

---

# MODELOWANIE MATEMATYCZNE I SYMULACJE KOMPUTEROWE

---

ANNA SZAFRAŃSKA

PROJEKT 2



### Informacje praktyczne

Ocenie podlega część symulacyjna oraz opisowa. Projekt należy załączyć na kursie przedmiotu za pomocą odpowiedniej aktywności w składowej *Projekt*.

Poniżej znajdują się szczegółowe informacje dotyczące wykonania obu części projektu.

#### 1. Część symulacyjna.

- Każdy kod załączony na eNauczaniu ma być **podpisany**.
- Proszę umieścić w kodzie wyraźne podziały na poszczególne części zadania.
- Należy sprawdzić kompilacje każdego fragmentu kodu, zwrócić uwagę na nazwy zmiennych (aby się nie nakładały).
- Nazwa pliku symulacyjnego: *ImieNazwisko\_Projekt1.rozszerzenie*.

#### 2. Część teoretyczna.

- Sprawozdanie powinno zawierać: imiona i nazwiska osób wchodzących w skład grupy pracujących nad danym projektem, specjalność, data wykonania.
- Sprawozdanie z projektu opisujemy w specjalistycznym środowisku matematycznym TeX lub LaTeX, i na kursie załączamy plik PDF o nazwie *ImieNazwisko\_Projekt1.pdf*.
- Proszę upewnić się, że w sprawozdaniu w każdym podpunkcie zamieszczone są wymagane informacje. Poza tym proszę umieścić wnioski i/lub opisy uzyskanych wyników czy wykresów. Proszę nie zapomnieć o podpisaniu każdego wykresu (w szczególności dla jakich parametrów został wykonany).
- Należy pamiętać, że ocenie podlega również sposób wykonania sprawozdania, więc nie należy zamieszczać “suchych” wyników lub wykresów.
- Do każdego punktu załączamy odpowiadający mu kod programu. Cały kod z opisem umieszczamy na końcu sprawozdania.

**Podstawowa literatura, przydatne pliki i linki:** podane na stronie kursu w module *Projekt*.

### Deterministyczne modele kompartmentowe - model SIR z demografią

Rozważamy matematyczny model dynamiki populacji otwartej (czyli bierzemy pod uwagę urodzenia i zgony osobników populacji).

Założmy, że mamy populację o ustalonej wielkości  $N$ , zatem urodzenia i zgony osobników populacji równoważą się. W sytuacji rozprzestrzeniania się infekcji populacja dzieli się na trzy grupy: podatnych (**S**usceptible), zainfekowanych (**I**nfected), wyzdrowiałych (**R**ecovered). W danym momencie czasu  $t$  mamy, że:

$$\begin{aligned} S(t) &= \text{liczbie osób jeszcze nie zainfekowanych ale podatnych,} \\ I(t) &= \text{liczbie osób zainfekowanych,} \\ R(t) &= \text{liczbie osób, które miały infekcję ale wyzdrowiały.} \end{aligned}$$

Zakładamy, że jeśli osobnik miał infekcję i wyzdrowiał, to nie może zarazić się ponownie. W każdym momencie czasu mamy zależność

$$S(t) + I(t) + R(t) = N, \quad (1)$$

która implikuje równość

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0. \quad (2)$$

Przy modelowaniu przyjmujemy następujące założenia

- Osobnicy podatni (grupa  $S$ ) mogą zostać zainfekowani w każdym momencie czasu, ich liczebność w danym momencie czasu zależy od liczby osób zmarłych, urodzonych i tych, którzy zostali zainfekowani (przeszli do grupy  $I$ );
- Osobnicy zainfekowani mogą z czasem wyzdrowieć (przejsć do grupy  $R$ ) lub umrzeć;
- Osobnicy wyzdrowiali nie mogą wrócić do grupy  $S$ , mogą pozostać w grupie  $R$  lub umrzeć.

Powyższe założenia wraz z (2) dają kompartmentowy model opisany układem równań różniczkowych

$$\begin{cases} \frac{d}{dt}S(t) = \mu N - \frac{\beta}{N}S(t)I(t) - \mu S(t), \\ \frac{d}{dt}I(t) = \frac{\beta}{N}S(t)I(t) - \gamma I(t) - \mu I(t), \\ \frac{d}{dt}R(t) = \gamma I(t) - \mu R(t) \end{cases} \quad (3)$$

gdzie  $\beta > 0$  współczynnik przejścia z grupy podatnych do grupy zainfekowanych,  $\gamma > 0$  współczynnik zdrowienia,  $\mu > 0$  współczynnik narodzin i jednocześnie śmiertelności (tyle ile osobników się narodzi, tyle umiera w każdej chwili czasu  $t$ ). Zakładamy również, że naturalna długość życia osobnika populacji wynosi  $\frac{1}{\mu}$  roku.

Korzystając z własności (1) zauważmy, że wystarczy rozpatrzyć model

$$\begin{cases} \frac{d}{dt}S(t) = \mu N - \frac{\beta}{N}S(t)I(t) - \mu S(t), \\ \frac{d}{dt}I(t) = \frac{\beta}{N}S(t)I(t) - \gamma I(t) - \mu I(t), \\ R(t) = N - S(t) - I(t). \end{cases} \quad (4)$$

W poniższych zadaniach będziemy przyjmować, że  $t \in [0, T]$ , gdzie dla  $t = 0$  mamy warunek początkowy modelu

$$S(0) = S_0, \quad I(0) = I_0.$$

Analizie poddamy rozwiązanie numeryczne powyższego problemu różniczkowego uzyskane nie-standardową metodą typu Mickensa'a.

— ★ —

### Analiza zachowania długoterminowego - symulacje rozwiązań

Dla modeli epidemiologicznych definiujemy tzw. podstawowy współczynnik reprodukcji  $R_0$  (*ang. basic reproduction number*), który można traktować jako oczekiwaną liczbę przypadków zakażeń bezpośrednio generowanych przez jednego chorego osobnika w populacji, w której wszystkie osobniki są podatne na zakażenie. Definicja współczynnika  $R_0$  odnosi się do sytuacji kiedy żadne inne osoby w populacji nie są zarażone ani zaszczepione.

Współczynnik  $R_0$  dla modelu SIR z demografią definiuje się wzorem

$$R_0 = \frac{\beta}{\mu + \gamma}.$$

Według wartości  $R_0$  możemy określić, czy dla konkretnych parametrów epidemia wybuchnie ( $R_0 > 1$ ), czy też szybko wygaśnie nie rozprzestrzeniając się ( $R_0 < 1$ ).

Ponadto model posiada dwa punkty równowagi:

- punkt równowagi **wolny od choroby** - STABILNY, gdy  $R_0 < 1$ ,
- punkt **endemiczny** - STABILNY, gdy  $R_0 > 1$ .

**Zadanie 1.** Wyznaczyć punkty równowagi modelu (3) (można również skorzystać z modelu (4)). Postać punktów zapisać w uproszczonej postaci, korzystając z definicji  $R_0$ .

**Zadanie 2.** Proszę przyjąć dowolną liczebność populacji  $N$  oraz  $T$  równe **co najmniej** 1000 dni (około 3 lat). Przyjąć krok obliczeniowy  $h$  (można wykonać poniższe punkty dla różnych wartości kroku czasowego), wówczas ilość wykonanych iteracji metody to  $n = T/h$ .

1. Wyprowadzić metodę niestandardową typu Mickensa dla modelu (4).
2. Dla każdego z zestawu parametrów

$$a) (S_0, I_0, \beta, \gamma, \mu) = (N - 1, 1, 0.5, 0.8, 0.4),$$

$$b) (S_0, I_0, \beta, \gamma, \mu) = (N - 1, 1, 0.2, 0.04, 0.01),$$

$$c) (S_0, I_0, \beta, \gamma, \mu) = (N - 1, 1, 0.8, 0.4, 0.01),$$

$$d) (S_0, I_0, \beta, \gamma, \mu) = (N - 1, 1, 0.8, 0.7, 0.01),$$

- wyznaczyć punkt równowagi oraz wartość  $R_0$ ;
- przedstawić na jednym wykresie rozwiązanie numeryczne modelu  $(S, I, R)$ ;
- wyznaczyć portrety fazowe dla par  $I \sim S$ ,  $S \sim R$ ,  $I \sim R$  wraz z zaznaczonymi punktami początkowymi oraz punktami równowagi (proszę zwrócić uwagę na odpowiedni dobór parametrów numerycznych, głównie  $T$  i  $h$ , aby osiągnąć punkt równowagi);
- wyznaczyć trójwymiarowy portret fazowy  $S \sim I \sim R$ . W programie R można skorzystać z funkcji `plot_ly` zawartej w pakiecie `plotly`, np.
 

```
install.packages("plotly")
library("plotly")
plot_ly(dane, x=~S, y=~I, z=~R, mode = 'lines', opacity = 1).
```

Co udało się Państwu zaobserwować z powyższych symulacji? Co można powiedzieć o otrzymanych wynikach z punktu widzenia doboru parametrów?

**Zadanie 3.** Kolejno dla zestawów parametrów z 2.2(b) – 2.2(d) wykonać symulacje grupy osobników zarażonych ( $I$ ) dla różnych warunków początkowych (co najmniej trzy kombinacje) i przedstawić wyniki na jednym wykresie. Zakładamy, że  $R(0) = 0$  (na początku epidemii nie mamy ozdrowieńców), zatem manipulujemy warunkami  $S_0$  i  $I_0$ , tak aby zawsze zachodziło  $I_0 + S_0 = N$ .

Czy uzyskane wyniki (trajektorie rozwiązania) zmieniają swój kształt w sposób wpływający na interpretację?

**Zadanie 4.** Przyjąć dowolną liczebność populacji  $N$ . Wyznaczyć rozwiązania modelu przeskalowane do liczebności populacji  $N = 1$ , wówczas rozwiązania modelu będą się mieścić (niezależnie od parametrów) w przedziale wartości  $[0, 1]$ .

Czy dla zestawów parametrów z Zadania 2.2 mamy tą samą charakterystykę trajektorii?

Dla jednego wybranego zestawu parametrów z Zadania 2.2 sprawdzić kształt trajektorii dla różnych liczebności populacji.

— ★ —

### Estymacja parametrów

Zainstalować pakiet `epimdr` zawierający potrzebny plik z danymi. Pakiet posiada również wiele funkcji odpowiadających różnym modelom epidemiologicznym oraz funkcje pomocne przy ich analizie.

```
install.packages("epimdr")
library("epimdr")
```

Rozważamy plik z danymi `niamey`, mieszczący się w pakiecie `epimdr`, który zawiera odnotowane cotygodniowo przypadki odry w trakcie wybuchu epidemii w Niamey w Nigerii na przełomie lat 2003/2004. Dla osób piszących w Pythonie, plik z danymi `Niamey.csv` zamieszczony jest na stronie kursu.

Plik zawiera między innymi informacje o liczbie zarażonych z odstępem tygodniowym (od 45 tygodnia 2003 r. do 23 tygodnia 2004 r. - 31 tygodni (zmienna `week`)) dla trzech dystryktów (zmienne `cases_1`, `cases_2`, `cases_3`).

Posiadane dane dotyczą trzech różnych dystryktów jednego miasta, dla których nie znamy liczebności populacji, znamy natomiast orientacyjną liczebność populacji w Niamey w latach 2003/2004, która wynosi  $N = 769454^1$ . Zatem w zadaniu będziemy rozważać **dane jako suma** informacji ze wszystkich trzech dystryktów.

Jako warunki początkowe dla populacji zainfekowanych  $I_0$  przyjmujemy pierwszy element ze zbioru danych, natomiast  $S_0 = N - I_0$ .

Średni wiek życia w Niamey to zaledwie 15 lat, więc przyjmujemy parametr

$$\mu = 1/(15 \cdot 365).$$

Zadaniem będzie zatem oszacowanie parametrów

$$(\beta, \gamma),$$

**Zadanie 5.** Wczytać z pakietu/pobrać plik z danymi. Zsumować dane z trzech dystryktów (zmienne `cases_1`, `cases_2`, `cases_3`) i przedstawić je na wykresie w przeciągu 31 **tygodni**.

**Zadanie 6.** Dla zestawu parametrów

$$(\beta, \gamma) = (0.79, 0.7)$$

wyznaczyć współczynnik  $R_0$  oraz rozwiązanie numeryczne modelu. Na jednym wykresie z danymi przedstawić uzyskane rozwiązanie numeryczne dla grupy zarażonych ( $I$ ).

**Zadanie 7.** Korzystając z metody estymacji najmniejszych kwadratów wyznaczyć  $(\hat{\beta}, \hat{\gamma})$ .

W tym celu minimalizujemy funkcję celu

$$BNK(p) = \sum_{j=1}^{31} \left( (I_{dane})_j - (I_{num})_j \right)^2,$$

gdzie  $p$  jest wektorem parametrów do estymacji,  $I_{dane}$  to odnotowane dane o przypadkach zarażonych, a  $I_{num}$  to rozwiązanie numeryczne dla grupy zarażonych  $I$ .

Dla minimalizacji funkcji  $BNK$  przyjąć warunki początkowe parametrów

$$(\beta, \gamma) = (0.79, 0.7).$$

W związku z tym, że parametry muszą być dodatnie, to rozważamy minimalizację z ograniczeniami, szukamy rozwiązania dla  $(\beta, \gamma) \in [0, 1]^2$ .

Minimalizację z ograniczeniami wykonać za pomocą dwóch algorytmów:

---

<sup>1</sup>R.F. Grais i reszta, *Time is of the essence: exploring a measles outbreak response vaccination in Niamey, Niger*, na stronie kursu

- *L-BFGS-B*, który w *R* jest metodą funkcji `optim()`; w *Pythonie* algorytm jest dostępny w funkcji `scipy.optimize.minimize()`;
- *NLOPT COBYLA*, który w *R* jest algorytmem funkcji `nloptr()` jako `NLOPT_LN_COBYLA`; w *Pythonie* algorytm jest dostępny w pakiecie `nlopt`;

*Przedstawić na wykresie wyniki działania obu algorytmów na tle danych.*

**Zadanie 8.** *Przeprowadzić analizę zmian wartości funkcji celu w zależności od doboru wartości korku czasowego metody numerycznej, wyciągnąć wnioski.*