Manuscript outline/draft

Ben Bolker   Daniel Park   David Earn

11:38 20 July 2016

Introduction

The goal is a *brief* (if possible) pedagogically oriented paper, aimed at biomathematicians/students coming for the first time to the idea of fitting curves to epidemic models. Without going too deeply into any one area (likelihood theory, optimization, ...), we aim to provide a primer and pitfall-avoidance guide for fits of simple ODE models (focusing particularly on SIR models) to epidemic data.

Potential topics

trajectory-matching vs. gradient-matching what's the difference? when does it matter? (this paper will focus on trajectory matching) (Ellner et al., 2002; Bolker, 2008)

early-epidemic vs. whole-epidemic lots of work on fitting early epidemics, e.g. Chowell et al. (2007), Ma et al. (2014); that's not what we're doing here ... also, we're only going to discuss general principles here, focus on simple SIR (maybe comment on SEIR), but not worry about extensions (waning immunity, Ebola/funeral-transmission ... etc.)

prevalence vs. incidence : need to distinguish these cases (typical cases will involve incidence data). Comment on pitfalls of cumulative-incidence approaches (Chowell) without appropriate corrections (maybe discussed elsewhere? King et al. (2015) ?) (Do we need to worry about mortality vs. incidence??)

least-squares vs. likelihood formulations equivalence of least-squares and likelihood approaches; advantages of likelihood in providing a framework for inference (confidence intervals). Mention Wald vs likelihood profile CIs.

optimization issues multiple maxima (if they exist?); ridges (Polansky et al., 2009). Optimization frameworks (e.g. Nelder-Mead vs quasi-Newton); integrating on $\log(I)$ scale; sensitivity equations (Raue et al., 2013). Starting points (auto-start methods)

identifiability/estimability issues (still in progress!) expected difficulties in optimization, especially with vague starting points. Something about solutions ... fixing parameters (with attendant dangers; Elderd et al. (2006). Bayesian priors, bounds ... ?

advanced methods *brief* pointers to relevant literature here. Dealing with combined process/measurement error (King et al., 2015); iterated filtering, TSIR, etc.. Bayesian methods (Stan, `debInfer`, ...)

To do

- what shall we use as a case study? sims; but also some real data set (pref. *not* Bombay data set, because of Bacar (2012) ...)

References

Bacar, N. (2012, February). The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality. *Journal of Mathematical Biology 64*(3), 403–422.

Bolker, B. M. (2008). *Ecological Models and Data in R*. Princeton, NJ: Princeton University Press.

Chowell, G., H. Nishiura, and L. M. Bettencourt (2007, February). Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *Journal of The Royal Society Interface 4*(12), 155–166.

Elderd, B. D., V. M. Dukic, and G. Dwyer (2006, October). Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *Proceedings of the National Academy of Sciences 103*(42), 15693 –15697.

Ellner, S. P., Y. Seifu, and R. H. Smith (2002). Fitting population dynamic models to time-series data by gradient matching. *Ecology 83*(8), 2256–2270.

King, A. A., M. D. d. Cells, F. M. G. Magpantay, and P. Rohani (2015, May). Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. B 282*(1806), 20150347.

Ma, J., J. Dushoff, B. M. Bolker, and D. J. D. Earn (2014). Estimating Initial Epidemic Growth Rates. *Bulletin of Mathematical Biology 76*(1), 245–260.

Polansky, L., P. de Valpine, J. O. Lloyd-Smith, and W. M. Getz (2009, August). Likelihood ridges and multimodality in population growth rate models. *Ecology 90*(8), 2313–2320.

Raue, A., M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, S. Hug, C. Kreutz, B. D. Harms, F. J. Theis, U. Klingmller, and J. Timmer (2013, September). Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. *PLOS ONE 8*(9), e74335.