2023年3月10日
電子情報通信学会 総合大会
企画講演セッション：Arimoto-Blahutアルゴリズムの50年

量子ABアルゴリズム再訪

Revisit to Quantum AB Algorithms

長岡浩司　（電気通信大学）

Hiroshi Nagaoka

0　はじめに

・・・私が解説できるのは四半世紀前に自分で提案した形の量子ABアルゴリズムに限ります。その考え方と応用例、いくつかの変種、提案当時の量子情報理論の様子などをお話しすることはできますが、それ以外の話題、特に

一方，量子版は，凸ではないのでABのような交互最適化あるいは勾配法のようなものでは最適解に収束しない

という問題に対処するためにその後に提案された種々のアルゴリズムについては不勉強であり、かつ今ひとつ興味が湧かないため、私は全く解説できません。それでも構わなければ・・・

（講演依頼に対する返信より）

参考文献

[1] H. Nagaoka, "Algorithms of Arimoto-Blahut type for computing quantum channel capacity," *Proc. of 1998 IEEE International Symposium on Information Theory*, p.354, 1998.

Arimoto, Blahut 1972 **から半世紀。そのちょうど中間の時期↑**

1　古典ABアルゴリズム

1.1　問題設定

- Given $\mathcal{X}$, $\mathcal{Y}$: **有限集合（入出力アルファベット）**
  - $\mathcal{P}(\mathcal{X}), \mathcal{P}(\mathcal{Y}) := \mathcal{X}, \mathcal{Y}$ **上の確率分布全体**
- $W = \{W(y|x)\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}}$ **：$\mathcal{X}$ から $\mathcal{Y}$ への通信路**
  - $W(y|x) \geq 0$ and $\sum_y W(y|x) = 1$
  - $W_x := W(\cdot|x) \in \mathcal{P}(\mathcal{Y})$
- **相互情報量：$\mathcal{X}$ 上の確率分布 $p \in \mathcal{P}(\mathcal{X})$ に対し、**

$$
\begin{aligned}
I(p) &:= \sum_{x\in\mathcal{X}} p(x) D(W_x \| W_p) = H(W_p) - \sum_{x\in\mathcal{X}} p(x) H(W_x) \\
&= I(X;Y) = H(Y) - H(Y|X)
\end{aligned}
$$

  - D: KL divergence,　H : Shannon entropy
  - $W_p := \sum_x p(x)\, W_x \in \mathcal{P}(\mathcal{Y})$
  - $p = p_X \sim X \to \boxed{W} \to Y \sim p_Y = W_p$

- capacity: $C := \max_p I(p)$

1.2 $I(p)$ の２変数化

- $$J(p, p') := -D(p\|p') + \sum_x p(x) D(W_x\|W_{p'})$$

  とおくと

  - $J(p,p) = I(p)$
  - $I(p) - J(p,p') = D(p\|p') - D(W_p\|W_{p'}) \geq 0$
  - $I(p) = \max_{p'} J(p,p')$ and $\arg\max_{p'} J(p,p') = p.$
  - $J(p,p') = H(p) + \sum_x p(x) \underbrace{\{\log p'(x) + D(W_x\|W_{p'})\}}_{=:F(x)}$
  - $\arg\max_p J(p,p') = \hat{p}$, where

$$\hat{p}(x) = \frac{1}{Z} \exp F(x) = \frac{1}{Z} p'(x) \exp D(W_x\|W_{p'})$$
$$Z = \sum_x p'(x) \exp D(W_x\|W_{p'})$$

1.3 アルゴリズム

- $p_{t+1} := \arg\max_p J(p, p_t), \quad t = 1, 2, \ldots,$ **とおくと**

$$I(p_t) = J(p_t, p_t) \leq J(p_{t+1}, p_t) \leq J(p_{t+1}, p_{t+1}) = I(p_{t+1})$$

1.4 写像 $\Gamma : \mathcal{P}(\mathcal{X}) \to \mathcal{P}(\mathcal{Y}),\ p \mapsto \Gamma(p) := W_p$ を用いると

- Γ **はアファイン写像：** $\Gamma(\lambda p + (1-\lambda)q) = \lambda\Gamma(p) + (1-\lambda)\Gamma(q).$
- $W_x = W(\delta_x)$
- $I(p) = \sum_x p(x) D(\Gamma(\delta_x) \| \Gamma(p))$
- $J(p, p') = -D(p\|p') + \sum_x p(x) D(\Gamma(\delta_x) \| \Gamma(p'))$
- $I(p) - J(p, p') = D(p\|p') - D(\Gamma(p) \| \Gamma(p')) \geq 0$

1.5 Arimoto, Blahut 1972 のオリジナル形について

- **逆向き通信路** $\varphi = \varphi(x|y)$ **を用いる。**
- $K(p,\varphi) := H(p) + \sum_{x,y} p(x)W(y|x)\log\varphi(x|y)$
- $\varphi_p^*(x|y) := \frac{p(x)W(y|x)}{W_p(y)}$ **とおくと**

$$p \sim X \to \boxed{W} \to Y \iff p \sim X \leftarrow \boxed{\varphi_p^*} \leftarrow Y$$

$$K(p,\varphi_p^*) = \underbrace{H(p)}_{H(X)} + \underbrace{\sum_{x,y} p(x)W(y|x)\log\varphi_p^*(x|y)}_{-H(X|Y)} = I(p)$$

- $$I(p) - K(p,\varphi) = \sum_{x,y} \underbrace{p(x)W(y|x)}_{W_p(y)\varphi_p^*(x|y)} \log\frac{\varphi_p^*(x|y)}{\varphi(x|y)}$$
$$= \sum_y W_p(y) D(\varphi_p^*(\cdot|y) \,\|\, \varphi(\cdot|y)) \geq 0$$

- $I(p) = \max_\varphi K(p,\varphi)$ and $\arg\max_\varphi K(p,\varphi) = \varphi_p^*$.

- $K(p,\varphi) = H(p) + \sum_x p(x) \underbrace{\sum_y W(y|x) \log \varphi(x|y)}_{=:G(x)}$
- $p_{t+1} := \arg\max_p K(p, \varphi^*_{p_t}), \quad t = 1, 2, \ldots,$ **とおくと**

$$I(p_t) = K(p_t, \varphi^*_{p_t}) \leq K(p_{t+1}, \varphi^*_{p_t}) \leq K(p_{t+1}, \varphi^*_{p_{t+1}}) = I(p_{t+1})$$

- $K(p, \varphi^*_{p'}) = J(p, p').$

1.6 比較 and φを用いた理由？

- $I(p) - J(p,p') = D(p\|p') - D(\Gamma(p)\|\Gamma(p')) \geq 0$ **：Dの単調性**
- $I(p) - K(p,\varphi) = \sum_y W_p(y) D(\varphi^*_p(\cdot|y) \,\|\, \varphi(\cdot|y)) \geq 0$ **：Dの正値性**
- $I(p) - K(p, \varphi^*_{p'}) = \sum_y W_p(y) D(\varphi^*_p(\cdot|y) \,\|\, \varphi^*_{p'}(\cdot|y))$

  $\|$

  $I(p) - J(p,p') = D(p\|p') - D(\Gamma(p)\|\Gamma(p'))$
- $D(p\|p') = D(\Gamma(p)\|\Gamma(p')) + \sum_y W_p(y) D(\varphi^*_p(\cdot|y) \,\|\, \varphi^*_{p'}(\cdot|y))$ **：Dの**chain rule

  $D(p_X\|p'_X) = D(p_{XY}\|p'_{XY}) = D(p_Y\|p'_Y) + \sum_y p_Y(y) D(p_{X|Y}(\cdot|y)\|p'_{X|Y}(\cdot|y))$

2 Quantum AB Algorithms

2.1 問題設定

- **有限次元ヒルベルト空間 $\mathcal{H}$ に対し、**
  - $\mathcal{L}(\mathcal{H})$：$\mathcal{H}$ **上の線形作用素の全体**
  - $\mathcal{S}(\mathcal{H})$：$\mathcal{H}$ **上の密度作用素の全体**

  とおく。
- Given **「通信路」** $\Gamma : \mathcal{S}_1 \to \mathcal{S}(\mathcal{H}_2)$ s.t.
  - $\mathcal{S}_1$ **は** $\mathcal{S}(\mathcal{H}_1)$ **の閉凸集合**
  - $\mathcal{H}_1, \mathcal{H}_2$ **は入力系と出力系を表すヒルベルト空間**
  - **アファイン性** $\Gamma(\lambda\sigma_1 + (1-\lambda)\sigma_2) = \lambda\Gamma(\sigma_1) + (1-\lambda)\Gamma(\sigma_2)$
  - Γ **は、**Trace Preserving Positive (TPP) Map $\mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ **の** $\mathcal{S}_1$ **への制限**
  - completely positive (TPCP) **であるという仮定は必要ない。**
  - **入力アンサンブルの集合：**$n = 2, 3, \ldots$ **に対し、**

$$\Pi_n := \{\pi = (\lambda_i, \sigma_i)_{i=1}^n \mid 0 \leq \lambda_i \in \mathbf{R},\ \sum_{i=1}^n \lambda_i = 1,\ \sigma_i \in \mathcal{S}_1\}.$$

- **ホレボー相互情報量：** $\pi = (\lambda_i, \sigma_i) \in \Pi_n$ **に対し、**

$$I(\pi) := \sum_{i=1}^{n} \lambda_i D(\Gamma\sigma_i \| \Gamma\sigma_\lambda) = H(\Gamma\sigma_\lambda) - \sum_{i=1}^{n} \lambda_i H(\Gamma\sigma_i),$$

where $\sigma_\lambda := \sum_i \lambda_i \sigma_i$, $\Gamma\sigma_i := \Gamma(\sigma_i)$, etc.

- D：**量子相対エントロピー、**H：von Neumann **エントロピー**
- cf. $I(p) = \sum_{x\in\mathcal{X}} p(x) D(\Gamma(\delta_x) \| \Gamma(p)) = H(\Gamma(p)) - \sum_{x\in\mathcal{X}} p(x) H(\Gamma(\delta_x))$

- **Γの（古典）通信路容量**
  - $C = \max_{n} \max_{\pi \in \Pi_n} I(\pi)$
  - $n \geq \dim \Gamma(\mathcal{S}_1) + 1 \quad \Rightarrow \quad C = \max_{\pi \in \Pi_n} I(\pi)$
  - C**は以下のような通信系の** Maximam Reliable Transmission Rate

送信メッセージ$m \in \{0, 1, \ldots, M-1\}$

$\rightarrow$ **符号化**f $\rightarrow$ **符号語**$f(m) = (\sigma_1, \ldots, \sigma_L)$, $\sigma_i \in \mathcal{S}_1$, L : code length

$\rightarrow$ **送信状態** $\sigma_1 \otimes \cdots \otimes \sigma_L$

$\rightarrow$ $\Gamma^{\otimes L}$ $\rightarrow$ **受信状態** $\Gamma(\sigma_1) \otimes \cdots \otimes \Gamma(\sigma_L)$

$\rightarrow$ **測定&復号化** $\rightarrow$ **復号メッセージ** $\hat{m}$

2.2 Some Remarks

- $\mathcal{S}_1 = \mathcal{S}(\mathcal{H}_1)$**の場合が主要な対象**
- **送信状態の候補**$\{\sigma_1, \ldots, \sigma_k\}$**が与えられている場合は、その凸包を**$\mathcal{S}_1$**とする。**
- Γ**を** c-q channel (c: **古典**, q: **量子**) **とみなした場合の（古典）通信路容量** (classical) capacity **が** C**。** Holevo capacity **とも呼ぶ。**

- Γが completely positive (TPCP) の場合は q-q channel ともみなせる。
  q-q channel には classical capacity と quantum capacity がある。
- q-q chanel の classical capacity と Holevo capacity は一致するとは限らない。
  これについては後述。
- $\pi = (\lambda_i, \sigma_i)_{i=1}^n$ を $\mathcal{S}_1$ 上の確率測度 $\sum_{i=1}^n \lambda_i \delta_{\sigma_i}$ とみなすなら、$I(\pi)$ は上に凸 (concave) な関数である。
- ただし、π の凸結合は n を保存しない。すなわち、Π_n は凸集合ではない。
  よって、固定した n のもとでの $\max_{\pi \in \Pi_n} I(\pi)$ は凸最適化問題ではない。
  $\Rightarrow$　以下で述べる逐次アルゴリズムの大域的収束性は保証されない。
- $\max_{\pi \in \Pi_n} I(\pi)$ を求めるとき、$\pi = (\lambda_i, \sigma_i)_{i=1}^n$ における n 個の状態 $\{\sigma_i\}$ はすべて $\mathcal{S}_1$ の端点と仮定してよい。
  特に、$\mathcal{S}_1 = \mathcal{S}(\mathcal{H}_1)$ の場合は、$\{\sigma_i\}$ はすべて純粋状態としてよい。

2.3 量子ABアルゴリズム・その１：境界型

- $I(\pi)$の2変数化

$$J(\pi,\pi') := -D(\lambda\|\lambda') + \sum_{i=1}^{n} \lambda_i \mathrm{Tr}\left[(\Gamma\sigma_i)\{\log(\Gamma\sigma'_i) - \log(\Gamma\sigma'_{\lambda'})\}\right]$$

where

$$\pi = (\lambda_i, \sigma_i),\ \pi' = (\lambda'_i, \sigma'_i),\quad \sigma'_{\lambda'} := \sum_i \lambda'_i \sigma'_i$$

- $J(\pi,\pi) = I(\pi)$
-

$$\begin{aligned} I(\pi) - J(\pi,\pi') &= D(\lambda\|\lambda') + \sum_i \lambda_i D(\Gamma\sigma_i\|\Gamma\sigma'_i) - D(\Gamma\sigma_\lambda\|\Gamma\sigma'_{\lambda'}) \\ &= D\Big(\bigoplus_i \lambda_i\,\Gamma\sigma_i \,\Big\|\, \bigoplus_i \lambda'_i\,\Gamma\sigma'_i\Big) - D(\Gamma\sigma_\lambda\|\Gamma\sigma'_{\lambda'}) \\ &\geq 0 \quad \leftarrow D\text{の単調性} \end{aligned}$$

where

$$\bigoplus_i \lambda_i \Gamma\sigma_i = \begin{pmatrix} \lambda_1\Gamma\sigma_1 & & O \\ & \ddots & \\ O & & \lambda_n\Gamma\sigma_n \end{pmatrix} \xrightarrow{\text{部分トレース}} \sum_i \lambda_i \Gamma\sigma_i = \Gamma\sigma_\lambda$$

- $\max_{\pi'} J(\pi, \pi') = I(\pi)$ and $\arg\max_{\pi'} J(\pi, \pi') = \pi$
- **上記の単調性は「逆向き通信路 + chain rule」では表せない。**
-

$$J(\pi, \pi') = H(\lambda) + \sum_{i=1}^{n} \lambda_i \Big\{ \log \lambda_i' + \underbrace{\mathrm{Tr}\left[(\Gamma\sigma_i)\{\log(\Gamma\sigma_i') - \log(\Gamma\sigma_{\lambda'}')\}\right]}_{=: a_i(\sigma_i)} \Big\}$$

- $\arg\max_\pi J(\pi, \pi') = \hat{\pi} = (\hat{\lambda}_i, \hat{\sigma}_i)_{i=1}^n$, where

$$\hat{\sigma}_i = \arg\max_{\sigma \in \partial_e \mathcal{S}_1} a_i(\sigma) \text{ and } \hat{\lambda}_i = \frac{1}{Z} \lambda'_i \exp a_i(\hat{\sigma}_i),$$

$$Z := \sum_i \lambda'_i \exp a_i(\hat{\sigma}_i),$$

$$\partial_e \mathcal{S}_1 = \text{the extreme boundary of } \mathcal{S}_1.$$

- **特に$\mathcal{S}_1 = \mathcal{S}(\mathcal{H}_1)$ の場合は、**
  - $\hat{\sigma}_i = \arg\max_{\sigma \in \partial_e \mathcal{S}_1} = |\psi_i\rangle\langle\psi_i|$, where

    $|\psi_i\rangle$ は、$\mathcal{H}_1$ 上の作用素 $\Gamma^*\{\log(\Gamma\sigma'_i) - \log(\Gamma\sigma'_{\lambda'})\}$ の最大固有値に対応する単位固有ベクトル。

  - **Γ^* は Γ の双対写像（転置写像）** $\mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_1)$：
    $\forall \sigma \in \mathcal{S}(\mathcal{H}_1), \forall X \in \mathcal{L}(\mathcal{H}_2)$, $\mathrm{Tr}\left[(\Gamma\sigma)X\right] = \mathrm{Tr}\left[\sigma(\Gamma^* X)\right]$.
- recursion: $\pi^{(t)} \to \pi^{(t+1)} = \arg\max_{\pi \in \Pi_n} J(\pi, \pi^{(t)}), \quad t = 1, 2, \ldots,$

- $\mathcal{S}_1 = \mathcal{S}(\mathcal{H}_1)$ とする。
- $I(\pi)$ の2変数化：

$$K(\pi, \pi') := J(\pi, \pi') - \sum_{i=1}^{n} \lambda_i D(\sigma_i \| \sigma'_i)$$

- $K(\pi, \pi) = I(\pi)$
- $I(\pi) - K(\pi, \pi') = I(\pi) - J(\pi, \pi') + \sum_i \lambda_i D(\sigma_i \| \sigma'_i) \geq 0.$
- $\max_{\pi'} K(\pi, \pi') = I(\pi)$ and $\arg\max_{\pi'} K(\pi, \pi') = \pi$
-

$$K(\pi, \pi') = H(\lambda) + \sum_i \lambda_i \Big(\log \lambda'_i + H(\sigma_i) + \mathrm{Tr}\left[\sigma_i \left(\log \sigma'_i + \Gamma^* \{\log(\Gamma \sigma'_i) - \log(\Gamma \sigma'_{\lambda'})\}\right)\right] \Big)$$

- $\arg\max_\pi K(\pi, \pi') = \hat{\pi} = (\hat{\lambda}_i, \hat{\sigma}_i)_{i=1}^n$, where

$$\hat{\lambda}_i = \frac{1}{Z}\mathrm{Tr}B_i \ \text{ and } \ \hat{\sigma}_i = \frac{1}{Z\hat{\lambda}_i}B_i,$$
$$B_i = \lambda_i' \exp\left(\log\sigma_i' + \Gamma^*\{\log(\Gamma\sigma_i') - \log(\Gamma\sigma_{\lambda'}')\}\right),$$
$$Z = \sum_i \mathrm{Tr}B_i.$$

- **一般化：定数$\alpha \geq 0$を用いて**

$$K_\alpha(\pi, \pi') := J(\pi, \pi') - \alpha\sum_{i=1}^n \lambda_i D(\sigma_i \| \sigma_i')$$

2.5　量子ABアルゴリズム・その３：状態固定／連続型

- $\lambda = \lambda(\sigma)d\sigma$: $\partial_e \mathcal{S}_1$ 上の確率測度
-

$$L(\lambda, \lambda') = -D(\lambda \| \lambda') + \int_{\partial_e \mathcal{S}_1} \lambda(\sigma)\, D(\Gamma\sigma \| \Gamma\sigma_{\lambda'})\, d\sigma$$

$$\sigma_{\lambda'} = \int_{\partial_e \mathcal{S}_1} \lambda'(\sigma)\, \sigma\, d\sigma.$$

- $\max_{\lambda'} L(\lambda, \lambda') = I(\lambda)$ and $\arg\max_{\lambda'} L(\lambda, \lambda') = \lambda$
- $\arg\max_{\lambda} L(\lambda, \lambda') = \hat{\lambda}$, where

$$\hat{\lambda}(\sigma) = \frac{1}{Z}\lambda'(\sigma) \exp\left(D(\Gamma\sigma \| \Gamma\sigma_{\lambda'})\right),$$

$$Z = \int_{\partial_e \mathcal{S}_1} \lambda'(\sigma) \exp\left(D(\Gamma\sigma \| \Gamma\sigma_{\lambda'})\right) d\sigma.$$

3　四方山話：発表時とその後について

- H. Nagaoka, "Algorithms of Arimoto-Blahut type for computing quantum channel capacity," *Proc. of 1998 IEEE International Symposium on Information Theory*, p.354, 1998.
- MIT, Boston
- ~~HWS~~ theorem $C = \max_{\pi} I(\pi)$: 1997 **前後**

HSW

Holevo-Schumacher-Westmoreland

MOC6 Quantum Information Theory

1998 ISIT 로마

THD4 Quantum Communications

1998 ISIT 202

Session WEAM3 Algebraic Coding Theory II 1997 ISIT

- **座長の** Holevo **曰く：**additivity problem **に適用してみたらどうか。**
- q-q channel Γ **の** classical capacity $C^*(\Gamma)$ **：**
  以下のような通信系の Maximam Reliable Transmission Rate

  送信メッセージ $m \in \{0, 1, \ldots, M-1\}$

  $\rightarrow$ **符号化** f $\rightarrow$ **符号語** $=$ **送信状態** $f(m) = \tilde{\sigma} \in \mathcal{S}(\mathcal{H}_1^{\otimes L})$: **エンタングルを許す**

  $\rightarrow$ $\Gamma^{\otimes L}$ $\rightarrow$ **受信状態** $\Gamma^{\otimes L}(\tilde{\sigma})$

  $\rightarrow$ **測定&復号化** $\rightarrow$ **復号メッセージ** $\hat{m}$

- $C^*(\Gamma) = \lim_{n\to\infty} \frac{1}{n} C(\Gamma^{\otimes n}) \geq C(\Gamma)$
-
- $C(\Gamma_1 \otimes \Gamma_2) \geq C(\Gamma_1) + C(\Gamma_2)$
- **加法性予想** additivity conjecture : **常に等号成立**
- **加法性予想** $\Rightarrow C^*(\Gamma) = C(\Gamma)$.

- S. Osawa and H. Nagaoka, "Numerical Experiments on the Capacity of Quantum Channel with Entangled Input States", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol.E84-A, No.10, pp.2583-2590 (2001). （加法性予想の検証）
- **しかし、実は・・・**
  Hastings, M.B.: Superadditivity of communication capacity using entangled inputs. Nature Phys. 5, 255 (2009)

ご清聴ありがとうございました。