信州大学
大学院総合理工学研究科

修士論文

離散無記憶通信路に対するコストあたり通信路容量の算出における精度保証

指導教員　西新 幹彦 准教授

学科　電子情報システム工学科
学籍番号　22w2089g
氏名　伏屋 麟

2024 年 03 月 01 日

目次

1　はじめに

通信システムとは，符号シンボルを送受信することによって情報を伝えるシステムである．一般に符号シンボルの送信にはコストがかかる．具体例として，符号シンボルとして，電圧や電磁波, 磁気などを用いる場合, いずれも電力がコストとして必要になる. また，1 つのシンボルを送信するのにかかる時間もコストとみなされうる.

情報理論における通信路符号化の問題では，通信路容量を求めることが基本的な問題となる．コストの概念のない通常の通信路符号化では，通信路容量は符号シンボルあたりの情報量という次元で定義される．これに対して, コスト付き通信路符号化問題と呼ばれる，符号シンボルにコストが定義されているような問題がある．この場合，コストあたりの情報量という次元で定義されるコストあたり通信路容量を求める問題を考えることになる.

コスト付き通信路符号化問題に対して，Gallager[3] は，コスト制約付き通信路容量を制約付き最適化問題の最適解として明らかにし，コストあたり通信路容量の算出アルゴリズムを，Blahut[4] が出している．そして，川合西新 [1] がより速い計算が期待できるコストあたり通信路容量を算出するためのアルゴリズムを提案している．川合西新提案のアルゴリズムは，有本 [2]，Blahut による，コストの概念のない通信路の通信路容量を求めるアルゴリズムを基礎にしている．有本，Blahut によるアルゴリズムは，算出値の精度保証を与える方法が知られている．しかしながら，川合西新提案のアルゴリズムには精度保証をするための手段が用意されていない.

本論文では，川合西新提案のコストあたり通信路容量を算出するアルゴリズムに対して，精度保証の手段を与える.

2　コストあたり通信路容量について

本研究では，通信路は定常無記憶であると仮定する．通信路は，符号シンボルを $x \in \mathcal{X}$, 入力確率変数と出力確率変数をそれぞれ X, Y としたとき，条件付き確率である

$$W(y|x) \triangleq \Pr\{Y = y|X = x\}, x \in \mathcal{X}, y \in \mathcal{Y} \tag{1}$$

で表される. ここで $\mathcal{X}, \mathcal{Y}$ はそれぞれ入力と出力のアルファベットである．このことから,

$$\sum_{y \in \mathcal{Y}} W(y|x) = 1, x \in \mathcal{X} \tag{2}$$

であると言える.

入力確率分布 p_x は

$$p_x \triangleq \Pr\{X = x\} \tag{3}$$

で定義され，

$$\sum_x p_x = 1, x \in \mathcal{X} \tag{4}$$

を満たす．ここで，p は

$$p = \{p_x\}_{x \in \mathcal{X}} \tag{5}$$

である．また，出力確率分布 q_y は

$$q_y \triangleq \sum_x p_x W(y|x) \tag{6}$$

である．

入力，出力確率変数である X, Y に対して，X, Y の間の相互情報量 $I(X;Y)$ は

$$I(X;Y) \triangleq \sum_{x,y} p_x W(y|x) \log \frac{W(y|x)}{q_y} \tag{7}$$

で定義される. 本研究では通信路は固定しているので，相互情報量は入力分布の関数であることから，$I(X;Y)$ は入力分布 p の関数として

$$I(p) = \sum_{x,y} p_x W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \tag{8}$$

と表すことが出来る．

入力分布を動かして相互情報量を最大化することで，誤りなしでの最大通信速度である通信路容量

$$C = \max_X I(X;Y) \tag{9}$$

を得られる [5].

ここで，符号シンボルに対してコストの概念を導入する．それぞれの符号シンボル $x \in \mathcal{X}$ に対して, 実数 $c(x) > 0$ が定められているとする．これを x のコストという.

また，$c(x)$ に関して,

$$\Gamma(X) \triangleq \sum_x c(x) p_x \tag{10}$$

が定義される.$\Gamma(X)$ はコストの期待値である.

コストの期待値を用いることによって,X, Y の間のコストあたり相互情報量は

$$I'(X;Y) \triangleq \frac{I(X;Y)}{\Gamma(X)} \tag{11}$$

で定義される. 式 (8) と同様に，$I'(X;Y)$ は入力分布 p の関数として

$$I'(p) = \frac{1}{\sum_x p_x c(x)} \sum_{x,y} p_x W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \tag{12}$$

と表すことが出来る．コストあたり相互情報量を入力分布に関して最大化した

$$C' = \sup_X I'(X;Y) \tag{13}$$

をコストあたり通信路容量という [5]. この量は誤りなしでのコストあたりの最大通信速度である．また，$I'(p)$ を用いてコストあたり通信路容量は

$$C' = \sup_p I'(p) \tag{14}$$

と表せる.

3　コストあたり通信路容量を算出するアルゴリズム

では, 川合西新提案のコストあたり通信路容量を算出するアルゴリズムについて紹介する [1].

まず，X のエントロピー $H(X)$ を,

$$H(X) \triangleq \sum_x p_x \log \frac{1}{p_x} \tag{15}$$

と定める. すると，X のエントロピーを用いて X, Y の間のコストあたり相互情報量は

$$I'(p) = \frac{1}{\Gamma(p)} \left\{ H(X) + \sum_{x,y} p_x W(y|x) \log \frac{p_x W(y|x)}{q_y} \right\} \tag{16}$$

と書ける. これを疑似的に 2 変数に拡張した概念として

$$I'(p;\varphi) = \frac{1}{\Gamma(p)} \left\{ H(X) + \sum_x p_x \sum_x W(y|x) \log \varphi(x|y) \right\} \tag{17}$$

を定義する．ただしパラメータ φ は

$$\varphi = \{\varphi(x|y)\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \tag{18}$$

であり，条件

$$\sum_x \varphi(x|y) = 1, y \in \mathcal{Y} \tag{19}$$

を満たす. 関数 $I'(p;\varphi)$ は, 変数を 2 つ持つ関数である．変数を 1 つ増やすことで，片方の変数を動かし, もう一方の変数を固定して段階的に算出値を最大化することができるようになる.

では，アルゴリズムの詳細な手順について説明する.

1. 初期入力分布 $p^{(1)} = \{p_x^{(1)}\}$ を内点にとる．一様分布

$$p_x^{(1)} = \frac{1}{|\mathcal{X}|} \tag{20}$$

としてよい.

2. $N=1$ とする．
3. $\sup_{\varphi} I'(p^{(N)};\varphi)$ を達成する φ を求める．具体的には，これを $\varphi^{(N)}=\{\varphi(x|y)^{(N)}\}$ とし，

$$\varphi(x|y)^{(N)}=\frac{p_x^{(N)}W(y|x)}{q_y^{(N)}} \tag{21}$$

とする．
4. $\sup_{p} I'(p;\varphi^{(N)})$ を達成する p を求める．具体的には，これを $p^{(N+1)}=\{p_x^{(N+1)}\}$ とし，

$$p_x^{(N+1)}=\exp\left\{\sum_y W(y|x)\log\varphi(x|y)^{(N)}-\alpha^{(N)}c(x)\right\} \tag{22}$$

とする．ただし $\alpha^{(N)}$ は，

$$\sum_x \exp\left\{\sum_y W(y|x)\log\varphi(x|y)^{(N)}-\alpha^{(N)}c(x)\right\}=1 \tag{23}$$

の実数解である．
5. N を 1 増やして 3. に戻る．

以上より得られる

$$\sup_{\varphi} I'(p^{(N)};\varphi)=I'(p^{(N)};\varphi^{(N)}) \tag{24}$$

および

$$\sup_{p} I'(p;\varphi^{(N)})=I'(p^{(N+1)};\varphi^{(N)}) \tag{25}$$

はコストあたり通信路容量 C' に収束する値となる [1]．したがって，ループの回数 N を十分大きくすることによって誤差の少ない値が得られる．

4　上界の導出

前章で述べたように，川合西新提案のアルゴリズムは，1 つの変数を固定しもう一つの変数を動かして算出値を最大化する工程を繰り返す．よって，ループ回数が大きくなるほど算出値も大きくなり，真のコストあたり通信路容量に近づく仕組みとなっている．つまり，算出値自身がコストあたり通信路容量の下界である．これに加えて，コストあたり通信路容量の上界を求めることによって，上界と下界の差を持って精度保証とすることが可能である．したがって，精度保証のためには, ループ回数が十分大きくなった時に，コストあたり通信路容量との誤差が小さくなるような上界を導出すればよい．

コストあたり通信路容量 C' は

$$\begin{aligned} C' &= \sup_p I'(p) \\ &= I'(p^*) \end{aligned} \tag{26}$$

とも書ける. ただし p^* は式 (26) を満たす入力分布とする．これに対応する出力の分布を q^* と表す．すなわち

$$q^*_y \triangleq \sum_x p^*_x W(y|x) \tag{27}$$

である．これらの分布は未知であることに注意する．

式 (26) を変形すると

$$\begin{aligned} C' &= I'(p^*) \\ &= \frac{1}{\Gamma(p^*)}\left\{\sum_{x,y} p^*_x W(y|x)\log\frac{W(y|x)}{\displaystyle\sum_x p^*_x W(y|x)}\right\} \\ &= \frac{1}{\Gamma(p^*)}\left\{\sum_x p^*_x \sum_y W(y|x)\log\frac{\sum_x p_x W(y|x)}{\sum_x p^*_x W(y|x)}\right. \\ &\qquad\qquad \left. + \sum_x p^*_x \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)}\right\} \qquad (28) \\ &= \frac{1}{\Gamma(p^*)}\left\{\sum_x p^*_x \sum_y W(y|x)\log\frac{q_y}{q^*_y} + \sum_x p^*_x \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)}\right\} \\ &= \frac{1}{\Gamma(p^*)}\left\{-D(q^* \parallel q) + \sum_x p^*_x \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)}\right\} \qquad (29) \end{aligned}$$

と変形する. 式 (28) に含まれる p_x は式 (22) で見られるアルゴリズム中に計算する値であるため, 未知数ではない. 式 (29) に含まれるダイバージェンス $D(q^* \parallel q)$ は非負より

$$C' \leqslant \frac{1}{\Gamma(p^*)}\left\{\sum_x p^*_x \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)}\right\} \tag{30}$$

となる. ここで

$$a(x) \triangleq \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)} \tag{31}$$

をおくと，

$$
\begin{aligned}
C' &\leqslant \frac{1}{\Gamma(p^*)}\left\{\sum_x p_x^* \sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)}\right\} \\
&= \frac{\displaystyle\sum_x p_x^* a(x)}{\Gamma(p^*)} \\
&= \frac{\displaystyle\sum_x p_x^* a(x)}{\displaystyle\sum_x p_x^* c(x)}
\end{aligned} \tag{32}
$$

と表すことができる. 式 (32) には p_x^* という未知の値が含まれている. ここで，p_x^* の代わりに新たなパラメータ μ_x を導入する． μ に関して最大化することで未知数 p^* によらない C' の上界を求める．式 (32) を μ の関数として

$$
F(\mu) = \frac{\sum_x \mu_x a(x)}{\sum_x \mu_x c(x)} \tag{33}
$$

のように定義する．ここで，μ_3 から μ_n までを定数として扱うと，

$$
\begin{cases}
\mu_1 : 0 \leqslant 1 - \mu_3 - \mu_4 \cdots - \mu_n \\
\mu_2 = 1 - \mu_1 - \mu_3 - \mu_4 \cdots - \mu_n \\
\mu_3 = \text{定数} \\
\mu_4 = \text{定数} \\
\vdots \\
\mu_n = \text{定数}
\end{cases} \tag{34}
$$

のように実質的に 1 変数の式として用いることができる.$F(\mu)$ に $\mu_2 = 1 - \mu_1 - \mu_3 - \mu_4 \cdots - \mu_n$ を代入したものを $F_{\mu_1\mu_2}(\mu_1)$ とすると，

$$
F_{\mu_1\mu_2}(\mu_1) = \frac{\mu_1 a(1) + (1 - \mu_1 - \sum_{x=3}^n \mu_x)a(2) + \sum_{x=3}^n \mu_x a(x)}{\mu_1 c(1) + (1 - \mu_1 - \sum_{x=3}^n \mu_x)c(2) + \sum_{x=3}^n \mu_x c(x)} \tag{35}
$$

となる. これを μ_1 で微分すると

$$
\begin{aligned}
\frac{\partial F_{\mu_1\mu_2}(\mu_1)}{\partial \mu_1} &= \frac{\left(1 - \sum_{x=3}^n \mu_x\right)\left(a(1)c(2) - a(2)c(1)\right)}{\left\{\mu_1 c(1) + (1 - \mu_1 - \sum_{x=3}^n \mu_x)c(2) + \sum_{x=3}^n \mu_x c(x)\right\}^2} \\
&+ \frac{\sum_{x=3}^n \left\{\left(a(1) - a(2)\right)c(x) - \left(c(1) - c(2)\right)a(x)\right\}}{\left\{\mu_1 c(1) + (1 - \mu_1 - \sum_{x=3}^n \mu_x)c(2) + \sum_{x=3}^n \mu_x c(x)\right\}^2}
\end{aligned} \tag{36}
$$

となる．分母に注目すると，2 乗の形となっている．また，分子は μ_1 によって符号変化しない値となっている．ここから，$F_{\mu_1\mu_2}(\mu_1)$ は，μ_1 に対して単調であることが分かる．関数 $F(\mu)$ の対称性より，同様の操作から，ある $x \in \mathcal{X}$ に対して，

$$\mu_x = 1 \tag{37}$$

となるような分布 μ の中に $F(\mu)$ の最大値を与えるものが存在する．つまり，コストあたり通信路容量は

$$\begin{aligned} C' \leqslant \max_{\mu} F(\mu) &= \max_{x\in\mathcal{X}} \frac{a(x)}{c(x)} \\ &= \max_{x\in\mathcal{X}} \frac{1}{c(x)} \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \right\} \end{aligned} \tag{38}$$

のように，上から抑えられるということが分かった。ここで，式 (38) の右辺を

$$I'_U(p) \triangleq \max_{x\in\mathcal{X}} \frac{1}{c(x)} \left\{ \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \right\} \tag{39}$$

と定義することで，

$$C' \leqslant I'_U(p) \tag{40}$$

を得る．

5 上界の収束

4 章で求めたコストあたり通信路容量の上界 $I'_U(p)$ を用いることで，アルゴリズムの繰り返しを途中で止めた時の算出値に精度保証を与えることができる．しかし，所望の精度がある場合，アルゴリズムを繰り返すことでその精度に達するかどうかは上記の結果からは保証できない．これを保証するためには上界 $I'_U(p)$ が繰り返し回数を大きくしたときに C' に収束することを示せばよい．一方，式 (39) の定義から $I'_U(p)$ は p に関して連続であることが分かる．したがって，最適な入力分布 p^* に対して $I'_U(p^*) = C'$ を示せば十分である．

$I'(p)$ を p_x で偏微分すると

$$
\begin{aligned}
\frac{\partial I'}{\partial p_x}(p) = & \frac{1}{\left\{\sum_x p_x c_x\right\}^2}\Bigg\{\left(\sum_x p_x c_x\right)\sum_y W(y|x)\log W(y|x) \\
& -c(x)\sum_{x,y} p_x W(y|x)\log W(y|x) - c(x)\sum_{x,y} p_x W(y|x)\log\sum_x p_x W(y|x) \\
& -\sum_x p_x c_x\left(\sum_y W(y|x)\log\sum_x p_x W(y|x) + \sum_y W(y|x)\frac{\sum_x p_x W(y|x)}{\sum_x p_x W(y|x)}\right)\Bigg\} \\
= & \frac{1}{\sum_x p_x c(x)}\sum_y W(y|x)\log W(y|x) \\
& -\frac{c(x)}{\left\{\sum_x p_x c(x)\right\}^2}\sum_{x,y} p_x W(y|x)\log W(y|x) \\
& +\frac{c(x)}{\left\{\sum_x p_x c(x)\right\}^2}\sum_{x,y} p_x W(y|x)\log\sum_x p_x W(y|x) \\
& -\frac{1}{\sum_x p_x c(x)}\left\{\sum_y W(y|x)\log\sum_x p_x W(y|x) + 1\right\} \\
= & \frac{1}{\sum_x p_x c(x)}\sum_y W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)} \\
& -\frac{c(x)}{\left\{\sum_x p_x c(x)\right\}^2}\sum_{x,y} p_x W(y|x)\log\frac{W(y|x)}{\sum_x p_x W(y|x)} \\
& -\frac{1}{\sum_x p_x c(x)}
\end{aligned} \tag{41}
$$

となる．式 (26) のような最適な入力をした際，$I'(p^*)$ はコストあたり相互情報量を最大化したものであるため，任意の実数 $\theta(0<\theta<1)$ と任意のベクトル p に対して

$$
0 \geqslant I'(\theta p + (1-\theta)p^*) - I'(p^*) \tag{42}
$$

が成立する. 式 (42) より

$$
\begin{aligned}
0 \geqslant & \frac{1}{\sum_x (\theta p_x + (1-\theta)p_x^*)c(x)} \sum_{x,y} (\theta p_x + (1-\theta)p_x^*) W(y|x) \log \frac{W(y|x)}{\sum_x (\theta p_x + (1-\theta)p_x^*) W(y|x)} \\
& - \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
= & \frac{\sum_{x,y} \theta p_x W(y|x) \log \frac{W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
& + \frac{\sum_{x,y} (1-\theta) p_x^* W(y|x) \log \frac{W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
& - \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
= & \frac{\sum_{x,y} \theta p_x W(y|x) \log \frac{W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
& + \frac{\left(\sum_x p_x c(x)\right) \sum_{x,y} (1-\theta) p_x^* W(y|x) \log \frac{\sum_x p_x^* W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\sum_x p_x^* c(x) \left(\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)\right)} \\
& - \frac{\left(\sum_x p_x c(x)\right) \sum_{x,y} \theta p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)}}{\sum_x p_x^* c(x) \left(\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)\right)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\theta \sum_{x,y} p_x W(y|x) \log \dfrac{W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad + \frac{\theta \sum_{x,y} p_x^* W(y|x) \log \left\{ \dfrac{\displaystyle\sum_x p_x^* W(y|x)}{\theta \displaystyle\sum_x p_x W(y|x) + (1-\theta) \displaystyle\sum_x p_x^* W(y|x)} \right\}^{\frac{1}{\theta}}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad - \frac{\theta \sum_{x,y} p_x^* W(y|x) \log \dfrac{\displaystyle\sum_x p_x^* W(y|x)}{\theta \displaystyle\sum_x p_x W(y|x) + (1-\theta) \displaystyle\sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad - \frac{\theta \left(\sum_x p_x c(x) \right) \sum_{x,y} p_x^* W(y|x) \log \dfrac{W(y|x)}{\sum_x p_x^* W(y|x)}}{\sum_x p_x^* c(x) \left(\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x) \right)}
\end{aligned} \tag{43}
$$

となる. 式 (43) について，$\theta(\theta > 0)$ で両辺を割ると

$$
\begin{aligned}
0 \geqslant\ & \frac{\sum_{x,y} p_x W(y|x) \log \dfrac{W(y|x)}{\theta \sum_x p_x W(y|x) + (1-\theta) \sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad + \frac{\sum_{x,y} p_x^* W(y|x) \log \left\{ \dfrac{\displaystyle\sum_x p_x^* W(y|x)}{\theta \displaystyle\sum_x p_x W(y|x) + (1-\theta) \displaystyle\sum_x p_x^* W(y|x)} \right\}^{\dfrac{1}{\theta}}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad - \frac{\sum_{x,y} p_x^* W(y|x) \log \dfrac{\displaystyle\sum_x p_x^* W(y|x)}{\theta \displaystyle\sum_x p_x W(y|x) + (1-\theta) \displaystyle\sum_x p_x^* W(y|x)}}{\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x)} \\
&\quad - \frac{\left(\sum_x p_x c(x) \right) \sum_{x,y} p_x^* W(y|x) \log \dfrac{W(y|x)}{\sum_x p_x^* W(y|x)}}{\sum_x p_x^* c(x) \left(\theta \sum_x p_x c(x) + (1-\theta) \sum_x p_x^* c(x) \right)}
\end{aligned} \tag{44}
$$

となる. ここで $\log\left\{\dfrac{\displaystyle\sum_x p_x^* W(y|x)}{\theta\displaystyle\sum_x p_x W(y|x) + (1-\theta)\displaystyle\sum_x p_x^* W(y|x)}\right\}^{\frac{1}{\theta}}$ について

$$\frac{1}{t} = \log \frac{\displaystyle\sum_x (p_x - p_x^*) W(y|x)}{\dfrac{1}{\theta}\displaystyle\sum_x p_x^* W(y|x) + \displaystyle\sum_x (p_x - p_x^*) W(y|x)} \tag{45}$$

とおくと,

$$\begin{aligned}\log\left\{\frac{\displaystyle\sum_x p_x^* W(y|x)}{\theta\displaystyle\sum_x p_x W(y|x) + (1-\theta)\displaystyle\sum_x p_x^* W(y|x)}\right\}^{\frac{1}{\theta}} &= \left(1-\frac{1}{t}\right)^{\frac{\sum_x (p_x - p_x^*) W(y|x)}{\sum_x p_x^* W(y|x)}(t-1)} \\ &= \left(\left(1-\frac{1}{t}\right)^{t}\left(1-\frac{1}{t}\right)^{-1}\right)^{\frac{\sum_x (p_x - p_x^*) W(y|x)}{\sum_x p_x^* W(y|x)}} \end{aligned} \tag{46}$$

となる. ここで, $\left(1-\dfrac{1}{t}\right)^{t}$ について $t \to \infty$ とすると,

$$\lim_{t\to+\infty}\left(1-\frac{1}{t}\right)^{t} = \frac{1}{e} \tag{47}$$

$$\begin{aligned}\lim_{t\to-\infty}\left(1-\frac{1}{t}\right)^{t} &= \lim_{s\to+\infty}\left(1-\frac{1}{s}\right)^{-s}, s=-t \\ &= \lim_{s\to+\infty}\left(1-\frac{1}{s+1}\right)^{s+1}\left(1+\frac{1}{s}\right) \\ &= \frac{1}{e}\end{aligned} \tag{48}$$

という結果が得られる．式 (45) から (48) より，式 (44) の式について,$\theta \to 0$ とすると

$$
\begin{aligned}
0 \geqslant & \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& - \frac{1}{\left\{ \sum_x p_x^* c(x) \right\}^2} \sum_x p_x c(x) \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
= & \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& - \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& - \frac{1}{\left\{ \sum_x p_x^* c(x) \right\}^2} \sum_x p_x c(x) \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& + \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
= & \sum_x (p_x - p_x^*) \left\{ \frac{1}{\sum_x p_x^* c(x)} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \right. \\
& \left. - \frac{c(x)}{\left\{ \sum_x p_x^* c(x) \right\}^2} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \right\} \\
= & \sum_x (p_x - p_x^*) \left\{ \frac{1}{\sum_x p_x^* c(x)} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \right. \\
& \left. - \frac{c(x)}{\left\{ \sum_x p_x^* c(x) \right\}^2} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} - \frac{1}{\sum_x p_x^* c(x)} \right\} \\
= & \sum_x (p_x - p_x^*) \frac{\partial I'}{\partial p_x}(p^*)
\end{aligned} \tag{49}
$$

となる. ただし，$\dfrac{1}{\sum_x p_x^* c(x)}$ は x について動かず，定数として扱うことが出来るため

$$
\sum_x (p_x - p_x^*) \left\{ \frac{1}{\sum_x p_x^* c(x)} \right\} = 0 \tag{50}
$$

である.

ここで，$p^* = \{p_x^*\}_{x \in \mathcal{X}}$ は，その成分の中に正のものが少なくとも 1 つある. それを

$p_1^*(p_1^* > 0)$ とおく．また，一般に第 a 成分 $(a \in \mathcal{X})$ が 1 で他の成分は 0 なるベクトル

$$e_a \triangleq \{e_a(x)\}_{x\in\mathcal{X}} \tag{51}$$

を定義する．そこで，$0 < \epsilon \leqslant p_1^*$ を満たす任意の ϵ に対してベクトル

$$p = p^* + \epsilon e_a - \epsilon e_1 \tag{52}$$

を定める. 式 (49),(52) より

$$\begin{aligned} 0 \geqslant & \sum_{x\in\mathcal{X}} \frac{\partial I'}{\partial p_x}(p^*)\Big(p_x^* + \epsilon e_a(x) - \epsilon e_1(x) - p_x^*\Big) \\ &= \frac{\partial I'}{\partial p_a}(p^*)\epsilon e_a(a) - \frac{\partial I'}{\partial p_1}(p^*)\epsilon e_1(1) \\ &= \epsilon\frac{\partial I'}{\partial p_a}(p^*) - \epsilon\frac{\partial I'}{\partial p_1}(p^*) \end{aligned} \tag{53}$$

となる. よって任意の $a \in \mathcal{X}$ に対して

$$\frac{\partial I'}{\partial p_a}(p^*) \leqslant \frac{\partial I'}{\partial p_1}(p^*) \tag{54}$$

といえる. ここで $p_a^* > 0$ の時，式 (52) において ϵ を $-p_a^* < \epsilon < 0$ とすることが出来る. その時，式 (53) から

$$\frac{\partial I'}{\partial p_a}(p^*) \geqslant \frac{\partial I'}{\partial p_1}(p^*) \tag{55}$$

が成り立つ. よって式 (54),(55) より，$p_a^* > 0$ なる任意の $a \in \mathcal{X}$ に対して

$$\frac{\partial I'}{\partial p_a}(p^*) = \frac{\partial I'}{\partial p_1}(p^*) \tag{56}$$

が成り立つ. ここで，a を x と書き直せば，式 (54),(56) より，$\dfrac{\partial I'}{\partial p_1}(p^*)$ を λ とおくと

$$\begin{cases} \dfrac{\partial I'}{\partial p_x}(p^*) = \lambda, \text{if } p_x^* > 0 \\ \dfrac{\partial I'}{\partial p_x}(p^*) \leqslant \lambda, \text{otherwise} \end{cases} \tag{57}$$

が得られる．ここで，λ について解く．式 (57) に注意すると

$$\begin{aligned}
\lambda &= \sum_x p_x^* \lambda \\
&= \sum_x p_x^* \frac{\partial I'}{\partial p_x}(p^*) \\
&= \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
&\quad - \frac{\sum_x p_x^* c(x)}{\left\{\sum_x p_x^* c(x)\right\}^2} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
&\quad - \frac{\sum_x p_x^*}{\sum_x p_x^* c(x)} \\
&= -\frac{1}{\sum_x p_x^* c(x)}
\end{aligned} \tag{58}$$

となる．よって，式 (57) は

$$\left\{ \begin{array}{l}
\dfrac{\partial I'}{\partial p_x}(p^*) = -\dfrac{1}{\sum_x p_x^* c(x)}, \text{if } p_x^* > 0 \\
\dfrac{\partial I'}{\partial p_x}(p^*) \leqslant -\dfrac{1}{\sum_x p_x^* c(x)}, \text{ otherwise}
\end{array} \right. \tag{59}$$

と書き換えられる．この時，式 (41) より任意の $x \in \mathcal{X}$ に対して

$$\begin{aligned}
-\frac{1}{\sum_x p_x^* c(x)} \geqslant& \frac{1}{\sum_x p_x^* c(x)} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& - \frac{c(x)}{\left\{\sum_x p_x^* c(x)\right\}^2} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
& - \frac{1}{\sum\limits_x p_x^* c(x)}
\end{aligned} \tag{60}$$

となる．両辺に $\sum\limits_x p_x^* c(x)$ をかけて変形すると

$$\begin{aligned}
\frac{1}{c(x)} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \leqslant& \frac{1}{\sum_x p_x^* c(x)} \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\
=& C'
\end{aligned} \tag{61}$$

となる. 式 (61) の左辺に対して，x についての最大値を取っても不等号の向きは変わらないため

$$C' \geqslant \max_x \frac{1}{c(x)} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \tag{62}$$

である. ところが，この右辺は $I'_U(p^*)$ であるため，

$$C' \geqslant I'_U(p^*) \tag{63}$$

を得る. 一方，式 (39) で $p = p^*$ とおけば

$$C' \leqslant I'_U(p^*) \tag{64}$$

となる. したがって，式 (63),(64) より

$$C' = I'_U(p^*) \tag{65}$$

となる. よって，導出した上界がコストあたり通信路容量に収束することが分かった.

6 まとめ

本研究では，川合西新提案のコストあたり通信路容量を算出するアルゴリズムにおけるバウンドの算出をすることによってアルゴリズムの算出値の精度保証を試みた. 結果として，アルゴリズム内のループを精度を保証しながら抜け出せるバウンドを導出することができた. だが，3 章で紹介した式 (23) を満たす $\alpha^{(N)}$ を算出するために，ループ構造を持つ計算をする必要がある. この $\alpha^{(N)}$ が本研究で行った精度保証に対して，どのような影響を及ぼすかは調査できていないため，今後の課題である.

謝辞

本研究を進めるにあたり，丁寧なご指導を頂いた西新幹彦准教授に感謝の意を表する.

参考文献

[1] 川合康太，西新幹彦「離散無記憶通信路に対するコストあたり通信路容量の算出に関する考察」，The 42nd Symposium on Information Theory and its Applications (SITA2019)，no2.2.3，pp.117–122, 2019.

[2] Suguru Arimoto, “An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels,” IEEE Transactions on Information Theory, vol.IT-18, no.1, pp.14–20, Jan. 1972.

[3] Robert G. Gallager, Information Theory and Reliable Communication, John Wiley & Sons, 1968.

[4] Richard E. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," IEEE Transactions on Information Theory, vol.IT-18, no.4, pp.460–473, Jul. 1972.

[5] Thomas M. Cover and Joy A. Thomas,"Elements of Information Theory second edition ", John Wiley & Sons, 2006.

付録 A

A.1 通信路容量を求めるアルゴリズムの上界の導出

コストに関係ない，通信路容量を求めるアルゴリズム [2] の上界の導出を示す．通信路容量 C は,

$$\begin{aligned} C &= \sup_p I(p) \\ &= I(p^*) \end{aligned} \tag{66}$$

とも書ける. ただし p^* は式 (66) を満たす入力分布とする．これに対応する出力の分布を q^* と表す．これらの分布は未知であることに注意する．通信路容量 C は

$$\begin{aligned} C &= \sum_{x,y} p_x^* W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\ &= \sum_x p_x^* \sum_y W(y|x) \log \frac{\sum_x p_x W(y|x)}{\sum_x p_x^* W(y|x)} + \sum_x p_x^* \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \\ &= -D(q^* \parallel q) + \sum_x p_x^* \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \end{aligned} \tag{67}$$

となる. なお，p_x は有本-Blahut アルゴリズム [2][4] 中に計算する値であるため，未知数ではない. ここに含まれるダイバージェンス $D(q^* \parallel q)$ は非負より

$$\begin{aligned} C &\leqslant \sum_x p_x^* \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \\ &\leqslant \max_{x \in \mathcal{X}} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \end{aligned} \tag{68}$$

となる. ここで，式 (68) の右辺を

$$I_U(p) \triangleq \max_{x \in \mathcal{X}} \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} \tag{69}$$

と定義することで,

$$C \leqslant I_U(p) \tag{70}$$

を得る.

A.2 入力確率分布が相互情報量を最大にする必要条件

式 (66) のような最適な入力をする際, $I(p^*)$ は相互情報量を最大化したものであるため, 任意の実数 $\theta(0<\theta<1)$ と任意のベクトル p に対して

$$0 \geqslant I(\theta p + (1-\theta)p^*) - I(p^*) \tag{71}$$

が成立する. ここで, 両辺を θ で割り, $\theta \to 0$ とすると

$$\begin{aligned} 0 \geqslant& \lim_{\theta\to 0} \frac{I(\theta p + (1-\theta)p^*) - I(p^*)}{\theta} \\ =& \sum_x (p_x - p_x^*) \frac{\partial I}{\partial p_x}(p^*) \end{aligned} \tag{72}$$

となる [1]. ただし, $\dfrac{\partial I}{\partial p_x}(p)$ は

$$\begin{aligned} \frac{\partial I}{\partial p_x}(p) =& \sum_x p_x \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} - \sum_x p_x \sum_{x,y} p_x W(y|x) \frac{W(y|x)}{\sum_x p_x W(y|x)} \\ =& \sum_x p_x \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x W(y|x)} - 1 \end{aligned} \tag{73}$$

である. ここで, $p^* = \{p_x^*\}_{x\in\mathcal{X}}$ は, その成分の中に正のものが少なくとも 1 つある. それを $p_1^*(p_1^*>0)$ とおく. また, 一般に第 a 成分 $(a\in\mathcal{X})$ が 1 で他の成分は 0 なるベクトル

$$e_a \triangleq \{e_a(x)\}_{x\in\mathcal{X}} \tag{74}$$

を定義する. そこで, $0<\epsilon\leqslant p_1^*$ を満たす任意の ϵ に対してベクトル

$$p = p^* + \epsilon e_a - \epsilon e_1 \tag{75}$$

を定める. 式 (72),(75) より

$$\begin{aligned} 0 \geqslant& \sum_{x\in\mathcal{X}} \frac{\partial I}{\partial p_x}(p^*) \Big(p_x^* + \epsilon e_a(x) - \epsilon e_1(x) - p_x^* \Big) \\ =& \frac{\partial}{\partial p_a} I(p^*) \epsilon e_a(a) - \frac{\partial}{\partial p_1} I(p^*) \epsilon e_1(1) \\ =& \epsilon \frac{\partial}{\partial p_a} I(p^*) - \epsilon \frac{\partial}{\partial p_1} I(p^*) \end{aligned} \tag{76}$$

となる. よって任意の $a \in \mathcal{X}$ に対して

$$\frac{\partial}{\partial p_a} I(p^*) \leqslant \frac{\partial}{\partial p_1} I(p^*) \tag{77}$$

といえる. ここで $p_a^* > 0$ の時，式 (75) において ϵ を $-p_a^* < \epsilon < 0$ とすることが出来る. その時，式 (76) から

$$\frac{\partial}{\partial p_a} I(p^*) \geqslant \frac{\partial}{\partial p_1} I(p^*) \tag{78}$$

が成り立つ. よって式 (77),(78) から，$p_a^* > 0$ なる任意の $a \in \mathcal{X}$ に対して

$$\frac{\partial I}{\partial p_a}(p^*) = \frac{\partial I}{\partial p_1}(p^*) \tag{79}$$

が成り立つ. 式 (77),(79) より，a を x と書き直し，$\dfrac{\partial I}{\partial p_1}(p^*)$ をλ とおくと

$$\begin{cases} \dfrac{\partial I}{\partial p_x}(p^*) = \lambda, \text{if } p_x^* > 0 \\ \dfrac{\partial I}{\partial p_x}(p^*) \leqslant \lambda, \text{otherwise} \end{cases} \tag{80}$$

が得られる. ここで，λ について解く. 式 (80) について注意すると

$$\begin{aligned} \lambda &= \frac{\partial I}{\partial p_x}(p^*) \\ &= \sum_x p_x^* \frac{\partial I}{\partial p_x}(p^*) \\ &= \sum_x p_x^* \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} - \sum_x p_x^* \sum_{x,y} p_x^* W(y|x) \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \\ &= \sum_x p_x^* \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} - 1 \\ &= -1 + I(p^*) \\ &= -1 + C \end{aligned} \tag{81}$$

となる. よって，式 (80) は

$$\begin{cases} \displaystyle\sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} = C, \text{if } p_x^* > 0 \\ \displaystyle\sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \leqslant C, \ \text{otherwise} \end{cases} \tag{82}$$

と書き替えられる.

A.3　通信路容量に対して上界の収束証明

A.1 において求めた上界 I_U が通信路容量に収束することを証明する．条件式 (82) より任意の p_x^* に対して

$$C \geqslant \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \tag{83}$$

である．式 (83) の右辺に対して x について最大値をとっても，不等号の向きは変わらない．よって，

$$C \geqslant \max_x \sum_y W(y|x) \log \frac{W(y|x)}{\sum_x p_x^* W(y|x)} \tag{84}$$

となる．ところが，右辺は $I_U(p^*)$ であるため

$$C \geqslant I_U(p^*) \tag{85}$$

を得る．一方，式 (69) で $p = p^*$ とおくと

$$C \leqslant I_U(p^*) \tag{86}$$

となる．したがって，式 (85),(86) より

$$C = I_U(p^*) \tag{87}$$

となる．よって，A.1 で求めた上界は通信路容量に収束することがわかる．