

ビッグデータと寡占競争*

加 藤 浩

1. はじめに

現代市場はデジタル化された環境の下で取引が行われており、企業は ICT を用いて商品を販売し、消費者は端末を経由して商品を購入している。取引の副産物としてデータが生み出され、企業はデータを収集して意思決定に活用する。データを活用して意思決定が行われる市場は、データ駆動型市場 (data-driven markets) と呼ばれている。近年は、高速通信が実現し、大容量保存が可能になり、さらにクラウド・コンピューティングや IoT (モノのインターネット化) といった新しい技術が普及した。これらの技術進歩により、企業はかつてないほどの大規模かつ多種多様なデータを収集できるようになった。このような流れもあり、多くの市場がデータ駆動型に転換されている。

現代産業においては、データが一種の資本となっている。とりわけデジタル・プラットフォームが享受している巨額の利益と市場支配力は、内部に蓄積された膨大なデータがその源泉となっている。プラットフォームには膨大な数のユーザーが集まり、膨大な数の取引が行われ、そこから膨大なデータ (すなわちビッグ・データ) が生み出され、プラットフォームに蓄積される。収集・蓄積したデータは、機械学習アルゴリズムを搭載したマシンに投入される。投入されるデータが多ければ多いほどマシンの学習機会が増えて、学習成果としてマシンが持つ予測やクラスタリングの能力が向上し、適切かつ効果的な意思決定を行うことがで

* 本稿は、九州経済法研究会第 122 回例会 (公正取引委員会事務総局九州事務所) で報告した『ビッグデータとプラットフォーム：経済学の観点から』をもとに、論文として書き下ろしたものである。

きる。例えば、需要予測の精度が高まれば、より多くの利益をもたらす最適な価格を設定することが可能となる。また、より繊細なクラスタリングは、消費者のごくわずかな違いも見逃さないで、消費者の選好により適合した商品の提供や、より効果的な販売促進活動につながる。

2面市場（two-sided market）を持つプラットフォームには、消費者サイド（無料サイド）のグループと、広告サイド（支払いサイド）のグループが集まる。プラットフォームは消費者にサービスやコンテンツを無料で提供し、それと交換で消費者データを収集する。収集したデータを活用することで高品質のサービスや優れたコンテンツが提供されて、消費者により高い効用を与える。高い効用を与えるプラットフォームに多くの消費者が加入し、プラットフォームはさらにデータを収集・蓄積して、より高品質のサービスやより優れたコンテンツを提供するために用いられる。データによってもたらされるこのような好循環は、データ駆動型ネットワーク効果と呼ばれる。

消費者サイドに加入者が増えると、広告サイドの加入者も増える。これが間接ネットワーク効果である。消費者サイドの加入者数が多いプラットフォームに広告を掲載すると、それだけ閲覧される可能性が高まる。また、豊富な消費者データと優れたアルゴリズムを所有するプラットフォームは、消費者の選好や関心について正確な情報を持っている。この情報を活用することで、消費者がクリックする可能性が高い効果的なターゲット広告を掲載することができる。それゆえ、データを多く蓄積しているプラットフォームに、広告サイドの加入者が集まり、プラットフォームは多額の広告収入を獲得する。

データ駆動型ネットワーク効果により、加入者の増加とデータの蓄積という好循環がもたらされ、市場はティッピング（tipping）し、プラットフォームは巨大化し、やがて市場を支配する。市場がティッピングするとは、市場シェアの大きな企業の市場シェアはますます増大し、市場シェアの小さな企業の市場シェアはますます縮小するという現象である¹⁾。

1) 市場がティッピングするという点で、伝統的ネットワーク効果と類似した効果を持つが、ティッピングを引き起こす要因が異なる。伝統的ネットワーク効果は、ネットワークに加入しているユーザー数が要因であるが、データ駆動型ネットワーク効果は企業が所有するデータが要因である。

本稿の関心は、データが企業に与える影響について、産業組織論の理論モデルを通じて明らかにすることにある。2節では、企業が収集・蓄積するデータをモデル化し、さらに、データを活用して意思決定を行う企業をモデル化する。また、「データはビッグであればあるほど良い」という巷間に流布している言説、産業組織論の観点から言い換えれば、収集・蓄積するデータが多いほど、企業に多くの利潤をもたらすという命題の妥当性について、理論的な視点から議論する。3節では、データが寡占競争に与える影響を明らかにする。データ駆動型市場の動学寡占モデルを取り上げ、市場をティッピングさせる原因となるデータ駆動型ネットワーク効果に注目して議論を展開する。4節では、まとめと今後の研究課題について言及する。

2. データと企業行動

2.1 ビッグデータとは

本稿は、次のような目的でデータを収集・蓄積する企業を想定している。企業はマシンを用いて意思決定を自動化しており、取引を通じて消費者からデータを収集する。収集したデータはマシンに搭載されているアルゴリズムのインプットとなり、処理・解析される。消費者の選好パラメータ θ は、その消費者の好みや関心を表す定数であり、消費者ごとに値が異なる。企業は正確な θ の値が分からないので、データを投入して、マシンに選好パラメータの推定値 $\hat{\theta}$ を計算させる。この推定値に基づいて、消費者が支払える最大水準の価格を設定したり、消費者の選好にフィットした商品のレコメンデーション（推奨）を提示したり、消費者の関心にマッチしたターゲット広告を掲載したりする。消費者の選好パラメータを正確に推定できれば、より適切な価格設定や、よりの確なレコメンデーションや、より効果的なターゲット広告が可能となり、企業の収入増加につながる。

マシンは消費者からデータ x を収集して θ の値を推定するので、データ x は θ に関する情報を伝えるシグナルでなければならない。ただし、データにはノイズも含まれている。ノイズ ϵ は確率変数であり、同じ情報源（センサ）からデータを収集しても、ノイズの実現値は収集するたびに変動する。したがって、データ

x も確率変数であり、

$$x = x(\theta, \varepsilon) \quad (1)$$

と表される。消費者の選好のシグナルとなるデータとして、消費者の属性データと消費者の行動データの2つが挙げられる。属性データは、年齢、所得、性別、居住地といった、消費者の選好と相関するデータである。行動データは、オンラインでの買い物行動を記録したものであり、消費者の選好を顕示するデータである。買い物行動を具体的に挙げると、購入した商品、購入した時間、Webサイトの訪問回数、閲覧したWebページ、Webページでの行動（クリック、検索ワード）といったものである。オンラインでは、cookie やアクセス・ログ解析により、これらのデータが容易に入手できる。

データがビッグであるとは、データ量（Volume）が多いことと、データが多様（Variety）であることの2つの意味が含まれている²⁾。Volume は標本数（インスタンス数）、Variety は特徴量の個数（次元数）で計測される³⁾。データ集合

$$\begin{pmatrix} \mathbf{x}^{(1)} \\ \vdots \\ \mathbf{x}^{(n)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & \cdots & x_m^{(1)} \\ \vdots & & \vdots \\ x_1^{(n)} & \cdots & x_m^{(n)} \end{pmatrix} \quad (2)$$

について、 $\mathbf{x}^{(i)} = (x_1^{(i)}, \dots, x_m^{(i)})$ はデータ点、 $x_j^{(i)}$ はデータである。Volume の指標は標本数 n 、Variety の指標は特徴量の個数 m となる。

収集・蓄積したデータが大量かつ多様であれば、消費者の選好パラメータをより正確に推定することができる。これがデータがビッグであることのメリットである。以降の小節では、このメリットが生じる理論的な背景について、Volume と Variety に分けて議論していく。

2) これに加えて、データが生成される速度（Velocity）も重要であり、企業が直面している状況をリアルタイムに把握できるデータであることもビッグデータの要件となる。例えば、ダイナミック・プライシングを実行するためには、最新の需要データ、在庫データ、位置データ、天候データが必要である。さらに、データが価値（Value）を持ち、データを活用することで利益が生み出されることも要求される。Volume, Variety, Velocity, Value は、ビッグデータを定義する「4つのV」と呼ばれる。このほかに、Veracity（正確さ）、Variability（変動性）、Visualization（可視化）を含めて「7つのV」とすることがある。

3) ここでの特徴量とは、消費者の選好を特徴づける変数であり、消費者の選好パラメータを目的変数とすると、特徴量は説明変数になる。

2.1.1 Volume

ビッグデータがもたらす Volume のメリットは、データの規模に関する収益性である。データが増えれば増えるほど、データはより情報提供的になり、推定の精度がそれだけ高まる。これが当てはまるデータの種類として、次の3つが挙げられる。

(1) データにノイズが多く含まれている

行動データは消費者の選好を顕示するシグナルとなるが、同時に多くのノイズが含まれている。しかし、ノイズが多いデータであっても、同じ情報源からデータを大量に集めて平均化すると、データに占めるノイズの比率は低下していく。以下では、簡単なモデルを用いてこの知見が正しいことを確認する。

同一の情報源（センサ）を通じて同じ消費者のデータを n 回収集する。これらのデータを $x^{(1)}, \dots, x^{(n)}$ とする。 i 回目のデータは次のようになる。

$$x^{(i)} = \theta + \varepsilon^{(i)}, \varepsilon^{(i)} \sim N(0, \eta^2) (i = 1, \dots, n). \quad (3)$$

η^2 はノイズの分散である。 η^2 が大きいほどノイズが酷く、データを抽出するたびにノイズの実現値が大きく変動するため、正確なデータを収集することが困難になる。同じセンサからデータを収集しているので、データを何回収集しても η^2 の値は一定である。また、 $\varepsilon^{(i)}$ と $\varepsilon^{(j)}$ は独立であり、前回抽出したデータの値が、今回抽出するデータの値に影響を与えることはない。

$$E(x^{(i)}) = \theta \quad (4)$$

となるから、 $x^{(i)}$ は θ の不偏推定量である。収集されたデータの平均値を選好パラメータの推定値とすると、

$$\hat{\theta} = \frac{\sum_{i=1}^n x^{(i)}}{n} = \theta + \frac{\sum_{i=1}^n \varepsilon^{(i)}}{n} \quad (5)$$

となり、

$$\frac{\sum_{i=1}^n \varepsilon^{(i)}}{n} \sim N\left(0, \frac{\eta^2}{n}\right) \quad (6)$$

であるから、標本数 n が大きいほど推定値 $\hat{\theta}$ の分散が小さくなる。推定値の精度 (precision) は分散の逆数で測られるから、標本数 n が大きくなるほど推定値の精度が上昇していく。

(2) 行動分布が重い裾 (heavy tail) を持つ

特徴量の個数 m を行動可能集合の要素数とする。消費者は行動可能集合 $A = \{a_1, \dots, a_m\}$ の中から行動 a_i を選択する。消費者が行動 a_i を選択したとき、データ x_i は非ゼロの値を取り、選択しないときにはゼロの値を取る。オンラインでは、詳細な消費者行動をデータとして捉えることができるので、行動データは細分化された特徴量 (fine-grained feature data) を持ち、したがって、行動可能集合の要素数 m は膨大である。しかも、行動 $\{a_1, \dots, a_m\}$ の分布は裾が重く (図 1)⁴⁾、裾に含まれる行動はまれにしか選ばれないものの、行動可能集合のかなりの部分を占めている⁵⁾。このようなことが原因で、行動データは疎 (スパース) である。疎なデータとは、あるデータ点 $(x_1^{(i)}, \dots, x_m^{(i)})$ について、殆どの要素がゼロの値を取り、

4) 対数正規分布やパレート分布は裾の重い分布である。

5) 具体的な事例として、検索エンジンと商品レビューを挙げる。いずれの例も、膨大なデータを有する企業に優位性が与えられる。専門性の強いマイナーな検索ワードは検索される確率が低いものの、検索ワード全体に占める割合が圧倒的に高い。どのような検索エンジンでも、メジャーな検索ワードの検索結果は精度が十分に高いため、競合する検索エンジンの中で差が出ない。それゆえ、マイナーな検索ワードについてユーザーが望む結果を表示することが、検索エンジンの優劣を決める。

EC サイトで販売されている商品の数は膨大であるが、その中で購入されるものはごくわずかである。それゆえ、売れ行きの悪いマイナーな商品のレビューはごくわずかしかない。しかし、取り扱う商品全体に占めるマイナー商品の数は圧倒的に多いため、個々の商品の販売量が少なくてもこれらの商品の多くが売れることで企業に大きな収入をもたらす。このことから、マイナー商品のレビューの多寡が、EC サイトの収入を左右する。

ほんの一部だけ非ゼロの値を取るようなデータのことである⁶⁾。選好パラメータの推定値の精度を上げるためには、裾に含まれる行動について非ゼロ値のデータが必要となる。標本数が小さいときは、このようなデータを入手できる確率が低いため、消費者の選好に関する有用な情報はほとんど得られない。しかし、標本数が大きくなると、十分な数の非ゼロ値のデータが観察されるので、データ集合には消費者の選好に関する重要な情報が含まれ、推定の精度が格段に上昇する。

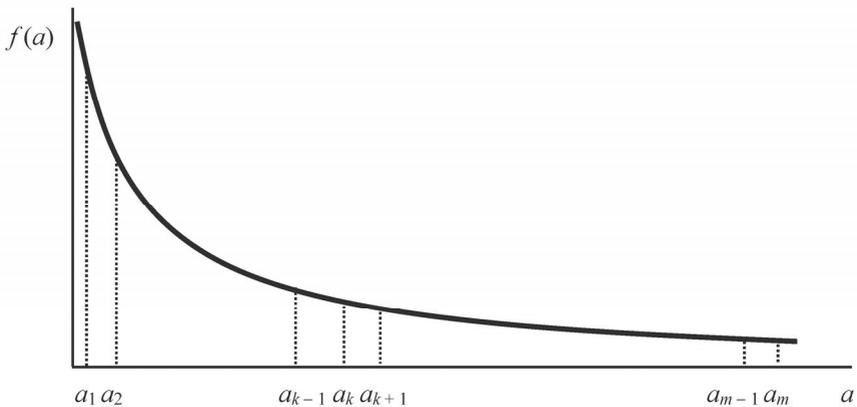


図1 重い裾を持つ行動分布

$f(a)$ は a が選択される確率
 m のオーダーは数万単位である

6) ビッグデータは、標本数と比べて特徴量の個数が非常に多くなる傾向にある。つまり、高次元データとなり、標本数 n と特徴量の個数 m について、 $n \ll m$ となる。 m は数万単位、 n は数百単位ということがざらである。疎性は、データ集合のデータ数 $m \times n$ に占めるゼロ値の割合で計測される。デジタル・プラットフォームで入手されるデータの疎性は、99.9%を超えることがほとんどである (de Fortuny, Martens&Provost (2013))。

(3) 機械学習の訓練データとして使用する

機械学習アルゴリズムを搭載したマシンは、投入したデータ（訓練データ）から規則性やパターンを自動的に学習する⁷⁾。学習した結果はマシンの推定能力に反映され、新たに入手したデータに対してマシンはより正確な推定値を出力する。学習成果は推定値の正確度（accuracy）で計測される。マシンに投入されるデータが多ければ多いほど、それだけ学習する機会が増えるため、優れた学習成果が期待できる⁸⁾。

訓練データの投入量 N と推定誤差 y の関係をグラフにしたものが、（誤差で測った）学習曲線である。データ点 $\mathbf{x} = (x_1, \dots, x_m)$ から計算された選好パラメータ θ の推定値を $\hat{\theta}(\mathbf{x})$ とすると、（汎化）誤差（generalization error）は、

$$y = E((\hat{\theta}(\mathbf{x}) - \theta)^2) \quad (7)$$

となり、この値が小さいほど推定値の正確度が高くなる⁹⁾。学習曲線はべき乗則に従うこと

7) 機械学習には教師あり学習と教師なし学習がある。教師あり学習の中心は予測モデルである。訓練データ $(x, y) = (x^{(1)}, y^{(1)}), \dots, (x^{(N)}, y^{(N)})$ を用いて分布 $p(y|x)$ を学習する。ここで、 x は独立変数、 y は従属変数である。学習プロセスでは、各 x に対して正解の値 y を与えることでマシンを訓練する。学習した $p(y|x)$ を用いて、新たに得られたデータ x から y の値を予測する。例えば、収集したデータ x を $p(y|x)$ に投入することで需要量 y の予測を確率として出力して、そこから最適な価格を計算する。

教師なし学習の中心となるモデルは、クラスタリング・モデルである。データ $x^{(1)}, \dots, x^{(N)}$ から構造やパターンを学習して、 $p(x^{(1)}, \dots, x^{(N)})$ をモデル化する。例えば、 N 人の消費者データ $x^{(1)}, \dots, x^{(N)}$ から選好の類似性を計算し、それぞれのデータを、事前に指定された数のクラスタに分割する。同じクラスタに含まれるデータどうしは選好が似ているが、別のクラスタに含まれるデータとは似ていない。データ点 $x^{(i)}, x^{(j)}$ の類似性とは、点の距離 $d(x^{(i)}, x^{(j)})$ で計測される。距離の近さが選好の類似性と解釈して、距離の近い2つのデータ点を同じクラスタに含める。クラスタリング・モデルはマーケット・セグメンテーションに使用され、クラスタのそれぞれがセグメントに対応する。従来のセグメンテーションと異なるのは、選好の類似性だけでクラスタが分割されており、各クラスタが性別や年齢や購入履歴といった、意味のある特徴付けを持たないことである。

8) 企業によってマシンに搭載されているアルゴリズムの質には優劣があるが、単純なアルゴリズムでも投入されるデータが多ければ、データが少ない複雑なアルゴリズムに能力の面で勝るという主張がある（Google のリサーチ・ディレクターの Peter Norvig の発言）。

9) 訓練誤差（training error）は、投入された訓練データから計算された推定値のみ取り上げて、その誤差について平均値を計算したものであるが、汎化誤差は、訓練では投入されていない標本データから計算された推定値も含めて、誤差の期待値を計算したものである。

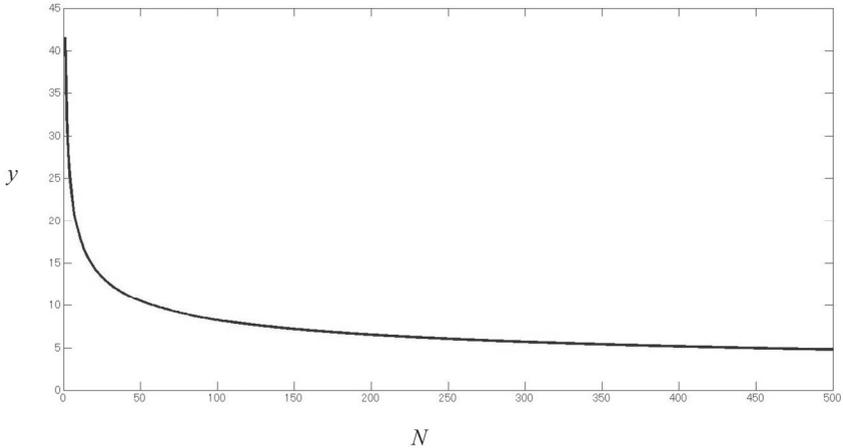


図2 学習曲線

$$y = 41.2N^{-0.36} + 0.39$$

208 個の隠れ層を持つニューラル・ネットワークで検証

出所：Hestness et al.(2017)

が数々の実証研究で示されており，学習曲線を表す式は次のようになる¹⁰⁾。

$$y = aN^{-b} + c. \quad (8)$$

a, b, c は正の定数である。 $y' < 0$, $y'' > 0$ であるから，データの投入量が多いほど誤差は減少するものの，減少の程度は逓減していく。つまり，学習成果はデータに関する収穫逓減を示す。特に， $N \rightarrow \infty$ のときは $y \rightarrow c$ となり， c は誤差の下限となり，データをどれだけ大量に投入しても，選好パラメータの真の値を推定することは不可能であることが分かる。収穫逓減の原因は，次のように説明されることが多い。投入したデータが少ないときに，追加してデータを投入すると，学習アルゴリズムはデータが持つ顕著な特徴を捉えて推測値に反映するため，誤差は大きく減少する。しかし，データを追加すればするほど目新しい情報が少なくなり，追加データからは微細な特徴を捉えて推測値の改善に反映するしかなく，誤差はあまり低下しなくなる。

10) Hestness et al.(2017), Rosenfeld et al.(2020)はニューラル・ネットワーク，John and Langley(1996)は単純ベイズ分類で学習曲線の形状を推計している。

2.1.2 Variety

ビッグデータがもたらす Variety のメリットは、データの範囲に関する収益性である。消費者の選好について完全な情報を与える単一の情報源（センサ）は存在しない。また、データの特徴量について消費者の選好を完全に決定づけるものではなく、ほとんどの特徴量が選好についてわずかな情報しか提供しない。そこで、多方面かつ様々な状況下にある複数のセンサ（multisensor）からデータを収集して、これらのデータを融合（data fusion）して消費者の選好を推定することが有効な手段となる¹¹⁾。それぞれのセンサから得られるデータが持つ情報価値がわずかであっても、異なる複数のセンサから集めた多種多様なデータを融合すると情報価値が高くなる。また、異なるセンサからデータを集めることで、選好と関連する特徴量の種類が増える。少ない特徴量で選好パラメータを推定するよりも、より多くの特徴量で推定したほうが、推定値の精度は高くなる¹²⁾。これまでは、消費者が所有する端末からデータを集中的に取得していたが、近年は IoT の普及により、分散したセンサからデータを収集することができるようになった。収集したデータは中央システム（サーバー）に蓄積されて、これらのデータは融合され、多種多様な特徴量を持つ 1 つのデータ点となる。以下では、単純なデータ融合モデルを通して、複数のセンサーから得られたデータを融合して推定値を計算する方法を示し、データ融合により精度の高い推定値が得られることを確認する。

いま、センサ $1, \dots, m$ は同じ消費者の選好パラメータ θ を計測している。センサ i からデータ x_i が得られる。データ x_1, \dots, x_m は θ に依存する確率変数であり、 θ しかなかった部分を持たない¹³⁾。センサ i から入手するデータを、

11) PC やスマートフォンといった端末から多種多様なデータが収集されるが、これらのデバイスから収集できるデータのタイプは限られているので、他のデバイスから異なるタイプのデータを収集して補完する。例えば、スマート・ウォッチから位置データや生体データ、スマート・スピーカから会話データ、電子マネーから買い物行動のデータ、スマート家電から生活スタイルのデータが収集できる。

12) 単一のセンサから集中的に取得するデータには、選好とは関連のない特徴量が多く含まれている。関連のない特徴量で推定すると過学習（overfitting）に陥り、推定誤差が増大する。

13) データ x_1, \dots, x_m が独立ならば、それぞれ無関係なデータとなるから、このようなデータを融合したところで、選好に関する価値のある情報は生み出されない。

$$x_i = \theta + \varepsilon_i, \quad \varepsilon_i \sim N(0, \eta_i^2) \quad (i = 1, \dots, m) \quad (9)$$

とする。 ε_i はセンサからデータを収集するときに付け加えられるノイズであり、 ε_i と ε_j は独立である。 $1/\eta_i^2$ はセンサ i の測定精度の指標となる。

まず、データを所有していない初期状態から出発する。 θ の推定は、実現可能な θ の値に対して割り振られる確率密度として表される。この確率密度関数は主観的なものであり、情報を得るたびに更新されることから、 θ の真の値に対する信念 (belief) として扱う。データを入手する前の推定、すなわち事前信念 $P(\theta)$ を消費者の人口分布とする。ここで言う人口分布は選好パラメータの分布のことであり、 $\theta \sim N(\theta_0, \sigma^2)$ とする。 θ の確率密度関数は次のようになる。

$$P(\theta) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{1}{2} \frac{(\theta - \theta_0)^2}{\sigma_0^2}\right\}. \quad (10)$$

σ^2 の値が小さくなると、より狭い区間 $[\theta_0 - a, \theta_0 + a]$ により高い確率を割り振ることになるので、推定の精度がより高くなることを意味する。したがって、 $1/\sigma^2$ は推定の精度を表す指標となる。

いま、センサ 1 からデータ x_1 を入手した。センサ 1 で θ を計測したときのデータ x_1 を入手する確率密度は $P(x_1 | \theta)$ である¹⁴⁾。 x_1, θ はともに正規分布に従うので、条件付き確率も正規分布に従い、確率密度関数は次のようになる。

$$P(x_1 | \theta) = \frac{1}{\sqrt{2\pi\eta_1^2}} \exp\left\{-\frac{1}{2} \frac{(x_1 - \theta)^2}{\eta_1^2}\right\}. \quad (11)$$

データ x_1 を入手すると θ に関する新しい情報が得られるので、 θ の推定が更新される。ここでは、ベイズ・ルールにしたがって、

14) θ の値を固定すると、 $P(x_1 | \theta)$ は x_1 の関数となり、特定の消費者を計測したときにデータ x_1 を出力する確率となる。したがって、 $P(x_1 | \theta)$ はセンサ 1 の測定モデルとしての特徴を表している。 θ の値を変えることは、測定対象の消費者を変えることを意味するので、 x_1 の分布も変わる。他方、ある値のデータ x_1 を入手したとき、 $P(x_1 | \theta)$ は θ の関数と見ることができ、ある特定の θ の値が真の値である確率と解釈することができる。したがって、 $P(x_1 | \theta)$ を尤度関数と見なして、 $L(\theta; x_1) = P(x_1 | \theta)$ と表すことができる。

$$P(\theta | x_1) = \frac{P(x_1 | \theta)P(\theta)}{P(x_1)} \quad (12)$$

と更新される。これより、事後信念は次のようになる。

$$P(\theta | x_1) = \frac{1}{P(x_1)} \times \frac{1}{2\pi\sigma_0^2\eta_1^2} \exp\left\{-\frac{1}{2} \frac{\sigma_0^2(x_1 - \theta)^2 + \eta_1^2(\theta - \theta_0)^2}{\sigma_0^2\eta_1^2}\right\}. \quad (13)$$

事後信念 $P(\theta | x_1)$ は正規分布となり、これを、

$$P(\theta | x_1) = \frac{1}{\sqrt{2\pi}\sigma_1^2} \exp\left\{-\frac{1}{2} \frac{(\theta - \theta_1)^2}{\sigma_1^2}\right\} \quad (14)$$

と置くと、

$$\theta_1 = \frac{\eta_1^2}{\sigma_0^2 + \eta_1^2} \theta_0 + \frac{\sigma_0^2}{\sigma_0^2 + \eta_1^2} x_1, \quad (15)$$

$$\sigma_1^2 = \frac{\sigma_0^2 \eta_1^2}{\sigma_0^2 + \eta_1^2} \quad (16)$$

という関係が成り立つ。

θ の推定値（点推定） $\hat{\theta}$ を最大事後確率推定値（MAP 推定量）とする。すなわち、

$$\begin{aligned} \hat{\theta} &= \arg \operatorname{Max}_{\theta} P(\theta | x_1) \\ &= \arg \min_{\theta} \{-\log P(\theta | x_1)\} \\ &= \arg \min_{\theta} \left\{ \frac{1}{2} \frac{(x_1 - \theta)^2}{\eta_1^2} + \frac{1}{2} \frac{(\theta - \theta_0)^2}{\sigma_0^2} \right\}. \end{aligned} \quad (17)$$

これより、 θ の推定値は、

$$\hat{\theta} = \frac{\eta_1^2}{\sigma_0^2 + \eta_1^2} \theta_0 + \frac{\sigma_0^2}{\sigma_0^2 + \eta_1^2} x_1 \quad (18)$$

となる。

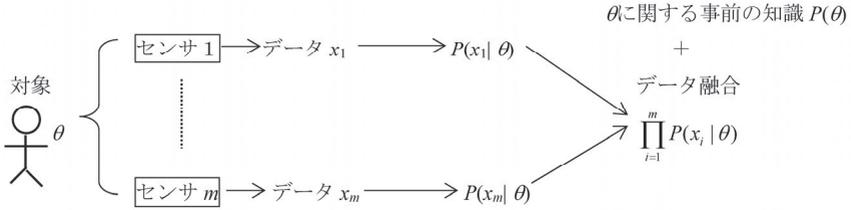


図3 複数のセンサとデータ融合

いま、センサ 1, ..., m からデータ x_1, \dots, x_m を同時に入手したとする。これを融合してデータ点 $\mathbf{x} = (x_1, \dots, x_m)$ を得ると、ベイズ・ルールから事後信念は、

$$P(\theta | x_1, \dots, x_m) = \frac{P(x_1, \dots, x_m | \theta)P(\theta)}{P(x_1, \dots, x_m)} \tag{19}$$

と計算される。 x_1, \dots, x_m は同じパラメータ θ を計測するセンサから得られたデータであり、なおかつセンサ 1, ..., m はそれぞれ独立して θ を計測するので、 θ の下で x_1, \dots, x_m は条件付き独立となる¹⁵⁾。すなわち、

$$P(x_1, \dots, x_m | \theta) = \prod_{i=1}^m P(x_i | \theta) \tag{20}$$

これより、事後信念は次式で与えられる (図3)。

$$P(\theta | x_1, \dots, x_m) = \frac{P(\theta) \prod_{i=1}^m P(x_i | \theta)}{P(x_1, \dots, x_m)} \tag{21}$$

再帰ベイズ更新ルールは、データを融合して推定値を計算するときに用いられる便利な公式である。これは、センサ 1 からセンサ m まで順番にデータを集めていき、データを新たに入手するたびにデータを融合し、 θ の推定値を更新するというものである。

既にセンサ 1, ..., k-1 からデータ x_1, \dots, x_{k-1} を入手しているとする。融合したデータ点 (x_1, \dots, x_{k-1}) から導かれる事前信念は $P(\theta | x_1, \dots, x_{k-1})$ である。いま新

15) 同じ θ を計測しているから、 x_1, \dots, x_m は互いに独立ではない。つまり、

$$P(x_1, \dots, x_m) \neq P(x_1) \cdots P(x_m).$$

θ の値が分かれば x_1, \dots, x_m は独立になる。

たにセンサ k からデータ x_k を入手した。データ点 (x_1, \dots, x_{k-1}) と融合してデータ点 (x_1, \dots, x_k) を得ると、事後信念は $P(\theta | x_1, \dots, x_k)$ と更新される。以下では、 $P(\theta | x_1, \dots, x_{k-1})$ を用いて $P(\theta | x_1, \dots, x_k)$ を計算するという再帰ベイズ更新ルールの式を導出する。

乗法公式より、

$$P(\theta, x_1, \dots, x_k) = P(\theta | x_1, \dots, x_k)P(x_1, \dots, x_k), \quad (22)$$

$$P(\theta, x_1, \dots, x_k) = P(x_1, \dots, x_k | \theta)P(\theta). \quad (23)$$

データ点 (x_1, \dots, x_{k-1}) とデータ x_k は θ の下で条件付き独立であるから、(23) 式は、

$$P(x_1, \dots, x_k | \theta)P(\theta) = P(x_k | \theta)P(x_1, \dots, x_{k-1} | \theta)P(\theta) \quad (24)$$

となる。(22), (24) 式より、

$$P(\theta | x_1, \dots, x_k)P(x_1, \dots, x_k) = P(x_k | \theta)P(x_1, \dots, x_{k-1} | \theta)P(\theta), \quad (25)$$

となり、再び乗法公式より、

$$P(x_1, \dots, x_{k-1} | \theta)P(\theta) = P(\theta | x_1, \dots, x_{k-1})P(x_1, \dots, x_{k-1}). \quad (26)$$

であるから、(25) 式は、

$$P(\theta | x_1, \dots, x_k) = \frac{P(x_k | \theta)P(\theta | x_1, \dots, x_{k-1})P(x_1, \dots, x_{k-1})}{P(x_1, \dots, x_k)} \quad (27)$$

となる。ここで、

$$\frac{P(x_1, \dots, x_k)}{P(x_1, \dots, x_{k-1})} = P(x_k | x_1, \dots, x_{k-1}). \quad (28)$$

これを (27) 式に代入することで、再帰ベイズ更新ルール

$$P(\theta | x_1, \dots, x_k) = \frac{P(x_k | \theta)P(\theta | x_1, \dots, x_{k-1})}{P(x_k | x_1, \dots, x_{k-1})} \quad (29)$$

を得る¹⁶⁾。

再帰ベイズ更新ルールは次のように適用される。消費者の人口分布 $\theta \sim N(\theta_0, \sigma_0^2)$ を初期信念 $P(\theta)$ とする。データ x_1 を入手すると、ベイズ・ルールに従って事後信念 $P(\theta | x_1)$ を計算する。続いてデータ x_2 を入手すると、事前信念 $P(\theta | x_1)$ を再帰ベイズ更新ルール(29)式に当てはめて、事後信念 $P(\theta | x_1, x_2)$ を計算する。さらにデータ x_3 を入手すると、事前信念 $P(\theta | x_1, x_2)$ を再帰ベイズ更新ルール(29)式に当てはめて、事後信念 $P(\theta | x_1, x_2, x_3)$ を計算する。このようにして、 $P(\theta | x_1, \dots, x_{k-1})$ まで計算したとする。これを事前信念として、その確率密度関数を、

$$P(\theta | x_1, \dots, x_{k-1}) = \frac{1}{\sqrt{2\pi}\sigma_{k-1}^2} \exp\left\{-\frac{1}{2} \frac{(\theta - \theta_{k-1})^2}{\sigma_{k-1}^2}\right\} \quad (30)$$

と置く。 σ_{k-1}^2 と θ_{k-1} はそれぞれ、融合されたデータ点 (x_1, \dots, x_{k-1}) から導かれた事前信念の分散と期待値である。尤度関数は次のようになる。

$$P(x_k | \theta) = \frac{1}{\sqrt{2\pi}\eta_k^2} \exp\left\{-\frac{1}{2} \frac{(x_k - \theta)^2}{\eta_k^2}\right\}. \quad (31)$$

事後信念を、

$$P(\theta | x_1, \dots, x_k) = \frac{1}{\sqrt{2\pi}\sigma_k^2} \exp\left\{-\frac{1}{2} \frac{(\theta - \theta_k)^2}{\sigma_k^2}\right\} \quad (32)$$

と置くと、再帰ベイズ更新ルール (29) 式から以下の関係が成り立つ。

$$\theta_k = \frac{\eta_k^2}{\sigma_{k-1}^2 + \eta_k^2} \theta_{k-1} + \frac{\sigma_{k-1}^2}{\sigma_{k-1}^2 + \eta_k^2} x_k, \quad (33)$$

$$\sigma_k^2 = \frac{\sigma_{k-1}^2 \eta_k^2}{\sigma_{k-1}^2 + \eta_k^2}. \quad (34)$$

σ_k^2 と θ_k はそれぞれ、融合したデータ点 (x_1, \dots, x_k) から導かれた事後信念の分散と期待値である。(34) 式より、 $\sigma_k^2 < \sigma_{k-1}^2$ であるから、異なるセンサから新たにデー

16) $P(x_k | x_1, \dots, x_{k-1})$ は θ の情報を与えるものではなく、データ x_k を入手した後は事後分布を正規化する定数に過ぎない。

タを入手して融合すると、推定の精度が上昇することが明らかになった。

θ の MAP 推定値は、

$$\hat{\theta} = \frac{\eta_k^2}{\sigma_{k-1}^2 + \eta_k^2} \theta_{k-1} + \frac{\sigma_{k-1}^2}{\sigma_{k-1}^2 + \eta_k^2} x_k \quad (35)$$

となる。(33), (34) 式を再帰的に解くことで次式を得る。

$$\theta_k = \frac{\sigma_k^2}{\sigma_0^2} \theta_0 + \frac{\sigma_k^2}{\eta_1^2} x_1 + \dots + \frac{\sigma_k^2}{\eta_k^2} x_k \quad (k = 1, \dots, m), \quad (36)$$

$$\sigma_k^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{1}{\eta_1^2} + \dots + \frac{1}{\eta_k^2}} \quad (k = 1, \dots, m). \quad (37)$$

これより、 $\sigma^2 > \sigma_1^2 > \sigma_2^2 > \dots > \sigma_{m-1}^2 > \sigma_m^2$ となることが分かる。

融合したデータ点 (x_1, \dots, x_m) から導かれる MAP 推定値は、

$$\hat{\theta} = \omega_0 \theta_0 + \sum_{k=1}^m \omega_k x_k \quad (38)$$

となる。ここで、ウエイトを、

$$\omega_0 = \frac{\sigma_m^2}{\sigma_0^2}, \quad (39)$$

$$\omega_k = \frac{\sigma_m^2}{\eta_k^2} \quad (k = 1, \dots, m) \quad (40)$$

と置と、ウエイト $\omega_k (i = 1, \dots, m)$ はデータの信頼性を表す定数となる。すなわち、 $\eta_h^2 > \eta_k^2$ となるとき、センサ h よりもセンサ k のほうが θ の計測器としての信頼性が高いので、ウエイトを $\omega_k > \omega_h$ と置く。

2.2 データと企業行動

データ駆動型市場で活動する企業をモデル化することに関心を移す。モデル化する上で明確にすべきことは、収集したデータをどのように用いて意思決定を行

うか、また、収集したデータの精度が利潤にどのような影響を与えるかということである。後者については、Acemoglu et al.(2022)が次のように定式化している。企業は消費者の選好パラメータの値が分からない¹⁷⁾。 L 人の消費者をターゲットとして、消費者 l ($l = 1, \dots, L$) の選好パラメータ θ_l に関するデータ x_l を収集する。データにはノイズが含まれており、

$$x_l = \theta_l + \varepsilon_l, \quad \varepsilon_l \sim N(0, \eta^2) \quad (l = 1, \dots, L) \quad (41)$$

となる。データに含まれるノイズは消費者間で独立である。 θ_l の推定値を $\hat{\theta}_l(x_l)$ とする。

新たにデータを入手すると推定値は更新されるが、消費者 l からデータを収集するたびに、企業はコスト c_l を負担しないとイケない。このコストは、データ市場やプラットフォームから金銭を支払ってデータを購入したり、データと交換でコンテンツやサービスを無料で消費者に提供したりすることで発生する¹⁸⁾。

選好パラメータを正確に推定できれば、それだけ消費者から多くの収入が得られる。つまり、企業の収入は推定誤差の減少関数となる。企業は、推定誤差とデータ収集コストの合計が最小になるように、各消費者から収集するデータの標本数を決める。つまり、

$$\min_{n_1, \dots, n_L} \sum_{l=1}^L \{E[(\hat{\theta}_l(x_l) - \theta_l)^2] + c_l n_l\} \quad (42)$$

ここで、 $x_l = (x_l^{(1)}, \dots, x_l^{(n_l)})$ は消費者 l から収集したデータ集合であり、 n_l は標本数である。データをたくさん収集すると推定値の精度が向上し収入が増加するが、

17) Acemoglu et al.(2022)は選好パラメータ自体が確率変数であり、消費者の間で選好が相関するケースを考えている。このとき、 θ_l の推定値は $\hat{\theta}_l(x_1, \dots, x_L)$ となる。したがって、ある消費者のデータを収集すると、他の消費者の選好がある程度明らかになるというデータ外部性が生じる。

18) 価値のあるデータは対価を支払って入手するのが当然である。サービスやコンテンツの供給に限界費用 c がかかるならば、サービスやコンテンツを無料で提供することは、企業が消費者に価格 c を支払ってデータを取得していることと同義になる。実際には、デジタル化されたサービスやコンテンツは複製が容易なので、限界費用はゼロであることが多い。したがって、データを価格ゼロで収集していることになる。

その代わりデータの収集コストも増加する¹⁹⁾。

続く小節では、消費者から収集したデータを用いて、パーソナライズド・プライシング（個別化された価格設定）、レコメンデーション、ターゲット広告といった意思決定を行う企業をモデル化して、推定誤差が小さいほど企業の収入が増加することを確認する。

2.2.1 パーソナライズド・プライシング

立地で差別化した商品を販売する企業 A、企業 B が市場で活動している。両企業の生産費用はゼロである。線形の商品空間 $[0, 1]$ において、企業 A は 0 地点、企業 B は 1 地点にそれぞれ立地している。消費者は $[0, 1]$ 上に一様に分布しており、消費者の選好パラメータ $\theta \in [0, 1]$ は、商品空間における消費者の住所となる。消費者が負担する単位移動費用を t とする。両企業が提供する商品は、どの消費者にも同じ価値 v を与える。企業 A、企業 B の価格をそれぞれ p_A, p_B とする。消費者は 1 単位だけ商品を購入する。住所 θ にいる消費者が企業 A で購入するときの余剰は、

$$V_A = v - t\theta - p_A, \quad (43)$$

企業 B で購入するときの余剰は、

$$V_B = v - t(1 - \theta) - p_B \quad (44)$$

となる。

企業 B はデータを収集せず、どの消費者にも一様価格 p_B を提示する。これに対して、企業 A は、消費者の住所に応じて価格を設定するというパーソナライズド・プライシングを実行する²⁰⁾。企業 A は直面している消費者の住所が正確に分

19) Acemoglu et al.(2022)は、消費者の効用関数を推定誤差の増加関数としている。この仮定は、消費者はプライバシーを重視しており、企業によって自己の嗜好を把握されることに苦痛を感じると解釈できる。

20) データの誤差により消費者に正の余剰を与えてしまうので、完全価格差別は実行できない。

からないので、この消費者の住所に関するデータを収集して、得られたデータに基づいて価格を設定する。企業 A は住所 θ の消費者から次のようなデータ x を入手する。

$$x = \theta + \varepsilon, \quad \varepsilon \sim \text{uni}[-\eta, \eta]. \quad (45)$$

η はノイズの幅であり、この値が小さいほどより正確な住所のデータが入手できる。企業 A は x の値を θ の推定値として、 x に依存した価格 $p_A(x)$ を設定する。

企業 A が設定する価格 $p_A(x)$ は次のようになる。企業 B で購入することで得られる余剰がゼロとなる消費者の住所は、

$$\theta_B = \frac{p_B - v + t}{t} \quad (46)$$

となるので、住所 $\theta \in [0, \theta_B]$ にいる消費者は企業 B から購入しない。このような消費者に対して、企業 A は余剰 V_A がゼロとなるように価格を設定する。つまり、

$$p_A(x) = v - tx. \quad (47)$$

住所 $\theta \geq \theta_B$ の消費者に対しては、どちらの企業で購入しても無差別となるように価格 $p_A(x)$ を設定する。すなわち、

$$tx + p_A(x) = t(1 - x) + p_B. \quad (48)$$

これより、

$$p_A(x) = t(1 - 2x) + p_B \quad (49)$$

となる²¹⁾。 $p_A(\theta) = 0$ となる消費者の住所は、

21) 住所 θ_B の消費者は、

$$V_B = v - t(1 - \theta_B) - p_B = 0$$

となるので、

$$v - t\theta_B = t(1 - 2\theta_B) + p_B$$

を満たす。つまり、この消費者に対しては(47)、(49)式のどちらの価格を設定しても同じである。

$$\theta_A = \frac{t + p_B}{2t} \tag{50}$$

であり、住所 $\theta \in (\theta_A, 1]$ にいる消費者に提示する価格はマイナスになるので、このような消費者には販売しない。住所 $\theta \in [\theta_B, \theta_A]$ の消費者に対して、(49) 式を満たす価格を提示する²²⁾。

θ について不正確なデータ $x \in (\theta, \theta + \eta]$ を入手したとき、消費者の住所を実際よりも遠くに推定してしまう。このときに消費者が得る余剰は、消費者 $\theta \in [0, \theta_B]$ について $V_A > 0$ 、消費者 $\theta \in [\theta_B, \theta_A]$ について $V_A > V_B$ となるので、企業 A で購入する。しかし、消費者が実際に支払える金額よりも低い価格を提示してしまうので、企業 A は本来得られたはずの利益を失う。 θ について不正確なデータ $x \in [\theta - \eta, \theta]$ を入手すると、消費者の住所を実際よりも近くに推定してしまい、高い価格を提示することで彼らへの販売を逃すことになる。このとき、消費者 $\theta \in [0, \theta_B]$ については $V_A < 0$ となり、どの企業からも購入しない。消費者 $\theta \in [\theta_B, \theta_A]$ については $V_A > V_B$ となるので、企業 B で購入する。

企業 B の期待利潤は次のようになる。

$$\begin{aligned} E\pi_B &= p_B(1 - \theta_A) + \int_{\theta_B}^{\theta_A} \int_{\theta - \eta}^{\theta} \frac{p_B}{2\eta} dx d\theta \\ &= \frac{p_B(2v + t - 3p_B)}{4t}. \end{aligned} \tag{51}$$

これより企業 B の最適価格は、

$$p_B^* = \frac{2v + t}{6} \tag{52}$$

となる²³⁾。企業 A の期待利潤は次のようになる。

22) $x \in (\theta_A, \theta_A + \eta]$ となるデータについて $p_A(x) < 0$ となるが、住所 θ_A にいる消費者である可能性を考慮して、負の価格で販売する。

23) $\theta_A < 1$ となるのは $v < 5t/2$ のとき、 $\theta_B < \theta_A$ となるのは $v > 7t/10$ のときである。このモデルは $v \in [7t/10, 5t/2]$ を前提としている。

$$\begin{aligned}
E\pi_A &= \int_0^{\theta_A} \int_{\theta}^{\theta+\eta} \frac{p_A(x)}{2\eta} dx d\theta \\
&= \int_0^{\theta_B} \int_{\theta}^{\theta+\eta} \frac{v-tx}{2\eta} dx d\theta + \int_{\theta_B}^{\theta_A} \int_{\theta}^{\theta+\eta} \frac{t(1-2x) + p_B}{2\eta} dx d\theta.
\end{aligned} \tag{53}$$

企業 A の均衡利潤は,

$$E\pi_A^* = \frac{1}{288t} (-72tv\eta - 28v^2 + 140tv - 13t^2) \tag{54}$$

となり, η の減少関数となる。したがって, 収集するデータの精度が高くなるほど均衡利潤も高くなることが示された²⁴⁾。

2.2.2 レコメンデーション

de Cornière&Taylor(2020), Guembel&Hege(2021)は, レコメンデーションを次のようにモデル化している。

市場には企業 A, 企業 B が活動しており, 両企業とも線形の商品空間 $[\underline{\theta}, \bar{\theta}]$ 上にあるどのような商品でも供給できる。どのような商品でも企業は同じ価格を提示し, 企業 A, 企業 B の価格をそれぞれ p_A, p_B とする。両企業の生産費用をゼロとする。消費者の選好パラメータ $\theta \in [\underline{\theta}, \bar{\theta}]$ は商品空間上の 1 点として表される。企業は θ の値が分からないので, θ に関するデータを収集する。企業 A, 企業 B が収集するデータをそれぞれ,

$$x_A = \theta + \varepsilon_A, \quad \varepsilon_A \sim \text{uni}[-\eta_A, \eta_A], \tag{55}$$

$$x_B = \theta + \varepsilon_B, \quad \varepsilon_B \sim \text{uni}[-\eta_B, \eta_B] \tag{56}$$

とする。 $\varepsilon_A, \varepsilon_B$ は互いに独立である。データ x を θ の推定値として, 商品空間上の $x \in [\underline{\theta}, \bar{\theta}]$ に位置する商品を消費者に推奨する。消費者は推奨された商品を 1 単位

24) Gu, Madio&Reggiani(2019), Montes, Sand-Zantman&Valletti(2019)は, パーソナライズド・プライシングを次のようにモデル化している。蓄積したデータから, 全体の消費者うち割合 a の消費者について住所を知っているので, 完全価格差別を実行する。割合 $1-a$ の消費者についてはデータがなく住所が分からないので, 一様価格を提示する。

だけ購入する。選好パラメータ θ とレコメンデーション x との乖離が大きいほど、消費者が受ける苦痛（負の効用）が大きくなる²⁵⁾。

企業 A、企業 B は、それぞれレコメンデーション x_A, x_B を消費者に提示する。レコメンデーションが明らかになる前の消費者余剰は確率変数であり、企業 A について、

$$V_A = v - (\theta - x_A)^2 - p_A = v - \varepsilon_A^2 - p_A, \quad (57)$$

企業 B について、

$$V_B = v - (\theta - x_B)^2 - p_B = v - \varepsilon_B^2 - p_B \quad (58)$$

となる。 v は消費者の選好に関係なく商品が消費者に与える価値であり、両企業とも同じ水準である²⁶⁾。 $(\theta - x)^2$ は選好と異なる商品を消費することで被る負の効用である。

$$\varepsilon_A^2 - \varepsilon_B^2 \sim \text{uni}[-\eta_B^2, \eta_A^2] \quad (59)$$

であるから、企業 A から購入する確率は、

$$\begin{aligned} \Pr(\varepsilon_A^2 - \varepsilon_B^2 \leq p_B - p_A) &= \int_{-\eta_B^2}^{p_B - p_A} \frac{1}{\eta_A^2 + \eta_B^2} d\varepsilon \\ &= \frac{p_B - p_A + \eta_B^2}{\eta_A^2 + \eta_B^2} \end{aligned} \quad (60)$$

となる。これより、企業 A の期待利潤は次のようになる。

$$\begin{aligned} E\pi_A &= \Pr(\varepsilon_A^2 - \varepsilon_B^2 \leq p_B - p_A) \times p_A \\ &= \frac{p_B - p_A + \eta_B^2}{\eta_A^2 + \eta_B^2} p_A. \end{aligned} \quad (61)$$

25) 推奨される商品が消費者の選好に近いほど満足が高まるので、レコメンデーションの精度が高くなることと、商品の品質が高くなることは同義である。

26) v は十分大きな値であり、すべての消費者がどのような商品を購入しても正の余剰が得られるものとする。

企業 A の最適反応関数は,

$$p_A = \frac{p_B + \eta_B^2}{2} \quad (62)$$

となる。同様に、企業 B の期待利潤は次式で与えられる。

$$\begin{aligned} E\pi_B &= \Pr(\varepsilon_A^2 - \varepsilon_B^2 \geq p_B - p_A) \times p_B \\ &= \frac{p_A - p_B + \eta_A^2}{\eta_A^2 + \eta_B^2} p_B. \end{aligned} \quad (63)$$

企業 B の最適反応関数は次のようになる。

$$p_B = \frac{p_A + \eta_A^2}{2}. \quad (64)$$

これよりナッシュ均衡価格が求まり,

$$p_A^* = \frac{\eta_A^2 + 2\eta_B^2}{3}, \quad (65)$$

$$p_B^* = \frac{\eta_B^2 + 2\eta_A^2}{3} \quad (66)$$

となる。均衡価格は η_A と η_B の増加関数である。すなわち、どちらかの企業が収集するデータの精度が上昇すると、価格競争が激化して均衡価格は低下する。

均衡利潤は,

$$E\pi_A^* = \frac{\eta_A^2 + 2\eta_B^2}{3(\eta_A^2 + \eta_B^2)}, \quad (67)$$

$$E\pi_B^* = \frac{2\eta_A^2 + \eta_B^2}{3(\eta_A^2 + \eta_B^2)} \quad (68)$$

となる。企業 A の均衡利潤は η_A^2 の減少関数、 η_B^2 の増加関数となる。つまり、自社が収集するデータの精度が上昇すると、自社の利潤は増加するが、競合企業が収集するデータの精度が上昇すると、自社の利潤は低下する。

2.2.3 ターゲット広告

Rutt(2012), de Cornière&Taylor(2020)は、データを用いたターゲット広告を次のようにモデル化している。

独占的なプラットフォームと、差別化された商品を販売する無数の企業群、および無数の消費者群を考える。企業群と消費者群の測度は1である。各企業は商品を独占的に販売している。消費者の商品に対する評価は、商品が選好にマッチするときは $V > 0$ であり、マッチしないときは0となる。消費者はマッチするときに商品を1単位だけ購入し、このとき商品価格は V となる。 θ は消費者が商品にマッチする確率（マッチング確率）であり、消費者の選好を顕示する定数である。どの商品についても、消費者群のマッチング確率は $[0, 1]$ 上の分布 $P(\theta)$ に従うものとする。個々の消費者について、商品のマッチング確率を低い順に並べると、商品群は $[0, 1]$ 上の一様分布を形成するものと仮定する。

どの企業も個々の消費者が持つマッチング確率は分からないが、マッチング確率の分布は知っているものとする。つまり、 θ の事前信念は $P(\theta)$ で表される。 $E(\theta) = \bar{\theta}$ として、この値を企業が算出するマッチング確率の推定値とする。期待収入は $V\bar{\theta}$ となる。

すべての消費者がプラットフォームのWebサイトを訪問・閲覧して、無料サービスや無料コンテンツを享受する。その過程でプラットフォームによってデータが収集される。消費者から収集するデータは、商品群のマッチング確率に関するもので、データにはノイズが含まれる。企業はプラットフォームから、自社商品のマッチング確率に関するデータを無料で提供される。企業は全消費者のデータを受け取り、データの値は $x \sim \text{uni}[0, 1]$ と分布している。企業はこれらのデータを基に、各消費者についてマッチング確率の推定値を更新する。データ x を入手したときの事後信念を $P(\theta | x, \eta)$ とする。ここで、 η はデータの精度の逆数である。すなわち、 η が小さいほど、データ x はマッチング確率に関してより精度の高い情報を伝える。条件付き期待値を $\hat{\theta}(x, \eta)$ として、データ x を入手したことで更新されたマッチング確率の推定値とする。したがって、データ x が得られた消費者について、そのマッチング確率は $\hat{\theta}(x, \eta)$ と推定して、期待収入は $V\hat{\theta}(x, \eta)$ と予想する。

$\hat{\theta}(x, \eta)$ は次のような性質を持つ。まず,

$$\hat{\theta}_x = \frac{\partial \hat{\theta}(x, \eta)}{\partial x} > 0. \quad (69)$$

つまり, より高い値を取るデータは, マッチング確率が高いことを伝える情報になるので, 推定値を上方に更新する。また,

$$\hat{\theta}_\eta = \frac{\partial \hat{\theta}(x, \eta)}{\partial \eta} \begin{cases} \geq 0 & (x \leq \tilde{x} \text{ のとき}) \\ < 0 & (x > \tilde{x} \text{ のとき}). \end{cases} \quad (70)$$

$\tilde{x} \in (0, 1)$ はマッチング確率に関する良い情報・悪い情報の境界値となり, データ $x > \tilde{x}$ は良い情報であり, データの精度が上昇したときにこのようなデータを入手すると, 推定値をより上方に更新する。他方, データ $x < \tilde{x}$ は悪い情報なので, データの精度が上昇したときにこのようなデータを入手すると, 推定値をより下方に更新する²⁷⁾。さらに,

$$\hat{\theta}_{x\eta} = \frac{\partial^2 \hat{\theta}(x, \eta)}{\partial \eta \partial x} \leq 0 \quad (71)$$

とする。

マッチング確率が高い消費者は, 自社商品の存在さえ知っていれば購入する確率が高い。そこで, この消費者が閲覧するプラットフォームの Web ページに, 自社商品のターゲット広告を掲載してもらう。広告スペース 1 単位当たりの料金は ϕ である²⁸⁾。ターゲット広告にはコストがかかるため, マッチング確率が高い消費者をターゲットにする。したがって, ターゲット広告を掲載するデータの境界値を \underline{x} を定めて, データ $x \in [\underline{x}, 1]$ が得られた消費者にターゲット広告を掲載するよ

27) 例えば, (18)式と類似した

$$\hat{\theta}(x, \eta) = \eta \bar{\theta} + (1 - \eta)x$$

を推定値とすると, 境界値は,

$$\tilde{x} = \bar{\theta}$$

となる。

28) プラットフォームが広告を掲載する企業を選別することがある。プラットフォームは消費者の選好パラメータ θ の値を知っており, 消費者 θ にターゲット広告を掲載したい企業を募る。プラットフォームはどの企業のターゲット広告を提示するかを, オークション等の選択メカニズムによって決める。

うにプラットフォームに注文する。 \underline{x} は期待利潤

$$\pi = \int_{\underline{x}}^1 \{V\hat{\theta}(x, \eta) - \phi\} dx \quad (72)$$

が最大になるように定められる。1階の条件より \underline{x} は、

$$V\hat{\theta}(\underline{x}, \eta) = \phi. \quad (73)$$

を満たす。企業がプラットフォームに発注する広告スペースは $n=1-\underline{x}$ となるから、広告に対する逆需要関数は、

$$P(n, \eta) = \phi = V\hat{\theta}(1-n, \eta) \quad (74)$$

となる。広告スペース n を広げると、より低いマッチング確率を持つ消費者にも広告を提示してしまい、広告の効果が低下するため、広告料金も低下する。また、 $n < 1-\tilde{x}$ のとき P は η の減少関数となるから、より正確なデータが得られるとき、広告料金は上昇する。

プラットフォームの利潤は次のようになる。

$$\Pi = nP(n, \eta) = nV\hat{\theta}(1-n, \eta). \quad (75)$$

プラットフォームは、利潤が最大になるように広告スペース n^* を用意する。1階の条件より n^* は、

$$\hat{\theta}(1-n^*, \eta) - n^* \hat{\theta}_{x\eta}(1-n^*, \eta) = 0 \quad (76)$$

を満たす。 $n^* < 1-\tilde{x}$ のとき、

$$\frac{dn^*}{d\eta} = \frac{\hat{\theta}_\eta - n^* \hat{\theta}_{x\eta}}{2\hat{\theta}_x - n^* \hat{\theta}_{xx}} < 0 \quad (77)$$

が成り立つ。つまり、データの精度が高まると、プラットフォームが提供する広告スペースは広がる。

消費者はプラットフォームのWebサイトを訪問して、無料コンテンツを閲覧したり、無料サービスを享受したりすることで正の効用を得る。同時に、Webページにあるターゲット広告を目にすることで負の効用（苦痛）を被る。消費者は、

正の効用と負の効用を考慮に入れて、Web サイトに訪問する頻度や Web ページを閲覧する時間を決める。この頻度や時間を閲覧強度（browsing intensity）と呼ぶことにする。消費者の効用関数は次のようになる。

$$U = u(z) - zn\varphi, \quad u' > 0, \quad u'' < 0. \quad (78)$$

z は閲覧強度、 $u(z)$ は無料コンテンツや無料サービスから得られる効用である。 φ は単位当たり広告スペースから受ける苦痛である。消費者が目にする広告の総量は、広告スペース×閲覧強度で測られる。最適な閲覧強度は、

$$z^* = u'^{-1}(n^*\varphi) \quad (79)$$

となる。

消費者の閲覧強度が高いほど、精度の高いデータが得られるものとする。つまり、

$$\eta = \eta(z), \quad \eta' < 0. \quad (80)$$

これより、

$$\frac{d\eta}{dn^*} = \frac{d\eta}{dz^*} \frac{dz^*}{dn^*} > 0 \quad (81)$$

であるから、Web ページの広告スペースが広がるほど消費者は苦痛を強く感じるため、閲覧強度は減少する。結果として、消費者から収集するデータの精度が低下する。

3. データ駆動型ネットワーク効果と寡占競争

データ駆動型寡占市場では、データを多く所有している企業が競争上有利になるため、競合する企業の間でデータの収集競争が繰り広げられる²⁹⁾。データを用いて商品開発を進めることで将来に優れたサービスを提供したり、データを分析して効果的な販売促進活動を行うことで将来の販売につなげたりする。すなわち、データの収集・蓄積とは、将来の競争を見据えた行動である。このことから、データ駆動型寡占市場におけるデータ収集競争は動学的な側面を持っている³⁰⁾。データ駆動型寡占市場を動学モデルとして構築し、企業間で展開されるデータ収集競争の長期的な帰結を明らかにした代表的なモデルが Prüfer&Schottmüller (2021)と Hagiü&Wright(2023)である。どちらのモデルもサービスを販売する複占企業を考えており、品質の高いサービスを提供する企業に需要が集まる³¹⁾。データを多く所有する企業は、品質の高いサービスを提供することができる。この意味で市場はデータ駆動型であり、なおかつデータ駆動型ネットワーク効果が発揮する。Hagiü&Wright モデルの想定するデータ駆動型ネットワーク効果は、蓄積したデータ量が増えるほど、より高品質なサービスを提供することができるというものである。Prüfer&Schottmüller モデルの想定するデータ駆動型ネットワーク効果は、データを活用すると、サービスの品質を改良する投資が効率的に行われるという

29) データに非競争性と非排除性があれば、データ収集競争は起こり得ない。データの非競争性とは、複数企業が同じデータを同時にアクセスできるので、自社がデータを取得しても、同じデータを他社が取得することを妨げないということである。データの非排除性とは、自社がデータを使用してもデータが消滅することはないし、デジタル化されたデータはコピーが容易であるので、同じデータを他社が使用することを妨げることができないということである。しかし、データの種類によっては収集に莫大なコストがかかるので、すべての企業が容易にアクセスできるわけではない。収集したデータを企業秘密にして、他社にアクセスできないようにすることも可能である。したがって、この種のデータには排除性がある。また、最初にアクセスした企業に大きな情報価値がもたらされるが、2番目以降にアクセスしても情報価値が低下するようなデータが存在する。このようなデータには競争性がある。

30) 動学寡占理論の研究を概観した文献として、Fudenberg&Tirole(2002)が挙げられる。

31) 消費者のニーズにフィットした検索結果やレコメンデーションを、品質の高いサービスと解釈する。

ものである³²⁾。以下では、これらのモデルの概要を解説する。

3.1 Hagiu&Wright モデル

市場では企業1と企業2が活動しており、無限期間に渡って測度1の消費者にサービスを提供している。両企業ともサービス生産の費用は同じであり、価格で競争する。毎期間、消費者はどちらかの企業から1単位だけサービスを購入する。企業はサービスを購入した顧客からデータを収集して蓄積する。データの収集・蓄積にはコストがかからない。また、蓄積したデータは劣化しないものとする³³⁾。マシンにたくさんのデータを投入すると、それだけマシンは学習するので、次期に提供されるサービスの品質が大きく向上する。この品質向上が、すべての消費者を満足させるのか、特定の消費者だけ満足させるのかで、想定するサービスが異なる³⁴⁾。したがって、以下の2つのケースを区別して分析する。

ユーザー間学習 (across-user learning) モデルでは、マシンがすべての顧客のデータを基に学習することで、顧客全体を満足させるサービスを提供する。これまでに企業 i で購入した累積顧客数 N_i ($i = 1, 2$) が企業 i の蓄積したデータ量になる。特定の消費者のデータを入手したか否かは、サービスの品質に影響を与えることはない。つまり、ある特定の消費者の購入決定が、将来提供されるサービスの品質に影響を与えることはないので、消費者は今期の余剰のみで購入先の企業を決める。

ユーザー内学習 (within-user learning) モデルでは、マシンが特定の顧客のデータを基に学習することで、その顧客だけを満足させるサービスを提供する。ある消費者が企業 i で購入した回数 N_i が、その消費者について蓄積したデータ量となる。消費者が企業 i で購入すればするほど、その消費者のデータが蓄積され、企

32) データ駆動型ネットワーク効果は、データを利用することで生じる供給の学習効果である。このことから、ユーザー間の繋がりが効用を生み出すことを意味する「ネットワーク効果」という言葉を使うことは不適切であるという意見がある (Casado&Lauten(2019))

33) 第 t 期までに蓄積されたデータ量を N_t 、第 t 期に収集するデータ量を D_t 、データの減耗率を ρ とする。第 $t+1$ 期に蓄積されるデータは次のようになる。

$$N_{t+1} = D_t + (1 - \rho)N_t.$$

34) 前者は検索エンジン、後者はレコメンデーション・エンジンが該当する。

業 i からより品質の高いサービスを受けることができる。一方で、企業 j が提供するサービスの品質は変化しない。それゆえ、企業 i で購入する回数が増えるほど、企業 i から企業 j へスイッチすることが益々困難になる。したがって、消費者は将来上昇するスイッチング・コストを見据えてサービスを購入する企業を選択する。

(1) 学習技術

企業 i のサービスから得られる効用は次のようになる。

$$u_i = v_i + f_i(N_i). \tag{82}$$

v_i は学習成果とは関係なく提供されるサービスの価値である。投入されたデータ量 N_i からマシンは学習して、その成果としてサービスの品質が上昇する。学習成果として生み出されたサービスの価値が $f_i(N_i)$ である。関数 f_i はマシンの学習技術を表す関数であり、機械学習の学習曲線に類似した次のような性質を持つ。学習の最大データ規模 \bar{N}_i が存在して、

$$i) \quad f_i(N_i + 1) > f_i(N_i), \quad \forall N_i < \bar{N}_i \tag{83}$$

$$ii) \quad f_i(N_i) = f_i(\bar{N}_i), \quad \forall N_i \geq \bar{N}_i \tag{84}$$

$$iii) \quad f_i(0) = 0 \tag{85}$$

を満たす。投入されたデータが多いほどそれだけ学習して、より優れたサービスを提供するが（性質 i)), 学習成果には上限があり、 \bar{N}_i を超えるデータ量を投入しても品質は上昇しない（性質 ii)),

より現実的な学習技術として、S 字型の学習曲線が考えられる。単純な S 字型の学習曲線は次式のように書ける（図 4）。

$$f_i(N_i) = \alpha_i \text{Max}\{\min\{N_i, \bar{N}_i\} - \underline{N}_i, 0\}. \tag{86}$$

パラメータ α_i はデータの限界的な学習成果であり、その大きさは学習速度を表し

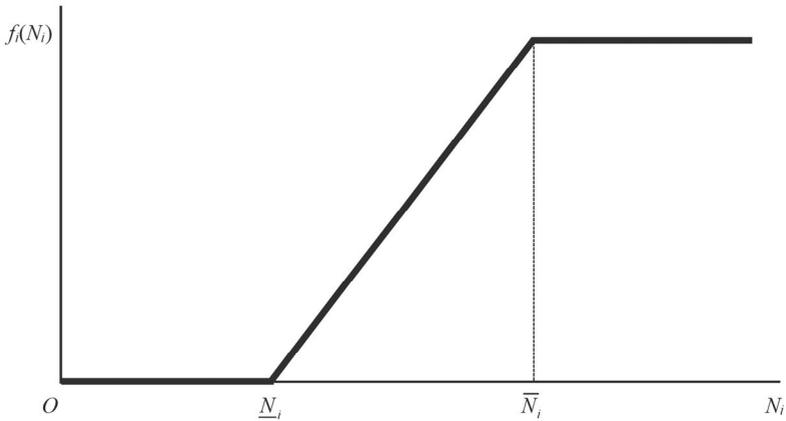


図4 S字型学習曲線

ている。投入されるデータが少なくても、学習が速ければ高い品質を持つサービスが生み出される。 \underline{N}_i は学習の最小データ規模であり、投入したデータ量がこの水準を超えないと学習成果が現れず、品質は上昇しない。

さらに、データの新規性 (= Velocity) が学習成果に影響を与えるような学習技術も考えることができる。これまで蓄積したデータ量を N_i とする。新たにデータを収集して蓄積するとデータ量が N'_i となった。したがって、データ量 N'_i のうち、古いデータは N_i 、新しいデータは $N'_i - N_i$ となる。データ量 N'_i から生み出されるサービスの価値は次のようになる。

$$\begin{aligned} \tilde{f}_i(N'_i) &= f_i(N_i) + \beta \{f_i(N'_i) - f_i(N_i)\} \\ &= \alpha_i [\beta \text{Max}\{\min\{N'_i, \bar{N}_i\} - \underline{N}_i, 0\} + (1 - \beta) \text{Max}\{\min\{N_i, \bar{N}_i\} - \underline{N}_i, 0\}] \end{aligned} \quad (87)$$

β は定数であり、この値が大きいくほど、新しいデータが学習成果に大きく寄与する³⁵⁾。

35) 通常、参入企業は既存企業よりも蓄積しているデータ量が少ない。参入企業は新しいデータを多く蓄積し、既存企業は古いデータを多く蓄積しているとき、 β が大きいく、なおかつ両企業ともデータ量が \underline{N}_i を超えているならば、参入企業は既存企業よりも高い品質を持つサービスを提供できる。

(2) 均衡

ユーザー間学習モデルであれ、ユーザー内学習モデルであれ、導かれる結果は同じである。関数 $\Delta(N_1, N_2)$ が存在して、

$$v_2 - v_1 > \Delta(N_1, N_2) \quad (88)$$

のときは、すべての期間で全消費者が企業 2 でサービスを購入する。

$$v_2 - v_1 < \Delta(N_1, N_2) \quad (89)$$

のときは、すべての期間で全消費者が企業 1 でサービスを購入する。ここで、

$$\Delta(N_1, N_2) = (1 - \delta) \left(\sum_{j=0}^{\infty} \delta^j f_1(N_1 + j) - \sum_{j=0}^{\infty} \delta^j f_2(N_2 + j) \right) \quad (90)$$

であり、 N_1 の増加関数、 N_2 の減少関数となる。 δ は割引因子である。

第 t 期に (89) 式が成り立つとき、すべての消費者が企業 1 からサービスを購入する。すると、 N_1 が増加して、 $\Delta(N_1, N_2)$ の値も増加する。再び (89) 式が成立するので、第 $t+1$ 期においてもすべての消費者が企業 1 から購入する。つまり、ある期に企業 1 が市場を支配しているならば、その期以降も企業 1 の市場支配が続くことが示された³⁶⁾。

3.2 Prüfer&Schottmüller モデル

市場では企業 1 と企業 2 が活動しており、測度 1 の消費者にサービスを提供している。両企業は無限期間に渡り投資を実行し、サービスの品質を上昇させる（以降、品質投資と呼ぶ）。投資は交互に行われ、奇数期に企業 1 が投資し、偶数期に企業 2 が投資する³⁷⁾。

36) $\Delta(N_1, N_2)$ は \bar{N}_1 の増加関数、 \bar{N}_2 の減少関数である。また、 f_1 が上方にシフトすると $\Delta(N_1, N_2)$ は増加し、 f_2 が上方にシフトすると $\Delta(N_1, N_2)$ は減少する。このことから、(89)式が成立して企業 1 が市場を支配していても、学習技術が変化することで、(88)式が成立して企業 2 が市場を支配することがある。

37) 両企業が同時に投資をするときは、純粋均衡戦略が存在しない。

毎期間、消費者はどちらかの企業からサービスを1単位だけ購入する。企業はサービスを購入した消費者から、コストをかけずにデータを収集することができる。データは蓄積されず、当該期で収集したデータだけ品質投資に使用される。企業1が提供するサービスの品質を q_1 、企業2の品質を q_2 として、両企業の品質差を

$$\Delta = q_1 - q_2 \quad (91)$$

とする。消費者は、品質差だけでサービスを購入する企業を決める³⁸⁾。したがって、企業1の需要関数は $D_1(\Delta)$ 、 $D_1' > 0$ 、企業2の需要関数は $D_2(\Delta)$ 、 $D_2' < 0$ となる。ここで、 $0 \leq D_i \leq 1$ である。収集できるデータ量は需要量に比例するので、企業 i が収集するデータ量を $D_i(\Delta)$ とする。企業が負担する費用は投資費用のみであり、サービス供給の費用はゼロとする。

(1) データ駆動型市場

投資量 z_i が多いほど、企業 i が提供するサービスの品質は前期と比べてより高くなる。そこで、第 t 期の投資量を、

$$z_{i,t} = q_{i,t} - q_{i,t-1} \quad (92)$$

とする。添え字の $t, t-1$ はそれぞれ第 t 期、第 $t-1$ 期を意味する。品質投資では、投資量 z_i に加えて、収集したデータ量 D_i もインプットとなる。データを活用して投資が行われるので、市場はデータ駆動型である。データそのものが品質を向上させることはないが、投入するデータ量が多いほど、同じ投資量 z_i でより効率的に投資を行うことができるので、投資の限界費用をそれだけ削減させる。具体的な企業 i の投資費用関数を、

38) 両企業の価格は固定されている。例えば、消費者にサービスやコンテンツを無料で提供している、あるいは、サブスクリプション方式で定額料金を徴収していると想定すればよい。

$$c(z_i, D_i) = \frac{\alpha z_i^2}{2} + \beta z_i \{1 - D_i(\Delta)\} \quad (93)$$

とする。 α, β は正の定数である。一般的には、

$$\frac{\partial^2 c}{\partial z_i \partial D_i} < 0 \quad (94)$$

が成り立つ投資費用関数を想定する。

導かれる結果は次のようなものである。第 t 期の品質差 Δ_t 、第 $t+1$ 期の品質差 Δ_{t+1} について、 $\Delta_t > 0$ のときは $\Delta_{t+1} > \Delta_t$ 、 $\Delta_t < 0$ のときは $\Delta_{t+1} < \Delta_t$ となる。 $\Delta_t > 0$ のとき、品質差により企業 1 の需要量が増加し、収集できるデータ量も増加する。このデータを投入すると投資の限界費用が減少するので、以前よりも投資量を増やすことが最適となる。これにより、企業 1 が提供するサービスの品質が今期よりも上昇する。一方、企業 2 の需要量は減少し、収集できるデータ量も減少する。すると、投資量が同じ水準であっても投資の限界費用が増加するので、以前よりも投資量を減らさざるを得ない。これにより、企業 2 が提供するサービスの品質は今期よりも低下する。こうして次期の品質差 Δ_{t+1} は上昇する。

この結果から、初期状態における品質差がどのような水準であっても、時間が経つにつれ品質差は広がり、品質の高い企業の市場シェアは拡大していき、長期的にはこの企業が市場を支配することが分かる。したがって、初期状態の品質差が $\Delta > 0$ のときは企業 1 が市場を支配し、 $\Delta < 0$ のときは企業 2 が市場を支配する。

(2) コネクティッド・マーケット

データ駆動型市場で支配的な企業が、データを活用する習慣がない伝統的市場に参入して、既存企業（以下、伝統的企業と呼ぶ）との間でサービス競争を繰り広げる。参入企業（以下、データ駆動型企業と呼ぶ）はデータ駆動型市場で収集したデータを伝統的市場の品質投資に活用する。伝統的企業はデータを用いずに品質投資を行う。

A 市場はデータ駆動型市場であり、B 市場は伝統的市場であるとする。A 市場で収集したデータを B 市場における品質投資のインプットとして使用できると

き、A市場とB市場はコネクティッドであると呼ぶ。投資量 z_B に加えて、A市場で収集したデータ量 D_A をB市場の品質投資に投入するとき、データ駆動型企業の投資費用関数は $c_B(z_B, D_A)$ となる。投資費用関数は次の関係を満たす。

$$\frac{\partial^2 c_B}{\partial z_B \partial D_A} < 0. \quad (95)$$

(94)式と類似した関係があるが、コネクティッド・マーケットではA市場のデータ量 D_A がB市場の投資の限界費用を削減させる。他方、伝統的企業は投資量 z_B だけでサービスの品質向上を図るので、投資費用関数は $c_B(z_B)$ となる。

A市場で集めた大規模データをB市場の品質投資に用いると、投資の限界費用が著しく低下する。すると、投資量 z_B を大幅に増やすことが最適となるので、伝統的企業よりも著しく品質の高いサービスを提供することになる。それゆえ、サービス競争はデータ駆動型企業に有利に働く。その結果、伝統的市場でも品質差が拡大していき、長期的にはデータ駆動型企業が市場を支配する。データを活用できない伝統的企業は市場から駆逐され、B市場で活動する企業はすべてデータ駆動型となる。こうして伝統的市場がデータ駆動型市場へと転換される。

さらに、B市場で支配的な地位を獲得した企業が、他の伝統的なC市場にも参入し、A市場とB市場で蓄積した膨大なデータを品質投資に投入する。その結果、C市場もデータ駆動型企業が支配し、データ駆動型市場に転換される。このように、膨大なデータを所有するデータ駆動型企業が伝統的市場に参入し、データ駆動型市場に転換するという流れが、ドミノ倒しのように起こる³⁹⁾。

4. まとめと今後の課題

需要を獲得するためには、消費者の選好を知る必要がある。しかし、選好そのものを把握することができないため、企業はデータを集めて分析して、選好に関

39) Google は、サーチ・エンジンを足掛かりにして、自動車、ヘルス・ケア、金融へと進出している。また、Amazon は書籍のオンライン直販から出発して、マーケットプレイス、クラウド・コンピューティング、医療、衛星通信へと進出している。

するより正確な情報を得ようとする。現代産業では、市場がデジタル化されているのでデータが入手しやすくなり、人工知能の技術が発展して高度なデータ分析が可能になったことから、企業間で繰り広げられるデータ収集競争はますます激しくなっている。収集・蓄積するデータが多ければ多いほど、新規顧客を開拓したり、競合企業から顧客を奪ったりすることが容易になる。顧客が増えれば増えるほど、それだけデータが企業内に流入し、蓄積されるデータ量が増える。つまり、データが多い企業に需要が集まり、需要が多い企業にデータが集まるのである。このようなデータ駆動型ネットワーク効果により企業は巨大化し、市場は特定の企業によって支配される。

データ駆動型市場の寡占競争を、動学寡占ゲームにより定式化することは有効な分析手法となるが、このような研究は緒に就いたばかりであり、今後のさらなる発展が期待される⁴⁰⁾。モデルの基本的な設定は次のようになる。操作変数は企業行動であり、行動を取る過程で副産物としてデータが発生する。行動の水準が高いほど多くのデータが生成されるが、それだけコストもかかるため、各期間で収集できるデータ量には限界がある。状態変数は蓄積されたデータ量である。蓄積されたデータは、アルゴリズムで規定される学習関数に投入され、出力結果は企業行動に反映される。この企業行動が次の期に生成されるデータ量を決める。データ駆動型寡占市場をこのようにモデル化して、数値計算によって具体的な均衡経路を導出する。その上で、様々な初期状態から経路を出発させたり、モデルのパラメータを連続的に変化させたり、ある時間でショックを発生させたりして、均衡経路の変化を観察する。このようなシミュレーション分析を通じて、データ駆動型市場における寡占競争が持つ固有の特徴が見出されるであろう。

本稿で取り上げたモデルは、入手できるデータ量が顧客数に比例すると仮定しているだけで、具体的にどのようなデータが入手できるのかについては明らかにしていない。実際のデータには様々なタイプがあり、データのタイプが異なると企業行動に与える影響も異なる。データのタイプを特定化して、その特性をモデルに明示した上で分析することで新たな知見が導かれ、データ駆動型市場に対す

40) 近年の研究として Kehoe, Larsen&Pastorino(2022)が挙げられる。

る理解がさらに深まることになる。様々なタイプのデータとして、以下の3つを挙げる。第一に、汎用的なデータと新規性が重要なデータとを区別する。属性データのような汎用的なものは、時間が経っても情報価値がなくなることがなく、なおかつ他の市場でも活用できる。これに対して、行動データは収集するタイミングによって常に変化するもので、新規性が重要となる。このようなデータをいくら蓄積しても、時間が経つにつれ陳腐化していき、価値ある情報が得られなくなる。第二に、中央集権的に収集されたデータと、分散した情報源から収集したデータとを区別する。データを取得する方法が、企業の意思決定に影響を与えるかどうかを検討することは、来るべきIoT時代では有意義な研究となるだろう。最後に、競争性・排除性のあるデータと、非競争性・非排除性を満たすデータとを明確に区別して分析する。この区別は非常に重要である。どちらの性質を持つデータを想定するかによって、データ収集競争が生じたり、あるいは、他の企業が収集したデータへのただ乗りが生じたりするので、探索すべき研究課題が大きく変わるからである。

参考文献

- [1] Acemoglu, D., A. Makhdomi., A. Malekian., and A. Ozdaglar. (2022), “Too Much Data : Prices and Inefficiencies in Data Markets.”, *American Economic Journal : Microeconomics*, Vol.14, pp.218-256.
- [2] Argenton, C. and J. Prüfer. (2012), ‘Search Engine Competition with Network Externalities.’, *Journal of Competition Law and Economics*, Vol. 8, pp.73-105.
- [3] Casado, M. and P. Luaten. (2019), “The Empty Promise of Data Moats.”, Andreessen Horowitz, <https://a16z.com/the-empty-promise-of-data-moats/>
- [4] Cornière, A. and G. Taylor. (2020), “Data and Competition : A General Framework with Applications to Mergers, Market Structure, and Privacy Policy.”, TSE Working Paper.
- [5] Durrant-Whyte, H. (2001), “Multi Sensor Data Fusion”, Australian Centre for Field Robotics, https://www.mit.bme.hu/system/files/oktatas/targyak/7357/MultiSensorDataFusionNotes_DurrantWhyte_2011.pdf
- [6] de Fortuny, E., D. Martens., and F. Provost. (2013), “Predictive Modeling with Big Data : Is Bigger Really Better? ”, *Big Data*, Vol. 1, pp.215-226.
- [7] Fudenberg, D. and J. Tirole. (2002), *Dynamic Models of Oligopoly*, Routledge.
- [8] Gu, Y., L. Madio and C. Reggiani. (2019), “Exclusive Data, Price Manipulation and Market Leadership”, CESifo working paper.
- [9] Guembel, A. and U. Hege. (2021), “Data, Product Targeting and Competition”, Working Paper.
- [10] Hagi, A. and J. Wright, (2023), “Data-enabled Learning, Network Effects and Competitive Advantage”, *RAND Journal of Economics*, Vol. 54, pp. 638-667.

- [11] Hestness, J., S. Narang., N. Ardalani., G. Damos., H. Jun., H. Kianinejad., M. Patwary., M. Ali., Y. Yang., and Y. Zhou. (2017), “Deep Learning Scaling is Predictable Empirically.”, Working Paper, arXiv:1712.00409.
- [12] John, G. and P. Langley. (1996), “Static Versus Dynamic Sampling for Data Mining.”, in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 367-370.
- [13] Katsov, I. (2017), *Introduction to Algorithmic Marketing : Artificial Intelligence for Marketing Operations.*, Ilia Katsov.
- [14] Kehoe, P., B. Larsen., and E. Pastorino(2022), “Dynamic Competition in the Era of Big Data”.Working Paper.
- [15] Mitchell, H. (2012), *Data fusion : Concepts and Ideas*, Springer.
- [16] Montes, R., W. Sand-Zantman., and T. Valletti. (2019), “The Value of Personal Information in Online Markets with Endogenous Privacy”, *Management Science*, Vol.65, pp.1342-1362.
- [17] OECD (2015), “Data-Driven Innovation : Big Data for Growth and Well-Being.”
- [18] Pradeep, A., A. Appel., and S. Sthanunathan. (2018), *AI for Marketing and Product Innovation : Powerful New Tools for Predicting Trends, Connecting with Customers, and Closing Sales.*, Wiley.
- [19] Prüfer, J. and C. Schottmüller. (2021), “Competing with Big Data.”, *Journal of Industrial Economics*, Vol. 69, pp. 967-1008.
- [20] Rosenfeld, J., A. Rosenfeld., Y. Belinkov., and N. Shavit. (2020), “A Constructive Prediction of the Generalization Error Across Scales”, Working Paper, arXiv:1909.12673.
- [21] Rutt, J. (2012), “Targeted Advertising and Media Market Competition.”, Working Paper, SSRN 2103061.
- [22] Schrage, M. (2020), *Recommendation Engines.*, MIT Press.
- [23] Stucke, M. and A. Grunes. (2016), *Big Data and Competition Policy.*, Oxford University Press.