

平成 27 年度 修士学位論文

劣化型放送通信路の通信路容量域計算アルゴリズム
ムについて

電気通信大学 大学院 情報理工学研究科
博士前期課程 情報・通信工学専攻

1331004 新井 敦

指導教員 大濱 靖匡 教授 川端 勉 教授

提出 平成 27 年 1 月 30 日

修 士 論 文 の 和 文 要 旨

| | | | |
|---|---------------------------------|------|---------------|
| 研究科・専攻 | 大学院 情報理工 学研究科 情報・通信工学 専攻 博士前期課程 | | |
| 氏 名 | 新井 敦 | 学籍番号 | 1 3 3 1 0 0 4 |
| 論 文 題 目 | 劣化型放送通信路の通信路容量域計算アルゴリズムについて | | |
| <p>要 旨</p> <p>情報理論の分野において、放送通信路の理論的モデルに対する通信の理論限界の解明に関する多数の研究が行われてきた。有本や Blahut によって、離散無記憶単一通信路に対する通信路容量を数値的に計算するアルゴリズムが得られているが、多端子通信路については通信路容量域を計算するアルゴリズムは確立されておらず、情報理論における基本的未解決問題の1つとなっている。</p> <p>そこで本論文では1入力2出力劣化型放送通信路(DBC)に対し、その容量域を数値的に求める問題を考察し、その容量域を逐次的に計算する新しいアルゴリズムを提案した。</p> <p>Arimoto-Blahut アルゴリズムを拡張して、DBC に適用させるアルゴリズムを用いると、最適値を与える分布に収束する条件が厳しく、通信路容量域を求める事が非常に困難である。</p> <p>一方、本研究で提案するアルゴリズムは、目的関数にペナルティ項を導入し、そのペナルティ項の値を非常に大きな値にすることで、最適値を与える分布に収束することが期待できる。</p> <p>実際に最適値を与える分布への収束の証明法については見当がついていないのだが、2元対称 DBC の理論値との比較実験を行い、本研究の提案アルゴリズムの性能評価を行ったところ、理論値と一致する事を確認する事ができた。</p> <p>このことから提案アルゴリズムの定める分布列は最適値を与える分布に収束するという事を確認でき、2元対称 DBC の通信路容量域を求める事ができた。</p> <p>DBC の通信路容量域に関する過去の研究で、実際に通信路容量域を描くという事はできておらず、おそらく本研究が初めてだといえる。このことから分かる通り、本研究は情報理論における未解決問題に大きな進展を与えたと考えられる。</p> | | | |



目次

| | | |
|----------|------------------------------------|-----------|
| 1 | はじめに | 3 |
| 2 | 2 端子通信路と通信路容量計算 | 5 |
| 2.1 | 2 端子通信路 | 5 |
| 2.2 | Arimoto-Blahut アルゴリズム | 6 |
| 2.3 | 最適分布への収束 | 9 |
| 3 | 劣化型放送通信路と容量域計算 | 11 |
| 3.1 | 劣化型放送通信路 | 11 |
| 3.2 | 容量域の支持直線表現 | 12 |
| 3.3 | DBC に関するこれまでの研究 | 14 |
| 4 | 提案アルゴリズム | 16 |
| 4.1 | Arimoto-Blahut アルゴリズムの拡張 | 16 |
| 4.2 | 最適分布への収束 | 19 |
| 4.3 | 提案する新しい反復アルゴリズム | 20 |
| 4.4 | 従来アルゴリズムとの相違 | 26 |
| 5 | 提案アルゴリズムの収束性 | 28 |
| 5.1 | 収束性に関する結果 | 28 |
| 5.2 | 最適分布への収束条件 | 30 |
| 6 | 数値実験 | 32 |
| 6.1 | 2 元対称 DBC の理論値 | 32 |
| 6.2 | 実験の設定 | 33 |
| 6.3 | 実験の結果 | 34 |
| 6.3.1 | 2 元の場合 | 34 |
| 6.3.2 | 3 元の場合 | 37 |

| | |
|------------------------|----|
| | 2 |
| 6.4 実験結果のまとめ | 39 |
| 7 おわりに | 40 |
| A 補足 | 43 |

第 1 章

はじめに

近年、コンピュータやインターネットなどの普及により、利用者が増加し、多種多様のデータが端末間でやり取りされている。そして、技術の向上に伴い、高品質のデータを用いる事が多くなったため、端末間でやり取りされるデータの通信量が年々増大している。したがって、膨大な情報を効率良く配信するために、同じ情報を同時に不特定多数の受信者に送信するブロードキャスト通信やマルチキャスト通信などが用いられている。特にブロードキャスト通信において異なる情報を複数の受信端末に送信する理論的なモデルである放送通信路の需要が高まっており、その解析が重要となっている。

情報理論の分野において、放送通信路の理論的モデルに対する、通信の理論限界の解明に関する多数の研究が行われてきた。有本 [1], Blahut [2] によって、離散無記憶単一通信路に対する通信路容量を数値的に計算するアルゴリズムが得られている。一方、多端子通信路については文献 [3] で述べられているいくつかの特別な場合を除き、通信路容量域を計算するアルゴリズムは確立されておらず、多端子情報理論における基本的未解決問題の 1 つとなっている。

本論文では 1 入力 2 出力劣化型放送通信路に対し、その容量域を数値的に求める問題を考察する。この問題は、文献 [4]-[6] で考察された。これらの研究では有本 [1], Blahut [2] のアルゴリズムの放送通信路への拡張アルゴリズムが提案され、そのアルゴリズムが定める分布列の領域の境界を与える分布への収束について議論されている。本研究では劣化型放送通信路に対し、その容量域を逐次的に計算する新しいアルゴリズムを提案する。従来の計算アルゴリズムが容量域の内点を与える分布から境界を与える分布に近づくのに対し、本研究で提案するアルゴリズムの定める分布列は、必ずしも容量域の内点を与えるとは限らない分布から、容量域の境界を与える分布へ近づいていくものである。この点で本研究の提案アルゴリズムは、従来のアルゴリズムとは異なるものになっている。また、本論文では提案したアルゴリズムの定める分布列の通信路容量域の境界を与える分布への収束の可能性を議論し、数値実験によりアルゴリズムの性能を評価する。

本論文の 2 章以降の構成は以下の通りである。第 2 章では、2 端子通信路の通信路容量を導出する Arimoto-Blahut アルゴリズムについて記す。第 3 章では劣化型放送通信路につ

いての定義や通信路容量を導出するための基本的な考え方, 劣化型放送通信路に関するこれまでの研究について記す. 第4章では, 提案アルゴリズムについて, Arimoto-Blahut アルゴリズムを自然に拡張したアルゴリズムと, 新しい提案アルゴリズムについて記す. 第5章では提案したアルゴリズムの定める分布列の通信路容量域の境界を与える分布への収束の可能性について記し, その最適解への収束条件について記す. 第6章では, 2元劣化型放送通信に対し, 提案アルゴリズムが最適値を与える分布に収束するかどうか理論値と比較実験を行い, 提案アルゴリズムの性能を評価する. 第7章ではまとめと今後の課題を記す.

第 2 章

2 端子通信路と通信路容量計算

この章では、多端子通信路の通信路容量について議論する前に、2 端子通信路の通信路容量について記述する。

2.1 2 端子通信路

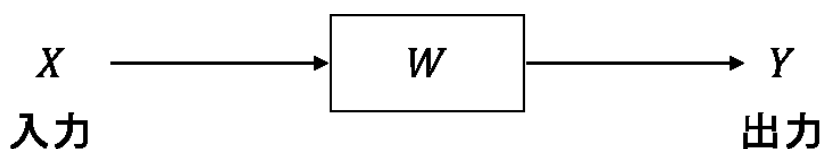


図 2.1: 2 端子通信路

本節では 2 端子通信路について記述する。本章では \mathcal{X}, \mathcal{Y} を有限集合とし、通信路の入力を表す確率変数を $X \in \mathcal{X}$ 、出力を表す確率変数を $Y \in \mathcal{Y}$ とする。また \mathcal{X}, \mathcal{Y} の各要素をそれぞれ x, y で表す。また、 $W = \{W(y|x)\}_{(x,y) \in \mathcal{X} \times \mathcal{Y}}$ を通信路を表す確率行列とし、 (X, Y) の同時分布を p_{XY} で表す。この同時分布は通信路の入力分布 p_X と W を用いると

$$\begin{aligned} p_{XY} &= \{p_{XY}(x, y)\}_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \\ &= \{p_X(x)W(y|x)\}_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \end{aligned} \quad (2.1)$$

で与えられる。入力確率変数 X と出力確率変数 Y との間の相互情報量 $I(X; Y)$ は、

$$I(X; Y) = \sum_{x,y} p_{XY}(x, y) \log \frac{W(y|x)}{p_Y(y)} \quad (2.2)$$

で与えられる。 $I(X; Y)$ が入力分布の関数であることを明示したい場合、 $I(X; Y) = I(p_X, W)$ と表す。 2端子通信路の通信路容量を $C(W)$ と表す。 これは以下の式で定義される。

$$C(W) \triangleq \max_{p_X} I(p_X, W) \quad (2.3)$$

$C(W)$ を求めるアルゴリズムは、 Arimoto[1] や Blahut[2] により独立に定義され、 Arimoto-Blahut アルゴリズムと呼ばれている。 このアルゴリズムは単一通信路の通信路容量を求める交互最大化アルゴリズムである。 次節では Arimoto-Blahut アルゴリズムについて記述する。

2.2 Arimoto-Blahut アルゴリズム

入力集合 \mathcal{X} 上の2つの確率分布を p_X, q_X とする。 通信路容量 $C(W)$ を求める為に以下の目的関数を定義する。

$$F(p_X, q_X) \triangleq I(q_X, W) + D(q_Y \| p_Y) - D(q_X \| p_X) \quad (2.4)$$

ここで p_Y, q_Y は、それぞれ分布 p_X, q_X で指定される確率変数 $X \in \mathcal{X}$ を通信路につないだときの出力 $Y \in \mathcal{Y}$ の確率分布である。 具体的には

$$p_Y = \{p_Y(y)\}_{y \in \mathcal{Y}}, \quad p_Y(y) = \sum_{x \in \mathcal{X}} p_X(x) W(y|x),$$

$$q_Y = \{q_Y(y)\}_{y \in \mathcal{Y}}, \quad q_Y(y) = \sum_{x \in \mathcal{X}} q_X(x) W(y|x)$$

で与えられる。 p_Y, q_Y がそれぞれ入力分布 p_X, q_X に対する W からの出力分布であることを明示するときは、 $q_Y = q_X W, p_Y = p_X W$ と記す。 このとき $F(p_X, q_X)$ は

$$F(p_X, q_X) \triangleq I(q_X, W) + D(q_X W \| p_X W) - D(q_X \| p_X)$$

と書ける。 この目的関数について次の補題が成り立つ。

補題 1 固定した q_X に対し、 $F(p_X, q_X)$ は $p_X = q_X$ のときに最大値

$$F(q_X, q_X) = I(q_X, W) \quad (2.5)$$

をとる。

(証明) ダイバージェンスに関する性質により、

$$D(q_Y \| p_Y) = D(q_X W \| p_X W) \leq D(q_X \| p_X) \quad (2.6)$$

が成り立つ。詳細は付録の定理Bを参照せよ。(2.6)より以下を得る。

$$F(p_X, q_X) = I(q_X, W) + D(q_X W \| p_X W) - D(q_X \| p_X) \leq I(q_X, W) \quad (2.7)$$

等号は $p_X = q_X$ のとき成り立つ。 ■

次に固定した p_X に対し考える。このとき、次の補題が成り立つ。

補題2 固定した p_X に対し $F(p_X, q_X)$ は

$$q_X(x) = \frac{1}{K} \exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{p_Y(y)} \right\} p_X(x)$$

のときに最大値 $\log K$ をとる。ここで K は規格化定数であり、以下で与えられる。

$$K = \sum_x \exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{p_Y(y)} \right\} p_X(x) \quad (2.8)$$

(証明) $F(p_X, q_X)$ は以下で与えられる。

$$\begin{aligned} F(p_X, q_X) &= \sum_{x,y} q_X(x) W(y|x) \log \frac{W(y|x)}{p_Y(y)} - \sum_x q_X(x) \log \frac{q_X(x)}{p_X(x)} \\ &= - \sum_x q_X(x) \log \frac{q_X(x)}{\exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{p_Y(y)} \right\} p_X(x)} \end{aligned} \quad (2.9)$$

式(2.9)の分母の量を確率分布にするために分母に $\frac{1}{K}$ をかけると以下を得る。

$$\begin{aligned} F(p_X, q_X) &= - \sum_x q_X(x) \log \frac{q_X(x)}{\frac{1}{K} \exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{p_Y(y)} \right\} p_X(x)} + \log K \\ &\leq \log K \end{aligned} \quad (2.10)$$

不等式(2.10)において、等号は

$$q_X(x) = \frac{1}{K} \exp \left\{ \sum_{x,y} p_X(x) W(y|x) \log \frac{W(y|x)}{p_Y(y)} \right\} p_X(x)$$

のとき成り立つ。 ■

次に補題1,2に基づき $C(W)$ を求めるアルゴリズムを導出する。 $t = 0, 1, 2, \dots$ に対し $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ 上の分布の列

$$q_X^{[t]} = \{q_X^{[t]}(x)\}_{x \in \mathcal{X}} \quad (2.11)$$

を次の様に定める。

分布更新アルゴリズム

1. $q_X^{[0]}$ を \mathcal{X} 上の適当な分布にとる.
2. $t = 0, 1, 2, \dots$ に対する分布更新式を以下で定める.

$$q_X^{[t+1]}(x) = \frac{1}{K_t} \exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{q_Y^{[t]}(y)} \right\} q_X^{[t]}(x), x \in \mathcal{X} \quad (2.12)$$

ここで K_t は規格化定数であり、以下で与えられる.

$$K_t = \sum_x \exp \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{q_Y^{[t]}(y)} \right\} q_X^{[t]}(x) \quad (2.13)$$

上記の分布更新アルゴリズムに対して、次の命題が成り立つ.

命題 1 $t = 0, 1, 2, \dots$ に対し、次の不等式が成り立つ.

$$\begin{aligned} & F(q_X^{[0]}, q_X^{[0]}) \\ & \stackrel{(a)}{\leq} F(q_X^{[0]}, q_X^{[1]}) \stackrel{(b)}{\leq} F(q_X^{[1]}, q_X^{[1]}) \leq \dots \\ & \stackrel{(a)}{\leq} F(q_X^{[t-1]}, q_X^{[t]}) \\ & \stackrel{(b)}{\leq} F(q_X^{[t]}, q_X^{[t]}) \end{aligned} \quad (2.14)$$

$$\begin{aligned} & \stackrel{(a)}{\leq} F(q_X^{[t]}, q_X^{[t+1]}) \\ & \stackrel{(b)}{\leq} F(q_X^{[t+1]}, q_X^{[t+1]}) \leq \dots \\ & \leq C(W) \end{aligned} \quad (2.15)$$

ここで不等式 (2.14) および (2.15) の右辺に現れる量に関してはそれぞれ以下が成り立つ.

$$F(q_X^{[t]}, q_X^{[t]}) = \sum_{x,y} q_X^{[t]}(x) W(y|x) \log \frac{W(y|x)}{q_Y^{[t]}(y)} \quad (2.16)$$

$$F(q_X^{[t]}, q_X^{[t+1]}) = \sum_{x,y} q_X^{[t+1]}(x) W(y|x) \left\{ \log \frac{W(y|x)}{q_Y^{[t]}(y)} + \log \frac{q_X^{[t]}(x)}{q_X^{[t+1]}(x)} \right\} \quad (2.17)$$

(証明) 命題 1 の不等式におけるステップ (a) と (2.17) は、補題 1 より従う. またステップ (b) と (2.16) は補題 2 より従う. ■

次の節で、Arimoto-Blahut アルゴリズムの定める分布列が最適分布へ収束することを証明する.

2.3 最適分布への収束

$F(q_X, q_Y)$ の最適値を与える分布を q_X^* とすると、補題1より

$$C(W) = \sum_{x,y} q_X^*(x) W(y|x) \log \frac{W(y|x)}{q_Y^*(y)} \quad (2.18)$$

Arimoto-Blahut アルゴリズムの収束性について次の命題が成り立つ。

命題2 $t = 0, 1, 2, \dots$ に対し以下が成り立つ。

$$\begin{aligned} & C(W) - F(q_X^{[t]}, q_X^{[t+1]}) \\ &= C(W) - \log K_t \\ &= \sum_{x,y} q_X^*(x) W(y|x) \log \frac{W(y|x)}{q_Y^*(y)} - \sum_x q_X^*(x) \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{q_Y^{[t]}(y)} \right\} + \sum_x q_X^*(x) \log \frac{q_X^{[t+1]}(x)}{q_X^{[t]}(x)} \\ &= -D(q_Y^* \| q_Y^{[t]}) + D(q_X^* \| q_X^{[t]}) - D(q_X^* \| q_X^{[t+1]}) \end{aligned}$$

命題2の結果より、次の定理が得られる。

定理1 Arimoto-Blahut アルゴリズムの定める分布列 $\{q_X^{[t]}\}_{t=0}^{+\infty}$ は $t \rightarrow +\infty$ のとき q_X^* へ収束する。

(証明) 命題2より、 $t = 0, 1, 2, \dots$ について以下が成り立つ。

$$C(W) - F(q_X^{[t]}, q_X^{[t+1]}) \leq D(q_X^* \| q_X^{[t]}) - D(q_X^* \| q_X^{[t+1]}) \quad (2.19)$$

そこで、 $\Delta_t = C(W) - F(q_X^{[t]}, q_X^{[t+1]})$ とおくと、

$$\Delta_t \leq D(q_X^* \| q_X^{[t]}) - D(q_X^* \| q_X^{[t+1]}) \quad (2.20)$$

$t = 0, 1, 2, \dots, T$ について和をとると、

$$\sum_{t=0}^T \Delta_t \leq D(q_X^* \| q_X^{[0]}) - D(q_X^* \| q_X^{[T+1]}) \quad (2.21)$$

命題1より $\{\Delta_t\}_{t=0}^T$ は単調減少列であるから(2.21)より、

$$T \Delta_T \leq \sum_{t=0}^T \Delta_t \leq D(q_X^* \| q_X^{[0]})$$

これより

$$0 \leq \Delta_T \leq \frac{1}{T} D(q_X^* \| q_X^{[0]}) \rightarrow 0 (T \rightarrow +\infty)$$

となり，分布更新アルゴリズムが最適値を与える分布へ収束することがわかる．以上で定理1は証明された． ■

以上より，Arimoto-Blahut アルゴリズムの定める分布は最適分布に収束し，2端子通信路の通信路容量を求める事ができる．次章からは劣化型放送通信路に対する通信路容量域計算アルゴリズムについて記述する．

第3章

劣化型放送通信路と容量域計算

3.1 劣化型放送通信路

本章から $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ を有限集合とし、通信路の入力を表す確率変数を $X \in \mathcal{X}$, 出力を表す確率変数を $Y \in \mathcal{Y}, Z \in \mathcal{Z}$ とする。また $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ の各要素をそれぞれ x, y, z で表すとする。

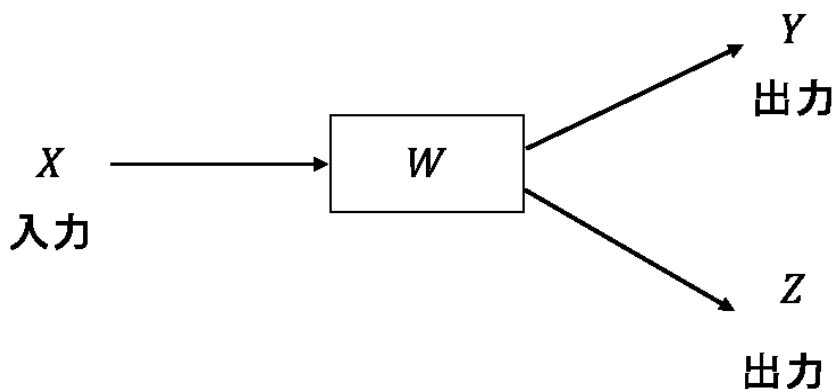


図 3.1: 対称放送通信路

その確率行列は以下で与えられる。

$$W = \{W(y, z|x)\}_{(x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}$$

本論文では特に、上記の確率行列が

$$W(y, z|x) = \{W_1(y|x)W_2(z|y)\}_{(x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} \quad (3.1)$$

なる条件を満たす場合を考える。

条件 (3.1) を満たす放送通信路は劣化型放送通信路 (Degraded Broadcast Channels (DBC)) とよばれる。以後、この通信路のことを DBC と記述する。通信路が劣化型であるとは、 X, Y, Z がマルコフ連鎖 $X \rightarrow Y \rightarrow Z$ をなすことと等価である。

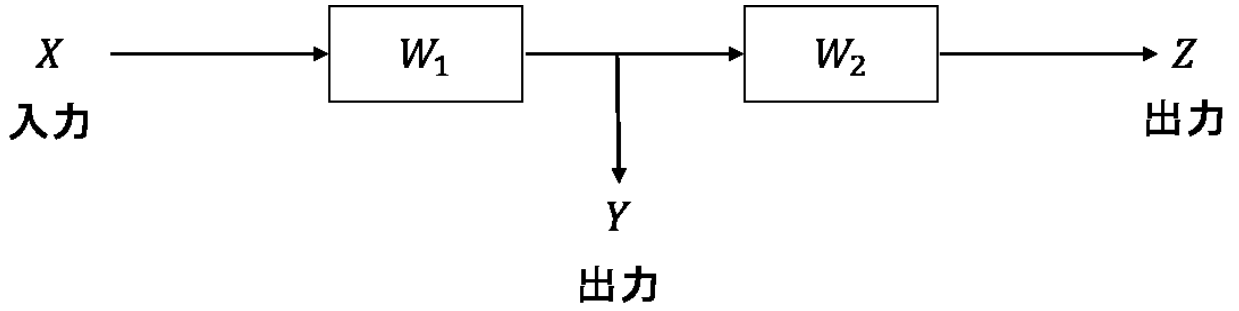


図 3.2: 劣化型対称放送通信路

DBC の容量域を $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ と表す. DBC の容量域を記述するために有限集合 \mathcal{U} に値をとる確率変数 U を導入し, (U, X, Y, Z) の同時分布が

$$p_{UXYZ}(u, x, y, z) = \{p_{UXYZ}(u, x, y, z)\}_{(u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} \quad (3.2)$$

で与えられる場合を考える. 式 (3.2) の条件は (U, X, Y, Z) がマルコフ連鎖 $U \rightarrow X \rightarrow Y \rightarrow Z$ をなすことと等価である. 確率変数 $X \in \mathcal{X}, Y \in \mathcal{Y}, Z \in \mathcal{Z}$ が (3.2) に従うときの条件付き相互情報量 $I_p(X; Y|U)$, $I_p(Z; U)$ と表す. すなわち,

$$I_p(X; Y|U) = \sum_{u, x, y} p_{UXY}(u, x, y) \log \frac{W_1(y|x)}{p_{Y|U}(y|u)}$$

$$I_p(Z; U) = \sum_{u, z} p_{UZ}(u, z) \log \frac{p_{Z|U}(z|u)}{p_Z(z)}$$

(U, X, Y, Z) の同時分布 $p = p_{UXYZ}$ の集合 $\mathcal{P}(W_1, W_2)$ を以下で定義する.

$$\mathcal{P}(W_1, W_2) \triangleq \left\{ p = p_{UXYZ} : p(u, x, y, z) = p(u, x)W_1(y|x)W_2(z|y), |\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|\} \right\}$$

このとき $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ は以下で与えられる. [7]

$$\mathcal{C}_{\text{DBC}}(W_1, W_2) = \left\{ (R_1, R_2) : R_1 \leq I_p(X; Y|U), R_2 \leq I_p(Z; U), \exists p \in \mathcal{P}(W_1, W_2) \right\}$$

本研究では容量域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ を数値的に計算するアルゴリズムを考察する. そのために $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の支持直線による表現を導出する. 次の節では支持直線について説明する.

3.2 容量域の支持直線表現

支持直線を導出するために, まず次の量を定義する.

$$C^{(\mu)}(W_1, W_2) \triangleq \max_{p \in \mathcal{P}(W_1, W_2)} \left\{ I_p(X; Y|U) + \mu I_p(Z; U) \right\} \quad (3.3)$$

このとき次の命題が成り立つ.

命題3

$$\mathcal{C}_{\text{DBC}}(W_1, W_2) = \bigcap_{\mu \geq 0} \left\{ (R_1, R_2) : R_1 + \mu R_2 \leq C^{(\mu)}(W_1, W_2) \right\}$$

命題3より、容量域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ は μ を動かしたときに得られる全ての支持直線の下側の領域となっている。支持直線は $C^{(\mu)}(W_1, W_2)$ を求めることで導出することができる。つまり容量域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ を数値的に求める問題は $C^{(\mu)}(W_1, W_2)$ を数値的に求める問題に帰着される。直線 $R_1 + \mu R_2 = C^{(\mu)}(W_1, W_2)$ は領域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の支持直線とよばれる。容量域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ とその支持直線を図3.3に示す。この図において直線 $R_1 + \mu R_2 = I_p(X; Y|U) + \mu I_p(Z; U)$ は $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ 内の点 $(I_p(X; Y|U), I_p(Z; U))$, $p \in \mathcal{P}(W_1, W_2)$ を通り、法線ベクトルを $(1, \mu)$ とする直線である。この図から分かる様に、支持直線 $R_1 + \mu R_2 = C^{(\mu)}(W_1, W_2)$ は、法線ベクトルを $(1, \mu)$ とする直線で領域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ と共有点をもつようなもののうち、その R_1 切片が最大となるものに等しい。以下では $C^{(\mu)}(W_1, W_2)$ を数値的に求める問

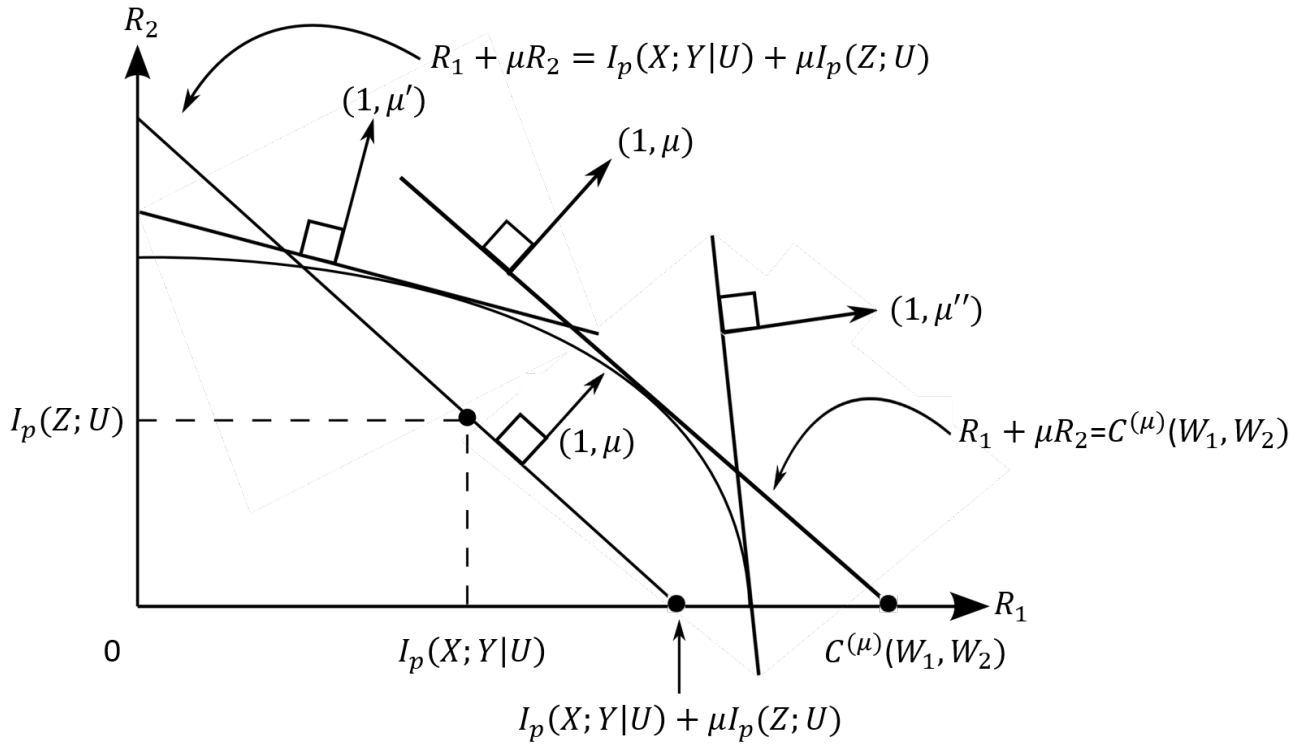


図 3.3: $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ とその支持直線

題を考察する。この問題は文献 [4]-[6] で考察され、単一通信路に対する有本 [1], Blahut [2] のアルゴリズムを DBC に拡張したアルゴリズムが提案されている。この拡張アルゴリズムが定める分布列は領域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の内点を与える分布から、その境界へ近づく性質をもつ。文献 [4]-[6] では分布列 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の境界を与える分布への収束について議論されている。次の節では DBC に関するこれまでの研究について記す。

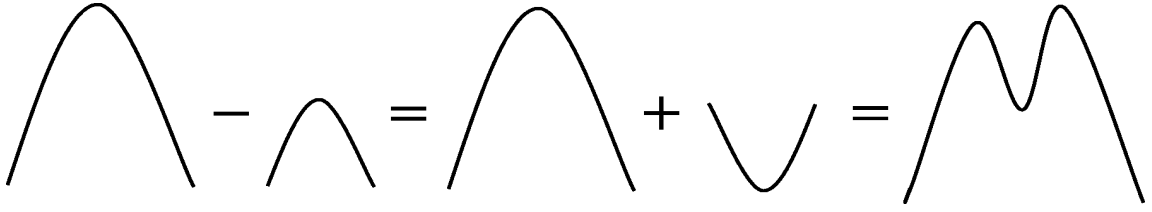


図 3.4: 差凸関数のイメージ

3.3 DBCに関するこれまでの研究

この節では DBC の容量域計算に関するこれまでの結果を述べる。

Calvo et. Al '2008[5] の研究

Calvo[5] は式 (3.3) で定義される $C^{(\mu)}(W_1, W_2)$, 即ち

$$C^{(\mu)}(W_1, W_2) = \max_{p \in \mathcal{P}(W_1, W_2)} \left\{ I_p(X; Y|U) + \mu I_p(Z; U) \right\}$$

を求める問題を考察し, DBC の容量域計算問題が差凸関数を目的関数とする非凸最適化問題となることを示した. 実際,

$$\begin{aligned} I_p(X; Y|U) + \mu I_p(Z; U) &= H_p(Y|U) - H_p(Y|X) + \mu \{H(Z) - H(Z|U)\} \\ &= H_p(Y|U) - H_p(Y|X) + \mu H(Z) - \mu H(Z|U) \end{aligned} \quad (3.4)$$

について考察した. (3.4) の $H_p(Y|U)$ は,

$$H_p(Y|U) = - \sum_{u,x,y} p_{UX}(u, x) W_1(y|x) \log \left[\sum_x p_{X|U}(x|u) W_1(y|x) \right]$$

よって, $H_p(Y|U)$ は分布 p_{UX} について上に凸である. 同様に $H(Z), H(Z|U)$ も分布 p_{UX} について上に凸である.

$$\begin{aligned} H(Z) &= - \sum_{u,x,y,z} p_{UX}(u, x) W_1(y|x) W_2(z|y) \log \left[\sum_{u,x,y} p_{UX}(u, x) W_1(y|x) W_2(z|y) \right] \\ H(Z|U) &= - \sum_{u,x,y,z} p_{UX}(u, x) W_1(y|x) W_2(z|y) \log \left[\sum_{x,y} p_{X|U}(x|u) W_1(y|x) W_2(z|y) \right] \end{aligned}$$

また, $H_p(Y|X)$ は,

$$H_p(Y|X) = - \sum_{u,x,y} p_{UX}(u, x) W_1(y|x) \log W_1(y|x)$$

分布 p_{UX} についての線形関数である。このことから式 (3.4) で与えられる目的関数は、凸関数 $H_p(Y|U) - H_p(Y|X) + \mu H(Z)$ と、凸関数 $\mu H(Z|U)$ の差で表される差凸関数となることがわかる。

Yasui and Matsushima '2010[6]の研究:

Yasui and Matsushima [6] は $C^{(\mu)}(W_1, W_2)$ の定める最適確率分布を求める反復型分布更新アルゴリズムを提案した。提案したアルゴリズムは Arimoto-Blahut アルゴリズムを DBC へ拡張したものになっている。そして、初期分布に関するある条件のもとで、提案アルゴリズムの定める分布列が最適分布へ収束することを示した。第4章では Arimoto-Blahut アルゴリズムを DBC へ自然な形で拡張したアルゴリズムが提案されている。しかし、このアルゴリズムでは最適分布への収束について非常に厳しい条件がつき、最適分布への収束が保証されないことがわかる。Yasui and Matsushima [6] は、 $C^{(\mu)}(W_1, W_2)$ を求める最適化問題の目的関数の変数の取り方に、ある変形を加え、さらに変形された目的関数に対する Arimoto-Blahut アルゴリズムを提案し、最適分布への収束条件を示した。

そこで本論文では、 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ を数値的に求める新しい逐次計算アルゴリズムを提案する。提案アルゴリズムについては次の章で述べる。

第 4 章

提案アルゴリズム

本章では DBC の通信路容量域を求める新しい反復アルゴリズムを提案する。提案アルゴリズムを記述する前に、2 端子通信路に対する Arimoto-Blahut アルゴリズムの自然な拡張として得られるアルゴリズムについて記述し、このアルゴリズムの問題点を述べる。

4.1 Arimoto-Blahut アルゴリズムの拡張

Arimoto-Blahut アルゴリズムを自然に拡張し、DBC の通信路容量を求める。通信路容量 $C^{(\mu)}(W_1, W_2)$ は

$$\begin{aligned} C^{(\mu)}(W_1, W_2) &= \max_{p: p=(p_{UX}, W_1, W_2)} \left\{ I_p(X; Y|U) + \mu I_p(U; Z) \right\} \\ &= \max_{p_{UX}} \left\{ I_{p_{UX}}(X; Y|U) + \mu I_{p_{UX}}(U; Z) \right\} \end{aligned}$$

ここで $p = (p_{UX}, W_1, W_2)$ について、 W_1, W_2 は固定されている事を明示するために

$$I_p(X; Y|U) = I_{p_{UX}}(X; Y|U)$$

と記した。 $C^{(\mu)}(W_1, W_2)$ を求めるために目的関数を定義する。

$$\begin{aligned} F(p_{UX}, q_{UX}) &\triangleq I_{q_{UX}}(X; Y|U) + D(q_{Y|U} \| p_{Y|U} | p_U) - D(q_{X|U} \| p_{X|U}) \\ &\quad + \mu I_{q_{UX}}(U; Z) + \mu D(q_Z \| p_Z) - \mu D(q_{Z|U} \| p_{Z|U} | p_U) - \mu D(q_U \| p_U) \quad (4.1) \end{aligned}$$

このとき次の補題が成り立つ。

補題 3 固定した q_{UX} に対し、 $F(p_{UX}, q_{UX})$ は $p_{UX} = q_{UX}$ のとき最大値

$$F(q_{UX}, q_{UX}) = I_{q_{UX}}(X; Y|U) + \mu I_{q_{UX}}(U; Z)$$

をとる。

(証明) ダイバージェンスに関する性質により

$$D(q_{Y|U} \| p_{Y|U}) \leq D(q_{X|U} \| p_{X|U}) \quad (4.2)$$

$$D(q_{Z|U} \| p_{Z|U}) \leq D(q_{Z|XU} \| p_{Z|XU}) \quad (4.3)$$

が成り立つ。詳細はそれぞれ付録の定理 B および定理 A を参照せよ。(4.2),(4.3) より、(4.1) は

$$\begin{aligned} F(p_{UX}, q_{UX}) &= I_{q_{UX}}(X; Y|U) + D(q_{Y|U} \| p_{Y|U} | p_U) - D(q_{X|U} \| p_{X|U}) \\ &\quad + \mu I_{q_{UX}}(U; Z) + \mu D(q_Z \| p_Z) - \mu D(q_{Z|U} \| p_{Z|U} | p_U) - \mu D(q_U \| p_U) \\ &\leq I_{q_{UX}}(X; Y|U) + \mu I_{q_{UX}}(U; Z) \end{aligned}$$

が成り立つ。等号は $p_{UX} = q_{UX}$ のとき成り立つ。 ■

同様に固定した p_{UX} に対し、 $F(p_{UX}, q_{UX})$ の最大値について、次の補題が成り立つ。

補題 4 固定した p_{UX} に対し、 $F(p_{UX}, q_{UX})$ は、 $q_{UX}(u, x)$ が

$$q_{UX}(u, x) = \frac{1}{K} \prod_y \left[\frac{W_1(y|x)}{p_{Y|U}(y|u)} \right]^{W_1(y|x)} \times \prod_z \left[\frac{p_{Z|U}(z|u)}{p_Z(z)} \right]^{\mu(W_1 W_2)(z|x)} \times p_{UX}(u, x) \quad (4.4)$$

のときに最大値 $\log K$ をとる。ここで K は規格化定数であり、以下で与えられる。

$$K = \sum_{u,x} \exp \left\{ \sum_y W_1(y|x) \log \frac{W_1(y|x)}{p_{Y|U}(y|u)} + \sum_z (W_1 W_2)(z|x) \log \frac{p_{Z|U}(z|u)}{p_Z(z)} \right\} p_{UX}(u, x) \quad (4.5)$$

(証明) $F(p_{UX}, q_{UX})$ の具体的な形は以下で与えられる。

$$\begin{aligned} F(p_{UX}, q_{UX}) &= \sum_{u,x} q_{UX}(u, x) \left[\sum_y W_1(y|x) \log \frac{W_1(y|x)}{p_{Y|U}(y|u)} \right. \\ &\quad \left. + \mu \left\{ \sum_z (W_1 W_2)(z|x) \log \frac{p_{Z|U}(z|u)}{p_Z(z)} \right\} \right] + \sum_{u,x} q_{UX}(u, x) \log \frac{p_{UX}(u, x)}{q_{UX}(u, x)} \quad (4.6) \end{aligned}$$

ここで上式の

$$A_{p_{UX}}(u, x) = \sum_y W_1(y|x) \log \frac{W_1(y|x)}{p_{Y|U}(y|u)} \quad (4.7)$$

$$B_{p_{UX}}(u, x) = \sum_z (W_1 W_2)(z|x) \log \frac{p_{Z|U}(z|u)}{p_Z(z)} \quad (4.8)$$

のようにおくと、下のように書き換えることができる。

$$\begin{aligned} F(p_{UX}, q_{UX}) &= \sum_{u,x} q_{UX}(u, x) \left[A_{p_{UX}}(u, x) + \mu B_{p_{UX}}(u, x) \right] + \sum_{u,x} q_{UX}(u, x) \log \frac{p_{UX}(u, x)}{q_{UX}(u, x)} \\ &= - \sum_{u,x} q_U(u) q_{X|U}(x|u) \log \frac{q_{UX}(u, x)}{\exp \left\{ A_{p_{UX}}(u, x) + \mu B_{p_{UX}}(u, x) \right\} p_{UX}(u, x)} \quad (4.9) \end{aligned}$$

分母を確率分布にするために、分母に $\frac{1}{K}$ をかける。

$$\begin{aligned} & F(p_{UX}, q_{UX}) \\ &= - \sum_{u,x} q_U(u) q_{X|U}(x|u) \log \frac{q_{UX}(u, x)}{\frac{1}{K} \exp\{A_{p_{UX}}(u, x) + \mu B_{p_{UX}}(u, x)\} p_{UX}(u, x)} + \log K \\ &\leq \log K \end{aligned} \quad (4.10)$$

等号は、

$$q_{UX}(u, x) = \frac{1}{K} \exp\{A_{p_{UX}}(u, x) + \mu B_{p_{UX}}(u, x)\} p_{UX}(u, x) \quad (4.11)$$

のときに成り立つ。 ■

補題に基づき、 $C^{(\mu)}(W_1, W_2)$ を求めるアルゴリズムを導出する。 $t = 0, 1, 2, \dots$ に対し、 $\mathcal{U} \times \mathcal{X}$ 上の分布の列

$$q^{[t]} = \{q^{[t]}(u, x)\}_{(u,x) \in \mathcal{U} \times \mathcal{X}} \quad (4.12)$$

を次の様に定める。

分布更新アルゴリズム

1. $q^{[0]}$ を $\mathcal{U} \times \mathcal{X}$ 上の適当な分布にとる。
2. $t = 0, 1, 2, \dots$ に対する分布更新式を以下で定める。

$$q^{[t+1]}(u, x) = \frac{1}{K_t} \exp\{A_{q^{[t]}}(u, x) + \mu B_{q^{[t]}}(u, x)\} q^{[t]}(u, x) \quad (4.13)$$

ここで K_t は規格化定数であり、以下で与えられる。

$$K_t = \sum_{u,x} \exp\{A_{q^{[t]}}(u, x) + \mu B_{q^{[t]}}(u, x)\} q^{[t]}(u, x) \quad (4.14)$$

上記の分布更新アルゴリズムに対して、次の命題が成り立つ。

命題 4 $t = 0, 1, 2, \dots$ に対し、次の不等式が成り立つ。

$$\begin{aligned} & F(q^{[0]}, q^{[0]}) \\ &\stackrel{(a)}{\leq} F(q^{[0]}, q^{[1]}) \stackrel{(b)}{\leq} F(q^{[1]}, q^{[1]}) \leq \dots \\ &\stackrel{(a)}{\leq} F(q^{[t-1]}, q^{[t]}) \\ &\stackrel{(b)}{\leq} F(q^{[t]}, q^{[t]}) \end{aligned} \quad (4.15)$$

$$\stackrel{(a)}{\leq} F(q^{[t]}, q^{[t+1]}) \quad (4.16)$$

$$\stackrel{(b)}{\leq} F(q^{[t+1]}, q^{[t+1]}) \leq \dots$$

$$\leq C^{(\mu)}(W_1, W_2)$$

ここで不等式 (4.15), (4.16) の右辺に現れる量に関しては以下が成り立つ.

$$F(q^{[t]}, q^{[t]}) = I_{q^{[t]}}(X; Y|U) + \mu I_{q^{[t]}}(U; Z) \quad (4.17)$$

$$\begin{aligned} F(q^{[t]}, q^{[t+1]}) &= \log K_t \\ &= A_{q^{[t]}}(u, x) + \mu B_{q^{[t]}}(u, x) + \log \frac{q^{[t]}}{q^{[t+1]}} \end{aligned} \quad (4.18)$$

(証明) 命題6の不等式におけるステップ(a)と式(4.18)は補題4より従う. またステップ(b)と式(4.17)は補題3より従う. ■

次節に, 拡張 Arimoto-Blahut アルゴリズムが, 最適値を与える分布へ収束するかについて記述する.

4.2 最適分布への収束

$F(p_{UX}, q_{UX})$ の最適値を与える分布を $q^*(u, x)$ とすると, 補題3から

$$\begin{aligned} C^{(\mu)}(W_1, W_2) &= \max_{p_{UX}} \left\{ I_{p_{UX}}(X; Y|U) + \mu I_{p_{UX}}(U; Z) \right\} \\ &= F(q^*, q^*) \\ &= I_{q^*}(X; Y|U) + \mu I_{q^*}(U; Z) \end{aligned} \quad (4.19)$$

収束性について次の命題が成り立つ.

命題5 $t = 0, 1, 2, \dots$ に対し, 以下が成り立つ.

$$\begin{aligned} &C^{(\mu)}(W_1, W_2) - F(q^{[t]}, q^{[t+1]}) \\ &= F(q^*, q^*) - F(q^{[t]}, q^{[t+1]}) \\ &= \sum_{u,x} q^*(u, x) \left[A_{q^*}(u, x) - A_{q^{[t]}}(u, x) + \mu \left\{ B_{q^*}(u, x) - B_{q^{[t]}}(u, x) \right\} \right] \\ &\quad + \sum_{u,x} q^*(u, x) \log \frac{q^{[t+1]}(u, x)}{q^{[t]}(u, x)} \\ &= -D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) - \mu D(q_Z^* \| q_Z^{[t]}) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]}) \\ &\quad + D(q^* \| q^{[t]}) - D(q^* \| q^{[t+1]}) \end{aligned} \quad (4.20)$$

もし全ての $t = 0, 1, 2, \dots$ について

$$-D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) - \mu D(q_Z^* \| q_Z^{[t]}) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]}) \leq 0 \quad (4.21)$$

が成り立つならば、拡張 Arimoto-Blahut アルゴリズムの定める分布列は最適分布へ収束する。しかし、

$$D(q_{Z|U}^* \| q_{Z|U}^{[l]}) - D(q_Z^* \| q_Z^{[l]}) \geq 0 \quad (4.22)$$

であるから、 μ が大きいとき成立が非常に難しくなる。つまり、Arimoto-Blahut アルゴリズムの拡張では、DBC の通信路容量を導出するのが困難となる。そこで、目的関数にペナルティ項を導入することで、この問題を解決する新しい提案アルゴリズムを次の節で記述する。

4.3 提案する新しい反復アルゴリズム

この節では新しい逐次計算アルゴリズムを提案する。まず初めに、 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ を数値的に求めるために次の量を定義する。

$$\omega_q^{(\alpha, \mu)}(u, x, y, z) \triangleq \log \frac{W_1(y|x)}{q_{Y|U}(y|u)} + \mu \log \frac{q_{Z|U}(z|u)}{q_Z(z)} + \alpha \log \frac{W_2(z|y)W_1(y|x)}{q_{ZY|XU}(z, y|x, u)} \quad (4.23)$$

$$F^{(\alpha, \mu)}(p, q) \triangleq (1 + \alpha) \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1 + \alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z) + \log \frac{p(u, x, y, z)}{q(u, x, y, z)} \right\} \quad (4.24)$$

関数 $F^{(\alpha, \mu)}(p, q)$ は、 (α, μ) をパラメータとする2つの分布 p, q の関数である。また、本論文ではパラメータ (α, μ) が以下の条件

$$0 \leq \mu \leq 1 + \alpha \quad (4.25)$$

を満たすと仮定して議論する。ここで次の補題が成り立つ。

補題 5 条件 (4.25) の仮定の下で、固定した q に対し、 $F^{(\alpha, \mu)}(p, q)$ は、 $p = q$ のとき最大値

$$\begin{aligned} F^{(\alpha, \mu)}(q, q) &= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \omega_q^{(\alpha, \mu)}(u, x, y, z) \right\} \\ &= I_q(X; Y|U) + \mu I_q(Z; U) - \alpha D(q_{ZY|XU} \| W_1, W_2 | q_{XU}) - D(q_{Y|XU} \| W_1 | q_{XU}) \end{aligned} \quad (4.26)$$

をとる。

(証明) $F^{(\alpha, \mu)}(p, q)$ に $\frac{1}{1+\alpha}$ をかけた量について、以下の一連の等式が成り立つ。

$$\begin{aligned}
& \frac{1}{1+\alpha} F^{(\alpha, \mu)}(p, q) \\
&= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1+\alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z) + \log \frac{p(u, x, y, z)}{q(u, x, y, z)} \right\} \\
&= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1+\alpha} \omega_q^{(\alpha, \mu)}(u, x, y, z) \right. \\
&\quad \left. + \frac{1}{1+\alpha} \left[\omega_p^{(\alpha, \mu)}(u, x, y, z) - \omega_q^{(\alpha, \mu)}(u, x, y, z) \right] + \log \frac{p(u, x, y, z)}{q(u, x, y, z)} \right\} \\
&= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1+\alpha} \omega_q^{(\alpha, \mu)}(u, x, y, z) \right. \\
&\quad \left. + \frac{1}{1+\alpha} \left[\log \frac{W_1(y|x)}{p_{Y|U}(y|u)} + \mu \log \frac{p_{Z|U}(z|u)}{p_Z(z)} + \alpha \log \frac{W_2(z|y)W_1(y|x)}{p_{ZY|XU}(z, y|x, u)} \right. \right. \\
&\quad \left. \left. - \log \frac{W_1(y|x)}{q_{Y|U}(y|u)} - \mu \log \frac{q_{Z|U}(z|u)}{q_Z(z)} - \alpha \log \frac{W_2(z|y)W_1(y|x)}{q_{ZY|XU}(z, y|x, u)} \right] + \log \frac{p(u, x, y, z)}{q(u, x, y, z)} \right\} \\
&= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1+\alpha} \omega_q^{(\alpha, \mu)}(u, x, y, z) + \frac{\alpha}{1+\alpha} \log \frac{q(y, z|x, u)}{p(y, z|x, u)} \right. \\
&\quad \left. + \frac{1}{1+\alpha} \log \frac{q(y|u)}{p(y|u)} + \frac{\mu}{1+\alpha} \log \frac{q(z)}{p(z)} + \frac{\mu}{1+\alpha} \log \frac{p(z|u)}{q(z|u)} + \log \frac{p(u, x, y, z)}{q(u, x, y, z)} \right\} \\
&= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{1}{1+\alpha} \omega_q^{(\alpha, \mu)}(u, x, y, z) \right\} + J_1 \tag{4.27}
\end{aligned}$$

ここで J_1 は以下で与えられる。

$$\begin{aligned}
J_1 &= \sum_{u, x, y, z} q(u, x, y, z) \left\{ \frac{\alpha}{1+\alpha} \log \frac{p(x|u)}{q(x|u)} + \frac{1}{1+\alpha} \log \frac{p(x, z|u, y)}{q(x, z|u, y)} + \frac{\mu}{1+\alpha} \log \frac{p(u|z)}{q(u|z)} \right. \\
&\quad \left. + \frac{1}{1+\alpha} [1 + \alpha - \mu] \log \frac{p(u)}{q(u)} \right\}
\end{aligned}$$

量 J_1 については以下の条件

$$0 \leq \mu \leq 1 + \alpha \tag{4.28}$$

のもとで $J_1 \leq 0$ が成り立つ。よって (4.28) の条件のもとで $\frac{1}{1+\alpha} F^{(\alpha, \mu)}(p, q)$ は、 $p = q$ のとき最大値をとる。 ■

補題5の式 (4.26) の、 $-\alpha D(q_{ZY|XU} \| W_1, W_2 | q_{XU}) - D(q_{Y|XU} \| W_1 | q_{XU})$ はペナルティ項となっている。次に以下の量を定義する。

$$C^{(\alpha, \mu)}(W_1, W_2) \triangleq \max_{(p, q)} F^{(\alpha, \mu)}(p, q) \tag{4.29}$$

このとき次の補題が成り立つ。

補題6

$$\begin{aligned}
C^{(\alpha,\mu)}(W_1, W_2) &= \max_q F^{(\alpha,\mu)}(q, q) \\
&= \max_q \left\{ I_q(X; Y|U) + \mu I_q(Z; U) \right. \\
&\quad \left. - \alpha D(q_{ZY|XU} \| W_1, W_2 | q_{XU}) - D(q_{Y|XU} \| W_1 | q_{XU}) \right\} \quad (4.30)
\end{aligned}$$

$C^{(\alpha,\mu)}(W_1, W_2)$ は $\alpha > 0$ に関して単調減少であり, 任意の $\alpha > 0$ に対し

$$C^{(\mu)}(W_1, W_2) \leq C^{(\alpha,\mu)}(W_1, W_2) \quad (4.31)$$

が成り立つ. 特に

$$\lim_{\alpha \rightarrow +\infty} C^{(\alpha,\mu)}(W_1, W_2) = C^{(\mu)}(W_1, W_2) \quad (4.32)$$

(証明) まず式(4.30)を示す. 補題1より, 固定した q に対し, $F^{(\alpha,\mu)}(p, q)$ は $p = q$ のときに最大値

$$\begin{aligned}
&F^{(\alpha,\mu)}(q, q) \\
&= \sum_{u,x,y,z} q(u, x, y, z) \{ \omega_q^{(\alpha,\mu)}(u, x, y, z) \} \\
&= I_q(X; Y|U) + \mu I_q(Z; U) + \alpha \sum_{u,x,y,z} q(u, x, y, z) \log \frac{W_2(z|y)W_1(y|x)}{q_{ZY|XU}(z, y|x, u)} \\
&= I_q(X; Y|U) + \mu I_q(Z; U) - \alpha D(q_{ZY|XU} \| W_1, W_2 | q_{XU}) - D(q_{Y|XU} \| W_1 | q_{XU})
\end{aligned}$$

をとる. よって式(4.30)が示された. 次に式(4.31), (4.32)を示す.

$$q^* = q_\alpha^* = \{q_\alpha(u, x, y, z)\}_{(u,x,y,z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} \quad (4.33)$$

を $C^{(\alpha,\mu)}(W_1, W_2)$ を達成する分布とすると

$$\begin{aligned}
C^{(\alpha,\mu)}(W_1, W_2) &= F^{(\alpha,\mu)}(q_\alpha^*, q_\alpha^*) \\
&= I_{q_\alpha^*}(X; Y|U) + \mu I_{q_\alpha^*}(Z; U) \\
&\quad - \alpha D(q_{ZY|XU,\alpha}^* \| W_1, W_2 | q_{XU,\alpha}^*) - D(q_{Y|XU,\alpha}^* \| W_1 | q_{XU,\alpha}^*) \quad (4.34)
\end{aligned}$$

が成り立つ. ここで, $q_{ZY|XU,\alpha}^*$, $q_{Y|XU,\alpha}^*$ および $q_{XU,\alpha}^*$ は分布 $q_\alpha^* = q_{UXYZ,\alpha}^*$ より誘導される条件分布である.

$$\Delta^{(\alpha,\mu)} \triangleq I_{q_\alpha^*}(X; Y|U) + \mu I_{q_\alpha^*}(Z; U) - C^{(\alpha,\mu)}(W_1, W_2)$$

とおくと、式(4.34)より

$$0 \leq D(q_{ZY|XU,\alpha}^* \| W_1, W_2 | q_{XU,\alpha}^*) = \frac{\Delta^{(\alpha,\mu)} - D(q_{Y|XU,\alpha}^* \| W_1 | q_{XU,\alpha}^*)}{\alpha} \leq \frac{\Delta^{(\alpha,\mu)}}{\alpha} \quad (4.35)$$

これより、 $\Delta^{(\alpha,\mu)} \geq 0$ であることがわかる。よって

$$C^{(\alpha,\mu)}(W_1, W_2) \leq I_{q_\alpha^*}(X; Y|U) + \mu I_{q_\alpha^*}(Z; U) \quad (4.36)$$

分布 $\hat{q}_\alpha = \hat{q}_{UXYZ,\alpha}$ を

$$\hat{q}_{UXYZ,\alpha}(u, x, y, z) = q_{XU,\alpha}^* W_1(y|x) W_2(z|y)$$

と定めると $\hat{q}_\alpha \in \mathcal{P}(W_1, W_2)$ であり、以下が成り立つ。

$$\begin{aligned} D(q_\alpha^* \| \hat{q}_\alpha) &= D(q_{UXYZ,\alpha}^* \| \hat{q}_{UXYZ,\alpha}) \\ &= D(q_{ZY|XU,\alpha}^* \| W_1, W_2) = \frac{\Delta^{(\alpha,\mu)}}{\alpha} \end{aligned}$$

$\alpha \rightarrow +\infty$ とすると $D(q_\alpha^* \| \hat{q}_\alpha) \rightarrow 0$ が成り立つ。すると $I_q(X; Y|Z)$ および $I_q(Z; U)$ の分布 q に関する連続性より

$$I_{q_\alpha^*}(X; Y|U) + \mu I_{q_\alpha^*}(Z; U) \leq I_{\hat{q}_\alpha}(X; Y|U) + \mu I_{\hat{q}_\alpha}(Z; U) + \tau(\alpha) \quad (4.37)$$

が成り立つ。ここで $\tau(\alpha)$ は $\lim_{\alpha \rightarrow +\infty} \tau(\alpha) = 0$ となる量である。このとき以下の一連の不等式が成り立つ。

$$\begin{aligned} C^{(\alpha,\mu)}(W_1, W_2) &\stackrel{(a)}{\leq} I_{q_\alpha^*}(X; Y|U) + \mu I_{q_\alpha^*}(Z; U) \\ &\stackrel{(b)}{\leq} I_{\hat{q}_\alpha}(X; Y|U) + \mu I_{\hat{q}_\alpha}(Z; U) + \tau(\alpha) \\ &\stackrel{(c)}{\leq} C^{(\mu)}(W_1, W_2) + \tau(\alpha) \end{aligned} \quad (4.38)$$

ここでステップ(a)は(4.36)より従う。またステップ(b)は(4.37)より従う。また、ステップ(c)は $\hat{q}_\alpha \in \mathcal{P}(W_1, W_2)$ と $C^{(\mu)}(W_1, W_2)$ の定義より従う。他方、 $C^{(\mu)}(W_1, W_2)$ を達成する分布を \tilde{q}_α とすると、 $\tilde{q}_\alpha \in \mathcal{P}(W_1, W_2)$ であることより、任意の $(u, x, y, z) \in (\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ に対し

$$\begin{aligned} \tilde{q}_{Y|XU,\alpha}(y|x, u) &= W_1(y|x), \\ \tilde{q}_{ZY|XU,\alpha}(z, y|x, u) &= W_2(z|y) W_1(y|x) \end{aligned}$$

が成り立つ。よって

$$D(\tilde{q}_{ZY|XU,\alpha} \| W_1, W_2 | \tilde{q}_{XU,\alpha}) = D(\tilde{q}_{Y|XU,\alpha} \| W_1 | \tilde{q}_{XU,\alpha}) = 0$$

このとき

$$\begin{aligned}
C^{(\mu)}(W_1, W_2) &= I_{\tilde{q}_\alpha}(X; Y|U) + \mu I_{\tilde{q}_\alpha}(Z; U) \\
&= I_{\tilde{q}_\alpha}(X; Y|U) + \mu I_{\tilde{q}_\alpha}(Z; U) \\
&\quad - \alpha D(\tilde{q}_{ZY|XU, \alpha} \| W_1, W_2 | \tilde{q}_{XU, \alpha}) - D(\tilde{q}_{Y|XU, \alpha} \| W_1 | \tilde{q}_{XU, \alpha}) \\
&\leq C^{(\alpha, \mu)}(W_1, W_2)
\end{aligned} \tag{4.39}$$

が成り立つ。式(4.38), (4.39)より以下を得る。

$$0 \leq C^{(\alpha, \mu)}(W_1, W_2) - C^{(\mu)}(W_1, W_2) \leq \tau(\alpha)$$

上式で $\alpha \rightarrow +\infty$ とすると式(4.32)を得る。 ■

補題6より、 $C^{(\mu)}(W_1, W_2)$ を求める問題は十分大きな α に対する $F^{(\alpha, \mu)}(p, q)$ の最大化問題に帰着されることになる。

$C^{(\mu)}(W_1, W_2)$ を求める方法

- α を十分大きい値に固定する。
- $F^{(\alpha, \mu)}(p, q)$ を大きくするように順次 p, q を更新していく。

以下ではこの最適化問題を考察する。固定した p に対する $F^{(\alpha, \mu)}(p, q)$ の最大値について、次の補題が成り立つ。

補題7 固定した p に対し、 $F^{(\alpha, \mu)}(p, q)$ は、

$$q(u, x, y, z) = \frac{1}{K} \exp\left\{\frac{1}{1+\alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z)\right\} p(u, x, y, z)$$

のとき最大値 $(1+\alpha) \log K$ をとる。ここで K は規格化定数であり、以下で与えられる。

$$K = \sum_{(u, x, y, z)} \exp\left\{\frac{1}{1+\alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z)\right\} p(u, x, y, z)$$

(証明) $\frac{1}{1+\alpha} F^{(\alpha, \mu)}(p, q)$ の具体的な形は以下で与えられる。

$$\frac{1}{1+\alpha} F^{(\alpha, \mu)}(p, q) = \sum_{u, x, y, z} q(u, x, y, z) \log \frac{\exp\left\{\frac{1}{1+\alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z)\right\} p(u, x, y, z)}{q(u, x, y, z)} \tag{4.40}$$

式(4.40)の分子を確率分布にするために規格化定数

$$K = \sum_{(u, x, y, z)} \exp\left\{\frac{1}{1+\alpha} \omega_p^{(\alpha, \mu)}(u, x, y, z)\right\} p(u, x, y, z)$$

を考え、分子に $\frac{1}{K}$ をかける。すると、次の等式を得る。

$$\begin{aligned} & \frac{1}{1+\alpha} F^{(\alpha,\mu)}(p, q) \\ = & \sum_{u,x,y,z} q(u, x, y, z) \log \frac{\frac{1}{K} \exp\{\frac{1}{1+\alpha} \omega_p^{(\alpha,\mu)}(u, x, y, z)\} p(u, x, y, z)}{q(u, x, y, z)} + \log K \end{aligned} \quad (4.41)$$

式(4.41)の右辺第1項について

$$\sum_{u,x,y,z} q(u, x, y, z) \log \frac{\frac{1}{K} \exp\{\frac{1}{1+\alpha} \omega_p^{(\alpha,\mu)}(u, x, y, z)\} p(u, x, y, z)}{q(u, x, y, z)} \leq 0$$

であるから、この式と式(4.41)より以下を得る。

$$F^{(\alpha,\mu)}(p, q) \leq (1+\alpha) \log K \quad (4.42)$$

式(4.42)により、等号は

$$q(u, x, y, z) = \frac{1}{K} \exp\left\{\frac{1}{1+\alpha} \omega_p^{(\alpha,\mu)}(u, x, y, z)\right\} p(u, x, y, z) \quad (4.43)$$

のとき成り立つ。 ■

補題5-7に基づき、十分大きな α に対する $C^{(\alpha,\mu)}(W_1, W_2)$ 求めるアルゴリズムを導出する。

$t = 0, 1, 2, \dots$ に対し $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ 上の分布の列

$$q^{[t]} = \{q^{[t]}(u, x, y, z)\}_{(u,x,y,z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}$$

を次の様に定める。

分布更新アルゴリズム

1. $q^{[0]}$ を $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ 上の適当な分布にとる。
2. $t = 0, 1, 2, \dots$ に対する分布更新式を以下で定める。

$$q^{[t+1]}(u, x, y, z) = \frac{1}{K_t} \exp\left\{\frac{1}{1+\alpha} \omega_{q^{[t]}}^{(\alpha,\mu)}(u, x, y, z)\right\} q^{[t]}(u, x, y, z) \quad (4.44)$$

ここで K_t は規格化定数であり、以下で与えられる。

$$K_t = \sum_{(u,x,y,z)} \exp\left\{\frac{1}{1+\alpha} \omega_{q^{[t]}}^{(\alpha,\mu)}(u, x, y, z)\right\} q^{[t]}(u, x, y, z)$$

上記の分布更新アルゴリズムに対して、次の命題が成り立つ。

命題 6 $t = 0, 1, 2, \dots$ に対し, 次の不等式が成り立つ.

$$\begin{aligned}
& F^{(\alpha, \mu)}(q^{[0]}, q^{[0]}) \\
& \stackrel{(a)}{\leq} F^{(\alpha, \mu)}(q^{[0]}, q^{[1]}) \stackrel{(b)}{\leq} F^{(\alpha, \mu)}(q^{[1]}, q^{[1]}) \leq \dots \\
& \stackrel{(a)}{\leq} F^{(\alpha, \mu)}(q^{[t-1]}, q^{[t]}) \\
& \stackrel{(b)}{\leq} F^{(\alpha, \mu)}(q^{[t]}, q^{[t]}) \tag{4.45}
\end{aligned}$$

$$\stackrel{(a)}{\leq} F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \tag{4.46}$$

$$\stackrel{(b)}{\leq} F^{(\alpha, \mu)}(q^{[t+1]}, q^{[t+1]}) \leq \dots$$

$$\leq C^{(\alpha, \mu)}$$

ここで不等式 (4.45) および (4.46) の右辺に現れる量に関しては以下が成り立つ.

$$\begin{aligned}
F^{(\alpha, \mu)}(q^{[t]}, q^{[t]}) &= \sum_{u, x, y, z} q^{[t]}(u, x, y, z) \{\omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z)\} \\
&= I_{q^{[t]}}(X; Y|U) + \mu I_{q^{[t]}}(Z; U) \\
&\quad - \alpha D(q_{ZY|XU}^{[t]} \| W_1, W_2 | q_{XU}^{[t]}) - D(q_{Y|XU}^{[t]} \| W_1 | q_{XU}^{[t]}) \tag{4.47}
\end{aligned}$$

$$\begin{aligned}
F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) &= (1 + \alpha) \log K_t \\
&= (1 + \alpha) \log \sum_{(u, x, y, z)} \exp\left\{\frac{1}{1 + \alpha} \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z)\right\} q^{[t]}(u, x, y, z) \tag{4.48}
\end{aligned}$$

(証明) 命題 6 の不等式におけるステップ (a) と式 (4.48) は補題 7 より従う. またステップ (b) と式 (4.47) は補題 5 より従う. ■

本研究で提案された分布更新アルゴリズムは, 新しいパラメータ α が導入されている点で従来のアルゴリズムとは異なるものになっている. 次の節でそのことについて記す.

4.4 従来アルゴリズムとの相違

Yasui and Matsushima[6] の研究:

$$C^{(\mu)}(W_1, W_2) \triangleq \max_{p \in \mathcal{P}(W_1, W_2)} \{I_p(X + Y|U) + \mu I_p(Z + U)\} \tag{4.49}$$

- 上記の $\max_{p \in \mathcal{P}(W_1, W_2)}$ のような制約ありの最適化問題に対し, Arimoto-Blahut アルゴリズムを適用した.

Arimoto-Blahut アルゴリズムの拡張:

$$\begin{aligned} C^{(\mu)}(W_1, W_2) &= \max_{p: p=(p_{UX}, w_1, w_2)} \left\{ I_p(X; Y|U) + \mu I_p(U; Z) \right\} \\ &= \max_{p_{UX}} \left\{ I_{p_{UX}}(X; Y|U) + \mu I_{p_{UX}}(U; Z) \right\} \end{aligned}$$

- Arimoto-Blahut アルゴリズムを自然に拡張し、制約ありの最適化問題を定式化している。
- $\mu > 0$ の場合、収束の保証がされない。

提案する新しい反復アルゴリズム:

$$\begin{aligned} C^{(\alpha, \mu)} &= \max_q \left\{ I_q(X; Y|U) + \mu I_q(Z; U) \right. \\ &\quad \left. - \alpha D(q_{ZY|XU} \| W_1, W_2 | q_{XU}) - D(q_{Y|XU} \| W_1 | q_{XU}) \right\} \end{aligned} \quad (4.50)$$

- 目的関数へのペナルティ項の導入より、上記の \max_q のような制約なしの最適化問題を定式化している。
- Arimoto-Blahut アルゴリズムとは異なる最適値計算アルゴリズムを提案している。

$C^{(\alpha, \mu)}(W_1, W_2) > C^{(\mu)}(W_1, W_2)$ であるから、分布更新アルゴリズムの定める分布列は領域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の内点を与えるものであるとは限らないことになる。しかし、パラメータ α の導入によりアルゴリズムの収束性を改善できる可能性がある。この提案したアルゴリズムが最適値を与える分布に収束するかどうかを、次章で議論する。

第 5 章

提案アルゴリズムの収束性

この章では提案アルゴリズムの領域 $\mathcal{C}_{\text{DBC}}(W_1, W_2)$ の境界を与える分布への収束について議論する。

5.1 収束性に関する結果

$F^{(\alpha, \mu)}(q, q)$ の最適値を与える分布を $q^*(u, x, y, z)$ とすると、補題 6 より以下が成り立つ。

$$\begin{aligned} C^{(\alpha, \mu)}(W_1, W_2) &= \max_q F^{(\alpha, \mu)}(q, q) \\ &= F^{(\alpha, \mu)}(q^*, q^*) \\ &= \sum_{u, x, y, z} q^*(u, x, y, z) \{\omega_{q^*}^{(\alpha, \mu)}(u, x, y, z)\} \end{aligned} \quad (5.1)$$

提案アルゴリズムの収束性について、次の命題が成り立つ。

命題 7 $t = 0, 1, 2, \dots$ に対し、以下が成り立つ。

$$\begin{aligned} &C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \\ &= F^{(\alpha, \mu)}(q^*, q^*) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \\ &= \sum_{u, x, y, z} q^*(u, x, y, z) \omega_{q^*}^{(\alpha, \mu)}(u, x, y, z) - (1 + \alpha) \log K_t \\ &= (1 + \alpha) \left[D(q^* \| q^{[t]}) - D(q^* \| q^{[t+1]}) \right] - D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) - \mu D(q_Z^* \| q_Z^{[t]}) \\ &\quad - \alpha D(q_{Z|Y|XU}^* \| q_{Z|Y|XU}^{[t]} | q_{XU}^*) - D(q_{Y|XU}^* \| q_{Y|XU}^{[t]} | q_{XU}^*) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]} | q_U^*) \end{aligned} \quad (5.2)$$

(証明) 分布更新アルゴリズムの式 (4.44) より、任意の (u, x, y, z) に対し、以下が成り立つ。

$$\begin{aligned} F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) &= (1 + \alpha) \log K_t \\ &= \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) - (1 + \alpha) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \end{aligned} \quad (5.3)$$

式(5.3)の両辺に $q^*(u, x, y, z)$ をかけて (u, x, y, z) について和をとると、以下の式を得る。

$$\begin{aligned}
 & F^{(\frac{1}{1+\alpha}, \alpha, \mu)}(q^{[t]}, q^{[t+1]}) \\
 = & \sum_{u, x, y, z} q^*(u, x, y, z)(1 + \alpha) \log K_t \\
 = & \sum_{u, x, y, z} q^*(u, x, y, z) \left\{ \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) - (1 + \alpha) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \right\} \quad (5.4)
 \end{aligned}$$

式(5.1), (5.4) より, $C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]})$ について以下の等式を得る。

$$\begin{aligned}
 & C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \\
 = & \sum_{u, x, y, z} q^*(u, x, y, z) \omega_{q^*}^{(\alpha, \mu)}(u, x, y, z) \\
 & - \sum_{u, x, y, z} q^*(u, x, y, z) \left\{ \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) - (1 + \alpha) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \right\} \\
 = & (1 + \alpha) \sum_{u, x, y, z} q^*(u, x, y, z) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \\
 & + \sum_{u, x, y, z} q^*(u, x, y, z) \left\{ \omega_{q^*}^{(\alpha, \mu)}(u, x, y, z) - \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) \right\} \\
 = & J_2 + J_3 \quad (5.5)
 \end{aligned}$$

ここで,

$$\begin{aligned}
 J_2 &= (1 + \alpha) \sum_{u, x, y, z} q^*(u, x, y, z) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \\
 J_3 &= \sum_{u, x, y, z} q^*(u, x, y, z) \left\{ \omega_{q^*}^{(\alpha, \mu)}(u, x, y, z) - \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) \right\}
 \end{aligned}$$

とおいた。上記の J_2 について,

$$\begin{aligned}
 J_2 &= (1 + \alpha) \sum_{u, x, y, z} q^*(u, x, y, z) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \times \log \frac{q^*(u, x, y, z)}{q^*(u, x, y, z)} \\
 &= (1 + \alpha) \sum_{u, x, y, z} q^*(u, x, y, z) \left[\log \frac{q^*(u, x, y, z)}{q^{[t]}(u, x, y, z)} - \log \frac{q^{[t+1]}(u, x, y, z)}{q^*(u, x, y, z)} \right] \\
 &= (1 + \alpha) \left[D(q^* \| q^{[t]}) - D(q^* \| q^{[t+1]}) \right] \quad (5.6)
 \end{aligned}$$

が成り立つ。次に J_3 について、以下が成り立つ。

$$\begin{aligned}
J_3 &= \sum_{u,x,y,z} q^*(u, x, y, z) \left\{ \log \frac{W_1(y|x)}{q_{Y|U}^*(y|u)} + \mu \log \frac{q_{Z|U}^*(z|u)}{q_Z^*(z)} + \alpha \log \frac{W_2(z|y)W_1(y|x)}{q_{ZY|XU}^*(z, y|x, u)} \right. \\
&\quad \left. - \log \frac{W_1(y|x)}{q_{Y|U}^*(y|u)} - \mu \log \frac{q_{Z|U}^*(z|u)}{q_Z^*(z)} - \alpha \log \frac{W_2(z|y)W_1(y|x)}{q_{ZY|XU}^*(z, y|x, u)} \right\} \\
&= \sum_{u,x,y,z} q^*(u, x, y, z) \left\{ \log \frac{q_{Y|U}^{[t]}(y|u)}{q_{Y|U}^*(y|u)} + \mu \log \frac{q_Z^{[t]}(z)}{q_Z^*(z)} + \mu \log \frac{q_{Z|U}^*(z|u)}{q_{Z|U}^{[t]}(z|u)} + \alpha \log \frac{q_{ZY|XU}^{[t]}(z, y|x, u)}{q_{ZY|XU}^*(z, y|x, u)} \right\} \\
&= -D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) - \mu D(q_Z^* \| q_Z^{[t]}) \\
&\quad - \alpha D(q_{ZY|XU}^* \| q_{ZY|XU}^{[t]} | q_{XU}^*) - D(q_{Y|XU}^* \| q_{Y|XU}^{[t]} | q_{XU}^*) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]} | q_U^*) \tag{5.7}
\end{aligned}$$

式(5.5) – (5.7) より以下の等式を得る。

$$\begin{aligned}
&C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \\
&= \sum_{u,x,y,z} q^*(u, x, y, z) \left\{ \omega_{q^*}^{(\alpha, \mu)}(u, x, y, z) - \omega_{q^{[t]}}^{(\alpha, \mu)}(u, x, y, z) \right\} \\
&\quad + (1 + \alpha) \sum_{u,x,y,z} q^*(u, x, y, z) \log \frac{q^{[t+1]}(u, x, y, z)}{q^{[t]}(u, x, y, z)} \\
&= (1 + \alpha) \left[D(q^* \| q^{[t]}) - D(q^* \| q^{[t+1]}) \right] - D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) \\
&\quad - \mu D(q_Z^* \| q_Z^{[t]}) - \alpha D(q_{ZY|XU}^* \| q_{ZY|XU}^{[t]} | q_{XU}^*) - D(q_{Y|XU}^* \| q_{Y|XU}^{[t]} | q_{XU}^*) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]} | q_U^*)
\end{aligned}$$

以上で命題7が証明された。 ■

この命題7が収束するための条件について、次の節で述べる。

5.2 最適分布への収束条件

この節では、提案アルゴリズムが最適分布に収束するための条件について記す。収束性について次の命題が成り立つ。

命題8 式(5.2)より、任意の $t = 0, 1, 2, \dots$ に対し

$$\begin{aligned}
&- D(q_{Y|U}^* \| q_{Y|U}^{[t]} | q_U^*) - \mu D(q_Z^* \| q_Z^{[t]}) \\
&- \alpha D(q_{ZY|XU}^* \| q_{ZY|XU}^{[t]} | q_{XU}^*) - D(q_{Y|XU}^* \| q_{Y|XU}^{[t]} | q_{XU}^*) + \mu D(q_{Z|U}^* \| q_{Z|U}^{[t]} | q_U^*) \leq 0 \tag{5.8}
\end{aligned}$$

なる不等式が成り立つと仮定する。このとき分布更新アルゴリズムの定める分布列 $\{q^{[t]}\}_{t=1}^{\infty}$ の $C^{(\alpha, \mu)}(W_1, W_2)$ を与える最適分布への収束する。

(証明) (5.8) の仮定が成り立つという条件のもとで

$$0 \leq C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \leq (1 + \alpha) \left[D(q^* \| q^{[t]}) - D(q^* \| q^{[t+1]}) \right] \quad (5.9)$$

が成り立つ. (5.9) の両辺について $t = 1$ から T まで和をとると

$$\sum_{t=1}^T \left\{ C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]}) \right\} = (1 + \alpha) [D(q^* \| q^{[1]}) - D(q^* \| q^{[T+1]})] \quad (5.10)$$

を得る. ここで,

$$\Delta_t = C^{(\alpha, \mu)}(W_1, W_2) - F^{(\alpha, \mu)}(q^{[t]}, q^{[t+1]})$$

とおくと命題2より, $\{\Delta_t\}_{t=1}^{\infty}$ は単調減少列であるから (5.10) より

$$\begin{aligned} T \Delta_T &\leq \sum_{t=1}^T \Delta_t \leq (1 + \alpha) [D(q^* \| q^{[1]}) - D(q^* \| q^{[T+1]})] \\ &\leq (1 + \alpha) D(q^* \| q^{[1]}) \end{aligned}$$

これより

$$0 \leq \Delta_T \leq \frac{1 + \alpha}{T} D(q^* \| q^{[1]}) \rightarrow 0 (T \rightarrow +\infty)$$

となり, 分布更新アルゴリズムが最適値を与える分布へ収束することがわかる. 以上で命題8は証明された. ■

この命題8の式(5.8)の仮定は成り立つかどうかは不明あり, 現在のところ収束性の証明はできていない.

- $\mu = 0$ のときは, この仮定が成り立つ. この場合は単一通信路の通信路容量を計算していることになる.
- $\mu > 0$ のときも, α が十分大きい値のときは, この仮定が成り立ちそうに思える. しかし, 現時点ではこのことを証明できていない.

そこで, 次章では理論値の求まっている放送通信路と比較実験を行い, アルゴリズムの収束性を確認する. そして提案アルゴリズムの性能を評価する.

第 6 章

数値実験

この章では、提案アルゴリズムの収束性が証明できていないため、数値実験を用いて理論値との比較を行い、提案アルゴリズムの性能を評価する。

6.1 2元対称DBCの理論値

本論文で比較実験に使用する通信路は、下図のような劣化型2元放送通信路である。

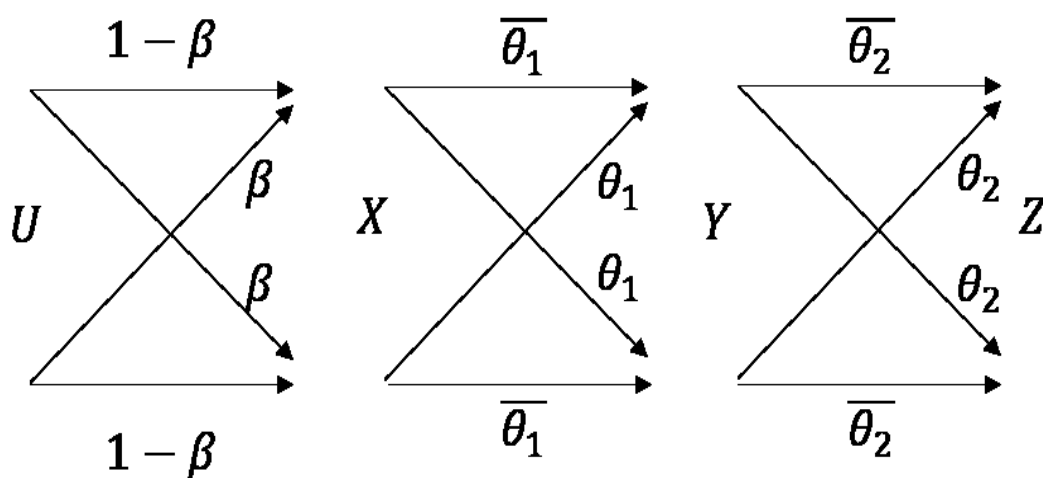


図 6.1: 2元対称 DBC

この放送通信路の理論値は以下のように導出することができる。

$$\begin{aligned}
I(X; Y|U) &= H(Y|U) - H(Y|X, U) \\
&= H(\beta * \theta_1) - H(\theta_1), \quad \beta * \theta_1 = \beta \bar{\theta}_1 + (1 - \beta)\theta_1 \\
I(U; Z) &= H(Z) - H(Z|U) \\
&= 1 - H(\beta * p), \quad \beta * p = \beta(1 - p) + (1 - \beta)p, \quad p = \theta_1 \bar{\theta}_2 + \bar{\theta}_1 \theta_2
\end{aligned}$$

6.2 実験の設定

比較実験で使用するパラメータを以下のように設定する。ここでは $C^{(\alpha, \mu)}(W_1, W_2)$ を求める提案アルゴリズムにより得られる $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ 上の分布列を $\{q_{\mu}^{[t]}\}_{t=0}^{+\infty}$ と記し、パラメータ μ を明示することにする。

パラメータの設定

$$\begin{aligned}
\alpha &= 10^6, & 0 \leq \mu \leq 1 + \alpha \\
\mu_k &= \delta_k, & k = 0, 1, 2, \dots, \lceil \frac{1+\alpha}{\delta} \rceil, & \delta = 10^{-4}
\end{aligned}$$

α は十分大きな値に固定する。そして μ は 10^{-4} ずつ大きくしていき、各 μ 毎に支持直線を導出していくようにする。分布更新の停止条件は以下のようにする。

分布更新の停止条件

任意の (u, x, y, z) のもとで

$$|q_{\mu_k}^{[t]}(u, x, y, z) - q_{\mu_k}^{[t-1]}(u, x, y, z)| \leq \epsilon, \quad \epsilon = 10^{-4}$$

を満たす最小の t (これを t_k^* とする) のときに停止する。

分布更新する前と後の差が ϵ 以下となったときに、分布更新を停止する。初期分布は以下のように設定する。

初期分布の設定

1. $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ 上の分布 $q_{\mu_0}^{[0]}$ をランダムに選ぶ。
2. $k = 0, 1, 2, \dots$ に対し, $q_{\mu_k}^{[0]} = q_{\mu_{k-1}}^{[t_{k-1}^*]}$ と設定する。

上記の設定は μ_{k-1} に対する収束分布を μ_k の初期分布にとるように設定することを意味する。

以上ような設定のもと、比較実験を行った。

6.3 実験の結果

6.3.1 2元の場合

赤線が実験値，緑線が理論値である．横軸が R_1 を表し，縦軸が R_2 を表している．確率遷移行列を以下のように設定した．

確率遷移行列の設定

$$W_1(y|x) = \begin{bmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}, \quad W_2(z|y) = \begin{bmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}$$

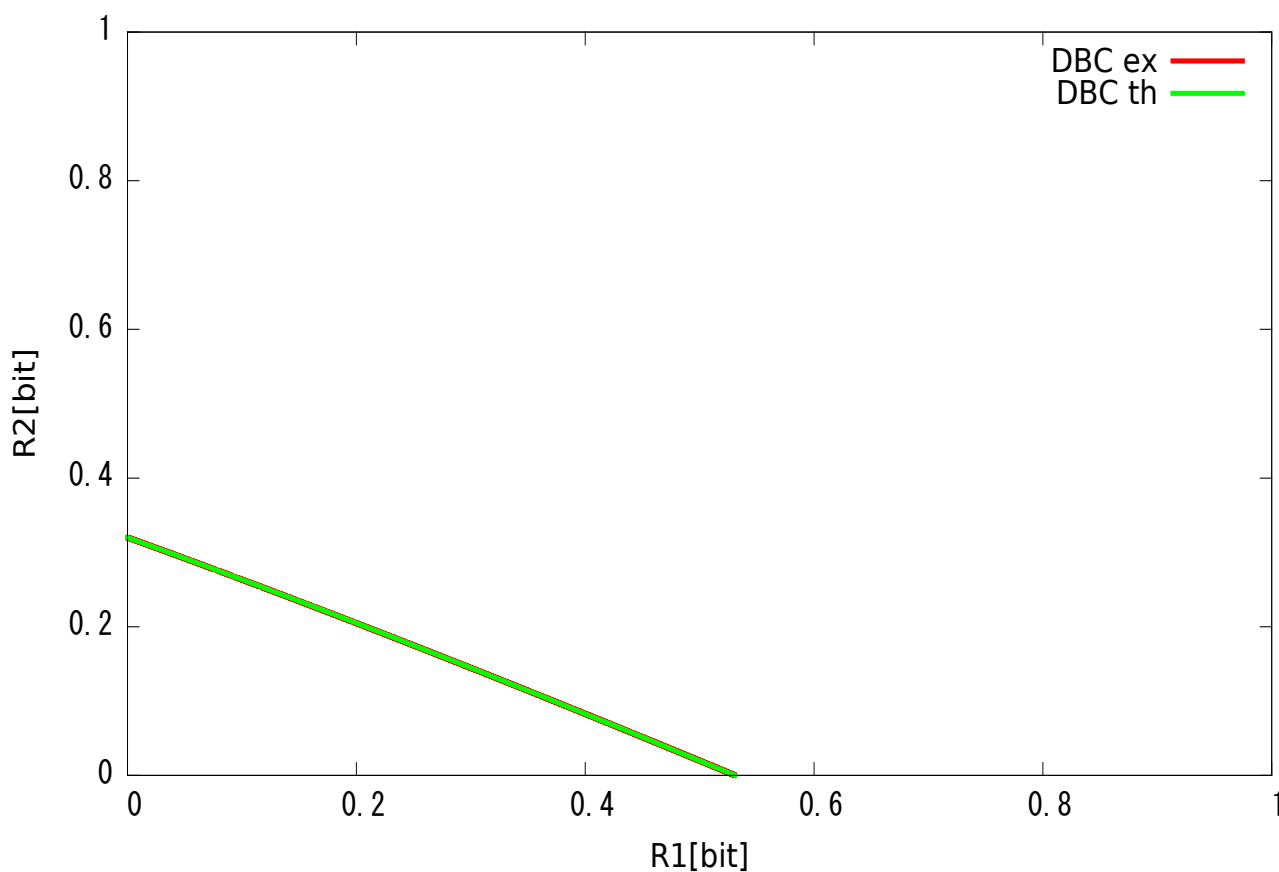


図 6.2: 実験値と理論値の比較結果

この確率遷移行列の場合，理論値と実験値がほとんど一致するということを確認することができた．次に $W_2(z|y)$ の値は変更せず， $W_1(y|x)$ の値のみを変更して比較実験を行った．

確率遷移行列を以下のように設定した。

確率遷移行列の設定

$$W_1(y|x) = \begin{bmatrix} \frac{99}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{99}{100} \end{bmatrix}, \quad W_2(z|y) = \begin{bmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}$$

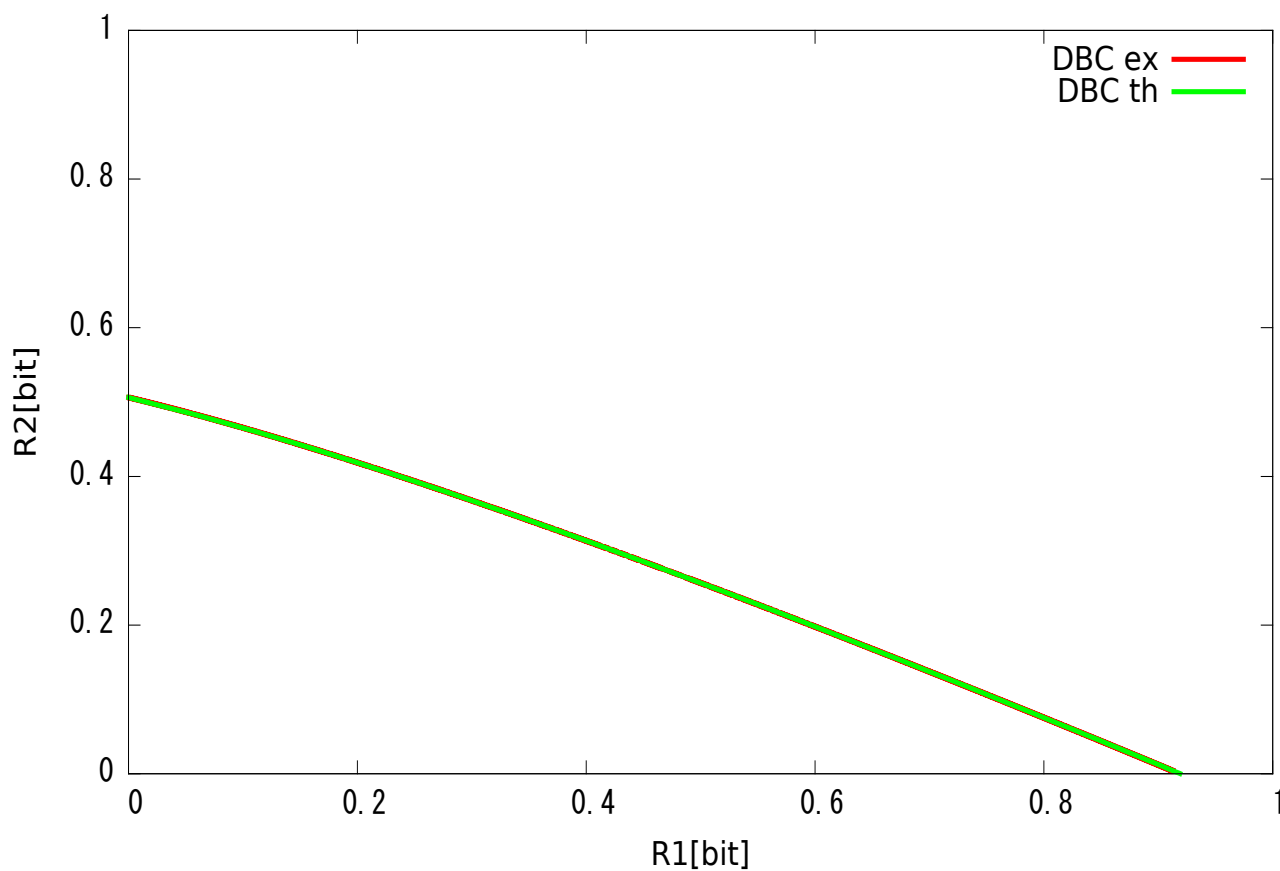


図 6.3: 実験値と理論値の比較結果

この確率遷移行列の場合も、理論値と実験値がほとんど一致するということを確認することができた。同様に $W_2(z|y)$ の値は変更せず、 $W_1(y|x)$ の値のみを変更して比較実験を行った。

確率遷移行列を以下のように設定した。

確率遷移行列の設定

$$W_1(y|x) = \begin{bmatrix} \frac{999}{1000} & \frac{1}{1000} \\ \frac{1}{1000} & \frac{999}{1000} \end{bmatrix}, \quad W_2(z|y) = \begin{bmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}$$

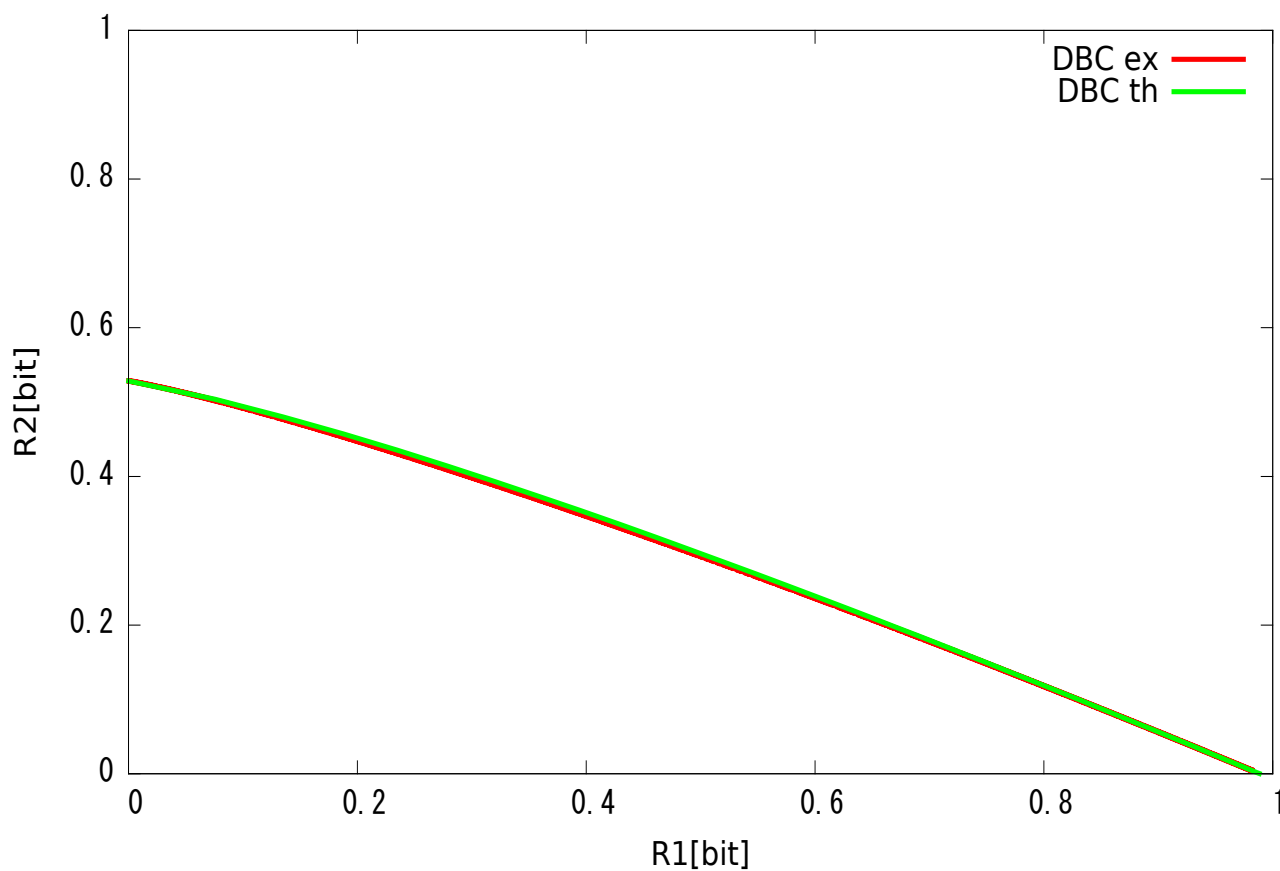


図 6.4: 実験値と理論値の比較結果

この確率遷移行列の場合、理論値と実験値が多少の誤差はあるものの、誤差は 10^{-4} 程度のものであるので、ほとんど一致しているといえる。以上の結果から提案アルゴリズムは、2元対称 DBC で、どのような初期分布を与えても最適値を与える分布に収束するということを比較実験により確認することができた。また、領域の具体形が陽に得られていない一般の場合にも提案アルゴリズムを適用できるかどうかを調べるために X, Y, Z がいずれも 3 元の場合に同様のパラメータの設定を用いて実験を行った。

6.3.2 3元の場合

確率遷移行列を以下のように設定した。

確率遷移行列の設定

$$W_1(y|x) = \begin{bmatrix} \frac{98}{100} & \frac{1}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{98}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{98}{100} & \frac{1}{100} \end{bmatrix}, \quad W_2(z|y) = \begin{bmatrix} \frac{6}{10} & \frac{2}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{6}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{2}{10} & \frac{6}{10} \end{bmatrix}$$

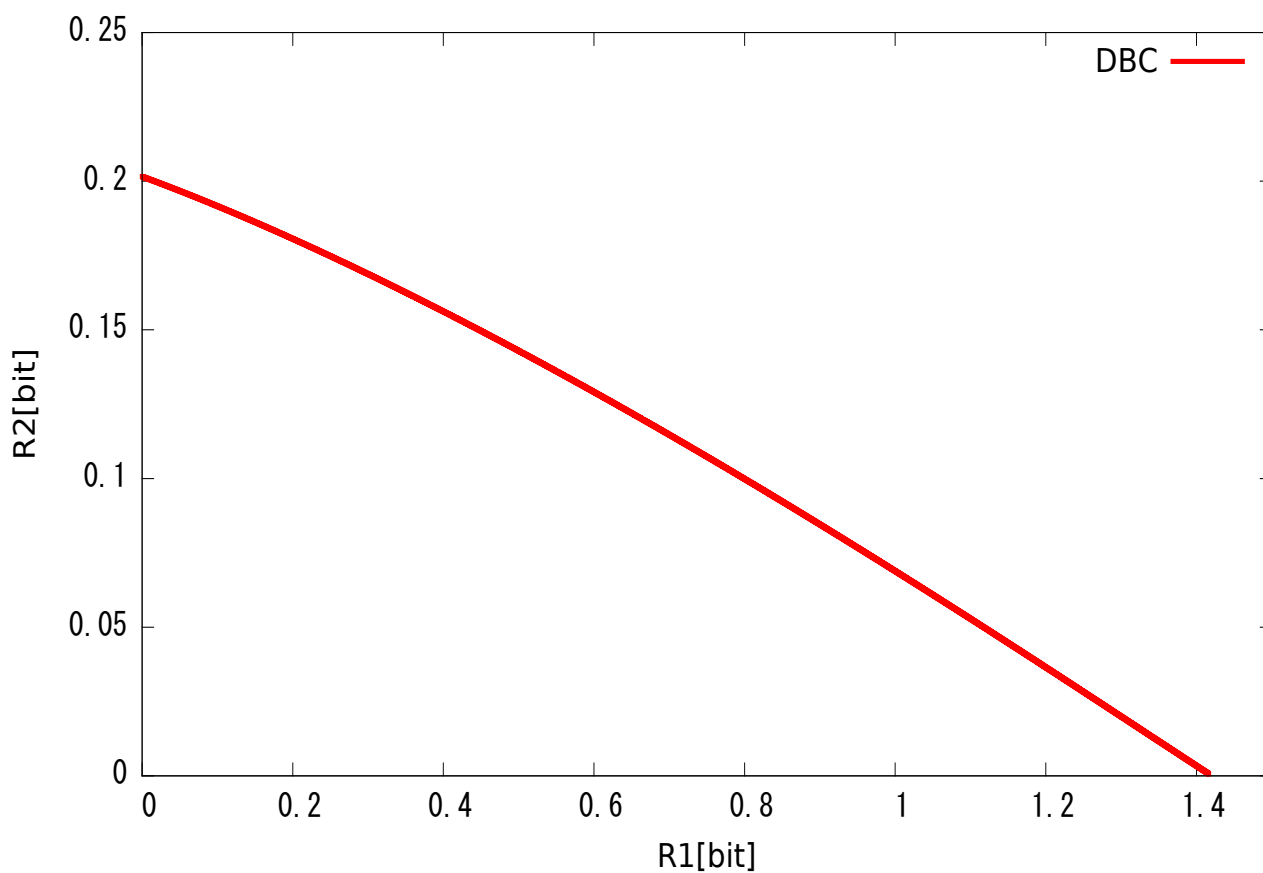


図 6.5: 3元の場合の実験値

3元の場合は理論値が求められていないため比較実験をすることができないのだが、2元の結果の類推から、提案アルゴリズムは3元の場合も通信路容量域を導出することができるのではないかと考えられる。しかし、2元の場合とは異なり、初期分布の与え方で、明らかに最適解ではない結果になることがある。その例を次に示す。

確率遷移行列の設定

$$W_1(y|x) = \begin{bmatrix} \frac{98}{100} & \frac{1}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{98}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{98}{100} & \frac{1}{100} \end{bmatrix}, \quad W_2(z|y) = \begin{bmatrix} \frac{6}{10} & \frac{2}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{6}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{2}{10} & \frac{6}{10} \end{bmatrix}$$

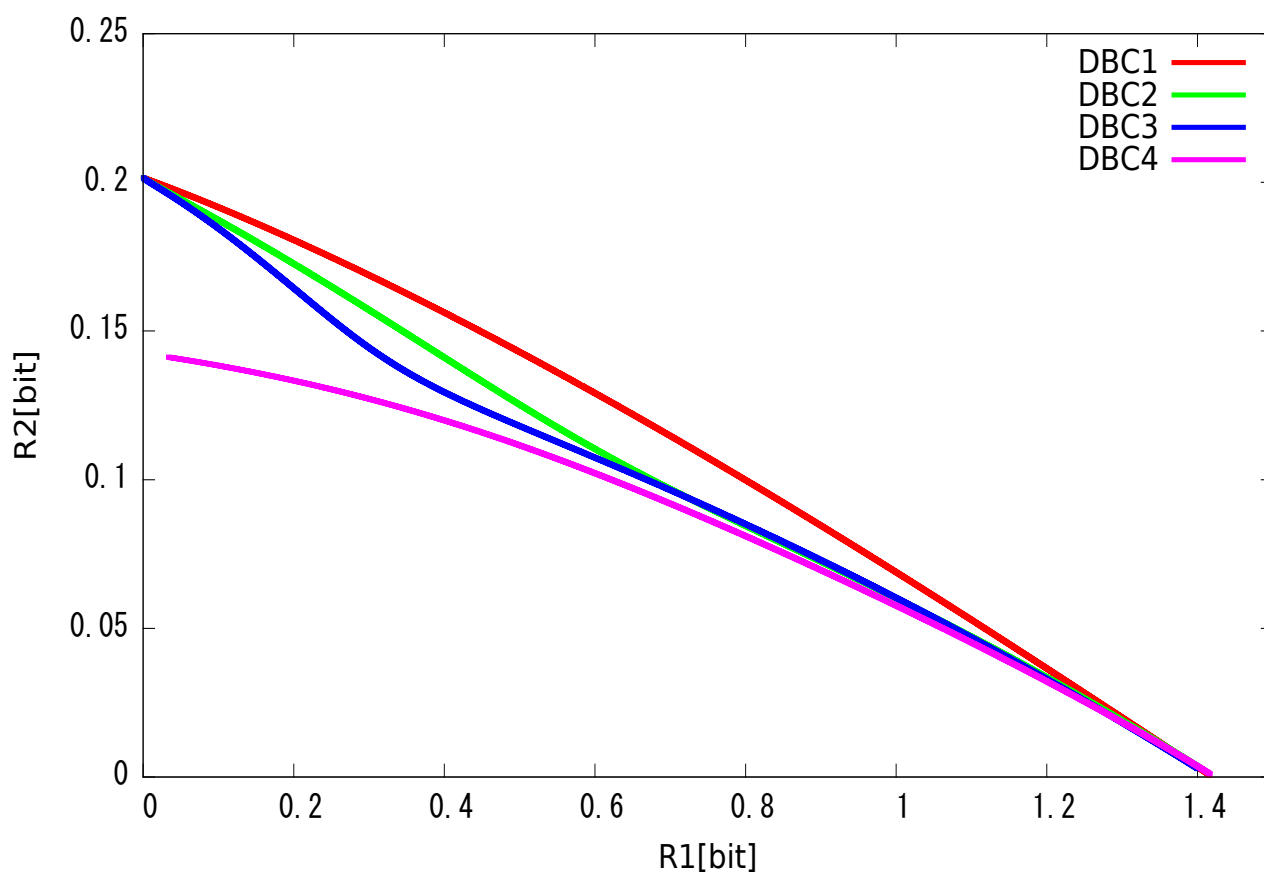


図 6.6: 3 元の場合の実験値

上図のように提案アルゴリズムは3元の場合、初期分布の与え方によって異なる結果を得た。これは初期分布の与え方で差凸関数の凸領域ではない部分に分布列が収束してしまうことが原因ではないかと考えられる。そこで各結果に対する初期分布を以下に記述する。

DBC1 の初期分布

$$q(u, x) = \begin{bmatrix} q(0, 0) = 0.115822 & q(0, 1) = 0.115890 & q(0, 2) = 0.106603 \\ q(1, 0) = 0.093810 & q(1, 1) = 0.132307 & q(1, 2) = 0.106580 \\ q(2, 0) = 0.096411 & q(2, 1) = 0.111829 & q(2, 2) = 0.120747 \end{bmatrix}$$

DBC2 の初期分布

$$q(u, x) = \begin{bmatrix} q(0, 0) = 0.128118 & q(0, 1) = 0.088595 & q(0, 2) = 0.145526 \\ q(1, 0) = 0.080623 & q(1, 1) = 0.123015 & q(1, 2) = 0.073763 \\ q(2, 0) = 0.113375 & q(2, 1) = 0.124842 & q(2, 2) = 0.122141 \end{bmatrix}$$

DBC3 の初期分布

$$q(u, x) = \begin{bmatrix} q(0, 0) = 0.114849 & q(0, 1) = 0.123280 & q(0, 2) = 0.100182 \\ q(1, 0) = 0.103137 & q(1, 1) = 0.129223 & q(1, 2) = 0.099822 \\ q(2, 0) = 0.153802 & q(2, 1) = 0.079472 & q(2, 2) = 0.096231 \end{bmatrix}$$

DBC4 の初期分布

$$q(u, x) = \begin{bmatrix} q(0, 0) = 0.118767 & q(0, 1) = 0.101191 & q(0, 2) = 0.094143 \\ q(1, 0) = 0.096188 & q(1, 1) = 0.121033 & q(1, 2) = 0.128425 \\ q(2, 0) = 0.124526 & q(2, 1) = 0.106493 & q(2, 2) = 0.109237 \end{bmatrix}$$

6.4 実験結果のまとめ

2元対称 DBC に対する理論値との比較実験で、理論値と実験値が一致した。つまり提案アルゴリズムは最適分布に収束し、通信路容量域を導出する事ができることを確認することができた。領域の具体形が陽に得られていない一般の場合にも提案アルゴリズムを適用できるかどうかを調べるために X, Y, Z がいずれも3元の場合で実験を行ってみた。3元の場合は理論値がわからないため、2元の結果からの類推しかできないが、最適値を与えていると期待できる結果を得る事ができた。しかし、明らかに最適値ではない結果を得る場合もあった。3元の場合は、初期分布の与え方で差凸関数の凸領域ではない部分に分布列が収束してしまうことが原因ではないかと考えられる。差凸関数の具体的な形の定量的評価が今後の課題である。

第7章

おわりに

本研究では劣化型放送通信路の容量域を数値的に求める新しい計算アルゴリズムを提案し、さらに提案アルゴリズムの定める分布の領域の境界を与える分布への収束の可能性を議論した。パラメータ α が十分大きいときに、分布更新アルゴリズムは最適分布に収束するというを示せるように思える。しかし現在のところ、この厳密な証明法はわかっていない。そこで理論値の求まっている2元対称DBCに対する数値実験により、提案するアルゴリズムが最適値を与える分布に収束するという事を確認できた。領域の具体形が陽に得られていない一般の場合にも提案アルゴリズムを適用できるかどうかを調べるために X, Y, Z がいずれも3元の場合で実験を行った。2元の結果の類推から最適値を与えていると期待できる結果を得る事ができた。ただし3元の場合、初期分布の与え方によって最適値を与える分布に収束しないことがある。おそらく初期分布の与え方で、差凸関数の凸領域ではない部分に収束してしまうことが原因ではないかと考えられる。今後の課題は、提案アルゴリズムの最適分布への収束性の証明と、差凸関数の具体的な形の定量的評価をしていく必要がある。

参考文献

- [1] S. Arimoto, “An algorithm for computing the capacity of arbitrary discrete memoryless channels,” *IEEE Trans. Inform Theory*, vol. 18, no. 1, pp 14-20, 1972.
- [2] R. E. Blahut, “Computation of channel capacity and rate-distortion function,” *IEEE Trans. Inform. Theory*, vol. 18, no. 4, pp. 460-473, 1972.
- [3] 安井 謙介 須子 統太 松嶋 敏泰, “拡張された有本-Blahut アルゴリズムの大域的収束性について,” *電子情報通信学会論文誌*, Vol. J91-A, No. 9, pp. 846-860, 2008.
- [4] G. Kumar and A. Thangaraj, “Computation of secrecy capacity for more-capable channel pair,” *Proceedings of the IEEE International Symposium on Information Theory*, Toronto, pp. 529-533, 2008.
- [5] E. Calvo, D. P. Palomar, J. R. Fonollosa and J. Vidai, “The computation of the capacity region of the discrete degraded BC is a non convex DC problem,” *Proceedings of the IEEE International Symposium on Information Theory*, Toronto, pp. 1721-1725, 2008.
- [6] K. Yasui T. Matsushima, “Toward computing the capacity region of degraded broadcast channel,” *Proceedings of the IEEE International Symposium on Information Theory*, Austin, pp. 570-574, 2010.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd ed.*, John Wiley & Sons, New Jersey, 2006.
- [8] 韓太舜・小林欣吾, *情報と符号化の数理*, 培風館, 1995.

謝辞

最後に，大変お忙しいのにも関わらず，本研究を進めるにあたり最後まで丁寧にご指導してくださった大濱靖匡教授に深く感謝致します。本当にありがとうございました。また，ゼミや学会など至る所で大変お世話になりました川端教授，八木准教授，竹内助教授に心より感謝致します。そして一緒に勉学に励んだ大濱研究室の学生の皆さんや，関連研究室の学生の皆さんにも感謝しています。3年間本当にありがとうございました。

付録A

補足

定理 A ダイバージェンス $D(P\|Q)$ は、 (P, Q) の下に凸な関数である。また、 $D(P\|Q)$ は P に関して狭義に下に凸である。[8]

(証明) $0 \leq \alpha \leq 1$ と X の上の 2 組の確率分布の対 (P_1, Q_1) と (P_2, Q_2) に対して、 $P = \alpha P_1 + (1 - \alpha)P_2, Q = \alpha Q_1 + (1 - \alpha)Q_2$ とおくと、

$$\alpha D(P_1\|Q_1) + (1 - \alpha)D(P_2\|Q_2) \geq D(P\|Q) \quad (\text{A.1})$$

を示せば良い。ところが、対数和不等式より

$$\alpha P_1(x) \log \frac{\alpha P_1(x)}{\alpha Q_1(x)} + (1 - \alpha)P_2(x) \log \frac{(1 - \alpha)P_2(x)}{(1 - \alpha)Q_2(x)} \geq P(x) \log \frac{P(x)}{Q(x)} \quad (\text{A.2})$$

が成り立っている。この不等式の両辺を x について和をとれば、不等式 (A.1) が得られる。 P に関する狭義の凸性はエントロピー $H(P)$ の狭義の凸性から導かれる。 ■

ダイバージェンスの凸性を用いると、 $D(p_{X|Z}\|p_{Y|Z}|p_Z) \geq D(p_X\|p_Y)$ であることがただちに導かれる。

定理 B 確率分布 P, Q をもつ 2 つの確率変数 X_1, X_2 が同一の通信路 W への入力であったときの出力をそれぞれ Y_1, Y_2 とすると、出力間のダイバージェンスは入力間のダイバージェンスより小さくなる。すなわち、

$$D(PW\|QW) \leq D(P\|Q) \quad (\text{A.3})$$

が成り立つ。[8]

(証明) 対数和不等式を用いると

$$\begin{aligned} D(P\|Q) &= \sum_x P(x) \log \frac{P(x)}{Q(x)} \\ &= \sum_{x,y} P(x)W(y|x) \log \frac{P(x)W(y|x)}{Q(x)W(y|x)} \\ &\geq \sum_y PW(y) \log \frac{PW(y)}{QW(y)} \\ &= D(PW\|QW) \end{aligned} \tag{A.4}$$

が導かれる。 ■