

Variational Inference Framework for Compartmental Disease Transmission Models

GitHub: <https://github.com/epidemic-research-team/>

June 1, 2020

1 Introduction

2 Model Goals Overview

The goal of this research is to utilise Variational Inference methods in order to increase the computational efficiency of the [7] paper.

1. Approximate the posterior
2. Apply VI methods on non-linear ODE system
3. Evaluate the results

3 Literature Review

1. Ghassemi, 2016 - Chapter 9

4 Method

4.1 ODE system model

The compartmental transmission model from [7] has a form as shown in Figure 1 [7, p.5].

The system of ODEs is then formulated as follow

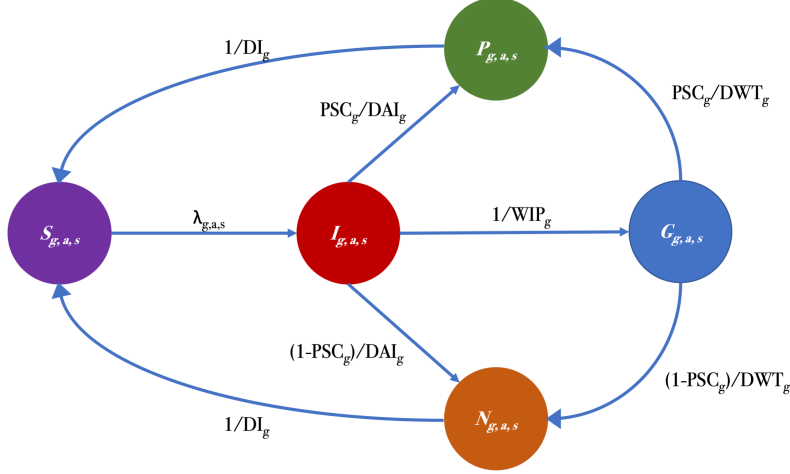


Figure 1: A compartmental Covid-19 transmission model.

$$\begin{aligned}
\dot{S}_{g,s,a} &= -\lambda_{g,s,a}(t)S_{g,s,a} + (P_{g,s,a} + N_{g,s,a})/DI_g + \frac{1}{r}S_{g,s,a-1} - \frac{1}{r}S_{g,s,a} + \\
&\quad \frac{1}{R} \sum_{g,s} (S_{g,s,20} + I_{g,s,20} + G_{g,s,20} + P_{g,s,20} + N_{g,s,20}) \times \\
&\quad \delta_1(a)(\pi_1\delta_1(s) + \pi_2\delta_2(s) + \pi_3\delta_3(s) + \pi_4\delta_4(s)) \\
\dot{I}_{g,s,a} &= \lambda_{g,s,a}(t)S_{g,s,a} - (1/WIP_g + 1/DAI_g)I_{g,s,a} + \frac{1}{r}I_{g,s,a-1} - \frac{1}{r}I_{g,s,a} \\
\dot{G}_{g,s,a} &= I_{g,s,a}/WIP_g - G_{g,s,a}/DWT_g + \frac{1}{r}G_{g,s,a-1} - \frac{1}{r}G_{g,s,a} \\
\dot{P}_{g,s,a} &= PSC_g(I_{g,s,a}/DAI_g + G_{g,s,a}/DWT_g) - P_{g,s,a}/DI_g + \frac{1}{r}P_{g,s,a-1} - \frac{1}{r}P_{g,s,a} \\
\dot{N}_{g,s,a} &= (1 - PSC_g)(I_{g,s,a}/DAI_g + G_{g,s,a}/DWT_g) - N_{g,s,a}/DI_g + \frac{1}{r}N_{g,s,a-1} - \frac{1}{r}N_{g,s,a}
\end{aligned}$$

From the Figure 1, the surveyed population is divided into the following non-overlapping compartments: S is for individuals who are at risk of Covid-19 infection, I is for infected individuals, G is for infected who de-

veloped serious symptoms and required ICU admission, P is for those who have recovered, and are seropositive and immune, N is for individuals who are recovered, immune but seronegative. The indices g , a and s indicate that every compartment is stratified by gender, age and level of social activity. Movements between compartments are occurring at per capita rates specified by the following parameters: **PSC** is the probability of becoming seropositive, **WIP** is the Covid-19 incubation period, **DAI** is the duration of asymptomatic (i.e. without Covid-19) infection symptoms, **DWT** is the duration of treatment for Covid-19, and **DI** is the duration of immunity. Subscripts denote stratification of parameters: for example, **DWT** _{g} means that in our model this parameter is gender-dependent. Finally, λ is the **force of infection** dependent on the proportion of individuals in I and is defined as

$$\lambda_{g,s,a} = \beta_g \sum_{s',\alpha'} \left\{ c_{g,s,s',\alpha,\alpha'}^* \frac{I_{g',s',\alpha'}}{S_{g',s',\alpha'} + I_{g',s',\alpha'} + G_{g',s',\alpha'} + P_{g',s',\alpha'} + N_{g',s',\alpha'}} \right\} \quad (1)$$

with β the probability of infection between individuals of different gender.

4.2 Model Inputs

Let X be a time dependent vector that includes the 5 states of the compartmental model, with $X_{g,s,a}(t) = [S_{g,s,a}(t), I_{g,s,a}(t), G_{g,s,a}(t), P_{g,s,a}(t), N_{g,s,a}(t)]$ and $t \in \{1, 2, \dots, T\}$. Also, assume that the observations of the model for time t are $O_{g,s,a}(t) = [D_{g,s,a}(t), Y_{g,s,a}(t)]$, with $D_{g,s,a}(t)$ the number of infected and $Y_{g,s,a}(t)$ represents the seroprevalence.

4.3 Priors

Next step, after they constructed the compartmental model, was to work on a Bayesian modelling framework that will involve the treatment the parameters of the non-linear system of ODE equations describing the epidemic as unknown random variables. In total, [7] model has 14 parameters with their relevant priors (Table 1)

Parameter	Interpretation	Prior
Transmission probability (males)	TRm	$\mathcal{U}[B_a, B_b]$
Transmission probability (females)	TRf	$\mathcal{U}[B_a, B_b]$
Average incubation period (males)	WIPm	$\mathcal{G}a(k_{WIPm}, \theta_{WIPm})$
Average incubation period (females)	WIPf	$\mathcal{G}a(k_{WIPf}, \theta_{WIPf})$
Average duration of treatment (males)	DWTm	$\mathcal{G}a(k_{DWTm}, \theta_{DWTm})$
Average duration of treatment (females)	DWTf	$\mathcal{G}a(k_{DWTf}, \theta_{DWTf})$
Average duration of asymptomatic (males)	DAIm	$\mathcal{G}a(k_{DAIm}, \theta_{DAIm})$
Average duration of asymptomatic (females)	DAIf	$\mathcal{G}a(k_{DAIf}, \theta_{DAIf})$
Average duration of immunity (males)	DIm	$\mathcal{U}(k_{DIm}, \theta_{DIm})$
Average duration of immunity (females)	DIf	$\mathcal{U}(k_{DIf}, \theta_{DIf})$
Probability of seroconversion (males)	PSCm	$\mathcal{B}e(\alpha_{PSCm}, \beta_{PSCm})$
Probability of seroconversion (females)	PSCf	$\mathcal{B}e(\alpha_{PSCf}, \beta_{PSCf})$
Observation error for incidence	σ	$inv\mathcal{G}a(k_\sigma, \theta_\sigma)$
Observation error scale for seroprevalence	A_Y	$\mathcal{G}a(k_{A_Y}, \theta_{A_Y})$

Table 1: Non-linear ODE Model Parameter Priors

4.4 Likelihood

Here, we present the likelihood model that we will use parameters vector θ includes the parameters in Table 1 and X includes the ODE states in Equation ???. For efficiency, we symbolise the priors in Table 1 as θ . The observed numbers of new diagnoses are assumed to be observed in Gaussian noise. This assumption is reasonable since the counts obtained are very large, so it is suitable to make a continuous distributional assumption. The likelihood model for the observed seroprevalences is assumed to be a beta distribution since it represents observations of proportions given the parameters.

$$\mathcal{L}(priors, X_{g,s,a}(1), \dots, X_{g,s,a}(t); O_{g,s,a}(1), \dots, O_{g,s,a}(t)) = \quad (2)$$

$$\Pi_a \Pi_s \Pi_g \Pi_t p(O_{g,s,a}(t) | X_{g,s,a}(t), priors) \quad (3)$$

$$\Pi_a \Pi_s \Pi_g \Pi_t \mathcal{N}(D_{g,s,a}(t) | \mu, \sigma) \mathcal{B}e(Y_{g,s,a}(t) | A_Y, B_Y) \quad (4)$$

$$(5)$$

with the normal distribution mean as

$$\mu = \frac{1}{WIP} \times \frac{I}{S + I + G + N + P} \times 1000 \quad (6)$$

and the beta distribution scale as A_Y and shape

$$B_Y = A_Y \left(\frac{1}{P} - 1 \right) \quad (7)$$

4.5 Posterior

The posterior of a Bayesian model has the form

$$p(\theta|X, y) \propto p(y|X, \theta)p(\theta|X) \quad (8)$$

by replacing the likelihood with Equation ?? and the priors (θ) with Table 1, we solve for

$$p(\theta|X, y) \propto \Pi_a \Pi_s \Pi_g \Pi_t \mathcal{N}(D_{g,s,a}(t)|\mu, \sigma) \mathcal{B}e(Y_{g,s,a}(t)|A_Y, B_Y) \times \Pi_{i=priors} \theta_i \quad (9)$$

References

- [1] M. Baguelin, S. Flasche, A. Camacho, N. Demiris, E. Miller, and W. J. Edmunds. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modelling study. *PLoS Med*, 10(10):e1001527, 2013.
- [2] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface*, 7(47):873–885, 2010.
- [3] L. Fumanelli, M. Ajelli, P. Manfredi, A. Vespignani, and S. Merler. Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. *PLOS Computational Biology*, 8(9):1–10, 09 2012.
- [4] H. J. Jones. Notes on \mathcal{R}_0 , 2007.
- [5] M. Keeling, P. Rohani, and P. U. Press. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.
- [6] P. Klepac, A. J. Kucharski, A. J. Conlan, S. Kissler, M. Tang, H. Fry, and J. R. Gog. Contacts in context: large-scale setting-specific social mixing matrices from the bbc pandemic project. *medRxiv*, 2020.

- [7] I. A. Korostil, G. W. Peters, J. Cornebise, and D. G. Regan. Adaptive markov chain monte carlo forward projection for statistical analysis in epidemic modelling of human papillomavirus. *Statistics in Medicine*, 32(11):1917–1953, 2013.
- [8] M. Li. *An Introduction to Mathematical Modeling of Infectious Diseases*. Mathematics of Planet Earth. Springer International Publishing, 2018.
- [9] S. Ma and Y. Xia. *Mathematical Understanding of Infectious Disease Dynamics*. Lecture Notes Series, Institute for Mathematical Sciences. World Scientific, 2009.
- [10] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine*, 5(3):1–1, 2008.
- [11] K. Prem, A. R. Cook, and M. Jit. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology*, 13:1–21, 09 2017.