

Computational Physics

Abschlussprojekt: Epidemieausbreitung

Abgabe bis einschließlich 27.07.2021

1 Einführung

1.1 Motivation

Die aktuelle Corona Pandemie demonstriert eindrücklich die ständige Gefahr, die von Infektionskrankheiten ausgehen. Faktoren wie Bevölkerungswachstum, Globalisierung und Klimawandel führen in der modernen Welt zur immer schnelleren und weitflächigeren Verbreitung von solchen Krankheiten. Deshalb ist es sehr wichtig, die Dynamik von Infektionskrankheiten in der Bevölkerung gut zu verstehen, um so aus der Kombination aller zur Verfügung stehenden Maßnahmen sinnvolle Strategien zur Eindämmung von Infektionskrankheiten zu erarbeiten.

In diesem Projekt soll die Dynamik von Epidemien auf einem Netzwerk simuliert werden. Die Knoten des Netzwerks werden durch Städte gebildet. Diese sind durch Kanten verbunden, auf denen Reiseverkehr stattfindet. Durch den Reiseverkehr durchmischen sich die Populationen. Die Infektionskrankheit wird mit dem SIR (Susceptible-Infected-Recovered) Modell beschrieben. Die räumliche Ausbreitung der Krankheit ergibt sich aus der Kombination des SIR Modells, das jeweils auf eine Stadt beschränkt ist, mit dem Reiseverkehr auf dem Netzwerk. Infizierte Personen reisen in eine Stadt, die bisher frei von Infektionen war, und stecken die dortig Bevölkerung mit der Krankheit an.

1.2 SIR Modell

Das SIR Modell beschreibt die Ausbreitung einer Infektionskrankheit in einer Bevölkerung. Das Modell funktioniert analog zum Lotka-Volterra Modell zur Beschreibung von Jäger-Beute Populationen. Die Differentialgleichungen sind gegeben durch

$$\frac{dS}{dt} = -\beta \frac{SI}{N} + \mu(N - S), \quad (1)$$

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I - \mu I, \quad (2)$$

$$\frac{dR}{dt} = \gamma I - \mu R, \quad (3)$$

mit $N(t) = S(t) + I(t) + R(t)$. Es berücksichtigt dabei die folgende Dynamik:

- Die S-Population wird von der I-Population mit der Krankheit infiziert. Die Infektionsrate ergibt sich aus dem Produkt der Anzahl der infizierbaren Personen, einem Ansteckungsparameter β und dem Anteil der Infizierten an der Gesamtbevölkerung. β ist charakteristisch für die Krankheit und ergibt sich aus der Wahrscheinlichkeit sich bei Kontakt mit einer infizierten Person anzustecken und der Anzahl entsprechender Kontakte.
- Individuen der I-Population genesen nach einer gewissen Zeit und sind danach gegen die Krankheit immun, gehen also in die R-Population über. Die Erholungsrate γ entspricht der inversen Dauer der Krankheit.
- Ab und zu sterben Mitglieder eines natürlichen Todes und durch Reproduktion kommen neue hinzu. Im einfachen SIR Modell ist die Sterberate gleich der Reproduktionsrate μ . Es wird davon ausgegangen, dass neue Populationsmitglieder gesund und nicht immun sind (das ist nicht für alle Krankheiten korrekt).

Eine nützliche Kenngröße des obigen SIR Modells ist die dimensionslose Basisreproduktionszahl R_0 , die angibt, wie viele (suszeptile) Personen im Mittel von einer infektiösen Person angesteckt werden. Hier ist die Basisreproduktionszahl gegeben durch

$$R_0 = \frac{\beta}{\mu + \gamma}. \quad (4)$$

Für $R_0 < 1$ gehen die Infektionen langfristig zurück, wohingegen $R_0 > 1$ einen kontinuierlichen Anstieg der Infektionen bedeutet.

Sie können das Modell im Rahmen der weiterführenden Aufgaben beliebig erweitern, für die Basisaufgaben soll jedoch nur diese einfache Version verwendet werden. Das SIR Modell und mögliche Erweiterungen werden in den folgenden Artikeln behandelt:

http://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology

http://en.wikipedia.org/wiki/Epidemic_model

<http://loe.org/images/content/091023/Zombie%20Publication.pdf>

1.3 Reiseverkehr auf dem Netzwerk

Da die zur Verfügung stehende Rechenleistung begrenzt ist, wird ein sehr einfaches Reisemodell verwendet. Alle Reisen laufen quasi instantan ab, d.h. in jedem Zeitschritt erfolgt eine sofortige Durchmischung, Ansteckungen in Verkehrsmitteln werden nicht extra berücksichtigt. Der Populationsaustausch durch Reisen ergibt sich durch den Reisestrom, der für jede Stadt berechnet wird, aufgeteilt auf die SIR Populationen. Der Reisestrom ergibt sich u.a. aus der Reisewahrscheinlichkeit p_t , der Population in der jeweiligen Stadt und der Population in den möglichen Reisezielen. Warum muss hier auch die Population der Reiseziele berücksichtigt werden? Welche Ströme heben sich auf und welche nicht?

1.4 Zeitentwicklung

Das Modell soll so ausgelegt werden, dass ein Zeitschritt genau einem Tag entspricht. Es wird also davon ausgegangen, dass alle Reisen einen Tag dauern und die Parameter des SIR Modells müssen entsprechend gewählt werden.

2 Aufgaben und Leitfragen

2.1 Basisaufgaben

a) Modell

Machen Sie sich mit dem Modell vertraut und überlegen Sie sich, wie realistisch es ist. Was wird korrekt abgebildet und was nicht?

b) Netzwerk und SIR Modell

Zunächst müssen Sie das Netzwerk für die Simulation erstellen. Schreiben Sie dazu eine Routine, die alle Städte mit mehr als $2 \cdot 10^5$ Einwohnern aus der bereitgestellten CSV-Datei „europe.csv“ einliest. Beachten Sie, dass die Datei Symbole aus dem UTF-8 Zeichensatz enthält. Wenn Sie beispielsweise eine numpy Methode zum Einlesen verwenden, übergeben Sie der Methode die richtige Kodierung indem Sie `encoding='utf-8'` als Parameter übergeben. Simulieren Sie den Landverkehr zwischen den Städten, indem Sie jede Stadt zu allen anderen Städten im 500 km Umkreis verbinden. Den Flugverkehr simulieren Sie durch direkte Verbindungen zwischen allen Städten mit 10^6 oder mehr Einwohnern. Da die Berechnung der Abstände aller Städte recht zeitaufwändig ist, empfiehlt es sich, das Ergebnis in einer Datei zu speichern und beim Starten des Programms die Datei einzulesen. Weiterhin ist es nützlich, **numpy** Arrays zu verwenden, um Operationen zu vektorisieren.

- Implementieren Sie Module, um die Änderung der Populationen in einer Stadt aufgrund der Reisebewegungen zu berechnen.
- Implementieren Sie drei Module, die die Änderungen der Populationen in einer Stadt aufgrund des SIR Modells berechnen, jedes mit einer anderen Strategie zur Lösung der Differentialgleichungen. Verwenden Sie die folgenden Verfahren:
 - ein Schritt des expliziten Euler Verfahrens,
 - ein Schritt des Runge-Kutta Verfahrens 4. Ordnung,
 - die Funktion `ode` (oder `odeint`) aus dem Modul `scipy.integrate`.
- Fassen Sie diese Routinen zusammen, um sie auf das gesamte Netzwerk anzuwenden. Halten Sie sich beim Ausführen eines Zeitschritts an die Reihenfolge Landverkehr, Luftverkehr, SIR Modell.

c) Graphische Darstellung

Stellen Sie die Ergebnisse graphisch auf einer Europakarte dar. Dafür können

Sie z.B. `matplotlib` in Verbindung mit dem Objekt `Basemap`¹ aus dem Modul `mpl_toolkits.basemap` nutzen (siehe Abbildungen 1 und 2 für zwei Beispiele). Als Alternative zum `Basemap` package kann z.B. auch `folium` verwendet werden, welches interaktive Karten ermöglicht.

d) **Test der Implementierung**

Anhand einiger einfacher Tests können Sie feststellen, ob Ihre Implementation grobe Fehler enthält:

- Wie verhält sich das Modell, wenn es keine infizierten Personen gibt?
- Können Sie anhand der Differentialgleichungen Grenzfälle für bestimmte Parameter erkennen? Können Sie diese in Ihrem Modell bestätigen?
- Betrachten Sie die Änderung der Gesamtbevölkerung N einer Stadt die durch Reisebewegungen verursacht wird, wenn es keine Infizierten gibt. Setzen Sie dazu die Reisewahrscheinlichkeit auf $p_t = 0.01$. Was beobachten Sie nach 30 Zeitschritten? Nach 365?
- Starten Sie eine Krankheit mit $\beta = 0.1 \text{ Person}^{-1} \text{ Tag}^{-1}$ und $\gamma = 0.07 \text{ Person}^{-1} \text{ Tag}^{-1}$ (um was für eine Krankheit könnte es sich handeln?) bei einer Reproduktionsrate von $\mu = 3 \cdot 10^{-5} \text{ Person}^{-1} \text{ Tag}^{-1}$ und einer Reisewahrscheinlichkeit von $p_t = 0.01$ mit 1000 Infizierten in London. Führen Sie 30 Zeitschritte aus. Führen Sie die Simulation für jede der drei Implementierungen des SIR Modells durch. Bereiten Sie sich darauf vor, beim Testat nach einzelnen Zahlenwerten dieser Simulation gefragt zu werden.
- Untersuchen Sie das Verhalten der drei Implementierungen des SIR Modells. Worin liegen die Stärken und Schwächen der Verfahren, wo ihre Grenzen? Welche der drei Implementierungen halten Sie für die geeignetste?

2.2 Weiterführende Untersuchungen

Teil e) muss auf jeden Fall bearbeitet werden, aus Teil g) sollten Sie mindestens zwei Punkte bearbeiten.

e) **Maßnahmen zur Eindämmung von Epidemien**

Beim Ausbrechen von Epidemien werden verschiedene Maßnahmen ergriffen, um die Ausbreitung zu verhindern. Je nach betroffener Spezies können diese Maßnahmen sehr drastisch ausfallen. Implementieren Sie ein paar Maßnahmen zur Eindämmung von Infektionen wie Quarantäne, Informations- oder Impfkampagnen, Reisebeschränkungen oder Flughafenschließungen. Diskutieren und bewerten Sie diese. Wie stark müssen die Maßnahmen umgesetzt werden, um einen Effekt zu erzielen?

¹<http://matplotlib.org/basemap/>

f) **Erweiterung des SIR Modells**

Implementieren sie tödliche Krankheiten und Inkubationszeit. Was ändert sich? Stellen Sie die Ergebnisse graphisch dar.

g) **Anregungen für weiterführende Untersuchungen**

- Betrachten Sie andere Krankheiten. Versuchen Sie aus den zugänglichen Informationen sinnvolle Parameter abzuleiten.
- Variieren Sie die Krankheitsparameter β und γ mit der Zeit, um saisonal auftretende Infektionen (Grippe) zu beschreiben.
- Erweitern Sie das Modell um Impfungen.
- Ergänzen Sie das einfache Reisemodell so, dass die Summe N der Populationen in einer Stadt zeitlich konstant ist. Fügen Sie dazu zusätzliche Reiseströme ein, welche die ursprünglichen Reiseströme ausgleichen.
- Machen Sie die Reisewahrscheinlichkeit vom Gesundheitszustand abhängig.
- Erstellen Sie ein weiteres Netzwerk für z.B. Nordamerika (siehe zusätzliche CSV-Dateien) und verbinden Sie es über große Flughäfen ($> 2 \cdot 10^6$ Einwohner) mit Europa.
- Modellieren Sie die „Zombieapokalypse“ (siehe Material im Abschnitt 1.2).
- Bringen Sie eigene Ideen ein.

3 Nomenklatur

Halten Sie sich für die Basisaufgaben und Tests unbedingt an die vorgegebene Nomenklatur, andernfalls wird das Projekt nicht testiert. Für die weiterführenden Fragen müssen Sie weitere Variablen einführen.

Die folgenden Variablen sollen global definiert werden. Übergeben Sie diese nicht in Routinen, sondern greifen Sie in den Routinen direkt auf die globalen Variablen zu. Speichern Sie die Eigenschaften des Netzwerks in (**numpy**) Arrays, der Index identifiziert dabei die Stadt. Verwenden Sie die folgenden Variablen:

<code>cities</code>	Namen der Städte
<code>population</code>	Bevölkerungszahl
<code>landConnections</code>	Liste der Landverbindungen zu anderen Städten
<code>airConnections</code>	Liste der Flugverbindungen zu anderen Städten

Legen Sie für jede SIR Population zweidimensionale Arrays an. Der erste Index repräsentiert die Evolutionschritte, der Zweite die Städte. Benennen Sie die Variablen mit **sus**, **inf** und **rec**.

Verwenden Sie für die restlichen Parameter die folgenden Bezeichnungen:

<code>pt</code>	Reisewahrscheinlichkeit p_t
<code>beta</code>	Kontakt-/Erkrankungsrate β der Krankheit
<code>gamma</code>	Erholungsrate γ der Krankheit
<code>mu</code>	Reproduktions- und Sterberate μ

Verwenden Sie die folgenden Namen für Ihre Routinen:

<code>travelLand</code>	Landreiseverkehr für eine Stadt
<code>travelAir</code>	Luftreiseverkehr für eine Stadt
<code>travelLandAll</code>	Landreiseverkehr für alle Städte
<code>travelAirAll</code>	Luftreiseverkehr für alle Städte
<code>infectRK4</code>	SIR Modell mit Runge-Kutta Verfahren für eine Stadt
<code>infectEuler</code>	SIR Modell mit Euler Verfahren für eine Stadt
<code>infectODEsolver</code>	SIR Modell mit <code>ode/odeint</code> für eine Stadt
<code>infectRK4All</code>	SIR Modell mit Runge-Kutta Verfahren für alle Städte
<code>infectEulerAll</code>	SIR Modell mit Euler Verfahren für alle Städte
<code>infectODEsolverAll</code>	SIR Modell mit <code>ode/odeint</code> für alle Städte

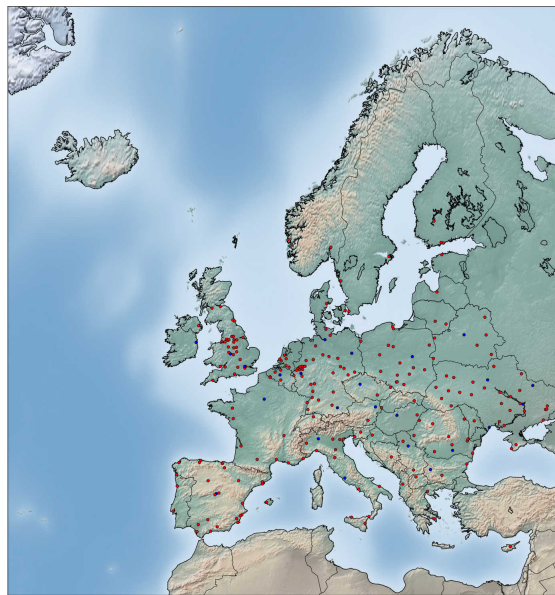


Abbildung 1: Karte von Europa. Rote Punkte kennzeichnen Städte mit mehr als $2 \cdot 10^5$ Einwohnern, blaue Punkte Städte mit mehr als 10^6 Einwohnern.

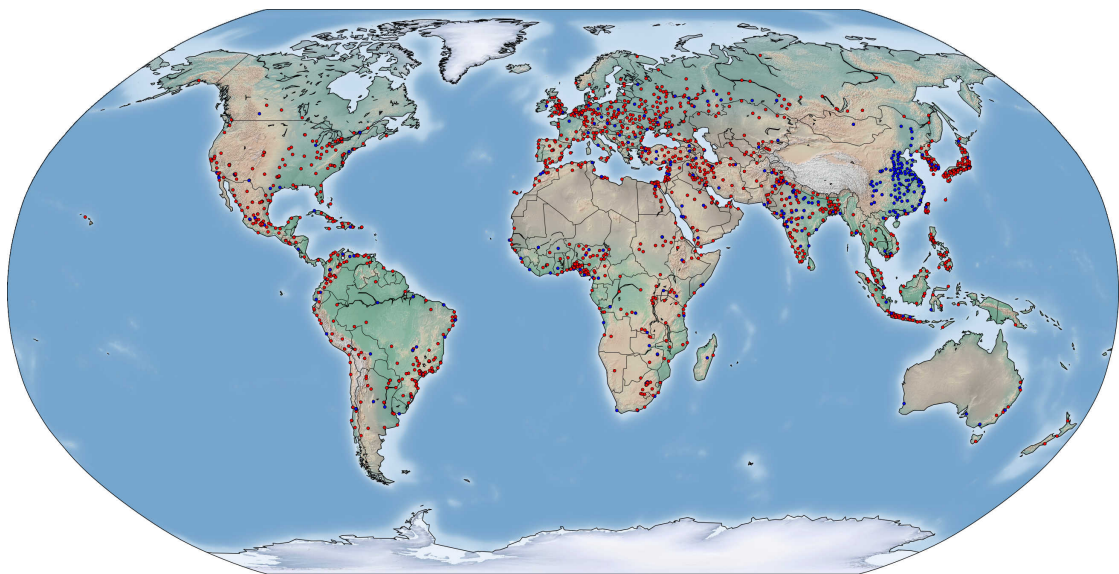


Abbildung 2: Weltkarte. Notation wie in Fig. 1.