Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$: 2312-9883
QJAE, Volume 25, Issue 2 (2023) ISSN$_{Print}$: 1816-9171

Bayesian Variable Selection For Single Index Model

Taha Hussain Ali Alshaybawee
Rasha Majeed Abed Alghanemi
Al-Qadisiyah University - College of Administration and Economics
Corresponding Author: Rasha Majeed Abed Alghanemi

Abstract **:** Nonparametric models with multivariate are suffer from the so called curse of dimensionality. Single index model one of the most attractive models use to overcome this problem, this model can reduce the dimensionality with holding a lot of flexibility of a nonparametric model, that what we need in each of econometrics and statistics. A Bayesian approach based on the asymmetric Laplace for estimation and variables selection was address by many researchers recently. In this paper, we proposed a new Bayesian lasso for single index model based on a scale mixture uniform. The Gaussian process prior have been considered to the nonparametric link function. Simulation examples (with different samples size and different modes) are considered, also real data for the proposed method show substantial improvements compare the other methods.

Keywords: nonparametric Bayesian inference, single index, semiparametric model, MCMC, scale mixture of uniform.

Introduction:

The term semi-parametric is attributed to the researcher (Oakes ,1981). (Powell ,1994) inappropriate application of parametric models may produce biased estimates and wrong inferences. On the contrary, non-parametric models have provided great flexibility to explore the shape of the function without prior knowledge in the data structure, Therefore, in recent years, non-parametric models have gained wide uses for them, as an alternative tool to avoid problems with the use of inappropriate parametric models. As it is known, non-parametric models suffer from the problem of (The curse of dimensionality) when there are a large number of explanatory variables. So the size of the space for the variables expands rapidly, so that the data becomes (sparsity) in the field and accordingly the estimation becomes unreliable and in this case nonparametric models can rarely be used to reach effective conclusions (Liu,2011). Therefore, single index model is an effective approach to deal with this problem.

The Single Index Model helps to reduce the dimensions by identifying the significant explanatory variables and put them in one indicator collectively. The single index model was studied by two researchers (Hardle & Stoker ,1989) (Ichimura ,1993), semi-parametric models can be used as a compromise between constrained parametric models and highly flexible nonparametric models (Lehrmann,2016).

The model have a finite parameter of unknown dimensions (β) , which represents the parametric part and an unknown link function g(.), this is what characterizes the semiparametric single index model as this model is of the most common semi-parametric. It can use in many different applied sciences for instance medical and economic.Whereas, the name of this model came from that all explanatory variables are summarized under one linear indicator ($\boldsymbol{X_i'\beta}$) (Stute & Zhu,2005). Single index model can get the most information about the relationship between the dependent variable response (Y) and variable independent (X) (Fan et al ,2018). The Single index models (SIMs) provide more elastic structures than the other of (GLM) models, by the overwhelming emphasis has been on good empirical and asymptotic properties of the point estimates of the index parameters and the link function *g (.)*. Most of the existing Bayesian approaches for (SIM) employ some basis representation such as splines and the wavelets (Antoniadis et al., 2004; Wang, 2009 ;Park et al., 2005), Wang (2009) used reversible jump Markov Chain Monte Carlo algorithms involving movable knots. (Gramercy and Lian, 2012) explain alternatives methods of Bayesian estimate by using a Gaussian process (GP) as prior to the unknown link function. New hierarchical representation for Lasso Bayesian was Suggested by (Mallick & Yi ,2014). They developed new Gibbs sampler method based on a mixture of uniform distribution (SMU) for Laplace density, by take the Laplace density as a scale mixture of uniform distribution :-

$$\frac{\lambda}{2} e^{-\lambda|\beta_j|} = \int_{s>|\beta|} \frac{1}{2s} \frac{\lambda^2}{\Gamma 2} s^{2-1}\, e^{-\lambda s}\, ds \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

In this paper, follow (Mallick & Yi ,2014), A new Bayesian lasso for single index model based on a scale mixture uniform. The Gaussian process prior have been considered to the nonparametric link function.

1. Bayesian single index model proposed method and prior assumption

Consider the following single index defined as (Ichamuri,1993)

$$y_i = g(X_i'\boldsymbol{\beta}) + \varepsilon_i \ , (i = 1,2,3 \ldots n)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Where $y_1, y_2 \dots\dots\dots\dots y_n$ are the dependent variables, and $\varepsilon_1, \varepsilon_2 \dots\dots\dots\dots \varepsilon_n$ are the errors and assumed to be (iid) normal distribution with (zero) mean and variance $\sigma^2{}_\varepsilon$, $\varepsilon_i \sim (0, \sigma^2{}_\varepsilon)$, $X_i = (x_1, x_2, \dots\dots x_p)'$ is p- dimensional predictive variable and *g(.)* is the unknown link function .
The likelihood function for error $\varepsilon_i = y_i - \mathrm{g}(X_i'\beta)$, (i=1,2,......,n) . Where (n) is the sample size can be shown as:-

$$\pi(\varepsilon/\sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(\varepsilon_i)^2\right\} \dots\dots\dots\dots\dots\dots\dots\dots (3)$$

$$\pi(\varepsilon/\mathrm{g}, \boldsymbol{\beta}, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{(y_i - \mathrm{g}(X_i'\boldsymbol{\beta}))^2}{\sigma^2}\right\} \dots\dots\dots \dots (4)$$

$$\pi(\varepsilon/\mathrm{g}, \boldsymbol{\beta}, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mathrm{g}(X_i'\boldsymbol{\beta}))^2\right\} \dots\dots\dots (5)$$

The Gaussian process distribution set as prior for the unknown link function $\mathbf{g}(.)$.
In the other word, the prior distribution for the $g(.)$ function is a Gaussian process with mean (zero) and square exponential covariance furcation for more detail (see Chio etal, 2011, and Gramcy &Lian 2012) , it on be shown as :-

$$\mathbf{g}(.) \sim GP\ (0, E(.,.)) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (7)$$

$$E(X_i, X_j) = \mathcal{S}\ exp\left\{-\frac{(X_i - X_j)^2}{d}\right\} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (8)$$

Where $\mathcal{S}$ and *d* hyper parameters .
Therefore, we can write the prior distribution for the unknown link function as follows:

$$\pi(\mathrm{g}/, \boldsymbol{\beta}, \sigma^2, \mathcal{S}) = det\ [E_n]^{-\frac{1}{2}} exp\ \left\{-\frac{\mathrm{g}_n' E_n \mathrm{g}_n}{2}\right\} \dots\dots\dots\dots\dots (9)$$

E_n with dimension (n×n) is denoted the covariance matrix and given:-

$$E(X_i\boldsymbol{\beta}, X_j'\boldsymbol{\beta}) = \mathcal{S} exp\left\{-(X_i - X_j)'\boldsymbol{\beta}\boldsymbol{\beta}'(X_i - X_j)\right\} \dots\dots\dots\dots\dots\dots (10)$$

(Gramacy and Lion ,2012) mention that the identifiable can be satisfy $\frac{\beta}{\sqrt{d}}$ with necessity for the constraint $\|\beta\| = 1$. When use Gaussian process as prior for non parametric link function , so that (β) will be use instead of $\frac{\beta}{\sqrt{d}}$ in the covariance function .

$$E(X_i'\boldsymbol{\beta}, X_j'\boldsymbol{\beta}) = \mathcal{S}\ exp\ \left\{-\left(X'_i\boldsymbol{\beta} - X'_j\boldsymbol{\beta}\right)^2\right\} \dots\dots\dots\dots\dots\dots\dots\dots (11)$$

For the hyperparamter $\mathcal{S}$, inverse Gamma will be set as a prior distribution, $\boldsymbol{\mathcal{S}} \sim IG(a_1, b_1)$, where a_1 and b_1 are the hyper parameters. Laplace distribution put as a prior distribution for the parameters index see (Tibshirani,1996. Park & Casella ,2008):-

$$\pi(\beta/\sigma^2) = \prod_{j=1}^{p} \frac{\gamma}{2(\sigma^2)^{\frac{1}{2}}} \exp\left\{-\frac{\gamma}{\sqrt{\sigma^2}}\|\boldsymbol{\beta}_j\|\right\} \dots\dots\dots\dots\dots\dots\dots\dots (12)$$

Where $(\gamma > 0)$ is penalty parameter .
In this paper, we will adapted (Mallick & Yi ,2014) new hierarchical representation for Lasso Bayesian. A new Gibbs sampler that developed the method of mixture uniform (SMU) of Laplace density, by take the Laplace density as a scale mixture of uniform distribution (SMUD) as in equation (1) for selection variable in single index model.

2. Scale mixture of uniform distribution (SMUD)

Follow Mallick and Yi (2014) Laplace distribution can be rewritten as a scale mixture of uniform and specific Gamma distribution, so that we can the prior will be shown as: Let $\gamma = \frac{\lambda}{\sqrt{o^2}}$

$$\pi(\boldsymbol{\beta}/\gamma) = \prod_{j=1}^{p} \frac{\gamma}{2} exp\{-\gamma|\boldsymbol{\beta}_j|\} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (13)$$

$$= \prod_{j=1}^{p} \int_{u>|\beta_j|} \frac{1}{2u_j} \frac{\gamma^2}{\Gamma(2)} u_j^{2-1} e^{-\gamma u_j} du_j \quad \dots\dots\dots\dots (14)$$

Therefore , the hierarchical model can be shown as :-

$$y/\beta,\ \sigma^2, \mathbf{g} \sim N(\mathbf{g}(\boldsymbol{X_i'}\boldsymbol{\beta}), \boldsymbol{\sigma^2})$$

$$\beta\Big/u_j \sim \prod_{j=1}^{p} unifiorm\ (-\sqrt{\sigma^2}.u_j, \sqrt{\sigma^2}.u_j)$$

$$u_j/\gamma \sim \prod_{j=1}^{p} Ga\ (2\ , \gamma\ )$$

$$\mathbf{g}/X_i, \beta,\ \mathcal{S} \sim N(0, E(.,.))$$
$$\boldsymbol{\mathcal{S}} \sim IG\ (a_{\mathcal{S}}, b_{\mathcal{S}})$$

$$\gamma \sim Ga\ (\ a_\gamma, b_\gamma)$$

$$\sigma^2 \sim Ga(a_{\sigma^2}, b_{\sigma^2}) \ldots\ldots\ldots\ldots\ \ldots\ldots\ldots.(15)$$

Where $a_{\mathcal{S}}, b_{\mathcal{S}},\ a_\gamma, b_\gamma, a_{\sigma^2}, b_{\sigma^2}$ are the hyper parameters.

4. The full conditional posterior distribution :

We can drive the full conditional posterior distribution follows :-

$$P(\mathbf{g_n}, \boldsymbol{\beta}, \boldsymbol{s}, \boldsymbol{\gamma}, \boldsymbol{\mathcal{S}}, \boldsymbol{\sigma^2}, \boldsymbol{y}) \propto$$
$$\left\{det[D]^{-1} exp\left\{-\frac{(y-\mathbf{g_n})'\boldsymbol{D^{-1}}(y-\mathbf{g_n})}{2}\right\}\right\} \times\ [\boldsymbol{E_n}]' \boldsymbol{exp}\ \left\{-\frac{\mathbf{g}_n' \boldsymbol{E_n} \mathbf{g_n}}{2}\right\} \times \prod_{j=1}^{p} \int_{u>\left|\boldsymbol{\beta}_j\right|} \frac{1}{2u_j} \frac{\gamma^2}{\Gamma(2)}\ u_j^{2-1} e^{-\gamma u_j} du_j\ \times$$
$$\prod_{i=1}^{n} \frac{1}{\sigma^2} exp\left(-\frac{s_j}{\sigma^2}\right) \times \prod_{j=1}^{p} (\gamma)^{a\gamma-1} \exp(-b_\gamma \gamma) \times (\boldsymbol{\mathcal{S}})^{-a\mathcal{S}-1} \times \exp\left(-\frac{b_{\mathcal{S}}}{\mathcal{S}}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(16)$$

4.1 The full conditional distribution for sampling $\boldsymbol{g_n}(.)$

$$\pi\left(\mathbf{g_n}/\boldsymbol{\mathcal{S}}, \beta, \gamma, \sigma^2, u_j, y\right) \propto \pi(y/\mathbf{g_n}, \beta, \sigma^2) \times \pi(\mathbf{g_n}/\boldsymbol{\beta})$$

$$\propto [\det(D)]^{-\frac{1}{2}} exp\left\{-\frac{(\boldsymbol{y}-\mathbf{g_n})'\boldsymbol{D^{-1}}(\boldsymbol{y}-\mathbf{g_n})}{\mathbf{2}}\right\} \times \det(\boldsymbol{E_n})^{-\frac{1}{2}} exp\left\{-\frac{\mathbf{g}_n' \boldsymbol{E_n} \mathbf{g_n}}{\mathbf{2}}\right\}$$

The full conditional distribution for $\mathbf{g_n}$ is $N(A_{\mathbf{g_n}}, B_{\mathbf{g_n}})$

$$A_{\mathbf{g_n}} = E_n(E_n + D)^{-1}(y)$$

$$B_{\mathbf{g_n}} = E_n(E_n + D)^{-1}(\ D)$$

4.2 The full conditional distribution for sampling β

$$\boldsymbol{\pi}(\boldsymbol{\beta}/\mathrm{y}, x, \mathbf{g_n}, u, \mathcal{S}, \gamma, \sigma^2) \propto \pi(y/\mathbf{g}_n, \boldsymbol{\beta}, x, u, \gamma, \mathcal{S}, \sigma^2) \times \pi(\mathrm{g}_n/\boldsymbol{\beta}, \mathcal{S}) \times \pi(\boldsymbol{\beta}/u, \gamma)$$

$$\propto exp\left\{-\frac{(y-\mathbf{g_n})'\boldsymbol{D^{-1}}(y-\mathbf{g_n})}{\mathbf{2}}\right\} \times [\det(D)]^{-\frac{1}{2}} \times \det(\boldsymbol{E_n})^{-\frac{1}{2}} exp\left\{-\frac{\mathbf{g}_n' \boldsymbol{E_n} \mathbf{g_n}}{\mathbf{2}}\right\} \times \prod_{\boldsymbol{j=1}}^{\boldsymbol{p}} \boldsymbol{I}\left\{\left|\boldsymbol{\beta_j}\right| < \boldsymbol{u_j}\right\}$$

The conditional distribution of (β) is truncated normal distribution, with:

$M = E_n\{E_n + D\}^{-1}(y - \mathbf{g}_n)\ I\{\left|\boldsymbol{\beta_j}\right| < u_j\}$

$V = \boldsymbol{E_n}\{\boldsymbol{E_n} + D\}^{-1}\ D$

4.3 The full conditional distribution for sampling $\boldsymbol{\mathcal{S}}$

$$\boldsymbol{\pi}(\mathcal{S}/\mathbf{g_n}, \boldsymbol{\beta}, x, \gamma, u, \sigma^2, y) \propto \pi(y/\mathbf{g}_n, \boldsymbol{\beta}, x, \gamma) \times \boldsymbol{\pi}(\mathrm{g}_n/\boldsymbol{\beta}, \mathcal{S}) \times \boldsymbol{\pi}(\boldsymbol{\mathcal{S}})$$
$$\propto exp\left\{-\frac{(\boldsymbol{y}-\mathbf{g_n})'\boldsymbol{D^{-1}}(\boldsymbol{y}-\mathbf{g_n})}{\mathbf{2}}\right\} [\det(D)]^{-\frac{1}{2}} \times \det(E_n)^{-\frac{1}{2}} exp\left\{-\frac{\mathbf{g}_n' E_n \mathbf{g_n}}{2}\right\} \times \left(\frac{1}{\mathcal{S}}\right)^{a_{\mathcal{S}}+1} exp\left\{-\frac{b_{\mathcal{S}}}{\mathcal{S}}\right\}$$

To sample ($\mathcal{S}$) Metropolis algorithm will be use.

4.4 The full conditional distribution for sampling $\boldsymbol{u_i}$

$$\pi(u/\mathbf{g_n}, \boldsymbol{\beta}, \gamma, \mathcal{S}, \sigma^2, y) \propto \pi(\beta/u) \times \pi(u)$$
$$\propto \prod_{j=1}^{p} exponential\ (\gamma)\ I(u_j > \left|\beta_j\right|)$$

Sample (u) from the left truncated exponential distribution .

4.5 The full conditional distribution for sampling $\boldsymbol{\sigma^2}$

$$\pi((\sigma^2/\mathbf{g_n}, \boldsymbol{\beta}, \mathcal{S}, u, y) \propto \pi(y/\mathbf{g_n}, \boldsymbol{\beta}, \sigma^2, x) \times \pi(\sigma^2)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} exp\left\{-\frac{1}{2\sigma^2}(y_i - \mathbf{g_n})^2\right\} \times \left(\frac{1}{\sigma^2}\right)^{a_{\sigma^2}} exp\left\{-\frac{b_{\sigma^2}}{\sigma^2}\right\}$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+a_{\sigma^2}+1} exp\left\{-\frac{1}{\sigma^2}\left\{\frac{1}{2}(y_i - \mathbf{g_n})^2 + b_{\sigma^2}\right\}\right\}$$

The conditional distribution for σ^2 is Inverse Gamma distribution.

5. Simulation study

The purpose of the simulation approach that we present in this section is to assess how Quality the proposed Bayesian method (BUSIReg), and to assess how it performs compare to other two existing method, Bayesian quantile regression single index model at 0.5 ,quantile ($BQSIM_{\tau=0.5}$) and classic single index regression (SIReg). We will consider two simulation examples. The effectiveness of the methods is assessed using median of mean absolute deviations, referred to as "MMAD" and mean absolute error referred to as "MAE". The first 3000 iterations of the 13000-iteration belong to Bayesian methods were removed as burn-in. All computations were done using R package.

5.1 Simulation example 1

In this simulation example, The data are generated by the following model

$$y = g(x_i^T\beta) + \sqrt{(sin\ \ x_i^T\beta + 1\,)\epsilon} \qquad where\ g(t) = 10\ \ sin\,(\,0.75t)$$

where $x_i, (i = 1,2,3,4,5,6,7,8,9,10)$ are identical independent distribution are simulated from normal distributions, which have mean 0 and variance $(0.25)^2$. $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9, \beta_{10}) = \frac{1}{\sqrt{3}}(1,1,0,0,1,0,0,0,0,0)$, and random errors term are generated from exponential distribution with shape parameter (0.5). six samples with different size are used in this simulation example. The results listed in Table (1) represents parameters estimation under estimation methods.

Table (1): The average of parameters estimates for BSIReg, $BQSIM_{\tau=0.5}$ and SIReg according to 100 replication for simulated Example1

	True parameter	0.577	0.577	0.000	0.000	0.577	0.000	0.000	0.000	0.000	0.000
n=25	SIReg	0.946	0.903	0.734	0.693	0.792	0.856	0.758	0.864	0.793	0.893
	$BQSIM_{\tau=0.50}$	0.172	0.677	0.572	0.846	0.125	0.711	0.842	0.508	0.634	0.652
	BSIReg	**0.671**	**0.502**	**0.147**	**0.169**	**0.601**	**0.065**	0.193	0.082	0.121	0.286
n=50	SIReg	0.746	0.852	0.481	0.613	0.892	0.572	0.624	0.672	0.581	0.542
	$BQSIM_{\tau=0.50}$	0.067	0.162	0.542	0.924	0.105	0.462	0.511	0.672	0.582	0.562
	BSIReg	**0.528**	**0.621**	**0.058**	**0.234**	**0.587**	**0.142**	0.261	0.134	0.256	0.184
n=100	SIReg	0.154	0.245	0.346	0.521	0.763	0.341	0.562	0.541	0.381	0.458
	$BQSIM_{\tau=0.50}$	0.431	0.562	0.562	0.452	0.833	0.562	0.475	0.582	0.432	0.452
	BSIReg	**0.531**	**0.602**	**0.106**	**0.163**	**0.491**	**0.193**	**0.078**	**0.131**	**0.097**	**0.110**
n=150	SIReg	0.734	0.694	0.477	0.394	0.842	0.482	0.393	0.623	0.403	0.393
	$BQSIM_{\tau=0.50}$	0.692	0.657	0.395	0.367	0.453	0.198	0.351	0.370	0.342	0.295
	BSIReg	**0.562**	**0.529**	**0.207**	**0.143**	**0.511**	**0.093**	**0.084**	**0.048**	**0.104**	**0.272**
n=200	SIReg	0.293	0.184	0.412	0.294	0.707	0.294	0.692	0.452	0.584	0.804
	$BQSIM_{\tau=0.50}$	0.373	0.451	0.389	0.292	0.562	0.319	0.382	0.346	0.532	0.527
	BSIReg	**0.505**	**0.611**	**0.219**	**0.184**	**0.233**	**0.056**	**0.127**	**0.148**	**0.136**	**0.073**
n=250	SIReg	0.193	0.246	0.307	0.526	0.430	0.294	0.429	0.526	0.394	0.584
	$BQSIM_{\tau=0.50}$	0.342	0.303	0.462	0.393	0.473	0.276	0.563	0.247	0.394	0.364
	BSIReg	**0.520**	**0.582**	**0.184**	**0.048**	**0.148**	**0.193**	**0.111**	**0.128**	**0.086**	**0.088**

The results shown in Table (1), it represent parameters estimation in direct way. The parameters estimates for the proposed method are clearly closer to the true parameter values compare to the other methods method. Therefore, the performance of our proposed method (BSIReg) appears to be quite good when compared to SIReg and $BQSIM_{\tau=0.5}$.

Figure 1 shows the parameter estimates for the SIReg , $BQSIM_{\tau=0.5}$ and our proposed method BSIReg in the first simulation example. The black line belong to the values of true parameters, while the blue line belong to value of parameters estimation by our proposed method . We can see the blue line is very closed from black line via different sample size. This figure show that our estimators are very close to the values of true parameter. Clearly, it can be seen that our method (BSIReg) outperforms on SIReg and $BQSIM_{\tau=0.5}$.

.

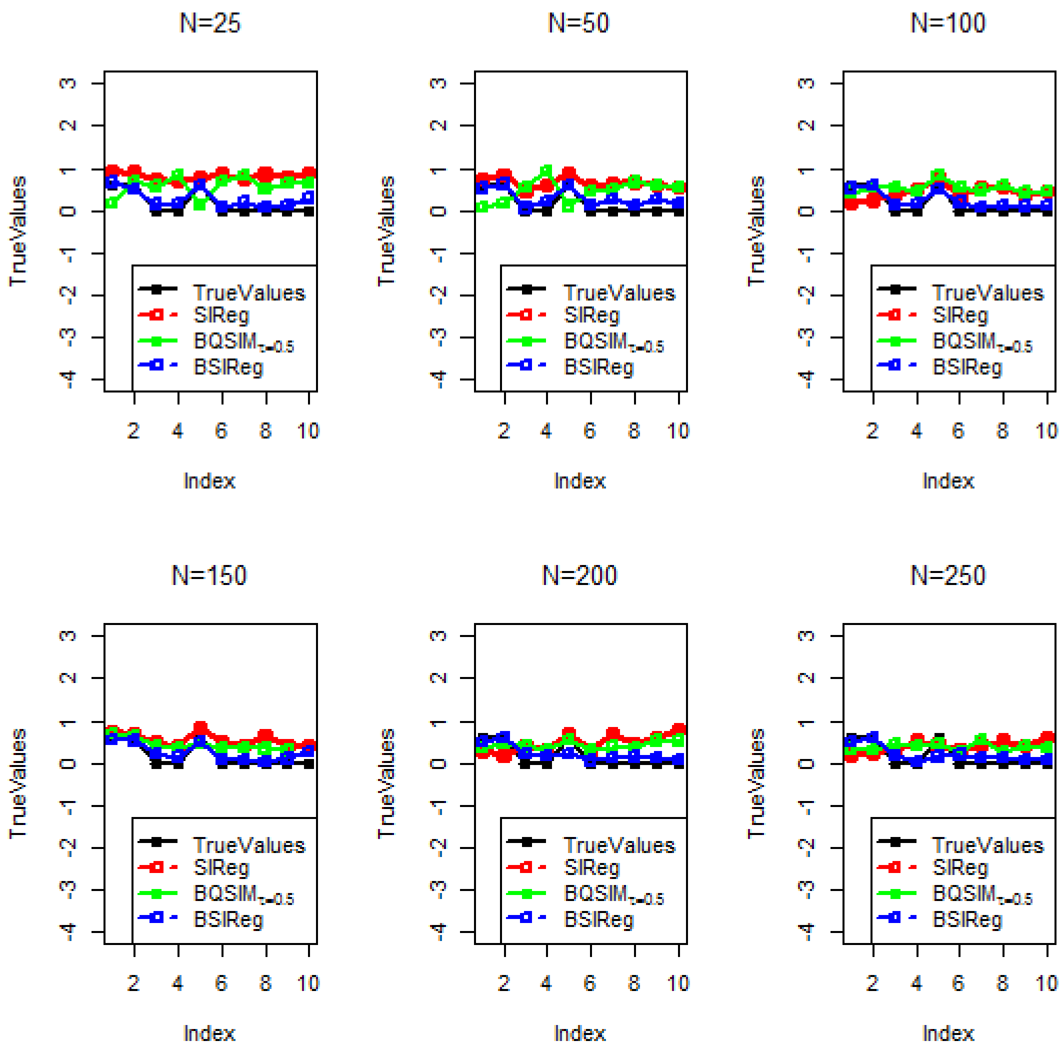

Figure 1. The estimates of the explanatory variable effects for Simulation 1, using SIReg , $BQSIM_{\tau=0.5}$ and our method (BSIReg) for $N \in \{25, 50, 100, 150, 200, 250\}$

5.2 Simulation example 2

For simulation example 2, the performance of the proposed method is demonstrated by following model

$$y = g(x_i^T \beta) + \epsilon \quad \text{where } g(t) = 5(-t^2)$$

where the independent variables $x_i, (x_1, x_2, x_3, x_4, x_5)$, these variables are simulated from standard normal distributions, with the following parameters $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) = \frac{1}{\sqrt{5}}(2,0,1,0,0)$. In this simulation example , we generated the random error from exponential distribution .In this simulation example five different samples size have been generated.

Table (2): The MAE and MMAD and SD for the simulation examples 2

Samples	Methods	MAE	MMAD
N=25	SIReg	0.935 (0.702)	0.842 (0.525)
	$BQSIM_{\tau=0.50}$	0.843 (0.561)	7.221 (0.589)
	BSIReg	0.326 (0.165)	0.497 (0.242)
N=50	SIReg	0.875 (0.534)	0.893 (0.642)
	$BQSIM_{\tau=0.50}$	0.654 (0.473)	0.623 (0.396)

	BSIReg	0.345 (0.142)	0.334 (0.131)
N=100	SIReg	0.743 (0.572)	0.643 (0.431)
	BQSIM$_{\tau=0.50}$	0.602 (0.423)	0.633 (0.491)
	BSIReg	0.338 (0.167)	0.262 (0.096)
N=150	SIReg	0.757(0.457)	0.738 (0.467)
	BQSIM$_{\tau=0.50}$	0.742 (0.445)	0.722 (0.473)
	BSIReg	0.356 (0.167)	0.334 (0.166)
N=200	SIReg	0.563 (0.353)	0.474 (0.279)
	BQSIM$_{\tau=0.50}$	0.535 (0.323)	0.431 (0.284)
	BSIReg	0.363 (0.182)	0.307 (0.183)
N=250	SIReg	0.531 (0.393)	0.517 (0.394)
	BQSIM$_{\tau=0.50}$	0.578 (0.384)	0.519 (0.372)
	BSIReg	0.272 (0.144)	0.205 (0.129)

Note: In the parentheses are standard deviation (S.D)

The results listed in Table (2), It is evident that the proposed method has the smallest MAE and MMAD compared with other method via different sample size, which demonstrates that our estimations are more accurate than the other two methods. Also ,we can see the our proposed method has smallest standard deviation (S.D) compared with other method via different sample size, which once more demonstrates that our estimations are more accurate than the other two methods.

6. Real data set

In this section, three methods are used to analyze prostate cancer data. These dataset first analyzed by (Stamey *et* al. 1989). The prostate cancer data is available in the package (‘bayesQR) in R. The dataset describes 97 individuals, eight independent variables and one response is the level of prostate antigen (y_i) referred to as (lpsa) , and the eight independent variables are: logarithm of cancer amount(x_1) referred to (lcavol), logarithm of the weight of prostate (x_2) referred to (lweight), age (x_3), logarithm of the volume of benign enlargement of the prostate(x_4) referred to as (lbph), seminal vesicle invasion (x_5) referred to as (svi), logarithm of Capsular penetration in prostate cancer (x_6) referred to (lcp), Gleason score (x_7) referred to (gleason) and percentage of Gleason scores 4 or 5,(x_8) referred to as (pgg45).

Similar to simulation, in this section we compare three methods: SIReg,$BQSIM_{\tau=0.5}$ and our proposed method BSIReg. The mean squared error (MSE), median of mean absolute deviations(MMAD) and mean absolute error(MAE) are used to evaluate the three methods, and 95% intervals. The results of the MAE,MMAD and MSE are listed in Table (3).

Table (3): show MAE,MMAD and MSE for the prostate cancer data

Methods	MAE	MMAD	MSE
SIReg	1.627	1.421	2.647
BQSIM$_{\tau=0.50}$	1.261	1.083	1.590
BSIReg	0.823	0.892	0.677

From Table (3), it can see MMAD,MSE and MAE are generated by our proposed method much smaller than (MMAD,MSE and MAE) are generated by other two methods. Therefore, the our proposed method is consider best than other two method. The 95% intervals of the proposed method (BSIReg) and the other competing methods are listed in table (4).

Table 4 Estimates 95% credible intervals of credible intervals parameters of the prostate cancer data.

Symbol of independent Variables	Parameters	SIReg Mean (95% CrI)	$BQSIM_{\tau=0.5}$ Mean (95% CrI)	BSIReg Mean (95% CrI)
lcavol	β_1	0.176 (-0.224, 0.974)	0.525 (0.359, 1.837)	0.617 (0.446, 1.638)
lweight	β_2	0.306 (-0.222, 0.690)	0.187 (-0.010, 0.405)	0.075 (-0.131, 0.422)
Age	β_3	0.001 (-0.011, 0.012)	0.004 (-0.020, 0.016)	0.003 (-0.028, 0.026)
lbph	β_4	-0.106 (-0.154, -0.067)	-0.212 (-0.357, -0.021)	-0.089 (-0.265, 0.013)
Svi	β_5	0.039 (-0.042, 0.334)	0.104 (-0.431, 0.207)	0.152 (-0.311, 0.262)
lcp	β_6	0.003 (-0.007, 0.008)	0.005 (-0.009, 0.056)	-0.015 (-0.102 , 0.274)
Gleason	β_7	-0.051 (-0.163, 0.423)	0.009 (-0.113, 0.232)	0.002 (-0.020, 0.024)
Pgg45	β_8	0.321 (0.000, 0.543)	0.210 (0.000, 0.461)	0.0007 (-0.178, 0.351)

From result listed in Table (4), it can see that the estimates according to our proposed method are very close to the $BQSIM_{\tau=0.5}$ estimated, and credible intervals computing according to our proposed method are much narrower than the credible intervals computing by the other two methods. Hence, the results in Table (4), shows support for using of the our proposed method to inference in single index regression. Also The credible intervals for the parameter estimates for the three methods shown in Figure (2). It is apparent from this figure that the BSIReg has a good narrower property compared to the other methods.

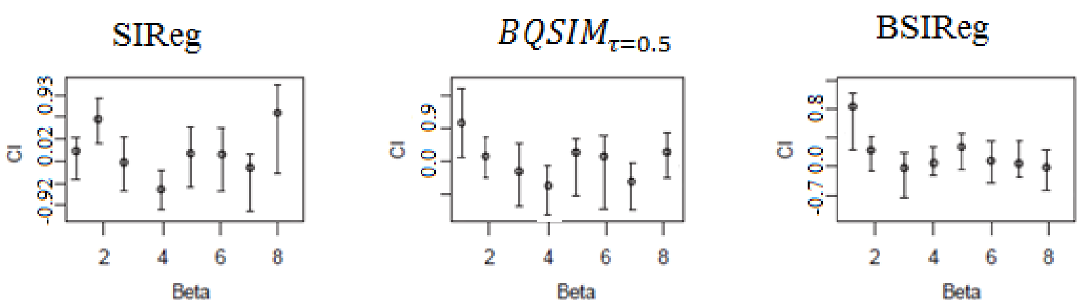

Figure 2. Credible intervals of the parameter estimates for the three methods under study (prostate cancer data).

7. Conclusion

In this paper, we proposed A new Bayesian lasso for single index model based on a scale mixture uniform. The Gaussian process prior have been considered to the nonparametric link function.

Simulation examples and real data were considered to illustrate and examine the performance of the proposed method compared with that of two other methods. The results that we reported in this paper demonstrated the superiority of our proposed method (BSIReg) over other two methods (SIReg and $BQSIM_{\tau=0.5}$.). therefore, we can conclude that the proposed method show substantial improvements compare the other methods.

References :-

1- Antoniadis, A., Grégoire, G., & McKeague, I. W. (2004). "**Bayesian estimation in single-index models**". Statistica Sinica, 1147-1164.

2- Choi, T., Shi, J. Q., & Wang, B. (2011). "**A Gaussian process regression approach to a single-index model**". Journal of Nonparametric Statistics, 23(1), 21-36.

3- Fan, Y., Härdle, W. K., Wang, W., & Zhu, L. (2018). "**Single-index-based CoVaR with very high-dimensional covariates**". Journal of Business & Economic Statistics, 36(2), 212-226.

4- Gramacy, R. B., & Lian, H. (2012). "**Gaussian process single-index models as emulators for computer experiments**". Technometrics, 54(1), 30-41.

5- Härdle, W., & Stoker, T. M. (1989). "**Investigating smooth multiple regression by the method of average derivatives".** Journal of the American statistical Association, 84(408), 986-995.

6- Ichimura, H. (1993). "**Semiparametric least squares (SLS) and weighted SLS estimation of single-index models**". Journal of econometrics, 58(1-2), 71-120.

7- Lehrmann, A. M. (2016). "**Non-parametric models for structured data and applications to human bodies and natural scenes**" ,(Doctoral dissertation, ETH Zurich).

8- Liu, X. (2011)." **Penalized Variable Selection for Semiparametric Regression Models**", (Doctoral dissertation, University of Rochester).

Mallick, H., & Yi, N. (2014). "**A new Bayesian lasso**". Statistics and its interface, 7(4), 571.

9- Oakes, D. (1981)." **Survival times: aspects of partial likelihood".** International Statistical Review/Revue Internationale de Statistique, 235-252.

10- Park, C. G., Vannucci, M., and Hart, J. D. (2005). "**Bayesian methods for wavelet series in single-index models.**" Journal of Computational and Graphical Statistics, 14(4).

11- Park, T., & Casella, G. (2008). "**The bayesian lasso**". Journal of the American Statistical Association, 103(482), 681-686.

12- Powell, J. L. (1994)." **Estimation of semi parametric models**". Handbook of econometrics, 4, 2443-2521.

13- Stamey, T. A., Kabalin, J. N., McNeal, J. E., Johnstone, I. M., Freiha, F., Redwine, E. A., & Yang, N. (1989). "**Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate**". II. Radical prostatectomy treated patients. The Journal of urology, 141(5), 1076-1083.

14- Stute ,W., and Zhu , L (2005)," **Non-parametric checks for single index models** " . The Annals of statistics , vol.33 , No.3 , pp. 1048-1083.

15- Tibshirani, R. (1996). "**Regression shrinkage and selection via the lasso**". Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

16- Wang, H.-B. (2009). "**Bayesian estimation and variable selection for single index models**." Computational Statistics & Data Analysis, 53(7): 2617–2627.