

メッセージ伝播法入門

後半第5回講義資料

状態発展法：厳密なアプローチ1

広島市立大学
令和4年9月26日～9月28日

豊橋技術科学大学
電気・電子情報工学系
准教授 竹内啓悟

状態発展方程式

以下の状態発展方程式[3-4]の厳密な証明を与える。

$$v^t = \sigma^2 + \frac{1}{\delta} \Psi_t(v^{t-1}), \quad v^0 = 1,$$

$$\Psi_{t+1}(v) = \mathbb{E}[\{x_1 - \eta_t(x_1 + \sqrt{v}\omega_1)\}^2], \quad \omega_1 \sim \mathcal{N}(0,1).$$

Bolthausenの方法[3-4, 5-1]

1. A を $\mathcal{N}(0,1/M)$ に従う独立な要素を持つ行列とする。
2. 反復 t 時点までのすべての推定値の履歴を条件として与えたときに、観測行列 A の条件付き分布を評価する。
3. 上記の条件付き分布を使って、反復 $t + 1$ での推定値の分布を厳密に計算する。

[5-1] E. Bolthausen, “An iterative construction of solutions of the TAP equations for the Sherrington-Kirkpatrick model,” *Commun. Math. Phys.*, vol. 325, no. 1, pp. 333–366, Jan. 2014.

誤差の発展方程式

AMP[3-1]

$$\mathbf{z}^t = \mathbf{y} - A\hat{\mathbf{x}}^t + \frac{\langle \eta'_{t-1}(\hat{\mathbf{x}}^{t-1} + A^T \mathbf{z}^{t-1}) \rangle}{\delta} \mathbf{z}^{t-1}, \quad \mathbf{z}^{-1} = \mathbf{0},$$

$$\hat{\mathbf{x}}^{t+1} = \eta_t(\hat{\mathbf{x}}^t + A^T \mathbf{z}^t).$$

誤差の発展方程式

$\mathbf{q}_t = \hat{\mathbf{x}}^t - \mathbf{x}$ 、 $\mathbf{h}_t = \hat{\mathbf{x}}^t + A^T \mathbf{z}^t - \mathbf{x}$ と誤差を定義する。
 $\mathbf{m}_t = \mathbf{z}^t$ 、 $\mathbf{b}_t = \mathbf{w} - \mathbf{z}^t$ とすると、

$$\mathbf{q}_{t+1} = \eta_t(\mathbf{x} + \mathbf{h}_t) - \mathbf{x}, \quad \mathbf{h}_t = A^T \mathbf{m}_t + \mathbf{q}_t,$$

$$\mathbf{m}_t = \mathbf{w} - \mathbf{b}_t, \quad \mathbf{b}_t = A\mathbf{q}_t - \lambda_{t-1}\mathbf{m}_{t-1}, \quad \lambda_t = \frac{\langle \eta'_t(\mathbf{x} + \mathbf{h}_t) \rangle}{\delta}.$$

誤差の発展方程式

$$\mathbf{B}_t = (\mathbf{b}_0, \dots, \mathbf{b}_{t-1}), \quad \mathbf{M}_t = (\mathbf{m}_0, \dots, \mathbf{m}_{t-1}),$$

$$\mathbf{H}_t = (\mathbf{h}_0, \dots, \mathbf{h}_{t-1}), \quad \mathbf{Q}_t = (\mathbf{q}_0, \dots, \mathbf{q}_{t-1}), \quad (\mathbf{0}, \mathbf{M}_0 \Lambda_0) = \mathbf{0},$$

$$\Lambda_t = \text{diag}\{\lambda_0, \dots, \lambda_{t-1}\}, \quad \mathbf{W}_t = (\mathbf{w}, \dots, \mathbf{w}) \in \mathbb{R}^{M \times t}.$$

上記の定義を使うと、

$$\mathbf{B}_t = \mathbf{A} \mathbf{Q}_t - (\mathbf{0}, \mathbf{M}_{t-1} \Lambda_{t-1}), \quad \mathbf{M}_t = \mathbf{W}_t - \mathbf{B}_t,$$

$$\mathbf{H}_t = \mathbf{A}^T \mathbf{M}_t + \mathbf{Q}_t, \quad \mathbf{q}_{t+1} = \eta_t(\mathbf{x} + \mathbf{h}_t) - \mathbf{x}.$$

特異値分解と射影行列

フルランク縦長行列 M の特異値分解

$$M = \Phi_M \begin{pmatrix} \Sigma_M \\ \mathbf{0} \end{pmatrix} \Psi_M^T, \quad \Phi_M = (\Phi_M^{\parallel}, \Phi_M^{\perp}).$$

一般化逆行列

$$M^{\dagger} = (M^T M)^{-1} M^T = \Psi_M \Sigma_M^{-1} (\Phi_M^{\parallel})^T.$$

M が張る空間への射影

$$P_M^{\parallel} = M M^{\dagger} = \Phi_M^{\parallel} (\Phi_M^{\parallel})^T.$$

直交補空間への射影

$$P_M^{\perp} = I - P_M^{\parallel} = \Phi_M^{\perp} (\Phi_M^{\perp})^T.$$

補題5.1 [3-4]

A を $\mathcal{N}(0,1/M)$ に従う独立な要素を持つ行列とする。

$X \in \mathbb{R}^{M \times t}$ 、 $Y \in \mathbb{R}^{N \times t'}$ 、 $U \in \mathbb{R}^{N \times t}$ 、 $V \in \mathbb{R}^{M \times t'}$ を縦長行列とし、以下の制約を満たすものとする。

$$X = AU, \quad Y = A^T V.$$

U と V がフルランクならば、 X 、 U 、 Y 、 V が与えられたときに、 A は以下の分布に従う。

$$\begin{aligned} A &\sim XU^\dagger + (V^\dagger)^T Y^T P_U^\perp + \Phi_V^\perp \tilde{A} (\Phi_U^\perp)^T \\ &\sim P_V^\perp XU^\dagger + (V^\dagger)^T Y^T + \Phi_V^\perp \tilde{A} (\Phi_U^\perp)^T. \end{aligned}$$

ただし、 \tilde{A} は A と独立で、 $\mathcal{N}(0,1/M)$ に従う独立な要素を持つ行列である。

補題5.1の証明[5-2]

$$U = \Phi_U \begin{pmatrix} \Sigma_U \\ \mathbf{0} \end{pmatrix} \Psi_U^T, \quad V = \Phi_V \begin{pmatrix} \Sigma_V \\ \mathbf{0} \end{pmatrix} \Psi_V^T.$$

上記の特異値分解を二つの制約式に代入すると、

$$\Phi_V^T X \Psi_U = \hat{A} \begin{pmatrix} \Sigma_U \\ \mathbf{0} \end{pmatrix}, \quad \Phi_U^T Y \Psi_V = \hat{A}^T \begin{pmatrix} \Sigma_V \\ \mathbf{0} \end{pmatrix}, \quad \hat{A} = \Phi_V^T A \Phi_U.$$

$$\hat{A} = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{pmatrix} \text{と分解すると、}$$

$$\begin{pmatrix} \hat{A}_{11} \\ \hat{A}_{21} \end{pmatrix} = \Phi_V^T X \Psi_U \Sigma_U^{-1}, \quad (\hat{A}_{11}, \hat{A}_{12}) = \Sigma_V^{-1} \Psi_V^T Y^T \Phi_U.$$

[5-2] K. Takeuchi, "Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 368-386, Jan. 2020.

補題5.1の証明

U と V が与えられたときに $A \sim \hat{A}$ なので、 \hat{A} の未知の要素 \hat{A}_{22} は $\mathcal{N}(0, 1/M)$ に従う独立な要素を持つ。

$$\begin{aligned} A &= \Phi_V \hat{A} \Phi_U^T = \Phi_V \begin{pmatrix} \Phi_V^T X \Psi_U \Sigma_U^{-1} & \Sigma_V^{-1} \Psi_V^T Y^T \Phi_U^\perp \\ & \hat{A}_{22} \end{pmatrix} \Phi_U^T \\ &= \Phi_V \begin{pmatrix} \Sigma_V^{-1} \Psi_V^T Y^T \Phi_U \\ (\Phi_V^\perp)^T X \Psi_U \Sigma_U^{-1} & \hat{A}_{22} \end{pmatrix} \Phi_U^T \end{aligned}$$

前者の表現を使って、

$$\begin{aligned} A &= X \Psi_U \Sigma_U^{-1} (\Phi_U^\parallel)^T + \Phi_V^\parallel \Sigma_V^{-1} \Psi_V^T Y^T P_U^\perp + \Phi_V^\perp \hat{A}_{22} (\Phi_U^\perp)^T \\ &= XU^\dagger + (V^\dagger)^T Y^T P_U^\perp + \Phi_V^\perp \hat{A}_{22} (\Phi_U^\perp)^T. \end{aligned}$$

もう一つの表現は後者から得られる。 ■

Steinの補題

$(Z_1, Z_2) \sim \mathcal{N}(\mathbf{0}, \Sigma)$ とすると、

$$\mathbb{E}[Z_1 f(Z_2)] = \mathbb{E}[Z_1 Z_2] \mathbb{E}[f'(Z_2)].$$

証明 固有分解 $\Sigma = V \Lambda V^T$ に対して、以下の変数変換を考える。

$$\begin{pmatrix} \tilde{Z}_1 \\ \tilde{Z}_2 \end{pmatrix} = V^T \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

$$\begin{aligned} \mathbb{E}[Z_1 f(Z_2)] &= \mathbb{E}[(V_{11} \tilde{Z}_1 + V_{12} \tilde{Z}_2) f(V_{21} \tilde{Z}_1 + V_{22} \tilde{Z}_2)] \\ &= V_{11} \mathbb{E}_{\tilde{Z}_2} \left[\mathbb{E}_{\tilde{Z}_1} [\tilde{Z}_1 f(V_{21} \tilde{Z}_1 + V_{22} \tilde{Z}_2)] \right] \\ &\quad + V_{12} \mathbb{E}_{\tilde{Z}_1} \left[\mathbb{E}_{\tilde{Z}_2} [\tilde{Z}_2 f(V_{21} \tilde{Z}_1 + V_{22} \tilde{Z}_2)] \right]. \end{aligned}$$

最後の等号は、 \tilde{Z}_1 と \tilde{Z}_2 は互いに独立なためである。

Steinの補題の証明

$\mathbb{E}[\tilde{Z}_1^2] = \lambda_1$ と $\mathbb{E}[\tilde{Z}_2^2] = \lambda_2$ に注意して、内側の期待値の評価に一変数のSteinの補題を使う。(第3回講義資料を参照)

$$\begin{aligned}\mathbb{E}[Z_1 f(Z_2)] &= \lambda_1 V_{11} V_{21} \mathbb{E}[f'(V_{21} \tilde{Z}_1 + V_{22} \tilde{Z}_2)] \\ &\quad + \lambda_2 V_{12} V_{22} \mathbb{E}[f'(V_{21} \tilde{Z}_1 + V_{22} \tilde{Z}_2)] \\ &= (\lambda_1 V_{11} V_{21} + \lambda_2 V_{12} V_{22}) \mathbb{E}[f'(Z_2)].\end{aligned}$$

共分散行列 $\Sigma = V \Lambda V^T$ の非対角成分を計算すると、

$$[\Sigma]_{12} = \lambda_1 V_{11} V_{21} + \lambda_2 V_{12} V_{22}.$$

したがって、Steinの補題を得る。

$$\mathbb{E}[Z_1 f(Z_2)] = \mathbb{E}[Z_1 Z_2] \mathbb{E}[f'(Z_2)]. \quad \blacksquare$$