

「拡張アンサンブル」イントロダクション

阪大院理 菊池誠¹

1 なぜ

”拡張アンサンブル”の考え方もだいぶ普及してきたので²、マルチカノニカルとか交換MC法とかいう名前をみても、びっくりする人はあまりいないことと思う（と期待して）。ここではまず、なぜ拡張アンサンブルなのか、から始めよう。熱力学量を求めたければカノニカルアンサンブル（あるいは状況に応じて、グランドカノニカルでも $T-p$ でも）を使えばいいだけなのに（しかもメトロポリス法はカノニカルアンサンブルを生成するように設計してあるのに）、なぜわざわざ違うアンサンブル、それも場合によっては見たことも聞いたこともないような奇妙なアンサンブルを使うのか。むろん、そこになんらかの益があるからやるはずなのだ³。

例えば、分子動力学で使われる能勢の方法を考えてみよう。場合には、素朴な分子動力学ではミクロカノニカルアンサンブルが生成されてしまうので、運動の本物らしさを犠牲にしてでもなんとかカノニカルを作りたい、という非常にわかりやすい（というか、切迫した）動機があった。結果として得られるアンサンブルは普通のカノニカルなので、”拡張アンサンブル”ではないかもしれないが、非常に巧妙な方法でミクロカノニカルをカノニカルに”拡張”しており、方向性は近い。

では、最近の拡張アンサンブルはなんのために使われるのか。ざっと考えてすぐに思いつくのは

1. 熱力学積分を実行する
2. 自由エネルギー障壁を越える（一次転移。rugged free-energy landscape）
3. 見たい変数を詳細に見る（自由エネルギー景観を描くなど）
4. 非物理的状态を使って緩和を速くする

といった目的だろう⁴。大雑把にわけると、初期の拡張アンサンブルは熱力学積分の遂行（つまり、広いパラメータ領域でのサンプリング）を目的とするものがほとんどだったのに対し、一方、最近の代表的方法である simulated tempering, 交換MC法, multicanonical などではむしろ2番目の目的、つまり自由エネルギー障壁を越えるために提案されたと言っ

¹ E-mail: kikuchi@phys.sci.osaka-u.ac.jp

² この日本語は伊庭氏が最初に使ったらしい。

³ 正直に言えば、単に面白いから、という理由も捨て難いのだが

⁴ もちろん、これらは互いに絡み合っているし、四つに分ける積極的な意味があるわけでもないけれど

ていいだろう。3 番目や 4 番目はより最近の傾向かなあ。

もちろん、結局は目的にあった計算法ならなんでもいいのだ、という当たり前の結論になるわけだが、欲しいものがカノニカルアンサンブルだからといって、メトロポリス流カノニカルモンテカルロに限る必要はないのだ（あるいは、積極的にカノニカルを踏み外したほうがいい場合もあるのだ）という点はあらためて強調しておいていいだろう。個々の事例を参考にいろいろ考えていただければ、と思う。

2 いろいろ

科学の常として、拡張アンサンブル法もいろいろな形で再発見の歴史をたどっている [1]。multicanonical には umbrella sampling [2] という先駆者がいるし、交換 MC 法は何度も再発見されているようだ⁵。simulated tempering にしても Marinari and Parisi とほぼ同時期（おそらく一年ばかり前）に Lyubartsev らによって expanded ensemble と言う名前で提案されている。分野や目的が違えば再発見も充分あり得るので、それ自体どうということはない。むしろ、前節でも書いたように、個々の分野のどういう文脈で発見されたかを見ておくことが重要だろう。たとえば、Marinari and Parisi がランダムスピン系を念頭に置いて最適化法として simulated tempering を提案したのに対し、Lyubartsev らは高分子や粒子系の自由エネルギー計算の手法として expanded ensemble を提案している。⁶

プログラムを書く立場で考えると、圧倒的に簡単なのが交換 MC 法だ。なにしろ、個々には単なるカノニカルアンサンブルで、時々温度を交換してやればいいだけなので、カノニカル用のプログラムをほとんどそのまま使うことができる。また multicanonical では各ステップで全エネルギーの計算を必要とするため、せっかくカノニカル用にベクトル化されたプログラムを書いてもベクトル化されなくなってしまい、悲しい気分させられてしまうが、交換 MC 法ならそんなことはない。並列計算機への移植も簡単にできる（なにしろもともと並列に計算しているのだ）。だからといって、どんな場合にも交換 MC 法が最強かといえ、むしろそういうわけではないので、目的と方法と簡単さの兼ね合いを考えなくてはならない。いろいろな方法の得意不得意については、伊庭氏の”まとめと展望”を参照のこと。

ちなみに、拡張アンサンブルとは全く違うコンセプトで提案されたものとして broad histogram 法 [4] とその改良版である flat histogram 法 [5] がある。flat histogram が作り出

⁵ 交換 MC 法にもいろいろな名前がつけられているが、最近しばしば使われる名前に parallel tempering というものがある。たしかに、式の上では simulated tempering の並列版と考えることができるので、そう呼びたい気持ちはわからないでもないのだけど、勝手に名付けるのはどうかと思う。交換 MC 法を発明した人達にもうちょっと敬意を払ってもいいんじゃないかという気もする。

⁶ Swendsen が、交換 MC 法は自分たちの replica Monte Carlo 法と同じだ、とか言ってるらしい。たしかによく考えると深い関係があり、“同じ”と言ってもいいのだが、そもそも Swendsen 自身も最近まで気づかなかったのだから、それで priority を主張するのはどうかと思う。一方、Ferrenberg and Swendsen の Histogram reweighting もまたそれ以前に多くの前例を持ち、古くは五〇年代にまで遡るのだが、後々に与えた影響の大きさでは F&S が圧倒的であり、その意味から高く評価されるべきだろう。[3]

す”アンサンブル”は長時間極限で multicanonical と完全に一致するので、これもまた、文脈を無視すれば単にパラメータ設定法だけの違いと考えることもできる⁷。

3 へん

非物理的状态を使う、というとは、なんだか変な気がするかもしれないが、実はそれほど変なアイデアではない。たとえば、格子気体（保存系）のシミュレーションを川崎ダイナミクスで行なう代わりに、普通のイジングモデル（非保存系）のシミュレーションを行なって、そこから磁化一定のアンサンブルを切り出してくる、というやりかただって、本来やりたいモデルから見れば、”非物理的状态”を使っていると言え言えるだろう。（いや、もちろん単なるグランドカノニカルですけど）。わざわざ状態空間をひろげていることになるわけだが、粒子数保存の条件は緩和を極端に遅くしてしまうので、非保存系にして状態空間をひろげてもその手間を越えるほど速い緩和が期待できるのなら、やる価値はある。

物理的アンサンブルを作るには、できあがったアンサンブルの中から非物理的状态を捨ててしまえばいいだけなので、手続きは簡単。非物理的状态を経由させることで緩和を速くするという目的を前面に押し出して提案された方法はまだ少ないかもしれないが（千見寺氏の報告を参照）、この方向にはまだまだ考える余地がありそうだ。

4 ほかに

たとえば、Lee and Choi は entropic sampling (multicanonical の変種) を巡回セールスマン問題に使い、multicanonical annealing と呼んだ⁸ [6]。simulated annealing で扱える問題なら、とりあえずは simulated tempering なり 交換 MC 法 なり multicanonical なりに焼き直すことができる。むろん、それでうまくいくかどうかはケースバイケースだろうし、最適化したいだけなのに全温度のデータを出すのはいかなものか、という気もしないでもないけれど⁹。温度を拡張するほかに適当に非物理的状态をとりまぜて緩和を速くするという戦略は最適化が目的の場合にも（というより、欲しいものが基底状態だけだとすると、途中はなんであっててもかまわないのだから、最適化の場合は特に）有効かもしれない。

数（エントロピー）を計算する手法としても工夫次第で役にたつ。たとえば、random walk から self-avoiding walk までを含む拡張アンサンブルを作れば self-avoiding walk の数が数えられる [7]。基底状態の縮退度がわかっているような格子ポリマー¹⁰ を拡張アンサン

⁷ Oliveira はもっと深いことを考えている節があるのだが、論文を読んでもそこらへんの哲学はあまりよくわからなかった。

⁸ 許される最長経路長をステップ数とともに下げるといことをやっているの、これは実際 annealing になっている

⁹ Lee の方法は、そういうわけで annealing である。しかし、逆にせつかく全温度（全エネルギー）を動いてくれるのだから、最適化目的であっても anneal しないほうが global minimum を見つけやすいのではないか、という気がする

¹⁰ そんなものが自由に作れるのかと思うかもしれないが、郷モデルという強い味方がいるので大丈夫なの

ブルによって全温度で計算すれば、やはり SAW の数が数えられる。前者はともかく、後者は難しい問題をもっと難しい問題に焼き直したことになるのかもしれないが、multi-self-overlap ensemble なら長さ 100 程度の郷モデルは計算できてしまうので実用にならないわけでもないだろう¹¹。

5 そして

というわけで、この研究会では拡張アンサンブルを使ったいろいろな研究について報告があった。いろいろといってもタンパクとランダム系じゃないか、という声もあるかもしれないが、それは今一番ホットなのがそのあたりの分野だから。拡張アンサンブルの考え方自体は一般的なものなので、今後広範な応用があるはずだ。また、基本的な問題として、拡張アンサンブルでの緩和はどうなっているのか、とかパラメータ学習の効率など、考えておきたい問題は多く残されている。

参考文献

- [1] 基本的な文献については伊庭氏の報告を参照。ここでは、それ以外の文献のみあげておくことにする。なお、拡張アンサンブルの歴史については、伊庭氏の博士論文がおそらく世界一詳しい¹²。
- [2] J.P. Valleau, J. Comput. Phys. 23 (1977) 187, umbrella sampling はこんなに昔に考えられていたのだ。
- [3] R.H Swendsen, J.-S. Wang, and A.M. Ferrenberg: in *The Monte Carlo Method in Condensed Matter Physics* ed. K. Binder (Springer), p.75, にヒストグラム法の歴史が書かれている。
- [4] P.M.C. de Oliveira, T.J.P. Penna, and H.J. Herrmann: Eur. Phys. J. B1 (1998) 205.
- [5] cond-mat/9909177
- [6] J. Lee and M. Y. Choi: Phys. Rev. E 50 (1994) R651.
- [7] P.N. Vorontsov-Velyaminov, A.V. Broukhno, T.V. Kuznetsova, and A.P. Lyubartsev: J. Chem. Phys. 100 (1996) 1153, 我々が multi-self-overlap ensemble で提案した self-avoiding 条件を緩めるというアイデアが、実はこの論文で既に議論されていることを最近知った（知らされた）。ただし、温度も含めた二方向にアンサンブルを拡張するところまでは議論されていないので、格子タンパクのようにエネルギーと self-avoidingness の両方が絡み合っている問題については multi-self-overlap ensemble でなければならないと思う。
- [8] K. Pinn and C. Wierczkowski: Int. J. Mod. Phys. C9 (1998) 541.

だ。このモデルの基底状態は定義からして縮退していない。

¹¹ 拡張アンサンブルで魔方陣の数を計算するという論文があり、画期的といえば画期的（一発屋の趣もある）[8]。ちなみに、これは普通の交換 MC 法を使っているが、数を数えるのが目的なら multicanonical のほうがいいだろう（multicanonical なら状態密度そのものが求められる）。また、全ての数を一度だけ使うという条件を緩めてしまえば、もっと速く計算できるかもしれない

¹² と書いたら、伊庭氏に、そんなことはない、と言われたので、そんなことはないのかもしれない