OPTIMIZING A HIGH-DIMENSION COMPARTMENT ODE MODEL

Jeremy Chiu,
APMA 923, Dr. JF Williams

December 2023

1 Project Overview

We consider the compartment ODE model presented by Guo-Wu[1], which describes the spread of Tuberculosis (TB) among the Canadian foreign-born population. Their ODE model is includes 5 dependent variables and 9 parameters, some of which are difficult to measure or estimate. Our goal is to find the parameters that give the global minimum of the model's prediction error (by comparing model output to real-world reported data).

1.1 Background and loss function

We consider the compartment ODE model presented by Guo-Wu[1], which describes the spread of Tuberculosis (TB) among the Canadian foreign-born population. Fundamentally, an SEIR-model is used, where the *exposed* category is partitioned into *early latent* and *late latent*. Figure 1 shows a flow diagram of their model.

After fixing initial conditions and parameter values, we use Matlab's `ode23` routine to solve the system of differential equations from [1], which returns population vs time. From the population, we compute the TB incidence

$$\text{Estimated TB Incidence } = \frac{100,000}{X(t)+E(t)+L(t)+T(t)+R(t)} \cdot (pwE(t)+vL(t)),$$

which can be compared to the reported TB incidence from Canadian reports[2] (see figure 2). The error is defined to be the difference between our computed (estimated) incidence and the reported incidence. We use parameters that are difficult to measure as input to the optimizer: q_1 and q_2 are the percentage of new immigrants that are undetected by Canadian TB screening, and the initial populations E_0, L_0, and R_0 are unknown; for a total of 5 variables to optimize across. Note that T_0 is reported, and X_0 is computed from the reported total foreign-born population, and all other parameters use estimates found from a literature search.

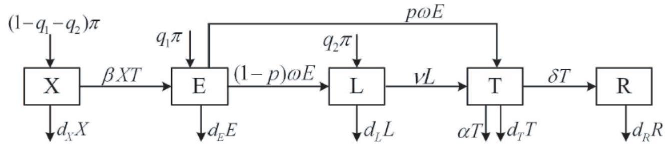

Figure 1: Flow diagram of the compartment ODE model used by [1].

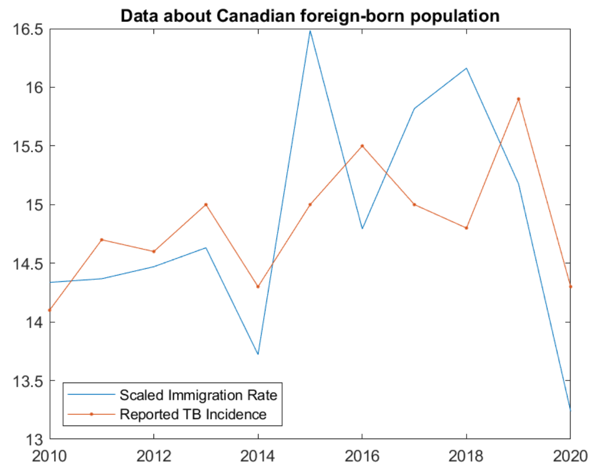

Figure 2: Reported incidence of Canadian foreign-born population. Annual immigration rate is also presented, scaled to be on same viewing rectangle. Notice how the number of immigrants drive the incidence rate.

1.2 Project Proposal Meeting

JF and I met on October 27 to discuss the project. Here were JF's suggestions about possible things to try:

- Do not over expand on problem background. Discuss how we solved it.
- Compare solvers.
- Try to find local sensitivity. Investigate derivatives with respect to parameters.
- Discuss `fmincon`.
- Global optimization toolbox.
- Find an upper bound on q_1 and q_2 using immigration data.
- Investigate convexity.

2 Finding the global min with `fmincon`

We used Matlab's constrained solver, `fmincon`. Our constraints were

$$q_1,\ q_2,\ E_0,\ L_0,\ R_0 \geq 0,\ \text{and}\ q_1 + q_2 \leq 1$$

Initial conditions were the original proportions presented in [1]. The solver does a decent job in minimization – figure 3 shows the estimated incidence after optimization, which is reasonably close to the reported incidence.

something I tried: – freeze thaw – optimize across q_1, q_2, and E_0, THEN L_0

References

[1] Hongbin Guo and Jianhong Wu. Persistent high incidence of tuberculosis among immigrants in a low-incidence country: Impact of immigrants with early or late latency. *Mathematical Biosciences and Engineering*, 8(3):695–709, 2011.

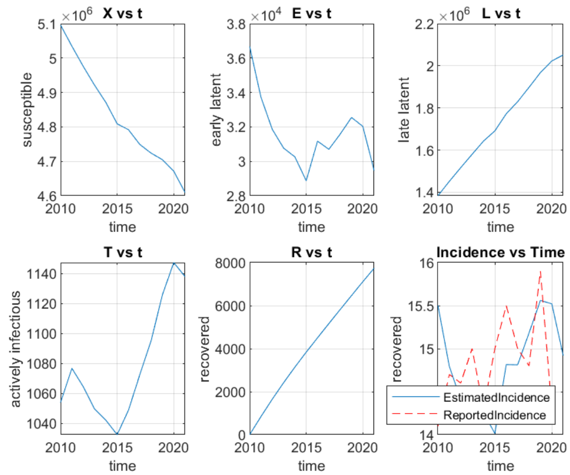

Figure 3: After optimization, the model reasonably estimate incidence.

[2] Mounchili A., Perera R., Henry C., Soualhine H., Sharma M., Yoshida C., Rabb M., and M. Carew. *Tuberculosis Surveillance in Canada 2010-2020 Summary Report.* Public Health Agency of Canada, 2022.