Inverse Heat Equations

Yunsheng Lu

April 25, 2024

Abstract

This dissertation explores inverse problems with a focus on the heat equation, primarily in one-dimensional settings, though it also touches upon multi-dimensional aspects. It provides a comprehensive overview of deterministic methods in inverse problems, such as techniques based on Truncated Singular Value Decomposition (TSVD) for both real matrices and compact operators, and Tikhonov regularizations. The dissertation provides detailed and accessible proofs of those theories, which, while commonly used, frequently lack thorough documentation in the literature on inverse problems. We then introduce traditional statistical methods, including Bayesian approaches applied to recovery on finite many points. It also introduces novel methods for reconstructing initial heat distributions, notably using the singular value decomposition of compact operators for infinite-dimensional reconstructions, and leveraging Gaussian process priors to achieve smooth solutions. Despite the potential novelty of these contributions, they may echo existing studies, albeit developed independently. These methods represent a creative synthesis of traditional and contemporary techniques to tackle inverse problems, particularly related to heat equations.

Contents

1 Introduction

Inverse problems involve determining unknown properties or configurations of a system based on observed outcomes. Unlike direct problems, where the cause is known and the effect is to be determined, inverse problems work the other way around: the effects are known, and the goal is to deduce the causes. This distinction introduces unique mathematical challenges, particularly in terms of the problems' well-posedness—a concept defined by Hadamard to mean that a solution exists, is unique, and depends continuously on the input data. We will cover more details about this concept later in this section1.1. Inverse problems often violate at least one of these conditions, making them "ill-posed" and thus requiring specialized approaches to obtain stable solutions.

In the study of partial differential equations, inverse problems related to the heat equation represent a crucial area of exploration, bridging theoretical mathematics and practical applications. The heat equation, $\frac{\partial u}{\partial t} - \Delta u = f$, where u represents the temperature distribution over space and time, Δ denotes the Laplacian operator, and f is a source term, fundamentally describes the distribution and evolution of heat in a given medium. Inverse problems, by their nature, aim to determine unknown properties or initial conditions of the heat equation from observed data. This typically involves recovering the initial temperature distribution $u(x, 0)$ from later observations $u(x, T)$, or identifying the source term $f(x, t)$ that leads to a particular temperature distribution. In our discussion, we assume $f = 0$, that is, we assume that there are no external sources of heat within the domain. Thus, our discussion focuses on the recovery of initial heat distribution. We will introduce detailed setup for our problem later in 1.2.

These problems are inherently ill-posed, as small errors in measurement can lead to significant discrepancies in the solution, as we will see in 1.2. To address these problems, we will first introduce deterministic methods in Section 2. We will first develop the theory of Truncated Value Decomposition (TSVD) for both real matrices and compact operators 2.1, and then introduce Tikhonov regularization 2.2. In the end of this section, we explore the applications of these techniuqes in heat equations. While the TSVD and Tikhonov regularizations of real matrices provide recovery of initial conditions of on finitely many points, those of compact operators gives us infinite dimensional recovery. Next, we introduce statistical methods in Section 3. We first introduce Discretized Bayesian Approach with Gaussian prior 3.1, which is essentially recovering information on finitely many points. Later in Section 3.2, we will introduce Gaussian process, which is the extension of multivariate Gaussian distribution over infinitely many points, and we will assume Gaussian process to be our prior, which could potentially generate smooth recovery of the initial data, with proper choices of the covariance function. However, there are also a lot of choices for prior besides Gaussian depending on the specific case we study. For complicated distributions, it's hard to derive a closed formula for the posterior mean, which is usually our estimation, as later we will introduce. For this purpose, in Section 3.5, we include Markov Chain Monte Carlo (MCMC), which is a class of algorithm that can sample almost any target distributions and then we can use law of large number to compute the posterior mean.

This thesis tackles inverse problems related to the heat equation, highlighting not just specific solutions but also the broader applicability of the methods used. These techniques are grounded in strong mathematical and computational principles, making them useful in other scientific and engineering areas facing similar challenges. Therefore, the approaches developed here can enhance our ability to solve complex problems in various fields, providing valuable tools that can be adapted and used widely in various fields, such as medical imaging, or in environmental science, predicting soil temperatures based on surface measurements.

1.1 Inverse Problems

We can abstractly formulate most (if not all) inverse problems as finding a solution, $u \in \mathcal{U}$, to the operator equation

$$\mathcal{G}(u) = f \tag{1}$$

Here, $\mathcal{G}$ is called the forward operator; u is the image or parameter and $f \in \mathcal{F}$ are the measurements. The forward operator, $\mathcal{G} : \mathcal{U} \to \mathcal{F}$, is a model of the underlying (physical) process and simulates the measurements for given u. It could represent for example an integral operator, a matrix, or a set of algebraic equations. The image/parameters, u, constitute a mathematical description of the quantity of interest. It could for example be a function or a vector. The measurements, f, are a mathematical representation

of the measured data. For the purposes of analysis, we may think of this as a function, but in practice measurements are always a finite-dimensional quantity.

We briefly mentioned the concept of Hadamard's well-posedness at the beginning, and here is its specific formulation:

1. The problem has a solution.
2. The solution is unique.
3. The solution's behavior changes continuously with the initial conditions.

We will explore how inverse problems deviate from the aforementioned well-posedness— also known as 'ill-posedness'— through the subsequent examples.

1.2 One-Dimensional Inverse Heat Equations

Consider the heat conduction equations on the rod:

$$\frac{\partial^2 v(x,t)}{\partial x^2} - \frac{\partial v(x,t)}{\partial t} = 0$$

where $1 \leq x \leq 1$ and $t > 0$, with Dirichlet boundary conditions, i.e., $v(0,t) = v(1,t) = 0$ for any $0 \leq t \leq 1$. Denote the intial condition to be u, and it has the Fourier expansion:

$$u(x) = \sum_{n=1}^{\infty} c_n \sin n\pi x$$

By separation of variables, we have:

$$u(x,t) = \sum_{n=1}^{\infty} c_n e^{-(n\pi)^2 t} \sin n\pi x$$

Denote the true underlying heat distribution at $t = 1$ as $\tilde{v}(x) = v(x,1)$. Notice that since we can only derive $\hat{v}$ through measurements, it inevitably contains error, so technically speaking $\hat{v}(x) = \tilde{v}(x) + \eta$. The forward problem is formulated as:

$$\mathcal{G}(u) = \hat{v}$$

The corresponding inverse problem is to find intial state u through the noisy measurements $\hat{v}$. Now consider two intial states differ by:

$$u_1 - u_2 = c_N \sin(N\pi x)$$

u_1 and u_2 result two different heat distributions at time t_1, denoted as v_1 and v_2, and they differ by:

$$\tilde{v}_1 - \tilde{v}_2 = c_N e^{-(N\pi)^2} \sin(N\pi x), \text{ for some } N \in \mathbb{N}$$

As $N \to \infty$, for any $x \in (0,1)$

$$\lim_{N\to\infty} \left| \frac{u_1(x) - u_2(x)}{\tilde{v}_1(x) - \tilde{v}_2(x)} \right| = \lim_{N\to\infty} e^{(N\pi)^2} = \infty$$

Thus in most of the settings, the inverse heat equation violates (3) in well-posedness. Moreover, the presence of measure error can make the situation even worse, as the minor difference between v_1 and v_2 could hide inside the noise. Methods like Truncated Singular Value Decomposition (TSVD) are aimed to deal with these problems.

1.3 Inverse Heat Equations over Hilbert spaces

Definition 1 (Lipschitz domain)**.** An open set Ω in $\mathbb{R}^d$ is said to have a Lipschitz boundary, if for some $L, a, r \in (0, \infty)$, for any $x_0 \in \partial\Omega$, there exist an orthogonal coordinate system with origin at $x_0 = 0$, a cylinder $K = K' \times (-a, a)$ centered at the origin, with K' open ball in $\mathbb{R}^{d-1}$ of radius r, and a function $\varphi : K' \to (-a, a)$, L-Lipschitz continuous with $\varphi(0) = 0$, and

$$\partial\Omega \cap K = \{(x', \varphi(x')); x' \in K'\},$$

$$\Omega \cap K = \{(x', x_N); x' \in K', x_N > \varphi(x')\}.$$

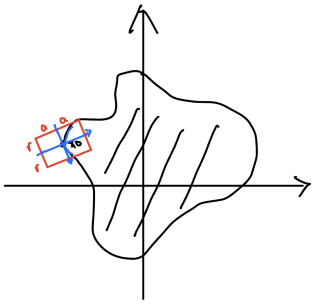

Figure 1: Lipschitz domain

Essentially, one can understand the Lipschitz boundary as a combination of graphs of Lipschitz continuous functions, and an example shown is as the following:
Now let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Then define the Hilbert space $H = L^2(D)$ and operator A to be the negative Laplacian, i.e., $A = -\Delta$. Furthermore, we define the domain of A to be $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$. Then we have the following results for the eigenvalue problem $A\phi_j = \alpha_j \phi_j$.

Theorem 1. *There exists a Hilbert basis $\{\phi_j\}_{j=1}^{\infty}$ of $L^2(\Omega)$ and a sequence $(\alpha_j)_{j=1}^{\infty}$ of reals with $\alpha_j > 0$ and $\lambda_n \to \infty$ such that $A\phi_j = \alpha_j \phi_j$.*

The proof refers to Section 9.8. in [1]. Moreover, for the sequence $\{\alpha_j\}_{j=1}^{\infty}$, we have:

Theorem 2. *For the sequence of eigenvalues $\{\alpha_j\}_{j=1}^{\infty}$, there exists C^- and C^+ so that:*

$$C^- j^{2/d} \leq \alpha_j \leq C^+ j^{2/d}$$

Proof. According to Weyl's Law in the Appendix 4,

$$N(\alpha) \sim \frac{Vol(D)}{(4\pi)^{d/2}\Gamma(1+d/2)} \alpha^{d/2}$$

where $N(\alpha)$ denotes to be the number of eigenvalues in the ball with radius α. Now to see the growth rate of j^{th} eigenvalue α_j, notice that if we order the eigenvalues from small to large, then there are j number of eigenvalues equal or less than α_j; equivalently speaking, we are trying to find the radius α so that $N(\alpha) = j$, and we have:

$$j = N(\alpha_j) \sim C\alpha_j^{d/2} \iff \alpha_j \sim C j^{2/d}$$

□

There is a throughout introduction of Sobolev space H^s and the Hilbert scales $\mathcal{H}^s$ in the Appendix 4. According to Theorem 1 and Theorem 2, the following definition of the norm of $\mathcal{H}^s$ is valid,

$$w = \sum_{j=1}^{\infty} \langle w, \phi_j \rangle \phi_j \ , \ \|w\|_{H^t}^2 = \sum_{j=1}^{\infty} j^{\frac{2t}{d}} |w_j|^2 = \sum_{j=1}^{\infty} j^{\frac{2t}{d}} |\langle w, \phi_j \rangle|^2$$

With the definition of H and A at the beginning of this subsection, we introduce the basic framework: consider the heat conduction equation on D, with Dirichlet boundary conditions, i.e.,

$$\frac{dv}{dt} + Av = 0, \quad v(0) = u$$

Lemma 1. *With the aforementioned setup, for every $u \in H$ and $s > 0$, there exists a unique solution v to the following equation:*

$$\frac{dv}{dt} + Av = 0, \quad v(0) = u$$

in the space $C([0,\infty); H) \cap C((0,\infty); \mathcal{H}^s)$, and we write:

$$v(t) = \exp(-At)u$$

sketch of proof. For any $u \in H$,

$$\begin{aligned}
\|v(t)\|_{H^s}^2 &= \|e^{-tA}\sum u_j\phi_j\|_{H^s}^2 \\
&= \|\sum u_j e^{-tA}\phi_j\|^2 \\
&= \|\sum u_j e^{-\alpha_j t}\phi_j\|^2 \\
&= \sum j^{2s/d}|u_j e^{-\alpha_j t}|^2 \\
&= \sum j^{2s/d}e^{-\alpha_j t}u_j^2 \\
&\asymp \sum \alpha_j^s e^{-\alpha_j y}u_j^2 = t^{-s}\sum \alpha_j^{st}e^{-\alpha_j y}u_j^2
\end{aligned}$$

In the second line, we used Spectral Mapping Theorem. Thus if $u \in H$, the series above converges, and $v \in H^s$. □

Similarly, the inverse problem is formulated as:

$$y = v(1) + \eta = \mathcal{G}(u) + \eta = e^{-A}u + \eta$$

where y is the measurement of heat distribution when $t = 1$ subject to noise η. The major concern we have is still that with the noise y may not lie in the image of H under $\mathcal{G}$. For example, we have shown in the previous lemma that $v \in \bigcap_{s>0}\mathcal{H}^s$. However, if η only has regularity H^b, where $b < s$, then $y \notin \bigcap_{s>0}\mathcal{H}^s$, which violates (1) of well-posedness.

Moreover, for both cases above, the measurements must only consist of finitely many points, while the intial state contains infinitely many points, so our inverse problems always violate (2) of well-posedness. The Bayesian framework will be introduced to address this issue.

2 Deterministic Methods to Inverse Heat Equations

In this section, we will introduce the two most frequently used methods in inverse problems and go over their applications to the one-dimensional heat equations.

2.1 Truncated Singular Value Decomposition (TSVD)

Definition 2 (Compact Operator)**.** Let V, W be Banach spaces. Then $T : V \to W$ be compact if the closure of the image of the unit ball B_V in V, i.e., $\overline{T(B_V)}$ is compact on W.

In this section, unless explicitly mentioned, A is compact. Since every compact operator is bounded, then by Theorem 23 from the Appendix, the adjoint of A, A^* always exists, and is also bounded. Moreover, we always assume the Hilbert spaces H_1 and H_2 to be real.

Theorem 3. *Let $(H_1, \langle\cdot,\cdot\rangle_1)$ and $(H_2, \langle\cdot,\cdot\rangle_2)$ be Hilbert spaces of finite or infinite dimensions, and $A : H_1 \to H_2$ is compact. Denote $Ker(A)$ to be the kernel of A, and $Ran(A)$ the range of A. Then:*

1. *The spaces H_j allow orthogonal decomposition:*

$$H_1 = Ker(A) \oplus (Ker(A))^\perp = Ker(A) \oplus \overline{Ran(A^*)}$$

$$H_2 = \overline{Ran}(A) \oplus (Ran(A))^\perp = \overline{Ran}(A) \oplus Ker(A^*)$$

   where $(Ker(A))^\perp$ denotes the orthogonal space of $Ker(A)$, and $\oplus$ denotes the direct sum.

2. *There exist orthonormal sets of $(v_n) \in H_1$, $(u_n) \in H_2$ and $\lambda_n \searrow 0^+$ so that:*

$$\overline{Ran(A)} = \overline{span}\{u_n | n \in \mathbb{N}\},\ Ker(A)^\perp = \overline{span}\{v_n | n \in \mathbb{N}\}$$

   and the operator A can be represented as:

$$Ax = \sum_n \lambda_n \langle x, v_n\rangle_1 u_n \tag{2}$$

   The tripe $(\lambda_n, v_n, u_n)_{n=1}^\infty$ in the expansion 2 is called the value singular value system of A.

The proof of the result above refers to Appendix A.2. in [6], which uses the spectrum decomposition of compact operators and the procedures are similar to SVD of real matrices. The classical Singular Value Decomposition of real matrices is introduced in the Appendix 4.

Corollary 1. *If A admits a singular value decomposition, i.e., there exists a singular value system (λ_n, v_n, v_n), then*

$$A^* y = \sum_n \lambda_n \langle y, u_n \rangle_2 v_n$$

Proof. From the decomposition above, for any $x \in H_1, y \in H_2$,

$$\begin{aligned}\langle x, A^* y \rangle_1 = \langle Ax, y \rangle_2 &= \langle \sum_n \lambda_n \langle x, v_n \rangle_1 u_n, y \rangle_2 \\ &= \sum_n \lambda_n \langle x, v_n \rangle_1 \langle u_n, y \rangle_2 \\ &= \langle x, \sum_n \lambda_n \langle u_n, y \rangle_2 v_n \rangle_1\end{aligned}$$

□

Now we discuss the existence of solution to the equation $Ax = y$. First define $P : H_2 \to \overline{Ran(A)}$ to be the following orthogonal projection onto the closure of $Span(\{u_j\})$:

$$P(y) = \sum_n \langle y, u_n \rangle_2 u_n \tag{3}$$

Theorem 4. *Let $A : H_1 \to H_2$ be a compact operator with sigular value system (λ_n, u_n, v_n). Then the following are equivalent: $Ax = y$ has a solution if and only if:*

$$y = \sum_n \langle y, u_n \rangle_2 u_n = Py, \ \sum_n \frac{1}{\lambda_n^2} |\langle y, u_n \rangle_2|^2 < \infty$$

The solution is of the form:

$$x = x_0 + \sum_n \frac{1}{\lambda_n^2} \langle y, u_n \rangle_2 v_n \tag{4}$$

where x_0 is any vector in the kernel $Ker(A)$.

Proof. Suppose $Ax = y$ has a solution, which means $y \in Ran(A)$, because P is a projection and $y \in \overline{Ran(A)}$, so $y = Py$. Now, because for each j:

$$\langle y, u_j \rangle_2 = \langle Ax, u_j \rangle_2 = \langle x, A^* u_j \rangle_1, v_j \rangle_1$$

Now by Corollary 1, we have $\langle x, A^* u_j \rangle_1 = \lambda_j \langle x, v_j \rangle_1$. Then by Bessel's inequality:

$$\sum_n \frac{1}{\lambda_n^2} |\langle y, u_n \rangle_2|^2 = \sum_n |\langle x, v_n \rangle_1|^2 \leq \|x\|^2 < \infty$$

For the converse, define $x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle v_n$, first notice that by assumption:

$$\|x\| \leq \|x_0\| + \sum_n \frac{1}{\lambda_n} |\langle y, u_n \rangle_2|^2 < \infty$$

and

$$\begin{aligned}Ax &= A(x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle_2 v_n) \\ &= Ax_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle_2 A v_n \\ &= \sum_n \langle y, u_n \rangle_2 u_n \\ &= Py = y\end{aligned}$$

Finally, to see that any solution of $Ax = y$ has the form as Equation 4. Let x be a solution of the form 4 and x' be another solution. Then:

$$A(x - x') = Ax - Ax' = y - y = 0 \implies x - x' \in Ker(A)$$

Alternatively speaking, x and x' only differ by an element in $Ker(A)$, so x' is also as the form of Equation 4. □

Corollary 2. *For any $y \in H_2$ so that $\sum_n \frac{1}{\lambda_n^2}|\langle y, u_n\rangle_2|^2 < \infty$, there exists a $x \in H_1$ so that:*

$$Ax = Py$$

We now introduce truncated singular value decomposition.

Let $P_k : H_2 \to Span(\{u_1, \ldots, u_k\})$ be the finite dimensional orthogonal projection, i.e., for and $y \in H_2$:

$$P_n(y) = \sum_{n=1}^{k} \langle y, u_n\rangle u_n$$

Definition 3 (TSVD)**.** Let $A : H_1 \to H_2$ be a compact operator with singular system (λ_n, u_n, v_n). For some $k \geq 1$, given $y \in H_2$, the solution of the following equation:

$$Ax = P_k y,\ x \perp Ker(A)$$

is called the truncated SVD approximation of the problem $Ax = y$.

In other words, the truncated SVD approximation is not only the solution to the equation $Ax = P_k y$, but it is also perpendicular to any vector in the kernel of A. The motivation behind the second requirement is that the solution we obtained has the minimized norm among all possible solutions and is unique. In particular, for TSVD, we have the following result:

Theorem 5. *For $k \geq 1$, the solution to TSVD problem, i.e.,*

$$Ax = P_k y,\ x \perp Ker(A)$$

has a unique solution x_k given by

$$x_k = \sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle_2 v_n$$

Moreover, recall the orthogonal projector P as 3, $\|Ax_k - y\| \to \|(1-P)y\|$ as $k \to \infty$

Proof. Notice that first $\sum_{n=1}^{k} \frac{1}{\lambda_n^2}|\langle y, u_n\rangle_2|^2 < \infty$ is always satisfied, and meanwhile, $P(P_x y) = P_x y$, then by Corollary 2, $Ax = P_k y$ has a solution of the form:

$$x_k = x_0 + \sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle_2 v_n$$

Because $x \perp Ker(A)$, by Theorem 3 $x_0 = 0$, and get the solution for the problem:

$$x_k = \sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle_2 v_n$$

To see uniqueness, notice that for any other solution of $Ax = P_k y$, say x', $x' \neq x$. By Theorem 4, any solution is of the form $x' = x_0' + \sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle_2 v_n$, where $x_0' \in Ker(A)$. Since $x' \neq x$, we must have $x_0' \neq 0$. Together with Theorem 3, we have:

$$\langle x', x_0'\rangle = \|x_0\| > 0$$

and it fails to satisfy the condition second requirement for TSVD, namely $x' \perp Ker(A)$. Finally, notice that:

$$\begin{aligned}
\|Ax_k - y\|^2 &= \|P_k y - y\|^2 \\
&= \|(1 - P + P - P_k)y\|^2 \quad (P \text{ is an orthogonal projector}) \\
&= \|(1-P)y\|^2 + \|(P - P_k)y\|^2 \\
&= \|(1-P)y\|^2 + \|\sum_{n=k+1}^{\infty} \langle y, u_n\rangle_2\|^2 \to \|(1-P)y\|^2 \text{ as } k \to \infty
\end{aligned}$$

where the convergence above is a result of monotone convergence theorem. □

The following is the motivation to include the requirement $x \perp Ker(A)$ as part of the problem.

Remark 1. 5 Notice that for solution $x_k = x_0 + \sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle$

$$\|x_k\|^2 = \|x_0\|^2 + \|\sum_{n=1}^{k} \frac{1}{\lambda_n}\langle y, u_n\rangle_2\|^2$$

Thus, x_k constructed in Theorem is the solution with minimum norm. What the theorem above shows, alternatively speaking, is that we have a unique TSVD solution with minimum norm for each chosen k.

Now consider the case the special case when $H_1 = \mathbb{R}^m$ and $H_2 = \mathbb{R}^n$, then $A \in \mathbb{R}^{m\times n}$. From Appendix 4 below, we know that A has singular value decomposition of the form: $U\Lambda V^T$, and in particular if $m < n$:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & & & & \\ 0 & \cdots & & \lambda_m & 0 & \cdots & 0 \end{bmatrix} = [\text{diag}(\lambda_1, \ldots, \lambda_m), 0]$$

If $m \geq n$, we have: $\Lambda = \begin{bmatrix} \text{diag}(w_1, \ldots, w_n) \\ 0 \end{bmatrix}$. Let $U = [u_1, \ldots, u_m]$ and $V = [v_1, \ldots, v_m]$.

Definition 4 (TSVD for real matrices)**.** Suppose the rank of Λ is p,then the k-th ($k \leq p$) truncated SVD approximation of $Ax = y$ is defined to be the solution to the following problem:

$$Ax = P_k y, \ x \perp Ker(A)$$

with the solution given by:

$$x_k = \sum_{j=1}^{k} \frac{1}{\lambda_j}\langle y, u_j\rangle v_j = \sum_{j=1}^{k} \frac{1}{\lambda_j} v_j(u_j^T y) = V\Lambda_k^\dagger U^T y$$

where V, U are given by the singular value decomposition of real matrix in Theorem 19 in the Appendix, and $\Lambda_k^\dagger$ is defined to be:

$$\Lambda_k^\dagger = \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 & & \cdots & 0 \\ 0 & 1/\lambda_2 & & & & & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & & & 1/\lambda_k & 0 & \cdots & 0 \\ 0 & \cdots & & & & & 0 \end{bmatrix}$$

Essentially, everything is the same from infinite dimensions to finite dimensions. The only difference is in finite-dimensional case, we have this notion of matrix.

Remark 2. When $k = p$, denote $A^\dagger = V\Lambda^\dagger U^T$, and we call $A^\dagger$ the Monroe-Penrose inverse of A.

As mentioned in the Introduction above, TSVD is aimed at providing stable inversion, because suppose there is an error ϵ in the measurement, then for large n, λ_n is large, and $1/\lambda_n$ in turn is small, so $A^\dagger$ could possibly amplify the noise. Now the question is how to select the best k, and the following is a rule of thumb:

Definition 5 (Discrepancy Principle)**.** Suppose y is a noisy measurement with y_0 be the underlying truth. Although the error is not explicitly known, but say we have an estimated error ϵ, i.e., $\|y - y_0\| \approx \epsilon$. Then the discrepancy principle states that we should select the smallest K so that:

$$\|y - Ax_k\| = \|y - P_k y\| \leq \epsilon$$

The motivation behind Discrepancy principle is that once we have $\|y - Ax_k\| > \epsilon$, we are essentially trying to transform back a part of the error.

2.2 Tikhonov Regularization

Essentially, Truncated SVD, together with discrepancy principle, is trying to control the residual $\|Ax - y\|$. However, it doesn't force any threshold on x itself. Recall that due to the presence of the error, the problem $Ax = \hat{y}$, where $\hat{y}$ is a noisy measuremnet of true y, is not well-posed, and thus when solving for x from $Ax = \hat{y}$, $\|x\|$ can possibly go to infinity. To address this problem, we introduce Tikhonov regularization, which controls $\|Ax - y\|$ and $\|x\|$ simultaneously and ensure the well-posedness of the problem.

Definition 6 (Tikhonov Regularization)**.** Let $\delta > 0$ be a given constant. The Tikhonov regularized solution $x_\delta \in H_1$ is the minimizer of the functional:

$$F_\delta = \|Ax - y\|^2 + \delta\|x\|^2$$

provided that a minimizer exists. The parameter $\delta > 0$ is called the regularization parameter.

Theorem 6. *Let $A : H_1 \to H_2$ be a compact operator with the singular system (λ_n, v_n, u_n). Then the Tikhonov regularized solution uniquely exists and is given by the formula:*

$$x_\delta = (A^*A + \delta I)^{-1}A^*y = \sum_n \frac{\lambda_n}{\lambda_n^2 + \delta}\langle y, u_n\rangle_2 v_n$$

Proof. We use the Lax-Milgram theorem from the Appendix 4. Define $B : H_1 \times H_1 \to \mathbb{R}$ for any $x, y \in H_1$:

$$B(x, y) = \langle x, (A^*A + \delta I)y\rangle_1$$

which is bilinear. Now to see B is bounded:

$$\begin{aligned}
|B(x,y)| &\le \|x\|\|(A^*A + \delta I)y\| \quad \text{by Cauchy-Schwartz's inequality} \\
&\le \|A^*A + \delta I\|\|x\|\|y\| \quad \text{by property of operator norm} \\
&\le (\|A^*A\| + \delta)\|x\|\|y\| \quad \text{by Triangle Inequality} \\
&= (\|A\|^2 + \delta)\|x\|\|y\|
\end{aligned}$$

So B is bounded. To see it's coercive:

$$\begin{aligned}
B(x,x) &= \langle x, (A^*A + \delta I)x\rangle_1 \\
&= \langle x, A^*Ax\rangle_1 + \langle x, \delta x\rangle_1 \\
&= \|Ax\|^2 + \delta\|x\|^2 \quad \text{by } \|A\| = \|A^*\| \\
&\ge \delta\|x\|^2
\end{aligned}$$

Thus $A^*A + \delta I$ is invertible and x_δ is well-defined. Now to see the singular system expression, first from the Corollary 1 above, we have:

$$\begin{aligned}
A^*Ax &= A^*(Ax) \\
&= A^*(\sum_n \lambda_n\langle x, v_n\rangle_1 u_n) \\
&= \sum_n \lambda_n\langle \sum_n \lambda_n\langle x, v_n\rangle_1 u_n, u_n\rangle_2 v_n \\
&= \sum_n \lambda_n^2\langle x, v_n\rangle_1 u_n
\end{aligned}$$

Now because $x_\delta = (A^*A + \delta I)^{-1}A^*y \iff (A^*A + \delta I)x_\delta = A^*y$, we have:

$$x_\delta = \frac{1}{\delta}A^*(y - Ax_\delta) \in \overline{Ran(A^*)} = \overline{Span(\{v_n\})} = \sum_n \langle x_\delta, v_n\rangle_1 v_n$$

Thus we have:

$$(A^*A + \delta I)x_\delta = A^*y \iff \sum_n(\lambda_n^2 + \delta)\langle x_\delta, v_n\rangle_1 v_n = \sum_n \lambda_n\langle y, u_n\rangle_2 v_n \implies \langle x_\delta, v_n\rangle_1 = \frac{\lambda_n}{\lambda_n^2 + \delta}\langle y, u_n\rangle_2$$

Finally, to see x_δ is indeed the minimizer of $F_\delta = \|Ax - y\|^2 + \delta\|x\|^2$, for any $z \in H_1$, we have:

$$\begin{aligned}
F_\delta(x_\delta + z) &= F_\delta(x_\delta) + 2\langle Az, Ax_\delta - y\rangle_2 + \langle Az, Az\rangle_2 + 2\delta\langle x_\delta, z\rangle_1 + \delta\langle z, z\rangle_1 \\
&= F_\delta(x_\delta) + 2\langle z, A^*Ax_\delta - A^*y\rangle_1 + \langle z, (A^*A + \delta I)z\rangle_1 + 2\langle z, \delta I x_\delta\rangle_1 \\
&= F_\delta(x_\delta) + \cancel{2\langle z, A^*Ax_\delta - (A^*A + \delta I)x_\delta\rangle_1} + \langle z, (A^*A + \delta I)z\rangle_1 + \cancel{2\langle z, \delta I x_\delta\rangle_1}
\end{aligned}$$

Because $A^*A + \delta I$ is positive definite, then the second term $\langle z, (A^*A + \delta I)z\rangle_1$ is nonnegative, and is equal to 0 if and only if $z = 0$. The fact that $F_\delta(x_\delta + z) > F_\delta(x_\delta)$ for any nonzero z also shows the uniqueness of Tikhonov regularized solution $\square$

Remark 3. When both H_1, H_2 are finite-dimensional, and A is a real matrix, we can write the Tikhonov functional as:

$$F_\delta(x) = \left\| \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix} x - \begin{bmatrix} y \\ 0 \end{bmatrix} \right\|^2$$

Then equivalently, the equation $(A^*A + \delta I)x = A^*y$ can be written as:

$$\begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix}^T \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix} x = \begin{bmatrix} A \\ \sqrt{\delta} I \end{bmatrix}^T \begin{bmatrix} y \\ 0 \end{bmatrix}$$

Similar to the number of selected dimensions K in TSVD problem 5, we now consider the choice of the tuning parameter δ in Tikhonov regularization. A commonly used technique is called *Morozov Discrepancy Principle*.

Definition 7 (Morozov Discrepancy Principle)**.** The Morozov Discrepancy Principle states that δ should be chosen from the condition that:

$$\|Ax_\delta - y\| = \epsilon$$

Theorem 7. *Let $P : H_2 \to Ker(A^*)$ be the orthogonal projector and $f(\delta) = \|Ax_\delta - y\|$, where x_δ is the unique Tikhonov regularized solution expressed in Theorem 6. Then:*

1. *$f(\delta)$ is monotonely increasing*
2. *$\|(1-P)y\| \le f(\delta) \le \|y\|$*
3. *There is a unique solution to the Morozov discrepancy principle if and only if $\|(1-P)y\| \le \epsilon \le \|y\|$*

Proof. We first compute $\|Ax_\delta - y\|$. Notice that the expansions of third equality is a result of Theorem 6 and the definition of the projection P 3 in the previous section.

$$\begin{aligned}
\|Ax_\delta - y\|^2 =& \|Ax_\delta - Py - (1-P)y\|^2 \\
=& \|Ax_\delta - Py\|^2 + \|(1-P)y\|^2 \quad \text{by Orthogonality} \\
=& \left\| \sum_i \lambda_i \langle \sum_j \frac{\lambda_j}{\lambda_j^2 + \delta} \langle y, u_j \rangle_2 v_j, v_i \rangle_1 u_i - \sum_i \langle y, u_i \rangle_2 u_i \right\|^2 + \|(1-P)y\|^2 \\
=& \left\| \sum_i \left(\frac{\lambda_i^2}{\lambda_i^2 + \delta} - 1 \right) \langle y, u_i \rangle_2 u_i \right\|^2 + \|(1-P)y\|^2 \\
=& \sum_i \left(\frac{\lambda_i^2}{\lambda_i^2 + \delta} - 1 \right)^2 \langle y, u_i \rangle_2^2 + \|(1-P)y\|^2 = \sum_i \left(\frac{\delta}{\lambda_i^2 + \delta} \right)^2 \langle y, u_i \rangle_2^2 + \|(1-P)y\|^2
\end{aligned}$$

Now for each n, we have:

$$\frac{d}{d\delta} \left(\frac{\delta}{\lambda_n^2 + \delta} \right)^2 = \frac{2\delta\lambda_n^2}{(\lambda_n^2 + \delta)^3} > 0$$

Since the derivative is always positive, (1) follows. Now from the singular value expansion, if $\delta \searrow 0^+$ (we require $\delta > 0$ in Tikhonov Regularization), then:

$$\lim_{\delta \to 0^+} \|Ax_\delta - y\| = \|(1-P)y\|^2 + \lim_{\delta \to 0^+} \delta^2 \sum \left(\frac{1}{\lambda_n^2 + \delta} \right)^2 \langle y, u_n \rangle_2^2 \le \|(1-P)y\|^2 + \sum \left(\frac{1}{\lambda_n^2} \right)^2 \langle y, u_n \rangle_2^2 \lim_{\delta \to 0^+} \delta^2$$

Now because $\sum (\frac{1}{\lambda_n^2})^2 \langle y, u_n \rangle_2^2 \le (\sum \frac{1}{\lambda_n^2} \langle y, u_n \rangle_2)^2$ and $\sum \frac{1}{\lambda_n^2} \langle y, u_n \rangle_2$ is bounded by Theorem 4, we have:

$$\lim_{\delta^+ \to 0} \|Ax_\delta - y\|^2 = 0 + \|(1-P)y\|^2 = \|(1-P)y\|^2$$

Now if $\delta \to \infty$, notice that:

$$\begin{aligned}
\lim_{\delta \to \infty} \|Ax_\delta - y\|^2 &= \lim_{\delta \to \infty} \sum (\frac{\delta^2}{\lambda_n^2 + \delta})^2 \langle y, u_n \rangle_2^2 + \|(1-P)y\|^2 \\
&= \sum \langle y, u_n \rangle_2^2 + \|(1-P)y\|^2 \\
&= \|Py\|^2 + \|(1-P)y\|^2 = \|y\|^2
\end{aligned}$$

where the second equality follows from the monotone convergence theorem and orthogonality between Py and $(1-P)y$. Now written compactly, we have:

$$\|Py\|^2 = \lim_{\delta^+ \to 0} \|Ax_\delta - y\|^2 \le \|Ax_\delta - y\|^2 \le \lim_{\delta \to \infty} \|Ax_\delta - y\|^2 = \|y\|^2$$

Thus, (2) is true. From (1) and (2) it follows that the Morozov disrepancy principle holds if and only if $\|Py\| \le \epsilon \|y\|$. □

2.3 Applications to Heat Equation: Finite-Dimensional Recovery

Suppose the temperatures are measured at $x_0 = 0, x_1 = 0.01, x_2 = 0.02, \ldots, x_{99} = 0.99, x_{100} = 1$ at time $t = 1$, and we have:

$$\frac{\partial_x^2 v(x_j, t)}{\partial x^2} \approx \frac{v(x_{j-1}, t) - 2v(x_j, t) + v(x_{j+1}, t)}{h^2},\ j = 1, \cdots, 99$$

Denote $V(t)$ to be the temperatures of at x_j at time t. More specifically, $V(1) = (V_j(1))_{j=1}^{99}$. Then:

$$\frac{\partial}{\partial t}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_{98} \\ V_{99} \end{bmatrix}(1) = \frac{1}{10^4}\begin{bmatrix} -2 & 1 & & & & & & \\ 1 & -2 & 1 & & & & & \\ & 1 & -2 & 1 & & & & \\ & & & \ddots & & & & \\ & & & & 1 & -2 & 1 & \\ & & & & & 1 & -2 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_{98} \\ V_{99} \end{bmatrix}(1)$$

Denote U to be the initial temperature vector at grid points $\{x_j\}$ chosen above and difference matrix to be B, then the forward problem is:

$$\dot{V}(t) = BV(t), V(0) = U$$

Let $A(t) = e^{tB}$. In Appendix 4.2, we will see $A(t)$ is well defined. Then the solution is given by:

$$V(t) = A(t)U$$

Moreover, in the case $t = 1$, we define

$$A = e^B = \sum_{k=0}^{\infty} \frac{1}{k!} B^k$$

Let $A_k^\dagger$ be the inverse of truncated SVD in the Definition 4 above with dimension k, y be the noisy measurement of $V(1)$, then the TSVD solution of the inverse heat equation in our case:

$$\hat{U} = A_k^\dagger y$$

Now consider the application of Tikhonov regularization of inverse heat equations. We inherit the notation from above. Then for a chosen δ, our solution is:

$$\hat{U} = (A^T A + \delta I)^{-1} A^T y$$

Moreover, k in Truncated SVD and δ in Tikhonov regularization can be chosen based on the Discrepancy Principles introduced, which requires an estimation for the error.

2.4 Applications to Heat Equation: Infinite-Dimensional Recovery

We now illustrate the application of TSVD and Tikhonov regularization of compact operators in inverse heat equations. Notice that in Section 2.3, both methods only give a recovery of initial distribution for finitely many points. However, consider $\tilde{A} : L^2([0,1]) \to \mathbb{R}^N$, where N is the number of measurements and in our previous cases $N = 99$. Then $\tilde{A}$ is a finite-rank operator, and thus compact. Then by Theorem 3, there exists (v_n) in $L^2([0,1])$ and (u_n) in $\mathbb{R}^N$ so that for any $x \in L^2([0,1])$: $Ax = \sum_n \lambda_n \langle x, v_n \rangle u_n$. Correspondingly,

1. the $k^{th} (k \leq N)$ truncated SVD approximation of the initial heat distribution is given by:

$$x_k = \sum_{j=1}^{k} \frac{1}{\lambda_j} \langle y, u_j \rangle v_j$$

2. For a given tuning parameter $\delta > 0$, the solution to the Tikhonov regularization is given by:

$$x_\delta = (A^* A + \delta I)^{-1} A^* y = \sum_n \frac{\lambda_n}{\lambda_n^2 + \delta} \langle y, u_n \rangle v_n$$

Different from using the traditional way, which uses the difference matrix, the inversion based on singular value decomposition of compact operators offers a continuous (in particular squared-integrable) recovery of initial heat distribution.

3 Statistical Methods to Inverse Heat Equations

If we know nothing about the initial heat distribution, then both methods introduced in Section 2.3 and Section 2.4 work. However, once we have prior information about the initial heat distribution, independent of its smoothness, or we have information about the temperature at certain points, we are unable to incorporate these knowledge to the methods we introduce. To address this problem, we further introduce the Bayesian approach in this section. We will first introduce Bayesian framework on finite-dimensions, and then introduce the Bayesian framework for heat equation defined on Hilbert space, based on the probability theory of infinite-dimensions.

3.1 Discretized Bayesian Approaches

In contrast with the previous "deterministic recovery" of the initial condition, we now assume that the prior is a random variable, denoted as X, and the forward problem can be summarized as:

$$Y = \mathcal{G}(X) + \eta$$

Here $\mathcal{G} : \mathbb{R}^n \to \mathbb{R}^J$ is measurable; η is an additive random noise so that η and X are independent, that is, $\eta \perp X$. Furthermore, we assume X has density $\pi_{\text{pr}}(x)$ with respect to Lebsgue measure on $\mathbb{R}^n$, and η is distributed according to the Lebesgue density $\pi_{noise}(e)$. We have the following lemma for the Lebesgue density of conditional random variable $Y|X$:

Lemma 2. *Suppose η is distributed according to Lebesgue density $\pi_{noise}(e)$. Then for a fixed $x \in \mathbb{R}^n$, the Lebesgue density $\pi_{Y|X}(y|x)$ for $Y|X$ is $\pi_{noise}(y - \mathcal{G}(x))$.*

Proof. For a given $x \in \mathbb{R}^n$, denote $\pi_{Y|X}$ to be the Lebesgue density of $\mu_{Y|X}$. For any measurable set in $A \subset \mathbb{R}^J$, we have:

$$\begin{aligned}
\int_A \pi_{Y|X}(y|x)dy &= \mu_{Y|X}(A) \\
&= \mathbb{P}[\mathcal{G}(X) + \eta \in A | X = x] \\
&= \mathbb{P}[\eta \in A - \mathcal{G}(x)] \\
&= \int_{A-\mathcal{G}(x)} \pi_{noise}(e)de \\
&= \int_A \pi_{noise}(y - \mathcal{G}(x))dy \quad \text{(by change of variables)}
\end{aligned}$$

The notation $A - \mathcal{G}(x)$ denotes the set A translated by $-\mathcal{G}(x)$. Since A is arbitrary, this shows $\pi_{Y|X}(y|x) = \pi_{noise}(y - \mathcal{G}(x))$. □

Corollary 3. *$(X, Y) \in \mathbb{R}^n \times R^J$ is a random variable with density $\pi_{noise}(y - G(x))\pi_{pr}(x)$.*

Proof. Denote the Lebesgue density of (X, Y) to be $\pi_{X,Y}(x, y)$. Then by Lemma 2, for any $x \in \mathbb{R}^n, y \in \mathbb{R}^J$, we have:

$$\pi_{X,Y}(x, y) = \pi_{Y|X}(y|x)\pi(x) = \pi_{noise}(y - \mathcal{G}(x))\pi_{\text{pr}}(x)$$

□

Now we are ready to establish the Bayes' formula for our case.

Theorem 8. *Give $y \in \mathbb{R}^J$, assume that:*

$$Z := \int_{\mathbb{R}^n} \pi_{noise}(y - \mathcal{G}(x))\pi_{pr}(x)dx > 0$$

Then for a given $Y = y$, $X|Y$ is a random variable with Lebesgue density $\pi_{X|Y}(x)$ given by:

$$\pi_{X|Y}(x|y) = \frac{1}{Z}\pi_{noise}(y - \mathcal{G}(x))\pi_{pr}(x)$$

Proof. First by Corollary 3:

$$\pi_{X,Y}(x, y) = \pi_{noise}(y - \mathcal{G}(x)\pi_{\text{pr}}(x)$$

By definition:

$$Z = \int_{\mathbb{R}^n} \pi_{noise}(y - \mathcal{G}(x)\pi_{\text{pr}}(x)dx = \int_{\mathbb{R}^n} \pi_{Y|X}(x, y)\pi_{\text{pr}}(x)dx = \int_{\mathbb{R}^n} \pi_{X,Y}(x, y)dx = \pi_Y(x)$$

Then by the conditional probability for continuous random variable, the Bayes' Formula follows and by assumption $\pi_Y(y) > 0$, the formula is well-defined:

$$\pi_{X|Y}(x|y) = \frac{\pi_{X,Y}(x,y)}{\pi_Y(y)}$$

□

Remark 4. In Bayes' formula above:

1. $\pi_{\text{pr}}(x)$ is called prior density.
2. $\pi(y - \mathcal{G}(x))$ is called the likelihood.
3. $\pi_{Y|X}$ is called the posterior density, and below we define it as π_{post}.
4. We define $\Phi(x, y) = -\log \pi_{noise}(y - \mathcal{G}(x))$, and Φ is called the potential.

Remark 5. Notice that Radon-Nikodym derivative is introduced in the Appendix 4. We can further relate our Bayes' formula to the Radon-Nikodym derivative.

1. $\mu_{Y|X}$ is absolutely continuous with respect to μ_X. Because for any measurable set A in $\mathbb{R}^n$:

$$\mu_{X|Y}(A) = \int_A \frac{1}{Z} \exp(-\Phi(x, y))\mu_X(dx)$$

   If $\mu_X(A) = 0$, then because the integral of any measurable function over measure zero set is 0, we have $\mu_{Y|X}(A) = 0$.

2. If we denote $\mu_{X|Y}(x)$ to be the posterior probability measure and $\mu_X(x)$ to be the prior, then for any measurable set A in $\mathbb{R}^n$. Since in (1) we have seen that $\mu_{Y|X}$ is absolute continuous with respect to μ_X, the Radon-Nikodym derivative of $\mu_{Y|X}$ with respect to μ_X is:

$$\frac{d\mu_{X|Y}}{d\mu_X}(x) = \frac{1}{Z} \exp(-\Phi(x, y))$$

3. By Theorem 25 introduced in the Appendix, for any measurable $f(x)$ so that $E^{\mu_{X|Y}}(f(X)) < \infty$, we have:

$$E^{\mu_{X|Y}}[f(X)] = E^{\mu_X}[(\frac{d\mu_{X|Y}}{d\mu_X} \cdot f)(X)]$$

4. In particular if we plug the result of (2) into (3), we have:

$$\begin{aligned} E^{\mu_{X|Y}}(f(X)) &= \int_{\mathbb{R}^n} f(x)\mu_{X|Y}(dx) \\ &= \int_{\mathbb{R}^n} \left(\frac{1}{Z} \exp(-\Phi(x; y))f(x)\right) \mu_X(dx) \\ &= \frac{\int_{\mathbb{R}^n} \exp(-\Phi(x; y))f(x)\mu_X(dx)}{\int_{\mathbb{R}^n} \exp(-\Phi(x; y))\mu_X(dx)} \end{aligned}$$

3.1.1 Prior Construction

Gaussian density is the most common choice for the prior. For more information about Gaussian prior in $\mathbb{R}^n$, please refer to the Appendix 4.

If we have n measurements on the rod, then the most direct way is to assume these n points follow a zero-mean n-variate Gaussian distribution, i.e., $X \sim \mathcal{N}(0, \Gamma)$ and its density is:

$$\pi_{prior} \propto \exp\big(-\frac{1}{2}(x - x_0)^T \Gamma^{-1}(x - x_0)\big)$$

However, in reality, this description might not be accurate, the temperature at one point could be correlated with its neighborhood. Thus, to capture this characteristic of our prior, another construction is introduced in [2].

Let $x : [0, 1] \to \mathbb{R}$ denote the original function we are trying to estimate. The function x could be the input signal in deconvolution problem or in our case, it's the initial heat distribution. Then the temperatures are measured at $\{s_0, s_1, \ldots, s_n\} \subseteq [0, 1]$, and $s_0 = 0, s_n = 1$. Notice that $x(s_0) = x(s_n)$ are boundary conditions, and are usually known in advance without any uncertainties. However, for interior

points, we interpolate that the middle value should be an average of its neighbors but with uncertainties. Let $X = (X_0, \ldots, X_{n+1})$ denote our prior for x at s_0 to s_n, so the system of equations can be written as:

$$X_0 = a, X_n = b,\ a, b \in \mathbb{R}$$

$$X_j = \frac{1}{2}(X_{j-1} + X_{j+1}) + W_j,\ W_j \sim \mathcal{N}(0, \gamma^2),\ j = 1, 2, \ldots, n \tag{5}$$

The second equation 5 can also be written as:

$$L_s X - \vec{b} = \frac{1}{2}\begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ 0 & & & -1 & 2 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{bmatrix} - \begin{bmatrix} a \\ 0 \\ \vdots \\ b \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_{n-1} \end{bmatrix} = W$$

where $L_s \in \mathbb{R}^{(n-1)\times(n-1)}$ and $W_n \sim \mathcal{N}(0, \gamma^2 I)$

Lemma 3. *Let $X \sim \mathcal{N}_k(\mu, \Sigma)$ be a multivariate Gaussian distribution. Then for $A \in R^{k\times k}$ and $b \in \mathbb{R}^k$, $Y = AX + b \sim \mathcal{N}(A\mu + b, A\Sigma A^T)$.*

Proof. The moment-generating function of a random vector X is

$$M_x(t) = \mathbb{E}[\exp(t^T x)]$$

and therefore the moment-generating function of the random vector Y is given by

$$\begin{aligned} M_y(t) &= \mathbb{E}[\exp(t^T(Ax + b))] \\ &= \mathbb{E}[\exp(t^T Ax) \cdot \exp(t^T b)] \\ &= \exp(t^T b) \cdot \mathbb{E}[\exp(t^T Ax)] \\ &\equiv \exp(t^T b) \cdot M_x(A^T t). \end{aligned}$$

The moment-generating function of the multivariate normal distribution is

$$M_x(t) = \exp\left(t^T\mu + \frac{1}{2}t^T\Sigma t\right)$$

and therefore the moment-generating function of the random vector Y becomes

$$\begin{aligned} M_y(t) &= \exp(t^T b) \cdot M_x(A^T t) \\ &= \exp(t^T b) \cdot \exp\left((A^T t)^T\mu + \frac{1}{2}(A^T t)^T\Sigma(A^T t)\right) \\ &= \exp\left(t^T(A\mu + b) + \frac{1}{2}t^T A\Sigma A^T t\right). \end{aligned}$$

Because moment-generating function and probability density function of a random variable are in one-to-one correspondence, this demonstrates that y is following a multivariate normal distribution with mean $A\mu + b$ and covariance $A\Sigma A^T$.

□

Then by Lemma 3, since $X = L_s^{-1}(W + \vec{b})$, we have:

$$X \sim \mathcal{N}(L_s^{-1}\vec{b}, \gamma^2(L_s^{-1}(L_s^{-1})^T))$$

Or:

$$\pi^s_{prior} \propto \exp(\frac{1}{2\gamma^2}\|L_s(x - \vec{b})\|^2)$$

Remark 6. In some rare cases, we are also uncertain about the boundary condition. In this case, we may take the following formulation of the prior. Say, we expect that $x(0) \approx a$ and $x(1) \approx b$. Then we let:

$$X_0 \sim \mathcal{N}(a, \frac{\gamma^2}{\delta_0^2}), X_n \sim \mathcal{N}(b, \frac{\gamma^2}{\delta_{100}^2})$$

and

$$L_r X - \vec{b} = \begin{bmatrix} \delta_0 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & \delta_{100} \end{bmatrix}\begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_n \end{bmatrix} - \begin{bmatrix} a \\ 0 \\ \vdots \\ b \end{bmatrix} = \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_n \end{bmatrix} = W$$

where $L_r \in \mathbb{R}^{(n+1)\times(n+1)}$ and $W_{n+1} \sim \mathcal{N}(0, \gamma^2 I)$.

Remark 7. In some other cases, we don't want smooth prior. Alternatively speaking, we don't assume the temperature at the middle point is the average of its neighbors. Assume we know boundary conditions of one side, then we will take the following framework:

$$X_0 = a, X_n = W_n$$

$$X_{j+1} = X_j + W_{j+1}, \quad j = 0, \ldots, n-1$$

and

$$L_n X - \vec{b} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} - \begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} = W$$

where $L_n \in \mathbb{R}^{n\times n}$ and $W_n \sim \mathcal{N}(0, \gamma^2 I)$.

3.1.2 Posterior

We have just introduced Gaussian prior models, and now will see how to combine the prior with our knowledge of observation (likelihood) to get posterior distribution. However, still, posterior is a probability distribution rather than a point estimation, so in the end, we will see how to get point estimation of the initial condition based on the posterior distribution.

Recall the additive model introduced at the beginning of the subsection 3.1 Discretized Bayesian Approaches. Suppose $\mathcal{G} : \mathbb{R}^{99} \to \mathbb{R}^{99}$ is a linear operator, which means that there exists a matrix $A \in \mathbb{R}^{99\times 99}$ so that $\mathcal{G}(X) = AX$. Now we try to derive the posterior distribution for the Gaussian smooth prior.

We assume the noise $E \sim \mathcal{N}(e_0, \Gamma_{noise})$. For convenience, denote $x_0 = L_s^{-1}\vec{b}$ and $\Gamma_{pr} = \gamma^2(L_s^{-1}(L_s^{-1})^T)$). Then by Lemma 2:

$$\begin{aligned} \pi_{post}(x|y) &\propto \pi_{\text{pr}}(x)\pi_{noise}(y - Ax) \\ &\propto \exp\left(-\frac{1}{2}(x - x_0)^T\Gamma_{\text{pr}}^{-1}(x - x_0)\right)\exp\left(-\frac{1}{2}(y - Ax - e_0)^T\Gamma_{\text{noise}}^{-1}(y - Ax - e_0)\right) \end{aligned} \tag{6}$$

However Equation 6 is not useful because it's difficult to identify the mean and variance. Instead, we should appeal to the result introduced in Appendix 4.

Theorem 9. *Assume that $X : \Omega \to \mathbb{R}^n$ and $E : \Omega \to \mathbb{R}^m$ are mutually independent Gaussian random variables,*

$$X \sim \mathcal{N}(x_0, \Gamma_{pr}), \quad E \sim \mathcal{N}(e_0, \Gamma_{noise}),$$

and $\Gamma_{pr} \in \mathbb{R}^{n\times n}$ and $\Gamma_{noise} \in \mathbb{R}^{k\times k}$ are positive definite. Assume further that we have a linear model $Y = AX + E$ for a noisy measurement Y, where $A \in \mathbb{R}^{k\times n}$ is a known matrix. Then the posterior probability density of X given the measurement $Y = y$ is

$$\pi_{X|Y}(x|y) \propto \exp\left(-\frac{1}{2}(x - \bar{x})^T\Gamma_{post}^{-1}(x - \bar{x})\right),$$

where

$$\bar{x} = x_0 + \Gamma_{pr}A^T(A\Gamma_{pr}A^T + \Gamma_{noise})^{-1}(y - Ax_0 - e_0),$$

and

$$\Gamma_{post} = \Gamma_{pr} - \Gamma_{pr}A^T(A\Gamma_{pr}A^T + \Gamma_{noise})^{-1}A\Gamma_{pr}.$$

Proof. Here we apply Theorem 26 from the Appendix. The first step is to derive the joint distribution of (X, Y). Notice that:

$$\mathbb{E}\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad y_0 = Ax_0 + e_0.$$

Since

$$\mathbb{E}\{(X - x_0)(X - x_0)^T\} = \Gamma_{\text{pr}},$$

$$\begin{aligned} \mathbb{E}\{(Y - y_0)(Y - y_0)^T\} &= \mathbb{E}\{((AX - Ax_0) + (E - e_0))((AX - Ax_0) + (\mathbb{E} - e_0))^T\} \\ &= A\Gamma_{\text{pr}}A^T + \Gamma_{\text{noise}} \end{aligned}$$

and, furthermore,

$$\mathbb{E}\{(X-x_0)(Y-y_0)^T\} = \mathbb{E}\{(X-x_0)((AX-Ax_0)+(\mathbb{E}-e_0))^T\} = \Gamma_{\text{pr}}A^T,$$

we find that

$$\text{cov}\begin{bmatrix} X \\ Y \end{bmatrix} = \mathbb{E}\begin{bmatrix} X-x_0 & Y-y_0 \\ X-x_0 & Y-y_0 \end{bmatrix} = \begin{bmatrix} \Gamma_{\text{pr}} & \Gamma_{\text{pr}}A^T \\ A\Gamma_{\text{pr}} & A\Gamma_{\text{pr}}A^T+\Gamma_{\text{noise}} \end{bmatrix}.$$

Hence, the joint probability density of X and Y is of the form

$$\pi(x,y) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}^T \begin{bmatrix} \Gamma_{\text{pr}} & \Gamma_{\text{pr}}A^T \\ A\Gamma_{\text{pr}} & A\Gamma_{\text{pr}}A^T+\Gamma_{\text{noise}} \end{bmatrix}^{-1} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}\right).$$

The expressions in the theorems then follow from Theorem 26 □

Remark 8. One can check that the following are equivalent expressions for the posterior mean and posterior covariance, and can enjoy some simplicity in future computation.

$$\Gamma_{\text{post}} = (\Gamma_{\text{pr}}^{-1} + A^T\Gamma_{\text{noise}}^{-1}A)^{-1}.$$

Furthermore, the posterior mean can be written as

$$\bar{x} = (\Gamma_{\text{pr}}^{-1} + A^T\Gamma_{\text{noise}}^{-1}A)^{-1}(A^T\Gamma_{\text{noise}}^{-1}(y-e_0) + \Gamma_{\text{pr}}^{-1}x_0).$$

However, notice that the posterior is essentially a probability distribution, and in most of the cases, we want a unique solution to our inverse problems. Thus, let $\pi(x|y)$ denote the posterior and the following are two common choices:

1. **Maximum a Posteriori (MAP)**: the MAP estimate x_{MAP} is defined to be:

$$x_{MAP} = \arg\max \pi(x|y)$$

2. **Conditional Mean**: the conditional mean of the unknown random variable X conditioned on y is:

$$x_{CM} = \mathbb{E}(X|Y) = \int x\pi(x|y)dx$$

To find x_{MAP}, this is an optimization problem and gradient-based methods are usually applied. To find the conditional mean, we will will introduce a very powerful technique called Markov Chain Monte Carlo (MCMC) in the next subsection.

Remark 9. If both the prior U and the noises E are Gaussian, then as Theorem 9 shows, the posterior is also Gaussian. Since Gaussian is unimodal and the peak is reached at the mean, we have:

$$x_{MAP} = x_{CM}$$

Now we introduce a special case when both the prior U and the noises E are *Gaussian white noises*, i.e., we assume:

$$X \sim \mathcal{N}(0, \gamma^2 I), \quad E \sim \mathcal{N}(0, \delta^2 I)$$

In this case, we have the following surprising result, which is formulated as a theorem:

Theorem 10. *Assume that $X : \Omega \to \mathbb{R}^n$ and $E : \Omega \to \mathbb{R}^m$ are mutually independent Gaussian random variables,*

$$X \sim \mathcal{N}(0, \gamma^2 I), \quad E \sim \mathcal{N}(0, \delta^2 I),$$

The posterior mean $\bar{x}$ is the solution of Tikhonov regularization to the equation:

$$Ax = y$$

with tuning parameter $\alpha = \delta^2/\gamma^2$.

Proof. By Theorem 9, the posterior mean is:

$$\begin{aligned} \bar{x} &= x_0 + \Gamma_{\text{pr}}A^T(A\Gamma_{\text{pr}}A^T + \Gamma_{\text{noise}})^{-1}(y - Ax_0 - e_0) \\ &= \gamma^2 A^T(\gamma^2 AA^T + \delta^2 I)^{-1}y \quad (x_0 = 0, e_0 = 0) \\ &= A^T(AA^T + \alpha I)^{-1}y \\ &= (A^TA + \alpha I)^{-1}A^Ty \end{aligned}$$

where $\alpha = \delta^2/\gamma^2$. To see the last equality, notice that:

$$\begin{aligned}(A^TA + \alpha I)A^T(AA^T + \alpha I)^{-1} &= (A^TAA^T + \alpha A^T)(AA^T + \alpha I)^{-1} \\ &= A^T(AA^T + \alpha I)(AA^T + \alpha I)^{-1} = A^T\end{aligned}$$

□

Remark 10. The correspondence between the posterior mean and the Tikhonov regularization should be canonical if we think of Maximum a Posteriori (MAP) as a bridge. Recall the notation of Potential $\Phi(x, y)$. Then by Corollary 3, we have:

$$\pi(x|y) \propto \exp(-\Phi(x, y), \quad \text{where } \Phi(x, y) = \frac{1}{2\gamma^2}\|x\|^2 + \frac{1}{2\delta^2}\|y - Ax\|^2$$

Then maximizing the posterior density is equivalent to minimizing the potential, which is a Tikhonov regularization problem.

3.2 Smooth Bayesian Approaches

Notice that in Discretized Bayesian Approaches 3.1, our prior is a random vector, and it provides no information about the temperature of other points. To this end, we introduce the following method based on theory of Gaussian process that provides smooth recovery on infinite dimensions. Essentially, we assume our prior distribution to be random functions on $[0, 1]$, which is infinite-dimensional.

3.2.1 Prior Construction

First we introduce Gaussian process, which is an extension of multivariate Gaussian to infinite dimensions. To this end, we introduce the following method based on theory of Gaussian process that provides smooth recovery on infinite dimensions.

Definition 8 (Gaussian Process)**.** A Gaussian process on $[0, 1]$ is a continuous-time stochastic process such that any finite subcollection of random variables has a multivariate Gaussian distribution.

In particular, a collection of random variables $\{F(x) : x \in \mathcal{X}\}$ is said to be drawn from a Gaussian process with mean function $m(\cdot)$ and covariance function $k(\cdot, \cdot)$ if for any finite set of elements $x_1, \ldots, x_m \in \mathcal{X}$, the associated finite set of random variables $F(x_1), \ldots, F(x_m)$ have distribution,

$$\begin{bmatrix} F(x_1) \\ \vdots \\ F(x_m) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_m) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_m) \\ \vdots & \ddots & \vdots \\ k(x_m, x_1) & \cdots & k(x_m, x_m) \end{bmatrix}\right).$$

We denote this Gaussian process using the following notation:

$$F(\cdot) \sim GP(m(\cdot), k(\cdot, \cdot)).$$

Observe that the mean function and covariance function are aptly named since the above properties imply that

$$m(x) = \mathbb{E}[F(x)]$$
$$k(x, x') = \mathbb{E}[(F(x) - m(x))(F(x') - m(x'))].$$

We also re-state Theorem 26 from Appendix 4, since this is one of the most theorem we will frequently use in both Section 3.2 and Section 3.4:

Theorem 11. *Given two jointly Gaussian vectors* $\mathbf{X}$ *and* $\mathbf{Y}$*, with the joint distribution specified by*

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}\right),$$

the conditional distribution of $\mathbf{Y}$ *given* $\mathbf{X} = \mathbf{x}$ *is a normal distribution with the mean and covariance matrix given by:*

$$\boldsymbol{\mu}_{Y|X} = \boldsymbol{\mu}_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(\mathbf{x} - \boldsymbol{\mu}_X),$$
$$\Sigma_{Y|X} = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}.$$

The following result is originally stated in [7]. We are usually not primarily interested in simply drawing random functions from the prior, but want to incorporate the knowledge we know about the function. First by simplicity, we are considering zero-mean Gaussian process, that is:

$$F \sim GP(0, k(\cdot, \cdot))$$

Suppose we have information on $\{(t_i, F(x_i))|t_i \in [0,1], i = 1, \cdots, n\}$, and denote $\vec{t} = (t_1, t_2, \ldots, t_n)$. Essentially, for whatever sample path we generate, it should pass through the collection of points $\{(x_i, f(x_i)\}_{i=1}^n$. Meanwhile, we want to generate samples at $t_{1_*}, t_{2_*}, \ldots, t_{n_*}$. Denote $\vec{F}$ be the random vector $[F_{t_1}, F_{t_2}, \ldots, F_{t_n}]^T$, and $\vec{F_*} = [F_{t_{1_*}}, F_{t_{2_*}}, \ldots, F_{t_{n_*}}]^T$. Now the joint distribution of $\vec{F}$ and $\vec{F_*}$ is:

$$\begin{bmatrix} \vec{F} \\ \vec{F_*} \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(\vec{t}, \vec{t}) & K(\vec{t}, \vec{t_*}) \\ K(\vec{t_*}, \vec{t}) & K(\vec{t_*}, \vec{t_*}) \end{bmatrix}\right)$$

Here $K(\vec{t}, \vec{t_*})$ is covariance matrix between $\vec{t}$ and $\vec{t_*}$, and we similarly define $K(\vec{t}, \vec{t}), K(\vec{t_*}, \vec{t}), K(\vec{t_*}, \vec{t_*})$. We now want to generate samples so that they pass through $(t_i, F(t_i))$, which is essentially the following conditional distribution:

$$\vec{F_*}|\vec{t_*}, \vec{t}, \vec{F} \sim \mathcal{N}(K(\vec{t_*}, \vec{t})K(\vec{t}, \vec{t})^{-1}F, K(\vec{t_*}, \vec{t_*}) - K(\vec{t_*}, \vec{t})K(\vec{t}, \vec{t})^{-1}K(\vec{t}, \vec{t_*}))$$

which is essentially an application of Theorem 11

Similarly, if the mean function $m(\cdot)$ is not 0, the posterior distribution of f_* is:

$$\vec{F_*}|\vec{t_*}, \vec{t}, \vec{F} \sim \mathcal{N}(m(\vec{t_*}) + K(\vec{t_*}, \vec{t})K(\vec{t}, \vec{t})^{-1}(\vec{F} - m(\vec{t})), K(\vec{t_*}, \vec{t_*}) - K(\vec{t_*}, \vec{t})K(\vec{t}, \vec{t})^{-1}K(\vec{t}, \vec{t_*})) \tag{7}$$

For convenience, in future notations, we denote $K = K(\vec{t}, \vec{t})$, $K_* = K(\vec{t}, \vec{t_*})$, $K_{**} = K(\vec{t_*}, \vec{t_*})$ and correspondingly, $K(\vec{t_*}, \vec{t}) = K^T$.

Now recall our general settings. Let $x : [0,1] \to \mathbb{R}$ denote the original function we try to estimate. The function x could be the input signal in deconvolution problem or in our case, it's the initial heat distribution. Then the temperatures are measured at $\{s_0, s_1, \ldots, s_n\} \subseteq [0,1]$, with $s_0 = 0, s_n = 1$, and denote the noisy measurement to be $\vec{y}$. Notice that $x(s_0)$ and $x(s_n)$ are boundary conditions, and are usually known in advance without uncertainties. First we assume the prior distribution of the original function x is distributed according to the Gaussian Process:

$$F \sim GP(m(\cdot), K(\cdot, \cdot))$$

Suppose we have Dirichlet boundary conditions with $x(0) = x(1) = 0$, then $\vec{F} = (F_0, F_1) = (0, 0)$. Meanwhile, $m(0) = m(1) = 0$. Denote the initial values on $\vec{s_*} = \{s_1, s_1, \ldots, s_{n-1}\}$ to be $\vec{F_*}$. Then we update the distribution of $\vec{F_*}$ according to equation 7, that is:

$$\vec{F_*} \sim \mathcal{N}(m(\vec{s_*}), K_{**} - K_* K^{-1} K_*^T)$$

We further denote $\tilde{K} = K_{**} - K_* K^{-1} K_*^T$.

3.2.2 Posterior

Similar as before, assume the noisy measurement of the output $\vec{y}$ and the initial values on those points $\vec{x} = (x(s_1), x(s_2), \ldots, x(s_{n-1}))$ has the following relationship:

$$\vec{y} = A\vec{x} + E$$

and we assume:

$$E \sim \mathcal{N}(0, \Gamma_{noise})$$

Now by Theorem 9, we have the posterior of $\vec{F_*}$ follows multivariate Gaussian distribution with posterior mean, which is also our estimates of x at $s_1, \ldots, s_{n-1}$:

$$[\hat{x}(s_1), \hat{x}(s_2), \hat{x}, \ldots, \hat{x}(s_{n-1}))]^T = [m(\vec{s_*}) + \tilde{K}A^T(A\tilde{K}A^T + \Gamma)^{-1}(y - Am(\vec{s_*})]^T$$

and posterior covariance:

$$\tilde{K} - \tilde{K}A^T(A\tilde{K}A^T + \Gamma)^{-1}(y - Am(\vec{s_*}))$$

Different from previous methods in Section 3.1, now we are able to estimate any point in $(0,1)$. For a new position $t \in (0,1)$ with $t \notin \{s_1, \ldots, s_{n-1}\}$, we try to estimate its initial value $x(t)$. Essentially, we estimate $x(t)$ as the posterior mean conditioned on:

$$\begin{aligned}[F_0, F_{s_1}, \ldots, F_{s_{n-1}}, F_1]^T &= [0, \hat{x}(s_1), \hat{x}(s_2), \hat{x}, \ldots, \hat{x}(s_{n-1}), 0]^T \\ &= [0, m(\vec{s_*}) + \tilde{K}A^T(A\tilde{K}A^T + \Gamma)^{-1}(y - Am(\vec{s_*})), 0]^T\end{aligned}$$

To estimate $x(t)$, similarly, we first write out the joint distribution of $[F_t, F_0, F_{s_1}, \ldots, F_{s_{n-1}}, F_1]$:

$$\begin{bmatrix} F_0 \\ F_{s_1} \\ \vdots \\ F_{s_{n-1}} \\ F_1 \\ F_t \end{bmatrix} \sim \mathcal{N}(m, \begin{bmatrix} K(\vec{s},\vec{s}) & K(\vec{s},t) \\ K(t,\vec{s}) & K(t,t) \end{bmatrix})$$

where $\vec{s}$ denotes $(s_0, s_1, \ldots, s_n)$, and $K(\vec{s}, t)$ denotes the covariance matrix between $\vec{s}$ and t, while $K(\vec{s},\vec{s})$, $K(\vec{s}, t)$, and , $K(t,t)$ are similarly defined. Again we apply Theorem 11, and get the posterior mean of F_t, which is also our estimation of $x(t)$:

$$\hat{x}(t) = m(t) + K(t,\vec{s})K(\vec{s},\vec{s})(\hat{m}(\vec{s}) - m(\vec{s})) \tag{8}$$

where:

$$\hat{x}(\vec{s}) = [0, m(\vec{s_*}) + \tilde{K}A^T(A\tilde{K}A^T + \Gamma)^{-1}(y - Am(\vec{s_*})), 0]^T \tag{9}$$

3.3 Application to Heat Equation: Finite-dimensional Recovery

Suppose that the known boundary condition is zero Dirichlet boundary, that is,

$$v(0,t) = v(1,t) = 0, \quad \text{for } t \in [0,1]$$

Recall the setting we defined in Section 2.3. Denote $V(1) = (V_j(1))_{j=1}^{99}$. Then:

$$\frac{\partial}{\partial t}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_{98} \\ V_{99} \end{bmatrix}(1) = \frac{1}{10^4}\begin{bmatrix} -2 & 1 & & & & & & \\ 1 & -2 & 1 & & & & & \\ & 1 & -2 & 1 & & & & \\ & & & \ddots & & & & \\ & & & & 1 & -2 & 1 & \\ & & & & & 1 & -2 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_{98} \\ V_{99} \end{bmatrix}(1) \tag{10}$$

Denote the initial condition to be U, and the difference matrix to be B, then the forward problem is:

$$\dot{V}(t) = BV(t), V(0) = U$$

Let $A = e^{tB}$. Then $V(1) = e^B U$, and alternatively:

$$A = e^B = \sum_{k=0}^{\infty} \frac{1}{k!} B^k$$

However, we now assume U to be a random variable, and $V = AU$ is also a random variable. First, we have the simplest case, in which assume that the heat of each point is independent of the others with mean zero, i.e.:

$$U \sim \mathcal{N}(0, \gamma^2 I)$$

Suppose the measurement error $\eta \sim \mathcal{N}(0, \sigma^2 I_{99})$. Then the estimation of the initial condition directly follows Theorem 9, i.e.,:

$$\hat{U} = (e^B(e^B)^T + \frac{\gamma^2}{\sigma^2} I)^{-1}(e^B)^T V(1)$$

However, we often assume the initial heat distribution is smooth, which corresponds to the case of smooth prior, Equation 5 we introduced in Section 3.1.1, i.e.:

$$L_s U \sim \mathcal{N}(0, \gamma^2 I_{99})$$

Again, suppose the measurement error $\eta \sim \mathcal{N}(0, \sigma^2 I_{99})$. Recall Φ is the potential defined in Remark 4, then by Remark 10,

$$\Phi(u, v) = \frac{1}{2\gamma^2}\|L_s u\|^2 + \frac{1}{2\delta^2}\|v - Au\|^2$$

Unfortunately, this problem no longer falls into the classical Tikhonov regularization, and we have to introduce the following theorem:

Theorem 12. *Consider the linear observation model $Y = AX + E$, $A \in \mathbb{R}^{m\times n}$, where $X \in \mathbb{R}^n$, $E \in \mathbb{R}^m$ are mutually independent random variables and E is Gaussian, $E \sim \mathcal{N}(0, \Gamma_{noise})$. Let $L \in \mathbb{R}^{n\times m}$ be a matrix such that $LX \sim \mathcal{N}(0, I_n)$. Moreover, if we assume $Ker(L) \cap Ker(A) = \{0\}$, then the function*

$$x \mapsto \pi_{pr}(x)\pi(y|x) \propto \exp\left(-\frac{1}{2}(\|Lx\|^2 + (y - Ax)^T\Gamma_{noise}^{-1}(y - Ax))\right) \tag{11}$$

defines a Gaussian density over $\mathbb{R}^n$ with mean given by:

$$\bar{x} = (L^T L + A^T\Gamma_{noise}^{-1}A)^{-1}A^T\Gamma_{noise}^{-1}y$$

and covariance given by the formulas:

$$\Gamma = (L^T L + A^T\Gamma_{noise}^{-1}A)^{-1}.$$

Proof. We start by showing that the matrix $G = L^T L + A^T\Gamma_{\text{noise}}^{-1}A \in \mathbb{R}^{n\times n}$ is invertible. Indeed, if $x \in \text{Ker}(G)$, then $x^T Gx = x^T(Gx) = x^T 0 = 0$. Meanwhile, we have

$$x^T Gx = xL^T Lx + x^T A^T\Gamma_{\text{noise}}^{-1}Ax = \|Lx\|^2 + (Ax)^T\Gamma_{\text{noise}}^{-1}Ax$$

Now since $\|Lx\| \geq 0$ and Γ_{noise} is positive definite, we must have $Lx = Ax = 0$, i.e., $x \in \text{Ker}(L) \cap \text{Ker}(A)$. Now because by assumption, their intersection is trivial, we have $x = 0$. Hence, $\text{Ker}(G) = \{0\}$ and G is invertible.

Consider the quadratic functional in the exponent of map 11. Reformulating it in terms of a positive definite quadratic form, we obtain

$$\begin{aligned}
\|Lx\|^2 + (y - Ax)^T\Gamma_{\text{noise}}^{-1}(y - Ax) &= x^T Gx - y^T\Gamma_{\text{noise}}^{-1}Ax - x^T A^T\Gamma_{\text{noise}}^{-1}y + y^T\Gamma_{\text{noise}}^{-1}y \\
&= x^T Gx - 2x^T A^T\Gamma_{\text{noise}}^{-1}y + y^T\Gamma_{\text{noise}}^{-1}y \\
&= x^T Gx - 2x^T G(G^{-1}A^T\Gamma_{\text{noise}}^{-1}y) + y^T\Gamma_{\text{noise}}^{-1}y \\
&= x^T Gx - 2x^T G\bar{x} + y^T\Gamma_{\text{noise}}^{-1}y
\end{aligned}$$

Now notice that since G is the sum of two symmetric matrices, it's again symmetric, so:

$$\begin{aligned}
x^T Gx - 2x^T G\bar{x} + y^T\Gamma_{\text{noise}}^{-1}y &= x^T Gx - x^T G\bar{x} - \bar{x}^T Gx + \bar{x}^T G\bar{x} + y^T\Gamma_{\text{noise}}^{-1}y - \bar{x}^T G\bar{x} \\
&= (x - \bar{x})^T G(x - \bar{x}) + c(y)
\end{aligned}$$

where $c(y)$ is a constant with respect to x.

□

To see now the theorem is applicable in our scenario, first notice that L_s is invertible. For $A = e^B$, we have the following lemma:

Lemma 4. *Let $B \in \mathbb{R}^{n\times n}$, then e^B is invertible.*

Proof. First we claim that if $A, B \in \mathbb{R}^{n\times n}$ commute, then $e^{A+B} = e^A e^B$. To see this:

$$\begin{aligned}
e^{A+B} &= \sum_{k=0}^{\infty}\frac{(A+B)^k}{k!} \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{A^i B^{k-i}}{k!}\binom{k}{i} \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{A^i B^{k-i}}{i!(k-i)!} \\
&= \sum_{k=0}^{\infty}\sum_{i+h=k}\frac{A^i B^h}{i!h!} \\
&= \sum_{i=0}^{\infty}\sum_{h=0}^{\infty}\frac{A^i B^h}{i!h!} = e^A e^B
\end{aligned}$$

Now as $(-A)A = -A^2 = A(-A)$, for e^A, its inverse is e^{-A}. □

Now as both L_s and A are invertible, their kernels are trivial and so does their kernel's intersection. Then Theorem 12 yield:

$$\hat{u} = \frac{1}{\delta^2}(\frac{1}{\gamma^2}L_s^T L_s + \frac{1}{\delta^2}A^T A)^{-1}A^T V(1)$$

Remark 11. We should notice that for L_r, L_n defined in Remark 6 and 7, with zero or near-zero (with uncertainty) Dirichlet boundary, Theorem 12 is also applicable, as both L_r and L_n are invertible.

3.4 Application to Heat Equation: Infinite-dimensional Recovery

Similar as our recovery of initial condition in Section 2.3, one limitation of our methods in Section 3.1 is that it only provides recovery of initial conditions on finitely many points. To estimate the initial value at each point $t \in (0, 1)$, we follow the same procedures from Section 3.2.

In general, the application of Gaussian process will be very favarable in two settings: we are either very confident about our prior knowledge, or we know nothing at all. In the first case, we can specify the mean function $m(x)$ to be whatever initial distribution we believe, and then update it with our observations. In contrast, if we have no prior information at all, then we let the Gaussian process to have zero mean, but even in this case, we can still have a reasonable estimation of the initial value for each point. Assume that the initial condition distributed according to:

$$F \sim GP(m(\cdot), K(\cdot, \cdot))$$

In our problem,

$$A = e^B = \sum_{k=0}^{\infty} \frac{1}{k!}B^k$$

where B is defined in Equation 10 as before. Then if we have noisy measurement at $0.01, 0.02, \ldots, 0.99$, and we denote $\vec{s_*} = [0.01, 0.02, \ldots, 0.99]^T$, according Equation 8 and Equation 9, the estimated initial values at $\vec{s_*}$ are:

$$\hat{x}(\vec{s}) = [0, m(\vec{s_*}) + \tilde{K}A^T(A\tilde{K}A^T + \Gamma)^{-1}(y - Am(\vec{s_*})), 0]^T$$

where

$$\tilde{K} = K(\vec{s_*}, \vec{s_*}) - K(\vec{s_*}, [0, 1]^T)K([0, 1]^T, [0, 1]^T)^{-1}K(\vec{s_*}, [0, 1]^T)^T$$

and for any given point t not in $\vec{s_*}$:

$$\hat{x}(t) = m(t) + K(t, \vec{s})K(\vec{s}, \vec{s})(\hat{m}(\vec{s}) - m(\vec{s}))$$

where

$$\hat{m}(\vec{s}) = [0, \hat{m}(\vec{s_*}), 0]^T$$

Remark 12. We can also apply Gaussian process to nonzero Dirichlet boundary. Essentially, if $x(0) = a$, $x(1) = b$, we specify the mean function with $m(0) = a$, $m(1) = b$, and conditioning on $X_0 = a$ and $X_1 = b$, so we update the mean function and covariance function accordingly using Equation 7.

Remark 13. So far we haven't specified the explicit covariance function. One great choice is the Radial Basis Function kernel, or RBF kernel, which is defined as:

$$C(s, t) = \exp(-\frac{\|s - t\|^2}{2\delta^2})$$

where we can choose δ^2 by ourselves. Generally we hope there is smoothness in the sample path we generate because for the initial condition, it would be better if it's twice-differentiable. According to Chapter 4 in [7], RBF generates smooth sample path almost surely.

3.5 Markov Chain Monte Carlo

Markov Chain Monte Carlo, is a class of algorithms using a stochastic process to generate samples from a probability distribution. In particular, we are constructing a Markov chain with asymptotic distribution being the probability distribution we want to draw our samples from. For basics about Stochastic Process, please refer to the Appendix 4. In the following subsection, we will first introduce Metropolis-Hastings Algorithm, the most classic MCMC algorithm. However, the classic Metropolis-Hastings algorithm is not efficient in higher dimensions, so we will introduce Gibbs' Sampler, which is an algorithm works in higher dimensions, and finally, we will include some MCMC algorithms that works in ultra high dimensions or even infinite dimensions.

First of all, we need to see why Markov Chain Monte Carlo is important to Bayesian statistics. Suppose our conditional random variable is continuous with probability density π. Then our estimate of the initial condition is either Maximum A Posteriori, which reduces to an optimization problem, or the conditional mean, which involves the computation of the integral:

$$E(X|Y=y) = \int_{\mathbb{R}^n} x\pi(x|Y=y)dx$$

However, one difficulty is to derive the explicit analytic expression of π. For some complex distributions, we usually get the information about the posterior distribution up to a constant. Moreover, in fact in Bayesian statistics, we usually assume the prior to be hierarchical model, meaning that we assume the parameters in the distribution also comes from some distributions. The following is a simple example illustrated this concept:

Example 1. In hierarchical Bayesian models, parameters are often assumed to come from distributions themselves. Consider data y_i for subject i:

$$y_i \sim \mathcal{N}(\mu, \sigma^2)$$

Here, μ and σ^2 are not fixed but are assumed to have their own prior distributions. The mean μ could be drawn from another normal distribution, and the variance σ^2 from an inverse-gamma distribution:

$$\mu \sim \mathcal{N}(\mu_0, \tau^2)$$

$$\sigma^2 \sim \text{Inverse-Gamma}(\alpha, \beta)$$

This approach allows the model to account for uncertainty in parameter estimates, providing a more flexible and robust framework for statistical inference.

Usually, hierarchical models don't have closed formula for their corresponding posterior mean or MAP, due to their complexity. In this case, we cannot directly apply traditional methods like using Riemann sum to approximate the integral. However, we will see how the Metropolis-Hastings algorithm circumvent this problem, and advanced algorithms introduced later are basically variants of Metropolis-Hastings algorithm.

3.5.1 Metropolis-Hastings Algorithm

The first algorithm we will introduce is Metropolis-Hastings Algorithm, denoted as MH-Algorithm, and is one of the first MCMC algorithm. MH-Algorithm is particularly useful when we can only know the targeted probability distribution up to a constant. In the following discussion, we assume the target distribution is continuous and has positive density everywhere, i.e., let $\pi(x) > 0$, for all $x \in \mathbb{R}^n$. Let P denote any transition kernel of a Markov kernel, and for any $x, y \in \mathbb{R}^n$, we have $p : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$:

$$\int_{\mathbb{R}^n} p(x, y)dy = 1 \tag{12}$$

Here p is understood as the "density", or the average probability that x move to an infinitesimal neighborhood around y.

Definition 9 (Reversibility)**.** A transition kernel P with density $p(x, y)$ is reversible with respect to a probability distribution μ with probability density $\pi(\cdot)$ on $\mathbb{R}^n$ if for each $x, y \in \mathbb{R}^n$:

$$\pi(x)p(x, y) = \pi(y)p(y, x)$$

where the equation above is called **detailed balanced equation**.

Corollary 4. *Let P be a probability transition kernel and let π be a probability measure. Then if P is reversible with respect to the probability measure μ with probability density π, then π is the density of an invariant probability measure of P.*

Proof. Let π be the density corresponding to μ, to see μ is invariant with respect to the transition kernel P, it's equivalent to show:

$$\mu P(B) = P(B) \iff \int_{\mathbb{R}^n} (\int_B p(x, y)dy)\pi(x)dx = \int_B \pi(y)dy$$

Now we expand both the left-handed side and right-handed side.

$$LHS = \int_{\mathbb{R}^n} \int_B p(x,y)dy\pi(x)dx = \int_{\mathbb{R}^n} \int_B p(x,y)\pi(x)dydx$$

$$RHS = \int_B \pi(y)(\int_{\mathbb{R}^n} p(y,x)dx)dy = \int_{\mathbb{R}^n} (\int_B p(y,x)\pi(y)dy)dx$$

To let LHS=RHS, we must have:

$$\int_{\mathbb{R}^n} (\int_B p(x,y)\pi(x)dy)dx = \int_{\mathbb{R}^n} (\int_B p(y,x)\pi(y)dy)dx \tag{13}$$

Since equation 13 holds for each $B \in \mathcal{B}$, we have:

$$\int_{\mathbb{R}^n} p(x,y)\pi(x)dx = \int_{\mathbb{R}^n} p(y,x)\pi(y)dx \tag{14}$$

The equation 14 is called **balanced equation**, and it holds if detailed balanced equation holds. □

Thus, if we can find an irreducible and aperiodic transition kernel with limiting distribution π be our target distribution, then we can generate samples from π by constructing a sequence of (X_n) with n being large, which is guaranteed by Theorem 28 in the Appendix. However, it's almost impossible to hit the nail on the head, so what we do is to first propose any candidate kernel Q with density $q(x,y) > 0$, for $y \neq x$, and:

$$\int_{\mathbb{R}^n} q(x,y)dy = 1$$

Now the Metropolis-Hasting algorithm is stated as the following:

Algorithm 1 Metropolis-Hastings MCMC Algorithm

1: Choose initial m_0
2: Draw a sample p from the proposal density $q(m_k, p)$
3: Compute $\pi(p), q(m_k,p), q(p,m_k)$
4: Compute the acceptance rate:

$$\alpha = \min(1, \frac{\pi(p)q(p,m_k)}{\pi(m_k)q(m_k,p)})$$

5: Accept $m_{k+1} = p$ with probability α or reject it with probability $1-\alpha$ and set $m_{k+1} = m_k$

If we consider $p(x,y) = \alpha(x,y)q(x,y)$, then $p(x,y)$ could be viewed as the corrected density from $q(x,y)$. If we denote P to be the probability transition kernel with density $p(x,y)$, then we have the following result:

Theorem 13. *The Metropolis-Hastings algorithm produces a Markov chain $\{X_n\}$ with π being the probability density of its invariant measure.*

Proof. Assume $x \neq y$, then:

$$\begin{aligned}
\pi(x)p(x,y) &= \pi(x)\alpha(x,y)q(x,y) \\
&= \pi(x)q(x,y)\min(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}) \\
&= \min(\pi(x)q(x,y), \pi(y)q(y,x)) \\
&= \pi(y)q(y,x)\min(1, \frac{\pi(x)q(x,y)}{\pi(y)q(y,x)})) \\
&= \pi(y)\alpha(y,x)q(y,x) = \pi(y)p(y,x)
\end{aligned}$$

Now if $x = y$, then the detailed balanced equation trivially holds and both sides are equal to $\pi(x)q(x,x)$ □

However, recall that the Ergodic Theorem 28 in the Appendix requires the transition kernel P to be both irreducible and aperiodic. First, to see irreducibility:

Theorem 14. *If the candidate kernel $q(\cdot,\cdot)$ is positive and continuous over $\mathbb{R}^d \times \mathbb{R}^d$, and π is finite everywhere, then the transition kernel P, whose density is $p(x,y) = \alpha(x,y)q(x,y)$, is irreducible.*

Proof. Let the target probability measure be μ with density π in $\mathbb{R}^n$. Let $A \in \mathcal{B}$ so that $Leb(A) > 0$. Now consider $P(x, A)$. We first divide A into the following disjoint sets:

$$S_1 = \{y \in A : \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \geq 1, \pi(x) \geq \pi(y)\}, \quad S_2 = \{y \in A : \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \geq 1, \pi(x) \leq \pi(y)\}$$

$$S_3 = \{y \in A : \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \leq 1, \pi(x) \geq \pi(y)\}, \quad S_4 = \{y \in A : \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \leq 1, \pi(x) \leq \pi(y)\}$$

Because $q(x, y)$ is assumed to be positive and continuous, we have:

$$\inf_{y \in A} \min(q(x,y), q(y,x)) \geq \epsilon > 0$$

Consider $P(x, A)$:

$$\begin{aligned}
P(x,A) &= \int q(x,y) \min(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}) dy \\
&\geq \int_{S_1} \epsilon \cdot 1 dy + \int_{S_2} \epsilon \cdot 1 dy + \int_{S_3} \frac{\pi(y)}{\pi(x)} \cdot q(y,x) dy + \int_{S_4} \frac{\pi(y)}{\pi(x)} \cdot q(y,x) dy \\
&\geq \int_{S_1} \epsilon \cdot 1 dy + \int_{S_2} \epsilon \cdot 1 dy + \frac{1}{\pi(x)} \int_{S_3} q(y,x) \cdot \pi(y) dy + \int_{S_4} 1 \cdot \epsilon dy \\
&\geq \epsilon \cdot Leb(S_1 \cup S_2 \cup S_4) + \frac{\epsilon}{\pi(x)} \cdot \mu(S_3)
\end{aligned}$$

Now if $Leb(S_1 \cup S_2 \cup S_3) > 0$, we are done. Otherwise $Leb(S_4) > 0$, and recall that we start off by assuming $\pi(x) > 0$ for every $x \in \mathbb{R}^n$, meaning that μ is absolutely continuous with respect to the Lebesgue measure, so $\mu(S_4) > 0$, from which we conclude that the Markov chain is irreducible. □

Now to see aperiodicity, we have:

Theorem 15. *If the candidate kernel $q(\cdot, \cdot)$ is positive and continuous over $\mathbb{R}^d \times \mathbb{R}^d$, and π is finite everywhere, then the transition kernel P, whose density is $p(x, y) = \alpha(x, y)q(x, y)$, is aperiodic.*

Proof. Suppose E_1 and E_2 are disjoint subsets of $\mathbb{R}^n$, and we prove by contradiction. Suppose for any $x \in E_1$, $\mathbb{P}(x, E_2) = 1$, then for any $x \in E_1$, $\mathbb{P}(x, E_1) = 0$. However for any $x \in E_1$, we also have:

$$\mathbb{P}(x, E_1) = \int_{y \in E_1} q(x,y)\alpha(x,y) dy$$

Since on the right hand side, it is the integral of a positive function over a set with positive measure, we have $\mathbb{P}(x, E_1) > 0$, contradicition. □

Thus we proved the aperiodicity and irreducibility of the transition kernel P and everything stated in Theorem 28 applies. Besides generating samples from the target distribution, we can also calculate the conditional mean using the sample:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} X_j = \int_{\mathbb{R}^n} x\mu(dx)$$

Remark 14. The following is some conventions in the practical use of MCMC:

1. First, the candidate kernel we propose is often symmetric, and acceptance rate can be simplified as:

   $$\alpha = \min(1, \frac{\pi(p)}{\pi(m_k)})$$

   Thus we somehow reduce the number of operations in the algorithm. For example, we usually use the Gaussian kernel, which is positive and symmetric:

   $$q(y,x) = \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left(-\frac{1}{2\gamma^2} \|m - p\|^2\right)$$

2. Since our target distribution is the asymptotic distribution of our Markov Chain, we often mannually delete some early sample points.

Remark 15. However, the assumption that the density $\pi(x) > 0$ may not always be true, and the following is some possible generalizations:

1. Suppose $\pi(x) = 0$ for some x, and very unfortunately the starting point m_0 has $\pi(m_0) = 0$. Consequently, the denominator of $\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}$ in the expression of $\alpha(x, y)$ is 0. In this case, we let $\alpha(x, y) = 1$. We can see that such a definition will not affect the proof of Theorem 15 and the inequalities in the proof of Theorem 14 still hold. As a result, the ergodicity still holds. Now starting from m_0, if the candidate kernel generate another point y and $\pi(y) = 0$, we will reject y deterministically, until the new sample fall into the support of π.
2. As we mentioned at the beginning of the section, it's possible that we know our distribution up to a constant, but intuitively, Metropolis-Hastings Algorithm still works because the constant in the denominator and numerator are cancelled out in $\alpha(x, y)$. We can easily adapt the results and proofs above to this case. For technical details, we refer to [8].

Remark 16. Although in our introduction of Metropolis-Hastings algorithm, we assume there is no "inertia", i.e., 12 holds. However, let P denote any transition kernel. Sometimes for a point $x \in \mathbb{R}^n$, x could stay (without going to other states) with positive probability, so let $B \in \mathcal{B}$, we usually decompose $P(x, B)$ as:

$$P(x, B) = \int_B K(x, y)dy + r(x)\chi_B(x) \tag{15}$$

In this case, Theorem 13 still holds, that is, for any proposed kernel Q, MH-Algorithm still produces a Markov chain with the target density π being the density of the invariant probability measure. To prove this, notice that if $y = x$, we simply change the notation $p(x, x)$ to $r(x)$. Finally we will show that, if the proposed kernel Q has decomposition as 15 above, the ergodic results still hold, i.e., our Metropolis-Hastings algorithm will still converge.

Theorem 16. *Let P be a probability transition kernel with decomposition as 15 and let π be a probability measure. Then if P is reversible with respect to the probability measure μ with probability density π, then π is the density of an invariant probability measure of P.*

Proof. In order to prove μ is an invariant measure, it's equivalent to show that:

$$\pi P(B) = P(B) \iff \int_{\mathbb{R}^n} P(x, B)\pi(x)dx = \int_B \pi(y)dy \tag{16}$$

Notice on the left-hand side:

$$\begin{aligned}
\int_{\mathbb{R}^n} P(x, B)\pi(x)dx &= \int_{\mathbb{R}^n} (\int_B K(x, y)dy + r(x)\chi_B(x))\pi(x)dx \\
&= \int_{\mathbb{R}^n} (\int_B K(x, y)\pi(x)dy)dx + \int_{\mathbb{R}^n} r(x)\chi_B(x)\pi(x)dx \\
&= \int_B (\int_{\mathbb{R}^n} K(x, y)\pi(x)dx)dy + \int_{\mathbb{R}^n} r(y)\chi_B(y)\pi(y)dy \\
&= \int_B (\int_{\mathbb{R}^n} K(x, y)\pi(x)dx)dy + \int_B r(y)\pi(y)dy \\
&= \int_B (\int_{\mathbb{R}^n} K(x, y)\pi(x)dx)dy + \int_B (1 - \int_{\mathbb{R}^n} K(y, x)dx)\pi(y)dy
\end{aligned}$$

Now to satisfy the equality in 16, we must have:

$$\int_B (\int_{\mathbb{R}^n} K(x, y)\pi(x)dx)dy = \int_B \int_{\mathbb{R}^n} (K(y, x)dx)\pi(y)dy \tag{17}$$

then by reversibility, the detailed balanced equation holds, and equation 17 also holds. □

Now consider the Metropolis-Hastings algorithm introduced in 3.5.1.

Theorem 17. *If the candidate kernel has decomposition as 15, then the modified transition kernel P is still irreducible and aperiodic and has μ as the asymptotic invariant measure.*

Proof. To see the irreducibility holds, let the target probability measure be μ with density π in $\mathbb{R}^n$. Let $A \in \mathcal{B}$ so that $Leb(A) > 0$. Now consider $P(x, A)$. First, if $x \in A$, then:

$$P(x, A) > r(x) = 0$$

Now if $x \notin A$, the proof follows from Theorem 14. Next to see the aperiodicity, similar as our modified proof of irreducibility, in the proof of Theorem 15, if $x \in E_1$ has $r(x) > 0$, then:

$$P(x, E_1) > r(x) > 0$$

so aperiodicity still holds. Finally to see π is indeed the invariant measure, the proof is identical to Theorem 13, except that if for some x so that $r(x) > 0$, if $x = y$, then the detailed balanced equation trivially holds and both sides are equal to $\pi(x)r(x)$. □

3.5.2 Gibbs Sampler

Although as we mentioned, Metropolis-Hastings algorithm works in the situation that we only know the probability density up to a constant multiplication C, itself is still not ideal to tackle with the sampling in high dimensions. Suppose the proposed transition kernel Q has density concentrated on different regions from the the true P with limiting distribution being π, the algorithm could have very slow convergence due to low acceptance rate.

In fact, the problem above belongs to the concept *curse of dimensionality*, introduced by Richard E. Bellman, which refers to the set of problems that arise when analyzing and organizing data in high-dimensional spaces. As the dimensionality n of a space increases, phenomena such as the exponential growth of volume lead to data becoming sparse, which is problematic for methods relying on numerical integration or optimization.
Thus, we introduce another type of MCMC that is in particular used to deal with the curse of dimensionality. Let π be the density of the distribution we want to sample from. Now suppose we have good methods to generate samples from some low-dimensional conditional distributions. We will make ourselves clear what "low-dim conditional distributions" are in the following discussion.

First, we need some notations to denote how we "partition" the dimensions n. Let $I = \{1, 2, \dots, n\}$ be the index set, and let $\{I_j\}_{j=1}^m$ be a collection of nonempty disjoint sets so that $I = \bigcup_{j=1}^m I_j$. Let k_j be the number of elements in I_j, then we may partition $x \in \mathbb{R}^n$ as:

$$x = [x_{I_1}, x_{I_2}, \dots, x_{I_m}] \in \mathbb{R}^{k_1} \times \mathbb{R}^{k_2} \times \cdots \mathbb{R}^{k_m}$$

Next, we define x_{-I_j} to be a vector in $\mathbb{R}^{n-k_j}$ so that

$$x_{-I_j} = [x_{I_1}, x_{I_2}, x_{I_{j-1}}, x_{I_{j+1}}, \dots, x_{I_m}]$$

Similarly, we define X_{I_j} to be a random vector in $\mathbb{R}^{k_j}$, and $X_{-I_j} = [X_{I_1}, X_{I_2}, X_{I_{j-1}}, X_{I_{j+1}}, \dots, X_{I_m}]$. Now we can write the conditional density as:

$$\pi(X_{I_j} = x_{I_j} | X_{-I_j} = x_{-I_j}) = \frac{\pi(x_{I_1}, x_{I_2}, \dots, x_{I_m})}{\int_{\mathbb{R}^{k_j}} \pi(x_{I_1}, \dots, x_{I_{j-1}}, x_{I_{j+1}}, \dots, x_{I_m}) dx_{I_j}} = C_j \pi(x_{I_1}, x_{I_2}, \dots, x_{I_m})$$

We first state the explicit algorithm of Gibbs sampler and then explain its mechanism:

Algorithm 2 Gibbs Sampler

1: Pick the initial value $x_1 \in \mathbb{R}^n$ and set $k = 1$
2: Set $x = x_k$. For $1 \le j \le m$, draw $y_{I_j} \in \mathbb{R}^{k_j}$ from $\pi(y_{I_j} | y_{I_1}, \cdots, y_{I_{j-1}}, x_{I_{j+1}}, \cdots, x_{I_m})$.
3: Set $x_{k+1} = y$. Stop when $k = K$, the desired sample size. Else, $k \to k + 1$ and repeat Step 2.

Remark 17. In Step (2) above, we didn't specify how to generate samples from $\pi(y_j | y_1, \cdots, y_{j-1}, x_{j+1}, \cdots, x_m$, and this is essentially what the "low-dimensional conditional distribution" really is, and we should assume that we already have good sampling methods at this step. If not, of course we can use MH algorithm each time, and this is known as *Metropolis within Gibbs*.

The Algorithm 2 above could be viewed as a Markov chain with transition kernel P whose density is:

$$p(x, y) = \prod_{j=1}^{m} \pi(y_{I_j} | y_{I_1}, \cdots, y_{I_{j-1}}, x_{I_{j+1}}, \cdots, x_{I_m}) \tag{18}$$

Theorem 18. *For P being a transition kernel with density p defined in equation 18, we have:*

1. $$\int_{\mathbb{R}^n} p(x, y) dy = 1, \quad \forall y \in \mathbb{R}^n$$
2. *Recall π is the target density, then π is the density of the invariant measure with respect to the transition kernel P.*

Proof. To see (1), first notice that:

$$\int_{\mathbb{R}^{k_i}} \pi(x_{I_i} | x_{I_1}, \dots, x_{I_{i-1}}, y_{I_{i+1}}, \dots, y_{I_m}) dx_i = 1$$

Now for all $i = 1, \ldots, m$. It follows that:

$$\begin{aligned}\int_{\mathbb{R}^{k_m}} p(y,x)dx_{I_m} &= \int_{\mathbb{R}^{k_m}} \prod_{i=1}^{m} \pi(x_{I_i}|x_1, \ldots, x_{I_{i-1}}, y_{I_{i+1}}, \ldots, y_{I_m})dx_{I_m} \\ &= \prod_{i=1}^{m-1} \pi(x_{I_i}|x_{I_1}, \ldots, x_{I_{i-1}}, y_{I_{i+1}}, \ldots, y_{I_m}) \int_{\mathbb{R}^{k_m}} \pi(x_{I_m}|y_{I_1}, \ldots, y_{I_{m-1}})dx_{I_m} \\ &= \prod_{i=1}^{m-1} \pi(x_{I_i}|x_{I_1}, \ldots, x_{I_{i-1}}, y_{I_{i+1}}, \ldots, y_{I_m})\end{aligned}$$

Inductively, by integrating always with respect to the last block of x, we obtain

$$\int_{\mathbb{R}^n} p(y,x)dx = 1.$$

To see that π is the density of an invariant measure, we need to prove:

$$\int_{\mathbb{R}^n} \pi(x)p(x,y)dx = \pi(y) \tag{19}$$

We will proceed by integrating component-wise using Fubini's Theorem. First by substituting p according to the definition, we see that:

$$\begin{aligned}\int_{\mathbb{R}^{k_1}} \pi(x)p(x,y)dx_{I_1} &= \int_{\mathbb{R}^{k_1}} \pi(x) \prod_{i=1}^{m} \pi(y_i|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})dx_{I_1} \\ &= \prod_{i=1}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m}) \int_{\mathbb{R}^{k_1}} \pi(x)d_{x_{I_1}} \\ &= \prod_{i=1}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\pi(x_{I_2}, \cdots, x_{I_m}) \\ &= \left(\prod_{i=2}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_1}|x_{I_2}, \cdots, x_{I_m})\pi(x_{I_2}, \cdots, x_{I_m}) \\ &= \left(\prod_{i=2}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_1}, x_{I_2}, \ldots, x_{I_m})\end{aligned}$$

Now we integrate with respect to x_{I_2}:

$$\begin{aligned}&\int_{R^{k_2}} \left(\prod_{i=2}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_1}, x_{I_2}, \ldots, x_{I_m})dx_{I_2} \\ =& \left(\prod_{i=2}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \int_{R^{k_2}} \pi(y_{I_1}, x_{I_2}, \ldots, x_{I_m})dx_{I_2} \\ =& \left(\prod_{i=2}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_1}, x_{I_3}, \cdots, x_{I_m}) \\ =& \left(\prod_{i=3}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_2}|y_{I_1}, x_{I_3}, \ldots, x_{I_m})\pi(y_{I_1}, x_{I_3}, \cdots, x_{I_m}) \\ =& \left(\prod_{i=3}^{m} \pi(y_{I_i}|y_{I_1}, \ldots, y_{I_{i-1}}, x_{I_{i+1}}, \ldots, x_{I_m})\right) \pi(y_{I_1}, y_{I_2}, x_{I_3} \ldots, x_{I_m})\end{aligned}$$

We proceed this inductively and finally we can see that the Equation 19 holds. □

3.6 Further Topics: Other Bayesian Approaches

Although the recovery of initial condition using Gaussian process in Section 3.4 gives us a smooth recovery for any point in $(0, 1)$, sometimes it's too "smooth", because generally the initial condition doesn't have to be a function that is differentiable for all degrees of differentiation. To this end, we naturally appeal to Sobolev spaces introduced in 4, which contain functions with the exact regularity we want. Then the construction of prior is essentially assigning a probability measure over Sobolev spaces. For more details, we refer to the work done by Dashti and Stuart[4]. Some set-up in Introduction 1 and concepts in the Appendix 4 would be helpful for the understanding.

4 Appendix

4.1 Linear Algebra

Theorem 19 (SVD for real matrices)**.** *Every matrix $A \in \mathbb{R}^{m\times n}$ allows a decomposition:*

$$A = UWV^T$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and if $m \le n$:

$$\Lambda = \begin{bmatrix} w_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & & & & \\ 0 & \cdots & & w_m & 0 & \cdots & 0 \end{bmatrix} = [diag(w_1, \ldots, w_m), 0]$$

where $w_1 \ge \cdots \ge w_n$ are called singular values. If $m \ge n$, then $W = \begin{bmatrix} diag(w_1, \ldots, w_n) \\ 0 \end{bmatrix}$ We can rewrite the decomposition as a sum of tensor products:

$$A = \sum_{j=1}^{n} w_j u_j v_j^T$$

Proof. (Existence of SVD) Let $A \in \mathbb{R}^{m\times n}$, and because A^TA is positive semidefinite (thus symmetric), we have the eigendecomposition:

$$A = Q\Lambda Q^T$$

where the diagonal of Λ are nonnegative eigenvalues of A^TA in decreasing order. Let $w_i = \sqrt{\lambda_i}$. Furthermore, because A^TA is symmetric, Q can choose to be orthonormal, and we let $V = Q$, and $W_{ij} = (\Lambda_{ij})^{1/2}$. We have found Λ and V. Now first suppose $r = m$, where r is the rank of Λ. we define the columns of U, $u_j = \frac{Av_j}{w_j}$, so there are totally m u_js. By our definition:

$$Av_i = w_j u_j, \forall j \iff AV = UW$$

Plug in the results above, we have:

$$UWV^T = AVV^T = A$$

Because V are constructed to be orthogonal. For the case when $r < m$, i.e., there are zero eigenvalues, we simply choose $m - r$ vectors to be the columns of U, and they are chosen to be the orthonormal basis of the orthogonal complement of $\text{Span}\{Av_j/w_j\}$ in $\mathbb{R}^m$. □

4.2 Functional Analysis

Definition 10. A normed space $(V, \|\cdot\|)$ is a real vector space V together with a norm, which is a function $\|\cdot\| : V \to [0, \infty)$ satisfying the following three properties:

(i) $\|v\| \ge 0$ for all $v \in V$ with equality if and only if $v = 0$ (positivity);

(ii) For all $\alpha \in \mathbb{R}$ and $v \in V$, we have that $\|\alpha v\| = |\alpha|\|v\|$ (homogeneity);

(iii) For all $v, w \in V$, we have that $\|v + w\| \le \|v\| + \|w\|$ (subadditivity).

Definition 11. An inner product space $(V, \langle\cdot,\cdot\rangle)$ is a real vector space V together with an inner product, which is a function $\langle\cdot,\cdot\rangle : V \to \mathbb{R}$ satisfying the following three properties:

(i) $\langle v, v\rangle \ge 0$ for all $v \in V$ with equality if and only if $v = 0$ (positive definiteness);

(ii) For all $v, w \in V$, we have that $\langle v, w\rangle = \langle w, v\rangle$ (symmetry);

(iii) For all $\alpha \in \mathbb{R}$ and $u, v, w \in V$, we have that $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$ and $\langle v + u, w\rangle = \langle v, w\rangle + \langle u, w\rangle$ (linearity).

Definition 12. A Banach space is a complete normed vector space.

Definition 13. A Hilbert space is a complete inner product space.

Theorem 20. *Let $A \in \mathbb{R}^{n\times n}$. Then:*

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

is well-defined, that is the series converges.

Proof. Notice if $\|\cdot\|$ defines a matrix norm, then for any $A, B \in \mathbb{R}^{n\times n}$ we have: $\|AB\| \leq \|A\|\|B\|$, and then by induction, we have:

$$\|\frac{A^n}{n!}\| = \frac{1}{n!}\|A^n\| \leq \frac{1}{n!}\|A\|^n$$

Since for any real number x, $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges:

$$\|e^A\| \leq \sum_{n=0}^{\infty} \frac{\|A\|^n}{n!} < \infty$$

Thus e^A converges in norm, and since A is a linear operator between finite dimensional space, then e^A converges to a point in $\mathbb{R}^{n\times n}$. □

Theorem 21 (Riesz's representation)**.** *Let $l : H \to \mathbb{R}$ be a bounded linear functional on a Hilbert space H, i.e.,*

$$|l(x)| \leq C\|x\|$$

Then there exists a unique point $y \in H$ so that:

$$l(x) = \langle x, y\rangle$$

An important result of Riesz's representation theorem is Lax-Milgram theorem, which is used in the proof of the existence of unique solution for Tikhonov regularization.

Theorem 22 (Lax-Milgram)**.** *Let H be a Hilbert space, and let $B : H \times H \to \mathbb{R}$ be a bilinear form such that:*

- *(Bounded) $|B(x, y)| \leq C\|x\|\|y\|, \forall x, y \in H$*
- *(Coercive) $B(x, x) \geq c\|x\|^2, \forall x \in H$*

for some constants $0 \leq c \leq C < \infty$. Then there exists a unique bounded invertible map $T : H \to H$ so that $B(x, y) = \langle x, Ty\rangle, \forall y \in H$.

Theorem 23. *Let H_1 and H_2 be real Hilbert spaces and suppose that $A : H_1 \to H_2$ is a bounded linear operator. Then there exists a unique bounded linear operator $A^* : H_2 \to H_1$, called the adjoint of A, satisfying $\langle Ax, y\rangle_{H_2} = \langle x, A^*y\rangle_{H_1}$. Moreover, $\|A\|_{H_1\to H_2} = \|A^*\|_{H_2\to H_1}$.*

Proof. Let $y \in H_2$ and consider $T_y : H_1 \to \mathbb{R}, x \mapsto \langle Ax, y\rangle_{H_2}$. Clearly, T_y is linear and bounded so by the Riesz representation theorem there exists a unique $z \in H_1$ s.t.

$$\langle Ax, y\rangle_{H_2} = T_y(x) = \langle x, z\rangle_{H_1} \text{ for all } x \in H_1.$$

Define $A^*y := z$. First we prove that the defined A^* is indeed linear. Let $a, b \in \mathbb{R}$ and $y_1, y_2 \in H_2$.

$$\begin{aligned}
\langle x, A^*(ay_1 + by_2)\rangle_{H_1} &= \langle Ax, ay_1 + by_2\rangle_{H_2} \\
&= a\langle Ax, y_1\rangle_{H_2} + b\langle Ax, y_2\rangle_{H_2} \\
&= a\langle x, A^*y_1\rangle_{H_1} + b\langle x, A^*y_2\rangle_{H_1} \\
&= \langle x, aA^*y_1 + bA^*y_2\rangle_{H_1}
\end{aligned}$$

Since $x \in H_1$ was arbitrary, $A^*(ay_1 + by_2) = aA^*y_1 + bA^*y_2$. To see A^* is bounded and $\|A\| = \|A^*\|$, notice that

$$\begin{aligned}
\|A^*\|_{H_2\to H_1} &= \sup_{\|y\|_{H_2}\leq 1} \|A^*y\|_{H_1} \\
&= \sup_{\|y\|_{H_2}\leq 1} \sup_{\|x\|_{H_1}\leq 1} |\langle A^*y, x\rangle| \\
&= \sup_{\|y\|_{H_2}\leq 1} \sup_{\|x\|_{H_1}\leq 1} |\langle y, Ax\rangle| \\
&= \sup_{\|x\|_{H_1}\leq 1} \|Ax\|_{H_2} = \|A\|_{H_1\to H_2} < \infty
\end{aligned}$$

Definition 14. Let H_1 and H_2 both be a Hilbert space. A linear operator $A : H_1 \to H_2$ is called positive if for any $x \in H_1$, $\langle Ax, x\rangle_V \geq 0$.

Definition 15. Let H_1 and H_2 both be a Hilbert space. A linear operator $A : H_1 \to H_2$ is called positive if for any $x \in H_1$, $y \in H_2$, $\langle Ax, y\rangle_{H_2} = \langle x, A^*y\rangle_{H_1}$.

□

4.3 Sobolev Spaces

We will introduce Hilbert scales and its connections to the Sobolev spaces. First, let's recall concepts of Sobolev spaces.

Definition 16 (Weak Derivative)**.** Let f be a locally integrable function on some domain $D \subset \mathbb{R}^n$. A function g is called the weak derivative of f if for every test function ϕ (smooth functions with compact support contained in D), the following condition holds:

$$\int_D f(x) \frac{\partial \phi}{\partial x_i} dx = -\int_D g(x)\phi(x) dx$$

Notice that the notion of weak derivative comes from integration by parts, because:

$$\int_D g(x)\phi(x) dx = [f(x)\phi(x)]_D - \int_D f(x) \frac{\partial \phi}{\partial x_i} dx$$

From the expressions above, it's clear that any derivative is a weak derivative.

Definition 17 (Sobolev space, $W^{r,p}(D)$)**.** Denote D^α to be the α-weak derivative of u. Then the Sobolev space consists of all L^p-integrable functions $u : D \mapsto \mathbb{R}$ whose α^{th} order weak derivatives exist and are L^p-integrable for all $|\alpha| \leq r$, i.e.,

$$W^{r,p} = \{u | D^\alpha u \in L^p(D) \text{ for } |\alpha| \leq r\}$$

with norm:

$$\|u\| = \sum_{|\alpha| \leq r} (\|D^\alpha u\|^p_{L^p(D)})^{1/p}$$

for $1 \leq p < \infty$ and for $p = \infty$:

$$\|u\| = \sum_{|\alpha| \leq r} \|D^\alpha u\|_{L^\infty(D)}$$

If $p = 2$, we denote $W^{r,2}(D)$ by $H^r(D)$.

Remark 18. Both H^1 and H^1_0 are separable Hilbert spaces.

Remark 19. Notice for H^1, we can simply express the inner product as:

$$\langle u, v \rangle_{H^1} = \langle u, v \rangle_{L^2(D)} + \langle \nabla u, \nabla v \rangle_{L^2(D)}$$

and for H^0_1, we can express the inner product as:

$$\langle u, v \rangle_{H^1_0} = \langle \nabla u, \nabla v \rangle_{L^2(D)}$$

4.4 Laplacian over $L^2(D)$

4.4.1 Weyl's Law

Basically, Weyl's law describes the asymptotic distribution of the number of eigenvalues within the ball with radius r with respect to the radius r.

Theorem 24 (Weyl's law)**.** *For a bounded domain $D \subset R^d$ and for the Laplacian operator with Dirichlet or Neumann boundary conditions, the number of eigenvalues less than or equal to a given value λ behaves asymptotically as:*

$$N(\lambda) \sim \frac{Vol(D)}{(4\pi)^{d/2}\Gamma(1 + d/2)} \lambda^{d/2}, \text{ as } \lambda \to \infty$$

Proof. For the proof, we refer to the proof in Chapter 6, Theorem 16 in [3], and the gist of the proof is to approximate the domain by triangles and rectangles, then obtain Weyl's law in the domain by using eigenvalue monotonicity results and Weyl's law for rectangles, which is established by explicit computation. □

4.4.2 Hilbert Scales

We introduce Hilbert scale because we try to describe where the solution to the heat equation for A over H lives in.

Definition 18 (Hilbert Scales $\mathcal{H}^s$)**.** Consider a positive self-adjoint operator A on an arbitrary bounded open set $D \subset \mathbb{R}^d$ with Lipschitz boundary. Define $H = L^2(D)$ and $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$. Moreover,we have the following assumptions on the eigenvalue problems: $A\phi = \alpha\phi$:

1. has countably many solutions.
2. $\{\phi_j\}$ satisfies the L^2-orthonormality and form a basis for H
3. $\alpha_j \asymp j^{2/d}$ (there exists C_1, C_2 so that $C_1 j^{2/d} \leq \alpha_j \leq C_2 j^{2/d}$)

Then for any $u \in H$, it allows the orthonormal expansion:

$$u = \sum_{j=1}^{\infty} \langle u, \phi_j \rangle \phi_j$$

Then we define the spaces $\mathcal{H}^s = \mathcal{D}(A^{s/2})$ for $s > 0$ to be:

$$\mathcal{H}^s = \left\{ u : D \to \mathbb{R} | \|u\|_{\mathcal{H}^s}^2 < \infty \right\}$$

where,

$$\|u\|_{\mathcal{H}^s}^2 = \sum_{j=1}^{\infty} j^{2s/d} |u_j|^2, \ u_j = \langle u, \phi_j \rangle.$$

In fact, $\mathcal{H}^s$ is a Hilbert space and if we define:

$$\langle u_j, v_j \rangle_{\mathcal{H}^f} = \sum_{j=1}^{\infty} u_j v_j$$

Then it's a subspace in the original Hilbert space H.

Lemma 5. *Let D be an bounded open set. Define $H = L^2(D)$ and $A = -\Delta$, with domain defined to be $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$. Then A is positive and self-adjoint.*

Proof. First consider any $u \in \mathcal{D}(A)$:

$$\langle -\Delta u, u \rangle = 0 - (-\int_D \nabla u \cdot \nabla u dx) = \int_D \nabla u \cdot \nabla u dx = \|\nabla u\|_{L^2(D)}^2 \geq 0$$

Now to see it's self-adjoint, first we see it's symmetric:

$$\langle -\Delta u, v \rangle = \int_D -\Delta u \cdot v dx = \int_D \nabla u \cdot \nabla v = \langle u, -\Delta v \rangle$$

To show $\mathcal{D}(A) = \mathcal{D}(A^*)$, we use results from *Extensions of Symmetric Operator* from Wikipedia page. □

From the lemma, we can well-define our Hilbert scales for negative Laplacian.

4.5 Probability Theory

4.5.1 Measure Theory

Definition 19 (Absolute Continuity)**.** Let μ, ν both be measures over the measurable space $(X, \mathcal{M})$. Then ν is absolutely continuous with respect to μ and denoted as:

$$\nu \ll \mu$$

If for every $E \in \mathcal{M}$ so that $\mu(E) = 0$, we have $\nu(E) = 0$.

Definition 20 (Radon-Nikodym derivative)**.** Let $f : X \to \mathbb{R}$ be a measurable function. Then f is called the Radon-Nikodym derivative of ν with respect μ if for every measurable set μ-measurable set $A \subset X$, we have:

$$\nu(A) = \int_A f(x) d\mu(x)$$

Theorem 25. *Suppose ν and μ be two probability measures on measurable space $(X, \mathcal{M})$ such that $\nu \ll \mu$. If $g \in L^1(\nu)$, then $g\frac{d\nu}{d\mu} \in L^1(\mu)$ and:*

$$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu$$

Proof. The proof refers to Proposition 3.9. in Folland's Real Analysis. [5] □

4.5.2 Gaussian measure on $\mathbb{R}^n$

Definition 21. Let $x_0 \in \mathbb{R}^n$ and $\Gamma \in \mathbb{R}^{n\times n}$ be a positive definite matrix (denote as $\Gamma > 0$) in the sequel. A (non-degenerate) Gaussian n-variate random variable X with mean x_0 and covariance Γ is a random variable with the probability density:

$$\pi(x) = (\frac{1}{2\pi|\Gamma|})^{n/2} \exp\big(-\frac{1}{2}(x-x_0)^T\Gamma^{-1}(x-x_0)\big)$$

where $|\Gamma| = \det(\Gamma)$. In such case, we use the notation: $X \sim \mathcal{N}(x_0, \Gamma)$

Remark 20. Notice that Gaussian random variable is often defined through Fourier transform, or in Probability Theory, we call it *Characteristic function*. More specifically, a random variable X is called Gaussian if:

$$E\{\exp(-i\xi^T X)\} = \exp(-i\xi^T x_0 - \frac{1}{2}\xi^T\Gamma\xi)$$

This formulation could be superfluous in finite-dimensions. However, in infinite-dimensional setting, as we will see later, there's no notion of probability density, meaning that Definition 21 fails. However, in separable Hilbert space, the definition according to Characteristic function is still valid.

Definition 22. Let

$$\Gamma = \begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix} \in \mathbb{R}^{n\times n}$$

be a positive definite symmetric matrix, where $\Gamma_{11} \in \mathbb{R}^{k\times k}$, $\Gamma_{22} \in \mathbb{R}^{(n-k)\times(n-k)}, k, n$, and $\Gamma_{21} = \Gamma_{12}^T$. We define the Schur Complements $\tilde{\Gamma}_{jj}$ of $\Gamma_{jj}, j = 1, 2$, by the formulas:

$$\tilde{\Gamma}_{22} = \Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}, \tilde{\Gamma}_{11} = \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}$$

Since Γ is positive-definite, it's invertible, and more specifically:

Lemma 6. *Let Γ be a matrix in Definition 22. Then:*

- *The Schur complements $\tilde{\Gamma_{jj}}$ are invertible for $j = 1, 2$*
- *The inverse of Γ is given by:*

$$\Gamma^{-1} = \begin{bmatrix}\tilde{\Gamma}_{22}^{-1} & -\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1}\\ -\tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \tilde{\Gamma}_{11}^{-1}\end{bmatrix}$$

Proof. The proof is referred to [6] Lemma 3.4. in Chapter 3. □

Theorem 26. *Let $X : \Omega \to \mathbb{R}^n$ and $Y : \Omega \to \mathbb{R}^k$ be two Gaussian random variables whose joint probability density $\pi : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}_+$ is of the form:*

$$\pi(x, y) \propto \exp(-\frac{1}{2}\begin{bmatrix}x-x_0\\ y-y_0\end{bmatrix}^T\begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix}^{-1}\begin{bmatrix}x-x_0\\ y-y_0\end{bmatrix}) \tag{20}$$

Then the probability distribution of X conditioned on $Y = y$, $\pi(x|y) : \mathbb{R}^n \to \mathbb{R}_+$ is of the form

$$\pi(x|y) \propto \exp\left(-\frac{1}{2}(x-\bar{x})^T\Gamma_{22}^{-1}(x-\bar{x})\right),$$

where

$$\bar{x} = x_0 + \Gamma_{12}\Gamma_{22}^{-1}(y - y_0).$$

Proof. The proof refers to Theorem 3.5. in [6], in Chapter 3. □

Theorem 27. *Let X and Y be Gaussian random variables with joint probability density given by 20. Then the marginal density of X is*

$$\pi(x) = \int_{\mathbb{R}^k} \pi(x, y)\, dy \propto \exp\left(-\frac{1}{2}(x-x_0)^T\Gamma_{11}^{-1}(x-x_0)\right).$$

Proof. The proof refers to Theorem 3.6. in [6], in Chapter 3. □

4.5.3 Stochastic Process

Definition 23 (Probability Transition Kernel)**.** Let $\mathcal{B} = \mathcal{B}(\mathbb{R}^n)$ denote the Borel sets over $\mathbb{R}^n$. A mapping $P : \mathbb{R}^n \times \mathcal{B} \to [0, 1]$ is called a probability transition kernel if:

1. For each $B \in \mathcal{B}$, the mapping $\mathbb{R}^n \to [0, 1]$, $x \mapsto P(x, B)$ is a measurable function;
2. For each $x \in \mathbb{R}^n$, the mapping $\mathcal{B} \to [0, 1]$, $B \mapsto P(x, B)$ is a probability distribution.

Finally, denote μ_{X_j} to be the probability measure of X_j.

Definition 24 (Discrete-Time Markov Chain)**.** A discrete-time stochastic process is an ordered set $\{X_j\}_{j=1}^{\infty}$ of random variables $X_j \in \mathbb{R}^n$.

Definition 25 (Time-homogeneous Markov Chain)**.** A time-homogeneous Markov Chain with the transition kernel P is a stochastic process $\{X_j\}_j$ with the properties:

$$\mu_{X_{j+1}}(B_{j+1}|x_1, \ldots, x_j) = \mu_{X_{j+1}}(B_{j+1}|x_j) = P(x_j, B_{j+1})$$

The notation $P^{(k)}$ is defined inductively through the expression:

$$\begin{aligned} P^{(k)}(x_j, B_{j+k}) &= \mu_{X_{j+k}}(B_{j+k}|x_j) \\ &= \int_{\mathbb{R}^n} P(x_{j+k-1}, B_{j+k}) P^{(k-1)}(x_j, dx_{j+k-1}), k \geq 2 \end{aligned}$$

where $P^{(k-1)}(x_j, dx_{j+k-1})$ represent the probability that x_j goes to an infinitesimal neighborhood around x_{j+k-1}. Moreover, this is the probability density of x_{j+k-1} conditioned on x_j, and in particular, for $k = 1$:

$$P^{(1)}(x_j, B_{j+1}) = P(x_j, B_{j+1})$$

Following the notation above: Suppose μ_{X_j} is the probability distribution of X_j, then the probability distribution of X_{j+1}, μ_j, is:

$$\mu_{X_{j+1}}(B_{j+1}) = \mu_{X_j} P(B_{j+1}) = \int_{\mathbb{R}^n} P(x_j, B_{j+1}) \mu_{X_j}(dx_j)$$

Definition 26 (Irreducibility)**.** A kernel is called irreducible (with respect to Lebesgue measure) if for any pair of $x \in \mathbb{R}^n$ and $B \in \mathcal{B}$ with positive Lebesgue measure, there exists $k > 0$, so that $P^{(k)}(x, B) > 0$.

Definition 27 (Periodicity)**.** Let P be an irreducible kernel. Then P is called peridoic if, for some integer $m \geq 2$, there is a collection of disjoint nonempty sets $\{E_1, \ldots, E_m\} \subset \mathbb{R}^n$ so that for any $j = 1, \ldots, m$ and $x \in E_j$:

$$P(x, E_{j+1}) = 1$$

Otherwise the kernel is called aperiodic.

Definition 28 (Invariant measure)**.** A probability measure is an invariant measure of a probability transition kernel $P(x_j, B_{j+1})$ if:

$$\mu P = P$$

that is:

$$\mu(B) = \int_{\mathbb{R}^n} P(x, B) \mu(dx)$$

Theorem 28. *Let μ be a probability measure over $\mathbb{R}^n$ and $\{X_j\}$ a time homogenous Markov chain with a transition kernel P. Assume further that μ is an invariant measure with the transition kernel P, and that P is irreducible and aperiodic. Then for all $x \in \mathbb{R}^n$,*

$$\lim_{N \to \infty} P^{(N)}(x, B) = \mu(B), \forall B \in \mathcal{B}$$

and for $f \in L^1_\mu(\mathbb{R}^n)$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} f(X_j) = \int_{\mathbb{R}^n} f(x) \mu(dx)$$

Proof. The result above is sometimes referred to as *Simple Ergodic Theorem*. A complete proof of the result above can be found in [8]. The original assumption is that the state space is equipped with countably generated σ-algebra, which apply to Borel sets $\mathcal{B}$ as $\mathcal{B}$ is generated by $\{(q, \infty) | q \in \mathbb{Q}\}$. □

References

[1] Haim Brezis and Haim Brézis. *Functional analysis, Sobolev spaces and partial differential equations*. Vol. 2. 3. Springer, 2011.

[2] Tan Bui-Thanh. "A gentle tutorial on statistical inversion using the Bayesian paradigm". In: *ICES Report* (2012), pp. 12–18.

[3] Richard Courant and David Hilbert. *Methods of mathematical physics: partial differential equations*. John Wiley & Sons, 2008.

[4] Masoumeh Dashti and Andrew M Stuart. "The Bayesian approach to inverse problems". In: *arXiv preprint arXiv:1302.6989* (2013).

[5] Gerald B Folland. *Real analysis: modern techniques and their applications*. Vol. 40. John Wiley & Sons, 1999.

[6] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*. Vol. 160. Springer Science & Business Media, 2006.

[7] Carl Edward Rasmussen and Christopher KI Williams. "Regression". In: (2005).

[8] Gareth O Roberts and Jeffrey S Rosenthal. "General state space Markov chains and MCMC algorithms". In: (2004).