

Detecting Parameter Redundancy in Covariate Models

D. J. Cole and B. J. T. Morgan

July 31, 2007

Abstract

Adding covariates to a parameter redundant model could reduce the number of parameters and result in a model that is no longer parameter redundant. However adding covariates also complicates calculation of the rank of the derivative matrix required to determine symbolically whether a model is parameter redundant or not. We show how it is possible to determine the number of estimable parameters, and hence whether a covariate model is parameter redundant or not, from the model without covariates, thereby simplifying calculation considerably. This is illustrated using examples in ring-recapture-recovery analysis, with Maple code available for the examples.

Keywords: computer algebra, identifiability, Maple, parameter redundant, recapture data, recovery data, symbolic algebra.

1 Introduction

A parameter redundant model is a model for which it is not possible to estimate all the parameters in the model, resulting in a non-identifiable model. A model that is not parameter redundant is termed full rank and is at least locally identifiable. One method of overcoming parameter redundancy is to add covariates to a model, which can result in a full rank model. This paper is concerned with determining if

a model containing covariates is full rank or not.

In Catchpole and Morgan (1997) and Catchpole *et al* (1998) a symbolic method is developed for detecting parameter redundancy within exponential family models. This method involves calculating the derivative matrix \mathbf{D} , with elements

$$D_{ij} = \frac{\partial \mu_j}{\partial \theta_i},$$

where μ_j is the j th mean of the exponential family model and θ_i is the i th parameter. If the row rank of \mathbf{D} is equal to the number of parameters, p , then the model is full rank. If the rank of \mathbf{D} is less than p the model is parameter redundant. The deficiency of a model is $d = \text{rank}(\mathbf{D}) - p$. The rank of \mathbf{D} can be determined using a symbolic algebra computer package, such as Maple (see Catchpole *et al*, 2002) The Maple code for the examples in this paper can be downloaded from <http://www.kent.ac.uk/ims/personal/djc24/maplecode.htm>.

The rank of \mathbf{D} is equal to how many parameters are estimable. It is possible to tell which, if any, of the original parameters are estimable by solving $\alpha(\theta)^T \mathbf{D}(\theta) = 0$. There are d solutions to $\alpha(\theta)^T \mathbf{D}(\theta) = 0$, denoted by $\alpha_j(\theta)$ for $j = 1, \dots, d$, with individual entries $\alpha_{ij}(\theta)$. Any $\alpha_{ij}(\theta)$ which are zero for j correspond to a parameter which is estimable (Catchpole *et al*, 1998). In order to find the parameter combinations, which are also estimable, we need to solve the system of linear first-order partial differential equations

$$\sum_{i=1}^p \alpha_{ij} \frac{\partial f}{\partial \theta_i} = 0 \text{ for } j = 1 \dots r$$

(Catchpole *et al*, 1998).

A limitation of Maple is that matrix dimensions cannot be specified symbolically: it is only possible to calculate the rank of a matrix of a given size, and so for large derivative matrices the computer will not have enough memory to be able to calculate the symbolic rank of the matrix. This problem is solved for product-multinomial models using an extension theorem. Catchpole and Morgan (1997) show that if a model defined by an $r \times c$ table is full rank then a corresponding

$r' \times c'$ ($r' \geq r$, $c' \geq c$) table will also be full rank, if the subsidiary derivative matrices from adding one extra row and one extra column are also full rank. Catchpole and Morgan (2001) show how to apply the extension theorem when a model is parameter redundant.

The Catchpole and Morgan approach can be used to determine parameter redundancy for models with covariates in the same way as for models without covariates. This is demonstrated below using a simple example of adding covariates to the Cormack-Jolly-Seber (CJS) model for capture-recapture data (Cormack, 1964, Seber, 1965 and Jolly, 1965) with 3 years of release and recapture. This model has 6 parameters $\theta = [\phi_1, \phi_2, \phi_3, \lambda_2, \lambda_3, \lambda_4]$, where ϕ_i is the probability that an animal survives from occasion i to $i + 1$, and λ_i is the probability an animal is recaptured on occasion i . The probabilities of recapture for the first time on occasion j , given release on occasion i are given by the (i, j) th entry of the probability matrix \mathbf{P} , where

$$\mathbf{P} = \begin{bmatrix} \phi_1\lambda_2 & \phi_1(1-\lambda_2)\phi_2\lambda_3 & \phi_1(1-\lambda_2)\phi_2(1-\lambda_3)\phi_3\lambda_4 \\ 0 & \phi_2\lambda_3 & \phi_2(1-\lambda_3)\phi_3\lambda_4 \\ 0 & 0 & \phi_3\lambda_4 \end{bmatrix}$$

This model is known to be parameter redundant with a deficiency of 1 as the parameters ϕ_3 and λ_4 only ever appear as $\phi_3\lambda_4$. If we add covariates to ϕ_i with

$$\phi_i = \frac{1}{1 + \exp(a + bx_i)},$$

then there are now 5 parameters with $\theta = [a, b, \lambda_2, \lambda_3, \lambda_4]$. The probability matrix then becomes

$$\mathbf{P} = \begin{bmatrix} \frac{\lambda_2}{1+\exp(a+bx_1)} & \frac{(1-\lambda_2)\lambda_3}{\{1+\exp(a+bx_1)\}\{1+\exp(a+bx_2)\}} & \frac{(1-\lambda_2)(1-\lambda_3)\lambda_4}{\{1+\exp(a+bx_1)\}\{1+\exp(a+bx_2)\}\{1+\exp(a+bx_3)\}} \\ 0 & \frac{\lambda_3}{1+\exp(a+bx_2)} & \frac{(1-\lambda_3)\lambda_4}{\{1+\exp(a+bx_2)\}\{1+\exp(a+bx_3)\}} \\ 0 & 0 & \frac{\lambda_4}{1+\exp(a+bx_3)} \end{bmatrix}.$$

Catchpole and Morgan (1997) show why it is sufficient to examine \mathbf{P} (or $\ln(\mathbf{P})$) in

product-multinomial models rather than μ . Therefore the derivative matrix is

$$\mathbf{D} = \frac{\partial \mathbf{P}}{\partial \theta},$$

and its rank can be found using Maple (see Maple sheet CJSexample.mw) The rank of \mathbf{D} is equal to 5, and therefore the model with covariates is full rank.

In this example if $x_1 = x_2$ this model becomes parameter redundant; Section 2 shows how it is possible to detect this and how to show whether a full rank model is always full rank, or whether it is near-parameter redundant.

The CJS model is a very simple model in terms of detecting parameter redundancy, so that adding covariates causes few problems. However for more complex models, Maple may not be able to calculate the rank of \mathbf{D} , because Maple runs out of memory. In Section 3 we examine how we can determine the parameter redundancy status of a model with covariates from the equivalent model without covariates. Finally in Section 4 the methods of the paper are illustrated with several examples.

2 Essentially and conditionally full rank and nearly parameter redundancy

Catchpole and Morgan (1997) distinguish between essentially and conditionally full rank models. This is particularly important when there are covariates in the model, as covariates being equal can often result in a conditionally full rank model. An essentially full rank model is a full rank model that is full rank for all θ and possible values of the covariates. A conditionally full rank model is a model that is not full rank for all θ and possible values of the covariates. Gimenez *et al* (2003) show that it is possible to distinguish between essentially and conditionally full rank models by considering decompositions of the derivative matrix. One method of decomposition is the LU decomposition, which can easily be executed in Maple using the code

```
(Q,L,U,R) := LUDecomposition(DD,output=['Q','L','U1','R']);
```

This gives $\mathbf{D} = \mathbf{QLUR}$, where \mathbf{Q} is a permutation matrix, \mathbf{L} is a lower triangular matrix with 1s on the diagonal, \mathbf{U} is an upper triangular matrix and \mathbf{R} is in reduced echelon form. As theorem 1.6.7 of Graybill (1969, page 13) states that if \mathbf{A} and \mathbf{B} are non-singular then $\text{rank}(\mathbf{C}) = \text{rank}(\mathbf{AC}) = \text{rank}(\mathbf{CB}) = \text{rank}(\mathbf{ACB})$, then \mathbf{R} can be used to determine the rank of \mathbf{D} as long as \mathbf{Q} , \mathbf{L} and \mathbf{U} are non-singular. As both \mathbf{Q} and \mathbf{L} have determinants of 1, all that is needed is to check that the determinant of \mathbf{U} is not equal to zero. If the determinant of \mathbf{U} is equal to zero for particular values of covariates (or θ) then a full rank model could be parameter redundant at those points. If $\mathbf{U} = 0$ at z and \mathbf{R} is undefined at z , which could occur if evaluating \mathbf{R} at z requires dividing by zero, then it is possible that the model is still full rank at z and further checks of the rank of \mathbf{D} at z and LU decompositions of \mathbf{D} at z are required (as shown at the end of the example below).

This method can also be used to check near parameter redundancy. If the determinant of \mathbf{U} is near to zero (or near to infinity) for particular values of covariates or θ then a full rank model is nearly parameter redundant at those points.

For the CJS model let $\mathbf{D} = \mathbf{QLUR}$. The determinant of \mathbf{U} is

$$\text{Det}(\mathbf{U}) = \frac{-(x_1 - x_2)(1 - \lambda_2)\lambda_3\lambda_4 \exp(a + bx_1) \exp(a + bx_2)}{\{1 + \exp(a + bx_1)\}^4 \{1 + \exp(a + bx_2)\}^4 \{1 + \exp(a + bx_3)\}^2}.$$

This is equal to 0 at $x_1 = x_2$ and \mathbf{R} is defined at this point, therefore this model is conditionally full rank, parameter redundant at $x_1 = x_2$ and near parameter redundant when x_1 is close to x_2 .

If the CJS model is extended to four years of recapture and recovery then this CJS model with covariates now has $p = 6$ parameters and rank 6, and so is full rank. Again letting $\mathbf{D} = \mathbf{QLUR}$, the determinant of \mathbf{U} is equal to 0 at $x_1 = x_2$. \mathbf{R} is not defined at $x_1 = x_2$ and still has rank 6 when $x_1 = x_2$. However if \mathbf{D}_{12} denotes \mathbf{D} evaluated at $x_1 = x_2$, and writing $\mathbf{D}_{12} = \mathbf{Q}_{12}\mathbf{L}_{12}\mathbf{U}_{12}\mathbf{R}_{12}$, the determinant of \mathbf{U}_{12} is zero at $x_3 = x_1$, and \mathbf{R}_{12} is defined at $x_3 = x_1$. Therefore this model is parameter redundant at $x_1 = x_2 = x_3$. (Maple code for the CJS model is in the worksheet CJSexample.mw).

3 Using Reparameterisation to simplify the derivative matrix

Consider first the model without covariates, which has p parameters θ . Using the Catchpole and Morgan method for detecting parameter redundancy the derivative matrix is $\mathbf{D} = \partial\mu/\partial\theta$. The rank of \mathbf{D} is q , so the model without covariates has q estimable parameters. The reduced parameter set of estimable parameters and parameter combinations is β (if $q = p$, $\beta = \theta$), and $\mathbf{D}_\beta = \partial\mu/\partial\beta$ is of full rank. Here we assume that β is a function of every θ_i . Next consider the model with covariates, which has p_c parameters θ_c . The derivative matrix is now $\mathbf{D}_c = \partial\mu_c/\partial\theta_c$, where $\mu_c(\theta_c) = \mu(\theta)$. Rather than calculate the rank of \mathbf{D}_c , the number of estimable parameters in the model with covariates is equal to $\min(p_c, q)$ (a proof is given in Appendix A).

The derivative matrix

$$\mathbf{D}_2 = \frac{\partial\beta(\theta_c)}{\partial\theta_c}$$

can be used to determine the set of estimable parameters and whether the model is essentially or conditionally full rank. (To find all boundary conditions at which the model is conditionally full rank an LU decomposition of \mathbf{D}_β is also required).

A parameter is said to be completely redundant when that parameter never appears in μ , this typically occurring if there are missing data (as shown in the example of Section 4.2). This results in the corresponding row of the derivative matrix having all zero entries. If β is not a function of every θ_i because θ_k is completely redundant then it is possible to apply Theorem 1 by first removing θ_k from θ .

We return again to the CJS model with covariates that has 3 years of release and 3 years of recapture, and has 5 parameters ($p_c = 5$). The corresponding model without covariates has 6 parameters and is parameter redundant with deficiency 1, therefore there are 5 estimable parameters in the model ($q = 5$). Hence $\text{rank}(\mathbf{D}_c) = \min(5, 5) = 5$, therefore the model with covariates is full rank.

The matrix $\mathbf{D}_2 = \partial\beta/\partial\theta$ can be used to check whether this is conditionally full

rank. Here

$$\beta = [\phi_1, \phi_2, \lambda_2, \lambda_3, \phi_3\lambda_4] = \left[\frac{1}{1 + \exp(a + bx_1)}, \frac{1}{1 + \exp(a + bx_2)}, \lambda_2, \lambda_3, \frac{\lambda_4}{1 + \exp(a + bx_3)} \right].$$

Let $\mathbf{D}_2 = \mathbf{QLUR}$. The determinant of \mathbf{U} is

$$\text{Det}(\mathbf{U}) = \frac{-(x_1 - x_2) \exp(a + bx_1) \exp(a + bx_2)}{\{1 + \exp(a + bx_1)\}^2 \{1 + \exp(a + bx_2)\}^2 \{1 + \exp(a + bx_3)\}},$$

again this is equal to 0 at $x_1 = x_2$ therefore this model is conditionally full rank, parameter redundant at $x_1 = x_2$ and near parameter redundant when x_1 is close to x_2 .

Suppose that we add a further covariate w_i so that

$$\phi_i = \frac{1}{1 + \exp(a + bx_i + cw_i)}.$$

There are now $p = 6$ parameters ($\theta = [a, b, c, \lambda_2, \lambda_3, \lambda_4]$), but the number of estimable parameters is still $q = 5$, so this model would be parameter redundant. However if we added another year of recapture and another year of recovery there are now $p = 7$ parameters ($\theta = [a, b, c, \lambda_2, \lambda_3, \lambda_4, \lambda_5]$), and there are $q = 7$ estimable parameters in the model without covariates ($\phi_1, \phi_2, \phi_3, \lambda_2, \lambda_3, \lambda_4, \phi_4\lambda_4$), so the model with covariates has at most 7 estimable parameters, and this model is full rank.

In general for k years of release and k years of recovery there are $q = (2k - 1)$ estimable parameters in the CJS model without covariates, so there are at most $(2k - 1)$ estimable parameters in the CJS model with covariates. As long as the model with covariates has fewer than $(2k - 1)$ parameters it will be full rank. Furthermore the estimable parameters are $\beta = [\phi_1, \dots, \phi_{k-1}, \lambda_2, \dots, \lambda_k, \phi_k\lambda_{k+1}]$ and if $\phi_i = 1/\{1 + \exp(a + bx_i)\}$ the model is full rank for $k \geq 3$, but only conditionally full

rank with it being parameter redundant at $x_1 = x_2 = \dots = x_{k-1}$. As

$$\mathbf{D}_2 = \frac{\partial \beta}{\partial \theta} = \begin{bmatrix} \frac{-\exp(a+bx_1)}{1+\exp(a+bx_1)} & \dots & \frac{-\exp(a+bx_{k-1})}{1+\exp(a+bx_{k-1})} & 0 & \dots & 0 & \frac{-\lambda_{k+1} \exp(a+bx_k)}{1+\exp(a+bx_k)} \\ \frac{-x_1 \exp(a+bx_1)}{1+\exp(a+bx_1)} & \dots & \frac{-x_{k-1} \exp(a+bx_{k-1})}{1+\exp(a+bx_{k-1})} & 0 & \dots & 0 & \frac{-\lambda_{k+1} x_k \exp(a+bx_k)}{1+\exp(a+bx_k)} \\ 0 & \dots & 0 & 1 & \dots & 0 & 0 \\ & & & & \ddots & & \\ 0 & \dots & 0 & 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & \frac{1}{1+\exp(a+bx_k)} \end{bmatrix}$$

it is easy to see that if $x_1 = x_2 = \dots = x_{k-1}$ then the row operation of row 2 = row 2 $-x_1 \times$ row 1, would produce a row with all zero entries apart from the last row. Hence the model would not be full rank.

4 Examples

In this section two further example of adding covariates to ecological models are examined.

4.1 Conditional analyses in ring recovery data

In ring recovery data the total number of birds ringed in each year may be unknown. In this case a model can only be fitted by conditioning on the number of birds recovered. The most commonly used model for such a conditional analysis assumes that there is a constant probability of survival, λ . For this model it is not possible to estimate λ (see Kgosi, 2000). However there is evidence that the reporting probability is decreasing over time (Baillie and Green, 1987). The resulting model is parameter redundant, as shown below. However it is also shown that introducing a logistic regression can give a full rank model.

It is assumed that the probability of surviving the first year is ϕ_1 , the probability of surviving other years is ϕ_a and the reporting probability, λ_j , is dependent on time. Thus the probability of being ringed in year i and recovered in year j conditional

on the number of ringings in a year is

$$P_{ij} = \begin{cases} \frac{(1 - \phi_1)\lambda_j}{(1 - \phi_1)\lambda_i + \sum_{k=i+1}^J \phi_1 \phi_a^{k-i-1} (1 - \phi_a)\lambda_k} & i = j \\ \frac{\phi_1 \phi_a^{j-i-1} (1 - \phi_a)\lambda_j}{(1 - \phi_1)\lambda_i + \sum_{k=i+1}^J \phi_1 \phi_a^{k-i-1} (1 - \phi_a)\lambda_k} & i < j \end{cases} \quad (1)$$

for $i = 1, \dots, I$ and $j = 1, \dots, J$.

If we consider 3 years of ringing and 3 years of release ($I = 3$ and $J = 3$), the derivative matrix is

$$\mathbf{D} = \frac{\partial \ln(\mathbf{P})}{\partial \theta},$$

where the parameters are $\theta = [\phi_1, \phi_a, \lambda_1, \lambda_2, \lambda_3]$. The derivative matrix and its rank are calculated in Maple (see worksheet conditionalexample.mw). The derivative matrix has rank 3 and as there are 5 parameters this model is parameter redundant with deficiency 2. The estimable parameter combinations are

$$\begin{aligned} \psi &= \frac{\phi_a - \phi_1}{\phi_a(1 - \phi_1)} \\ \tau_2 &= \frac{\phi_a \lambda_2}{\lambda_1} \\ \tau_3 &= \frac{\phi_a^2 \lambda_3}{\lambda_1}. \end{aligned}$$

In general by multiplying the numerator and denominator of equation (1) by $\frac{\phi_a^{i-1}}{\lambda_1(1 - \phi_1)}$ it can be shown that the probability matrix can be reparameterised to give

$$P_{ij} = \begin{cases} \frac{\tau_j}{\tau_j + (1 - \psi) \sum_{k=i+1}^J \tau_k} & i = j \\ \frac{(1 - \psi)\tau_j}{\tau_j + (1 - \psi) \sum_{k=i+1}^J \tau_k} & i < j \end{cases}$$

with $\tau_j = \phi_a^{j-1} \lambda_j / \lambda_1$ and $\tau_1 = 1$. It is possible to show that this reparameterisation is always full rank by considering again the case of 3 years of ringing and 3 years of release with parameters $\theta = [\psi, \tau_2, \tau_3]$. The rank of the derivative matrix for this reparameterised model is 3, so this model is full rank. Using the extension theorem (Catchpole and Morgan, 1997) this model is therefore always full rank if $I \geq 3$ and $J \geq 3$.

As the estimable set of parameters for this model is $\beta = \left[\frac{\phi_a - \phi_1}{\phi_a(1 - \phi_1)}, \frac{\phi_a \lambda_2}{\lambda_1}, \dots, \frac{\phi_a^{J-1} \lambda_J}{\lambda_1} \right]$, if we add covariates to the model as long as the total number of parameters for the covariate model is no more than J , the covariate model will be full rank. For example, consider λ as a logistic regression on time with $\lambda_j = \frac{1}{1 + \exp(\lambda_a + \lambda_b j)}$; this would allow the consideration of whether λ is decreasing with time. This model has 4 parameters $\theta_c = [\phi_1, \phi_a, \lambda_a, \lambda_b]$. If $J = 3$ this covariate model is still parameter redundant, however if $J \geq 4$ this covariate model is full rank.

4.2 Penguin ring recovery

This example examines a subset of recovery data of Little Penguins (*Eudyptula minor*) on Philip Island. The numbers of penguins found dead in year j , given that they were tagged as chicks in year i , is given by (i, j) th element of the matrix

$$\mathbf{N} = \begin{bmatrix} 40 & 1 & 0 & 2 & 3 & 2 & 0 & 0 \\ & 17 & 5 & 1 & 2 & 1 & 3 & 3 \\ & & 3 & 1 & 0 & 0 & 0 & 0 \\ & & & 14 & 9 & 2 & 3 & 1 \\ & & & & 16 & 1 & 0 & 2 \\ & & & & & 14 & 4 & 1 \\ & & & & & & 18 & 1 \\ & & & & & & & 7 \end{bmatrix}$$

where $i = 1$ is the year 1996 and $i = 8$ is the year 2003 (Sidhu, 2007).

The model considered here has the reporting probability dependent on time and the

probability of survival dependent on age and time. We wish to examine the effect that the assumption of time varying covariates has on the parameter redundancy status of this model.

The probability of a penguin being recovered dead between time t_j and t_{j+1} given that it was marked at time t_i is

$$P_{i,j} = \begin{cases} (1 - \phi_{j-i+1,j})\lambda_j & i = j \\ \left(\prod_{k=1}^{j-i} \phi_{k,j} \right) (1 - \phi_{j-i+1,j})\lambda_j & i < j \end{cases},$$

where $\phi_{a,j}$ is the probability a penguin aged a alive at time t_j survives until time t_{j+1} .

The derivative matrix is

$$\mathbf{D}_{i,j} = \frac{\partial \kappa_j}{\partial \theta_i}$$

where κ is a vector with all the non-zero elements of $N_{i,j} \ln(P_{i,j})$ (again see Catchpole and Morgan, 1997 for why we can consider $N_{i,j} \ln(P_{i,j})$ rather than μ). The parameters are $\theta = [\phi_{1,1}, \phi_{1,2}, \dots, \phi_{8,8}, \lambda_1, \dots, \lambda_8]$.

It is important to note that the parameters $\phi_{3,3}$, $\phi_{7,7}$ and $\phi_{8,8}$ are completely redundant (this is because these parameters only appear in $P_{1,3}$, $P_{1,7}$, $P_{1,8}$ respectively and $N_{1,3} = N_{1,7} = N_{1,8} = 0$). These parameters are ignored (as they only result in zero rows in \mathbf{D}). The rank of the derivative matrix is 28. The model is therefore parameter redundant (with a deficiency of 13). (Calculation for this example can be found in Maple worksheet `penguinexample.mw`).

If we want to add covariates to this model, as long as there are 28 or fewer parameters in the covariates model this model will be full rank. For example, Sidhu (2007) considers the time varying covariates: chicks per pair (cpp), average weight at banding (bw), mean laying date (mld) and southern oscillation index (SOI). If ϕ was a logistic function of constant + age + time + chpp + bw + mld + SOI, then the model would have 12 parameters and would be full rank.

5 Discussion

We have shown that the parameter redundancy status of a model with covariates can be determined from simply the same model without covariates. As adding covariates to a model can make derivative matrix calculations much more complex, this method simplifies the calculation considerably.

It is likely that covariate models are full rank, especially in recovery or recapture models when considering many years of data. However it is possible that the models become parameter redundant at specific values of the covariates, for example in the CJS model when $x_1 = \dots = x_{k-1}$. An LU decomposition provides a useful method of checking for these conditionally full rank models. Near parameter redundancy can also occur when the covariates are close to these specific values, resulting in problems estimating parameters.

Further work by Cole and Morgan (unpublished) looks at generalising the reparameterisation method to simplify more complicated derivative matrix in a more general setting.

6 References

- Baillie, S. R. and Green, R. E. (1987) The importance of variation in recovery rates when estimating survival rates from ringing recoveries, *Acta Ornithol*, **23**, 41-60.
- Catchpole, E. A. and Morgan, B. J. T (1997) Detecting parameter redundancy. *Biometrika*, **84**, 187-196.
- Catchpole, E. A., Morgan, B. J. T and Freeman, S. N. (1998) Estimation in parameter redundant models *Biometrika*, **85**, 462-468.
- Catchpole, E. A. and Morgan, B. J. T (2001) Deficiency of parameter redundant models. *Biometrika*, **88**, 593-598.
- Catchpole, E. A., Morgan, B. J. T and Viallefont, A. (2002) Solving problems in parameter redundancy using computer algebra. *Journal of Applied Statistics*, **29**,

625-636.

Cormack, R. M (1964) Estimates of survival from the sightings of marked animals. *Biometrika*, **51**, 429-438.

Gimenez, O., Choquet, R. and Lebreton, J. (2003) Parameter Redundancy in Multistate Capture-Recapture Models. *Biometrical Journal*, **45**, 704-722.

Graybill, , F. A. (1969) Introduction to Matrices with Applications in Statistics. Wadsworth Publishing Company, Inc., Belmont, California.

Kgosi, P. M. (2000) Models for Avian Demography. Ph.D. thesis, University of Kent, Canterbury, UK.

Jolly, G. M. (1965) Explicit estimates from capture-recapture data with both death and immigration. *Biometrika*, **52**, 225-247.

Sidhu, L. A. (2007) Analysis of recovery-recapture data for little penguins. Ph.D. thesis, University of New South Wales, Australia.

Seber, G. A. F. (1965) A note on the multiple recapture census. *Biometrika*, **52**, 249-259.

A Proof that the model without covariates can be used to determine whether a model with covariates is parameter redundant

The number of estimable parameters in the model with covariates is equal to $\min(p_c, q)$ because by the chain rule

$$\mathbf{D}_c = \frac{\partial \beta}{\partial \theta_c} \frac{\partial \mu}{\partial \beta} = \mathbf{D}_2 \mathbf{D}_\beta.$$

Now the β are independent (as \mathbf{D}_β is full rank) and the rank of \mathbf{D}_β is q , therefore if $p_c \geq q$ the rank of \mathbf{D}_2 is q and if $p_c < q$ the rank of \mathbf{D}_2 is p_c . Therefore $\text{rank}(\mathbf{D}_2) = \min(p_c, q)$. As \mathbf{D} is a $q \times n$ matrix with rank q the rank of $\mathbf{D}_c = \mathbf{D}_2 \mathbf{D}$ has the same

rank as \mathbf{D}_2 . Therefore $\text{rank}(\mathbf{D}_c) = \min(p_c, q)$.