

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC - KỸ THUẬT MÁY TÍNH



MÔ HÌNH HOÁ TOÁN HỌC

Nhóm: HNPUS — Bài tập lớn

Mô hình SIR trong dự báo COVID-19

GVHD: Nguyễn An Khương
SV thực hiện: Tô Duy Hưng – 1810198
Võ Hoàng Hải Nam – 1810340
Võ Thanh Phong – 1712633
Huỳnh Thị Uyên – 1810648
Lê Thành Sơn – 1810481

Tp. Hồ Chí Minh, Tháng 6/2020



Bảng phân công công việc

| Họ và tên | Nhiệm vụ |
|------------------|---|
| Lê Thành Sơn | Câu 1. Kiến thức và kết quả chuẩn bị |
| Tô Duy Hưng | Câu 2. Xây dựng chương trình sử dụng thuật toán Euler tìm nghiệm hệ SIR |
| Võ Hoàng Hải Nam | Câu 2. Tìm hiểu thuật toán Runge-Kutta và so sánh với thuật toán Euler |
| Huỳnh Thị Uyên | Câu 3. Lấy mẫu cho hệ số β và γ sử dụng thuật toán Metropolis-Hastings |
| Võ Thanh Phong | Câu 4. Ước lượng giá trị R_0 ở các khu vực |

Nhật ký làm việc nhóm

| Thời gian | Nội dung |
|------------|--|
| 02/07/2020 | Lập nhóm và gặp mặt phân tích bài tập lớn |
| 06/07/2020 | Hoàn thiện sơ bộ phần lý thuyết cho câu 1 |
| 08/07/2020 | Hoàn thiện chương trình sử dụng thuật toán Euler |
| 09/07/2020 | Hoàn thiện chương trình sử dụng thuật toán Runge-Kutta |
| 12/07/2020 | Họp nhóm lần thứ 2 để bàn phương pháp giải quyết các câu còn lại |
| 16/07/2020 | Hoàn thiện chương trình lấy mẫu sử dụng thuật toán Metropolis-Hastings |
| 18/07/2020 | Hoàn thiện chương trình ước lượng R_0 và phân tích các khu vực |
| 19/07/2020 | Họp nhóm để đánh giá và chỉnh sửa báo cáo. Bàn luận về việc sử dụng Neuron Network trong việc dự đoán các hệ số $[\beta, \gamma, \mu]$ |

Nhật ký chỉnh sửa báo cáo

| Thời gian | Nội dung | Người chỉnh sửa |
|------------|-------------------------------------|------------------|
| 06/07/2020 | Thêm phần 1 | Lê Thành Sơn |
| 08/07/2020 | Thêm phần 2.1 | Tô Duy Hưng |
| 09/07/2020 | Chỉnh sửa phần 1 | Lê Thành Sơn |
| 10/07/2020 | Thêm phần 2.2 | Võ Hoàng Hải Nam |
| 17/07/2020 | Thêm phần 3 | Huỳnh Thị Uyên |
| 18/07/2020 | Chỉnh sửa phần 3 | Huỳnh Thị Uyên |
| 19/07/2020 | Thêm phần 4 | Võ Thanh Phong |
| 20/07/2020 | Chỉnh sửa nội dung các phần 2 | Lê Thành Sơn |
| 21/07/2020 | Chỉnh sửa định dạng và lỗi chính tả | Lê Thành Sơn |
| 27/07/2020 | Thêm phần 1.2.1 và 1.2.2 | Tô Duy Hưng |

Mục lục

| | | |
|----------|---|-----------|
| 1 | Kiến thức và kết quả chuẩn bị | 4 |
| 1.1 | Lý thuyết về mô hình SIR | 4 |
| 1.1.1 | Thiết lập mô hình SIR | 4 |
| 1.1.2 | Mô hình SIR với biến rời rạc | 4 |
| 1.1.3 | Mô hình SIR với biến liên tục | 6 |
| 1.1.3.a | Hệ phương trình vi phân | 6 |
| 1.2 | Một số vấn đề liên quan đến mô hình SIR | 8 |
| 1.2.1 | Trạng thái cân bằng | 8 |
| 1.2.2 | Hệ động lực tuyến tính | 9 |
| 1.2.3 | Tổng quát hơn nữa hệ liên tục đã xây dựng | 10 |
| 1.2.4 | Hệ số lây nhiễm cơ bản R_0 | 12 |
| 1.2.5 | Mô hình SIRD | 12 |
| 1.2.6 | Một số mô hình khác | 13 |
| 2 | Xây dựng chương trình sử dụng thuật toán Euler tìm nghiệm của hệ SIR | 14 |
| 2.1 | Phương pháp Euler | 14 |
| 2.1.1 | Tổng quan lý thuyết | 14 |
| 2.1.2 | Thực hiện | 14 |
| 2.2 | Phương pháp Runge-Kutta | 17 |
| 3 | Lấy mẫu cho hệ số β và γ | 21 |
| 3.1 | Tổng quan lý thuyết | 21 |
| 3.1.1 | Phân phối chuẩn | 21 |
| 3.1.2 | Thuật toán Metropolis-Hastings | 21 |
| 3.2 | Hiện thực | 22 |
| 4 | Ước lượng giá trị R_0 ở các khu vực | 26 |
| 4.1 | Tổng quan lý thuyết | 26 |
| 4.2 | Hiện thực | 26 |
| 4.3 | Phân tích | 27 |
| 4.3.1 | Italy | 27 |
| 4.3.2 | Việt Nam | 27 |
| 5 | Kết luận | 29 |
| | Tài liệu | 30 |



Danh sách hình vẽ

| | | |
|----|---|----|
| 1 | Kết quả tính toán và phác hoạ đồ thị | 15 |
| 2 | Kết quả tính toán và phác hoạ đồ thị | 16 |
| 3 | Kết quả tính toán và phác hoạ đồ thị | 16 |
| 4 | Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp | 18 |
| 5 | Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp | 19 |
| 6 | Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp | 20 |
| 7 | Các phân bố mẫu của β và γ cùng với 50000 lần lặp | 23 |
| 8 | Phân phối chung cho (β, γ) với 50000 lần lặp | 23 |
| 9 | Các phân bố mẫu của β và γ cùng với 100000 lần lặp | 24 |
| 10 | Phân phối chung cho (β, γ) với 50000 lần lặp | 24 |
| 11 | Biểu đồ thay đổi giá trị R_0 của Italy | 27 |
| 12 | Biểu đồ thay đổi giá trị R_0 của Việt Nam | 28 |
| 13 | Các giai đoạn và các mốc thời gian cụ thể về Covid-19 ở Việt Nam | 28 |

LỜI NÓI ĐẦU

Bài báo cáo này trình bày về mô hình SIR trong dự báo cho dịch bệnh Covid-19 từ các bước để xây dựng mô hình đó, từ các lý thuyết đến xây dựng mô hình bằng các cách khác nhau bằng ngôn ngữ lập trình Python.

1 Kiến thức và kết quả chuẩn bị

1.1 Lý thuyết về mô hình SIR

Trong phần này chúng tôi giới thiệu các lý thuyết về mô hình SIR, cách xây dựng mô hình SIR cho trường hợp biến rời rạc và liên tục cùng một số vấn đề liên quan đến mô hình này.

1.1.1 Thiết lập mô hình SIR

Xét một dịch bệnh đang xảy ra tại một quốc gia. Chúng ta sẽ chia tập dân cư của nước đó ra thành 3 loại: có khả năng nhiễm bệnh (S), bị nhiễm bệnh (I) và đã khỏi bệnh (R). Bên cạnh đó, ta sẽ chỉ xét trong trường hợp cực kỳ tối ưu:

1. Không một ai trong nước đó đi ra nước khác cũng như có người từ nước khác nhập cảnh vào nước đó; và dịch bệnh chỉ xảy ra trong cộng đồng nước đó, không xảy ra bên ngoài nước khác. Hay nói cách khác, cộng đồng đang cách ly, tổng số dân trong cộng đồng là cố định không đổi.
2. Mỗi người chỉ có thể thuộc một trong 3 nhóm: có khả năng nhiễm (S); đang bị nhiễm (I) và có khả năng lây bệnh cho người khác; và đã khỏi (R) gồm những người đã khỏi bệnh khoẻ mạnh và những người chết vì dịch, những người đã khỏi bệnh sẽ không mắc bệnh lại.
3. Ban đầu mỗi người có thể bị bệnh hoặc đã khỏi bệnh.
4. Một người đã mắc bệnh sẽ không thể mắc bệnh này (nghĩa là người khỏi bệnh sẽ hình thành kháng thể đặc hiệu chống lại bệnh này hoàn toàn).

Ngoài ra, chúng ta sẽ đặt thêm một số điều kiện:

1. Thời gian trung bình của dịch bệnh là η tuần, trong thời gian này, một người đã nhiễm bệnh có thể lây nhiễm cho người khác, sau thời gian này sẽ được tính là đã khỏi bệnh.
2. Chu kỳ mỗi lần tính là tuần.

Khi đó mô hình này được gọi là mô hình SIR.

1.1.2 Mô hình SIR với biến rời rạc

Ta định nghĩa một số hàm như sau:

- $S(n)$: số người trong cộng đồng nghi nhiễm sau chu kỳ thứ n .
- $I(n)$: số người bị nhiễm sau chu kỳ thứ n .
- $R(n)$: số người đã khỏi bệnh sau chu kỳ thứ n .

Biến rời rạc chính là thời gian n , n là các số tự nhiên.

Ta sẽ bắt đầu xây dựng mô hình này. Lưu ý rằng, theo thực tế và các điều kiện của mô hình, một người nghi nhiễm có thể bị nhiễm bệnh rồi chuyển sang khỏi bệnh, không thể chuyển từ nghi nhiễm sang khỏi bệnh hay từ bị nhiễm sang nghi nhiễm hoặc ngược lại.

Đầu tiên sẽ bắt đầu với hàm $R(n)$. Ta đã giả sử một người bị bệnh sẽ bị trong η tuần, và sẽ hết bệnh sau đó. Cho nên mỗi tuần, sẽ có $1/\eta$ người đang nhiễm khỏi bệnh (theo quy tắc tam suất). Xét ở tuần thứ $n + 1$, số người đã khỏi bệnh ghi nhận ở tuần thứ n là $R(n)$ và sẽ có thêm $\frac{I(n)}{\eta}$ người khỏi bệnh mới như lập luận trên. Do đó, số người khỏi bệnh được ghi nhận ở tuần thứ $n + 1$ được thể hiện bởi công thức:

$$R(n + 1) = R(n) + \frac{I(n)}{\eta}$$

Ta sẽ tiếp tục xây dựng công thức hồi quy cho $I(n)$. Theo định nghĩa, ta có nhận xét: $I(n)$ sẽ giảm mỗi tuần so với trước đó, và số lượng giảm sẽ chính là số người đã khỏi bệnh trong tuần đó, nếu xét ở tuần thứ n chính là $\frac{I(n)}{\eta}$ như trên.

Ngoài ra, $I(n)$ còn tăng mỗi tuần so với trước đó, số lượng tăng bằng số người có khả năng nhiễm bệnh trong cộng đồng bị nhiễm mới. Như đã đặt ra ở giả thiết, chỉ người đang trong thời kỳ nhiễm bệnh mới có thể lây nhiễm cho người có khả năng nhiễm. Ngoài ra, một thực tế là không phải bất cứ ai đang nghi nhiễm cũng bị nhiễm ngay lập tức nếu tiếp xúc với người nhiễm. Do đó, ta định nghĩa thêm một hệ số $\alpha \leq 1$ để thể hiện xác suất một người nghi nhiễm có thể bị nhiễm bệnh khi tiếp xúc người bệnh. α thể hiện khả năng lây nhiễm của dịch bệnh, hay còn gọi là hệ số truyền nhiễm và α chính là một hệ số theo xác suất. Lại đặt một giả sử trong trường hợp tối ưu: α là không đổi theo thời gian và nó đã được xác định ngay từ ban đầu khi xảy ra dịch bệnh. Từ các định nghĩa trên, số người bị nhiễm mới trong tuần thứ n sẽ là $\alpha S(n)I(n)$, trong đó $S(n)I(n)$ thể hiện việc người có khả năng nhiễm tiếp xúc với người đang nhiễm bệnh.

Do đó, sau tuần thứ $n + 1$, số người bị nhiễm mới sẽ được thể hiện bởi công thức:

$$I(n + 1) = I(n) - \frac{I(n)}{\eta} + \alpha S(n)I(n)$$

Ta sẽ xây dựng hàm $S(n)$. Rõ ràng, số người nghi nhiễm mỗi tuần sẽ luôn luôn giảm, bởi vì cộng đồng này đang cách ly nên số người trong cộng đồng không thể thay đổi. Số người nghi nhiễm giảm mỗi tuần bằng số lượng người nhiễm mới mỗi tuần. Nếu xét tại cuối tuần $n + 1$, kết hợp với các lập luận trên, số người nghi nhiễm tại tuần đó được thể hiện bởi công thức:

$$S(n + 1) = S(n) - \alpha S(n)I(n)$$

Theo mô hình, ban đầu chỉ có người bệnh và người nghi nhiễm, không có người hồi phục, do đó, ta có các giá trị khởi đầu như sau: $S(0) = a, I(0) = b, R(0) = 0$. Ngoài ra, do số người trong cộng đồng là cố định và không đổi nên trong mọi tuần thứ n , nếu gọi N là tổng số dân trong cộng đồng đó, ta luôn có

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

Kết hợp với các công thức thiết lập được như trên, ta có được hệ như sau:

$$\begin{aligned}S(n+1) &= S(n) - \alpha S(n)I(n) \\I(n+1) &= I(n) - \frac{I(n)}{\eta} + \alpha S(n)I(n) \\R(n+1) &= R(n) + \frac{I(n)}{\eta} \\S(n) + I(n) + R(n) &= N \forall n \in N \\S(0) &= a, I(0) = b, R(0) = 0 \\n, a, b, N &\in N \\\alpha, \eta &> 0; \alpha \leq 1\end{aligned}\tag{1}$$

Hệ trên có thể giải được và biểu diễn thành đồ thị. Với các giá trị cụ thể, ta có thể tính toán các giá trị S, I, R sau các tuần.

1.1.3 Mô hình SIR với biến liên tục

Trên thực tế, số lượng người nghi nhiễm (S), bị nhiễm bệnh (I) và khỏi bệnh (R) thay đổi theo thời gian thực. Do đó, ta sẽ cố gắng thể hiện các số liệu thông qua các hàm với biến liên tục.

Vấn xét mô hình SIR với các điều kiện đã được thiết lập như trong phần 1.1.1. Ta sẽ định nghĩa một số hàm như sau:

- $S(t)$: số người trong cộng đồng nghi nhiễm sau thời gian t .
- $I(t)$: số người bị nhiễm sau thời gian t .
- $R(t)$: số người đã khỏi bệnh sau thời gian t .

Ngoài ra, ta có thể gọi N chính là tổng số dân trong cộng đồng. Theo các giả thiết của mô hình đã đưa ra, N là số tự nhiên và là hằng số.

Biến liên tục ở đây chính là thời gian t , trong đó t dương. Ta sẽ xây dựng mô hình tính toán này.

1.1.3.a Hệ phương trình vi phân

Đầu tiên sẽ bắt đầu với hàm $R(t)$. Ta đã giả sử một người bị bệnh sẽ bị trong η tuần, và sẽ hết bệnh sau đó. Cho nên mỗi tuần, sẽ có $1/\eta$ người đang nhiễm khỏi bệnh (theo quy tắc tam suất). Để đánh giá được độ thay đổi của hàm số $R(t)$ theo thời gian t , ta sử dụng khái niệm đạo hàm. Xét số người khỏi bệnh tại 2 thời điểm t và $t + \delta t$. Theo quy tắc tam suất, kết hợp với lập luận như trên, số người khỏi bệnh tại thời điểm $t + \delta t$ sẽ tăng thêm $\delta t \cdot \frac{I(t)}{\eta}$ so với thời điểm t trước đó. Cho nên $R(t + \delta t) = R(t) + \delta t \cdot \frac{I(t)}{\eta}$. Khi đó, ta sẽ có:

$$\frac{R(t + \delta t) - R(t)}{\delta t} = \frac{R(t) + \delta t \cdot \frac{I(t)}{\eta} - R(t)}{\delta t} = \frac{I(t)}{\eta}$$

Cho nên:

$$\lim_{\delta t \rightarrow +\infty} \frac{R(t + \delta t) - R(t)}{\delta t} = \frac{I(t)}{\eta}$$

Hay:

$$\frac{dR}{dt} = \frac{I(t)}{\eta}$$

Giá trị $\frac{1}{\eta}$ được định nghĩa là tỉ lệ khỏi bệnh mỗi tuần. Giá trị này thể hiện tỉ lệ người khỏi bệnh mỗi tuần và có thể tính toán từ các dữ liệu thực.

Ta sẽ tiếp tục xây dựng công thức cho $I(t)$. Theo định nghĩa, ta có nhận xét: $I(t)$ sẽ giảm mỗi tuần so với trước đó, và số lượng giảm sẽ chính là số người đã khỏi bệnh trong tuần đó, nếu xét ở thời gian t chính là $\frac{I(t)}{\eta}$ như trên. Ngoài ra, $I(t)$ còn tăng so với trước đó, số lượng tăng bằng số người có khả năng nhiễm bệnh trong cộng đồng bị nhiễm mới. Như đã đặt ra ở giả thiết, chỉ người đang trong thời kỳ nhiễm bệnh mới có thể lây nhiễm cho người có khả năng nhiễm. Ngoài ra, một thực tế là không phải bất cứ ai đang nghi nhiễm cũng bị nhiễm ngay lập tức nếu tiếp xúc với người nhiễm. Do đó, ta định nghĩa thêm một hệ số $\alpha \leq 1$ để thể hiện xác suất một người nghi nhiễm có thể bị nhiễm bệnh khi tiếp xúc người bệnh. α thể hiện khả năng lây nhiễm của dịch bệnh, hay còn gọi là hệ số truyền nhiễm và α chính là một hệ số theo xác suất. Lại đặc một giả sử trong trường hợp tối ưu: α là không đổi theo thời gian và nó đã được xác định ngay từ ban đầu khi xảy ra dịch bệnh. Từ các định nghĩa trên, số người bị nhiễm mới sau thời gian t sẽ là $\alpha S(t)I(t)$, trong đó $S(t)I(t)$ thể hiện việc người có khả năng nhiễm tiếp xúc với người đang nhiễm bệnh. Xét số người khỏi bệnh tại 2 thời điểm t và $t + \delta t$. Từ các lập luận trên, theo quy tắc tam suất, ta có được:

$$I(t + \delta t) = I(t) + \left(-\frac{I(t)}{\eta} + \alpha S(t)I(t)\right) \cdot \delta t$$

Cho nên:

$$\frac{I(t + \delta t) - I(t)}{\delta t} = \frac{I(t) + \left(-\frac{I(t)}{\eta} + \alpha S(t)I(t)\right) \cdot \delta t - I(t)}{\delta t} = -\frac{I(t)}{\eta} + \alpha S(t)I(t)$$

Từ đó ta có được:

$$\lim_{\delta t \rightarrow +\infty} \frac{I(t + \delta t) - I(t)}{\delta t} = -\frac{I(t)}{\eta} + \alpha S(t)I(t)$$

Hay:

$$\frac{dI}{dt} = -\frac{I(t)}{\eta} + \alpha S(t)I(t)$$

Cuối cùng ta thực hiện xây dựng công thức cho hàm $S(t)$. Rõ ràng, số người nghi nhiễm mỗi tuần sẽ luôn luôn giảm, bởi vì cộng đồng này đang cách ly nên số người trong cộng đồng không thể thay đổi. Số người nghi nhiễm giảm mỗi tuần bằng số lượng người nhiễm mới mỗi tuần. Nếu xét tại 2 thời điểm t và $t + \delta t$, kết hợp với các lập luận trên, số người nghi nhiễm tại thời điểm đó được thể hiện bởi công thức:

$$S(t + \delta t) = S(t) - \alpha S(t)I(t) \cdot \delta t$$

Từ đó ta có được:

$$\frac{S(t + \delta t) - S(t)}{\delta t} = \frac{S(t) - \alpha S(t)I(t) \cdot \delta t - S(t)}{\delta t} = -\alpha S(t)I(t)$$

Cho nên:

$$\lim_{\delta t \rightarrow +\infty} \frac{S(t + \delta t) - S(t)}{\delta t} = -\alpha S(t)I(t)$$

Hay:

$$\frac{dS}{dt} = -\alpha S(t)I(t)$$

Theo mô hình, ban đầu chỉ có người bệnh và người nghi nhiễm, không có người hồi phục, do đó, ta có các giá trị khởi đầu như sau: $S(0) = a, I(0) = b, R(0) = 0$. Ngoài ra, do cộng đồng đang cách ly nên dân số không đổi tại mọi thời điểm t , do đó, ta có:

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

Kết hợp với các công thức thiết lập được như trên, với một N đủ lớn, ta có được hệ như sau:

$$\begin{aligned}\frac{dR}{dt} &= \frac{I(t)}{\eta} \\ \frac{dI}{dt} &= -\frac{I(t)}{\eta} + \alpha S(t)I(t) \\ \frac{dS}{dt} &= -\alpha S(t)I(t) \\ S(t) + I(t) + R(t) &= N \forall t \in \mathbb{R}^+ \\ S(0) &= a, I(0) = b, R(0) = 0 \\ a, b &\in \mathbb{N} \\ \alpha, \eta &> 0; \alpha \leq 1\end{aligned}\tag{2}$$

Hệ trên là một hệ phương trình vi phân có thể giải được và biểu diễn bằng đồ thị.

1.2 Một số vấn đề liên quan đến mô hình SIR

1.2.1 Trạng thái cân bằng

Xét mô hình SIR liên tục. Trạng thái (S, I, R) được gọi là trạng thái cân bằng của hệ SIR nếu

$$\frac{dS(t)}{dt} = \frac{dI(t)}{dt} = \frac{dR(t)}{dt} = 0, \forall t \geq 0\tag{3}$$

Từ hệ (2) và (3), ta có được:

$$-\alpha S(t)I(t) = -\frac{I(t)}{\eta} + \alpha S(t)I(t) = \frac{I(t)}{\eta} = 0, \forall t \geq 0\tag{4}$$

Biểu thức trên có nghiệm khi $\frac{I(t)}{\eta} = 0, \forall t \geq 0$, hay $I(t) = 0, \forall t \geq 0$. Kiểm tra lại thấy kết quả này thỏa mãn các phương trình còn lại và là nghiệm duy nhất thỏa mãn hệ (4).

Vậy, mô hình SIR đạt trạng thái cân bằng khi và chỉ khi không có người nào mắc bệnh.

1.2.2 Hệ động lực tuyến tính

- Cơ sở lý thuyết: Cho hai hàm số thực f và g phụ thuộc vào hai biến $x, y \in \mathbb{R}$. Gọi (x_*, y_*) là nghiệm của các phương trình $f(x, y) = g(x, y) = 0$. Sự **tuyến tính hóa hệ động lực**

$$\begin{aligned}\frac{dx}{dt} &= f(x, y) \\ \frac{dy}{dt} &= g(x, y)\end{aligned}$$

xung quanh điểm (x_*, y_*) cho bởi hệ động lực tuyến tính

$$\begin{aligned}\frac{dx}{dt} &= f_x(x_*, y_*)x + f_y(x_*, y_*)y \\ \frac{dy}{dt} &= g_x(x_*, y_*)x + g_y(x_*, y_*)y\end{aligned}$$

trong đó f_x là đạo hàm riêng của hàm f theo biến x .

- Vận dụng: Xét hệ tuyến tính hóa cho hệ SIR liên tục (chỉ xét hai phương trình cho S và I)

$$\begin{aligned}\frac{dS}{dt} &= f(S, I) = -\alpha S(t)I(t) \\ \frac{dI}{dt} &= g(S, I) = -\frac{I(t)}{\eta} + \alpha S(t)I(t)\end{aligned}$$

xung quanh điểm $(S(0), I(0)) = (N, 0)$ với N là tổng số dân trong cộng đồng và N là hằng số. Khi đó, ta có được hệ động lực tuyến tính cho hai hàm S và I lần lượt là

$$\begin{aligned}\frac{dS}{dt} &= f_S(S(0), I(0))S(t) + f_I(S(0), I(0))I(t) \\ \frac{dI}{dt} &= g_S(S(0), I(0))S(t) + g_I(S(0), I(0))I(t)\end{aligned}$$

Cho nên:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha I(0)S(t) - \alpha S(0)I(t) \\ \frac{dI}{dt} &= \alpha I(0)S(t) + \left(-\frac{1}{\eta} + \alpha S(0)\right)I(t)\end{aligned}$$

Hay:

$$\frac{dS}{dt} = -\alpha N I(t) \quad (5)$$

$$\frac{dI}{dt} = \left(-\frac{1}{\eta} + \alpha N\right)I(t) \quad (6)$$

Từ đó, với những giả thiết giúp xây dựng được hệ động lực tuyến tính như trên, ta có thể tính toán được nghiệm chính xác của S và I theo biến thời gian $t \geq 0$. Đầu tiên, từ phương trình vi phân (6), ta suy ra được:

$$\frac{dI}{I(t)} = \left(-\frac{1}{\eta} + \alpha N\right)dt$$

Lấy nguyên hàm 2 vế, ta được:

$$\ln(I(t)) = \left(-\frac{1}{\eta} + \alpha N\right)t + C$$

Hay:

$$I(t) = Ce^{\left(-\frac{1}{\eta} + \alpha N\right)t}$$

Với $t = 0$, ta có $I(t) = C = I(0)$. Khi đó, nghiệm chính xác của I là $I(t) = I(0)e^{\left(-\frac{1}{\eta} + \alpha N\right)t}$. Kết quả này giúp ta đánh giá được tốc độ nhiễm bệnh sẽ tăng nhanh theo tốc độ hàm mũ khi $-\frac{1}{\eta} + \alpha N > 0$, tương đương khi hệ số $R_0 = \frac{\alpha N}{\frac{1}{\eta}} > 1$

Sau đó thay nghiệm I vừa tìm được vào phương trình vi phân (5), ta suy ra được:

$$dS = -\alpha NI(0)e^{\left(-\frac{1}{\eta} + \alpha N\right)t} dt$$

Lấy nguyên hàm 2 vế, ta được:

$$S(t) = \frac{-\alpha NI(0)e^{\left(-\frac{1}{\eta} + \alpha N\right)t}}{-\frac{1}{\eta} + \alpha N} + C$$

Với $t = 0$, ta có $S(t) = \frac{-\alpha NI(0)}{-\frac{1}{\eta} + \alpha N} + C = S(0)$. Khi đó, nghiệm chính xác của S là

$$S(t) = \frac{-\alpha NI(0)e^{\left(-\frac{1}{\eta} + \alpha N\right)t}}{-\frac{1}{\eta} + \alpha N} + S(0) + \frac{\alpha NI(0)}{-\frac{1}{\eta} + \alpha N}$$

1.2.3 Tổng quát hơn nữa hệ liên tục đã xây dựng

Trong phần 1.1.2 và 1.1.3, ta đã xây dựng mô hình SIR với biến t liên tục và biểu diễn dưới dạng các công thức. Ta sẽ tiếp tục sử dụng các ký tự với ý nghĩa như đã trình bày trong phần trước đó. Trong các hệ (1) và (2), ta có giả sử rằng các hằng số η định nghĩa thời gian do được kể từ lúc một người mắc bệnh cho đến khi được cho là "khỏi bệnh" và từ đó $\gamma = \frac{1}{\eta}$ định nghĩa tỷ lệ người khỏi bệnh mỗi tuần; α thể hiện xác suất một người có thể bị nhiễm khi tiếp xúc với một người bệnh.

Với định nghĩa của α có thể hiểu là xác suất trung bình một người khoẻ mạnh bị nhiễm bệnh, và do đó, nếu gọi β là ước lượng số tiếp xúc của người trong nhóm S với người trong nhóm I thì α có thể được tính bởi $\alpha = \frac{\beta}{N}$.

Một điều có thể dễ dàng nhận thấy là ta đều giả sử η , α hay N đều là các hằng số trong suốt thời gian dịch bệnh trong các mô hình trên. Tuy nhiên, trên thực tế, các con số này có thể thay đổi theo thời gian dựa theo tình hình của cộng đồng đó và mức độ của dịch bệnh. Do đó, bên cạnh các hàm $S(t)$, $I(t)$, $R(t)$ như định nghĩa trong phần 1.1.3, ta định nghĩa thêm một số hàm như sau:

- $\eta(t)$: thời gian từ lúc một người bị bệnh cho đến lúc người đó được cho là "khỏi bệnh" tính toán được tại thời điểm t .
- $\gamma(t)$: tỷ lệ hồi phục khi mắc bệnh được tính toán tại thời điểm t .
- $\alpha(t)$: xác suất trung bình một người khoẻ mạnh bị nhiễm bệnh được tính toán tại thời điểm t .
- $\beta(t)$: ước lượng tiếp xúc của người trong nhóm $S(t)$ với người trong nhóm $I(t)$ tại thời điểm t .
- $N(t)$: dân số trong cộng đồng đang xét tại thời điểm t .

Tương tự như khi là hằng số, ta cũng có một số quan hệ giữa 5 hàm số trên như $\gamma(t) = \frac{1}{\eta(t)}$ và $\alpha(t) = \frac{\beta(t)}{N(t)}$.

Một điều có thể được để ý nữa đó chính là với các lập luận tương tự như trong phần 1.1.3, các kết quả nhận được sẽ không thay đổi nếu ta xét với α , η là các hàm số theo biến t . Do đó, với những suy luận tương tự, ta có thể viết lại hệ (2) như sau:

$$\begin{aligned}\frac{dR}{dt} &= \frac{I(t)}{\eta(t)} \\ \frac{dI}{dt} &= -\frac{I(t)}{\eta(t)} + \alpha(t)S(t)I(t) \\ \frac{dS}{dt} &= -\alpha(t)S(t)I(t) \\ S(t) + I(t) + R(t) &= N(t) \forall t \in \mathbb{R}^+ \\ S(0) &= a, I(0) = b, R(0) = 0 \\ a, b &\in \mathbb{N} \\ \alpha(t), \eta(t) &> 0; \alpha(t) \leq 1 \forall t \in \mathbb{R}^+\end{aligned}$$

Hay:

$$\begin{aligned}\frac{dR}{dt} &= \gamma(t)I(t) \\ \frac{dI}{dt} &= -\frac{I(t)}{\eta(t)} + \frac{\beta(t)}{N(t)}S(t)I(t) \\ \frac{dS}{dt} &= -\frac{\beta(t)}{N(t)}S(t)I(t) \\ S(t) + I(t) + R(t) &= N(t) \forall t \in \mathbb{R}^+ \\ S(0) &= a, I(0) = b, R(0) = 0 \\ a, b &\in N \\ \gamma(t), \beta(t), N(t) &> 0 \forall t \in \mathbb{R}^+\end{aligned}\tag{7}$$

Với các hàm đã được định nghĩa như trên.

1.2.4 Hệ số lây nhiễm cơ bản R_0

Một trong những đại lượng quan trọng nhất trong các mô hình dịch bệnh là hệ số lây nhiễm cơ bản, hay còn gọi là "hệ số R_0 ". Xét trong hệ (2), R_0 có thể được tính như sau: $R_0 = \alpha\eta$. Trong đó, α thể hiện số người khỏe mạnh trung bình mà một người mắc bệnh có thể bị lây trong thời gian nhiễm bệnh η . Tại một thời điểm, dịch bệnh sẽ được coi là dập tắt nếu $R_0 < 1$, nếu $R_0 > 1$ thì dịch vẫn đang bùng phát. Điều này có thể thấy tương đối đúng bằng trực giác: nếu một người bệnh lây cho nhiều hơn một người thì số người bị bệnh sẽ tăng lên (theo cấp số nhân) và dịch bệnh bùng; còn nếu một người bị bệnh lây ít hơn cho một người bệnh khác thì số người bị bệnh đang giảm và dịch đang tắt.

1.2.5 Mô hình SIRD

Một điểm hạn chế của mô hình SIR đó chính là ta đã gộp số người chết và số người hồi phục vào nhóm R. Do đó, nếu ta muốn tính toán số người chết hoặc khỏi bệnh nhưng không chết, ta cần mở rộng mô hình SIR đã tính toán.

Với các định nghĩa $S(t), I(t)$ của mô hình SIR như đã trình bày ở 1.1.3, ta định nghĩa thêm hàm $D(t)$ thể hiện số người chết sau thời gian t , đồng thời $R(t)$ chỉ thể hiện số người hồi phục tại thời điểm t , không bao gồm người chết như xây dựng trên.

Gọi, N là tổng số dân của cộng đồng; $\gamma = \frac{1}{\eta}$ là tỉ lệ hồi phục; β là ước lượng số tiếp xúc của người trong nhóm S với người trong nhóm I thì α có thể được tính bởi $\alpha = \frac{\beta}{N}$. Như đã lập luận trong phần 1.1.3, ta sẽ có:

$$S(t) + I(t) + R(t) + D(t) = N \forall t \in \mathbb{R}^+$$

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

Và:

$$\frac{dS}{dt} = -\alpha S(t)I(t) = -\frac{\beta}{N}S(t)I(t)$$

Ta xây dựng hàm $D(t)$. Gọi khả năng một người không hồi phục được (chết) là μ , hay còn gọi là tỷ lệ tử vong.

Xét số người chết tại 2 thời điểm t và $t + \delta t$. Theo quy tắc tam suất, kết hợp với lập luận như trên, số người chết tại thời điểm $t + \delta t$ sẽ tăng thêm $\delta t \mu I(t)$ so với thời điểm t trước đó. Cho nên $R(t + \delta t) = R(t) + \delta t \mu I(t)$. Khi đó:

$$\frac{D(t + \delta t) - D(t)}{\delta t} = \frac{D(t) + \delta t \mu I(t) - R(t)}{\delta t} = \mu I(t)$$

Cho nên:

$$\lim_{\delta t \rightarrow +\infty} \frac{D(t + \delta t) - D(t)}{\delta t} = \mu I(t)$$

Hay:

$$\frac{dD}{dt} = \mu I(t)$$

Lại để ý một điều nữa, số người nhiễm mới tại thời điểm t sẽ giảm một lượng bằng tổng số người chết và hồi phục tại thời điểm t . Từ các nhận xét trên, ta có thể thay thế phương trình 2 của hệ (2) trong trường hợp này thành:

$$\frac{dI}{dt} = \alpha S(t)I(t) - \mu I(t) - \gamma I(t) = \frac{\beta}{N} S(t)I(t) - \mu I(t) - \gamma I(t)$$

Kết hợp lại các công thức đã thiết lập, ta được hệ:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N} S(t)I(t) \\ \frac{dI}{dt} &= \frac{\beta}{N} S(t)I(t) - \mu I(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t) \\ \frac{dD}{dt} &= \mu I(t) \\ S(t) + I(t) + R(t) + D(t) &= N \forall t \in \mathbb{R}^+ \\ S(0) &= a, I(0) = b, R(0) = 0, D(0) = 0 \\ a, b &\in N \\ \beta, \gamma, \mu, N &> 0 \end{aligned} \tag{8}$$

Hệ trên là mô hình SIRD mở rộng của SIR, các giá trị của các hàm số tại thời điểm t có thể tính được nếu biết giá trị của chúng tại một thời điểm t_0 trước đó.

Hệ số lây nhiễm cơ sở R_0 trong hệ SIRD có thể được xác định bởi công thức $R_0 = \frac{\beta}{\gamma + \mu}$.

1.2.6 Một số mô hình khác

Bên cạnh mô hình SIR hay SIRD, tùy vào các điều kiện được đưa vào xem xét trong mô hình như miễn dịch tự nhiên, ủ bệnh, khả năng tái nhiễm bệnh,... mà người ta có thể đưa ra các mô hình khác như SIS, SEIR, SIRS

2 Xây dựng chương trình sử dụng thuật toán Euler tìm nghiệm của hệ SIR

2.1 Phương pháp Euler

2.1.1 Tổng quan lý thuyết

Xét bài toán:

$$\begin{cases} y'(t) = f(t, y(t)), a \leq t \leq b \\ y(a) = y_0 \end{cases} \quad (9)$$

Hàm $y(t)$ khả vi trên đoạn $[a; b]$, y_0 là giá trị ban đầu cho trước của $y(t)$ tại $t = a$. Bài toán trên được gọi là bài toán Cauchy.

Đối với bài toán Cauchy, ta chỉ có thể tìm được nghiệm gần đúng của một số phương trình đơn giản, còn đối với trường hợp hàm $f(t, y(t))$ được cho ở dạng bất kỳ, nói chung là không có phương pháp giải, ngoài ra một số nghiệm giải được tương đối phức tạp. Cho nên việc tính gần đúng các giá trị có vai trò tương đối quan trọng.

Phương pháp Euler có thể được dùng để xấp xỉ giá trị của hàm $y(t)$ bằng dãy $\{y(t_n)\}$ thoả:

$$\begin{cases} y(t_0) = y_0 \\ y(t_{k+1}) = y(t_k) + hf(t_k, y(t_k)), k = 0, 1, \dots, n-1 \end{cases} \quad (10)$$

Tham số h có thể được xây dựng bởi công thức $h = \frac{b-a}{n}$ trong đó n là số đoạn nhỏ bằng nhau trong đoạn $[a; b]$.

Phương pháp trên còn có thể mở rộng cho hệ các phương trình vi phân với công thức tương tự.

2.1.2 Thực hiện

Cho bài toán: Dùng mô hình SIR để miêu tả một loại cúm trong một cộng đồng có các đặc điểm được quan sát như sau:

- Cộng đồng này đang bị cách ly, không ai được ra vào.
- Loại cúm này có thời gian từ khi phát bệnh cho đến khi hồi phục là η tuần không đổi theo thời gian.
- Một người khi mắc bệnh và hồi phục thì không còn mắc bệnh này lần thứ 2.
- Sau một thời gian điều tra, tỷ lệ mắc bệnh khi có tiếp xúc với người bệnh ở mức α sau một tuần tiếp xúc và giả sử tỷ lệ này cũng không đổi theo thời gian.
- Tại thời điểm ban đầu, số người có khả năng nhiễm bệnh là a người, số người mắc bệnh là b người và số ca hồi phục khi ấy không có.

Sử dụng các ký hiệu và các công thức như đã trình bày ở phần 1, ta được mô hình:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha S(t)I(t) = -\frac{\beta}{N}S(t)I(t) \\ \frac{dI}{dt} &= -\frac{I(t)}{\eta} + \alpha S(t)I(t) = \frac{\beta}{N}S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt} &= \frac{I(t)}{\eta} = \gamma I(t) \\ S(t) + I(t) + R(t) &= N \forall t \in R^+ \\ S(0) &= a, I(0) = b, R(0) = 0 \\ a, b &\in N\end{aligned}$$

Ta hiện thực chương trình giải quyết bài toán trên bằng Python. Nội dung chi tiết hiện thực của thuật toán nằm trong file Euler_method_SIR.ipynb được đính kèm. Phần hiện thực công thức Euler được cho như dưới đây.

```
for step in range(1, self.eons):
    S_to_I = (self.rateSI * Susceptible[-1] * Infectious[-1]) / self.numIndividuals
    I_to_R = Infectious[-1] * self.rateIR
    Susceptible.append(Susceptible[-1] - self.h*S_to_I)
    Infectious.append(Infectious[-1] + self.h*(S_to_I - I_to_R))
    Recovered.append(Recovered[-1] + self.h*I_to_R)
```

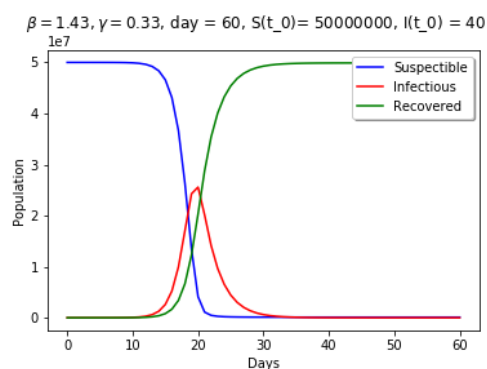
Kết quả ở một số trường hợp có thể được thể hiện ở các hình dưới đây (nhóm chọn đơn vị là ngày).

Với các tham số $t, \beta, \gamma, S(t_0), I(t_0)$ nhập vào ban đầu lần lượt là 60, 1.43, 0.33, 50 000 000, 40, ta được kết quả như hình dưới, trong đó, kết quả ước tính ở ngày thứ 60 của S, I, R lần lượt là 109428.3482, 4.2857, 49890607.3661.

Statistical table of first 60 days:

| Day | Susceptible | Infectious | Recovered |
|-----|-------------|------------|-----------|
| 0. | 50000000. | 40. | 0. |
| 1. | 49999942.8 | 84. | 13.2 |
| 2. | 49999822.68 | 176.4 | 40.92 |
| 3. | 49999570.43 | 370.44 | 99.13 |
| 4. | 49999040.71 | 777.92 | 221.38 |
| 5. | 49997928.31 | 1633.6 | 478.09 |
| 6. | 49995592.36 | 3430.46 | 1017.18 |
| 7. | 49990687.24 | 7203.53 | 2149.23 |
| 8. | 49980388.12 | 15125.49 | 4526.39 |
| 9. | 49958767.17 | 31755.03 | 9517.81 |
| 10. | 49913394.96 | 66648.07 | 19996.96 |
| 11. | 49818253.37 | 139795.8 | 41990.83 |
| 12. | 49619072.19 | 292844.36 | 88123.44 |
| 13. | 49203495.49 | 611782.43 | 184762.08 |
| 14. | 48342583.73 | 1270805.99 | 386650.28 |

(a) Các giá trị tính toán được trong các ngày



(b) Phác hoạ kết quả thành đồ thị

Hình 1: Kết quả tính toán và phác hoạ đồ thị

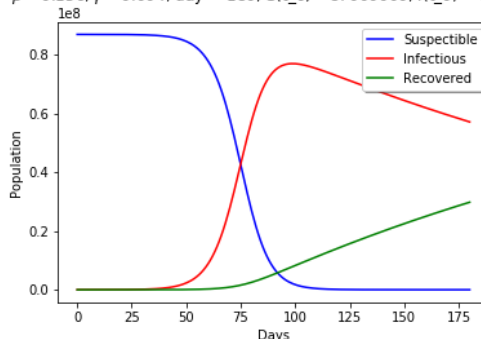
Với các tham số $t, \beta, \gamma, S(t_0), I(t_0)$ nhập vào ban đầu lần lượt là 180, 0.156, 0.004, 87 000 000, 2000 ta được kết quả như hình dưới, trong đó, kết quả ước tính ở ngày thứ 180 của S, I, R lần lượt là 57.7249, 57181583.8143, 29820358.4608.

Statistical table of first 180 days:

| Day | Susceptible | Infectious | Recovered |
|-----|-------------|------------|-----------|
| 0 | 87000000 | 2000 | 0 |
| 1 | 86999688.01 | 2303.99 | 8 |
| 2 | 86999328.59 | 2654.19 | 17.22 |
| 3 | 86998914.55 | 3057.61 | 27.83 |
| 4 | 86998437.58 | 3522.35 | 40.06 |
| 5 | 86997888.12 | 4057.73 | 54.15 |
| 6 | 86997255.14 | 4674.48 | 70.38 |
| 7 | 86996525.96 | 5384.96 | 89.08 |
| 8 | 86995685.96 | 6203.42 | 110.62 |
| 9 | 86994718.3 | 7146.27 | 135.43 |
| 10 | 86993603.58 | 8232.4 | 164.02 |
| 11 | 86992319.44 | 9483.61 | 196.95 |
| 12 | 86990840.17 | 10924.95 | 234.88 |
| 13 | 86989136.09 | 12585.32 | 278.58 |
| 14 | 86987173.07 | 14498 | 328.93 |

(a) Các giá trị tính toán được trong các ngày

$\beta = 0.156, \gamma = 0.004, \text{day} = 180, S(t_0) = 87000000, I(t_0) = 2000$



(b) Phác hoạ kết quả thành đồ thị

Hình 2: Kết quả tính toán và phác hoạ đồ thị

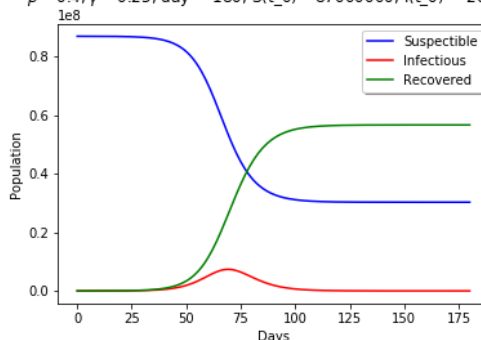
Với các tham số $t, \beta, \gamma, S(t_0), I(t_0)$ nhập vào ban đầu lần lượt là 180, 0.4, 0.25, 87 000 000, 2000 ta được kết quả như hình dưới, trong đó, kết quả ước tính ở ngày thứ 180 của S, I, R lần lượt là 30303676.0952, 56.9696, 56698266.9352.

Statistical table of first 180 days:

| Day | Susceptible | Infectious | Recovered |
|-----|-------------|------------|-----------|
| 0 | 87000000 | 2000 | 0 |
| 1 | 86999200.02 | 2299.98 | 500 |
| 2 | 86998280.06 | 2644.95 | 1075 |
| 3 | 86997222.12 | 3041.65 | 1736.23 |
| 4 | 86996005.53 | 3497.83 | 2496.64 |
| 5 | 86994606.49 | 4022.4 | 3371.1 |
| 6 | 86992997.67 | 4625.63 | 4376.7 |
| 7 | 86991147.61 | 5319.28 | 5533.11 |
| 8 | 86989020.16 | 6116.91 | 6862.93 |
| 9 | 86986573.77 | 7034.08 | 8392.16 |
| 10 | 86983760.63 | 8088.69 | 10150.68 |
| 11 | 86980525.83 | 9301.32 | 12172.85 |
| 12 | 86976806.23 | 10695.6 | 14498.18 |
| 13 | 86972529.23 | 12298.7 | 17172.08 |
| 14 | 86967611.41 | 14141.83 | 20246.75 |

(a) Các giá trị tính toán được trong các ngày

$\beta = 0.4, \gamma = 0.25, \text{day} = 180, S(t_0) = 87000000, I(t_0) = 2000$



(b) Phác hoạ kết quả thành đồ thị

Hình 3: Kết quả tính toán và phác hoạ đồ thị

Các ví dụ trên có thể thể hiện các loại dịch bệnh khác nhau. Trong đó, ví dụ 1 thể hiện một loại bệnh nhẹ và dễ lây lan, dịch bệnh sẽ kết thúc khi mọi người trong cộng đồng đều nhiễm; ví dụ 2 thể hiện một dịch bệnh tương đối mạnh, khó khỏi và kéo dài rất lâu; ví dụ 3 thể hiện một loại bệnh bình thường, nó sẽ kết thúc khi một số lượng người bị nhiễm khỏi tạo nên miễn dịch bảo vệ các người khác (miễn dịch cộng đồng).

2.2 Phương pháp Runge-Kutta

Nhược điểm của phương pháp Euler là bậc của độ chính xác giảm dần. Muốn có độ chính xác cao đòi hỏi h phải rất nhỏ, điều này sẽ tăng thời gian tính toán lên cao - tăng độ phức tạp của giải thuật.

Phương pháp Runge-Kutta giải quyết được tình trạng này bằng cách sử dụng các điểm trung gian giữa các bước lặp (khai triển Taylor nghiệm $y(x)$ tại x_i tại nhiều số hạng hơn). Trong thực tế, người ta sẽ dùng công thức Runge-Kutta bậc 4 vì nó có độ chính xác cao mà lại không quá phức tạp.

Công thức Runge-Kutta tổng quát:

$$y_{t+h} = y_t + h \cdot \sum_{i=1}^s a_i k_i + \mathcal{O}(h^{s+1})$$

Khi bỏ qua $\mathcal{O}(h^5)$, ta thu được công thức Runge-Kutta bậc 4:

- $y_0 = y(x_0)$
- $k_1^{(i)} = hf(x_i, y_i)$
- $k_2^{(i)} = hf(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2})$
- $k_3^{(i)} = hf(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2})$
- $k_4^{(i)} = hf(x_i + h, y_i + k_3^{(i)})$
- $y_{i+1} = y_i + \frac{1}{6}(k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)})$ $i = 0, 1, \dots, n-1$

Hiện thực chương trình Runge-Kutta bậc 4 bằng Python.

```
for i in range(1, self.eons):
    Si = Susceptible[-1]
    Ii = Infectious[-1]

    Sk1 = self.dSdt(Si, Ii)
    Ik1 = self.dIdt(Si, Ii)

    Sk2 = self.dSdt(Si + self.dt * Sk1/2, Ii + self.dt * Ik1/2)
    Ik2 = self.dIdt(Si + self.dt * Sk1/2, Ii + self.dt * Ik1/2)

    Sk3 = self.dSdt(Si + self.dt * Sk2/2, Ii + self.dt * Ik2/2)
    Ik3 = self.dIdt(Si + self.dt * Sk2/2, Ii + self.dt * Ik2/2)

    Sk4 = self.dSdt(Si + self.dt * Sk3, Ii + self.dt * Ik3)
    Ik4 = self.dIdt(Si + self.dt * Sk3, Ii + self.dt * Ik3)

    nextSi = Si + self.dt * (Sk1+2*Sk2+2*Sk3+Sk4)/6
    nextIi = Ii + self.dt * (Ik1+2*Ik2+2*Ik3+Ik4)/6
    nextRi = self.numIndividuals - nextSi - nextIi
```

Hiện thực chi tiết nằm trong file RK4_method.ipynb. Kết quả khi thử với các ví dụ như trong phần 2.1.2

Day: 60

Susceptible: $S = 50\,000\,000$ people

Infectious: $I = 40$ people

Recovered: $R = 0$

Rate of the susceptible contact with the infectious: $\beta = 1.43$

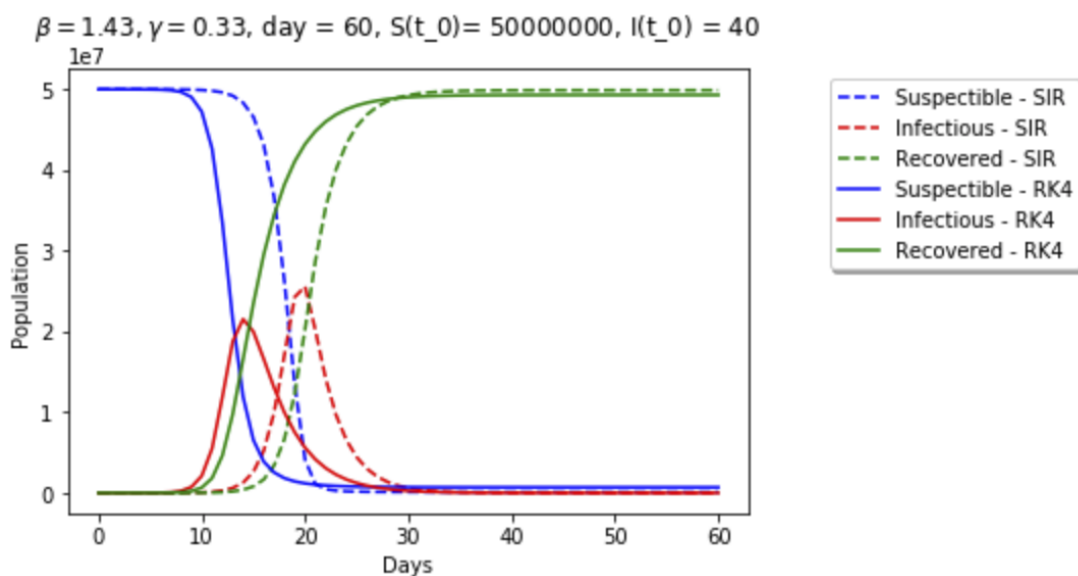
Rate of the recovered from the infectious: $\gamma = 0.33$

Giá trị ước lượng tại thời gian nhập vào của S, I, R lần lượt là 698903.117, 24.0122 và 49301112.8708.

Statistical table of first 60 days:

| Day | Susceptible | Infectious | Recovered |
|-----|-------------------|-------------------|-------------------|
| 0. | 50000000. | 40. | 0. |
| 1. | 49999896.63275002 | 119.51322484 | 23.85402514 |
| 2. | 49999587.79096325 | 357.08339991 | 95.12563683 |
| 3. | 49998665.04713713 | 1066.88240176 | 308.07046111 |
| 4. | 49995908.26572797 | 3187.4480006 | 944.28627143 |
| 5. | 49987673.48405226 | 9521.57784481 | 2844.93810294 |
| 6. | 49963087.4473495 | 28431.07820852 | 8521.47444199 |
| 7. | 49889790.08467562 | 84788.407293 | 25461.50803138 |
| 8. | 49672223.18974074 | 251923.75014895 | 75893.06011031 |
| 9. | 49034703.3673951 | 740369.80674815 | 224966.82585674 |
| 10. | 47235154.16642953 | 2108300.54589786 | 656585.28767261 |
| 11. | 42648939.00937086 | 5514735.80310257 | 1836365.18752657 |
| 12. | 33468875.15136261 | 11893785.36177734 | 4637379.48686004 |
| 13. | 21477742.72787583 | 18766683.06396878 | 9755614.20815539 |
| 14. | 11927653.90477021 | 21528042.22586821 | 16544343.86936158 |

(a) Các giá trị tính toán được trong các ngày



(b) Phác họa kết quả thành đồ thị

Hình 4: Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp

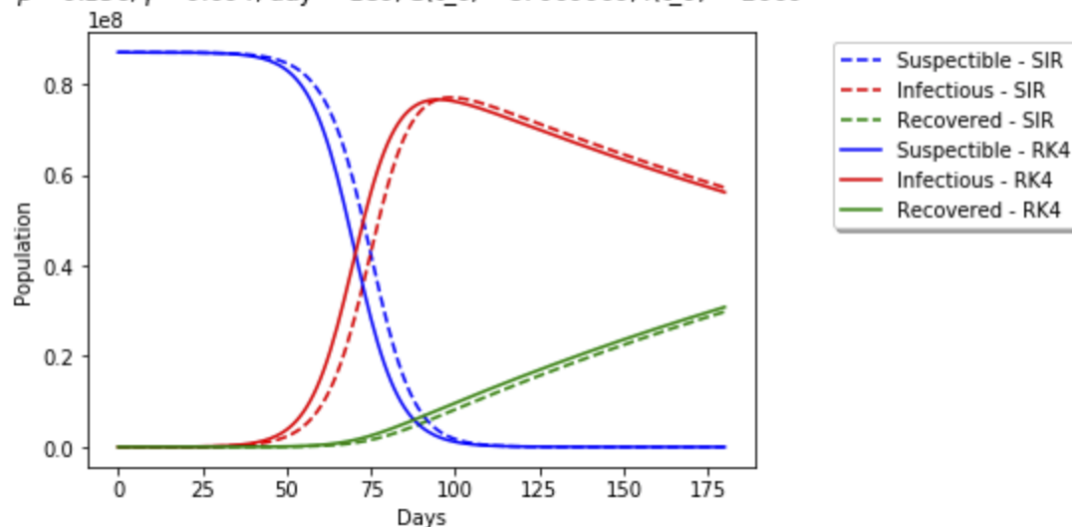
Day: 180
Susceptible: $S = 87\,000\,000$ people
Infectious: $I = 2000$ people
Recovered: $R = 0$
Rate of the susceptible contact with the infectious: $\beta = 0.156$
Rate of the recovered from the infectious: $\gamma = 0.004$
Giá trị ước lượng tại thời gian nhập vào của S, I, R lần lượt là 86.1752, 56160676.5129 và 30841237.3119.

Statistical table of first 180 days:

| Day | Susceptible | Infectious | Recovered |
|-----|-------------------|----------------|--------------|
| 0. | 87000000. | 2000. | 0. |
| 1. | 86999663.0499902 | 2328.31005061 | 8.63995919 |
| 2. | 86999270.78971489 | 2710.5120773 | 18.69820781 |
| 3. | 86998814.14064342 | 3155.45180017 | 30.40755641 |
| 4. | 86998282.5343371 | 3673.42663474 | 44.03902816 |
| 5. | 86997663.66801573 | 4276.42385438 | 59.90812989 |
| 6. | 86996943.220051 | 4978.3977971 | 78.3821519 |
| 7. | 86996104.51882738 | 5795.59250745 | 99.88866517 |
| 8. | 86995128.15734063 | 6746.91724723 | 124.92541214 |
| 9. | 86993991.54466097 | 7854.38351949 | 154.07181954 |
| 10. | 86992668.38394342 | 9143.61365782 | 188.00239876 |
| 11. | 86991128.06499083 | 10644.43266627 | 227.50234289 |
| 12. | 86989334.95742865 | 12391.55689139 | 273.48567996 |
| 13. | 86987247.58829238 | 14425.3953073 | 327.01640032 |
| 14. | 86984817.68521094 | 16792.98174451 | 389.33304455 |

(a) Các giá trị tính toán được trong các ngày

$\beta = 0.156, \gamma = 0.004, \text{day} = 180, S(t_0) = 87000000, I(t_0) = 2000$



(b) Phác họa kết quả thành đồ thị

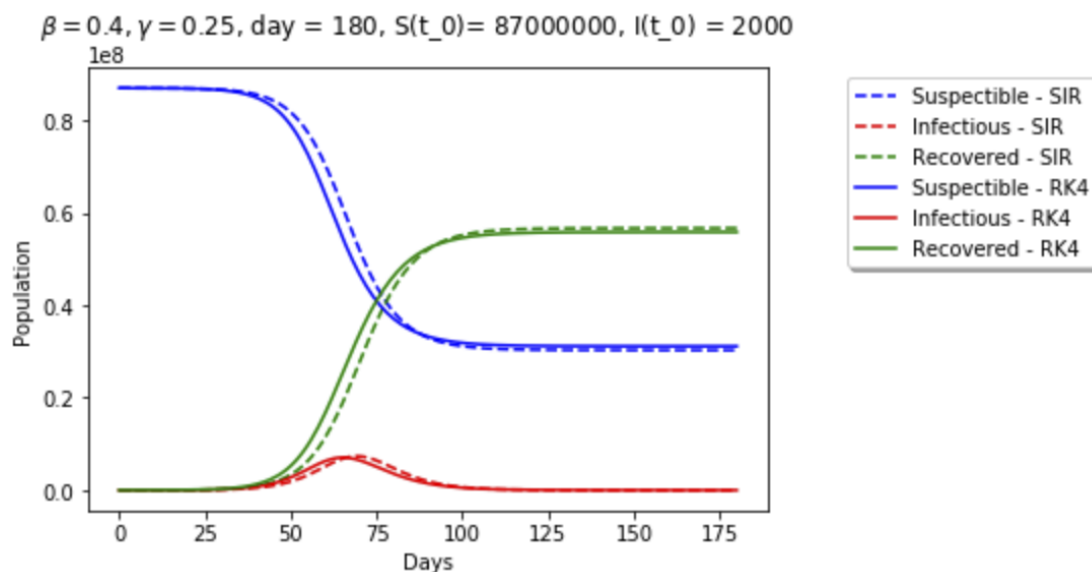
Hình 5: Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp

Day: 180
Susceptible: $S = 87\,000\,000$ people
Infectious: $I = 2000$ people
Recovered: $R = 0$
Rate of the susceptible contact with the infectious: $\beta = 0.4$
Rate of the recovered from the infectious: $\gamma = 0.25$
Giá trị ước lượng tại thời gian nhập vào của S, I, R lần lượt là 31146807.9348, 106.0921 và 55855085.9731.

Statistical table of first 180 days:

| Day | Suspectible | Infectious | Recovered |
|-----|-------------------|----------------|----------------|
| 0. | 87000000. | 2000. | 0. |
| 1. | 86999136.91625755 | 2323.64132668 | 539.44241578 |
| 2. | 86998134.18062969 | 2699.64295025 | 1166.17642007 |
| 3. | 86996969.20405483 | 3136.47193587 | 1894.32400929 |
| 4. | 86995615.74622376 | 3643.96312438 | 2740.29065186 |
| 5. | 86994043.32630669 | 4233.53958889 | 3723.13410442 |
| 6. | 86992216.53889519 | 4918.46845135 | 4864.99265346 |
| 7. | 86990094.26002385 | 5714.15766962 | 6191.58230653 |
| 8. | 86987628.72575833 | 6638.50027502 | 7732.77396665 |
| 9. | 86984764.46309952 | 7712.27353622 | 9523.26336426 |
| 10. | 86981437.0498058 | 8959.60166396 | 11603.34853024 |
| 11. | 86977571.6761246 | 10408.49197015 | 14019.83190525 |
| 12. | 86973081.47729033 | 12091.4558698 | 16827.06683987 |
| 13. | 86967865.60092412 | 14046.22778435 | 20088.17129154 |
| 14. | 86961806.96809605 | 16316.59688445 | 23876.4350195 |

(a) Các giá trị tính toán được trong các ngày



(b) Phác họa kết quả thành đồ thị

Hình 6: Kết quả tính toán và đồ thị so sánh với phương pháp Euler trong cùng trường hợp

Các lệnh vẽ đồ thị so sánh nằm trong file combine.ipynb. Các hình vẽ của từng cách tính riêng biệt sẽ được xuất ra file result.png.

Dựa vào đồ thị so sánh, ta có thể thấy có sự chênh lệch số liệu giữa 2 phương pháp trong một khoảng thời gian, nhưng chúng sẽ cùng hội tụ tại các điểm tương tự nhau. Thực tế người ta thấy phương pháp Runge-Kutta bậc 4 cho kết quả hơn phương pháp Euler truyền thống.

3 Lấy mẫu cho hệ số β và γ

3.1 Tổng quan lý thuyết

3.1.1 Phân phối chuẩn

Phân phối chuẩn, hay còn gọi là phân phối Gauss, là một phân phối xác suất rất quan trọng, được áp dụng trong nhiều lĩnh vực. Nó là một họ các phân phối, khác nhau về giá trị trung bình (tham số vị trí) μ và phương sai σ^2 .

Hàm mật độ xác suất của một biến ngẫu nhiên X với trung bình μ và phương sai σ^2 (hoặc độ lệch chuẩn σ) là một ví dụ của hàm Gauss và được cho như sau:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (11)$$

Một biến ngẫu nhiên X có phân phối này thì ta gọi X là có phân phối chuẩn và ký hiệu $X \sim N(\mu, \sigma^2)$.

Đồ thị của hàm mật độ xác suất phụ thuộc vào μ và σ , μ xác định vị trí trung tâm của đồ thị, còn σ xác định chiều dài và chiều rộng của đồ thị. Do đồ thị của hàm mật độ xác suất có dạng hình chuông nên đôi khi, phân phối chuẩn còn được gọi là đường cong chuông. Do tính chất đó của đồ thị hàm mật độ xác suất phân phối chuẩn, đây là phân phối rất phù hợp với các mô hình dịch bệnh khi thể hiện đầy đủ các giai đoạn từ ủ bệnh, tăng nhanh và đạt đỉnh cho đến khi giảm dần và kết thúc.

3.1.2 Thuật toán Metropolis-Hastings

Marko Chain Monte Carlo (MCMC) cho phép ta vẽ ra được một phân phối ngay cả nếu ta không tính được nó dựa trên các quan sát và các giả thiết trước đó. MCMC là một lớp các cách thức và thuật toán Metropolis-Hastings là một hiện thực của MCMC.

Xét bài toán: sử dụng thuật toán Metropolis-Hastings để lấy mẫu với tham số đầu vào là phân bố xác suất tiên nghiệm $\pi(\beta, \gamma)$ cho trước. Giá trị trả về là một mẫu gồm các cặp β và γ có phân bố xác suất $\pi(\beta, \gamma)$.

Như đã phân tích ở trên, ta sẽ sử dụng phân bố chuẩn cho các phân bố xác suất tiên nghiệm cho trước. Thuật toán Metropolis-Hastings được thực hiện như sau:

- Bước 1: Ta giả sử β, γ là 2 biến độc lập. Chọn phân bố xác suất tiên nghiệm cho β, γ là hàm phân bố chuẩn:

$$\begin{aligned} \pi(\beta) &= f(\beta, \mu_\beta, \sigma_\beta^2) = \frac{1}{\sigma_\beta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\beta-\mu_\beta}{\sigma_\beta}\right)^2} \\ \pi(\gamma) &= f(\gamma, \mu_\gamma, \sigma_\gamma^2) = \frac{1}{\sigma_\gamma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\gamma-\mu_\gamma}{\sigma_\gamma}\right)^2} \end{aligned}$$

- Bước 2: Khởi tạo mẫu $\beta = \beta_0, \gamma = \gamma_0$ ban đầu từ $\pi(\beta), \pi(\gamma)$.
- Bước 3: Khởi tạo β^*, γ^* ngẫu nhiên từ hàm phân phối chuẩn với giá trị β, γ cho trước:

$$\beta^* = \beta + N(0, 1)$$

hay

$$\begin{aligned}\beta^*|\beta &\sim N(\beta, 1) \\ \gamma^* &= \gamma^* + N(0, 1)\end{aligned}\tag{12}$$

hay

$$\gamma^*|\gamma \sim N(\gamma, 1)\tag{13}$$

Từ (8) và (9), kết hợp với (7), ta có:

$$\begin{aligned}p(\beta^*|\beta) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\beta^*-\beta)^2} \\ p(\gamma^*|\gamma) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\gamma^*-\gamma)^2}\end{aligned}$$

- Bước 4: Vì β, γ là 2 biến độc lập nhau nên

$$\begin{cases} p(\beta, \gamma) = p(\beta)p(\gamma) \\ \pi(\beta, \gamma) = \pi(\beta)\pi(\gamma) \end{cases}$$

Mà $p(\beta^*|\beta) = p(\beta|\beta^*), p(\gamma^*|\gamma) = p(\gamma|\gamma^*)$ cho nên: $p(\beta^*, \gamma^*|\beta, \gamma) = p(\beta, \gamma|\beta^*, \gamma^*)$.

Nên $p(\beta, \gamma)$ được chọn ở đây là hàm đối xứng. Do đó:

$$r = \min\left(1, \frac{\pi(\beta^*, \gamma^*)}{\pi(\beta, \gamma)}\right) = \left(1, \frac{\pi(\beta^*)\pi(\gamma^*)}{\pi(\beta)\pi(\gamma)}\right)$$

với r được gọi là xác suất giữ lại β^*, γ^*

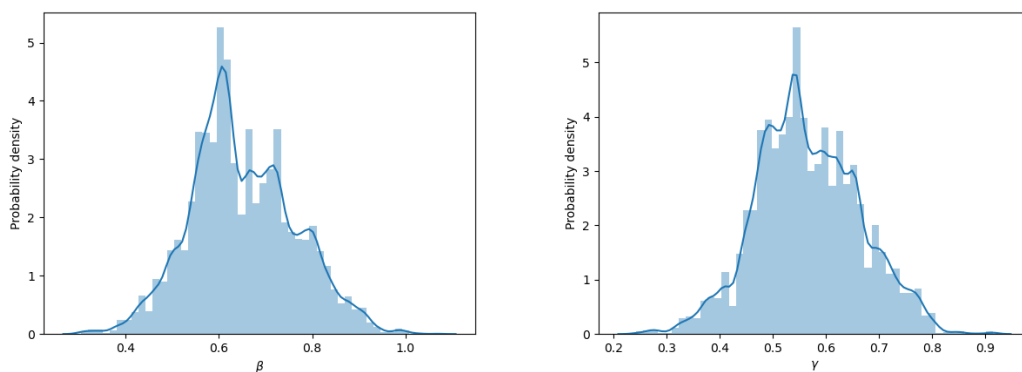
- Bước 5: Lấy ngẫu nhiên giá trị q từ phân phối đều liên tục $U(0, 1)$
 - Nếu $q < r$ chấp nhận mẫu (β^*, γ^*) , gán $\beta = \beta^*, \gamma = \gamma^*$
 - Ngược lại, giữ nguyên giá trị β, γ

Lặp lại từ Bước 2 cho đến khi đủ kích cỡ mẫu.

3.2 Hiện thực

Trong ví dụ này, từ bài báo [Muk], ta thấy μ là giá trị trung bình nên ta chọn nó bằng giá trị β cho trong bài và giá trị ngẫu nhiên. Ta đi tạo mẫu (β, γ) từ phân bố xác suất tiên nghiệm $\pi(\beta)$ với $\mu_\beta = 0.66, \sigma_\gamma = 0.11$ và $\pi(\gamma)$ với $\mu_\beta = 0.573, \sigma_\gamma = 0.1$. Với số lượng mẫu càng lớn, phân bố ngẫu nhiên tạo ra sẽ càng gần với phân bố được cho trong tiên nghiệm, do đó ta sẽ càng có kết quả chính xác hơn.

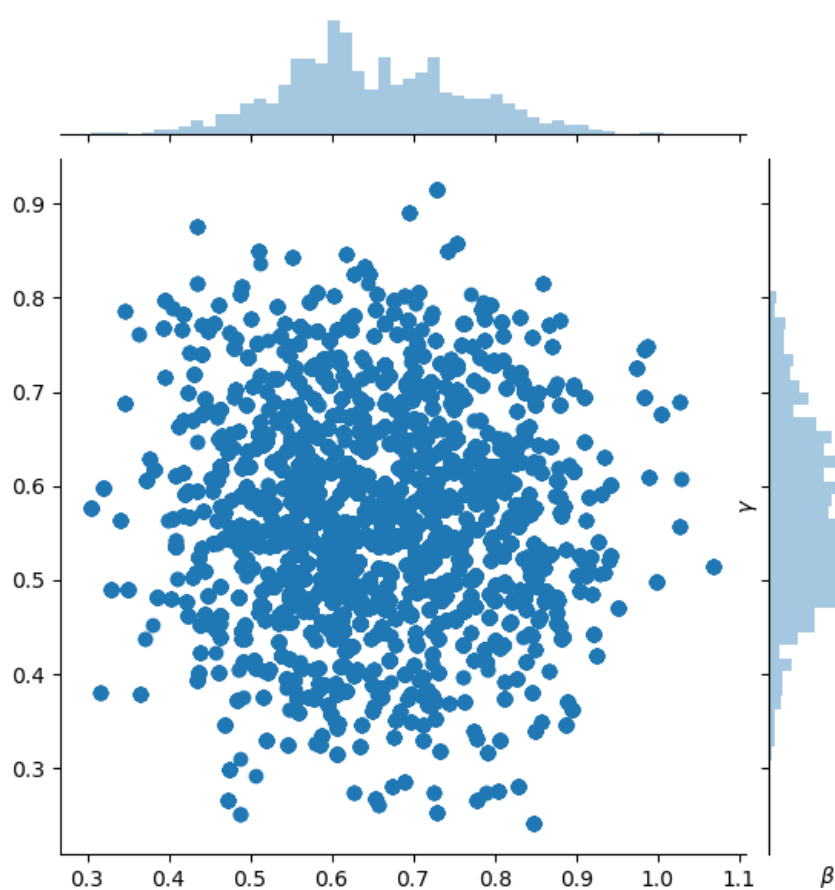
Kết quả với số lượng mẫu $m = 50000$:



(a) Phân bố mẫu $\beta \sim N(0.66, 0.11)$

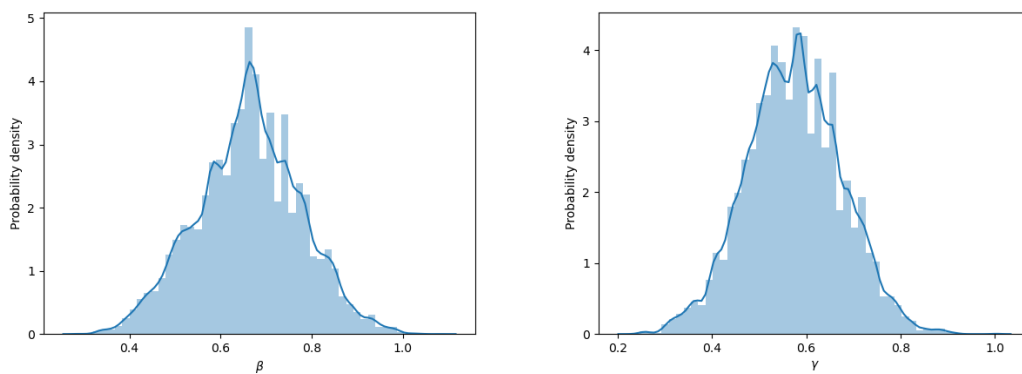
(b) Phân bố mẫu $\gamma \sim N(0.573, 0.1)$

Hình 7: Các phân bố mẫu của β và γ cùng với 50000 lần lặp



Hình 8: Phân phối chung cho (β, γ) với 50000 lần lặp

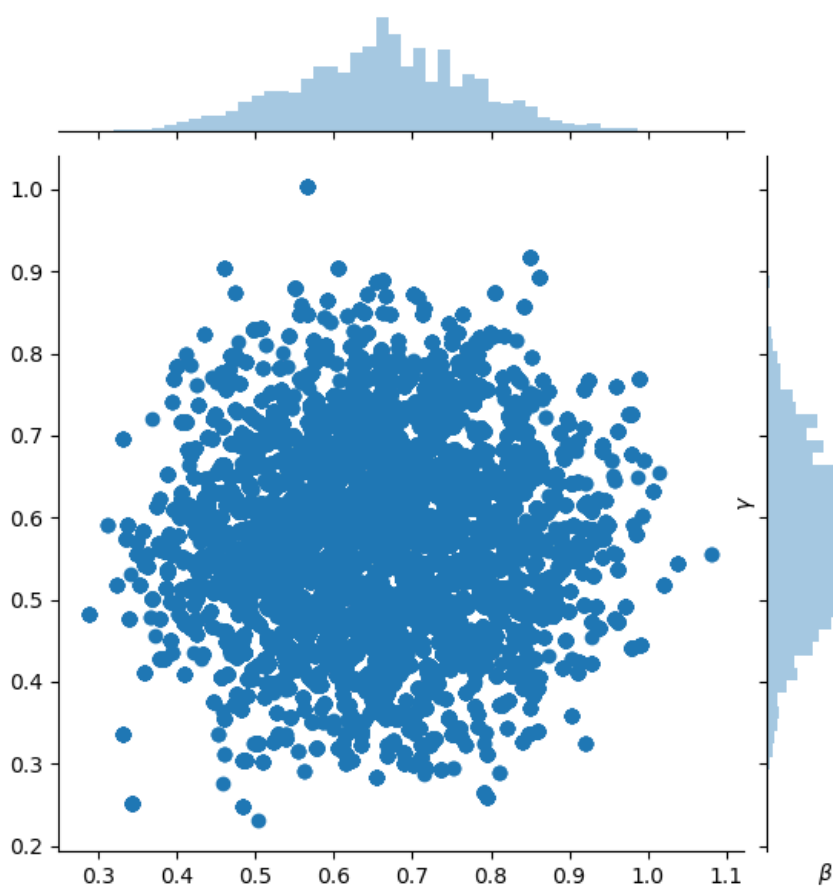
Kết quả với số lượng mẫu $m = 100000$:



(a) Phân bố mẫu $\beta \sim N(0.66, 0.11)$

(b) Phân bố mẫu $\gamma \sim N(0.573, 0.1)$

Hình 9: Các phân bố mẫu của β và γ cùng với 100000 lần lặp



Hình 10: Phân phối chung cho (β, γ) với 50000 lần lặp



Trong các hình trên, các chấm xanh thể hiện các điểm thoả mãn (β, γ) có phân bố xác suất (chuẩn) $\pi(\beta, \gamma)$ cho trước. Các đường xanh trong các phân bố mẫu β, γ được tạo tự động để ước lượng đường phân bố liên tục gần nhất. Nội dung chi tiết hiện thực nằm trong file sample.py, các kết quả được xuất ra các file beta.png, gamma.png, sample.txt và samples.png.

4 Ước lượng giá trị R_0 ở các khu vực

4.1 Tổng quan lý thuyết

Từ cơ sở lý thuyết ở Phần 3 vừa được đề cập, ta áp dụng Thuật toán Metropolis–Hastings; khi đó ta có thể dễ tạo được một mẫu (β, γ) dựa trên phân bố xác suất $\pi(\beta, \gamma)$. Ta tiến hành tính hệ số ước lượng R_0 .

$$R_0 = \frac{\beta}{\gamma}$$

Khi $R_0 < 1$ thì ta có thể kết luận không có đợt bùng phát dịch bệnh xảy ra do tỷ lệ tiếp xúc người mắc bệnh β nhỏ hơn tỷ lệ hồi phục.

Với kết quả ngược lại, nếu $R_0 > 1$, ta có cơ sở kết luận rằng dịch bệnh sẽ bùng phát trong tương lai gần.

Giá trị trung bình của R_0 được tính bằng công thức:

$$E(R_0) = \int \pi(\beta, \gamma | X) R_0(\beta, \gamma) d(\beta, \gamma) \quad (14)$$

Giá trị trung bình này ta có thể ước lượng được vì $\pi(X|\beta, \gamma)$ có thể tính được dễ dàng từ công thức:

$$\pi(X|\beta, \gamma) = \prod_{i=1}^n f(X(t_i)|\beta, \gamma) = \prod_{i=1}^n \frac{\gamma^\beta}{\Gamma(\beta)} X(t_i)^{\beta-1} \exp\{-\gamma X(t_i)\} \quad (15)$$

Tuy nhiên, giá trị trung bình của hệ số R_0 không thể tính trực tiếp được thông qua lý thuyết tích phân do độ phức tạp trong việc tính toán quá lớn. Thay vào đó, ta sử dụng công thức xấp xỉ:

$$E(R_0) = \int \pi(\beta, \gamma | X) R_0(\beta, \gamma) d(\beta, \gamma) \propto \int \pi(X|\beta, \gamma) R_0(\beta, \gamma) d(\beta, \gamma) \approx \sum_{i=1}^m \pi(X|\beta_i, \gamma_i) \frac{\beta_i}{\gamma_i} \quad (16)$$

Từ giá trị xấp xỉ đạt được, ta sẽ phân tích và dự đoán được tình hình dịch bệnh sẽ có xu hướng bùng phát hay đã được khống chế trong tương lai gần.

4.2 Hiện thực

Thực hiện việc truy xuất cơ sở dữ liệu, với mỗi file *.csv*, thực hiện query 2 cột là *Confirmed* và *Recovered* với tên của quốc gia được chọn. Ví dụ, trong bài báo cáo này chọn hai quốc gia là Italy và Việt Nam.

Sau đó, ta sẽ tính giá trị biến X , với X được hiểu là biến ngẫu nhiên quan sát số ca mắc bệnh và số ca hồi phục tại từng thời điểm $t > 0$.

Do đó: $X(t_i) = \text{Confirmed}_{t_i} + \text{Recovered}_{t_i}$.

Từ các lý thuyết đã trình bày như trên, kết hợp với chương trình đã viết ở phần 3, ta sẽ ước lượng giá trị trung bình của R_0 . Trong đó, ở công thức (11), ta lựa chọn π là phân phối chuẩn, với các giá trị σ, μ của β và γ được chọn với lý do như phần 4. Đối với mỗi quốc gia, ta có được một giá trị R_0 trung bình tương ứng.

Cuối cùng, ta thực hiện tính toán ước lượng tại 2 quốc gia đã chọn là Italy và Việt Nam. Ở cả hai quốc gia này, ta đều được kết quả là giá trị R_0 nhỏ hơn 1. Nội dung hiện thực nằm trong file Bai4.py được đính kèm.

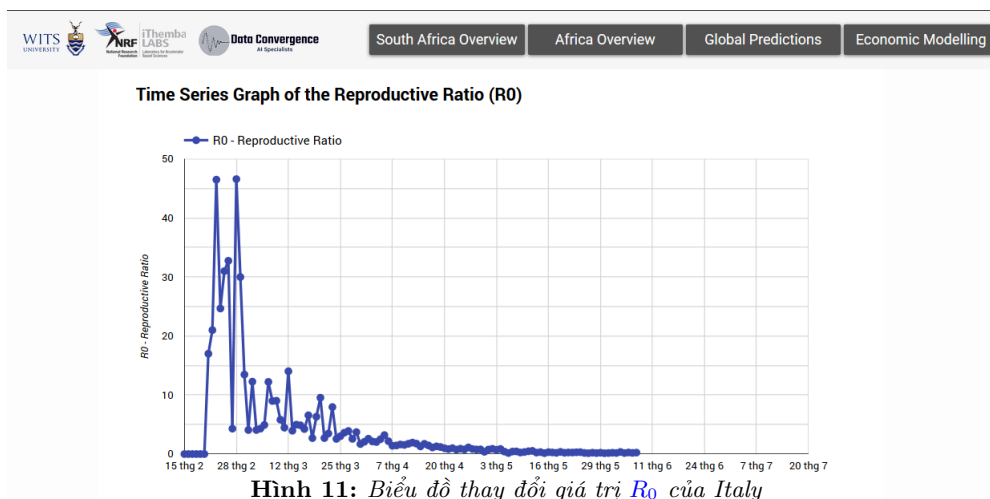
4.3 Phân tích

4.3.1 Italy

Dịch Covid-19 bùng phát tại trung tâm tài chính Milan của Italy từ ngày 21-2, sau đó bắt đầu lan ra khu vực giàu có và đông dân ở miền bắc nước này. Italy nhanh chóng trở thành ổ dịch lớn nhất thế giới với số ca mắc lên đến 245000 người, tổng số trường hợp tử vong lên đến 35000 tính đến ngày 23/7/2020.

Đứng trước làn sóng Covid-19 diễn biến mạnh mẽ, ngay sau khi xuất hiện ca tử vong đầu tiên, quan chức cấp cao nước này đã ban hành lệnh giới nghiêm ở các vùng chịu ảnh hưởng nặng nề, sau đó là ban hành lệnh giới nghiêm trên phạm vi cả nước. Cụ thể: từ ngày 10-3 đến 3-4 là thời gian thi hành lệnh giới nghiêm, người dân khi ra khỏi nhà chỉ được đi trong bán kính 2km, đeo khẩu trang, phải có giấy xác nhận của cơ quan hành chính và chỉ một người trong nhà được đi ra. Điều này nhằm hạn chế thấp nhất giá trị β và từ đó giảm thiểu R_0 ở Italy. Đến nay mới có các kế hoạch khôi phục kinh tế.

Sắc lệnh mới sẽ có hiệu lực trên khắp Italy, trong đó có đảo Sicily và Sardinia, và tác động đến khoảng 60 triệu dân.



Hình 11: Biểu đồ thay đổi giá trị R_0 của Italy

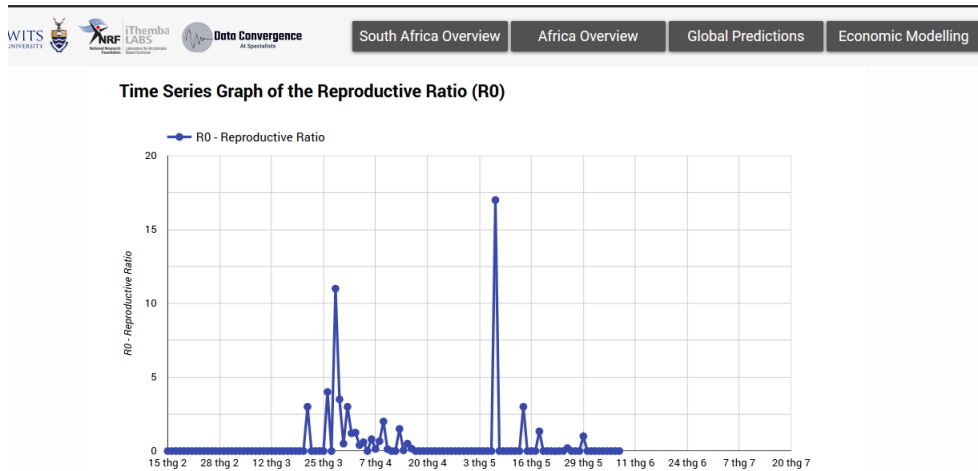
Nhờ chính sách cách ly và hạn chế đi lại được ban hành, chỉ số R_0 từ mức rất cao đến nay đã đang dần tiến về 0. Điều đó cho thấy việc cách ly và hạn chế đi lại đã ảnh hưởng rất lớn đến sự thay đổi của R_0 .

4.3.2 Việt Nam

Có thể nói Việt Nam chúng ta hiện tại đang là một trong số ít các nước kiểm chế dịch bệnh Covid-19 hiệu quả. Tính tới thời điểm hiện tại (ngày 23 tháng 7 năm 2020), tổng số ca mắc bệnh là 412 ca, bình phục 365 ca và đặc biệt là không có trường hợp nào tử vong. Giáp biên giới Trung Quốc, nơi đầu tiên bùng phát dịch bệnh, nhưng Việt Nam lại có thể kiểm soát tốt tình hình lây nhiễm và được bạn bè quốc tế đánh giá cao.

Đại dịch COVID-19 do virus SARS-CoV-2 gây ra được xác nhận lần đầu tiên tại Việt Nam vào ngày 23 tháng 1 năm 2020.

Việt Nam đã thực hiện các biện pháp cách ly, theo dõi và hạn chế người đến từ vùng có dịch, đóng cửa biên giới và triển khai việc thực hiện khai báo y tế. Nhiều hoạt động tập trung đông người tại các địa phương bị hạn chế, đồng thời nhiều nơi thực hiện các biện pháp như đo thân



Hình 12: Biểu đồ thay đổi giá trị R_0 của Việt Nam

nhật, trang bị chất sát khuẩn, phát khẩu trang miễn phí ở các nơi công cộng, siết chặt kiểm soát. Việc đi lại, buôn bán trong nước cũng bị hạn chế.

| # | Thời gian | Số ca nhiễm | Mô tả |
|---|----------------|-------------|--|
| 1 | 23/1-25/2/2020 | 16 | Các ca nhiễm được xác nhận từ BN có lịch sử di chuyển qua Trung Quốc. |
| 2 | 6-19/3/2020 | 69 | Virus lây lan toàn cầu, nhiều ca nhiễm từ các quốc gia khác nhưng vẫn dễ dàng truy tìm nguồn gốc và cách ly. |
| 3 | 20/3-21/4/2020 | 183 | Các ca nhiễm lây lan trong cộng đồng, nhiều ổ bệnh bắt đầu xuất hiện ở các vùng mật độ dân cư cao. Nguồn gốc lây nhiễm chưa thể truy tìm hết được. |
| 4 | 22/4/2020-nay | 113+ | Việt Nam đã kiểm soát tốt dịch bệnh, việc phòng chống dịch đại hoi và căn cơ hơn đồng thời phát triển kinh tế xã hội. |

Dòng thời gian: Tháng 1 đến tháng 3 • Tháng 4 đến tháng 6 • Tháng 7 đến tháng 12

Hình 13: Các giai đoạn và các mốc thời gian cụ thể về Covid-19 ở Việt Nam

Phân tích các mốc thời gian cụ thể để làm rõ biểu đồ R_0 tại Việt Nam:

Giai đoạn 1: gồm 16 ca bệnh COVID-19 đầu tiên. Hai trường hợp xác nhận nhiễm bệnh đầu tiên đã nhập viện vào Bệnh viện Chợ Rẫy, Thành phố Hồ Chí Minh, bao gồm một người đàn ông Trung Quốc 66 tuổi đi từ Vũ Hán đến Hà Nội để thăm con trai sống ở Việt Nam, và con trai 28 tuổi, người được cho là đã bị lây bệnh từ cha mình khi họ gặp gỡ tại Nha Trang. Vào ngày 1 tháng 2, một người phụ nữ 25 tuổi được xác định nhiễm virus corona tại tỉnh Khánh Hòa. Đáng chú ý, đây là trường hợp truyền nhiễm nội địa đầu tiên tại Việt Nam, dẫn đến việc thủ tướng Nguyễn Xuân Phúc công bố dịch tại Việt Nam và ra quyết định thắt chặt biên giới, thu hồi giấy phép hàng không và hạn chế thị thực. Ngày 25 tháng 2, trường hợp số 16 được tuyên bố hồi phục và xuất viện. Đây cũng là ca cuối cùng trong 16 ca đầu tiên ở Việt Nam xuất viện. Các biện pháp cách ly và xét nghiệm giúp phát hiện sớm virus được áp dụng nhanh và mạnh mẽ trong giai đoạn này.

Giai đoạn 2: gồm các ca bệnh xâm nhập từ nước ngoài. Tối ngày 6 tháng 3, Hà Nội đã công bố trường hợp đầu tiên dương tính với virus corona, là một phụ nữ 26 tuổi. Tối ngày 19 tháng 3, tổng số bệnh nhân trên cả nước lên 85. Hà Nội đã khoanh vùng khu vực có bệnh nhân, lập chốt tại 2 đầu khu phố Trúc Bạch, đóng cửa các hàng quán tại khu vực. Các đội điều tra dịch tễ lập danh sách những người tiếp xúc gần với bệnh nhân và những người tiếp xúc với người tiếp xúc gần tại khu vực nhà bệnh nhân; tại nơi bệnh nhân đến khám ban đầu; tại Bệnh viện Nhiệt đới cơ sở 2 và hồi cứu quá trình nhập cảnh tại sân bay Nội Bài của người này.

Giai đoạn 3: với nguy cơ lây lan trong cộng đồng. Chiều 20/3, Bộ Y tế công bố 2 BN COVID-19 thứ 86 và 87 là 2 nữ điều dưỡng Bệnh viện Bạch Mai (Hà Nội) với tiền sử dịch tễ

không cho thấy nguồn lây khi cả 2 không có lịch sử tiếp xúc với các BN COVID-19. Ngày 21/3, Việt Nam tạm dừng nhập cảnh đối với tất cả người nước ngoài từ 0 giờ ngày 22/3, đồng thời thực hiện cách ly tập trung 14 ngày đối với mọi trường hợp nhập cảnh. Từ 0 giờ ngày 1 tháng 4, Việt Nam thực hiện cách ly xã hội trong vòng 15 ngày. Cùng ngày, Thủ tướng Nguyễn Xuân Phúc công bố dịch COVID-19 trên phạm vi cả nước, thay thế cho quyết định công bố dịch trước đó vào ngày 1 tháng 2 năm 2020.

Giai đoạn 4: là giai đoạn Việt Nam đã kiểm soát tốt dịch bệnh. Từ ngày 23 tháng 4, cả nước cơ bản dừng cách ly xã hội nhưng vẫn tiếp tục đảm bảo các quy tắc phòng chống dịch. Ngày 25 tháng 4, Thủ tướng ban hành chỉ thị 19 nhằm tiếp tục các biện pháp phòng, chống dịch COVID-19 trong tình hình mới.

Có thể trong mỗi giai đoạn có nguy cơ bùng dịch ở nước ta, chỉ số R_0 từ mức dưới 1 tăng vọt lên đến 15, nhưng nhanh chóng quay về mức ban đầu. Đó là do các biện pháp cách ly, khoanh vùng và giãn cách xã hội do Chính phủ ban hành. Các biện pháp diễn ra nhanh chóng, hiệu quả, kịp thời nên R_0 tăng vọt nhưng cũng xuống rất nhanh.

5 Kết luận

Mô hình SIR/SIRD được sử dụng trong báo cáo là mô hình tương đối tiêu chuẩn và đơn giản để mô phỏng lại đại dịch Covid 19. Với các số liệu được cập nhật hằng ngày, mô hình có thể đưa ra con số tương đối để phỏng đoán tình hình trong các ngày tiếp theo. Tuy nhiên, do ở mức đơn giản, không xét đến các biến số quan trọng khác như tình hình cách ly của người dân, tỉ lệ trang bị y tế được cung cấp, các nguồn hỗ trợ, ... nên mô hình chỉ mang tính tham khảo / cảnh báo, không nên sử dụng như số liệu chính thống cho các việc quan trọng như ban bố chính sách,...

Hiện nay, với các công cụ hiện đại hơn như Machine Learning và sức mạnh tính toán lớn của các siêu máy tính, việc thêm / cập nhật các biến số trở nên khả thi thì mô hình thu được sẽ có sai số nhỏ so với thực tế. Đây là triển vọng rất lớn trong việc phân tích và đưa ra các chính sách phù hợp hơn nhằm đẩy lùi dịch bệnh, khôi phục kinh tế và trạng thái bình thường.

Tài liệu

- [Dal] Dalgaard, P. *Introductory Statistics with R*. Springer 2008.
- [K-Z] Kenett, R. S. and Zacks, S. *Modern Industrial Statistics: with applications in R, MINITAB and JMP*, 2nd ed., John Wiley and Sons, 2014.
- [Ker] Kerns, G. J. *Introduction to Probability and Statistics Using R*, 2nd ed., CRC 2015.
- [VINIF] Nguyễn Hoàng Thạch, Phan Thị Hà Dương *Tìm hiểu về một Mô hình dự báo dịch Covid-19 từ Vũ Hán*
- [A-J] Abhijit Chakraborty, Jiaying Chen *Analyzing Covid-19 Data using SIRD Models*
- [F-S] Frank R. Giordano, William P. Fox, Steven B. Horton *A First Course in Mathematical* 5th ed., 2013.
- [Muk] Mukesh Jakhar, P. K. Ahluwalia and Ashok Kumar *COVID-19 Epidemic Forecast in Different States of India using SIR Model*, May 14, 2020
- [NDH] NDH COVID-19 Corona Virus South African Resource The Johns Hopkins University <https://www.covid19sa.org>
- [ZN] ZingNew *Toàn cảnh 100 ngày chống dịch Covid-19 tại Việt Nam*, May 4, 2020
- [ND] Báo Nhân Dân *Italy thành ổ dịch Covid-19 lớn thứ hai thế giới, Thủ tướng lệnh phong tỏa toàn quốc*, March 11, 2020
- [VE] VnExpress *Cả thứ 6 Việt Nam nhiễm virus Corona*, February 1, 2020
- [TT] Tuổi Trẻ *Dừng toàn bộ chuyến bay giữa Việt Nam và Trung Quốc từ chiều nay*, February 1, 2020