IMPACT: Journal of Modern Developments in Natural Sciences Research (IMPACT: JMDNSR)
Vol. 1, Issue 1, Jun 2017,37-46

SPLINE ADDITIVE REGRESSION MODELS

AMMAR MUSLIM ABDULHUSSEIN[1], AMEERAJABER MOHAISEN[2] & FIRAS A. AL-SAADAWI[3]

[1,3] Department of Mathematics, The Open Educational College in Basrah, Iraq

[2] Department of Mathematics, AL-Basrah University, College of Education for Pure Science, Iraq

ABSTRACT

In this paper, we study additive regression models with spline smoothing, and determining the numbers of knots and their locations by using some statistical criteria.

KEYWORDS: Spline, Penalized Spline, Mixed Models, Additive Regression, Knots, Cross-Validation

INTRODUCTION

Regression analysis is a statistical tool that utilizes the relation between two or more quantitative variables so that one variable can be predicted from the other, or others. For example, if one knows the relation between advertising expenditures and sales, one can predict sales by regression analysis once the level of advertising expenditures has been set.

Linear regression is a statistical modeling technique that relates the change in one variable to other variables(see[12]).

A simple linear regression line has an equation of the form $y = \beta_0 + \beta_1 x + \varepsilon$, where x is the explanatory variable and y is the dependent variable. The slop of the line isβ_1, β_0 is the intercept, and ε isan error term

(see[14]).

In many applications in different fields, we need to use one of a collection of models for correlated data structures, for example, multivariate observations clustered data, repeated measurements, longitudinal data and spatially data. Often random effects are used to describe the correlation structure in this type of this data. Mixed models are an extension of regression models that allow for the incorporation of random effects. However, they also turn out to be closely related to smoothing (see [16]).

In this paper we study Additive models with spline smoothing, and we present the definition, properties of the statistical models, estimation method. Also we present the number of knots and their locations.

NONPARAMETRIC REGRESSION

Given data of the form$(x_1 , y_1) , (x_2 , y_2) , \dots , (x_n , y_n)$. Let themodel(see[5]):

$$y = g(x) + \varepsilon \qquad (1)$$

Where the noise term ε satisfies the usual conditions assumed for simple linear regression, we seek an estimate of the regression function $g(x)$ satisfying the model (1). There are several approaches to this problem, we will describe methods involving splines.

SPLINES

The discovery that piecewise polynomials or splines could be used in place of polynomials occurred in the early twentieth century. Splines have since become one of the most popular ways of approximating nonlinear functions. Splines are essentially defined as piecewise. Let k be any real number, then can define a p^{th}degree truncated power function as(see[2,3,4,7,8,9,10]):

$$(x - k)_+^p = (x - k)^p I_{\{x>k\}}(x) \quad (2)$$

As a function of x, this function takes on the value 0 to the left of k, and it takes on the value $(x - k)^p$ to the right of k. The number k is called a knot.

The above truncated power function is a basic example of a spline. It is a member of the set of basis functions for the space of splines.

Let us consider a general p^{th} degree splinewith a single knot at k. Let $P(x)$ denote an arbitrary p^{th} degree polynomial.

$$P(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_p x^p$$

Then:

$$S(x) = P(x) + \beta_{p+1}(x - k)_+^p \quad (3)$$

Takes on the value $P(x)$ for any $x \leq k$, and it takes on the value

$P(x) + \beta_{p+1}(x - k)^p$for any$x > k$

Thus, restricted to each region, the function is a p^{th} degree polynomial. As a whole, this function is a p^{th} degree piecewise polynomial; there are two pieces.

Note that require $p + 2$ coefficients to specify this piecewise polynomial. This is a result of the addition of the truncated power function specified by the knot at k. In general, we may add K truncated power function specified at $k_1, k_2, \ldots, k_K$, each multiplied by different coefficients. Thus would result in $p + K + 1$ degree of freedom.

An important property of splines is their smoothness. Polynomials are very smooth, possessing all derivatives everywhere. Splines possess all derivatives only at points which are not knots. The number of derivatives at a knot depends on the degree of the spline, consider the spline by (3),we can show that$S(x)$ is continuous at k,when $p > 0$ by noting that:

$$S(k) = P(k)$$

And$lim_{x \to k} \beta_{p+1}(x - k)_+^p = 0$

So that $lim_{x \to k} S(x) = P(k)$

Can argue similarly for the first $p - 1$ derivatives

$S^{(j)}(k) = P^{(j)}(k),\ j = 1, 2, \ldots, p\text{-}1$

And

$$lim_{x\to k}\,\beta_{p+1}\;p(\,p-1)\;\ldots\;(p-j+1)\;(x-k)^{p-\,j}\;=\;0$$

So that$lim_{x\to k}\,S^{(j)}(x)\;=\;P^{(j)}(k)$ Thep^{th} derivative behaves differently:

$$S^{(p)}(t)\;=\;p!\,\beta_p$$

And$lim_{x\to k}\,S^{(p)}(x)=\;p!\,\beta_p+\;p!\,\beta_{p+1}$

So usually there is a discontinuity in the p^{th} derivative. Thus p^{th} degree splines are usually said to have no more than $(p-1)$ continuous derivatives.

The discussion below (3) indicates that can represent any piecewise polynomials of degree p in the following way:

$$S(x)\;=\;\beta_0\;+\;\beta_1\,x\;+\;\ldots\;+\;\beta_p x^p\;+\;\beta_{p+1}(\,x-\,k_1)_+^p\;+\;\ldots\;+\;\beta_{p+K}(\,x-\,k_K)_+^p \qquad (4)$$

Any piecewise polynomial can be expressed as a linear combination of truncated power functions and polynomial of degree p

$$S(x)\;=\;\begin{cases}\beta_0+\beta_1\,x\;+\;\ldots\;+\;\beta_p x^p & ,\quad x\;\le\;k_1\\ \beta_0+\beta_1\,x\;+\;\ldots\;+\;\beta_p x^p\;+\;\beta_{p+1}(\,x-\,k_1)^p & ,\quad k_1<x\le\;k_2\\ \beta_0+\beta_1\,x\;+\;\ldots\;+\;\beta_p x^p\;+\;\beta_{p+1}(\,x-\,k_1)^p\;+\;\beta_{p+2}(\,x-\,k_2)^p & ,\quad k_2<x\le\;k_3\\ \vdots & \\ \beta_0+\beta_1\,x\;+\;\ldots\;+\;\beta_p x^p\;+\;\beta_{p+1}(\,x-\,k_1)^p\;+\;\ldots\;+\;\beta_{p+k}(\,x-\,k_K)^p & ,\quad x>k_K\end{cases}$$

In the other words,

$$\{\,1\,,x\,,x^2\,,\ldots\,,x^p\,,(\,x-\,k_1)_+^p\,,(\,x-\,k_2)_+^p\,,\ldots,(\,x-\,k_K)_+^p\,\}$$

Is a basis for the space ofp^{th} degree splines possessing knots at $k_1,k_2\,,\ldots,k_K$. By adding a noise term to (4), we can obtain a splines regression model relating a response

$$Y=S(x)+\varepsilon \qquad (5)$$

To the predictor x.

Penalized Splines

Let us consider the model (1) with linear spline (knots) as(see[1,8,9,10,15]):

$$S(x)\;=\;\beta_0\;+\;\beta_1\,x\;+\;\textstyle\sum_{j=1}^{q}\beta_{1j}\,(\,x-\,k_j)_+$$

Then the ordinary least squares fit can be written:$\hat{Y}\;=\;X\,\hat{\beta}$,

Where$\hat{\beta}$ minimizes $\|Y-X\beta\|^2$, with $\beta=(\,\beta_0\,,\beta_1\,,\beta_{11}\,,\beta_{12}\,,\ldots,\beta_{1q})^T$and with

$$X\;=\begin{bmatrix}1 & x_1 & (\,x_1-\,k_1\,)_+ & \cdots & (\,x_1-\,k_q\,)_+\\ 1 & x_2 & (\,x_2-\,k_1\,)_+ & \cdots & (\,x_2-\,k_q\,)_+\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & (\,x_n-\,k_1\,)_+ & \cdots & (\,x_n-\,k_q\,)_+\end{bmatrix}$$

Unconstrained estimation of $\beta_{11}, \beta_{12}, \ldots, \beta_{1q}$ leads to a wiggly fit. For judicious choice of C, a constraint of the type:

$$\sum_{j=1}^{q} \beta_{ij}^2 < C$$

If we define the $(q+2) \times (q+2)$ matrix.

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 0_{2\times 2} & 0_{2\times q} \\ 0_{q\times 2} & I_{q\times q} \end{bmatrix}$$

The minimization problem can be written as: $Min\ \|Y - X\beta\|^2$

Subject to

$$\beta^T D\ \beta < C$$

It can be shown, using a Lagrange multiplier argument, that this is equivalent to choosing β to minimize:

$$\|Y - X\beta\|^2 + \lambda^2 \beta^T D\beta \quad (6)$$

For some $\lambda \geq 0$. This has the solution

$$\hat{\beta}_\lambda = (X^T X + \lambda^2 D\ )^{-1} X^T Y \quad (7)$$

The term $\lambda^2 \beta^T D\beta$ is called a roughness penalty because it penalizes fits that are too rough, thus yielding a smoother result. The amount of smoothing is controlled by λ(the smoothing parameter).

When the value of the smoothing parameter (λ) is very large then $\beta_{1k} \to 0$ leads to the estimator is polynomials of degree q only, while if the $\lambda = 0$ then leads to no exist roughness penalty.

Number and Position of Knots

If the number of knots too small, then the bias can be large in estimator, and if the number too large it is, preferred, we can use all the observations as knots.

Literature proposes several aapproaches to automatic knot selection. Many of them are based on stepwise regression ideas. Although most of theautomatic knot selection procedures proposed exhibit good performance they are each quite complicated and computationally intensive. In penalized spline the number of knots (K) that usually works well is:

$K = min\left(\frac{1}{4}\ number\ of\ unique\ x_i, 35\right)$, (see[10,15,17]):

As the position of knots determine from the $\left(\frac{k+1}{K+2}\right)^{th}$ sample quantile of the unique x_i for $k = 1, 2, \ldots, K$.

Cross Validation (CV)

Let $\hat{m}(x, \lambda)$ denote the regression estimate at a point x with smoothing parameter λ.One of the most common

measures for the goodness of fit of a regression curve to a scatter plot is the residual sum of squares *(RSS)*:

$$RSS\ (\lambda\ )\ =\ \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

With$\hat{y}_i = \hat{m}(x_i, \lambda)$.However, since $RSS$is minimized at the interpollant($\hat{y}_i = y_i,\ i = 1, 2, \dots, n$),minimization of this criterion will lead to the smooth that is closest to interpolation. For penalized spline this corresponds to a zero smoothing parameter. Cross validation gets around this problem. The cross validationcriterion is(see[5,14]):

$$CV(\lambda)\ = \sum_{i=1}^{n}[y_i - \hat{m}_{-i}\ (x_i, \lambda)]^2 \tag{9}$$

Where$\hat{m}_{-i}$ denotes the regression estimator applied to the data but with (x_i, y_i)deleted. The CV choiceofλ, $\hat{\lambda}_{CV}$is the one that minimizes$CV(\lambda)$over $\lambda \geq 0$.

Generalized Cross Validation (GCV)

Efficient algorithms for computation of $CV(\lambda)$were developed in the mid1980s.Before that time,the difficulties surrounding computation of the cross-validation criterion led to the proposal of a simplified version. This simplified criterion is known a generalized cross-validation.

$$GCV(\lambda) = \frac{n\ \sum_{i=1}^{n}(\{(I - S_\lambda)\ y\ \}_i)^2}{(\ tr\ (I - S_\lambda))^2} = \frac{n\ \sum_{i=1}^{n}\ (y_i - \hat{m}\ (x_i, \lambda)\ )^2}{(\ tr\ (I - S_\lambda))^2} = \frac{n\ RSS\ (\lambda\ )}{(\ tr\ (I - S_\lambda))^2} \tag{10}$$

Where S_λ be the smoother matrix associated with$\hat{m}$and S_λ satisfy $\hat{Y} = S_\lambda Y$(see[5,10,15]).

Mixed Models

Mixed models are an extension of regression models that allow for the incorporation of random effects. A more contemporary application of mixed models is the analysis of longitudinal data, clustered data repeated measurements and spatially correlated data.The general form of a linear mixed model is given as follows(see[15]):

$$Y_i = X_i\beta + \sum_{j=1}^{r} Z_{ij}u_{ij} + \epsilon_i \tag{11}$$

$$u_{ij} \sim N(0, G_j),\ \epsilon_i \sim N(0, R_i)$$

Where the vector Y_i has length m_i, X_i and Z_{ij} are, respectively, a $m_i \times p$ design matrix and a $m_i \times q_i$ design matrix of fixed and random effects. $\beta$is a p-vector of fixed effects and u_{ij} are the q_i-vectors of random effects. The variance matrix G_j is a $q_i \times q_i$ matrix and R_i is a $m_i \times m_i$ matrix.

We assume that the random effects $\{\ u_{ij}\ ;\ i\ =\ 1, \dots, n\ ;\ j\ =\ 1, \dots, r\}$ and the set of error terms $\{\epsilon_1, \dots, \epsilon_n\}$ are independent. In matrix notation,

$$Y = X\beta + Zu + \epsilon \tag{12}$$

Here $Y = (Y_1, \dots ., Y_n)^T$ has length $N\ = \sum_{i=1}^{n} m_i$, $X = (X_1^T, \dots, X_n^T)^T$is a $N \times p$ design matrix of fixed effects, Z is a $N \times q$ block diagonal design matrix of random effects, $q = \sum_{j=1}^{r} q_j$, $u = (u_1^T, \dots, u_r^T)^T$ is a q-vector of random effects, $R = diag(R_1, \dots, R_n)$ is a $N \times N$ matrix and $G = diag(G_1, \dots, G_r)$ is a $q\ \times\ q$ block diagonal matrix.

We nowthat treat estimation ofβ, prediction of u, and estimation of the parameters in G and R, one way to drive an estimate of β is to rewrite (12) as:

$$Y = X\beta + \varepsilon^*, \text{ where} \varepsilon^* = Zu + \varepsilon$$

This is just a linear model with correlated error, since:

$$cov(\varepsilon^*) \equiv V = ZGZ^T + R$$

For given V, the estimator of β is:

$$\tilde{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y \quad (13)$$

And is sometime referred to as generalized linear squares (*GLS*). For Y having a general distribution (13) can be shown to be the best linear unbiased estimator (*BLUE*) for β. Alternatively, if Y is multivariate normal, then the right hand said of (13) is both the maximum likelihood estimator (*MLE*) and the uniformly minimum variance unbiased estimator (*UMVUE*).

The latter is the estimator that has the best(smallest) possible variance of any unbiased estimator regardless of the parameters values [15].

The random effects vector can be predicted via best linear prediction.

$$\tilde{u} = BLP(u) = GZ^T V^{-1}(Y - X\tilde{\beta}) \quad (14)$$

Then the *BLUP* of (β, u) can also be written as:-

$$\begin{bmatrix}\tilde{\beta} \\ \tilde{u}\end{bmatrix} = (C^T R^{-1} C + B)^{-1} C^T R^{-1} Y \quad (15)$$

Where

$$C \equiv [\,X \;\; Z\,] and B = \begin{bmatrix}0 & 0 \\ 0 & G^{-1}\end{bmatrix}$$

The fitted values are then:

$$BLUP(Y) = X\tilde{\beta} + Z\tilde{u} = C(C^T R^{-1} C + B)^{-1} C^T R^{-1} Y = HY \quad (16)$$

Where$H$ called Hat matrix or smoother matrix,

The Log – likelihood of$Y$ under the model $Y \sim N(X\beta, V)$ is:-

$$L(\beta, V) = -\frac{1}{2}\{\, n\, log(2\pi) + log|V| + (Y - X\beta)^T V^{-1}(Y - X\beta)\} \quad (17)$$

By substitution (13) in (17) we obtain the profile log – likelihood for V:

$$L_p(V) = -\frac{1}{2}\{\, log|V| + (Y - X\tilde{\beta})^T V^{-1}\big(Y - X\tilde{\beta}\big) + n\, log(2\pi)\}$$

$$= -\frac{1}{2}\{\, log|V| + Y^T V^{-1}[\, I - X(X^T V^{-1} X\,)^{-1} X^T V^{-1}\}Y] - \frac{n}{2} log(2\pi) \quad (18)$$

Penalized Spline as BLUPs

The penalized spline fitting criterion (6) ,when divided by σ_ε^2 can then be written as(see[15]):

$$\frac{1}{\sigma_\varepsilon^2}\|Y - X\beta - Zu\|^2 + \frac{\lambda^2}{\sigma_\varepsilon^2}\|u\|^2 \quad (19)$$

Notice that this can be made to equal the BLUP criterion by treading the u as a set of random coefficients with:

$cov(u) = \sigma_u^2 I$, where $\sigma_u^2 = \frac{\sigma_\varepsilon^2}{\lambda^2}$

Putting all of this together yields the mixed model representation of the regression spline

$Y = X\beta + Zu + \varepsilon$and$f = X\beta + Zu$

$$cov\begin{bmatrix} u \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \sigma_u^2 I & 0 \\ 0 & \sigma_\varepsilon^2 I \end{bmatrix} \quad (20)$$

Note that the fitted values $\tilde{f}$ can be rewritten as:

$$\tilde{f} = C(C^T C + \lambda^2 D)^{-1} C^T Y \quad (21)$$

Additive Models

Let the Model:-

$$Y_i = \sum_{j=0}^{p} \beta_j X_{ji} + m_l(X_{p+l,i}) + \varepsilon, i = 1, \dots, n \quad (22)$$

We call (22) the additive models it has a parametric component $\sum_{j=0}^{p} \beta_j X_{ji}$ and nonparametric components$m_l(X_{p+1,i})$.

In this paper will get this additive model $Y_i = \sum_{j=0}^{p} \beta_j X_{ji} + m(X_{p+l,i}) + w(X_{p+2,i}) + \varepsilon$

By using penalized spline of degree q to first nonparametric component and s to second nonparametric component, get:

$$y_i = \sum_{j=0}^{p} \beta_j x_{ji} + \sum_{j=1}^{q} \beta_{p+j} x_{p+1,i}^j + \sum_{k=1}^{K_q} \{u_k (x_{p+1,i} - k_k)_+^q + \sum_{j=1}^{s} \beta_{p+q+j} x_{p+1,i}^j + \sum_{k=1}^{K_s} \{u_{K_q+k} (x_{p+1,i} - k_k)_+^s + \epsilon_i \quad (23)$$

Where$k_1, \dots, k_{K_q}$and $k_1, \dots, k_{K_s}$are inner knots $a < k_1 <,,< k_{K_q} < b$ and $a < k_1 <,,< k_{K_s} < b$.

By using a convenient connection between penalized splines and mixed models. Model (23) is rewritten as follows(see[6,8,9,13,15,16,17])

$$Y = X\beta + Zu + \varepsilon \quad (24)$$

Where

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \\ \beta_{p+1} \\ \vdots \\ \beta_{p+q} \\ \beta_{p+q+1} \\ \vdots \\ \beta_{p+q+s} \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ \vdots \\ u_{K_q} \\ u_{K_q+1} \\ \vdots \\ u_{K_s} \end{bmatrix}$$

$$, Z = \begin{bmatrix} (x_{p+1,1}-k_1)_+^q & \cdots & (x_{p+1,1}-k_{K_q})_+^q & (x_{p+2,1}-k_1)_+^s & \cdots & (x_{p+2,1}-k_{K_s})_+^s \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (x_{p+1,n}-k_1)_+^q & \cdots & (x_{p+1,n}-k_{K_q})_+^q & (x_{p+2,n}-k_1)_+^s & \cdots & (x_{p+2,n}-k_{K_s})_+^s \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{p1} & x_{p+1,1} & \dots & x_{p+1,1}^q & x_{p+2,1} & \dots & x_{p+2,1}^s \\ 1 & x_{12} & \dots & x_{p2} & x_{p+1,2} & \dots & x_{p+1,2}^q & x_{p+2,2} & \dots & x_{p+2,2}^s \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{pn} & x_{p+1,n} & \dots & x_{p+1,n}^q & x_{p+2,n} & \dots & x_{p+2,n}^s \end{bmatrix}$$

Assume that u and ε are independent and normally distributed as $u \sim N(0, G)$, $\varepsilon \sim N(0, R)$, where $R = diag(\sigma_{\varepsilon_1}, \dots, \sigma_{\varepsilon_n})$ is a $n \times n$ matrix and $G = diag(\sigma_{u_1}, \dots, \sigma_{u_{K_s}})$

The estimation of the parameters β and u entails minimizing the penalized least squares criterion

$$\| Y - X\beta - Zu \|^2 + \lambda^2 u^T D\, u\,; \tag{25}$$

Where D¸ is penalty matrix. For a given smoothing parameter matrix D, the penalized least squares estimators from (25)are :

$$\begin{pmatrix} \hat{\beta} \\ \hat{u} \end{pmatrix} = \begin{pmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + D \end{pmatrix}^{-1} \begin{pmatrix} X^T \\ Z^T \end{pmatrix} Y \tag{26}$$

And the fitted values are$\hat{Y} = X\hat{\beta} + Z\hat{u} = HY$, where H is the smoothingmatrix given by

$$H = (X \;\; Z) \begin{pmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + D \end{pmatrix}^{-1} \begin{pmatrix} X^T \\ Z^T \end{pmatrix} \tag{27}$$

Where H is smoothing matrix.

CONCLUSIONS

We can representation additive model as mixed model by using a convenient connection between penalized splines and mixed models.

REFERENCES

1. Claeskens, G., krivobokova, T. and Opsomer, J.D.(2009) "Asymptotic properties of penalized spline " Biometrika, 96,529-544.
2. de Boor , C. (1978) " Aproactical Guide to splines " springer, New York.
3. Eubak , R.(1990) " Smoothing splines and nonparametric regression " Marcel Dekker, New york.
4. Fan, J. and Gijbels, I. (1996) " Local polynomial modeling and it's applications " Chapman and Hall, London.
5. Green, P. and silverman, B. (1994) " Nonparametric regression and generalized linear models " Chapman and Hall, London.
6. Hastie, T. and Tibshirani, R. (1990) " Generalized additive models " Chapman and Hall, London.
7. Langs, K. (2010) " Numerical Analysis for statisticians " second edition, springer.

8. Muslim, A. and Muhaisn, A. (2014) " Fuzzy sets and penalized spline in Bayesian semiparametric regression " , LAMBERT Academic Publishing.
9. Muslim, A. and Muhaisn, A. (2014) "Spline semiparametric Regression Models regression ", Al Kufauniversity .
10. Motair, Hafed M. (2011) " A comparison of some nonparametric regression smoothing methods using simulation " M.Sc thesis, university of al-Qadisiyah, college of computer sciences and mathematics.
11. Montgomery, D.,C. and Peck, E.,A. (1982) " Introduction to linear regression analysis "John Wiley & Sons.
12. Natio, k. (2002) " Semiparametric regression with multiplicative adjustment " Communications in statistics, Theory and methods 31 2289-2309.
13. Neter, J. and, Wasserman, W. (1974) " Applied linear statistical models, regression, analysis of variance and experimental designs " Richard. D. IRWIN, INC.
14. Ruppert, D. Wand, M.P. and Carroll, R.J. (2003) " Semiparametric regression ", Cambridge University press.
15. Tarmaratram, K. (2011) " Robust Estimation and model selection in Semiparametric regression models ", proefschriftvoorgedragen tot het behalen van de grad van Doctor.
16. Yuan, A. and DE Gooijer, J. (2007) " Semiparametric regression with kernel error model " Scandinavian Journal of statistics.
17. Wand, M.P. (2009) " Semiparametric regression and graphical models " Aust, N.Z, J. Stat. 51(1), 9-41.