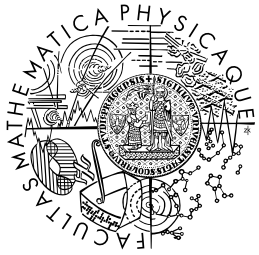

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Josef Martínek

Parameter optimization in COVID-19 epidemiological models

Department of Numerical Mathematics

Supervisor of the bachelor thesis: doc. RNDr. Václav Kučera, Ph.D.
Study programme: Mathematics
Study branch: General Mathematics

Prague 2021

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ....

Author's signature

I would like to express my gratitude to my supervisor doc. Václav Kučera for his continuous support and valuable advice.

Title: Parameter optimization in COVID-19 epidemiological models

Author: Josef Martínek

Department: Department of Numerical Mathematics

Supervisor: doc. RNDr. Václav Kučera, Ph.D., Department of Numerical Mathematics

Abstract: This work is concerned with modelling of the spread of infectious diseases with emphasis on the current COVID-19 pandemic. Our goal is to estimate unknown parameters in epidemiological models from real data on the spread of the disease in the Czech Republic. To model the evolution of the epidemic, we consider compartmental models, which lead to a system of ordinary differential equations. We then formulate a non-linear least squares problem for the optimization of the model parameters to fit the model outcome to the observed data. We numerically optimize by the Levenberg–Marquardt method, which requires the Jacobian of the vector of residuals. This is obtained by deriving and solving the sensitivity equations corresponding to the considered model. We test the method on noisy artificial data and on a well documented English boarding school influenza epidemic. Finally, we apply the method to Czech COVID-19 data and discuss the results. One of the conclusions of this work is the introduction of the concept of effective population size, to overcome the unrealistic assumption of complete homogeneity of the population. Thus the population size is not apriori given, but is an unknown parameter to be optimized. This leads to much better agreement of the models and real data. This appears to be a new concept.

Keywords: COVID-19, parameter optimization, non-linear least squares, epidemiological models, effective population size

Contents

Introduction

Mathematical modelling has become a real focus of interest lately due to the global pandemic of COVID-19. Epidemiological modelling, formerly studied only by a handful of experts, is now in the center of attention of the general public. The scientific results and predictions are one of the fundamental sources of information for adopting restrictive measures against the spread of the disease.

There are many approaches for addressing the problem of modelling infectious diseases. A variety of stochastic or discrete time models have been used to treat this problem, cf. [1], [2] and [3] for an overview. The focus of this work is on the standard compartmental models, which is presumably the most widely used category of epidemiological models. Due to the COVID-19 pandemic, even many non-experts have heard of the simplest compartmental model – the SIR model – introduced in the pioneering work of Kermack and McKendrick in 1927, [4]. However, the aim of the work lies beyond the classical approach of numerical simulations of various models. The compartmental models depend on several parameters characterizing the epidemic, estimated usually by means of medical research. Our goal is to estimate the parameters from observational epidemiological data using some numerical methods.

Mathematically, compartmental models are formulated as a system of ordinary differential equations for the number of individuals in each epidemiological category. These systems of ordinary differential equations contain parameters which must be tuned so that the outcome of the model best fits the observed data. This problem thus fits into the broader framework of numerical data fitting, cf. [5]. Specifically, we consider a nonlinear least squares formulation of the parameter estimation problem, where we minimize a least squares functional measuring the discrepancy between the model outcome and measured data. From the many possible approaches to tackle such a problem, in this thesis we choose a technique based on the so-called sensitivity equations, cf. [6] and [5]. These differential equations describe how the solution of the original system of ordinary differential equations depends on the chosen parameters. This in turn allows us to optimize with respect to these parameters using standard numerical techniques such as the Gauss-Newton or Levenberg-Marquardt methods, cf. [7].

One of the major issues of the standard epidemiological models we encounter in this thesis lies in the unrealistic assumptions made both on the population and on the disease. The models are derived providing complete homogeneity of the population, which is clearly not satisfied in practice. To deal with this problem, a concept of the effective population size reflecting the assumptions of the models is introduced in this thesis. Effectively, we treat the population size as an unknown parameter rather than an apriori given constant. This approach appears to improve significantly the accuracy of the models.

The thesis is divided into four chapters. In the first chapter, the compartmental models are introduced and a commentary on their assumptions is given. The second chapter provides the reader with the necessary mathematical background of the parameter optimization algorithms – selected topics on the theory of ordinary differential equations with an emphasis on the derivation of the sensitivity equations. In the third chapter, the nonlinear least squares problem is introduced

and the optimization algorithms themselves are described along with a method for the numerical solution of the ordinary differential equations. The program implemented for the numerical algorithms is introduced as well. Finally, the fourth chapter presents the results of these methods for various epidemiological data along with an interpretation and discussion of the results.

1. Epidemiological models

1.1 Introduction

The origin of mathematical modelling in epidemiology dates back to the second half of the 18th century when the Swiss mathematician and physicist Daniel Bernoulli studied and mathematically analysed the increase in life expectancy caused by inoculation against smallpox. In his paper, published in 1776, he presented the earliest mathematical model of this particular disease. However, predictive modelling was not given any special attention until the beginning of the 20th century. A great contribution in this direction was made by Kermack and McKendrick who published a paper in 1927 [4] in which they described the so called *compartmental models*, the models we use in this thesis. Epidemiological modelling went through a dramatic expansion in the second half of the 20th century, but it has become a real focus of attention recently owing to the pandemic of COVID-19.

Mathematical models in epidemiology may be sorted into various categories according to different criteria – discretisation of time (models with discrete intervals and continuous time models), allowing for randomness (stochastic and deterministic models), structure of the population etc. In this thesis we take into account exclusively *deterministic*, *continuous time* models and the population is assumed to be a *homogeneous continuum*. Presumably the most widely known representatives of this kind of models are the standard compartmental models, some of which are derived and described in this chapter.

The compartmental models are based on the principle of dividing the population into several labeled *compartments* (eg. Infectious, Recovered etc.) under certain simplifying assumptions. The development of the epidemic in the population is then determined by certain relations describing the flow between the compartments. Each relation indicates the rate of flow between a pair of compartments. The model is formulated mathematically by means of a *system of ordinary differential equations*. In the subsequent sections we address some individual models, which will be later used in practice.

1.2 SIR model

In this section, we introduce the *SIR model*, which is the basis for the more sophisticated models we use. This simple model can be used to illustrate some of the fundamental principles. We do not present here the analysis of the model from the perspective of the theory of ordinary differential equations, see [2] for further information.

1.2.1 Derivation

In this section, we follow the book [1]. In order to derive the SIR model, it is necessary to make some simplifying assumptions regarding both the population and the spread of the disease. On one hand, the model must be sufficiently simple for practical application, on the other hand, the model should take into account

the specificities of the disease in question. Let us describe the compartments of the SIR model and the flow between them. Let $T > 0$. The epidemic is considered on the time interval $[0, T]$. The population is divided into three groups, each group a function of time:

- *Susceptible* (S) — those who have not come across the disease and can fall ill if they come into contact with an infectious person. Afterwards, they become infectious themselves.
- *Infectious* (I) — those who spread the disease among the susceptible population. After recovery they move to the compartment R:
- *Recovered* (R) — those who are removed from the compartment I either due to recovery or due to death.

The relations between the presented compartments are based on four fundamental assumptions:

1. The vital dynamics is neglected and the size of the population is supposed to be constant, we denote it by N, $N > 0$.
2. The population is assumed to be a homogeneous continuum, i.e. all people have an equal number of contacts, the probability of the transmission of the disease between a susceptible and an infectious person during their contact remains constant and the infectious people are equally distributed among the population.
3. The rate of flow between the compartments I and R is directly proportional to the size of the compartment I.
4. The recovered people acquire immunity and cannot spread the infection. Those who fall victims to the disease are treated as recovered.

Let us denote by r the number of contacts of a person per unit time and let $p \in (0, 1)$ be the probability of the transmission between an infectious and a susceptible person when they meet. It is desired to find the number of people an infectious person infects per unit time. The fraction of susceptible population within the total population is $\frac{S}{N}$. Therefore, the infectious person meets a total of $r\frac{S}{N}$ susceptible people per unit time. It follows that the number of infected susceptible people per infectious person per unit time is $pr\frac{S}{N}$. It proves convenient to define a new constant $\beta = pr$. Because the total number of infectious people is equal to I, it can be concluded that the total number of infected people per unit time is $\beta I\frac{S}{N}$.

We now determine the relation between compartments I and R. As stated in the assumption 2, the rate of flow between the compartments I and R is directly

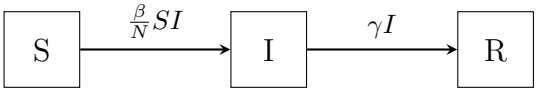

Figure 1.1: SIR model

proportional to the size of the compartment I. Denote by γ the coefficient of proportionality. The rate of flow is then equal to γI. For visualization of the compartments and relations between them, see Figure 1.1.

In order to complete the derivation of the model, it remains to define the initial conditions. Let $I_0 > 0$ and $R_0 \geq 0$ such that $N - R_0 - I_0 > 0$. We set

$$\begin{aligned} S(0) &= N - R_0 - I_0, \\ I(0) &= I_0, \\ R(0) &= R_0. \end{aligned} \tag{1.1}$$

It is clear that there are no recovered people at the beginning of the epidemic, i.e. $R_0 = 0$. However, the model may not always be applied from the beginning of an epidemic and in this case, we allow $R_0 > 0$.

The development of the model is described formally by a system of ordinary differential equations. In conclusion, we obtain the SIR model in the form of an initial value problem

$$\begin{aligned} S' &= -\frac{\beta}{N} SI, \\ I' &= \frac{\beta}{N} SI - \gamma I, \\ R' &= \gamma I, \end{aligned} \tag{1.2}$$

with the initial conditions given by (1.1). The resulting model is shown schematically in Figure 1.1.

As presented in [1], the value $\frac{1}{\gamma}$ is equal to the expected amount of time spent in the compartment I. We define the basic reproduction number $R_0 = \frac{\beta}{\gamma}$, which represents the number of infected people from a single infectious person in a population where all people are susceptible.

1.2.2 Consequences of assumptions of the SIR model

Let us make a few remarks on the consequences of the assumptions of the SIR model on its practical application.

The population is assumed to be distributed into three compartments. It is clear that this pattern is satisfied for very specific disease outbreaks as it neglects many important factors such as the latent period of the disease, quarantine, case fatality rate, etc. More advanced models involving some of these factors are presented in Sections 1.3 and 1.4.

The relations between the three compartments are based on the assumptions 1–4 stated above. These assumptions may limit applications of the model considerably. The condition of a constant population size is usually satisfied if we restrict ourselves to epidemics lasting a short period of time. As regards the homogeneity of the population, this assumption may cause some difficulties. If the considered epidemic consists of several small local outbreaks, this condition is clearly not satisfied. This consideration leads us to the definition of an effective population size. The idea is to use a reduced population size which reflects the assumption of homogeneity. However, before the detailed description of this notion is given, the reader must be provided with some theoretical background.

We therefore postpone the detailed description to Section 4.1. The third condition involves the rate of flow between the compartments I and R. It turns out that the directly proportional rate of flow does not often correspond with the real situation and it can be set in a more realistic way, see [8]. We however restrict ourselves to this simple case. The satisfaction of the last condition regarding the acquired immunity depends on the properties of the disease. As a rule, at least temporary immunity is acquired by the recovered population considering most of the common diseases. In this model, those who fall victims to the disease are treated as recovered. This does not pose a problem when the case fatality rate of the disease is zero or close to zero.

1.3 SIQR model

We move on to a more advanced model implementing quarantine. This model, adapted from [9], is based on the additional assumption that every infectious subject is quarantined after the infection is detected.

Let us briefly introduce the model. It is built on the standard SIR model introduced in Section 1.2. In addition to the compartments S, I, and R, we define a new compartment called *Quarantined* and denoted by Q. The infectious move from the compartment I to the compartment Q with a rate of flow directly proportional to the size of I. Analogously, the quarantined leave the compartment Q and move on to the compartment R with a rate of flow directly proportional to the size of Q. The coefficients of proportionality are denoted by α and δ, respectively.

One more supplementary modification must be made in the SIR model so as to obtain the SIQR model. We need to take into account that the quarantined people are not able to interact with the rest of the population (the so called *active* population). Therefore, the rate of flow between the compartments S and I in the SIR model (see Figure 1.1) has to be modified in an appropriate manner. Since the size of the active population can be expressed as $N - Q$, we replace the expression $\frac{\beta}{N}SI$ with $\frac{\beta}{N-Q}SI$. The resulting SIQR model is illustrated in Figure 1.2.

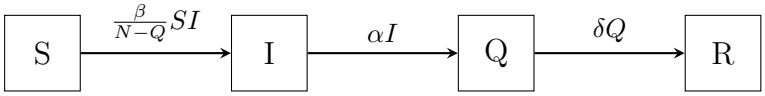

Figure 1.2: SIQR model

For completeness, we define the initial conditions of the model. They are analogous to the initial conditions of the SIR model (1.1). Let $I_0 > 0$, $Q_0 \geq 0$, and $R_0 \geq 0$ such that $N - R_0 - Q_0 - I_0 > 0$. We set

$$\begin{aligned} S(0) &= N - R_0 - Q_0 - I_0, \\ I(0) &= I_0, \\ Q(0) &= Q_0, \\ R(0) &= R_0. \end{aligned} \tag{1.3}$$

To conclude, we obtain the initial value problem corresponding to the SIQR model:

$$\begin{aligned} S' &= -\frac{\beta}{N-Q}SI, \\ I' &= \frac{\beta}{N-Q}SI - \alpha I, \\ Q' &= \alpha I - \delta Q, \\ R' &= \delta Q, \end{aligned} \tag{1.4}$$

with the initial conditions (1.3). The basic reproduction number for this model is given by $R_0 = \frac{\beta}{\alpha}$. Analogously to the SIR model, the value $\frac{1}{\alpha}$ is equal to the expected amount of time one spends in the compartment I and $\frac{1}{\delta}$ equals the expected amount of time spent in the compartment Q.

1.4 Other advanced models

The SIR and SIQR models presented in Sections 1.2 and 1.3 are applied on the COVID-19 epidemiological data in Chapter 4 followed by thorough discussion of the results. Thanks to our own implementation of the optimization algorithm described in Section 3.3, we were able to test a wider range of compartmental models. Since the focus of this work is not on the models themselves, we describe these more sophisticated models only briefly. The outcome is interesting – these models gave us results almost identical with the two previously described models. A concise discussion of the results is therefore provided.

These advanced models are introduced schematically in the form of diagrams, a short description is given below.

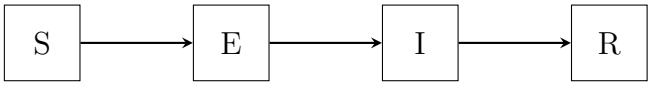

Figure 1.3: SEIR model

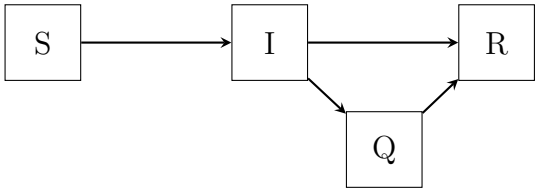

Figure 1.4: SIQR model No. 2

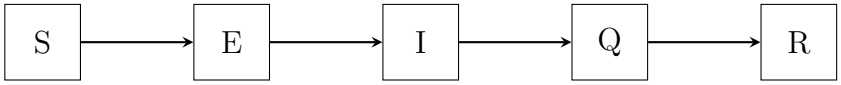

Figure 1.5: SEIQR model

The SEIR model 1.3 adds the latency period to the standard SIR model. The letter E stands for the word *Exposed*, this compartment contains infected

people who are not infectious yet. Development of the compartment I is therefore delayed in comparison with the SIR model.

A different approach to the quarantine is shown by the second SIQR model 1.4. This approach allows us to model the case when some infectious are not detected and avoid the quarantine. Finally, the SEIQR model combines the latency period with quarantine.

Results of the application of these model to the COVID-19 epidemiological data from the Czech Republic are reviewed briefly in Section 4.5.

2. Sensitivity equations

The main goal of this thesis is the estimation of the parameters in models presented in Chapter 1. This cannot be accomplished without introducing the necessary mathematical background regarding the theory of the ordinary differential equations. In this chapter we derive the equations describing the sensitivity of a system of ordinary differential equations (we write ODE for short) with respect to a parameter. Then we give some illustrative examples and present the derivation of the sensitivity equations for the SIR model.

2.1 Derivation and definition

2.1.1 Ordinary differential equations and notation

For our purposes, we use the notion of a system of differential equations in the following way:

Notation. Let $n \in \mathbb{N}$, let $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ be a non-empty open set and $f_i : \Omega \to \mathbb{R}$ for $i \in \{1, \dots, n\}$. By a *system of differential equations* we mean any system of the form

$$\begin{aligned} y_1' &= f_1(y_1, \dots, y_n, t), \\ y_2' &= f_2(y_1, \dots, y_n, t), \\ &\vdots \\ y_n' &= f_n(y_1, \dots, y_n, t). \end{aligned} \tag{2.1}$$

We write it in vector notation $y' = f(t, y(t))$ for brevity. For our purposes $\Omega = \mathbb{R} \times \mathbb{R}^n$ if not stated otherwise.

Definition 1. By a *solution to the system* (2.1) we mean a vector-valued function $y = (y_1, \dots, y_n)^T$ defined on an open interval I such that for all $t \in I$ and for every $i \in \{1, \dots, n\}$ the condition (2.1) holds.

Definition 2. By an *initial value problem* we mean a system of differential equations $y' = f(t, y(t))$ together with a point $(t_0, y^0) \in \Omega$ called the *initial condition*. A function y is said to be a *solution to the initial value problem* if y is a solution to the system of differential equations and satisfies $y(t_0) = y^0$.

Since the systems of ODEs corresponding to the epidemiological models depend on some parameters, we need to formalise the notion of a function dependent on a parameter. Consider a function $g : I \times G \to \mathbb{R}^n$, where $I \subset \mathbb{R}$ is an open interval and $G \subset \mathbb{R}$ is an open set (for our purposes $I = G = \mathbb{R}$ if not stated otherwise). Then g is a function of two variables, we write $g = g(t, c)$, where $c \in \mathbb{R}$. By $g(\cdot, c)$ we mean a function of one variable (the variable is denoted by the dot) with a fixed value of the parameter c. In other words we define $g(\cdot, c) = h$, where $h : I \to \mathbb{R}^n$ is given by $h(t) = g(t, c)$. In order to simplify the notation, for some fixed value of the parameter c we will sometimes omit the second argument and write $g(t, c) = g(t)$. Analogously, we write $g'(t, c) = g'(t) = \frac{\partial g}{\partial t}(t, c)$ if the right-hand side is defined. This will simplify the notation for ordinary differential equations, where t is the relevant variable and c is only a parameter.

2.1.2 Sensitivity equations – simple ODE

In this section we follow the paper of Dickinson and Gelinas [6] and the monograph [5] by Schittkowski. Let us consider an initial value problem

$$y'(t,c) = f\Big(y(t,c),t,c\Big), \quad y(0,c) = y^0, \tag{2.2}$$

which depends on a real parameter c. The initial value problem may be represented either by one equation or by a system of equations, thus y and f are either scalar-valued or vector-valued functions and we do not distinguish between these two cases in notation. We now assume that the initial condition $y(0,c) = y^0 \in \mathbb{R}^n$ does not depend on the parameter c. Let $y = (y_1,\ldots,y_n)^T$ be a solution of (2.2). Then y can be treated as a function of two variables, t and c, we write $y = y(t,c)$. In order to optimize the parameters in our models we need to define and determine the so called *sensitivity* of the system with respect to the parameter c. We also introduce the notion of a sensitivity equation.

Definition 3. Suppose that for every $i \in \{1,\ldots,n\}$ and for all t and c of the domain of y there exists $\frac{\partial y_i}{\partial c}(t,c)$. We define the *sensitivity of the i-th variable with respect to the parameter c* by

$$z_i(t,c) = \frac{\partial y_i}{\partial c}(t,c).$$

The sensitivities defined above can be obtained as a solution of a system of differential equations called the sensitivity equations which we derive now. Let $i \in \{1,\ldots,n\}$. We assume that the partial derivatives $\frac{\partial y_i}{\partial c}$ and $\frac{\partial y_i}{\partial t}$ are sufficiently smooth functions. Then we obtain by Definition 3 and the rule for interchanging the order of differentiation

$$\frac{\partial z_i}{\partial t}(t,c) = \frac{\partial}{\partial t}\left(\frac{\partial y_i}{\partial c}(t,c)\right) = \frac{\partial}{\partial c}\left(\frac{\partial y_i}{\partial t}(t,c)\right).$$

By using (2.2), the chain rule for differentiation and Definition 3 we have

$$\begin{aligned}\frac{\partial z_i}{\partial t}(t,c) &= \frac{\partial}{\partial c}\Big[f_i\Big(y_1(t,c),\ldots,y_n(t,c),t,c\Big)\Big] \\ &= \frac{\partial f_i}{\partial c}\Big(y_1,\ldots,y_n,t,c\Big) + \sum_{j=1}^{n}\frac{\partial f_i}{\partial y_j}\Big(y_1,\ldots,y_n,t,c\Big)\frac{\partial y_j}{\partial c}(t,c) \\ &= \frac{\partial f_i}{\partial c}\Big(y_1,\ldots,y_n,t,c\Big) + \sum_{j=1}^{n}\frac{\partial f_i}{\partial y_j}\Big(y_1,\ldots,y_n,t,c\Big)\cdot z_j(t,c).\end{aligned} \tag{2.3}$$

We obtain what will be referred to as the sensitivity equations. These are a system of n differential equations which can be solved simultaneously with the original system (2.2). We now determine the initial condition of the sensitivity equations. Since the initial condition of the original system (2.2) does not depend on the parameter c, we have for $i \in \{1,\ldots,n\}$ by Definition 3

$$z_i(0,c) = \frac{\partial y_i}{\partial c}(0,c) = \frac{\partial y_i^0}{\partial c} = 0.$$

For some fixed value of the parameter c we write simply $\frac{\partial z_i}{\partial t}(t,c) = z_i'(t)$. Summing up, we have the following definition:

Definition 4. Let $y'(t,c) = f(y(t,c),t,c),\ y(0,c) = y^0$ be an initial value problem of the form (2.2) and suppose that the initial condition does not depend on the parameter $c \in \mathbb{R}$. We define the *sensitivity equations* by

$$z_i'(\cdot,c) = \frac{\partial f_i}{\partial c}\Big(y_1,\dots,y_n,t,c\Big) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial y_j}\Big(y_1,\dots,y_n,t,c\Big)\cdot z_j(\cdot,c), \quad z_i(0,c) = 0,$$

for $i \in \{1,\dots,n\}$.

2.1.3 Multiple parameters and parameter in initial condition

Until now we have discussed the case when the initial condition does not depend on the parameter. However, a parameter may appear both in the equation and in the initial condition and we need to derive the sensitivity equations for this situation as well. Consider the following initial value problem:

$$y'(t,c) = f\Big(y(t,c),t,c\Big), \quad y(0,c) = (c, y_2^0,\dots,y_n^0)^T, \tag{2.4}$$

where $c \in \mathbb{R}$ and $(y_2^0,\dots,y_n^0)^T \in \mathbb{R}^{n-1}$. The derivation of the sensitivity equation itself is identical with the first case (2.3). We obtain the equation of the same form as presented in Definition 4:

$$z_i' = \frac{\partial f_i}{\partial c} + \sum_{j=1}^{n} \frac{\partial f_i}{\partial y_j} z_j,\ i \in \{1,\dots,n\}.$$

We now compute the corresponding initial conditions. For the first variable z_1 we have by Definition 3

$$z_1(0,c) = \frac{\partial y_1}{\partial c}(0,c) = \frac{\partial}{\partial c} c = 1.$$

For $i \in \{2,\dots,n\}$ we get

$$z_i(0,c) = \frac{\partial y_i}{\partial c}(0,c) = \frac{\partial y_i^0}{\partial c} = 0,$$

which completes the derivation. Note that the only difference from the case with the parameter-independent initial condition from Definition 4 is in the value of the initial condition of z_i.

In our models there can often be found more parameters than one. Thus, we generalise our derivation of the sensitivity equations for the case of multiple parameters. The corresponding initial value problem is stated as follows:

$$y'(t,c) = f\Big(y(t,c),t,c\Big), \quad y(0,c) = y^0 \in \mathbb{R}^n,\ c = (c_1,\dots,c_m)^T \in \mathbb{R}^m. \tag{2.5}$$

Notation 1. Consider the initial value problem (2.5). Suppose that for every $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$ and for all t and c of the domain of y there exists $\frac{\partial y_i}{\partial c_j}(t,c)$. We define the *sensitivity of the i-th variable with respect to the parameter* c_j by

$$z_i^j(t,c) = \frac{\partial y_i}{\partial c_j}(t,c).$$

Let $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. Similarly to the derivation in the previous case (2.3) we obtain the sensitivity equation for the sensitivity of the i-th variable with respect to the parameter c_j:

$$\begin{aligned}\frac{\partial z_i^j}{\partial t}(t,c) &= \frac{\partial}{\partial t}\left(\frac{\partial y_i}{\partial c_j}(t,c)\right) = \frac{\partial}{\partial c_j}\left(\frac{\partial y_i}{\partial t}(t,c)\right) = \frac{\partial}{\partial c_j}\Big[f_i\Big(y_1(t,c),\dots,y_n(t,c),t,c\Big)\Big] \\ &= \frac{\partial f_i}{\partial c_j}\Big(y_1,\dots,y_n,t,c\Big) + \sum_{k=1}^{n}\frac{\partial f_i}{\partial y_k}\Big(y_1,\dots,y_n,t,c\Big)\frac{\partial y_k}{\partial c_j}(t,c) \\ &= \frac{\partial f_i}{\partial c_j}\Big(y_1,\dots,y_n,t,c\Big) + \sum_{k=1}^{n}\frac{\partial f_i}{\partial y_k}\Big(y_1,\dots,y_n,t,c\Big)\cdot z_k^j(t,c).\end{aligned}$$

Since the initial condition y^0 does not depend on the parameters $c_1, \dots, c_m$, we get in total $m \cdot n$ sensitivity equations of the form

$$(z_i^j)' = \frac{\partial f_i}{\partial c_j} + \sum_{k=1}^{n}\frac{\partial f_i}{\partial y_k} z_k^j,$$

along with the initial conditions $z_i^j(0,c) = 0$.

2.2 Examples and derivation for the SIR model

2.2.1 Simple ODE

Example 1. Consider the equation $y' = cy,\ y(0) = K,\ c, K \in \mathbb{R}$. Find the sensitivity equation and solve both equations.

Solution. The given ODE is separable, the solution is therefore $y(t) = Ke^{ct},\ t \in \mathbb{R}$. The sensitivity equation is derived as follows in accordance with Definition 4:

$$z' = \frac{\partial}{\partial c}(cy) + \frac{\partial}{\partial y}(cy)\cdot z = y + cz = Ke^{ct} + cz.$$

We obtain the sensitivity equation which is a linear differential equation with constant coefficients and a special right-hand side

$$z' - cz = Ke^{ct}.$$

Its fundamental system is $\{e^{ct}\}$. The form of the right-hand side gives us a particular solution of the form pte^{ct} for some $p \in \mathbb{R}$. Comparing the coefficients on the left- and right-hand side yields $p = K$. Therefore, the general solution is

$$z(t) = Kte^{ct} + \alpha e^{ct},\ t \in \mathbb{R},\ \alpha \in \mathbb{R}.$$

Our solution satisfies the condition $z(0) = 0$, hence $\alpha = 0$. Summing up, we get

$$\begin{aligned}y(t) &= Ke^{ct}, \\ z(t) &= Kte^{ct},\ t \in \mathbb{R}\end{aligned}$$

which is our solution. Note that since we are able to solve the given equation analytically, the sensitivity can be computed directly from Definition 3 and the results are identical. ♣

Example 2. Consider the equation $y' = cy,\ y(0) = c,\ c \in \mathbb{R}$. Find the sensitivity equation and solve both equations.

Solution. The solution of the given ODE is (analogously to the previous example) $y(t) = ce^{ct},\ t \in \mathbb{R}$. The sensitivity equation is derived as follows:

$$z' = \frac{\partial}{\partial c}(cy) + \frac{\partial}{\partial y}(cy) \cdot z = y + cz = ce^{ct} + cz.$$

Similarly as in Example 1 we obtain $z' - cz = ce^{ct}$, which is a linear differential equation with constant coefficients and a special right-hand side. It has the general solution of the form

$$z(t) = cte^{ct} + \alpha e^{ct},\ t \in \mathbb{R},\ \alpha \in \mathbb{R}.$$

The initial condition $z(0) = 1$ (the equation is of the form (2.4)) yields $\alpha = 1$, therefore

$$z(t) = e^{ct} + cte^{ct},$$

which is the formula for sensitivity. ♣

2.2.2 SIR model

For our models, the sensitivity equations used in the optimization algorithm are computed by our program using MATLAB. For completeness, we present here the derivation of the sensitivity equations for the SIR model. Consider the equations of the SIR model (1.2) along with the initial conditions (1.1).

This system of differential equations depends on two parameters, β and γ. The resulting initial value problem is thus of the form (2.5). According to 1, the sensitivity equations can be derived either with respect to β or γ. We introduce the following notation: $f_1(S, I, R, t, \beta, \gamma) = -\frac{\beta}{N}SI$, $f_2 = \frac{\beta}{N}SI - \gamma I$, $f_3 = \gamma I$. The sensitivity equations are derived following the results from the last paragraph of Section 2.1.3. We begin with the sensitivity equations with respect to the parameter β:

$$\begin{aligned}
(z_S^\beta)' &= \frac{\partial f_1}{\partial \beta} + \frac{\partial f_1}{\partial S}z_S^\beta + \frac{\partial f_1}{\partial I}z_I^\beta + \frac{\partial f_1}{\partial R}z_R^\beta = -\frac{1}{N}SI - \frac{\beta}{N}Iz_S^\beta - \frac{\beta}{N}Sz_I^\beta,\\
(z_I^\beta)' &= \frac{\partial f_2}{\partial \beta} + \frac{\partial f_2}{\partial S}z_S^\beta + \frac{\partial f_2}{\partial I}z_I^\beta + \frac{\partial f_2}{\partial R}z_R^\beta = \frac{1}{N}SI + \frac{\beta}{N}Iz_S^\beta + \Big(\frac{\beta}{N}S - \gamma\Big)z_I^\beta,\\
(z_R^\beta)' &= \frac{\partial f_3}{\partial \beta} + \frac{\partial f_3}{\partial S}z_S^\beta + \frac{\partial f_3}{\partial I}z_I^\beta + \frac{\partial f_3}{\partial R}z_R^\beta = \gamma z_I^\beta.
\end{aligned}$$

The derivation with respect to the parameter γ follows. We present here only the results as the process is analogous.

$$\begin{aligned}
(z_S^\gamma)' &= -\frac{\beta}{N}Iz_S^\gamma - \frac{\beta}{N}Sz_I^\gamma,\\
(z_I^\gamma)' &= -I + \frac{\beta}{N}Iz_S^\gamma + \Big(\frac{\beta}{N}S - \gamma\Big)z_I^\gamma,\\
(z_R^\gamma)' &= I + \gamma z_I^\gamma.
\end{aligned}$$

Since the initial conditions of the standard SIR model do not depend on the parameters, all the initial conditions of the sensitivity equations are equal to 0.

3. Numerical methods

Numerical methods are used in two ways in this thesis. First of all, we need to find an approximate numerical solution to differential equations we are not able to solve analytically. We address this problem by using the fourth-order Runge–Kutta method. Then we proceed to the formulation of our parameter estimation problem and we find a suitable numerical method for its solution.

3.1 Runge–Kutta method

In order to approximate a solution to a given system of ODEs, we use the standard fourth-order Runge–Kutta method, which provides us with adequate accuracy. We give here a brief description of this method adapted from the book [10]. Consider the following initial value problem:

$$y'(t) = f(y(t), t), \quad y(0) = y^0,$$

where $y^0 \in \mathbb{R}^n$. Let us denote by y the solution to this problem. We start at a given time t_0 (in our case $t_0 = 0$) and construct a finite sequence $\{(t_j, y_j)\}_{j=0}^N$ such that $y_j \approx y(t_j)$ for $j \in \{1, \ldots, N\}$ (here $y_j \in \mathbb{R}^n$). Let $h > 0$ be a fixed step and $y_0 = y(0)$. For $j = 0, 1, 2, \ldots$ define

$$\begin{aligned} t_{j+1} &= t_j + h, \\ y_{j+1} &= y_j + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \end{aligned} \tag{3.1}$$

where $k_1, \ldots, k_4 \in \mathbb{R}^n$ are increments computed at each step and defined by

$$\begin{aligned} k_1 &= f(y_j, t_j), \\ k_2 &= f\Big(y_j + \tfrac{1}{2}hk_1, t_j + \tfrac{1}{2}h\Big), \\ k_3 &= f\Big(y_j + \tfrac{1}{2}hk_2, t_j + \tfrac{1}{2}h\Big), \\ k_4 &= f\Big(y_j + hk_3, t_j + h\Big). \end{aligned}$$

For our purposes, it proves sufficient to set $h = \frac{1}{100}$. More details and derivation of this method can be found in [10].

3.2 Algorithms for parameter optimization

In this section we formulate the non-linear least squares problem in the context of our models. Then we introduce two algorithms suitable for solving our problem. The described methods are tested on the function from Example 1. At the end, we give the description of our program implementing the Levenberg–Marquardt algorithm.

3.2.1 Problem formulation

We now address the problem of optimizing the parameters in the models introduced in Chapter 1. The parameters are sought in such a way that the models

correspond as precisely as possible with the observational data. There are many possibilities how to approach this problem, see [5]. Our approach is the following: The resulting function obtained as a solution to the considered model fits the measured data in the least squares sense. More precisely, consider the initial value problem

$$y'(t,c) = f\Big(y(t,c),t,c\Big), \quad y(0,c) = y^0 \in \mathbb{R}^n, \tag{3.2}$$

which depends on m parameters $c = (c_1,\dots,c_m)^T \in \mathbb{R}^m$. The function y is represented in component form as $y = (y_1,\dots,y_n)^T$. The initial condition in (3.2) is given here in such a way that it does not depend on the parameters $c_1,\dots,c_m$. However, if it does depend on the parameters, the process is completely identical.

Suppose we have a set of data points $\{(t_j, Y^j) \in \mathbb{R}^{n+1},\ j = 0,\dots,M\}$. We want to find a vector of parameters $c_{min} \in \mathbb{R}^m$ such that it satisfies the condition

$$c_{min} = \underset{c\in\mathbb{R}^m}{\arg\min} \sum_{j=0}^{M} \|y(t_j,c) - Y^j\|^2, \tag{3.3}$$

where $y(\cdot,c),\ c = (c_1,\dots,c_m)^T \in \mathbb{R}^m$, denotes the solution to (3.2) and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^n$. The expression in (3.3) may be rewritten

$$\sum_{j=0}^{M} \|y(t_j,c) - Y^j\|^2 = \sum_{j=0}^{M}\sum_{i=1}^{n} \Big(y_i(t_j,c) - Y_i^j\Big)^2 = \sum_{j=0}^{M}\sum_{i=1}^{n} \Big(r_{ij}(c)\Big)^2.$$

Here

$$r_{ij}(c) = y_i(t_j,c) - Y_i^j \tag{3.4}$$

are the residuals. Let us denote for $c \in \mathbb{R}^m$

$$F(c) = \sum_{j=0}^{M} \|y(t_j,c) - Y^j\|^2.$$

Now, we are able to rewrite (3.3) in the form

$$c_{min} = \underset{c\in\mathbb{R}^m}{\arg\min} \sum_{j=0}^{M}\sum_{i=1}^{n} \Big(r_{ij}(c)\Big)^2 = \underset{c\in\mathbb{R}^m}{\arg\min}\, F(c). \tag{3.5}$$

The problem of minimizing a function of the form (3.5) is called the nonlinear least squares problem (for further information see the book [7]). In the case when the functions $r_{ij}(c)$ depend linearly on the parameter vector c, the problem reduces to (linear) least squares. Since we are not able to find analytic solutions to most of the equations obtained from the epidemiological models, we cannot write the explicit formulae for the functions $r_{ij}(c)$. It is therefore required to find suitable numerical algorithms in order to solve this problem. In spite of the fact that we do not have the explicit formulae for $r_{ij}(c)$, it is possible to compute the partial derivatives

$$\frac{\partial r_{ij}}{\partial c_k}(c), \quad k \in \{1,\dots,m\}$$

by using the sensitivity equations derived in Chapter 2. We use these observations in the numerical algorithms presented in the subsequent sections.

3.2.2 Gauss–Newton algorithm

First, we give the description of the Gauss–Newton algorithm in the context of our problem introduced in Section 3.2.1. We present here the derivation of the Gauss–Newton method from the standard Newton method for the problem (3.5) adapted from [7].

It is desired to find a vector c_{min} satisfying the condition (3.5). The considered functions are assumed to be sufficiently smooth. The Newton method iteratively constructs a sequence $c^{(1)}, c^{(2)}, \dots$ of estimates of the vector c_{min}. Let $c^{(0)}$ be our initial estimate of the minimum. The Newton method proceeds iteratively by setting

$$c^{(l+1)} = c^{(l)} - \Big(H(c^{(l)})\Big)^{-1} \nabla F(c^{(l)}), \tag{3.6}$$

where

$$H(c^{(l)}) = \begin{pmatrix} \frac{\partial^2 F}{\partial c_1^2}(c^{(l)}) & \cdots & \frac{\partial^2 F}{\partial c_1 \partial c_m}(c^{(l)}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial c_m \partial c_1}(c^{(l)}) & \cdots & \frac{\partial^2 F}{\partial c_m^2}(c^{(l)}) \end{pmatrix}$$

is the Hessian matrix of F (it is assumed to be invertible) and

$$\nabla F(c^{(l)}) = \begin{pmatrix} \frac{\partial F}{\partial c_1}(c^{(l)}) & \cdots & \frac{\partial F}{\partial c_m}(c^{(l)}) \end{pmatrix}^T$$

is the gradient of F. For simplicity of notation, we sometimes omit the argument $c^{(l)}$. The α-th element of ∇F is equal to

$$\Big(\nabla F\Big)_\alpha = \frac{\partial F}{\partial c_\alpha} = \sum_{j=0}^{M} \sum_{i=1}^{n} \frac{\partial}{\partial c_\alpha} (r_{ij})^2 = \sum_{j=0}^{M} \sum_{i=1}^{n} 2 r_{ij} \frac{\partial r_{ij}}{\partial c_\alpha} = 2(J^T r)_\alpha, \tag{3.7}$$

where r denotes the vector of the residua, i.e.

$$r = (r_{1,0}, \dots, r_{1,M}, r_{2,0}, \dots, r_{2,M}, \dots, r_{n,0}, \dots, r_{n,M})^T,$$

and the Jacobian matrix J is given by

$$J = \begin{pmatrix} \frac{\partial r_{1,0}}{\partial c_1} & \frac{\partial r_{1,0}}{\partial c_2} & \cdots & \frac{\partial r_{1,0}}{\partial c_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{1,M}}{\partial c_1} & \frac{\partial r_{1,M}}{\partial c_2} & \cdots & \frac{\partial r_{1,M}}{\partial c_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{n,0}}{\partial c_1} & \frac{\partial r_{n,0}}{\partial c_2} & \cdots & \frac{\partial r_{n,0}}{\partial c_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{n,M}}{\partial c_1} & \frac{\partial r_{n,M}}{\partial c_2} & \cdots & \frac{\partial r_{n,M}}{\partial c_m} \end{pmatrix}.$$

Let us compute the element at the α-th row and β-th column of matrix $H(c^{(l)})$:

$$\begin{aligned} H_{\alpha,\beta} = \frac{\partial^2 F}{\partial c_\alpha \partial c_\beta} &= \frac{\partial^2}{\partial c_\alpha \partial c_\beta} \sum_{j=0}^{M} \sum_{i=1}^{n} (r_{ij})^2 = \sum_{j=0}^{M} \sum_{i=1}^{n} \frac{\partial^2}{\partial c_\alpha \partial c_\beta} (r_{ij})^2 \\ &= \sum_{j=0}^{M} \sum_{i=1}^{n} 2\Big(\frac{\partial r_{ij}}{\partial c_\alpha} \frac{\partial r_{ij}}{\partial c_\beta} + r_{ij} \frac{\partial^2 r_{ij}}{\partial c_\alpha \partial c_\beta}\Big). \end{aligned}$$

The main idea of the Gauss–Newton method is to neglect the second partial derivatives $\frac{\partial r_{ij}}{\partial c_\alpha \partial c_\beta}$. By setting them to zero we obtain

$$H_{\alpha,\beta} = 2\sum_{j=0}^{M}\sum_{i=1}^{n}\frac{\partial r_{ij}}{\partial c_\alpha}\frac{\partial r_{ij}}{\partial c_\beta} = 2(J^T J)_{\alpha,\beta}. \tag{3.8}$$

According to (3.7) and (3.8), the iteration (3.6) may be written as

$$c^{(l+1)} = c^{(l)} - (J^T J)^{-1} J^T r(c^{(l)}).$$

It is assumed that $(M+1)n \geq m$, which is a necessary condition for invertibility of the matrix $J^T J$. The matrix $J^T J$ is a Gram matrix of the column vectors of J. It is therefore invertible if and only if the column vectors of J are linearly independent. In the case when $(M+1)n < m$, it is evident that the column vectors of J are linearly dependent, because the dimension of J is $(M+1)n \times m$.

The matrix J can be expressed by computing the values of the partial derivatives $\frac{\partial r_{i,j}}{\partial c_k}$, $k \in \{1, \ldots, m\}$. Using (3.4) and following Notation 1 we obtain

$$\frac{\partial r_{i,j}}{\partial c_k}(c^{(l)}) = \frac{\partial}{\partial c_k}\left(y_i(t_j, c^{(l)}) - Y_i^j\right) = \frac{\partial}{\partial c_k} y_i(t_j, c^{(l)}) = z_i^k(t_j, c^{(l)}).$$

The matrix J may now be written in the form

$$J = \begin{pmatrix} z_1^1(t_0) & z_1^2(t_0) & \cdots & z_1^m(t_0) \\ \vdots & \vdots & \ddots & \vdots \\ z_1^1(t_M) & z_1^2(t_M) & \cdots & z_1^m(t_M) \\ \vdots & \vdots & \ddots & \vdots \\ z_n^1(t_0) & z_n^2(t_0) & \cdots & z_n^m(t_0) \\ \vdots & \vdots & \ddots & \vdots \\ z_n^1(t_M) & z_n^2(t_M) & \cdots & z_n^m(t_M) \end{pmatrix}.$$

This derivation leads us to the following statement of the algorithm. Let $N \in \mathbb{N}$ be the number of iterations. We start by defining an initial estimate $c^{(0)}$. The iterations are then given for $l \in \{1, \ldots, N\}$ by

$$c^{(l+1)} = c^{(l)} - (J^T J)^{-1} J^T r(c^{(l)}).$$

In our implementation we compute the term $(J^T J)^{-1} J^T r(c^{(l)})$ as a solution to a system of linear equations. Denote $x = (J^T J)^{-1} J^T r(c^{(l)})$. Then we have

$$(J^T J)x = J^T r(c^{(l)}). \tag{3.9}$$

This system has a unique solution under the assumption that the column vectors of J are linearly independent. In conclusion, we get the algorithm in the form of pseudocode (see Algorithm 1).

3.2.3 Levenberg–Marquardt algorithm

The Levenberg–Marquardt algorithm is a more robust and frequently used algorithm used for solving the non-linear least squares problem. In practice, one

Algorithm 1: Gauss–Newton algorithm

Input: Initial condition y^0; (t_j, Y^j), $j = 0, \ldots, M$; $c^{(0)}$
Output: $c^{(N)}$

for $i = 1$ **to** N **do**
  for $j = 0$ **to** M **do**
    compute $y(t_j, c^{(i-1)})$ and $z(t_j, c^{(i-1)})$ using the Runge–Kutta method;
  end
  construct matrix J and vector $r(c^{(i-1)})$;
  solve $(J^T J)x = J^T r(c^{(i-1)})$;
  let $c^{(i)} = c^{(i-1)} - x$;
end

often encounters problems with the Gauss–Newton algorithm caused by the near-singularity of the matrix $J^T J$ in (3.9). The Levenberg-Marquardt algorithm fixes this problem using a simple modification of the matrix $J^T J$. We do not present here the details and the derivation of this algorithm, for further information see the books [7] and [5]. In the Gauss–Newton algorithm we solve the following equation for x (see (3.9)):

$$(J^T J)x = J^T r(c^{(l)}). \tag{3.10}$$

We now replace this equation by an equation of the form

$$(J^T J + \lambda^{(l)} I)x = J^T r(c^{(l)}), \tag{3.11}$$

where I is the identity matrix and $\lambda^{(l)} > 0$. Our strategy for the choice of the factor $\lambda^{(l)}$ is taken from the paper [11]. We proceed directly to the pseudocode, see Algorithm 2. In the statement of the algorithm we use the notation introduced in Section 3.2.1.

3.2.4 Numerical experiments

We now perform some numerical experiments with the algorithms presented in Sections 3.2.2 and 3.2.3 on the equation from Example 1. Due to the fact that we are able to find the exact solution to this problem, we can analyse the error and estimate the order of convergence and the rate of convergence of the numerical methods. We introduce these notions in the following definitions.

Consider an iterative numerical method estimating the solution (denoted by c_{min}) to the problem (3.5). Let $c^{(1)}, c^{(2)}, \ldots$ be a sequence of estimates obtained from the method. Let us denote by e_n the *error in the n-th iteration*, i.e. $e_n = c^{(n)} - c_{min}$.

Algorithm 2: Levenberg–Marquardt algorithm

Input: Initial condition y^0; $(t_j, Y^j),\ j = 0, \ldots, M$; $c^{(0)}$
Output: $c^{(N)}$

let $\nu > 1$;
let $\lambda^{(0)} > 0$;
let $F^{(0)} = F(c^{(0)})$;
for $i = 1$ **to** N **do**
    for $j = 0$ **to** M **do**
        compute $y(t_j, c^{(i-1)})$ and $z(t_j, c^{(i-1)})$ using the Runge–Kutta method;
    end
    construct matrix J and vector $r(c^{(i-1)})$;
    solve $(J^T J + \lambda^{(i-1)} I)x_a = J^T r(c^{(i-1)})$;
    solve $(J^T J + \frac{\lambda^{(i-1)}}{\nu} I)x_b = J^T r(c^{(i-1)})$;
    let $c_a = c^{(i-1)} - x_a$;
    let $c_b = c^{(i-1)} - x_b$;
    compute $F(c_a)$;
    compute $F(c_b)$;
    if $F(c_b) \leq F^{(i-1)}$ **then**
        let $\lambda^{(i)} = \frac{\lambda^{(i-1)}}{\nu}$;
        let $F^{(i)} = F(c_b)$;
        let $c^{(i)} = c_b$;
    else if $f(c_a) \leq f^{(i-1)}$ **then**
        let $\lambda^{(i)} = \lambda^{(i-1)}$;
        let $F^{(i)} = F(c_a)$;
        let $c^{(i)} = c_a$;
    else
        {multiply λ by ν until for some smallest k it holds $F(c_c^{(k)}) \leq F^{(i-1)}$, where $c_c^{(k)}$ is obtained by solving $(J^T J + \lambda^{(i-1)}\nu^k I)x_c^{(k)} = J^T r(c^{(i-1)})$, $c_c^{(k)} = c^{(i-1)} - x_c^{(k)}$};
        let $\lambda^{(i)} = \lambda^{(i-1)}\nu^k$;
        let $F^{(i)} = F(c_c^{(k)})$;
        let $c^{(i)} = c_c^{(k)}$;
end

Definition 5. The numerical method is said to be *convergent* for a given problem of the form (3.5) if

$$\lim_{n\to\infty} \|e_n\| = 0.$$

A real number $q \geq 1$ is said to be the *order of convergence* of the numerical method if

$$\lim_{n\to\infty} \frac{\|e_{n+1}\|}{\|e_n\|^q} = C$$

for some $C > 0$ called the *rate of convergence*.

In applications, the order and the rate of convergence are estimated using a system of algebraic equations

$$C_n = \frac{\|e_n\|}{\|e_{n-1}\|^{q_n}}, \quad C_n = \frac{\|e_{n-1}\|}{\|e_{n-2}\|^{q_n}}.$$

Example 3. Let $c \in \mathbb{R}$. Consider the equation

$$y' = cy, \quad y(0) = 1 \tag{3.12}$$

from Example 1 on the interval $[0, 3]$ and a set of data points $D = \{(j\tau, Y^j)^T \in \mathbb{R}^2,\ \tau = 0.1,\ j = 0, \ldots, 30\}$. The exact description of the set D is given below. It is desired to solve the corresponding non-linear least squares problem stated in Section 3.2.1. According to (3.3), we seek $c_{min} \in \mathbb{R}$ such that

$$c_{min} = \arg\min_{\tilde{c}\in\mathbb{R}} \sum_{j=0}^{30} |y(j\tau, \tilde{c}) - Y^j|^2. \tag{3.13}$$

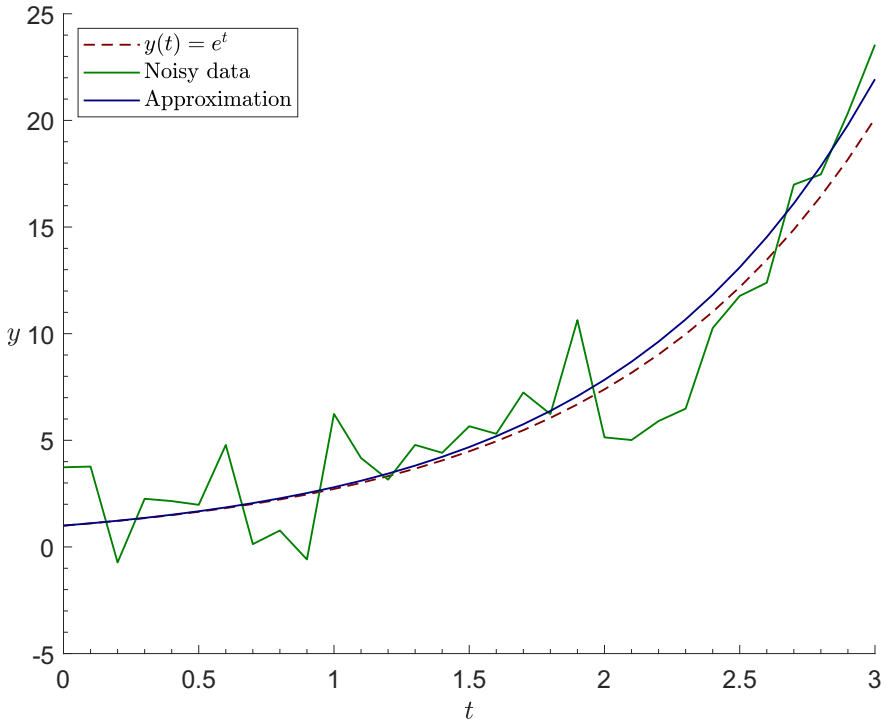

Figure 3.1: Original data and approximation found by Gauss-Newton method

	GN method		LM method	
n	c_n	e_n	c_n	e_n
1	3.65776	2.64153	3.65801	2.64177
2	3.31452	2.29828	3.31494	2.29871
3	2.97020	1.95397	2.97076	1.95452
4	2.62502	1.60878	2.62566	1.60942
5	2.28000	1.26377	2.28069	1.26446
6	1.93865	0.92242	1.93935	0.92312
7	1.61092	0.59469	1.61159	0.59536
8	1.32152	0.30528	1.32207	0.30583
9	1.11611	0.09988	1.11643	0.10020
10	1.02930	0.01307	1.02938	0.01315
11	1.01657	0.00034	1.01657	0.00034
12	1.01624	0.00000	1.01624	0.00000

Table 3.1: Results of Gauss–Newton (GN) and Levenberg–Marquardt (LM) method

n	q_n	C_n
3	1.14253	0.75757
4	1.16599	0.74050
5	1.19769	0.72122
6	1.24176	0.70024
7	1.30442	0.67969
8	1.39415	0.66556
9	1.51904	0.67230
10	1.67557	0.72928
11	1.82024	0.86585
12	1.80015	0.82671

Table 3.2: Rate and order of GN method

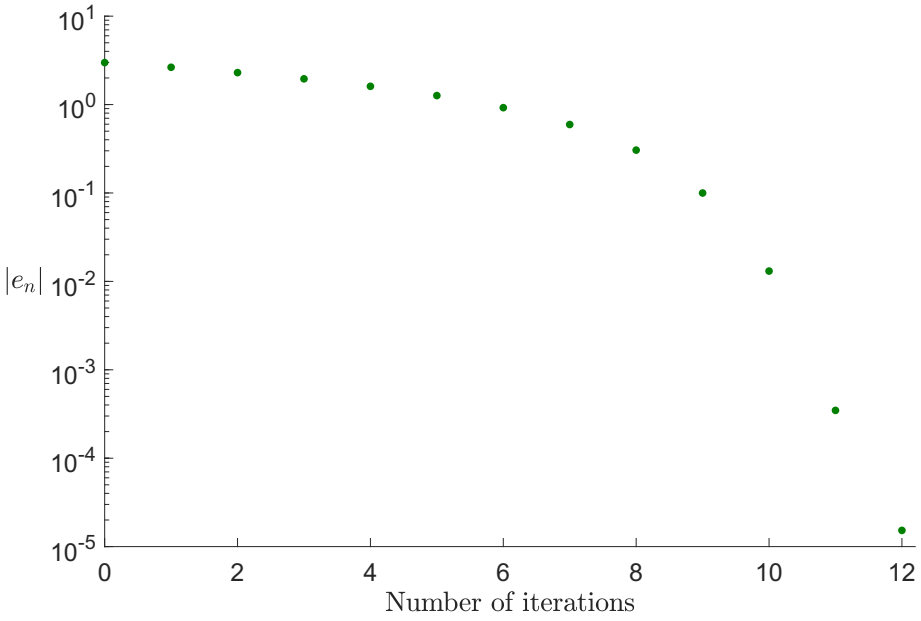

Figure 3.2: Error of Gauss-Newton method

We now describe the set D more precisely. The data points are given in such a way that they represent statistical errors in some observational data. At first, we take the exact solution to the equation (3.12) with $c = 1$, i. e. $y(t) = e^t,\ t \in [0, 3]$ (see Example 1). We then define $Y^j = y(j\tau) + w_j$, where w_j, $j \in \{0, \dots, 30\}$, are uniformly distributed random numbers in the interval $(-4, 4)$. To give a sense of scale, Figure 3.1 shows the resulting so called *noisy data*.

Note that the exact equality $c = c_{min}$ does not hold in general due to the properties of the noisy data. In this example we have on one hand $c = 1$, on the other hand $c_{min} \approx 1.0162$. The latter can be computed in this specific case

from (3.13) using the exact solution to Equation (3.12). These two values are distinguished in this example. In practice, it is desired to find the value of c, while the value of c_{min} is found by the numerical method (if the convergence is successful) and it is assumed $c \approx c_{min}$. Since this simple problem can be completely solved analytically, we used it to verify the correctness of our implementation.

Let us define an initial estimate $c^{(0)} = 4$. The stopping criterion is given by the total number of iterations, we set $N = 12$. The fourth-order Runge–Kutta method described in Section 3.1 is implemented in both algorithms for solving the differential equation with the step $h = \frac{1}{100}$. In the Levenberg–Marquardt algorithm we set $\nu = 3$ and $\lambda^{(0)} = \frac{1}{100}$ (see Algorithm 2). Comparison of both algorithms based on Table 3.1 shows that they are equally successful in terms of convergence considering this simple equation. One can observe that the coefficient $\lambda^{(n)}$ in the Levenberg–Marquardt algorithm tends to 0 as n increases. This indicates why the results of both methods are so similar – the Levenberg–Marquardt method "approaches" the Gauss–Newton method, cf. (3.10) and (3.11). The corresponding error $\|e_n\|$ of the Gauss–Newton method is illustrated in Figure 3.2. The estimates of the convergence order q_n and the convergence rate C_n for the same method are shown in Table 3.2. From the presented results it follows that the convergence is superlinear, i. e. $q_n \in (1, 2)$. ♣

3.3 Program description

For the considered epidemiological models, the Levenberg–Marquardt method proved to be superior to the Gauss–Newton method, the latter often failing or exhibiting slow convergence. The Levenberg–Marquardt method was therefore chosen for implementation in our final program using MATLAB. It has been designed to provide the user with multiple useful options and to reduce manual computations to an essential minimum.

On the input of the program there are only the necessary data – a system of ODEs corresponding to our epidemiological model, an initial estimate of the parameters of our model and a set of data points for optimization. The output contains the desired estimate of the parameters.

As for the algorithm itself, the Levenberg-Marquardt method has been implemented in accordance with Algorithm 2. After performing many numerical experiments, we settled on the values $\lambda^{(0)} = \frac{1}{100}$ and $\nu = 3$. The sensitivity equations are derived automatically using MATLAB. The user is provided with several options: It is not necessary to optimize the model with respect to all its parameters, one can select only the relevant ones. This is a useful tool in applications – the values of some of the parameters are known more precisely and can therefore be fixed.

In addition, in the optimization it is not necessary to take into account all the compartments of the model in question. This is practical, since the reliability of the data from certain compartments may be questionable. Moreover, some compartments cannot be measured at all in practice. In these cases it is possible to include in the minimized function F (see (3.5)) only the compartments we want – those with reliable data. Examples to elucidate this approach are presented in the next chapter.

The program can also handle the situation when a parameter occurs in the

initial condition of the system of equations. As described in Section 1.2.2, it turns out that it may be convenient to optimize the models with respect to the total population size N. Since N appears typically in the initial conditions of the models, we need to include this feature in our program.

4. Application to epidemiological data

In this chapter, the models and methods developed previously are applied to various epidemiological data. At first, we experiment with the properties and limitations of the presented epidemiological models using computer-generated data. Then we proceed to real epidemiological data. Before the models are applied to the COVID-19 epidemiological data, we present here one simpler case, where the conditions are easier to analyse – the case of an influenza epidemic in an English boarding school.

4.1 Numerical experiments

In this section we perform some numerical experiments with the epidemiological models presented in Chapter 1. The computer-generated data are given in such a way that they represent real statistical errors. In the first example of the SIR model, the problem is formulated precisely and in the rest of the examples we proceed analogously.

4.1.1 SIR model

Example 4. Let $\beta > 0$ and $\gamma > 0$. Consider the initial value problem given by the SIR model (1.2) for the population of $N = 2000$ people on the time interval $[0, 80]$ with the initial conditions

$$\begin{aligned} R_0 &= 0, \\ I_0 &= 1, \\ S_0 &= 1999. \end{aligned} \tag{4.1}$$

It is required to solve the non-linear least squares problem stated in Section 3.2.1 corresponding to a set of data points D, which is given as follows:

The set of data points D is defined in such a manner that it represents some statistical data obtained from an influenza epidemic in a population of 2000 people. To acquire the data set D, we first set the values of parameters β and γ. Let $\beta = 0.4$ and $\gamma = 0.25$.[1] Let $\tau = 1$ be the data interval corresponding to one day. We proceed by solving the corresponding system (1.2) with the initial conditions (4.1) using the fourth-order Runge–Kutta method with sufficient accuracy, the results are visualized in Figure 4.1. Let us define

$$\begin{aligned} S^j &= S(j\tau) + \xi_j, \\ I^j &= I(j\tau) + \eta_j, \\ R^j &= R(j\tau) + \zeta_j, \end{aligned}$$

where ξ_j, η_j and ζ_j for $j \in \{0, \dots, 80\}$ are some random numbers from the normal distribution with the expected value $\nu = 0$ and the standard deviation $\sigma = 100$.

[1] This can be computed using the fact that $\frac{\beta}{\gamma} = R_0$, see Section 1.2. It can be estimated $R_0 \approx 1.6$, see [12]. We assume the infectious period to be 4 days long, i.e. $\gamma \approx 0.25$.

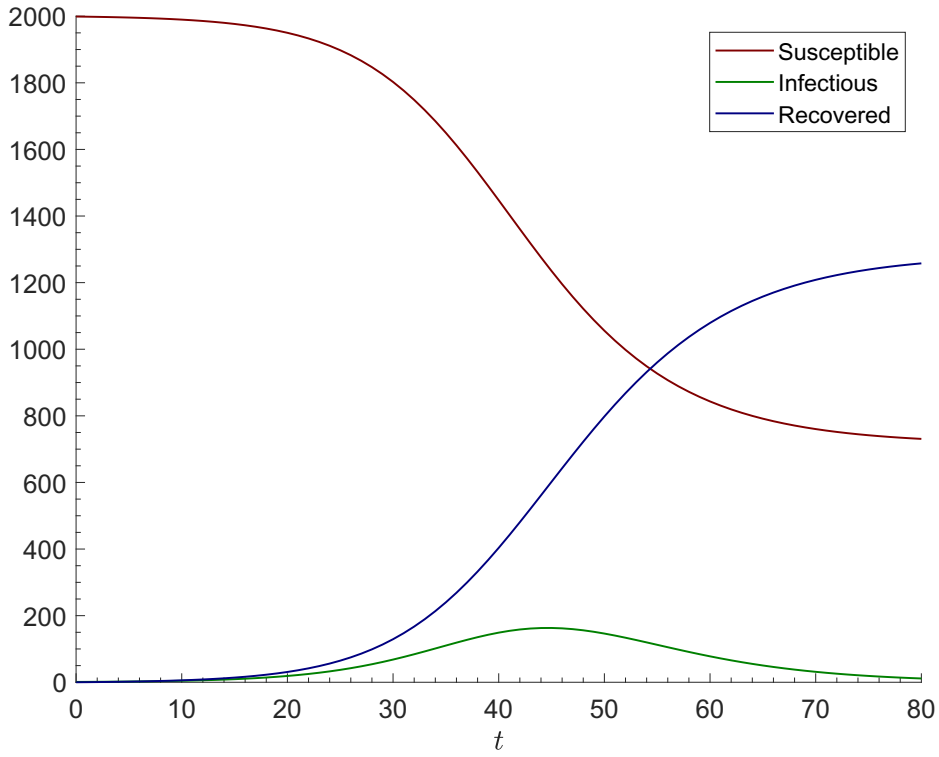

Figure 4.1: SIR model, $\beta = 0.4$, $\gamma = 0.25$

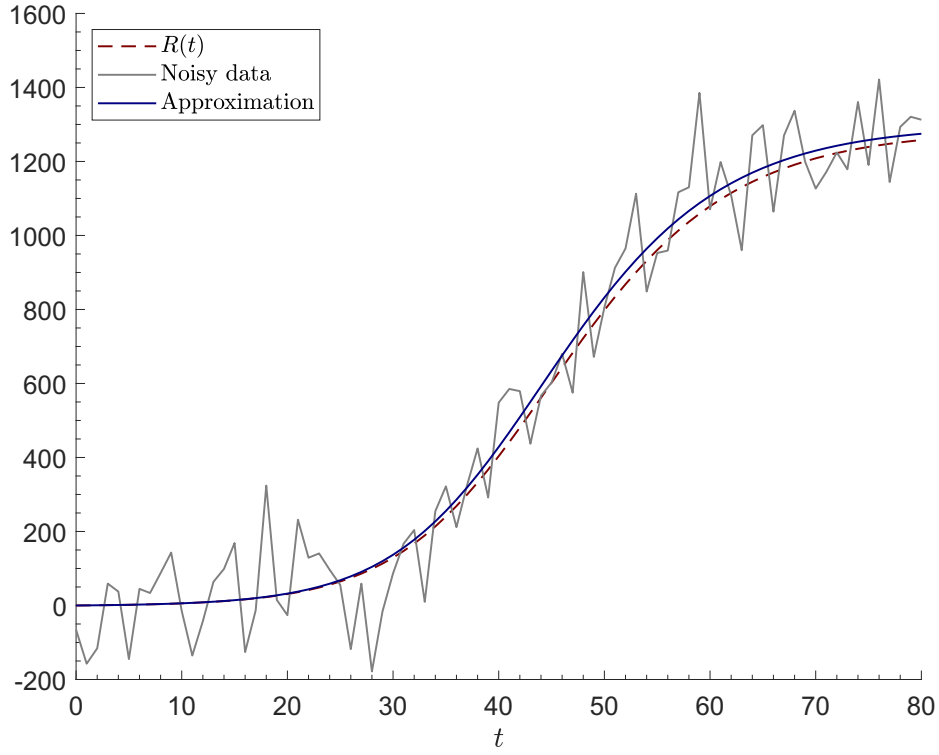

Figure 4.2: Data from compartment R and the resulting approximation

See Figure 4.2 for illustration of the compartment R with the resulting noise. Finally, we set $D = \{(j\tau, S^j, I^j, R^j)^T \in \mathbb{R}^4, j \in \{0, \dots, 80\}\}$ representing measured data on individual days.

We are now able to formulate the non-linear least squares problem precisely. In accordance with (3.3), it is desired to find a vector of parameters $(\beta_m, \gamma_m)^T$ such that

$$(\beta_m, \gamma_m)^T = \underset{(\tilde{\beta},\tilde{\gamma})^T \in \mathbb{R}^2}{\arg\min} \sum_{j=0}^{80} \left(\left| \tilde{S}(j\tau) - S^j \right|^2 + \left| \tilde{I}(j\tau) - I^j \right|^2 + \left| \tilde{R}(j\tau) - R^j \right|^2 \right), \quad (4.2)$$

where $\tilde{S} = S(\cdot, \tilde{\beta}, \tilde{\gamma})$, $\tilde{I} = I(\cdot, \tilde{\beta}, \tilde{\gamma})$ and $\tilde{R} = R(\cdot, \tilde{\beta}, \tilde{\gamma})$ denote the solution of the SIR model equations (1.2) with $(\beta, \gamma)^T = (\tilde{\beta}, \tilde{\gamma})^T$ and the initial conditions (4.1).

The program described in Section 3.3 is used to solve the problem (4.2) considering the initial estimate $(\beta^{(0)}, \gamma^{(0)})^T = (1, 1)^T$ and the stopping criterion given by the increment size

$$\|(\beta_m, \gamma_m)^T - (\beta_{m+1}, \gamma_{m+1})^T\|_\infty < 10^{-5}, \quad (4.3)$$

where $\|\cdot\|_\infty$ is the maximum norm on $\mathbb{R}^2$. The stopping criterion of this form is used in the whole chapter. The desired estimate computed by the program is $(\beta_m, \gamma_m)^T \approx (0.4022, 0.2493)^T$, which is a good approximation of the true values $\beta = 0.4$ and $\gamma = 0.25$. It can be observed that the convergence is fast in spite of the inaccurate initial guess of the minimum. To get a better perspective of the solution, we solve the SIR model equations (1.2) taking $(\beta, \gamma)^T = (0.4022, 0.2493)^T$. The result is shown in Figure 4.2 for compartment R. ♣

Example 5. In practice, one does not often posses the data from all three compartments. In many cases, only the data regarding compartment I are available for our computations. It is therefore necessary to test the accuracy of the minimization algorithm using this incomplete data.

Let us consider the SIR model along with the initial conditions from Example 4 with the same value of parameters $\beta = 0.4$ and $\gamma = 0.25$. The noisy data are obtained the same way as in Example 4 with the exception that only the compartment I is considered now, i.e. we have the noisy data of the form

$$I^j = I(j\tau) + \eta_j,$$

where $\tau = 1$ and η_j for $j \in \{0, \dots, 80\}$ are taken from Example 4. In this case, we set $D = \{(j\tau, I^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 80\}\}$. On the whole, we have the minimization problem of the form

$$(\beta_m, \gamma_m)^T = \underset{(\tilde{\beta},\tilde{\gamma})^T \in \mathbb{R}^2}{\arg\min} \sum_{j=0}^{80} \left| \tilde{I}(j\tau) - I^j \right|^2, \quad (4.4)$$

where the notation follows the notation in Equation 4.2. The initial guess is again set as $(\beta^{(0)}, \gamma^{(0)})^T = (1, 1)^T$ and the stopping criterion is given by the increment size (see (4.3)), which is achieved after $N_I = 30$ iterations. The final estimate found by Levenberg–Marquardt algorithm is $(\beta, \gamma)^T = (0.4151, 0.2635)^T$ The results for compartment I may be seen in Figure 4.3. Unexpectedly, the incomplete data do not affect the accuracy of the algorithm to a large degree, although more iterations are required for successful convergence. ♣

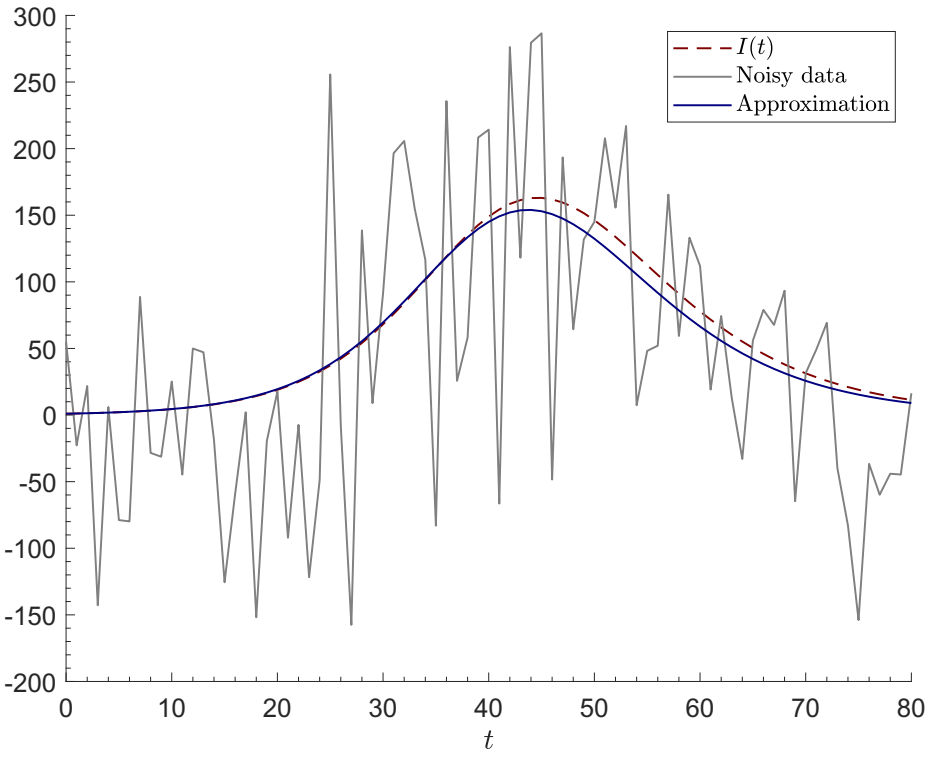

Figure 4.3: Data from compartment I and the resulting approximation

Example 6. As we suggested in Section 1.2.2, it may be convenient to introduce the notion of an effective population, as the real population size is often exaggerated for the purposes of the epidemiological models. Informally speaking, the real population size is reduced in order to ensure the homogeneity of the population in accordance with the assumptions of the SIR model. The question is how to determine the size of the effective population. Our approach is simple – the population size N will not be considered a fixed constant (as it has been until now). Instead, N will be treated as a parameter. In other words, we now consider our epidemiological data (number of infectious etc.) corresponding to an unknown population size and we want to determine this size. Formally, the change is that instead of the parameter vector $(\beta, \gamma)^T$ for the SIR model, we have now an extended parameter vector $(\beta, \gamma, N)^T$. The following example is given to elucidate this notion:

Let us take the data set $D = \{(j\tau, S^j, I^j, R^j)^T \in \mathbb{R}^4, j \in \{0, \ldots, 80\}\}$ from Example 4. Note that this data set corresponds to a population of 2000 people. In accordance with the explanation above, N is now considered a parameter. The optimization is thus performed with respect to three parameters – β, γ and N. Let us remark that the parameter N appears now in the initial conditions of the SIR equations, see (1.1). This does not pose a problem since the theory regarding parameters in initial conditions has been presented in Section 2.1.3 and this feature has been implemented in our program as described in Section 3.3.

The optimization problem is of the form

$$(\beta_m, \gamma_m, N_m)^T = \underset{(\tilde{\beta},\tilde{\gamma},\tilde{N})^T \in \mathbb{R}^3}{\arg\min} \sum_{j=0}^{80} \left(\left|\tilde{S}(j\tau) - S^j\right|^2 + \left|\tilde{I}(j\tau) - I^j\right|^2 + \left|\tilde{R}(j\tau) - R^j\right|^2 \right),$$

where $\tilde{S} = S(\cdot, \tilde{\beta}, \tilde{\gamma}, \tilde{N})$, $\tilde{I} = I(\cdot, \tilde{\beta}, \tilde{\gamma}, \tilde{N})$ and $\tilde{R} = R(\cdot, \tilde{\beta}, \tilde{\gamma}, \tilde{N})$ denote the solution of the SIR model equations (1.2) taking $(\beta, \gamma, N)^T = (\tilde{\beta}, \tilde{\gamma}, \tilde{N})^T$ with the initial conditions (4.1).

Let us perform the optimization using the initial guess $(\beta^{(0)}, \gamma^{(0)}, N^{(0)})^T = (1, 1, 4000)^T$ and the stopping criterion given by the increment size, analogously to (4.3). The tolerance 10^{-5} is achieved after $N_I = 16$ iterations. The resulting estimate is $(\beta_m, \gamma_m, N_m)^T \approx (0.4030, 0.2499, 2008)^T$, which is an accurate approximation of the original values. ♣

4.2 Influenza epidemic in a boarding school

The SIR model was derived under certain conditions made on the population and the disease itself. This may significantly affect the accuracy of the model. It is no more than wishful thinking to assume that these conditions are satisfied in most practical situations. We present here one case, which is apparently as close as possible to satisfying those conditions. This is the case of an influenza outbreak in an English boarding school. We give here just a brief description of the epidemic based on the information from [13]. The data for optimization are taken from [3].

In total, 763 boys were present at school at the beginning of the epidemic. From 15 to 18 January, one boy had an influenza-like illness. Over the next fortnight, a total of 512 boys developed similar symptoms spending between three and seven days in the college infirmary. It is desired to estimate the values of parameters β and γ from the SIR model (1.2) corresponding to this epidemic.

Let us make a few remarks on the conditions specified in Section 1.2.1 in this particular case. It is clear that the population remains constant over the whole period, i.e. $N = 763$. As for the homogeneity of the population, the contacts of the pupils were limited to the people in school, the students and the staff can therefore be considered a closed community – it seems that the population is as homogeneous as possible.

The disease itself seems to satisfy the assumptions of the SIR model. Dividing the population into three groups appears to be convenient since the conditions for spread of the disease were ideal. The presymptomatic period is short, no death cases occurred and the recovered people acquire sufficiently long immunity.

One problem concerning the available data may occur. As a rule, in practical cases we do not possess the data which fit into the structure of the SIR model precisely. The data we possess consist of the number of students confined to bed each day. Following [1], we assume the data to be from the compartment I.

We proceed to the parameter estimation. Having only the data from compartment I, we get the optimization problem of the form (4.4). For visualisation of the measured data, see Figure 4.4. The initial estimate is given by $(\beta^{(0)}, \gamma^{(0)})^T = (1, \frac{1}{7})^T$ and the stopping criterion by the increment size, see (4.3).

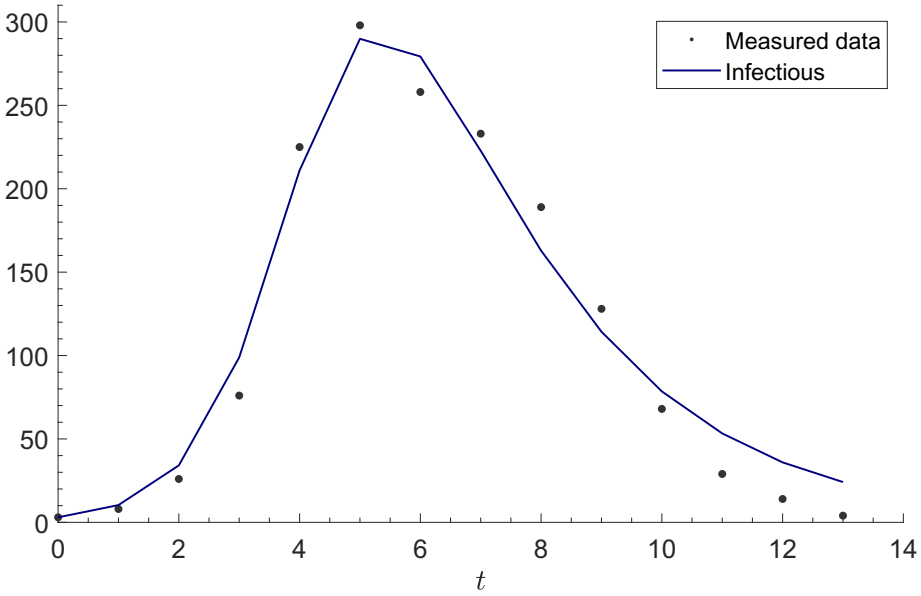

Figure 4.4: Measured data of the flu epidemic and the estimate of compartment I

The desired tolerance 10^{-5} is achieved after $N_I = 6$ iterations. The resulting estimate of the parameters is $(\beta_m, \gamma_m)^T \approx (1.6998, 0.4469)^T$. Figure 4.4 shows that the estimated values of the compartment I are in good agreement with measured data.

However, we may find some discrepancies if we examine the case more closely. Let us solve the SIR model (1.2) with the obtained $(\beta, \gamma)^T = (1.6998, 0.4469)^T$ and the initial conditions corresponding to this influenza epidemic. We find out that the results do not quite correspond with the available data. Namely, the SIR model shows that the total number of people who suffered from the illness is 744, whereas the number stated in [13] is 512. In addition, we know from Chapter 1 that the value $\frac{1}{\gamma} \approx 2.24$ represents the expected time (in days) one spends in the Infectious compartment. It can be seen that this value is somewhat less then the observed value, which is three to seven days. These considerations suggest that even in this simple case some unexpected issues limiting the accuracy of the model occur. This is a consequence of several facts. As stated above, the available data do not fit the model precisely – a person diagnosed with the illness has limited possibilies of spreading the disease because their contacts with the susceptible population are restricted. In addition, the pattern of the SIR model may not be entirely convenient for this particular disease. Nevertheless, in order to adjust the model in accordance with the disease we need additional medical information. These are not available since the epidemic was small and was not given any special significance.

4.3 COVID-19 – the SIR model

To conclude our work, we apply the presented numerical methods to the COVID-19 epidemiological data from the Czech Republic. Multiple approaches presented in Examples 4, 5, and 6 are used and compared in this section. At first, let us

discuss the selection of the data.

4.3.1 Data description

The source of epidemiological data for our computations are the data sets provided by the Ministry of Health of the Czech republic, see [14]. Note that it is not the purpose of this thesis to analyse the methodology of data collection. For further information regarding the methodology, see [15]. First, we need to select a time period for modelling of the epidemic. It follows from the derivation of the SIR model in Chapter 1 that the parameters β and γ in the SIR model are specified by the properties of the disease itself. However, the values of the parameters depend also on some external influences. For instance, adopting some restrictive measures against the spread of the disease decreases the value of parameter β, because the number of contacts of a person is reduced. Thus, the chosen time period should satisfy the condition that the values of the parameters β and γ remain constant within that period. We choose the period from 13 March 2020 to 24 May 2020. The reasons are the following: On 13 March, the key measure forbidding retail sales and the sales of services in business premises came into effect and on 25 May the crucial part of the restrictive measures ended. However, this choice of the time period presents us with a challenge – there is no need to examine the epidemiological data in detail to see that the number of infected population in this time period is very small in proportion to the total population of the Czech Republic which is considered to be $N = 1.065 \cdot 10^7$. This allows us to test the possibilities of the models and the numerical methods in extreme conditions.

Let us describe the available data. We operate with the data set [16] obtained from the web page of the Ministry of Health of the Czech Republic. From this data set, we can extract the cumulative number of infected people, recovered people and people who died of the illness. One problem must be addressed now: It is necessary to adjust the data in order to fit the pattern of the SIR model. In the SIR model, the case fatality rate is neglected and therefore the compartment for the people who died of the illness is not included. As presented in Section 1.2, we include the number of the people who died of the illness in the compartment R.

It follows from the description of the data that the available number of infected people, i.e. the cumulative number, does not correspond to the compartment I. This problem does not occur in case of the compartment R – for this compartment it is possible to use directly the given data. Nonetheless, the data corresponding to the compartment I can be computed very simply. Denote by $I_c(t)$ the cumulative number of infected people at time t. The value $I(t)$ can be obtained from the following equation:

$$I(t) = I_c(t) - R(t).$$

It remains to determine the values of the compartment S. The value $S(t)$ of the compartment S at time t is given by

$$S(t) = N - I(t) - R(t),$$

where N is the total population size.

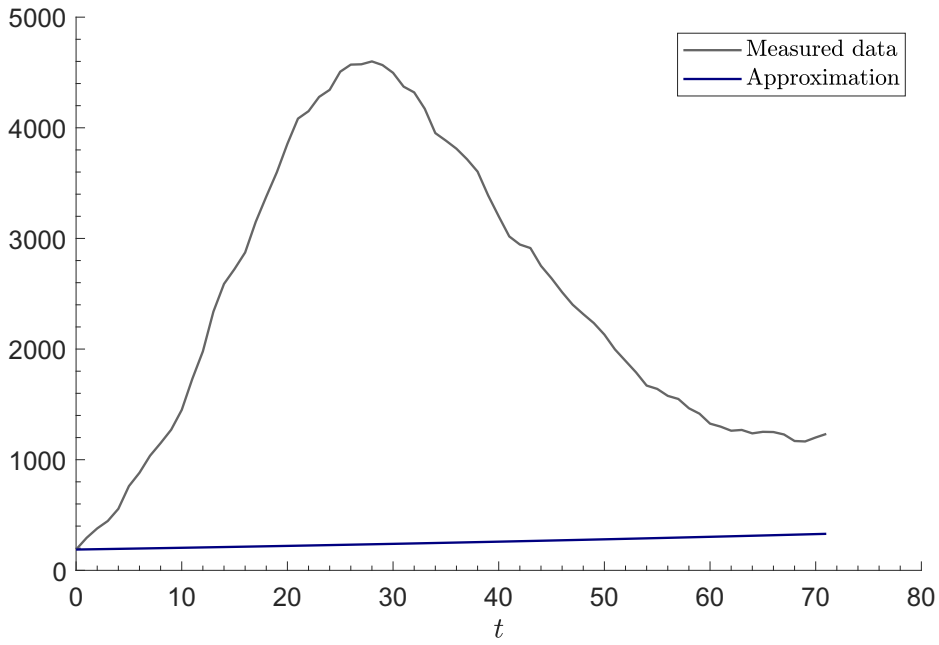

Figure 4.5: SIR: Results of the full-data approach for compartment I

Taking into account the methodology of the data collection, it is questionable, whether the computed data (from the compartment I for instance) fit the definition of I precisely. Since the SIR model is very simple, it leaves us no other option. Using results of medical research, the data could be slightly modified or scaled, for example. Nevertheless, this is not what we focus on.

4.3.2 Numerical experiments

We proceed directly to the numerical experiments. The first (so-called full-data) approach uses all the available data for the optimization and optimizes with respect to both β and γ. In this case, the form of the optimization problem is completely analogous to Example 4. To sum up, we minimize a function of the form

$$(\beta_m, \gamma_m)^T = \underset{(\tilde{\beta},\tilde{\gamma})^T \in \mathbb{R}^2}{\arg\min} \sum_{j=0}^{71} \left(\left| \tilde{S}(j\tau) - S^j \right|^2 + \left| \tilde{I}(j\tau) - I^j \right|^2 + \left| \tilde{R}(j\tau) - R^j \right|^2 \right)$$

with the data set $D = \{(j\tau, S^j, I^j, R^j)^T \in \mathbb{R}^4, j \in \{0, \dots, 71\}\}$ described in Section 4.3.1. The data interval $\tau = 1$ corresponds to one day. The epidemic is modelled on the time interval $[0, 71]$, which is the length of the considered time period in days. The initial guess of the parameters is given by $(\beta^{(0)}, \gamma^{(0)})^T = (1, 1)^T$ and the initial conditions $S(0)$, $I(0)$, and $R(0)$ are given completely by the data set. The tolerance 10^{-5} in the stopping criterion given by the increment size (4.3) is achieved after $N_I = 10$ iterations with the resulting estimate $(\beta_m, \gamma_m)^T \approx (0.5162, 0.5081)^T$. To determine the accuracy of the model, it suffices to visualise the results for the compartment I, see Figure 4.5.

It is clear that the results are utterly inapplicable in practice. This is due to the obvious fact that the epidemic of COVID-19 in the Czech Republic at that

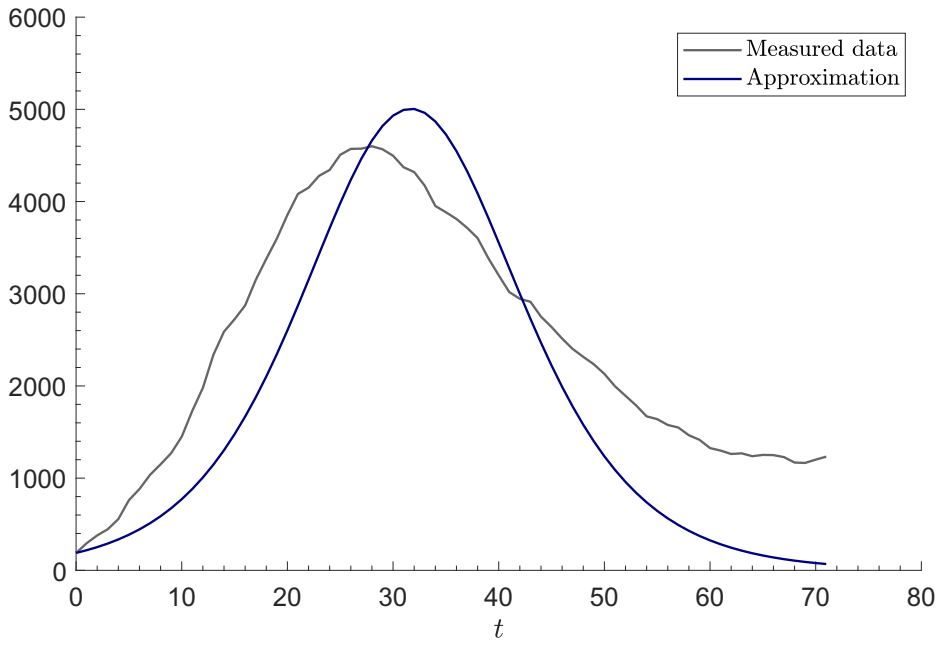

Figure 4.6: SIR: Results of the incomplete-data approach for compartment I

time period consisted of small local outbreaks and thus the assumptions of the SIR model discussed in Section 1.2 are not satisfied. The measured data from the compartment I are on the level of a statistical error in comparison with the population size $N = 1.065 \cdot 10^7$.

The second (so-called incomplete-data) approach is based on the idea of optimizing the parameters β and γ using the data from the compartment I only. It seems that the data from the compartment I are crucial for determining the behaviour of the epidemic, which leads us to this approach. In this case, the optimization problem follows Example 5, i.e. the function to minimize is of the form (4.4) with the data set $D = \{(j\tau, I^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 71\}\}$ obtained as described in Section 4.3.1. The initial guess of the parameters is again given by $(\beta^{(0)}, \gamma^{(0)})^T = (1, 1)^T$ and the initial conditions are the same as in the previous case (of the full-data approach). Again, the tolerance in the stopping criterion (4.3) was achieved after $N_I = 10$ iterations. The computed estimate is $(\beta_m, \gamma_m)^T \approx (4.6687, 4.5244)^T$. It can be observed from Figure 4.6 that the values from the compartment I are approximated more accurately compared to the previous approach as expected. However, if we examine the results in more detail, we find some discrepancies. From the computed estimate it follows that the expected time a person remains infectious is $\frac{1}{\gamma_m} \approx 0.22$ days, which is clearly an unrealistic value based on the results of meta-analysis [17]. From this it can be deduced that the approximation of the values of the remaining two compartments S and R is not good either. Indeed, the model shows that the total number of recovered people at the and of the considered time interval is approximately $6.15 \cdot 10^5$, which is about a hundred times higher than the actual value 7750.

The third approach is based on the idea of the effective population proposed in Example 6. In addition, we consider only the data from the compartment I for the optimization (similarly to the second, incomplete-data approach). As a

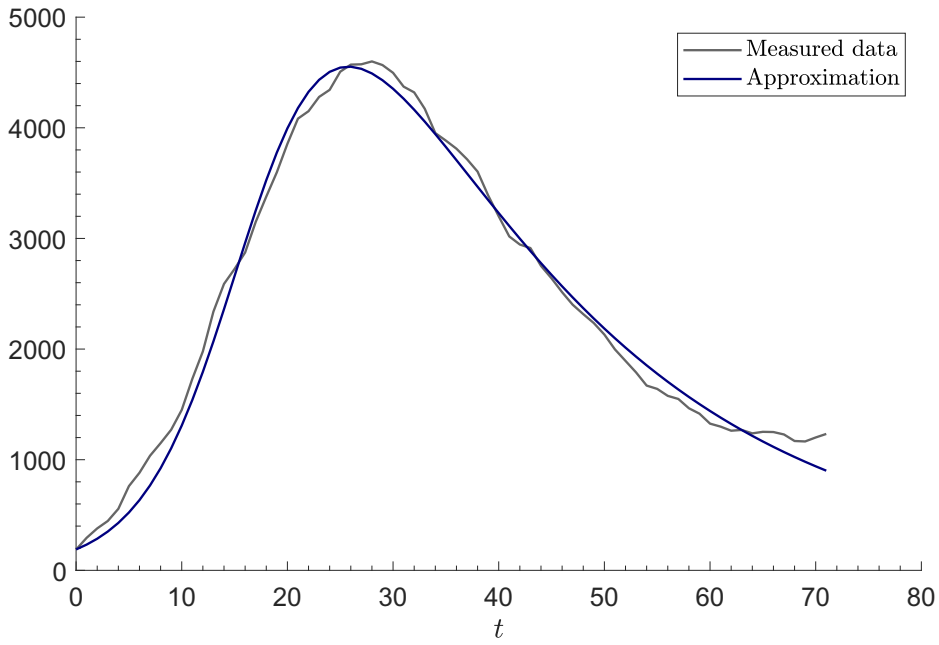

Figure 4.7: SIR: Results of effective population approach for compartment I

result, we obtain the optimization problem of the form

$$(\beta_m, \gamma_m, N_m)^T = \underset{(\tilde{\beta},\tilde{\gamma},\tilde{N})^T \in \mathbb{R}^3}{\arg\min} \sum_{j=0}^{71} \left| \tilde{I}(j\tau) - I^j \right|^2$$

with the same data set $D = \{(j\tau, I^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 71\}\}$ obtained as described in Section 4.3.1. Since we are not able to estimate the size of the effective population in advance, we set the initial guess $(\beta^{(0)}, \gamma^{(0)}, N^{(0)})^T = (1, 1, 10^6)^T$.

The computed results are $(\beta_m, \gamma_m, N_m)^T \approx (0.2587, 0.0444, 8593)$, the stopping criterion is given by the increment size with the tolerance 10^{-5} analogously to (4.3). It proved sufficient to set the number of iterations $N_I = 50$ to achieve the desired tolerance. The approximation of the measured values from the compartment I is accurate as illustrated in Figure 4.7. The other important characteristics have been improved as well. The estimated total number of recovered people at the and of the considered time interval is 7636, which is a good approximation of the true value 7750. The expected length of the infectious period is in this case approximately 22 days. This is close to the length of the potential maximal infectious period (in other words illness duration) estimated in meta-analysis [17], which is from 15 to 21 days. The estimate of the basic reproduction number $R_0 = \frac{\beta_m}{\gamma_m} \approx 5.8$ exceeds the values in the interval from 2.4 to 3.4 estimated by the meta-analysis [18].

To conclude, the presented method of the effective population considerably increased the accuracy of the basic SIR model in the situation when the SIR model itself failed due to high inconsistency of the measured data with the assumptions of the model.

4.4 COVID-19 – the SIQR model

The second model we want to analyse thoroughly is the SIQR model introduced in Section 1.3. The data set used for the parameter optimization remains the same and was described in Section 4.3.1. We proceed in the same manner as with the SIR model, although wider discussion regarding the data interpretation is now required.

Let us recapitulate the properties of the data set we possess. The data set, as described in Section 4.3.1, corresponds roughly to the pattern of the SIR model. From the original data, we have computed the values corresponding to all the three compartments S, I, and R. However, having examined the data set in more detail, we observe that the interpretation of the measured data corresponding to the SIQR model is different. The measured data originally belonging to the compartment I represent the number of positively tested subjects. It is evident that a positively tested person is immediately quarantined in real situation. It is therefore natural to assume that the numbers of positively tested people correspond to the compartment Q instead of I in the pattern of the SIQR model. All in all, we now have the data corresponding to the compartments Q and R. In this case, we are not able to compute the values corresponding to the compartment S, as the numbers regarding the compartment I are unknown in advance. This is why we skip the full-data approach from Section 4.3.2 and proceed directly to the incomplete-data approach.

In order to set up the input of the Algorithm 2 implemented in our program, we need to determine the initial conditions. This was trivial in the case of the SIR model, because the data set fitted precisely the pattern of the model. Considering the SIQR model, we have to determine the values of $S(0)$ and $I(0)$. It suffices to find the value of $I(0)$, the former can then be expressed as

$$S(0) = N - I(0) - Q(0) - R(0).$$

It is clear that value of $I(0)$ must be greater than the value of $Q(0)$, since the epidemic is at its beginning and it grows. Our choice is based on the following considerations: The length of the presymptomatic infectious period is estimated to be from 1 to 4 days, see [17]. We add two days for the testing procedure getting in total around 5 days. It follows that the value $I(0)$ corresponds roughly to the value Q^5 from the data set $D = \{(j\tau, Q^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 71\}\}$, it is however with some uncertainty. This is why we test and distinguish two cases here. In the end, we find out that the choice of the initial condition $I(0)$ does not affect the result to a high extent in case of the effective population approach.

First, let us take $I(0) \approx Q^5$. The procedure in now similar to the incomplete-data approach from Section 4.3.2 with the exception that we have three parameters now, i.e. α, β, and δ, corresponding to the SIQR model (1.4). The data set is modified as well, as described above. The functional to be optimized is of the form

$$(\alpha_m, \beta_m, \delta_m)^T = \underset{(\tilde{\alpha},\tilde{\beta},\tilde{\delta})^T \in \mathbb{R}^3}{\arg\min} \sum_{j=0}^{71} \left| \tilde{Q}(j\tau) - Q^j \right|^2$$

with the data set $D = \{(j\tau, Q^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 71\}\}$. The initial guess is given by $(\alpha^{(0)}, \beta^{(0)}, \delta^{(0)})^T = (1, 1, 0.01)^T$. The stopping criterion defined by the increment size is, mutatis mutandis, of the form (4.3).

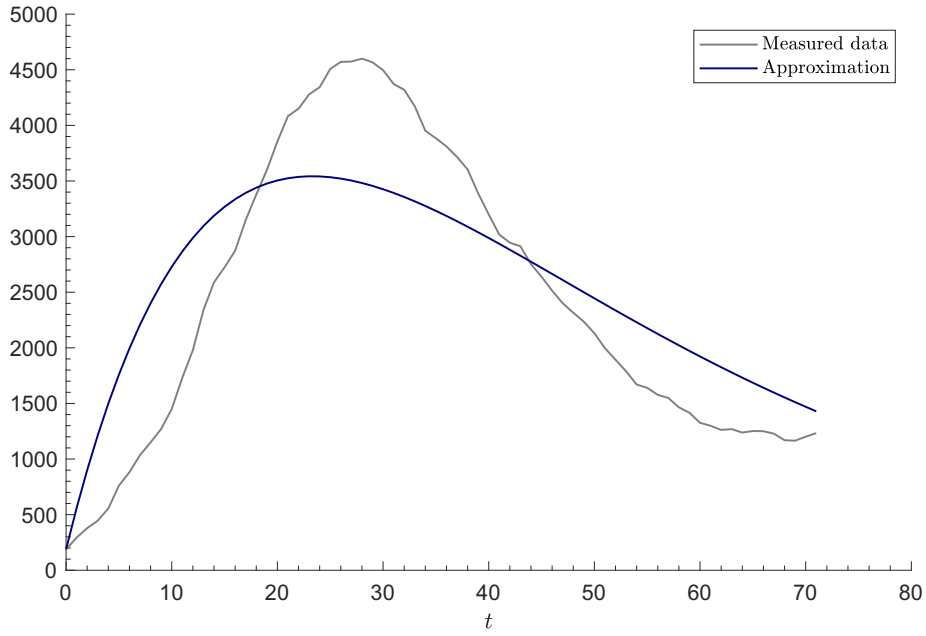

Figure 4.8: SIQR: Results of incomplete data approach for compartment Q

The convergence is slow, it takes approximately $N_I = 500$ iterations to achieve the result $(\alpha_m, \beta_m, \delta_m)^T \approx (0.7512, 0.7925, 0.0423)^T$ with the desired accuracy. As illustrated in Figure 4.8, the resulting approximation for compartment Q is only roughly in agreement with the measured data. This is why we proceed immediately to the effective population approach to see whether the results get better.

We distinguish two cases here. First, we set $I(0) \approx Q^5$ (the same as in the previous paragraph). The process is similar to the effective population approach from Section 4.3.2. It is desired to solve the following optimization problem:

$$(\alpha_m, \beta_m, \delta_m, N_m)^T = \underset{(\tilde{\alpha},\tilde{\beta},\tilde{\delta},\tilde{N})^T \in \mathbb{R}^4}{\arg\min} \sum_{j=0}^{71} \left| \tilde{Q}(j\tau) - Q^j \right|^2 .$$

Again, we have the data set $D = \{(j\tau, Q^j)^T \in \mathbb{R}^2, j \in \{0, \dots, 71\}\}$ and the same stopping criterion specified by the increment size. The initial guess is defined by $(\alpha^{(0)}, \beta^{(0)}, \delta^{(0)}, N^{(0)})^T = (1, 1, 0.01, 10^5)^T$. Let us remark that certain accuracy of the initial guess is required in this more sophisticated case, otherwise a blow-up often occurs.

The desired tolerance is achieved after $N_I = 90$ iterations and the final estimate is $(\alpha_m, \beta_m, \delta_m, N_m)^T \approx (0.1633, 0.3369, 0.0497, 10061)^T$. Figure 4.9 shows the resulting approximation for the compartment Q, which is in a good agreement with the observed data. It is however necessary to discuss other indicators as well. The total number of recovered people at the end of the considered time period is 8664, which exceeds the true value 7750. The maximal potential infectious period (the duration of the illness) is estimated to be $\frac{1}{\alpha_m} + \frac{1}{\delta_m} \approx 26$ days. The actual value estimated in [17] varies from 15 to 21 days. The basic reproduction number is predicted to be $R_0 = \frac{\beta_m}{\alpha_m} \approx 2.1$. The interval for the basic reproduction number estimated in [18] is from 2.4 to 3.4.

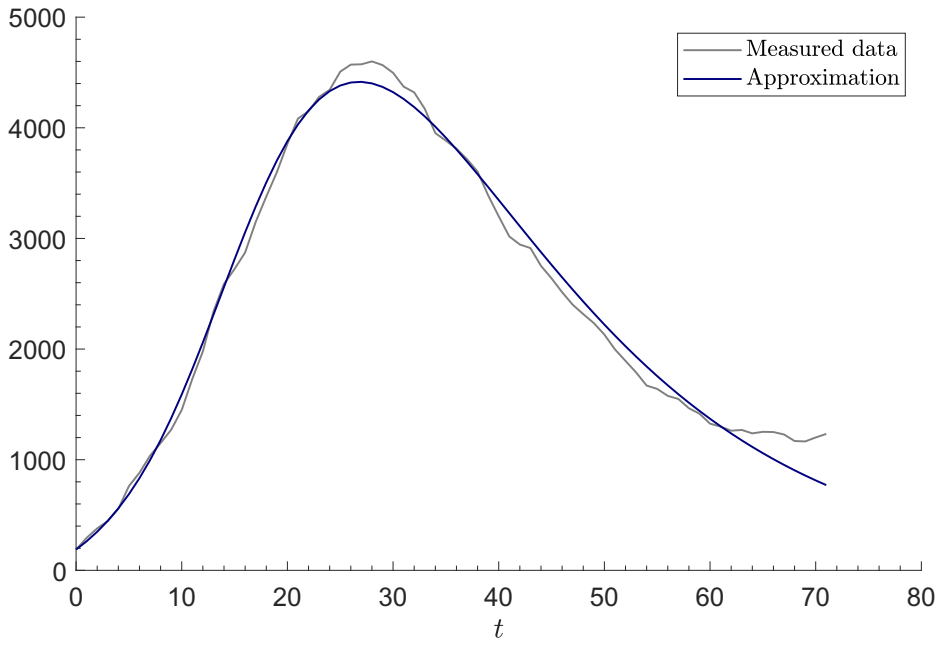

Figure 4.9: SIQR: Results of effective population approach for compartment Q

To demonstrate that the outcome is not sensitive to changes of the initial condition $I(0)$, we increase the value of $I(0)$ by 50 % of its original size. The other input parameters remain the same as in the previous case, i.e. we take the initial estimate $(\tilde{\alpha}^{(0)}, \tilde{\beta}^{(0)}, \tilde{\delta}^{(0)}, \tilde{N}^{(0)})^T = (1, 1, 0.01, 10^5)^T$ with the data set and the stopping criterion described in the previous paragraph. The resulting estimate is $(\tilde{\alpha}_m, \tilde{\beta}_m, \tilde{\delta}_m, \tilde{N}_m)^T \approx (0.1152, 0.3033, 0.0574, 10095)^T$. By comparison with the original estimate $(\alpha_m, \beta_m, \delta_m, N_m)^T \approx (0.1633, 0.3369, 0.0497, 10061)^T$, it can be concluded that the change in the initial condition does not affect the result to a high extent all the important properties being preserved.

We observe that the effective population approach improved the accuracy of the SIQR model. We have computed the important characteristics of the disease and we have found out that the computed values are in agreement with the medically observed values, although some small discrepancies occur.

4.5 Other advanced models

Apart from the basic SIR and the simple SIQR model, three other models were described in Section 1.4. In the previous Sections 4.3.2 and 4.4, we provided an extensive discussion of the results of the SIR and SIQR model. It turns out that the three other advanced models do not present us with additional information on top of the data obtained by the SIR and SIQR models. In other words, the results proved to be almost identical in terms of the effective population approach. Let us explain this in a more detailed manner.

Consider the SEIR model 1.3. When applied to the COVID-19 epidemiological data using the effective population approach, the model seems to "converge" to the SIR model: The rate of flow between the compartments E and I is more than a hundred times higher than the rate of flow between the compartments

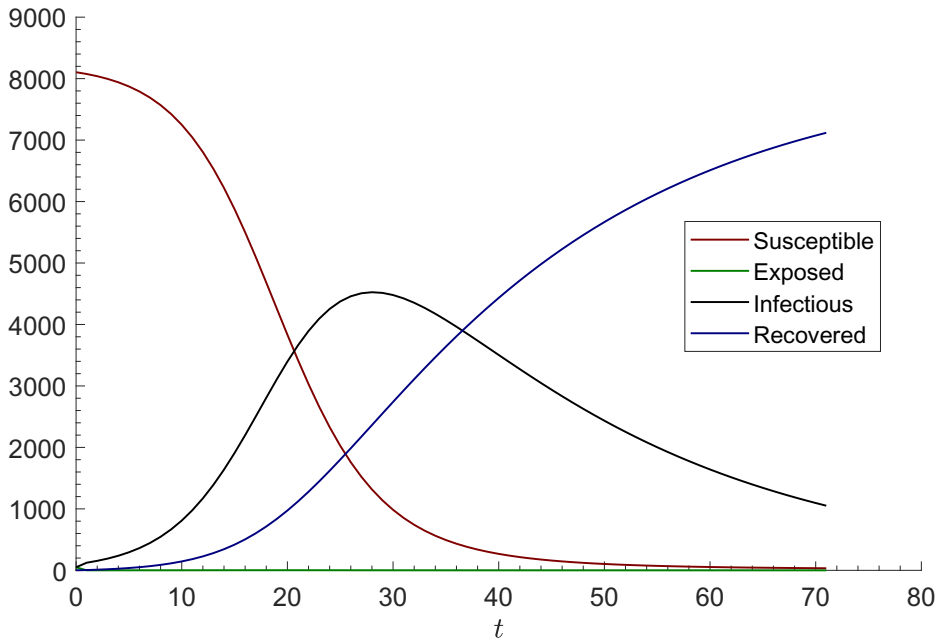

Figure 4.10: SEIR: Results of the effective population approach for the Czech Republic COVID-19 data

S and E. This causes the latent period to be extremely short. The SEIR model thus "approaches" the SIR model. To illustrate this phenomenon, Figure 4.10 shows the resulting estimate of the SEIR model. We can explain this phenomenon by the fact that our data set described in 4.3.1 neglects the latency period of the disease.

The other two models – SEIQR and SIQR No. 2 behave in a similar way. The model SIQR No. 2 (see Figure 1.4) "converges" to the first SIQR model 1.2. This is due to the fact that the rate flow between the compartments I and R tends to zero. The SEIQR model from Figure 1.5 "approaches" the SIQR model 1.2 as well. This is caused by the extremely high rate of flow between the compartments E and I.

Conclusion

On the topic of mathematical modelling in epidemiology, many papers were written from both the mathematical and epidemiological point of view. It is without any doubt that joint work of experts among both of those fields is the key to address this problem successfully. In this thesis, we have brought together some considerations from these two approaches.

The main goal of the thesis has been to estimate the parameters characterizing the coronavirus disease COVID-19 from available epidemiological data from the Czech Republic. In order to give insight into this complex issue, the reader was provided with the necessary theoretical background. We gave a brief introduction to compartmental models – a specific category of epidemiological models – with special focus on the basic SIR model. We formulated the parameter estimation problem mathematically as a non-linear least squares optimization problem, where we minimize a least squares functional measuring the discrepancy between the model outcome and measured data.

Since the considered compartmental models lead to a system of ordinary differential equations, we chose the sensitivity equation approach of [6] as a basis of the optimization technique. These equations describe the derivative of the solutions to the original differential equations with respect to its parameters, which can then be applied in the Gauss–Newton or Levenberg–Marquardt optimization algorithms. Several illustrative examples were given in order to elucidate these notions.

The presented methods were implemented in a complex program in MATLAB allowing us to perform a wide range of numerical experiments for various models. Before we proceeded to the COVID-19 epidemiological data, the methods had been extensively tested on fabricated data and on a well documented influenza epidemic in an English boarding school, [13], in order to examine their performance and convergence properties with respect to noisy or incomplete data.

Application to the influenza and COVID-19 epidemiological data has lead us to interesting conclusions. One of the crucial outcomes is the introduction of the effective population size. As presented, it is necessary to reduce the size of the population in which the epidemic is modelled in order to satisfy the unrealistic assumptions of the simple epidemiological models. Considering this reduced population size, considerably better accuracy of the estimates was accomplished. In effect, it is thus necessary to include the population size with the other unknown parameters to be optimized. The estimated parameters were then much more in accordance with the actual values based on the comparison with the medically observed data.

We were not able to find in the literature this approach of adding the population size to the list of unknown model parameters to be optimized. Some cases appeared, when the researchers cut down the population to a fraction of its actual size without any further commentary. This thesis provides a basis for these considerations along with a thorough discussion. We consider the concept of the effective population size to be a great possibility for further research from both mathematical and epidemiological perspective.

Bibliography

[1] James D. Murray. *Mathematical Biology I. An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer, New York, 3 edition, 2002.

[2] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Rev.*, 42(4):599–653, 2000.

[3] Gerda de Vries, Johannes Müller, Thomas Hillen, Birgitt Schönfisch, and Mark Lewis. *A course in mathematical biology – quantitative modeling with mathematical and computational methods*, volume 12 of *Mathematical modeling and computation*. SIAM, 2006.

[4] William O Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, 115(772):700–721, 1927.

[5] Klaus Schittkowski. *Numerical Data Fitting in Dynamical Systems: A Practical Introduction with Applications and Software*. Kluwer Academic Publishers, USA, 2002.

[6] Robert P. Dickinson and Robert J. Gelinas. Sensitivity analysis of ordinary differential equation systems - a direct method. *Journal of computational physics*, 21:123–143, 1976.

[7] John E. Dennis, Jr. and Robert B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, volume 16 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, USA, 1996.

[8] Olga Krylova and J. D. David Earn. Effects of the infectious period distribution on predicted transitions in childhood disease dynamics. *Journal of the Royal Society, Interface*, 10(81):20130098, 2013.

[9] Nuño M., Castillo-Chavez C., Z. Feng, and M. Martcheva. *Mathematical Models of Influenza: The Role of Cross-Immunity, Quarantine and Age-Structure*, pages 349–364. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[10] John Butcher. *Numerical Methods for Ordinary Differential Equations*. Wiley, Chichester, 2003.

[11] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[12] Christophe Fraser et al. Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science (New York, N.Y.)*, 324:1557–61, 05 2009.

[13] Anonymous report. Influenza in a boarding school. *British Medical Journal*, page 587, 4 March 1978.

[14] M. Komenda, M. Karolyi, V. Bulhart, J. Žofka, T. Brauner, J. Hak, et al. Onemocnění aktuálně. COVID-19: overview of current situation in the Czech Republic. URL: https://onemocneni-aktualne.mzcr.cz/covid-19. [Accessed 24-4-2021].

[15] Martin Komenda, Vojtěch Bulhart, Matěj Karolyi, Jiří Jarkovský, Jan Mužík, et al. Complex reporting of the COVID-19 epidemic in the Czech Republic: Use of an interactive web-based app in practice. *J Med Internet Res*, 22(5):e19367, May 2020.

[16] M. Komenda, M. Karolyi, V. Bulhart, J. Žofka, T. Brauner, J. Hak, et al. COVID-19: Celkový (kumulativní) počet osob s prokázanou nákazou dle krajských hygienických stanic včetně laboratoří, počet vyléčených, počet úmrtí a provedených testů (v2). https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19/nakazeni-vyleceni-umrti-testy.json. [Accessed 24-4-2021].

[17] Andrew William Byrne, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*, 10(8):e039856, 2020.

[18] M.A. Billah, M. M. Miah, and M. N. Khan. Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence. *PLoS One*, 15(11):e0242128, November 2020.

List of Figures

List of Tables