計算天文学 II 第 9 回 最適化 (2)

牧野淳一郎

2001 年 1 月 15 日

1 最適化 (2)

今日は、確率的最適化手法を扱う。

確率的な方法といってもいろいろあるが、現在応用や研究がさかんなのはシミュレーテッドアニーリング (SA) と遺伝的アルゴリズム (GA) である。どちらも、数学的な最適化手法というよりは、物理現象 (SA) や生物の進化 (GA) を真似することで、まあまあの解が得られたらうれしいなという方法である。で、GA はこの 2 つのなかでも新しい方法であり、理論的裏づけもいろいろはっきりしないところがあるので、今日は主に SA の話をする。

2 シミュレーテッドアニーリングの考え

アニーリングとは「焼きなまし」のことである。こういう言葉を知ってるかどうか知らないけど、理論的にはなにかを加熱したあと、ゆっくり冷やすことで熱平衡状態を実現するということである。急に冷やすと熱平衡ではないところで固まってしまう。例えば、安定状態が結晶であるものでも、急速に冷やすと欠陥が多い結晶になったり、あるいはアモルファスで固まったりするわけである。

ここで、熱平衡状態というのは、熱力学的にはエネルギー最低の状態 (何が最低になっているかはもちろん境界条件、例えば圧力一定か体積一定かによる) になっているわけで、つまり、世の中のあらゆる物質というのは、単に「ゆっくり温度を変える」というだけで、「エネルギー最低の状態を実現する」という最適化問題を解いていることになる。特に、温度 0 での熱平衡状態は、系のポテンシャルエネルギーを最小化したものになっている。これを実現するためには、系を熱平衡状態に保ちながらゆっくり冷やしていけばよい。

というわけで、そんな風にやろうというのが SA の基本的な考え方ということになる。

3 熱平衡状態の実現 – メトロポリス・モンテカルロ

さて、熱平衡状態を実現しないといけないわけであるが、それにはどうすればいいだろうか ?とりあえず、古典多粒子系の場合を例に考えてみる。

統計力学的には、温度・体積一定の系でのある物理量 x の値は、アンサンブル平均によって与えられると考えられる。アンサンブル平均は、系のとりうるすべての状態に対して、その出現確率で重みつけして平均をとる。

ここで、ハミルトニアンが $H = T(\boldsymbol{v}) + \Psi(\boldsymbol{r})$ という風に位置の関数と速度（運動量）の関数にわけられる（大抵そうである）場合を考えて、さらに求めたい量が位置だけの関数である場合を考えると、速度については積分を落せるので以下の式がなりたつ。

$$< x > = \frac{\int \exp[-\beta\Psi(\boldsymbol{r})] x dr}{\int \exp[-\beta\Psi(\boldsymbol{r})] dr} \tag{1}$$

これが実際に求められればある量 x が求まるということになる。ここで$\beta = 1/kT$ である。

求めたいのは最適解とか最小値で平均でもなんでもないので、こんなのが求まってもしょうがないのではと思うかもしれないけど、もうちょっと我慢して欲しい。

ここで問題なのは、実際には上の積分は有限の計算量では評価できないということである。実際の物理系では自由度がアボガドロ数くらいあるので全く論外だが、例えば原子の数が 100 とか 1000 くらいにしたところでこれは 300 とか 3000 次元での数値積分になる。しかも、ほとんどのところでは出現確率 $\exp[-\beta\Psi(\boldsymbol{r})]$ が非常に小さく、まじめに計算してもしょうがない。

ここで、確率的積分というのを考えるわけである。確率的積分といっても、例えばランダムに粒子をばらまいて、出現確率にしたがって重みをつけて積分していっても、やはりほとんどのところで確率が非常に小さいのであまりうまくいかない。確率が低くないところだけを重点的に拾うような方法が必要である。

これを実現するのがメトロポリス・モンテカルロ法というものである。ちなみにメトロポリスは人の名前で、ロスアラモス研究所の研究スタッフであった。モンテカルロ法を提案した論文は 1953 年のもので、エドワード・テラー他 4 人との共著論文である。

と、そんなことはさておき、メトロポリス法の大事なところは、全くランダムに粒子をばらまくのではなく、今ある分布からちょっと変えてみてエネルギーの変化をみて、それからどうするか決めるということである。

つまり、具体的には、以下のような手順で計算を進める

1. ランダムに粒子を一つ選ぶ
2. 粒子を動かす向き、大きさをランダムに決める
3. 動かした前と後のエネルギー差 ΔE を計算する。
4. $\Delta E < 0$ なら新しい配置を採用する。そうでなければ、確率 $\exp(-\beta\Delta E)$ で新しい配置を採用する。
5. 配置が新しくなったら、それを記録するなり、それを使って求めたい物理量を計算するなりしておく
6. step 1 に戻る。

ここで、粒子を動かす向き、大きさが「適当」でもいいというのが大事なところで、一般には、詳細釣り合いがなりたっていれば、つまり、相互に移動可能な 2 つの状態 1 と 2 について、相互の遷移確率 P_{12} と P_{21}に以下の関係があれば

$$\exp(-\beta\Phi_1)P_{12} = \exp(-\beta\Phi_2)P_{21} \tag{2}$$

上のやりかたで平均をとったものは極限で統計力学的なアンサンブル平均に収束するということが証明されている。あ、もう一つ、自明な必要条件として、任意の状態から任意の状態に上の手続きを有限回くりかえすことでいける必要があるというのがある。

したがって、どのように動かすかというのは、上の対称性が成り立っているかぎりなんでもいいことになる。例えば、ある半径の球内の一様乱数でもいいし、ガウシアンとか、もっと変なものでも構わない。計算の手間と収束性を考えると、動かすのをあまり大きくするとほとんど採用されなくなって無駄であるし、逆にあまり小さいと今度は時間をかけてもあまり配置が変わらないことになってやはり無駄である。というわけで、採用される確率が $1/2$ くらいになるようにうまくとるのがいいということになっている。

4　SA の実現

さて、メトロポリス法はいいとして、それが最適化とどう関係するかという話に戻る。メトロポリス法では、とにかくある温度でのアンサンブル平均を実行できた。しかし、最適化によって我々がやりたいのは、ある関数の最小化であって別に平均でもなんでもない。

ここで、温度 0 の極限を考えると、熱力学的にはもちろん温度 0 の極限はポテンシャルエネルギーの極小値に対応する。が、温度 0 で機械的にメトロポリス法をやってもうまくいくとは限らない。というのは、温度 0 は$\beta = \infty$ に対応し、少しでもエネルギーが増えるような移動はしないということに対応する。したがって、計算を始めたところの近くに局所的な極小値があれば、そこに落ち込んで計算が止まってしまうからである。

局所的な極小値に落ち込まないようにするには、最初は高い温度にしておいて、ある程度位相空間全体を動けるようにしてメトロポリス法を適用し、それから温度を下げてはまたメトロポリス法で動かすというのを繰り返すという方法が考えられる。これが SA 法である。

実際上は、

- 最初の温度をどれくらいにするか
- どれくらいずつ温度を下げるか
- 一つの温度でどれくらいモンテカルロ計算を続けるか

といったことを考えないといけない。この辺は理論がないわけではないが、それは例えば収束性を保証するためには温度は $T_k = T_1/\log(k)$ というふうに繰り返し数 k の対数でしか温度を下げられないというようなものなのであまり役に立たない。実際には、 $T_k = T_0 r^k$ という風に指数関数的に温度を下げても意外にうまくいってしまう。そういうわけで、実際に問題を持ってきた時に上のようなパラメータ (総称してアニーリング・スケジュールということが多い) をどう決めるかは、いろいろ実験してみる必要がある。

5　組合わせ論的最適化への SA の応用

さて、もともとのメトロポリス法では、原理として考えていたのは多粒子系の熱平衡とかそういうもので、あんまり組合わせ論的最適化という感じはしない。

組合わせ論的最適化というと、代表的なのは前回にあげた巡回セールスマン問題 (TSP) のような、$n!$ 個の組合せを調べあげないと答がわからないようなものである。この他にも、ナップザック問題とか最大充足問題とかいろんなものがあるが、 SA を使うという観点からすればどれも同じようなものである。

つまり、SA という観点からすれば、

- ある近似解について、目的関数を計算する方法
- ある近似解の「近く」の解を「ランダム」に発生する方法

の 2 つを与えることができれば、あとはアニーリング・スケジュールを決めれば答がでる。さらに、前に述べた議論から、近くの解を発生する方法は、対称性さえ満たしていれば適当でいい。収束の速さとかを考えると適当にするよりいろいろ考えた方がいいのは当然だが、それは収束の速さに影響するだけだしその影響もあまり大きくないのが SA のいいところである。これは、条件が悪いと実際上収束しなくなる最急降下法なんかとはだいぶ違う。

5.1 SA で TSP の近似解を求める

TSP は、以下のようにかける。

n 個の点 $p_1, ...p_n$ に対して、点間の移動コスト c_{ij} が与えられているとする。n 個の点をすべて通ってもとに戻るような巡回路でコストの合計が最小になるようなものを見い出せ。

巡回路自体は、n 個の点（の番号）の並び $q_1, ...q_n$ で与えられる。この並びは 1 から n までを並べ変えたもの、つまり、ここで $1 \geq q_i \geq n, \quad q_i \neq q_j (i \neq j)$ ということになる。

目的関数は、

$$C = \sum_{i=1}^{n} c_{q_i, q_{i+1}} \tag{3}$$

（但し、$q_{n+1} = q_1$ とする）であるので簡単に計算できる。近くの解の発生のさせ方だが、もっとも簡単なのは q_i と q_{i+1} を入れ換えてみるというものである。もっと大きくつなぎ変えたければ、適当に 2 箇所を選んでそこで切ってひっくり返してつなぐというのも考えられる。こちらのほうが、局所的な最適解に落ちる可能性が少ないという意味では安全である。

定義域が連続的な関数とは違って、動かす量を調整することで採用確率を調整するのは難しいというか、組合わせ論的最適化の場合にはそういうのを考えてもしょうがない。

実際にプログラムを作る時には、定義式にしたがって目的関数を計算するのではなく、つなぎ変えたことによる変化だけを計算すればいい。これによって一回のモンテカルロ反復に対する計算量を $O(1)$ にすることができる。

例えば、温度の変え方を指数型で $r = 0.9$ とし、一温度での反復は $100n$ とか、採用されたのが $10n$ とか適当に決めて、大抵の場合にちゃんとした答がでるようである。

5.2 SA のプログラムの一般的な方針

原理的には、近似解候補を与えた時に目的関数を計算する関数（手続き）と、今の近似解からランダムに新しい近似解を作る手続きの 2 つがあればプログラムは作れる。が、多くの場合に、「新しい近似解と今の近似解との差」を計算するほうが計算量が減る。TSP の場合であれば、これは $O(n)$ と $O(1)$ で非常に大きい。また、古典粒子系のモンテカルロのように目的関数が全粒子間のポテンシャルエネルギーの和であれば、全部計算すれば $O(n^2)$ だが動かしたことによる変化は $O(n)$ である。このように、変化分を高速に計算する方法を工夫することが極めて重要になる。

5.3 もっと高速化する方法

SA は、まあとにかく答がでるのが利点とはいえ、「ゆっくり冷やさないといけない」という条件がつくので計算量は多い。また、メトロポリス法では配置を作っては乱数と比べるというのの繰り返しなので、普通にやったんでは並列化して沢山計算機を使って速くするのも難しい。

最近、さまざまな並列化手法が提案されている。特に注目されているのは温度並列 SA (TPSA) というもので、計算機毎に違う温度でモンテカルロをやり、時々違う温度のもの同士を交換するという方法である。温度がもっとも低いもののところに最終的な結果が求まる。これは「レプリカ交換法」とかいろんな名前がついているが、基本的な発想はみんな同じである。

次週予告

来週はデータ構造となってますね。一応、簡単なツリー構造のデータを扱ってみることにする。

6 レポート課題

1 は必須。後はオプション。プログラムが必要なものはプログラムを提出すること。

1. TSP について、参考プログラム

   `http://grape.astron.s.u-tokyo.ac.jp/~makino/`
   `kougi/keisan_tenmongakuII/programs/index.html`

   にある `anneal.C` を動かしてみて、単位正方形内にランダムに 100 個程度の点をばらまいた時にどんな答がでるか見てみよ。また、アニーリングスケジュールをいろいろ変えてみて、答がどのように変わるか調べよ。
2. 規則的に点を並べた場合には、TSP の厳密解が全数探索をしなくても求まる。100 点程度の場合にそのような厳密解がある場合を作ってみて、SA で厳密解に到達できるかどうか調べよ。
3. 参考プログラムではコストがユークリッド距離であるとしている。これをマンハッタン距離 $(|x| + |y|)$ にして、答の変化を見てみよ
4. $x = 0.5$ のところに川があって、川を渡るのには λ のコストが余計に掛かるとして答の変化を見てみよ。もっともらしいか ?また、どのようにしてもっともらしいと判断したか ?
5. 「単位正方形の中に n 個の円を重ならないように詰め込めるような円の最大半径はいくつか」という問題を、「円の詰め込み問題」という。いくつかの n については厳密解が知られているが、一般の n について解かれているわけではない。これを SA で解いてみて、どれくらいもっともらしい答がでるか調べよ。目的関数として、n 個の点間の距離と壁への距離の最小値をとればよい。知られている厳密解のいくつかは以下にまとめられているので、比べてみること。

   `http://www.cg.inf.ethz.ch/~peikert/personal/CirclePackings/`