The posterior mean and covariance operator in linear Gaussian inverse problems

Alen Alexanderian*

Abstract

We consider two different representations of the mean and posterior covariance operator in infinite-dimensional linear Gaussian inverse problems. The results considered here are known. The goal is to provide some clear and self-contained derivations.

1 Introduction

Let $\mathcal{M}$ be an infinite-dimensional real separable Hilbert space. Consider a Bayesian inverse problem of estimating $m \in \mathcal{M}$ using the observation model

$$y = \mathrm{F}m + \eta. \tag{1.1}$$

Here, $\mathrm{F} : \mathcal{M} \to \mathbb{R}^d$ is a continuous linear transformation, and $\eta \in \mathbb{R}^d$ models observation error and is distributed according to $N(0, \Gamma)$, where $\Gamma \in \mathbb{R}^{d\times d}$ symmetric positive definite. We assume a Gaussian prior $N(0, \mathrm{C}_0)$. The discussions that follow can be easily adapted to the case of a nonzero prior mean m_0 that is sufficiently regular.[1]

[1] *More precisely, m_0 must belong to the Cameron–Martin space, which is defined as $\mathcal{E} := \mathrm{Range}(\mathrm{C}_0^{1/2})$.*

We also define,

$$\tilde{\mathrm{H}} := \mathrm{C}_0^{1/2}\mathrm{F}^*\Gamma^{-1}\mathrm{F}\mathrm{C}_0^{1/2}.$$

This is the prior-preconditioned data misfit Hessian. The present terminology is motivated from a variational perspective. Namely, $\tilde{\mathrm{H}}$ is a symmetrically preconditioned version of the data-misfit Hessian in the optimization problem characterizing the MAP point.[2]

[2] *Recall that in linear Gaussian inverse problems the MAP point and the posterior mean coincide.*

In the present setup, it is known [2] that the posterior is a Gaussian measure $\mu_{\text{post}} = N(m^*, \mathrm{C})$, with

$$\mathrm{C} = \mathrm{C}_0 - \mathrm{C}_0\mathrm{F}^*(\mathrm{F}\mathrm{C}_0\mathrm{F}^* + \Gamma)^{-1}\mathrm{F}\mathrm{C}_0 \tag{1.2}$$

$$m^* = \mathrm{C}_0\mathrm{F}^*(\Gamma + \mathrm{F}\mathrm{C}_0\mathrm{F}^*)^{-1}y. \tag{1.3}$$

In what follows, we seek to justify the alternative formulas for m^* and C. Specifically, we seek to prove the following results:

Theorem 1.1. *The following holds*

$$\mathrm{C} = \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{H}})^{-1}\mathrm{C}_0^{1/2}. \tag{1.4}$$

Theorem 1.2. *The map point satisfies*

$$m^* = \mathrm{C}\mathrm{F}^*\Gamma^{-1}y. \tag{1.5}$$

*North Carolina State University, Raleigh, NC, USA. E-mail alexanderian@ncsu.edu

Last revised: December 8, 2025

Note that the covariance operator on the left-hand side of (1.4) is the well-known "Woodbury form" of the posterior covariance operator, which avoids explicitly inverting C_0. The operator on the right-hand side also avoids C_0^{-1} and is convenient for computations. In particular, this form facilitates using a low-rank spectral decomposition of $\tilde{\mathrm{H}}$ for fast computations, a structure that commonly arises in ill-posed inverse problems [1]

2 The covariance expression

We first prove some technical lemmas. These lemmas are not meant to be general and concern only the specific problem under study.

Lemma 2.1. *Let* $\mathrm{G} : \mathcal{M} \to \mathbb{R}^d$ *be a bounded linear operator. Then* $\mathrm{I} + \mathrm{G}^*\mathrm{G}$ *and* $\mathrm{I} + \mathrm{G}\mathrm{G}^*$ *are bounded linear operators that have a bounded inverse.*

Proof. Note that $K = \mathrm{G}^*\mathrm{G}$ is a positive selfadjoint compact operator and $\mathrm{I} + K$ is injective. Hence, by the Fredholm alternative, $\mathrm{I} + K$ has a bounded inverse. An analogous argument shows that $\mathrm{I} + \mathrm{G}\mathrm{G}^*$ has a bounded inverse as well. □

Lemma 2.2. *Let* $\mathrm{G} : \mathcal{M} \to \mathbb{R}^d$ *be bounded linear operator. The following holds.*

$$\mathrm{I} - \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} = (\mathrm{I} + \mathrm{G}^*\mathrm{G})^{-1}.$$

Proof. The result follows from a direct calculation:

$$\begin{aligned}
&(\mathrm{I} + \mathrm{G}^*\mathrm{G})\big(\mathrm{I} - \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G}\big) \\
&\qquad = \mathrm{I} - \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} + \mathrm{G}^*\mathrm{G} - \mathrm{G}^*\mathrm{G}\,\mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} \\
&\qquad = \mathrm{I} + \mathrm{G}^*\mathrm{G} - \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} - \mathrm{G}^*(\mathrm{G}\mathrm{G}^*)(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} \\
&\qquad = \mathrm{I} + \mathrm{G}^*\mathrm{G} - \mathrm{G}^*\big((\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G} + (\mathrm{G}\mathrm{G}^*)(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}\mathrm{G}\big) \\
&\qquad = \mathrm{I} + \mathrm{G}^*\mathrm{G} - \mathrm{G}^*[(\mathrm{I} + \mathrm{G}\mathrm{G}^*)(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}]\mathrm{G} \\
&\qquad = \mathrm{I} + \mathrm{G}^*\mathrm{G} - \mathrm{G}^*\mathrm{G} \\
&\qquad = \mathrm{I}.
\end{aligned}$$
□

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Let the operator $\tilde{\mathrm{F}}$ be defined by $\tilde{\mathrm{F}} := \Gamma^{-1/2}\mathrm{F}\mathrm{C}_0^{1/2}$. Note that $\tilde{\mathrm{F}}$ is a finite-dimensional operator. Furthermore, the following are straightforward to note

$$\tilde{\mathrm{F}}^*\tilde{\mathrm{F}} = \tilde{\mathrm{H}} \tag{2.1}$$

$$\Gamma^{1/2}(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)\Gamma^{1/2} = \mathrm{F}\mathrm{C}_0\mathrm{F}^* + \Gamma. \tag{2.2}$$

Subsequently,

$$\begin{aligned}
\mathrm{C} &= \mathrm{C}_0 - \mathrm{C}_0\mathrm{F}^*(\mathrm{F}\mathrm{C}_0\mathrm{F}^* + \Gamma)^{-1}\mathrm{F}\mathrm{C}_0 \\
&= \mathrm{C}_0 - \mathrm{C}_0\mathrm{F}^*\big(\Gamma^{1/2}(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)\Gamma^{1/2}\big)^{-1}\mathrm{F}\mathrm{C}_0 && \text{by (2.2)} \\
&= \mathrm{C}_0 - \mathrm{C}_0\mathrm{F}^*\Gamma^{-1/2}(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)^{-1}\Gamma^{-1/2}\mathrm{F}\mathrm{C}_0 \\
&= \mathrm{C}_0^{1/2}\Big[\mathrm{I} - \mathrm{C}_0^{1/2}\mathrm{F}^*\Gamma^{-1/2}(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)^{-1}\Gamma^{-1/2}\mathrm{F}\mathrm{C}_0^{1/2}\Big]\mathrm{C}_0^{1/2} \\
&= \mathrm{C}_0^{1/2}\Big[\mathrm{I} - \tilde{\mathrm{F}}^*(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)^{-1}\tilde{\mathrm{F}}\Big]\mathrm{C}_0^{1/2} \\
&= \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{F}}^*\tilde{\mathrm{F}})^{-1}\mathrm{C}_0^{1/2} && \text{by Lemma 2.2} \\
&= \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{H}})^{-1}\mathrm{C}_0^{1/2}. && \text{by (2.1)}
\end{aligned}$$
□

3 The m^* expression

We first record the following technical lemma.

Lemma 3.1. *Let* $\mathrm{G} : \mathcal{M} \to \mathbb{R}^d$ *be a bounded linear operator. Then* $(\mathrm{I} + \mathrm{G}^*\mathrm{G})^{-1}\mathrm{G}^* = \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)^{-1}$.

Proof. By Lemma 2.1, both $\mathrm{I} + \mathrm{G}^*\mathrm{G}$ and $\mathrm{I} + \mathrm{G}\mathrm{G}^*$ are bounded operators that admit bounded inverses. Next, note that $(\mathrm{I} + \mathrm{G}^*\mathrm{G})\mathrm{G}^* = \mathrm{G}^*(\mathrm{I} + \mathrm{G}\mathrm{G}^*)$. Right-multiplying this equation by $(\mathrm{I}+\mathrm{G}\mathrm{G}^*)^{-1}$ and then left-multiplying the resulting equation by $(\mathrm{I} + \mathrm{G}^*\mathrm{G})^{-1}$ yields the desired identity. □

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. As in the proof of Theorem 1.1, we define the operator $\tilde{\mathrm{F}} := \Gamma^{-1/2}\mathrm{F}\mathrm{C}_0^{1/2}$. Consider the expression (1.3) for m^*. We have

$$\begin{aligned} m^* &= \mathrm{C}_0\mathrm{F}^*(\Gamma + \mathrm{F}\mathrm{C}_0\mathrm{F}^*)^{-1}y \\ &= \mathrm{C}_0\mathrm{F}^*\Gamma^{-1/2}(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)^{-1}\Gamma^{-1/2}y \\ &= \mathrm{C}_0^{1/2}\tilde{\mathrm{F}}^*(\mathrm{I} + \tilde{\mathrm{F}}\tilde{\mathrm{F}}^*)^{-1}\Gamma^{-1/2}y, \end{aligned}$$

where the second equality follows from (2.2). Therefore, using Lemma 3.1 and (2.1),

$$\begin{aligned} m^* &= \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{F}}^*\tilde{\mathrm{F}})^{-1}\tilde{\mathrm{F}}^*\Gamma^{-1/2}y \\ &= \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{H}})^{-1}\tilde{\mathrm{F}}^*\Gamma^{-1/2}y \\ &= \mathrm{C}_0^{1/2}(\mathrm{I} + \tilde{\mathrm{H}})^{-1}\mathrm{C}_0^{1/2}\mathrm{F}^*\Gamma^{-1}y \\ &= \mathrm{C}\mathrm{F}^*\Gamma^{-1}y. \quad \square \end{aligned}$$

References

[1] Tan Bui-Thanh, Omar Ghattas, James Martin, and Georg Stadler. A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion. *SIAM Journal on Scientific Computing*, 35(6):A2494–A2523, 2013.

[2] Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.