Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$: 2312-9883

QJAE, Volume 24, Issue 3 (2022) ISSN$_{Print}$: 1816-9171

Manuscript ID:

A proposed Penalty Method for Processing High-Dimensional data in a Factorial Experiment (2^4) with two Replication.

Wisam wadullah saleem **Dr.Asmaa GHalib Jaber**

wisam-stat@uomosul.edu.iq drasmaa.ghalib@coadec.uobaghdad.edu.iq

College of Computer Science and Mathematics / Department of Statistics and Informatics / University of Mosul

Abstract: In this paper, a method is proposed to model the mathematical model of a factorial experiment (2^4) applied in a complete random design suffering from high-dimensional data to a multiple regression model using the general linear model where the matrix (X) contains (81) columns and (32) rows, The columns represent the effect of the levels of each factor and the effect of the interaction between the levels of these factors in addition to the effect of the general arithmetic mean, and the rows represent the number of experiment observations (sample size), then the selection of factors and interactions of important factors using the penalty methods (Lasso, adaptive lasso) and the proposed method on the lasso adaptive weight It was compared with the methods mentioned . A simulation study are conducted to investigate the performance of the proposed method. We employed the mean square errors, R-square, Adjusted R-square and Mean absolute deviation to select the best model and the results show that the proposed method using the lasso adaptive weight variable selection performs well.

Key words: Two- Level Factorial Design, Regression Experiment Design, Variable Selection, Interaction, Lasso, Adaptive lasso.

INTRODUCTION:

In recent years, interest in the topic of high-dimensional data analysis has increased, especially in which the number of explanatory variables is greater than the sample size, as it has attracted many researchers and within various fields, including applied mathematics, electronic engineering, genetics and others.

When designing experiments, the problem of high dimensionality of data in the matrix (X) arises after converting the mathematical model of the design into a multiple regression model through the use of the general linear model by transforming the influence of factor levels and the influence of interactions between the levels of these factors into independent variables, so that the number of columns (The effect of factor levels and the effect of interactions between the levels of these factors) is greater than the number of rows (sample size), which means that ($p > n$) for the experiment. so classical methods cannot be used to estimate parameters. Penal methods were used and a penalty method was proposed on the adaptive lasso weights through five new values of Gamma (1, 1.5, 3, 5, 10) as follows:

Study Problem:

The problem of high dimensionality of the data affect the estimation of parameters, the statement of important parameters and the estimation of error in regression analysis. As for the design of experiments, the problem of high dimensions of data in a matrix (X) arises after transforming the mathematical model of the design into a multiple regression model through the use of the general linear model by transforming the levels of factors and the interactions between the levels of these factors into explanatory variables, so that the number of columns (factor levels and interactions between the levels of these factors) is greater than the number of rows (sample size), meaning that (p>n) for the factorial experiment applied in the complete random design, where a new approach was followed that combines experiment design and regression analysis and through use of punitive methods the most important factors are assessed and selected.

Study Objectives:

The study objectives could be summarized in the light of the following:

1. The research aims to address the problem of high-dimensional data in the design matrix of an applied factor experiment using a complete random design after transforming the mathematical model for the design of experiments into a regression model by representing the levels of factors and the interaction between the levels of these factors into explanatory variables and choosing the important parameters using penal methods.
2. Research the effect of the high-dimensional data problem on the design matrix in addition to its effect according to the model size and sample size on the results.
3. Research the effect of the proposed penalty method on the results of selecting the important factors of the factorial experiment and comparing them with the penalty methods used in order to find out the optimal method through comparison criteria for selecting variables and the standard prediction accuracy using the simulation method and applying real data as well.

Study significance and contributions:

1. The importance of the study lies in the processing of high-dimensional data in factor experiments (2^4) by converting the mathematical model of the factorial experiment into a regression model.
2. Suggest a regression model consisting of the levels of factors and the interaction between the levels of these factors.
3. Statement of the importance of each level of the factors and the importance of the interaction between the levels of factors.

1. Factorial Experiments

These are experiments in which the interest is to study the effect of two or more factors in one experiment, using all possible combinations between many different levels of the factors to be studied. These experiments are used to study the main effects of each factor individually, as well as study the interaction at the same time. Al-Rawi, [1].

2. Four-factor Interaction Model:

In factorial experiments in which each factor has only two levels, the number of treatments will be equal to 2^n. If factor (A) has two levels, Factor (B) has two levels, Factor C has two levels, and Factor D has two levels. So, the number of processors used in the experiment ($2^4 = 16$). Mohammed, B. K [9]. ; Montgomery, D.C. [11].

mathematical model of the factorial experiment is:

$$y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_m + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\delta)_{im} + (\beta\gamma)_{jk} + (\beta\delta)_{jm} + (\gamma\delta)_{km} + (\alpha\beta\gamma)_{ijk} + (\alpha\beta\delta)_{ijm} + (\alpha\gamma\delta)_{ikm} + (\beta\gamma\delta)_{jkm} + (\alpha\beta\gamma\delta)_{ijkm} + \varepsilon_{ijkml} \quad \ldots (1)$$

$i = 1,2 \quad , \quad j = 1,2 \quad , \quad k = 1,2 \quad , \quad m = 1,2$

3. Least Absolute Shrinkage and Selection Operator (LASSO)

The most frequently employed penalty term is ($L1{-}norm$) which is named as least absolute shrinkage and selection operator, abbreviated as LASSO. Tibshirani (1996) proposed LASSO by imposing ($L1{-}nor$) to the residual sum of squares. LASSO gets its popularity and became as a basic for other penalized methods. This is because of its ability in performing both continuous shrinkage and variable selection simultaneously. It can exactly set the regression coefficients to zero. Tibshirani, [14].

The regression coefficients of the penalized linear regression using LASSO are obtained as:

$$PLR(\beta.\lambda)^{LASSO} = (y - X\beta)^T(y - X\beta) + \lambda \sum_{j=1}^{p} |\beta_j| \quad \ldots (2)$$

where the penalty term $\sum_{j=1}^{p} |\beta_j|$ represents the ($L1{-}norm$) and (λ) is a positive tuning parameter. As $\lambda \rightarrow 0$, the OLS estimates are obtained and for (λ) sufficiently increases, more coefficient estimates are shrunk to be equal zero.

The regression coefficients are obtained from the partial linear regression using Lasso as follows:

$$\hat{\beta}_{PLR}^{LASSO} = \operatorname{argmin}_{\beta} \left\{ \sum_{i=1}^{n} (y - X\hat{\beta})^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \quad \ldots (3)$$

4. Adaptive LASSO (ALASSO)

The adaptive LASSO (ALASSO) was proposed by Zou (2006) [17]. to alleviate the consistency problems of the LASSO, in which a data-driven weight is used with ($L1{-}norm$). Simply, ALASSO replaces the ($L1{-}norm$) by a weighted ($L1{-}norm$). The fundamental idea behind ALASSO is that, by allowing a relatively large amount of shrinkage for the zero coefficients and small amount for the nonzero coefficients, it may be possible to reduce the bias of the estimation and improve the variable selection consistency.

The penalized linear regression using ALASSO is defined by

$$PLR(\beta.\lambda)^{ALASSO} = (y - X\beta)^T(y - X\beta) + \lambda \sum_{j=1}^{p} w_j |\beta_j| \quad \ldots (4)$$

where $w = (w_j; j = 1.2. \ldots . p y_{1 \ldots\ldots} y_n)^T \in R^p$ is a given data-driven weight vector and it is obtained by

$$\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-\gamma}. j = 1.2. \dots . p \quad \dots (5)$$

where (γ) is a positive constant and$(\hat{\beta}_j^{initial})$ initial is an initial value, which is represented as an estimator for (β) obtained, for example, from OLS estimates or ridge regression estimates are considered as initial values. For fixed value of (n) and larger values of (λ) in Eq. (5) lead to higher bias and less variability in predicted values Hui et al, [6].

Further theoretical studies of ALASSO are performed, when $p > n$, and several initial weights are proposed. Huang et al, [5]. studied the asymptotic properties of the ALASSO estimators, where the initial weight is chosen by the marginal regression estimator. The distribution of the ALASSO estimator was studied by Potscher and Schneider, [13]. In addition, the LASSO estimator itself can be used as an initial estimator and Van De Geer, [3]. Moreover, Wang and Wang, [15]. proposed to use the maximum marginal likelihood estimator as an initial weight in high dimensional GLM.

The ALASSO estimator is defined as follows:

$$\hat{\beta}_{PLR}^{ALASSO} = \mathrm{argmin}_{\beta}\left\{(y - X\beta)^T(y - X\beta) + \lambda \sum_{j=1}^{p} \hat{w}_j|\beta_j|\right\} \quad \dots (6)$$

5. Proposed Initial Weight in Adaptive LASSO.

Choosing the initial weight is very important because it plays a vital role in the consistency of adaptive penalized likelihood methods. When $p < n$, Zou, [17]. Proved that ALASSO achieves oracle properties by suggesting to use the OLS estimator as an initial weight, i.e, $\hat{w}_j = \left[abs(\hat{\beta}_j^{OLS})\right]^{-\gamma}$. Similarity, when GLM is used, the initial weight be $\hat{w}_j = \left[abs(\hat{\beta}_j^{MLE})\right]^{-\gamma}$. However, both $\hat{\beta}^{OLS}$ and $\hat{\beta}^{MLE}$ are neither correct in using when there is high correlation among the explanatory variables nor available when $p > n$.

As an remedy from these flaws, LASSO estimator in Eq. (5) was used as an initial weight, that is , $\hat{w}_j = \left[abs(\hat{\beta}_j^{OLS})\right]^{-\gamma}$, Buhlmann and Van De Geer [3]. ; Lian, [7]. In fact, using LASSO estimator as initial weight in ALASSO may be not appropriate in either $p < n$ or $p > n$. This is because LASSO suffers from inconsistency in selection variables and does not encourage the selection of the correlated explanatory variables together.

Motivated by these points, several weights have been proposed by assuming several values of (γ) which are (1, 1.5, 3, 5, 10) at $p > n$, as follows:

1. When $(\gamma = 1)$ Weight is defined as: $\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-1}$…(7)
2. When $(\gamma = 1.5)$ Weight is defined as: $\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-1.5}$…(8)
3. When $(\gamma = 3)$ Weight is defined as: $\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-3}$…(9)
4. When $(\gamma = 5)$ Weight is defined as: $\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-5}$…(10)
5. When $(\gamma = 10)$ Weight is defined as: $\hat{w}_j = \left[abs(\hat{\beta}_j^{initial})\right]^{-10}$…(11)

The proposed PALASSO estimator is defined as follows:

$$\hat{\beta}_{PLR}^{ALASSO} = \text{argmin}_{\beta}\left\{(y - X\beta)^T(y - X\beta) + \lambda \sum_{j=1}^{p} \hat{w}_j|\beta_j|\right\} \quad \dots (12)$$

where $w = (w_j; j = 1,2, \dots, py_{1\dots\dots}y_n)^T \in R^p$ is a given data-driven weight vector and it is obtained by

$$\hat{w}_j = \left[\text{abs}(\hat{\beta}_j^{\text{initial}})\right]^{-\gamma}, j = 1,2, \dots, p \ , (-\gamma = 1 \text{ or } 1.5 \text{ or } 3 \text{ or} 5 \text{ or } 10\ ) \dots (13)$$

6. **Regression Experiment Design**

The classical methods use the approach of multiple regression to define the concept of main effects and interaction effects for two-level designs Mason et. al [8] ; Williams, [16]. The analysis of experiments can be formulated in the form of the general linear regression where we have (n) observation on a dependent variable (y) and predictors $(X_1, X_2, \dots, X_P)$, and

$$Y = X\beta + \varepsilon \quad \dots (14)$$

Where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $\beta = (\beta_{1\dots\dots}\beta_p)$ Throughout this work, we center each input variable so that the observed mean is zero and standard deviation is one. Mohammed, Odah & BAGER, [10].

7. **Regression Model of the Factorial Design:**

Factorial design is important to study models built using two or more factors, and factorial designs also have the advantage of being able to measure the interaction effect between all factors through the interaction between factors, the response variable measurement is affected by changing the levels of the factor, the effect of the main factor is completely different compared to its interaction with other factors. Another possible approach is the possibility of using the main effects and the interference effects of the two-level factorial designs as a regression model after transforming the levels of factors and the interactions between the levels of these factors into new variables called explanatory variables, assuming that we have a factorial experiment applied in a complete random design (2^4) i.e. we have four factors (A, B, C, D) and each factor has two levels That is, the mathematical model of the factorial experiment (2^4) will contain (80) effect, in addition to the effect of the general arithmetic mean, as follows: Bahr et al.[2].

(8) Main effects of the levels of each factor.
(24) Effect of binary overlap between factor levels.
(32) Effect of tripartite overlap between factor levels.
(16) Effect of interquartile interaction between factor levels.

8. **The proposed method (Wisam's Model in Regression Model of the Factorial Design)**

This method was proposed to transform the mathematical model of a factorial experiment (2^4) into a regression model, and then use the penalty methods and the proposed penalty method to address the problem of high-dimensional data.

This proposal is done by transforming each level of the factors and the interaction between the levels of these factors into an explanatory variable, and the number of these levels and the interactions between the levels represented by the explanatory variables are as follows:

Using equation (15) the number of prime factors for each level is found and the number of interaction between the binary, ternary and quaternary factors for all levels of factors Onyiah [12].

$$C_1^t = \frac{t!}{1!(t-1)!} \quad \dots (15)$$

1. $C_1^4 = \frac{4!}{1!(4-1)!} = 4$

(4) represents the number of main factors of the experiment which are $(\alpha, \beta, \gamma, \delta)$, and since each factor has two levels, the number of main effects for each level of the four factors is (8) which is $(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \delta_1, \delta_2)$.

2. $C_2^4 = \frac{4!}{2!(4-2)!} = 6$

6) epresents the number of binary interactions for the experiment, which are $(\alpha\beta, \alpha\gamma, \alpha\delta, \beta\gamma, \beta\delta, \gamma\delta)$, and since each factor has two levels, the number of binary interactions for each two factors is (4) according to the levels of each factor for each interference as follows:

$\alpha_i\beta_j \Rightarrow (\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2\beta_2)$.

$\alpha_i\gamma_k \Rightarrow (\alpha_1\gamma_1, \alpha_1\gamma_2, \alpha_2\gamma_1, \alpha_2\gamma_2)$.

$\alpha_i\delta_m \Rightarrow (\alpha_1\delta_1, \alpha_1\delta_2, \alpha_2\delta_1, \alpha_2\delta_2)$.

$\beta_j\gamma_k \Rightarrow (\beta_1\gamma_1, \beta_1\gamma_2, \beta_2\gamma_1, \beta_2\gamma_2)$.

$\beta_j\delta_m \Rightarrow (\beta_1\delta_1, \beta_1\delta_2, \beta_2\delta_1, \beta_2\delta_2)$.

$\gamma_k\delta_m \Rightarrow (\gamma_1\delta_1, \gamma_1\delta_2, \gamma_2\delta_1, \gamma_2\delta_2)$.

We have (24) interactions resulting from (6) binary interactions, and each binary interactions have four interactions according to the level of each factor.

3. $C_3^4 = \frac{4!}{3!(4-3)!} = 4$

(4) represents the number of triple interactions for the experiment, which are $(\alpha\beta\gamma, \alpha\beta\delta, \alpha\gamma\delta, \beta\gamma\delta)$ and since each factor has two levels, the number of triple interactions is (8) according to the levels of each factor and for each interactions as follows:

$\alpha_i\beta_j\gamma_k \Rightarrow (\alpha_1\beta_1\gamma_1, \alpha_1\beta_1\gamma_2, \alpha_1\beta_2\gamma_1, \alpha_1\beta_2\gamma_2, \alpha_2\beta_1\gamma_1, \alpha_2\beta_1\gamma_2, \alpha_2\beta_2\gamma_1, \alpha_2\beta_2\gamma_2).$

$\alpha_i\beta_j\delta_m \Rightarrow (\alpha_1\beta_1\delta_1, \alpha_1\beta_1\delta_2, \alpha_1\beta_2\delta_1, \alpha_1\beta_2\delta_2, \alpha_2\beta_1\delta_1, \alpha_2\beta_1\delta_2, \alpha_2\beta_2\delta_1, \alpha_2\beta_2\delta_2).$

$\alpha_i\gamma_k\delta_m \Rightarrow (\alpha_1\gamma_1\delta_1, \alpha_1\gamma_1\delta_2, \alpha_1\gamma_2\delta_1, \alpha_1\gamma_2\delta_2, \alpha_2\gamma_1\delta_1, \alpha_2\gamma_1\delta_2, \alpha_2\gamma_2\delta_1, \alpha_2\gamma_2\delta_2).$

$\beta_j\gamma_k\delta_m \Rightarrow (\beta_1\gamma_1\delta_1, \beta_1\gamma_1\delta_2, \beta_1\gamma_2\delta_1, \beta_1\gamma_2\delta_2, \beta_2\gamma_1\delta_1, \beta_2\gamma_1\delta_2, \beta_2\gamma_2\delta_1, \beta_2\gamma_2\delta_2).$

We have (32) an interactions resulting from (4) three interactions, and each triple interactions have eight overlaps according to the level of each factor.

4. $C_4^4 = \frac{4!}{4!(4-4)!} = 1$

(1) represents the number of interquartile interactions for the experiment, which is ($\alpha\beta\gamma\delta$) and since each factor has two levels, the

number of interquartile interactions is (16) according to the levels of each factor as follows.

$\alpha_i\beta_j\gamma_k\delta_m \Rightarrow (\alpha_1\beta_1\gamma_1\delta_1, \alpha_1\beta_1\gamma_2\delta_2, \alpha_1\beta_1\gamma_2\delta_1, \alpha_1\beta_1\gamma_2\delta_2, \alpha_1\beta_2\gamma_1\delta_1, \alpha_1\beta_2\gamma_1\delta_2,$

$\alpha_1\beta_2\gamma_2\delta_1, \alpha_1\beta_2\gamma_2\delta_2, \alpha_2\beta_1\gamma_1\delta_1, \alpha_2\beta_1\gamma_2\delta_2, \alpha_2\beta_1\gamma_2\delta_1, \alpha_2\beta_1\gamma_2\delta_2, \alpha_2\beta_2\gamma_1\delta_1, \alpha_2\beta_2\gamma_1\delta_2,$

$\alpha_2\beta_2\gamma_2\delta_1, \alpha_2\beta_2\gamma_2\delta_2).$

The total number of the main factors and the two-, three-, and four-way interactions using the levels of each factor is the product of summing the number of factors and interactions for the previous cases are (8 + 24 + 32 + 8 = 80) which represent the explanatory variables that correspond to the factors as shown in the table (1).

Table (1): Representation of factor levels and the interaction between factor levels with explanatory variables (X)

Factors level	The explanatory variables they correspond	Factors level	The explanatory variables they correspond	Factors level	The explanatory variables they correspond	Factors level	The explanatory variables they correspond
α_1	x_{11}	$\beta_1\gamma_1$	$x_{21}x_{31}$	$\alpha_1\beta_1\delta_1$	$x_{11}x_{21}x_{41}$	$\beta_2\gamma_1\delta_1$	$x_{22}x_{31}x_{41}$
α_2	x_{12}	$\beta_1\gamma_2$	$x_{21}x_{32}$	$\alpha_1\beta_1\delta_2$	$x_{11}x_{21}x_{42}$	$\beta_2\gamma_1\delta_2$	$x_{22}x_{31}x_{42}$
β_1	x_{21}	$\beta_2\gamma_1$	$x_{22}x_{31}$	$\alpha_1\beta_2\delta_1$	$x_{11}x_{22}x_{41}$	$\beta_2\gamma_2\delta_1$	$x_{22}x_{32}x_{41}$
β_2	x_{22}	$\beta_2\gamma_2$	$x_{22}x_{32}$	$\alpha_1\beta_2\delta_2$	$x_{11}x_{22}x_{42}$	$\beta_2\gamma_2\delta_2$	$x_{22}x_{32}x_{42}$
γ_1	x_{31}	$\beta_1\delta_1$	$x_{21}x_{41}$	$\alpha_2\beta_1\delta_1$	$x_{12}x_{21}x_{41}$	$\alpha_1\beta_1\gamma_1\delta_1$	$x_{11}x_{21}x_{31}x_{41}$
γ_2	x_{32}	$\beta_1\delta_2$	$x_{21}x_{42}$	$\alpha_2\beta_1\delta_2$	$x_{12}x_{21}x_{42}$	$\alpha_1\beta_1\gamma_2\delta_2$	$x_{11}x_{21}x_{32}x_{42}$
δ_1	x_{41}	$\beta_2\delta_1$	$x_{22}x_{41}$	$\alpha_2\beta_2\delta_1$	$x_{12}x_{22}x_{41}$	$\alpha_1\beta_1\gamma_2\delta_1$	$x_{11}x_{21}x_{32}x_{41}$
δ_2	x_{42}	$\beta_2\delta_2$	$x_{22}x_{42}$	$\alpha_2\beta_2\delta_2$	$x_{12}x_{22}x_{42}$	$\alpha_1\beta_1\gamma_2\delta_2$	$x_{11}x_{21}x_{32}x_{42}$
$\alpha_1\beta_1$	$x_{11}x_{21}$	$\gamma_1\delta_1$	$x_{31}x_{41}$	$\alpha_1\gamma_1\delta_1$	$x_{11}x_{31}x_{41}$	$\alpha_1\beta_2\gamma_1\delta_1$	$x_{11}x_{22}x_{31}x_{41}$
$\alpha_1\beta_2$	$x_{11}x_{22}$	$\gamma_1\delta_2$	$x_{31}x_{42}$	$\alpha_1\gamma_1\delta_2$	$x_{11}x_{31}x_{42}$	$\alpha_1\beta_2\gamma_1\delta_2$	$x_{11}x_{22}x_{31}x_{42}$
$\alpha_2\beta_1$	$x_{12}x_{21}$	$\gamma_2\delta_1$	$x_{32}x_{41}$	$\alpha_1\gamma_2\delta_1$	$x_{11}x_{32}x_{41}$	$\alpha_1\beta_2\gamma_2\delta_1$	$x_{11}x_{22}x_{32}x_{41}$
$\alpha_2\beta_2$	$x_{12}x_{22}$	$\gamma_2\delta_2$	$x_{32}x_{42}$	$\alpha_1\gamma_2\delta_2$	$x_{11}x_{32}x_{42}$	$\alpha_1\beta_2\gamma_2\delta_2$	$x_{11}x_{22}x_{32}x_{42}$
$\alpha_1\gamma_1$	$x_{11}x_{31}$	$\alpha_1\beta_1\gamma_1$	$x_{11}x_{21}x_{31}$	$\alpha_2\gamma_1\delta_1$	$x_{12}x_{31}x_{41}$	$\alpha_2\beta_1\gamma_1\delta_1$	$x_{12}x_{21}x_{31}x_{41}$
$\alpha_1\gamma_2$	$x_{11}x_{32}$	$\alpha_1\beta_1\gamma_2$	$x_{11}x_{21}x_{32}$	$\alpha_2\gamma_1\delta_2$	$x_{12}x_{31}x_{42}$	$\alpha_2\beta_1\gamma_2\delta_2$	$x_{12}x_{21}x_{32}x_{42}$
$\alpha_2\gamma_1$	$x_{12}x_{31}$	$\alpha_1\beta_2\gamma_1$	$x_{11}x_{22}x_{31}$	$\alpha_2\gamma_2\delta_1$	$x_{12}x_{32}x_{41}$	$\alpha_2\beta_1\gamma_2\delta_1$	$x_{12}x_{21}x_{32}x_{41}$
$\alpha_2\gamma_2$	$x_{12}x_{32}$	$\alpha_1\beta_2\gamma_2$	$x_{11}x_{22}x_{32}$	$\alpha_2\gamma_2\delta_2$	$x_{12}x_{32}x_{42}$	$\alpha_2\beta_1\gamma_2\delta_2$	$x_{12}x_{21}x_{32}x_{42}$
$\alpha_1\delta_1$	$x_{11}x_{41}$	$\alpha_2\beta_1\gamma_1$	$x_{12}x_{21}x_{31}$	$\beta_1\gamma_1\delta_1$	$x_{21}x_{31}x_{41}$	$\alpha_2\beta_2\gamma_1\delta_1$	$x_{12}x_{22}x_{31}x_{41}$
$\alpha_1\delta_2$	$x_{11}x_{42}$	$\alpha_2\beta_1\gamma_2$	$x_{12}x_{21}x_{32}$	$\beta_1\gamma_1\delta_2$	$x_{21}x_{31}x_{42}$	$\alpha_2\beta_2\gamma_1\delta_2$	$x_{12}x_{22}x_{31}x_{42}$
$\alpha_2\delta_1$	$x_{12}x_{41}$	$\alpha_2\beta_2\gamma_1$	$x_{12}x_{22}x_{31}$	$\beta_1\gamma_2\delta_1$	$x_{21}x_{32}x_{41}$	$\alpha_2\beta_2\gamma_2\delta_1$	$x_{12}x_{22}x_{32}x_{41}$
$\alpha_2\delta_2$	$x_{12}x_{42}$	$\alpha_2\beta_2\gamma_2$	$x_{12}x_{22}x_{32}$	$\beta_1\gamma_2\delta_2$	$x_{21}x_{32}x_{42}$	$\alpha_2\beta_2\gamma_2\delta_2$	$x_{12}x_{22}x_{32}x_{42}$

By converting the levels of factors and the interaction between factors into explanatory variables, the regression equation will contain (80) an explanatory variable and (80) a parameter in addition to a parameter (β_0) as in the following equation (16) :

$$y_i = \beta_0 + \beta_{11}x_{11} + \cdots + \beta_{42}x_{42} + \beta_{1121}x_{11}x_{21} + \cdots + \beta_{3242}x_{32}x_{42} + \beta_{112131}x_{11}x_{21}x_{31} + \cdots + \beta_{223242}x_{22}x_{32}x_{42} + \beta_{11213141}x_{11}x_{21}x_{31}x_{41} + \cdots + \beta_{12223242}x_{12}x_{22}x_{32}x_{42} + u_i \ldots (16)$$

y_i : response variable $(i = 1.2. \ldots . n)$.

$(x_{11}, \ldots, x_{44}, x_{11}x_{21}, \ldots, x_{12}x_{22}x_{32}x_{42})$: Independent variables.

β_0 : intersection parameter.

$(\beta_{11}, \ldots, \beta_{12223242})$ Regression parameters $(j = 1.2. \ldots . 2^N)$

(N) : Number of factors.

u_i : Random error follows a normal distribution with mean (0) and variance (σ^2).

Figure (1) shows the method of encoding explanatory variables in addition to encoding parameters:

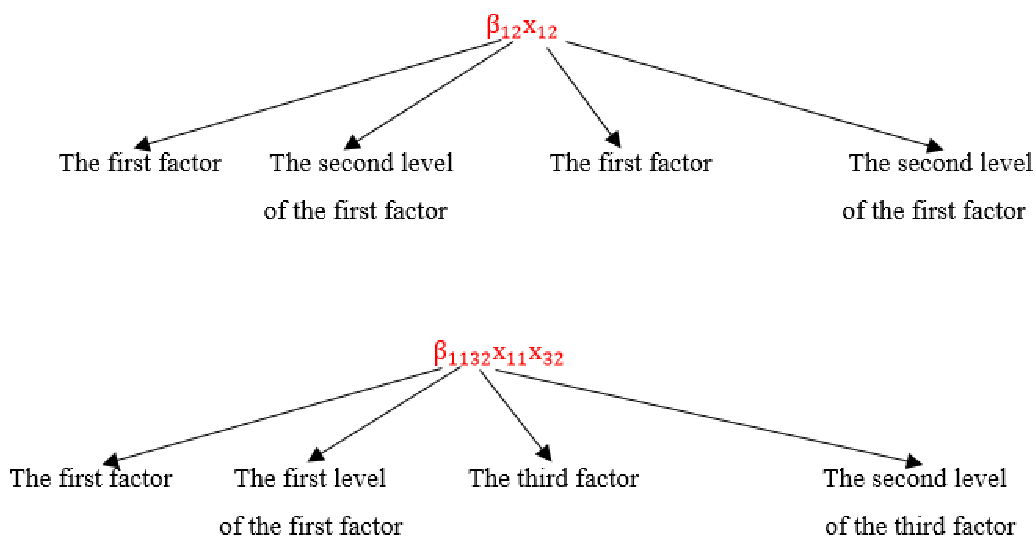

Figure (1): shows the method of encoding explanatory variables in addition to encoding parameter

After building and modeling the regression equation (16), the matrix (X) consists of (81) a column and each column represents the effect of the levels of the main factors for each factor in addition to the effect of the interaction between all levels of the factors and that the number of rows is (32) for a factorial experiment (2^4) with two repetitions, each row represents observations or response variable and that the number of rows represents the sample size, so the dimensions of the matrix (X) is (32*81) We note that (p>n) so the matrix (X) suffers from high-dimensional data and the high dimensions will be processed through the use of penalty methods (Lasso, adaptive lasso, and the proposed method on adaptive Lasso weights) And by integrating the regression equation (16) with the mentioned penalty methods, a model will be built that consists of the sum of the squares of the residual values extracted from the equation (16) and the penalty term for each of the Lasso method, the adaptive Lasso method, and the proposed method on the adaptive Lasso weights for the purpose of estimating and selecting the coefficients the task in the experiment, then compare the penalty methods with the proposed penalty method to get the best model represented by the selected important variables, then compare the results with the analysis of variance for a factorial experiment applied using a complete random design, as follows:

9. Lasso Factorial Design Regression Model

By adding the penalty (LASSO) from equation (2) to equation (16), we get:

$$y_i = \beta_0 + \beta_{11}x_{11} + \cdots + \beta_{42}x_{42} + \beta_{1121}x_{11}x_{21} + \cdots + \beta_{3242}x_{32}x_{42} + \beta_{112131}x_{11}x_{21}x_{31} + \cdots + \beta_{223242}x_{22}x_{32}x_{42} + \beta_{11213141}x_{11}x_{21}x_{31}x_{41} + \cdots + \beta_{12223242}x_{12}x_{22}x_{32}x_{42} + u_i + \lambda[|\beta_0| + |\beta_{11}| + \cdots |\beta_{12}| + |\beta_{21}| + |\beta_{1121}| + \cdots + |\beta_{1122}| + |\beta_{11213141}| + \cdots + |\beta_{12223242}|] \quad \ldots (17)$$

10. Adaptive Lasso Factorial Design Regression Model

By adding the penalty (ALASSO) from equation (4) to equation (16), we get:

$$y_i = \beta_0 + \beta_{11}x_{11} + \cdots + \beta_{42}x_{42} + \beta_{1121}x_{11}x_{21} + \cdots + \beta_{3242}x_{32}x_{42} + \beta_{112131}x_{11}x_{21}x_{31} + \cdots + \beta_{223242}x_{22}x_{32}x_{42} + \beta_{11213141}x_{11}x_{21}x_{31}x_{41} + \cdots + \beta_{12223242}x_{12}x_{22}x_{32}x_{42} + u_i + \lambda[w_0|\beta_0| + w_{11}|\beta_{11}| + \cdots + w_{42}|\beta_{42}| + w_{1121}|\beta_{1121}| + \cdots + w_{3242}|\beta_{3242}| + w_{112131}|\beta_{112131}| + \cdots + w_{223242}|\beta_{223242}| + w_{11213141}|\beta_{11213141}| + \cdots + w_{12223242}|\beta_{12223242}|] \ldots (18)$$

11. Proposed Initial Weight in Adaptive LASSO Factorial Design Regression Model

Several weights have been proposed assuming several values of (γ) which are (1, 1.5, 3, 5, 10) at $p > n$, as in the equations (7) (8) (9) (10) and (11).

Applying the previous equation (18) but with the suggested weights of the adapted Lasso method at the assumed (γ) values of (1, 1.5, 3, 5, 10).

Then the estimation method was applied to each method to obtain the parameter values estimated by (R) software to obtain the results and compare the methods.

12. Application:

14.1 Simulation study:

Researchers that have worked with this subject have come up with a few different definitions of simulation over the years. The purpose of simulation is to create a representation or imitation of the actual reality of the system using specific models written in accordance with a software method and with specific iterations in order to obtain experimental results whose goal is to validate the results obtained from the theoretical side in order to select the most appropriate method for analysis. Fishman, G. S, [4].

The Monte Carlo approach was used in the simulation, which was designed and implemented using the R programming language and based on packages (devtools, hqreg, robustHD, parcor, gglasso, lasso2, MASS, glmnet, HDeconometrics). The simulation is designed and implemented using these packages. In order to evaluate the performance of the proposed method for transforming and modeling the mathematical model in the case of experimenting with factors (24) that suffer from a high-dimensional data problem, into a proposed regression equation that contains (80) an explanatory variable that represents the effect of the levels of each factor in addition to the effect of the interaction between the levels of these factors in addition to the effect of The general arithmetic mean represented by (β_0), and the penalty methods (Lasso, Adaptive Lasso) were used to address the estimation problem in the case of high-dimensional data and to propose new weights for the adaptive Lasso method by suggesting (5) values for Gamma (γ) which are (-1, -1.5, -3, -5 , -10) to change the weight of the adaptive Lasso method according to the proposed Gamma (γ) values and compare it with the penalty methods (Lasso, ALasso) to estimate parameters and select important variables based on several criteria, including the following (Rsq, ARsq, MSE, Mad AIC, EFF) . And the use of several sample sizes (50,75,100,150,200) where the frequency (500) was applied. The simulation study was presented as follows:

Step 1: Generate the response variable based on normal distribution with known mean and known variance $y \sim N(\mu, \sigma)$.

Step 2: Generate the random error term based on normal distribution with known mean and known variance $e \sim N(\mu, \sigma)$.

Step 3: generate design matrix with four factors and two levels (high level [+1] and low level [-1]).

Step 4: generate different number of sample in each iteration with (N= 50, 75, 100, 150, and 200).

Step 5: compute the main effects and interaction effects according to equation (2-121).

Step 6: applying lasso, adaptive lasso and suggested adaptive lasso according to equations (2-122), (2-123) and (2-123) but with the suggested weights of the adaptive Lasso method at the assumed (γ) values of (1, 1.5, 3, 5, 10), according to the equations (2-124), (2-125), (2-126), (2-127) and (2-128) respectively.

Step 7: for the suggested method we use different values of lambda and w_j according to equations (2-124), (2-125), (2-126), (2-127) and (2-128)

Step 8: calculate the mean square error (MSE), determination of coefficient (R-square) for methods of study.

Step 9: compute the efficiency (EFF) of methods respect to suggested method as follows

$$\mathrm{EFF} = \frac{\mathrm{MSE\ (suggeste\_method)}}{\mathrm{MSE\ (classical\_method)}}$$

Values of EFF that less than 1 indicate that the suggested method is more efficiency than the classical method.

Step 10: compute the AIC for methods of study according equation (2-28), where the method that has minimum value of AIC it has the best performance.

12.2 Simulation Results

Using the simulation method that was performed on the proposed regression model that suffers from the problem of high-dimensional data, in this regard it has been relied on comparing the punitive methods with the proposed penalty method described in the second chapter on the basis of the following cases:

1. **In case (γ=1)**

   The following table (2) provides a comparison between the results obtained using the penalty methods (Lasso, Adaptive Lasso) and the proposed method on adaptive lasso weights for the case (**γ=1**) and for all sample sizes using the criteria (Rsq, ARsq, MSE, Mad AIC, EFF) to evaluate the effectiveness of the proposed method in relation to the penalty methods.

Table (2) The results of the comparison criteria after estimation using the Lasso method, adaptive Lasso method and the proposed method at (γ=1) and for all sample sizes

γ	N	Methods	Rsq	AdRsq	MSE	MAD	AIC	Eff.	Lambda
1	50	Lasso	0.4466	0.4014	0.8902	1.2512	103.181	0.6107	0.2144
		ALasso	0.4515	0.4061	0.8584	1.2381	105.166	0.6333	
		ALasso.Sug	0.7260	0.6721	0.5436	1.0564	1.934	1	
	75	Lasso	0.4392	0.3942	0.7906	1.1619	73.382	0.6549	0.2786
		ALasso	0.3568	0.3144	0.9140	1.3055	56.579	0.5665	
		ALasso.Sug	0.6897	0.6369	0.5178	0.9854	1.210	1	
	100	Lasso	0.2864	0.2462	0.6337	1.0675	40.868	0.6009	0.2407
		ALasso	0.2786	0.2387	0.6465	1.0768	39.574	0.5889	
		ALasso.Sug	0.6126	0.5622	0.3808	0.8599	10.428	1	
	150	Lasso	0.1797	0.1428	0.7794	1.2435	22.343	0.5587	0.5814
		ALasso	0.2050	0.1673	0.7173	1.2193	26.984	0.6070	
		ALasso.Sug	0.5641	0.5152	0.4354	0.8689	14.069	1	
	200	Lasso	0.1515	0.1155	0.8228	1.2331	20.417	0.5698	0.5511
		ALasso	0.1422	0.1065	0.8425	1.2280	18.674	0.5565	
		ALasso.Sug	0.5431	0.4949	0.4688	0.9250	12.964	1	

From Table (2) when Gama = 1, we note that the proposed method (ALasso.Sug) was able to choose the best model by increasing the values of the coefficient of determination and the corrected coefficient of determination with the lowest value of the Akiaki coefficient with lower values of estimation error (the lowest values of mse and mad). In addition, we can clearly see that the relative efficiency of the lasso and adlasso methods did not exceed 0.6333 and 0.6549 respect to suggested method (ALasso.Sug), respectively. In general, through all the comparison criteria, we can clearly notice that the proposed method has the supreme performance compared to other methods and at all sample sizes, as shown in the figures below.

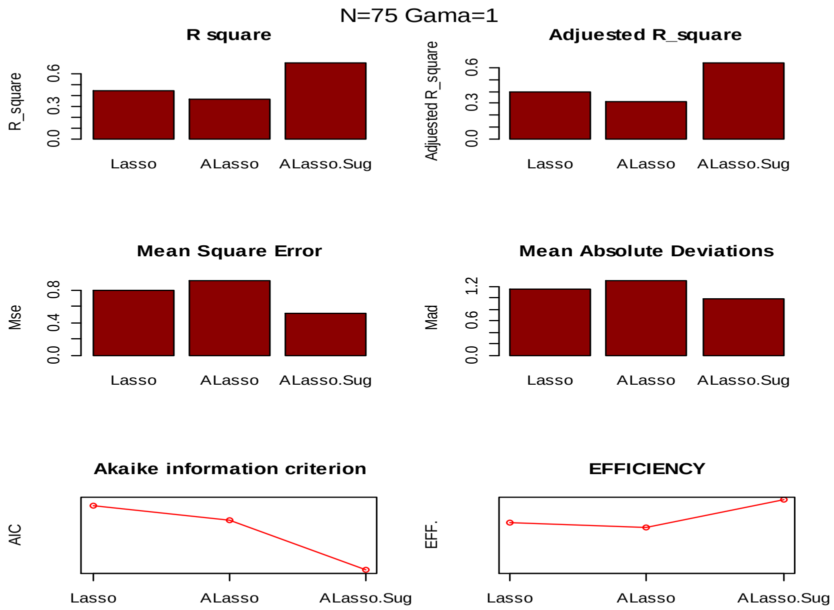

Figure (2) Comparison of penalty methods with the proposed method using the criteria at (γ=1) and (N=75)

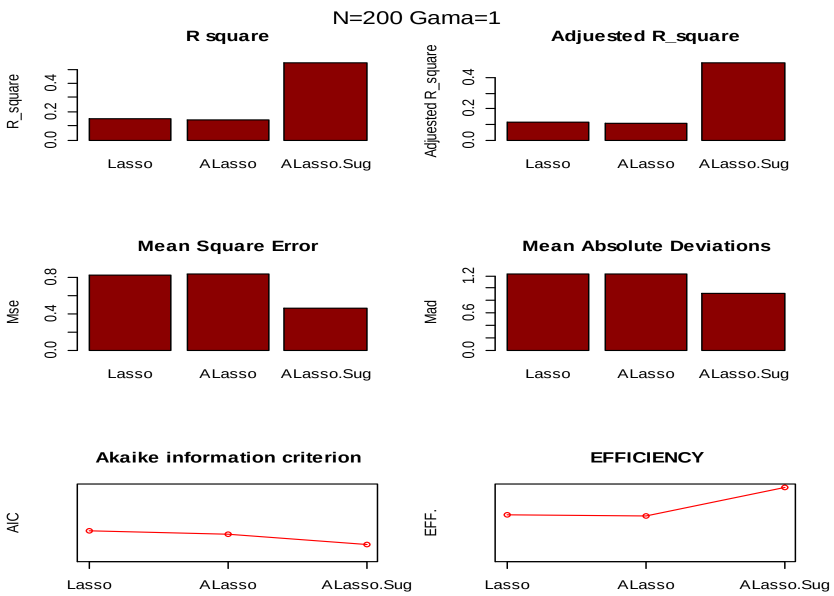

Figure (3) Comparison of penalty methods with the proposed method using the criteria at (γ=1) and (N=200)

Two cases were taken to plot at the sample size (N=75) and (N=150) and for the other Gamma values, the results of each case will be displayed

And the same interpretation for the rest of the tables for the rest of the proposed gamma values.

2. **In case (γ=1.5)**

Table (3) The results of the comparison criteria after estimation using the Lasso method, adaptive Lasso method and the proposed method at (γ=1.5) and for all sample sizes

γ	N	Methods	Rsq	AdRsq	MSE	MAD	AIC	Eff.	Lambda
1.5	50	Lasso	0.4501	0.4048	1.0166	1.2678	98.052	0.5155	0.0484
		ALasso	0.4431	0.3980	0.9550	1.2937	96.562	0.5487	
		ALasso.Sug	0.7328	0.6786	0.5240	1.0814	21.337	1	
	75	Lasso	0.2012	0.1637	1.0569	1.3674	37.220	0.5313	0.2792
		ALasso	0.2225	0.1843	0.9557	1.3321	40.307	0.5876	
		ALasso.Sug	0.5991	0.5491	0.5616	0.9598	9.692	1	
	100	Lasso	0.2041	0.1664	0.8897	1.2961	30.574	0.4982	0.2838
		ALasso	0.2491	0.2101	0.7951	1.2527	36.954	0.5574	
		ALasso.Sug	0.6280	0.5771	0.4432	0.9124	9.257	1	
	150	Lasso	0.1251	0.0900	1.1389	1.4347	25.058	0.6348	0.3225
		ALasso	0.0876	0.0536	1.2112	1.5257	18.099	0.5969	
		ALasso.Sug	0.4928	0.4462	0.7230	1.1930	9.192	1	
	200	Lasso	0.1419	0.1062	0.8417	1.3322	18.158	0.5595	0.4632
		ALasso	0.1808	0.1439	0.8066	1.2532	24.105	0.5838	
		ALasso.Sug	0.5468	0.4984	0.4709	0.9563	13.056	1	

3. **In case (γ=3)**

Table (4) The results of the comparison criteria after estimation using the Lasso method, adaptive Lasso method and the proposed method at (γ=3) and for all sample sizes

γ	N	Methods	Rsq	AdRsq	MSE	MAD	AIC	Eff.	Lambda
3	50	Lasso	0.3773	0.3342	0.8915	1.1267	71.008	0.5403	0.3251
		ALasso	0.4475	0.4023	0.7327	1.0976	84.818	0.6574	
		ALasso.Sug	0.7190	0.6653	0.4817	0.8885	36.708	1	
	75	Lasso	0.1858	0.1488	0.9884	1.5175	34.817	0.5306	0.3155
		ALasso	0.2073	0.1695	0.9273	1.4591	38.051	0.5655	
		ALasso.Sug	0.6105	0.5602	0.5244	0.9451	9.158	1	
	100	Lasso	0.1052	0.0707	1.3850	1.5905	19.082	0.5923	0.1974
		ALasso	0.0957	0.0614	1.3418	1.5716	20.327	0.6113	
		ALasso.Sug	0.4853	0.4388	0.8203	1.1586	7.119	1	
	150	Lasso	0.1598	0.1235	1.2181	1.4507	30.500	0.5893	0.4797
		ALasso	0.1285	0.0933	1.2976	1.5061	24.890	0.5532	
		ALasso.Sug	0.5526	0.5040	0.7178	1.2476	6.357	1	
	200	Lasso	0.0983	0.0640	1.2247	1.4629	16.810	0.6278	0.5183
		ALasso	0.1003	0.0660	1.2644	1.4830	16.186	0.6081	
		ALasso.Sug	0.4597	0.4141	0.7689	1.2829	3.892	1	

4. **In case (γ=5)**

Table (5) The results of the comparison criteria after estimation using the Lasso method, adaptive Lasso method and the proposed method at (γ=5) and for all sample sizes

γ	N	Methods	Rsq	AdRsq	MSE	MAD	AIC	Eff.	Lambda
5	50	Lasso	0.6570	0.6052	0.4881	1.0079	133.602	0.6502	0.1617

		ALasso	0.5958	0.5459	0.6137	1.1012	120.741	0.6205	
		ALasso.Sug	0.8085	0.7520	0.3808	0.9224	5.301	1	
	75	Lasso	0.1411	0.1055	1.3463	1.5600	27.941	0.5290	0.4333
		ALasso	0.1507	0.1148	1.3624	1.5101	28.585	0.5227	
		ALasso.Sug	0.5458	0.4975	0.7122	1.1075	4.291	1	
	100	Lasso	0.1543	0.1182	1.0919	1.3096	26.186	0.5403	0.4965
		ALasso	0.1661	0.1297	1.1209	1.2714	26.822	0.5263	
		ALasso.Sug	0.5850	0.5355	0.5900	1.0980	10.543	1	
	150	Lasso	0.2863	0.2461	0.6250	1.0244	39.131	0.5803	0.2408
		ALasso	0.2800	0.2400	0.6001	0.9881	37.721	0.6043	
		ALasso.Sug	0.6085	0.5583	0.3627	0.8094	10.519	1	
	200	Lasso	0.1433	0.1076	0.8226	1.1535	17.751	0.5157	0.4876
		ALasso	0.1531	0.1171	0.7962	1.1132	19.814	0.5328	
		ALasso.Sug	0.5549	0.5063	0.4242	0.8185	12.193	1	

5. **In case (γ=5)**

Table (6) The results of the comparison criteria after estimation using the Lasso method, adaptive Lasso method and the proposed method at (γ=10) and for all sample sizes

γ	N	Methods	Rsq	AdRsq	MSE	MAD	AIC	Eff.	Lambda
10	50	Lasso	0.3331	0.2914	1.5915	1.5587	78.050	0.5567	0.1489
		ALasso	0.2607	0.2213	1.8325	1.7212	63.770	0.4835	
		ALasso.Sug	0.6582	0.6064	0.8860	1.2494	7.916	1	
	75	Lasso	0.2701	0.2304	1.3011	1.5130	55.107	0.5524	0.1315
		ALasso	0.2334	0.1948	1.2657	1.5129	51.099	0.5679	
		ALasso.Sug	0.6236	0.5728	0.7188	1.1621	16.383	1	
	100	Lasso	0.1935	0.1562	1.0214	1.4684	35.615	0.5678	0.7262
		ALasso	0.1682	0.1317	1.0573	1.5866	31.972	0.5485	
		ALasso.Sug	0.5985	0.5486	0.5800	1.1047	2.563	1	
	150	Lasso	0.1460	0.1102	1.1085	1.5170	22.448	0.4482	0.8950
		ALasso	0.0991	0.0648	1.1428	1.5511	15.800	0.5647	
		ALasso.Sug	0.5266	0.4789	0.6454	1.1878	1.232	1	
	200	Lasso	0.1380	0.1024	0.9111	1.2598	20.008	0.6029	0.5070
		ALasso	0.1006	0.0662	0.9898	1.3103	15.088	0.5550	
		ALasso.Sug	0.4842	0.4378	0.5493	1.0657	6.897	1	

13. Conclusions

1. We draw the conclusion from the simulation results that all of the suggested penalty method's outcomes, when compared to penalty methods, were better for all gamma values and for all sample sizes.

2. We conclude that the relative efficiency of Lasso and ALasso methods did not exceed (0.67) with respect to the proposed method (ALasso.Sug).

3. We conclude that the best model in the case of the proposed value (γ=1) is with a sample size (N=100) as it has the least value for the mean square error (Mse=0.381) and mean absolute deviation (MAD=0.86)

4. We conclude that the best model in the case of the proposed value (γ=1.5) is with a sample size (N=100) as it has the least value for the mean square error (Mse=0.443) and mean absolute deviation (MAD=0.91)

5. We conclude that the best model in the case of the proposed value (γ=3) is with a sample size (N=50) as it has the least value for the mean square error (Mse=0.481) and mean absolute deviation (MAD=0.88)

6. We conclude that the best model in the case of the proposed value (γ=5) is with a sample size (N=150) as it has the least value for the mean square error (Mse=0.362) and mean absolute deviation (MAD=0.81)

7. We conclude that the best model in the case of the proposed value (γ=10) is with a sample size (N=200) as it has the least value for the mean square error (Mse=0.549) and mean absolute deviation (MAD=1.06)

References:

1. Al-Rawi, K. M. and Khalaf A. A. A. K. M. (2000). Design and Analysis of Agricultural Experiments. second edition. Dar Al-Kutub Directorate for Printing and Publishing, University of Mosul.
2. Bahr, K. M., Odah, M. H., & Bager, A. S. M. (2018). Application of Factorial Experiments with Asymmetric Responses for Determine Important Factors Affecting on Production of Dates. LUMEN Proceedings, 3, 64-77.
3. Buhlmann, P. and Van De Geer, S. (2011). Statistics for High-Dimensional Data: Methods, Theory and Applications. NY: Springer.
4. Fishman, G. S. (1973). Concepts and methods in discrete event digital simulation. Wiley.
5. Huang, J., Ma, S. and Zhang, C.-H. (2008). Adaptive Lasso for sparse high-dimensional regression models. Statistica Sinica. 18, 1603–1618.
6. Hui, F. K. C., Warton, D. I. and Foster, S. D. (2015). Tuning parameter selection for the adaptive lasso using ERIC. Journal of the American Statistical Association. 110(509), 262–269.
7. Lian, H. (2012). Variable selection in high-dimensional partly linear additive models. Journal of Nonparametric Statistics. 24(4), 825–839.
8. Mason, R. L., Gunst, R. F., & Hess, J. L. (2003). Statistical design and analysis of experiments: with applications to engineering and science (Vol. 474). John Wiley & Sons.
9. Mohammed, B. K. (2018). Robust Lasso Variable Selection for Factorial Experiments Analysis with Application. International Journal of Statistics and Applications, 8(2), 79-87.

10. Mohammed, B. K., Odah, M. H., & BAGER, A. S. M. (2017). Determining the most important factors affecting dairy products during the storage period using factorial experiments: case study in Iraq. In the Proceedings of the International Management Conference (Vol. 11, No. 1, pp. 26-34). Faculty of Management, Academy of Economic Studies, Bucharest, Romania.
11. Montgomery, D.C. (2017), Design and Analysis of Experiments. 9th Edition. John Wily & Sons. Inc New York.
12. Onyiah, L.C. ; (2008), " Design and Analysis of Experiments" ; Taylor & Francis Group, LLC , New York.
13. Potscher, B. M. and Schneider, U. (2009). On the distribution of the adaptive LASSO estimator. Journal of Statistical Planning and Inference. 139(8), 2775–2790.
14. Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". Journal of the Royal Statistical Society. Series B (Methodological). 58(1), 267–288.
15. Wang, X. and Wang, M. (2015). Variable selection for high-dimensional generalized linear models with the weighted elastic-net procedure. Journal of Applied Statistics, In press.
16. Williams, J. D. (1970). “A regression approach to experimental design”. The Journal of Experimental Education, 39(2), 83-90.
17. Zou, H. (2006). “The Adaptive lasso and Its Oracle properties” Journal of the American statistical Association (JASA) 101 No.476., PP .1418 -1429.