Statistics and Probability Letters 106 (2015) 1–4

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A note on the Karhunen–Loève expansions for infinite-dimensional Bayesian inverse problems

Jinglai Li

Institute of Natural Sciences, Department of Mathematics, and MOE Key Laboratory of Scientific and Engineering Computing, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 9 March 2015
Received in revised form 25 June 2015
Accepted 25 June 2015
Available online 3 July 2015

Keywords:
Karhunen–Loève expansion
Bayesian inverse problems
Gaussian measure

ABSTRACT

In this note, we consider the truncated Karhunen–Loève expansion for approximating the maximum a posteriori (MAP) solutions to infinite dimensional Bayesian inverse problems. In particular, we derive an a priori error bound between a MAP solution and its finite dimensional approximation.

1. Introduction

A large class of inverse problems consist of estimating infinite dimensional parameters such as functions of time or space (Kaipio and Somersalo, 2005; Hanke and Brühl, 2003). The Bayesian inference approaches (Stuart, 2010) which treat the unknown function as a random field have become a very popular tool to solve such problems, mainly due to their ability to quantify uncertainties in the estimation results. In practice, the inference problems often do not have closed form solutions and have to be solved numerically, for example, with Markov Chain Monte Carlo (MCMC) simulations. To this end, a common practice is to first project the unknown into a finite-dimensional subspace and then solve the resulting finite-dimensional problem numerically.

When the prior measure is Gaussian, the Karhunen–Loève (K–L) expansion (Papoulis and Pillai, 2002, Chapter 11) can be used to construct the finite-dimensional approximation. Namely one represents the unknown function by a finite expansion of the eigenfunctions of the covariance operator of the prior measure. The K–L method has been long used to reduce the dimensionality in both forward (Ghanem and Spanos, 2003; Xiu, 2010) and inverse stochastic problems (McLaughlin and Townley, 1996; Li and Cirpka, 2006; Marzouk and Najm, 2009). There are several works on theoretical error analysis (e.g., Charrier, 2012) for the KLE in the forward problems; however, the use of it is never rigorously justified for inverse problem, to the best of my knowledge. To be specific, it is unclear whether the reduced fixed-dimensional problem can well approximate the original one. In this note, we partially address the problem by proving that at least the maximum a posteriori (MAP) estimator defined as a minimizer of the Onsager–Machlup functional, is well approximated by the truncated K–L expansion. In particular we derive an *a priori error bound* between the MAP estimator and its finite-dimensional approximation.

2. Problem setup

We assume the state space X is a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_X$. Our goal is to estimate $u \in X$ from some data y. The Bayes' formula in this setting should be interpreted as providing the Radon–Nikodym derivative between

E-mail address: jinglaili@sjtu.edu.cn.

http://dx.doi.org/10.1016/j.spl.2015.06.025

the posterior measure μ and the prior measure μ_0 (Cotter et al., 2009; Dashti et al., 2013):

$$\frac{d\mu}{d\mu_0}(u) = \exp(-\Phi(u)), \tag{1}$$

where $\exp(-\Phi(u))$ is the likelihood function. A typical example is to assume that the unknown u is mapped to the data y via a forward model $y = G(u) + \zeta$, where $G : X \to R^d$ and ζ is a d-dimensional Gaussian noise with mean zero and covariance C. In this case $\Phi(u) = |C^{-\frac{1}{2}}(Gu - y)|_2^2$.

Next we assume a Gaussian prior is used. Namely we let μ_0 be a zero-mean Gaussian measure defined on X with covariance operator Q. Note that Q is symmetric positive and of trace class. $E = Q^{\frac{1}{2}}(X)$ is a Hilbert space with inner product

$$\langle \cdot, \cdot \rangle_E = \langle Q^{-\frac{1}{2}} \cdot, Q^{-\frac{1}{2}} \cdot \rangle_X,$$

which is known as the Cameron–Martin space associated with measure μ_0. Often we are only interested in a point estimate of u, rather than the posterior measure μ itself. To this end, as is shown in Cotter et al. (2009), Dashti et al. (2013), the maximum a posteriori (MAP) estimator of u can be defined as the minimizers of the Onsager–Machlup functional over E:

$$\min_{u \in E} I(u) := \Phi(u) + \|u\|_E^2, \tag{2}$$

where $\|u\|_E^2 = \langle u, u \rangle_E$. Note that Eq. (2) can also be understood as a classic inverse problem where the cost function $\Phi(\cdot)$ is minimized with a Tikhonov regularization in the Hilbert space E (Bissantz et al., 2004).

3. Karhunen–Loève representation

For convenience sake, we introduce the substitution $u = Q^{\frac{1}{2}}x$ and rewrite Eq. (2) as

$$\min_{x \in X} J(x) := \Phi(Q^{\frac{1}{2}}x) + \|x\|_X^2. \tag{3}$$

The following proposition states the equivalence of the two optimization problems.

Proposition 3.1. *If x minimizes $J(x)$ over X, $u = Q^{\frac{1}{2}}x$ minimizes $I(u)$ over E, and if u minimizes $I(u)$ over E, $x = Q^{-\frac{1}{2}}u \in X$ minimizes $J(x)$ over X.*

Proof. We prove the proposition by contradiction. First it is easy to verify that, for any $x \in X$ and $u \in E$ satisfying $u = Q^{\frac{1}{2}}x$, we have $I(u) = J(x)$. Let x be a minimizer $J(\cdot)$ over X, and assume $u = Q^{\frac{1}{2}}x$ is not a minimizer of $I(\cdot)$ over E. Namely, there exists an $u' \in E$ such that $I(u') < I(u)$. It follows directly that $x' = Q^{-\frac{1}{2}}u' \in X$ and $J(x') < J(x)$, which contradicts that x is a minimizer of J over X. Thus we have proved the first part of the proposition. The second part can be proved by following the same argument.

Now we introduce the K–L expansion to reduce the dimensionality of Eq. (3). We start with the following lemma (Da Prato, 2006, Chapter 1).

Lemma 3.2. *There exists a complete orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ on X and a sequence of non-negative numbers $\{\lambda_k\}_{k\in\mathbb{N}}$ such that $Qe_k = \lambda_k e_k$ and $\sum_{k=1}^{\infty} \lambda_k < \infty$, i.e., $\{e_k\}_{k\in\mathbb{N}}$ and $\{\lambda_k\}_{k\in\mathbb{N}}$ being the eigenfunctions and eigenvalues of Q respectively.*

The basic idea of the K–L method is to solve the optimization problem in a finite-dimensional subspace of X:

$$\min_{x \in X_n} J(x) := \Phi(Q^{\frac{1}{2}}x) + \|x\|_X^2, \tag{4}$$

where X_n be the space spanned by $\{e_k\}_{k=1}^n$ for a given $n \in \mathbb{N}$. In numerical implementation Eq. (4) can be recast as

$$\min_{(\xi_1, \ldots, \xi_n) \in R^n} \Phi\left(\sum_{k=1}^{n} \xi_k \sqrt{\lambda_k} e_k\right) + \sum_{k=1}^{n} \xi_k^2 \tag{5}$$

which is the usual MAP estimator in the finite dimensional setting. As is mentioned earlier, a critical question here is whether the finite subspace X_n can provide good approximation to the solutions of Eq. (3). Our main results regarding this problem are presented in the following theorem:

Theorem 3.3. *Suppose $\Phi(u)$ is locally Lipschitz continuous, i.e., for every $r > 0$, there exists a constant $L_r > 0$ such that for all $z_1, z_2 \in X$ with $\|z_1\|_X, \|z_2\|_X < r$, we have*

$$|\Phi(z_1) - \Phi(z_2)| < L_r \|z_1 - z_2\|_X.$$

Let $\{e_k\}_{k\in\mathbb{N}}$ and $\{\lambda_k\}_{k\in\mathbb{N}}$ be the eigenfunctions and eigenvalues of Q as defined in Lemma 3.2. There exists a constant $L > 0$ such that, for any $x \in \arg\min_{x\in X} J(x)$, we have

$$\|x - x_n\|_X < L\sqrt{\lambda_n^*},$$

where $x_n = \sum_{k=1}^n \langle x, e_k\rangle_X e_k$, and $\lambda_n^* = \max_{k>n} \lambda_k$. Moreover, let $u = Q^{\frac{1}{2}}x$ and $u_n = Q^{\frac{1}{2}}x_n$, and we have $\|u - u_n\|_X < L\lambda_n^*$.

Proof. Let $x \in X$ be a minimizer of Eq. (3). Since $\{e_k\}$ is a complete orthonormal basis for X, x can be written as

$$x = \sum_{k=1}^{\infty} \xi_k e_k,$$

where $\xi_k = \langle x, e_k\rangle_X$. Let

$$x_n = \sum_{k=1}^{n} \xi_k e_k.$$

As x is a minimizer of $J(\cdot)$, take $r = \Phi(0) + 1$ and so we have $J(x) < r$, which implies that $\|x_n\|_X \le \|x\|_X < r$. $Q^{\frac{1}{2}}$ is bounded, and so we have $\|Q^{\frac{1}{2}}x_n\|_X, \|Q^{\frac{1}{2}}x\|_X < \|Q^{\frac{1}{2}}\|r$. Now recall that $\Phi(\cdot)$ is locally Lipschitz continuous, and so there exists a constant $L > 0$ such that

$$\left|\Phi\left(Q^{\frac{1}{2}}x\right) - \Phi\left(Q^{\frac{1}{2}}x_n\right)\right| < L\|Q^{\frac{1}{2}}x - Q^{\frac{1}{2}}x_n\|_X.$$

Since x minimizes $J(\cdot)$, we have $J(x) \le J(x_n)$ which implies

$$\begin{aligned}\|x - x_n\|_X^2 &\le \left|\Phi\left(Q^{\frac{1}{2}}x\right) - \Phi\left(Q^{\frac{1}{2}}x_n\right)\right| < L\|Q^{\frac{1}{2}}x - Q^{\frac{1}{2}}x_n\|_X \\ &= L\langle x - x_n, Q(x - x_n)\rangle_X^{\frac{1}{2}} = L\left\langle \sum_{k=n+1}^{\infty} \xi_k e_k, \sum_{k=n+1}^{\infty} \xi_k\lambda_k e_k \right\rangle_X^{\frac{1}{2}} \\ &= L\sqrt{\sum_{k=n+1}^{\infty} \lambda_k \xi_k^2} \le L\sqrt{\lambda_n^*}\|x - x_n\|_X.\end{aligned}$$

It then follows immediately that

$$\|x - x_n\|_X \le L\sqrt{\lambda_n^*}, \quad \text{and} \quad \|u - u_n\|_X < L\lambda_n^*.$$

Finally we want to know if the finite-dimensional MAP estimator, i.e. the solution to Eq. (4) or equivalently Eq. (4), is a good approximation to the infinite dimensional one. Regarding this we have the following results:

Corollary 3.4. *Let $x_n' \in \arg\min_{x\in X_n} J(x)$ and we have*

$$\min_{x\in X} J(x) \le J(x_n') \le \min_{x\in X} J(x) + L^2\lambda_n^*.$$

The corollary follows directly from Theorem 3.3 and so proof is omitted.

4. Concluding remarks

We theoretically study the truncated K–L expansions for approximating the solutions of infinite-dimensional Bayesian inverse problems. We show that the error between a MAP estimator and its projection on a pre-determined finite-dimensional space is bounded by the eigenvalues of the covariance operator of the prior.

Acknowledgment

The work was supported by the NSFC under grant number 11301337.

References

Bissantz, N., Hohage, T., Munk, A., 2004. Consistency and rates of convergence of nonlinear tikhonov regularization with random noise. Inverse Problems 20 (6), 1773.
Charrier, Julia, 2012. Strong and weak error estimates for elliptic partial differential equations with random coefficients. SIAM J. Numer. Anal. 50 (1), 216–246.
Cotter, S.L., Dashti, M., Robinson, J.C., Stuart, A.M., 2009. Bayesian inverse problems for functions and applications to fluid mechanics. Inverse Problems 25 (11), 115008.
Da Prato, G., 2006. An Introduction to Infinite-dimensional Analysis. Springer.
Dashti, M., Law, K.J.H., Stuart, A.M., Voss, J., 2013. MAP estimators and their consistency in Bayesian nonparametric inverse problems. Inverse Problems 29 (9), 095017.
Ghanem, Roger G., Spanos, Pol D., 2003. Stochastic Finite Elements: A Spectral Approach. Courier Corporation.
Hanke, M., Brühl, M., 2003. Recent progress in electrical impedance tomography. Inverse Problems 19 (6), S65.
Kaipio, Jari, Somersalo, Erkki, 2005. Statistical and Computational Inverse Problems, Vol. 160. Springer.
Li, W., Cirpka, O.A., 2006. Efficient geostatistical inverse methods for structured and unstructured grids. Water Resour. Res. 42 (6).
Marzouk, Y.M., Najm, H.N., 2009. Dimensionality reduction and polynomial chaos acceleration of bayesian inference in inverse problems. J. Comput. Phys. 228 (6), 1862–1902.
McLaughlin, D., Townley, L.R., 1996. A reassessment of the groundwater inverse problem. Water Resour. Res. 32 (5), 1131–1161.
Papoulis, A., Pillai, S.U., 2002. Probability, Random Variables, and Stochastic Processes. Tata McGraw-Hill Education.
Stuart, A.M., 2010. Inverse problems: a Bayesian perspective. Acta Numer. 19, 451–559.
Xiu, Dongbin, 2010. Numerical Methods for Stochastic Computations: A Spectral Method Approach. Princeton University Press.