

# Non-linear maximum likelihood estimation of Beta and HOF response models

Jari Oksanen      Peter R. Minchin

May 2, 2002

## Abstract

Peter Minchin and I have a manuscript (Oksanen and Minchin, 2002) where we compare HOF models and some of their alternatives for estimating skewness of species response functions. We compared Gaussian, HOF, GAM and Beta response models. In this appendix we explain how we fitted the HOF models and Beta response models without fixing the endpoints (Oksanen, 1997). This appendix was too technical for inclusion with the article, but we made a brief version available here.

## 1 Introduction

Non-linear maximum likelihood estimation is relatively simple with modern software. In this appendix we explain the methods we used in the R software. We do not give general functions that would be applicable in most cases, but only the solutions we used, and these must be adapted for the use with other data sets and error models.

There are three components in a non-linear maximum likelihood model:

1. Function to give fitted values `fv` as a function of explanatory variable(s) and parameters.
2. Starting values for iterative search of best parameter values.
3. Likelihood function to be maximized.

We deal with points 1 and 2 separately for each model, but the last point, likelihood function is similar to both models, and so we discuss it here.

The purpose of the maximum likelihood estimation is to find the parameter values that give the best fitted values in the sense that they maximize the likelihood function. In least squares regression we minimize the sum of squared differences between fitted and observed values. In maximum likelihood regression we assess the probability of obtaining observed values given a set of fitted values, and the most likely fitted values are those that maximize this probability. In maximum likelihood regression, we usually minimize negative log-likelihood, which gives identical estimates to maximizing direct likelihood. We fitted binomial models with denominator 1, and so the following R statement gives the negative log-likelihood `ll` for a current estimate of fitted values `fv` and observed values `y`:

```
ll <- sum(-dbinom(y, 1, fv, log=TRUE))
```

In R we can use function `nlm` (non-linear minimizer) to find the best estimates for `ll`. This is different from S-plus where you use function `ms` with slightly different syntax.

## 2 Beta response function

We used the logarithm of Beta response function, so that exp-transformation would give the correct fitted values. Because we dealt with binary observations (presences and absences) we wanted to guarantee that the fitted values remain within the acceptable limits  $0 \dots 1$ , and therefore we used the logistic function. Moreover, we analysed both the general case where  $\alpha \neq \gamma$  with five parameters, and the 'symmetric' case where  $\alpha = \gamma$  with four parameters. The following function can identify these cases and produces the fitted values for either:

```
beta.fun <- function(p, x)
{
  zero <- .Machine$double.eps
  one <- 1 - zero
  if (length(p)==4) p <- c(p,p[4])
  eta <- p[1] + p[4]*log(x-p[1]) + p[5]*log(p[2]-x)
  fv <- plogis(eta)
  fv[x <= p[1]] <- ifelse(p[4] < 0, one, zero)
  fv[x >= p[2]] <- ifelse(p[5] < 0, one, zero)
  fv
}
```

Here `x` is the gradient (explanatory variable) and the parameters are given in the vector `p` in the order  $[p_1, p_2, k, \alpha, \gamma]$ . Please note that observed values `y` do not appear here.

As starting values, it is natural to use the coefficients from the corresponding GLM with fixed endpoints:

```
beta.start <- function(x, y, extend=50, symmetric=FALSE)
{
  p1 <- min(x[y>0]) - extend
  p2 <- max(x[y>0]) + extend
  inc <- x > p1 & x < p2
  if (symmetric)
    p <- coef(glm(y ~ I(log(x-p1)+log(p2-x)), family=binomial, subset=inc))
  else
    p <- coef(glm(y ~ log(x-p1) + log(p2-x), family=binomial, subset=inc))
  p <- c(p1,p2,p)
  p
}
```

Now we have everything we need for the maximum likelihood estimation, and we combine the calculation of fitted values and evaluation of its likelihood into a single function;

```

beta <- function(p, x, y)
{
  fv <- beta.fun(p, x)
  -sum(dbinom(y, 1, fv, log=TRUE))
}

```

The following fits the example in this data set (*Bauera rubioides* on the altitude gradient):

```
sol <- nlm(beta, p=beta.start(Altitude, BAUERUBI), x=Altitude, y=BAUERUBI)
```

Here we did not call function `fv` explicitly, but it was found by R scoping rules in the calling environment; this is probably different in S-plus, and it may be better to write a separate function to call both `fv` and `ll` and return the negative log-likelihood.

### 3 HOF response model

We have a series of five alternative HOF models. The following function takes the `model` as an arabic number and returns the fitted values:

```

HOF.fun <- function(p,model,x,M=1)
{
  p0 <- min(x)
  ra <- max(x) - p0
  x <- (x-p0)/ra
  switch(model,
    fv <- rep(M/(1+exp(p[1])),length(x)),
    fv <- M/(1+exp(p[1]+p[2]*x)),
    fv <- M/(1+exp(p[1]+p[2]*x))/(1+exp(p[3])),
    fv <- M/(1+exp(p[1]+p[2]*x))/(1+exp(p[3]-p[2]*x)),
    fv <- M/(1+exp(p[1]+p[2]*x))/(1+exp(p[3]-p[4]*x))
  )
  fv
}

```

To simplify the fitting, the gradient is scaled into  $0 \dots 1$ , which must be remembered when looking at the parameter estimates.

We do not have as a natural choice for starting values as in Beta response. However, model 4 is very similar to Gaussian response, and we found equations that translate the Gaussian optimum ( $u$ ) and height ( $h$ ) into corresponding parameters in HOF model IV, and an empirical relation with the Gaussian width ( $t$ ) and so far unsolved parameters in the HOF model IV. We start models IV and V with this values, that is, we start model V with a symmetric response as well. Model II is trivially related to logistic model, and we may start it with a GLM (and with identical error distribution, the starting values will be the final values as well). HOF model I is only a complicated way of writing an equation for constant abundance, and we find it directly as well. Model III is started with model II parameters as well. These are fairly bad starting values in many cases, and we may need to adjust the starting values. Putting all together, we get:

```

HOF.start <- function (x,y,model=5)
{
  x <- scale01(x)
  if (model >= 4) {
    k <- coef(glm(y ~ x + I(x^2), family=binomial))
    u <- -k[2]/2/k[3]
    h <- plogis(k[1] - k[2]^2/4/k[3])
    r <- log(1/h*(-2*h+2*sqrt(h))/2)
    b <- 5.07 - 0.227*k[3]
    a <- -b*u+r
    c <- b*u+r
  }
  else {
    k <- coef(glm(y ~ x, family=binomial))
    a <- -k[1]
    b <- -k[2]
    c <- 0
  }
  switch(model,
    p <- c(a=a),
    p <- c(a=a,b=b),
    p <- c(a=a,b=b,c=c),
    p <- c(a=a,b=b,c=c),
    p <- c(a=a,b=b,c=c,d=b)
  )
  p
}

```

Then we combine calculation the fitted values and evaluation of its likelihood:

```

HOF <- function(x, y, p, model, M=1)
{
  fv <- HOF.fun(p, model, x, M=M)
  -sum(dbinom(y, M, fv, log=TRUE))
}

```

and the example species can be fitted with:

```

p <- HOF.start(Altitude, BAUERUBI, model=5)
sol <- nlm(HOF, p, x=Altitude, y=BAUERUBI, model=5)

```

## References

- Oksanen, J., 1997. Why the beta-function cannot be used to estimate skewness of species responses. *Journal of Vegetation Science* 8, 147–152.
- Oksanen, J., Minchin, P. R., 2002. Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling* In press.