A binomial generalized additive model with a cubic spline smoother fitted by maximum likelihood

Generalized additive models (GAMs) have traditionally been estimated by maximizing a penalized likelihood. However, reparameterizing the smoother coefficients as a combination of fixed and random effects and estimating by maximum marginal likelihood has gained interest more recently (Wood 2006; Wager et al. 2007; Reiss and Ogden 2009; Wood 2011). The general GAM model for the mean of observation i with a single smoother is

$$g(\mu_i) = \beta_0 X_0(z_i) + \beta_1 X_1(z_i) + \cdots \beta_p X_p(z_i) = \mathbf{X}_i^T \boldsymbol{\beta}$$

where $\mathbf{X}_i(z_i)$ is the basis of the smoother (dimension p) of a continuous covariate z_i, $X_0(z_i) = 1$ and $g(\mu_i)$ is the link function that transforms the mean response to a linear function of the covariates.

In this example we use the approach described by Wood (2006, Section 6.6.1) to estimate a binomial GAM with a cubic spline smoother by maximum likelihood. Breifly, the basis matrix $\mathbf{X}$ and the corresponding penalty matrix $\mathbf{S}$ are obtained using gam() from the mgcv package in R (R Core Team 2012). The eigenvectors corresponding to the zero and non-zero eigenvalues ($\mathbf{U}_f$ and $\mathbf{U}_r$, respectively) are used to transform $\mathbf{X}$ into design matrices for the fixed and random effects, $\mathbf{X}_f = \mathbf{X}\mathbf{U}_f$ and $\mathbf{X}_r = \mathbf{X}\mathbf{U}_r$. The mean conditional on the random effects is

$$g(\mu_i|\mathbf{b}) = \log\left(\frac{\mu_i}{1-\mu_i}\right) = \mathbf{X}_{fi}^T \boldsymbol{\beta}_f + \mathbf{X}_{ri}^T \mathbf{b}$$

and the joint distribution of the random effects is $\mathbf{b} \sim N(\mathbf{0}, \mathbf{D}_+^{-1}/\lambda)$ where $\mathbf{D}_+$ is a diagonal matrix of the non-zero eigenvalues and λ is the smoothing parameter to be estimated.

Transformed coefficients that are comparable to those given by the gam function in R/mgcv are $\boldsymbol{\beta} = \mathbf{U}_f \boldsymbol{\beta}_f + \mathbf{U}_r \mathbf{b}$.

Comparison of ADMB-RE and R/mgcv results

There are two fixed effects, one smoothing parameter and 8 random effects in the example. The maximized log-likelihood values are identical for ADMB and R/mgcv (Table 1). There are slight differences in the estimated coefficients and corresponding standard errors due to the differences in estimation procedures and definitions of variance, but the predicted smoothers and 95% confidence intervals are virtually identical (Figure 1). The standard errors of the coefficients provided by ADMB are a bit greater due presumably to the second term in Eq. 3.1 of the manual on the random effects module which is a delta method term of the random effects as functions of any true parameters estimated by maximum likelihood.

References

R Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Reiss, P. T. and Ogden, R. T. 2009. Smoothing parameter selection for a class of semiparametric linear models. Journal of the Royal Statistical Society. Series B **72**(2): 505–523.

Wager, C., Vaida, F., and Kauermann, G. 2007. Model selection for penalized spline smoothing using Akaike Information Criteria. Australian & New Zealand Journal of Statistics **49**(2): 173–190.

Wood, S. N. 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall, Boca Raton, Florida, 392 pp.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society. Series B **73**(1): 3–36.

Table 1. Results from ADMB-RE and R/mgcv fits.

	ADMB-RE	R/mgcv
Maximized Log-likelihood	852.77	852.77
$\hat{\beta}_0$	0.83187	0.8294285
$\hat{\beta}_1$	0.76542	0.7643446
$\hat{\beta}_2$	0.18484	0.1875419
$\hat{\beta}_3$	-0.66980	-0.6644185
$\hat{\beta}_4$	-1.1888	-1.1823850
$\hat{\beta}_5$	-0.99387	-0.9889921
$\hat{\beta}_6$	-0.70880	-0.7067093
$\hat{\beta}_7$	-0.63130	-0.6309558
$\hat{\beta}_8$	-0.85946	-0.8582687
$\hat{\beta}_9$	-0.92880	-0.9265529
$\hat{\lambda}$	18947.74	18961.75
$\widehat{SE}\left(\hat{\beta}_0\right)$	0.10952	0.1075556
$\widehat{SE}\left(\hat{\beta}_1\right)$	0.16377	0.1385211
$\widehat{SE}\left(\hat{\beta}_2\right)$	0.20360	0.1717913
$\widehat{SE}\left(\hat{\beta}_3\right)$	0.22398	0.2172634
$\widehat{SE}\left(\hat{\beta}_4\right)$	0.25068	0.2463415
$\widehat{SE}\left(\hat{\beta}_5\right)$	0.24049	0.2120539
$\widehat{SE}\left(\hat{\beta}_6\right)$	0.20237	0.1824219
$\widehat{SE}\left(\hat{\beta}_7\right)$	0.30164	0.2999042
$\widehat{SE}\left(\hat{\beta}_8\right)$	0.60214	0.5817717
$\widehat{SE}\left(\hat{\beta}_9\right)$	1.0404	0.9597099

Figure 1. Predicted smoother from R/mgcv (black) and ADMB-RE (red) with 95% confidence intervals (dashed).