

# SIRモデルを用いた感染症流行の数値解析

Numerical analysis of an infectious disease epidemic using SIR model

石原広夢, 平社和也, 水野貴敏

Hiromu Ishihara, Kazunari Hirakoso and Takatoshi Mizuno

玉川大学工学部エンジニアリングデザイン学科, 194-8610 東京都町田市玉川学園6-1-1

Department of Engineering Design, College of Engineering, Tamagawa University,  
6-1-1 Tamagawagakuen Machida-shi Tokyo 194-8610

## Abstract

Using SIR model, an influenza epidemic in a boy boarding school have been analyzed numerically. The simulation results show that the value and the time to peak of infection numbers has matched with the actual data. In addition, the numerical analysis has suggested the possibility of the accurate prediction of an infectious disease epidemic.

Keywords: SIR model, influenza epidemic, prediction of epidemics

## 1. はじめに

2019年12月に新型コロナウイルス感染症 (COVID-19) が中国湖北省武漢市で確認<sup>1)</sup>されて以降、瞬く間に世界的な大流行 (パンデミック) を引き起こし、現在も収束する気配が見えない。このように、人類は常に感染症の脅威にさらされてきた。例えば、14世紀のペスト (黒死病) はヨーロッパ人口の1/3以上を死に至らしめ、1918年のスペイン風邪 (インフルエンザ) のパンデミックでは、全世界で4000万人以上の死者を出した<sup>2)</sup>。

感染症には、その原因が明らかになっても完全な予防方法や治療方法が確立されていないものも多く、安全な日常生活を保持するためには、それらの流行の動向を予測し流行を制御する必要がある。そのために、感染症の流行過程を解明する数理的研究が行われてきた。最初の数理的研究は、18世紀の数学者ベルヌーイの天然痘死亡率の寿命への影響に関する研究だと言われている<sup>3)</sup>、<sup>4)</sup>。そして、20世紀初頭にハーマーが、伝染病の流行の推移が感染者と感受性者の接触率に依存す

ると仮定し、この仮定がそれ以降の数理モデルに大きな影響を与えた<sup>3)</sup>。1920年代から1930年代にかけてカーマックとマッケンドリックが微分方程式を用いた数理モデル・SIRモデルを確立し<sup>3)</sup>、<sup>4)</sup>、これが現在の感染症数理モデルの基礎となっている。近年では、2009年の新型インフルエンザや現在も猛威を振るう新型コロナへの対策においても数理モデルによる数値解析が活用されている<sup>1)</sup>、<sup>2)</sup>。

本研究では、SIRモデルを用いて数値解析コードの作成と、そのコードによる、あるインフルエンザの流行データを対象とした数値解析を行った結果について報告する。

## 2. SIRモデル<sup>3)</sup>、<sup>5)</sup>

カーマックとマッケンドリックが確立したSIRモデルとは、次のような考え方に基づく。まず、対象となる人口集団を感受性者 (Susceptible)、感染者 (Infectious)、回復者 (Recovered) の3つの状態に分割する。ここで、回復者とは回復後

に免疫を獲得し再度感染をしない人を指し、除去者 (Reduced) とも言う。そして、経過時間  $t$  における感受性者数を  $S(t)$ 、感染者数を  $I(t)$ 、回復者数を  $R(t)$  とおくと、SIRモデルでは感染症の流行が次の微分方程式で記述される。

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) \quad (1)$$

$$\frac{dI(t)}{dt} = -\gamma I(t) + \beta S(t)I(t) \quad (2)$$

$$\frac{dR(t)}{dt} = \gamma I(t) \quad (3)$$

ここで、 $\beta$  は感染率、 $\gamma$  は回復率と呼ばれる定数で、感染率  $\beta$  は単位時間あたりに一人の感染者が一人の感受性者に感染させる割合を示し、回復率  $\gamma$  は単位時間あたりに一人の感染者が回復する割合を示す。(2)式より感染者の増減は、感染者の増加率  $\beta S(t)I(t)$  と減少率  $\gamma I(t)$  のバランスにより決まることがわかる。言い換えると、 $\beta S(t)I(t) > \gamma I(t)$  であるときに感染者数が増加し、逆に  $\beta S(t)I(t) < \gamma I(t)$  のときに感染者数が減少するというわけである。感染率  $\beta$  と回復率  $\gamma$  が定数であることに注意すると、このモデルでは、感受性者数の減少が感染者数を増加から減少へ転じさせる要因となるであろうことが容易に推測される。

次に、このモデルにおける感染拡大の初期段階について考察していく。多くの感染症では、ごく少数の感染者から感染が拡大していく。初期 ( $t=0$ ) の感受性者数  $S_0 (=S(0))$  が初期の感染者数  $I_0 (=I(0))$  に対して十分に大きい場合、感染拡大の初期段階では  $S(t) \approx S_0$  と近似することができる。この近似を用いると、(2)式は次のように線形化できる。

$$\frac{dI(t)}{dt} = (\beta S_0 - \gamma)I(t) = \lambda_0 I(t)$$

ここで、 $\lambda_0$  は初期成長率と呼ばれる定数である。この微分方程式を解くと、感染者数  $I(t)$  が次のように与えられる。

$$I(t) = I_0 e^{\lambda_0 t} \quad (4)$$

これは「感染初期段階では感染者数が指数関数的に増加する」というマルサスの法則を表している。このマルサスの法則の下で、感染症が流行する条件は、

$$\lambda_0 = \beta S_0 - \gamma > 0 \quad (5)$$

である。この(5)式は次のように書き直すことができる。

$$R_0 = \frac{\beta S_0}{\gamma} > 1 \quad (6)$$

ここで定義された定数  $R_0$  を基本再生産数と呼ぶ。この基本再生産数  $R_0$  は、一人の感染者が単位時間に感染させる人数  $\beta S_0$  と一人の感染者が回復までに要する時間  $1/\gamma$  の積である。したがって、(6)式は、一人の感染者が回復するまでに感染させる人数が一人より多いことを意味する。

続いて、対象となる人口集団の全人口  $N$  について考える。全人口  $N$  を感受性者、感染者、回復者の3つに分けたので、

$$N = S(t) + I(t) + R(t)$$

が成り立つことは自明である。そして、(1)式～(3)式の和を取ると、

$$\frac{d}{dt}[S(t) + I(t) + R(t)] = \frac{dN}{dt} = 0$$

となる。つまり、全人口  $N$  は常に一定である。また、初期の回復者  $R(0) = 0$  であることに注意すると、感染の初期:  $t=0$  においては、

$$N = S_0 + I_0 \quad (7)$$

を満たすことがわかる。

### 3. 事例：寄宿学校でのインフルエンザの流行

今回、1978年1月22日から2月4日にかけて英国北部の寄宿舎で発生したインフルエンザの流行<sup>6)</sup>を解析していく。まずは、参考文献[6]で報告されている内容のうち、インフルエンザの流行を理解する上で前提となる事項を以下にまとめておく。

- 10歳～18歳の少年763名が在籍していた。

- 30名を除く全員が寮生活を送っていた。
- 1月15日から18日にかけて一過性の熱性疾患にかかっていた。
- 1月22日（日曜日）に3人の学生が医務室を訪れた。

図1は、1月22日から2月4日の間の寝たきりで療養していた学生の推移である。最終的に、全体の約67%にあたる512名が3日から7日の間、授業を休み療養した<sup>6)</sup>。

今回は、この寄宿学校でのインフルエンザの流行をサンプルデータとして、数値解析を行っていく。なお、この寄宿学校でのインフルエンザの流行は、SIRモデルを用いた数値解析が行われており<sup>5)</sup>、よく合致する例として紹介されている。その際の計算のパラメータとしては、次の値が用いられている。

$S_0 = 762$ ,  $I_0 = 1$ ,  $\beta = 2.18 \times 10^{-3}$ ,  $\gamma/\beta = 202$   
この値から求めた $\lambda_0$ と $R_0$ はそれぞれ、

$$\lambda_0 = 1.22, \quad R_0 = 3.77$$

となる。

#### 4. 解析方法

SIRモデルを用いた感染流行の数値解析を行うにあたり、ルンゲクッタ法<sup>7)</sup>により(1)式～(3)式の微分方程式を解く数値解析コードを作成した。

これらの微分方程式を解くにあたり、(1)式～(3)式の中に現れる二つの定数、感染率 $\beta$ と回復率 $\gamma$ と、初期の感染者数 $I_0$ を決定する必要がある。なお、対象とする全人口 $N$ は既知なので、(7)式より、初期の感染者数 $I_0$ が決まれば、初期の感受性者数 $S_0$ は自動的に決まる。

感染率 $\beta$ は、本来、感染症ごとに固有の値を有すると思われるが、その値を直接的に知ることは難しい。もちろん、実験を行って $\beta$ を測定するなどということは不可能であるし、実効的な $\beta$ の値自体が、感染者の隔離や感受性者の予防対策の実施などの様々な要因により変わると考えられる。

一方、回復率 $\gamma$ については、その逆数 $1/\gamma$ に注

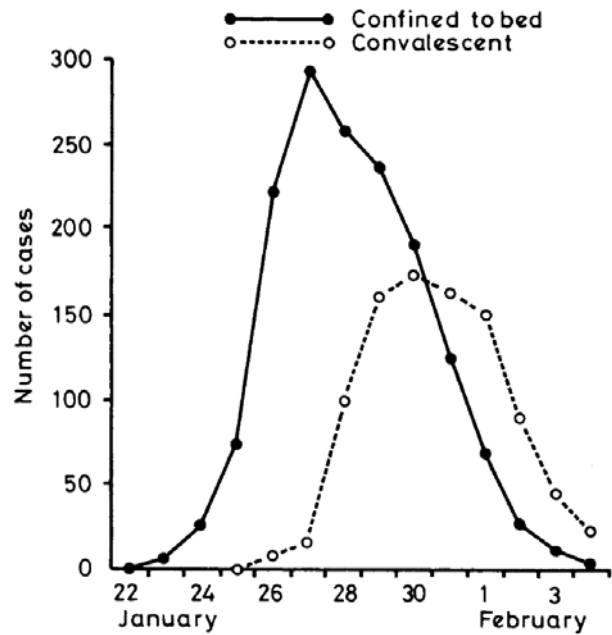


図1 1/22 から 2/4 の医務室の患者数の推移<sup>6)</sup>

目する。インフルエンザのような既存の感染症であれば、一人の感染者が回復するのに要する時間 $1/\gamma$ の値を、ある程度は経験的に推定することが可能である。しかし、未知の感染症の流行が起こったとき、即座に $1/\gamma$ の値を推定することは難しいだろう。つまり、感染症流行を解析するにあたり、必要なパラメータである $\beta$ と $\gamma$ の値を直接的に決定することは難しいと言わざるを得ない。

同様に、初期の感染者数 $I_0$ を決定することも困難である。元を辿れば、 $I_0 = 1$ となる可能性はあるが、例えば、海外から持ち込まれた感染症が日本国内で流行する場合などを考えると、必ずしも $I_0 = 1$ であるとは断定できない。また、仮に、顕在化している初期の感染者数 $I_0$ が $I_0 = 1$ であっても、その他に潜在的な初期の感染者がいる可能性は否定できない。

以上のことを踏まえて、数値解析に必要なパラメータ $\beta$ 、 $\gamma$ および $I_0$ をそれぞれ独立に決定するのではなく、例えば、感染症流行のデータに基づいて、推定していくべきだと考える。

そこで、感染率 $\beta$ と回復率 $\gamma$ ではなく、初期成長率 $\lambda_0$ と基本再生産数 $R_0$ に着目する。数値解析に必要なパラメータ $\beta$ と $\gamma$ は、 $\lambda_0$ 、 $R_0$ および $I_0$ を

用いて、次のように表すことができる。

$$\gamma = \frac{\lambda_0}{R_0 - 1}$$

$$\beta = \frac{\gamma R_0}{S_0} = \frac{\gamma R_0}{N - I_0}$$

したがって、初期成長率  $\lambda_0$ 、基本再生産数  $R_0$  および初期の感染者数  $I_0$  がわかれば、数値解析が可能となる。

ここからは、数値解析の手順について説明していく。はじめに、初期成長率  $\lambda_0$  と初期の感染者数  $I_0$  は、マルサスの法則に基づいて推定する。具体的には、初期段階における感染者数の推移を(4)式でフィッティングすることにより、 $\lambda_0$  と  $I_0$  の値を決定できる。ただし、残りの基本再生産数  $R_0$  をデータから推定することはできないので、 $R_0$  を任意に仮定する。このように決められた  $\lambda_0$ 、 $I_0$ 、 $R_0$  から  $\beta$ 、 $\gamma$  を求め、数値解析により(1)式～(3)式の解を求める。得られた解のうち、感染者数の時間変化  $I$  と、感染者数の推移の実データ  $I^*$  の残差の二乗平均  $\bar{\delta}$  を求める。残差の二乗平均  $\bar{\delta}$  は次式で与えられる。

$$\bar{\delta} = \sqrt{\frac{\sum (I - I^*)^2}{n}}$$

ここで、 $n$  は、感染者数の推移の実データのデータ数である。例えば、感染者数の1日ごとの推移のデータを5日分用いた場合は、 $n=5$  となる。

1stステップでは、 $\lambda_0$  と  $I_0$  を固定し、 $R_0$  を変動させて、パラメータサーベイを行い、 $\bar{\delta}$  が最小となる  $R_0$  を決める。次に、 $I_0$  と  $R_0$  を固定し、 $\lambda_0$  を変動させて、パラメータサーベイを行い、 $\bar{\delta}$  が最小となる  $\lambda_0$  を決める。これを2ndステップとする。そして、最後に、3rdステップとして  $\lambda_0$  と  $R_0$  を固定し、 $I_0$  を変動させて、パラメータサーベイを行い、 $\bar{\delta}$  が最小となる  $I_0$  を決める。この1stステップから3rdステップまでを1セットとして、 $R_0$ 、 $\lambda_0$ 、 $I_0$  の有効数字の範囲内で、 $\bar{\delta}$  が変化しなくなるまでパラメータサーベイを繰り返し行い、 $R_0$ 、 $\lambda_0$ 、 $I_0$  の推定を行う。この作業を  $R_0$ 、 $\lambda_0$ 、 $I_0$  の「最

適化」と呼ぶことにする。以上が今回実施した数値解析の手順である。

## 5. インフルエンザ流行の数値解析

解析を行うにあたり、初めに寄宿学校でのインフルエンザの流行における1月22日から1月24日までの初めの3日間のデータを、(4)式で表される指数関数でフィッティングし、 $\lambda_0$ 、 $I_0$  を以下のように推定した。

$$\lambda_0 = 1.06, \quad I_0 = 0.92$$

このパラメータからスタートして、 $R_0$ 、 $\lambda_0$ 、 $I_0$  の順で最適化を行うと、以下のように、

$$I_0 = 0.80, \quad \lambda_0 = 1.25, \quad R_0 = 3.77$$

既存の数値解析結果<sup>3)・6)</sup>とほぼ同じ結果を得た。

図2に、感染者数および感受性者数の推移の数値解析結果を示す。実線と破線はそれぞれ感染者数と感受性者数の数値解析結果を示し、点は実データを示す。残差の二乗平均は、

$$\bar{\delta} = 17.32$$

となり、実データとのずれが1日あたり17名程度となっていることがわかる。

感染者数は6日目付近で最大値およそ291名に到達し、そこから減少していくことがわかり、実際の感染状況と同じ傾向を示している。

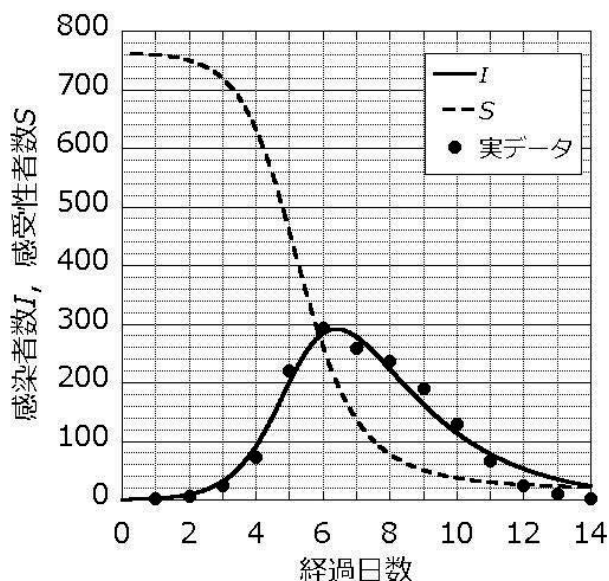


図2 寄宿学校のインフルエンザ流行を対象とした数値解析結果

感受性者は5日目付近に初期値の半分に到達、逆に言えば、全体の約半数が感染したことがわかる。そして、感受性者数が半分を切るとともに、感染者数の指数関数的な増加が終わり、感染者数の増加率が減少し、6日目付近にピークを迎えて感染者数の減少へとつながっていく。前述のとおり、SIRモデルでは、感染者数の増加率は感受性者数に比例するため、感染者数の増加が減少に転じる要因が感受性者数の減少であるという推測と合致している。

また、14日目の感受性者数はおよそ22名であるので、全体の97%にあたる741名が感染していることになる。これは、実際の寄宿学校での流行とは30%もずれている。その要因は恐らく授業を欠席するなどの隔離や感染者の拡大に伴って行われたであろう拡大防止の対策の成果ではないかと考えられる。その効果は、数値解析において感染率 $\beta$ に影響を与えると考えられるので、より正確な数値解析解を得るためには、一定値として扱っている感染率 $\beta$ を動的に変化させる必要があると考えられる。

## 6. インフルエンザの流行予測

感染症流行の数値解析を行う目的の一つは、流行予測である。前述のとおり、2009年の新型インフルエンザや2020年の新型コロナでも数値解析による流行予測が実際に行われている<sup>4), 5)</sup>。そこで、やはり、英国の寄宿学校のインフルエンザ流行のデータを用いて流行予測を試みた。そのために、先ほどと同じく1日目から3日目のデータを指数関数でフィッティングして得られた $\lambda_0$ ,  $I_0$ の推定値 $I_0 = 0.92$ ,  $\lambda_0 = 1.06$ を基に、 $R_0$ ,  $\lambda_0$ ,  $I_0$ の順で最適化を行った。先ほどと異なるのは、残差の二乗平均 $\bar{\delta}$ を求める際に用いる感染者数の実データが1日目から3日目のデータのみであるという点である。

実際にパラメータサーベイを行うと、最初の $R_0$ の段階で、 $R_0$ を大きくすればするほど残差の二乗

平均が小さくなるという結果を得た。この傾向は4日目までのデータを用いた場合や5日目までのデータを用いた場合でも同様に、指数関数的に増加している段階では、 $R_0$ ,  $\lambda_0$ ,  $I_0$ を全て最適化することができないと考えられる。そこで、 $R_0$ は固定し、 $\lambda_0$ ,  $I_0$ のみを最適化してみた。図3に、 $R_0 = 2.00, 3.00, 5.00$ で固定したときの、それぞれの感染者数の時間変化の解析結果を示す。感染の初期段階では、 $R_0$ ,  $\lambda_0$ ,  $I_0$ を全て最適化することはできないが、図3のように、 $R_0$ の値ごとに解析を行うことで、 $R_0$ の値の違いによる流行パターンの変化を俯瞰することができる。

一方、指数関数的な増加が止まった6日目までのデータを用いると、 $R_0$ ,  $\lambda_0$ ,  $I_0$ をそれぞれ、

$$I_0 = 0.80, \lambda_0 = 1.24, R_0 = 4.06$$

と最適化でき、全データを用いた場合とほぼ同じ結果が得られた。図4に、全データを用いた数値解析と6日目までのデータを用いた予測の数値解析の結果を示す。予測の結果の方がやや最大値が大きく、ピークを迎える時期も少し遅れているが、ほぼ同じ結果が得られている。つまり、指数関数的な増加が止まった辺りのデータを用いれば、感染者数の推移をある程度予測することが可能であると考えられる。

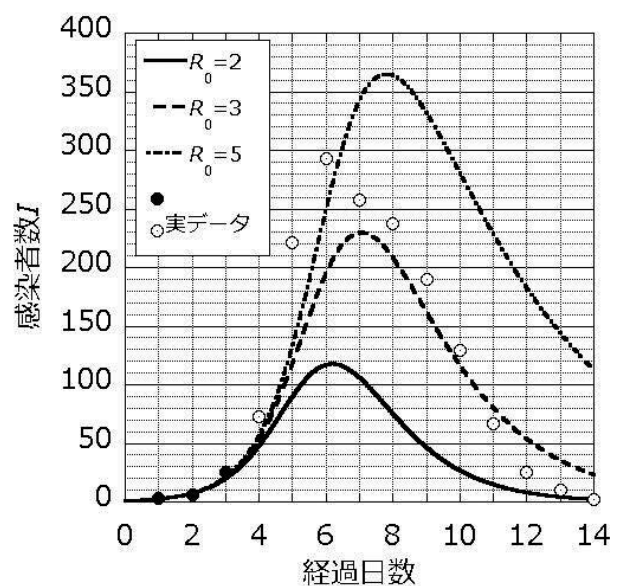


図3 寄宿学校のインフルエンザ流行を対象とした流行予測

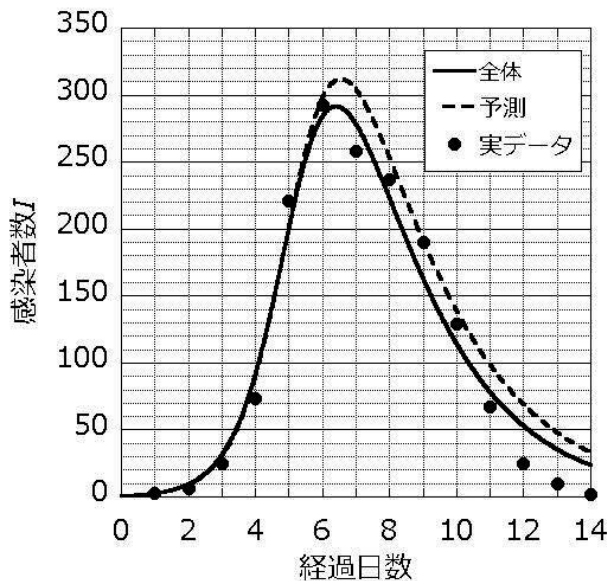


図4 全データを用いた数値解析と6日目までデータを用いた数値解析の比較

## 7. まとめと今後の課題

今回、最も基礎的なSIRモデルを用いて、インフルエンザの流行の数値解析を行った。基礎的なSIRモデルでもインフルエンザの流行についてはその傾向を把握するのに十分な結果が得られることがわかった。また、ある程度のデータがそろえば、数値解析により、全体の傾向が予測できる可能性があることも確認できた。

今回は、小規模で寄宿学校という比較的閉じた空間で起こった流行を対象として解析したが、例えば、最近の国内のインフルエンザの流行のような大規模な流行を対象として解析を行い、このモデルでの解析が可能かを検討したい。また、ワクチンの効果、予防対策の効果、あるいは新型コロナのように潜伏期を考慮するなどモデルの改良も今後の課題である。

## 参考文献

- 1) 栗田順子：新型コロナ感染症第一波のパンデミックシミュレーション，技術評論社 (2020).
- 2) 大日康史，菅原民枝：パンデミックシミュレーション感染症数理モデル，技術評論社

(2009).

- 3) 稲葉 寿：感染症の数理モデル，培風館 (2008).
- 4) 佐藤 總夫：自然の数理と社会の数理Ⅱ，日本評論社 (1987).
- 5) J. D. Murray : *Mathematical Biology Third Edition, I. An Introduction*, Springer (2013)
- 6) *British Medical Journal*, 4 March 1978
- 7) 夏目 雄平，小川 建吾：計算物理Ⅰ，朝倉書店 (2002)

2021年12月25日原稿受付，2021年12月27日採録決定  
Received, December 25th, 2021; accepted, December 27th, 2021