

Hamiltonian Monte Carlo

Course : Stochastic Simulation

Students : Aude Maier & Tara Fjellman

Fall 2024

1 Introduction

2 Section ...

3 (c)

3.1 Gibbs distribution invariance

Under the assumption that there is no numerical error, we want to prove that the Gibbs measure is invariant for the chain generated by the hamiltonian dynamics.

This is equivalent to saying that the Gibbs measure π is the same before and after an evolution of t seconds from the hamiltonian dynamics. To prove this we first introduce hamiltonian dynamics operators φ, Φ acting respectively on the phase space and the Gibbs measure : $\varphi_t(q_s, p_s) = (q_{s+t}, p_{s+t}); \Phi_t[\pi_s] = \pi_{t+s} \quad \forall t \in \mathbb{R}$. The statement we want to prove can then be expressed as

$$\Phi_t[\pi_s](D) = \pi_{s+t}(D) \quad \forall D \in \mathcal{B}(\Omega), \forall s, t \in \mathbb{R}, \quad (1)$$

with Ω the phase space.

We can now write the left hand side of the equation as

$$\Phi_t[\pi_s](D) = \int_D \pi_{s+t}(q, p) \, dq dp \quad (2)$$

$$= \int_{\varphi_{-t}(D)} \pi_s(q, p) \, dq dp \quad (3)$$

$$= \pi_s(\varphi_{-t}(D)). \quad (4)$$

The final result is obtained using the fact that volumes in phase space are preserved by the hamiltonian dynamics (in conservative systems). This result is known as Liouville's theorem, but is mentioned as theorems 2.3 in [cite].

This implies specifically that $q_k \sim \pi$ for all $k \in \mathbb{N}$ if $q_0 \sim \pi$.

If the dynamics is discretised with the Velocity Verlet algorithm, the volume in phase space is preserved up to a small error, which is why the algorithm is used in practice [cite wikipedia].

3.2 What happens for discrete evolution ?

4 (f)

As an alternative to HMC we consider rejection sampling. We therefore want to find a function $g(q)$ and a constant C such that the following inequality holds for all q :

$$\tilde{f}(q) = e^{q^T X^T (y - 1_n)} e^{-1_n^T \log[1 + \exp(-x_i^T q)]_{n \times 1}} e^{-\frac{1}{2} q^T \Sigma^{-1} q} \leq C g(q), \quad (5)$$

where we have denoted $\Sigma = \text{Diag}(\sigma_1^2, \dots, \sigma_p^2)$. Given that

$$e^{-\sum_i \log[1 + \exp(-x_i^T q)]} = \prod_i \frac{1}{1 + \exp(-x_i^T q)} < 1, \quad (6)$$

we can simplify the problem to finding a function $g(q)$ such that

$$\tilde{f}(q) \leq e^{q^T b} e^{-\frac{1}{2} q^T \Sigma^{-1} q} =: Cg(q), \quad (7)$$

with $b = X^T(y - 1_n)$.

By completing the square in the exponent of $Cg(q)$, we can write it in terms of a Multivariate Gaussian distribution with mean $\mu = \Sigma b$ and covariance Σ . Indeed :

$$e^{-\frac{1}{2}(q-\mu)^T \Sigma^{-1}(q-\mu)} = e^{-\frac{1}{2}\mu^T \Sigma^{-1}\mu} e^{q^T \Sigma^{-1}\mu} e^{-\frac{1}{2}q^T \Sigma^{-1}q} \quad (8)$$

$$\implies \tilde{f}(q) \leq e^{\frac{1}{2}\mu^T \Sigma^{-1}\mu} e^{-\frac{1}{2}(q-\mu)^T \Sigma^{-1}(q-\mu)}. \quad (9)$$

Using now the normalisation constant of the Multivariate Gaussian distribution

$$\sqrt{(2\pi)^p |\Sigma|} = \int_{\mathbb{R}^p} e^{-\frac{1}{2}(q-\mu)^T \Sigma^{-1}(q-\mu)} dq, \quad (10)$$

we can define g and C as

$$g(q) = \frac{1}{\sqrt{(2\pi)^p |\Sigma|}} e^{-\frac{1}{2}(q-\mu)^T \Sigma^{-1}(q-\mu)}, \quad (11)$$

$$C = e^{\frac{1}{2}\mu^T \Sigma^{-1}\mu} \sqrt{(2\pi)^p |\Sigma|} = \sqrt{(2\pi)^p |\Sigma|} e^{\mu^T \Sigma^{-1}\mu}. \quad (12)$$

5 Section ...

6 Conclusion

Aknowledgements

References