SPATIAL STATISTICS MODELS FOR STOCHASTIC INVERSE PROBLEMS IN HEAT CONDUCTION

J. Wang and N. Zabaras
Materials Process Design and Control Laboratory
Sibley School of Mechanical and Aerospace Engineering
188 Frank H.T. Rhodes Hall
Cornell University, Ithaca, NY 14853-3801
jw332@cornell.edu, zabaras@cornell.edu

Abstract

A Bayesian statistical inference approach is presented herein for the solution of stochastic inverse problems in heat conduction. Spatial statistics models, in particular Markov random fields (MRF), are used to model the prior distributions of unknown thermal quantities (boundary heat flux or heat source). The posterior distribution of the unknown is derived from Bayes' formula and explored using Markov chain Monte Carlo (MCMC) algorithms and in particular the Gibbs sampler. Hierarchical Bayesian formulation is used to automatically select the regularization parameter and to estimate the statistics of measurement noise simultaneously with the unknown quantities.

Introduction

Inverse problems in heat conduction, including boundary heat flux reconstruction and heat source reconstruction, have been studied over several decades due to the significance in a variety of scientific and engineering applications (Alifanov, 1994; Beck et al., 1985). However, the majority of these studies were conducted in a deterministic manner without considering the statistical characteristics of uncertainties (Wang and Zabaras, 2004).

A Bayesian inference approach (Besag et al., 1995; Congdon, 2001) using spatial statistics priors (Besag et al., 1993), in particular, Markov random fields (MRF), is presented in this study for the solution of stochastic inverse problems in heat conduction. The Bayesian approach provides not just point estimate but also the probability density function (pdf) of the unknown (Congdon, 2001). In addition, it can provide flexible regularization to the ill-posed inverse problems through spatial prior models (Wang and Zabaras, 2004). The non-trivial problem of selecting the regularization parameter in other inverse techniques is resolved since the Bayesian method treats the inverse problem as a well-posed problem in an expanded stochastic space. Even when seeking a point estimate, the regularization parameter can be automatically selected through a hierarchical Bayesian formulation. Finally, the associated rich family of sampling strategies (Andrieu et al., 2003; Geman et al., 1984) enable the numerical exploration of high dimensional and implicit distributions.

The organization of this paper is as follows. At first, the formulation of a posterior probability density function (PPDF) for inverse problems in heat conduction is presented. A hierarchical Bayesian model is introduced to relax the prior assumptions. The prior distributions of temporal-spatially varying quantities (heat flux and heat source) are modelled using Markov random fields (MRF) (Besag et al., 1993; Wang and Zabaras). We proceed with the design of the MCMC sampler to explore the posterior state space in the current Bayesian formulations. Finally, numerical examples are presented to demonstrate the methodologies.

Bayesian Formulation of Inverse Problems in Heat Conduction

In general, inverse problems in heat conduction can be defined through the following equations,

$$\rho C_p \frac{\partial T}{\partial t} = \bigtriangledown \cdot (k \bigtriangledown T) + f(\mathbf{x}, t), \quad \text{in} \quad \Omega, \quad t \in [0, t_{\max}] \tag{1}$$

$$k\frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_h, \quad \text{on} \quad \Gamma_h, \quad k\frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_0, \quad \text{on} \quad \Gamma_0, \qquad t \in [0, t_{\max}] \tag{2}$$

$$T(\mathbf{x}, t) = T_g, \quad \text{on} \quad \Gamma_g, \quad t \in [0, t_{\max}], \tag{3}$$

$$T(\mathbf{x}, 0) = T_0(\mathbf{x}), \quad \text{in} \quad \Omega, \tag{4}$$

where ρ, C_p, k denote the density, heat capacity and thermal conductivity, respectively and f is the heat source. T_g, T_0 and q_h are the known temperature condition on boundary Γ_g, known initial temperature condition, and known heat flux condition on boundary Γ_h, respectively. In the above equations, the missing information can be either the heat source function $f(\mathbf{x}, t)$ or the unknown boundary heat flux q_0. The missing information will herein be deduced from temperature measurements at thermocouple locations within the domain. For simplicity of the discussion, in the following presentation of methodology, the case where q_0 is unknown is considered. All discussions are applicable to the heat source reconstruction as well.

To introduce the Bayesian approach, $q_0(\mathbf{x}, t)$ is first discretized as follows:

$$q_0(\mathbf{x}, t) = \sum_{i=1}^{m} \theta_i w_i(\mathbf{x}, t), \tag{5}$$

where w_i's are prescribed linear finite element basis functions, m is the number of basis functions, and θ_i's (or in a vector form θ) are the unknown variables to be estimated. In fact, each component θ_i is the heat flux at a nodal point on the discretized spatial-to-temporal lattice.

Different from other data-driven inverse techniques that aim at computing a point estimate, the primary objective of Bayesian estimation is to deduce the posterior distribution of θ, according to the Bayes' formula:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \propto p(Y|\theta)p(\theta), \tag{6}$$

where $p(\theta|Y)$ is the posterior distribution (PPDF), $p(Y|\theta)$ is the likelihood function and $p(\theta)$ is the prior distribution. Once the PPDF is known, various point estimators can be defined such as the 'Maximum A Posteriori' (MAP) estimator $\hat{\theta}_{\mathrm{MAP}} = augmax_\theta \ \ p(\theta|Y)$, and the posterior mean estimator $\hat{\theta}_{\mathrm{postmean}} = E \ \ \theta|Y$.

For the above introduced inverse problem, the functional relation between the unknown heat flux, measurement noise and data can be written as follows:

$$Y = F(\theta) + \omega \tag{7}$$

where F is a numerical solver of Eqs. (1) - (4) that computes the temperatures at thermocouple locations for a given value of θ, and the vector ω contains the measurement errors, which are assumed as independent identically-distributed (i.i.d) Gauss random noise with mean 0 and variance v_T. Subsequently, the likelihood can be written as,

$$p(Y|\theta) = \frac{1}{(2\pi)^{n/2} v_T^{n/2}} \exp\{-\frac{(Y - F(\theta))^T (Y - F(\theta))}{2v_T}\}, \tag{8}$$

where n is the number of measurements. The MRF used for prior distribution of θ is of the following form:

$$p(\theta) \propto \lambda^{m/2} \exp(-\frac{1}{2}\lambda\theta^T W\theta). \tag{9}$$

The entries of the $m \times m$ matrix W are determined as, $W_{ij} = n_i$ if i=j, $W_{ij} = -1$ if nodes (on the finite element lattice) i and j are adjacent, and as 0 otherwise. n_i is the number of nodes adjacent to node i and λ is a scaling parameter. Due to the inherent difference between length scales in time and space, in multi-dimensional problems we use a two-scale MRF prior model by the following procedure: (i) formulating a MRF matrix W_s by defining only spatially adjacent nodes as adjacent nodes; (ii) formulating a MRF matrix W_t by defining only temporally adjacent nodes as adjacent nodes; (iii) W is the summation of W_s and γW_t, where γ is the ratio of non-dimensional time step length to space step length in the discretization of heat flux.

With the above likelihood and prior distribution, the PPDF (assuming fixed λ and v_T) can be evaluated as,

$$p(\theta|Y) \propto \exp\{-\frac{(Y - F(\theta))^T (Y - F(\theta))}{2v_T}\} \cdot \exp\{-\frac{1}{2}\lambda\theta^T W\theta\}. \tag{10}$$

A more reasonable approach is to relax the prior assumptions on λ and v_T by treating these hyper-parameters as random variables. A hierarchical Bayesian PPDF is hence used,

$$p(\theta, \lambda, v_T|Y) \propto v_T^{-n/2} \exp\{-\frac{(Y - F(\theta))^T (Y - F(\theta,))}{2v_T}\} \cdot \lambda^{m/2} \exp\{-\frac{1}{2}\lambda\theta^T W\theta\}$$

$$\cdot \lambda^{\alpha_0 - 1} \exp\{-\beta_0\lambda\} \cdot v_T^{-(1+\alpha_1)} \exp\{-\beta_1 v_T^{-1}\}, \tag{11}$$

where Gamma and inv-Gamma distributions are assigned to λ and v_T as priors, respectively.

Numerical Exploration

To explore the PPDF introduced above, Markov chain Monte Carlo (MCMC) simulation method is used. The idea of Monte Carlo simulation is to draw an i.i.d set of samples $\{\theta^{(i)}\}_{i=1}^{L}$ from a target distribution $p(\theta)$ (PPDF in current study) defined on a high dimensional space $\mathcal{R}^m$. Then, the expectation of any function of θ can be estimated using the sample mean. MCMC method is developed as a strategy to generate samples $\theta^{(i)}$ while exploring the state space of θ using a Markov chain mechanism.

The Metropolis-Hastings (MH) algorithm is the most basic form of all MCMC algorithms. For a target distribution $p(\theta)$, the MH algorithm samples a candidate $\theta^{(*)}$ from a proposal distribution $q(\theta^{(*)}|\theta^{(i)})$ at each iteration, where $\theta^{(i)}$ is the sample at the previous iteration. The chain moves to the next state with an acceptance probability A defined as $A(\theta^{(*)}, \theta^{(i)}) = min\{1, \frac{p(\theta^{(*)})q(\theta^{(i)}|\theta^{(*)})}{p(\theta^{(i)})q(\theta^{(*)}|\theta^{(i)})}\}$. Gibbs sampler is a type of MCMC algorithm that updates one component of θ at each iteration. The sampler designed for the current PPDF is a combination of the basic MH and Gibbs samplers, and it is summarized below:

1. Initialize $\theta^{(0)}$, $\lambda^{(0)}$ and $v_T^{(0)}$
2. For $i = 0 : \text{Nmcmc} - 1$
   — sample $\theta_1^{(i+1)} \sim p(\theta_1|\theta_2^{(i)}, ..., \theta_m^{(i)}, \lambda^{(i)}, v_T^{(i)})$
   — sample $\theta_2^{(i+1)} \sim p(\theta_2|\theta_1^{(i+1)}, \theta_3^{(i)}, ..., \theta_m^{(i)}, \lambda^{(i)}, v_T^{(i)})$
   — $\vdots$
   — sample $\theta_m^{(i+1)} \sim p(\theta_m|\theta_1^{(i+1)}, ..., \theta_{m-1}^{(i+1)}, \lambda^{(i)}, v_T^{(i)})$
   — sample $u \sim U(0,1)$
   — sample $\lambda^{(*)} \sim q_\lambda(\lambda^{(*)}|\lambda^{(i)})$
   — if $u < A(\lambda^{(*)}, \lambda^{(i)})$
   $\lambda^{(i+1)} = \lambda^{(*)}$
   — else
   $\lambda^{(i+1)} = \lambda^{(i)}$,
   — sample $u \sim U(0,1)$
   — sample $v_T^{(*)} \sim q_v(v_T^{(*)}|v_T^{(i)})$
   — if $u < A(v_T^{(*)}, v_T^{(i)})$
   $v_T^{(i+1)} = v_T^{(*)}$
   — else
   $v_T^{(i+1)} = v_T^{(i)}$,

where q_λ and q_v are proposal distributions for λ and v_T, respectively, which are taken as normal distributions with means $\lambda^{(i)}$ and $v_T^{(i)}$ and variances $5\%\lambda^{(i)}$ and $5\%v_T^{(i)}$, respectively.

Numerical Experiments

Example I

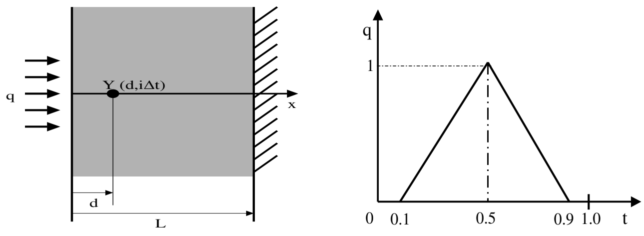

Figure 1. The fig. on the left is the schematic of the 1D heat conduction. The fig. on the right is the profile of the true boundary heat flux used to generate simulated sensor data.

The first example studied is the reconstruction of the boundary heat flux in 1D heat conduction. Let us consider the experiment in Fig. 1: the body has zero initial temperature and is insulated at the right end (x=L). An unknown heat flux $q(t)$ is applied at the left end ($x = 0$). Temperatures are recorded at $x = d$. To simplify the discussion, the numerical study is conducted in a dimensionless manner as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}, \quad 0 < t < 1, \;\; 0 < x < 1 \tag{12}$$

$$T(x,0) = 0, \;\; 0 \leq x \leq 1; \quad \frac{\partial T}{\partial x}|_{x=1} = 0, \;\; \frac{\partial T}{\partial x}|_{x=0} = q(t), \;\; 0 < t < 1. \tag{13}$$

A triangular test heat flux profile is used to generate the simulation data by adding Gauss random noise with mean 0 and variance v_T (standard deviation σ_T) to the computed temperatures at d.

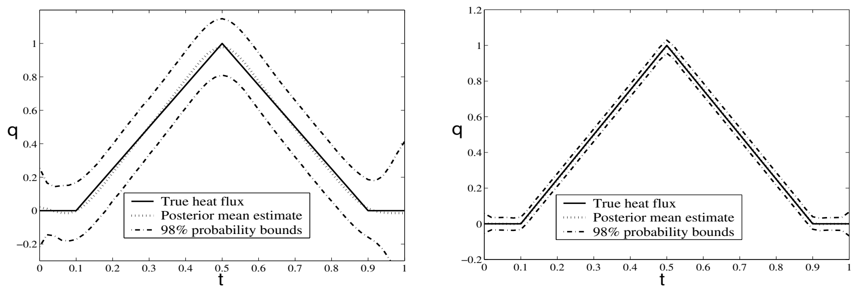

Figure 2. Posterior mean estimates of heat flux and the 98% probability bounds of the posterior distributions. The figure on the left is obtained when $\sigma_T = 0.01$ and $d = 0.5$. The figure on the right is obtained when $\sigma_T = 0.001$ and $d = 0.1$.

The posterior mean estimates and the 98% probability bounds of the posterior distributions using different v_T and d are plotted in Fig. 2. In both cases, 100 measurements are taken at the sensor location (sampling time interval $\Delta t = 0.01$). In the discretization of $q(t)$, 51 basis functions are used for each case. The effective regularization parameter λ is treated as random variable and updated in each MCMC step. For all cases, 50000 MCMC samples are generated and the results are obtained from the last 25000 samples.

Example II

In this example, we consider a heat source reconstruction problem as follows,

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + f(x,y,t), \quad 0 < t < 1, \;\; 0 < x, y < 1 \tag{14}$$

$$T(x,y,0) = 0, \quad 0 \leq x, y \leq 1 \tag{15}$$

$$\frac{\partial T}{\partial x}|_{x=0} = \frac{\partial T}{\partial y}|_{y=0} = \frac{\partial T}{\partial x}|_{x=1} = \frac{\partial T}{\partial y}|_{y=1} = 0 \quad t > 0. \tag{16}$$

In the above equations, the heat source $f(x,y,t)$ is unknown. The problem is to reconstruct this temporal-spatially varying function from temperature measurements at a number of locations within the domain and boundary.

As mentioned previously, all formulations in addressing boundary heat flux reconstruction can be applied to heat source reconstruction as well. This example is solved using the formulation of Eq. (11) by replacing the heat flux with the heat source. The two-scale MRF model is used in prior modelling of the heat source.

We consider 25 uniformly distributed thermocouples inside the domain and on the boundary. At each sensor location, 20 measurements are taken at equal time interval from $t = 0$ to $t = 0.1$. The heat source is reconstructed using a discretization of 16×16 grid in space and 11 basis functions in time. The testing heat source used to generate the data has a profile of the following form:

$$f(x,y,t) = \exp(-t)\frac{50}{2\pi \cdot 0.25^2}\exp\{-\frac{(x-0.5)^2+(y-0.5)^2}{2 \cdot 0.25^2}\} \tag{17}$$

The simulated random errors have standard deviation of 0.01. The reconstructed heat source profiles at different times are plotted in Fig. 3. It is seen that the posterior mean estimates are rather accurate overall although the deviations at initial time and final time are slightly larger.

Conclusions

A Bayesian inference approach using hierarchical Bayes formulation and spatial statistical priors is presented for the solution of stochastic inverse problems in heat conduction. It is demonstrated through numerical examples that Bayesian computation provides the means to quantify uncertainties and to deduce accurate probabilistic description of inverse solutions. The presented approach is applicable to various stochastic optimization problems in continuum systems in the context of estimation, design, control and design of experiments.

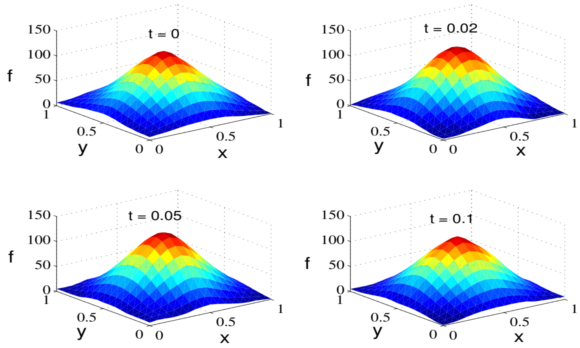

Figure 3. Posterior mean estimates of the heat source profiles.

Acknowledgements

This work was partially supported by NASA, Office of Biological and Physical Sciences Research (98-HEDS-05), the Air Force Office of Scientific Research (grant FA9550-04-1-0070) and by the National Science Foundation (grant DMI-0113295). This research was conducted using the resources of the Cornell Theory Center.

References

Alifanov, O. (1994), *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin.

Andrieu, C., N. Freitas, A. Doucet, and M. Gordan (2003), "An introduction to MCMC for machine learning," *Machine Learning*, **50**, 5-43.

Beck, J., B. Blackwell, and C. St-Clair Jr. (1985), *Inverse Heat Conduction: Ill-posed Problems*, Wiley-Inter Science, New York.

Besag, J. and P.J. Green (1993), "Spatial Statistics and Bayesian Computation," *Journal of the Royal Statistical Society, Series B, Methodological*, **55**, 25-37.

Besag, J., P. Green, D. Higdon, and K. Mengersen (1995), "Bayesian Computation and Stochastic Systems," *Statistical Science*, **10**, 3-41.

Congdon, P. (2001), *Bayesian Statistical Modelling*, John Wiley Sons, New York.

Geman, S. and D. Geman (1984), "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721 - 741.

Wang, J. and N. Zabaras (2004), "A Bayesian Inference Approach to the Stochastic Inverse Heat Conduction Problem," *International Journal of Heat and Mass Transfer*, in press.

Wang, J. and N. Zabaras, "Hierarchical Bayesian Models for Inverse Problems in Heat Conduction," *Inverse Problems*, submitted for publication.