

外出人数と COVID-19 実効再生産数の Scaling 関係による 人-人接触項の検証

島田 尚

¹ 東京大学 大学院工学系研究科 システム創成学専攻

² 東京大学 数理・情報教育研究センター

概要

携帯電話の位置情報から得られた東京主要駅の乗降客数の変化と、東京都における実効再生産数との関係を解析した。まず、「接触機会が外出人数の増加関数である」という自然で簡単な仮定から、感染から感染報告までの平均遅れ時間について約 15 日という現実的な推定値を得た。また、この報告遅れ時間に基づいた接触機会と外出人数の関係性のフィッティングから、接触機会の密度依存性がロトカボルテラ型より緩やかであることが見出された。さらに、この全体的傾向からのずれとして行動変容などの効果を伺うことができること、また、特に目立つ急峻な変化が各種要請の発出や解除のタイミングと一致していることなども見出された。

Estimation of human contact rate from the scaling relation between the effective reproduction number of COVID-19 and the size of active crowd in Tokyo area

Takashi Shimada^{1,2}

¹ Department of Systems Innovation, Graduate School of Engineering, The University of Tokyo

² Mathematics and Informatics Center, The University of Tokyo

Abstract

The relation between the number of passengers in the main stations and the infection rate of COVID19 in Tokyo area is studied. Based on a natural assumption that the human contact rate is an increasing function of active population density, we first determine average time from the day of infection to the report day as 15 days. Furthermore, from the scaling relation between the density of active population and the infection rate, we find the dependency of the human contact rate on the population density is slower than that of Lotka-Volterra type model.

1 はじめに

COVID-19 などの感染症の拡大防止の一般的対策として行われる外出の制限（要請）は、これによる人-人接触機会の低減を目的としている。このような制限はその有効性は疑いようが無い一方で文化・社

会・経済などへの負の影響も大きいため、外出人数のような把握が容易なマクロな量と接触機会との関係の検証が重要である。生態・社会学の理論の側からも、ロトカボルテラ方程式や SIR モデルなど、ポピュレーションの積で単純にモデル化された接触機会の項が人の行動のモデルとしてどれだけ当てはま

るかを理解・検証することは意義深い。

本稿では、携帯電話の位置情報から推定した東京主要駅の乗降客数の変化と同じく東京都における COVID-19 の実効再生産数（感染者の相対人数が非常に少ないことに基づけば、接触機会と比例する）との関係から人-人接触機会の密度依存性を調べた我々の研究 [1] を紹介し、さらにこの解析結果について議論する。

2 解析方法とデータ

2.1 解析の方針

本研究では未感染状態、感染状態、免疫獲得状態の3状態による簡単なコンパートメントモデル：

$$\begin{aligned}\Delta S_t &= S_{t+1} - S_t = -\beta C(S_t, I_t), \\ \Delta I_t &= I_{t+1} - I_t = \beta C(S_t, I_t) - \gamma I_t, \\ \Delta R_t &= R_{t+1} - R_t = \gamma I_t,\end{aligned}\quad (1)$$

に基づいて解析を行う。ここで S_t, I_t, R_t はそれぞれ考えている地域における未感染状態、感染状態、免疫獲得状態の人数の比（もしくは人口密度）である。 β は未感染者と感染者の接触機会当たりの感染率を表す定数であり、一般に病原体や人々の生活様式によって決まる量である。 γ は単位時間当たりの治癒率で、ここでは疫学研究的報告から $\gamma = 14^{-1}$ [day $^{-1}$] 程度にとる。後の解析のため、もっとも一般的な SIR モデルで $\beta S_t I_t$ と書かれる感染プロセスの項の接触機会頻度を表す $S_t I_t$ の部分を

$$C(S_t, I_t) = S_t I_t \cdot A(S_t, I_t, R_t). \quad (2)$$

と一般化して書いてあることに留意いただきたい。ここに $A(S_t, I_t, R_t)$ は接触機会頻度の関数形の $S_t I_t$ からのずれを表す修正項である。

感染拡大の趨勢の評価に最も大事な量である実効再生産数（その時点における、感染者1人が治癒もしくは死亡までに感染させる人数の平均）が各日におけるこの地域での新規感染報告数の時系列 O_t から次のように求められる：

$$\mathcal{R}_t \equiv \frac{-\Delta S_t}{\Delta R_t} = \frac{O_{t+t_d}}{\gamma \left(\sum_{\tau=0}^t O_{\tau+t_d} e^{-\gamma(t-\tau)} \right)} \quad (3)$$

ただしここで、 t_d は真の感染日から感染報告日までの時間遅れの平均であり、一般には疫学的な知見などから決定する必要がある。

さて今、この地域が（我々が経験したように）外出制限などの結果として「感染機会となり得る、住居外での活動をした人数」が大きく変化した場合を考える。この場合、各時点での外出人数の通常時 \bar{N} との比を $n_t = N_t/\bar{N}$ と書けば、感染プロセス $\beta C(S_t, I_t)$ 中の S_t, I_t はそれぞれ $\tilde{S}_t = n_t S_t, \tilde{I}_t = n_t I_t$ で置き換えられなければならない。一方で治癒率の項 γI_t はそのままであるので、ここから新規感染報告数と実効再生産数が

$$O_{t+t_d} = -\Delta S_t = (\beta S_t I_t) n_t^2 A(S_t, I_t, R_t) \quad (4)$$

$$\mathcal{R}_t = \frac{O_{t+t_d}}{\gamma I_t} \approx \left(\frac{\beta}{\gamma} \right) n_t^2 A(S_t, I_t, R_t) \quad (5)$$

と n_t に依存することになる。言い換えれば、外出制限などで外出人数（実効的な活動を行なった人数）が大きく変動した期間の実効再生産数 \mathcal{R}_t の変化の振る舞いが n_t の2乗からずれていれば、それは人-人接触項の SIR モデル型（もしくはロトカ-ボルテラや化学反応型）からの修正を示唆するということになる。

2.2 データ

本研究では、匿名化された携帯電話の位置情報から得られた東京主要駅（東京、新宿、池袋、渋谷）における乗降人数の各日推移のデータを用いる [2]。具体的には、上述の「通常時に対する外出人数の比」 n_t をこれら主要駅の乗降人数の合計の比で置き換える。外出人数に直接に比例すると考えられるこのデータの存在は我々の上述の解析にとって決定的に重要である。また、新規感染者報告数については東京都より公表されているデータを用いた。なお、よく知られているように新規感染者報告数は感染実態とは関連が薄いと思われる曜日依存性が大きいいため、前後各1週間の2週間移動平均をとったものを O_t として扱った。また、これに対応させて乗降人数の時系列についても2週間の移動平均をとった。

3 結果

3.1 報告遅れ時間 t_d の推定

コンパートメントモデルを基盤とした解析の利点として、通常は疫学的な調査から別途定める必要のある「感染日から報告日までのずれ」 t_d を推定することができる点も重要である。上述したように我々は人と人との接触機会について SIR モデルなどとはずれた関数形も視野に入れて解析するわけである

が、その上で「接触機会が外出人数の増加関数である」という仮定をおく。この仮定に基づくと、適当に選んだ t_d について計算した n_t と R_t の二次元プロットが最も良く一価の関数で表されるような t_d を選べば良いことになる。実際、図 1 に示すように、 $n_t - R_t$ プロットは遅れ時間 t_d を (非現実的な) 0 日とした場合には大きくループを描いて一価の関係にならない一方で他の研究報告の値に近い 2 週間程度とした場合はほぼ一価の関数関係にあることが分かる。「一価の関数関係の良さ」に基づく最適な t_d は (もちろん目視によってでなく) n_t と R_t の順位相関を取ることによって求められ、その値は $t_d = 15$ 日である。こうして決める最適値が、他の知見から天下り的に与えたパラメータである γ の値にあまり依存性しないことも確かめられている。

この解析に用いた仮定は、観測期間においてある一定の数理モデルが適用可能であるとすれば非常に自然でロバスタな仮定と言える。一方で、観測期間内の人々の行動変容やウイルスの strain の変化の効果、またさらにはそもそもこのような巨視的量に基づいたモデリングへの疑義まで含めれば、特に一価の関数関係が成り立つことは全く自明では無い。この点からは、今回解析対象とした 2020 年の 2 月から 9 月という緊急事態宣言や就労・通学の様式の激変があった長い期間に渡り n_t と R_t の関係性が関数関係でほぼ一定に保たれていたという結果はむしろ驚きに値すると言える。また、緊急事態宣言による外出の最大限の抑制期間の前後で R_t が同じ曲線に乗って上昇したことは行動変容の効果が小さかったこと (もしくはこの期間の前から人々の感染防止行動がかなり最適に近かったこと) をも示している。

3.2 人-人接触項の評価

報告遅れ時間 t_d と、実効的な人口密度 n_t と R_t (人-人接触項 C に比例) との関数関係の良さが見出されたので、次にこの関数関係についてフィッティングを行う。フィッティングには、SIR モデルが良く成り立つ場合に対応する

$$R_t = R_0 n_t^2 \quad (6)$$

と、同じく SIR モデルの接触項の仮定を保持しつつ、感染プロセスに関わるポピュレーションには主要駅での乗降客数からは捕捉されない定数項 n_0 が

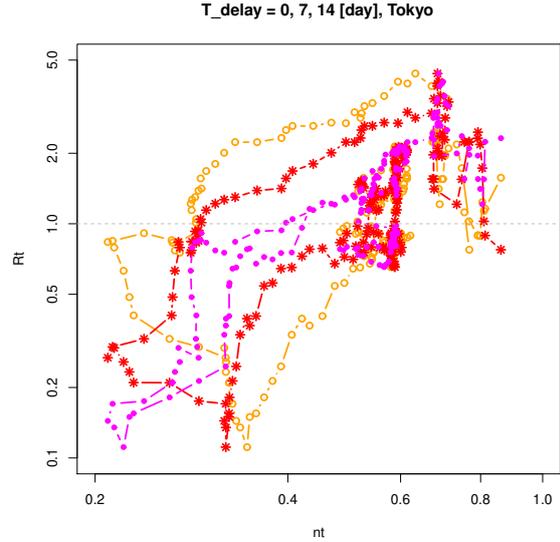


図 1: 感染日から感染報告日までの平均遅れ時間 t_d を 0 日、一週間、二週間とそれぞれとった場合の n_t と R_t の関係。順位相関から求まる最適点に近い $t_d = 2$ 週間の場合には n_t と R_t はほぼ一価の関係となる。

あるとした関数

$$R_t' = R_0 \left(\frac{n_0 + n_t}{n_0 + 1} \right)^2 \quad (7)$$

また最後に人と人との接触項の密度依存性が SIR モデルと異なるとした関数を用いた:

$$R_t^{pow} = R_0 n_t^{2-\alpha}. \quad (8)$$

この結果、最後の関数 R_t^{pow} が最も良いフィットを与えることが分かった。ここで R_t^{pow} の形状は人の接触項について SIR モデルから以下のような修正がされたことに対応する:

$$C(S_t, I_t) = S_t I_t A(S_t, I_t, R_t) = \frac{C_0 S_t I_t}{n_t^\alpha} \quad (9)$$

ので、 α はいわば「人の活動の主体性を示すパラメータ」と言えることに注意したい。すなわち、 $\alpha = 0$ は例えば街行く人々同士が近接する確率はその濃度の二乗に比例するのであるからあたかも原子分子のように「意志無く」振る舞う場合に相当する。一方 $\alpha > 0$ は分子運動に比べて低密度では衝突頻度が下がり、また高密度では衝突頻度の上がり方が緩やかな場合に対応し、これは人々は外出するそもその目的が人との接触機会 (仕事、買い物、面会) と関連しており、また移動経路上など可能な場面では人

と距離を取ろうとするものと考えたと自然な補正方向である。反対に $\alpha < 0$ は衝突項のより急峻な密度依存性に対応するが、これは例えばいつもは距離を保とうとするものの、駅や店内、電車内などで高密度の場所が発生した場合に受動的に近接してしまうという状況に対応する。接触頻度が密度の増加関数でなくなってしまうというのは非現実的であると考え、この「人の行動の主体性」パラメータの自然な領域は $\alpha < 2$ であると言える。

実データのベストフィッティングから得られる指数は $\alpha \sim 1$ であり、これは人々の行動の主体的な面がとらえられたと考えることができる。また、 $\alpha \sim 1$ すなわち接触項が密度の1乗に近いという結果は式(1)のコンパートメントモデルに戻って考えると別の意味で示唆深い。今回の解析ではある地域における感染プロセスに関わる n_t が時間的に大きく変化する場合を扱ったが、今度は異なる色々の地域の平均の人口密度比を n_t によりモデリングしていると考えてみる。この場合接触(感染)項に加えて治癒項 γI_t も n_t 倍に修正されるので、コンパートメントモデルの差分を表す各項が n_t 倍されるだけ：

$$\begin{aligned}\Delta S_t &= -n_t \cdot \beta S_t, I_t, \\ \Delta I_t &= n_t \cdot (\beta S_t, I_t - \gamma I_t), \\ \Delta R_t &= n_t \cdot \gamma I_t,\end{aligned}\quad (10)$$

となり、ここから計算される実効再生産数 R_t が人口密度に依らないということになる。この面からは、今回のフィッティングで得られた修正接触項は、平均の人口密度の大きく異なる世界の各地域で一様に感染拡大が起こっている現象の良い巨視モデルとなっていると言える。

4 まとめと展望：Scaling 関係を越えて

以上、東京主要駅の乗降客数の変化と、東京都における実効再生産数との関係についてのマクロモデルを基盤とした解析から、真の感染から感染報告までの平均遅れ時間という調査の難しい量についての良い推定を得たほか、人と人との接触機会の密度依存性を評価することができた。

また最後に、これらの期間全体についての解析を基にしつつ、全体の傾向からのずれから窺えることについても定性的に議論したい。図2に示す通り、期間全体に渡り良い相関を示す n_t と R_t の時系列

をより細かくみると、 n_t がさほど変化していないが R_t が大きく変化している時点があることが分かる。興味深いことにこれらの変化は、緊急事態宣言の全国への適用時 (R_t 減少)、緊急事態宣言の一部解除 (R_t 上昇)、東京アラートの宣言 (R_t 減少) など今回対象の東京圏には直接の制限などが及ばない自粛要請や解除のタイミングと良く一致している。これらの、今回のモデルに沿えば β の変化に対応すると思われるイベントについての解析には他のデータとの照合が必要と思われる。

さらに、今回得られた接触機会の人密度依存性が地域や社会をまたいでどれほど普遍的であるかについても検証が必要である。その試みの一例として関西圏の動向についての黒田らによる解析を参照されたい(MSTF2020 発表予定)。本研究は JSPS KAKENHI 18K03449 の支援を受けたものである。

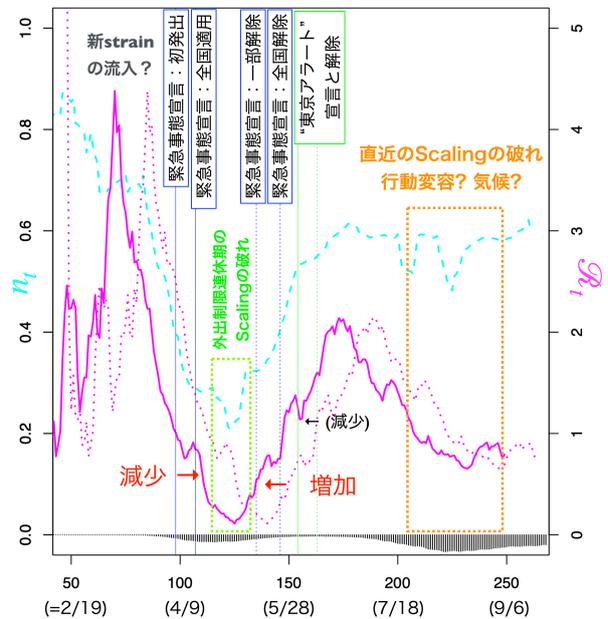


図2: 本解析から得られた実効再生産数(マゼンタ実線、点線は t_d の補正無しの場合)と主要駅での乗降客数比 n_t (シアン破線)の時系列の比較。スケーリング関係で確認した通りこれらの曲線は全体に良い相関を示すが、緊急事態宣言などの時期に対応したタイミングで乗降客数の変化に依らない R_t の急峻な変化が見られる。

参考文献

- [1] T. Shimada, Y. Suimon, and K. Izumi, <https://arxiv.org/abs/2008.07791>.
- [2] <https://k-locationanalyzer.com>.