知 識 天 地

馬氏蒙地卡羅的理論分析

黃啓瑞研究員，許順吉研究員，黃基志博士後（數學研究所）

簡介

在應用領域中常需考慮高維度或大樣本空間的機率分佈（Probability Distribution，以 P 表示）。直接從這類機率分佈抽樣，實際上並不可行，尤其是不知道所對應密度（Density）的歸一常數（Normalizing Constant）時。利用以 P 爲平衡分佈的馬氏過程（Markov Process）來逼近 P，這稱作動態（Dynamic）或馬氏蒙地卡羅（Markov Chain\Process Monte Carlo, MCMC）。本文簡介相關的數學概念，例如，評估逼近的好壞，定義不同的比較準則，找出最好的馬氏過程，如何加速收斂等。

有限的樣本空間

如何比較統計和影像處理上常用的算則 Metropolis Algorithm 和 Gibbs Sampler？先看兩個對立的直觀論點：Metropolis Algorithm 有較大的轉移機率（Transition Probability），所以收斂到 P 的速度較快；然而 Gibbs Sampler 的轉移機率在局部已達到平衡當然收斂得快。以上兩個直觀想法都錯。現定義比較準則爲：轉移矩陣（Transition Matrix）的固有值去掉 1，取絕對值後的最大值。考慮 P 是二維具週期邊界的 Ising model 及一群包括 Metropolis Algorithm 和 Gibbs Sampler 的算則。我們證明 Metropolis Algorithm 在低溫時最好，在高溫時最差；在任何溫度下可在該群中挑到比 Gibbs Sampler 好的算則[5]。

大樣本空間馬氏鏈蒙地卡羅法（轉移矩陣 Q）計算隨機變數 f 對 P 的期望值，其對應的漸近變異數（Asymptotic Variance）以 v（f,Q）表示。以 v（f,Q）的 Worst Case Analysis（w.r.t. f）爲比較準則，考慮所有對 P 可逆的（Reversible or in Detailed Balance w.r.t. P）馬氏鏈，我們找到最好的 Q。雖然實際上無法直接由最好的 Q 造出對應的馬氏鏈，但利用它的數學結構可改進馬氏場（Markov Random Field）的抽樣算則[4]。當然也可以 Average Case Analysis 爲比較準則或考慮所有以 P 爲平衡分佈的馬氏鏈，但目前尚無滿意結果。一般常用的算則都是可逆的馬氏鏈，我們證明加上局部對 P 的反對稱擾動可均勻改進 v（f,Q）[8]。

連續的樣本空間

當 P 有正連續密度時，其對應的擴散過程的隨機微分方程可看成理論上的算則。我們固定擴散（Diffusion）係數爲根號 2，漂移（Drift）係數可以變動，考慮一群馬氏蒙地卡羅。其中漂移係數爲 Gradient（Log(p)）是唯一可逆的也是常用的擴散過程，其他的漂移係數要再加上 P 加權後散度爲零（Weighted Divergence Free）的向量場（對 P 的反對稱擾動）。若以擴散過程的無限小生成算子（Infinitesimal generator）的譜距（Spectral Gap）爲比較準則，可逆的擴散過程的收斂速度最慢[7]。我們亦可用時間 t 的轉移機率和 P 之間的變分模（Variational Norm）的收斂指數準則，一樣可逆的擴散過程最慢[7]。

是否有最好的反對稱擾動？當 P 爲高斯分佈並只考慮線性反對稱擾動，答案是肯定的[6]。一般情形，固定反對稱擾動的大小，尚不知如何選最佳擾動方向。在歐氏空間是否能將譜距推到無窮大仍不清楚，Torus 上就可做到[3,9]。

固定擾動方向加快擾動速度不一定能加速收斂[6]。但是否存在擾動方向使得加快擾動速度後收斂速度趨近無窮大，這和相關的一個一階微分算子是否有平滑的固有函數有關，也和是否存在弱混流（Weakly Mixing Flow）有關[1,2,3]。

結論

藉由理論分析，釐清馬氏蒙地卡羅在應用上的一些不正確的概念，並改進或設計新算則。這研究在數學上也產生有意思的新課題。

參考文獻

1. H. Berestycki, F. Hamel, N. Nadirashvili (2005) Elliptic eigenvalue problems with large drift and applications to nonlinear propagation phenomena, Commun. Math. Phys., 253, 451-480.
2. P. Constantin, A. Kislev, L. Ryzhik, A. Zlatos (2007) Diffusion and mixing in fluid flows, to appear in Ann. Math.
3. B. Franke, C.-R. Hwang, S.-J. Sheu. and H.-M. Pai (2007) The behaviour of the spectral gap under growing drift, manuscript.
4. A. Frigessi, C.-R. Hwang, and L. Younes (1992) Optimal spectral structure of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields, Ann. Applied Probab., 2, 610-628.
5. A. Frigessi, C.-R. Hwang, S.-J. Sheu, and P. di Stefano (1993) Convergence rate of the Gibbs sampler, the Metropolis algorithm, and other single-site updating dynamics, J. R. Statist. Soc. B 55, 205-219.
6. C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu (1993) Accelerating Gaussian diffusions, Ann. Applied Probab., 3, 897-913.
7. C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu (2005) Accelerating diffusions, Annals Appl. Probab., 15, 1433-1444.
8. C.-R. Hwang, S.-Y. Hwang-Ma (2006) Variance reduction of dynamic Monte Carlo method by antisymmetric perturbations, manuscript.
9. C.-R. Hwang, H.-M. Pai (2006) Blowing up spectral gap of Laplacian on torus by antisymmetric perturbations, manuscript.