전염병 감염 확산 모형 (SIR model)

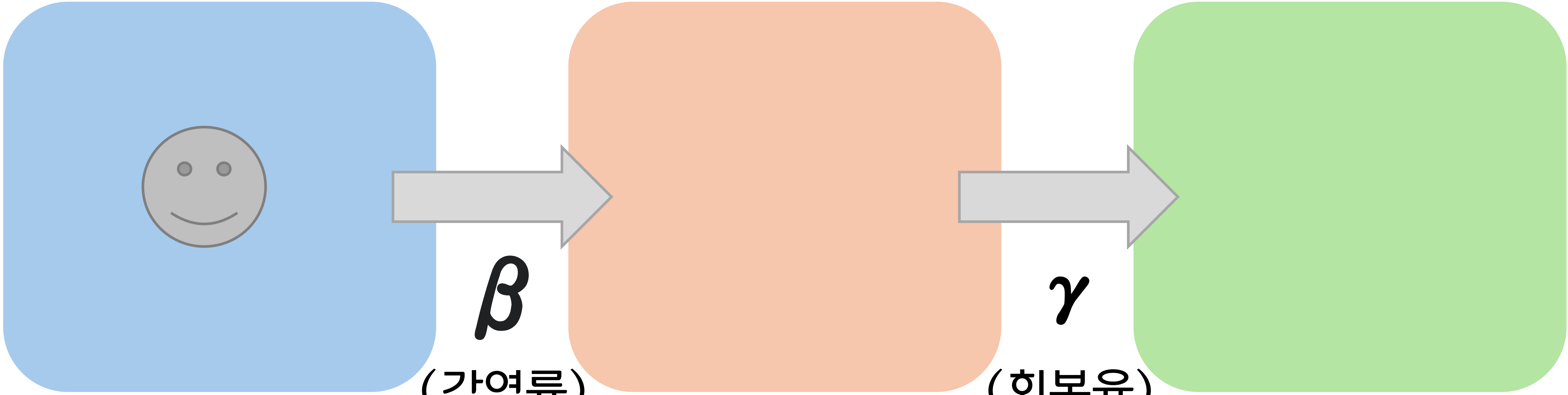
; 코로나 1차 대유행 시기 데이터를 통하여
Euler 방법과 odeint 함수를 비교하는 방식으로

1930015 이은지

▪ SIR model – Introduction

총 인구 N명,

Susceptible (미감염자) **I**nfected (감염자) **R**ecovery (회복자)



β (감염률)

γ (회복율)

기초감염재생산지수 $R_0 = \frac{\beta}{\gamma} > 1$

SIR model – Theory

초기 조건 $S(0) = 1$ / $I(0) = 0$ / $R(t) = 0$

$$N = \underset{\text{미감염자}}{s(t)} + \underset{\text{감염자}}{i(t)} + \underset{\text{회복자}}{r(t)} \rightarrow S(t) = \frac{s(t)}{N},\ I(t) = \frac{i(t)}{N},\ R(t) = \frac{r(t)}{N}$$

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

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

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

(1) ÷ (3)

$$\frac{dS}{dR} = -\frac{\beta \cdot S \cdot I}{\gamma \cdot I} = -R_0 \cdot S \qquad \left(R_0 = \frac{\beta}{\gamma}\right)$$

$$\Rightarrow S = a \cdot \exp(-R_0 \cdot S)$$

초기조건 대입 시, $a = 1$

$$\therefore S = \exp(-R_0 \cdot S)$$

▪ SIR model – Theory

초기 조건 $S(0) = 1$ / $I(0) = 0$ / $R(t) = 0$

$$N = \underbrace{s(t)}_{\text{미감염자}} + \underbrace{i(t)}_{\text{감염자}} + \underbrace{r(t)}_{\text{회복자}} \rightarrow S(t) = \frac{s(t)}{N},\ I(t) = \frac{i(t)}{N},\ R(t) = \frac{r(t)}{N}$$

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

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

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

(2) ÷ (3)

$(S = \exp(-R0 \cdot S))$

$$\frac{dI}{dR} = \frac{\beta}{\gamma} \cdot S - 1 = R_0 \cdot S - 1 = R_0 \cdot \exp(-R_0 \cdot S) - 1$$

$$\Rightarrow I = b - R - \exp(-R_0 \cdot S)$$

초기조건 대입 시, $b = 1$

& $\exp(-x) = 1 + (-x) + (\frac{1}{2!}x^2) + \cdots$ 를 이용,

$$\therefore I = 1 - R - \left(1 - R_0 \cdot S + \frac{R_0^2 \cdot S^2}{2}\right)$$

$$= \frac{R_0^2}{2}\left[\frac{2(R_0 - 1)}{R_0^2} - R\right] \cdot R$$

SIR model – Theory

$$I = \frac{R_0^2}{2}\left[\frac{2(R_0-1)}{R_0^2} - R\right]\cdot R \quad \& \quad \frac{dR}{dt} = \gamma \cdot I \quad \text{---} (3)$$

$$t' = \gamma \cdot t \quad \rightarrow \quad \frac{d}{dt} = \frac{d}{dt'} \cdot \frac{dt'}{dt} = \frac{d}{dt'} \cdot \gamma$$

$$\frac{dR}{dt'} = I(t')$$

$$\therefore \frac{dR}{dt'} = \frac{R_0^2}{2}\left[\frac{2(R_0-1)}{R_0^2} - R\right]\cdot R$$

$$= A\,[\,B - R\,]\cdot R$$

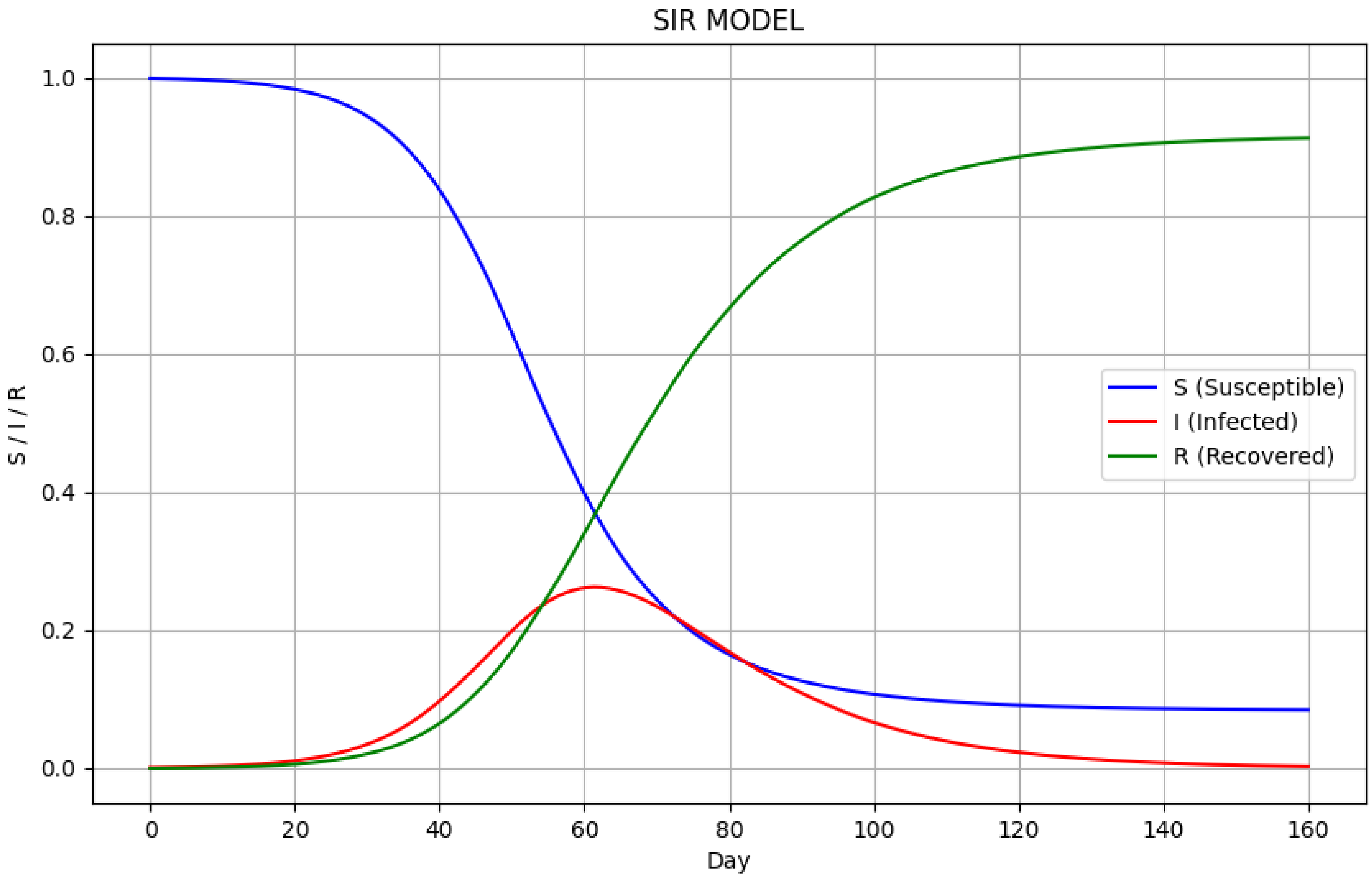
Logistic function

$$f(x) = \frac{e^x}{e^x+1} = \frac{1}{1+e^{-x}}$$

$$\frac{d}{dx}f(x) = f(x)\big(1-f(x)\big)$$

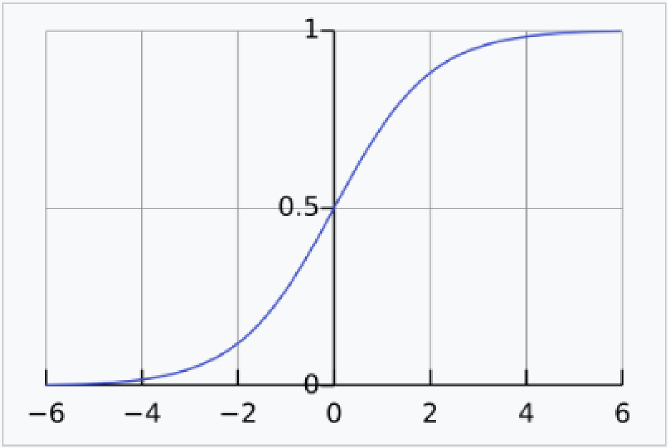
SIR model – Theory

$$\therefore \frac{dR}{dt'} = \frac{R_0^2}{2}\left[\frac{2(R_0-1)}{R_0^2} - R\right]\cdot R$$



Logistic function

$$\frac{d}{dx}f(x) = f(x)\big(1 - f(x)\big)$$

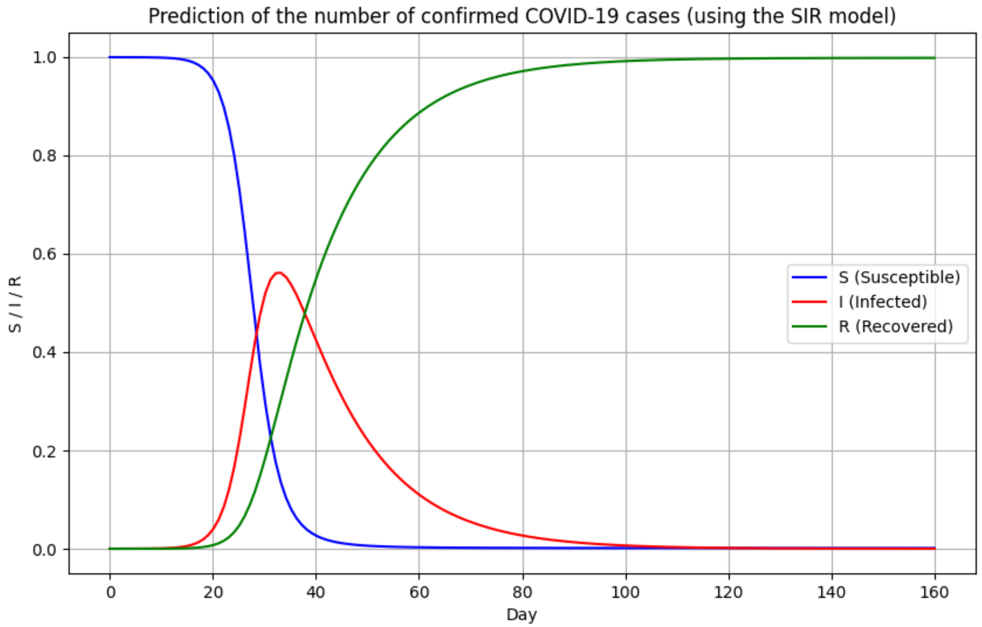


Standard logistic function where $L = 1, k = 1, x_0 = 0$

▪ SIR model – ODEint

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.integrate import odeint

# SIR 모델 정의
def sir_model(y, t, beta, gamma):
    S, I, R = y

    dSdt = -beta * S * I

    dIdt = beta * S * I - gamma * I

    dRdt = gamma * I

    return [dSdt, dIdt, dRdt]

# 초기값 설정
I0 = 31

N = 2436488
S0 = (N-I0)/N     #미감염자(Susceptible)
I0 = 34/N         #감염자(Infected)
R0 = 12/N         #회복자(Recovered)

y0 = [S0, I0, R0]

beta = 0.471      #감염률
gamma = 1/14      #회복율
r0 = 6.6          #감염재생산지수(R0)

t = np.linspace(0, 160, 160)
```

```python
solution = odeint(sir_model, y0, t, args=(beta, gamma))
S, I, R = solution.T

#일별 변화 출력
for day, s, i, r in zip(t, S, I, R):
    print(f"Day {int(day)}: S = {s*N:.4f}, I = {i*N:.4f}, R = {r*N:.4f}")

#결과 시각화
plt.figure(figsize=(10, 6))
plt.plot(t, S, label="S (Susceptible)", color="blue")
plt.plot(t, I, label="I (Infected)", color="red")
plt.plot(t, R, label="R (Recovered)", color="green")
plt.title("Prediction of the number of confirmed COVID-19 cases (using the SIR model)")
plt.xlabel("Day")
plt.ylabel("S / I / R")
plt.legend()
plt.grid(True)
plt.show()
```

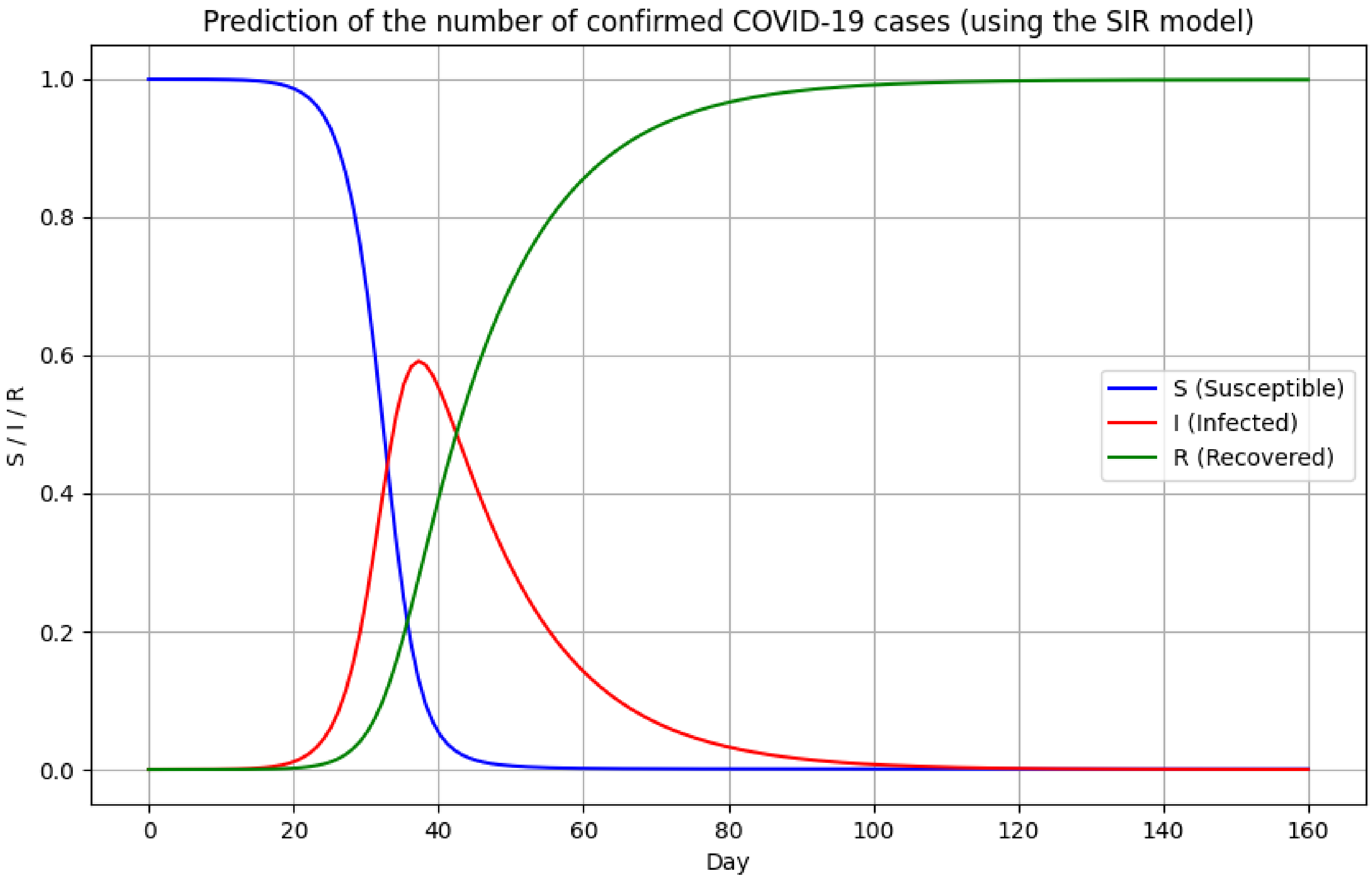


▪ SIR model – Euler

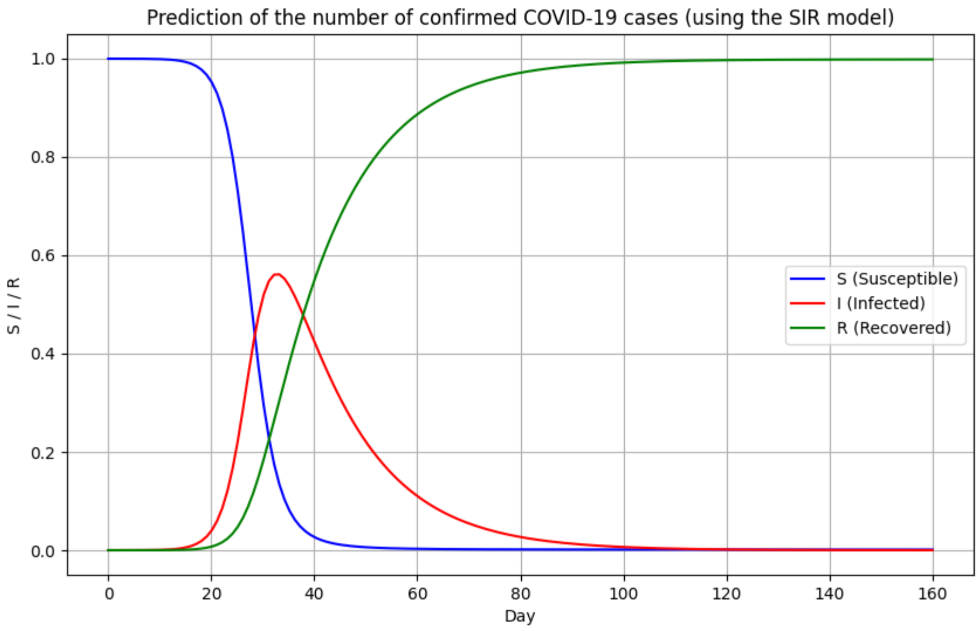
```python
import numpy as np
import matplotlib.pyplot as plt

#SIR 모델 정의
def sir_model(y, t, beta, gamma):
    S, I, R = y

    dSdt = -beta * S * I

    dIdt = beta * S * I - gamma * I

    dRdt = gamma * I

    return [dSdt, dIdt, dRdt]

#초기값 설정
I0 = 31

N = 2436488
S0 = (N-I0)/N    #미감염자(Susceptible)
I0 = 34/N        #감염자(Infected)
R0 = 12/N        #회복자(Recovered)

beta = 0.471     #감염률
gamma = 1/14     #회복율
r0 = 6.6         #감염재생산지수(R0)\

y0 = [S0, I0, R0]
t = np.linspace(0, 160, 160)
```

```python
#오일러 방식을 이용
def euler_method(sir_model, y0, t, beta, gamma):

    S = [S0]
    I = [I0]
    R = [R0]
    dt = t[1] - t[0]

    for i in range(1, len(t)):

        dS, dI, dR = sir_model([S[i-1], I[i-1], R[i-1]], t[i-1], beta, gamma)

        S.append(S[i-1] + dS * dt)
        I.append(I[i-1] + dI * dt)
        R.append(R[i-1] + dR * dt)

    return np.array(S), np.array(I), np.array(R)

S, I, R = euler_method(sir_model, y0, t, beta, gamma)

#일별 변화 출력 (chat gpt에게 각각의 데이터 값을 출력하는 코드를 짜달라고 했습니다.)
for day, s, i, r in zip(t, S, I, R):
    print(f"Day {int(day)}: S = {s * N:.4f}, I = {i * N:.4f}, R = {r * N:.4f}")

#결과 시각화
plt.figure(figsize=(10, 6))
plt.plot(t, S, label="S (Susceptible)", color="blue")
plt.plot(t, I, label="I (Infected)", color="red")
plt.plot(t, R, label="R (Recovered)", color="green")
plt.title("Prediction of the number of confirmed COVID-19 cases (using the SIR model)")
plt.xlabel("Day")
plt.ylabel("S / I / R")
plt.legend()
plt.grid(True)
plt.show()
```

- SIR model – Euler



VS OEDint



▪ SIR model vs Real Data

Real Data (I)

일자	계(명)	국내발생(명)	해외유입(명)	
누적(명)	34,572,554	34,492,629	79,925	
2020-02-18	2	2	-	
2020-02-19	34	34	-	Day 0
2020-02-20	16	16	-	
2020-02-21	74	74	-	.
2020-02-22	190	190	-	
2020-02-23	210	209	1	.
2020-02-24	207	206	1	.
2020-02-25	130	129	1	
2020-02-26	253	252	1	.
2020-02-27	449	447	2	.
PEAK 2020-02-28	427	427	-	
2020-02-29	909	909	-	Day 10
2020-03-01	595	593	2	
2020-03-02	686	684	2	
2020-03-03	600	599	1	
2020-03-04	516	516	-	
2020-03-05	438	437	1	

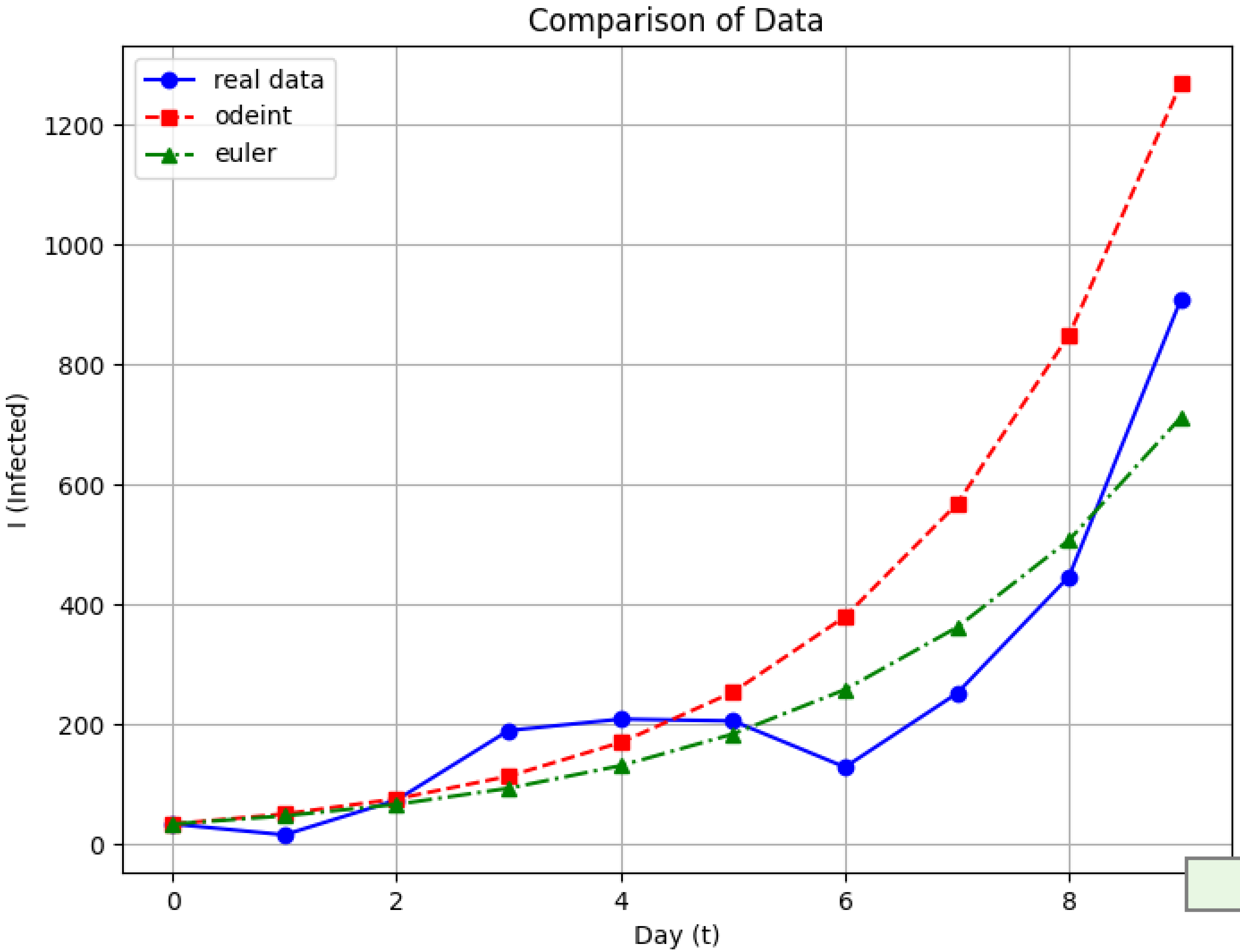
ODEint

```
Day 0: S = 2436457.0000, I = 34.0000, R = 12.0000
Day 1: S = 2436437.1200, I = 50.8651, R = 15.0149
Day 2: S = 2436407.4321, I = 76.0506, R = 19.5173
Day 3: S = 2436363.0291, I = 113.7195, R = 26.2514
Day 4: S = 2436296.6496, I = 170.0317, R = 36.3187
Day 5: S = 2436197.4805, I = 254.1600, R = 51.3595
Day 6: S = 2436049.2043, I = 379.9463, R = 73.8494
Day 7: S = 2435827.5750, I = 567.9572, R = 107.4678
Day 8: S = 2435496.3098, I = 848.9680, R = 157.7222
Day 9: S = 2435001.2945, I = 1268.8745, R = 232.8310
Day 10: S = 2434261.6523, I = 1896.2621, R = 345.0856
Day 11: S = 2433156.8524, I = 2833.3243, R = 512.8233
Day 12: S = 2431507.2750, I = 4232.3108, R = 763.4143
Day 13: S = 2429045.7683, I = 6319.5680, R = 1137.6637
Day 14: S = 2425376.1056, I = 9430.5877, R = 1696.3067
Day 15: S = 2419912.6111, I = 14060.7908, R = 2529.5981
```

Euler

```
Day 0: S = 2436457.0000, I = 34.0000, R = 12.0000
Day 1: S = 2436440.8855, I = 47.6707, R = 14.4438
Day 2: S = 2436418.2918, I = 66.8379, R = 17.8703
Day 3: S = 2436386.6141, I = 93.7114, R = 22.6745
Day 4: S = 2436342.2003, I = 131.3895, R = 29.4102
Day 5: S = 2436279.9303, I = 184.2154, R = 38.8542
Day 6: S = 2436192.6266, I = 258.2781, R = 52.0952
Day 7: S = 2436070.2274, I = 362.1129, R = 70.6597
Day 8: S = 2435898.6290, I = 507.6834, R = 96.6876
Day 9: S = 2435658.0644, I = 711.7569, R = 133.1787
Day 10: S = 2435320.8334, I = 997.8283, R = 184.3383
Day 11: S = 2434848.1270, I = 1398.8130, R = 256.0600
Day 12: S = 2434185.5886, I = 1960.8078, R = 356.6036
Day 13: S = 2433257.1179, I = 2748.3399, R = 497.5422
Day 14: S = 2431956.2358, I = 3851.6774, R = 695.0868
Day 15: S = 2430134.0815, I = 5396.9816, R = 971.9369
```

▪ SIR model – ODEint vs Euler



ODEint

- Runge-Kutta 4차 방법, LSODA, BDF, Adams-Bashforth와 같은 여러 알고리즘을 사용
- 속도가 느리지만 매우 정확
- 고차 미분 방정식과 복잡한 시스템을 다루는 데 적합

Euler Method

- 1차 방법을 사용하므로 상대적으로 정확도가 낮지만 계산 속도가 빠름
- 비교적 간단한 문제를 다루는 데 적합

Why ODEint < Euler ?

▪ SIR model - Conclusion

1. SIR 모델을 통해 구한 값이 Real Data와 다른 이유?

- 코로나19 1차 대유행 당시 마스크 착용 권고, 사회적 거리두기 등 다양한 방역정책 시행
- Real Data는 정확한 감염자의 수가 아님

(∵ 자가진단키트 사용 후 보건소에 가서 검사를 받고 양성 판정 받은 사람 수가 기록된 것

→ 자가진단키트 불량 및 보건소 검사 오류 가능성有, 감염 후 바로 확진 판정을 받는 것 X (즉, 감염자≠확진자),

다양한 사회적 이유로 확진자 낙인에 대한 두려움 때문에 검사받지 않은 감염자 등 다양한 변수 존재)

2. ODEint보다 Euler Method로 구한 값이 더 정확해 보이는 이유?

그냥 그렇게 보일 뿐,

Real Data는 감염자(코로나19에 감염된 사람) 수가 아닌

확진자(코로나 확진 '판정'을 받은 사람)를 구한 것이므로,

실제 코로나19에 감염된 사람 수는 사실은 ODEint에 좀 더 가까울 것!