

# 相互最尤推定の収束性について

○ 津村 幸治 (東京大学)

## Convergence on mutual maximum likelihood estimation

\* K. Tsumura (The University of Tokyo)

**Abstract**– In this paper, we deal with a bilateral automatic negotiation system composed of two players. Each player proposes offers to the other mutually based on its own utility and strategy with the estimation on those information of the other's and the estimation is realized by the observation on the series of the offers by the other player. Therefore, the total system has a feedback structure between two players and the agreement of the negotiation is regarded as a convergence to their equilibrium states. We furthermore suppose that the offers are generated probabilistically based on the utility and optimal likelihood estimation is utilized for the opponent estimation. Then, in order to discuss the condition for the agreement, we analyze the convergence of the mutual optimal likelihood estimation between two players.

**Key Words:** Automatic negotiation, Feedback system, Optimal likelihood estimation

### 1 はじめに

近年、理工学が対象とする大きな問題の一つは、物理システムと情報システムといった多数の異質なシステムが互いに作用するヘテロな大規模システムをどのように調和させるのかというものである。この問題はサイバーフィジカルシステム (CPS) や Society5.0 といった名称で取り組まれている。

この課題に寄与する研究対象の一つが自動交渉である。従来「交渉」は人同士や企業・団体同士の間で行われるものであったが、ヘテロな大規模システムにおける調和の実現には、異なる目的を持つ人や企業といったシステム間、あるいは自動運転された自動車というシステム間、さらには人とモノとの間で高速で自動的に折り合う技術、つまり自動交渉システムの開発・実現が必要である。

このような背景をもとに自動交渉に関する研究は活発に行われており、例えば自動交渉アルゴリズム同士を競わせるコンペティションが毎年開催されている<sup>1)</sup>。一般に交渉は<sup>2,3)</sup>、2者間の場合、playerの一方から相手のplayerへの条件案のoffer、offerを受けたplayerによる案のaccept、あるいはrejectし対案のoffer、および以上の繰り返しや打ち切りなどからなると考えられている。offerは自身の効用と戦略および相手のそれらの推定結果によって決定する。また1対多、多対多など、様々な交渉の形態が考えられている。コンペティションではこのようなアルゴリズム開発と制度設計を目標とされているが、学術的な観点での戦略の最適性、交渉成立に至る交渉過程の解析、環境変化や外乱へのロバストネスなどに関する研究は必ずしも十分ではない。

以上の課題を解決するために本研究グループでは、自動交渉を制御システムの問題として取り組んでいる。つまり自動交渉システムをplayer間のofferの提示と、それを受けての代案提示の繰り返しという、動的システム同士のフィードバックシステムととらえる。そしてこのようなフィードバックシステムにおける交渉条件の収束性、戦略の最適性、あるいは自動交渉システムの頑健性について、制御理論の方法論に従って解析する

ことを目的としている。以上の考えに基づき本研究グループでは、文献<sup>4)</sup>において基本となる2者間自動交渉システムの数理モデルを導出した。そのモデルでは、各playerの動作は相手のofferを入力とし、相手の戦略の推定を通して、自身の戦略に基づく最適なofferを対案として出力するという入出力システムとしてモデリングされ、システム全体は2者の入出力システムからなるフィードバックシステムとして与えられている。

本研究ではこの数理モデルに基づき、自動交渉において最も基本となるシステムの収束性を解析する。具体的には相手の戦略の推定を最尤推定とし、一方offerはその推定値と自身の戦略に基づき、確率的に生成されるものとするにより、システム全体を相互最尤推定を行うフィードバックシステムととらえる。そしてこの相互最尤推定の収束性を解析し、システム全体の収束性を論じる。

### 2 2者間自動交渉のモデリング

ここでは文献<sup>4)</sup>に従って、2者間 (player 1, player 2) の交渉モデルを導入する。各playerの振る舞いを表す信号、関数およびその記号を次に与える。

player 1 (player 2 も同様)

- $z_1[t]$ : player 1 の戦略を特徴付ける変数
- $\tilde{z}_1[t]$ : player 2 による  $z_1[t]$  の推定値
- $z_2[t]$ : player 2 の戦略を特徴付ける変数  $z_2[t]$  の推定値
- $y_1[t]$ : player 1 の提案
- $u_1[t]$ : player 2 からの提案、つまり  $u_1[t] = y_2[t]$
- $f_1(\cdot)$ :  $\tilde{z}_2$  を与える推定器
- $h_1(\cdot)$ :  $\tilde{z}_1$  を与える推定器
- $g_1(\cdot)$ :  $z_1, \tilde{z}_2$  に基づく player 1 の提案  $y_1$  を生成するサブシステム

時刻  $t$  における player 1, player 2 の動作を次とする。

player 1

$$\begin{aligned}\tilde{z}_2[t] &= f_1(\tilde{z}_1[1:t], u_1[1:t](= y_2[1:t])) \\ \tilde{z}_1[t+1] &= h_1(y_1[1:t], \tilde{z}_2[1:t]) \\ y_1[t] &= g_1(z_1, \tilde{z}_2[t])\end{aligned}\quad (1)$$

player 2

$$\begin{aligned}\tilde{z}_1[t] &= f_2(\tilde{z}_2[1:t], u_2[1:t](= y_1[1:t])) \\ \tilde{z}_2[t+1] &= h_2(y_2[1:t], \tilde{z}_1[1:t]) \\ y_2[t] &= g_2(z_2, \tilde{z}_1[t])\end{aligned}\quad (2)$$

player 1 において,  $g_1$  は自身の戦略  $z_1$  と相手の戦略の推定値  $\tilde{z}_2$  に基づき, 提案  $y_1$  を生成するサブシステムである. 次に  $f_1$  は, 相手の提案の観測履歴  $u_1[1:t](= y_2[1:t])$  に基づき, それを生成した  $g_2$  で用いられている  $z_2$  の推定器である. 注意する点は,  $g_2$  で用いられている  $\tilde{z}_1[t]$  は, ノイズやモデルの不確かさがなければ player 1 でも計算可能であり, サブシステム  $h_1$  で与えられ,  $f_1$  で用いられる. また player 2 でも同様である.

### 3 等価システム

ここでは先述の自動交渉システムの等価システムを導く. まずシステム全体の構成を Fig. 1 に示す.

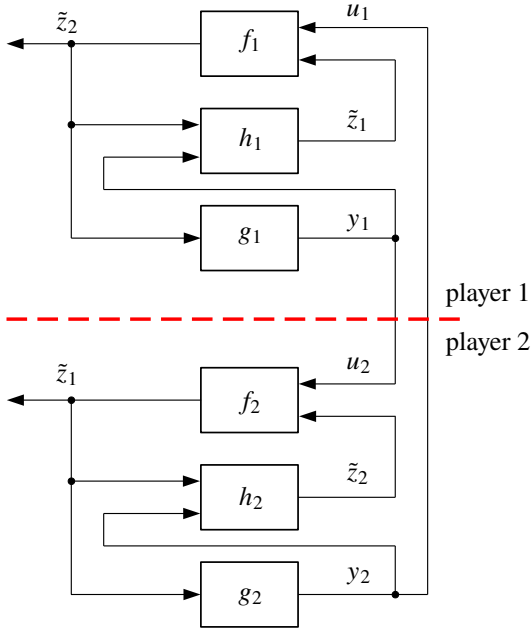


Fig. 1: Structure of the bilateral negotiation system

この図からわかるように player 1 と player 2 は信号  $y_1(= u_2)$ ,  $y_2(= u_1)$  のみをやりとりしている. つまり player 1 内の計算を進めるには, 外部からは信号  $y_2(= u_1)$  のみを必要とする. 一方,  $y_2$  は player 2 のサブシステム  $g_2$  において, 入力信号としては  $\tilde{z}_1$  のみを用いて計算される. またサブシステム  $h_1$  と  $f_2$  (あるいは  $h_2$  と  $f_1$ ) は互いに既知の情報を用いて同じ計算が行われている. そこで Fig. 1 と等価で縮小されたシステムとして次の Fig. 2 が導かれる.

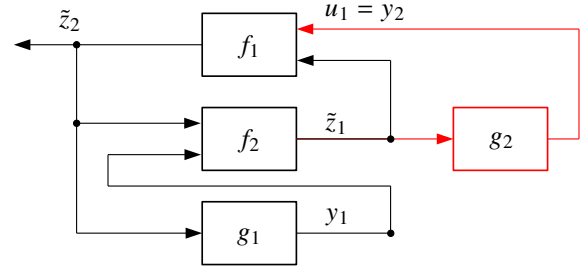


Fig. 2: Equivalent reduced system (I)

また Fig. 2 は見やすいように Fig. 3 と変形できる.

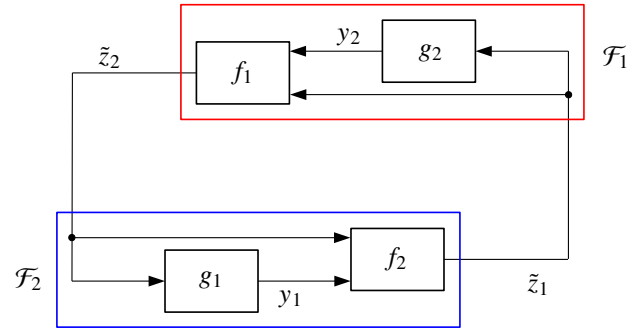


Fig. 3: Equivalent reduced system (II)

Fig. 3 からシステム全体は, サブシステム  $f_1, g_2$  からなる  $\mathcal{F}_1$  と, サブシステム  $f_2, g_1$  からなる  $\mathcal{F}_2$  との縮小されたフィードバック系とみなせることがわかる.

### 4 収束性

本章では Fig. 3 のシステムの収束性について順を追って議論する.

#### 4.1 サブシステム $g_1, g_2, f_1, f_2$ と相互最尤推定

以降では式 (1) のサブシステム  $g_1$  における offer  $y_1[t]$  の導出を具体的に定める. まず player 1 の効用関数

$$\begin{aligned}p_1(y_1, z_1, \tilde{z}_2) \\ p_1 \geq 0, \int p_1(y_1, z_1, \tilde{z}_2) dy_1 = 1\end{aligned}\quad (3)$$

が存在し, offer  $y_1[t]$  はその効用関数を確率密度関数と同等視し, その確率密度関数に従い独立に選ばれるものとする.  $p_1(y_1, z_1, \tilde{z}_2)$  はさらに次を満たすものとする.

正則条件<sup>5)</sup>

- (1)  $\{y_1 : p_1(y_1, z_1, \tilde{z}_2) > 0\}$  は  $z_1, \tilde{z}_2$  に依存しない.
- (2)  $\{(z_1, \tilde{z}_2)\}$  は開集合で有限.
- (3)  $p_1(y_1, z_1, \tilde{z}_2)$  は  $z_1$  について連続.
- (4)  $(z_1, \tilde{z}_2) \neq (z'_1, \tilde{z}'_2)$  なら  $p_1(y_1, z_1, \tilde{z}_2) \neq p_1(y_1, z'_1, \tilde{z}'_2)$ .

サブシステム  $g_2$  における offer  $y_2[t]$  も同様の正則条件下で選ばれるものとする. これらに従い, 式 (1) のサブシステム  $f_1$  での計算は,

$$\tilde{z}_2 = \arg \max_{z_2} \prod_{t=1}^T p_2(y_2[t]; z_2, \tilde{z}_1[t])$$

と表せる。対数尤度

$$\ell_2(y_2[t]; z_2, \tilde{z}_1[t]) := \log(p_2(y_2[t]; z_2, \tilde{z}_1[t])) \quad (4)$$

を用いると、上式は

$$\tilde{z}_2 = \arg \max_{z_2} \sum_{t=1}^T \ell_2(y_2[t]; z_2, \tilde{z}_1[t])$$

と表せる。サブシステム  $f_2$  も同様とする。

Fig. 3 のサブシステム  $\mathcal{F}_1$  は、このような  $g_2, f_1$  から構成されているが、 $g_1$  は  $y_2$  を発生する信号源、 $f_1$  はサンプル列  $y_2[1 : T]$  を用いた  $z_2$  の最尤推定器とみなせる。スタンダードな最尤推定と異なる点は、 $g_2, f_1$  ともに毎時刻、共通する  $\tilde{z}_1[t]$  によって密度関数  $p_2$  が変化する、ということである。 $z_{21} := (z_2, \tilde{z}_1)$  とすると密度関数  $p_2$  は  $z_{21}$  により関数が定まり、 $f_1$  においてはパラメタ  $z_{21}$  のうちの一部、つまり  $z_2$  のみが未知パラメタであり、 $\tilde{z}_1$  は毎時刻変化するが既知のパラメタとなる。サブシステム  $\mathcal{F}_2$  も同様である。

以下ではこのような互いに最尤推定し合う  $\mathcal{F}_1, \mathcal{F}_2$  からなる閉ループ系の収束、つまり相互最尤推定の収束性について議論する。まず  $\mathcal{F}_1$  の収束性について解析する。

#### 4.2 問題設定

上記収束性の問題を一般化した定式化のもとで進める。まず次の時刻  $t$  における対数尤度関数を定義する。

$$Q^t(\alpha; X^t, \beta^t) := \frac{1}{t} \sum_{i=1}^t \ell(x_i, \alpha, \beta_i) \quad (5)$$

$$\ell(x_i, \alpha, \beta_i) := \log p(x_i, \alpha, \beta_i)$$

ここに  $x_i$  は時刻  $i$  における観測値、 $\alpha$  は推定すべきパラメタ、 $\beta_i$  は時刻  $i$  での既知のパラメタ、 $\bullet^t$  は  $\bullet$  の時刻  $t$  までの数列とする。

この  $Q^t(\alpha; X^t, \beta^t)$  の期待値は次で計算される。

$$\begin{aligned} & E_o[Q^t(\alpha; X^t, \beta^t)] \\ &= \int \frac{1}{t} \sum_{i=1}^t \ell(x_i, \alpha, \beta_i) \prod_{i=1}^t p(x_i, \alpha, \beta_i) dx_i \\ &= \frac{1}{t} \sum_{i=1}^t \int \ell(x_i, \alpha, \beta_i) p(x_i, \alpha, \beta_i) dx_i \\ &= \frac{1}{t} \sum_{i=1}^t E_o[\ell(x_i, \alpha, \beta_i)] =: \frac{1}{t} \sum_{i=1}^t Q_{oi}(\alpha; \beta_i) \\ &=: Q_o^t(\alpha; \beta^t) \end{aligned} \quad (6)$$

ただし  $\alpha_o$  は  $\alpha$  の真値とする。

ここで次の3つの問題を考える。

#### 問題 4.1

Q1 式 (6) の  $Q_o^t(\alpha; \beta^t)$ 、つまり時刻  $t$  における真の期待値としての対数尤度関数は、観測値の数  $t$  の増

加に伴い何らかの関数にどのように収束するのか、つまり

$$Q_o^t(\alpha; \beta^t) \xrightarrow{t \rightarrow \infty} ?$$

の右辺の存在、存在する場合の収束の種類、およびその関数はなにか。

Q2 式 (6) の  $Q_o^t(\alpha; \beta^t)$  の意味での最尤推定値

$$\bar{\alpha}_t := \arg \max_{\alpha} Q_o^t(\alpha; \beta^t) \quad (7)$$

は、 $t$  の増加に伴い真値  $\alpha_o$  に収束するのか、つまり

$$\bar{\alpha}_t \xrightarrow{t \rightarrow \infty} \alpha_o ?$$

が成り立つか否か。

Q3 推定で用いられる対数尤度、つまり式 (5) の  $Q^t(\alpha; X^t, \beta^t)$  に基づく最尤推定値

$$\tilde{\alpha}_t := \arg \max_{\alpha} Q^t(\alpha; X^t, \beta^t) \quad (8)$$

は、 $t$  の増加に伴い真値  $\alpha_o$  に収束するのか、つまり

$$\tilde{\alpha}_t \xrightarrow{t \rightarrow \infty} \alpha_o ?$$

か

$$\tilde{\alpha}_t \xrightarrow{t \rightarrow \infty} \bar{\alpha}_t \xrightarrow{t \rightarrow \infty} \alpha_o ?$$

が成り立つか否か。

以下では問題 Q1-3 を順を追って議論する。

#### 4.3 Q1 について

パラメタ  $\beta_i$  の存在範囲を

$$\beta_i \in \mathcal{B}, \mathcal{B} : \text{bounded} \quad (9)$$

と表す。このとき、任意に固定された  $\alpha$  に対して、写像

$$Q_{oi}(\alpha; \beta_i) : \beta_i \rightarrow Q_{oi}(\alpha; \beta_i) \quad (10)$$

は有界である。以下では  $Q_{oi}(\alpha; \beta_i)$  の値域を  $\mathcal{K}_{\alpha}$  と表す。

以上の準備のもと、固定された  $\alpha$  における  $Q_o^t(\alpha; \beta^t)$  の数列の差分について次が成り立つ。

$$\begin{aligned} & |Q_o^{t+1}(\alpha; \beta^{t+1}) - Q_o^t(\alpha; \beta^t)| \\ &= \left| \frac{1}{t+1} \sum_{i=1}^{t+1} Q_{oi}(\alpha; \beta_i) - \frac{1}{t} \sum_{i=1}^t Q_{oi}(\alpha; \beta_i) \right| \\ &= \frac{1}{t+1} \left| Q_{o(t+1)}(\alpha; \beta_{t+1}) - \frac{1}{t} \sum_{i=1}^t Q_{oi}(\alpha; \beta_i) \right| \\ &\leq \frac{1}{t+1} \kappa_{\alpha} \end{aligned} \quad (11)$$

ただし

$$\kappa_{\alpha} := \max_{Q_{oj}(\alpha; \beta_j), Q_{ok}(\alpha; \beta_k) \in \mathcal{K}_{\alpha}} |Q_{oj}(\alpha; \beta_j) - Q_{ok}(\alpha; \beta_k)| \quad (12)$$

である。

さらにある  $\kappa > 0$  が存在して、 $\kappa_\alpha \leq \kappa, \forall \alpha$  を仮定する。このとき

$$\sup_{\alpha} |Q_o^{t+1}(\alpha; \beta^{t+1}) - Q_o^t(\alpha; \beta^t)| \leq \frac{1}{t+1} \kappa \xrightarrow{t \rightarrow \infty} 0 \quad (13)$$

が成り立つ。よって

$$\lim_{t \rightarrow \infty} Q_o^t(\alpha; \beta^t) =: Q_o(\alpha; \beta^\infty) \text{ (uniform)} \quad (14)$$

が帰結される。

#### 4.4 Q2 について

まず

$$\arg \max_{\alpha} Q_{oi}(\alpha; \beta^i) = \alpha_o, \forall \beta^i \in \mathcal{B} \quad (15)$$

が成り立つ。よって

$$\alpha_o = \arg \max_{\alpha} Q_o(\alpha; \beta^\infty) \quad (16)$$

である。つまり

$$\bar{\alpha}_t = \alpha_o, \forall t \quad (17)$$

が成り立つ。

#### 4.5 Q3 について

まず  $\mathcal{B}$  の部分集合として、 $\beta$  を中心とし、半径  $\delta$  の  $\delta$ -ball を

$$B(\beta, \delta) := \{\tilde{\beta} \mid \|\tilde{\beta} - \beta\| \leq \delta, \tilde{\beta} \in \mathcal{B}\} \quad (18)$$

と定義し、 $M$  個の  $\delta$ -ball で  $\mathcal{B}$  を被覆する、つまり

$$\bigcup_{j=1}^M B(\beta_{(j)}, \delta) \supset \mathcal{B} \quad (19)$$

$$\beta_{(j)} \in \mathcal{B}$$

とすることを考える。また各  $\delta$ -ball に含まれる  $\beta_i$  の個数を

$$\#B(\beta_{(j)}, \delta) := \text{number of } \{\beta_i\} \text{ included in } B(\beta_{(j)}, \delta) \quad (20)$$

と表し、一般性を失わず、

$$\#B(\beta_{(1)}, \delta) \geq \#B(\beta_{(j)}, \delta), j \neq 1 \quad (21)$$

とインデックス付けする。ここに集合  $\{\beta_{(j)}\}_{j=1}^M$  は各  $\delta$ -ball の中心点の集合である。さらに集合  $\{\beta_{(j)}\}_{j=1}^M$  の部分集合  $\{\beta_{(k)}\}$  を、

$$0 < \lim_{t \rightarrow \infty} \frac{\#B(\beta_{(k)}, \delta)}{\#B(\beta_{(1)}, \delta)} = w_k \quad (22)$$

を満たすものとする。つまり各  $\delta$ -ball 内に「十分多く」の  $\beta_i$  を含む  $\delta$ -ball の中心点の集合である。そのような

$\{\beta_{(k)}\}$  を、インデックスを  $k = 1, 2, \dots, M(\delta)$  と置き直して

$$\{\beta_{(k)}\} = \{\beta_{(k)}\}_{k=1}^{M(\delta)} \quad (23)$$

と表す。ただし  $M(\delta)$  は条件 (22) を表す  $\beta_{(k)}$  の数を表す。さらに  $\beta_i$  のインデックスを

$$\{\beta_{(1)}^1, \beta_{(1)}^2, \dots\} \in B(\beta_{(1)}, \delta),$$

$$\{\beta_{(2)}^1, \beta_{(2)}^2, \dots\} \in B(\beta_{(2)}, \delta),$$

$$\dots,$$

$$\{\beta_{(M(\delta))}^1, \beta_{(M(\delta))}^2, \dots\} \in B(\beta_{(M(\delta))}, \delta),$$

$$\beta_1, \beta_2, \dots \quad (24)$$

とつけ直す。

次に時刻  $t$  において  $B(\beta_{(j)}, \delta)$  に含まれる  $\beta_i$  の数を  $t_j$  とおき、

$$Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j}) := \frac{1}{t_j} \sum_{i=1}^{t_j} \ell(x_{(j)}^i, \alpha, \beta_{(j)}^i) \quad (25)$$

を定義する。これは次の誤差関数

$$\Delta(x_{(j)}^i, \alpha, \beta_{(j)}^i, \beta_{(j)}) := \ell(x_{(j)}^i, \alpha, \beta_{(j)}^i) - \ell(x_{(j)}^i, \alpha, \beta_{(j)}) \quad (26)$$

を用いると、

$$Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j})$$

$$= \frac{1}{t_j} \sum_{i=1}^{t_j} \left( \ell(x_{(j)}^i, \alpha, \beta_{(j)}) + \Delta(x_{(j)}^i, \alpha, \beta_{(j)}^i, \beta_{(j)}) \right)$$

$$= Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}) + \frac{1}{t_j} \sum_{i=1}^{t_j} \Delta(x_{(j)}^i, \alpha, \beta_{(j)}^i, \beta_{(j)}) \quad (27)$$

$$Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}) := \frac{1}{t_j} \sum_{i=1}^{t_j} \ell(x_{(j)}^i, \alpha, \beta_{(j)}) \quad (28)$$

と表せる。

確率密度関数についての正則条件から、誤差関数についてある  $\phi_\delta > 0$  が存在し

$$|\Delta(x_{(j)}^i, \alpha, \beta_{(j)}^i, \beta_{(j)})| \leq \phi_\delta \quad (29)$$

が成り立つ。よって

$$|Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j}) - Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)})| \leq \phi_\delta \quad (30)$$

が導かれる。

一方、正規条件から

$$Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}) \xrightarrow{t_j \rightarrow \infty} \mathbb{E}_o[\ell(x_{(j)}^i, \alpha, \beta_{(j)})]$$

$$=: Q_{o(j)}(\alpha; \beta_{(j)}) \text{ (uniform)} \quad (31)$$

が成り立つ。上式の収束は次で表現される。つまり任意に与えた  $\epsilon, \eta > 0$  に対してデータ数  $N_1(\epsilon, \eta) > 0$  が存在し、

$$\Pr(\sup_{\alpha} |Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}) - Q_{o(j)}(\alpha; \beta_{(j)})| > \epsilon) < \eta$$

$$\text{when } t_j > N_1(\epsilon, \eta) \quad (32)$$

が成り立つ。よってこれより直ちに、

$$\begin{aligned} & \Pr(\sup_{\alpha} |Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j}) - Q_{o(j)}(\alpha; \beta_{(j)})| > \phi_{\delta} + \epsilon) \\ & < \eta \\ & \text{when } t_j > N_1(\epsilon, \eta) \end{aligned} \quad (33)$$

が導かれる。

また式 (29) と同様に、ある  $\xi_{\delta} > 0$  が存在し

$$|Q_{o(j)}(\alpha; \beta_{(j)}) - Q_{o(j)}(\alpha; \beta_{(j)}^i)| < \xi_{\delta} \quad (34)$$

が成り立ち、これより

$$\left| Q_{o(j)}(\alpha; \beta_{(j)}) - \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{o(j)}(\alpha; \beta_{(j)}^i) \right| < \xi_{\delta} \quad (35)$$

となる。

以上の結果を元に、時刻  $t$  における尤度  $Q^t(\alpha; X^t, \beta^t)$  が、 $t \rightarrow \infty$  で  $Q_o(\alpha; \beta^{\infty})$  に一様収束することを示す。まず  $Q^t(\alpha; X^t, \beta^t)$  と  $Q_o^t(\alpha; \beta^t)$  を、式 (24) に基づく  $\{\beta_i\}$  のグループ分けに伴い、次の様に分解する。

$$\begin{aligned} & Q^t(\alpha; X^t, \beta^t) \\ &= \frac{1}{t} \left[ \sum_{j=1}^{M(\delta)} \left\{ \sum_{i=1}^{t_j} \ell(x_{(j)}^i, \alpha, \beta_{(j)}^i) \right\} + \sum_{i=1}^{t_{\infty}} \ell(x_i, \alpha, \beta_i) \right] \\ &= \sum_{j=1}^{M(\delta)} \left\{ \frac{t_j}{t} Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j}) \right\} + \frac{1}{t} \sum_{i=1}^{t_{\infty}} \ell(x_i, \alpha, \beta_i) \end{aligned} \quad (36)$$

$$\begin{aligned} & Q_o^t(\alpha; \beta^t) \\ &= \sum_{j=1}^{M(\delta)} \frac{t_j}{t} \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{oi}(\alpha; \beta_{(j)}^i) + \frac{1}{t} \sum_{i=1}^{t_{\infty}} Q_{oi}(\alpha; \beta_i) \end{aligned} \quad (37)$$

式 (37) からさらに

$$\begin{aligned} & Q_o(\alpha; \beta^{\infty}) \\ &= \lim_{t \rightarrow \infty} Q_o^t(\alpha; \beta^t) \\ &= \sum_{j=1}^{M(\delta)} \lim_{t \rightarrow \infty} \frac{t_j}{t} \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{oi}(\alpha; \beta_{(j)}^i) + \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^{t_{\infty}} Q_{oi}(\alpha; \beta_i) \end{aligned} \quad (38)$$

を得る。よって式 (36), (38) より次を得る。

$$\begin{aligned} & Q^t(\alpha; X^t, \beta^t) - Q_o(\alpha; \beta^{\infty}) \\ &= \sum_{j=1}^{M(\delta)} \frac{t_j}{t} \left( Q^{t_j}(\alpha; X_{(j)}^{t_j}, \beta_{(j)}^{t_j}) - Q_{o(j)}(\alpha; \beta_{(j)}) \right) \\ & \quad + \sum_{j=1}^{M(\delta)} \frac{t_j}{t} \left( Q_{o(j)}(\alpha; \beta_{(j)}) - \lim_{t \rightarrow \infty} \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{oi}(\alpha; \beta_{(j)}^i) \right) \\ & \quad + \sum_{j=1}^{M(\delta)} \left( \left( \frac{t_j}{t} - \lim_{t \rightarrow \infty} \frac{t_j}{t} \right) \lim_{t \rightarrow \infty} \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{oi}(\alpha; \beta_{(j)}^i) \right) \\ & \quad + \left( \frac{1}{t} \sum_{i=1}^{t_{\infty}} \ell(x_i, \alpha, \beta_i) - \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^{t_{\infty}} Q_{oi}(\alpha; \beta_i) \right) \end{aligned} \quad (39)$$

ここで右辺第 1 項の大きさは式 (33) より評価でき、右辺第 2 項の大きさは式 (35) より評価できる。第 3 項は確率密度関数に関する正則条件より

$$|Q_{oi}(\alpha; \beta_{(j)}^i)| < \psi_1, \exists \psi_1 > 0 \quad (40)$$

が、また

$$\begin{aligned} & \left| \frac{t_j}{t} - \lim_{t \rightarrow \infty} \frac{t_j}{t} \right| < \psi_2 \\ & \text{when } N_3(\psi_2) < t, \exists N_3(\psi_2) > 0 \end{aligned} \quad (41)$$

が成り立つ。よって次が得られる。

$$\begin{aligned} & \left| \left( \frac{t_j}{t} - \lim_{t \rightarrow \infty} \frac{t_j}{t} \right) \lim_{t \rightarrow \infty} \frac{1}{t_j} \sum_{i=1}^{t_j} Q_{oi}(\alpha; \beta_{(j)}^i) \right| < \psi(\psi_2) \\ & \text{when } N_3(\psi_2) < t, \psi(\psi_2) := \psi_1 \psi_2 \end{aligned} \quad (42)$$

第 4 項については、式 (22) に基づく  $\{\beta_i\}$  のグループ分けからまず

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^{t_{\infty}} Q_{oi}(\alpha; \beta_i) = 0 \quad (43)$$

となる。また確率密度関数に関する正則条件より次が成り立つ。

$$\begin{aligned} & \left| \frac{1}{t} \sum_{i=1}^{t_{\infty}} \ell(x_i, \alpha, \beta_i) \right| < \tau \\ & \text{when } N_4(\tau) < t, \exists N_4(\tau) > 0 \end{aligned} \quad (44)$$

以上をまとめて次を得る。

**定理 4.2** 次が成り立つ。

$$\begin{aligned} & \Pr \left( \sup_{\alpha} |Q^t(\alpha; X^t, \beta^t) - Q_o(\alpha; \beta^{\infty})| > W \right) < \eta \\ & W := \sum_{i=1}^{M(\delta)} \left[ \frac{t_j}{t} (\phi_{\delta} + \epsilon + \xi_{\delta}) + \psi(\psi_2) \right] + \tau \\ & \text{when } t > \max\{N_1(\epsilon, \eta), N_3(\psi_2), N_4(\tau)\} \end{aligned} \quad (45)$$

さらに  $W$  は、 $\delta$  を任意に小さな値に設定し、それに応じて十分大な  $t$  において任意に小さな値に設定できる。以上より、 $Q^t(\alpha; X^t, \beta^t)$  は  $Q_o(\alpha; \beta^{\infty})$  に一様に収束する。

最後に  $Q^t(\alpha; X^t, \beta^t)$  の  $Q_o(\alpha; \beta^{\infty})$  への一様収束により、 $\tilde{\alpha}_t$  が  $\alpha_o$  の一致推定量であることが帰結される。

#### 4.6 閉ループ系の収束性

定理 4.2 を  $\tilde{z}_2[t] \leftrightarrow \beta_i, x_i \leftrightarrow y_2[t], \tilde{\alpha}_t \leftrightarrow \tilde{z}_2[t]$  と対応させ、サブシステム  $\mathcal{F}_1$  に適用することにより、式 (9) にあるように  $\beta_i$  の定義域が有界、つまり  $\tilde{z}_1$  の値域が有界ならば、よって  $\tilde{z}_2$  の収束が導かれる。以上より、 $\mathcal{F}_1$  と  $\mathcal{F}_2$  からなる閉ループ系で、 $z_1, z_2$  の定義域がそれぞれ有界で互いに既知であるならば、閉ループ系は平衡点に収束することが導かれる。

## 5 数値実験

ここでは簡単な状況で数値実験を行う。まず次の混合ガウス分布のモデルを考える (Fig. 4 参照)。

$$p(y; z_1, z_2, \sigma_1^2, \sigma_2^2) = \frac{1}{2}N(z_1, \sigma_1^2) + \frac{1}{2}N(z_2, \sigma_2^2)$$

ただし  $\sigma_1^2, \sigma_2^2$  は固定され互いに既知とする。このモデルは player 1 は自身の元々の効用として右辺第 1 項を持ち、player 2 は右辺第 2 項を持つ、という考えに基づいている。それぞれの戦略パラメタは  $z_1, z_2$  であり、それらは混合ガウス分布における各ガウス分布の中心点である。

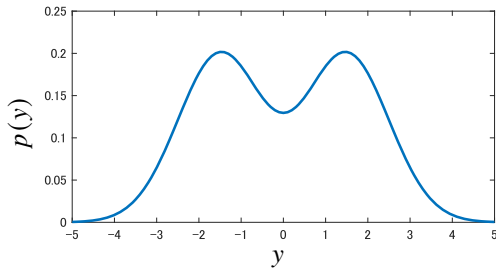


Fig. 4: Mixed Gaussian model for generating offers

このモデルを元に、player 1 は上式の  $z_2$  を推定値  $\tilde{z}_2$  に置き換え、 $p(y; z_1, \tilde{z}_2, \sigma_1^2, \sigma_2^2)$  に基づいて自身の offer  $y = y_1$  をランダムに生成し、相手に提示する。一方 player 2 は  $p(y; \tilde{z}_1, z_2, \sigma_1^2, \sigma_2^2)$  に基づき  $y = y_2$  をランダムに生成する。次に player 1 は player 2 より受け取った  $y_2[t]$  の時系列データに基づき、 $p(y; \tilde{z}_1, z_2, \sigma_1^2, \sigma_2^2)$  のモデルに従い、 $z_2$  を最尤推定する。player 2 も同様である。

以上の設定のもと、次の条件で数値実験を行った。 $\sigma_1^2 = \sigma_2^2 = 1$ ,  $z_1 = -1.5$ ,  $z_2 = 1.5$ ,  $\tilde{z}_1[0] = 0$ ,  $\tilde{z}_2[0] = 0$ 。信号のやりとりの繰り返し回数は 500 回である。

Fig. 5 は最尤推定値  $\tilde{z}_1$  (赤実線) と  $\tilde{z}_2$  (青破線) の時間推移である。図よりそれぞれ真値に収束しており、相互最尤推定の収束性が確認される。また Fig. 6 および Fig. 7 は相互最尤推定の過程で生成された  $y_1$  および  $y_2$  のデータのヒストグラムである。

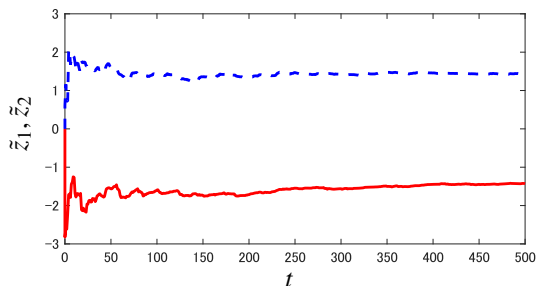


Fig. 5: Time transition of  $\tilde{z}_1$  and  $\tilde{z}_2$ :  $\tilde{z}_1$ : red-solid line,  $\tilde{z}_2$ : blue-dashed line

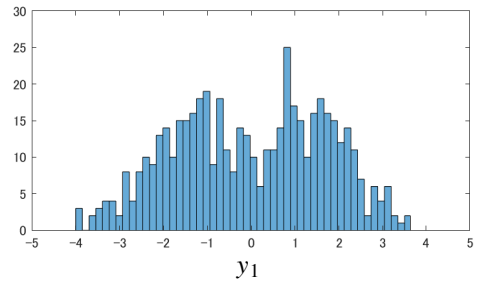


Fig. 6: Histograms of the generated offers  $y_1$

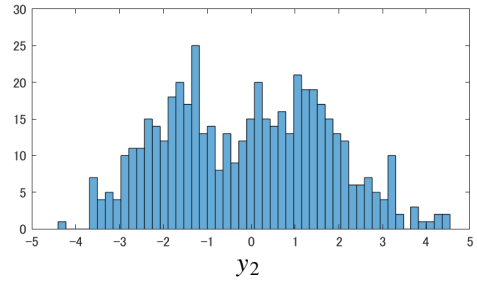


Fig. 7: Histograms of the generated offers  $y_2$

## 6 まとめ

本論文では 2 者間の自動交渉システムのモデルに着目し、その収束性が相互最尤推定の収束性に帰着できることを説明した。また幾つかの条件下で、毎時刻変化する部分的パラメタの情報の元での最尤推定値が真値に収束することを示した。さらに戦略のパラメタの定義域が有界である条件下で、相互最尤推定が収束すること、つまり自動交渉システムが収束することを示した。

なお本研究は、NEDO SIP、「ビッグデータ・AIを活用したサイバー空間基盤技術」のサポートを受けて行われた。

## 参考文献

- 1) C. M. Jonker, R. Aydođan, T. Baarslag, K. Fujita, T. Ito, K. Hindriks: Automated Negotiating Agents Competition (ANAC), Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)
- 2) T. Baarslag, M. J. C. Hendriks, K. V. Hindriks, C. M. Jonker: Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques, Autonomous Agents and Multi-agent Systems, 30, pp. 849–898 (2016)
- 3) 藤田, 森, 伊藤: ANAC: Automated Negotiating Agents Competition (国際自動交渉エージェント競技会), 人工知能, 31-2, pp. 237–247 (2016)
- 4) 津村: 自動交渉のフィードバックシステムによる定式化と解析, 第 62 回自動制御連合講演会 (2019)
- 5) 野田, 宮岡: 数理統計学の基礎, 共立出版 (1992)