

# メッセージ伝播法入門

## 第3回講義資料

### 一般化近似的メッセージ伝播法

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

# ベイズ最適AMP

更新式

$$\mathbf{z}^t = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}^t + \frac{\langle \eta'(\hat{\mathbf{x}}^{t-1} + \mathbf{A}^T \mathbf{z}^{t-1}; v^{t-1}) \rangle}{\delta} \mathbf{z}^{t-1},$$

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

$$\hat{\mathbf{x}}^{t+1} = \eta(\hat{\mathbf{x}}^t + \mathbf{A}^T \mathbf{z}^t; v^t).$$

AWGN観測モデル  $\bar{x}_n = x_n + v^{1/2} \omega_n, \quad \omega_n \sim \mathcal{N}(0, 1).$

事後平均  $\eta(\bar{x}_n; v) = \mathbb{E}[x_n | \bar{x}_n, v].$

最小平均二乗誤差 (MMSE)  $\text{MMSE}(v) = \mathbb{E}[\{x_n - \eta(\bar{x}_n; v)\}^2].$

信号の事前分布が既知である必要がある。

# AMP[3-1]

更新式

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

$$\hat{\mathbf{x}}^{t+1} = \eta_t(\hat{\mathbf{x}}^t + A^T \mathbf{z}^t).$$

初期値

$$\mathbf{x}^0 = \mathbf{0}, \quad \mathbf{z}^{-1} = \mathbf{0}.$$

Denoiserを関数列 $\{\eta_t\}$ に一般化

[3-1] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

## オンサーガ(Onsager)項

### オンサーガ項

大システム極限において、誤差の漸近ガウス性を成立させるための補正項

$$\mathbf{z}^t = \mathbf{y} - A\hat{\mathbf{x}}^t + \frac{\xi_{t-1}}{\delta} \mathbf{z}^{t-1}, \quad \xi_{t-1} = \langle \eta'_{t-1}(\hat{\mathbf{x}}^{t-1} + A^T \mathbf{z}^{t-1}) \rangle$$

### L. Onsager (1903—1976)

オンサーガ相反関係(不可逆過程の熱力学の基礎)を発見した貢献により、1968年にノーベル化学賞を受賞

### 名前の由来

項の名前は、液体論に関わるオンサーガの1936年の論文で提案された方法論に由来する。

物理学者は、彼の方法論に従って導出される反跳場(reaction field)を  
**オンサーガ反跳場**と呼ぶ。

# オンサーガのキャビティ(Cavity)アプローチ

## キャビティ場(Cavity field)

$n$ 番目の信号を取り除いた $M \times (N - 1)$ キャビティ観測システムにおいて、  
干渉除去後に残っている干渉雑音

AWGNとみなせることを仮定する。

## 反跳場(Reaction field)

キャビティ観測システムに $n$ 番目の信号を追加したときに、  
追加した信号がキャビティ場に与える無視できない影響

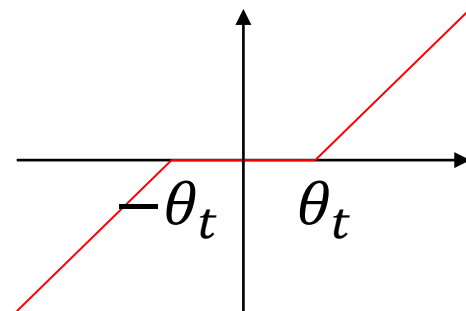
$[A^T(\delta^{-1}\xi_{t-1}z^{t-1})]_n$ のことだと思えばよい。

真の残留干渉 $[\hat{x}^t + A^T(y - A\hat{x}^t)]_n - x_n$ に反跳場による影響の補正を  
加えた $[\hat{x}^t + A^Tz^t]_n - x_n$ は、 $x_n$ と独立なキャビティ場とみなせるはず。

# Denoiserの例

## Soft thresholding (ST)

$$\eta_t(x) = \eta_{\text{ST}}(x; \theta_t) = \begin{cases} x - \theta_t & \text{for } x > \theta_t \\ 0 & \text{for } |x| \leq \theta_t \\ x + \theta_t & \text{for } x < -\theta_t. \end{cases}$$



## ミニマックス最適性[3-2]

任意の  $p, q > 0$  に対して、

$$\lim_{a \downarrow 0} \frac{r_{\text{ST}}(a)}{r_{\text{opt}}(a)} = 1.$$

$$r_{\text{ST}}(a) = \inf_{\theta} \sup_{p(x_n): (\mathbb{E}[|x_n|^p])^{1/p} \leq a} \mathbb{E} \left[ |x_n - \eta_{\text{ST}}(x_n + \sqrt{v}\omega_n; \theta)|^q \right],$$

$$r_{\text{opt}}(a) = \inf_{\eta} \sup_{p(x_n): (\mathbb{E}[|x_n|^p])^{1/p} \leq a} \mathbb{E} \left[ |x_n - \eta(x_n + \sqrt{v}\omega_n)|^q \right].$$

[3-2] D. L. Donoho and I. M. Johnstone, "Minimax risk over  $l_p$ -balls for  $l_q$ -error," *Probab. Theory Relat. Fields*, vol. 99, no. 2, pp. 277-303, Jun. 1994.

# ミニマックス最適性の意義

## モーメント制約

$$p(x_n) = (1 - \rho)\delta(x_n) + \rho\pi(x_n), \quad \int |x|^p \pi(x) dx = M_{p,\rho} < \infty.$$

$\rho < a^p / M_{p,\rho}$ を満たす分布は、モーメント制約を満たす。

$M_{p,\rho}$ が $\rho$ に依存しない信号分布では、信号密度 $\rho$ が0に収束する。

## ミニマックス最適性

- モーメント制約を満たす分布の中で**最悪な信号分布**を想定
- パラメータ $\theta$ を**最適値**に設定
- ST関数は $L_q$ 誤差を**最小化**

最適解の唯一性は主張していない。

# 状態発展法[3-3]

## 状態発展方程式

$$\bar{v}^t = \sigma^2 + \frac{1}{\delta} \Psi(\bar{v}^{t-1}, \lambda), \quad \bar{v}^0 = 1.$$

$$\Psi(v, \lambda) = \mathbb{E}[\{x_1 - \eta_{\text{ST}}(x_1 + \sqrt{v}\omega_1; \lambda\sqrt{v})\}^2].$$

閾値を $\theta_t = \lambda\sqrt{\bar{v}^t}$ とおいた。

## 誤差の漸近評価

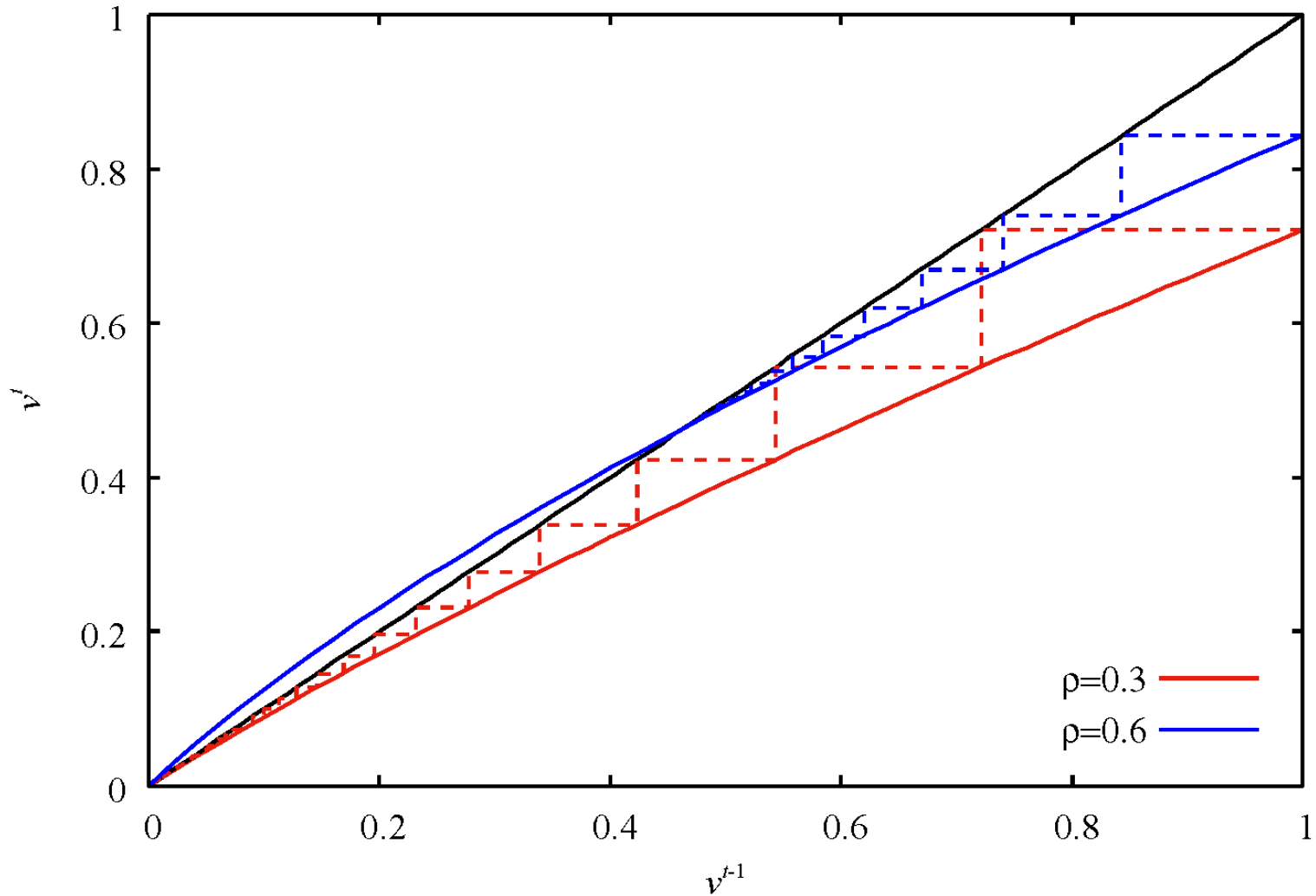
$$\lim_{M=\delta N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\|\mathbf{x} - \hat{\mathbf{x}}^t\|^2] = \Psi(\bar{v}^{t-1}, \lambda).$$

状態発展法の解説は、第4～6回講義資料を参照

[3-3] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.



# 数値的評価



信号密度 $\rho$ のBG事前分布、 $\sigma^2 = 0$ 、 $\delta = 0.65$ 、最適な $\lambda$

# 一般化線形観測モデル

## 観測モデル

$$\mathbf{y} = f(\mathbf{z}, \mathbf{w}), \quad \mathbf{z} = \mathbf{A}\mathbf{x}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$$

## 例1: AWGN観測

$$f(\mathbf{z}, \mathbf{w}) = \mathbf{z} + \mathbf{w}$$

## 例2: 非理想的増幅器

$$f(\mathbf{z}, \mathbf{w}) = \begin{cases} \theta & \text{for } \mathbf{z} + \mathbf{w} > \theta \\ \mathbf{z} + \mathbf{w} & \text{for } |\mathbf{z} + \mathbf{w}| \leq \theta \\ -\theta & \text{for } \mathbf{z} + \mathbf{w} < -\theta \end{cases}$$

## 例3: 量子化

$$f(\mathbf{z}, \mathbf{w}) = \begin{cases} 1 & \text{for } \mathbf{z} + \mathbf{w} \geq 0 \\ -1 & \text{for } \mathbf{z} + \mathbf{w} < 0 \end{cases}$$

# 一般化AMP (GAMP) [3-4]

更新式

$$\mathbf{z}_t = \mathbf{A}\hat{\mathbf{x}}_t + \frac{\xi_{\text{in},t-1}}{\delta\xi_{\text{out},t-1}}\hat{\mathbf{z}}_{t-1}$$

$$\hat{\mathbf{z}}_t = \eta_{\text{out},t}(\mathbf{z}_t, \mathbf{y}), \quad \xi_{\text{out},t} = \left\langle \frac{\partial}{\partial \mathbf{z}_t} \eta_{\text{out},t}(\mathbf{z}_t, \mathbf{y}) \right\rangle$$

$$\mathbf{x}_t = \hat{\mathbf{x}}_t - \frac{1}{\xi_{\text{out},t}} \mathbf{A}^T \hat{\mathbf{z}}_t,$$

$$\hat{\mathbf{x}}_{t+1} = \eta_{\text{in},t}(\mathbf{x}_t), \quad \xi_{\text{in},t} = \langle \eta'_{\text{in},t}(\mathbf{x}_t) \rangle$$

初期値

$$\hat{\mathbf{x}}_0 = \mathbf{0}, \quad \hat{\mathbf{z}}_{-1} = \mathbf{0}$$

[3-4] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. 2011 IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, Aug. 2011, pp. 2168–2172.

# 状態発展法[3-4]

状態発展法で定まるパラメータ $\bar{\zeta}_{out,t}$ 及び $\bar{\zeta}_t$ 、 $\sigma_t^2$ を用いて

$\eta_{in,t}$ 用の

AWGN観測モデル

$$X_t = \frac{\bar{\zeta}_t}{\bar{\zeta}_{out,t}} X + H_t, \quad H_t \sim \mathcal{N}(0, \sigma_t^2)$$

$X_t$ を観測したときに、未知の信号 $X \sim x_1$ を推定せよ。

$\eta_{out,t}$ 用の

AWGN観測モデル

$$Y = f(Z, W), \quad W \sim \mathcal{N}(0, \sigma^2)$$

$$Z_0 = 0, \quad Z_t = \frac{\bar{\zeta}_t}{\bar{\zeta}_{out,t}} Z + B_t$$

$$(Z, Z_t) \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{1}{\delta} \begin{bmatrix} \mathbb{E}[X^2] & \mathbb{E}[X\eta_{in,t-1}(X_{t-1})] \\ \mathbb{E}[X\eta_{in,t-1}(X_{t-1})] & \mathbb{E}[\eta_{in,t-1}^2(X_{t-1})] \end{bmatrix} \right)$$

$Z_t$ と $Y$ を観測したときに、信号 $Z$ を推定せよ。

# 状態発展法[3-4]

状態発展方程式 前述の確率変数 $Z_t$ 、 $Z$ 、 $W$ を用いて、

$$\bar{\xi}_{\text{out},t} = \mathbb{E} \left[ \frac{\partial}{\partial Z_t} \eta_{\text{out},t}(Z_t, f(Z, W)) \right],$$

$$\bar{\zeta}_t = -\mathbb{E} \left[ \frac{\partial}{\partial Z} \eta_{\text{out},t}(Z_t, f(Z, W)) \right], \quad \sigma_t^2 = \frac{\mathbb{E}[\eta_{\text{out},t}^2(Z_t, f(Z, W))]}{\bar{\xi}_{\text{out},t}^2}$$

これらを用いて前述の確率変数 $X_t$ を定め、 $Z_{t+1}$ を更新する。

誤差の漸近評価

$$\lim_{M=\delta N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\|\mathbf{x} - \hat{\mathbf{x}}_t\|^2] = \mathbb{E} \left[ \{\eta_{\text{in},t-1}(X_{t-1}) - X\}^2 \right].$$

状態発展法の解説は、第4～6回講義資料を参照

# ベイズ最適なDenoiser

$\eta_{in,t}$ の場合

$$\eta_{in,t}(X_t) = \mathbb{E}[X|X_t]$$

AMPの場合と同様に定めればよい。

$\eta_{out,t}$ の場合

AWGN観測 $X_t$ の信号対雑音比(SNR)を最大化せよ。

$$\text{snr}_t = \left( \frac{\bar{\zeta}_t}{\bar{\xi}_{out,t}} \right)^2 \frac{1}{\sigma_t^2} = \frac{\bar{\zeta}_t^2}{\mathbb{E}[\eta_{out,t}^2(Z_t, f(Z, W))]}$$

# 補題3.1 [3-5]

AWGN観測モデル  $Y = f(Z, W), \quad W \sim \mathcal{N}(0, \sigma^2)$

$$Z_0 = 0, \quad Z_t = \frac{\bar{\zeta}_t}{\bar{\zeta}_{\text{out},t}} Z + B_t$$

$Z_t$ と $Y$ を観測したときの $Z$ の事後平均推定量 $\hat{Z}_t$ を以下で定義する。

$$\hat{Z}_t(z_t, y) = \frac{\int z p_{Y|Z}(y|z) e^{-\frac{(z-z_t)^2}{2v_t}} dz}{\int p_{Y|Z}(y|z) e^{-\frac{(z-z_t)^2}{2v_t}} dz}, \quad v_t = \frac{\mathbb{E} \left[ \left\{ \eta_{\text{in},t-1}(X_{t-1}) - X \right\}^2 \right]}{\delta}$$

ただし、 $p_{Y|Z}$ は $Z$ が与えられたときの $Y$ の確率密度関数を表す。

任意定数 $C \in \mathbb{R}$ に対して、以下の $\eta_{\text{out},t}$ は $\text{snr}_t$ を最大化する。

$$\eta_{\text{out},t}(Z_t, Y) = C \frac{Z_t - \hat{Z}_t(Z_t, Y)}{v_t}$$

[3-5] K. Takeuchi, "Decentralized generalized approximate message-passing for tree-structured networks," submitted to 2024 IEEE Int. Conf. Acoust. Speech Signal Process., Seoul, Korea, Apr. 2024.

# Zの条件付き分布

事後平均推定量の性質  $\mathbb{E}[\eta_{\text{in},t-1}(X_{t-1})\{\eta_{\text{in},t-1}(X_{t-1}) - X\}] = 0$  から、

$$\mathbb{E}[ZZ_t] = \frac{1}{\delta} \mathbb{E}[X\eta_{\text{in},t-1}(X_{t-1})] = \frac{1}{\delta} \mathbb{E}[\eta_{\text{in},t-1}^2(X_{t-1})] = \mathbb{E}[Z_t^2]$$

$Z_t$  が与えられたときの  $Z$  は次のように表現できる。

$$Z \sim Z_t + N_t, \quad N_t \sim \mathcal{N}(0, v_t)$$

**証明** 以下の  $Z$  と  $Z_t$  に関する共分散を確かめよ。

$$\begin{aligned} \mathbb{E}[ZZ_t] &= \mathbb{E}[(Z_t + N_t)Z_t] = \mathbb{E}[Z_t^2] \\ \mathbb{E}[Z^2] &= \mathbb{E}[Z_t^2] + \mathbb{E}[N_t^2] = \frac{1}{\delta} \mathbb{E}[\eta_{\text{in},t-1}^2(X_{t-1})] + v_t \\ &= \frac{1}{\delta} \{ \mathbb{E}[X^2] + 2\mathbb{E}[\eta_{\text{in},t-1}(X_{t-1})(\eta_{\text{in},t-1}(X_{t-1}) - X)] \} = \frac{1}{\delta} \mathbb{E}[X^2] \end{aligned}$$

■



# 線形観測の場合

$f(z, w) = z + w$ の場合にベイズ最適な $\eta_{\text{out},t}$ を導出する。

$Y = Z + W$ および $Z \sim \mathcal{N}(Z_t, v_t)$ から

$$\hat{Z}_t(Z_t, Y) = Z_t + \frac{v_t}{v_t + \sigma^2} (Y - Z_t) \quad \text{確かめよ。}$$

それゆえ  $\eta_{\text{out},t}(Z_t, Y) = \frac{C}{v_t + \sigma^2} (Z_t - Y)$   $C$ は任意

$$\xi_{\text{out},t} = \left\langle \frac{\partial}{\partial z_t} \eta_{\text{out},t}(\mathbf{z}_t, \mathbf{y}) \right\rangle = \frac{C}{v_t + \sigma^2}, \quad \bar{\xi}_t = \frac{C}{v_t + \sigma^2}$$

ベイズ最適なGAMPでは、一般に $\bar{\xi}_{\text{out},t} = \bar{\xi}_t$ [3-4]が言える。

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

線形観測に対するGAMPはAMPに帰着する。