Florian Maire,
Département de mathématiques et de statistique,
Université de Montréal,
maire@dms.umontreal.ca

Dereversibilizing Metropolis-Hastings: simple implementation of non-reversible MCMC method

Context: When using Markov chain Monte Carlo in Bayesian inference, the Metropolis-Hastings (MH) algorithm is often the first tool used to learn from the posterior of interest π. Little is known about a simple trick that can improve MH, sometimes dramatically: the Guided Walk (GW) proposed by Gustafson (1998) breaks the reversibility of the MH Markov chain by introducing a momentum variable which imposes a direction of propagation. This idea is neat in unidimensional problems where only two directions are possible. We generalize the Guided Walk algorithm to higher dimensional setups by introducing privileged directions of propagation.

Notations:

- Probability space $(\mathsf{X} \subset \mathbb{R}^d, \mathcal{X}, \mathbb{P})$
- Distribution of interest π on $(\mathsf{X}, \mathcal{X})$
- Markov chain $\{X_t,\, t > 0\}$ on $(\mathsf{X}, \mathcal{X})$
- Directional variable $\theta \in \Theta_d := \{-1, 1\}^d$
- Markov kernel $Q(x, \mathrm{d}x') = q_x(x - x')\mathrm{d}x'$

Algorithm 1 Random Walk (MH)	**Algorithm 2** Guided Walk (GW)
set $X_{t+1} = X_t$	set $X_{t+1} = X_t$ and $\theta_{t+1} = -\theta_t$
draw $\zeta \sim q_{X_t}$, $U \sim \mathrm{unif}(0,1)$	draw $\zeta \sim q_{X_t}$, $U \sim \mathrm{unif}(0,1)$
set $X = X_t + \zeta$	set $X = X_t + \theta_t\lvert\zeta\rvert$
calculate the acceptance probability	calculate the acceptance probability
$\alpha(X_t, X) = 1 \wedge \dfrac{\pi(X)q_X(X - X_t)}{\pi(X_t)q_{X_t}(X_t - X)}$	$\alpha(X_t, X) = 1 \wedge \dfrac{\pi(X)q_X(X - X_t)}{\pi(X_t)q_{X_t}(X_t - X)}$
if $U \le \alpha(X_t, X)$, set $X_{t+1} = X$	if $U \le \alpha(X_t, X)$, set $X_{t+1} = X$ and $\theta_{t+1} = \theta_t$

Generalized Guided Walk (GGW)

Consider r Markov kernels $P_1, P_2, \ldots, P_r$

Each moves according to $(\mathcal{H}_i, \boldsymbol{e}_i)$ with momentum $\theta_{t,i} \in \{-1, 1\}$

If there is $I \subseteq (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r)$ s.t. $\mathsf{X} = \mathrm{span}(I)$

The Markov chain on $\mathsf{X} \times \{-1, 1\}^d$

$$(X_{t+1}, \theta_{t+1}) \sim P_{1+\mathrm{mod}(t,r)}((X_t, \theta_t), \cdot\,),$$

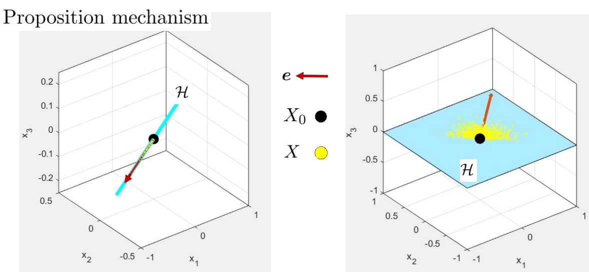
converges marginally to π.

Proposition 1 *Consider $\mathsf{X} \subseteq \mathbb{R}^d$ with $d > 1$, a subspace $\mathcal{H} \subseteq \mathsf{X}$ and a vector $\boldsymbol{e} \in \mathcal{H}$. Let $\{(X_t, \theta_t),\, t \in \mathbb{N}\}$ be the Markov chain evolving as in Alg. 2 with proposal mechanism*

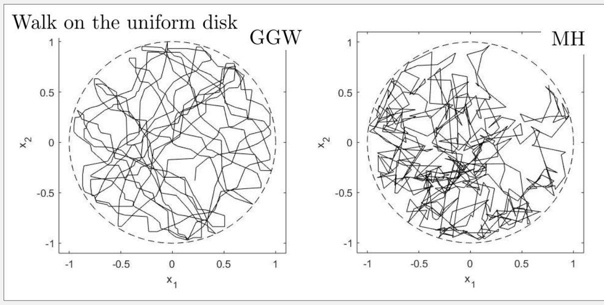
$$X = X_t + \theta_t \rho_{\mathcal{H}}(\zeta)\,, \qquad \zeta \sim q_{X_t}(\cdot)\mathbb{1}_{\rho_{\mathcal{H}}(\zeta)^T \boldsymbol{e} > 0}(\zeta)\,,$$

and accept/reject X with probability $\alpha(X_t, X)$ switching the momentum θ_t in case of rejection. It holds that $X_t \sim \pi \Rightarrow X_{t+1} \sim \pi$.

Proposition mechanism



Walk on the uniform disk



$$\pi = (1/4)\sum_{k=1}^{4} \mathcal{N}([0\ 0]; \Gamma_k) \qquad \Gamma_k = [\boldsymbol{e}_k\ \boldsymbol{e}_k^{\perp}]\mathrm{diag}(1, 1/1000)[\boldsymbol{e}_k\ \boldsymbol{e}_k^{\perp}]^{\mathrm{t}}$$