Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Statistical Planning and Inference 136 (2006) 3007–3031
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Local regression for vector responses

A.H. Welsh[a,*], T.W. Yee[b]

[a]*Centre for Mathematics and its Applications, The Australian National University, Canberra ACT 0200, Australia*
[b]*Department of Statistics, University of Auckland, New Zealand*

Received 5 December 2001; accepted 18 January 2004
Available online 16 March 2005

Abstract

We explore a class of vector smoothers based on local polynomial regression for fitting nonparametric regression models which have a vector response. The asymptotic bias and variance for the class of estimators are derived for two different ways of representing the variance matrices within both a seemingly unrelated regression and a vector measurement error framework. We show that the asymptotic behaviour of the estimators is different in these four cases. In addition, the placement of the kernel weights in weighted least squares estimators is very important in the seeming unrelated regressions problem (to ensure that the estimator is asymptotically unbiased) but not in the vector measurement error model. It is shown that the component estimators are asymptotically uncorrelated in the seemingly unrelated regressions model but asymptotically correlated in the vector measurement error model. These new and interesting results extend our understanding of the problem of smoothing dependent data.

MSC: Primary 62G07; Secondary 62J02

Keywords: Asymptotic bias and variance; Kernel estimator; Local regression; Non-parametric regression; Seemingly unrelated regressions; Vector generalized additive models; Vector smoothing; Weighted multivariate regression

* Corresponding author. Tel.: +61 2 6125 2960; fax: +61 2 6125 5549.
E-mail address: alan.welsh@anu.edu.au (A.H. Welsh).

0378-3758/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2004.01.024

1. Introduction

The purpose of this paper is to explore the properties of local linear estimators of the mean functions in a general nonparametric vector regression model in which the response is a vector of correlated observations. We consider several simple methods of incorporating the correlation structure into the estimators and conclude that, in the homoscedastic case, the estimators we consider show no gain in asymptotic efficiency but that, in the heteroscedastic case, gains in asymptotic efficiency from incorporating the correlation structure can be achieved.

We consider the vector regression model in which observations $y_i^{(j)}$ are related to known, univariate explanatory variables $x_i^{(j)}$, $i=1,\dots,n,\ j=1,\dots,M$, by

$$
\begin{aligned}
y_i^{(1)} &= f_1(x_i^{(1)}) + \varepsilon_i^{(1)} \\
&\vdots \\
y_i^{(M)} &= f_M(x_i^{(M)}) + \varepsilon_i^{(M)},
\end{aligned}
\tag{1}
$$

where $f_1,\dots,f_M$ are unknown regression functions and $\boldsymbol{\varepsilon}_i=(\varepsilon_i^{(1)},\dots,\varepsilon_i^{(M)})^{\mathrm{T}}$ are independent random vectors with $E(\boldsymbol{\varepsilon}_i)=\mathbf{0}$, and $\mathrm{Var}(\boldsymbol{\varepsilon}_i)=\boldsymbol{\Sigma}_i$ or $\mathrm{Var}(\boldsymbol{\varepsilon}_i)=\boldsymbol{\Sigma}(\mathbf{x}_i)$ with $\mathbf{x}_i=(x_i^{(1)},\dots,x_i^{(M)})^{\mathrm{T}}$. We represent $\mathrm{Var}(\boldsymbol{\varepsilon}_i)$ in two different ways because the estimators we consider have different asymptotic properties in the following two cases.

Case A: Following Ruckstuhl et al. (2000) and Lin and Carroll (2000), we treat the elements of $\mathrm{Var}(\boldsymbol{\varepsilon}_i)=\boldsymbol{\Sigma}_i$ as unknown constants, writing the diagonal elements as σ_{ji}^2 and the off-diagonal elements as $\sigma_{ji}\sigma_{ki}\rho_{jki}$. There are $nM(M+1)/2$ nuisance parameters so they cannot all be unknown.

Case B: Following Ruppert and Wand (1992), we model the elements of $\mathrm{Var}(\boldsymbol{\varepsilon}_i)=\boldsymbol{\Sigma}(\mathbf{x}_i)$ as smooth functions of the covariates and write the diagonal elements as $\sigma_j(x^{(j)})^2$ and the off-diagonal elements as $\sigma_j(x^{(j)})\sigma_k(x^{(k)})\rho_{jk}(x^{(j)},x^{(k)})$.

Since the estimators we consider have the same structure in both cases except for notational differences, we will use the Case A notation for both cases when no distinction needs to be drawn and use the separate notations only when the distinction matters. In either case, we are interested in estimating the regression functions $f_1,\dots,f_M$ in the presence of the nuisance variance parameters (Table 1).

Table 1
A table showing the estimator of the regression or derivative function with the smaller asymptotic variance in the different cases

Case		SUR Not all covariates equal	VME Common covariates
A	Variance unrelated to covariates	$\rho_i=\rho$ Weighted $\rho_i\neq\rho$ Either	Weighted
B	Variance a function of covariates	$\rho(x)=\rho$ Equal $\rho(x)\neq\rho$ Either	Equal