type of global basis functions as aforementioned discussion. In this work we use the *Proper Orthogonal Decomposition* or short POD-Basis function. This POD-method is also known *Principal component analysis (PCA)*, *Kahunen-Loeve transform (KLT)* or *Hotelling transform*. This technique uses some given empirical data to extract a basis function, which hold the important feature from the given data set. More information about POD-basis function can be found in [Ast04] and [New96]. Suppose we possess a set of possible heat source distribution function data as shown in the figure 4.8. These data are described with the equation

$$Q(x) = 100e^{-b}(1 + e^{(x-a)^2}), \text{for } a = \{2, 3, 4\}, b = \{0, 0.1, ..., 6\} \tag{4.3}$$

which we presumably do not know. These data represent known learn input data.

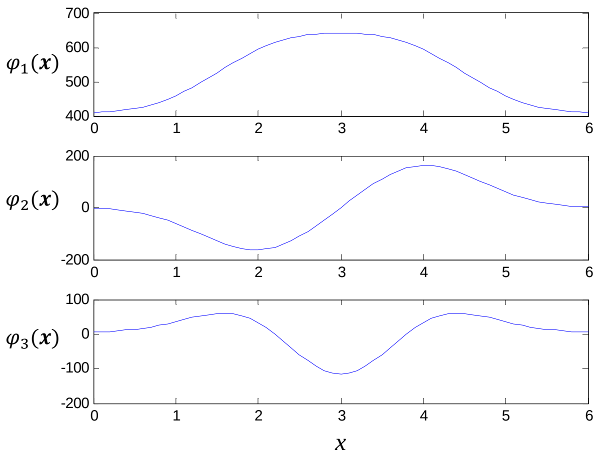


Figure 4.9: the first three POD-basis function

From these data we can extract the global basis functions by using POD-method. According to the equation (2.4) this possible source distribution can be approximated by using the first three important basis function as shown in figure 4.9.

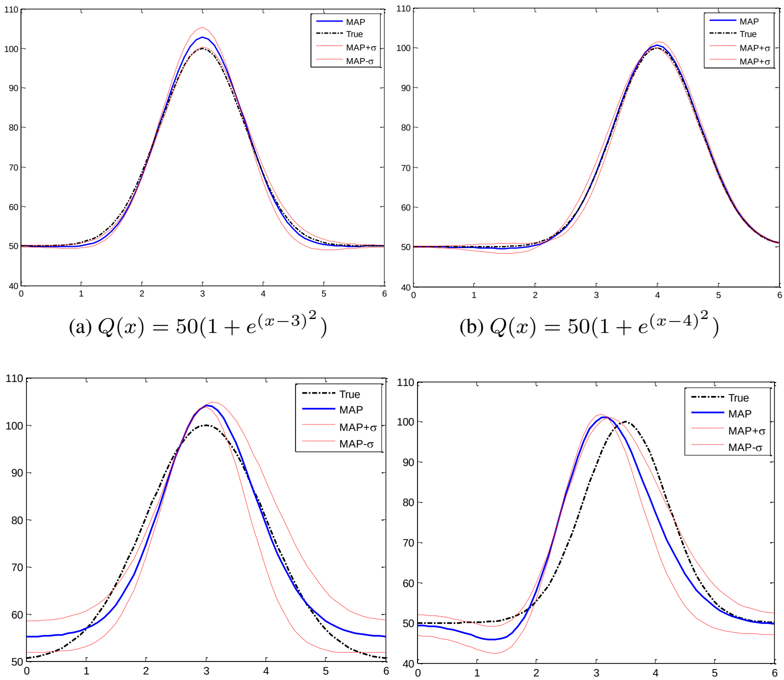
With this three global basis function, we need to estimate only three unknown parameters. In this example we test our approach with 4 heat source distribution as following equations .

$$Q(x) = 50(1 + e^{(x-3)^2}) \tag{4.4a}$$

$$Q(x) = 50(1 + e^{(x-4)^2}) \tag{4.4b}$$

$$Q(x) = 50(1 + e^{-\frac{1}{2}(x-3)^2}) \tag{4.4c}$$

$$Q(x) = 50(1 + e^{(x-3.5)^2}) \tag{4.4d}$$



(a) $Q(x) = 50(1 + e^{(x-3)^2})$ (b) $Q(x) = 50(1 + e^{(x-4)^2})$

(c) $Q(x) = 50(1 + e^{-\frac{1}{2}(x-3)^2})$ (d) $Q(x) = 50(1 + e^{(x-3.5)^2})$

Figure 4.10: the reconstructed source distribution using POD basis function in the example 4.3

The figure 4.10 shows the reconstruction of heat source distribution functions with the POD-basis function. It can be seen that the estimation is very accurate for cases a) and b). There are also no large deviation on the right boundary due to the property of the global basis function. On the other hand these three basis functions cannot reconstruct the heat source functions of case c) and d) satisfyingly. The reason is that both heat source functions in case c) and d) do not belong to defined possible heat source distributions (equation (4.3)).

5 Conclusion and Future work

In this report, the solving PDEs inverse problem using Bayesian statistical inference method is presented. The unsolvable infinite dimension problem is approximated into finite dimension problem with different type of spatial basis functions. Using MCMC simulation the posterior distribution the unknown variables are computed from the prior informations and the likelihood functions. The Bayesian inference method is illustrated by heat source estimation problems in one dimensional stationary heat conduction as study examples. From the examples, it is found that choosing spatial basis function has a considerable influence in the solution. Using the local basis function is simple, but it needs a high number of unknown variables. The global spatial basis function such as POD-basis function can reconstruct the heat source distribution function with only a low number of variables, but it also need a set of learn data to generate the basis function. Each type of spatial basis function has different advantage and disadvantage in different situation. The exact effect of spatial basis function in this method will be further researched in the future work.

Bibliography

[AD01] A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview, 2001.

[Ast04] Patricia Astrid. *Reduction of process simulation models: a proper orthogonal decomposition approach.* Technische Universiteit Eindhoven, Eindhoven, 2004.

[CO10] Karen Willcox Chad Lieberman and Omar Ghattas. Parameter and state model reduction for large-scale statistical inverse problems. *SIAM Journal on Scientific Computing*, 32(5):2523–2542, 2010.

[Had23] Jacques Hadamard. *Lectures on Cauchy's problem in linear partial differential equations, by Jacques Hadamard,..* Yale University press, New Haven, 1923.

[HLMS06] Heikki Haario, Marko Laine, Antonietta Mira, and Eero Saksman. Dram: Efficient adaptive mcmc. *Statistics and Computing*, 16:339–354, 2006.

[Jin06] Jingbo Wang Ph. D. *Bayesian computational techniques for inverse problems in transport processes*. 2006.

[KS05] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*. Springer, New York, 2005.

[KS07] Jari Kaipio and Erkki Somersalo. Statistical inverse problems: Discretization, model reduction and inverse crimes: Applied computational inverse problems. *Journal of Computational and Applied Mathematics*, 198(2):493–504, 2007.

[Li11] Han-Xiong Li. *Spatio-tenporal modeling of nonlinear distributed parameter systems: A time/space separation based approach*. Springer, [S.l.], 2011.

[Lie09] Chad Eric Lieberman. *Parameter and state model reduction for Bayesian statistical inverse problems*. 2009.

[New96] Andrew J. Newman. Model reduction via the karhunen-loeve expansion part i: An exposition. Technical report, Tech. Rep. T.R. 96-32, Inst. Systems Research, 1996.

[SD00] Bruno Sudret and Armen DerKiureghian. *Stochastic finite element methods and reliability: A state-of-the-art report*, volume 2000,08. Dept. of Civil Engineering Univ. of California, Berkeley, 2000.