A mathematical introduction to Neural Networks and Neural Ordinary Differential Equations

Argimiro Arratia
argimiro@cs.upc.edu
http://www.cs.upc.edu/~argimiro

CS, Universitat Politécnica de Catalunya

Yet another Python package for NODE: JAX

JAX: https://pypi.org/project/jax/
And also check this Deep learning perspective (with examples in JAX): **Deep Implicit Layers**:
http://implicit-layers-tutorial.org/
Neural ODE as deep implicit layers model. see Chapter 3.

JAX

Google JAX is a machine learning framework for transforming numerical functions. It brings together a modified version of Autograd (automatic obtaining of the gradient function through differentiation of a function) and TensorFlow's XLA (Accelerated Linear Algebra)
Official Guide: https://pypi.org/project/jax/

JAX basic functions

JAX is designed to follow the structure and workflow of NumPy and works with various existing frameworks such as TensorFlow and PyTorch. The primary functions of JAX are:

- grad: automatic differentiation
- jit: Just-in-time compilation
- vmap: auto-vectorization
- pmap: parallelization for matrix multiplication

Check a summary in
https://en.wikipedia.org/wiki/Google_JAX

Building a deep learning model with Python and JAX

see the notebook: NODEwJAXbasics.ipynb

NODE Applications: Epidemic Modelling

Epidemic models

Epidemic models attempt to capture the dynamics in the spreading of a disease (or idea, computer virus, product adoption).

Central questions they try to answer are:

- How do contagions spread in populations?
- Will a disease become an epidemic?
- Who are the best people to vaccinate?
- Will a given YouTube video go viral?
- What individuals should we market to for maximizing product penetration?

Full mixing in classic epidemiological models

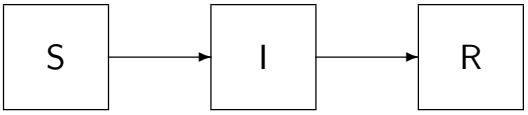
Full mixing assumption

In classic epidemiology, it is assumed that every individual has an equal chance of coming into contact with every other individual in the population

The SIR model
Allowing *recovery* and *immunity*

In the SIR model, individuals can be in one of two states:

- *susceptible* (S), or
- *infective* (I),or
- *recovered* (R)



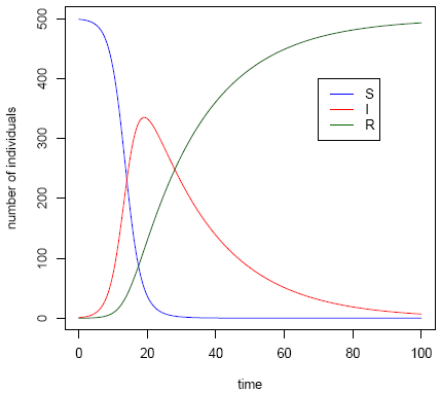
The parameters of the SIR model are

- β infection rate: probability of contagion after contact per unit time
- γ recovery rate: probability of recovery from infection per unit time

The SIR model
Dynamics

$$\frac{dS}{dt} = -\beta SI \qquad \frac{dI}{dt} = \beta SI - \gamma I \qquad \frac{dR}{dt} = \gamma I$$

The solution to this system (with $S + I + R = 1$, think of proportions of the population N) is not analytically tractable, but solutions look like the following (for ($\beta < \gamma$):



Solving SIR with NODE

$$NNet\left(\left[\frac{dS}{dt}, \frac{dI}{dt}, \frac{dR}{dt}\right] = [-\beta S(t)I(t), \beta S(t)I(t) - \gamma I(t), \gamma I(t)]\right)$$

We apply the neural network ODE paradigm with the SIR(θ) dynamics with parameters $\theta = [\beta, \gamma]$. This means:

- Give arbitrary (initial) values to θ
- Repeat for a number of epochs:
- Forward evaluation: resolve with numerical method the SIR(θ)
- Evaluate the loss with respect to final values of some given true trajectory of I (or S or R)
- backpropagation of gradients: to calibrate θ.

see the notebook: NODEjaxSIRmodel.ipynb

The SIR model I
A threshold phenomenon

Now we are interested in considering the *fraction of the population that will get sick* (i.e. size of the epidemic), basically captured by $R(t)$ as $t \to \infty$

Substituting $dt = \frac{dR}{\gamma I}$ from the third equation into $dS = -\beta SIdt$ and solving for S (assuming $R(0) = 0$), we obtain that

$$S(t) = S(0)e^{-\frac{\beta}{\gamma}R}$$

and so

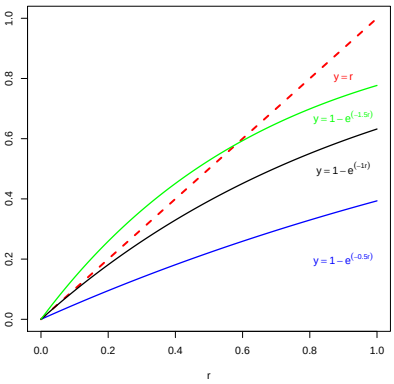
$$\frac{dr}{dt} = \gamma(1 - R - S(0)e^{-\frac{\beta}{\gamma}R})$$

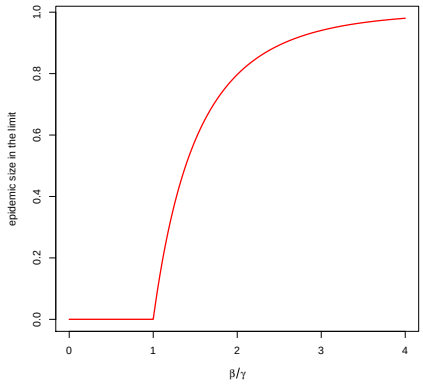
As $t \to \infty$, we get that $R(t)$ stabilizes and so $\frac{dR}{dt} = 0$, thus:

$$R = 1 - S(0)e^{-\frac{\beta}{\gamma}R}$$

The SIR model II
A threshold phenomenon

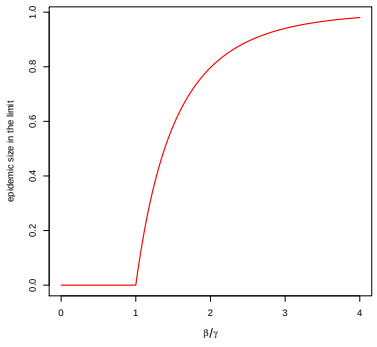
Assume that $S(0) \approx 1$, since typically we start with a small nr. of infected individuals and we are considering large populations, and so $R = 1 - e^{-\frac{\beta}{\gamma}R}$





The SIR model III
A threshold phenomenon

- if $\frac{\beta}{\gamma} \leq 1$ then no epidemic occurs
- if $\frac{\beta}{\gamma} > 1$ then epidemic occurs
- $\beta = \gamma$ is the *epidemic transition*



The SIR model
The basic reproduction number R_0

Basic reproduction number R_0

R_0 is the average number of additional people that a newly infected person passes the disease onto before they recover[a].

[a] It is defined for the early stages of the epidemic and so one can assume that most people are in the susceptible state.

- $R_0 > 1$ means each infected person infects more than 1 person and hence the epidemic grows exponentially (at least at the early stages)
- $R_0 < 1$ makes the epidemic shrink
- $R_0 = 1$ marks the *epidemic threshold* between the growing and shrinking regime

In the SIR model, $R_0 = \frac{\beta}{\gamma}$

Another Epidemic model using NODE

Raj Dandekar, Chris Rackauckas and George Barbastathis (2020). A Machine Learning-Aided Global Diagnostic and Comparative Tool to Assess Effect of Quarantine Control in COVID-19 Spread. *Patterns, v1 (9)*