

DOI: <http://dx.doi.org/10.21123/bsj.2020.17.2.0550>

## Proposing Robust LAD-Atan Penalty of Regression Model Estimation for High Dimensional Data

Ali Hameed Yousif<sup>1</sup>

Omar Abdulmohsin Ali<sup>2\*</sup>

Received 18/8/2019, Accepted 9/12/2019, Published 1/6/2020



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

### Abstract:

The issue of penalized regression model has received considerable critical attention to variable selection. It plays an essential role in dealing with high dimensional data. Arctangent denoted by the Atan penalty has been used in both estimation and variable selection as an efficient method recently. However, the Atan penalty is very sensitive to outliers in response to variables or heavy-tailed error distribution. While the least absolute deviation is a good method to get robustness in regression estimation. The specific objective of this research is to propose a robust Atan estimator from combining these two ideas at once. Simulation experiments and real data applications show that the proposed LAD-Atan estimator has superior performance compared with other estimators.

**Key words:** Atan penalty, High dimensional data, Least absolute deviation, Robust regression, Variable selection.

### Introduction:

Variable selection is one of the issues that have been gaining popularity in recent years as many studies require dealing with high-dimensional data such as sonar, genetics, and others. As known, the presence of many covariates will lead to a very difficult process of model building, especially, in interpretation and with large variance. The variable selection is one of the most important problems in statistics as stated by Brad Efron and that was a single problem, which is merely variable selection associated with regression (1). That problem consists of choosing variables from a lot of candidate variables, estimating parameters for those variables, and then make the rest of inferences. The ordinary least square method and other traditional methods cannot deal with these problems. Penalized regression has been widely used with models including large explanatory variables that improve predictive accuracy as well as the selection of important variables in the model. This method is based on the minimization of the objective function which is formed of two parts: the first is a loss function, and the second is a penalty function based on the  $\lambda$  penalty parameter.

While the penalty function makes a tradeoff between bias and variance as ridge regression (2), the latest does not exclude undesirable variable because its coefficient is not set to zero.

Tibshirani (3) proposed the least absolute shrinkage and selection operator: LASSO method based on the principle of estimation and selection of variables in the simultaneous approach. LASSO has a fascinating advantage that it makes the value of some regression coefficients  $\beta$  which are non-significant to be zero. Detail discussion interested with LASSO and its consequences can be found in (4). Some generalizations of LASSO can be found in (5). Recent contributions have been made by Kim (6) and Uraibi (7).

Fan and Li (8) suggest smoothly clipped absolute deviation: SCAD penalty function to keep the continuity of the penalty function at threshold points and to smooth it to be as the effect of quadratic splines at node points ( $\lambda$ ) and ( $a\lambda$ ). Useful medical application has been introduced by Fang et al. (9). However, Fang et al. listed some other properties of SCAD penalized regression which can be summarized in the target function to be a high-dimensional non-concave function, singular at the origin and does not have continuous second-order derivatives.

Zhang (10) suggested nearly unbiased variable selection under the minimax concave penalty (MCP) penalty function. Wang et al. (11)

<sup>1</sup>Department of Statistics, College of Administration and Economics, Wasit University, Kut, Iraq.

[ahameed@uowasit.edu.iq](mailto:ahameed@uowasit.edu.iq)

<sup>2</sup>Department of Statistics, College of Administration and Economics, University of Baghdad, Baghdad, Iraq.

\*Corresponding author: [dromar72@coadec.uobaghdad.edu.iq](mailto:dromar72@coadec.uobaghdad.edu.iq)

\*ORCID ID: <https://orcid.org/0000-0003-0274-9325>

introduced a wide comparison among LASSO, SCAD and MCP methods and then took their distribution-transformation under consideration to be applied. Wang and Zhu (12) proposed an arctangent type penalty which very closely resembles  $l_o$  penalty call it Atan penalty

**Atan Penalized Least Square Method:**

Suppose the following linear model is considered:

$$y_i = x_i' \beta + \epsilon_i, \quad i = 1, 2, 3, \dots, n \quad \dots (1)$$

Where  $x_i = (x_{i1}, \dots, x_{ip})'$  is  $p$ -dimensional covariate,  $\beta = (\beta_1, \dots, \beta_p)$ , is an unknown parameter vector,  $p$  is the number of covariates,  $n$  is the size of sample,  $y = (y_1, y_2, \dots, y_n)'$  is  $(n \times 1)$  response vector, and  $\epsilon_i$  is random error vector where  $E(\epsilon_i), var(\epsilon_i) = \sigma^2$ .

Then, the penalized least square method can be got from the solution of the following equation:

$$\min \left\{ \frac{1}{2n} \|y - X\beta\|^2 + \sum_{j=1}^p P_\lambda(|\beta_j|) \right\} \quad \dots (2)$$

Where  $X = (x_1, \dots, x_n)$  is  $(n \times p)$  design matrix.  $\|\cdot\|$  represents  $L_2$ - norm and  $P_\lambda(\cdot)$  is the penalty function that based on penalty parameter  $\lambda > 0$ .

There are many penalties used in penalized least square as SCAD penalty and has three merits in keeping continuity in data, unbiasedness in parameter estimator, and sparsity for small parameter estimators to put them as zero and defined by continuously differentiable function described as follow:

$$P_\lambda(|\beta|) = \lambda \left[ I(|\beta| \leq \lambda) + \frac{(a\lambda - \beta)_+}{(a-1)} I(|\beta| > \lambda) \right] \quad \dots (3)$$

Where assuming the value of  $(a = 3.7)$ .

Zhang (10) proposed concave penalty function called: (MCP) which is defined by:

$$P_\lambda(|\beta|) = \frac{(\gamma\lambda - |\beta|)_+}{\gamma} \quad \dots (4)$$

With  $\lambda \geq 0$  and  $\gamma > 1$ .

Wang and Zhu (12) proposed an arctangent type penalty which very closely resembles  $l_o$  penalty call it Atan penalty which defined as follow:

$$P_{\lambda, \gamma}(|\beta|) = \lambda \left( \gamma + \frac{2}{\pi} \right) \arctan \left( \frac{|\beta|}{\gamma} \right) \quad \dots (5)$$

With  $\lambda \geq 0$  and  $\gamma > 0$ .

The first derivative of the Atan penalty defined as follow:

$$P'_{\lambda, \gamma}(|\beta|) = \lambda \frac{\gamma(\gamma + 2/\pi)}{\gamma^2 + \beta^2} \quad \dots (6)$$

Where the optimal value of  $(\gamma = 0.005)$

**Robust LAD-Atan Regression:**

Atan estimator is not robust, which means it is very sensitive to the presence of outlying observation, to deal with this problem, a robust loss function can be used to get the robust Atan method. In this research, it will be proposed the combining the Arctangent (Atan) penalty function with the LAD absolute loss function.

$$\min \left\{ \sum_{i=1}^n |y_i - x_i' \beta| + \lambda \sum_{j=1}^p \left( \gamma + \frac{2}{\pi} \right) \arctan \left( \frac{|\beta_j|}{\gamma} \right) \right\} \quad \dots (7)$$

Atan estimator can be improved by combining it with a LAD part to get LAD-Atan estimator in a similar method to that used in illustrating LAD proposed by Wang et al. (13) and as follow:

Suppose the augment dataset  $\{(y_i^*, x_i^*)\}$  and

$$(y_i^*, x_i^*) = \begin{cases} (y_i, x_i) & \text{for } i = 1, \dots, n \\ (0, n\lambda_j e_j) & \text{for } i = n + 1, \dots, n + p \end{cases} \quad \dots (8)$$

Where is  $e_j$  the unit vector with the  $j^{\text{th}}$  elements equal to one and all other equal to zero. Then the LAD-Atan estimator can be obtained by minimizing

$$\sum_{i=1}^{n+p} |y_i^* - x_i^* \beta| \quad \dots (9)$$

It can be made the benefit of (rq function in quantreg package of R) in (14) to get the LAD-Atan estimator by easy.

**Theoretical Properties:**

To identify the regression model correctly, the LAD-Atan estimator has to satisfy the following properties which are combined from (LAD) part as consistency, sparsity properties according to Wang et al. (13). Furthermore, (Atan) part satisfied regularity and oracle properties according to Wang and Zhu (12).

- **$\sqrt{n}$  - Consistency:**

Suppose that  $\beta = (\beta_a, \beta_b)$  and  $\beta_a = (\beta_1, \dots, \beta_{p_0})$  and  $\beta_b = (\beta_{0+1}, \dots, \beta_p)$

Let us define:  $a_n = \max \{ \lambda_j, 1 \leq j \leq p_0 \}$  and

- $b_n = \min \{ \lambda_j, p_0 \leq j \leq p \}$

where  $\lambda_j$  is some function in terms  $(n)$ .

$\sqrt{n}$  - Consistent if:  $\sqrt{n} a_n \rightarrow 0$  which

requires the following assumptions to be satisfied:

- The error term  $\epsilon_i$  behaves as a continuous probability distribution at the origin.
- The covariance matrix  $cov(x_i) = \Sigma$  exists and positive definite.

- Sparsity:**  
 According to the previous consistency condition, another condition will be satisfied which is  $\sqrt{nb_n} \rightarrow \infty$  then estimator must verify  $\hat{\beta}_b$  to equal zero with probability tends to one.
- Regularity Conditions:**  
 It is important to apply the following conditions on the proposed penalty:
  - $n \rightarrow \infty$  and  $p\sigma^2/n \rightarrow 0$ , where  $n$  is the sample size and  $p$  is the number of regression parameters.
  - $\rho\sqrt{n/(p\sigma^2)} \rightarrow \infty$  where  $\rho = \min_{j \in A} |\beta_j^*|$
  - $\lambda = O(1)$ ,  $\lambda\sqrt{n/(p\sigma^2)} \rightarrow \infty$  as  $n \rightarrow \infty$ , and  $\gamma = O(p^{1/2}\sigma^3n^{-3/2})$
  - Identifiability conditions:  

$$\lim_{n \rightarrow \infty} n^{-1} \max_{1 \leq i \leq n} \sum_{j=1}^p x_{ij}^2 = 0$$

$$E\left(\varepsilon_i / \sigma\right)^{2+\delta} < M \text{ for some } \delta \text{ and } M < \infty$$
- Oracle property:**  
 If the estimator is consistent for  $\sqrt{na_n} \rightarrow 0$  and has sparsity  $\sqrt{nb_n} \rightarrow \infty$  of, then this estimator must verify that:  $\Pr\{\hat{\beta}_b = 0\} \rightarrow 1$  and  

$$\sqrt{n}(\hat{\beta}'_a - \beta'_a) \rightarrow N(0, 0.25\Sigma_0^{-1} / f^2(0)),$$
 where  $\Sigma_0 = \text{cov}(x_{ia})$ , and  $f(t)$  is the probability distribution of  $\varepsilon_i$ .

**Selection of Penalty Parameter:**

The process of selecting the penalty parameter is one of the important steps in the penalized regression, it controls the amount of transaction reduction and the selection of the sub-variables included in the final model. Tibshirani (3) and Fan and Li (8) used generalized cross validation (GCV) to select penalty parameters. In this research the suggestion to use generalized cross validation to select the penalty parameter is introduced through minimizing the following formula:

$$GCV(\lambda) = \frac{\sum_{i=1}^n |y_i - \hat{x}_i \beta|}{\left(1 - \frac{df(\lambda)}{n}\right)^2} \quad \dots (10)$$

Where  $df(\lambda)$  represent the degree of freedom,  
 $df(\lambda) = \text{tr}[X(X'X + nQ)^{-1}X']$   
 It represents the number of non-zero estimated parameters.

$$Q = \text{diag} \left( \frac{p'_{\lambda(|\hat{\beta}_1|)}}{|\hat{\beta}_1|}, \frac{p'_{\lambda(|\hat{\beta}_2|)}}{|\hat{\beta}_2|}, \dots, \frac{p'_{\lambda(|\hat{\beta}_p|)}}{|\hat{\beta}_p|} \right)$$

A penalty parameter  $\lambda$  is selected that makes minimizing this GCV.

**Simulation Experiments:**

In this section, simulation experiments were used to show the performance of the proposed estimator. The proposed LAD-Atan estimator was compared with each of the following estimators: SCAD and MCP estimators based on (15). While the Atan penalty was based on (12).

Simulation experiments were conducted by R program with package simFrame introduced by Koenker (16), and 200 replicates. Data generating were represented by the following linear regression model:

$$y = x'_i \beta + \varepsilon \quad \dots (11)$$

Where  $\beta = \{3, 1.5, 0, 0, 2, 0, \dots, 0\}$

Considering error distributions are standard normal distribution and t-distribution with three degrees of freedom,  $x \sim N_p(0, \Sigma)$ , with covariance matrix  $\Sigma_{ij} = \rho^{|i-j|}$ ,  $\rho = 0.5$ .

Taking the contaminated data with vertical outliers under consideration with the proportion of (0 %, 5%, and 10%) and the experiments as two kinds

**First Experiment:** Setting number of covariates ( $p=8$ ), sample sizes ( $n=60$  and  $100$ ), and standard deviations ( $\sigma = 1$  and  $3$ ) respectively.

**Second Experiment:** Setting number of covariates ( $p=50$ ), sample sizes ( $n=60$  and  $100$ ), and standard deviations ( $\sigma = 1$  and  $3$ ) respectively.

Measuring of performance to compression between the estimators via mean square error (MSE) criterion which represented as followed:

$$MSE(\hat{\beta}) = \frac{1}{p} \sum_{j=1}^p (\hat{\beta}_j - \beta_j)^2 \quad \dots (12)$$

Also, computing false positive rate (FPR) represents the hypothesized ( $\beta$ ) equal zero in the actual model, while it is estimated one as nonzero. Conversely, false-negative rate (FNR) represents the hypothesized ( $\beta$ ) nonzero in the actual model, while it is estimated as zero. This definition according to Alfons et al. (17) as the following formulas:

$$FPR(\hat{\beta}) = \frac{|\{j \in \{1, \dots, p\}: \hat{\beta}_j \neq 0 \cap \beta_j = 0\}|}{|\{j \in \{1, \dots, p\}: \beta_j = 0\}|} \quad (13)$$

$$FNR(\hat{\beta}) = \frac{|\{j \in \{1, \dots, p\}: \hat{\beta}_j = 0 \cap \beta_j \neq 0\}|}{|\{j \in \{1, \dots, p\}: \beta_j \neq 0\}|} \quad (14)$$

**Table 1. Simulation results for the first experiment with standard normal error**

n, $\sigma$	Method	No contamination			Vertical outliers 5%			Vertical outliers 10%		
		MSE	FPR	FNR	MSE	FPR	FNR	MSE	FPR	FNR
n=60 $\sigma = 1$	SCAD	1.8735	0.36	0.03	2.5939	0.44	0.15	2.9531	0.52	0.17
	MCP	1.8278	0.37	0.02	2.5155	0.46	0.11	2.8946	0.53	0.14
	Atan	1.7377	0.33	0.00	2.4319	0.38	0.03	2.8548	0.50	0.05
	LAD-Atan	1.6996	0.26	0.00	1.7068	0.27	0.00	1.7048	0.25	0.00
n=100 $\sigma=3$	SCAD	0.3094	0.36	0.12	1.0218	0.58	0.26	1.8344	0.65	0.27
	MCP	0.2021	0.39	0.09	0.9943	0.63	0.23	1.8357	0.73	0.23
	Atan	0.1205	0.36	0.02	0.9636	0.59	0.07	1.8394	0.66	0.08
	LAD-Atan	0.0711	0.35	0.01	0.1101	0.31	0.06	0.1357	0.26	0.06

From Table 1, the results of the simulation for standard normal error with no contamination, Atan and LAD-Atan estimators are the best because of their smallest MSE, FPR, and FNR respectively

and for both sample sizes  $n = 60, 100$ . For vertical outliers, the LAD-Atan estimator shows superiority according to its smallest MSE, FPR, and FNR and for both sample sizes  $n= 60, 100$ .

**Table 2. Simulation results for the first experiment with t-distribution error**

n, $\sigma$	Method	No contamination			Vertical outliers 5%			Vertical outliers 10%		
		MSE	FPR	FNR	MSE	FPR	FNR	MSE	FPR	FNR
n=60 $\sigma = 1$	SCAD	1.6537	0.23	0.00	2.5794	0.47	0.13	3.0154	0.47	0.17
	MCP	1.6484	0.22	0.00	2.4664	0.53	0.10	2.9488	0.51	0.15
	Atan	1.6338	0.20	0.00	2.3807	0.72	0.03	2.9543	0.78	0.06
	LAD-Atan	1.6550	0.21	0.00	1.6651	0.25	0.00	1.6848	0.28	0.00
n=100 $\sigma=3$	SCAD	0.1146	0.32	0.06	0.9726	0.61	0.23	1.7860	0.67	0.29
	MCP	0.0703	0.33	0.04	0.9194	0.68	0.17	1.8305	0.72	0.19
	Atan	0.0342	0.28	0.00	0.8811	0.55	0.06	1.77329	0.68	0.07
	LAD-Atan	0.0546	0.35	0.00	0.0764	0.33	0.01	0.1190	0.27	0.02

From Table 2, the results of the simulation for t-distribution error with no contamination shows that Atan and LAD-Atan estimators are the best results because of their smallest MSE, FPR, and FNR respectively. While for vertical outliers, the LAD-Atan estimator shows superiority results because of its smallest criteria: MSE, FPR, and FNR respectively.

**Real Data Application:**

Real data ( $n = 30$ ) were collected to detect the chemical properties of the soil and studying their effect on date palm crops in Iraq. The application focuses on investigating the effect of 16 variables which were: Potential of Hydrogen (pH), Organic Matter (OM), Electrical Conductivity (EC), Lime, Gypsum, Cation Exchange Capacity (CEC), Calcium (Ca), Magnesium (Mg), Potassium (K), Sodium (Na), Chlorine (Cl), Sulphate ( $SO_4$ ), Hydrocarbons ( $HCO_3$ ), Phosphorus (P), Exchangeable Sodium Ratio (ESR), Exchange Sodium Percentage (ESP) as covariates. While the response variable was represented by the date palm crop after taking its logarithmic (log). The results of the practical application were as in the following table below.

**Table 3. Nonzero coefficients for the log(palm crop)**

Predictor	SCAD	MCP	Atan	LAD-Atan
pH	0	0	0.72412	0.75309
OM	0	0	0	0
EC	-0.03431	-0.09274	-0.05188	0
Lime	0	0	0	0
Gypsum	0	0	0	0
CEC	0	0	0	0
Ca	0	0	0	0
Mg	0	0	0	0
K	0	-0.01470	0	0
Na	0	0	0	0
Cl	0.01586	0.03322	0.02862	0
$SO_4$	-0.00053	0	0	0
$HCO_3$	0	-0.26780	-0.26618	-0.44225
P	0.46026	0	0.99188	0.86162
ESR	0	0	0	0
ESB	0	0	0	0

From Table 3 above, it shows that SCAD method selected 4 variables (EC, Cl,  $SO_4$ , P), MCP method selected 4 variables (EC, K, Cl,  $HCO_3$ ), Atan method selected 4 variables (PH, EC, Cl,  $HCO_3$ , P), and the proposed robust method (LAD-Atan) selected 3 variables (PH,  $HCO_3$ , P) respectively.

To compare the estimators, the mean square error was used and the results were obtained as follow:

**Table 4. Mean square error for every method**

Method	MSE	Variable Selection
SCAD	0.1737	4
MCP	0.1688	4
Atan	0.1315	4
LAD-Atan	0.0198	3

From Table 4 it easy to notice that the LAD-Atan is the best estimator which combining estimation robustness (indicated by smaller MSE), in addition to being minimum selection variables (indicated by three variables only).

### Conclusions:

Despite the fact that the LAD part is used in obtaining a robust estimator, it is not suitable to select the variable. Therefore, this problem is solved by combining the Atan penalty to the objective function. It is shown that from the simulation experiments, the LAD-Atan estimator gives the best performance for both estimation and variable selection rather than other methods. It can be noticed that LAD-Atan approves its robustness as long as increasing the contamination proportions for t-distribution, while it has good competition sometimes with Atan and MCP for normal distribution. Furthermore, the application results coincide with simulation results in showing the superiority of the proposed robust LAD-Atan estimator where are outliers in real data under consideration. The proposed method has more reduction in variable selection associated with more accuracy than the other three methods.

### Acknowledgment:

The authors would like to thank the editor and anonymous reviewers of Baghdad Science Journal for their constructive comments and suggestions, which have greatly improved this paper.

### Authors' declaration:

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are mine ours. Besides, the Figures and images, which are not mine ours, have been given the permission for re-publication attached with the manuscript.
- Ethical Clearance: The project was approved by the local ethical committee in University of Baghdad.

### References:

1. Hesterberg T., Choi N., Meier L., Fraley C., Least angle and  $\ell_1$  penalized regression: A review. *Statistics Surveys*, 2008; 2: 61 - 93.
2. Hoerl E, Kennard RW. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 1970; 12(1): 55–67.
3. Tibshirani R., Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B: Methodological*, 1996; 58(1): 267-288.
4. Tibshirani R., Least angle regression. *The Annals of Statistics*, 2004; 32(2): 407–499.
5. Tibshirani R., Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society, Series B: Statistical Methodology (J.R.Statist.Soc.B)*, 2011; 73(3): 273-282.
6. Kim J., Ryu J., Lee S., Jun S., Penalized regression models for patent keyword analysis. *Model Assisted Statistics and Applications (MASA)*, 2017; 12(1): 239–24.
7. Uraibi H. S., Weighted Lasso subsampling for high dimensional regression. *Electronic Journal of Applied Statistical Analysis (EJASA)*, 2019; 12(01): 69-84.
8. Fan J., Li R., Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association (JASA)*, 2001; 96(456): 1348-1360.
9. Fang Y., Kong Y., Jiao Y., Optimization Techniques for SCAD Variable Selection in Medical Research. *Biomedical Journal of Scientific and Technical Research (BJSTR)*, 2018; 8(2): 1-2.
10. Zhang C.M., Nearly Unbiased Variables Selection Under Minimax Concave Penalty. *The Annals of Statistics*, 2010; 38(2): 894-942.
11. Wang T., Xu P., Zhu L., Non-convex penalized estimation in high-dimensional models with single-index structure. Elsevier Inc., *Journal of Multivariate Analysis (JMVA)*, 2012; 109: 221-235.
12. Wang Y., Zhu L., Variable Selection and Parameter Estimation with the Atan Regularization Method. Hindawi Publishing Corporation, *Journal of Probability and Statistics*, 2016; Article ID 6495417: 1-12.
13. Wang H., Li G., Jiang G., Robust Regression Shrinkage and Consistent Variable Selection through the LAD-Lasso. *Journal of Business & Economic Statistics (JBES)*, 2007; 25(3): 347-355.
14. Koenker R., *Quantreg: Quantile regression*. R package; 2019.
15. Breheny P., Huang J., Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *The Annals of Applied Statistics*, 2011; 5(1): 232–253.
16. Alfons A., *SimFrame: Simulation framework*. R package; 2015.
17. Alfons A., Croux C., Gelper S., Sparse Least Trimmed Squares Regression for Analyzing High-Dimensional Large Data Sets. *The Annals of Applied Statistics*, 2013; 7(1): 226–248.

## اقتراح تقدير الجزاء الحصين LAD-Atan لنموذج انحدار بيانات عالية البُعدية

عمر عبدالمحسن علي<sup>2</sup>

علي حميد يوسف<sup>1</sup>

<sup>1</sup> قسم الاحصاء، كلية الادارة والاقتصاد، جامعة واسط، الكوت، العراق.  
<sup>2</sup> قسم الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد، بغداد، العراق.

### الخلاصة:

لاقت قضية نموذج الانحدار اهتمامًا بالغ الأهمية لاختيار المتغيرات، إذ انه يؤدي دورًا أساسيًا في التعامل مع البيانات ذات الابعاد العالية. يتم استخدام معكوس الظل الذي يشير إليه دالة جزاء Atan في كل من التقدير والاختيار المتغير كطريقة فعالة. ومع ذلك ، فإن دالة الجزاء Atan حساسة جدًا للقيم الشاذة لمتغيرات الاستجابة أو توزيع ملتوي للأخطاء أو توزيع ذو ذيل ثقيل. بينما LAD هي وسيلة جيدة للحصول على حصانة تقدير الانحدار. ان الهدف الاساس من هذا البحث هو اقتراح مُقدّر Atan يجمع بين هاتين الفكرتين في أن واحد. لقد اظهرت تجارب المحاكاة وتطبيق البيانات الحقيقية أن مقدّر LAD-Atan المقترح هو الافضل مقارنة بالمقدرات الأخرى.

**الكلمات المفتاحية:** اختيار المتغير، انحدار حصين، بيانات عالية البُعدية، جزاء Atan، مطلق الفروق الاصغر.