メッセージ伝播法入門
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５年９月２５日～９月２６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

Boltausenの方法[3-3, 5-1]

GAMPの状態発展方程式[3−3, 3−4]の厳密な証明を与える。

1. $\boldsymbol{A}$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$\boldsymbol{A}$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

GAMP[3–4]　$\hat{\boldsymbol{x}}_0=\boldsymbol{0}$と$\hat{\boldsymbol{z}}_{-1}=\boldsymbol{0}$として、

$$\boldsymbol{z}_t=\boldsymbol{A}\hat{\boldsymbol{x}}_t+\frac{\xi_{\mathrm{in},t-1}}{\delta\xi_{\mathrm{out},t-1}}\hat{\boldsymbol{z}}_{t-1},\qquad \hat{\boldsymbol{z}}_t=\eta_{\mathrm{out},t}(\boldsymbol{z}_t,\boldsymbol{y}),$$

$$\boldsymbol{x}_t=\hat{\boldsymbol{x}}_t-\frac{1}{\xi_{\mathrm{out},t}}\boldsymbol{A}^{\mathrm{T}}\hat{\boldsymbol{z}}_t,\qquad \hat{\boldsymbol{x}}_{t+1}=\eta_{\mathrm{in},t}(\boldsymbol{x}_t)$$

誤差　$\boldsymbol{q}_t=\hat{\boldsymbol{x}}_t-\bar{\zeta}_t\bar{\xi}_{\mathrm{out},t}^{-1}\boldsymbol{x}$、$\boldsymbol{h}_t=\boldsymbol{x}_t-\bar{\zeta}_t\bar{\xi}_{\mathrm{out},t}^{-1}\boldsymbol{x}$を定義する。
$\boldsymbol{m}_t=\xi_{\mathrm{out},t}^{-1}\hat{\boldsymbol{z}}_t$、$\boldsymbol{b}_t=\boldsymbol{z}_t-\bar{\zeta}_t\bar{\xi}_{\mathrm{out},t}^{-1}\boldsymbol{z}$とし、$\boldsymbol{z}=\boldsymbol{A}\boldsymbol{x}$を使って

$$\boldsymbol{q}_{t+1}=\eta_{\mathrm{in},t}\left(\frac{\bar{\zeta}_t}{\bar{\xi}_{\mathrm{out},t}}\boldsymbol{x}+\boldsymbol{h}_t\right)-\frac{\bar{\zeta}_{t+1}}{\bar{\xi}_{\mathrm{out},t+1}}\boldsymbol{x},\quad \boldsymbol{h}_t=\boldsymbol{q}_t-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{m}_t,$$

$$\boldsymbol{m}_t=\frac{1}{\xi_{\mathrm{out},t}}\eta_{\mathrm{out},t}\left(\frac{\bar{\zeta}_t}{\bar{\xi}_{\mathrm{out},t}}\boldsymbol{z}+\boldsymbol{b}_t,\boldsymbol{y}\right),\quad \boldsymbol{b}_t=\boldsymbol{A}\boldsymbol{q}_t+\frac{\xi_{\mathrm{in},t-1}}{\delta}\boldsymbol{m}_{t-1}$$

ただし、$\boldsymbol{q}_0=-\bar{\zeta}_0\bar{\xi}_{\mathrm{out},0}^{-1}\boldsymbol{x}$、$\boldsymbol{m}_{-1}=\boldsymbol{0}$

誤差の発展方程式

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

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

$$\boldsymbol{\Lambda}_t = \frac{1}{\delta}\mathrm{diag}\{\xi_{\mathrm{in},0}, \dots, \xi_{\mathrm{in},t-1}\},$$

上記の定義を使うと、

$$\boldsymbol{B}_t = \boldsymbol{A}\boldsymbol{Q}_t + (\boldsymbol{0}, \boldsymbol{M}_{t-1}\boldsymbol{\Lambda}_{t-1}), \quad \boldsymbol{m}_t = \eta_{\mathrm{out},t}\left(\frac{\bar{\zeta}_t}{\bar{\xi}_{\mathrm{out},t}}\boldsymbol{z} + \boldsymbol{b}_t, \boldsymbol{y}\right),$$

$$\boldsymbol{H}_t = \boldsymbol{Q}_t - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{M}_t, \qquad \boldsymbol{q}_{t+1} = \eta_{\mathrm{in},t}\left(\frac{\bar{\zeta}_t}{\bar{\xi}_{\mathrm{out},t}}\boldsymbol{x} + \boldsymbol{h}_t\right) - \frac{\bar{\zeta}_{t+1}}{\bar{\xi}_{t+1}}\boldsymbol{x},$$

特異値分解と射影行列

フルランク縦長行列$\boldsymbol{M}$の特異値分解

$$\boldsymbol{M} = \boldsymbol{\Phi}_{\boldsymbol{M}} \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{M}} \\ \boldsymbol{O} \end{pmatrix} \boldsymbol{\Psi}_{\boldsymbol{M}}^{\mathrm{T}}, \qquad \boldsymbol{\Phi}_{\boldsymbol{M}} = \left(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}, \boldsymbol{\Phi}_{\boldsymbol{M}}^{\perp}\right).$$

一般化逆行列

$$\boldsymbol{M}^{\dagger} = \left(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\right)^{-1}\boldsymbol{M}^{\mathrm{T}} = \boldsymbol{\Psi}_{\boldsymbol{M}}\boldsymbol{\Sigma}_{\boldsymbol{M}}^{-1}\left(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}\right)^{\mathrm{T}}.$$

$\boldsymbol{M}$が張る空間への射影

$$\boldsymbol{P}_{\boldsymbol{M}}^{\parallel} = \boldsymbol{M}\boldsymbol{M}^{\dagger} = \boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}\left(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}\right)^{\mathrm{T}}.$$

直交補空間への射影

$$\boldsymbol{P}_{\boldsymbol{M}}^{\perp} = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{M}}^{\parallel} = \boldsymbol{\Phi}_{\boldsymbol{M}}^{\perp}\left(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\perp}\right)^{\mathrm{T}}.$$

補題5.1[3-3]

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

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

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{U}, \qquad \boldsymbol{Y} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}.$$

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

$$\begin{aligned}\boldsymbol{A} &\sim \boldsymbol{X}\boldsymbol{U}^{\dagger} + \left(\boldsymbol{V}^{\dagger}\right)^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}_{\boldsymbol{U}}^{\perp} + \boldsymbol{\Phi}_{\boldsymbol{V}}^{\perp}\widetilde{\boldsymbol{A}}\left(\boldsymbol{\Phi}_{\boldsymbol{U}}^{\perp}\right)^{\mathrm{T}} \\ &\sim \boldsymbol{P}_{\boldsymbol{V}}^{\perp}\boldsymbol{X}\boldsymbol{U}^{\dagger} + \left(\boldsymbol{V}^{\dagger}\right)^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}} + \boldsymbol{\Phi}_{\boldsymbol{V}}^{\perp}\widetilde{\boldsymbol{A}}\left(\boldsymbol{\Phi}_{\boldsymbol{U}}^{\perp}\right)^{\mathrm{T}}.\end{aligned}$$

ただし、$\widetilde{\boldsymbol{A}}$は$\boldsymbol{A}$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]

$$\boldsymbol{U} = \boldsymbol{\Phi}_U \begin{pmatrix} \boldsymbol{\Sigma}_U \\ \boldsymbol{O} \end{pmatrix} \boldsymbol{\Psi}_U^{\mathrm{T}}, \qquad \boldsymbol{V} = \boldsymbol{\Phi}_V \begin{pmatrix} \boldsymbol{\Sigma}_V \\ \boldsymbol{O} \end{pmatrix} \boldsymbol{\Psi}_V^{\mathrm{T}}.$$

上記の特異値分解を二つの制約式に代入すると、

$$\boldsymbol{\Phi}_V^{\mathrm{T}} \boldsymbol{X} \boldsymbol{\Psi}_U = \widehat{\boldsymbol{A}} \begin{pmatrix} \boldsymbol{\Sigma}_U \\ \boldsymbol{O} \end{pmatrix}, \qquad \boldsymbol{\Phi}_U^{\mathrm{T}} \boldsymbol{Y} \boldsymbol{\Psi}_V = \widehat{\boldsymbol{A}}^{\mathrm{T}} \begin{pmatrix} \boldsymbol{\Sigma}_V \\ \boldsymbol{O} \end{pmatrix}, \qquad \widehat{\boldsymbol{A}} = \boldsymbol{\Phi}_V^{\mathrm{T}} \boldsymbol{A} \boldsymbol{\Phi}_U.$$

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

$$\begin{pmatrix} \widehat{\boldsymbol{A}}_{11} \\ \widehat{\boldsymbol{A}}_{21} \end{pmatrix} = \boldsymbol{\Phi}_V^{\mathrm{T}} \boldsymbol{X} \boldsymbol{\Psi}_U \boldsymbol{\Sigma}_U^{-1}, \qquad \left(\widehat{\boldsymbol{A}}_{11}, \widehat{\boldsymbol{A}}_{12}\right) = \boldsymbol{\Sigma}_V^{-1} \boldsymbol{\Psi}_V^{\mathrm{T}} \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{\Phi}_U.$$

[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

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

$$\boldsymbol{A} = \boldsymbol{\Phi}_{\boldsymbol{V}} \widehat{\boldsymbol{A}} \boldsymbol{\Phi}_{\boldsymbol{U}}^{\mathrm{T}} = \boldsymbol{\Phi}_{\boldsymbol{V}} \begin{pmatrix} \boldsymbol{\Phi}_{\boldsymbol{V}}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{\Psi}_{\boldsymbol{U}} \boldsymbol{\Sigma}_{\boldsymbol{U}}^{-1} & \boldsymbol{\Sigma}_{\boldsymbol{V}}^{-1} \boldsymbol{\Psi}_{\boldsymbol{V}}^{\mathrm{T}} \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{\Phi}_{\boldsymbol{U}}^{\perp} \\ & \widehat{\boldsymbol{A}}_{22} \end{pmatrix} \boldsymbol{\Phi}_{\boldsymbol{U}}^{\mathrm{T}}$$

$$= \boldsymbol{\Phi}_{\boldsymbol{V}} \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{V}}^{-1} \boldsymbol{\Psi}_{\boldsymbol{V}}^{\mathrm{T}} \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{\Phi}_{\boldsymbol{U}} & \\ \left(\boldsymbol{\Phi}_{\boldsymbol{V}}^{\perp}\right)^{\mathrm{T}} \boldsymbol{X} \boldsymbol{\Psi}_{\boldsymbol{U}} \boldsymbol{\Sigma}_{\boldsymbol{U}}^{-1} & \widehat{\boldsymbol{A}}_{22} \end{pmatrix} \boldsymbol{\Phi}_{\boldsymbol{U}}^{\mathrm{T}}$$

前者の表現を使って、

$$\boldsymbol{A} = \boldsymbol{X} \boldsymbol{\Psi}_{\boldsymbol{U}} \boldsymbol{\Sigma}_{\boldsymbol{U}}^{-1} \left(\boldsymbol{\Phi}_{\boldsymbol{U}}^{\parallel}\right)^{\mathrm{T}} + \boldsymbol{\Phi}_{\boldsymbol{V}}^{\parallel} \boldsymbol{\Sigma}_{\boldsymbol{V}}^{-1} \boldsymbol{\Psi}_{\boldsymbol{V}}^{\mathrm{T}} \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{P}_{\boldsymbol{U}}^{\perp} + \boldsymbol{\Phi}_{\boldsymbol{V}}^{\perp} \widehat{\boldsymbol{A}}_{22} \left(\boldsymbol{\Phi}_{\boldsymbol{U}}^{\perp}\right)^{\mathrm{T}}$$

$$= \boldsymbol{X} \boldsymbol{U}^{\dagger} + \left(\boldsymbol{V}^{\dagger}\right)^{\mathrm{T}} \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{P}_{\boldsymbol{U}}^{\perp} + \boldsymbol{\Phi}_{\boldsymbol{V}}^{\perp} \widehat{\boldsymbol{A}}_{22} \left(\boldsymbol{\Phi}_{\boldsymbol{U}}^{\perp}\right)^{\mathrm{T}}.$$

もう一つの表現は後者から得られる。 ■

TOYOHASHI UNIVERSITY OF TECHNOLOGY

Steinの補題

$Z \sim \mathcal{N}(0, \sigma^2)$とすると、

$$\mathbb{E}[Zf(Z)] = \sigma^2 \mathbb{E}[f'(Z)].$$

証明

$$p(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2\sigma^2}},$$

$$\mathbb{E}[Zf(Z)] = \int zf(z)p(z)dz \; = -\sigma^2 \int f(z)p'(z)dz$$

$$= \sigma^2 \int f'(z)p(z)dz = \sigma^2 \mathbb{E}[f'(Z)]. \qquad \blacksquare$$

Steinの補題の一般化

$(Z_1, Z_2, Z_3) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$とすると、

$$\mathbb{E}[Z_1 f(Z_2, Z_3)] = \mathbb{E}[Z_1 Z_2]\mathbb{E}\left[\frac{\partial f}{\partial Z_2}(Z_2, Z_3)\right] + \mathbb{E}[Z_1 Z_3]\mathbb{E}\left[\frac{\partial f}{\partial Z_3}(Z_2, Z_3)\right].$$

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

$$\begin{pmatrix} \tilde{Z}_1 \\ \tilde{Z}_2 \\ \tilde{Z}_3 \end{pmatrix} = \boldsymbol{V}^{\mathrm{T}} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

$$\begin{aligned} \mathbb{E}[Z_1 f(Z_2, Z_3)] &= \mathbb{E}\left[\left(V_{11}\tilde{Z}_1 + V_{12}\tilde{Z}_2 + V_{13}\tilde{Z}_3\right) f(Z_2, Z_3)\right] \\ &= V_{11}\mathbb{E}\left[\mathbb{E}_{\tilde{Z}_1}[\tilde{Z}_1 f(Z_2, Z_3)]\right] + V_{12}\mathbb{E}\left[\mathbb{E}_{\tilde{Z}_2}[\tilde{Z}_2 f(Z_2, Z_3)]\right] \\ &\quad + V_{13}\mathbb{E}\left[\mathbb{E}_{\tilde{Z}_3}[\tilde{Z}_3 f(Z_2, Z_3)]\right]. \end{aligned}$$

最後の等号は、$\{\tilde{Z}_1, \tilde{Z}_2, \tilde{Z}_3\}$は独立なためである。

Steinの補題の証明

$$Z_2 = V_{21}\tilde{Z}_1 + V_{22}\tilde{Z}_2 + V_{23}\tilde{Z}_3, \quad Z_3 = V_{31}\tilde{Z}_1 + V_{32}\tilde{Z}_2 + V_{33}\tilde{Z}_3$$

$\mathbb{E}\left[\tilde{Z}_i^2\right] = \lambda_i$に注意し、内側の期待値に一変数のSteinの補題を使う。

$$\begin{aligned}\mathbb{E}[Z_1 f(Z_2, Z_3)] = &\lambda_1 V_{11}\left(V_{21}\mathbb{E}\left[\frac{\partial f}{\partial Z_2}\right] + V_{31}\mathbb{E}\left[\frac{\partial f}{\partial Z_3}\right]\right)\\ &+\lambda_2 V_{12}\left(V_{22}\mathbb{E}\left[\frac{\partial f}{\partial Z_2}\right] + V_{32}\mathbb{E}\left[\frac{\partial f}{\partial Z_3}\right]\right)\\ &+\lambda_3 V_{13}\left(V_{23}\mathbb{E}\left[\frac{\partial f}{\partial Z_2}\right] + V_{33}\mathbb{E}\left[\frac{\partial f}{\partial Z_3}\right]\right)\end{aligned}$$

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

$$\mathbb{E}\left[Z_1 Z_j\right] = [\boldsymbol{\Sigma}]_{1j} = \lambda_1 V_{11} V_{j1} + \lambda_2 V_{12} V_{j2} + \lambda_3 V_{13} V_{j3}.$$

よって、

$$\mathbb{E}[Z_1 f(Z_2, Z_3)] = \mathbb{E}[Z_1 Z_2]\mathbb{E}\left[\frac{\partial f}{\partial Z_2}\right] + \mathbb{E}[Z_1 Z_3]\mathbb{E}\left[\frac{\partial f}{\partial Z_3}\right] \quad \blacksquare$$

TOYOHASHI UNIVERSITY OF TECHNOLOGY