Parameter Estimation via Adjoint Functions in Epidemiological Reaction-Diffusion Models

Peter Heidrich and Thomas Götz

Abstract The current pandemic situation due to COVID-19 demonstrates the need for epidemiologic models to represent infection events as accurately as possible. An important factor is the mobility of the affected individuals which can be investigated with discrete or continuous spatial models. In this contribution, parameter estimation via adjoint functions is presented to fit a reaction-diffusion PDE system with epidemiological SIS model to data sets. For this purpose static and dynamic optimization methods are used to solve an L^2-norm based least squares problem. An artificial data set is generated to test the accuracy of the procedure. Subsequently, the PDE system is adapted to this data set using methods of optimal control theory. Unknown parameters like diffusivity and transmission rate can be determined. The noise in the data set is also taken into account by fitting the initial conditions. The results show that the method is well suited for this purpose and should be further used with real data sets.

1 Introduction

The example of the current COVID-19 pandemic clearly shows the significant influence of mobility on the spread of a disease. Mathematical-epidemiological models can address this using various techniques. The movement of people between separate patches such as airports, islands, cities etc. can be represented using Lagrangian movement for short-term stays or Eulerian movement for long-term

P. Heidrich (✉)
Mathematical Institute, University of Koblenz-Landau, Campus Koblenz, Koblenz, Germany

Magister Laukhard IGS Herrstein-Rhaunen, Herrstein, Germany
e-mail: heidrich@uni-koblenz.de

T. Götz
Mathematical Institute, University of Koblenz-Landau, Campus Koblenz, Koblenz, Germany
e-mail: goetz@uni-koblenz.de

migrations [5]. The modelling here is done via ordinary differential equations (ODE). However, since this point-by-point distribution of pathogens does not reflect reality on its own, spatial spread need to be taken into account as well. This can be achieved with reaction-diffusion systems which contain partial differential equations (PDE) [2, 10]. Consequently, we consider a system of the form

$$\partial_t u = \kappa \Delta u + f(u),$$

$$u = u_0, \quad t = 0,$$

$$\partial_\nu u = 0, \quad x \in \partial\Omega.$$

The goal is to fit this model to data sets. Unfortunately, several parameters are unknown in the epidemiological context, such as the transmission rate or even the parameters describing mobility. In addition, noisy data may be expected, for this the initial value condition shall be adjusted.

In this contribution, a parameter estimation via adjoint functions is tested. This corresponds to techniques from static and dynamic optimization. To investigate the accuracy of the method, we consider an artificially generated data set. Numerical simulations are performed to fit the model to this data set.

2 Model

In the following we consider the set $\Omega = (0, a) \times (0, b)$ as spatial coordinates and a time axis $(0, T)$ with resulting domain $V = \Omega \times (0, T)$. To model a spatial spread of an infectious disease, we use an epidemiological SIS model. The resulting reaction-diffusion system reads as

$$\partial_t S = \kappa_S \Delta S - \frac{\beta}{N} SI + \gamma I, \tag{1a}$$

$$\partial_t I = \kappa_I \Delta I + \frac{\beta}{N} SI - \gamma I, \tag{1b}$$

$$S = S_0, \ I = I_0, \quad t = 0, \tag{1c}$$

$$\partial_\nu S = \partial_\nu I = 0, \quad x \in \partial\Omega. \tag{1d}$$

The functions $S, I, N \in C^{2,1}(V)$ represent the densities of the compartments of susceptible (S) and infected (I) individuals and the total population density $N = S + I$ in coordinate x at time t.

Here, e.g. $\partial_t S = \frac{\partial S}{\partial t}$ stands for the time derivative of S and $\Delta S = \text{div}(\text{grad } S) = \frac{\partial^2 S}{\partial x_1^2} + \frac{\partial^2 S}{\partial x_2^2}$ stands for the Laplace operator. For the two compartments initial value conditions are given by $S_0, I_0 \in C^2(\Omega)$. At the boundary $\partial\Omega$ Neumann boundary conditions are implied, whereby $\partial_\nu S$ denotes the derivative of S to the direction of

the unit outward normal ν. In context, this means that none of the individuals leaves the region Ω. We also assume that $\int_\Omega I(x, 0)\, dx > 0$ holds with $S_0, I_0 \geq 0$. We define

$$\overline{N} := \int_\Omega N(x, 0)\, dx\,,$$

which stands for the total number of individuals at time $t = 0$. Due to the Neumann boundary conditions the Gauss's theorem delivers

$$\frac{\partial}{\partial t} \int_\Omega N(x, t)\, dx = \int_\Omega \kappa_S \Delta S + \kappa_I \Delta I\, dx = \int_{\partial\Omega} \kappa_S \partial_\nu S + \kappa_I \partial_\nu I\, d\omega = 0\,.$$

Thus, the total population is constant with value $\overline{N}$.

The parameters $\beta, \gamma > 0$ represent the transmission and recovery rates of the corresponding disease and $\kappa_S, \kappa_I > 0$ the diffusivity of the corresponding compartments. For simplicity, we assume that $\kappa_s = \kappa_I$ holds and β, γ are constants independent of x. For the derivation of such a model in one dimensional case and the operation of epidemiological models, we refer to [5].

SIS-based reaction-diffusion systems as in (1) have already been studied in [1, 3, 6–9]. The existence of a global and unique solution is shown, also for cases in which $\kappa_S \neq \kappa_I$ holds and β, γ are Hölder continuous functions over Ω. In [1] a Basic Reproduction Number is established on Sobolev space $H^1(\Omega)$ by

$$\mathcal{R}_0 = \sup_{\substack{\varphi \in H^1(\Omega) \\ \varphi \neq 0}} \left(\frac{\int_\Omega \beta\varphi^2}{\int_\Omega \kappa_I |\nabla\varphi|^2 + \gamma\varphi^2} \right)\,. \tag{2}$$

There is shown, that if $\mathcal{R}_0 < 1$ holds, the unique disease-free equilibrium DFE $= \left(\frac{\overline{N}}{|\Omega|}, 0 \right)$ is globally asymptotically stable and unstable for $\mathcal{R}_0 > 1$. The expression $|\Omega|$ here stands for the corresponding measure. On the other hand, for $\mathcal{R}_0 > 1$ the existence of a unique endemic equilibrium EE is shown.

Furthermore, we set $\kappa := \kappa_S = \kappa_I$ and substitute $S = N - I$. If we additionally define $u := \frac{I}{N}$, we receive a reduced system with $f(u) := \beta(1 - u)u - \gamma u$

$$\partial_t u = \kappa \Delta u + f(u)\,, \tag{3a}$$

$$u = u_0, \quad t = 0\,, \tag{3b}$$

$$\partial_\nu u = 0, \quad x \in \partial\Omega\,. \tag{3c}$$

The simplifying assumptions and the normalization are used to test the presented parameter fitting via adjoint functions. It is clear that in realistic situations much more complex models should be used.

3 Adjoint System

We now want to fit model (3) to data sets using adjoint functions known from optimal control theory. In the epidemiological context, this means parameter estimation of the transmission rate $\beta > 0$ and diffusivity $\kappa > 0$. The recovery rate $\gamma > 0$ can be assumed to be the reciprocal of the average infection duration and thus does not need to be fitted. Furthermore, we assume that the data is noisy and therefore the initial condition $u_0 \in C^2(\Omega)$ has to be adjusted. In the following, the function u^{DATA} contains the available data points and u_0^{DATA} the supposedly noisy initial value of the data set at $t = 0$.

We introduce an objective function $J : \mathbb{R}^2 \times C^2(\Omega) \to \mathbb{R}$

$$J(\beta, \kappa, u_0) := w_0 \|u - u^{DATA}\|_{L_V^2}^2 + w_1(\beta^2 + \kappa^2) + w_2 \|u_0 - u_0^{DATA}\|_{L_\Omega^2}^2 . \qquad (4)$$

The function u stands for the solution of the reaction-diffusion PDE system (3). The objective function includes the L^2-norm $\|g\|_{L_Y^2} := \left(\int_Y g(y)^2 \, dy\right)^{1/2}$ and corresponding normalizing weights $w_0 := 1/\|u^{DATA}\|_{L_V^2}^2$ respectively $w_2 := 1/\|u_0^{DATA}\|_{L_\Omega^2}^2$. The convex and radially unbounded regularization term $w_1(\beta^2 + \kappa^2)$ depends on a very small choosen weight w_1 whose influence is investigated in the subsequent simulations. Assuming one already has initial guess $\hat{\beta}, \hat{\kappa}$ for the parameters, a term of the form $w_1\left((\beta - \hat{\beta})^2 + (\kappa - \hat{\kappa})^2\right)$ can be used alternatively.

This leads to a minimization problem with dynamic constraints

$$\min_{\beta, \kappa, u_0} \; J(\beta, \kappa, u_0) \quad \text{subject to PDE system (3)}. \qquad (5)$$

A Lagrange function is introduced containing adjoint functions $z \in C^{2,1}(V)$

$$L(\beta, \kappa, u_0, u, z) := \int_V g \, dx dt + \psi + \int_V z \left(f(u) + \kappa \Delta u - \partial_t u\right) dx dt , \qquad (6)$$

whereby $g := w_0 \left(u - u^{DATA}\right)^2$ and $\psi := w_1(\beta^2 + \kappa^2) + w_2 \int_\Omega \left(u_0 - u_0^{DATA}\right)^2 dx$.

The necessary condition for a minimum $(\beta^*, \kappa^*, u_0^*, u^*, z^*)$ is fulfilled, if

$$0 = \nabla L := \left(\partial_\beta L, \partial_\kappa L, \partial_{u_0} L, \partial_u L, \partial_z L\right)$$

holds true. It should be noted that Gâteaux derivatives are needed for the derivatives of L to the directions u_0, u and z. This leads to the following system in $(\beta^*, \kappa^*, u_0^*, u^*, z^*)$:

(i) $0 = \partial_\beta \psi + \int_V z \partial_\beta f \, dx dt$, (Optimality Condition)

$0 = \partial_\kappa \psi + \int_V z \Delta u \, dx dt$,

(ii) $u_0 = u_0^{DATA} - \frac{z(x,0)}{2w_2}$, (Optimal Initial Condition)

(iii) $\partial_t z = -\partial_u g - z \partial_u f - \kappa \Delta z$, (Adjoint Equation)

$z = 0, \quad t = T$, (Transversality Condition)

$\partial_\nu z = 0, \quad x \in \partial\Omega$, (Adjoint Neumann Boundary Condition).

When L is derived in the z direction, the original PDE system (3) is recovered.

4 Numerical Simulations

From the analysis in Sect. 3, the gradient of L with respect to β and κ reads

$$\partial_\beta L = 2w_1\beta + \int_V z(1-u)u \, dx dt \tag{7a}$$

$$\partial_\kappa L = 2w_1\kappa + \int_V z \Delta u \, dx dt \tag{7b}$$

and we obtain the adjoint equation

$$\partial_t z = -2w_0 \left(u - u^{DATA}\right) - z(\beta(1-2u) - \gamma) - \kappa \Delta z. \tag{8}$$

The latter must be solved backward in time t due to the transversality condition. This is done using the forward-backward sweep method, see [4]. The performed algorithm can be found in Appendix 1. Solving the PDEs is done using finite differences

$$\Delta u_{i,j}^n \approx \frac{1}{h^2} \left(u_{i-1,j}^n + u_{i,j-1}^n - 4u_{i,j}^n + u_{i+1,j}^n + u_{i,j+1}^n\right) \tag{9}$$

and an explicit Euler-scheme

$$u_{i,j}^{n+1} = u_{i,j}^n + \tau(\kappa \Delta u_{i,j}^n + f(u_{i,j}^n)) \tag{10}$$

on the domain $V = \Omega \times (0,T)$ with $\Omega = (0,a) \times (0,b)$. The Neumann boundary conditions are implemented by $u_{k+1,j}^n = u_{k,j}^n$ etc., if index (k,j) stands for a point at the rectangular boundary $\partial\Omega$. In the following simulations we use the setting

- $h := 0.1, \tau := 0.001, a := 3, b := 2, T := 1$
- $x_1^i = ih$: $i = 0, \ldots, 30$ $x_2^j = jh$: $j = 0, \ldots, 20$ $t^n = n\tau$: $n = 0, \ldots, 1000$.

To test the procedure an artificial data set is generated with initial condition

$$u_0^{DATA}(x_1^i, x_2^j) := 0.02\delta_{(0.4, 0.6)}(x_1^i, x_2^j) + 0.1\delta_{(2,1)}(x_1^i, x_2^j) \tag{11}$$

whereby $\delta_{(\tilde{x}_1, \tilde{x}_2)}(x_1^i, x_2^j) = 1$, if $(x_1^i, x_2^j) = (\tilde{x}_1, \tilde{x}_2)$ and else $\delta_{(\tilde{x}_1, \tilde{x}_2)} = 0$. Subsequently, the state variable PDE (3) is solved with $\beta := 0.3$, $\kappa := 0.2$ and $\gamma := 0.1$. The received solution is called $\bar{u}$ in the following. To simulate noisy data, a normally distributed $q_{i,j}^n \sim \mathcal{N}(0, \sigma^2)$ is generated, so that the desired data set is calculated by

$$u^{DATA}(x_1^i, x_2^j, t^n) := \max\left(0, (1 + q_{i,j}^n) \cdot \bar{u}(x_1^i, x_2^j, t^n)\right). \tag{12}$$

5   Results and Conclusions

The application of the presented method is tested in three simulations with different initial values β_0, κ_0. The initial value for the initial condition u_0 is taken from the desired data set u^{DATA}. The resulting Table 1 and Fig. 1 in Appendix 2 show adequate parameter estimates. A test run without artificial noise on the data set resulted in the original values $\beta = 0.3$ and $\kappa = 0.2$. The simulations also show the effect of the weight w_1 of the regularization term on the minimization of the objective function J. Despite this disturbance, better results are obtained than without it. The prerequisite for this is a correspondingly small choice for w_1 which influences the convexity of the objective function in the respective parameters.

The present simulations show that the applied method works very well in this toy problem with self-generated data set. In principle, the procedure is suitable to perform such parameter estimations. In the next step, the method should be tested with real data sets. Depending on the disease, much more sophisticated epidemiological models may also be required. Mobility movements between patches, such as daily commuting or travelling, should also be added to the model. With respect to the PDE solution, other solution methods should also be tested, since the simple Euler method may be numerically unstable. In addition, a simple rectangular area was assumed in our example. In real cases, appropriate adjustments are necessary here.

Appendix 1

Algorithm 1 Pseudocode for the parameter estimation via adjoint functions

---

1: $\beta, \kappa, u^{DATA}, u_0^{DATA} \leftarrow$ load initial values and data
2: $u, z \leftarrow$ solve PDE for state variable and adjoint function
3: $J, \nabla J \leftarrow$ compute objective function and gradient regarding β and κ
4: $s_1 \leftarrow$ compute search direction for β and κ (Quasi-Newton (BFGS))
5: $s_2 \leftarrow (\tilde{u}_0 - u_0)$ compute search direction for u_0 with $\tilde{u}_0 = u_0^{DATA} - \frac{z(x,0)}{2w_2}$
6: **repeat**
7:     $J_{old} \leftarrow J$
8:     $\theta \leftarrow 1$
9:     $(\beta, \kappa) \leftarrow (\beta, \kappa) + \theta s_1$
10:     $u_0 \leftarrow u_0 + \theta s_2$
11:     $u, J \leftarrow$ update
12:     **repeat**
13:         $\theta \leftarrow 0.5\theta$
14:         $(\beta, \kappa) \leftarrow (\beta, \kappa) + \theta s_1$
15:         $u_0 \leftarrow u_0 + \theta s_2$
16:         $u, J \leftarrow$ update
17:     **until** $J \leq J_{old} + 0.001\theta s^T \nabla J_{old}$ (Armijo Rule)
18:     $z, \nabla J, s_1, s_2 \leftarrow$ update
19: **until** $\frac{\|J - J_{old}\|_2}{\|J_{old}\|_2} <$ TOL

Appendix 2

Table 1 The recovery rate is fixed with $\gamma := 0.1$. The algorithm stops with tolerance TOL $:= 10^{-6}$. The original parameters of the artificial data set are $\beta := 0.3$ and $\kappa := 0.2$. The artificial noise is generated with standard deviation $\sigma := 0.1$

Simulation	β	κ	J	w_1	Iterations
$\beta_0 := 0.5, \kappa_0 := 0.4$					
Best fit	0.2796	0.1994	$1.16 \cdot 10^{-4}$	10^{-07}	242
	0.2787	0.1994	$1.17 \cdot 10^{-4}$	10^{-08}	364
	0.2496	0.1988	$1.65 \cdot 10^{-4}$	10^{-09}	197
	0.2609	0.1990	$1.42 \cdot 10^{-4}$	0	240
$\beta_0 := 0.1, \kappa_0 := 0.5$					
Best fit	0.2659	0.1991	$1.33 \cdot 10^{-4}$	10^{-08}	181
	0.2251	0.1983	$2.36 \cdot 10^{-4}$	10^{-09}	364
	0.2857	0.1995	$1.11 \cdot 10^{-4}$	10^{-10}	345
	0.2753	0.1993	$1.20 \cdot 10^{-4}$	0	388
$\beta_0 := 1.0, \kappa_0 := 1.0$					
Best fit	0.2499	0.1988	$1.64 \cdot 10^{-4}$	10^{-10}	314
	0.2824	0.1994	$1.14 \cdot 10^{-4}$	10^{-11}	655
	0.2836	0.1994	$1.11 \cdot 10^{-4}$	10^{-12}	555
	0.2540	0.1988	$1.55 \cdot 10^{-4}$	0	304
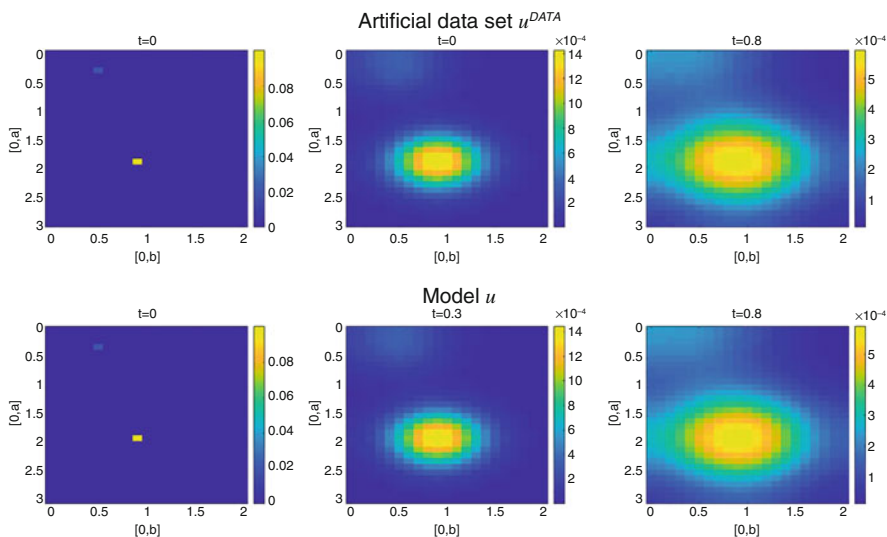
Fig. 1 Graphical Results for the Simulation with $\beta_0 := 1$, $\kappa_0 := 1$, $w_1 := 10^{-12}$

References

1. Allen, L.J.S., Bolker, B.M., Lou, Y., Nevai, A.L.: Asymptotic profiles of the steady states for an SIS epidemic reaction–diffusion model. Discrete Contin. Dynam. Syst., **21**, 1–20 (2008)
2. Britton, N.F.: Reaction-Diffusion Equations and Their Applications to Biology. Academic Press, London (1986)
3. Deng, K., Wu, Yixiang: Dynamics of a susceptible–infected–susceptible epidemic reaction–diffusion model. Proceedings of the Royal Society of Edinburgh: Section A Mathematics, **146(5)**, 929–946 (2016)
4. Lenhart, S., Workman, J.T.: Optimal control applied to biological models. CRC Press, 2007.
5. Martcheva, M.: An Introduction to Mathematical Epidemiology. Springer, New York (2015)
6. Peng, R.: Asymptotic profiles of the positive steady state for an SIS epidemic reaction–diffusion model. J. Diff. Eqns, **247**, 1096–1119 (2009)
7. Peng, R., Liu, S.: Global stability of the steady states of an SIS epidemic reaction–diffusion model. Nonlin. Analysis, **71**, 239–247 (2009)
8. Peng, R., Yi, F.: Asymptotic profile of the positive steady state for an SIS epidemic reaction–diffusion model: effects of epidemic risk and population movement. Physica D, **259**, 8–25 (2013)
9. Peng, R., Zhao, X.: A reaction–diffusion SIS epidemic model in a time-periodic environment. Nonlinearity, **25**, 1451–1471 (2012)
10. Webb, G.F.: A Reaction-Diffusion Model for a Deterministic Diffusive Epidemic. J. Math. Anal. Appl., **84**, 150–161 (1986)