Mathematical Biosciences 233 (2011) 77–89

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Numerical algorithms for estimation and calculation of parameters in modeling pest population dynamics and evolution of resistance

Mingren Shi [a,b,*], Michael Renton [a,b,c]

[a] School of Plant Biology, University of Western Australia, 35, Stirling Highway, Crawley, WA 6009, Australia
[b] Cooperative Research Centre for National Plant Biosecurity, Australia
[c] CSIRO Ecosystem Sciences, Underwood Avenue, Floreat, WA 6014, Australia

ARTICLE INFO

Article history:
Received 8 March 2011
Received in revised form 19 May 2011
Accepted 20 June 2011
Available online 13 July 2011

Keywords:
Parameter estimation
Offspring genotype table
Probit models
Mortality estimation
Population dynamics
Resistance evolution

ABSTRACT

Computational simulation models can provide a way of understanding and predicting insect population dynamics and evolution of resistance, but the usefulness of such models depends on generating or estimating the values of key parameters. In this paper, we describe four numerical algorithms generating or estimating key parameters for simulating four different processes within such models. First, we describe a novel method to generate an offspring genotype table for one- or two-locus genetic models for simulating evolution of resistance, and how this method can be extended to create offspring genotype tables for models with more than two loci. Second, we describe how we use a generalized inverse matrix to find a least-squares solution to an over-determined linear system for estimation of parameters in probit models of kill rates. This algorithm can also be used for the estimation of parameters of Freundlich adsorption isotherms. Third, we describe a simple algorithm to randomly select initial frequencies of genotypes either without any special constraints or with some pre-selected frequencies. Also we give a simple method to calculate the "stable" Hardy–Weinberg equilibrium proportions that would result from these initial frequencies. Fourth we describe how the problem of estimating the intrinsic rate of natural increase of a population can be converted to a root-finding problem and how the bisection algorithm can then be used to find the rate. We implemented all these algorithms using MATLAB and *Python* code; the key statements in both codes consist of only a few commands and are given in the appendices. The results of numerical experiments are also provided to demonstrate that our algorithms are valid and efficient.

1. Introduction

Estimating parameters based on measured empirical data is a critical issue in biosecurity models, such as simulation models of population dynamics and evolution of resistance in stored-grain insect pests [12]. These simulation models are based on integrating sub-models representing different key biological processes, such as genetic recombination and mortality due to pesticides. Various parameters for different sub-models must be calculated or estimated before these models are used to predict the effects of different possible management strategies. These parameters include: the chance of certain genotypes being produced as the result of the mating of certain parent genotypes (which we call *offspring genotype tables*), initial frequencies of genotypes, mortalities of insect pests under various pesticide doses, and the intrinsic rate of natural increase of an insect population. These are important parameters within the sub-models for simulating genetic recombination and thus determining the genotype of offspring, initialisation of the population, simulating the effects of pesticide applications and calculating the number of eggs produced by each insect, respectively.

By an *offspring genotype table* we mean a table that lists all possible combinations of parental genotypes, and, for each possible parental combination, gives the expected proportions of offspring genotypes (see Hedrick' book [19, p. 76] for an example of this kind of table, although no formal name is provided in this or other literature). Such a table is indispensable for a genetic model simulating evolution of resistance, or other traits. We develop a novel method to generate the offspring genotype table for a one-locus genetic model: quantifying all possible genotypes of parents and offspring and then using a block-matrix multiplication approach to generate the full table describing the chance of certain genotypes being produced as the result of the mating of each and every possible combination of parent genotypes. The offspring genotype tables for more than one locus are then produced recursively, with the table for a model with a higher number of loci produced from the tables for lower numbers of loci. This algorithm for the one- and

---

* Corresponding author at: School of Plant Biology, University of Western Australia, 35, Stirling Highway, Crawley, WA 6009, Australia. Tel.: +61 8 6488 1992; fax: +61 8 6488 1108.

E-mail addresses: Shi.mingren@gmail.com (M. Shi), mrenton@cyllene.uwa.edu.au (M. Renton).

0025-5564/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2011.06.005

two-locus cases is given in Section 2.1. We also explain how this algorithm can be extended for models with more than two loci.

Many problems of quantitative inference in biological and technological research concern the relation between a stimulus (e.g. phosphine fumigation dose) and a binomial response (e.g. mortality of insect pests). A binomial generalized linear model, with a link function such as the probit function (the inverse of cumulative distribution function), is usually used to analyse the empirical data. Normally, maximum likelihood estimation or chi-square approximation is applied to fitting the parameters of such probit models. In fact, however, in such probit models the probit is a linear function of parameters or metameter (e.g. log) of parameters and the corresponding equations with respect to the parameters form an over-determined linear system. We used a generalized inverse matrix method to find the least-squares solution of the regularization equations. We describe the method in Section 2.2. This method has advantages over other methods [4] if we only need to estimate parameters without other statistical information such as significance or confidence intervals for the estimates: it is simple with only one key command, provides a more accurate estimate of parameters, and even if the coefficient matrix of the over-determined linear system is not numerically (column) full ranked it will still work and yield a solution with minimum error in the L_2 norm sense [4].

In some situations, we may wish to randomly select some or all of the initial frequencies of genotypes for a biological or genetic model. These frequencies must satisfy two simple constraints: each frequency is in the range $[0,1]$ and the sum is equal to 1. In Section 2.3, we describe how we select the initial frequencies either without any extra conditions, or with some pre-selected frequencies, or with linear equality and inequality constraints. Also we give a simple block-matrix multiplication method to calculate the equilibrium proportions that should result from these initial frequencies according to the Hardy–Weinberg Principle [29].

The intrinsic rate of natural increase (or development rate) is an important parameter in modeling the dynamics of an insect population. In Section 2.4, we describe how we converted the problem of estimating this parameter into a root-finding problem and used a *bisection method* to find the rate to any desired accuracy.

All the above algorithms are implemented using MATLAB (www.mathworks.com) and *Python* (www.python.org) code, using the Scientific Python library (www.scipy.org), and the key statements and results of numerical experiments are given in Section 3 demonstrating that our algorithms are valid and efficient.

2. Methods

2.1. Quantification of genotypes through block-matrix multiplication algorithm for creation of offspring genotype table

We developed a novel *quantification of genotypes through block-matrix multiplication* algorithm to generate the offspring genotype tables for a one-locus genetic model. In this section we describe how this algorithm can be used to induce the two-locus table from the one-locus table by block-matrix multiplication, and then how this algorithm can recursively be extended to generate the offspring genotype tables for models with more than two loci. Based on assumptions of random mating and no dependence of inheritance on gender, this algorithm now makes it relatively straightforward to express genotype frequencies of an insect population as the proportion of offspring from all possible parental unions that belong to each genotype. Note that we developed the method in this paper only for diploid species, i.e. where each locus has two alleles, but the idea for developing this algorithm is also suitable for constructing algorithms for species where each locus has more than two alleles.

2.1.1. One-locus case

To use computational methods for generating the one-locus offspring genotype table, we need to quantify the parental and offspring's genotypes first. In the one-locus case, the two alleles, dominant "*A*" and recessive "*a*", are distributed among offspring in the usual, binomial ratios. Each mating of "female parent × male parent" will produce four possible combinations: [each of 2 alleles of female parent (F_1,F_2)] × [each of 2 alleles of male parent (M_1,M_2)]. For example, the mating $Aa \times Aa$, will produce $F_1 \times M_1 : AA$, $F_2 \times M_1 : aA\ (=Aa)$, $F_1 \times M_2 : Aa$ and $F_2 \times M_2 : aa$. This process can be obtained by a schematic or a diagrammatic method, known as the *Punnett square*, or by constructing a tree diagram [31]. The Punnett square, named after the geneticist Reginald C. Punnett, for the above case is shown in Table 1.

Hence the proportions of offspring are equal to $2/4 = 0.5$ for genotype *Aa*, $1/4 = 0.25$ for *aa* and also $1/4 = 0.25$ for *AA*. It is important to note that Punnett squares give probabilities only for *genotypes*, not *phenotypes*. The way in which the *A* and *a* alleles interact with each other to affect the phenotype of the offspring depends on how the gene products (proteins) interact. For classical dominant/recessive genes, like that which determines whether a rat has black hair (*A*) or white hair (*a*), the dominant allele will mask the recessive one. Thus in the example above 75% of the offspring will be black (*AA* or *Aa*) while only 25% will be white (*aa*). The ratio of the phenotypes is 3:1.

The proportion of each genotype in the offspring can be calculated by hand by counting the number of this genotype in the Punnett square or by calculating the probability using a multiplication rule in the tree diagram [31]. Our more efficient computer-based method to do this works as follows. First we use numbers to denote the genotypes of parents: "1" for the allele *A* and "2" for the allele *a*. Then the *Aa* genotype of female and male parents can be expressed by the following matrices respectively:

$$F_{Aa} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad M_{Aa} = [1, 2]. \tag{2.1}$$

The genotypes and numbers of four possible combinations of their offspring can be generated by matrix multiplication:

$$F_{Aa} M_{Aa} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [1, 2] = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}. \tag{2.2}$$

In the product, which can be regarded as a *digitized* or *quantified* Punnett square, "1" stands for the genotype *AA* (as $1 \times 1 = 1$), "2" for *Aa* ($1 \times 2 = 2 \times 1 = 2$) and "4" for *aa* ($2 \times 2 = 4$). We do not need to produce all of the "products" of the different genotypes one by one, instead, the whole offspring genotype table can be obtained at once by the following process:

(i) Let *M* be a 1×6 matrix (or a 1×3 block-matrix) representing the three possible genotypes of the male parent:

$$M = \begin{array}{cccccccc} A & A & & A & a & & a & a \\ {[1} & 1 & \vdots & 1 & 1 & \vdots & 2 & 2] \end{array} \tag{2.3}$$

and $F = M^T$ (transpose of *M*) be a 3×1 block-matrix representing the three genotypes of the female parent.

(ii) Then the block-matrix product *FM* is a 3×3 block-matrix with each block being a 2×2 sub-matrix where

$$FM = \left[\begin{array}{cc:cc:cc} 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 & 2 \\ \hdashline 1 & 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 2 & 4 & 4 & 4 \\ \hdashline 2 & 2 & 2 & 4 & 4 & 4 \\ 2 & 2 & 2 & 4 & 4 & 4 \end{array}\right]. \tag{2.4}$$