ON THE USEFULNESS OF SET-MEMBERSHIP ESTIMATION IN THE EPIDEMIOLOGY OF INFECTIOUS DISEASES

Andreas Widder

ORCOS, Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology
Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

Abstract. We present a method, known in control theory, to give set-membership estimates for the states of a population in which an infectious disease is spreading. An estimation is reasonable due to the fact that the parameters of the equations describing the dynamics of the disease are not known with certainty. We discuss the properties of the resulting estimations. These include the possibility to determine best- or worst-case-scenarios and identify under which circumstances they occur, as well as a method to calculate confidence intervals for disease trajectories under sparse data. We give numerical examples of the technique using data from the 2014 outbreak of the Ebola virus in Africa. We conclude that the method presented here can be used to extract additional information from epidemiological data.

1. **Introduction.** Infectious diseases are an ever-current topic in human society. The impact on life expectancy, quality of life, and the economy of the affected communities can be immense. This is true not only of decades-long lethal pandemics such as HIV/AIDS ([6]) and violent disease outbreaks such as the recent Ebola virus epidemic ([17]), but also less threatening but persistent diseases such as influenza ([14]). Even diseases that affect only animals, such as foot and mouth disease, can have a noticeable impact on the economy ([7]).

Given the importance of the subject, it is consequential that many different, powerful mathematical modelling techniques have been developed that are widely used (for an introduction see for example [9, 11]). For the sake of simplicity we will presently only deal with models described by ordinary differential equations (ODEs), which in particular includes compartmental models. In these models the population is divided into several sub-populations (so called compartments) and the dynamics are described by giving the rates at which individuals progress from one compartment to another. As an example, among the simplest such models are susceptible-infected-recovered-models (SIR-models) where the population is divided into the three mentioned compartments and the dynamics are given by

$$\dot{S}(t) = -\beta I(t) S(t),$$
$$\dot{I}(t) = \beta I(t) S(t) - \gamma I(t),$$
$$\dot{R}(t) = \gamma I(t).$$

In order for this model to be able to correctly describe the progression of a disease, the two parameters β and γ, which describe the force of infection and recovery rate respectively, need to be known. The identification of the parameter vector $p = (\beta, \gamma)$ for a specific disease is carried out by statistical analysis of epidemiological data. Due to the importance of estimating parameters in epidemiological models there is a vast literature concerning this subject. We only mention [3, 8, 13, 18] and refer to references therein.

The result is that knowledge about the parameter p always entails some uncertainty. For example, the information may be given in the form of confidence intervals for each parameter component. The aim is then to extract as much information as possible about the disease from the available data. We present here a technique that can in some instances be used to gain additional information that is not obtained by other methods. If information about some parameter is only available as a member of a certain set, it is natural to ask whether we also can describe the possible outcomes as some set-membership. This idea is well established in control theory (see e.g. [12, 16]) and techniques have be developed to estimate these set-memberships. We show one of these techniques and discus how this method can be used in an epidemiological context to gain additional insight from given data, that is relevant for both researchers and policy makers.

The paper is organised as follows. In Section 2 we present the set-estimation procedure in a general setting. In Section 3 we discuss this procedure and explore some of its properties. In Section 4 we demonstrate the technique in some numerical examples by applying it to data from the Ebola virus outbreak in 2014.

2. Set-membership estimation.

2.1. Formulation of the problem.
We consider an ODE system

$$\dot{x}(t) = f(t, x(t), p), \quad x(0) = x_0. \tag{1}$$

Here, $x \in \mathbb{R}^n$ are the state variables, $x_0 \in \mathbb{R}^n$ is the known initial condition, $p \in \mathbb{R}^m$ is a vector of parameters, and the function $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is measurable in t, differentiable in x and p with Lipschitz continuous partial derivatives. The solution to (1) at time t while using the specific vector p will be denoted by $x[p](t)$.

We assume that the parameter p lies in a compact and convex set $\mathcal{P} \subset \mathbb{R}^m$. For each fixed time $T > 0$ we now want to estimate what values $x(T)$ can take, given that $p \in \mathcal{P}$. The best possible estimate is given by the reachable set

$$\mathcal{R}(T) = \{x \in \mathbb{R}^n : \exists p \in \mathcal{P} \text{ so that } x = x[p](T)\}.$$

To put this in an epidemiological context using the simple example given in the introduction, the vector $x(t) = (S(t), I(t), R(t))$ is composed of the individual compartments used in the model. The function $f(t, x, p)$ describes its dynamics and the vector $p = (\beta, \gamma)$ contains all parameters. The set $\mathcal{P}$ describes our knowledge about the parameter vector by denoting a region in which we assume the parameter p may lie. The reachable set $\mathcal{R}(T)$ then tell us which vectors (S, I, R) may describe the system at time T given our assumptions.

We will in the following describe how an estimate $\mathcal{E}(T) \supseteq \mathcal{R}(T)$ can be calculated. In many cases not all components of x are of immediate interest, but only the components lying in a subspace $L \subset \mathbb{R}^n$, e.g. if we are only interested in the number of infected individuals but not the recovered ones. We will therefore also consider

estimates $\mathcal{E}_L(T)$ that fulfil $\mathcal{E}_L(T) \supseteq \mathrm{pr}_L(\mathcal{R}(T))$, where pr_L is the projection operator onto L.

2.2. **Set-membership estimation.** Here we present a method to find approximations to the exact set-membership approximation $\mathcal{R}(T)$ that is well known in control theory ([5, 12]).

For a fixed time T and a unit vector $\ell \in \mathbb{R}^n$ we consider the optimisation problem

$$J_\ell(p) := \langle \ell, x[p](T) \rangle \to max, \tag{2}$$

where $x[p](T)$ satisfies (1) and $p \in \mathcal{P}$ ($\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^n$). We denote by x_ℓ an ε-solution of (2), i.e. there exists a $p \in \mathcal{P}$ such that $x_\ell = x[p](T)$ and x_ℓ fulfils

$$\langle \ell, x_\ell \rangle \geq \sup_{p \in \mathcal{P}} \langle \ell, x[p](T) \rangle - \varepsilon.$$

Obviously, $\mathcal{R}(T)$ lies in the half-space $\{x \in \mathbb{R}^n : \langle \ell, x - x_\ell \rangle \leq \varepsilon\}$.

Let $\Lambda = \{\ell_1, \ldots, \ell_k\}$ be a set of unit vectors for which we have ε-solutions. Then

$$\mathcal{R}(T) \subseteq \bigcap_{i=1,\ldots,k} \{x \in \mathbb{R}^n : \langle \ell_i, x - x_{\ell_i} \rangle \leq \varepsilon\} =: \mathcal{E}_\Lambda(T).$$

Consequently, $\mathcal{E}_\Lambda(T)$ is a set-membership estimation of $x(T)$. As an intersection of convex sets $\mathcal{E}_\Lambda(T)$ is itself convex. It is easy to show that if Λ is a δ-net on the unit sphere and both δ and ε converge to zero, then $\mathcal{E}_\Lambda(T)$ converges to the convex hull of $\mathcal{R}(T)$ in the Hausdorff sense.

If we are only interested in a set-membership estimation in the subspace $L \subset \mathbb{R}^n$, then we can simplify the above definition by requiring that the collection of unit vectors Λ belongs to the space L from which we can define the estimate

$$\mathrm{pr}_L(\mathcal{R}(T)) \subseteq \mathrm{pr}_L \left(\bigcap_{i=1,\ldots,k} \{x \in \mathbb{R}^n : \langle \ell_i, x - x_{\ell_i} \rangle \leq \varepsilon\} \right) =: \mathcal{E}_L(T).$$

For example, if in an SIR-model we are only interested in the number of infected individuals, we can choose as subspace L the space spanned by the vector $(0, 1, 0)$, which contains all the desired information about the vector $x = (S, I, R)$. The set Λ will then consist of the two vectors $(0, 1, 0)$ and $(0, -1, 0)$. Note that in order to calculate an estimate $\mathcal{E}(T)$ we need to solve (at least approximately) the optimisation problem (2).

2.3. **The optimisation problem.** For $\ell \in \mathbb{R}^n$ we want to maximise

$$J_\ell(p) = \langle \ell, x[p](T) \rangle.$$

Note that the solution to this problem gives the scalar projection of the vector $x[p](T)$ onto the unit vector ℓ. To continue the example given in the introduction where $x = (S, I, R)$, if $\ell = (0, -1, 0)$ then $J_\ell(p) = -I(T)$. So by maximising $J_\ell(p)$ we can deduce what the minimum value is that $I(T)$ can take for a $p \in \mathcal{P}$. We now prove a theorem that is crucial to solving this maximisation problem.

Denote by A^* the transpose of the matrix A. For a given $p \in \mathcal{P}$ and corresponding state trajectory $x[p](t)$ we define the adjoint function $\lambda[p](t)$ as the solution to the terminal value problem

$$\dot{\lambda}[p](t) = -f'_x(t, x[p](t), p)^* \lambda[p](t), \quad \lambda[p](T) = -\ell. \tag{3}$$

Theorem 2.1. *The functional J_ℓ is differentiable and the derivative is given by*

$$J'_\ell(p) = -\int\limits_0^T f'_p(t, x[p](t), p)^* \lambda[p](t)\, \mathrm{d}t. \tag{4}$$

Proof. We have

$$J_\ell(p) - J_\ell(\hat{p}) = \langle \ell, x[p](T) - x[\hat{p}](T) \rangle$$

and

$$\int\limits_0^T \left\langle \dot{\lambda}[\hat{p}](t), x[p](t) - x[\hat{p}](t) \right\rangle \mathrm{d}t = \langle \lambda[\hat{p}](T), x[p](T) - x[\hat{p}](T) \rangle$$

$$- \langle \lambda[\hat{p}](0), x[p](0) - x[\hat{p}](0) \rangle - \int\limits_0^T \langle \lambda[\hat{p}](t), \dot{x}[p](t) - \dot{x}[\hat{p}](t) \rangle\, \mathrm{d}t.$$

Using that $\lambda[\hat{p}](T) = -\ell$ and that $x[p](0) = x[\hat{p}](0)$ we can add these two equations to get

$$J_\ell(p) - J_\ell(\hat{p}) = -\int\limits_0^T \left\langle \dot{\lambda}[\hat{p}](t), x[p](t) - x[\hat{p}](t) \right\rangle + \langle \lambda[\hat{p}](t), \dot{x}[p](t) - \dot{x}[\hat{p}](t) \rangle\, \mathrm{d}t$$

$$= -\int\limits_0^T \left\langle \dot{\lambda}[\hat{p}](t), x[p](t) - x[\hat{p}](t) \right\rangle + \langle \lambda[\hat{p}](t), f(t, x[p](t), p) - f(t, x[\hat{p}](t), \hat{p}) \rangle\, \mathrm{d}t$$

$$= -\int\limits_0^T \left\langle \dot{\lambda}[\hat{p}](t), x[p](t) - x[\hat{p}](t) \right\rangle + \langle \lambda[\hat{p}](t), f'_x(t, x[\hat{p}](t), \hat{p})(x[p](t) - x[\hat{p}](t)) \rangle$$

$$+ \left\langle \lambda[\hat{p}](t), f'_p(t, x[\hat{p}](t), \hat{p})(p - \hat{p}) \right\rangle \mathrm{d}t + o(|p - \hat{p}|)$$

$$= \left\langle -\int\limits_0^T f'_p(t, x[\hat{p}](t), \hat{p})^* \lambda[\hat{p}](t)\, \mathrm{d}t, p - \hat{p} \right\rangle + o(|p - \hat{p}|).$$

The last equality is due to (3), while the fact that the remaining terms in the Taylor approximation are $o(|p - \hat{p}|)$ follows from Grönwall's Lemma, which tells us that $|x[p_1](t) - x[p_2](t)| \le c|p_1 - p_2|$. The result follows obviously from the above equality. $\qquad\square$

2.4. **Numerical procedure.** Using the result of the previous section we can derive a numerical gradient projection procedure for solving the problem (2). Starting with a point $p_0 \in \mathcal{P}$ we first solve the initial value problem (1) and then the terminal value problem (3). Using the formula in Theorem 2.1 we can calculate the gradient $J'_\ell(p_0)$. We then choose a step size α that maximises $J_\ell\left(\mathrm{pr}_\mathcal{P}\left(p_0 + \alpha J'_\ell(p_0)\right)\right)$ and set $p_1 = \mathrm{pr}_\mathcal{P}\left(p_0 + \alpha J'_\ell(p_0)\right)$. Iterating this process yields convergence to a critical point $\hat{p}$ so that after enough iterations the calculated value will be an ε-solution. For more detailed explanation of gradient projection algorithms, in particular the choice of the step size α, see for example [15]. Algorithm 1 below shows the pseudo code for a possible implementation of this procedure.

The optimisation problem (2) has to be solved for many different times T_n. If $\hat{p}$ maximises J_ℓ for T_n then $\hat{p}$ is a sensible initial guess for the maximisation of J_ℓ

at the time T_{n+1} which can make the algorithm more efficient. Furthermore, the adjoint equation (3) is linear which is another property that might in some cases simplify the calculation.

Another possible difficulty in this procedure is the calculation of the projection onto the set $\mathcal{P}$, which can be very complicated depending on its nature. However, if the set $\mathcal{P}$ is given by box constraints then this task is very simple. Such constraints arise for example if the set is described by confidence intervals of the individual parameter components. This is a common way the information about the parameter is available.

---

input : functions $f(t, x, p)$, $f'_x(t, x, p)$, and $f'_p(t, x, p)$; projection operator $\mathrm{pr}_{\mathcal{P}}$; initial condition x_0; initial guess p_0; tolerance ε; stopping condition $\tau < 1$; unit vectors $\ell_1, \ldots, \ell_k$; time discretisation $0 = T_0, \ldots, T_r = T$

for $j = 1, \ldots, k$ **do**

    $p_{new} = p_0$;

    for $i = 1, \ldots, r$ **do**

        $error = 1$;

        while $error \geq \tau$ **do**

            $p_{old} = p_{new}$;

            calculate $x[p_{old}](t)$ on $[0, T_i]$ using (1);

            calculate $\lambda[p_{old}](t)$ on $[0, T_i]$ with terminal condition $\lambda(T_i) = -l_j$ using (3);

            calculate $J'_{\ell_j}(p_{old})$ using (4) with T_i as the upper integral bound;

            choose step size α;

            $p_{new} = \mathrm{pr}_{\mathcal{P}} \left(p_{old} + \alpha J'_{\ell_j}(p_{old}) \right)$;

            $error = \|p_{new} - p_{old}\|$ or $\|J_{\ell_j}(p_{new}) - J_{\ell_j}(p_{old})\|$ or similar;

        end

        $x_{\ell_j}(T_i) = x[p_{new}](T_i)$;

    end

end

for $i = 1, \ldots, r$ **do**

    $\mathcal{E}(T_i) = \bigcap_{j=1,\ldots,k} \left\{ x \in \mathbb{R}^n : \langle \ell_j, x - x_{\ell_j}(T_i) \rangle \leq \varepsilon \right\}$;

end

output: set-membership estimation $\mathcal{E}(t)$ of $\mathcal{R}(t)$ at each point T_i

Algorithm 1: Set-membership estimation

---

3. **Discussion of the estimation.** Here we want to discuss some attributes of the set-membership estimation that is described in Section 2.

3.1. **Convexity of the estimate.** As already noted, the set-membership estimation $\mathcal{E}(T)$ is always convex. However, epidemiological models are usually non-linear, and the reachable sets of non-linear equations do not have to be convex ([4]). The convex hull of $\mathcal{R}(T)$ (which our procedure approximates) is therefore usually strictly larger than the reachable set itself, which limits the accuracy of our approach. We therefore present two ways in which this limitation can be circumvented.

The first way is to consider only estimations of a single component of x. Since $x[p](T)$ depends continuously on p and $\mathcal{P}$ is connected, $\mathcal{R}(T)$ is also connected. The projection onto a one dimensional subspace is therefore necessarily convex. The

estimate $\mathcal{E}_L(T)$ for a one dimensional space L therefore approximates the reachable set in L itself since it coincides with its convex hull. A further advantage of this approach is that the unit sphere in a one dimensional space is given by the two vectors $\{-1, 1\}$ and so no further discretisation is necessary.

If an approximation of the non-convex set $\mathcal{R}(T)$ in a higher dimensional space is needed, the procedure presented in Section 2 can be used in the following way. Create a partition of the set $\mathcal{P}$ into two convex sets $\mathcal{P}_A$ and $\mathcal{P}_B$. Calculate for each of them their respective set-membership estimation $\mathcal{E}_{\mathcal{P}_A}(T)$ and $\mathcal{E}_{\mathcal{P}_B}(T)$, i.e. the estimates for $x(T)$ given that $p \in \mathcal{P}_A$ and $\mathcal{P}_B$ respectively. Since $\mathcal{P}_A \subsetneq \mathcal{P}$, the diameter of $\mathcal{E}_{\mathcal{P}_A}(T)$ cannot be larger than the diameter of the estimate $\mathcal{E}_{\mathcal{P}}(T)$. Obviously, the same reasoning can be applied to $\mathcal{P}_B$. On the other hand, the union of $\mathcal{E}_{\mathcal{P}_A}(T)$ and $\mathcal{E}_{\mathcal{P}_B}(T)$ necessarily contains $\mathcal{R}(T)$. If necessary, the sets $\mathcal{P}_A$ and $\mathcal{P}_B$ can now themselves be partitioned further until we arrive at a partition $\mathcal{P}_1, \ldots, \mathcal{P}_k$ of $\mathcal{P}$ such that the diameter of $\mathcal{E}_{\mathcal{P}_i}(T)$ is smaller than some given tolerance $\eta > 0$. The union of these estimates is then an estimate of $\mathcal{R}(T)$ with a distance (in the Hausdorff sense) from $\mathcal{R}(T)$ of at most η.

This procedure may be numerically costly so that it might be more reasonable to consider the numerically cheaper estimates for the convex hull of $\mathcal{R}(T)$.

There are some special cases in which other procedures may be more effective to calculate the reachable set. For example, if the parameter set is one dimensional, i.e. given by an interval $[a, b]$, then the reachable set can be approximated by discretising the interval and calculating the solution for each point in the discretisation. Similar approaches become infeasible if the parameter vector is high dimensional.

3.2. **Identifying extremal values.** Set-membership estimates can be used to calculate worst-case-scenarios (or best-case-scenarios). If the quantity in question is one of the components of $x(T)$, then the set-membership estimate in the one dimensional space of this component calculates the maximum and minimum value this component can take. This allows us for example to determine the maximum number of people that could become infected during a disease outbreak, given the parameter constraints.

If the quantity in question is not one of the components of x, but its derivative can be computed, we can of course extend the vector x by the variable in question. For example, assume we are interested in the prevalence of the infection given by $\frac{I(t)}{N(t)}$, where $I(t)$ denotes the number of infected individuals and $N(t)$ the population size. Assume that $I(t)$ and $N(t)$ are components of x but that the prevalence is not. Since we are able to calculate $\dot{I}(t)$ and $\dot{N}(t)$, we can add the variable $x_{n+1}(t)$ with initial conditions $x_{n+1}(0) = \frac{I(0)}{N(0)}$ and derivative $\frac{\dot{I}(t)N(t) - I(t)\dot{N}(t)}{N(t)^2}$. The set-membership estimation can then be used to calculate the maximal prevalence that can occur during an outbreak, given the parameter constraints.

Furthermore, if we are interested in a quantity $y(x(T))$ for which the above is not possible, the set-membership estimate can still be useful. Given an estimate $\mathcal{E}(T)$ for $x(T)$ we can calculate $\max_{x \in \mathcal{E}(T)} y(x)$ to get an estimate for the maximal possible value of y.

Another useful aspect of this calculations is that not only do we get information about the extremal values of x, but also for which parameters p this values are obtained. In many cases it is not obvious which choice of parameters will maximise or minimise a certain variable. Identifying them can be of importance in constructing policies to minimise the impact of a disease.

3.3. **Confidence intervals for the trajectories.** A common way (see e.g. [2]) to obtain pointwise confidence intervals for the trajectories of various variables is the following: Calculate the maximum likelihood estimate $\hat{p}$ of the unknown model parameters and the associated variance-covariance matrix Σ. Then simulate large number or parameters with a normal distribution with mean $\hat{p}$ and variance matrix Σ. Calculate the trajectory for each of the simulated parameters and use e.g. the 2.5% and 97.5% quantile to construct the 95% confidence interval.

This is based on the fact that under certain circumstances the maximum likelihood estimate is asymptotically normally distributed. In particular this approach is only valid if enough data is available for the estimator to approach this asymptotic distribution. Set-membership estimates can be used to describe confidence intervals in a different way, without any need of further assumptions.

Under the model assumption, i.e. that the trajectories follow equation (1), every trajectory is given by $x[p](t)$ for some parameter vector p. Let $\mathcal{P}$ be a 95% confidence region for p, i.e. there is a 95% chance that the set $\mathcal{P}$ contains the true parameter $\hat{p}$[1]. Let $\mathcal{R}(T)$ be the reachable set of (1) at time T under the assumption that $p \in \mathcal{P}$. Then, letting $\mathbb{P}(A)$ describe the probability of the event A, we have

$$\begin{aligned}
\mathbb{P}\left(x[\hat{p}](T) \in \mathcal{R}(T)\right) =& \mathbb{P}\left(x[\hat{p}](T) \in \mathcal{R}(T) \mid \hat{p} \in \mathcal{P}\right) \mathbb{P}\left(\hat{p} \in \mathcal{P}\right) \\
& + \mathbb{P}\left(x[\hat{p}](T) \in \mathcal{R}(T) \mid \hat{p} \notin \mathcal{P}\right) \mathbb{P}\left(\hat{p} \notin \mathcal{P}\right).
\end{aligned} \tag{5}$$

By the definition of the reachable set the first probability on the right hand side is 1. Due to the fact that $\mathcal{P}$ describes the 95% confidence region we have $\mathbb{P}\left(\hat{p} \in \mathcal{P}\right) = 0.95$. The value of the second addend is unknown since we have in general no information about the probability that a trajectory of the system, with a parameter that lies outside of $\mathcal{P}$, will at time T be in the reachable set for the system with parameter values that lie in $\mathcal{P}$. However, since it is certainly non-negative we have

$$\mathbb{P}\left(x[\hat{p}](T) \in \mathcal{R}(T)\right) \geq 0.95.$$

This probability is of course to be interpreted in the sense of a confidence region. That is, if the reachable set is calculated from a data set, there is at least a 95% chance that the reachable set contains the true trajectory.

Thus the reachable set describes a confidence region for the trajectory of at least the confidence level of the region for the parameter. The same is of course true if we look at the trajectory of components of $x(T)$ in the subspace $L \subset \mathbb{R}^n$ and the respective reachable sets $\mathcal{R}_L(T)$. In particular, if we look at a single component of $x(T)$, then the set-estimate $\mathcal{E}_L(T)$ approximates a confidence interval (i.e. a one dimensional confidence region) with a level of at least that of $\mathcal{P}$. This approximation can be done to arbitrary accuracy and does not depend on any additional assumptions. This means for example that, unlike the method described at the beginning of this section, the procedure is valid even for small data sets.

4. **Numerical examples.** In this section we use a model and data presented by Althaus in [1] for the outbreak of the Ebola virus in 2014 in Guinea. The model

---

[1]We use here frequentist connotation. However, the same argument can be made in Bayesian terms, if $\mathcal{P}$ is replaced by a 95% credible region, i.e. a set such that the probability that the parameter p lies in $\mathcal{P}$ is 95%.
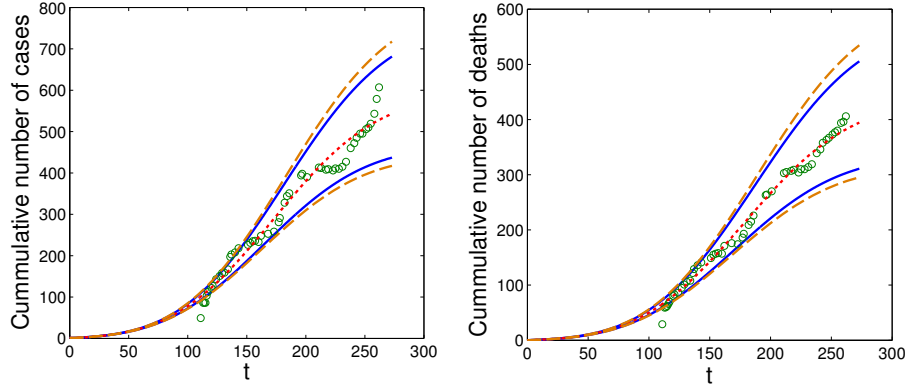
FIGURE 1. A comparison with the set membership estimation with
the data. The dotted line is calculated using the maximum likeli-
hood estimate. The solid lines are the set-membership estimation
calculated by using the 95% confidence intervals. The dashed lines
are calculated from a 95% confidence region using Bonferroni's
method.

equations are

$$
\begin{aligned}
\dot{S}(t) &= -\beta e^{-kt} \frac{S(t)I(t)}{N(t)}, \\
\dot{E}(t) &= \beta e^{-kt} \frac{S(t)I(t)}{N(t)} - \sigma E(t), \\
\dot{I}(t) &= \sigma E(t) - \gamma I(t), \\
\dot{R}(t) &= (1-f)\gamma I(t),
\end{aligned}
\tag{6}
$$

together with

$$
\begin{aligned}
\dot{C}(t) &= \sigma E(t), \\
\dot{D}(t) &= f\gamma I(t).
\end{aligned}
$$

Here, S denotes the susceptible individuals, E the exposed class, I the infectious
individuals, and R the recovered. The total number of individuals is given by
$N = S + E + I + R$. The additional two variables denote the cumulative number of
infected cases C and deaths D. The parameters $\frac{1}{\sigma}$ and $\frac{1}{\gamma}$ are the average durations
of incubation and infectiousness, respectively. The transmission rate is denoted by
β and the fatality rate by f. The factor k determines the exponential decay of
the transmission rate due to intervention measures. The initial conditions for the
vector (S, E, I, R, C, D) are $(10^6 - 1, 0, 1, 0, 1, 0)$. The two parameters σ and γ are
fixed at $\frac{1}{5.3}$ and $\frac{1}{5.61}$ respectively. The remaining three parameters are estimated
from data. The maximum likelihood estimate (and the 95% confidence interval) for
these parameters are[2]

---

[2]The values presented here are slightly different than the one in [1]. There, rounded values
have been given. The more accurate values presented here were kindly provided by the author of
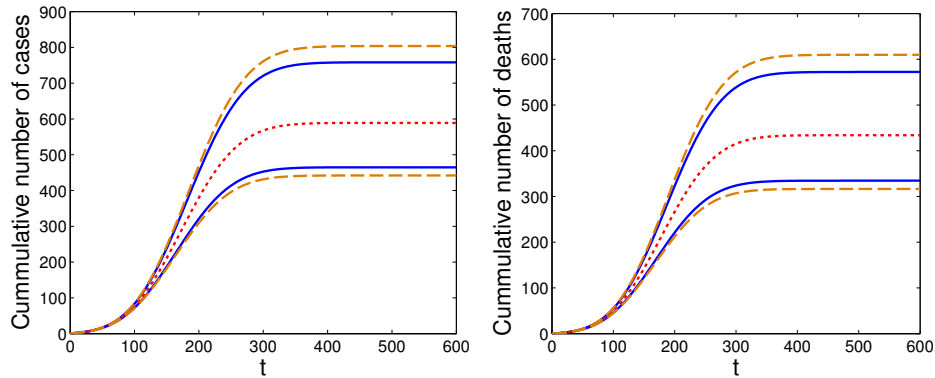[1] in personal correspondence.

FIGURE 2. A long term calculation of the maximum likelihood estimate (dotted line), and estimations using the 95% confidence intervals (dashed lines), and a 95% confidence region (solid lines).

$$\beta = 0.268574 \, (0.266924, 0.270189),$$
$$f = 0.7368011 \, (0.719586, 0.754707),$$
$$k = 0.0023379 \, (0.0022597, 0.002415).$$

The box described in $\mathbb{R}^3$ by these three intervals can be used as the set $\mathcal{P}$ of possible parameters. As an initial guess p_0 for the parameter we simply use the centre of the box, although other guesses such as the maximum likelihood estimate itself are of course also possible. With respect to the set $\mathcal{P}$ we note, however, that this box will not describe a 95% confidence region. There are different interpretations of the result when using this set. One is to say that this set describes the best knowledge we have about the location of the parameters, and the set-membership estimation therefore yields the best forecast that is possible, given our information. This interpretation may be particularly justified in cases where the estimates for different parameters come from different sources (and in particular different data sets). In cases such as the present one, where the data can be analysed, we may interpret the box described by the confidence intervals as a confidence region using the Bonferroni method ([10]). According to this method, to get a $100 - \alpha\%$ confidence region for an m-dimensional parameter, we calculate the $100 - \frac{\alpha}{m}\%$ confidence intervals for each component. The box described by these intervals is then a confidence region of the desired level. Therefore, in our case the 95% intervals provided yield an 85% confidence region for the parameter. To get an 95% confidence region we need to calculate the $98.\dot{3}\%$ confidence intervals for the individual components of the parameter vector. In Figure 1 we compare the two resulting set-membership estimations to the actual data[3]. In Figure 2 we calculate the estimations over a longer time period. Doing this for more confidence levels results in a graduation of confidence regions for the trajectory.

Extremal parameters. In simple models it is sometimes obvious what parameter configuration maximises (or minimises) certain compartments. For example, looking

---

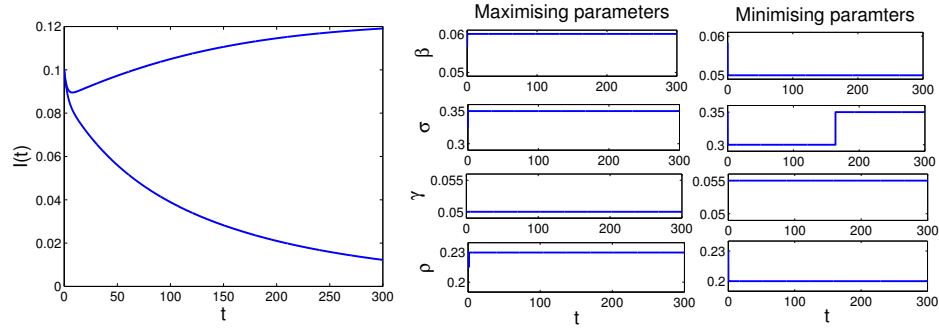[3]The calculations here were carried out in Matlab. The code has been made public in [21].

FIGURE 3. Set-membership estimates for $I(t)$ along with the parameter combinations that maximise and minimise $I(t)$. Even in the simple model presented here this combination may change with time.

at the equations in (6) we see that an increase in β will lead to a decrease of $S(t)$. So to maximise $S(t)$ we certainly need to minimise β. In more complicated models such relations are often hard to discern. As mentioned in Section 3, the set-membership estimation technique allows us to calculate for each t which parameter combination will yield a maximum for certain variables.

In order to capture at least one time dependent effect, we now consider a slight variation of Equations (6). We now assume that recovered individuals may return into the susceptible class and let ρ denote the rate at which this happens. We furthermore assume that the disease is not fatal and that no intervention is made. The resulting equations are then given by

$$\dot{S}(t) = -\beta \frac{S(t)I(t)}{N(t)} + \rho R(t),$$
$$\dot{E}(t) = \beta \frac{S(t)I(t)}{N(t)} - \sigma E(t), \tag{7}$$
$$\dot{I}(t) = \sigma E(t) - \gamma I(t),$$
$$\dot{R}(t) = \gamma I(t) - \rho R(t).$$

We assume that our knowledge about the parameters can be described by the inclusions

$$\beta \in [0.05, 0.06],$$
$$\sigma \in [0.3, 0.35],$$
$$\gamma \in [0.05, 0.055],$$
$$\rho \in [0.2, 0.23].$$

These numbers are not based on data and are chosen to highlight the dynamical features. In Figure 3 we show the set-membership estimation for $I(t)$ based on this information. We also show the parameter combinations that maximise and minimise $I(t)$. We see that even for the simple model we have chosen, the parameter combination that minimises $I(t)$ changes with time. For more complex models the time dependence of extremal parameter configurations may be more pronounced. Also note that it is not necessary that the extremal parameter combination is on the boundary of the interval for each component.

5. **Conclusion.** In this paper we discuss a technique to calculate set-membership estimations of trajectories for dynamical systems in which some parameters are not known. We apply this technique, which is well known in control theory, to epidemiological models, and emphasise certain aspects that are of particular use in this context.

Firstly, we describe how the estimates can be interpreted as a confidence region for the trajectories if the knowledge about the parameters is given in form of confidence intervals, as is often the case. Since this is a topic of great interest, further analysis of this method is desired. In particular it would be of interest to study the efficiency of the presented method compared with other established ones.

We furthermore explain how this method can be used to find parameters which yield the highest or lowest value for a state variable of choice. The technique presented here identifies the best and worst case scenario given all available information, and indicates which parameters should be influenced to achieve a more preferable outcome. This may be of particular interest to policy makers.

In this paper we focus on ODE models in which the dynamics depend on a finite dimensional parameter vector. However, this method can be altered for more general settings. For example, Theorem 2.1 can easily be adapted to include the case that the initial condition x_0 is itself an unknown parameter, or cases where the dynamics do not depend smoothly on the parameter p

It can also be applied to heterogeneous models with structured populations, where the dynamics are described by partial differential equations. In such models, initial conditions are given by a density function in an infinite dimensional space. These densities are usually not known. Also, parameters may depend on the population structure and therefore lie in some infinite dimensional function space. Some results have been presented to deal with unknown initial conditions ([19, 20]), and further research may provide results regarding unknown parameter functions.

Set-membership estimation is shown to be a flexible technique that can be applied in many different situations. It is not meant to supplant any existing technique, but to supplement them, and to help extract the maximum amount of information from epidemiological data.

REFERENCES

[1] C. L. Althaus, Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa, *PLOS Currents Outbreaks*, (2014).

[2] C. L. Althaus, N. Low, E. O. Musa, F. Shuaib and S. Gsteiger, Ebola virus disease outbreak in Nigeria: Transmission dynamics and rapid control, *Epidemics*, **11** (2015), 80–84.

[3] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*, Lecture Notes in Statistics, 151. Springer-Verlag, New York, 2000.

[4] R. Baier and M. Gerdts, A computational method for non-convex reachable sets using optimal control, in *2009 European Control Conference (ECC)*, IEEE, 2009, 97–102.

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, 1995.

[6] A. Bhargava and F. Docquier, HIV pandemic, medical brain drain, and economic development in Sub-Saharan Africa, *The World Bank Economic Review*, **22** (2008), 345–366.

[7] A. Blake, M. T. Sinclair and G. Sugiyarto, Quantifying the impact of foot and mouth disease on tourism and the UK economy, *Tourism Economics*, **9** (2003), 449–465.

[8] A. Capaldi, S. Behrend, B. Berman, J. Smith, J. Wright and A. L. Lloyd, Parameter estimation and uncertainty quantication for an epidemic model, *Mathematical Biosciences and Engineering*, **9** (2012), 553–576.

[9] O. Diekmann, H. Heesterbeek and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*, Princeton University Press, 2013.

[10] J. Hsu, *Multiple Comparisons: Theory and Methods*, CRC Press, 1996.

[11] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, 2008.

[12] A. B. Kurzhanski and P. Varaiya, *Dynamics and Control of Trajectory Tubes*, Springer, 2014.

[13] P. E. Lekone and B. F. Finkenstädt, Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study, *Biometrics*, **62** (2006), 1170–1177.

[14] N.-A. M. Molinari, I. R. Ortega-Sanchez, M. L. Messonnier, W. W. Thompson, P. M. Wortley, E. Weintraub and C. B. Bridges, The annual impact of seasonal influenza in the US: measuring disease burden and costs, *Vaccine*, **25** (2007), 5086–5096.

[15] E. Polak, *Computational Methods in Optimization: A Unified Approach*, vol. 77, Academic Press, New York-London, 1971 .

[16] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Second edition. Texts in Applied Mathematics, 6. Springer-Verlag, New York, 1998.

[17] M. R. Thomas, G. Smith, F. H. Ferreira, D. Evans, M. Maliszewska, M. Cruz, K. Himelein and M. Over, The economic impact of Ebola on Sub-Saharan Africa: Updated estimates for 2015, *World Bank Group*.

[18] T. Toni, D. Welch, N. Strelkowa, A. Ipsen and M. P. Stumpf, Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, *Journal of the Royal Society Interface*, **6** (2009), 187–202.

[19] T. Tsachev, V. M. Veliov and A. Widder, Set-membership estimation for the evolution of infectious diseases in heterogeneous populations, *Journal of Mathematical Biology*, **74** (2017), 1081–1106.

[20] V. Veliov and A. Widder, Modelling and estimation of infectious diseases in a population with heterogeneous dynamic immunity, *Journal of Biological Dynamics*, **10** (2016), 457–476.

[21] A. Widder, Matlab files for set-membership estimation, *Zenodo*, 2016.

E-mail address: andreas.d.widder@alumni.tuwien.ac.at