RT1: Modelling Dynamical Systems

Harry Newton

Overview

Animal herd movements, population dynamics, predator-prey interactions—these processes and animal behaviours are all of immense interest to the Ecologist; and are also of great importance in our efforts to understand the effects of climate change. They are also all examples of complex dynamical systems. Dynamic means a system's behaviour evolves or changes over time; the complexity arises from the large number of components (e.g. animals, species, etc.) involved, and the myriad interactions between these components.

Studying these systems is quite a challenge, not just because of their complexity; depending on the context, data collection is also difficult. Fieldwork allows for the most control over the data collection process, in which catch and release methods are often employed. However, these studies are costly and so usually only employ small teams of researchers for a short time, limiting the amount of data that can be collected. Old company records can also be a useful source on animal populations, although, of greater quantity, lack of, or partial knowledge of how the data was collected can lead to seemingly counterfactual phenomena. As is the case with (in)famous Hudson's Bay Company Hare and Lynx pelt data, which seemed to suggest that hares eat lynxes. Citizens science is another alternative, and it too has the potential for a greater quantity of data collection. Here, though, we encounter more room for uncertainty with observer difference, false-positives etc.

Models can be of great use to Ecologists as they provide convenient mathematical descriptions of the 'real world' behaviour of these complex dynamical systems. Due to the aforementioned challenges, models are often initially derived from partial information on the system, based upon theoretical knowledge or patterns identified from the empirical data. Therefore, to accurately describe the system, a model must be calibrated to fit the data.

Once fit, the model can give an Ecologist many useful insights on the system; particular laws and mechanisms (e.g. population growth rates, predator-prey interactions, infection rate) associated with a system are embodied within the model's parameters. Censored data, i.e. the data unable to be collected due to the reasons just mentioned, can be indirectly uncovered by extrapolating the model backwards or forwards in time. The latter is particularly useful for conservation decisions, predictions as to the likelihood of a species extinction.

Many considerations must be given to determine an appropriate choice of modelling method. One can choose generality, adaptability and robustness, but suffer from a greater computational burden and a reduced understanding of the system's underlying mechanisms (Stochastic Models). On the contrary, one may pursue simplicity and greater mechanistic detail, but at the cost of flexibility (Mean-Field Models). Such choices are influenced by a system's properties, such as its scale (can we exploit averaged behaviour) and degree of stochasticity. No approach is perfect; only excelling if we specify what we want from the model and correctly identify the conditions in which it is most useful to us.

1   Introduction

Many Ecological processes such as population dynamics, herd movements and so on fall into the category of dynamic systems, whose output evolves over time; with the potential to undergo radical, qualitative shifts in their overall behaviour. Within the Ecological sciences, it is desirable to develop accurate models of these Ecological processes so as to supplement their description, aid in hypothesis validation, provide refinements in experimental design (Optimal Design) (Overstall et al., 2020) and predict the evolving behaviour of a system to a reasonable degree of accuracy. Building models is difficult as there is very little knowledge on the exact laws and behaviours of these systems, due to their complexity, or the limited availability of empirical data. This gap between a system's data and its suitable mathematical description is called the **Inverse Problem**. It is usually approached with a combination of modelling and statistical inference, where we construct a model with unknown parameters and then fit this model to the data to obtain these unknown parameters.

In this report, we will cover two approaches towards model fitting for a dynamical system, one stochastic model; the Linear Gaussian State Space Model (LG-SSM) and one deterministic; Ordinary Differential Equations (ODEs). We will introduce them in Sections 2 and 3, and then in Section 4 we discuss parameter inference techniques which can be applied to fit each model. Finally, we will conclude in Section 5 by discussing the relative merits of each method, and the appropriate contexts in which they should be used.

1.1   Modeling Approaches

We can apply two classifications to the models we may use to solve an Inverse Problem- *stochastic* and *deterministic*. For deterministic mathematical models, uncertainty is not explicitly accounted for in the model formulation. This may seem counter-intuitive when dealing with the systems described in the introduction, which include many sources of uncertainty. But, when we are dealing with large-scale problems, in which populations are the focus rather than groups, then the interactions between certain components can be 'averaged out'. This stymies the influence of stochastic effects; allowing us to recognise and quantify underlying patterns within the system. Dynamical models of this category are usually composed of many differential equations or ODEs. When factoring in spatio-temporal behaviour PDEs may be used instead. Deterministic models, usually require a great deal more prior knowledge in their design compared to stochastic methods.

Stochastic approaches on the other hand directly incorporate our uncertainty, or limited prior knowledge of a system's behaviour into its structure. They assume that the underlying behaviour of a system is dictated by some kind of **Stochastic Process**. A stochastic model of a particular system can be obtained by conditioning a stochastic process on available empirical data, to find the states the system was while not observed. Or, if the model is detailed enough, the parameters which describe certain useful features of the system, e.g. population growth, infection rates etc. The current model *du jour* in Statistical Ecology is the State Space Model (SSM), which falls into this category, but Artificial Neural Networks and Stochastic Differential Equations (SDEs) have also been applied to the study of ecological time-series data.

2   State Space Models (SSM)

The **SSM** is actually a general description for a class of models with varying degrees of complexity. This is an extremely vast topic deserving of its own report, which can be found here (Newman et al., 2023)-so we will only cover a basic introduction to SSM.

An SSM is built on the assumption that an individual observation of a system's behaviour-taken at a time t, denoted y_t- reflects the outcome or mapping of a **hidden state** of that system α_t. This mapping can be described by the general function $h(\cdot)$. Of course, this observation will have some **uncertainty** attributed to the act of measurement; which can be described with an additive Gaussian noise term $\eta \sim N(0, \sigma_\eta^2)$

$$y_t = h(\alpha_t, \theta) + \eta \tag{1}$$

Extending this assumption in time, the overall temporal behaviour of that system-for its period of study $t = 1 : T$-is a consequence of a hidden chain of causality. Where movement between system states is dictated by some kind of Markovian process i.e. the current state of the system α_t, inferred from observation $\mathbf{y}_t$ is dictated only by the previous hidden state α_{t-1} [1]. A basic illustration of an SMM, as just described, can be found in Figure 1.

$$\alpha_{t+1} = f(\alpha_t, \theta) + \nu \tag{2}$$

Equation 2 describes the hidden process, Where $f(\cdot)$ is our state transition function and ν is our stochastic term, for the Markov chain. The nature of the state transition function is informed by our prior knowledge of the system's mechanics. It is context-dependent, it may be a standard function (linear, quadratic etc.) or as is the case with more complex models, some form of differential equation.

When we fit an SSM to some observed data $y_{1:t}$ our main objective is to do inference on the unknown parameters θ. However, it is highly likely that the hidden states $\alpha_{1:t}$ are also unknown, in this case, we can determine the nature of these through **Smoothing** or **Filtering**.

Filtering The distribution $p(\alpha_t | y_{1:t}, \theta)$ for α_t is conditioned on observations available up to time t.

Smoothing The distribution $p(\alpha_{0:T} | y_{1:T}, \theta)$ for α_t is based upon all available observations and states up to $T < t$.

These two concepts are quite similar and can easily be confused with one another. Both of them wish to isolate a given state from the system's stochastic noise, but Smoothing is ideal if one wishes to preserve the underlying trends in the hidden process. To help understand the system on a mechanistic level.

2.1 Linear Gaussian: Applied to Population Dynamics

To properly explore the concepts we have just described, let us now build an SSM to model a standard Ecological inverse problem; population dynamics. In this example, We will apply the Linear Gaussian SSM (LG-SSM), which is the most general and simplified method applied to statistical ecology problems.

Let us consider some data on the varying size of an animal population for a series of time intervals $\mathbf{t} = (1, \ldots, T)$. We assume these observations $\mathbf{y} = (y_1, y_2, \ldots, y_T)$ are unbiased, and represent a true measure of the population size at a given time t. We can define the measurement process (Observation Model) like so:

$$y_t = \alpha_t + \eta \quad t = 1, \ldots, T \tag{3}$$

---

[1]Strictly speaking this is a 1st order Markov chain. For nth ($n > 1$) order Markov chains, state transitions are not only dependent on the previous state-as is the case with 1st order Markov chains, but n other transitions prior to that
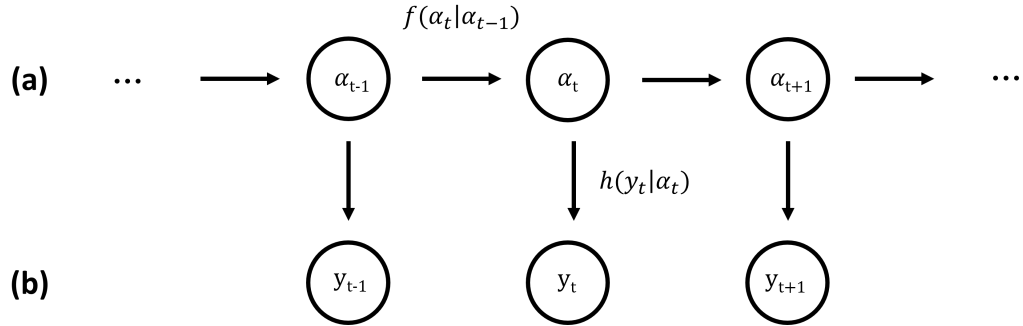
Figure 1: An illustration of a Hidden Markov Model, (a) is the **Hidden Process**, with a transition probability given by $f(\alpha_t|\alpha_{t-1})$ and (b) is the **Observed Process** with a mapping $h(y_t|\alpha_t)$

When very little information is known on the mechanisms underlying an Ecological system's dynamics, a general assumption to make on our model is that the hidden state transitions are dictated by a **Gaussian Process** (GP), so that $h(\cdot)$ is a linear function and $\nu \sim N(0, \sigma_\nu^2)$. Modifying Equation 2 with our assumptions provides us with an expression for the **Hidden Model**:

$$\alpha_{t+1} = \rho\alpha_t + \nu \quad t = 1, \ldots, T \tag{4}$$

To complete our model it must fit our population data, in doing so, the unknown parameter(s) $\theta = (\rho, \sigma_\nu, \sigma_\eta)$-which in this case is the population growth rate ρ, measurement uncertainty σ_η, and variance for the GP σ_ν-can then be determined.

The correct model fitting method to apply to this problem is dependent on the complexity of the model, the tractability of our likelihood function $L(\rho, \sigma_\nu, \sigma_\eta; y)$, the identifiability of the parameters, and so on. We will cover this topic with greater detail in Section 4.

3 Mean-Field Model

The SSM approaches just discussed are appropriate for Ecological systems at a smaller scale, such as the spatio-temporal study of herd movements or local population growth, i.e. when the study of individual animal behaviours is more desirable. If we attempt to apply SSM methods to inverse problems of a significantly larger scale we can quickly fall prey to issues of increasing model dimensionality; resulting in a too-complex model, with an abundance of parameters to estimate.

In certain cases mean-field models allow us to partially alight this complexity problem by selectively aggregating system components so that the overall behaviour of the model output is averaged out. In essence, instead of modelling individuals, we have increased our scope to population densities or averages. The upshot of this scale adjustment is that the system's dynamics may now be described with a *deterministic* mathematical model; composed of a set of parameterised Partial/Ordinary Differential Equations (PDE/ODEs). From this standpoint, the inverse problem is another model fitting exercise, to obtain values for the parameters θ which reduce the distance between the mathematical model's solution (prediction) $\mathcal{G}(\hat{x}(t), \theta)$ and the observed data $y(t)$

$$y(t) = \mathcal{G}(\hat{x}(t), \theta) + \eta \tag{5}$$

4

Again, we also include an additive Gaussian Noise term $\eta \sim N(0, \sigma_\eta^2)$ to account for uncertainty on measurements.

3.1 Lotka-Volterra Equations: Predator-Prey Model

The exact compositions of the mathematical model are context-dependent; influenced by existing theoretical work on Ecological dynamics, or borrowing suitable models from other areas of the Natural Sciences. To provide a more concrete and relevant example let us apply the mean-field approach to an example in Ecology; in this particular inverse problem, we shall consider the interacting dynamics between predator and prey populations. We define the average number of prey at a given time t with $x(t)$; the average predator population is defined as $y(t)$. The time-evolving behaviour of $x(t)$ and $y(t)$ can be described by a system of coupled ODEs known as the Lotka-Volterra equations, which can be found below.

$$\frac{dx(t)}{dt} = \alpha x(t) - \beta x(t) y(t) \tag{6}$$

$$\frac{dy(t)}{dt} = \gamma x(t) y(t) - \delta y(t) \tag{7}$$

Given time series data on average prey and predator numbers there are numerous methods which can be applied to fit our model and calculate the unknown parameters $\theta = (\alpha, \beta, \gamma, \delta)$, we will discuss these other methods in Section 4. Here we will demonstrate one method which begins with a re-parameterisation of the model; α and β are population growth rate multipliers-the same as ρ in Section 2-δ and β are coupling parameters which link the growth rate of the prey/predator to their opposite i.e. if the prey population increases, the predators have access to more food, they will therefore experience an increase in their numbers, which will have a knock-on effect on prey numbers. As α and δ describe uncoupled growth rates we can redefine our system of dynamical equations to a simpler model which is more tractable.

$$\frac{dx(t)}{dt} = \alpha x(t) - x(t) y(t) \tag{8}$$

$$\frac{dy(t)}{dt} = x(t) y(t) - \delta y(t) \tag{9}$$

We can obtain estimators $\hat{\beta}$ and $\hat{\gamma}$ by deriving Maximum likelihood Estimators (MLE). To begin, we assume our observed data of the predator/prey populations can be treated as Random Variables X_{t_i} and Y_{t_i}. We assume they are Poisson distributed with a mean defined on the integrated average number of predators/prey divided by the respective coupling parameter. And so we can derive an expression for the likelihood, conditioning our γ and β on the available predator-prey data X and Y:

$$L(\gamma, \beta; X, Y) = \prod_i f\left(\gamma^{-1} \int_{t_i/t_1}^{t_{i+1}/t_1} x\, dx;\right) f\left(\beta^{-1} \int_{t_i/t_1}^{t_{i+1}/t_1} y\, dy\right) \tag{10}$$

Our estimators for the coupled parameters dependent on α and δ and observed data for points t_i for $\forall i \in [1, T]$ can be calculated explicitly from Equation 10 by taking derivatives of w.r.t β and γ:

$$\hat{\beta} = \frac{\sum_i \int_{t_i/t_1}^{t_{i+1}/t_1} y\, dy}{\sum_i Y_i} \tag{11}$$

$$\hat{\gamma} = \frac{\sum_i \int_{t_i/t_1}^{t_{i+1}/t_1} x\, dx}{\sum_i X_i} \tag{12}$$

5

As we have an explicit formula for our estimators, we can apply an **Expectation Maximisation** algorithm to iteratively converge on the correct parameters. This works by (1) initialising our parameters $\theta^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \gamma^{(0)}, \delta^{(0)})$ with a guess, or appropriate assumption. (2) we obtain values for $\alpha^{(1)}$ and $\delta^{(1)}$ which maximise Equation 10 for $\beta^{(0)}$ and $\gamma^{(0)}$. Solve the dynamical model with some numerical solver, such as RK4, and use the solution in Equations 11 and 12 to estimate values for $\beta^{(1)}$ and $\gamma^{(1)}$. Rinse and repeat for n iterations until the distance between the model's output and the data converges to a reasonable tolerance ϵ.

4   Parameter Inference

The **exact** approach to **Parameter Inference** is based upon likelihood calculations which condition the unknown parameters θ on the observed data $y_{1:T}$. Estimates for the parameters based upon the likelihood can be obtained through either Optimisation or Bayesian Inference. However, when the likelihood is intractable or too computationally costly, parameter estimates can be obtained through likelihood Approximation or likelihood free methods. We will cover both approaches in this section

4.1   Exact Methods

The classic approach involves some optimisation of a parameter's MLE, the idea being that the correct parameters should minimise the distance between a model's output and the empirical data to a reasonable degree of tolerance ϵ. Due to uncertainty on the data and the hidden process, the estimators will only converge on the correct parameter values. For both deterministic and stochastic methods steps must be taken to validate parameter robustness these values are accurate through **Sensitivity Analysis** and other uncertainty quantification techniques. Optimising MLEs for deterministic systems requires solutions to the ODE/PDEs, in rare instances an analytical closed-form expression exists, but in most cases solving differential equations has to be done via an iterative numerical method such as Euler's or RK4. Numerical Solvers are often computationally costly, this becomes a greater problem if the model-fitting process involves other iterative algorithms like MCMC. For complex ODE-based models simplification of certain model elements, or the treatment of a larger model as independent components, as seen in (Campillo-Funollet et al., 2021) can also make parameter inference more tractable. Gradient-based methods can be used to find the maxima of the parameter space which minimises some error functions between the model and the data (Deng, 2018).

Bayesian Inference differs from MLE methods as we are instead trying to determine the parameter's distribution. In this case, we assume some prior $p(\theta)$ on the parameters. This requires partial knowledge on the nature of the parameters which can be sourced from the scientific context of the problem. If this knowledge is limited, a uniform distribution may be used instead (Roda, 2020). From the given prior, the posterior distribution of the parameter space $p(\theta|y)$ can be obtained using Bayes theorem:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{13}$$

Where $p(y|\theta)$ is the likelihood. Markov Chain Monte Carlo (MCMC) methods are used to sample from $p(\theta|y)$, given enough time an MCMC algorithm will converge on a sample path that has arrived at a stationary process; effectively covering the target's posterior (Roda, 2020).

4.2 In-Exact Methods

Many in-exact methods are Bayesian based but the **Laplace approximation** allows for MLE optimisation but assumes the likelihood can be approximated with a multivariate Gaussian distribution. This is quite a useful technique when we wish to determine the marginal distributions of individual parameters from a likelihood. **Approximate Bayesian Computation** (ABC) methods sample from an approximated posterior, by rejecting samples in which the distance between the observed data y and the model's output z is greater than an accepted tolerance ϵ (Alahmadi et al., 2020)

$$p_\epsilon(\theta, z|y) \propto \mathbb{1}(\rho(z, y) \leq \epsilon)p(\theta)f(z|\theta) \tag{14}$$

Finally **Sequential Monte Carlo**, also known as the particle filter, in which individual particles represent a set of potential parameters for a model (SSM or ODE). The process starts of with a number of particles whose parameter values are randomly generated or sampled from prior distributions. The next stage is the filter step, for SSM use **Filtering** to obtain a value for the next hidden state α_t (SSM). For the ODES we can obtain solutions for each particle numerically up to and including the next observation. We then obtain weights for each particle based on how their predictions compare to our empirical data. We then sample from these particles, their respective weights improve their acceptance chances. We then update our model parameters informed on the particles. If our model fits the data to reasonable tolerance we stop, otherwise we repeat the process again with particles we the previous particles. Conveniently, ABC can be used in SMC to update a model's parameters from the particles.

5 Strengths and Weaknesses

Deterministic models work best in problems in which a large quantity of data has already been collected and the scale of the system is sufficient for a mean-field approximation of the interactions being studied. Once a model of this type has been fit and has been shown to represent the system to a reasonable degree of accuracy, it can be used as a forecasting tool. This is because the parameters and model itself represent a mechanistic description of the system; we could call these characteristics laws. Providing the system still obeys these laws up into the immediate future no further inference is required to update this model. The weakness of Deterministic models is their rigidity; should the data we collect on the system begin to show patterns contrary to the model, or we become aware of some change within the system, for example, in a population dynamics model, where the habitat we are observing experiences a forest fire. Then our model must be fit once again, as such a violent event has severely disrupted the stable equilibrium of the system. On its face, such a scenario is undesirable. Because of the intrinsic context associated with the model's parameters, conclusions can be drawn and hypotheses validated by how animal behaviours may have been altered by such a violent catastrophe. In comparison, SSMs explicitly incorporate process and measurement uncertainty in their model, which makes them considerably more robust and flexible when handling noisy or limited data. Systems in which spatial or temporal heterogeneity is the focus, such as the study of individual animal/group foraging patterns, migration paths etc. These small details are impossible to model with the deterministic approach. SSM flexibility allows for constant updates on a model's parameters and hidden states in cases where data on a system is consistently updated with new measurements. A property with potential use in capture re-capture studies, as forecasts on the whereabouts of individuals can help researchers locate them again. For the reasons just mentioned, SSMs are extremely popular in Statistical Ecology, but

their usefulness diminishes if made too complex. In such cases, the quantity of parameters becomes too much to estimate without a great amount of computational effort, and questions should be asked about whether more parsimonious models could be used instead. Troublingly enough, one study (Auger-Méthé et al., 2016) has demonstrated issues with parameter and state estimation, where parameter estimates from an LG-SSM possessed significant bias. They demonstrated that these erroneous estimates occurred primarily when measurement error was greater than process error.

References

Alahmadi Amani A., Flegg Jennifer A., Cochrane Davis G., Drovandi Christopher C., Keith Jonathan M. A comparison of approximate versus exact techniques for Bayesian parameter inference in nonlinear ordinary differential equation models // Royal Society Open Science. III 2020. 7, 3. 191315.

Auger-Méthé Marie, Field Chris, Albertsen Christoffer M., Derocher Andrew E., Lewis Mark A., Jonsen Ian D., Mills Flemming Joanna. State-space models' dirty little secrets: even simple linear Gaussian models can have estimation problems // Scientific Reports. V 2016. 6, 1. 26677.

Campillo-Funollet Eduard, Van Yperen James, Allman Phil, Bell Michael, Beresford Warren, Clay Jacqueline, Dorey Matthew, Evans Graham, Gilchrist Kate, Memon Anjum, Pannu Gurprit, Walkley Ryan, Watson Mark, Madzvamuse Anotida. Predicting and forecasting the impact of local outbreaks of COVID-19: use of SEIR-D quantitative epidemiological modelling for healthcare demand and capacity // International Journal of Epidemiology. VIII 2021. 50, 4. 1103–1113.

Deng Bo. An Inverse Problem: Trappers Drove Hares to Eat Lynx // Acta Biotheoretica. IX 2018. 66, 3. 213–242.

Newman Ken, King Ruth, Elvira Víctor, De Valpine Perry, McCrea Rachel S., Morgan Byron J. T. State-space models for ecological time-series data: Practical model-fitting // Methods in Ecology and Evolution. I 2023. 14, 1. 26–42.

Overstall Antony M., Woods David C., Parker Ben M. Bayesian Optimal Design for Ordinary Differential Equation Models With Application in Biological Science // Journal of the American Statistical Association. IV 2020. 115, 530. 583–598.

Roda Weston C. Bayesian inference for dynamical systems // Infectious Disease Modelling. 2020. 5. 221–232.