Statistica Sinica **14**(2004), 449-455

ROOT-N CONSISTENCY OF PENALIZED SPLINE ESTIMATOR FOR PARTIALLY LINEAR SINGLE-INDEX MODELS UNDER GENERAL EUCLIDEAN SPACE

Yan Yu and David Ruppert

University of Cincinnati and Cornell University

Abstract: Single-index models are important in multivariate nonparametric regression. In a previous paper, we proposed a penalized spline approach to a partially linear single-index model where the mean function has the form $\eta_0(\boldsymbol{\alpha}_0{}^T\mathbf{x}) + \boldsymbol{\beta}_0{}^T\mathbf{z}$. This approach is computationally stable and efficient in practice. Furthermore, it yields a root-n consistent estimate of the single-index parameter $\boldsymbol{\alpha}$ and the partially linear parameter $\boldsymbol{\beta}$ with a nontrivial smoothing parameter under the assumption of a compact parameter space. In this paper, we relax the compactness assumption and prove the existence and root-n consistency of the constrained penalized least squares estimators. We expect our proof technique to be useful for establishing asymptotic properties of the penalized spline approach to other model fitting.

Key words and phrases: Asymptotics, compact, inference, nonparametric, ridge regression.

1. Introduction

Consider the partially linear single-index model

$$y_i = \eta_0(\boldsymbol{\alpha}_0^\top \mathbf{x}_i) + \boldsymbol{\beta}_0^\top \mathbf{z}_i + \epsilon_i, \tag{1}$$

where

(i) $y_i \in \mathbf{R}$ is the dependent variable, and $\mathbf{x}_i \in \mathbf{R}^d$, $\mathbf{z}_i \in \mathbf{R}^{d_z}$ are fixed observed predictor vectors;

(ii) the unknown single-index parameter $\boldsymbol{\alpha}_0$ is in $\mathbf{R}^d$, $\|\boldsymbol{\alpha}_0\| = 1$ and the first nonzero element of $\boldsymbol{\alpha}_0$ is positive (for identifiability); the unknown linear parameter $\boldsymbol{\beta}_0$ is in $\mathbf{R}^{d_z}$, and $\eta_0 : \mathbf{R} \to \mathbf{R}$, is an unknown univariate function;

(iii) $\{\epsilon_i\}$ is a mean zero independent error process with variance σ_0^2.

In the recent literature there is voluminous research on fitting single-index models using kernel, local linear, and average derivatives methods. We refer readers to Yu and Ruppert (2002) for a more complete literature review. In addition, Chong (1999) uses sliced inverse regression to obtain a single-index parameter estimate first and then follows a partially linear model with a local polynomial smoother.

While all of these approaches to fitting single-index models have demonstrated promise, there are some potential weaknesses (see Yu and Ruppert (2002) for details). Yu and Ruppert (2002) have proposed a computationally expedient penalized spline approach by modeling $\eta_0(\cdot)$ as a spline. Here $\eta_0(\cdot) = \boldsymbol{\delta}_0^{\mathsf{T}}\mathbf{B}(\cdot)$, with a spline coefficient vector $\boldsymbol{\delta}$ and some spline basis functions $\mathbf{B}(\cdot)$. Then model (1) can be written as

$$y_i = \boldsymbol{\delta}_0^{\mathsf{T}}\mathbf{B}\Big(\boldsymbol{\alpha}_0^{\mathsf{T}}\mathbf{x}_i\Big) + \boldsymbol{\beta}_0^{\mathsf{T}}\mathbf{z}_i + \epsilon_i.$$

The penalized least squares estimator $\boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \boldsymbol{\delta}\end{pmatrix}$ is obtained by minimizing the constrained penalized sum of squared errors

$$n^{-1}\sum_{i=1}^{n}\Big\{y_i - \Big\{\boldsymbol{\delta}^{\mathsf{T}}\mathbf{B}\Big(\boldsymbol{\alpha}^{\mathsf{T}}\mathbf{x}_i\Big) + \boldsymbol{\beta}^{\mathsf{T}}\mathbf{z}_i\Big\}\Big\}^2 + \lambda_n\boldsymbol{\delta}^{\mathsf{T}}\mathbf{D}\boldsymbol{\delta}$$

with constraint $\|\boldsymbol{\alpha}\| = 1$, where $\mathbf{D}$ is an appropriate positive semi-definite symmetric matrix and $\lambda_n \geq 0$ is a penalty parameter.

In Yu and Ruppert (2002), we have shown that this penalized spline approach offers a number of computational advantages. Furthermore, assuming a fixed but potentially large number of knots, we have shown the $\sqrt{n}$-consistency of estimates of the single-index parameter $\boldsymbol{\alpha}_0$, the partially linear coefficient $\boldsymbol{\beta}_0$, and the spline coefficient vector $\boldsymbol{\delta}_0$ with a nontrivial smoothing parameter under the assumption of a *compact* parameter space Θ of $\boldsymbol{\theta}$.

The compactness (Assumption 1 in Yu and Ruppert (2002)) is rather restrictive, because it implicitly requires that there be *known* bounds on the true parameter values of the partially linear coefficient $\boldsymbol{\beta}$ and the spline coefficient $\boldsymbol{\delta}$. Neither of these assumptions is natural and it is desirable to investigate if the boundedness assumptions can be relaxed.

Without a penalty, the penalized spline approach to partially linear single-index models becomes a constrained nonlinear least squares problem with a finite number of parameters to estimate. Past literature on existence and consistency of nonlinear least squares estimates, e.g., Jennrich (1969), Malinvaud (1970) and Wu (1981), all assume a compact parameter space. Malinvaud (1970) gives counterexamples when the parameter space is not compact, while Newey and McFadden (1994) comment that it is useful in practice to be able to drop the compactness restriction.

In this paper, we relax the compactness assumption and obtain the existence and root-n consistency of the constrained penalized least squares estimators. For practical applications, refer to Yu and Ruppert (2002). We handle the constraint

$\|\boldsymbol{\alpha}\| = 1$ by reparametrization. The main ingredient in our proof is that the constraint $\|\boldsymbol{\alpha}\| = 1$ restricts the reparametrized single-index parameter $\boldsymbol{\phi}$ to be in a compact subspace Φ. That is, only the compactness of Φ is needed and the partially linear coefficients $\boldsymbol{\beta}$ and the spline coefficients $\boldsymbol{\delta}$ can be handled separately through linear ridge regression analysis. We expect some of the techniques employed here to be useful for proving asymptotic properties for the penalized spline approach to other model fitting. For instance, an immediate extension is to the large sample theory for the penalized spline approach to multiple-index models and generalized partially linear single-index models (Carroll, Fan, Gijbels and Wand (1997)). In general, this proof technique can be useful when model parameters can be expressed separately as linear regression coefficients and parameters of interest that are constrained in a unit ball due to identifiability.

We organize the remainder of the paper as follows. Section 2 contains the main results of root-n consistency of our constrained penalized least squares estimators, emphasizing differences in the assumptions with those in Yu and Ruppert (2002). The proof is in Section 3.

2. Model and Main Results

2.1. Model and estimators

For the partially linear single-index models (1), the unknown univariate function $\eta_0(\cdot)$ is estimated by penalized splines $\eta_0(\cdot) = \boldsymbol{\delta}_0^{\mathsf{T}}\mathbf{B}(\cdot)$. Define $\mathbf{v}_i = \begin{pmatrix}\mathbf{x}_i \\ \mathbf{z}_i\end{pmatrix}$, the mean function $m(\mathbf{v}_i;\boldsymbol{\theta}) = \boldsymbol{\delta}^{\mathsf{T}}\mathbf{B}\Big(\boldsymbol{\alpha}^{\mathsf{T}}\mathbf{x}_i\Big)+\boldsymbol{\beta}^{\mathsf{T}}\mathbf{z}_i$, and the average sum of squared errors

$$Q_n(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^{n}\Big\{y_i - m(\mathbf{v}_i;\boldsymbol{\theta})\Big\}^2. \tag{2}$$

The penalized least squares estimator of $\boldsymbol{\theta}$ minimizes the penalized sum of squared errors

$$\begin{aligned} Q_{n,\lambda_n}(\boldsymbol{\theta}) :&= n^{-1}\sum_{i=1}^{n}\Big\{y_i - m(\mathbf{v}_i;\boldsymbol{\theta})\Big\}^2 + \lambda_n\boldsymbol{\delta}^{\mathsf{T}}\mathbf{D}\boldsymbol{\delta} \\ &= Q_n(\boldsymbol{\theta}) + \lambda_n\boldsymbol{\delta}^{\mathsf{T}}\mathbf{D}\boldsymbol{\delta}, \end{aligned} \tag{3}$$

where $\lambda_n \geq 0$ is a penalty parameter and $\mathbf{D}$ is some general positive semi-definite symmetric matrix.

The constraint $\|\boldsymbol{\alpha}_0\| = 1$ on the d-dimensional single-index parameter $\boldsymbol{\alpha}$ is handled by reparametrization. That is, let $\boldsymbol{\phi}$ be a $d-1$ dimensional parameter and $\boldsymbol{\alpha}(\boldsymbol{\phi}) = \begin{pmatrix} \sqrt{1-(\phi_1^2+\cdots+\phi_{d-1}^2)} \\ \phi_1 \\ \vdots \\ \phi_{d-1} \end{pmatrix}$. The true parameter $\boldsymbol{\phi}_0$ is assumed to satisfy $\|\boldsymbol{\phi}_0\| < 1$.

2.2. Main results

For the penalized spline approach to partially linear single-index models $y_i = {\boldsymbol{\delta}_0}^{\mathsf{T}}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}_0)\mathbf{x}_i\right) + \boldsymbol{\beta}_0^{\mathsf{T}}\mathbf{z}_i + \epsilon_i$, denote $\boldsymbol{\delta}_z = \begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\beta} \end{pmatrix}$ and $\mathbf{B}_z(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{x}_i) = \begin{pmatrix} \mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{x}_i\right) \\ \mathbf{z}_i \end{pmatrix}$. Then the mean function of our model can be written as

$$m(\mathbf{v}_i;\boldsymbol{\theta}) = {\boldsymbol{\delta}_z}^{\mathsf{T}}\mathbf{B}_z\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{x}_i\right). \tag{4}$$

We can show that our consistency results hold under Assumption 2′ below. Here the compactness of the parameter space Θ of $\boldsymbol{\theta}$ in Assumption 1 of Yu and Ruppert (2002) is dropped. Continuity of the mean function holds in our spline model. Conditions on mean functions in Assumption 2 of Yu and Ruppert (2002) can be explicitly written in terms of basis functions in Assumption 2′, which may be easier to check.

Assumption 2′ $1/n\sum_{i=1}^{n}\mathbf{B}_z\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{x}_i\right)\mathbf{B}_z\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}^*)\mathbf{x}_i\right)^{\mathsf{T}}$ *converges uniformly in* $\boldsymbol{\phi},\boldsymbol{\phi}^* \in \Phi$, *and the limit* $R(\boldsymbol{\phi},\boldsymbol{\phi})$ *is positive definite for all* $\boldsymbol{\phi}\in\Phi$, *where*

$$R(\boldsymbol{\phi},\boldsymbol{\phi}^*) = \lim_n \frac{1}{n}\sum_{i=1}^{n}\mathbf{B}_z\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{x}_i\right)\mathbf{B}_z\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}^*)\mathbf{x}_i\right)^{\mathsf{T}}, \tag{5}$$

$$P(\boldsymbol{\phi}) = R(\boldsymbol{\phi}_0,\boldsymbol{\phi}_0) - R(\boldsymbol{\phi}_0,\boldsymbol{\phi})R^{-1}(\boldsymbol{\phi},\boldsymbol{\phi})R(\boldsymbol{\phi},\boldsymbol{\phi}_0) \tag{6}$$

has a unique zero at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$.

Theorem 1′ *Under Assumption* 2′, *if the smoothing parameter* λ_n *satisfies* $\lambda_n = o(1)$, *then a sequence of penalized least squares estimators* $(\widehat{\boldsymbol{\theta}}_{n,\lambda_n})$ *minimizing* (3) *exists and is a consistent estimator of* $\boldsymbol{\theta}_0$.

Root-n consistency and asymptotic normality as in Theorem 2 in Yu and Ruppert (2002) are also established under the relaxed Assumption 2′ and Assumption 4 using the previously proved consistency. Assumption 3 of Yu and Ruppert (2002) is dropped since $\|\boldsymbol{\phi}_0\| < 1$ and the partially linear parameter $\boldsymbol{\beta}_0$ and spline coefficient vector $\boldsymbol{\delta}_0$ lie in a general Euclidean space. The regularity condition of the mean function in Assumption 4 obviously holds in our spline model. The proof is essentially the same as that in Yu and Ruppert (2002) using standard Taylor expansion techniques and is thus omitted.

Note that in Yu and Ruppert (2002) and this paper, we assume that $\{(\mathbf{x}_i, \mathbf{z}_i)\}$ are given observed constants as in many regression problems (e.g., Wu (1981)). If a random design is desired instead, then the condition that $\{\epsilon_i\}$ is independent of predictor variables $\{(\mathbf{x}_i, \mathbf{z}_i)\}$ should be added in the model (1). Mode of convergence (almost surely convergence) should be specified whenever the limits involve design variables $\{(\mathbf{x}_i, \mathbf{z}_i)\}$ (e.g., equations (5)). Proofs should similarly follow (Jennrich (1969)). Also note that the assumption that $\mathbf{x}$ and $\mathbf{z}$ are independent is not necessary in the penalized spline approach to partially linear single-index models.

3. Proof of Theorem 1′

(1) Proof of existence.

Without loss of generality, we consider the model $y_i = \boldsymbol{\delta}_0{}^{\mathsf{T}}\mathbf{B}(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}_0)\,\mathbf{x}_i)+\epsilon_i$. The same proof is valid for partially linear single-index models when we replace $\mathbf{B}(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i)$ by $\mathbf{B}_z(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i)$ and $\boldsymbol{\delta}$ by $\boldsymbol{\delta}_{\mathbf{Z}}$.

Penalized least squares estimators $\widehat{\boldsymbol{\delta}}_n$ and $\hat{\boldsymbol{\phi}}_n$ minimize

$$Q_{n,\lambda_n}(\boldsymbol{\delta},\boldsymbol{\phi}) = n^{-1}\sum_{i=1}^{n}\Big\{y_i - \boldsymbol{\delta}^{\mathsf{T}}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\Big\}^2 + \lambda_n\boldsymbol{\delta}^{\mathsf{T}}\mathbf{D}\boldsymbol{\delta}. \tag{7}$$

Thus $\widehat{\boldsymbol{\delta}}_n$ solves the first order equations $\partial Q/\partial\boldsymbol{\delta} = 0$, which gives explicit solution to the penalized least squares estimators $\widehat{\boldsymbol{\delta}}_n$ in terms of $\boldsymbol{\phi}$:

$$\widehat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}) = \left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)^{\mathsf{T}} + \lambda_n\mathbf{D}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)y_i. \tag{8}$$

We need to show that penalized least squares estimators $\widehat{\boldsymbol{\delta}}_n$ and $\hat{\boldsymbol{\phi}}_n$ exist. Formally, we need to show that there exists $\hat{\boldsymbol{\phi}}_n$ from $\mathcal{Y}$ to Φ such that, for all y in $\mathcal{Y}$, $Q_n(\hat{\boldsymbol{\phi}}_n) = \inf_{\boldsymbol{\phi}\in\Phi} Q_n(\boldsymbol{\phi})$, where

$$Q_n(\boldsymbol{\phi}) = n^{-1}\sum_{i=1}^{n}\Big\{y_i - \widehat{\boldsymbol{\delta}}_n^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\Big\}^2. \tag{9}$$

Note that the penalty term $\lambda_n\boldsymbol{\delta}^{\mathsf{T}}\mathbf{D}\boldsymbol{\delta}$ in (7) is dropped in (9) since the penalty is on the spline coefficients $\boldsymbol{\delta}$ only. The penalty term is implicit in $\widehat{\boldsymbol{\delta}}_n(\boldsymbol{\phi})$, which is given at (8). Here $\boldsymbol{\phi}\in\Phi$ is compact. The existence of a sequence of least squares estimators $\hat{\boldsymbol{\phi}}_n$ is guaranteed by Lemma 1 in Jennrich (1969). Then $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n)$ given at (8) is the penalized least squares estimates.

(2) Proof of consistency.

To prove the consistency of the penalized least squares estimators $\hat{\boldsymbol{\phi}}_n$ and $\hat{\boldsymbol{\delta}}_n$, we go through two steps: (i) show $\hat{\boldsymbol{\phi}}_n \to \boldsymbol{\phi}_0$ in probability, where $Q_n(\hat{\boldsymbol{\phi}}_n) = \inf_{\boldsymbol{\phi}\in\Phi} Q_n(\boldsymbol{\phi})$ is defined at (9), and Φ is compact; (ii) show $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) \to \boldsymbol{\delta}_0$ in probability, when (i) holds.

Proof. (i) The least squares estimators $\hat{\boldsymbol{\phi}}_n$ minimize $Q_n(\boldsymbol{\phi})$, which can be expanded as

$$\begin{aligned} Q_n(\boldsymbol{\phi}) &= n^{-1}\sum_{i=1}^{n}\Big\{y_i - \widehat{\boldsymbol{\delta}}_n^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\Big\}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}{\epsilon_i}^2 + \frac{2}{n}\sum_{i=1}^{n}\Big\{{\boldsymbol{\delta}_0}^{\mathsf{T}}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}_0)\,\mathbf{x}_i\right) - \widehat{\boldsymbol{\delta}}_n^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\Big\}\epsilon_i \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\Big\{{\boldsymbol{\delta}_0}^{\mathsf{T}}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}_0)\,\mathbf{x}_i\right) - \widehat{\boldsymbol{\delta}}_n^{\mathsf{T}}(\boldsymbol{\phi})\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\Big\}^2 \\ &= C_1 + C_2 + C_3. \end{aligned}$$

The following limits are taken when $n \to \infty$ unless otherwise stated. Similar to the proof of Theorem 1 in Yu and Ruppert (2002), $C_1 \to \sigma_0^2$ for almost every $\epsilon > 0$. We then show that for almost every ϵ, $C_2 \to 0$ uniformly for all $\boldsymbol{\phi}$ in Φ. Lastly, we prove that C_3 converges almost surely to a limit that has a unique zero at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$. The rest of proof is similar to the proof of Theorem 1 of Yu and Ruppert (2002).

Term C_2 can be further expanded as

$$\begin{aligned} C_2 &= 2{\boldsymbol{\delta}_0}^{\mathsf{T}}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi}_0)\,\mathbf{x}_i\right)\epsilon_i\Big\} - 2\widehat{\boldsymbol{\delta}}_n^{\mathsf{T}}(\boldsymbol{\phi})\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\epsilon_i\Big\} \\ &= C_{21} + C_{22}. \end{aligned}$$

We show that under Assumption $2'$, $(1/n)\sum_{i=1}^{n}\mathbf{B}\left(\boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{\phi})\,\mathbf{x}_i\right)\epsilon_i \to 0$ uniformly almost surely over Φ, using a proof similar to Theorem 4 in Jennrich (1969). Note that an important fact used is that $\boldsymbol{\phi}$ is a finite dimensional vector. We refer interested readers to a technical report Yu and Ruppert (2003) for a detailed proof.

Next, using a standard expansion, we have for the smoothing parameter $\lambda_n = o(1)$, $\widehat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}) \to R^{-1}(\boldsymbol{\phi},\boldsymbol{\phi})R(\boldsymbol{\phi},\boldsymbol{\phi}_0)\boldsymbol{\delta}_0$ uniformly almost surely in $\boldsymbol{\phi}$, where $R(\boldsymbol{\phi},\boldsymbol{\phi}^*)$ is defined in (5). Thus, $C_{22} \to 0$ and $C_2 \to 0$ uniformly for all $\boldsymbol{\phi}$ in Φ.

Term C_3 can be further expanded. Again by Assumption $2'$, we have $C_3 \to {\boldsymbol{\delta}_0}^{\mathsf{T}}P(\boldsymbol{\phi})\boldsymbol{\delta}_0$ almost surely, where $P(\boldsymbol{\phi}) = R(\boldsymbol{\phi}_0,\boldsymbol{\phi}_0) - R(\boldsymbol{\phi}_0,\boldsymbol{\phi})R^{-1}(\boldsymbol{\phi},\boldsymbol{\phi})R(\boldsymbol{\phi},\boldsymbol{\phi}_0)$ has a unique zero at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ by assumed positive definiteness of R. The rest of proof is similar to the proof of Theorem 1.

(ii) We now show that $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) \to \boldsymbol{\delta}_0$ in probability, using the consistency results obtained in (i). We can write $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) - \boldsymbol{\delta}_0 = [\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) - \hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}_0)] + [\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}_0) - \boldsymbol{\delta}_0]$. Term $\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}_0) - \boldsymbol{\delta}_0$ reduces to the linear ridge regression case when the single-index parameter $\boldsymbol{\phi}_0$ is known. By Lemma 1 in Yu and Ruppert (2002) of the consistent results for linear ridge regression, $\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}_0) - \boldsymbol{\delta}_0 \to 0$ in probability. $\{\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi})\}$ is equicontinuous on Φ since $\{\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi})\}$ converges uniformly to a continuous function on Φ. By the consistency result of $\hat{\boldsymbol{\phi}}_n$ obtained from (i) that $\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0 \to 0$ in probability and the equicontinuity of $\{\hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi})\}$, we have $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) - \hat{\boldsymbol{\delta}}_n(\boldsymbol{\phi}_0) \to 0$. Consequently, $\hat{\boldsymbol{\delta}}_n(\hat{\boldsymbol{\phi}}_n) \to \boldsymbol{\delta}_0$ in probability.

Acknowledgement

This research was partially supported by NSF Grant DMS-9804058. This manuscript is based in part on the Ph.D. thesis of the first author at Cornell University.

References

Carroll, R. J., Fan, J., Gijbels, I. and Wand, M. P. (1997). Generalized partially linear single-index models. *J. Amer. Statist. Assoc.* **92**, 477-489.

Chong, Y. S. (1999). Dimension reduction methods for discrete and continuous covariates. Unpublished Ph.D. thesis, University of California at Davis.

Jennrich, R. I. (1969). Asymptotic properties of nonlinear least squares estimatiors. *Ann. Math. Statist.* **40**, 633-643.

Malinvaud, E. (1970). The consistency of nonlinear regressions. *Annals of Mathematical Statistics* **41**, 956-969.

Newey, W. K. and McFadden, D. L. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics* **4** (Edited by R. Engle and D. L. McFadden), 2133-2136. Elsevier Science B. V., Amsterdam.

Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* **9**, 501-513.

Yu, Y. and Ruppert, D. (2002). Penalized spline estimation for partially linear single-index models. *J. Amer. Statist. Assoc.* **97**, 1042-1054.

Yu, Y. and Ruppert, D. (2003). Root-n consistency of penalized spline estimator for partially linear single-index models under general Euclidean space. Technical report #1385, Cornell University, downloadable from www.orie.cornell.edu/orie_template.php3?select=TechReps.

Department of Quantitative Analysis and Operations Management, College of Business Administration, University of Cincinnati, PO Box 210130, Cincinnati, Ohio 45221-0130, U.S.A.

E-mail: yuy@statqa.cba.uc.edu

School of Operations Research and Industrial Engineering, 225 Rhodes Hall, Cornell University, Ithaca, NY, 14853, U.S.A.

E-mail: dr24@cornell.edu

(Received October 2002; accepted September 2003)