

The paper is organized as follows: In section 2, we introduce Bayesian inversion framework and the Fisher information adaptive MALA. We outline the finite-dimensional Bayesian approach, describe the Fisher adaptive MALA process, and develop the convergence analysis. Section 3 presents numerical experiments for Bayesian inverse problems that demonstrate the effectiveness of this algorithm and compare its performance with other methods. Section 4 concludes the paper with a summary of the main findings and a discussion of potential directions for future research.

2 Fisher Adaptive MALA in Bayesian inversion

In this section, we review the Bayesian approach to inverse problems. Then, we apply the Fisher adaptive MALA to sample from the Bayesian posterior distributions arising in inverse problems.

2.1 Bayesian inversion framework

Throughout this work, we consider the inverse problem of finding an unknown parameter $x \in \mathbb{R}^d$ from data $y \in \mathbb{R}^n$, where the relationship between x and y is described by the following model

$$y = \mathcal{F}(x) + \eta, \quad (2.1)$$

where $\mathcal{F} : \mathbb{R}^d \rightarrow \mathbb{R}^n$ is referred as forward operator, and η represents the measurement noise, which is assumed to follow an n -dimensional Gaussian distribution. In the Bayesian framework, the vectors x , η and y are treated as random variables. Assuming that x and η are independent with π_0 as the prior distribution and $\eta \sim N(0, \Sigma)$ as the Gaussian noise. The joint distribution of (x, y) is expressed as

$$\pi(x, y) = \pi(y | x) \pi_0(x).$$

Furthermore, applying Bayes' rule, the posterior distribution of the unknown parameter x from the observed data y is given by

$$\pi(x | y) = \frac{\pi(y|x)\pi_0(x)}{Z(y)} \propto \exp(-\Phi(x))\pi_0(x),$$

where, $Z(y)$ is the normalization constant and

$$\Phi(x) := \frac{1}{2} \|\mathcal{F}(x) - y\|_{\Sigma}^2 = \frac{1}{2} (\mathcal{F}(x) - y)^{\top} \Sigma^{-1} (\mathcal{F}(x) - y), \quad (2.2)$$

is the data fidelity term. For the prior distribution $\pi_0(x)$, we primarily assume a Gaussian prior with zero mean and covariance C , i.e. $\pi_0(x) = N(0, C)$ and

$$\pi_0(x) \propto \exp\{-\frac{1}{2}\|x\|_C^2\} = \exp\{-\frac{1}{2}x^{\top}C^{-1}x\}.$$

Thus the posterior distribution can be expressed as:

$$\pi(x | y) \propto \exp\{-\Phi(x) - \frac{1}{2}\|x\|_C^2\}. \quad (2.3)$$

Let $\pi(x) = \pi(x | y)$ (here omit the conditioning of data y), then the gradient of the log posterior is given by

$$\nabla \log \pi(x) = -C^{-1}x - \nabla \Phi(x) \text{ and } \nabla \Phi(x) = (\nabla \mathcal{F}(x))^{\top} \Sigma^{-1} (\mathcal{F}(x) - y), \quad (2.4)$$

where $\nabla \mathcal{F}(x)$ is the *Fréchet* derivative of x .

A special case appears if $\mathcal{F}(x)$ is a linear mapping, such as $\mathcal{F}(x) = Fx$. Then it is known from [19] that posterior distribution is Gaussian, i.e. $\pi(x | y) = N(\mu_{post}, C_{post})$ and the mean μ_{post} and covariance matrix C_{post} are given by

$$C_{post} = (C^{-1} + F^{\top} \Sigma^{-1} F)^{-1}, \quad \mu_{post} = C_{post} F^{\top} \Sigma^{-1} y, \quad (2.5)$$

and

$$\nabla \Phi(x) = F^{\top} \Sigma^{-1} (Fx - y). \quad (2.6)$$