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

統計数理研究所　伊庭幸人

〒 106　港区 南麻布 4-6-7　(iba@ism.ac.jp)

訂正と追加

以下の点について訂正あるいは追加する。

- 図 1a-e, 図 2a-e の結果について

  図 1a-e, 図 2a-e は Caucy 分布による平滑化 ・変化点検出において損失関数がどのように結果に影響するかを調べた数値実験の結果を示したものである。すでに論文の中でも図 2c と図 2d,e の大きな差は数値計算の artifact かもしれないと書いたが、その後より精度の高い計算 ("縦軸" の離散化の際のメッシュを細かくした計算) を試みたところ、図 2c のまっすぐな部分にも僅かな曲がりが認められ、精度をあげるにしたがって増えるようであった。従って、精度無限大の極限 (連続極限) では図 2c と図 2d の違いは少なくなる可能性がある。なお、周辺分布の計算結果 (図 2d,e) は、離散化誤差に対してよりロバストである。

- "損失関数による推定値の違い" に関するそれ以外の論点について

  図 1a-e, 図 2a-e に関連した別の問題は、(14) 式で表される Caucy 分布の範囲だけで問題を考えることが妥当かということである。北川がもとの論文で行なっているように Gauss 分布と Caucy 分布の両方を含む族を考え、そこで最適な hyper parameter を探し、それについて損失関数の違いによる比較を行なったら、結果は違ってくるかもしれない。一般に良いモデル族を考えるほど、損失関数の違いによる差は減るように思われる。

  また、雑音のある伝送路での畳み込み符号の問題のように、あらかじめ最適な hyper parameter(この場合、雑音の大きさ) が知られている場合は良いが、一般の画像などではそれもデータから推定しないといけない。そこで、hyper parameter の推定法と損失関数による違いの問題が絡みあって、大変面倒になる。そのような状況を考えようとした研究の例は、Dinten et al.(1991) にみられる (結論はやはりはっきりしないのだが)。

結局、筆者としては、“損失関数による推定値の違い” は実用上はあまり重要な問題ではないのではないか、という方向に傾いている。“有限温度” の利点が重要になるとしたら、むしろ hyper parameter の推定に関してであろう (もとの論文の 3 章及び本号に掲載の伊庭 (1996) を参照)。また、もうひとつの方向として、次のデータをどのように獲得するかという計画を立てるために、事後分布に含まれる不確定さの情報を使うという考えがある。これについては多くのひとが論じているようである (たとえば MacKay(1992c)、伊庭 (1994))。

- 構成上のミス

  3 章は本来、3.1 節、4 章は 3.2 節になるはずであった。したがって、4 章の “歴史” は “hyper parameter の推定” に関する歴史であって、2.2 節 (“損失関数と optimal estimator” に関する歴史) と対になるのが正しい。次のように訂正されたい。

  1 章と 2 章 → そのまま。

  3. hyper parameter の推定
     3.1 hyper parameter の推定問題
     [もとの 3 章の全体]
     3.2 歴史
     [もとの 4 章の全体]

  5 章、5.1 節、5.2 節 → それぞれ、4 章、4.1 節、4.2 節

- 文献の訂正

  誤
  Kitagawa,G.(1987)
  Journal of the American Statistical Association 79 1032-1063.
  正
  Kitagawa,G.(1987)
  Journal of the American Statistical Association **82** 1032-1063.
  誤
  Geiger,D. and Girosi,F.(1991)
  IEEE Transactions on Pattern Analysis and Machine Intelligence 12 401-412.
  正
  Geiger,D. and Girosi,F.(1991)
  IEEE Transactions on Pattern Analysis and Machine Intelligence **13** 401-412.

- 用語のミス

  誤 マルコフ鎖モンテカルロ法

  正 マルコフ連鎖モンテカルロ法

文献の追加

この機会に、その後、出たり、知ったりした文献を追加しておきたい。簡単な解説をつけるが、詳しい内容については十分理解できていないものも多いことをお断りしておく。

- 2 章 (損失関数と optimal estimator) に関係のある文献として Rue(1995) が出た。画像処理における損失関数の重要性と設計法が論じられている。Besag(1986) の 2.4 節には (Marroquin 以前についての) 歴史的な内容を含む記述がある。

- 3 章 (hyper parameter の推定) の関係では MacKay(1995) がある。parameter の方を先に積分する方法 (ここでいう温度 1 の方法、周辺尤度最大化法) と hyper parameter の方を先に積分する方法 (ここでいう温度零の方法とは少し違うが、たぶん本質的に同じ) を比較して前者が良いと主張している。伊庭 (1996) でいう “場合の数の効果” が強調されているようである。MacKay についてはあとの項目も参照されたい。Dinten et al. (1991) には、(32)(33) 式では hyper parameter の推定がうまくいかない簡単な例が出ている。

- 訂正前の 5 章 (有限温度のアルゴリズム、特に拡張アンサンブル法) の関係では、ベイズ統計における拡張アンサンブル (交換法 ・Simulated Tempering 法) の応用を扱った Geyer and Thompson (1995) が出版された。“温度並列シミュレーテッド ・アニーリング法” の系統では石川 ・金久 (1993)、小西ほか (1995) がある。福島 (1994) の研究も近く出版されると思われる。まだアブストラクトしか見ていないが、Neal (1994) というのも関係ありそうである。大きさの違う系をまとめた拡張アンサンブルの試みについては伊庭 (1995a)、及び、Besag et al.(1995) に対する Wong のコメント (Besag et al. の p.52-53) と返事 (p.64-65) を参照されたい。

- Bruce and Saad (1994) は 3 章でいう温度 1 の方法 (周辺尤度最大化法、ABIC 法、経験ベイズ法) について統計物理の観点から議論している。また、イジング模型による画像処理を統計物理の観点から調べた論文として、Pryce and Bruce (1995) が出た。

- 統計学における (動的) モンテカルロ法の応用については既に多量の文献があるが、ここでは最新の報告として Besag et al.(1995) をあげておく。伊庭 (1995b) とその参考文献も参考にされたい。

- 画像処理におけるマルコフ場とモンテカルロ法の応用についての教科書として Winkler(1995) が出た。論文集として Possolo(1991) がある。また、Journal of Applied Statistics 16 No.2 (1989) は統計的画像処理の特集号で、マルコフ場やベイズ統計、モンテカルロ法の応用についての論文を含んでいる。この続編が出た (同じ雑誌の 21 No.1(1994)) が、まだ入手していない。

- 画像処理におけるクラスター近似、tree 近似の類についても論文が出た (Wu and Doerschuk (1995a,1995b), Tanaka and Morita (1995))。Qian と Titterington もいろいろな形でクラスター的な扱い (1 次元の chain に分解する扱い) を考えているようである (Qian and Titterington(1989,1991))。

- MacKay は学位論文で Neural Network におけるベイズ的な枠組を扱い、一連の論文として発表している (MacKay(1992a,b,c,d))。MacKay は周辺尤度のことを evidence と呼んでいる。画像や時系列の smoothing への応用は (たとえば) 統計数理研究所の研究者が先行しているが、パーセプトロンへの適用は新味がある。MacKay の www は http://131.111.48.24/pub/mackay/homepage.html。

- 比較的新しい動きとして、物体の輪郭線や連続変形に対するベイズ的なモデル化があるが、これに関しては、Phillips and Smith(1994), Stovik(1994), Grenandar and Keenan (1989), Amit et al.(1991), Grenandar and Miller(1995) などがある。画像 ・パーセプトロン以外では、データから遺伝子の伝搬や染色体上の位置関係を推定する問題が面白そうである (Thompson (1991, 1994), Kong(1991))。

- 大規模モデルや imformative prior を意識したものではないが、周辺化によるメリット (有限温度のメリット) に関係した理論の論文がいくつか出ている。Clarke and Barron (1994) は、情報圧縮における周辺化のメリットと事前分布の選択についてミニマックスの立場で論じている。Komaki(1994) は予測分布について情報幾何の見地から分析している。

追加した文献のリスト

Pryce and Bruce (1995) 及び Tanaka and Morita(1995) を教えてくださった田中和之氏、MacKay の論文について教えてくださった松本隆氏、松井淳氏に感謝します。坂野鋭氏にも文献を教えていただきました。

Amit,Y.,Grenander,U. and Piccioni,M. (1991)
Structural image restoration through deformable templates
Journal of the American Statistical Association 86 376-387.

Besag,J. (1986)
On the statistical analysis of dirty pictures
The Journal of the Royal Statistical Society B 48 259-302 (with discussion).

Besag,J., Green,P., Higdon,D. and Mengersen,K. (1995)
Bayesian computation and stochastic systems
Statistical Science 10 3-66 (with discussion).

Bruce,A.D. and Saad,D. (1994)
Statistical mechanics of hypothesis evaluation
Journal of Physics A (Mathematical and General) 27 3355-3363.

Clarke,B. and Barron,A.R. (1994)
Jeffreys' prior is asymptotically least favorable under entropy risk
Journal of Statistical Planning and Inference 41, 37-60.

Dinten,J.M., Guyon,X., and Yao,J.F. (1991)
On the choice of the regularization parameter: The case of binary images in the Bayesian restoration framework
in Spatial Statistics and Imaging Ed. A. Possolo
LECTURE NOTES-MONOGRAPH SERIES Vol.20
Institute of Mathematical Statistics, Hayward, California.

Geyer,C.J. and Thompson,E.A. (1995)
Annealing Markov chain Monte Carlo with applications to ancestral inference
Journal of the American Statistical Association 90 909-920.

Grenandar,U. and Keenan (1989)
Towards automated image understanding
Journal of Applied Statistics 16 207-221.

Grenandar,U. and Miller,I.M. (1994)
Representations of knowledge in complex systems
Journal of Royal Statistical Society B, 56 549-603 (with discussion).

福島孝治 (1994)
物理学会 秋の分科会 (1994)(及び、学位論文 (1996)).

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.

伊庭幸人 (1995a)
統計数理研究所 年度末発表会 (1994 年度) 要旨
マルコフ連鎖モンテカルロ法の改良
統計数理 43 No.1 235-236.

伊庭幸人 (1995b)
マルコフ連鎖モンテカルロ法とその統計学への応用
ISM Research Memorandum No.570
(第 1 版を 「統計数理」に投稿中).

伊庭幸人 (1996)
学習と階層 — ベイズ統計の立場から —
ISM Research Memorandum No.558
(「物性研究」のこの号に掲載予定).

石川幹人、金久實 (1993)
文字列を比較し並べる
in 小特集 「生命の設計図に迫る!」, 日本物理学会誌 Vol.48 No.5.

Kitagawa,G.(1987)
Non-Gaussian state space modeling of nonstationary time series (with discussion)
Journal of the American Statistical Association 82 1032-1063.

Komaki,F.(1994)
On Asymptotic properties of predictive distributions
Technical Report METR 94-21
Dept. of Mathematical Engineering and Information Physics, The University of Tokyo.

Kong,A.(1991)
Analysis of pedigree data using methods combining peeling and Gibbs sampler
in Computing Science and Statistics:
Proc. of the 23rd symposium on the interface(Ed. E.M.Keramides)
pp.379-385 Interface Foundation, Fairfax Station, Va.

小西健三、瀧和男、木村宏一 (1995)
温度並列シミュレーテッド・アニーリング法とその評価
情報処理学会論文誌 Vol.36 No.4 797-807.

MacKay, D.J.C. (1992a)
Bayesian Interpolation
Neural Computation 4 415-447.

MacKay, D.J.C. (1992b)
A Practical Bayesian Framework for Backprop Networks
Neural Computation 4 448-472.

MacKay, D.J.C. (1992c)
Information-based objective functions for active data selection
Neural Computation 4 589-603.

MacKay, D.J.C. (1992d)
The evidence framework applied to classification networks
Neural Computation 4 698-714.

MacKay, D.J.C. (1995)
Hyperparameters: optimize, or integrate out?
submitted to Neural Computation.

Neal, R.M. (1994)
Sampling from multimodal distributions using tempered transitions, Technical Report No. 9421, Dept. of Statistics, University of Toronto,
ftp://ftp.cs.toronto.edu/pub/radford/www/homepage.html.

Phillips,D.B. and Smith, A.F.M.(1994)
Bayesian faces via hierarchical template modeling
Journal of the American Statistical Association 89 1151-1163.

Possolo,A.(ed.)(1991)
Spatial Statistics and Imaging
LECTURE NOTES-MONOGRAPH SERIES Vol.20
Institute of Mathematical Statistics, Hayward, California.

Pryce,J.M. and Bruce,A.D.(1995)
Statistical mechanics of image restoration
Journal of Physics A (Mathematical and General) 28 511-532.

Qian,W. and Titterington,D.M.(1989)
On the use of Gibbs Markov chain models in the analyisis of images based on second-order pairwise interactive distributions
Journal of Applied Statistics 16 No.2 267-282.

Qian,W. and Titterington,D.M.(1991)
Stochastic relaxations and EM algorithms for Markov random fields
J. Statist. Comput. Simul. 40 55-69.

Rue,H. (1995)
New loss functions in Bayesian imaging
Journal of the American Statistical Association 90 900-908.

Storvik,G.(1994)
A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing
IEEE Transactions on Pattern Analysis and Machine Intelligence 16 976-986.

Tanaka,K. and Morita,T. (1995)
Cluster variation method and image restoration problem
Physics Letters A203 122-128.

Thompson,E.A.(1991)
Probabilities on complex pedigrees; the Gibbs sampler approach
in Computing Science and Statistics:
Proc. of the 23rd symposium on the interface(Ed. E.M.Keramides)
pp.371-378 Interface Foundation, Fairfax Station, Va.

Thompson,E.A.(1994)
Monte Carlo likelihood in genetic mapping
Statistical Science 9 355-366.

Winkler,G.(1995)
Image analysis, Random fields and Dynamic Monte Carlo Methods
A Mathematical Introduction
Springer-Verlag, Berlin.

Wu,Chi-hsin and Doerschuk,P.C. (1995a)
Cluster expansions for deterministic computation of Bayesian estimators based on Markov random fields
IEEE Transactions on Pattern Analysis and Machine Intelligence 17 275-293.

Wu,Chi-hsin and Doerschuk,P.C. (1995b)
Tree approximations to Markov random fields
IEEE Transactions on Pattern Analysis and Machine Intelligence 17 391-402.