Gaussian Processes for PDEs

Yanni Papandreou

November 2019

1 Notes 1

Let Ω be a domain with boundary $\partial\Omega$ and let $\mathcal{L}$ be a suitable linear differential operator. We consider the following Dirichlet problem:

$$\begin{cases} \mathcal{L}u = f & \text{on } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \tag{1}$$

where the functions f, g may be noisy, typically Gaussian. The operator $\mathcal{L}$ may also be random. For simplicity we will take $g = 0$ and will assume that $\mathcal{L}$ is deterministic. We seek the solution in some appropriate Hilbert space of functions $\mathcal{H}$. Formally we have $u = \mathcal{L}^{-1}f$.

Remark. If f is Gaussian, i.e. $f \sim \mathcal{N}(\bar{f}, K)$, then (u, f) is joint Gaussian:

$$\begin{pmatrix} f \\ u \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \bar{f} \\ \mathcal{L}^{-1}\bar{f} \end{pmatrix}, \begin{pmatrix} K & K(\mathcal{L}^{-1})^* \\ \mathcal{L}^{-1}K & \mathcal{L}^{-1}K(\mathcal{L}^{-1})^* \end{pmatrix} \right) \tag{2}$$

provided that $\mathcal{L}^{-1}$ is bounded.

Let us consider placing a Gaussian prior on u, $u \sim \mathcal{N}(0, V)$, where we control V so that u is almost surely in some appropriate subspace $\mathcal{U} \subset \mathcal{H}$ where the linear operator $\mathcal{L}$ restricted to this subspace is bounded. I.e. $\mathcal{L} : \mathcal{U} \to \mathcal{U}'$ is bounded where $\mathcal{U} \subset \mathcal{H} \subset \mathcal{U}'$.

Remark. We shall assume here that the subspace $\mathcal{U}$ is reflexive so that $\mathcal{U}''$ can be identified with $\mathcal{U}$.

Now consider the following information operators:

$$I_j \cdot = \int_\Omega \psi_j \cdot \mathrm{d}x, \quad j = 1, \ldots, J$$

where $\psi_j \in \mathcal{F}_h$ for $j = 1, \ldots, J$ where $\mathcal{F}_h$ is some discretisation of the function space $\mathcal{H}$ (and h is the mesh size).

Remark. Think of the $\{\psi_j\}$ as a basis for the finite element spaces.

Let $\mathcal{I} = (I_1, \ldots, I_J)^T$. We have,

$$\mathcal{I}\mathcal{L}u = (I_1\mathcal{L}u, \ldots, I_J\mathcal{L}u) \in \mathbb{R}^J$$

and under our assumptions on V we have that,

$$\begin{pmatrix} u \\ \mathcal{I}\mathcal{L}u \end{pmatrix} = \begin{pmatrix} I \\ \mathcal{I}\mathcal{L} \end{pmatrix} u \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \mathbf{0}_J \end{pmatrix}, \begin{pmatrix} V & V\mathcal{L}^*\mathcal{I}^* \\ \mathcal{I}\mathcal{L}V & \mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^* \end{pmatrix} \right) \tag{3}$$

Under the assumption that (1) holds we have $\mathcal{I}\mathcal{L}u = \mathcal{I}f =: F$. Assuming now that we fix a realisation of f (and so F is fixed) we have by conditioning (see Notes 3 for a justification of why this is valid) that:

$$u | \{\mathcal{I}\mathcal{L}u = F, f\} \sim \mathcal{N}(a, \Sigma) =: \mu_{a,\Sigma} \tag{4}$$

where,

$$a = V\mathcal{L}^*\mathcal{I}^*(\mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^*)^{-1}F \tag{5}$$

$$\Sigma = V - V\mathcal{L}^*\mathcal{I}^*(\mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^*)^{-1}\mathcal{I}\mathcal{L}V \tag{6}$$

Remark. We use the notation $\mu_{m,A}$ to denote a Gaussian measure (in arbitrary dimensions) with mean m and covariance A. The space (and hence the dimension) should be clear from the domain of the covariance operator.

We will now assume that f is distributed as $f \sim \mathcal{N}(\bar{f}, K)$ and we will seek to marginalize over f in (4) so as to obtain the "average" conditional distribution over all possible values of f. To do so we first note:

$$f \sim \mathcal{N}(\bar{f}, K) =: \mu_{\bar{f}, K} \implies \mathcal{I}f = F \sim \mathcal{N}(\mathcal{I}\bar{f}, \mathcal{I}K\mathcal{I}^*) =: \mu_{\bar{F}, K_{\mathcal{I}}} \tag{7}$$

where $\bar{F} := \mathcal{I}\bar{f}$ and $K_{\mathcal{I}} := \mathcal{I}K\mathcal{I}^*$. In order to specify the "average" conditional distribution it suffices to compute the expectation of arbitrary bounded cylindrical test functions $\phi(u^N) := \phi(u(x_1), \ldots, u(x_N))$ since the $\sigma-$algebra generated by cylinder sets coincides with the Borel $\sigma-$algebra (see for instance Theorem 2.1.1 in [1]). We must thus compute:

$$\int \int \phi(u^N) \mu_{a, \Sigma}(\mathrm{d}u) \mu_{\bar{f}, K}(\mathrm{d}f) \tag{8}$$

Note that $u^N = Pu$ where the bounded linear operator $P : \mathcal{H} \to \mathbb{R}^N$ is defined by $Ph := (h(x_1), \ldots, h(x_N))^T$ for any function h. Thus, u^N is multivariate normal, i.e., $u^N \sim \mathcal{N}(Pa, P\Sigma P^*) =: \mu_{Pa, \Sigma_N}$ where $\Sigma_N := P\Sigma P^*$ is the $N \times N$ covariance matrix of u^N. We thus have:

$$\int \int \phi(u^N) \mu_{a, \Sigma}(\mathrm{d}u) \mu_{\bar{f}, K}(\mathrm{d}f) = \int \int \phi(u^N) \mu_{Pa, \Sigma_N}(\mathrm{d}u^N) \mu_{\bar{f}, K}(\mathrm{d}f) \tag{9}$$

We note that $Pa = PAF$ where $A := V\mathcal{L}^*\mathcal{I}^*(\mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^*)^{-1}$. Since the conditional distribution of u (4) and hence of u^N only depends on f through $F \in \mathbb{R}^J$ we can thus write:

$$\int \int \phi(u^N) \mu_{Pa, \Sigma_N}(\mathrm{d}u^N) \mu_{\bar{f}, K}(\mathrm{d}f) = \int \int \phi(u^N) \mu_{Pa, \Sigma_N}(\mathrm{d}u^N) \mu_{\bar{F}, K_{\mathcal{I}}}(\mathrm{d}F) \tag{10}$$

Both measures in the above integral are now multivariate normal and so we have:

$$\int \int \phi(u^N) \mu_{Pa, \Sigma_N}(\mathrm{d}u^N) \mu_{\bar{F}, K_{\mathcal{I}}}(\mathrm{d}F) =$$

$$= \frac{1}{Z_u} \int \int \phi(u^N) \exp\left(-\frac{1}{2}\left\langle u^N - PAF, \Sigma_N^{-1}(u^N - PAF)\right\rangle\right) \mu_{\bar{F}, K_{\mathcal{I}}}(\mathrm{d}F)\mathrm{d}u^N$$

$$= \frac{1}{Z_u Z_f} \int \int \phi(u^N) \exp\left(-\frac{1}{2}\left\langle u^N - PAF, \Sigma_N^{-1}(u^N - PAF)\right\rangle\right) \exp\left(-\frac{1}{2}\left\langle F - \bar{F}, K_{\mathcal{I}}^{-1}(F - \bar{F})\right\rangle\right) \mathrm{d}F\mathrm{d}u^N \tag{11}$$

where the normalization constants are $Z_u := (2\pi)^{N/2} \det(\Sigma_N)^{1/2}$ and $Z_f := (2\pi)^{J/2} \det(K_{\mathcal{I}})^{1/2}$. We now need to compute the integral over F. In order to do so we combine the exponents in (11) into a quadratic in F in order to be able to use the well-known formula of a multidimensional Gaussian integral. Going through the algebra we obtain:

$$\frac{1}{Z_u Z_f} \int \int \phi(u^N) \exp\Bigg(-\frac{1}{2}\Big(\left\langle u^N, \Sigma_N^{-1}u^N\right\rangle - 2\left\langle \Sigma_N^{-1}PAF, u^N\right\rangle + \left\langle F, A^*P^*\Sigma_N^{-1}PAF\right\rangle$$

$$+ \left\langle F, K_{\mathcal{I}}^{-1}F\right\rangle - 2\left\langle K_{\mathcal{I}}^{-1}\bar{F}, F\right\rangle + \left\langle \bar{F}, K_{\mathcal{I}}^{-1}\bar{F}\right\rangle\Big)\Bigg)\mathrm{d}F\mathrm{d}u^N =$$

$$= \frac{1}{Z_u Z_f} \int \phi(u^N) \exp\left(-\frac{1}{2}\left(\left\langle u^N, \Sigma_N^{-1}u^N\right\rangle + \left\langle \bar{F}, K_{\mathcal{I}}^{-1}\bar{F}\right\rangle\right)\right) \cdot$$

$$\left(\int \exp\left(-\frac{1}{2}\left\langle F, BF\right\rangle + \left\langle A^*P^*\Sigma_N^{-1}u^N + K_{\mathcal{I}}^{-1}\bar{F}, F\right\rangle\right)\mathrm{d}F\right)\mathrm{d}u^N \tag{12}$$

where $B := \left(A^*P^*\Sigma_N^{-1}PA + K_{\mathcal{I}}^{-1}\right)$. Computing the inner integral over F we thus obtain:

$$\frac{(2\pi)^{J/2}}{Z_u Z_f \det(B)^{1/2}} \int \phi(u^N) \exp\Bigg(\frac{1}{2}\left\langle A^*P^*\Sigma_N^{-1}u^N + K_{\mathcal{I}}^{-1}\bar{F}, B^{-1}\left(A^*P^*\Sigma_N^{-1}u^N + K_{\mathcal{I}}^{-1}\bar{F}\right)\right\rangle$$

$$-\frac{1}{2}\left(\left\langle u^N, \Sigma_N^{-1}u^N\right\rangle + \left\langle \bar{F}, K_{\mathcal{I}}^{-1}\bar{F}\right\rangle\right)\Bigg)\mathrm{d}u^N \tag{13}$$

We now focus on the terms in the exponent and simplify these as follows:

$$\frac{1}{2}\left\langle A^*P^*\Sigma_N^{-1}u^N + K_\mathcal{I}^{-1}\bar{F}, B^{-1}\left(A^*P^*\Sigma_N^{-1}u^N + K_\mathcal{I}^{-1}\bar{F}\right)\right\rangle - \frac{1}{2}\left(\langle u^N, \Sigma_N^{-1}u^N\rangle + \langle \bar{F}, K_\mathcal{I}^{-1}\bar{F}\rangle\right) =$$

$$= -\frac{1}{2}\Big(\langle u^N, \Sigma_N^{-1}u^N\rangle + \langle \bar{F}, K_\mathcal{I}^{-1}\bar{F}\rangle - \langle A^*P^*\Sigma_N^{-1}u^N + K_\mathcal{I}^{-1}\bar{F}, B^{-1}\left(A^*P^*\Sigma_N^{-1}u^N + K_\mathcal{I}^{-1}\bar{F}\right)\rangle\Big)$$

$$= -\frac{1}{2}\Big(\langle u^N, \Sigma_N^{-1}u^N\rangle + \langle \bar{F}, K_\mathcal{I}^{-1}\bar{F}\rangle - \langle u^N, \Sigma_N^{-1}PAB^{-1}A^*P^*\Sigma_N^{-1}u^N\rangle$$

$$-2\langle u^N, \Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}\rangle - \langle \bar{F}, K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1}\bar{F}\rangle\Big)$$

$$= -\frac{1}{2}\Big(\langle u^N, (\Sigma_N^{-1} - \Sigma_N^{-1}PAB^{-1}A^*P^*\Sigma_N^{-1})u^N\rangle - 2\langle u^N, \Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}\rangle + \langle \bar{F}, (K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1})\bar{F}\rangle\Big)$$

$$= -\frac{1}{2}\Big(\langle u^N, \Sigma_\mathcal{I}^{-1}u^N\rangle - 2\langle u^N, \Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}\rangle + \langle \bar{F}, (K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1})\bar{F}\rangle\Big) \qquad (14)$$

where $\Sigma_\mathcal{I} := \Sigma_N + PAK_\mathcal{I}A^*P^*$. The inverse of $\Sigma_\mathcal{I}$ is indeed the coefficient matrix for the quadratic term in u^N in (14). This can be seen by utilizing the Woodbury matrix identity as follows:

$$\Sigma_\mathcal{I}^{-1} = (\Sigma_N + PAK_\mathcal{I}A^*P^*)^{-1}$$
$$= \Sigma_N^{-1} - \Sigma_N^{-1}PA(K_\mathcal{I}^{-1} + A^*P^*\Sigma_N^{-1}PA)^{-1}A^*P^*\Sigma_N^{-1}$$
$$= \Sigma_N^{-1} - \Sigma_N^{-1}PAB^{-1}A^*P^*\Sigma_N^{-1}$$

We now complete the square (in terms of u^N) to obtain:

$$-\frac{1}{2}\Big(\langle u^N - h^N, \Sigma_\mathcal{I}^{-1}(u^N - h^N)\rangle + \langle \bar{F}, (K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1})\bar{F}\rangle - \langle \Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}, \Sigma_\mathcal{I}\Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}\rangle\Big) =$$

$$= -\frac{1}{2}\Big(\langle u^N - h^N, \Sigma_\mathcal{I}^{-1}(u^N - h^N)\rangle + \langle \bar{F}, (K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1})\bar{F}\rangle\Big)$$
$$(15)$$

where $h^N := \Sigma_I\Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}\bar{F}$. We now show that the term quadratic in $\bar{F}$ vanishes by showing that the coefficient matrix of $\bar{F}$ is equal to the zero matrix. Note that this coefficent matrix can be rewritten as:

$$K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1} =$$
$$= K_\mathcal{I}^{-1}K_\mathcal{I}K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}K_\mathcal{I}^{-1} - K_\mathcal{I}^{-1}B^{-1}A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PAB^{-1}K_\mathcal{I}^{-1}$$
$$= K_\mathcal{I}^{-1}(K_\mathcal{I} - B^{-1} - B^{-1}A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PAB^{-1})K_\mathcal{I}^{-1}$$
$$= K_\mathcal{I}^{-1}B^{-1}(BK_\mathcal{I}B - B - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA)B^{-1}K_\mathcal{I}^{-1}$$

and so showing that it is the zero matrix is equivalent to showing that $(BK_\mathcal{I}B - B - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA)$ is the zero matrix. This can be shown as follows:

$$BK_\mathcal{I}B - B - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA =$$
$$= (A^*P^*\Sigma_N^{-1}PA + K_\mathcal{I}^{-1})K_\mathcal{I}B - B - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA$$
$$= A^*P^*\Sigma_N^{-1}PAK_\mathcal{I}B + B - B - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA$$
$$= A^*P^*\Sigma_N^{-1}PAK_\mathcal{I}(A^*P^*\Sigma_N^{-1}PA + K_\mathcal{I}^{-1}) - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA$$
$$= A^*P^*\Sigma_N^{-1}PAK_\mathcal{I}A^*P^*\Sigma_N^{-1}PA + A^*P^*\Sigma_N^{-1}PA - A^*P^*\Sigma_N^{-1}\Sigma_\mathcal{I}\Sigma_N^{-1}PA$$
$$= A^*P^*\Sigma_N^{-1}(PAK_\mathcal{I}A^*P^* + \Sigma_N - \Sigma_\mathcal{I})\Sigma_N^{-1}PA = 0$$

where the last equality follows by the definition of $\Sigma_\mathcal{I}$. Thus, our integral simplifies to:

$$\frac{(2\pi)^{J/2}}{Z_u Z_f \det(B)^{1/2}}\int \phi(u^N)\exp\Big(-\frac{1}{2}\langle u^N - h^N, \Sigma_\mathcal{I}^{-1}(u^N - h^N)\rangle\Big)\mathrm{d}u^N \qquad (16)$$

We now focus on simplifying the normalizing constants infront of the integral:

$$\frac{(2\pi)^{J/2}}{Z_u Z_f \det(B)^{1/2}} = \frac{(2\pi)^{J/2}}{(2\pi)^{N/2}\det(\Sigma_N)^{1/2}(2\pi)^{J/2}\det(K_\mathcal{I})^{1/2}\det(B)^{1/2}} =$$
$$= \frac{1}{(2\pi)^{N/2}\det(\Sigma_N)^{1/2}\det(K_\mathcal{I})^{1/2}\det(B)^{1/2}}$$

3

To proceed we note that $\det B$ can be rewritten as follows:

$$
\begin{aligned}
\det(B) &= \det(A^* P^* \Sigma_N^{-1} P A + K_{\mathcal{I}}^{-1}) \\
&= \det(K_{\mathcal{I}}^{-1}(I + K_{\mathcal{I}} A^* P^* \Sigma_N^{-1} P A)) \\
&= \det(K_{\mathcal{I}}^{-1}) \det(I + (K_{\mathcal{I}} A^* P^*)(\Sigma_N^{-1} P A))) \\
&= \det(K_{\mathcal{I}})^{-1} \det(I + \Sigma_N^{-1} P A K_{\mathcal{I}} A^* P^*)
\end{aligned}
$$

where we have utilized Sylvester's determinant theorem (*note: the identity matrices in the last two lines are of different sizes*). We can now finish up the simplification of the constants outside the integral:

$$
\begin{aligned}
&\frac{1}{(2\pi)^{N/2} \det(\Sigma_N)^{1/2} \det(K_{\mathcal{I}})^{1/2} \det(B)^{1/2}} \\
&= \frac{1}{(2\pi)^{N/2} \det(\Sigma_N)^{1/2} \det(K_{\mathcal{I}})^{1/2} \det(K_{\mathcal{I}})^{-1/2} \det(I_N + \Sigma_N^{-1} P A K_{\mathcal{I}} A^* P^*)^{1/2}} \\
&= \frac{1}{(2\pi)^{N/2} \det(\Sigma_N + P A K_{\mathcal{I}} A^* P^*)^{1/2}} \\
&= \frac{1}{(2\pi)^{N/2} \det(\Sigma_{\mathcal{I}})^{1/2}}
\end{aligned}
$$

Thus, our integral becomes:

$$
\begin{aligned}
\int \phi(u^N) \frac{1}{(2\pi)^{N/2} \det(\Sigma_{\mathcal{I}})^{1/2}} \exp\left(-\frac{1}{2}\left\langle u^N - h^N, \Sigma_{\mathcal{I}}^{-1}(u^N - h^N) \right\rangle\right) \mathrm{d}u^N = \\
= \int \phi(u^N) \mu_{h^N, \Sigma_I}(\mathrm{d}u^N)
\end{aligned}
\tag{17}
$$

from which we see that we have obtained the expectation of ϕ w.r.t. a multivariate Gaussian with mean and covariance given by:

$$
h^N := \Sigma_I \Sigma_N^{-1} P A B^{-1} K_{\mathcal{I}}^{-1} \bar{F}
\tag{18}
$$

$$
\Sigma_{\mathcal{I}} := P(\Sigma + A K_{\mathcal{I}} A^*) P^*
\tag{19}
$$

Thus, we conclude that "averaging" over f gives the following Gaussian posterior:

$$
N\left((\Sigma + A K_{\mathcal{I}} A^*) P^* \Sigma_N^{-1} P A B^{-1} K_{\mathcal{I}}^{-1} \bar{F}, \Sigma + A K_{\mathcal{I}} A^*\right)
\tag{20}
$$

2 Notes 2

We will now start focusing on a specific PDE. We will consider here the standard elliptic stationary conductivity problem:

$$
\begin{aligned}
\mathcal{L}u &:= -\nabla \cdot (a(x) \nabla u(x)) = f(x), \quad x \in \Omega \\
u &= 0, \quad x \in \partial\Omega
\end{aligned}
\tag{21}
$$

where f is stochastic. In this case, standard energy estimates will give that $\mathcal{L}$ is a bounded operator from H^1 to H^{-1} (!!!!CHECK!!!!!). Thus, we seek a solution $u \in H_0^1(\Omega)$.

Once we assume this, we can be more precise about the choice of Gaussian prior on u. In particular, we want, at the very least, that the Cameron-Martin space of $\mathcal{N}(0, V)$ lies in H^1. Let $\mathcal{A} = -\Delta$ (with Dirichlet or Neumann boundary conditions). $\mathcal{A}$ satisfies Assumptions 2.9(i)-(iii) of [2]. Further, on page 474 of [2] it states that for such a choice of $\mathcal{A}$ the spaces $\mathcal{H}^s$ in (2.29) are contained in the usual Sobolev spaces H^s. Thus, provided we choose $\alpha > 1 + d/2$, Lemma 6.27 of [2] gives us that $u \sim \mathcal{N}(0, V)$, where $V = \mathcal{A}^{-\alpha}$, is in $\mathcal{H}^s$ almost surely for any $s \in [0, \alpha - d/2)$. In particular, since we have assumed $\alpha > 1 + d/2$ we have that $\alpha - d/2 > 1$ and so $u \sim \mathcal{N}(0, V)$ is in $\mathcal{H}^1$ almost surely. This gives us that u is almost surely in H^1 since $\mathcal{H}^1 \subset H^1$ for our choice of $\mathcal{A}$.

We will now connect this to FEM. We start by introducing a choice of information operators. These operators must be consistent with the spaces we are working with. In particular, I_j should lie in $\mathcal{U}'' = U$. i.e. I_j maps H^{-1} to $\mathbb{R}$. Given a finite element basis $\{\psi_i\}$, we can now choose to define I_j as:

$$I_j g := \langle g, \psi_j \rangle \tag{22}$$

where $\langle , \rangle$ denotes thes $H^{-1}(\Omega), H_0^1(\Omega)$ duality pairing. This can be identified with an inner product on L^2 using the Riesz Representation theorem. It can be shown that for this choice we have:

$$I_j \mathcal{L} u = \int_\Omega a(x) \nabla u(x) \cdot \nabla \psi_j(x) \mathrm{d}x \tag{23}$$

Next we make a particular choice of prior covariance V (*note: this is not related to the choice discussed above and is just to relate the formulation to FEM*). We define $V : H^{-1} \to H^1$ as follows:

$$V u(x) := \sum_{i=1}^J \lambda_i \psi(x) \int_\Omega \psi_i(y) u(y) \mathrm{d}y \tag{24}$$

In order to proceed it will prove helpful to express V in a different form as follows. We note that we can write:

$$V u(x) = \Phi(x)^* \Lambda \int_\Omega \Phi(y) u(y) \mathrm{d}y$$

where $\Phi(x) := (\psi_1(x), \ldots, \psi_J(x))^T$ is a vector in $\mathbb{R}^J$ and $\Lambda := \mathrm{diag}\{\lambda_i\}_{i=1}^J$ is a $J \times J$ matrix. We now introduce the following operators, $\{T_i\}_{i=1}^J$ defined by:

$$T_i u := \int_\Omega \psi_i(y) u(y) \mathrm{d}y$$

If we now let $T := (T_1, \ldots, T_J)^T$ we have:

$$V u(x) = \Phi(x)^* \Lambda T u$$

Thus,

$$V = \Phi^* \Lambda T$$

We note that $\Lambda T u$ is a vector in R^J which gives the coefficients of the expansion of the function $V u$ in terms of the finite element basis. Noting this will enable us to make sense what happens to (5) and (6). Before we go through this calculation we first introduce the following operators: $G_j := I_j \mathcal{L} : H^1 \to \mathbb{R}$ for $j = 1, \ldots, J$. We also let $G = (G_1, \ldots, G_J)^T$ so that $G = \mathcal{I}\mathcal{L}$. Looking at the mean we have:

$$a = V\mathcal{I}^* \mathcal{L}^* (\mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^*)^{-1} F$$
$$= V G^* (GVG^*)^{-1} F$$

Now, G is a map from H^1 to $\mathbb{R}^J$ and so G^* is a map from $\mathbb{R}^J$ to H^{-1}. Thus, GVG^* is a map from $\mathbb{R}^J$ to $\mathbb{R}^J$, i.e. a $J \times J$ matrix. We first compute this matrix. We have $GV = G\Phi^*\Lambda T$. We observe that:

$$G\Phi^* = \begin{pmatrix} G_1 \\ \vdots \\ G_J \end{pmatrix} \begin{pmatrix} \psi_1 & \ldots & \psi_J \end{pmatrix}$$

is a $J \times J$ matrix. The $ijth$ entry of this matrix is:

$$\begin{aligned} (G\Phi^*)_{ij} &= G_i \psi_j \\ &= I_i \mathcal{L} \psi_j \\ &= \int_\Omega a(x) \nabla \psi_i(x) \cdot \nabla \psi_j(x) \mathrm{d}x \\ &=: M_{ij} \end{aligned}$$

This matrix M is the standard Galerkin matrix which appears in the finite element method. Continuing we thus have,

$$GV = G\Phi^*\Lambda T = M\Lambda T$$
$$\implies GVG^* = M\Lambda TG^*$$

To figure out TG^* we express it as $TG^* = (GT^*)^*$ where the adjoint of T is $T^* = (T_1^*, \ldots, T_J^*)$. To figure out the adjoints T_i^* we consider for $\alpha \in \mathbb{R}$ and for any function u the following:

$$\alpha T_i u = \alpha \int_\Omega \psi_i(y)u(y)\mathrm{d}y$$
$$= \int_\Omega \alpha\psi_i(y)u(y)\mathrm{d}y$$

which suggests that $T_i^*\alpha = \alpha\psi_i$. I.e. T_i^* is the operator which maps a real number α to the function $\alpha\psi_i$. Thus,

$$GT^*\alpha = \begin{pmatrix} G_1 \\ \vdots \\ G_J \end{pmatrix} \begin{pmatrix} T_1^* & \ldots & T_J^* \end{pmatrix} \alpha$$

and this has $ijth$ entry given by:

$$(GT^*\alpha)_{ij} = G_i T_j^*\alpha$$
$$= G_i(\alpha\psi_j)$$
$$= \alpha G_i(\psi_j)$$
$$= \alpha M_{ij}$$

which gives us that $GT^* = M$. Thus, $TG^* = (GT^*)^* = M^*$. We thus have that $GVG^* = M\Lambda TG^* = M\Lambda M^*$ and so we can now finish the simplification of a:

$$a = VG^*(GVG^*)^{-1}F$$
$$= VG^*(M\Lambda M^*)^{-1}F$$
$$= \Phi^*\Lambda TG^*(M\Lambda M^*)^{-1}F$$
$$= \Phi^*\Lambda M^*(M\Lambda M^*)^{-1}F$$

Now using standard finite element arguments, one can identify conditions under which the matrix M is non-singular. We will expand on these conditions in the next section. Under the assumption that these conditions hold everything drastically simplifies and we obtain:

$$a = \Phi^*\Lambda M^*(M\Lambda M^*)^{-1}F$$
$$= \Phi^*\Lambda M^*(M^*)^{-1}\Lambda^{-1}M^{-1}F$$
$$= \Phi^*M^{-1}F$$

We thus see that when expressed in terms of the finite element basis the mean function has coefficients given by the vector $\hat{a} := M^{-1}F$ just like FEM gives. We now turn to the covariance (6) and compute:

$$\Sigma = V - V\mathcal{L}^*\mathcal{I}^*(\mathcal{I}\mathcal{L}V\mathcal{L}^*\mathcal{I}^*)^{-1}\mathcal{I}\mathcal{L}V$$
$$= V - VG^*(GVG^*)^{-1}GV$$
$$= V - \Phi^*\Lambda TG^*(M\Lambda M^*)^{-1}GV$$
$$= V - \Phi^*\Lambda M^*(M^*)^{-1}\Lambda^{-1}M^{-1}GV$$
$$= V - \Phi^*M^{-1}GV$$
$$= V - \Phi^*M^{-1}M\Lambda T$$
$$= V - \Phi^*\Lambda T$$
$$= V - V = 0$$

We can thus see that when we choose our information operators according to (22) and our prior covariance according to (24) that the posterior (for a fixed realisation of F) collapses to a point measure located at the FEM solution to the PDE.

We now move on to average over the distribution of f. We note that we cannot just utilize (20) as in this expression we require the inverse of Σ_N which under our assumptions is $\Sigma_N = P\Sigma P^* = 0$ as $\Sigma = 0$. We must instead redo the calculation of the expectation of an arbitrary bounded cylindrical test function which we performed in section 1. As before we are interested in computing:

$$\int \int \phi(u^N)\mu_{a,\Sigma}(\mathrm{d}u)\mu_{\bar{f},K}(\mathrm{d}f) \tag{25}$$

where now $a = \Phi^* M^{-1} F$ and $\Sigma = 0$. Thus, $\mu_{a,\Sigma} = \delta_a$, that is to say that the measure is in fact a Dirac point mass at $a = \Phi^* M^{-1} F$. We thus have:

$$\int \int \phi(u^N)\mu_{a,\Sigma}(\mathrm{d}u)\mu_{\bar{f},K}(\mathrm{d}f) = \int \int \phi(u^N)\delta_{\Phi^* M^{-1}F}(\mathrm{d}u)\mu_{\bar{f},K}(\mathrm{d}f)$$

Now since the posterior for u for a fixed realisation of f is a point mass we have that the posterior for $u^N = Pu$ is $\delta_{P\Phi^* M^{-1}F}$. Thus, we can write:

$$\int \int \phi(u^N)\delta_{\Phi^* M^{-1}F}(\mathrm{d}u)\mu_{\bar{f},K}(\mathrm{d}f) = \int \int \phi(u^N)\delta_{P\Phi^* M^{-1}F}(\mathrm{d}u^N)\mu_{\bar{f},K}(\mathrm{d}f)$$

It is now important to point out that $P\Phi^*$ is to be interpreted[1] as the following $N \times J$ matrix:

$$\begin{aligned}
P\Phi^* &= P(\psi_1, \ldots, \psi_J) \\
&= (P\psi_1, \ldots, P\psi_J) \\
&= \begin{pmatrix} \psi_1(x_1) & \ldots & \psi_J(x_1) \\ \vdots & & \vdots \\ \psi_1(x_N) & \ldots & \psi_J(x_N) \end{pmatrix} =: \Psi
\end{aligned}$$

i.e. the matrix with ijth entry $\Psi_{i,j} := \psi_j(x_i)$ for $i \in \{1, \ldots, N\}, j \in \{1, \ldots, J\}$. We thus have that $Y := P\Phi^* M^{-1} F = \Psi M^{-1} F \sim \mathcal{N}(\Psi M^{-1}\bar{F}, \Psi M^{-1} K_{\mathcal{I}} M^{-1} \Psi^*)$. Together with this and with the fact that the posterior of u^N only depends on f through F we can write:

$$\begin{aligned}
\int \int \phi(u^N)\delta_{P\Phi^* M^{-1}F}(\mathrm{d}u^N)\mu_{\bar{f},K}(\mathrm{d}f) &= \int \int \phi(u^N)\delta_{\Psi M^{-1}F}(\mathrm{d}u^N)\mu_{\bar{F},K_{\mathcal{I}}}(\mathrm{d}F) \\
&= \int \int \phi(u^N)\delta_Y(\mathrm{d}u^N)\mu_{\Psi M^{-1}\bar{F}, \Psi M^{-1}K_{\mathcal{I}}M^{-1}\Psi^*}(\mathrm{d}Y) \\
&= \int \phi(u^N)\left(\int \delta_Y(\mathrm{d}u^N)\mu_{\Psi M^{-1}\bar{F}, \Psi M^{-1}K_{\mathcal{I}}M^{-1}\Psi^*}(\mathrm{d}Y)\right) \\
&= \int \phi(u^N)\mu_{\Psi M^{-1}\bar{F}, \Psi M^{-1}K_{\mathcal{I}}M^{-1}\Psi^*}(\mathrm{d}u^N)
\end{aligned}$$

which we recognize as a finite dimensional projection of a Gaussian measure. Thus, we conclude that "averaging" over f gives the following Gaussian posterior under these choices:

$$\mathcal{N}(\Phi^* M^{-1}\bar{F}, \Phi^* M^{-1} K_{\mathcal{I}} M^{-1}\Phi) \tag{26}$$

This Gaussian measure directly links to what is found at the end of section 2 in the Statistical FEM paper.

3   Notes 3

We now provide details for some of the claims in Notes 1 and 2 above.

To start we provide justification for why it is valid to condition as we did in (4). The justification for this is essentially given by Theorem 3.3 in [3]. To apply this theorem to our problem we note

---

[1] By this we mean that $P\Phi^*$ is actually interpreted as $P \otimes \Phi^*$

that our pair $(u, \mathcal{ILL}u)^T$ lies in the orthogonal direct sum $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ of the separable Hilbert spaces $\mathcal{H}_1 := H^1(\Omega) \times \{0\}$ and $\mathcal{H}_2 := \{0\} \times \mathbb{R}^J$. The corresponding Gaussian measure we are dealing with on this $\mathcal{H}$ is given by (3). Theorem 3.3 states that the conditional distribution is a Gaussian measure with covariance operator being the short of the covariance operator of the distribution (3) to $\mathcal{H}_2$. This operator,

$$\begin{pmatrix} V & V\mathcal{L}^*\mathcal{I}^* \\ \mathcal{ILV} & \mathcal{ILV}\mathcal{L}^*\mathcal{I}^* \end{pmatrix} \tag{27}$$

is already written in its $(\mathcal{H}_1, \mathcal{H}_2)$ partition representation. In the section on shorted operators in [3] it is stated that provided the lower right partition of the covariance operator of (3) corresponding to the covariance in $\mathcal{H}_2$, i.e. $\mathcal{ILV}(\mathcal{IL})^*$ is invertible then the shorted operator is given by:

$$\begin{pmatrix} V - V\mathcal{L}^*\mathcal{I}^*(\mathcal{ILV}\mathcal{L}^*\mathcal{I}^*)^{-1}\mathcal{ILV} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \tag{28}$$

We assume the invertiblity of this partition. We do note however that this part of the partition reduces to the $J \times J$ matrix $M\Lambda M^*$ when we choose the FEM prior and FEM information operator. The invertibility of this matrix is then equivalent to the invertibility of the Galerkin Matrix, M (if we assume that the entries of Λ are non-zero).

We thus see that the conditional distribution (4) has the correct covariance operator.

Theorem 3.3 also goes on to give the mean of the conditional distribution in the case that the covariance operator of the joint distribution is compatible with $\mathcal{H}_2$. For our problem this is in fact the case as we now show. We first denote the covariance operator of the joint distribution by:

$$\begin{pmatrix} V & V\mathcal{L}^*\mathcal{I}^* \\ \mathcal{ILV} & \mathcal{ILV}\mathcal{L}^*\mathcal{I}^* \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \tag{29}$$

i.e. $C_{11} = V, C_{12} = V\mathcal{L}^*\mathcal{I}^*, C_{21} = \mathcal{ILV}, C_{22} = \mathcal{ILV}\mathcal{L}^*\mathcal{I}^*$. If we now define the following bounded operator $Q : \mathcal{H} \to \mathcal{H}$ by

$$Q = \begin{pmatrix} 0 & 0 \\ C_{22}^{-1}C_{21} & I \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \hat{Q} & I \end{pmatrix} \tag{30}$$

where $\hat{Q} := C_{22}^{-1}C_{21} : \mathcal{H}_1 \to \mathcal{H}_2$ we can easily check that such an operator Q is an (C-symmetric) oblique projection onto $\mathcal{H}_2$. Since Q is not 0 we have by Theorem 3.3 that the mean of the conditional distribution (4) is given by $\hat{Q}^*F$ where the adjoint is defined by the relation $\left\langle \hat{Q}^*h_2, h_1 \right\rangle_{\mathcal{H}_1} = \left\langle h_2, \hat{Q}h_1 \right\rangle_{\mathcal{H}_2}$ for all $h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2$. Since we have that C_{22} and hence C_{22}^{-1} is self-adjoint and further that $C_{12}^* = C_{21}$ we can easily deduce that $\hat{Q}^* = C_{12}C_{22}^{-1}$ and so the mean is given by

$$\hat{Q}^*F = C_{12}C_{22}^{-1}C_{21}F = V\mathcal{L}^*\mathcal{I}^*(\mathcal{ILV}\mathcal{L}^*\mathcal{I}^*)^{-1}F = a \tag{31}$$

where a is what we had in (5).

We can now move on to provide conditions for the Galerkin matrix M defined in Notes 2 to be invertible. A sufficient condition to ensure this invertibility is the following assumption (see Assumption 2.34 in [4]):

Assumption (regularity of coefficients) The diffusion coefficient $a(x)$ satisfies:

$$0 < a_{\min} \le a(x) \le a_{\max} < \infty \text{ for almost all } x \in \Omega \tag{32}$$

for some real constants $a_{\min}, a_{\max}$. In particular $a \in L^\infty(\Omega)$.

If we assume that this assumption holds then it is a simple exercise to show that the following bilinear form from $H_0^1(\Omega) \times H_0^1(\Omega)$ to $\mathbb{R}$ defined by:

$$\tilde{a}(u, w) := \int_\Omega a(x)\nabla u(x) \cdot \nabla w(x)\mathrm{d}x \tag{33}$$

8

defines a norm $|\cdot|_E$ on $H_0^1(\Omega)$ via $|u|_E := \tilde{a}(u,u)^{1/2}$. To show that the Galerkin matrix M is invertible we will show that it is (strictly) positive definite. To this end let $\mathbf{v} \in \mathbb{R}^J \backslash \{0\}$ and define $v = \sum_{i=1}^J v_i \psi_i$ which is a function in $H_0^1(\Omega)$. We now have:

$$
\begin{aligned}
\mathbf{v}^T(M)\mathbf{v} &= \sum_{i,j=1}^J v_j(M_{ij})v_i \\
&= \sum_{i,j=1}^J v_j \left(\int_\Omega a(x) \nabla \psi_i(x) \cdot \nabla \psi_j(x) \mathrm{d}x \right) v_i \\
&= \sum_{i,j=1}^J v_j \tilde{a}(\psi_i, \psi_j) v_i \\
&= \tilde{a} \left(\sum_{i=1}^J v_i \psi_i, \sum_{j=1}^J v_j \psi_j \right) \\
&= \tilde{a}(v,v) \\
&= |v|_E^2 > 0
\end{aligned}
$$

where the strict inequality follows since $\tilde{a}(v,v) = 0$ iff $|v|_E = 0$ iff $v = 0$ (since $|\cdot|_E$ is a norm on $H_0^1(\Omega)$) iff $\mathbf{v} = 0$ since the $\{\psi_i\}$ are linearly independent.

4   Notes 4

In Notes 1 the true joint distribution of (u, f) given by (2) gives us that the "true" distribution of the solution u to the problem (21) is $\nu_1 = \mathcal{N}(\mathcal{L}^{-1}\bar{f}, \mathcal{L}^{-1}K(\mathcal{L}^{-1})^*)$. Meanwhile, in Notes 2 we worked out that "averaged" posterior for u (for the FEM prior) is $\nu_2 = \mathcal{N}(\Phi^*M^{-1}\bar{F}, \Phi^*M^{-1}K_\mathcal{I}M^{-1}\Phi)$ (see (26)). In this section we will consider quantifying how close these two distributions ν_1, ν_2 are. This will be achieved by obtaining an upperbound for the Wasserstein distance between ν_1, ν_2 using the connection between the Wasserstein distance between Gaussian measures and the Procrustes Metric on covariance operators (see [5]). We start by first giving the definition of the Wasserstein distance between two probability measures μ, ν on $\mathcal{H}$, $W(\mu, \nu)$:

$$
W^2(\mu, \nu) = \inf_{\pi \in \Gamma(\mu, \nu)} \int_{\mathcal{H} \times \mathcal{H}} \|x - y\|^2 \mathrm{d}\pi(x, y)
$$

where $\Gamma(\mu, \nu)$ is the set of couplings of μ and ν, i.e:

$\Gamma(\mu, \nu) = \{$Borel probability measures π on $\mathcal{H} \times \mathcal{H} | \pi(E \times \mathcal{H}) = \mu(E)$ and $\pi(\mathcal{H} \times F) = \nu(F)$ for all Borel $E, F \subset \mathcal{H}\}$

When μ and ν are both Gaussian measures an explicit expression can be obtained for the Wasserstein distance. Suppose $\mu = \mathcal{N}(m_1, \Sigma_1)$ and $\nu = \mathcal{N}(m_2, \Sigma_2)$. One has (see [5]),

$$
W^2(\mu, \nu) = \|m_1 - m_2\|_\mathcal{H}^2 + \mathrm{tr}\,(\Sigma_1) + \mathrm{tr}\,(\Sigma_2) - 2\,\mathrm{tr}\,\sqrt{\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2}}
$$

This formula is true in both the finite and infinite dimensional cases. The term $\mathrm{tr}\,\sqrt{\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2}}$ is difficult to analyse in our situation. As such we will make use of Proposition 3 from [5] which states:

Proposition 4.1. The Procrustes distance between two trace-class operators Σ_1 and Σ_2 on $\mathcal{H}$ coincides with the Wasserstein distance between two second-order Gaussian processes $\mathcal{N}(0, \Sigma_1)$ and $\mathcal{N}(0, \Sigma_2)$ on $\mathcal{H}$,

$$
\Pi(\Sigma_1, \Sigma_2) := \inf_{R: R^*R = I} \|\Sigma_1^{1/2} - R\Sigma_2^{1/2}\|_2 = W(\mathcal{N}(0, \Sigma_1), \mathcal{N}(0, \Sigma_2))
$$

where $\|\cdot\|_2$ is the Hilbert-Schmidt norm defined by $\|A\|_2 = \sqrt{\mathrm{tr}(A^*A)}$.

Using this result one can obtain a simple upperbound on the Wasserstein distance by choosing $R = I$ in the infimum:

$$W(\mathcal{N}(0, \Sigma_1), \mathcal{N}(0, \Sigma_2)) = \inf_{R: R^* R = I} \|\Sigma_1^{1/2} - R\Sigma_2^{1/2}\|_2 \leq \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_2$$

Now since in our case we have un-centered Gaussian measures ν_1, ν_2 we must first link the Wasserstein distance between ν_1, ν_2 to the Wasserstein distance of the centred measures ν_1^*, ν_2^* using a general result mentioned in [6]:

$$W^2(\nu_1, \nu_2) = \|m_1 - m_2\|^2 + W^2(\nu_1^*, \nu_2^*)$$

Note: here m_1, m_2 are the means of ν_1, ν_2 respectively.

For our particular case, we will consider the underlying Hilbert space to be $\mathcal{H} = L^2(\Omega)$. As such the norm for the difference in means is the L^2 norm However, note that the solution u to our boundary value problem lies in $H_0^1(\Omega)$, a subspace of $L^2(\Omega)$. Since we are interested in seeking an upper bound on the Wasserstein distance we can focus on the $|\cdot|_{H_0^1(\Omega)}$ norm[2] on $H_0^1(\Omega)$: $|m_1 - m_2|_{H^1(\Omega)}^2 = \|\nabla(m_1 - m_2)\|_{L^2(\Omega)}^2$. Since this norm is equivalent to the usual $H^1(\Omega)$ norm on $H_0^1(\Omega)$, which is an upper bound for the L^2 norm, we can go about finding an upper bound for the $|\cdot|_{H_0^1(\Omega)}$ norm. If we now denote the covariance operators of ν_1, ν_2 as Σ_1, Σ_2 respectively we have:

$$W^2(\nu_1, \nu_2) = \|m_1 - m_2\|_{L^2(\Omega)}^2 + W^2(\mathcal{N}(0, \Sigma_1), \mathcal{N}(0, \Sigma_2))$$

We now go about obtaining an upperbound on each of these two terms. We want to control each term by the FEM mesh size h (which is inversely proportional to the number of finite elements J). For brevity we will assume now that $\Omega \subset \mathbb{R}^2$ so we have a 2 dimensional problem. The analysis follows in almost exactly the same way for $\mathbb{R}^d$. We will assume that Ω is a convex polygonal domain (the polygonal assumption can easily be relaxed). The convexity assumption gives us that the $H^2(\Omega)$ norm of the variational solution of our PDE is controlled by the L^2 norm of the RHS. We now take our FEM mesh to be a triangulation of Ω with h being the maximum side length of any triangle in the triangulation. We require a further technical assumption that the meshes we consider remain regular in the sense that as we refine the mesh by decreasing h to zero the angles of all the triangles are bounded below independently of h. Since $m_1 = \mathcal{L}^{-1}\bar{f}$ is the solution to the following elliptic boundary value problem:

$$-\nabla \cdot (a(x)\nabla v(x)) = \bar{f}(x), \quad x \in \Omega$$
$$v = 0, \quad x \in \partial\Omega$$

and since $m_2 = \Phi^* M^{-1} \bar{F} = \Phi^* M^{-1} \mathcal{I}\bar{f}$ is the FEM solution to the variational formulation of the above problem the error analysis of FEM transfers over to allow us to bound the norm of the difference of the means as follows:

$$\|m_1 - m_2\|_{L^2(\Omega)} \leq \|m_1 - m_2\|_{H^1(\Omega)} \leq C|m_1 - m_2|_{H^1(\Omega)} \leq \tilde{C}h\|\bar{f}\|_{L^2(\Omega)} \tag{34}$$

for some constant $C, \tilde{C} > 0$ (!!ask about which value to use for this!!) The assumptions and error analysis is taken from Chapter 5 of [7].

We now move on to getting an upperbound for the second term. Using the link with the Procrustes distance discussed above we have:

$$W^2(\mathcal{N}(0, \Sigma_1), \mathcal{N}(0, \Sigma_2)) \leq \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_2^2$$

The RHS of the above is still difficult to deal with so we make use of Lemma 4.1 from [8] to obtain:

$$W^2(\mathcal{N}(0, \Sigma_1), \mathcal{N}(0, \Sigma_2)) \leq \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_2^2 \leq \|\Sigma_1 - \Sigma_2\|_1$$

where $\|\cdot\|_1$ is the trace norm or nuclear norm defined by $\|A\|_1 = \text{tr}(\sqrt{A^*A})$.

---

[2]This is a semi-norm on $H^1(\Omega)$.

We now investigate this term:

$$\|\Sigma_1 - \Sigma_2\|_1 = \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}K_{\mathcal{I}}M^{-1}\Phi\|_1$$
$$= \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K\mathcal{I}^*M^{-1}\Phi\|_1$$
$$= \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\Phi^*M^{-1}\mathcal{I})^*\|_1$$

where we have used the defintion of $K_{\mathcal{I}}$ and the fact that M and hence M^{-1} is self-adjoint. From this simplification we can see that we can control how "close" the two covariance operators are because we can control how "close" $\Phi^*M^{-1}\mathcal{I}$ is to $\mathcal{L}^{-1}$. To be more precise:

$$\|\Sigma_1 - \Sigma_2\|_1 = \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\Phi^*M^{-1}\mathcal{I})^*\|_1$$
$$= \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\mathcal{L}^{-1})^* + \Phi^*M^{-1}\mathcal{I}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\Phi^*M^{-1}\mathcal{I})^*\|_1$$
$$\leq \|\mathcal{L}^{-1}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\mathcal{L}^{-1})^*\|_1 + \|\Phi^*M^{-1}\mathcal{I}K(\mathcal{L}^{-1})^* - \Phi^*M^{-1}\mathcal{I}K(\Phi^*M^{-1}\mathcal{I})^*\|_1$$
$$= \|(\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I})K(\mathcal{L}^{-1})^*\|_1 + \|\Phi^*M^{-1}\mathcal{I}K(\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I})^*\|_1$$
$$\leq \|\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I}\|_\infty\|K\mathcal{L}^{-1}\|_1 + \|\Phi^*M^{-1}\mathcal{I}K\|_1\|(\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I})^*\|_\infty$$
$$\leq \|\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I}\|_\infty(\|K\mathcal{L}^{-1}\|_1 + \|\Phi^*M^{-1}\mathcal{I}\|_1\|K\|_1)$$

where we have utilized Holder's inequality and the sub-multiplicativity of Schatten-p norms (*note: the $\|\cdot\|_\infty$ is the operator norm here and is a special case of the Schatten-p norm for p being infinity*). As mentioned, we can control how "close" $\Phi^*M^{-1}\mathcal{I}$ is to $\mathcal{L}^{-1}$. To do this we note that (34) holds for all $\bar{f}$ in $L^2(\Omega)$, i.e. we have:

$$|m_1 - m_2|_{H^1(\Omega)} = |(\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I})\bar{f}|_{H^1(\Omega)} \leq C_1h\|\bar{f}\|_{L^2(\Omega)} \quad \forall \bar{f} \in L^2(\Omega)$$

where $C_1 := \tilde{C}/C > 0$.

This implies that we can bound the operator norm of $\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I}$ by C_1h, i.e.,

$$\|\mathcal{L}^{-1} - \Phi^*M^{-1}\mathcal{I}\|_\infty \leq C_1h \tag{35}$$

Utilising this upper bound together with the fact that $\|\Phi^*M^{-1}\mathcal{I}\|_1$ is bounded (!!do I need to expand on this!!) we have:

$$W^2(\mathcal{N}(0,\Sigma_1), \mathcal{N}(0,\Sigma_2)) \leq \|\Sigma_1 - \Sigma_2\|_1 \leq C_1h\|K\mathcal{L}^{-1}\|_1 \tag{36}$$

Combining (34) and (36) we now have:

$$W^2(\nu_1,\nu_2) = \|m_1 - m_2\|_{L^2(\Omega)}^2 + W^2(\mathcal{N}(0,\Sigma_1), \mathcal{N}(0,\Sigma_2))$$
$$\leq \tilde{C}^2h^2\|\bar{f}\|_{L^2(\Omega)}^2 + C_1h\|K\mathcal{L}^{-1}\|_1$$

We thus have an upper bound on the Wasserstein distance between ν_1, ν_2 in terms of the FEM mesh size h:

$$W(\nu_1,\nu_2) \leq h^{1/2}\sqrt{C_1\|K\mathcal{L}^{-1}\|_1 + \tilde{C}^2h^2\|\bar{f}\|_{L^2(\Omega)}^2} \leq \gamma h^{1/2} \tag{37}$$

where $\gamma > 0$ is a constant.

5   Notes 5

We now have two distributions for the solution u to our boundary value problem, $\nu_i = \mathcal{N}(m_i, \Sigma_i)$ for $i = 1, 2$. We now assume we also have sensor data which give us noisy observations of the value of u at some points $y_1, \ldots, y_s$ in Ω. We now want to update our belief in the distribution of u using this sensor data. We can use ν_i as prior distributions. Our goal will be to investigate how different the two resulting posteriors are.

Let $S : \mathcal{H} \to \mathbb{R}^s$ be the operator which maps a function $h \in \mathcal{H}$ to $(h(y_1), \ldots, h(y_s))^T$ in $\mathbb{R}^s$. Assuming that the sensors make observations at each y_j with a normally distributed error we have that the likelihood for our sensor readings, $\mathbf{v}$, is:

$$\mathbf{v}|u \sim \mathcal{N}(Su, \epsilon^2 I) \tag{38}$$

The goal is to now find the posterior $u|\mathbf{v}$ when our prior is $u \sim \nu_i \equiv \mathcal{N}(m_i, \Sigma_i)$, $i = 1, 2$. In order to do so we will first find the joint distribution of $(u, \mathbf{v})$. To do this we note our assumption that $\mathbf{v}|u$ is distributed according to (38) is equivalent to the assertion that:

$$\mathbf{v} = Su + \boldsymbol{\delta}$$

where $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \epsilon^2 I)$ is independent of u. From this we can see that the joint distribution of $(u, \boldsymbol{\delta})$ is:

$$\begin{pmatrix} u \\ \boldsymbol{\delta} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} m \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \epsilon^2 I \end{pmatrix} \right)$$

We now note that $(u, \mathbf{v})$ can be expressed as:

$$\begin{pmatrix} u \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} I & 0 \\ S & I \end{pmatrix} \begin{pmatrix} u \\ \boldsymbol{\delta} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}$$

Since this is just a linear transformation of $(u, \boldsymbol{\delta})$ we have that the joint distribution of $(u, \mathbf{v})$ is given by:

$$\begin{pmatrix} u \\ \mathbf{v} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} m \\ Sm \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma S^* \\ S\Sigma & \epsilon^2 I + S\Sigma S^* \end{pmatrix} \right) \tag{39}$$

From this it is simply a matter of conditioning to obtain the posterior distribution of $u|\mathbf{v}$. We thus obtain:

$$u|\mathbf{v} \sim \mathcal{N}(m_{u|\mathbf{v}}^{(i)}, \Sigma_{u|\mathbf{v}}^{(i)}) \tag{40}$$

where,

$$m_{u|\mathbf{v}}^{(i)} := m_i + \Sigma_i S^* (\epsilon^2 I + S\Sigma_i S^*)^{-1} (\mathbf{v} - Sm_i) \tag{41}$$

$$\Sigma_{u|\mathbf{v}}^{(i)} := \Sigma_i - \Sigma_i S^* (\epsilon^2 I + S\Sigma_i S^*)^{-1} S\Sigma_i \tag{42}$$

Since the two posteriors are still Gaussian we can quantify how differnty they are by using the Wasserstein Distance again. We have,

$$W^2 \left(\mathcal{N}(m_{u|\mathbf{v}}^{(1)}, \Sigma_{u|\mathbf{v}}^{(1)}), \mathcal{N}(m_{u|\mathbf{v}}^{(2)}, \Sigma_{u|\mathbf{v}}^{(2)}) \right) = \|m_{u|\mathbf{v}}^{(1)} - m_{u|\mathbf{v}}^{(2)}\|_{L^2(\Omega)}^2 + W^2 \left(\mathcal{N}(0, \Sigma_{u|\mathbf{v}}^{(1)}), \mathcal{N}(0, \Sigma_{u|\mathbf{v}}^{(2)}) \right)$$

$$\leq \|m_{u|\mathbf{v}}^{(1)} - m_{u|\mathbf{v}}^{(2)}\|_{L^2(\Omega)}^2 + \|\Sigma_{u|\mathbf{v}}^{(1)} - \Sigma_{u|\mathbf{v}}^{(2)}\|_1$$

We now proceed to obtain an upper-bound on these terms. For convenience we will use the notation $B_i := \epsilon^2 I + S\Sigma_i S^*$ for $i = 1, 2$. It will be helpful to first work out an upper-bound on the operator norm of the difference $B_1^{-1} - B_2^{-1}$. To this end we first compute:

$$\|B_1 - B_2\|_\infty = \|S(\Sigma_1 - \Sigma_2)S^*\|_\infty$$
$$\leq \|S\|_\infty^2 \|\Sigma_1 - \Sigma_2\|_\infty$$
$$\leq \|\Sigma_1 - \Sigma_2\|_1$$
$$\leq C_1 h \|K\mathcal{L}^{-1}\|_1$$

where we have used the fact that $\|S\|_\infty \leq 1$. Since B_1 (and B_2) is invertible and for sufficiently small h we have that $\|B_1 - B_2\|_\infty < 1/\|B_1^{-1}\|_\infty$

6 Questions

Question 1: I still do not see how defining I_j by $I_j := \langle g, \psi_j \rangle$ in (22) implies that $I_j \mathcal{L} u = -\int_\Omega a \nabla u \cdot \nabla \psi_j \mathrm{d}x$. The reason I am confused is because of the fact that I_j is a map from $H^{-1}(\Omega)$ to $\mathbb{R}$. I do see that if I assume that I_j acts on functions g via:

$$I_j g = \int_\Omega \psi_j(x) g(x) \mathrm{d}x$$

then for $g = \mathcal{L} u$ I would have:

$$
\begin{aligned}
I_j \mathcal{L} u &= \int_\Omega \psi_j(x) \nabla \cdot (a(x) \nabla u(x)) \mathrm{d}x \\
&= -\int_\Omega a(x) \nabla u(x) \cdot \nabla \psi_j(x) \mathrm{d}x + \int_{\partial\Omega} \psi_j(x) a(x) \nabla u(x) \cdot \mathbf{n} \mathrm{d}S \\
&= -\int_\Omega a(x) \nabla u(x) \cdot \nabla \psi_j(x) \mathrm{d}x
\end{aligned}
$$

since the second integral is 0 as ψ_j is zero on the boundary $\partial\Omega$. This calculation however only makes sense provided the functions g (and so $\mathcal{L}u$) is integrable (L^2 would let this work). I have tried to follow the argument using the Riesz Representation theorem in Evans in the section about H^{-1} and apply it to our problem here but I cannot seem to connect the two.

Question 2: In Notes 2 I defined the operators T_i by $T_i u := \int_\Omega \psi_i(y) u(y) \mathrm{d}y$. I then wanted to figure out the adjoint of this operator T_i^*. If I view T_i as an operator from $L^2(\Omega)$ to $\mathbb{R}$ then the adjoint is clearly given by $T_i^* \alpha = \alpha \psi_i$. My question is basically the following: when we introduced the FEM prior V it was as a map from H^{-1} to H^1. This suggest that T_i is also a map from $H^{-1} \to \mathbb{R}$. But then wouldn't the adjoint be a map from $\mathbb{R} \to H_0^1(\Omega)$ defined by:

$$\alpha T_i g = \langle T_i^* \alpha, g \rangle_{H_0^1, H^{-1}}$$

for all $g \in H^{-1}$ and for all $\alpha \in \mathbb{R}$? If this is the case I cannot seem to show that T_i^* is defined by $T_i^* \alpha = \alpha \psi_i$.

Question 3: Since the particular PDE we are working with has the boundary condition that $u = 0$ on $\partial\Omega$ does this mean that everywhere we mention the space H^1 we are talking about $H_0^1(\Omega)$? This is still a separable Hilbert space right?

References

[1] Alessandra Lunardi, Michele Miranda, and Diego Pallara. Infinite dimensional analysis. In *19th Internet Seminar*, volume 2016, 2015.

[2] Andrew M Stuart. Inverse problems: a bayesian perspective. *Acta numerica*, 19:451–559, 2010.

[3] Houman Owhadi and Clint Scovel. Conditioning gaussian measure on hilbert space. *arXiv preprint arXiv:1506.04208*, 2015.

[4] Gabriel J Lord, Catherine E Powell, and Tony Shardlow. *An introduction to computational stochastic PDEs*, volume 50. Cambridge University Press, 2014.

[5] Valentina Masarotto, Victor M Panaretos, and Yoav Zemel. Procrustes metrics on covariance operators and optimal transportation of gaussian processes. *Sankhya A*, 81(1):172–213, 2019.

[6] JA Cuesta-Albertos, C Matrán-Bea, and A Tuero-Diaz. On lower bounds for the l2-wasserstein metric in a hilbert space. *Journal of Theoretical Probability*, 9(2):263–283, 1996.

[7] Stig Larsson and Vidar Thomée. *Partial differential equations with numerical methods*, volume 45. Springer Science & Business Media, 2008.

[8] Robert T Powers and Erling Størmer. Free states of the canonical anticommutation relations. *Communications in Mathematical Physics*, 16(1):1–33, 1970.