Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (2) (2024), 524 – 536
DOI : 10.15672/hujms.1303627

Research Article

Principal simple linear regression

Heydar Ali Mardani-Fard*

Department of Mathematics and Statistics, Yasouj University, Yasouj, Iran

Abstract

In this paper, we propose a new method called *principal simple linear regression* for predicting a continuous response variable, Y using a single continuous predictor, X by utilizing multiple regression lines. This method is based on the theory of principal points and offers several advantages over classical simple linear regression methods, such as the ability to predict central, dispersion, and distributional tendencies of Y on X, and simultaneous estimation. We provide the main properties, inferences, and limiting behavior of the estimators. Additionally, we conduct a comprehensive simulation study to validate our theoretical results. The model is also applied to real datasets to demonstrate its effectiveness.

Mathematics Subject Classification (2020). 62J05, 62F30, 62J20

Keywords. Simple linear regression, principal points, quantile regression

1. Introduction

Simple linear regression (SLR) has long been a cornerstone in statistical analysis, providing insights into the relationship between a continuous response variable, Y and a single continuous predictor variable, X. Traditionally, SLR focuses on determining the central tendency of Y given X. However, this emphasis on central tendency often leads to an oversight of critical aspects in the Y-X relationship, such as dispersion and distributional tendencies.

In response to this limitation, we introduce a novel method named principal simple linear regression (PSLR) in this paper. PSLR leverages the theory of principal points as outlined by [7] and extends from seminal works such as [16]. These principal points allow us to utilize multiple regression lines, enabling a more comprehensive exploration of the Y-X relationship. This approach stands out from conventional SLR and other methodologies, including those focused solely on dispersion tendencies such as variance components [4], ARCH [5], and GARCH [3] models. Furthermore, it distinguishes itself from multiple quantile regression [10], a method requiring separate model fittings and, consequently, distinct error variance estimations. Our work builds upon the foundational principles laid out by [6] and [17], exploring the estimation of principal points and their self-consistency in elliptical distributions. Subsequent research by [15] delves into the estimation of principal points for univariate distributions. The uniqueness of two principal points for univariate

*Corresponding Author.
Email addresses: h_mardanifard@yahoo.com (H.A. Mardani–Fard)
Received: 27.05.2023; Accepted: 27.11.2023

location mixtures is examined by [18], while Stampfer and Stadlober [14] propose methods for estimating principal points. Expanding beyond univariate settings, Matsuura and Kurata [11] explore principal points of a multivariate mixture distribution. Matsuura et.al. [12] present optimal estimators for minimizing expected mean squared distance. Yamashita [19] contributes a doctoral dissertation studying principal points for a multivariate binary distribution, and the comparison of model selection methods for their estimation is presented by [20]. The utility of principal points extends to diverse applications, as seen in the work of [9], who apply them to partition functional gene expression data. Yu [21] contributes to our understanding by investigating the uniqueness of principal points with respect to p-order distance for a class of univariate continuous distributions.
The paper is organized as follows: an overview of the definition and properties of principal points is given in the subsequent part of this section. Section 2 contains the main theoretical results, including the definition and properties of PSLR, methods of parameter estimation, and asymptotic behavior of the estimators. In Section 3, a real data analysis is performed using the PSLR method. This section includes a specific focus on comparing the outcomes with those obtained from multiple quantile regression, presented exclusively through figures for clarity. Further, in Section 4, comprehensive simulations are conducted to illustrate the asymptotic behavior of the estimators. The results are presented to provide a thorough understanding of the method's performance. Finally, Section 5 encapsulates our conclusions.

1.1. Principal points

In this section, we will provide an overview of the concept of principal points, which is the foundation of the proposed PSLR method. Principal points were first introduced by [7] as a way to minimize the expected squared distance between a random vector $\mathbf{X}$ and a set of points $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k\}$:

Definition 1.1. [7] Let $\mathbf{X}$ be a random vector with finite second moment. the k-principal points of $\mathbf{X}$ is the set of $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_k\}$, which minimizes the expected (squared) distance between $\mathbf{X}$ and $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k\}$:

$$\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_k\} = \underset{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k}{\operatorname{argmin}} \mathbb{E}\left[d^2\left(\mathbf{X}|\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k\right)\right], \tag{1.1}$$

in which

$$d^2\left(\mathbf{X}|\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k\right) = \min\left\{\left(\mathbf{X}-\mathbf{p}_j\right)'\left(\mathbf{X}-\mathbf{p}_j\right); j = 1, 2, \ldots, k\right\}.$$

The minimum value of this expected squared distance is denoted by $P_{\mathbf{X}}(k)$. If $k = 1$, then the 1-principal point of $\mathbf{X}$ is equivalent to the expectation of $\mathbf{X}$. Principal points can be seen as a generalization of the mean, which is the minimizer of the expected squared distance between $\mathbf{X}$ and a single point:

$$\mathbb{E}(\mathbf{X}) = \underset{\mathbf{p}}{\operatorname{argmin}}\, \mathbb{E}\left\{\left(\mathbf{X}-\mathbf{p}\right)'\left(\mathbf{X}-\mathbf{p}\right)\right\}.$$

Previous research has focused on extending and applying principal points in various fields. For example, methods of estimating principal points [6], properties and characterizations of principal points for continuous univariate distributions [22], principal points for elliptical distributions [17], and functional principal points [2] have been studied. Additionally, it was shown by [7] that principal points are a linear operator, meaning that if we have a set of principal points for a random vector $\mathbf{X}$ and we apply a linear transformation to $\mathbf{X}$ and its principal points, then the new set of points will also be principal points for the transformed vector. More precisely, if $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_k$ be the k-principal points of $\mathbf{X}$

and $\mathbf{Y} = \boldsymbol{\mu} + \sigma H\mathbf{X}$, where $\boldsymbol{\mu} \in \mathcal{R}^d$, $\sigma \in \mathcal{R}^+$, and H is an orthogonal matrix, then the k-principal points of $\mathbf{Y}$ will be:

$$\boldsymbol{\eta}_j = \boldsymbol{\mu} + \sigma H\boldsymbol{\xi}_j, \qquad j = 1, 2, \ldots, k. \tag{1.2}$$

Flury [7] proved that for univariate symmetric continuous distributions the two principal points are: $\xi_{1,2} = \pm\mathbb{E}\,|X|$ whenever $f\,(0)\,\mathbb{E}\,|X| < 0.5$. This holds for the standard normal distribution and t_ν distribution with $\nu \geq 3$.
For the standard normal distribution, Z, the k-principal points are provided in Table 1 for $1 \leq k \leq 5$.

Table 1. The $k = 1, \cdots, 5$ principal points of the standard normal distribution

k	Principal points	$P_\Phi(k)$
1	0.0	1.000
2	$\pm\sqrt{2/\pi}$	$1 - 2/\pi \approx 0.3634$
3	0.0, $\pm$1.227	0.1900
4	$\pm$0.451, $\pm$1.507	0.1170
5	0.0, $\pm$0.754, $\pm$1.707	0.0800

Additionally, using the linear operator property of principal points, it is easy to compute the principal points for non-standard normal distributions by using equation (1.2). It is also important to note that for normal distributions, the 2-principal points fully determine the distribution:

$$\left.\begin{array}{rcl} \mu &=& (\xi_1 + \xi_2)\,/2 \\ \sigma &=& \sqrt{\pi/2}\,(\xi_2 - \xi_1)\,/2 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{rcl} \xi_1 &=& \mu - \sigma\sqrt{2/\pi} \\ \xi_2 &=& \mu + \sigma\sqrt{2/\pi} \end{array}\right. \tag{1.3}$$

Furthermore, it is worth noting that for a t_ν distribution, the 2-principal points fully determine the distribution, and for a χ^2_ν distribution, the 1-principal point, which is the expected value, is equal to ν. Additionally, it has been shown by [7] that for univariate symmetric distributions, the principal points are also symmetric.

2. Principal simple linear regression (PSLR)

In this section, we focus on the principal points of $Y|X = x$ that can be written as functions of x, known as the conditional principal points, represented by $\{\xi_1(x), \xi_2(x), \ldots \xi_k(x)\}$. We will specifically focus on the case when the conditional principal points are linear in x, which will be the main topic of the rest of this paper.

2.1. Definition and properties

We define the concept of PSLR as linear functions of conditional principal points of Y given $X = x$. PSLR can be written as:

$$\xi_j(x) = \alpha_j + \beta_j x; \qquad j = 1, 2, \ldots, k. \tag{2.1}$$

This is an extension of SLR which only focuses on the linear conditional central tendency of $Y|X = x$. However, PSLR takes into account both linearity in central and dispersion tendencies. For example, when $k = 2$, PSLR can model both the conditional mean and variance of $Y|X = x$, simultaneously. With higher values of k, other distributional conditional tendencies such as skewness can also be modeled as a function of x.
When $k = 2$, homogeneity of variance (constant variance) is equivalent to equality of slopes: $\beta_1 = \beta_2$. For instance, in the case of the regular conditions of SLR, i.e., when

$y = \alpha + \beta x + \varepsilon$, with *iid* ε's of individuals with common distribution $N(0, \sigma^2)$, the 2-principal points of $Y|X = x$ will be:

$$\xi_j(x) = \alpha_j + \beta x, \qquad j = 1, 2. \tag{2.2}$$

When β_j in (2.1) are not equal, the model would be heterogeneous (especially the conditional standard deviation is linear in x). More precisely, let $Y|X = x \sim N(a+bx, (c+dx)^2)$, and suppose that $c + dx \geq 0$. Then from (1.3) the 2-principal points are:

$$\xi_{1,2}(x) = (a + bx) \pm \sqrt{2/\pi}(c + dx),$$

which can be written as:

$$\xi_{1,2}(x) = \alpha_{1,2} + \beta_{1,2}x \coloneqq \left(a \pm \sqrt{2/\pi}c\right) + \left(b \pm \sqrt{2/\pi}d\right) x.$$

Therefore, when the β_j values in equation (2.1) are not equal, the model will be characterized by heterogeneous variation (specifically linearity of standard deviation in x); this type of model can be useful when the scatter plot of y and x is in the shape of a funnel. Additionally, using PSLR allows for modeling multiple lines, simultaneously, as opposed to separate quantile regression models. This can provide a more comprehensive understanding of the relationship between X and Y. Furthermore, by constraining the slopes to be equal in the PSLR model, homogeneity of variance can be tested and accounted for in the analysis.

2.2. Parameter estimation

Flury [6] devised four methods to estimate principal points. Among these, three are applicable to univariate cases: Maximum Likelihood Estimation (MLE), Constrained Minimization subject to symmetry, and Unconstrained Minimization. The MLE method requires assuming a known distribution family for the data. To apply unconstrained and constrained optimization, based on the definition of principal points, they can be estimated by minimizing the sample version of the expected squared distance between X and a set of points $\{p_1, p_2, \ldots, p_k\}$, which is given by the following equation:

$$\left\{\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_k\right\} = \underset{p_1, p_2, \ldots, p_k}{\operatorname{argmin}} \sum_{j=1}^{n} d^2\left(x_j | p_1, p_2, \ldots, p_k\right). \tag{2.3}$$

The unconstrained optimization method yields the immediate outcome of (2.3). Constrained optimization occurs when considering symmetry for principal points, and in this case, the optimization involves points that exhibit symmetry around a single point.This point is either the mid-principal point if k is odd or the average of two centroid principal points if k is even.

For estimating the parameters of PSLR, we consider a set of n observations, each consisting of a predictor variable, x and a response variable, y. Assumptions are made that both x and y are continuous variables. Alongside the discussion on estimating principal points, there are three different methods for estimating the parameters of the model. The ML method can be used when the distribution of $Y|X = x$ is known and has a few unknown parameters. When only the symmetry of the distribution is known, or when the variance of the model is assumed to be constant, the constrained method can be applied. Lastly, when there is no information about the population distribution, the unconstrained optimization method can be used to estimate the principal lines.

Maximum likelihood estimation. The ML method is based on our knowledge of the true distribution of $Y|X = x$. If the conditional distribution of $Y|X = x$ is known, we can estimate the parameters by maximizing the log-likelihood function. For instance, we consider two especial cases:

Normal distribution with homogeneous variances. If $Y|X = x \sim N(a + bx, \sigma^2)$ follows a normal distribution with homogeneous variances, it can be seen from equation (1.3) that the 2-principal points of $Y|X = x$ are given by

$$\xi_{1,2}(x) = \underbrace{a \pm \sqrt{2/\pi}\sigma}_{\alpha_{1,2}} + \underbrace{bx}_{\beta x}. \tag{2.4}$$

Since the ML estimators are equivariant under one-to-one transformations, the ML estimators of the 2-principal points are given by

$$\hat{\alpha}_{1,2} = \hat{a} \mp \sqrt{2/\pi}\hat{\sigma}, \text{ and } \hat{\beta} = \hat{b}, \tag{2.5}$$

where $\hat{a}$, $\hat{b}$, and $\hat{\sigma}$ are the ML estimators for ordinary SLR.

Normal distribution with heterogeneous variances. In this case, we consider the conditional distribution of $Y|X = x$ to be of the form $N\left(a + bx, (c + dx)^2\right)$, where c and d are chosen such that $c + dx$ is non-negative within the range of x. By analyzing the distribution, it can be shown that the 2-principal points of $Y|X = x$ are given by the following equation:

$$\xi_{1,2}(x) = \underbrace{a \pm \sqrt{2/\pi}c}_{\alpha_{1,2}} + \underbrace{\left(b \pm \sqrt{2/\pi}d\right)x}_{\beta_{1,2}x}. \tag{2.6}$$

In this case, there are no regular ML estimators for the parameters and they must be found numerically, taking into account the condition of non-negativity of $c + dx$ in the range of x. It's worth noting that in this case, the scatter plot of y and x would have a funnel shape.

Unconstrained and constrained estimations. When the true form of the population distribution is not known, the sample version of the equation in (2.3) can be minimized to estimate the parameters:

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{n} d^2\left(y_i | \alpha_j + \beta_j x_i;\ j = 1, 2, \ldots, k\right), \tag{2.7}$$

where $\boldsymbol{\theta}$ is the parameter vector containing α_j's and β_j's.
This is known as the non-parametric approach, where the parameters are estimated without assuming any specific distribution for the data. However, if information about symmetry of the population is available, the minimization can be constrained to a set of symmetric points. If k is even, the principal lines are symmetric around their mean-line, and if k is odd, the principal lines are symmetric around the central principal line.

2.3. Asymptotic behavior

In this section, we discuss the asymptotic properties of the estimators for PSLR. Let $\psi(\theta)$ be the function that is minimized in (2.7) and ψ_i be the related summand:

$$\psi(\boldsymbol{\theta}) = \sum_{i=1}^{n} \psi_i(\boldsymbol{\theta}) \coloneqq \sum_{i=1}^{n} d^2\left(y_i | \alpha_j + \beta_j x_i;\ j = 1, 2, \ldots, k\right). \tag{2.8}$$

When the exact form of the population distribution is unknown, we can use the sample version of (2.3) to estimate the parameters. If we know that the population distribution is symmetric, we can constrain the minimization to values of $\boldsymbol{\theta}$ for which the principal lines are symmetric around the central principal line.

Theorem 2.1. *Under regular conditions, where the lines do not cross in the range of x, the unconstrained and constrained estimators of* (2.7) *are consistent and asymptotically normal with a mean of $\boldsymbol{\theta}$ and a covariance structure that can be approximated numerically as the inverse of the Hessian matrix.*

Proof. The proof for these results can be deduced from well-known limit theorems. Remember from (2.8) that $\psi_i(\boldsymbol{\theta}) = \min_{1\le j\le k}(y_i - \alpha_j - \beta_j x_i)^2$; hence, ψ_i is twice differentiable at $\boldsymbol{\theta}$ until the $y_i - \alpha_j - \beta_j x_i$; $j = 1, 2, \cdots, k$ are not vanish. So, if the minimization is taken on $\boldsymbol{\theta}$'s for which the lines does not meet in the range of x, then the score function is satisfied in *regular conditions* and so the asymptotic results hold. □

3. Real data analysis

In this section, we implement the PSLR method on two real datasets sourced from the R package `datasets`: `trees` and `mtcars`. Additionally, we apply quantile regression (QR) with $\tau = 0.25$ and 0.75 to the same data, facilitating a comparative analysis between the results of QR and PSLR.

Example 3.1. `mtcars` dataset.[1, 13] This data was extracted from the 1974 Motor Trend US magazine and includes information on fuel consumption and various aspects of automobile design and performance for 32 automobiles (1973-74 models). In this example, we examine the relationship between `mpg` and `wt`. Figure 1 shows the scatter plot and Table 2 contains the estimated parameters for SLR, QR, and PSLR. In reviewing the results presented in Table 2, it becomes evident that all parameters across the three models demonstrate statistical significance. Absence of significant evidence of heteroscedasticity suggests parallelism between the two principal lines. The confidence intervals, approximately defined with β_1 within the range of $(-7.082, -6.222)$ and β_2 within the range of $(-6.536, -5.800)$ using a confidence interval length of 2 standard errors, exhibit overlapping. Despite similar findings regarding homo/heteroscedasticity with QR, the estimates show higher standard deviations. Moreover, the visual representation in Figure 1 indicates non-parallelism of the fitted quantile lines.

Table 2. SLR, QR, and PSLR parameter estimations for `mtcars` dataset

parameter	SLR		QR		PSLR	
α_1	37.29	(1.878)	36.91	(0.947)	39.64	(0.658)
β_1	−5.344	(0.559)	−6.083	(0.483)	−6.652	(0.215)
α_2			38.14	(3.987)	43.75	(0.726)
β_2			−5.114	(1.145)	−6.168	(0.184)

Example 3.2. `trees` dataset.[8] This data set includes information on the girth, height, and volume of 31 felled black cherry trees. In this example, we examine the relationship between the girth and volume of the trees. Figure 2 shows the scatter plot and Table 3 contains the estimated parameters for SLR, QR, and PSLR. As observed in Table 3, all parameters for the three models exhibit statistical significance. Non-parallelism of the principal lines suggests the presence of heteroscedasticity in the data. The confidence intervals, approximately within β_1 range of $(0.719, 0.819)$ and β_2 range of $(2.068, 2.216)$, do not overlap. The results from QR confirm the presence of homo/heteroscedasticity, yet the estimated standard deviations are notably higher compared to PSLR. Furthermore, the visual representation in Figure 2 reveals a meaningful crossing of the fitted first and third quartile lines, a phenomenon not observed in PSLR. It is noteworthy that unlike PSLR, quantile lines in QR are estimated separately, potentially contributing to the crossing of quartile lines.

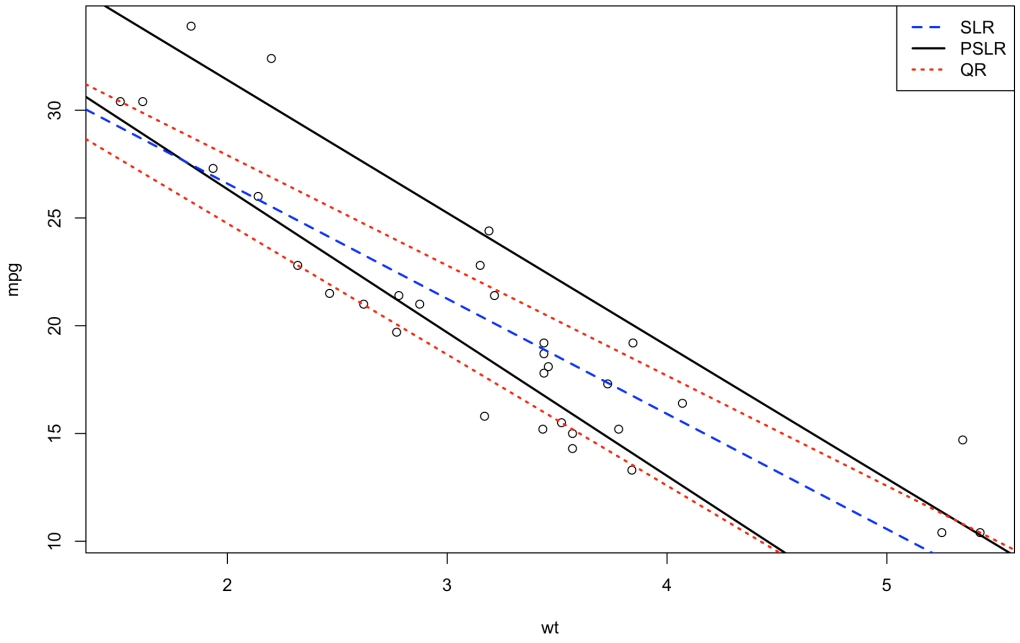

Figure 1. SLR, QR, and PSLR fits to `mtcars` dataset

Table 3. SLR, QR, and PSLR parameter estimations for `trees` dataset

parameter	SLR		QR		PSLR	
α_1	−87.12	(29.27)	−37.49	(17.45)	−36.69	(1.875)
β_1	1.543	(0.384)	0.755	(0.240)	0.769	(0.025)
α_2			−141.00	(43.14)	−117.31	(2.880)
β_2			2.400	(0.601)	2.142	(0.037)

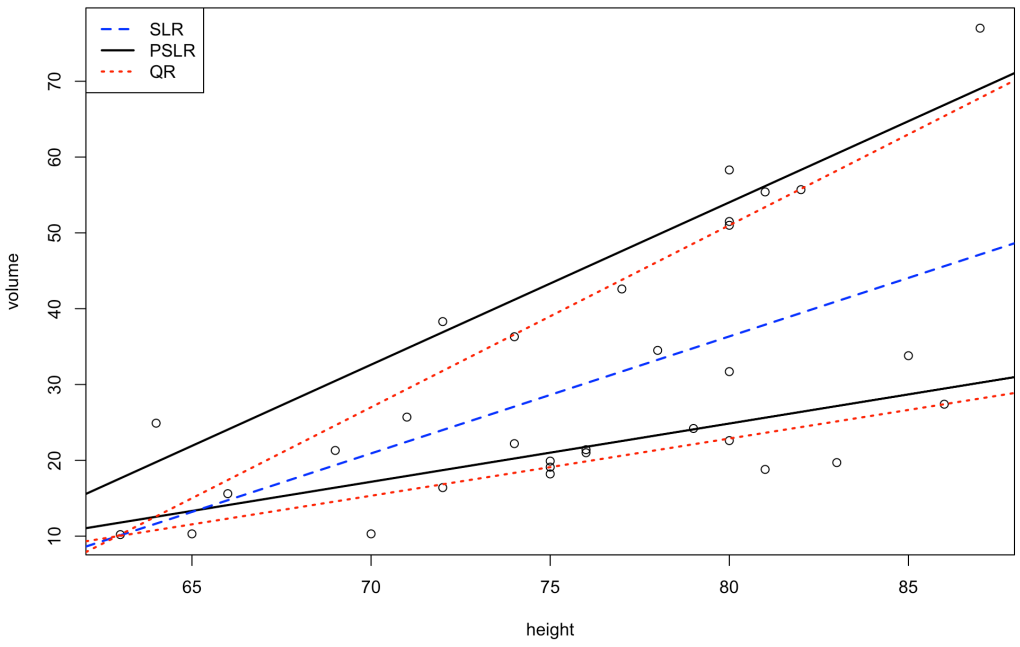

Figure 2. SLR, QR, and PSLR fits to `trees` dataset

4. Simulation study

In this section, we conduct simulations to evaluate the performance of the PSLR method. We considered three different cases, with varying levels of heteroscedasticity, as described in the previous section as follows:

Case (i). $y = 1 + 1.5x + \sqrt{\pi/2}\varepsilon,$

Case (ii). $y = 1 + 1.5x + 0.5\sqrt{\pi/2}x\varepsilon,$

Case (iii). $y = 1 + 0.5\sqrt{\pi/2}x\varepsilon,$

In Case (i), we generate $y = 1 + 1.5x + \sqrt{\pi/2}\varepsilon$, where x is randomly generated from a uniform distribution between 1 and 5 and ε is generated from a normal distribution with mean 0 and standard deviation 1.

In Case (ii), we generate $y = 1 + 1.5x + 0.5\sqrt{\pi/2}x\varepsilon$, where x and ε are generated in the same way as Case (i). In Case (iii), we generate $y = 1 + 0.5\sqrt{\pi/2}x\varepsilon$, where x and ε are generated in the same way as Case (i). Case (i) represents an ordinal linear model without heteroscedasticity, Case (ii) represents a linear model with heteroscedasticity, and Case (iii) represents a model where only the variance is affected by x. In all cases, the conditional two principal points of y given x are linear. The true values of the parameters are shown in Table 4. Figure 3 shows a schematic view of the simulated data in each case with 1000 simulated data points. The blue dashed-line shows the mean line and two solid black lines show the two-principal lines.

Table 4. Table of true values of coefficients for the two-principal lines of the conditional distribution of $y|x$ for cases (i)-(iii)

	$\xi_1(y\|x) = \alpha_1 + \beta_1 x$		$\xi_2(y\|x) = \alpha_2 + \beta_2 x$	
Case	α_1	β_1	α_2	β_2
(i)	0	1.5	2	1.5
(ii)	1	1	1	2
(iii)	1	-0.5	1	0.5

For each case, we generated data with sample sizes of $n = 20$, 50, 100, and 500, and repeated the simulation 10,000 times to compute the distribution and other properties of the estimators. Table 5 shows the mean, standard error, and mean squared error (MSE) of the estimated parameters. The results, summarized in Table 5, show that as the sample size increases, the standard errors and MSE of the estimators decrease. Specifically, in case (i) which has no heteroscedasticity, the MSE and standard errors are significantly smaller compared to cases (ii) and (iii) which have varying levels of heteroscedasticity. Additionally, the results of cases (ii) and (iii) are similar, with similar standard errors and MSEs, especially when the sample size is large. These findings suggest that the PSLR method performs well in estimating the principal points, even in the presence of heteroscedasticity. To evaluate the normality of the estimators, we used R's well-known non-parametric density estimation function, `density`, to plot the density of the estimators and compare them to a normal density with corresponding parameters. Figure 4, Figure 5, and Figure 6 show the results for cases (i), (ii), and (iii), respectively. As the sample size, n, increases, the distribution of the estimators becomes increasingly similar to a normal distribution. This indicates that the estimators are asymptotically normal, as predicted by Theorem 2.1.

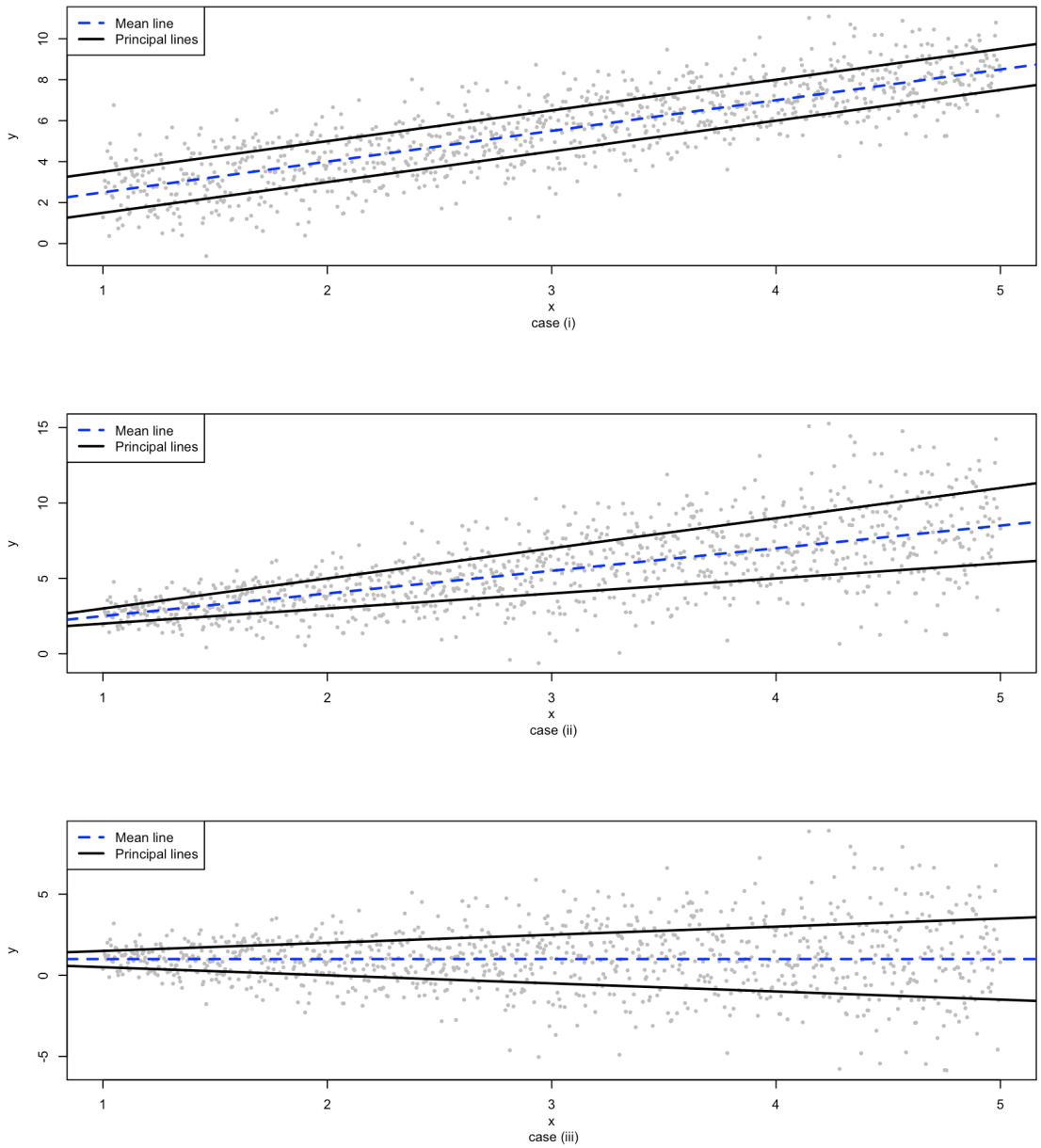

Figure 3. A visual representation of 1000 simulated data points (shown in gray) along with the two-principal lines (solid black lines) and the mean line (dashed blue line) for cases (i)-(iii)

5. Conclusion

In this paper, we introduced PSLR, an innovative extension of traditional SLR. PSLR overcomes SLR's limitations by incorporating the theory of principal points, enabling simultaneous modeling of central, dispersion, and distributional tendencies of a response variable, Y with respect to a single continuous predictor, X. The method was validated through real data analyses on `mtcars` and `trees` datasets, showcasing its superiority over SLR and QR in capturing heteroscedasticity and providing nuanced insights into the Y-X relationship.

Table 5. Comparison of estimation results for cases (i)-(iii) in terms of mean, standard error, and MSE of the estimated parameters

Case	n	α_1			β_1			α_2			β_2		
		mean	se	MSE	mean	se	MSE	mean	se	MSE	mean	se	MSE
(i)	20	0.037	1.188	1.411	1.470	0.38	0.145	1.997	1.205	1.450	1.526	0.393	0.155
	50	-0.040	0.795	0.633	1.500	0.252	0.064	1.978	0.774	0.599	1.519	0.247	0.061
	100	-0.060	0.581	0.341	1.509	0.181	0.033	1.970	0.566	0.320	1.514	0.181	0.033
	500	-0.017	0.276	0.076	1.506	0.088	0.008	1.997	0.272	0.074	1.504	0.086	0.007
(ii)	20	0.985	2.459	6.039	0.962	0.763	0.582	1.118	1.769	3.138	1.990	0.666	0.443
	50	0.934	1.176	1.385	1.001	0.441	0.194	1.048	1.065	1.1360	2.003	0.416	0.173
	100	0.955	0.758	0.576	0.995	0.301	0.090	1.045	0.741	0.550	1.990	0.301	0.091
	500	0.988	0.362	0.131	1.004	0.149	0.022	1.013	0.353	0.125	2	0.145	0.021
(iii)	20	1.021	1.779	3.161	-0.539	0.659	0.436	1.134	1.762	3.121	0.487	0.657	0.432
	50	0.938	1.138	1.297	-0.498	0.428	0.183	1.048	1.055	1.114	0.503	0.409	0.167
	100	0.961	0.745	0.556	-0.505	0.295	0.087	1.042	0.740	0.549	0.492	0.299	0.090
	500	0.984	0.366	0.134	-0.494	0.149	0.022	1.010	0.352	0.124	0.502	0.144	0.021

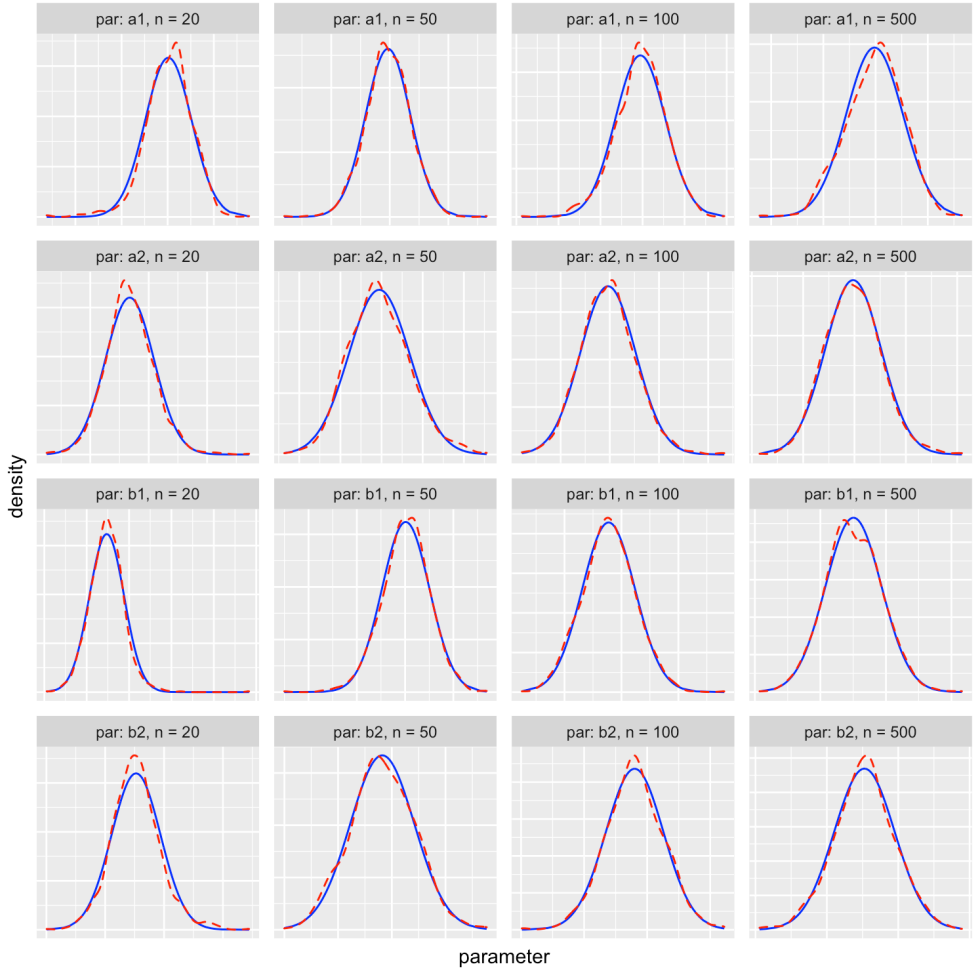

Figure 4. Comparison of non-parametric density estimates and theoretical normal densities for parameter estimations in case (i) for different parameters and sample sizes.

A comprehensive simulation study demonstrated PSLR's robustness, especially in scenarios with varying levels of heteroscedasticity. As sample size increased, the estimators exhibited asymptotic normality, confirming the method's reliability. Overall, PSLR offers

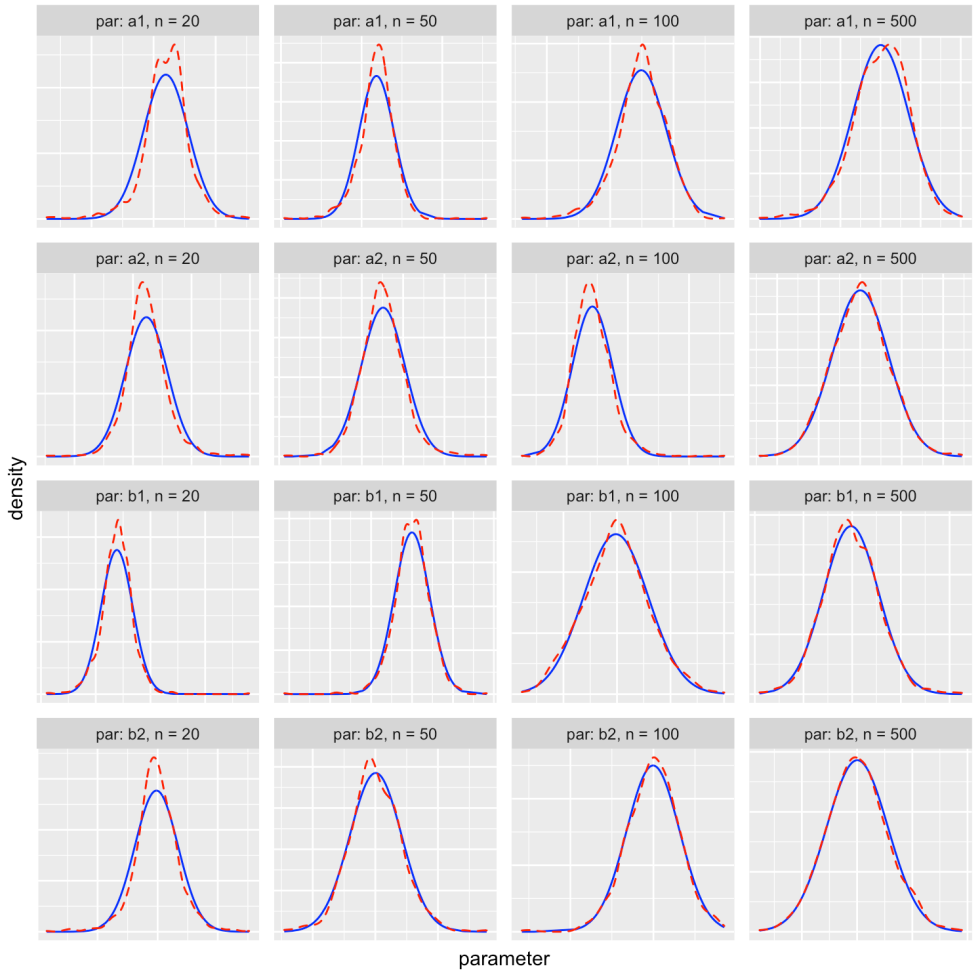

Figure 5. Comparison of non-parametric density estimates and theoretical normal densities for parameter estimations in case (ii) for different parameters and sample sizes.

a flexible and powerful approach for researchers seeking a comprehensive understanding of their data, laying the groundwork for future exploration in multivariate settings and diverse applications.

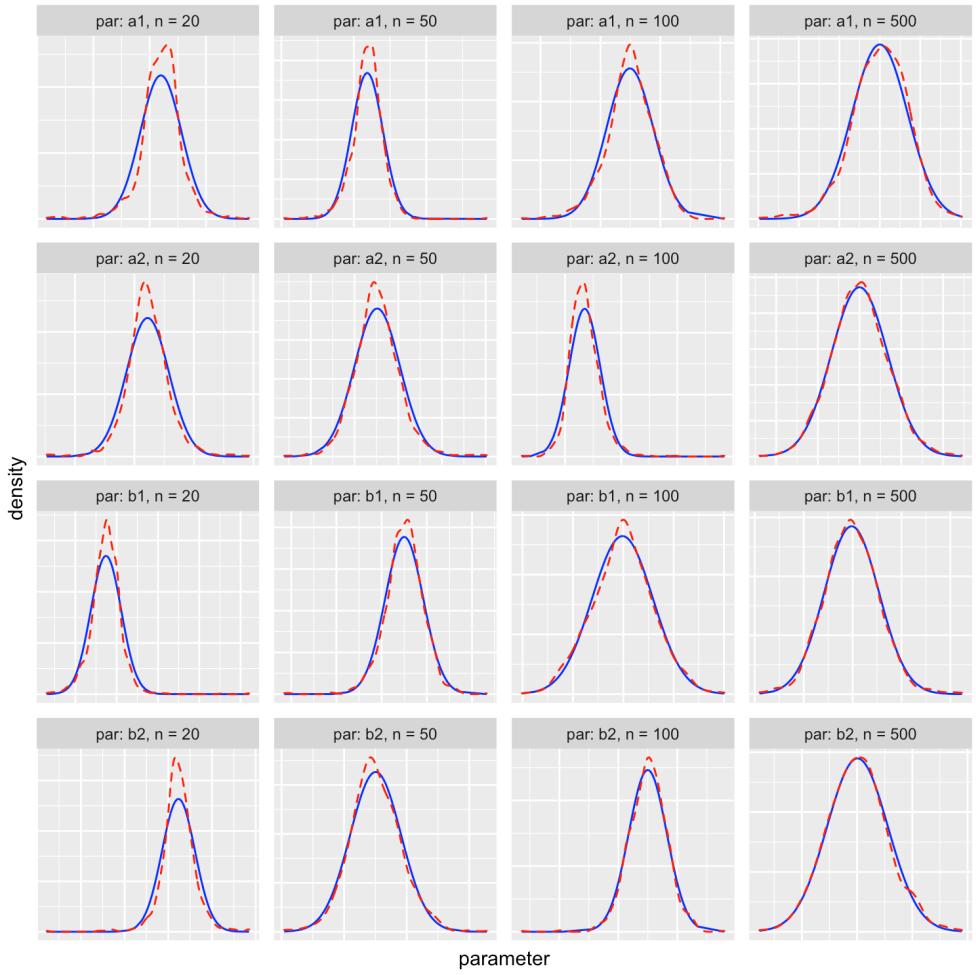

Figure 6. Comparison of non-parametric density estimates and theoretical normal densities for parameter estimations in case (iii) for different parameters and sample sizes.

References

[1] A.C. Atkinson, *Plots, Transformations and Regression*, Oxford University Press, 1985.

[2] J.L. Bali and G. Boente, *Principal points and elliptical distributions from the multivariate setting to the functional case*, Stat. Probab. Lett. **79**, 1858–1865, 2009.

[3] T. Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, J. Econom. **31** (3), 307–327, 1986.

[4] K.G. Brown, *Estimation of variance components using residuals*, J. Am. Stat. Assoc. **73** (361), 141–146, 1978.

[5] R.F. Engle, *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*, Econometrica **50** (4), 987–1007, 1982.

[6] B. Flury, *Estimation of principal points*, J. R. Stat. Soc., C: Appl. Stat. **42** (1), 139–151, 1993.

[7] B. Flury, *Principal points*, Biometrica **77** (1), 33–41, 1990.

[8] H.V. Henderson and P.F. Velleman, *Building multiple regression models interactively*, Biometrics **37**, 391-411, 1981.

[9] J. Kim and H. Kim, *Partitioning of functional gene expression data using principal points*, BMC Bioinformatics **18**, 1–17, 2017.

[10] R. Koenker, *Quantile Regression*, Cambridge University Press, 2005.

[11] S. Matsuura and H. Kurata, *Principal points of a multivariate mixture distribution*, J. Multivar. Anal. **102** (2), 213–224, 2011.

[12] S. Matsuura, H. Kurata and T. Tarpey, *Optimal estimators of principal points for minimizing expected mean squared distance*, J. Stat. Plan. Inference **167**, 102–122, 2015.

[13] T.A. Ryan, B.L. Joiner and B.F. Ryan, *The Minitab Student Handbook*, Duxbury Press, 1976.

[14] E. Stampfer and E. Stadlober, *Methods for estimating principal points*, Commun. Stat. B: Simul. Comput. **31** (2), 261–277, 2002.

[15] T. Tarpey, *Estimating principal points of univariate distributions*, J. Appl. Stat. **24** (5), 499–512, 1997.

[16] T. Tarpey, Principal points, PhD Thesis, ProQuest, 1992.

[17] T. Tarpey, L. Li and B. Flury, *Principal points and self-consistent points of elliptical distributions*, Ann. Stat. **23** (1), 103–112, 1995.

[18] W. Yamamoto and N. Shinozaki, *On uniqueness of two principal points for univariate location mixtures*, Stat. Probab. Lett. **46** (1), 33–42, 2000.

[19] H. Yamashita, A Study on Principal Points for a Multivariate Binary Distribution, PhD Thesis, 2015.

[20] H. Yamashita, S. Matsuura and H. Suzuki, *Comparison of model selection methods for the estimation of principal points for a multivariate binary distribution*, Total Qual. Sci. **1** (1), 22–31, 2015.

[21] F. Yu, *Uniqueness of principal points with respect to p-order distance for a class of univariate continuous distribution*, Stat. Probab. Lett. **183**, 109341, 2022.

[22] A. Zoppe, *Principal points of univariate continuous distributions*, Stat. Comput. **5**, 127–132, 1995.