A note on determining parameter redundancy in age-dependent tag return models for estimating fishing mortality, natural mortality and selectivity

Diana J. Cole and Byron T. J. Morgan
University of Kent, Canterbury, UK.

January 21, 2009

Abstract

Jiang *et al* (2007, JABES p177-194) present models for tag return data on fish. They examine whether the models are parameter redundant, but need to resort to numerical methods as symbolic methods were sometimes found to be intractable. Also, their results are only applicable for a specified number of years of tagging data. Here we show how symbolic methods can in fact be used and also how conclusions apply to any number of years of tagging data.

Keywords: derivative matrix; ecology; exhaustive summaries; identifiability; Maple

1 Introduction

Jiang et al (2007) present new models for tag return data on fish. The important advance in the paper is the incorporation of age-dependence in the models. They are based on the probability that a fish tagged at age k, released in year i, are harvested and returned in year j has the form,

$$P_{ijk} = \begin{cases} (1-\Phi_{ijk})\dfrac{F_j \mathrm{Sel}_k \lambda}{F_j \mathrm{Sel}_k + M} & \text{when } j = i \\ \left(\displaystyle\prod_{v=i}^{j-1} \Phi_{ivk}\right)(1-\Phi_{ijk})\dfrac{F_j \mathrm{Sel}_{k+j-i} \lambda}{F_j \mathrm{Sel}_{k+j-i} + M} & \text{when } j > i, \end{cases}$$

where Φ_{ijk} is the conditional probability of surviving year j, given that it is alive at the start of the year, for a fish tagged at age k in year i, given by $\Phi_{ijk} = \exp(-F_j \mathrm{Sel}_{k+j-i} - M)$. The parameters are: F_j the instantaneous fishing mortality rate for fully recruited fish, Sel_a the selectivity coefficient for fish aged a (with fish being fully recruited at age a_c, so that $\mathrm{Sel}_a = 1$ for $a > a_c$), M the instantaneous natural mortality rate and λ the reporting probability for dead fish. The parameters M and λ can depend on year and age, in which case $M_{y,a} = M_y^Y M_a^A$ with $M_1^Y = 1$, with a similar parameterisation for λ. Henceforth we shall refer to their paper by"$JPBH^3$".

The motivating data for the work of the paper arise from a study of Chesapeake Bay striped bass, *Morone saxatilis*, tagged between 1991 and 2002. The release and tag-return data, stratified by age, are presented in $JPSH^3$, and are available from `http://www.amstat.org/publications/jabes/data.shtml/`.

In their paper $JPBH^3$ examine whether their models are parameter redundant; it is not possible to estimate all of the parameters by classical inference in parameter redundant models. In order to examine the parameter redundancy status of the model they use the symbolic algebra method of Catchpole and Morgan (1997). This involves calculating the derivative matrix $\mathbf{D} = \left[\frac{\partial P_\ell}{\partial \theta_i}\right]$, where P_ℓ refers to each of the non-zero P_{ijk} taken in turn. The symbolic rank of $\mathbf{D}$ is equal to the number of estimable parameters in the model (Catchpole and Morgan, 1997). This symbolic rank may in principle be calculated using a symbolic algebra package such as Maple. However it is not possible to calculate the rank of the derivative matrix for several of their models; they are structurally too complex and Maple runs out of memory trying to calculate the symbolic rank. In these cases, numerical methods are used, for particular values of parameters.

Recently, Cole and Morgan (2009) present a more general approach to determining parameter redundancy, in which the derivative matrix arises from differentiating what is called an *exhaustive summary*. An exhaustive summary is simply a parameter vector that uniquely defines the model, so that the vector $\mathbf{P}$ consisting of the non-zero P_{ijk} taken in turn is an example of an exhaustive summary for $JPBH^3$'s model. However we know that this exhaustive summary results in a derivative matrix that is too structurally complex for Maple to be able to calculate the rank. A simpler exhaustive summary can be found by first reparameterising the model and then finding a new derivative matrix with respect to this new reparameterisation rather than the original parameters. If the rank of the new derivative matrix is the same as the number of terms in the reparameterisation, then the reparameterisation forms what is called a *reduced-form* exhaustive summary. Otherwise a new reduced-form exhaustive summary can be found by solving an appropriate set of partial differential equations.

In this paper we show how symbolic algebra can be used to determine whether all of the models of $JPBH^3$ are parameter redundant, by using this reparameterisation method.

2 Determining the parameter redundancy status of the model

We start by considering the case when there are 2 age classes with $a_c = 2$, 4 years of tagging and 4 years of recovery. The full model in this instance has 16 parameters:

$$\boldsymbol{\theta} = [F_1, F_2, F_3, F_4, \mathrm{Sel}_1, \mathrm{Sel}_2, M_2^Y, M_3^Y, M_4^Y, M_1^A, M_2^A, \lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_1^A, \lambda_2^A].$$

An exhaustive summary (assuming no missing data) consists of the non-zero entries of $\mathbf{P}$. However Maple cannot calculate the rank of the derivative matrix

$\mathbf{D} = \left[\frac{\partial P_\ell}{\partial \theta_i}\right]$. Instead we find a reduced-form exhaustive summary using the reparameterisation method. The new exhaustive summary is given by

$$\mathbf{r} = \begin{bmatrix} F_1\mathrm{Sel}_1 + M_1^A \\ F_1\mathrm{Sel}_2 + M_2^A \\ F_2\mathrm{Sel}_1 + M_2^Y M_1^A \\ F_2\mathrm{Sel}_2 + M_2^Y M_2^A \\ F_2 + M_2^Y M_2^A \\ F_3\mathrm{Sel}_1 + M_3^Y M_1^A \\ F_3\mathrm{Sel}_2 + M_3^Y M_2^A \\ F_3 + M_3^Y M_2^A \\ F_1\mathrm{Sel}_1\lambda_1^A \\ F_1\mathrm{Sel}_2\lambda_2^A \\ F_1\lambda_2^A \\ F_2\mathrm{Sel}_1\lambda_2^Y\lambda_1^A \\ F_3\mathrm{Sel}_1\lambda_3^Y\lambda_1^A \\ F_4\mathrm{Sel}_1\lambda_4^Y\lambda_1^A\{1-\exp(-F_4\mathrm{Sel}_1 - M_4^Y M_1^A)\}/(F_4\mathrm{Sel}_1 + M_4^Y M_1^A) \\ F_4\mathrm{Sel}_2\lambda_4^Y\lambda_2^A\{1-\exp(-F_4\mathrm{Sel}_2 - M_4^Y M_2^A)\}/(F_4\mathrm{Sel}_2 + M_4^Y M_2^A) \\ F_4\lambda_4^Y\lambda_2^A\{1-\exp(-F_4 - M_4^Y M_2^A)\}/(F_4 + M_4^Y M_2^A) \end{bmatrix} \tag{1}$$

which can then be used in place of the original exhaustive summary. Taking parameters in the order presented in $\boldsymbol{\theta}$ above, we obtain the derivative matrix

$$\mathbf{D} = \left[\frac{\partial r_\ell}{\partial \theta_i}\right] = \begin{bmatrix} \mathrm{Sel}_1 & \mathrm{Sel}_2 & 0 & 0 & \cdots \\ 0 & 0 & \mathrm{Sel}_1 & \mathrm{Sel}_2 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ F_1 & 0 & F_2 & 0 & \\ \vdots & & & & \end{bmatrix}.$$

As the derivative matrix is structurally much simpler than previously, using the original parameter set, we can now calculate the symbolic rank using Maple. The value obtained is 16, and therefore this model is full rank and not parameter redundant.

This result can be extended using the extension theorem of Cole and Morgan (2009) to show that for 4 or more years of tagging and recovery and for 2 or more age classes, all of the parameters in this model are estimable (details are given in the Appendix).

To consider whether any submodels of the full model are parameter redundant, we can use Theorems 4 and 5 of Cole and Morgan (2009). This involves a modified PLUR decomposition of $\mathbf{D}$, writing $\mathbf{D} = \mathbf{PLUR}$, where $\mathbf{P}$ is a permutation matrix, $\mathbf{L}$ is a lower diagonal matrix with 1s on the diagonal, $\mathbf{U}$ is an upper triangular matrix and $\mathbf{R}$ is a matrix in reduced echelon form. Any parameter redundant submodels appear as solutions to $\mathrm{Det}(\mathbf{U}) = 0$. In this instance

$$\mathrm{Det}(\mathbf{U}) = \mathrm{Sel}_1^4 F_1^2 F_2 M_2 \mathrm{Sel}_2 (F_2\mathrm{Sel}_2 M_3^Y - F_2 M_3^Y + F_3 M_2^Y - F_3 M_2^Y \mathrm{Sel}_2)\ldots,$$

where the omitted terms do not naturally factorise, as shown in the Maple code. Therefore the model is parameter redundant if $F_i = F$ and $M_i^Y = 1$, except

Model	Number of parameters	Deficiency of Model
$F_y, M_{y\times a}, \mathrm{Sel}_a, \lambda_{y\times a}$	$3N+3K-2$	0
$F_y, M_y, \mathrm{Sel}_a, \lambda_y$	$3N+K$	0
$F_y, M, \mathrm{Sel}_a, \lambda_y$	$2N+K+1$	0
$F, M_y, \mathrm{Sel}_a, \lambda_y$	$N+K+2$	0
$F, M, \mathrm{Sel}_a, \lambda_y$	$N+K+2$	0
$F_y, M_a, \mathrm{Sel}_a, \lambda_a$	$N+3K$	0
$F_y, M, \mathrm{Sel}_a, \lambda_a$	$N+2K+1$	0
$F_y, M, \mathrm{Sel}_a, \lambda$	$N+K+2$	0
$F, M_a, \mathrm{Sel}_a, \lambda_a$	$2K+2$	1
$F, M, \mathrm{Sel}_a, \lambda_a$	$2K+2$	1
$F, M, \mathrm{Sel}_a, \lambda$	$K+3$	0
F, M, λ	3	1

Table 1: The deficiency of a number of submodels. In each of the models the subscript y refers to year dependence of a parameter, the subscript a refers to age dependence of a parameter and the subscript $y \times a$ refers to age and year dependence of a parameter. These results apply for $N \geq 4$ years of tagging and return and $K \geq 2$ age classes.

in the case when $M_i^A = M$, $\lambda_i^Y = 1$ and $\lambda_i^A = \lambda$ which is again full rank, and excluding the model with just single values for F, M and λ, which has deficiency unity. When $M_i^A = M$, $\lambda_i^Y = 1$ and $\lambda_i^A = \lambda$ equation (1) is no longer an exhaustive summary and this case needs to be considered separatley, as shown in the Maple code.

3 Discussion

$JPBH^3$ only considered 3 years of tagging and recovery and 3 age classes. They found their most general model to be parameter redundant with deficiency 1. However this was evidently not the complete result for this model. In fact, the general model is not parameter redundant for 4 or more years of tagging and recovery. For the submodels that they consider, we agree with the results of $JPBH^3$; all the parameter redundancy results for the submodels considered by them are given in general in Table 1.

Numerical methods are good at providing an initial idea of whether a model is parameter redundant or not. However numerical methods will only apply to the specific set of parameter values and for the specified number of years of tagging and recovery and age classes adopted. However, for a more detailed understanding of the parameter redundancy status of a model in general, symbolic methods are necessary. We have shown in this paper how it is now possible to use symbolic methods rather than having to use numerical methods to detect parameter redundancy for members of the important class of models in $JPBH^3$.

The Maple code used in the paper can be found at `http://www.kent.ac.uk/ims/personal/djc24/maplecode.htm`.

4 References

Catchpole, E. A. and Morgan, B. J. T. (1997) Detecting parameter redundancy. *Biometrika*, **84**, 187-196.

Cole, D. J. and Morgan, B. J. T. (2009) Determining the Parametric Structure of Non-linear Models. Submitted for publication.

Jiang, H., Pollock, K. H., Brownie, C., Hightower, J. E., Hoenig, J. M., Hearn, W. S. (2007) Age-dependent tag return models for estimating fishing mortality, natural mortality and selectivity. *JABES*, **12**, 177-194.

5 Appendix

This appendix gives more detail on how the parameter redundancy status of $JPBH^3$'s models was determined.

In the case when there are 4 years of tagging and recovery and 2 age classes, a possible reparameterisation of the general form of the models is

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \\ s_{11} \\ s_{12} \\ s_{13} \\ s_{14} \\ s_{15} \\ s_{16} \\ s_{17} \end{bmatrix} = \begin{bmatrix} F_1\mathrm{Sel}_1 + M_1^A \\ F_1\mathrm{Sel}_2 + M_2^A \\ F_2\mathrm{Sel}_1 + M_2^Y M_1^A \\ F_2\mathrm{Sel}_2 + M_2^Y M_2^A \\ F_2 + M_2^Y M_2^A \\ F_3\mathrm{Sel}_1 + M_3^Y M_1^A \\ F_3\mathrm{Sel}_2 + M_3^Y M_2^A \\ F_3 + M_3^Y M_2^A \\ F_4\mathrm{Sel}_1 + M_4^Y M_1^A \\ F_4\mathrm{Sel}_2 + M_4^Y M_2^A \\ F_4 + M_4^Y M_2^A \\ F_1\mathrm{Sel}_1\lambda_1^A \\ F_1\mathrm{Sel}_2\lambda_2^A \\ F_1\lambda_2^A \\ F_2\mathrm{Sel}_1\lambda_2^Y\lambda_1^A \\ F_3\mathrm{Sel}_1\lambda_3^Y\lambda_1^A \\ F_4\mathrm{Sel}_1\lambda_4^Y\lambda_1^A \end{bmatrix}$$

This reparameterisation is not unique (as it has 17 entries but there are only 16 parameters) The original exhaustive summary, formed from the non-zero entries of $\mathbf{P}$, is rewritten in terms of $\mathbf{s}$ to give:

$$\boldsymbol{\kappa}(\mathbf{s}) = \begin{bmatrix} s_{12}\{1-\exp(-s_1)\}/s_1 \\ s_{13}s_{15}\{1-\exp(-s_4)\}\exp(-s_1)/(s_{12}s_4) \\ \vdots \\ s_{13}s_{17}\{1-\exp(-s_{10})\}/(s_{12}s_{10}) \end{bmatrix}.$$

The derivative matrix $\mathbf{D}_s = \dfrac{\partial\boldsymbol{\kappa}(\mathbf{s})}{\partial\mathbf{s}}$ has rank 16. We find a new reduced-form exhaustive summary by first solving $\boldsymbol{\alpha}\mathbf{D}_s = 0$. This reveals that $s_1, \ldots, s_8, s_{12}, \ldots, s_{16}$ are estimable, but $s_9, s_{10}, s_{11}, s_{17}$ are not estimable. The reparameterisation theorem of Cole and Morgan (2009) then suggests that you solve a particular set of linear partial differential equations in order to obtain expressions for the estimable parameters. Frequently Maple can provide the solution to this set of partial differential equations, but interestingly in this instance it fails to do so. Instead $\boldsymbol{\kappa}(\mathbf{s})$ can be examined visually, in order to see that a reparameterisation of length 16 is that of equation (1). The derivative matrix formed with respect to equation (1), $\mathbf{D}_r = \dfrac{\partial \boldsymbol{\kappa}(\mathbf{r})}{\partial \mathbf{r}}$, also has rank 16, but this time the reparameterisation is unique so we can apply all of the reparameterisation theorem. As $\mathbf{D}_r$ is full rank $\mathbf{r}$ is a reduced form exhaustive summary and there are 16 estimable parameters in the model. This can be further confirmed by finding the rank of derivative matrix $\mathbf{D} = \left[\dfrac{\partial r_\ell}{\partial \theta_i}\right]$, which also has rank 16.

By inspection, we can deduce that a general reduced form exhaustive summary for N years of tagging and recovery and K age classes is

$$\mathbf{r}^{ge} = \begin{bmatrix} F_1 \mathrm{Sel}_1 + M_1^A \\ F_1 \mathrm{Sel}_2 + M_2^A \\ \vdots \\ F_1 \mathrm{Sel}_K + M_K^A \\ F_2 \mathrm{Sel}_1 + M_2^Y M_1^A \\ F_2 \mathrm{Sel}_2 + M_2^Y M_2^A \\ \vdots \\ F_2 \mathrm{Sel}_K + M_2^Y M_K^A \\ F_2 + M_2^Y M_K^A \\ \vdots \\ F_{N-1} \mathrm{Sel}_1 + M_{N-1}^Y M_1^A \\ F_{N-1} \mathrm{Sel}_2 + M_{N-1}^Y M_2^A \\ \vdots \\ F_{N-1} \mathrm{Sel}_K + M_{N-1}^Y M_K^A \\ F_{N-1} + M_{N-1}^Y M_K^A \\ F_1 \mathrm{Sel}_1 \lambda_1^A \\ F_1 \mathrm{Sel}_2 \lambda_2^A \\ \vdots \\ F_1 \mathrm{Sel}_K \lambda_K^A \\ F_1 \lambda_2^A \\ F_2 \mathrm{Sel}_1 \lambda_2^Y \lambda_1^A \\ \vdots \\ F_{N-1} \mathrm{Sel}_1 \lambda_{N_1}^Y \lambda_1^A \\ F_N \mathrm{Sel}_1 \lambda_N^Y \lambda_1^A \{1 - \exp(-F_N \mathrm{Sel}_1 - M_N^Y M_1^A)\}/(F_N \mathrm{Sel}_1 + M_N^Y M_1^A) \\ F_N \mathrm{Sel}_2 \lambda_N^Y \lambda_2^A \{1 - \exp(-F_N \mathrm{Sel}_2 - M_N^Y M_2^A)\}/(F_N \mathrm{Sel}_2 + M_N^Y M_2^A) \\ \vdots \\ F_N \mathrm{Sel}_K \lambda_N^Y \lambda_K^A \{1 - \exp(-F_N \mathrm{Sel}_K - M_N^Y M_K^A)\}/(F_N \mathrm{Sel}_K + M_N^Y M_K^A) \\ F_N \lambda_N^Y \lambda_K^A \{1 - \exp(-F_N - M_N^Y M_K^A)\}/(F_N + M_N^Y M_K^A) \end{bmatrix}.$$

In order to generalise the result that this model is full rank to $N \geq 4$ years of tagging and recovery we use the extension theorem of Cole and Morgan (2009), which states if the derivative matrix formed from the extra exhaustive summary terms is full rank as well as the original derivative matrix, then the extended model is also full rank. However when we go from 4 years to 5 years of tagging and recovery the last three terms of the new exhaustive summary do not appear in the exhaustive summary for 5 years of tagging and recovery. Instead we split the exhaustive summary into two parts, as shown in the Maple worksheet. As both derivative matrices are full rank, the extended model is full rank - this can be generalised to N years of tagging and recovery using induction. To generalise the result that this model is full rank for $K \geq 2$ age classes we first need to find a new reduced-form exhaustive summary (the details of which are given in the Maple worksheet). We can therefore say by the extension theorem that for $N \geq 4$ years of tagging and recovery and $K \geq 2$ age classes this model is always full rank.