Profiling the parameters of models with linear predictors

The profileModel R package

Ioannis Kosmidis
I.Kosmidis@warwick.ac.uk

CRiSM
Research Fellow
Department of Statistics
THE UNIVERSITY OF WARWICK

useR! 2008, Dortmund

Outline

A variety of estimation methods

- Deviations from maximum likelihood:
  - $\longrightarrow$ Firth (1993) for penalized likelihoods and adjusted scores.
  - $\longrightarrow$ Lindsay (1988) for composite likelihoods.
- Estimating equations:
  - $\longrightarrow$ Wedderburn (1974); McCullagh (1983) for quasi-likelihoods.
  - $\longrightarrow$ Liang & Zeger (1986) for generalized estimating equations.

A variety of estimation methods (cont.)

- Appropriate objectives (inference functions) can be profiled:

  $\longrightarrow$ Heinze & Schemper (2002); Bull et al. (2007) for profiles of penalized likelihoods.

  $\longrightarrow$ Lindsay & Qu (2003) for profiles of appropriate quadratic score functions.

The profileModel R package

The *profileModel* **R** package has been developed to

- calculate,
- plot, and
- construct confidence intervals from

the profiles of user-defined objectives (via "plug-in" functions) for arbitrary *glm*-like fitted objects (`object`) with linear predictor.

Supported fitted objects

Fitted objects constructed according to Chambers & Hastie (1991, Chapter 2):

- The fitting procedure which results in `object` accepts `offset` in `formula`.
- `object$call` is the call that resulted in `object`.
- `object$terms` exists with the same meaning as for *lm*/*glm* objects.
- `coef(object)` returns a *vector* of coefficients with each component corresponding to a column of

  ```
  model.matrix(object)
  ```

The profileModel objective functions

- Restricted fit: Fix a parameter at a value and estimate the remaining parameters (using `offset`).
- The profiles of the objective are obtained/extracted from restricted fits.

The profileModel objective functions (cont.)

For example,

- `object` is the result of a `glm` call.
- Interest on the profiles of the log-likelihood (use deviance).

→ An appropriate *profileModel* objective is

```
profObj <- function(restrFit, dispersion)
  restrFit$deviance/dispersion
```

- Within the `profileModel` function:

→ the restricted fits for a grid of parameter values are obtained, and

→ for each restricted fit the difference

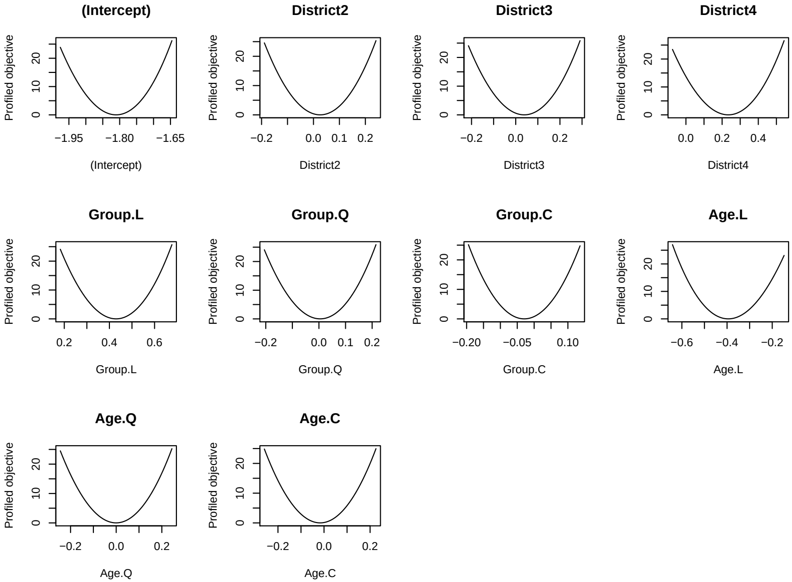
```
profObj(restrFit) - profObj(object)
```

is calculated.

Profiling some standard deviations away from the estimate

e.g.

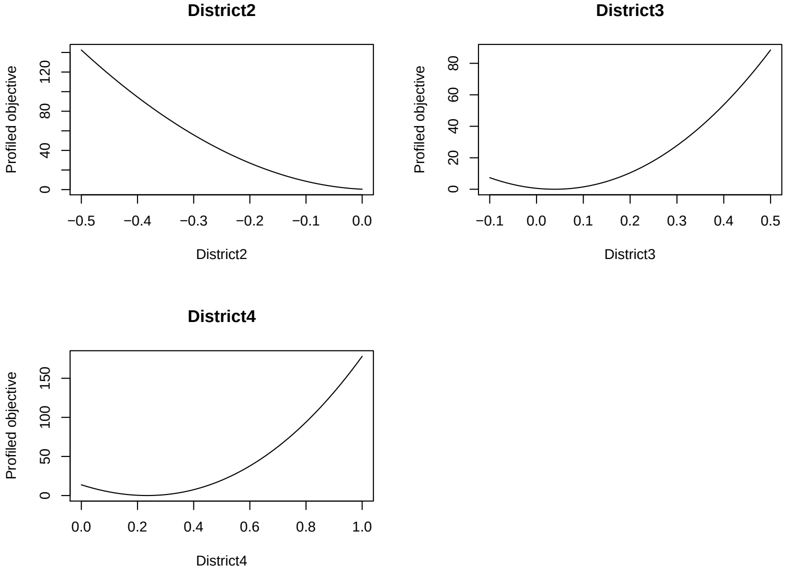
```
> library(MASS)
> m1 <- glm(Claims ~ District + Group + Age +
+     offset(log(Holders)), data = Insurance,
+     family = poisson)
> prof1 <- profileModel(m1, objective = profObj,
+     dispersion = 1)
> plot(prof1)
```


(Intercept)
Profiled objective
0
10
20
−1.95
−1.80
−1.65
(Intercept)
District2
Profiled objective
0
10
20
−0.2
0.0
0.1
0.2
District2
District3
Profiled objective
0
10
20
−0.2
0.0
0.2
District3
District4
Profiled objective
0
10
20
0.0
0.2
0.4
District4
Group.L
Profiled objective
0
10
20
0.2
0.4
0.6
Group.L
Group.Q
Profiled objective
0
10
20
−0.2
0.0
0.1
0.2
Group.Q
Group.C
Profiled objective
0
10
20
−0.20
−0.05
0.10
Group.C
Age.L
Profiled objective
0
10
20
−0.6
−0.4
−0.2
Age.L
Age.Q
Profiled objective
0
10
20
−0.2
0.0
0.2
Age.Q
Age.C
Profiled objective
0
10
20
−0.2
0.0
0.2
Age.C

Profiling over a grid of values.

e.g.

```
> prof2 <- update(prof1,
+     which = paste("District", 2:4, sep=""),
+     grid.bounds = c(-0.5, 0 , -0.1, 0.5, 0, 1))
> plot(prof2)
```

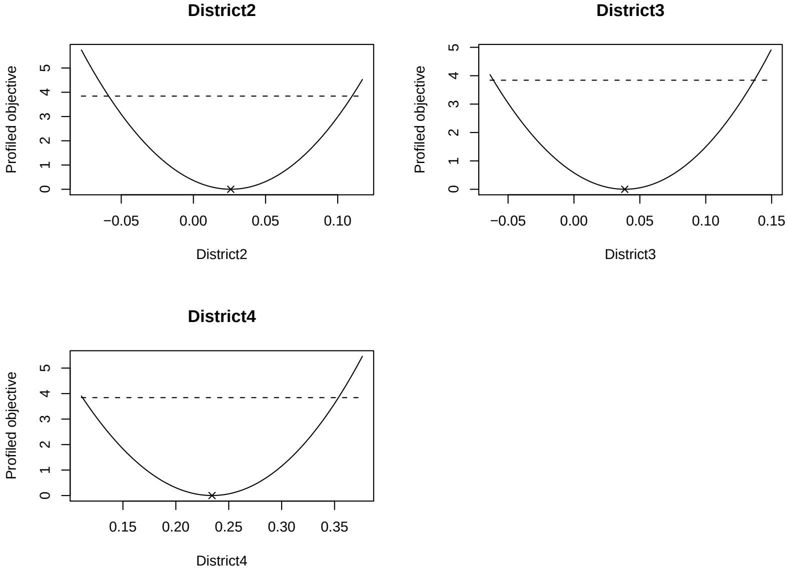

District2
Profiled objective
0
40
80
120
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
District2
District3
Profiled objective
0
20
40
60
80
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
District3
District4
Profiled objective
0
50
100
150
0.0
0.2
0.4
0.6
0.8
1.0
District4

Profiling until the profiles reach a certain value

- Construction of asymptotic confidence intervals.
- This procedure, currently, depends on the convexity of the objective.

e.g.

```
> prof3 <- update(prof2,
+     grid.bounds = NULL,
+     quantile = qchisq(0.95, 1))
> plot(prof3)
```


District2
Profiled objective
0
1
2
3
4
5
−0.05
0.00
0.05
0.10
District2
District3
Profiled objective
0
1
2
3
4
5
−0.05
0.00
0.05
0.10
0.15
District3
District4
Profiled objective
0
1
2
3
4
5
0.15
0.20
0.25
0.30
0.35
District4

Asymptotic confidence intervals based on the profiles

- Using spline smoothing.
  - It is fast.
  - Useful for routine use.
- Using a binary search.
  - It is slower than smoothing but it returns accurate (up to a tolerance) endpoints.
  - Useful when the spline does not approximate well the profile (large departures from quadratic behaviour or asymptotes) and for empirical coverage studies.

Profile likelihood for survreg objects

- An appropriate objective for *survreg* objects is

```
> profLogLik <- function(restrFit) {
+     -2*restrFit$loglik[2]
+ }
```

- Then,

```
> library(survival)
> m3 <- survreg(
+     Surv(futime, fustat) ~ ecog.ps + rx,
+     ovarian, dist= "weibull", scale = 1)
> prof.m3 <- profileModel(m3,
+     quantile=qchisq(0.95,1),
+     objective = profLogLik,
+     stdErrors = summary(m3)$table[,2])
```

Profile likelihood for survreg objects (cont.)

- The 95% asymptotic profile confidence intervals are

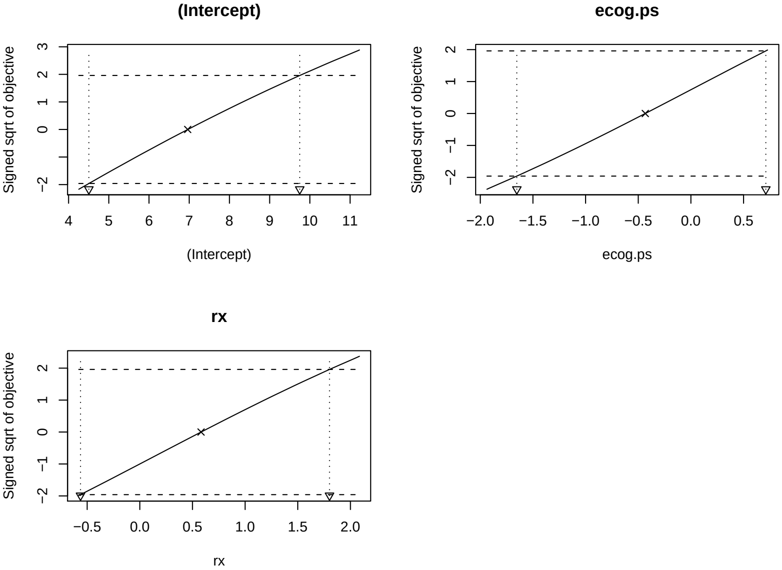
```
> ci.m3 <- profConfint(prof.m3)
> ci.m3
                 Lower     Upper
(Intercept)  4.5040322 9.7478129
ecog.ps     -1.6530027 0.7115067
rx          -0.5631386 1.8013708
```

- The 95% Wald asymptotic confidence intervals are

```
> confint(m3)
                 2.5 %    97.5 %
(Intercept)  4.3710056 9.5526696
ecog.ps     -1.5836210 0.7173517
rx          -0.5689836 1.7319891
```

- The confidence intervals are similar because the profiles are almost quadratic.

```
> plot(prof.m3, signed = TRUE, cis = ci.m3)
```

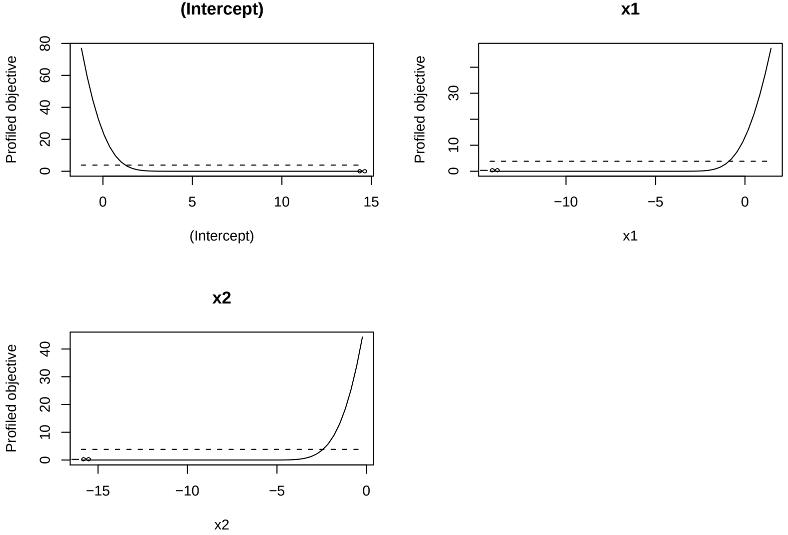

(Intercept)
Signed sqrt of objective
(Intercept)
ecog.ps
Signed sqrt of objective
ecog.ps
rx
Signed sqrt of objective
rx

Infinite maximum likelihood estimates

Data:

X_1	X_2	Successes	Totals
0	0	16	16
	1	1	13
1	0	12	20
	1	0	18

```
> x1 <- c(0, 0, 1, 1)
> x2 <- c(0, 1, 0, 1)
> y <- c(16, 1, 12, 0)
> tots <- c(16, 13, 20, 18)
> m2 <- glm(y/tots ~ x1 + x2,
+    weights = tots,
+    family=binomial(probit))
> coef(m2)
(Intercept)          x1          x2
   6.649437   -6.396090   -8.075514
> coef(summary(m2))[,"Std. Error"]
(Intercept)          x1          x2
   5914.617    5914.617    5914.617
```


(Intercept)
Profiled objective
0
20
40
60
80
0
5
10
15
(Intercept)
x1
Profiled objective
0
10
30
−10
−5
0
x1
x2
Profiled objective
0
10
20
30
40
−15
−10
−5
0
x2

Infinite maximum likelihood estimates (cont.)

- Default `profile` method

```
> confint(m2)
Waiting for profiling to be done...
                 2.5 %   97.5 %
(Intercept)  -511.1173       NA
x1                  NA 506.1762
x2          -2561.4923 382.1928
```

- The *profileModel*'s method for confidence intervals.

```
> confintModel(m2, quantile = qchisq(0.95, 1),
+    stepsize = 0.2, objective = profObj,
+    dispersion = 1, method = "zoom")
               Lower      Upper
(Intercept) 1.245953        Inf
x1              -Inf -0.8845107
x2              -Inf -2.4205613
```

Documentation and conclusions

- Package and complementary material
  - Package available on CRAN (`http://cran.r-project.org`).
  - For more examples and further features see `?profileModel` and `?confintModel`, and
  - complementary material for *profileModel* on `http://go.warwick.ac.uk/kosmidis/software`.
- Key features
  - It allows developers to have access to profiling capabilities by merely authoring a function for the objective to be profiled
    $\longrightarrow$ see `?RaoScoreStatistic` for an implementation of the quadratic score statistic for *glm*-like objects.
  - It provides an alternative to already implemented methods for profiling.
  - In its current version (0.5-4), it has been tested and it is known to work with objects resulting from `lm`, `glm`, `polr`, `gee`, `geeglm`, `brglm`, `BTm` and `survreg`.

Future development

- Profiling objectives for pairs of parameters and a method for plotting the contours of the profile.
- Quantile-based profiling and confidence intervals for non-convex objectives.
- Implementation using parallel computing.

Bull, S. B., Lewinger, J. B. & Lee, S. S. F. (2007). Confidence intervals for multinomial logistic regression in sparse data. *Statistics in Medicine* **26**, 903–918.

Chambers, J. M. & Hastie, T. (1991). *Statistical Models in S*. Chapman & Hall.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika* **80**, 27–38.

Heinze, G. & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine* **21**, 2409–2419.

Liang, K.-Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Lindsay, B. G. (1988). Composite likelihood methods. In *Statistical Inference from Stochastic Processes*, N. U. Prabhu, ed. American Mathematical Society.

Lindsay, B. G. & Qu, A. (2003). Inference functions and quadratic score tests. *Statistical Science* **18**, 394–410.

McCullagh, P. (1983). Quasi-likelihood functions. *The Annals of Statistics* **11**, 59–67.

Venables, W. N. & Ripley, B. D. (2002). Statistics complements to Modern Applied Statistics with S. URL `http://www.stats.ox.ac.uk/pub/MASS4/VR4stat.pdf`.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 439–447.