状態発展法入門
～メッセージ伝播法を設計しよう～

SITA2023　ワークショップ
2023年11月29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

状態発展法の新規参入障壁1

条件付き期待値に関する厳密な表現

Xを確率空間$(\Omega, \mathcal{F}, P)$上の可積分な実確率変数とする。
実確率変数Yによって生成されるσ-加法族$\sigma(Y) \subset \mathcal{F}$が与えられたときの$X$の条件付き期待値$\mathbb{E}[X|\sigma(Y)]$を計算せよ。

条件付き期待値に関する素朴な表現

$\mathbb{E}[X|Y]$を計算せよ。

離散分布ならば
$$\mathbb{E}[X|Y=y] = \sum_x x P_{X|Y}(x|y)$$

絶対連続分布ならば
$$\mathbb{E}[X|Y=y] = \int x p_{X|Y}(x|y) dx$$

特異分布ならば　考えることを放棄

状態発展法の新規参入障壁2

実確率変数列$\{X_n\}_{n=1}^N$は、次数$k \geq 2$の疑似リプシッツの意味で確率変数Xに経験収束する。

$$\{X_1, \ldots, X_N\} \xrightarrow{\mathrm{PL}(k)} X$$

$\Longleftrightarrow$ 次数kの任意の疑似リプシッツ関数$f: \mathbb{R} \to \mathbb{R}$に対して、

$$\frac{1}{N}\sum_{n=1}^{N} f(X_n) \xrightarrow{\mathrm{a.s.}} \mathbb{E}[f(X)]$$

次数kの疑似リプシッツ関数　$L > 0$をリプシッツ定数として、

$$|f(x) - f(y)| \leq L|x - y|\left(1 + |x|^{k-1} + |y|^{k-1}\right) \text{ for all } x, y \in \mathbb{R}$$

経験収束の素朴な理解

$\{X_n\}$は、有界k次モーメント、組ごとに独立、同一な確率変数列

状態発展法の新規参入障壁3

- 学部教養レベルの数学（線形代数、微分積分学、確率論）に関する膨大な量の基礎的演算を行う。（本資料の後半参照）

- 論文では、上記の基礎的演算に関して、新規参入障壁1・2で述べた理学部数学科レベルの難解な説明が行われ、議論の本質が見えにくくなっている。

⇒ 「賢明な」研究者は、手出しすべきでない研究と判断する？

本ワークショップのねらい

過度に難解な説明を排除し、状態発展法の本質を解説する。

信号再構成問題

ノイズありの線形観測モデル

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}, \qquad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_M)$$

$\boldsymbol{y} = (y_1, \dots, y_M)^{\mathrm{T}} \in \mathbb{R}^M$: M次元観測ベクトル

$\boldsymbol{x} = (x_1, \dots, x_N)^{\mathrm{T}} \in \mathbb{R}^N$: N次元信号ベクトル（$M \le N$）

$\boldsymbol{w} = (w_1, \dots, w_M)^{\mathrm{T}} \in \mathbb{R}^M$: M次元ノイズベクトル

$\boldsymbol{A} = [A_{mn}] \in \mathbb{R}^{M \times N}$: M行N列の観測行列

目標

$\boldsymbol{y}$、$\boldsymbol{A}$、未知の確率変数の統計的性質に関する情報を使って、
多項式時間かつベイズ最適な性能で信号ベクトル$\boldsymbol{x}$を推定せよ。

信号推定アルゴリズム

直交/ベクトル近似的メッセージ伝播法(OAMP/VAMP)[1, 2]

モジュールA（線形推定）

$$\boldsymbol{x}_{\mathrm{A},t} = f_{\mathrm{A},t}\left(\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}, v_{\mathrm{B}\to\mathrm{A},t}; \boldsymbol{A}, \boldsymbol{y}\right),$$

$$\eta_{\mathrm{A},t} = \left\langle \eta_{\mathrm{A},t}\left(\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}, v_{\mathrm{B}\to\mathrm{A},t}; \boldsymbol{A}, \boldsymbol{y}\right)\right\rangle,$$

$$\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t} = \frac{\boldsymbol{x}_{\mathrm{A},t} - \eta_{\mathrm{A},t}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}}{1-\eta_{\mathrm{A},t}},$$

$$v_{\mathrm{A}\to\mathrm{B},t} = \frac{\eta_{\mathrm{A},t} v_{\mathrm{B}\to\mathrm{A},t}}{1-\eta_{\mathrm{A},t}}.$$

モジュールB（非線形推定）

$$\boldsymbol{x}_{\mathrm{B},t+1} = f_{\mathrm{B},t}\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}, v_{\mathrm{A}\to\mathrm{B},t}\right),$$

$$\eta_{\mathrm{B},t} = \left\langle \eta_{\mathrm{B},t}\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}, v_{\mathrm{A}\to\mathrm{B},t}\right)\right\rangle,$$

$$\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\boldsymbol{x}_{\mathrm{B},t+1} - \eta_{\mathrm{B},t}\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}}{1-\eta_{\mathrm{B},t}}$$

$$v_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\eta_{\mathrm{B},t} v_{\mathrm{A}\to\mathrm{B},t}}{1-\eta_{\mathrm{B},t}}.$$

表記　$f: \mathbb{R} \to \mathbb{R}$に対して、

$[f(\boldsymbol{x})]_n = f([\boldsymbol{x}]_n)$ 要素ごとに適用

$\langle \boldsymbol{a} \rangle = \frac{1}{N}\sum_{n=1}^{N} a_n$ 算術平均

[1] J. Ma and L. Ping, “Orthogonal AMP,” *IEEE Access*, vol. 5, pp. 2020–2033, Jan. 2017.

[2] S. Rangan, P. Schniter, and A. K. Fletcher, “Vector approximate message passing,” *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6664–6684, Oct. 2019.

モジュールAの詳細

線形フィルタ　　特異値分解$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$を使って、

$$f_{\mathrm{A},t}\left(\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}, v_{\mathrm{B}\to\mathrm{A},t}; \boldsymbol{A}, \boldsymbol{y}\right) = \boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t} + \boldsymbol{V}\boldsymbol{D}_t\boldsymbol{U}^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t})$$

$$\boldsymbol{D}_t = \begin{bmatrix} \widetilde{\boldsymbol{D}}_t \\ \boldsymbol{O} \end{bmatrix} \in \mathbb{R}^{N\times M}, \qquad \widetilde{\boldsymbol{D}}_t = \mathrm{diag}\left\{d_t\left(\sigma_m, v_{\mathrm{B}\to\mathrm{A},t}\right)\right\}$$

ただし、特異値 $\sigma_m > 0$は$\boldsymbol{\Sigma}$のm番目の対角成分を表す。

線形最小平均二乗誤差(LMMSE)フィルタ

$$\boldsymbol{D}_t = v_{\mathrm{B}\to\mathrm{A},t}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\left(\sigma^2\boldsymbol{I}_M + v_{\mathrm{B}\to\mathrm{A},t}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)^{-1}\boldsymbol{U}$$

$$= v_{\mathrm{B}\to\mathrm{A},t}\boldsymbol{\Sigma}^{\mathrm{T}}\left(\sigma^2\boldsymbol{I}_M + v_{\mathrm{B}\to\mathrm{A},t}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}}\right)^{-1}$$

$$= \begin{bmatrix} \widetilde{\boldsymbol{D}}_t \\ \boldsymbol{O} \end{bmatrix}, \qquad d_t(\sigma_m, v_{\mathrm{B}\to\mathrm{A},t}) = \frac{v_{\mathrm{B}\to\mathrm{A},t}\sigma_m}{\sigma^2 + v_{\mathrm{B}\to\mathrm{A},t}\sigma_m^2}.$$

OAMP/VAMPの設計方針

1. 一般化誤差モデルを定義する。

2. 状態発展法を使って、一般化誤差モデルを解析する。

   結果を主定理にまとめる。証明は後述する。

3. OAMP/VAMPの誤差モデルが一般化誤差モデルに含まれるように、オンサーガ補正の係数$\eta_{\mathrm{A},t}$と$\eta_{\mathrm{B},t}$を設計する。

TOYOHASHI UNIVERSITY OF TECHNOLOGY

一般化誤差モデル

二つの関数$\phi_t\colon \mathbb{R}^3 \to \mathbb{R}$と$\psi_t\colon \mathbb{R}^2 \to \mathbb{R}$を使って、

$$\boldsymbol{b}_t = \boldsymbol{V}^{\mathrm{T}} \boldsymbol{q}_t,$$

$$\boldsymbol{m}_t = \phi_t(\boldsymbol{b}_t, \widetilde{\boldsymbol{w}}, \boldsymbol{\sigma}) - \langle \phi_t'(\boldsymbol{b}_t, \widetilde{\boldsymbol{w}}, \boldsymbol{\sigma}) \rangle \boldsymbol{b}_t,$$

$$\boldsymbol{h}_t = \boldsymbol{V} \boldsymbol{m}_t,$$ モジュールA出力の誤差

$$\boldsymbol{q}_{t+1} = \psi_t(\boldsymbol{h}_t, \boldsymbol{x}) - \langle \psi_t'(\boldsymbol{h}_t, \boldsymbol{x}) \rangle \boldsymbol{h}_t.$$ モジュールB出力の誤差

ただし、微分は第一変数に関する偏微分を表す。

定義

$$\widetilde{\boldsymbol{w}} = \begin{bmatrix} \boldsymbol{U}^{\mathrm{T}} \boldsymbol{w} \\ \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^N,$$

$n > M$に対して$\sigma_n = 0$として、$\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)^{\mathrm{T}}$

仮定のまとめ

仮定１　信号ベクトル$\boldsymbol{x}$は、独立同一分布に従う要素を持つ。

仮定２　大システム極限($\delta = M/N$を固定して$M, N \to \infty$)を取る。

右直交不変　行列$\boldsymbol{M}$は、$\boldsymbol{M}$と独立な任意の直交行列$\boldsymbol{\Phi}$に対して、以下が成り立つとき、右直交不変と呼ばれる。

$$\boldsymbol{M\Phi} \sim \boldsymbol{M}$$

左直交不変性も同様に定義される。

ハール直交行列　右直交不変な直交行列$\boldsymbol{V}$をハール直交行列と呼ぶ。

注意：ハール直交行列は、左直交不変でもある。

仮定３　特異値分解$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathrm{T}}$において、$\boldsymbol{V}$は$\boldsymbol{U\Sigma}$と独立なハール直交行列である。$\boldsymbol{A}$の経験特異値分布は、大システム極限でサポートが有界な決定論的分布に概収束する。

表記のまとめ

∥と⊥

縦長行列$\boldsymbol{M}$の特異値分解$\boldsymbol{M}=\boldsymbol{\Phi}_{\boldsymbol{M}}\boldsymbol{\Sigma}_{\boldsymbol{M}}\boldsymbol{\Psi}_{\boldsymbol{M}}^{\mathrm{T}}$、$\boldsymbol{\Phi}_{\boldsymbol{M}}=(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel},\boldsymbol{\Phi}_{\boldsymbol{M}}^{\perp})$で、$\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}$と$\boldsymbol{\Phi}_{\boldsymbol{M}}^{\perp}$は、それぞれ非零特異値と零特異値に対応する左特異ベクトルから成る。

$\boldsymbol{M}$の列ベクトルが張る空間への射影行列

$$\boldsymbol{P}_{\boldsymbol{M}}^{\parallel}=\boldsymbol{M}\left(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\right)^{-1}\boldsymbol{M}^{\mathrm{T}}$$

特異値分解$\boldsymbol{M}=\boldsymbol{\Phi}_{\boldsymbol{M}}\boldsymbol{\Sigma}_{\boldsymbol{M}}\boldsymbol{\Psi}_{\boldsymbol{M}}^{\mathrm{T}}$を使うと、$\boldsymbol{P}_{\boldsymbol{M}}^{\parallel}=\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}\left(\boldsymbol{\Phi}_{\boldsymbol{M}}^{\parallel}\right)^{\mathrm{T}}$

$\boldsymbol{M}$の直交補空間への射影行列

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

主定理（モジュールA）

$\boldsymbol{B}_t = [\boldsymbol{b}_0, \dots, \boldsymbol{b}_{t-1}]$、$\boldsymbol{M}_t = [\boldsymbol{m}_0, \dots, \boldsymbol{m}_{t-1}]$、$\boldsymbol{H}_t = [\boldsymbol{h}_0, \dots, \boldsymbol{h}_{t-1}]$、$\boldsymbol{Q}_t = [\boldsymbol{q}_0, \dots, \boldsymbol{q}_{t-1}]$とする。大システム極限において、

(A1) $\{\boldsymbol{B}_\tau, \boldsymbol{M}_\tau, \boldsymbol{H}_\tau, \boldsymbol{Q}_\tau\}$に関する条件下で、

$$\boldsymbol{b}_\tau \sim \boldsymbol{B}_\tau \left(\boldsymbol{Q}_\tau^{\mathrm{T}} \boldsymbol{Q}_\tau\right)^{-1} \boldsymbol{Q}_\tau^{\mathrm{T}} \boldsymbol{q}_\tau + \boldsymbol{\Phi}_{[\boldsymbol{B}_\tau, \boldsymbol{M}_\tau]}^{\perp} \widetilde{\boldsymbol{V}} \left(\boldsymbol{\Phi}_{[\boldsymbol{Q}_\tau, \boldsymbol{H}_\tau]}^{\perp}\right)^{\mathrm{T}} \boldsymbol{q}_\tau + \boldsymbol{o}(1)$$

ただし、$\widetilde{\boldsymbol{V}}$は他の確率変数と独立なハール直交行列

(A2) $\{B_\tau\}$を平均0で共分散が$\mathbb{E}[B_\tau B_{\tau'}] - N^{-1} \boldsymbol{q}_\tau^{\mathrm{T}} \boldsymbol{q}_{\tau'} \xrightarrow{\text{a.s.}} 0$を満たすガウス確率変数とする。適切な関数$\phi: \mathbb{R}^{\tau+3} \to \mathbb{R}$に対して、

$$\langle \phi(\boldsymbol{b}_0, \dots, \boldsymbol{b}_\tau, \widetilde{\boldsymbol{w}}, \boldsymbol{\sigma}) \rangle \xrightarrow{\text{a.s.}} \mathbb{E}[\phi(B_0, \dots, B_\tau, W, \Sigma)], \quad W \sim \mathcal{N}(0, \sigma^2)$$

Σは、他の確率変数と独立で、$\boldsymbol{A}$の漸近特異値分布に従う。

(A3) 任意の$\tau' \in \{0, \dots, \tau\}$に対して、

$$\frac{1}{N} \boldsymbol{b}_{\tau'}^{\mathrm{T}} \boldsymbol{m}_\tau \xrightarrow{\text{a.s.}} 0.$$

OAMPの名称の由来

TOYOHASHI UNIVERSITY OF TECHNOLOGY

主定理（モジュールB）

(B1) $\{\boldsymbol{B}_{\tau+1}, \boldsymbol{M}_\tau, \boldsymbol{H}_\tau, \boldsymbol{Q}_{\tau+1}\}$に関する条件下で、

$$\boldsymbol{h}_\tau \sim \boldsymbol{H}_\tau\left(\boldsymbol{M}_\tau^{\mathrm{T}}\boldsymbol{M}_\tau\right)^{-1}\boldsymbol{M}_\tau^{\mathrm{T}}\boldsymbol{m}_\tau + \boldsymbol{\Phi}^{\perp}_{[\boldsymbol{H}_\tau, \boldsymbol{Q}_{\tau+1}]}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}^{\perp}_{[\boldsymbol{M}_\tau, \boldsymbol{B}_{\tau+1}]}\right)^{\mathrm{T}}\boldsymbol{m}_\tau + \boldsymbol{o}(1)$$

ただし、$\widetilde{\boldsymbol{V}}$は他の確率変数と独立なハール直交行列

(B2) $\{H_\tau\}$を平均0で共分散が$\mathbb{E}[H_\tau H_{\tau'}] - N^{-1}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{m}_{\tau'} \overset{\text{a.s.}}{\to} 0$を満たすガウス確率変数とする。適切な関数$\psi: \mathbb{R}^{\tau+2} \to \mathbb{R}$に対して、

$$\langle \psi(\boldsymbol{h}_0, \dots, \boldsymbol{h}_\tau, \boldsymbol{x}) \rangle \overset{\text{a.s.}}{\to} \mathbb{E}[\psi(H_0, \dots, H_\tau, X)]$$

Xは、他の確率変数と独立で、x_1と同じ分布に従う。

モジュールA出力の誤差が漸近的にガウス分布する。

(B3) 任意の$\tau' \in \{0, \dots, \tau\}$に対して、

$$\frac{1}{N}\boldsymbol{h}_{\tau'}^{\mathrm{T}}\boldsymbol{q}_{\tau+1} \overset{\text{a.s.}}{\to} 0.$$

OAMPの名称の由来

TOYOHASHI UNIVERSITY OF TECHNOLOGY

OAMP/VAMPの誤差モデル

モジュールB $\boldsymbol{h}_t = \boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t} - \boldsymbol{x}$と$\boldsymbol{q}_{t+1} = \boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} - \boldsymbol{x}$を定義すると、

$$\boldsymbol{q}_{t+1} = \frac{\boldsymbol{x}_{\mathrm{B},t+1} - \boldsymbol{x} - \eta_{\mathrm{B},t}\boldsymbol{h}_t}{1-\eta_{\mathrm{B},t}} = \psi_t(\boldsymbol{h}_t, \boldsymbol{x}) - \frac{\eta_{\mathrm{B},t}}{1-\eta_{\mathrm{B},t}}\boldsymbol{h}_t,$$

ただし、

$$\psi_t(\boldsymbol{h}_t, \boldsymbol{x}) = \frac{f_{\mathrm{B},t}\left(\boldsymbol{x} + \boldsymbol{h}_t, v_{\mathrm{A}\to\mathrm{B},t}\right) - \boldsymbol{x}}{1-\eta_{\mathrm{B},t}}.$$

モジュールA

$$\boldsymbol{h}_t = \frac{\boldsymbol{x}_{\mathrm{A},t} - \boldsymbol{x} - \eta_{\mathrm{A},t}\boldsymbol{q}_t}{1-\eta_{\mathrm{A},t}}.$$

線形フィルタ$f_{\mathrm{A},t}$の定義から

$$\begin{aligned}\boldsymbol{x}_{\mathrm{A},t} - \boldsymbol{x} &= \boldsymbol{q}_t + \boldsymbol{V}\boldsymbol{D}_t\boldsymbol{U}^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}) \\ &= \boldsymbol{q}_t + \boldsymbol{V}\boldsymbol{D}_t\boldsymbol{U}^{\mathrm{T}}\left(-\boldsymbol{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}\boldsymbol{q}_t + \boldsymbol{w}\right) = (1-\eta_{\mathrm{A},t})\boldsymbol{V}\phi_t,\end{aligned}$$

ただし、

$$[\phi_t(\boldsymbol{b}_t, \tilde{\boldsymbol{w}}, \boldsymbol{\sigma})]_n = \frac{\left\{1 - d_t\left(\sigma_n, v_{\mathrm{B}\to\mathrm{A},t}\right)\sigma_n\right\}b_{n,t} + d_t\left(\sigma_n, v_{\mathrm{B}\to\mathrm{A},t}\right)\tilde{w}_n}{1-\eta_{\mathrm{A},t}},$$

$$d_t\left(\sigma_n, v_{\mathrm{B}\to\mathrm{A},t}\right) = 0 \text{ for all } n > M.$$

したがって、

$$\boldsymbol{h}_t = \boldsymbol{V}\boldsymbol{m}_t, \qquad \boldsymbol{m}_t = \phi_t(\boldsymbol{b}_t, \tilde{\boldsymbol{w}}, \boldsymbol{\sigma}) - \frac{\eta_{\mathrm{A},t}}{1-\eta_{\mathrm{A},t}}\boldsymbol{b}_t.$$

OAMP/VAMPの設計

OAMP/VAMPの誤差モデルが一般化誤差モデルに含まれるように、オンサーガ補正項の係数$\eta_{\mathrm{A},t}$と$\eta_{\mathrm{B},t}$を設計する。

一般化誤差モデル

$$\boldsymbol{m}_t = \phi_t(\boldsymbol{b}_t, \tilde{\boldsymbol{w}}, \boldsymbol{\sigma}) - \langle \phi_t'(\boldsymbol{b}_t, \tilde{\boldsymbol{w}}, \boldsymbol{\sigma}) \rangle \boldsymbol{b}_t,$$

$$\boldsymbol{q}_{t+1} = \psi_t(\boldsymbol{h}_t, \boldsymbol{x}) - \langle \psi_t'(\boldsymbol{h}_t, \boldsymbol{x}) \rangle \boldsymbol{h}_t.$$

OAMP/VAMPの誤差モデル

$$\boldsymbol{m}_t = \phi_t(\boldsymbol{b}_t, \tilde{\boldsymbol{w}}, \boldsymbol{\sigma}) - \frac{\eta_{\mathrm{A},t}}{1-\eta_{\mathrm{A},t}} \boldsymbol{b}_t.$$

$$\boldsymbol{q}_{t+1} = \psi_t(\boldsymbol{h}_t, \boldsymbol{x}) - \frac{\eta_{\mathrm{B},t}}{1-\eta_{\mathrm{B},t}} \boldsymbol{h}_t.$$

設計結果　前ページのϕ_tとψ_tの定義から、初等的な計算によって

$$\eta_{\mathrm{A},t} = \left\langle f_{\mathrm{A},t}'\left(\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}, v_{\mathrm{B}\to\mathrm{A},t}; \boldsymbol{A}, \boldsymbol{y}\right)\right\rangle, \quad \eta_{\mathrm{B},t} = \left\langle f_{\mathrm{B},t}'\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}, v_{\mathrm{A}\to\mathrm{B},t}\right)\right\rangle.$$

ただし、微分は第一変数に関する偏微分を表す。

主定理の証明の流れ

数学的帰納法による証明

$\tau = 0$の場合の証明は省略する。

仮定

$\tau < t$に対する(A1)～(A3)、
$\tau < t$に対する(B1)～(B3)

証明

$\tau = t$に対する(A1)～(A3)

- (B3)→(A1)
- (A1)→(A2)
- (A2)→(A3)

説明省略

仮定

$\tau \leq t$に対する(A1)～(A3)、
$\tau < t$に対する(B1)～(B3)

証明

$\tau = t$に対する(B1)～(B3)

- (A3)→(B1)
- (B1)→(B2)
- (B2)→(B3)

順に説明

Bolthausenの方法[2,3,4]

性質(B1)の証明方針

1. 反復t以前までのすべてのメッセージの履歴を条件として与えたときに、ハール直交行列$\boldsymbol{V}$の条件付き分布を評価する。

2. 上記の条件付き分布を使って、反復tでのメッセージの分布を計算する。

[3] 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.

[4] 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.

補題1[2, 3]

$\boldsymbol{X} \in \mathbb{R}^{N\times t}$を観測行列として、$\boldsymbol{X}$と独立な$N \times N$ハール直交行列$\boldsymbol{V}$をノイズなしの理想的な観測をしたときの出力行列を $\boldsymbol{Y} \in \mathbb{R}^{N\times t}$ とする。

$$\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{X}$$

主張　$t < N$として$\boldsymbol{X}$がフルランクならば、$\widetilde{\boldsymbol{V}}$を$\boldsymbol{X}$、$\boldsymbol{Y}$と独立なハール直交行列として、$\boldsymbol{X}$と$\boldsymbol{Y}$に関する条件下で、

$$\begin{aligned}\boldsymbol{V} &\sim \boldsymbol{Y}\left(\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Phi}_{\boldsymbol{Y}}^{\perp}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}_{\boldsymbol{X}}^{\perp}\right)^{\mathrm{T}} \\ &= \boldsymbol{Y}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Phi}_{\boldsymbol{Y}}^{\perp}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}_{\boldsymbol{X}}^{\perp}\right)^{\mathrm{T}}.\end{aligned}$$

証明　省略

確認　$\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{X}$から$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$が成り立つ。$\left(\boldsymbol{\Phi}_{\boldsymbol{X}}^{\perp}\right)^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{O}$なので、

$$\left[\boldsymbol{Y}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Phi}_{\boldsymbol{Y}}^{\perp}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}_{\boldsymbol{X}}^{\perp}\right)^{\mathrm{T}}\right]\boldsymbol{X} = \boldsymbol{Y}$$

制約式$\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{X}$を満たす。

性質(B1)の証明

$\boldsymbol{h}_t = \boldsymbol{V}\boldsymbol{m}_t$の更新直前で、$\boldsymbol{B}_{t+1}$、$\boldsymbol{M}_t$、$\boldsymbol{H}_t$、$\boldsymbol{Q}_{t+1}$は計算済みである。

制約式$[\boldsymbol{H}_t, \boldsymbol{Q}_{t+1}] = \boldsymbol{V}[\boldsymbol{M}_t, \boldsymbol{B}_{t+1}]$に対して補題1を適用すると、

$$\boldsymbol{V} \sim [\boldsymbol{H}_t, \boldsymbol{Q}_{t+1}]\left(\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}} \\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}[\boldsymbol{M}_t, \boldsymbol{B}_{t+1}]\right)^{-1}\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}} \\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}$$

$$+\boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}^{\perp}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}.$$

$\tau \leq t$に関する帰納法の仮定(A3)を使うと、

$$\boldsymbol{h}_t \sim \boldsymbol{H}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t + \boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}^{\perp}\widetilde{\boldsymbol{V}}\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t + \boldsymbol{o}(1)$$

$$\because \left(\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}} \\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}[\boldsymbol{M}_t, \boldsymbol{B}_{t+1}]\right)^{-1}\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}} \\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}\boldsymbol{m}_t \approx \begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}_{t+1}^{\mathrm{T}}\boldsymbol{B}_{t+1}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t \\ \boldsymbol{0}\end{bmatrix}$$

TOYOHASHI
UNIVERSITY OF TECHNOLOGY

補題2

$\boldsymbol{V}$をハール直交行列とする。$\boldsymbol{V}$と独立な任意のベクトル$\boldsymbol{a}$に対し、

$$\boldsymbol{Va} \sim \frac{\|\boldsymbol{a}\|}{\|\boldsymbol{z}\|}\boldsymbol{z}, \quad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$$ は$\boldsymbol{a}$と独立な標準ガウス確率ベクトル

$\boldsymbol{Va}$は、振幅を除いてほぼガウス確率ベクトルとみなせる。

証明　$\boldsymbol{Va}$を特異値分解すると、$\boldsymbol{Va} = \|\boldsymbol{a}\|\boldsymbol{u}$を得る。（$\|\boldsymbol{u}\| = 1$）

$\boldsymbol{Va}$は左直交不変なので、$\boldsymbol{u}$も左直交不変である。

同様に、$\boldsymbol{z}$を特異値分解すると、$\boldsymbol{z}$も左直交不変なので、
$\boldsymbol{v}$を直交不変な単位ベクトルとして、$\boldsymbol{z} = \|\boldsymbol{z}\|\boldsymbol{v}$を得る。

同一分布性$\boldsymbol{u} \sim \boldsymbol{v}$を使うと、

$$\boldsymbol{Va} = \|\boldsymbol{a}\|\boldsymbol{u} \sim \|\boldsymbol{a}\|\boldsymbol{v} = \frac{\|\boldsymbol{a}\|}{\|\boldsymbol{z}\|}\boldsymbol{z} \qquad \blacksquare$$

説明の簡単化

近似　$\tilde{\boldsymbol{z}}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$として、性質(B1)に補題2を適用する。

$$\boldsymbol{h}_t \sim \boldsymbol{H}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t + \frac{\left\|\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t\right\|}{\|\tilde{\boldsymbol{z}}_t\|}\boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}^{\perp}\tilde{\boldsymbol{z}}_t$$

$$\approx \boldsymbol{H}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t + \frac{\left\|\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t\right\|}{\sqrt{N}}\boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}^{\perp}\tilde{\boldsymbol{z}}_t$$

$$\approx \boldsymbol{H}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t + \frac{\left\|\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t\right\|}{\sqrt{N}}\boldsymbol{z}_t, \quad \boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$$

最後の近似の妥当性

- $\boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}^{\perp}$は$\boldsymbol{\Phi}_{[\boldsymbol{H}_t,\boldsymbol{Q}_{t+1}]}$から$2t+1$本の列を取り除いた行列
- 標準ガウス確率ベクトルは左直交不変

(B2)の厳密な証明　文献[3]のLemma 1を使えば良い。

性質(B2)の証明

$\boldsymbol{\alpha}_t = \left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t$、$v_t = N^{-1}\left\|\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t\right\|^2$とおくと、

$$\boldsymbol{h}_\tau \sim \sum_{\tau'=0}^{\tau-1}[\boldsymbol{\alpha}_\tau]_{\tau'}\boldsymbol{h}_{\tau'} + \sqrt{v_\tau}\boldsymbol{z}_\tau \quad \text{for all } \tau \in \{0,\dots,t\}$$

ただし、$\{\boldsymbol{z}_\tau \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I})\}_{\tau=0}^t$は独立な標準ガウス確率ベクトル。

定義から$\{\boldsymbol{h}_\tau\}_{\tau=0}^t$は平均$\boldsymbol{0}$で$\mathbb{E}[\boldsymbol{h}_{\tau'}\boldsymbol{h}_\tau^{\mathrm{T}}] = \Sigma_{\tau,\tau'}\boldsymbol{I}$のガウス確率ベクトル。

したがって、仮定1と大数の強法則から

$$\langle\psi(\boldsymbol{h}_0,\dots,\boldsymbol{h}_t,\boldsymbol{x})\rangle \overset{\text{a.s.}}{\to} \mathbb{E}[\psi(H_0,\dots,H_t,X)]$$

ただし、$\{H_\tau\}_{\tau=0}^t$は平均0で$\mathbb{E}[H_\tau H_{\tau'}] = \Sigma_{\tau,\tau'}$のガウス確率変数。

よって、以下を帰納法で証明すればよい。

$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_t - \frac{1}{N}\boldsymbol{m}_\tau^{\mathrm{T}}\boldsymbol{m}_t \overset{\text{a.s.}}{\to} 0 \quad \text{for all } \tau \in \{0,\dots,t\}$$

性質(B2)の証明

$\tau < t$のとき　$\boldsymbol{h}_t \sim \boldsymbol{H}_t \boldsymbol{\alpha}_t + \sqrt{v_t} \boldsymbol{z}_t$を使って、

$$\frac{1}{N} \boldsymbol{h}_\tau^{\mathrm{T}} \boldsymbol{h}_t \sim \frac{1}{N} \boldsymbol{h}_\tau^{\mathrm{T}} \boldsymbol{H}_t \boldsymbol{\alpha}_t + \frac{1}{N} \sqrt{v_t} \boldsymbol{h}_\tau^{\mathrm{T}} \boldsymbol{z}_t.$$

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$N^{-1} \boldsymbol{h}_\tau^{\mathrm{T}} \boldsymbol{H}_t - N^{-1} \boldsymbol{m}_\tau^{\mathrm{T}} \boldsymbol{M}_t \overset{\text{a.s.}}{\to} 0$と$\boldsymbol{\alpha}_t$の定義を使うと、

$$\frac{1}{N} \boldsymbol{h}_\tau^{\mathrm{T}} \boldsymbol{h}_t \overset{\text{a.s.}}{\to} \frac{1}{N} \boldsymbol{m}_\tau^{\mathrm{T}} \boldsymbol{M}_t \left(\boldsymbol{M}_t^{\mathrm{T}} \boldsymbol{M}_t \right)^{-1} \boldsymbol{M}_t^{\mathrm{T}} \boldsymbol{m}_t = \frac{1}{M} \boldsymbol{m}_\tau^{\mathrm{T}} \boldsymbol{m}_t.$$

$$\because \boldsymbol{m}_\tau^{\mathrm{T}} \boldsymbol{M}_t \left(\boldsymbol{M}_t^{\mathrm{T}} \boldsymbol{M}_t \right)^{-1} = \boldsymbol{e}_\tau^{\mathrm{T}} \boldsymbol{M}_t^{\mathrm{T}} \boldsymbol{M}_t \left(\boldsymbol{M}_t^{\mathrm{T}} \boldsymbol{M}_t \right)^{-1} = \boldsymbol{e}_\tau^{\mathrm{T}}$$

$\tau = t$のとき　$\boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$と大数の強法則から、

$$\frac{1}{N} \boldsymbol{h}_t^{\mathrm{T}} \boldsymbol{h}_t \overset{\text{a.s.}}{\to} \frac{1}{N} \boldsymbol{\alpha}_t^{\mathrm{T}} \boldsymbol{H}_t^{\mathrm{T}} \boldsymbol{H}_t \boldsymbol{\alpha}_t + v_t.$$

性質(B2)の証明

帰納法の仮定$N^{-1}\boldsymbol{H}_t^{\mathrm{T}}\boldsymbol{H}_t - N^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t \overset{\text{a.s.}}{\to} 0$と$v_t$の定義を使うと、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_t \overset{\text{a.s.}}{\to} \frac{1}{N}\boldsymbol{\alpha}_t^{\mathrm{T}}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\boldsymbol{\alpha}_t + \frac{1}{N}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}\boldsymbol{m}_t$$

$\boldsymbol{P}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp} = \boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\left(\boldsymbol{\Phi}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\right)^{\mathrm{T}}$として、

$$\frac{1}{N}\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{h}_t \overset{\text{a.s.}}{\to} \frac{1}{N}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{M}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t + \frac{1}{N}\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{P}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp}\boldsymbol{m}_t = \frac{\boldsymbol{m}_t^{\mathrm{T}}\boldsymbol{m}_t}{N}$$

$$\because\ \boldsymbol{\alpha}_t = \left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t$$

最後の等号は、$\boldsymbol{P}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\perp} = \boldsymbol{I} - \boldsymbol{P}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\parallel}$および以下から従う。

$$\boldsymbol{P}_{[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]}^{\parallel}\boldsymbol{m}_t = [\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]\left(\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}}\\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}[\boldsymbol{M}_t,\boldsymbol{B}_{t+1}]\right)^{-1}\begin{bmatrix}\boldsymbol{M}_t^{\mathrm{T}}\\ \boldsymbol{B}_{t+1}^{\mathrm{T}}\end{bmatrix}\boldsymbol{m}_t$$

$$\overset{\text{a.s.}}{\to} \boldsymbol{M}_t\left(\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^{\mathrm{T}}\boldsymbol{m}_t. \quad \because\ \text{性質(B1)の証明参照} \quad \blacksquare$$

Steinの補題

$(Z_1, Z_2) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \rho \\ \rho & \sigma_2^2\end{bmatrix}\right)$として、

$$\mathbb{E}[Z_1 f(Z_i)] = \mathbb{E}[Z_1 Z_i]\mathbb{E}[f'(Z_i)] \ \text{for}\ i = 1, 2$$

証明　　最初に、$i = 1$の場合を証明する。

$$p(z_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{z_1^2}{2\sigma_1^2}},$$

$$\mathbb{E}[Z_1 f(Z_1)] = \int z_1 f(z_1) p(z_1) dz_1 = -\sigma_1^2 \int f(z_1) p'(z_1) dz_1$$

$$= \sigma_1^2 \int f'(z_1) p(z_1) dz_1 = \sigma_1^2 \mathbb{E}[f'(Z_1)] = \mathbb{E}[Z_1^2]\mathbb{E}[f'(Z_1)].$$

$$\because p'(z_1) = -\frac{z_1}{\sigma_1^2} p(z_1)$$

Steinの補題の証明

次に、$i=2$の場合を証明する。

Z_1と独立な$W \sim \mathcal{N}\left(0, \sigma_2^2 - \frac{\rho^2}{\sigma_1^2}\right)$を使って、$Z_2 = \frac{\rho}{\sigma_1^2} Z_1 + W$

$$\because \ \mathbb{E}[Z_1 Z_2] = \frac{\rho}{\sigma_1^2}\mathbb{E}[Z_1^2] = \rho, \quad \mathbb{E}[Z_2^2] = \frac{\rho^2}{\sigma_1^4}\mathbb{E}[Z_1^2] + \sigma_2^2 - \frac{\rho^2}{\sigma_1^2} = \sigma_2^2.$$

Z_1とWは独立なので、

$$\mathbb{E}[Z_1 f(Z_2)] = \mathbb{E}_W\left[\mathbb{E}_{Z_1}\left[Z_1 f\left(\frac{\rho}{\sigma_1^2}Z_1 + W\right)\right]\right]$$

$$= \frac{\rho}{\sigma_1^2}\mathbb{E}_W\left[\mathbb{E}[Z_1^2]\mathbb{E}_{Z_1}\left[f'\left(\frac{\rho}{\sigma_1^2}Z_1 + W\right)\right]\right] \quad \because i = 1\text{の場合}$$

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

性質(B3)の証明

$$\frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{q}_{t+1} = \frac{1}{N}\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{\psi}_t(\boldsymbol{h}_t, \boldsymbol{x}) - \frac{1}{N}\langle\boldsymbol{\psi}_t'(\boldsymbol{h}_t, \boldsymbol{x})\rangle\boldsymbol{h}_\tau^{\mathrm{T}}\boldsymbol{h}_t$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}[H_\tau\psi_t(H_t, X)] - \mathbb{E}[\psi_t'(H_t, X)]\mathbb{E}[H_\tau H_t]$$

$$= \mathbb{E}[H_\tau H_t]\mathbb{E}[\psi_t'(H_t, X)] - \mathbb{E}[\psi_t'(H_t, X)]\mathbb{E}[H_\tau H_t] = 0.$$

∵ 性質(B2)

∵ Steinの補題 ■