MODELLING TRAJECTORIES AND DENSITIES ACROSS TIME

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

Rajeev Rajaram
Brian Castellani
Kent State University

Outline

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \bigtriangledown \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Case based density approach
- Ordinary differential equations
- Notion of transport in case trajectories- advection equation
- Modelling cases at the individual level – the vector space approach
- Modelling cases at the aggregate level – the density approach
- Uniqueness of the approach
- Applications

Case based density approach

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Every case has a certain number of measurements.
- Cases (and hence the measurements) evolve as a function of time.
- The movement of each case is its case trajectory.
- The movement of a distribution (or density) of cases is the density trajectory.
- Our approach aims to capture the majority and minority trends of case trajectories, and also models the motion of densities.

Ordinary Differential equations

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Traditionally used to model mechanical, electrical, chemical, biological and ecological processes.
- Examples are population growth, predator-prey models, passive electronic circuits, mechanical stresses and strains, first and second order chemical reactions, rate of forgetting etc.
- First order differential equation – Given velocity find position. Need initial conditions.
- Solution is a trajectory which is a function of time (just one independent variable)

ODEs

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- For our purposes, ODEs can be used to model case trajectories.

$$\dot{x} = f(x); x(0) = x_0$$

- Need velocity information of cases in order to compute case trajectories by solving the ODE.
- Velocity can be a function of the current case profile, and current time.
- Solving an ODE amounts to computing the case trajectory given the velocity information

Advection equation – transport of density of cases

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Transforms the motion of individual cases to the motion of a density of cases.

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho|_{\Gamma_i} = 0; \rho(x,0) = \rho_0(x)$$

- Requires the initial distribution of case profiles, and the velocity vector field of cases (same as the one used in the ODE), and can compute the motion of the initial density assuming that the total number of cases is a constant (called mass conservation property).

- Used in modeling of transport phenomena such as fluid dynamics (oil spill), traffic on streets.

Advection equation – transport of density of cases

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- This is a partial differential equation (PDE) since the density is a function of the case profile and time (more than one independent variable for the state).
- Prior work – Using the advection PDE to study the aggregate motion of states instead of individual trajectories to coin a new notion of stability called **almost everywhere uniform stability.**
- Motion of individual states are described by nonlinear ODEs, however the motion of aggregates is described by a linear PDE!!

Advection equation – transport of density of cases

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Notion of transport is applicable to a variety of topics in sociology such as residential mobility and health trajectories.

- Residential mobility – variables are actual geographical ones. Trajectories are in physical coordinate space.

- Health trajectories– Variables are biological, sociological markers – state space is more abstract


AP

Modelling cases at the individual level – the vector space approach

- Each case is a vector in a k dimensional vector space
- The motion of a case across time is equivalent to the motion of it equivalent vector across time
- The collection of cases (or k-dimensional vectors) is a vector space
- Hence, a case trajectory is represented as a k-dimensional vector trajectory in a vector space (also sometimes called the state space)
- The motion of individual cases is modeled by an ODE

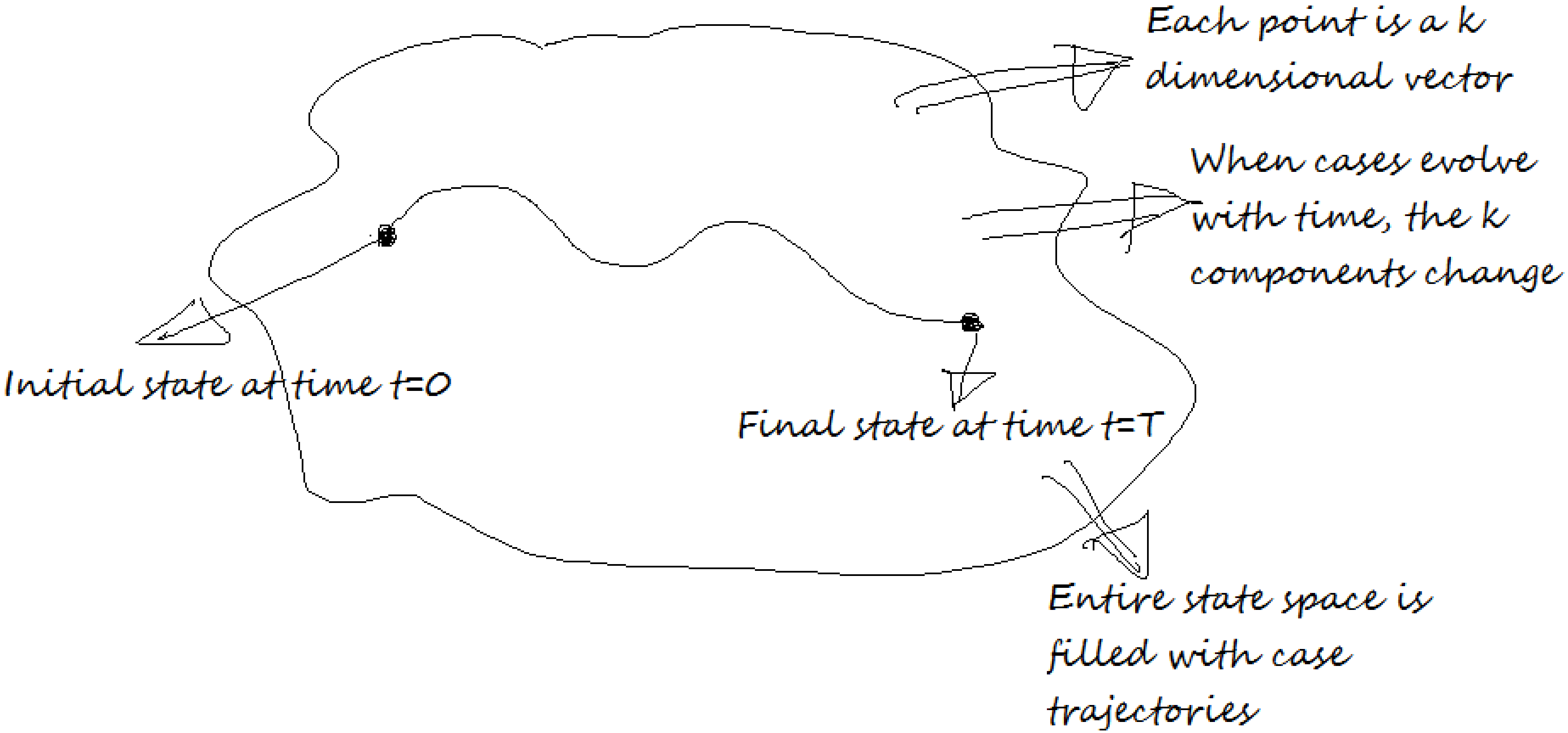
$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

State space - the case trajectory viewpoint



Modelling cases at the aggregate level- the density approach

- An aggregate of cases forms a distribution (or a density) in state space

- As the cases evolve, the density of cases also evolves

- The initial density can be chosen to have more cases with a certain profile and vice versa. Any initial distribution can be chosen.

- The motion of densities with mass conservation is modeled by the advection equation which is a PDE

$$\dot{x} = f(x); x(0) = x_0$$

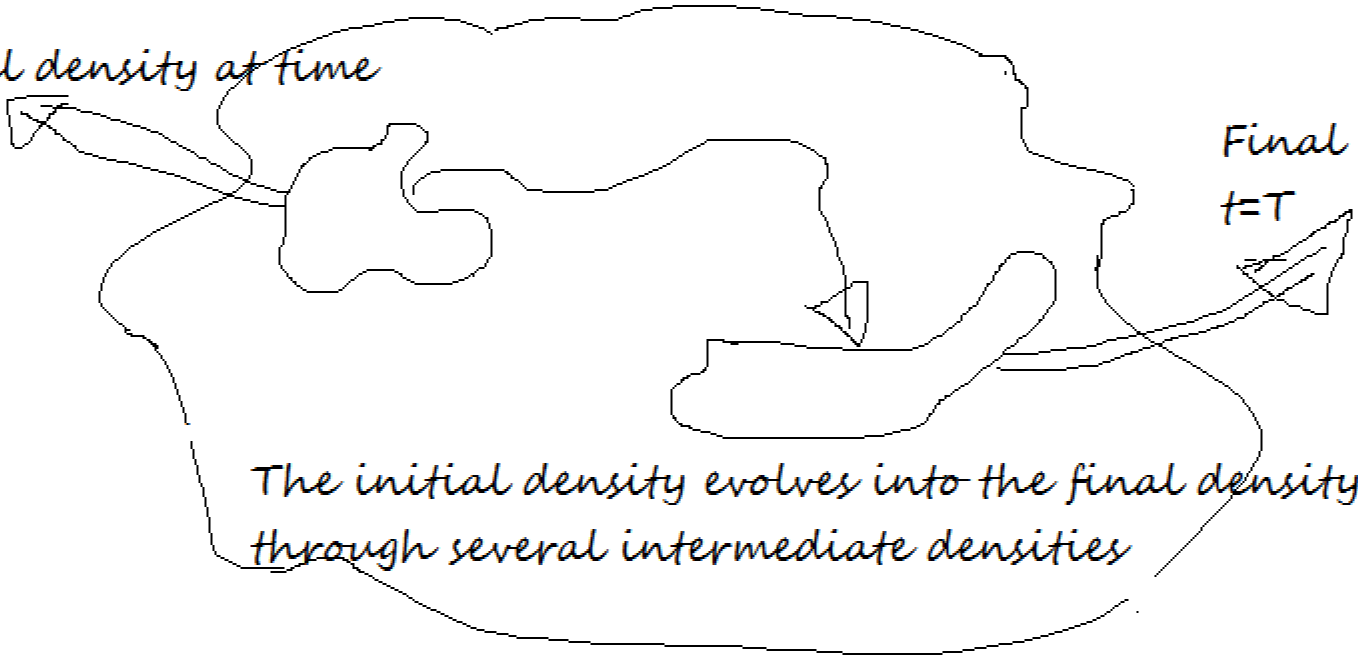
$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

State space – the density trajectory view point

Initial density at time t=0



Final density at time t=T

The initial density evolves into the final density after going through several intermediate densities

Modelling aggregate dynamics

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- At the microscopic level, we have nonlinear and complex dynamics but at the macroscopic level, we have lower order and slow dynamics
- The macroscopic aggregate dynamics are akin to Haken's order parameters.
- Our approach aims at extracting slower macroscopic dynamics and classifying them as major and minor trends.
- We expect the trends to be a fundamental defining characteristic of complexity in the system.

Uniqueness of our approach

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Continuous time modeling
- Deterministic modeling
- Differential equations (both ODE and PDE)
- Gradation of state space based on velocity of motion
- Non-equilibrium clustering using the Lyapunov density plot

Strengths

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Prediction of longitudinal evolution of cases with multiple variables across time
- Studying complexity in dynamical motion of cases in the form of saddles, sources, sinks, or periodic orbits
- Gradation of the state space into regions where cases move faster (or slower) from the velocity contour plot
- Non-equilibrium clustering of trajectories from the Lyapunov density plot (higher values mean more trajectories have squeezed through)

Strengths

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Prediction of majority trends in trajectories for novel choices of initial profiles or densities
- Multiple models to describe the same phenomena allowing for a choice of better ones
- Ease of incorporation of new data into the modeling process to fit the database as it grows

Simplifying Assumptions

Our approach (which combines what is known in physics and applied mathematics as the inverse and direct problem) is novel in four important ways: first, we take a unique, data-driven view of the cases in a cohort, which we define as K dimensional vectors, where the velocity vector for each case is computed according to its particular measurements on some set of empirically defined social, psychological, or biological variables.

Second, we translate the data-driven, nonlinear trajectories of these microscopic cohort constituents (cases) into the linear movement of macroscopic trajectories, which take the form of densities.

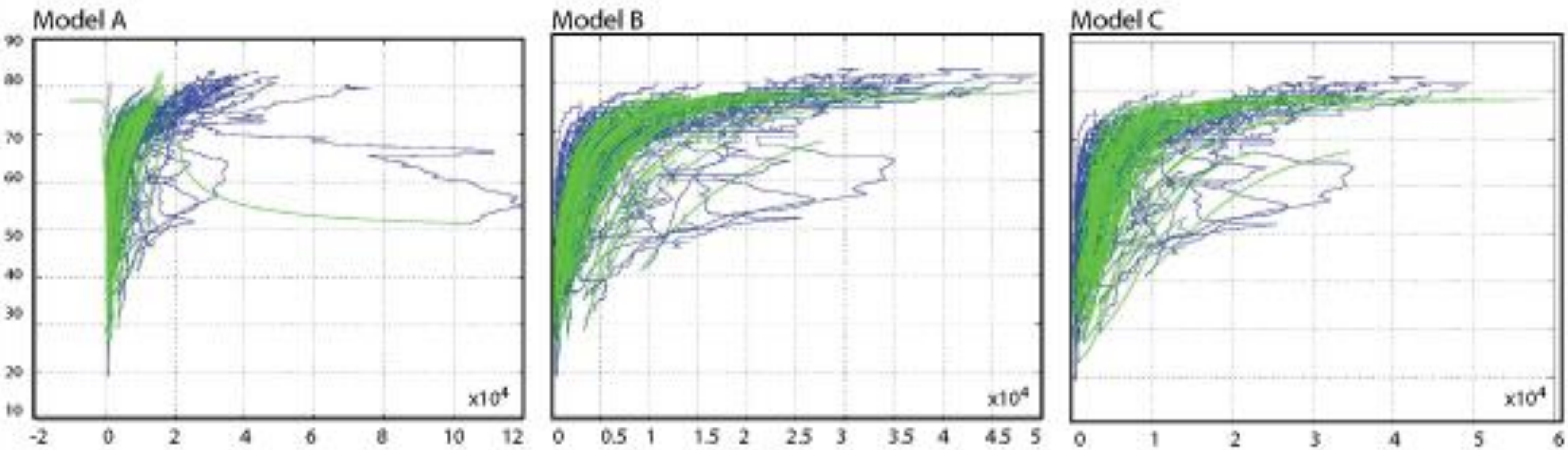
Here, we are drawing on Haken's synergetics and the idea that self-organizing macroscopic trajectories are less dynamic, generally speaking, than microscopic trajectories, which are high dynamic, out of which the former emerge.

Simplifying Assumptions

For our empirical case, we drew our data from the Gapminder website. The Gapminder website (created by Ola Rosling, Anna Rosling Rnnlund, and Hans Rosling) provides researchers, teachers, students, and the general public a wealth of time-series data (often starting back in the early 1900s) on the economic, political, cultural, social, biomedical, and health development of countries throughout the world, which it converts into a series of two-dimensional (2D) animations and interactive graphics (see http://www.gapminder.org/). For the sake of demonstration, therefore, we consider a database with two variables ($K = 2$) from Gapminder; namely, per capita GDP ($x_1(t)$) and life expectancy ($x_2(t)$) for 156 countries over 63 years (t).

FIGURE 4

State Space Fit for Best Model



State space fit for the best model.

State space fit for the best model without Kuwait or Luxembourg.

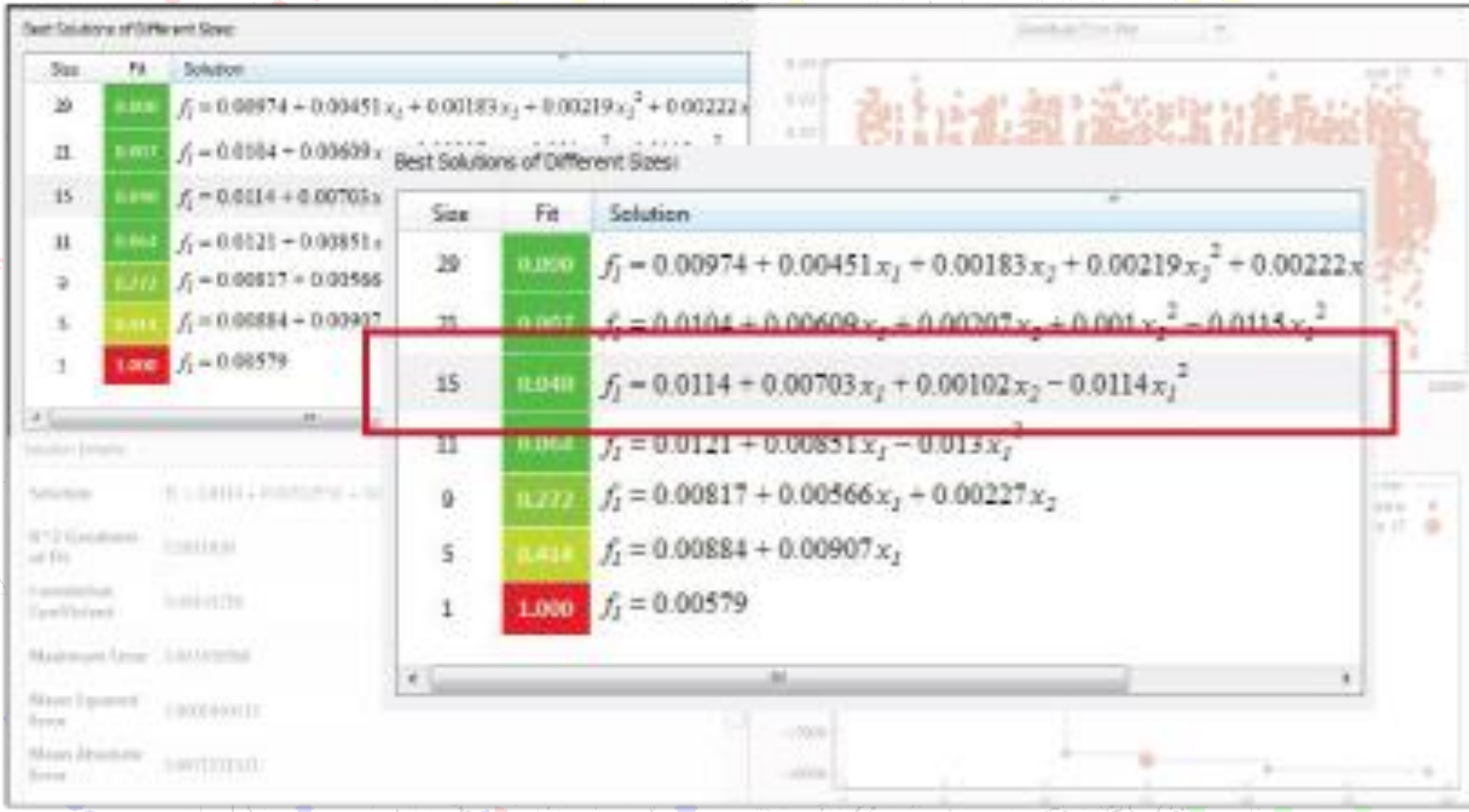
State space fit for the best model without Kuwait or Luxembourg, but with *time* as an independent variable.

Shown here are several computed Matlab models for the first component of velocity vector f1. Models were created using the ordinary differential equation solution from Eureqa. In all three models, the X-axis represents ***GDP***; and the Y-axis represents ***Life Expectancy***. In the models, the blue trajectories are from the data; green trajectories are the fitted model

Simplifying Assumptions

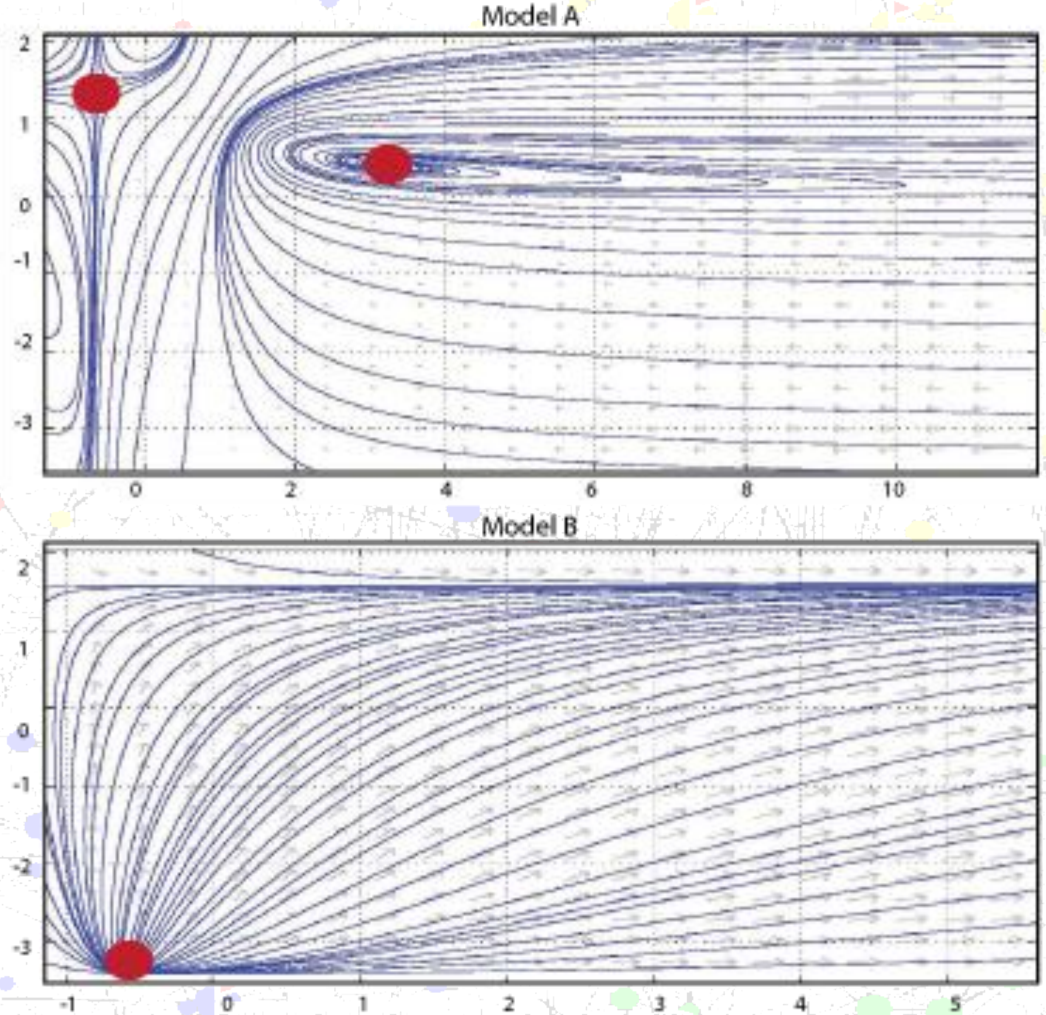
Third, we perform this translation by fitting the time trajectories of these cases using an autonomous (and, in some instances nonautonomous) ordinary differential equation (ODE) (1). In most cohort studies, be they network studies or otherwise, the laws governing their macroscopic dynamics are not known [17,18]. Fitting functions with an autonomous ODE must, therefore, be entirely data driven and based on a “goodness of fit” model. Our unique solution to this data-driven problem is to employ a genetic algorithm, as it does not require any a priori knowledge of the laws governing the data [19]. Instead, it uses the data to evolve to an optimal solution. It can do so because a genetic algorithm is a “brute force” search, but in an efficient way. Also, it finds global minima (as opposed to local minima), hence, it is a global optimization routine. As such, a genetic algorithm allows researchers to find the novel, mechanical laws of motion for social science and biomedical cohort data—with the knowledge that, often, each new dataset presents a new search for new laws, hence the study of complexity [20].

FIGURE 2



*Eureqa gives multiple models for the vector field of velocities. Figure 2 shows several computed models for the first component of velocity vector f1. The best fit model (#15 in our case, shown above) is usually the one that has a mid-level complexity in terms of number of polynomial symbols and the error values in the mid range amongst all models.

FIGURE 3



Shown here is are the state space trajectories for two of the models we settled on using Eureqa. In both models (A and B), the X-axis represents ***GDP*** (converted to z-scores); and the Y-axis represents ***Life Expectancy*** (converted to z-scores). In both models, the arrows show the direction of the trajectories; the larger the arrow the higher the vector's velocity. ***Model A:*** In this model, all countries are included; the red dot located in the top-left sectionof Figure 3 shows a saddle point; and the red dot located in the top-middle sectionof Figure 3 shows a spiraling source. ***Model B:*** In this model, the minority trajectories of Luxembourg and Kuwait were removed; the red dot here is a source.

Simplifying Assumptions

Fourth, using the vector field thus obtained, we use the advection PDE to simulate the evolution of a distribution of cases (as densities) across time (2). The advection PDE has been used extensively in fluid mechanics and electromagnetism to model the transport of physical quantities such as mass and charge, respectively [21,22].

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

Residential mobility

Global Health of Nations paper

Motion of density

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

Questions??