メッセージ伝播法入門
第６回講義資料
状態発展法：厳密なアプローチ２

九州大学
平成３０年１０月１０日～１０月１２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

主定理[3-4]

大システム極限で、以下が成り立つ。

(a) $\boldsymbol{b}_t \sim \boldsymbol{B}_t \boldsymbol{\beta}_t + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim \mathcal{N}(\boldsymbol{0}, M^{-1}\|\boldsymbol{q}_t^{\perp}\|^2 \boldsymbol{I}_M).$

(b) $\boldsymbol{h}_t \sim \boldsymbol{H}_t \boldsymbol{\alpha}_t + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim \mathcal{N}(\boldsymbol{0}, M^{-1}\|\boldsymbol{m}_t^{\perp}\|^2 \boldsymbol{I}_N).$

(c) $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'} - \frac{1}{M}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{m}_{t'} \to 0, \quad \frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{b}_{t'} - \frac{1}{\delta N}\boldsymbol{q}_t^{\mathrm{T}}\boldsymbol{q}_{t'} \to 0.$

(d) $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} - \frac{\delta\lambda_{t'-1}}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'-1} \to 0.$

(e) $\frac{1}{N}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{w} \to 0, \quad \frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{x} \to 0.$

$$\boldsymbol{\alpha}_t = \boldsymbol{M}_t^{\dagger}\boldsymbol{m}_t, \quad \boldsymbol{\beta}_t = \boldsymbol{Q}_t^{\dagger}\boldsymbol{q}_t = (\beta_0, \dots, \beta_{t-1})^{\mathrm{T}},$$

$$\boldsymbol{m}_t^{\perp} = \boldsymbol{P}_{\boldsymbol{M}_t}^{\perp}\boldsymbol{m}_t, \quad \boldsymbol{q}_t^{\perp} = \boldsymbol{P}_{\boldsymbol{Q}_t}^{\perp}\boldsymbol{q}_t.$$

主定理の証明の流れ

(e)の証明　　(d)の証明と同じ。(省略)

(c)の証明　　(a)(b)を使う。

(d)の証明　　(b)(e)を使う。

(a)の証明　　(c)(d)を使う。

(b)の証明　　(a)を使う。(省略)

性質(c)の証明[3-4]

前者のみを帰納法により証明する。

$t < \tau, t' \leq t$に対して、 $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'} - \frac{1}{M}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{m}_{t'} \to 0$ を仮定して

$t = \tau, t' \leq t$の場合を証明する。

$t' < t$のとき　　性質(b)を使って、

$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_{t'} \sim \frac{1}{N}\boldsymbol{\alpha}_\tau^{\mathrm{T}}\boldsymbol{H}_\tau^{\mathrm{T}}\boldsymbol{h}_{t'} + \frac{1}{N}\boldsymbol{\zeta}_\tau^{\mathrm{T}}\boldsymbol{h}_{t'}.$$

大数の強法則から、第二項は0に概収束する。
第一項に帰納法の仮定と$\boldsymbol{\alpha}_\tau = \boldsymbol{M}_\tau^{\dagger}\boldsymbol{m}_\tau$を使うと、

$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_{t'} \to \frac{1}{M}\boldsymbol{\alpha}_\tau^{\mathrm{T}}\boldsymbol{M}_\tau^{\mathrm{T}}\boldsymbol{m}_{t'} = \frac{1}{M}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{P}_{\boldsymbol{M}_\tau}^{\parallel}\boldsymbol{m}_{t'} = \frac{1}{M}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{m}_{t'}.$$

性質(c)の証明[3-4]

$t' = t$のとき
$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_\tau \to \frac{1}{N}\boldsymbol{\alpha}_\tau^{\mathrm{T}}\boldsymbol{H}_\tau^{\mathrm{T}}\boldsymbol{H}_\tau\boldsymbol{\alpha}_\tau + \frac{1}{N}\boldsymbol{\zeta}_\tau^{\mathrm{T}}\boldsymbol{\zeta}_\tau.$$

帰納法の仮定と$\boldsymbol{\zeta}_\tau \sim \mathcal{N}(\boldsymbol{0}, M^{-1}\|\boldsymbol{m}_\tau^{\perp}\|^2\boldsymbol{I}_N)$を使うと、

$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_\tau \to \frac{1}{M}\boldsymbol{\alpha}_\tau^{\mathrm{T}}\boldsymbol{M}_\tau^{\mathrm{T}}\boldsymbol{M}_\tau\boldsymbol{\alpha}_\tau + \frac{1}{M}\|\boldsymbol{m}_\tau^{\perp}\|^2$$
$$= \frac{1}{M}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{P}_{\boldsymbol{M}_\tau}^{\parallel}\boldsymbol{m}_\tau + \frac{1}{M}\|\boldsymbol{m}_\tau^{\perp}\|^2 = \frac{1}{M}\left\|\boldsymbol{m}_\tau^{\parallel}\right\|^2 + \frac{1}{M}\|\boldsymbol{m}_\tau^{\perp}\|^2.$$

ただし、 $\boldsymbol{m}_\tau^{\parallel} = \boldsymbol{P}_{\boldsymbol{M}_\tau}^{\parallel}\boldsymbol{m}_\tau.$

$\left(\boldsymbol{m}_\tau^{\parallel}\right)^{\mathrm{T}}\boldsymbol{m}_\tau^{\perp} = 0$なので、 $\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_\tau \to \frac{1}{M}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{m}_\tau.$ ■

Steinの補題

$(Z_1, Z_2) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$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$\mathbf{\Sigma} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}$に対して、以下の変数変換を考える。

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

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

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)

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

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、

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

したがって、Steinの補題を得る。

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

性質(d)の証明[3-4]

定義$\boldsymbol{q}_{t'} = \eta_{t'-1}(\boldsymbol{x} + \boldsymbol{h}_{t'-1}) - \boldsymbol{x}$と性質(e)を使うと、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} \to \frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\eta_{t'-1}(\boldsymbol{x} + \boldsymbol{h}_{t'-1})$$

$$\to \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[[\boldsymbol{h}_t]_n\eta_{t'-1}(x_n + [\boldsymbol{h}_{t'-1}]_n)\right].$$

最後の表現は、大数の強法則のためである。性質(b)から$\boldsymbol{h}_{t'}$はi.i.d.ガウス要素を持つため、Steinの補題を使うと、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} \to \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[[\boldsymbol{h}_t]_n[\boldsymbol{h}_{t'-1}]_n\right]\mathbb{E}\left[\eta'_{t'-1}(x_n + [\boldsymbol{h}_{t'-1}]_n)\right].$$

性質(d)の証明[3-4]

$\boldsymbol{h}_t$はi.i.d.な要素を持つため、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} \to \frac{1}{N}\mathbb{E}\left[\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'-1}\right]\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[\eta'_{t'-1}(x_n + [\boldsymbol{h}_{t'-1}]_n)\right].$$

定義　$\lambda_{t'-1} = \dfrac{\langle \eta'_{t'-1}(\boldsymbol{x} + \boldsymbol{h}_{t'-1})\rangle}{\delta}$　を使うと、

大数の強法則から、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} - \frac{\delta\lambda_{t'-1}}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'-1} \to 0. \qquad \blacksquare$$

性質(a)の証明[3-4]

$$\boldsymbol{B}_t + (\boldsymbol{0}, \boldsymbol{M}_{t-1}\boldsymbol{\Lambda}_{t-1}) = \boldsymbol{A}\boldsymbol{Q}_t, \qquad \boldsymbol{H}_t - \boldsymbol{Q}_t = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{M}_t.$$

上記の制約式の下で、補題5.1を使うと、

$$\boldsymbol{A} \sim \{\boldsymbol{B}_t + (\boldsymbol{0}, \boldsymbol{M}_{t-1}\boldsymbol{\Lambda}_{t-1})\}\boldsymbol{Q}_t^{\dagger} + \left(\boldsymbol{M}_t^{\dagger}\right)^{\mathrm{T}}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{P}_{\boldsymbol{Q}_t}^{\perp} + \boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\widetilde{\boldsymbol{A}}\left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp}\right)^{\mathrm{T}}.$$

ただし、$\boldsymbol{Q}_t^{\mathrm{T}}\boldsymbol{P}_{\boldsymbol{Q}_t}^{\perp} = \boldsymbol{O}$を使った。

定義$\boldsymbol{b}_t = \boldsymbol{A}\boldsymbol{q}_t - \lambda_{t-1}\boldsymbol{m}_{t-1}$に上記を代入して、

$$\boldsymbol{b}_t \sim \boldsymbol{B}_t\boldsymbol{\beta}_t + \sum_{t'=0}^{t-2}\beta_{t'+1}\lambda_{t'}\boldsymbol{m}_{t'} + \boldsymbol{M}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{q}_t^{\perp}$$

$$-\lambda_{t-1}\boldsymbol{m}_{t-1} + \boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\widetilde{\boldsymbol{A}}\left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp}\right)^{\mathrm{T}}\boldsymbol{q}_t.$$

性質(a)の証明[3-4]

$\boldsymbol{G}_t = M^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t$とし、$\boldsymbol{q}_t^{\perp} = \boldsymbol{q}_t - \boldsymbol{Q}_t\boldsymbol{\beta}_t$を使って第三項を評価する。

$$\boldsymbol{M}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{q}_t^{\perp} = \frac{1}{\delta N}\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{q}_t - \frac{1}{\delta N}\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{Q}_t\boldsymbol{\beta}_t.$$

性質(c)(d)から、

$$\frac{1}{\delta N}\left[\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{q}_t\right]_{t'} = \frac{1}{\delta N}\boldsymbol{h}_{t'}^{\mathrm{T}}\boldsymbol{q}_t \to \frac{\lambda_{t-1}}{N}\boldsymbol{h}_{t'}^{\mathrm{T}}\boldsymbol{h}_{t-1} \to \frac{\lambda_{t-1}}{M}\boldsymbol{m}_{t'}^{\mathrm{T}}\boldsymbol{m}_{t-1}.$$

上記を第一項に適用すると、

$$\frac{1}{\delta N}\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{q}_t \to \frac{\lambda_{t-1}}{M}\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_{t-1} = \lambda_{t-1}\boldsymbol{m}_{t-1}.$$

性質(a)の証明[3-4]

第二項に関して、

$$\frac{1}{\delta N}\left[\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{Q}_t\boldsymbol{\beta}_t\right]_{t'} = \frac{1}{\delta N}\boldsymbol{h}_{t'}^{\mathrm{T}}\boldsymbol{Q}_t\boldsymbol{\beta}_t = \frac{1}{\delta N}\sum_{\tau=0}^{t-1}\beta_\tau\boldsymbol{h}_{t'}^{\mathrm{T}}\boldsymbol{q}_\tau\,.$$

性質(c)(d)を使って、

$$\frac{1}{\delta N}\left[\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{Q}_t\boldsymbol{\beta}_t\right]_{t'} \to \frac{1}{N}\sum_{\tau=1}^{t-1}\beta_\tau\lambda_{\tau-1}\boldsymbol{h}_{t'}^{\mathrm{T}}\boldsymbol{h}_{\tau-1} \to \frac{1}{M}\sum_{\tau=1}^{t-1}\beta_\tau\lambda_{\tau-1}\boldsymbol{m}_{t'}^{\mathrm{T}}\boldsymbol{m}_{\tau-1}\,.$$

これを第二項に適用すると、

$$\frac{1}{\delta N}\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{Q}_t\beta_t \to \frac{1}{M}\sum_{\tau=0}^{t-2}\beta_{\tau+1}\lambda_\tau\boldsymbol{M}_t\boldsymbol{G}_t^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_\tau = \sum_{\tau=0}^{t-2}\beta_{\tau+1}\lambda_\tau\boldsymbol{m}_\tau\,.$$

性質(a)の証明[3-4]

以上の結果をまとめると、

$$\boldsymbol{b}_t \to \boldsymbol{B}_t \boldsymbol{\beta}_t + \boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp} \widetilde{\boldsymbol{A}} \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t.$$

$\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp} \widetilde{\boldsymbol{A}} \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t$は平均$\boldsymbol{0}$のガウス分布に従う。

$\widetilde{\boldsymbol{A}}_0 \in \mathbb{R}^{t \times (N-t)}$を$\mathcal{N}(0,1/M)$に従う独立な成分を持つ行列として、

$$\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp} \widetilde{\boldsymbol{A}} \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t = \boldsymbol{\Phi}_{\boldsymbol{M}_t} \begin{pmatrix} \widetilde{\boldsymbol{A}}_0 \\ \widetilde{\boldsymbol{A}} \end{pmatrix} \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t - \boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\parallel} \widetilde{\boldsymbol{A}}_0 \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t.$$

大システム極限で第二項は無視できるので、補題4.1を使うと、

$\left\| \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t \right\|^2 = \boldsymbol{q}_t^{\mathrm{T}} \boldsymbol{P}_{\boldsymbol{Q}_t}^{\perp} \boldsymbol{q}_t = \| \boldsymbol{q}_t^{\perp} \|^2$から、

$$\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp} \widetilde{\boldsymbol{A}} \left(\boldsymbol{\Phi}_{\boldsymbol{Q}_t}^{\perp} \right)^{\mathrm{T}} \boldsymbol{q}_t \sim \mathcal{N}\left(\boldsymbol{0}, M^{-1} \| \boldsymbol{q}_t^{\perp} \|^2 \boldsymbol{I}_M \right). \qquad \blacksquare$$