テーマ２：平均時計算量の解析（その２）

本日の講義の目的

- 信念伝播法：最尤解計算問題に対する一つの手法
- 信念伝播法の「計算」からながめてみると…

2.3. 信念伝搬法：Bisection 問題を例に

これまで議論してきた「最尤解計算問題」に対する一つの有力な解法として信念伝搬法が考えられる．今日の講義では Bisection 問題を例に信念伝搬法について説明する．まず，Bisection 問題の復習から．

Bisection 問題
入力： 無向グラフ $G = (V, E)$.
仕事： G の 等数頂点分割 (V_0, V_1) で，カット辺数が最小となるものを求めよ．

我々は，この Bisection 問題に対して次のような平均のシナリオ（確率分布モデル）を考えていた．

Bisection 問題に対する平均のシナリオ（パラメータ $p > r$ に対して）
与えられた頂点数 $2n$ に対し
1. 頂点数が n 個の集合 V_0, V_1 を 1 つ定める．
2. V_i（$i \in \{0, 1\}$）の各頂点間に対し，独立に確率 p で辺をはる．
3. V_0 と V_1 の各頂点間に対し，独立に確率 r で辺をはる．

このシナリオのもとで生成されるランダムグラフを以下では $G_{\mathrm{BIS}(p,r)}(V_0, V_1)$（もしくは，もっと簡単に $G_{\mathrm{BIS}(p,r)}$）と表すことにする．また，$G_{\mathrm{BIS}(p,r)}$ の生成確率を最大にする等分割を，$G_{\mathrm{BIS}(p,r)}$ に対する 最尤分割（すなわち最尤解）と定義する．すると，この最尤解計算問題は，次の意味で Bisection 問題と同値になる．

定理 2.2 パラメータ p, r が $p > 3r$ で $p \geq \Omega(\log n / n)$ のとき以下が成り立つ．

$$\Pr[\ G_{\mathrm{BIS}(p,r)} \text{ に対する最尤分割が Bisection 問題の唯一の解である }] \ = \ 1 - o(1).$$

ちなみに，同じ条件のもとで $G_{\mathrm{BIS}(p,r)}(V_0, V_1)$ 最尤分割は，実は埋め込み解 V_0, V_1 に等しいことも証明できる．

このような最尤解を計算するときに効果を発揮するのが 信念伝播法（もしくは確率伝搬法とも呼ばれる）[1] である．これは，もともとはある確率モデル（ベイジアンネット

ワークと呼ばれるもの）の上での周辺確率分布を求める手法なのだが，それを用いて最尤解を求めるのである．

具体的には，以下に示すアルゴリズムが信念伝播法のレシピ [2] に基づいて導出したものである（ただし，議論を簡単にするために単純化した点が 2 つほどある）．

```
procedure BPforBisection (G, p, r);
begin
  set all b_i to 0;
  b_1 ← +∞;
  repeat MAXROUND times do {
    for each i ∈ V do in parallel {
      b_i ← Σ_{j∈N_i} f_+(b_j) − Σ_{j∉N_i} f_−(b_j);
    }
    if all b_i's get stabilized then break;
  }
  output (+1, sg(b_2), ..., sg(b_{2n}));
end-procedure
```

parameters & functions (for $r < p < 0.5$)

$$h_- = \frac{p+r}{2}, \qquad h_+ = \frac{2-(p+r)}{2},$$
$$c_- = \frac{1-p}{1-r}, \qquad c_+ = \frac{p}{r},$$
$$\theta_- = \frac{4h_-h_+^2(-\ln c_-)}{(p-r)^2}, \quad \theta_+ = \frac{4h_+h_-^2 \ln c_+}{(p-r)^2},$$
$$f_\pm(x) \;=\; \begin{cases} h_\pm \cdot \theta_\pm, & \text{if } \theta_\pm < x, \\ h_\pm \cdot x, & \text{if } -\theta_\pm \le x \le \theta_\pm, \\ -h_\pm \cdot \theta_\pm, & \text{if } x < -\theta_\pm, \end{cases}$$

(By, e.g., $f_\pm$, we mean f_+ and f_- resp.)

$\mathrm{sg}(z) =$ the sign of z, and

$N_i =$ the set of v_i's neighbors.

Bisection 問題に対する信念伝播法に基づくアルゴリズム

この信念伝播法にはいろいろな分野の人がいろいろな解析をしている．ここでは，それを「計算のメカニズム」として見るとおもしろいつながりが見えてくることを紹介しよう．(以下，詳細は講義中のスライドの説明にかえる．以下の話の出典は [3] なので，詳しくはそれを参照して欲しい.）

2.5. もっと単純な方法でも！メッセージ伝搬法

信念伝搬法の計算にはそれなりの理由があるのだが，単に最尤解（あるいは，我々のシナリオにおける埋め込み解）を求めるだけであれば，もっと単純な方法でもよい．確率値（信念）の代わりに単に ±1（もしくは 0）のメッセージを送り合うだけでよいのである．その仕組みを直観的に説明するには，Bisection 問題をは以下に説明するような（非充足型の）連立方程式の最適解計算問題と考えるとよい．

まずは，今回考える（二値）線形連立方程式に関する用語の説明から．これは二値（±1 の値を取る）変数に対し，たとえば

$$x_1 \oplus x_3 \oplus x_{11} = +1$$

のような線形方程式を考える．ただし，$\oplus$ は通常の整数の掛け算（排他的論理輪和もしくは mod 2 での計算）とする．このような式を k-変数線形方程式（略して k-変数式）とい

う．ただし，k は式に現われる変数の数である．k-変数線形連立方程式 (kLIN) とは，こうした k-変数式の集合である．(この授業では $k=2$ に固定した場合を中心に考える．)

こうした連立方程式に対する最尤解計算問題の一つのパターンは，最も多くの式を成立させる解を求める問題である．つまり一般的には次のような問題として定式化できる．

k-変数方程式の最大充足解計算問題 (MAX-kLIN)
入力： n 個の変数から成る k-変数方程式 E.
仕事： E 中の最も多くの式を満たす解を求めよ．

ただし，同じ MAX-2LIN 問題でも平均のシナリオ（つまり式を生成する確率モデル）によって最尤解計算問題は異なる．最尤解計算問題を考える場合，元となる解（埋め込み解）がある．つまり $\boldsymbol{x}=x_1,...,x_n$ に対する値の割り当て $\boldsymbol{a}=(a_1,...,a_n)$ である．これに対し，右辺の値が正しい値になっている式を（$\boldsymbol{a}$ に対し）無矛盾な式といい，その逆になっている式を 誤った式 と呼ぶことにする．

今回の Bisection 問題では，$2n$ 個の変数からなる 2-変数式すべてからなる方程式を考える．また，埋め込み解として，半分の変数に $+1$ を，もう半分の変数に -1 を割り当てた解を考える．以下では，簡単のため，$x_1,...,x_n$ に $+1$ を，$x_{n+1},...,x_{2n}$ に -1 を割り当てることにする．また，$V_+=\{1,...,n\}$, $V_-=\{n+1,...,2n\}$ とする．そして，この解に基づいて式の右辺を決めるのだが，式 $x_i \oplus x_j$ の右辺値を，i,j が同じ V_+（あるいは V_-）に入る場合には確率 $1-p$ で，そうでない場合には確率 $1-r$ で逆転させる（誤った値にする）．このように生成された問題に対する最尤解が Bisection 問題の解に対応するのである．

なお，このように式を作ると，元の問題で辺があるときに右辺が $+1$，無いときに右辺が -1 になる．したがって，Bisection 問題のグラフから連立方程式を作る際には，辺の有無で右辺値を決めればよい．

さて，こうした 2LIN に対して次ページ示すのがメッセージ伝搬アルゴリズムである．

参考文献

[1] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc., 1988.

[2] R. McEliece, D. MacKay, and J. Cheng, Turbo decoding as an instance of Pearl's "Belief Propagation" algorithm, in *IEEE J. on Selected Areas in Comm.* 16(2), 1998.

[3] M. Onsjö and O. Watanabe, A simple message passing algorithm for graph partition problem, in *Proc. 17th Int'l Sympos. on Algorithms and Computation* (ISAAC'06), LNCS 4288, 507–516, 2006.

procedure Message Passing Algorithm for the Perturbed 2-LIN Problem
input $L = \{eq_1, ..., eq_m\}$;
  set $\mathtt{x} := +1$, and set the others all 0;
  repeat `MAXSTEP` times **do** {
    for all $\mathtt{x}_i$, $1 \le i \le n$, **in paralell do** {
      $\mathtt{x}_i := \sum_{eq \in E_i} m_{eq \to i}$;   — $(*)$
    }
    if stabilized **then exit**;
  }
  output $(\mathrm{sign}(\mathtt{x}_1), ..., \mathrm{sign}(\mathtt{x}_n))$;
end-procedure

Remarks:

- E_i = the set of equations containing $\mathtt{x}_i$.
- $m_{eq \to i} \stackrel{\mathrm{def}}{=} \mathtt{x}_j$ if $b = +1$, and $m_{eq \to i} \stackrel{\mathrm{def}}{=} -\mathtt{x}_j$ if $b = -1$, where eq = "$x_i \oplus x_j = b$".
- $\mathrm{sign}(z) = +1$ (resp., $= -1$) if $z > 0$ (resp., $z < 0$), and it is 0 if $z = 0$.

A simple message passing algorithm for the Perturbed 2-LIN Problem

テーマ＃２での課題（再掲）（〆切：７月２７日（金）４時）

次のどれかを選択してレポートを作成して欲しい．

1. 定理 2.1 を証明せよ．この定理は正確には次のような定理である．(ヒント（？）：そう簡単ではない．ランダムな H と $\boldsymbol{n}_*$ に対して，$\boldsymbol{c}$ となるような別のノイズベクトルの重みが（確率的に）どうなるかを議論すればよいと思う．論文 [4] がその手助けになるだろう．)

定理 2.1 与えれれた n に対して，一様ランダムに $(3,4)$-疎行列 H を生成することを仮定する．また（あらかじめ決められている）パラメータ $p << 1$ に対して，ノイズベクトル $\boldsymbol{n}_* \in \{0,1\}^n$ を各ビットごと独立に確率 p で 1 に，$1-p$ で 0 となるように生成する．こうしたランダムモデルのもとで

$$\Pr_{H,\boldsymbol{n}_*}[\ \boldsymbol{n}_* \text{ が } \boldsymbol{c} = G(\boldsymbol{n}_*; H) \text{ のただ一つの最尤解である }\ ]$$

という確率が n を大きくしたときに 0 に収束する．

2. 定理 2.2 を証明せよ．(定理自体は前回と同じだが，今回のプリントの方がより数学的に記述されている．グラフの expander 性などの解析手法を使うと証明できる．)

3. 今回は Bisection 問題に対応する特殊な形の連立方程式（とその確率モデル）について説明したが，メッセージ伝播法は，もっと自然な次のような連立方程式の生成モデルでも平均的に有効に働くことが知られている．

MAX-2LIN 問題に対する平均のシナリオ（パラメータ $p > c/n$ と $q < 1/2$ に対して）
与えられた n に対し

1. n 個の変数に対する ± 1 の割り当て $(a_1, ..., a_n)$ を 1 つ定める．
2. 各 $i < j$ に対して，式 $x_i \oplus x_j = b$ を連立方程式に入れるか否かを確率 p で独立に定める．
3. 式として入れる場合には右辺値 b を確率 $1 - q$ で $a_i \oplus a_j$ に，確率 q で逆の値（誤った値）に設定し，それを式に入れる．

どのような確率パラメータ p, q だとうまく動くかを実験し，その結果・解析・考察を述べよ．

参考：データならびにサンプルプログラムは

`http://www.is.titech.ac.jp/~watanabe/class/maxlin/` から．

また理論的な解析は以下を参照．

[4] U. Feige and E. Ofek, Spectral techniques applied to sparse random graphs, *Random Structures and Algorithms* 27, 251–275, 2005.

4. 疎なランダム $0, 1$-行列の最大固有ベクトルの解析．たとえば $n \times n$ のランダム行列で，確率 p で 1，確率 $1 - p$ で 0 となるものの最大固有値，最大固有ベクトルはどうなるか？$p = c/n$ のように小さい場合が難しい．これに関しては以下を参照．

[5] T. Ando, Y. Kabashima, H. Takahashi, .O. Watanabe, M. Yamamoto, Spectral analysis of random sparse matrices, IEICE Transactions on Information and Systems, E94-D(6), 1247–1256, 2011.