Bayesian inference for Gaussian models: Inverse problems and evolution equations
Yan, D.

Author: Yan, D.
Title: Bayesian inference for Gaussian models: Inverse problems and evolution equations
Issue Date: 2020-03-03

Summary

This dissertation studies the Asymptotics of Bayesian nonparametric inference for Gaussian linear models. The models are in the form of

$$\text{Observations} = \underset{\text{e.g. a smoothed parameter}}{\text{Transformed signal}} + \text{Gaussian noise},$$

and the goal is to recover the original signal using Bayesian methods.

In Part I we collect the materials that are essential for the development of this thesis. In particular, we provide theory for the transformed-signal-in-noise model. The goal is to estimate the (original) signal with as few assumptions on the parameter space as possible, and to recover the signal with fast rates. The parameter (i.e., signal) is assumed to possess a certain level of smoothness, which will be quantified in this part of the thesis. The noise structure of the statistical model is also of importance, because it largely determines the difficulty of the recovery problem. To recover the signal we will employ Bayesian methods, which are formally defined in Chapter 4.

In Chapter 2 smoothness classes are introduced to quantify the measure of regularity of the parameter space. As first example we consider smoothness scales. A smoothness scale $\{H_s\}_{s\in\mathbb{R}}$ is a collectionof Hilbert spaces H_s that are nested by their norms, and this collection has an additional property which is referred to as a norm duality. A particular interesting subclass of smoothness scales are Hilbert scales. A Hilbert scale is a smoothness scale where the index s naturally defines a generating operator. The resulting generating operator establishes a link between the Hilbert scale and the covariance operator of Gaussian priors and the Gaussian noise. Another type of a smoothness class is used to describe the anisotropic smoothness in higher dimensions, which are used to study stochastic evolution equations, see Part III. We also define approximation numbers, which describe the approximation properties of the smoothness classes, and connect them to metric entropies.

In Chapter 3 we formally introduce Gaussian measures in infinite dimensional spaces. Before doing so, we first review general probability measures on Banach spaces. Once Gaussian measures are formally defined, we show that they have desirable properties. Of particular interest is the covariance structure of Gaussian measures, which we demonstrate with several examples. In addition, we also briefly discuss cylindrical measures and radonification in this chapter.

In Chapter 4 we formally introduce Bayesian inference in infinite dimensional spaces. First Bayes' rule on general function spaces is defined, and the basic

concepts of asymptotic analysis such as consistency and contraction rates are introduced. Bayes rule naturally induces an estimation procedure, which will be applied to Gaussian linear models. In particular, we consider the case with continuous, and the case with discrete observations. This chapter concludes with a general theorem on posterior contraction rates, which provides the general approach throughout this thesis.

Part II investigates linear inverse problems with Gaussian noise, which allows us to use the theory developed for Gaussian linear models. In the setting of inverse problems, one observes a smoothed signal contaminated by noise. The smoothing makes the problem difficult, and naive methods to recover the unsmoothed signal become inappropriate, because the noise dominates due to the inverse operator. One way to overcome this problem is by regularising the inverse operator, which can be accomplished by using Bayesian methods. In this part we investigate inverse problems under two observation schemes: continuously and discretely observed inverse problems, which are commonly known as the white noise model and regression.

Chapter 5 serves as an introduction to linear inverse problems. These problems are introduced both heuristically and mathematically rigorously, and several examples are given. This chapter concludes with the Galerkin method, a relatively general projection method, which is a simple but convenient tool for demonstrating contraction rates for inverse problems used in the remainder of this part.

In Chapter 6, we study the asymptotic performance of Bayesian nonparametric methods for linear inverse problems with continuous observations. A general theorem for continuously observed inverse problems is proved by modifying the proof of the general theorem on contraction rates that is derived in Chapter 4. In general adaptivity is not guaranteed. The additional conditions necessary for adaptivity are given in the second theorem in this chapter. The first theorem is then used to study the performance of two types of priors: random series priors and Gaussian priors. Random series priors lead to posteriors that contract at the minimax rate, up to a logarithmic factor, whereas Gaussian priors are in general suboptimal. To improve on the latter case, we introduce a Gaussian mixture prior, by setting a hyperprior on Gaussian priors, resulting in a procedure that is both adaptive and optimal in the minimax sense, which is shown with the second general theorem in this chapter.

Chapter 7 is our first attempt to tackle discretely observed inverse problems, by using Gaussian conjugacy in linear problems. The word 'discretely' refers to the fact that the observations are only recorded at certain points, in contrast to the continuous case, where the entire trajectory is observed. A sequence formulation for discretely observed inverse problems is derived based on the singular value decomposition of the smoothing operator. The main results include the contraction rates, and the credible sets for mildly and extremely ill-posed inverse problems. The results are further exemplified by simulations.

Chapter 8 is our second effort to analyse the discretely observed inverse problems. In this chapter, we use the results of Chapter 6 to examine the model of interest. For this we consider a signal reconstruction technique that builds continuous trajectories by interpolating discrete points, for which we also derive error estimates. This is followed by extending the general theorems from Chapter 6 that are valid for inverse problems with continuous observations to the discretely observed models. This leads to derivation for the contraction rates of random series priors, Gaussian priors, and Gaussian mixtures.

Part III studies the Bayesian nonparametric inference for stochastic evolution equations. The state space of the stochastic evolution equation is assumed to be infinite-dimensional and this dynamical system is assumed to be affected by a random noise process. In this part the dynamics are assumed to be driven by a deterministic component described by a linear partial differential equation, and a stochastic component described by additive space-time Gaussian noise. This model can be considered as a lifting of the white noise model to infinite-dimensional state space. The goal in this part is to derive theory for Bayesian inference concerning the drift and initial condition of the evolution equations.

Chapter 9 provides a general overview of stochastic evolution equations. We start with defining $\mathcal{Q}$-Wiener processes, a generalisation of Brownian motion to infinite-dimensional space. These $\mathcal{Q}$-Wiener processes allow us to formulate a class of stochastic integrals in Hilbert spaces, which is extended to a larger class of integrands. These integrals combined with the theory of deterministic evolution equations are then used to define stochastic evolution equations formally. We also demonstrate that the solutions to these equations can be represented by stochastic processes.

Chapter 10 studies the asymptotic performance of Gaussian priors in the recovery of the initial conditions and drifts of stochastic evolution equations. To obtain the results on posterior contraction, we modify the general approach established in Chapter 4. For the recovery of initial conditions, we introduce a spatial Gaussian prior and derive the rate at which the posterior distribution concentrates around the true initial condition. This task can be viewed as a generalisation of the extremely ill-posed problem from Chapter 7. The recovery of the drift component using Bayesian methods is also investigated for which we construct another type of Gaussian priors. These priors are developed to identify the anisotropic smoothness of signals in higher dimensions. A crucial insight is that we can apply a transformation to the observations so the model is converted into a multi-indexed Gaussian sequence model. This allows us to derive results for the contraction rates.