メッセージ伝播法入門
後半第6回講義資料
状態発展法：厳密なアプローチ2

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

主定理[3-4]

大システム極限で、以下が成り立つ。

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

$$\boldsymbol{\beta}_t = \boldsymbol{Q}_t^{\dagger}\boldsymbol{q}_t = (\beta_0, \dots, \beta_{t-1})^{\mathrm{T}}, \qquad \boldsymbol{q}_t^{\perp} = \boldsymbol{P}_{\boldsymbol{Q}_t}^{\perp}\boldsymbol{q}_t.$$

(A−2) $\frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{b}_{t'} - \frac{1}{\delta N}\boldsymbol{q}_t^{\mathrm{T}}\boldsymbol{q}_{t'} \to 0.$ (A−3) $\frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{w} \to 0,$

(A−4) $\frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{m}_{t'} + \frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{b}_{t'} \to 0.$

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

$$\boldsymbol{\alpha}_t = \boldsymbol{M}_t^{\dagger}\boldsymbol{m}_t, \qquad \boldsymbol{m}_t^{\perp} = \boldsymbol{P}_{\boldsymbol{M}_t}^{\perp}\boldsymbol{m}_t,$$

(B−2) $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'} - \frac{1}{M}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{m}_{t'} \to 0,$ (B−3) $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{x} \to 0.$

(B−4) $\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{q}_{t'} - \frac{\delta\lambda_{t'-1}}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_{t'-1} \to 0.$

状態発展方程式の導出

性質(B−1)から、誤差$\boldsymbol{h}_t$はガウス分布に従うi.i.d.成分からなることが言える。

以下を仮定して、$\mathrm{N}^{-1}\boldsymbol{q}_{t+1}^{\mathrm{T}}\boldsymbol{q}_{t+1}$を評価する。

$$\frac{1}{N}\boldsymbol{q}_t^{\mathrm{T}}\boldsymbol{q}_t \to \Psi_t(v^{t-1}), \qquad v^t = \frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_t.$$

性質(B−2)(A−3)(A−2)、$\boldsymbol{m}_t = \boldsymbol{w} - \boldsymbol{b}_t$、大数の強法則を使って、

$$v^t \to \frac{1}{M}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{m}_t \to \sigma^2 + \frac{1}{M}\boldsymbol{b}_t^{\mathrm{T}}\boldsymbol{b}_t \to \sigma^2 + \frac{1}{\delta}\Psi_t(v^{t-1}).$$

$$\frac{1}{N}\boldsymbol{q}_{t+1}^{\mathrm{T}}\boldsymbol{q}_{t+1} = \frac{1}{N}\sum_{n=1}^{N}\{\eta_t(x_n + [\boldsymbol{h}_t]_n) - x_n\}^2 \to \Psi_{t+1}(v^t).$$

主定理の証明の流れ

数学的帰納法によって、証明する。

$t = \tau$に対する(A−1)～(A−4)

仮定

$t < \tau$に対する(A−1)～(A−4)、
$t < \tau$に対する(B−1)～(B−4)

証明

- (B−2)(B−4)→(A−1)
- (A−1)→(A−2)
- (A−3)の証明は(A−4)と同じ
- (A−1)(A−3)→ (A−4)

$t = \tau$に対する(B−1)～(B−4)

仮定

$t \leq \tau$に対する(A−1)～(A−4)、
$t < \tau$に対する(B−1)～(B−4)

証明

- (A−2)(A−4)→(B−1)
- (B−1)→(B−2)
- (B−3)の証明は(B−4)と同じ
- (B−1)(B−3)→ (B−4)

性質(B-1)の証明[3-4]

$\boldsymbol{h}_t$を更新する直前までの反復を固定する。

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

上記の制約式の下で、補題5.1を使うと、

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

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

定義$\boldsymbol{h}_t = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{m}_t + \boldsymbol{q}_t$に上記を代入して、

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

TOYOHASHI UNIVERSITY OF TECHNOLOGY

性質(B-1)の証明[3-4]

$\boldsymbol{G}_t = N^{-1}\boldsymbol{Q}_t^{\mathrm{T}}\boldsymbol{Q}_t$とし、$\boldsymbol{m}_t^{\perp} = \boldsymbol{m}_t - \boldsymbol{M}_t\boldsymbol{\alpha}_t$を使って第一項を評価する。

$$\boldsymbol{Q}_{t+1}\left(\boldsymbol{Q}_{t+1}^{\mathrm{T}}\boldsymbol{Q}_{t+1}\right)^{-1}\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{m}_t^{\perp} = \frac{\delta}{M}\boldsymbol{Q}_{t+1}\boldsymbol{G}_{t+1}^{-1}\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{m}_t - \frac{\delta}{M}\boldsymbol{Q}_{t+1}\boldsymbol{G}_{t+1}^{-1}\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t.$$

性質(A−4) から、

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

$\because$ 性質(A−2)

上記を第一項に適用すると、

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

性質(B-1)の証明[3-4]

第二項に関して、

$$\frac{\delta}{M}\left[\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t\right]_{t'} = \frac{\delta}{M}\boldsymbol{b}_{t'}^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t = \frac{\delta}{M}\sum_{\tau=0}^{t-1}\alpha_\tau\boldsymbol{b}_{t'}^{\mathrm{T}}\boldsymbol{m}_\tau\,.$$

性質(A−4)(A−2)を使って、

$$\frac{\delta}{M}\left[\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t\right]_{t'} \to -\frac{\delta}{M}\sum_{\tau=0}^{t-1}\alpha_\tau\boldsymbol{b}_{t'}^{\mathrm{T}}\boldsymbol{b}_\tau \to -\frac{1}{N}\sum_{\tau=0}^{t-1}\alpha_\tau\boldsymbol{q}_{t'}^{\mathrm{T}}\boldsymbol{q}_\tau\,.$$

これを第二項に適用すると、

$$-\frac{\delta}{M}\boldsymbol{Q}_{t+1}\boldsymbol{G}_{t+1}^{-1}\boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t \to \frac{1}{N}\sum_{\tau=0}^{t-1}\alpha_\tau\boldsymbol{Q}_{t+1}\boldsymbol{G}_{t+1}^{-1}\boldsymbol{Q}_{t+1}^{\mathrm{T}}\boldsymbol{q}_\tau$$

$$= \sum_{\tau=0}^{t-1}\alpha_\tau\boldsymbol{q}_\tau = \boldsymbol{Q}_t\boldsymbol{\alpha}_t.$$

性質(B-1)の証明[3-4]

以上の結果をまとめると、

$$\boldsymbol{h}_t \to \boldsymbol{H}_t \boldsymbol{\alpha}_t + \boldsymbol{\Phi}_{\boldsymbol{Q}_{t+1}}^{\perp} \widetilde{\boldsymbol{A}}^{\mathrm{T}} \left(\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\right)^{\mathrm{T}} \boldsymbol{m}_t.$$

$\boldsymbol{\Phi}_{\boldsymbol{Q}_{t+1}}^{\perp} \widetilde{\boldsymbol{A}}^{\mathrm{T}} \left(\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\right)^{\mathrm{T}} \boldsymbol{m}_t$は平均$\boldsymbol{0}$のガウス分布に従う。

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

$$\boldsymbol{\Phi}_{\boldsymbol{Q}_{t+1}}^{\perp} \widetilde{\boldsymbol{A}}^{\mathrm{T}} \left(\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\right)^{\mathrm{T}} \boldsymbol{m}_t = \boldsymbol{\Phi}_{\boldsymbol{Q}_{t+1}} \begin{pmatrix} \widetilde{\boldsymbol{A}}_0^{\mathrm{T}} \\ \widetilde{\boldsymbol{A}}^{\mathrm{T}} \end{pmatrix} \left(\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\right)^{\mathrm{T}} \boldsymbol{m}_t - \boldsymbol{\Phi}_{\boldsymbol{Q}_{t+1}}^{\parallel} \widetilde{\boldsymbol{A}}_0^{\mathrm{T}} \left(\boldsymbol{\Phi}_{\boldsymbol{M}_t}^{\perp}\right)^{\mathrm{T}} \boldsymbol{m}_t.$$

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、

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

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

性質(B-2)の証明[3-4]

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−1)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(B-2)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、 $\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_\tau \to \frac{1}{M}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{m}_\tau.$ ■

性質(B-4)の証明[3-4]

定義$\boldsymbol{q}_{t'} = \eta_{t'-1}(\boldsymbol{x} + \boldsymbol{h}_{t'-1}) - \boldsymbol{x}$と性質(B−3)を使うと、

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

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

最後の表現は、大数の強法則のためである。性質(B−1)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、

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

性質(B-4)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、

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