Integral Projection Models for Species with Complex Demography

Paper by Stephen P. Ellner and Mark Rees

UNL Intro to Math Literature Course
Summer 2009
Eric Eager

Outline

- Introduction of matrix models/linear operators
- Explanation of the important aspects of matrix models.
- Introduction to the often more realistic world of IPMs
- Showing/deriving analogies between the two types of models.
- Motivating the tools with an example of a plant called Onopordum Illyricum.
- Analysis of this particular model
- Sensitivity and Elasticity
- Stability of equilibrium populations
- Results
- Conclusion

Initial Motivation

- Biologists, Ecologists, Mathematicians and scientists in general study populations of a various species in order to understand their impact on the environment we are interested in.
- Once one understands the behavior of a population they can use the results to predict outcomes/implement of policy change, for example.
- The various reasons for studying populations is seemingly endless.

Some Background

- We consider for now populations of biologically relevant species.
- Suppose one can split the population of interest into different “stages.” For example, a population of humans into children, pre-teen, teenager, adult and senior.
- One could then characterize the population by a vector *v*, where if the modeler considers the *i*th component of the population to be teenagers, for example, then the *i*th component of *v* is the number of people in the population that are teenagers.
- When one also considers populations that change over time, then the vector *v(t)* is the population vector at time *t.*
- One can also equip this population analysis by using the norm:

$$\|v\| = \sum_{i=1}^{n} |v_i|$$

How Does One Move From One Time step to Another?

- Suppose for example that one had a population with two stages, young and old and we start with $x(0) = \begin{bmatrix} 0.42 \\ 0.58 \end{bmatrix}$ as the initial population vector.
- Now suppose that young either die or graduate to old and the probability of graduation is 0.5.
- Also suppose that old survive and reproduce to the young stage, survival probability being 0.1 and fecundity being 2 youths per surviving adults.
- The modeler can set up a system of linear equations now, with these equations being:

$$x_1(t+1) = 2x_2(t)$$

$$x_2(t+1) = 0.5x_1(t) + 0.1x_2(t)$$

- The modeler can make this into a matrix!

Population Projection Matrix

- In this example, the population projection matrix, *A*, is given by:

$$A = \begin{bmatrix} 0 & 2 \\ 0.5 & 0.1 \end{bmatrix}$$

- To get to the *nth* timestep all one need to do is apply this linear operator *A* *n* times, i.e. $x(n) = A * A * A * A * \ldots A * x(0) = A^n x(0)$
- One could want to know if they can characterize this iteration with a single growth rate.
- Luckily, for a certain class of matrices, the modeler can!

What Kind of Matrices?

- For *A* a non-negative (like most biological models), primitive (there exists an integer *k* so that the *kth* iteration of *A* yields a completely positive matrix) then *A* has a simple, positive dominant eigenvalue with a positive eigenvector.
- This eigenvalue ends up being the asymptotic growth rate and the eigenvector, when normalized, is the asymptotic discrete probability distribution of the population.

Other Important Aspects of Matrix Models

- We can analyze the effect of changes in parameters/matrix entries on the leading eigenvalue λ with what is called "sensitivity analysis," which we will address later.
- We can also address the proportional changes in parameters/matrix entries on the leading eigenvalue with what is called "elasticity analysis," which we will also address later.
- These analyses are, in some contexts, more important than the asymptotic growth rate/population distributions.

What's Wrong With This Kind of Model?

- Sometimes nothing.
- But often, among other things:
  - Modelers are often discretizing a state variable when it is characterizing something continuous, like size, age, etc.
  - It is linear, where most biological phenomena deals with interactions between individuals and crowding out due to scarcity of environment and hence non-linearities (see the logistic equation, predator-prey models, SIR models, etc.)

This paper addresses the often arbitrary discretizing of the state variable

- The new type of model remains discrete in the time variable, since most biologist collect data in discrete times, i.e. we can still write the model as: $n(t+1) = An(t)$, where *A* is an integral operator and *n(t)* is a continuous population distribution (when normalized it is a continuous probability distribution).
- If Ms is the value of the largest stage and ms is the value of the smallest stage, then the model can we written as:

$$n(x,t+1) = \int_{ms}^{Ms} K(x,y)n(y,t)dy$$

Analysis of This Model

- Note that $n(x,t)ds$ is the number of individuals in the range $[x,x+dx]$.
- The kernel, K, is often written as

$$K(x,y)=P(x,y)+F(x,y),\text{ where}$$

P is the survival kernel and F is the fecundity kernel. For example $P(w,z)dwdz$ is the probability of survival from state w to state z in one time step.

Analogues to Matrix Models

- *K(x,y)* is analogous to an "*i,jth*" component of a matrix.
- Under similar assumptions to matrix models this IPM predicts a population growth rate λ with associated eigenvectors (which in this case end up being continuous functions) and state-dependent sensitivity and elasticity functions (which are also continuous functions)

What Makes Up the Kernel?

- The functions that make up the kernel are often (always?) found using statistical regression.
- The kernel components are often the product of independent probability distributions and fecundity functions from these statistical regressions.
- Example: $F(x,y)=PeJ(y)s(x)f(x)S(x)$, where
  - s(x) is the survival probability at stage x.
  - f(x) is the probability of reproduction at stage x.
  - S(x) is the number of newborns produced at size x.
  - J(y) is the stage distribution of newborns.
  - Pe is the probability of seedling establishment.

What Makes Up the Kernel?

- Thus, to reproduce one must survive, reproduce, determine how many offspring there are and their distribution in stage space.
- An example of a statistically derived function as a part of the kernel would be that $J(y)$ is often found to be a normal distribution, i.e. offspring obey a bell-curve with respect to a stage such as size.
- The survival, reproduction probability distributions usually obey a log-transformed regression with respect to the stage variable.

Complex Demography

- This paper is now considering individuals who are characterized by more than one quantity, i.e. age and size, or age and “quality”.
- Often times this complexity can lead to nonlinearity, for example some biological populations have survival being density dependent (think logistic equation).
- Suppressing these complexities/nonlinearities can lead to an overestimation of the growth rate.

Complex Demography

- But often times the extra complexity comes at a relatively cheap price. For example a size-age dependent IPM for Carlina vulgaris required only one extra parameter to describe the effect of age on flowering probability.

Onopordum Illyricum

Onopordum Illyricum

- This paper's focus species is the thistle O. Illyricum, which has size and age dependent demography as well as substantial latent between-individual variation in survival unrelated to size or age.
- O. Illyricum is a monocarpic perennial (reproduction is fatal) whose reproduction occurs only by seeds, which form a seed bank with typical half-life of 2-3 years.
- The paper's goal is to unify the development that makes IPMs a practical alternative to deterministic matrix models for structured populations with continuous trait variables (stages).

Some Data Analysis

- Plant sizes were in the natural log of rosette area.
- One of the statistically significant points of the data analysis is that the survival probabilities increases with size and decreased with age, as one would expect.
- Seedling establishment probability is a different story than the other statistically derived functions. Since there is a seed bank (i.e. seeds can lay dormant) one cannot just estimate this number (in the density independent case) by just using the ratio between recruitment and seed production. Thus one uses the observed rate of population increase ($\lambda = 1.026$).

Some More Data Analysis

- In the density dependent model the paper notes that intraspecific competition with neighbors have very little influence on growth and survival, thus D.D. influences only the seedling establishment.
- This means that the more that the area of interest is “crowded out” the less likely it is for a seed to become a plant (“establish itself”).

Putting the IPM Together

- The fate of O. Illyricum plants is influenced by their size, age and quality. Size is a continuous variable while age and quality are discrete.
- There are a set of functions, $n_{a,k}(x,t)$, that gives the distribution of size for individuals of age *a* (*a = 0,…,7*) and quality class *k* (*k=1-Q*)
- There is also a set of survival-growth and fecundity kernel components that specifies the fate and fecundity of individuals of each (*a,k*) combination.

Putting the IPM Together

- Thus survival takes individuals from (*a,k*) to (*a+1,k*), which means that the survival-growth kernel is derived from the probability of survival and the size distribution of survivors.
- Fecundity then project individuals from (*a,k*) to (*0,j*), or *0* age and *j* quality.

Again, Why Not Use a Matrix?

- Even in, for example, a *5x5* population projection matrix, there are ~10 parameters to be estimated, with the huge lack of being able to model the continuous nature of the stage variable.
- There are only seven parameters in this particular continuous IPM model, which is already less than a 5x5 matrix model
- Also, in words, there is “too much going on” between the arguably arbitrary stage cut-offs to do any justice in picking an average fecundity parameter, for instance, anyway.

A General Integral Model

- This is a general model to accommodate species with complex demography.
- The space of individual states *X* can be a set of discrete points and a set of continuous domains.
- Each continuous domain is either a closed interval or a closed rectangle in *d*-dimensional space.
- Each of these sets is called a component and is denoted as Ω_j.

A General Integral Model

- The state of the population is described by a function that gives $n(x,t) \geq 0$ the distribution of individual states *x* at time *t*.
- *n* consists of discrete values $n_j = n(x_j)$ and continuous functions $n_j(x)$.
- To describe transitions within and among components there is a set of kernel components $K_{i,j}(x,y)$, $1 \leq i, j \leq N$, where $K_{i,j} \neq 0$, whenever individuals in Ω_j contribute to next year's Ω_i.
- In general $n_i(x, t+1) = \sum_{j=1}^{N} \int_{\Omega_j} K_{i,j}(x, y) n_j(y, t) dy$.

Examples Kernel Components

- Number of two-leaf seedlings next year per one-leaf seedlings this year
- Size distribution of recruits produced by seeds that happen to germinate
- In words, $K_{i,j}(x,y)$ is a function (often made up of many individual functions) that gives the contribution of state *y* individuals in component *j* to state *x* individuals in component *i*.

Density Dependent

- The models previously described are all density independent.
- Density dependent model are composed by merely adding an extra dependency on the kernel, i.e. the kernel turns into $K = K(x, y, t)$, or for some measure *N* of total population density, $K = K(x, y, N)$.

How to Actually Implement a General Integral Model.

- We usually need to evaluate these integrals' characteristics numerically
- This paper recommends using the midpoint rule.
- The midpoint rule changes the analysis back to matrix analysis without having the parameter-estimation and accuracy-of-stages issues that a matrix model has.

- The numerical analysis follows the same as a matrix model as:
  - If the model is density independent, we have a leading eigenvalue which is the asymptotic growth rate and a leading “eigenvector”, which, when normalized, ends up being a continuous probability distribution of the population.
  - If the model is density dependent, the population settles down to an equilibrium population distribution, independent of (non-zero) initial population distribution (WOW).
  - Both of these can we approximated (very-well) by discretizing the model using matrices/vectors via the mid-point rule.

Some Results

- The analytic, long-term behavior of the integral model is identical to that of a power-positive matrix model (i.e. the same theorems apply).
- One can calculate the net reproductive rate R_0 which is the long-term generation-to-generation population growth rate. So that if g_k is the total number of offspring of generation *k* individuals, then $g_k / R_0^k \longrightarrow G$.

Analogue to Sensitivity Analysis

- Sensitivity analysis is also nearly identical to the matrix case, under appropriate definition of a perturbation of the kernel, i.e. how much does the leading eigenvalue change when we perturb the kernel.
- One need to do a slightly different computation in the continuous case.

Stability Analysis For Density Dependent Models

- If one assumes the kernel is of the form *K(x,y,N)*, where *N* is a weighted population size, if one perturbs an assumed equilibrium population distribution one can derive stability criterion for this equilibrium population that is analogous to a matrix model.
- In this paper, they actually proved that for the class of kernels than Onopordum Illyricum is in that, if only one positive equilibrium exists it is always locally stable.

Model Results for Onopordum Illyricum

- Because the model has finite maximum age and the distribution of offspring states is independent of parent state the density independent model satisfies the mixing at birth assumption and therefore has a unique dominant eigenvalue and associated eigenvector.
- This result verifies the field study preceded the mathematical analysis.
- Age 3 plants make the greatest contribution to the long-term growth rates (by the left e-vector).
- Elasticity analysis shows that survival-growth has a larger influence on the growth-rate than reproduction by 3-1.

More Model Results

- In the density-dependent case, if one assumes that in each year the probability of establishment was equal to the observed average number of recruits divided by the total seed production, there was a smooth convergence to an equilibrium density of 4.8 plants per square-meter.
- These numerical results confirmed the analytical results of the unique stable equilibrium in the Appendix.

Conclusion

- Ellner and Rees provide a solid introduction to the wonderful analytical tool of IPMs as both a contrast and a close-cousin of PPMs.
- This paper uses the plant Onopordum Illyricum to show how both density independent and density dependent IPMs can be implemented to analyze the long-term asymptotic behavior more efficiently than a matrix model with a relatively large number of stages.
- This overview effectively teaches these tools in the context of matrices and thus allows for a student/reseacher to effectively grasp the material for implementation to their own models.

Thank You!