Using Spline Approximation and Local Polynomial Methods to Estimate the Additive Partial Linear Model

Hayder Raaid Talib
University of Sumer, Dhi Qar Al-Rifai, Iraq
h.raid @ uos.edu.iq

Dr. Munaf Yousif Hmood
University of Baghdad, Baghdad, Iraq
munaf.yousif@coadec.uobaghdad.edu.iq

Article Info
Page Number: 5046 – 5059
Publication Issue:
Vol 71 No. 4 (2022)

Article History
***Article Received**: 25 March 2022*
***Revised**: 30 April 2022*
***Accepted**: 15 June 2022*
***Publication**: 19 August 2022*

Abstract
The Additive partial linear model (APLM) are very useful in multivariate nonparametric regression. This model is used to study various phenomena, including financial, agricultural, economic, medical, and others.It is characterized by flexibility in dealing with it in terms of the number, type and nature of the studied variables in terms of whether they are linear or non-linear variables. It is used to estimate the effects of some linear and nonlinear explanatory variables on the response variable, and two different methods were used to estimate the model, which are the B-Spline approximation method and the Local polynomial estimators method, using the backfitting algorithm. The comparison between the two methods was done by using comparison criteria represented by mean squares error squares. Through the results of the analysis of simulated experiments, it was noted that the Spline approximate method is more efficient than the method of local polynomial estimators.

Key words: APLM, Spline Approximation, Local Polynomial.

1- Additive Partial Linear Model (APLM)

Semi-parametric regression can play a major role in solving complex scientific problems. This Kind of models can be used to reduce complex data sets. If it is applied correctly while retaining the basic features of the data, it will help in making the right decisions. [3][6] [11]

The Additive Partial Linear Model (APLM) is one of the semi-parametric models, that received Special attention recently because it combines both the nonparametric and parametric components. that is, the response variable depends on some explanatory variables in a linear and another non-linear manner. Hence, this model can put clearer explanations and is preferred over nonparametric, parametric models or non-parametric additive models, so it can be considered as a special case of Additive Generalized non-parametric regression Models. [2][4]

This model is More flexible Compared with linear model and more Efficient compared with non-parametric regression model because it reduces a problem known as the multidimensionality. [2]

The APLM model can be written as:

$$Y = X^T\beta + \sum_{k=1}^{K} g_k(Z_k) + \varepsilon \qquad . \ . \ . \qquad (1)$$

where $\{(X_i, Z_i, Y_i)\}_{i=1}^{n}$ denote to an independent and identical distributed random samples of size n, where

$X = (X_1, X_{2,\dots}, X_d)^T$ denote to the linear variables.

$z = (Z_1, Z_{2,\dots}, Z_k)^T$ denote to the nonlinear variables.

$(g_1, g_{2,\dots}, g_k)$ refers to unknown smoothing function.

$B = (B_1, B_2, \dots, B_d)^T$ indicates a vector of un known parameters.

ε is random error of the model and is independent of each (x, z).

with the mean equal zero and variance equal to σ^2 and $E\big((\varepsilon|X,Z)\big) = 0$

To ensure that the functions non-parametric are defined, we assume that:

$$E[g_k(Z_k)] = 0 \qquad \text{for} \quad K = 1,2,\dots,k$$

APLM is characterized by the simplicity of the General Linear Model (GLM) and the flexibility of the General Additional Model (GAM).

It combines non-parametric and parametric components. [6]

In many cases, the non-parametric components are often the troublesome Parameters and the Main Interest is the linear Part, therefore, our aim is to select the variables in the parametric part. [4]

The natural extension of the APLM model is the Generalized additive partial Linear model (GAPLM).

$$E(Y_i \mid X_i, Z_i) = g\left\{X_i^T\beta + \sum_{k=1}^{K} g_k(Z_{k,i})\right\} \qquad . \ . \ . \qquad (2)$$

where $g(.)$ is a un known link function.

We note that the model (1) is a special case of the model GAPLM in (2) , so when (K = 1) we get a model (1) which called Partial Linear Model: [2][6] [7][9]

$$Y = X^{T}\beta + g(Z) + \varepsilon \qquad . \quad . \quad . \qquad (3)$$

Rong Liu and Lijian Yang (2010) Spline- Backfitted kernel preparation of an additive coefficient model was studied. The additive coefficients model is a flexible regression tool to overcome the Curse dimensionality problem. It was suggested that the spline-back fitted kernel (SBk) and spline-back fitted local linear (SBLL) estimators for functions, components in the model, additive coefficients, It has these estimates: **(i)**It is computationally suitable and usable for high-dimensional data analysis.**(ii)**Theoretically, a complex function can be inferred with confidence. (**iii)**It is easy to use and has high accuracy results.

Li Wang, et al. (2011) This study aims to build generalized additive partial linear models. They suggest using polynomial spline smoothing to estimate nonparametric functions and derive quasi-likelihood for linear parameters. As well as the study of the approximate normality of estimators of parameter components. All large systems of equations are handled as in kernel-based procedures; Hence, it results in gains in computational simplicity. The procedures of the selection were studied with the linear explanatory variables through the use of non-concave penal quasi-likelihood. It has approximate oracle properties. Xin Yi Li et al. (2019) In this research, the additive partial linear regression model (APLM) was studied to analyze high-dimensional data. In other words, the number of linear and nonlinear components is larger than the sample size. The method of estimating, selecting, and inference for the compounds in the APLM model has been suggested at the same time.

2- Estimation Methods for the Additive Partial Linear Model

The Additive partial Linear Model (APLM) consists of two components, the parametric regression function, and the additive nonparametric regression function. when the model (APLM) is to be estimated, it has been estimated or selected the simple parametric part, multiple linear, or non-linear by one of the parametric methods, and the nonparametric part could be estimated according to some of the following methods:

1-2 Spline Approximation

$$g_k = g_{01}(Z_1) + \cdots + g_{0K}(Z_K) \quad . \quad . \quad . \quad (4)$$

Where it is denoted to an additive function assuming that the explanatory variables Z_k are distributed over closed intervals:

$[b_L, C_L], L = 1, 2, \ldots, l$ such that $(b_L < C_L)$ and without loss of generality.

$$[b_L, C_L] = [0, 1] \ ; \text{all } L = 1, 2, \ldots, l \quad . \quad . \quad . \quad (5)$$

With some assumptions, (g_{0k}) can be approximated by spline functions. assuming that (Sn) refers to the space of the Polynomial Splines defined over the period $[0, 1]$ of degree $\rho \geq 1$, assuming there are J_n interior knots, The number of interior knots increases when the size of the sample increases n, as the number of internal knots achieves the following inequality:

$\left[n^{1/(2p)} \ll J_n \ll n^{1/3}\right]$ that is, in the case of P=2 the number of internal knots is equal to $\left[J_n \sim n^{1/4} \log n\right]$ and S_n consists of functions ξ achieves the following conditions: [5][6]

[1] ξ is polynomial of degree ρ which is defined over All of the sub periods.

$$k_j = [I_j, I_{j+1}), \quad k_{J_n} = [I_{J_n}, 1], \quad j = 0, \ldots, j_n - 1$$

[2] In the case of $\rho \geq 2$, ξ is differentiable to the degree of $(\rho - 1)$ and is continuously defined over the interval $[0, 1]$.

The knots of equal distance will be adopted, with the possibility of using knots of unequal distances. the distance between adjacent knots of equal distances can be calculated by:

$$h = \frac{1}{(J_n + 1)}$$

To estimate the additive spline $\hat{g}$ of g_0 it is based on the independent random sample (X_i, Y_i, Z_i) all $i = 1, \ldots, n$ and assuming that g_n is a set of function g of an additive where:

$$g(z) = g_1(z_1) + g_2(z_2) + \ . \ . \ . + g_K(z_K)$$

as each function $(g_K \in Sn)$ and $\{\sum_{i=1}^{n} g_K(Z_{iK}) = 0\}$ the goal of estimating the model is to find each of the function $(g \in G_n)$ as well as the values of β by minimizing the function of sum squares of error [4] [6]

$$L(g,\beta) = \frac{1}{2}\sum_{i=1}^{n}\left[Y_i - \{g(Z_i) + X_i^T\beta\}\right]^2 \quad g_K \in G_n \qquad . . . (6)$$

for the explanatory variable Z_K is assumed $b_{j,k}(Z_K)$ represents the base functions of B-spline of degree ρ to any $(g \in G_n)$ it is possible to write:

$$\begin{aligned} g(Z) &= \gamma^T b_{(Z)} \\ b(z) &= \{b_{j,k}(z_K), \quad j = -\rho, \dots, J_n \ ; K = 1, \dots, k\}^T \end{aligned} \qquad . . . (7)$$

also, the vector γ represents the spline coefficients

$$\gamma = \{\gamma_{j,k} \quad ; K = 1, \dots, k \quad ; j = -\rho, \dots, J_n\}^T$$

the problem of minimizing equation No. (6) is equivalent to finding the values of γ and β by reducing or minimizing the following expression:

$$L(\gamma,\beta) = \frac{1}{2}\sum_{i=1}^{n}\left[Y_i - \{\gamma^T b(Z_i) + X_i^T\beta\}\right]^2 \qquad . . . (8)$$

the estimators resulting from minimizing the above equation are indicated $\hat{\Gamma} = \{\hat{\gamma}_{j,k}, j = -\rho, \dots, J_n, k = 1, \dots, K\}^T$ and that the spline estimator of the function g_0 would be $\left(\hat{g} = \hat{\gamma}^T b(z)\right)$ where the component's central Spline estimator (g_K):

$$\hat{g}_{(K)}(Z_K) = \sum_{j=-\rho}^{J_n} \hat{\gamma}_{j,k}\, b_{j,k}(Z_K) - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=-\rho}^{J_n} \hat{\gamma}_{j,k}\, b_{j,k}(Z_{iK}) \qquad . . . (9)$$

for any measurable function (ϕ_2, ϕ_1) over the interval $[0,1]^k$, Empirical inner product and the corresponding norm are defined as follows:

$$\left.\begin{aligned} \langle\phi_1,\phi_2\rangle_n &= n^{-1}\sum_{i=1}^{n}\{\phi_1(\mathbf{Z}_i)\phi_2(\mathbf{Z}_i)\} \\ \|\phi\|_n^2 &= n^{-1}\sum_{i=1}^{n}\phi^2(\mathbf{Z}_i) \end{aligned}\right\} \qquad . . . (10)$$

If (ϕ_2, ϕ_1) are integrable L^2, then the theoretical inner product and corresponding norm can be defined as follows:

$$\left.\begin{aligned} &\langle \phi_1, \phi_2 \rangle = E\{\phi_1(z)\phi_2(z)\} \\ \Longrightarrow\ &\langle \phi_1, \phi_2 \rangle = \int_{[0,1]^k} \phi_2(z)\phi_1(z) f_{(Z)} d_z \\ &\| \phi \|_2^2 = E\,\phi^2(Z_i) \end{aligned}\right\} \quad . . . \quad (11)$$

as $\left(\| \phi \|_{nk}^2 ; \| \phi \|_{2k}^2\right)$ they are the empirical and theoretical criteria for a univariate function (ϕ) over the period $(0,1)$ and they can defined as follows:[5][10]

$$\left.\begin{aligned} \| \phi \|_{nk}^2 &= n^{-1} \sum_{i=1}^{n} \phi^2(Z_{ik}) \\ \| \phi \|_{2k}^2 &= E\,\phi^2(Z_k) \qquad\qquad 0 < Z < 1 \\ &= \int_{\text{all } Z} \phi^2(z_k) f_k(z_k) d_{Z_k} \end{aligned}\right\} \quad . . . \quad (12)$$

Since (f_k) is a density function of $(Z_k, k = 1, \dots, K)$ and by defining basis of the central spline as follows:

$$b_{j,k}^*(z_k) = b_{j,k}(z_k) - \frac{\|b_{j,k}\|_{2k}}{\|b_{j-1,k}\|_{2k}} b_{j-1,k}(z_k) \quad . . . \quad (13)$$

$\forall\, k = 1, \dots, K\ ;\ j = -\rho + 1, \dots, J_n$

and the standard form of $B_{j,k}(Z_k)$; all $K = 1, \dots, k$

$$B_{j,k}(Z_k) = \frac{b_{j,k}^*(Z_k)}{\|b_{j,k}^*\|_{2k}} \qquad \text{all } j = -\rho + 1, \dots, J_n \quad . . . \quad (14)$$

It was noted that finding (γ, β) to minimize equation (6) is matheamatically Equivalent to the finding of (γ, β) to minimize the following expression:

$$\frac{1}{2} \sum_{i=1}^{n} \left[Y_i - \{\gamma^T B(Z_i) + X_i^T \beta\} \right]^2 \quad . . . \quad (15)$$

$$B(z) = \{B_{j,k}(Zk), \quad j = -\rho + 1, \dots, J_n, K = 1, \dots, k\}^T$$

The spline estimator of the function $(g_{(O)})$ will be $\left[\hat{g}_{(K)} = \hat{\gamma}^T B(z)\right]$ and the estimator of the central spline for each component will be:

$$\hat{g}_{(K)}(Z_K) = \sum_{j=-\rho+1}^{J_n} \hat{\gamma}_{j,k}\, B_{j,k}(Z_K) - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=-\rho+1}^{J_n} \hat{\gamma}_{j,k} B_{j,k}(Z_{iK}) \quad . \; . \; . (16)$$

Let $T = (X, Z)$

$m_o(T) = g_0(Z) + X^T\beta_0$

$\Gamma_{(Z)} = E(X|Z = z)$ and $\tilde{X} = X - \Gamma_Z$

under the conditions below:

[1] Each function achieves $[\, g_{0k} \in H \,]$ all $K = 1, \ldots, k$

where H is the set a g functions in $[0,1]$.

[2] The Z distribution is continuous and the probability density function f is always determined fram zero over $(0, 1)^K$.

[3] Random Variable vectar X achieve and vector $(W \in R^d)$

$\{ C\, \|w\|^2 \leq w^T E(XX^T|Z = z)\, w \leq C\, \|w\|^2 \}$

[4] The Number of knots entering J_n achieves the following inequality $[n^{1/(2p)} << J_n << n^{1/3}]$ in the case of P=2, the number of internal knots is equal to $[J_n \sim n^{1/4} \log n]$

The estimations $\hat{\beta}$ of β are consistent with a root-n. It has an approximately normal distribution. the although rate of convergence of the estimations of the non-parametric compounds g_0 is Slower than the root-n, meaning that: [1][5]

$$\sqrt{n}\,(\hat{\beta} - \beta) \xrightarrow{D} N(0, R^{-1}\Sigma R^{-1}) \qquad . \; . \; . \quad (17)$$

$R = E[X^{\otimes 2}]\; ; \; \Sigma = E[\varepsilon^2 X^{\otimes 2}]$

In addition to that if ε and (X, Z) independent:

$$\sqrt{n}\,(\beta - \hat{\beta}) \xrightarrow{D} N(0, \sigma^2 R^{-1}) \qquad . \; . \; . \quad (18)$$

$\sigma^2 = E(\varepsilon^2)$

2-2-<u>Local Polynomial estimation</u>

Local Polynomial regression is considered as a good smoothing method with high efficiency compared to other various smoothers, and assuming that (k = 2) in the model (1), i.e. that:

$$Y = X^T\beta + \sum_{k=1}^{2} g_k(Z_k) + \varepsilon$$

$$Y = X^T\beta + g_1(Z_1) + g_2(Z_2) + \varepsilon \qquad ... \qquad (19)$$

The additive functions can be written as follows:

$$g_1 = \{g_1(Z_{11}), g_1(Z_{21}) \ldots g_1(Z_{n1})\}^T$$

$$g_2 = \{g_2(Z_{12}), g_1(Z_{22}) \ldots g_1(Z_{n2})\}^T$$

so, the model can be rewritten:

$$Y = X^T\beta + g_1 + g_2 + \varepsilon \qquad ... \qquad (20)$$

the backfitting algorithm is used for the model (20) assuming that $\left[s^T_{1,Z_1}, s^T_{2,Z_2}\right]$ is the equivalence of kernel functions for the local linear regression at (Z_1, Z_2) then: [7][9]

$$s^T_{1,Z_1} = e_1^T \left(Z_1^T \, \Omega_1 \, Z_1\right)^{-1} Z_1^T \, \Omega_1 \qquad ... \qquad (21)$$

$$s^T_{2,Z_2} = e_1^T \left(Z_2^T \, \Omega_2 \, Z_2\right)^{-1} Z_2^T \, \Omega_2 \qquad ... \qquad (22)$$

$$e_1 = (1,0)^T$$

$$\Omega_1 = \mathrm{Diag}\left\{\frac{1}{h_1} K\left(\frac{Z_{11} - Z_1}{h_1}\right), \ldots \frac{1}{h_1} K\left(\frac{Z_{n1} - Z_1}{h_1}\right)\right\}$$

$$\Omega_2 = \mathrm{Diag}\left\{\frac{1}{h_1} K\left(\frac{Z_{12} - Z_2}{h_1}\right), \ldots \frac{1}{h_2} K\left(\frac{Z_{n2} - Z_2}{h_2}\right)\right\}$$

where $K(.)$ Kernel function

h_2, h_1 band widths and Z_2, Z_1 they are design matrices with dimensions $(n \times 2)$ and they are defined as follows:

$$Z_1 = \begin{bmatrix} 1 & Z_{11} - Z_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & Z_{n1} - Z_1 \end{bmatrix} \quad ; \quad Z_2 = \begin{bmatrix} 1 & Z_{12} - Z_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & Z_{n2} - Z_2 \end{bmatrix}$$

(S_1, S_2) they are smooth matrices that represent or equivalence to the kernels of the observations

$(Z_{11},\dots Z_{n1})^T$ and $(Z_{12},\dots Z_{n2})^T$ ie:

$$S_1 = \begin{bmatrix} S_1 Z_{11} \\ \vdots \\ S_1 Z_{n1} \end{bmatrix} \quad ; \quad S_2 = \begin{bmatrix} S_2 Z_{12} \\ \vdots \\ S_2 Z_{n2} \end{bmatrix}$$

$\{S_1^C = (I - 11^T/n)S_1\}$ denotes the central smoothing matrix of S_1.

$\{S_2^C = (I - 11^T/n)S_2\}$ denotes the central smoothing matrix of S_2

(1) vector with dimension n whose elements are units.

using the backfiting algorithm of an APLM "Additive Partial Linear model" to estimate both parametric and nonparametric components:[8]

$$\hat{\beta} = (X^T X)^{-1} X^T \left(Y - \sum_{k=1}^{2} \hat{g}_k \right) \quad . \; . \; . \quad (23)$$

$$\left.\begin{aligned} \hat{g}_1^{(m)} &= S_1^C \left(Y - X\hat{\beta} - g_2^{(m-1)} \right) \\ \hat{g}_2^{(m)} &= S_2^C \left(Y - X\hat{\beta} - g_1^{(m-1)} \right) \end{aligned}\right\} \quad . \; . \; . \quad (24)$$

$\left(g_1^{(m)}\right)$ and $\left(g_2^{(m)}\right)$ They represent the estimators in the M^{th} stage of the backfitting, as a result, non-iterative estimators of β are of the form:

$$\hat{\beta}_{or} = \{X^T(I - S_{12})X\}^{-1} X^T(I - S_{12})Y \quad . \; . \; . \quad (25)$$

$$\begin{aligned} S_{12} = &\left\{I - \left(I - S_1^C S_2^C\right)^{-1} \left(I - S_1^C\right)\right\} \\ &+ \left\{I - \left(I - S_2^C S_1^C\right)^{-1} \left(I - S_2^C\right)\right\} \end{aligned} \quad . \; . \; . \quad (26)$$

To ensure that $\hat{\beta}_{or}$ is a consistent estimate of root-n within the necessary smoothing by removing the constraint using a profile likelihood, the basic idea Can be deescribed as follows:

Let $\left(\hat{g}_1(\beta, Z_1), \hat{g}_2(\beta, Z_2)\right)$ backfitting estimators of $g_1(Z_1), g_2(Z_2)$ on respectively as in formula (24) Except for Replacing $\hat{\beta}$ by β and $(\hat{g}_1\ ,\ \hat{g}_2)$ can be expressed as:

$$\left.\begin{aligned} \hat{g}_1(\beta) &= \left\{I - \left(I - S_1^C S_2^C\right)^{-1} \left(I - S_1^C\right)\right\}(Y - X\beta) \\ \hat{g}_2(\beta) &= \left\{I - \left(I - S_2^C S_1^C\right)^{-1} \left(I - S_2^C\right)\right\}(Y - X\beta) \end{aligned}\right\} \dots (27)$$

Substituting $\left(\hat{g}_1(\beta), \hat{g}_2(\beta)\right)$ in the model (24) and Using the Least-Squares principle, we get the estimators based on the form of profile β for:

$$\hat{\beta}_{Pb} = \{X^T(I - S_{12})^{*2} X\}^{-1} X^T(I - S_{12})^{*2} Y \quad ... \quad (28)$$

as discussed by "Tibshirani & Hastic (1990) and Ruppert & Opsomer (1999)" the centering of all of $[S_1^C, S_2^C]$ is necessary to ensure the convergence of the algorithm and the estimator $\hat{\beta}_{Pb}$ and it is defined well by the assumption that: [6][8]

$$\sum_{i=1}^{n} g_1(Z_{i1}) = \sum_{i=1}^{n} g_2(Z_{i2}) = 0 \quad . \quad . \quad . \quad (29)$$

the difference between $\hat{\beta}_{or}$ and $\hat{\beta}_{Pb}$ is that there is a factor $(I - S_{12})$ that appears in $\hat{\beta}_{Pb}$ and this tearm avoids an under smoothing restriction for APLM "Additive Partial Linear Model".

the bias of the estimator $\hat{\beta}_{Pb}$ can be found as

$[X^T(I - S_{12})^{*2} \ (g_1 + g_2)]$ each element has the order

$[h_1^4 + h_2^4 + O(h_1^4) + O(h_2^4)]$.

usually, the optimum bandwidth is $(n^{-1/5})$ this means that the estimator $\hat{\beta}_{Pb}$ is consistent for $\sqrt{n}$.

3- Mean Average Squared errors (MASE)

Many criteria measure the amount of efficiency in estimating the regression function that has been traded theoretically for the additive partial linear model (APLM), One of these criteria is MASE where

$$MASE = \frac{1}{n} E \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \quad . \quad . \quad . \quad (30)$$

4- <u>Simulation part</u>

In this section, the performance of the finite sample of the procedure proposed by. is investigated by Montecarlo Simulations. The methods used in Part 2, are compared, as they use the quadratic approximation of the nonparametric functions. using the model(1) where k=2

$$Y = X^T\beta + \sum_{k=1}^{2} g_k(Z_k) + \varepsilon \quad . \quad . \quad . \quad (31)$$

Let $\begin{cases} g_1(Z_1) = 8 \cos(7\pi Z_1) \\ g_2(Z_2) = 30 \{e^{-4.75 Z_2} - 2 e^{-8.1 Z_2} + 5 e^{-3.25 Z_2}\} \end{cases}$

As the experiment is repeated 1000 with different sample sizes (n= 20 ,50, 100 and 150) and assuming there are 60 explanatory linear variables and two nonlinear variables.

let $\beta = (2,1.2,0.8,-0.9,1.7,3,5.1,-6.1,4.2,-2.1,0\ ,0,\ldots,0,1.3,-1.7,1.9,3.2,-3.1)$ and $(\sigma = 1\ ,2\ ,6)$. the variables X and ε are Normal distribution. X and ε are independents. the relationship between (X_j) and (X_i) is $[\rho^{|i-j|}]$

with a $\rho = (0.3,\ 0.5 \text{ and } 0.8)$ (Z_1) and (Z_2) are independents and uniformly distributed on [0, 1]. and B-splines Cubic used to approximate the non-parametric functions and the number of knots in the approximation for each non-parametric component ranges from 2 to 10. and based on a program written in R language, and the results were as follows:

Table 1: represents the results of Simulation the values of Mean Average Squared errors (MASE) with repeated of 1000, sample sizes (n= 20, 50, 100, 150) and correlation coefficient $(\rho = 0.3)$

n	**Methods**	***σ***		
		1	**2**	**6**
20	**Spline**	7.3367	8.5122	13.3531
	Local Polynomial	9.0995	11.2172	14.0014
50	**Spline**	6.0333	6.9999	10.9808
	Local Polynomial	7.4829	9.2244	11.5139
100	**Spline**	4.7924	5.5602	8.7224
	Local Polynomial	5.9439	7.3272	9.1458
150	**Spline**	3.7303	4.3280	6.7894
	Local Polynomial	4.6266	5.7034	7.1190

Table 2: represents the results of Simulation the values of Mean Average Squared errors (MASE) with repeated of 1000, sample sizes (n= 20, 50, 100, 150) and correlation coefficient $(\rho = 0.5)$

n	**Methods**	***σ***		
		1	**2**	**6**
20	**Spline**	7.9190	9.1878	14.4129
	Local Polynomial	9.8217	12.1075	15.1126

50	**Spline**	6.5121	7.5555	11.8523
	Local Polynomial	8.0768	9.9564	12.4277
100	**Spline**	5.1727	6.0015	9.4146
	Local Polynomial	6.4156	7.9087	9.8717
150	**Spline**	4.0264	4.6715	7.3282
	Local Polynomial	4.9938	6.1560	7.6840

Table 3: represents the results of Simulation the values of Mean Average Squared errors (MASE) with repeated of 1000, sample sizes (n= 20, 50, 100, 150) and correlation coefficient $(\rho = 0.8)$

n	**Methods**	σ		
		1	**2**	**6**
20	**Spline**	8.2683	9.5931	15.0487
	Local Polynomial	10.2550	12.6416	15.7794
50	**Spline**	6.7994	7.8888	12.3752
	Local Polynomial	8.4331	10.3957	12.9760
100	**Spline**	5.4009	6.2663	9.8300
	Local Polynomial	6.6986	8.2576	10.3072
150	**Spline**	4.2040	4.8776	7.6515
	Local Polynomial	5.2141	6.4276	8.0230

5-Conclusions

1. Through the results of the analysis shown in Tables (1, 2, and 3), which are the Mean Average Squared errors the Spline approximation method is more efficient than the Local Polynomial estimation method, so it is preferred to use it to estimate the components of the additive partial linear model.
2. It is possible to apply this APLM in a number of fields, including health, educational, social, financial, economic, and others, which have high flexibility in their inclusion on more than one linear explanatory variable as well as more than one non-linear parametric function. These features make the model highly flexible in various applications.
3. Using the additive partial linear model (APLM) to reduce a high-dimensional problem.

References

[1]. Clark, T. & Weist, K. (2006)" **Using out of-sample mean squared predication errors to test the martingale difference hypothesis**" Journal of Econometrics.

[2]. Cook, R. Dennis (1998) " **Regressions Graphics Ideas for studying Regressions therough Graiphics**" New York, wily.

[3]. Daowen Zhang and Xh. Ng Lin (2003) “**Hypothesis testing in semi parametric additive Mixed Models** " Biostatistics ;4,1; PP.57 -74

[4]. Degui Li, Yuan Ke and Wenyang Zhaing (2015) "**Models selection and structure specification in Ultra High dimensional Generalised semivarying coefficient Model**" The annials of statistics Vol.43, No.6 PP. 2676-2705.

[5]. Deng Guahua and Liang Hua (2010) "**Models averaging for Semi Parametric adiditive Partial linears Models**" Science china mathemiatics Vol.53 No.5 pp. 1363-1376.doi: http://dx.doi.org/10.5705/ss.202014.0151

[6]. Graciela Boente, Xuming Jinhua Zhou and He (2006) "**Robust estimateis in Generalizead Partially linear Models** "The Annials of statistics Vol.34, No. 6, pp. 2856-2878.

[7]. Hmood, M. Y. and Mousa, M. J. (2016), “**Using the levy model to estimate the stock returns for some Iraqi banks** ”, Journal of economics and administrative sciences, Vol. 89, No. 108, pp. 232-247.

[8]. Hmood, M. Y. and Saleh, T. A. (2016), “**Comparison of some semi-parametric methods in the analysis of the single index model**”, Journal of economics and administrative sciences, Vol. 91, No. 22, pp. 367-388.

[9]. Horowitz, J, Huang, J. L, Ma, S (2008) "**Asymptoitic Properties of bridge estimators in sparse highdimension regressions Model** " The Annials of statistics; 36 (2), pp.587 – 613.

[10]. Hua Liang and Runze Li (2009) "**Variables Selection for Partiailly linear Model with Measureament Errors**" Jamer Statisit Assocs; 104: 234- 248.

[11]. Hurdle, Wolfgang and Liang, Hua and Gao, Jiti (2000) "**partially Linear Models**", https://mpra.ub.uni-muenchen.de/39562/

[12]. Li wang, Xiang Liu, Hua Liang and Ray Nond J. Carroll (2011) " **Estimation and Variable Selection for Generalized Additive Partial Linear Models"** Ann stat. 39 (4) P.P 1827 - 1851.

[13]. Rong Liu and Lijian Yang (2010) "**Spline -Backfitted Kernel soothing of additive Coefficient Model**" Econometric Theory, Vol. 26 PP. 29-59.

[14]. Xia, YTong, H، Li، WandZhu, Li (2002)" **Anadaptive estimation of dimension reductionspace**" Jourinal of the Royal Sitatistical sorely, series B 64 Parts, P.P. 363-410

[15]. XiaoNi, Daowen Zhang and Helen Zhang Hao (2010) " **Variables Selection for semi-parametric Mixed Models in Longitudinal Studiies**" Biometricis, Vol. 66 No. (1) pp. 79 -99.

[16]. XiaoNi, Helen Zhang Hao and Daowen Zhang (2009) " **Automatic Models selection for Partially lineair Models**"Journial of Multivariatie Analysis 100 PP 2100- 2111.

[17]. Xinyi Li, Li Wang and Dan Nettleton (2019) "**Additive partially linear Models for ultra-high-dimensional regression**" Stat, Vol. 8, No,1.