メッセージ伝播法の入門から最先端まで
第4回講義資料
状態発展法：発見的アプローチ

九州大学
令和元年10月10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

AMPと状態発展方程式

観測モデル

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

AMP[3-1]

$$\boldsymbol{z}^t = \boldsymbol{y} - \boldsymbol{A}\widehat{\boldsymbol{x}}^t + \frac{\langle \eta'_{t-1}(\widehat{\boldsymbol{x}}^{t-1} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^{t-1}) \rangle}{\delta} \boldsymbol{z}^{t-1},$$

$$\widehat{\boldsymbol{x}}^{t+1} = \eta_t(\widehat{\boldsymbol{x}}^t + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^t).$$

状態発展方程式[3-4]

$$v^t = \sigma^2 + \frac{1}{\delta}\Psi_{t-1}(v^{t-1}), \qquad v^{-1} = 1,$$

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

$$\lim_{M=\delta N \to \infty} \frac{1}{N} \mathbb{E}[\|\boldsymbol{x} - \widehat{\boldsymbol{x}}^{t+1}\|^2] = \Psi_t(v^t).$$

発見的導出の手順[3-4]

1. 観測モデルを反復回数tごとに独立な仮想的モデルに置き換える。

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

$\{\boldsymbol{A}^t\}$と$\{\boldsymbol{w}^t\}$はそれぞれ$\boldsymbol{A}$と$\boldsymbol{w}$に従うi.i.d.系列

2. AMPからオンサーガ項を取り除く。

$$\boldsymbol{z}^t = \boldsymbol{y}^t - \boldsymbol{A}^t \hat{\boldsymbol{x}}^t,$$

$$\hat{\boldsymbol{x}}^{t+1} = \eta_t\left(\hat{\boldsymbol{x}}^t + (\boldsymbol{A}^t)^{\mathrm{T}} \boldsymbol{z}^t\right).$$

3. 大システム極限で、性能解析を行う。

大数の強法則[4-1]

$\{X_n\}_{n=1}^N$を二次モーメントが有界な確率変数列とし、
$S_N = \sum_{n=1}^N X_n$を定義する。

$$\sum_{N=1}^{\infty} \frac{\sqrt{\mathbb{V}[S_N]}}{N^2} < \infty.$$

上記を満たすならば、以下の大数の強法則が従う。

$$T_N = \frac{S_N - \mathbb{E}[S_N]}{N} \to 0 \ \text{ almost surely as } N \to \infty.$$

注意

無相関な確率変数列は条件を満たす。

[4-1] R. Lyons, “Strong laws of large numbers for weakly correlated random variables,” *Michigan Math. J.*, vol. 35, no. 3, pp. 353–359, 1988.

中心極限定理[4-2]

$\{X_n\}$を平均μ_n分散σ_n^2の独立な確率変数列とし、
$s_N^2 = \sum_{n=1}^N \sigma_n^2$を定義する。

$$\lim_{N\to\infty} \frac{1}{s_N^2} \sum_{n=1}^{N} \mathbb{E}[(X_n - \mu_n)^2 1(|X_n - \mu_n| > \epsilon s_N)] = 0 \ \text{ for any } \epsilon > 0.$$

上記の条件を満たすならば、以下の中心極限定理が従う。

$$\frac{1}{s_N} \sum_{n=1}^{N} (X_n - \mu_n) \to \mathcal{N}(0,1) \text{ in distribution as } N \to \infty.$$

注意

ある$\alpha > 0$に対して、X_nの$2 + \alpha$次モーメントが存在すればよい。

K体i.i.d.性の定義

ある自然数$K(\leq N)$に対して、N次元確率ベクトル$\boldsymbol{v} \in \mathbb{R}^N$が以下の性質を満たすとき、$\boldsymbol{v}$を平均$\mu$分散$\sigma^2$の$K$体i.i.d.な確率ベクトルと呼ぶ。

- $\boldsymbol{v}$から任意のK個の異なる要素を取り出してできるベクトルはi.i.d.要素を持ち、各要素は平均μ分散σ^2である。

注意

さらに、各要素がガウス分布に従う場合、$\boldsymbol{v}$はK体i.i.d.なガウス確率ベクトルと呼ばれる。

中心極限定理に関する注意

大数の強法則は無相関等の弱い仮定で主張できるが、中心極限定理を主張するためには確率変数列の独立性が必要である。

中心極限定理の反例[4-2]

$N-1$体i.i.d.標準ガウス確率変数列$\{X_n\}_{n=1}^{N}$で、和$Y=N^{-1/2}\sum_{n=1}^{N}X_n$の分布が$\mathcal{N}(0,1)$に収束しない例が存在する。

[4-2] K. Takeuchi, "A family of counterexamples to the central limit theorem based on binary linear codes," *IEICE Trans. Fundamentals*., vol. E102-A, no. 5, pp. 738-740, May 2019.

補題4.1

$\boldsymbol{w} \in \mathbb{R}^M$を任意の決定論的なベクトルとする。

$\boldsymbol{A} \in \mathbb{R}^{M \times N}$を平均0分散$1/M$のi.i.d.要素を持つ行列とする。

$M^{-1}\|\boldsymbol{w}\|^2$は極限$M \to \infty$で$\sigma^2 > 0$に確率収束する。

ベクトル$\boldsymbol{v} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}$は以下を満たす。

- $\boldsymbol{v}$は大システム極限で平均$\boldsymbol{0}$共分散行列$\sigma^2\boldsymbol{I}_N$のガウス確率ベクトルに分布収束する。

注意

本資料では、例えば$M^{-1}\|\boldsymbol{w}\|^2$が収束する等のような技術的な仮定に関する議論を省略する。

補題4.1の証明

$\boldsymbol{w}$が与えられたときに$\boldsymbol{v}$が平均0のi.i.d.要素を持つことは、$\boldsymbol{A}$の列ベクトルの独立性から従う。

$\boldsymbol{v}$の最初の要素を評価すると、

$$[\boldsymbol{v}]_1 = \sum_{m=1}^{M} w_m A_{m1} .$$

中心極限定理より、$[\boldsymbol{v}]_1$の分布は大システム極限で$\mathcal{N}(0, \sigma^2)$に収束する。 ■

TOYOHASHI UNIVERSITY OF TECHNOLOGY

補題4.2

$\boldsymbol{u} \in \mathbb{R}^N$を任意の決定論的なベクトルとする。

極限$M^{-1}\|\boldsymbol{u}\|^2 \to a$が存在する。

$\boldsymbol{A} \in \mathbb{R}^{M \times N}$を平均$0$分散$1/M$のi.i.d.要素を持つ行列とする。

ベクトル$\boldsymbol{v} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} - \boldsymbol{I}_N\right)\boldsymbol{u}$は以下を満たす。

- 任意の有限な$K \in \mathbb{N}$に対して、$\boldsymbol{v}$は大システム極限で平均0分散aのK体i.i.d.なガウス確率ベクトルに分布収束する。

補題4.2の証明

一般性を失うことなく、$\boldsymbol{v}$の最初のK個の要素からなるベクトル$\boldsymbol{v}_1 \in \mathbb{R}^K$に注目する。

$$\boldsymbol{v}_1 = \boldsymbol{A}_1^{\mathrm{T}}(\boldsymbol{A}_1, \boldsymbol{A}_2)\boldsymbol{u} - \boldsymbol{u}_1 = \left(\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1 - \boldsymbol{I}_K\right)\boldsymbol{u}_1 + \boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{u}_2,$$

$$\boldsymbol{A} = (\boldsymbol{A}_1, \boldsymbol{A}_2) \in \mathbb{R}^{M\times K} \times \mathbb{R}^{M\times(N-K)}. \quad \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix} \in \mathbb{R}^K \times \mathbb{R}^{N-K}.$$

大数の強法則から、$M \to \infty$において$\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1$は$\boldsymbol{I}_K$に概収束するため、右辺第一項も$\boldsymbol{0}$に概収束する。

第二項はi.i.d.要素を持つベクトルなので、$\boldsymbol{v}_1$も漸近的にi.i.d.要素を持つベクトルである。

それゆえ、第二項の一番目の要素のガウス性を示せばよい。

TOYOHASHI UNIVERSITY OF TECHNOLOGY

補題4.2の証明

$$
\left[\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{u}_2\right]_1 = \frac{1}{\sqrt{M}}\sum_{n=K+1}^{N} u_n X_n, \qquad \frac{X_n}{\sqrt{M}} = \left[\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\right]_{1n} = \sum_{m=1}^{M} A_{m1}A_{mn}.
$$

$\{A_{m1}\}$の条件の下で、$\{X_n\}$は平均0のi.i.d.確率変数列である。

$$
\mathbb{E}[X_n^2|\{A_{m1}\}] = M\sum_{m=1}^{M} A_{m1}^2\mathbb{E}[A_{mn}^2] = \sum_{m=1}^{M} A_{m1}^2 \to 1.
$$

特に上記の性質から、$\{X_n\}$はi.i.d.標準確率変数列である。

$$
\mathbb{V}\left[\frac{1}{\sqrt{M}}\sum_{n=K+1}^{N} u_n X_n\right] = \frac{1}{M}\sum_{n=K+1}^{N} u_n^2 \to a.
$$

中心極限定理より、$\left[\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{u}_2\right]_1$の漸近ガウス性を得る。 ■

TOYOHASHI UNIVERSITY OF TECHNOLOGY

状態発展方程式の導出

大システム極限で$N^{-1}\|\boldsymbol{x}-\hat{\boldsymbol{x}}^t\|^2 \to \Psi_{t-1}(v^{t-1})$を仮定して、$N^{-1}\|\boldsymbol{x}-\hat{\boldsymbol{x}}^{t+1}\|^2 \to \Psi_t(v^t)$を示す。

観測モデルから、 $\boldsymbol{z}^t = \boldsymbol{y}^t - \boldsymbol{A}^t\hat{\boldsymbol{x}}^t = \boldsymbol{A}^t(\boldsymbol{x}-\hat{\boldsymbol{x}}^t) + \boldsymbol{w}^t.$

閾値関数η_tの入力は、

$$\hat{\boldsymbol{x}}^t + (\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{z}^t = \boldsymbol{x} + \{(\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{A}^t - \boldsymbol{I}\}(\boldsymbol{x}-\hat{\boldsymbol{x}}^t) + (\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{w}^t.$$

任意の$K \in \mathbb{N}$に対して、補題4.1と補題4.2から、右辺の第二項と第三項の和は、平均0分散v^tのK体i.i.d.なガウス確率ベクトルに分布収束する。

$$v^t = \frac{1}{\delta}\frac{1}{N}\|\boldsymbol{x}-\hat{\boldsymbol{x}}^t\|^2 + \frac{1}{M}\|\boldsymbol{w}\|^2 \to \sigma^2 + \frac{1}{\delta}\Psi_{t-1}(v^{t-1}).$$

状態発展方程式の導出

$$\boldsymbol{\xi} = \{(\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{A}^t - \boldsymbol{I}\}(\boldsymbol{x} - \hat{\boldsymbol{x}}^t) + (\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{w}^t.$$

前ページの結果と大数の強法則とから、

$$\frac{1}{N}\|\boldsymbol{x} - \hat{\boldsymbol{x}}^{t+1}\|^2 = \frac{1}{N}\left\|\boldsymbol{x} - \eta_t\left(\hat{\boldsymbol{x}}^t + (\boldsymbol{A}^t)^{\mathrm{T}}\boldsymbol{z}^t\right)\right\|^2$$

$$= \frac{1}{N}\sum_{n=1}^{N}\{x_n - \eta_t(x_n + \xi_n)\}^2$$

$$\to \mathbb{E}\left[\left\{x_1 - \eta_t\left(x_1 + \sqrt{v^t}\omega_1\right)\right\}^2\right] = \Psi_t(v^t). \quad \blacksquare$$