

マルコフ確率場のハイパーパラメータ分布推定

Distribution estimation of hyperparameters in Markov random field

中西 (大野) 義典 *1*2 永田 賢二 *1 庄野 逸 *3 岡田 真人 *1*4
 Yoshinori Nakanishi-Ohno Kenji Nagata Hayaru Shouno Masato Okada

*1 東京大学 大学院新領域創成科学研究科 Graduate School of Frontier Sciences, The University of Tokyo
 *2 日本学術振興会 特別研究員 Research Fellow of Japan Society for the Promotion of Science

*3 電気通信大学 大学院情報理工学研究科
 Graduate School of Informatics and Engineering, The University of Electro-Communications

*4 理化学研究所 脳科学総合研究センター
 RIKEN Brain Science Institute

We proposed a method of distribution estimation of hyperparameters in Markov random field (MRF). This study was motivated by various kinds of image data in natural sciences thanks to recent advances in measurement techniques. The MRF model is used for image processing, and its hyperparameters must be adjusted to improve the performance. It is important that the hyperparameters appearing in data analysis represent physical quantities such as diffusion coefficients. Indeed, many frameworks of hyperparameter estimation have been proposed, but most are point estimation that is susceptible to stochastic fluctuations. Then, distribution estimation is required because it allows you to evaluate the confidence you have had in point estimates. We used a solvable MRF model to investigate the performance of our distribution estimation method in simulations. By using our method, we discussed the trade-off problem between the number of pixels and the number of observations caused by the limited observation time.

1. 序論

自然科学では計測技術の進歩により様々な画像データが得られている。こうした画像データからの情報抽出は重要な課題である。画像データの解析には、しばしばマルコフ確率場 (MRF) モデルが用いられる [Geman 84, Pryce 95, Tanaka 02]。MRF のハイパーパラメータ調整は、解析結果に大きく影響することが知られている [Demoment 89, Tanaka 07, Ohno 12]。したがって、ハイパーパラメータをデータから推定する様々な手法が議論されてきた。しかし、そのほとんどは最尤推定や最大事後確率 (MAP) 推定のような点推定に留まる。点推定は、依然としてノイズの影響を無視できない最先端計測の場では脆弱である。

本研究では、点推定値の性能を評価できる分布推定を提案し、その性能を評価する。はじめに、画像データやノイズの性質を MRF で記述する。ここで重要なのは、MRF モデルのハイパーパラメータは拡散係数に対応する点である。ハイパーパラメータが、画像処理のための単なる調整パラメータというだけでなく、観測対象の従う物理を記述する重要なパラメータであることから、点推定値を安易に信じることは危険である。次に、ベイズ推定の枠組みを用いることにより事後分布を求める [Bishop 06]。ベイズ推定の枠組みを用いるためにも、確率モデルである MRF を用いることが重要である。こうして求めた事後分布の広がり点が点推定値の信頼度に対応する。とりわけガウス MRF モデルを用いるため、容易に事後分布の解析解を導出できる。さらに、数値実験により、導出した事後分布による分布推定の性能を評価する。最後に、解像度と観測回数とのトレードオフに関して議論する。

本稿の構成は次の通りである。2 節で MRF モデルを用いて

連絡先: 岡田真人, 東京大学大学院新領域創成科学研究科, 277-8561 千葉県柏市柏の葉 5-1-5, okada@k.u-tokyo.ac.jp

画像データの生成過程を記述する。3 節でベイズ推定の枠組みを用いて分布推定を定式化する。4 節で数値実験と、それに対する考察を行う。5 節で結論を述べる。

2. モデル

MRF を用いて、画像データの生成過程を記述する。周期的な境界条件を持つ 2 次元画像を考える。観測対象 $\mathbf{u} = \{u_i\} (u_i \in \mathbb{R}, i = (i_1, i_2))$ の事前分布を次式で与える、

$$p(\mathbf{u}|\mu, \alpha) = \frac{1}{Z_{\text{pri}}} \exp \left[-\frac{\alpha}{2} \sum_{\langle i, j \rangle} (u_i - u_j)^2 \right] \delta \left[\frac{1}{n^2} \sum_i u_i - \mu \right]. \quad (1)$$

ここで、 μ, α がハイパーパラメータであり、 μ が観測対象の平均値、 α は観測対象の連続性を表す。 n^2 は画素数を表す。 Z_{pri} は規格化定数を表す。 $\langle i, j \rangle$ はすべての隣接サイト対に関して和をとることを表す。 δ はディラックのデルタ関数を表す。簡単のため、 α は画像全体で一定の値をとるものとし、不連続点を考慮しないものとする。この事前分布で表される系は、次に示す拡散方程式を離散化したものに対応する、

$$\begin{aligned} \frac{\partial u}{\partial t} &= D \nabla^2 u \\ &= -\frac{\delta D}{\delta u} \frac{D}{2} \int dx \|\nabla u\|^2. \end{aligned} \quad (2)$$

ここで重要なのはハイパーパラメータ α が拡散係数 D と対応することである。画像修復において重要な役割を果たすハイパーパラメータが、実は観測対象を支配する物理においても本質的なパラメータであるということは注目に値する。

対象 \mathbf{u} を観測して画像データ $\mathbf{v} = \{v_i\} (v_i \in \mathbb{R}, i = (i_1, i_2))$ を得る過程を次の確率分布で与える,

$$p(\mathbf{v}|\mathbf{u}, \beta) = \frac{1}{Z_{\text{like}}} \exp \left[-\frac{\beta}{2} \sum_i (v_i - u_i)^2 \right]. \quad (3)$$

ここで, β がハイパーパラメータであり, 観測精度を表す. Z_{like} は規格化定数を表す. 簡単のため, それぞれの画素に独立同分布のガウスノイズが重畳されるとした. 同じ対象を複数回独立に観測することを考慮して, T 枚の画像データ $\{\mathbf{v}^t\} = \{\mathbf{v}^1, \dots, \mathbf{v}^T\}$ が生成される過程は次式で与えられる,

$$p(\{\mathbf{v}^t\}|\mu, \alpha, \beta) = \int d\mathbf{u} \left(\prod_t p(\mathbf{v}^t|\mathbf{u}, \beta) \right) p(\mathbf{u}|\mu, \alpha). \quad (4)$$

3. 定式化

ベイズ推定の枠組みを用いて, 分布推定を行う. ベイズの定理により, ハイパーパラメータの事後分布は次式で与えられる,

$$p(\mu, \alpha, \beta|\{\mathbf{v}^t\}) \propto p(\{\mathbf{v}^t\}|\mu, \alpha, \beta)p(\mu, \alpha, \beta). \quad (5)$$

簡単のため, ハイパーパラメータの事前分布に関しては,

$$p(\mu, \alpha, \beta) = \text{const.} \quad (6)$$

が無情報を表すとし, これを用いる. 一般的にハイパーパラメータ推定は計算困難な多重積分を伴う. 本研究で扱っているモデルも \mathbf{u} に関する n^2 重積分を伴う. しかし, 周期的境界条件およびガウス性を利用することにより, 本モデルでは陽に事後分布を求めることができる. 式 (1), (3), (4) および, 離散フーリエ変換

$$\tilde{v}_k^t = \frac{1}{n} \sum_i v_i^t e^{-2\pi i i \cdot k/n} \quad (7)$$

を用いて,

$$\begin{aligned} & p(\{\mathbf{v}^t\}|\mu, \alpha, \beta) \\ &= \int d\mathbf{u} \left(\prod_t \prod_k \sqrt{\frac{\beta}{2\pi}} \exp \left[-\frac{\beta}{2} |\tilde{v}_k^t - \tilde{u}_k|^2 \right] \right) \\ & \left(\prod_{k \neq (0,0)} \sqrt{\frac{\alpha \lambda_k}{2\pi}} \exp \left[-\frac{\alpha \lambda_k}{2} |\tilde{u}_k|^2 \right] \right) \\ & \delta [\tilde{u}_{(0,0)} - n\mu] \end{aligned} \quad (8)$$

と書ける. ここで,

$$\lambda_k = \sum_{d=1}^2 4 \sin^2(\pi k_d/n) \quad (9)$$

である. 式 (6), (8) を式 (5) に代入し, ガウス積分の公式

$$\int dx \exp \left[-\frac{a}{2} x^2 \right] = \sqrt{\frac{2\pi}{a}} \quad (10)$$

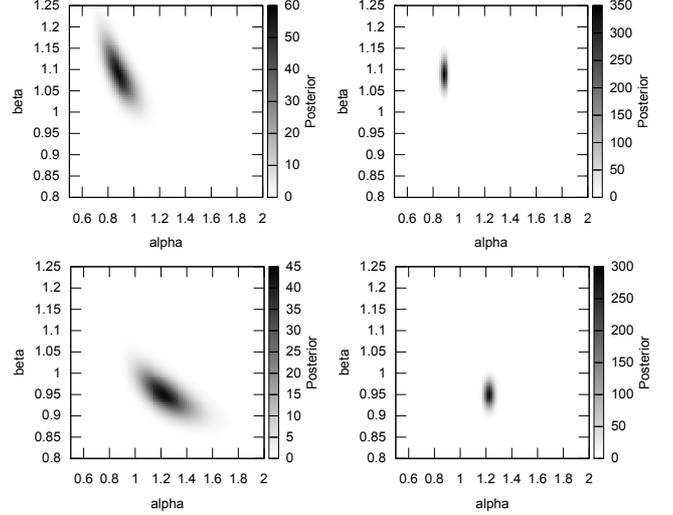


図 1: 異なる 2 組のデータに対する事後分布 $p(\alpha, \beta|\{\mathbf{v}^t\})$. $n^2 = 64^2$, $T = 1$, $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$. 左列が厳密解, 右列が MF 解. 同行の図は同じデータに対する結果.

を用いて, 事後分布

$$\begin{aligned} & p(\mu, \alpha, \beta|\{\mathbf{v}^t\}) \\ & \propto \prod_{k \neq 0} \left(\frac{\beta}{2\pi} \right)^{\frac{T}{2}} \left(\frac{\alpha \lambda_k}{\beta T + \alpha \lambda_k} \right)^{\frac{1}{2}} \\ & \exp \left[-\frac{\beta}{2} \sum_t |\tilde{v}_k^t|^2 + \frac{\beta}{2} \frac{\beta}{\beta T + \alpha \lambda_k} \left| \sum_t \tilde{v}_k^t \right|^2 \right] \\ & \left(\frac{\beta}{2\pi} \right)^{\frac{T}{2}} \exp \left[-\frac{\beta}{2} \sum_t |\tilde{v}_0^t - n\mu|^2 \right]. \end{aligned} \quad (11)$$

を得る.

4. 数値計算

数値計算により分布推定を行い, その性能を評価する. 式 (11) の事後分布は, μ に関して単純な形, 具体的にはガウス分布であり, 以降の議論では μ に関して周辺化した分布を議論する,

$$\begin{aligned} & p(\alpha, \beta|\{\mathbf{v}^t\}) = \int d\mu p(\mu, \alpha, \beta|\{\mathbf{v}^t\}) \\ & \propto \prod_{k \neq 0} \left(\frac{\beta}{2\pi} \right)^{\frac{T}{2}} \left(\frac{\alpha \lambda_k}{\beta T + \alpha \lambda_k} \right)^{\frac{1}{2}} \\ & \exp \left[-\frac{\beta}{2} \sum_t |\tilde{v}_k^t|^2 + \frac{\beta}{2} \frac{\beta}{\beta T + \alpha \lambda_k} \left| \sum_t \tilde{v}_k^t \right|^2 \right] \\ & \left(\frac{\beta}{2\pi} \right)^{\frac{T}{2}} \left(\frac{2\pi}{\beta T n^2} \right)^{\frac{1}{2}} \\ & \exp \left[-\frac{\beta}{2} \sum_t |\tilde{v}_0^t|^2 + \frac{\beta}{2T} \left| \sum_t \tilde{y}_0^t \right|^2 \right]. \end{aligned} \quad (12)$$

はじめに, 式 (12) により得られる事後分布を実際に図 1 に示す. ハイパーパラメータの真値は $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$ で

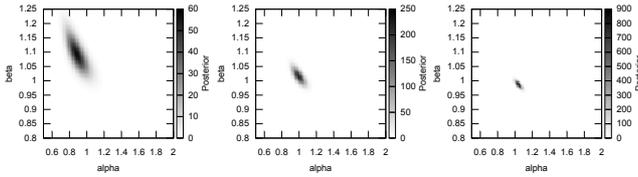


図 2: 画素数 n^2 を変えたときの事後分布 $p(\alpha, \beta | \{v^t\})$. 左から $n^2 = 64^2, 128^2, 256^2$. $T = 1$. $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$.

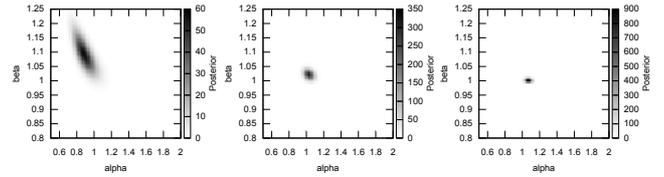


図 4: 観測回数 T を変えたときの事後分布 $p(\alpha, \beta | \{v^t\})$. 左から $T = 1, 4, 16$. $n^2 = 64^2$. $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$.

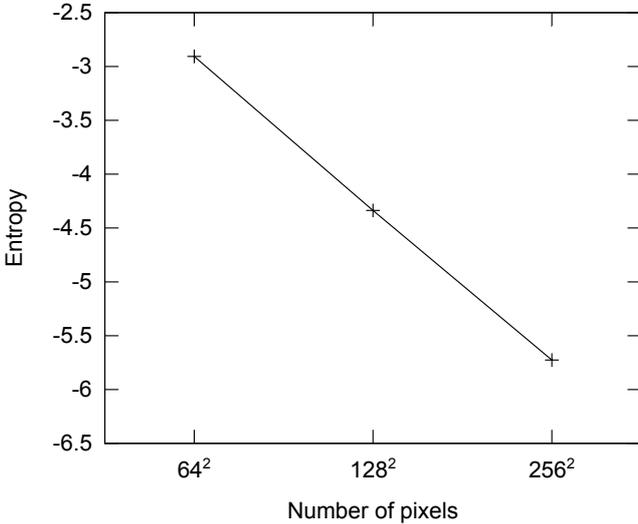


図 3: 画素数を変えたときの事後分布のエントロピー. 100 試行間の中央値を示した. $T = 1$. $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$.

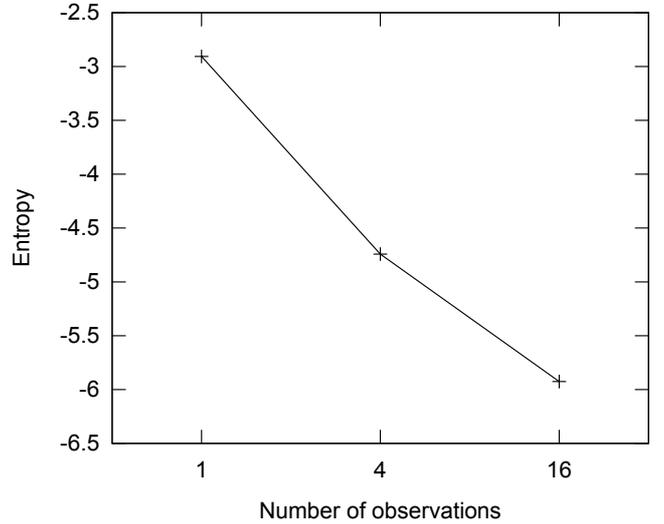


図 5: 観測回数を変えたときの事後分布のエントロピー. 100 試行間の中央値を示した. $T = 1$. $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$.

ある. 画素数は $n^2 = 64^2$ である. 観測回数は $T = 1$ である. 図 1 には式 (12) により得られる厳密解に加えて, 平均場近似 (MF) 解を示した. ここで MF 解とは,

$$q_{\text{MF}}(\mathbf{x}, \mu, \alpha, \beta) = q(\mathbf{x})q(\mu)q(\alpha)q(\beta) \quad (13)$$

で表される確率分布のうち, カルバック・ライブラー情報量

$$D_{\text{KL}}[q||p] = \int dx q(x) \ln \frac{q(x)}{p(x)} \quad (14)$$

の意味で厳密解に最も近いものである. MF 解は変分ベイズ法を用いることにより比較的容易に求めることができ, 多重積分が伴うハイパーパラメータ推定の近似解として頻繁に用いられる [Attias 99]. 図 1 から分かることは, 同じハイパーパラメータから生成されたデータであるが, その MAP 値はデータ依存で確率的に大きく揺らぐことである. したがって, 点推定値の信頼度を評価できる分布推定を行うことは重要である. 左列に示した厳密な事後分布は, その広がりから MAP 値が真値から離れていることを読み取ることができる. また, MAP 値と離れているとはいえ, 真値にも十分な確率密度が与えられていることが分かる. 一方で, 点推定では頻繁に用いられる MF 解は, 分布推定では十分な役割を果たしていないことが分かる. なぜなら, MF 解は分布の広がりを小さく見積もっているためである. 分布の広がりにはエントロピー

$$H[p] = - \int dx p(x) \ln p(x) \quad (15)$$

により定量化できる. 分散が σ^2 であるガウス分布のエントロピーは $\ln(\sigma\sqrt{2\pi e})$ であり, エントロピーが大きいほど確率分

布が広がっていることを表す. 100 組のデータに対する事後分布のエントロピーを計算し, その中央値を求めると, 厳密解は -2.907 であり, MF 解は -4.767 である. MF 解は, MAP 値が真値から離れているにもかかわらず, 信頼度を過大評価する危険性を伴う.

次に画素数 n^2 を変えて分布推定を行う. 画像データの画素数を多くすると, ハイパーパラメータ推定の信頼性が増すことが予想される. 図 2 に $n^2 = 64^2, 128^2, 256^2$ とした場合の事後分布を示す. ハイパーパラメータの真値は $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$ である. 観測回数は $T = 1$ である. 図 2 に示したように, 画素数を多くした場合に事後分布が狭まった. 事後分布の広がりを定量的に評価するために, それぞれの場合について事後分布のエントロピーを求めた. 図 3 に, 100 試行間の中央値を示した. 画素数を増やすことにより, ハイパーパラメータ事後分布の広がりが小さくなることを定量的に確かめられた. 解像度の向上に伴う推定信頼度の向上を分布推定が捉えられることが分かった.

それから観測回数 T を変えて分布推定を行う. 観測回数を多くすれば, ハイパーパラメータ推定の信頼性が増すことは予想される. 図 4 に $T = 1, 4, 16$ とした場合の事後分布を示す. ハイパーパラメータの真値は $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$ である. 画素数は $n^2 = 64^2$ である. 図 4 に示したように, 画素数を多くした場合に事後分布が狭まった. 図 2 と図 4 とを比較すると, 分布の狭まり方が異なることが分かる. 図 4 では, 解像度を一定にして観測回数を増やした. これは, 実質的に β に関するデータを増やすことに対応するため, β 方向の分布が狭まっている. 一方で, 図 2 では, 観測回数を一定にして, 解像

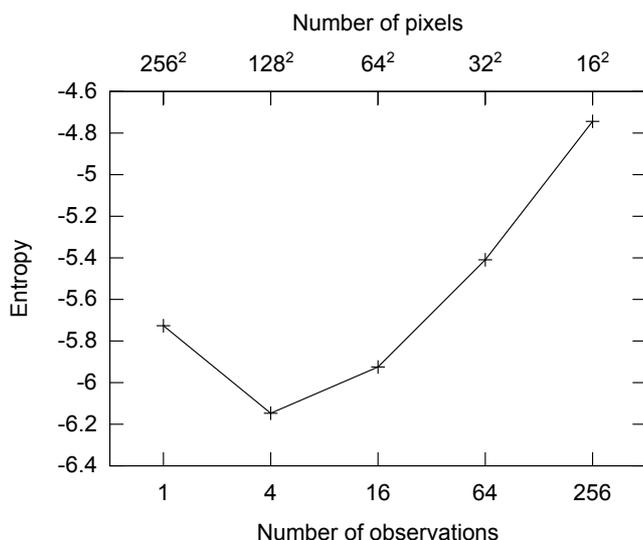


図 6: 観測時間が $n^2T = 2^{16}$ であるときの事後分布のエントロピー. $(\mu_0, \alpha_0, \beta_0) = (0, 1, 1)$.

度を上げた。これは、 β に関するデータと α に関するデータとの双方を増やすことに対応する。確率分布に関して言えば、動径方向に分布が狭まっている。事後分布の広がりを定量的に評価するために、それぞれの場合について事後分布のエントロピーを求めた。図 5 に、100 試行間の中央値を示した。観測回数を増やすことにより、ハイパーパラメータ事後分布の広がりが小さくなること定量的に確かめられた。観測回数の増加に伴う推定信頼度の向上を分布推定が捉えられることが分かった。

最後に画素数と観測回数とのトレード・オフについて議論する。撮像時間が限られているとき、解像度を上げると観測回数を減らさなければならず、観測回数を増やすためには解像度を落とさなければならない。画素数と観測回数との積が一定である時、どのような条件を選べば信頼度の高い観測が行えるかは非自明な問題である。実際に $n^2T = 2^{16}$ で固定し、 (n^2, T) を変えて事後分布のエントロピーを求めた。それぞれの場合について 100 試行し、その中央値を図 6 に示す。事後分布のエントロピーが小さいほど推定の信頼度が高いことを思い出すと、図 6 中の最適な観測は $(n^2, T) = (128^2, 4)$ であることが分かる。安易に解像度を上げたからといって推定の信頼度が上がるとは限らず、多少解像度を犠牲にしても観測回数を増やしたほうが良いことが示された。

5. 結論

本研究では、ベイズ推定の枠組みを用いてハイパーパラメータの分布推定を行う手法を提案した。また、解析的に事後分布が導出できるモデルを用いて提案手法の性能を評価した。提案手法は点推定値の信頼度を評価することができた。点推定で頻繁に用いられる MF 近似は分布推定では不十分であることを示した。最後に、画素数と観測回数とのトレード・オフの問題を論じた。

謝辞

本研究は、科学研究費補助金、特別研究員奨励費 (13J04920, 中西), 新学術領域研究 (25106506, 永田; 25120009, 岡田), 基盤研究 (B) (25280090, 岡田), 基盤研究 (C) (25330283,

永田; 25330285, 庄野) の助成を受けたものである。

参考文献

- [Geman 84] Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 6, pp. 721-741 (1984).
- [Pryce 95] Pryce, J. M. and Bruce, A. D.: Statistical mechanics of image restoration, *J. Phys. A: Math. Gen.*, Vol. 28, pp. 511-532 (1995).
- [Tanaka 02] Tanaka, K.: Statistical-mechanical approach to image processing, *J. Phys. A: Math. Gen.*, Vol. 35 R81-R150 (2002).
- [Demoment 89] Demoment, G.: Image reconstruction and restoration: overview of common estimation structures and problems, *IEEE Trans. Acoust. Speech Signal Process.*, Vol. 37, pp. 2024-2036 (1989).
- [Tanaka 07] Tanaka, K. and Titterton, D. M.: Statistical trajectory of an approximate EM algorithm for probabilistic image processing, *J. Phys. A: Math. Theor.*, Vol. 40, pp. 11285-11300 (2007).
- [Ohno 12] Ohno, Y., Nagata, K., Kuwatani, T. Shouno, H., and Okada, M.: Deterministic algorithm for nonlinear Markov random field model, *J. Phys. Soc. Japan*, Vol. 81, pp. 064006-1-064006-6 (2012).
- [Bishop 06] Bishop, C. M: *Pattern Recognition and Machine Learning* (New York: Springer) (2006).
- [Attias 99] Attias, H: Inferring parameters and structure of latent variable models by variational Bayes, in *Proc. 15th Conf. on Uncertainty in Artificial Intelligence* (San Francisco, CA: Morgan Kaufmann) pp. 21-30 (1999).