Comment

J. L. FRENCH, E. E. KAMMANN, and M. P. WAND

1. INTRODUCTION

Semiparametric nonlinear mixed models are a useful addition to the regression modeling arsenal, and Ke and Wang have done an admirable job in laying out their formulation and implementation. Overall, the article reads well and is fairly accessible to both statistician and interface readers alike. Model formulation such as that given in (6) will be easily digested by those familiar with mixed-effects regression models. However, we believe that there is room for improvement in the formulation of f, at least in the case where f is defined on $\mathbb{R}^d$ with a $\int (D^2 f)^2$ penalty. The authors claim naturalness and uniformity but neglect to use the well-established mixed model representations of f in this case. The implementation described in the article requires the S-PLUS library nlme and the Fortran library RKPACK, when really nlme appears to be all that is required for most applications. In the next section we describe a particularly simple mixed model representation of smoothing splines, along with a useful generalization that permits much more digestible model specification.

2. SIMPLE MIXED MODEL REPRESENTATION OF SMOOTHING SPLINES

Let $(x_i, y_i) \in \mathbb{R}^2$, $1 \le i \le n$, be a set of predictor/response pairs and let $\hat{f}_{m,\lambda}(\cdot)$ be the smoothing spline on $\mathbb{R}$ of order m and with smoothing parameter $\lambda > 0$. There are several ways which $\hat{f}_{m,\lambda}(\cdot)$ can be represented as a mixed model (e.g., Wahba 1978; Speed 1991; Verbyla 1994; Cantoni and Hastie 2000). In this section we describe a particularly simple formulation $\hat{f}_{m,\lambda}(\cdot)$ that requires little more than familiarity with linear mixed models.

For simplicity of presentation, for now, we deal just with $m = 1$ (linear splines). With $\lambda = \sigma_\epsilon^2/\sigma_u^2$, $\hat{f}_{1,\lambda}(\cdot)$ corresponds to best linear unbiased prediction (BLUP) in the linear mixed model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_{\text{p}}\mathbf{u} + \varepsilon, \tag{1}$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & x_i \end{bmatrix}_{1\le i\le n}, \qquad \mathbf{Z}_{\text{p}} = \underset{1\le i,j\le n}{[-|x_i - x_j| + |x_i| + |x_j|]}, \tag{2}$$

and

$$E\begin{bmatrix}\mathbf{u}\\ \varepsilon\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \qquad \text{cov}\begin{bmatrix}\mathbf{u}\\ \varepsilon\end{bmatrix} = \begin{bmatrix}\sigma_u^2 \mathbf{Z}_p^- & \mathbf{0}\\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I}\end{bmatrix}. \tag{3}$$

There are numerous expressions for the BLUP in a general linear mixed model (e.g., Robinson 1991). One such expression involves minimization of

$$\frac{1}{\sigma_\varepsilon^2}\|\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}_{\text{p}}\mathbf{u}\|^2 + \mathbf{u}^{\text{T}}\text{cov}(\mathbf{u})^{-1}\mathbf{u}$$
$$= \frac{1}{\sigma_\varepsilon^2}\{\|\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}_{\text{p}}\mathbf{u}\|^2 + \lambda \mathbf{u}^{\text{T}}\mathbf{Z}_{\text{p}}\mathbf{u}\}.$$

In the following section, we show that the BLUP in (1)–(3) is equivalent to the minimizer of

$$\|\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}_{\text{G}}\mathbf{u}\|^2 + \lambda \mathbf{u}^{\text{T}}\mathbf{Z}_{\text{G}}\mathbf{u}, \tag{4}$$

where

$$\mathbf{Z}_{\text{G}} = \underset{1\le i,j\le n}{[-|x_i - x_j|]}.$$

This is the radial basis function representation of smoothing splines given by, for example, Green and Silverman (1994, pp. 142–143) and Nychka (2000). Note that

$$\text{cov}(\mathbf{Z}_{\text{p}}\mathbf{u}) = \sigma_u^2 \mathbf{Z}_{\text{p}},$$

and that the right side is a proper covariance matrix (hence the subscript "P"). On the other hand, $\mathbf{Z}_{\text{G}}$ is not a proper covariance matrix, although in the geostatistical literature (e.g., Kitanidis 1997) it is sometimes referred to as a *generalized* covariance matrix.

The value of λ may be chosen using GCV or REML. But once this choice has been made the values of $\hat{f}_{m,\lambda}(\cdot)$ depend only on $\mathbf{Z}_{\text{G}}$. The contributions from $\mathbf{Z}_{\text{p}} - \mathbf{Z}_{\text{G}} = [|x_i| + |x_j|]$ vanish. The proof is given in the following section. For general $m = 1, 2, \ldots,$

$$\mathbf{Z}_{\text{G}} = \underset{1\le i,j\le n}{[(-1)^m |x_i - x_j|^{2m-1}]},$$

with $m = 2$ corresponding to cubic smoothing splines. The form of $\mathbf{Z}_{\text{p}}$ for $m > 1$ may be worked out from equation (2.4.25) of Wahba (1990), although because only $\mathbf{Z}_{\text{G}}$ matters for smoothing spline construction, $\mathbf{Z}_{\text{p}}$ is of mainly academic interest. Nychka (2000) provided an interesting perspective on this representation in relation to the geostatistical method known as kriging. There are also connections with the penalized least squares and Green's function perspective described by Solo (2000).

This mixed model representation is also intuitive in the sense that it is immediately apparent that for general $x \in \mathbb{R}$, $\hat{f}_{m,\lambda}(x)$ is of the form

$$\hat{f}_{m,\lambda}(x) = a_0 + a_1 x + \sum_{j=1}^{n} b_j |x - x_j|^{2m-1},$$

J. L. French is Postdoctoral Research Fellow, E. E. Kammann is Graduate Student, and M. P. Wand is Associate Professor, Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115. The authors are grateful to Doug Nychka for pointing out some of the connections used in Sections 2.1 and 2.2, and to Trevor Hastie for sending the details on smooth.spline( ).

which is a degree $2m-1$ spline with knots at the data and linear beyond the data.

Note that we do not advocate using this representation of smoothing splines for actual computation, but rather see it as a useful way of formulating the model that is easily understood by a more general readership. However, in Section 2.2 we discuss some low-rank approximations to smoothing splines, in the spirit of Eilers and Marx (1996) and Hastie (1996), that do have readily computable forms.

2.1 Equivalence of Best Linear Unbiased Prediction Using $\mathbf{Z}_\mathrm{p}$ and $\mathbf{Z}_\mathrm{G}$

The equivalence of the BLUP in (1)–(3) and the minimizer of (4) can be established through examination of the criterion for BLUP given by, for example, McCulloch and Searle (2000). For each $1 \le i \le n$, we seek ℓ for which $\ell^\mathsf{T}\mathbf{y}$ is unbiased for $(\mathbf{X}\beta+\mathbf{Z}_\mathrm{p}\mathbf{u})_i$. This is equivalent to

$$\mathbf{X}^\mathsf{T}\ell = \mathbf{X}^\mathsf{T}\mathbf{e}_i,$$

where $\mathbf{e}_i$ is the $n\times 1$ vector with 1 in the ith position and 0's elsewhere. For $\mathbf{X}$ as given in (2), the unbiasedness condition becomes

$$\ell^\mathsf{T}\mathbf{1}=1 \quad \text{and} \quad \ell^\mathsf{T}\mathbf{x}=x_i,$$

where $\mathbf{x}=[x_1,\dots,x_n]^\mathsf{T}$.

Let m_0 and m_1 be Lagrange multipliers that impose the unbiasedness conditions. For ℓ to minimize the variance of the prediction error $\ell^\mathsf{T}\mathbf{y}-(\mathbf{X}\beta+\mathbf{Z}_\mathrm{p}\mathbf{u})_i$, we need to minimize

$$\begin{aligned}\mathcal{C}(\ell,m_0,m_1) &\equiv \mathrm{var}\{\ell^\mathsf{T}\mathbf{y}-(\mathbf{X}\beta+\mathbf{Z}_\mathrm{p}\mathbf{u})_i\}+m_0(\ell^\mathsf{T}\mathbf{1}-1)\\ &\quad +m_1(\ell^\mathsf{T}\mathbf{x}-x_i)\\ &= \ell^\mathsf{T}(\sigma_u^2\mathbf{Z}_\mathrm{p}+\sigma_\varepsilon^2\mathbf{I})\ell-2\sigma_u^2\ell^\mathsf{T}\mathbf{Z}_\mathrm{p}\mathbf{e}_i+\sigma_u^2(\mathbf{Z}_\mathrm{p})_{ii}\\ &\quad +m_0(\ell^\mathsf{T}\mathbf{1}-1)+m_1(\ell^\mathsf{T}\mathbf{x}-x_i)\\ &= \ell^\mathsf{T}(\sigma_u^2\mathbf{Z}_\mathrm{G}+\sigma_\varepsilon^2\mathbf{I})\ell-2\sigma_u^2\ell^\mathsf{T}\mathbf{Z}_\mathrm{G}\mathbf{e}_i+\sigma_u^2(\mathbf{Z}_\mathrm{G})_{ii}\\ &\quad +m_0(\ell^\mathsf{T}\mathbf{1}-1)+m_1(\ell^\mathsf{T}\mathbf{x}-x_i)\\ &\quad +2\sigma_u^2(\ell^\mathsf{T}\mathbf{1}-1)(\ell^\mathsf{T}|\mathbf{x}|-|x_i|).\end{aligned}$$

It is clear from this that the first unbiasedness constraint will cause the last term to vanish and that minimization over ℓ is unaffected by replacement of $\mathbf{Z}_\mathrm{p}$ by $\mathbf{Z}_\mathrm{G}$. Therefore, for fixed $\lambda=\sigma_\varepsilon^2/\sigma_u^2$, only $\mathbf{Z}_\mathrm{G}$ matters.

Note that similar arguments have appeared in the literature to demonstrate equivalences between smoothing splines and kriging (e.g., Kimeldorf and Wahba 1971; Duchon 1976; Cressie 1990).

2.2 Low-Rank Smoothing Splines

Low-rank, or reduced knot, smoothers have risen to prominence in recent years because of the work of such authors as Eilers and Marx (1996), Hastie (1996), and Rice and Wu (2001). However, the idea of using fewer than n basis functions in spline smoothing had been discussed earlier by others, including Parker and Rice (1985), O'Sullivan (1986, 1988), Gray (1992), and Kelly and Rice (1990). Some mention was made of low-rank spatial smoothing by Nychka, Haaland, O'Connell, and Ellner (1998) and Nychka and Saltzman (1998). Finally, the S function `smooth.spline( )` uses low-rank approximations for $n>50$ (see Sec. 2.3).

This push toward low-rank smoothing is driven by a number of factors. The main factor is the reduction in computational overhead due to large n and having to do several simultaneous smooths, as required for additive models (e.g., Hastie 1996) and curve data models (e.g., Rice and Wu 2001). A major payoff is the ability to use standard mixed model software.

The mixed model representation of smoothing splines can be extended to the low-rank case as follows. The exact smoothing spline is

$$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}_\mathrm{p}\mathbf{u}+\varepsilon,$$

where

$$\mathrm{cov}(\mathbf{u})=\sigma_u^2\mathbf{Z}_\mathrm{p}^{-1}=\sigma_u^2(\mathbf{Z}_\mathrm{p}^{-1/2})(\mathbf{Z}_\mathrm{p}^{-1/2})^\mathsf{T} \tag{5}$$

and $\mathbf{Z}_\mathrm{p}^{-1/2}$ is based on the singular value decomposition of $\mathbf{Z}_\mathrm{p}$ (e.g., Searle 1982, pp. 316–317). The second equality in (5) holds because $\mathbf{Z}_\mathrm{p}$ is symmetric and positive definite. An approximate smoothing spline based on $\mathbf{Z}_\mathrm{G}$ is

$$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}_\mathrm{G}\mathbf{u}+\varepsilon,$$

where

$$\mathrm{cov}(\mathbf{u})=\sigma_u^2(\mathbf{Z}_\mathrm{G}^{-1/2})(\mathbf{Z}_\mathrm{G}^{-1/2})^\mathsf{T}.$$

This involves a linear combination of the radial basis functions

$$r_j(x)=(-1)^m|x-x_j|^{2m-1}, \qquad 1\le j\le n.$$

Let $\kappa_1,\dots,\kappa_K$ be a set of knot locations. Their choice is described in the next section. Then an approximation based on the smaller set of basis functions,

$$r_k^K(x)=(-1)^m|x-\kappa_k|^{2m-1}, \qquad 1\le k\le K,$$

arises from fitting the mixed model

$$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}_K\mathbf{u}+\varepsilon, \qquad \mathrm{cov}(\mathbf{u})=\sigma_u^2(\Omega_K^{-1/2})(\Omega_K^{-1/2})^\mathsf{T},$$

where

$$\mathbf{Z}_K=[(-1)^m|x_i-\kappa_k|^{2m-1}]_{\substack{1\le i\le n\\ 1\le k\le K}}$$

and

$$\Omega_K=[(-1)^m|\kappa_k-\kappa_{k'}|^{2m-1}]_{1\le k,k'\le K},$$

and we note that $\mathbf{u}$, is now a $K\times 1$ random vector. Using the transformation $\mathbf{Z}=\mathbf{Z}_K\Omega_K^{-1/2}$, we can write the final model as

$$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}\mathbf{u}+\varepsilon, \qquad \mathrm{cov}\begin{bmatrix}\mathbf{u}\\ \varepsilon\end{bmatrix}=\begin{bmatrix}\sigma_u^2\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \sigma_\varepsilon^2\mathbf{I}\end{bmatrix}. \tag{6}$$

This form allows fitting through standard mixed model software (see Sec. 2.3).

Note that the generalized covariance function used here,

$$C(r)=(-1)^m|r|^{2m-1},$$

can be replaced by any of the proper covariance functions used in Kriging (e.g., Cressie 1993; O'Connell and Wolfinger 1997; Stein 1999).

2.3 Choice of Knots

The choice of knots for low-rank smoothing splines has already been thought through by the developers of the `smooth.spline( )` function in the S language. The rule used there is essentially

$$\kappa_k = (k/K)\text{th sample quantile of unique } x_i\text{'s}, \quad 1 \le k \le K,$$

and K is chosen according to

$$K = \begin{cases} n, & n \le 50 \\ 100, & n = 200 \\ 140, & n = 800 \\ 200 + (n - 3200)^{1/5}, & n > 3200. \end{cases}$$

Other values of n between 50 and 3200 are handled via a logarithmic interpolation.

Ruppert (2001) and Kammann and Wand (2001) have investigated the choice of K in depth. Their results show that much smaller K than used by `smooth.spline( )` can be adequate. As mentioned earlier, this is important for more complex models.

Figure 1 shows the results of a small simulation study designed to illustrate the effect of K on the performance of low-rank smoothing splines. One hundred replications (x_i, y_i), $1 \le i \le 200$, were generated according to

$$y_i = f(x_i) + .10\varepsilon_i,$$

where the x_i and ε_i are random samples from the uniform distribution on (0, 1) and the standard normal distribution. The mean function f is

$$f(x) = 1.5\phi\left(\frac{x - .35}{.15}\right) - \phi\left(\frac{x - .8}{.04}\right),$$

where ϕ is standard normal density function.

Estimates for f based on (6) with $m = 2$ were obtained for $K = 25, 50, 100$ using S-PLUS function `lme( )` with REML

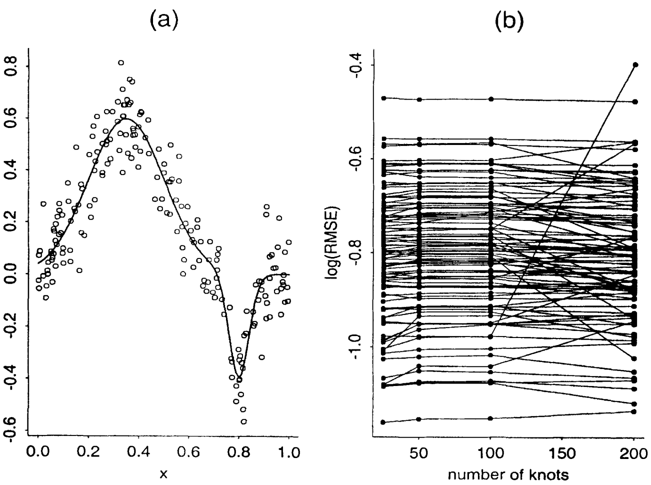

Figure 1. Summary of a Small Simulation Study. (a) Typical realization of data and true mean function $\hat{f}$; (b) $\log_{10}$(RMSE) values for $K = 25$, 50, 100, 200, with lines indicating results obtained with the same sample.

variance component estimation. For data vectors x and y and number of knots K, the $\mathbf{X}$ and $\mathbf{Z}$ matrices were obtained using the S-PLUS commands:

```
knots <- quantile(unique(x),seq(0,1,length=(K+2))[-c(1,(K+2))])
x <- cbind(rep(1,length(x)),x)
svd.Omega <- svd(abs(outer(knots,knots,"-"))^3)
matrix.sqrt.Omega <- t(svd.Omega$v%*%(t(sva.Omega$u)
                         *sqrt(svd.Omega$d)))
Z <- t(solve(matrix.sqrt.Omega),t(abs(outer(x,knots,"-")^3))))
```

Full-rank smoothing spline estimates ($n = K = 200$) were obtained using the S-PLUS function `smooth.spline( )` with `all.knots=T`. Note that these full-rank smooths are not completely comparable with the corresponding low-rank smooths, because the former uses GCV rather than REML.

Figure 1(b) shows $\log_{10}$(RMSE) for each estimate, with lines joining values obtained form the same sample. The RMSE is the square root of the average squared differences between the estimate and the true function, evaluated at the x_i's. We see from this graphic that in most cases, there is little or no improvement in the quality of the estimate as K increases. Paired Wilcoxon signed rank tests support this. Obviously, this is just one simulation setting out of several that could be considered so general recommendations cannot be deduced from it alone. Ruppert (2001) and Kammann and Wand (2001) provide much more evidence.

2.4 Higher-Dimensional Extension

For $\mathbf{x}_i \in \mathbb{R}^d$, $1 \le i \le n$, and $\kappa_k \in \mathbb{R}^d$, $1 \le k \le K$, then higher dimension approximate smoothing splines (also called *thin plate* splines) can be obtained by using the design matrices

$$\mathbf{X} = \begin{bmatrix} 1 & \mathbf{x}_i^{\mathrm{T}} \end{bmatrix} \quad \text{and} \quad \mathbf{Z} = [C(\|\mathbf{x}_i - \kappa_k\|)][C(\|\kappa_k - \kappa_{k'}\|)]^{-1/2},$$

where

$$C(\mathbf{r}) = \begin{cases} \|\mathbf{r}\|^{2m-d}, & d \text{ odd} \\ \|\mathbf{r}\|^{2m-d}\log\|\mathbf{r}\|, & d \text{ even.} \end{cases}$$

The choice of κ_k is somewhat more challenging, although good advice about this is provided by Nychka and Saltzman (1998).

2.5 Inference

Section 5 of Ke and Wang describes how inference concerning f can be done using Gaussian process formulations and Bayesian confidence intervals. However, the estimates of f described here afford the use of mixed model inferential tools.

3. OTHER HILBERT SPACES AND PENALTIES

The previous section dealt exclusively with f defined on $\mathbb{R}^d$ and the $\int(D^2 f)^2$ penalty. Ke and Wang consider f defined on more general Hilbert spaces, and penalties of the form $\int(Lf)^2$ for a linear operator L. Because of time limitations for preparing this comment, we have not been able to explore the connections between mixed model representations and some of these extensions. The pursuit of this would be worthwhile.

4. MAIN POINTS

It is our hope that this comment makes the following main points:

1. For f on $\mathbb{R}^d$ with a $\int(D^2f)^2$ penalty, simple mixed model representations of f exist. These circumvent the need for the complex Hilbert space terminology used in the article.
2. Good low-rank approximations exist that allow for mixed model software to perform the entire fitting algorithm. This circumvents the need to use RKPACK.
3. Inference for f can be performed within the mixed model framework.

ADDITIONAL REFERENCES

Cantoni, E., and Hastie, T. (2000), "Degrees of Freedom Tests for Smoothing Splines," unpublished manuscript.

Cressie, N. (1990), "Reply to Wahba's Letter," *The American Statistician*, 44, 256–258.

———, (1993), *Statistics for Spatial Data*, New York: Wiley.

Duchon, J. (1976), "Fonctions-Spline et Esperances Conditionnelles de Champs Gaussiens," *Ann. Sci. Univ. Clermont Ferrand II Math.*, 14, 19–27.

Green, P. J., and Silverman, B. W. (1994), *Nonparametric Regression and Generalized Linear Models*, London, Chapman and Hall.

Gray, B. (1992), "Flexible Methods for Analyzing Survival Data Using Splines, With Application to Breast Cancer Prognosis," *Journal of the American Statistical Association*, 87, 942–951.

Hastie, T. J. (1996), "Pseudosplines," *Journal of the Royal Statistical Society*, Ser. B, 58, 379–396.

Kammann, E. E., and Wand, M. P. (2001), "Low-Rank Radial Smoothing/Kriging via Restricted Maximum Likelihood," unpublished manuscript.

Kelly, C., and Rice, J. (1990), "Monotone Smoothing With Application to Dose-Response Curves and the Assessment of Synergism," *Biometrics* 46, 1071–1085.

Kimeldorf, G., and Wahba, G. (1971), "Some Results on Tchebycheffian Spline Functions," *Journal of Mathematical Analysis and Its Application*, 33, 82–95.

Kitanidis, P. K. (1997), *Introduction to Geostatistics*, Cambridge, U.K: Cambridge University Press.

McCulloch, C. E., and Searle, S. R. (2000), *Generalized, Linear, and Mixed Models*, New York: Wiley.

Nychka, D. W. (2000), "Spatial Process Estimates as Smoothers" in *Smoothing and Regression*, M. Schimek, ed., Heidelberg: Springer-Verlag.

Nychka, D., Haaland, P., O'Connell, M., and Ellner, S. (1998), "FUNFITS, Data Analysis and Statistical Tools for Estimating Functions," in *Case Studies in Environmental Statistics*, D. Nychka, W. W. Piegorsch, and L. H. Cox, eds., New York: Springer-Verlag.

Nychka, D., and Saltzman, N. (1998), "Design of Air Quality Monitoring Networks," in *Case Studies in Environmental Statistics*, D. Nychka, L. Cox, W. Piegorsch, eds., New York: Springer-Verlag.

O'Connell, M. A., and Wolfinger, R. D. (1997), "Spatial Regression Models, Response Surfaces, and Process Optimization," *Journal of Computational and Graphical Statistics*, 6, 224–241.

O'Sullivan, F. (1986), "A Statistical Perspective on Ill-Posed Inverse Problems" (with discussion), *Statistical Science*, 1, 505–527.

———, (1988), "Fast Computation of Fully Automated Log-Density and Log-Hazard Estimators," *SIAM Journal on Scientific and Statistical Computing*, 9, 363–379.

Parker, R. L., and Rice, J. A. (1985), Discussion of "Some Aspects of the Spline Smoothing Approach to Nonparametric Curve Fitting" by B. W. Silverman, *Journal of the Royal Statistical Society*, Ser. B, 47, 40–42.

Rice, J. A., and Wu, C. O. (2001), "Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curves," *Biometrics*, 57, 253–259.

Robinson, G. K. (1991), "That BLUP is a Good Thing: The Estimation of Random Effects," *Statistical Science*, 6, 15–51.

Ruppert, D. (2001), "Selecting the Number of Knots for Penalized Splines," unpublished manuscript.

Solo, V. (2000), "A Simple Derivation of the Smoothing Spline," *The American Statistician*, 54, 40–43.

Searle, S. R. (1982), *Matrix Algebra Useful for Statistics*, New York: Wiley.

Speed, T. (1991), Comment on "That BLUP is a Good Thing: The Estimation of Random Effects," by G. K. Robinson, *Statistical Science*, 6, 42–44.

Stein, M. L. (1999), *Interpolation of Spatial Data: Some Theory for Kriging*, New York: Springer.

Verbyla A. P. (1994), "Testing Linearity in Generalized Linear Models," *Contributed Pap. 17th Int. Biometric Conf., Hamilton, Aug. 8–12*, 177.

Wahba, G. (1978), "Improper Priors, Spline Smoothing and the Problem of Guarding Against Model Errors in Regression," *Journal of the Royal Statistical Society*, Ser. B, 40, 364–372.

Comment

Xihong LIN and Daowen ZHANG

Wang and Ke are to be congratulated for an interesting article that introduces nonparametric regression to nonlinear mixed models for the analysis of longitudinal data. The proposed double-penalized likelihood method is computationally attractive, because it requires only existing software. The article makes an interesting contribution to the developing literature on nonparametric regression in longitudinal data. The proposed SNMMs present several important theoretical and computational challenges that will stimulate further research in this area.

Different from the existing nonparametric mixed models, a key attractive feature of SNMMs is that Wang and Ke allow the random effects **b** to enter into the nonparametric function $f(\cdot)$. On the other hand, this formulation presents an intrinsic difficulty in fitting this class of models, raises a challenging issue on what the smoothing spline knots are in this class of models, and leads to the failure of many standard results in the smoothing spline and mixed model literature. Our discussion focuses on contrasting these results and highlighting the subsequent challenges. In particular, we show that unlike in standard smoothing spline regression, under SNMMs when the random effects $\mathbf{b}_i$ enter into $f(\cdot)$, (1) the maximum penalized marginal likelihood estimator of $f(\cdot)$ might not exist,

Xihong Lin is Associate Professor, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109 (E-mail: *xlin@sph.umich.edu*). Her research was supported by National Cancer Institute grant CA-76404. Daowen Zhang is Assistant Professor, Department of Statistics, North Carolina State University, Raleigh, NC 27695 (E-mail: *dzhang2@stat.ncsu.edu*). His research was supported by National Institute of Health grant R01-CA85848.

Journal of the American Statistics Assocation
December 2001, Vol. 96, No. 456, Theory and Methods

LINKED CITATIONS

You have printed the following article:

[Semiparametric Nonlinear Mixed-Effects Models and Their Applications]: Comment
J. L. French; E. E. Kammann; M. P. Wand
Journal of the American Statistical Association, Vol. 96, No. 456. (Dec., 2001), pp. 1285-1288.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28200112%2996%3A456%3C1285%3A%5BNMMAT%3E2.0.CO%3B2-7

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

Additional References

Flexible Methods for Analyzing Survival Data Using Splines, With Applications to Breast Cancer Prognosis
Robert J. Gray
Journal of the American Statistical Association, Vol. 87, No. 420. (Dec., 1992), pp. 942-951.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28199212%2987%3A420%3C942%3AFMFASD%3E2.0.CO%3B2-9

Monotone Smoothing with Application to Dose-Response Curves and the Assessment of Synergism
Colleen Kelly; John Rice
Biometrics, Vol. 46, No. 4. (Dec., 1990), pp. 1071-1085.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28199012%2946%3A4%3C1071%3AMSWATD%3E2.0.CO%3B2-B

Spatial Regression Models, Response Surfaces, and Process Optimization
Michael A. O'Connell; Russell D. Wolfinger
Journal of Computational and Graphical Statistics, Vol. 6, No. 2. (Jun., 1997), pp. 224-241.
Stable URL:
http://links.jstor.org/sici?sici=1061-8600%28199706%296%3A2%3C224%3ASRMRSA%3E2.0.CO%3B2-B

LINKED CITATIONS

A Statistical Perspective on Ill-Posed Inverse Problems
Finbarr O'Sullivan
Statistical Science, Vol. 1, No. 4. (Nov., 1986), pp. 502-518.
Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28198611%291%3A4%3C502%3AASPOII%3E2.0.CO%3B2-B

Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curves
John A. Rice; Colin O. Wu
Biometrics, Vol. 57, No. 1. (Mar., 2001), pp. 253-259.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28200103%2957%3A1%3C253%3ANMEMFU%3E2.0.CO%3B2-1

That BLUP is a Good Thing: The Estimation of Random Effects
G. K. Robinson
Statistical Science, Vol. 6, No. 1. (Feb., 1991), pp. 15-32.
Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28199102%296%3A1%3C15%3ATBIAGT%3E2.0.CO%3B2-%23

A Simple Derivation of the Smoothing Spline
V. Solo
The American Statistician, Vol. 54, No. 1. (Feb., 2000), pp. 40-43.
Stable URL:
http://links.jstor.org/sici?sici=0003-1305%28200002%2954%3A1%3C40%3AASDOTS%3E2.0.CO%3B2-0

[That BLUP is a Good Thing: The Estimation of Random Effects]: Comment
Terry Speed
Statistical Science, Vol. 6, No. 1. (Feb., 1991), pp. 42-44.
Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28199102%296%3A1%3C42%3A%5BBIAGT%3E2.0.CO%3B2-N