統計学と統計物理

統計数理研究所　伊庭幸人

物理学関係で情報処理というと, すぐに非平衡系ということになるが, 確率分布を基礎とした手法にもまだまだ面白い部分がある. 平衡系であっても, 揺らぎやランダムネスの効果が有益であるような情報処理, 最適解でなく経路積分が意味を持つような世界があってもおかしくない. こうした話題は対称結合のニューラルネットの分野にも見出すことができるが, それ以前に統計学という古典的な大分野がある. ここ 15 年の工学的なニューラルネットや確率ベースの人工知能の進展は統計学的概念の再発見やその重要性の再認識を導き, 同時に統計学の分野への大きな刺激を与えた. 特に重要なのは, 統計物理で常用される動的モンテカルロ法 (メトロポリス法や熱浴法) が統計学の分野に導入されたことである. 本講演では, このような背景を踏まえて, 統計学, 特にベイズ統計と統計物理の関係について説明した. 内容全般については [2][4] 及び筆者のホームページにある原稿 (たとえば [7]) を, 動的モンテカルロ法の統計学への応用については, [12], [5] を参照されたい.

概論の部分の要点は以下の通りである.

- ベイズ統計と統計物理の類似性
  - データ解析
    ベイズ統計ではパラメータの確率分布 (事後分布) ($\Leftrightarrow$ 熱平均) が主役である. この分布からのサンプリングや期待値の計算に統計物理の近似アルゴリズムが使える. 例としては, (動的) モンテカルロ法, 平均場近似などがある.
  - 理論
    理論的な解析では, これにデータについての平均 ($\Leftrightarrow$ ランダム平均) が加わる. ここから生じる 2 重平均の計算にランダム系の統計力学の手法, たとえば, レプリカ法が使える.
- 最適化 .vs. 確率を扱う計算法
  - 最適化一般
    一般的な解釈では, 温度の概念は local optima を抜ける手段と

してのみ存在する. (絶対零度, あるいは, $T \to 0$ の世界). アルゴリズムとしては, simulated annealing 法や Hopfield -Tank 法 の世界.

- 統計物理・統計学・情報理論
  温度の概念そのものは物理以外にはないが, 確率分布が問題になる点は共通である (有限温度, あるいは, $T = 1$ の世界). アルゴリズムとしては, (分布からのサンプリング法としての) 動的モンテカルロ法 や (周辺分布の近似法としての) 平均場近似の世界.

講演では, ベイズ統計の枠組みによる情報処理を “有限温度の情報処理” として位置づけ, 事後分布を活用する推論の利点をイジング模型による画像再構成 (あるいは画像の特徴づけ) という例を通じて説明した. 具体的には, (a) 有限温度での推定の最適性 [2], (b) 不確実さの情報の利用, 特に, acive なサンプリング [3], (c) 雑音の含まれるパターンからの結合定数の推定 [1], などについて議論した. また, この分野の今後の展望について, 次のような例をあげて議論した.

- ベイズ統計・統計学一般の範囲内で
  - 変数選択 (pruning) とベイズ的手法 [6] [8]
  - 誤り訂正符号
  - 位相に関する推定問題
- ちょっと変えて...
  - タンパク質 (ヘテロポリマー) の設計 [9]
  - 論理回路の誤り検出
- 方向を逆にして： 統計 $\to$ 物理 **etc.**
  - データからの準安定状態の検出 [10]
  - 準位統計への応用 (ex. 非定常化) [11]
  - ニューロンの同時計測の解析

参考文献

[1] 伊庭幸人 (1991)
メトロポリス的モンテカルロ法の巨視的パラメータ推定への応用, – 2 次元イジング模型の場合 –, 統計数理 39 No.1 1-21.

[2] 伊庭幸人 (1993, 1996)
ベイズ統計と統計物理 -有限温度での情報処理- , 物性研究 60-6 (1993 年 9 月号) 677-699.
ベイズ統計と統計物理 (物性研究 1993 年 9 月号) への訂正と追加 , 「物性研究」65-5(1996 年 2 月号)678-685.

[3] 伊庭幸人 (1994)
統計数理研究所 年度末発表会 (1993 年度) 要旨, マルコフ連鎖モンテカルロ法と適応的デザイン, 統計数理 42 No.2 343-345.

[4] 伊庭幸人 (1996)
学習と階層 — ベイズ統計の立場から — , 「物性研究」65-5(1996 年 2 月号) 657-677.

[5] 伊庭幸人 (1996)
マルコフ連鎖モンテカルロ法とその統計学への応用, 統計数理 44 No.1 49-84.

[6] 伊庭幸人 (1996)
回帰分析の変数選択におけるベイズ的方法, – 平均場近似に基づくアプローチ –, ISM Reseach Memorandum No.624, October, 1996, SITA 96(Dec. 1996, 箱根) 予稿. (http://www.ism.ac.jp/~iba/ で入手可能.)

[7] 伊庭幸人 (1997)
統計科学 文献案内, ISM Reseach Memorandum No.630, January, 1997 (http://www.ism.ac.jp/~iba/で入手可能.)

[8] 統計数理研究所研究会 重回帰分析・ニューラルネットワークにおける変数選択とベイズ的方法, 1998 年 1/30-1/31
プログラム http://www.ism.ac.jp/~iba/prog.htm (将来撤去の可能性あり)

[9] 菊池誠氏 (阪大理), 時田恵一郎氏 (阪大理) と共同研究中 (物理学会 (1997 春, 秋) などで発表). 論文準備中. 本研究会のポスターセッションでも, 黒板にトラペを張ってアドリブで発表した。

[10] 福島孝治氏 (物性所) と共同研究中. 物理学会 (1997 秋) で発表.

[11] 物理学会 (1997 秋) で発表したがタコだった. 再起すべく検討中.

[12] Gilks, W., Spiegelhalter, D. and Richardson, S.(eds)(1996), Markov Chain Monte Carlo in practice, Chapman and Hall, London, New York.