Research Article

Solution of the Ill-Posed Semiparametric Regression Model Based on Singular Value Modification Restriction

Yan Zhou ¹, Fengxiang Jin ², and Depeng Ma ³

¹Department of Resources and Civil Engineering, Shandong University of Science and Technology, Taian 271019, China

²Shandong Jianzhu University, Jinan 250101, China

³School of Water Conservancy and Civil Engineering, Shandong Agricultural University, Taian 271018, China

Correspondence should be addressed to Yan Zhou; zhouyansdust@163.com and Depeng Ma; mdp123@163.com

Received 15 February 2019; Revised 3 May 2019; Accepted 10 June 2019; Published 22 July 2019

Academic Editor: Francesca Vipiana

Copyright © 2019 Yan Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a solution of the ill-posed semi-parametric regression model based on singular value modification restriction, aimed at the ill-posed problem of the normal matrix which may occur in the process of solving the semiparametric regression model. First, the coefficient matrix is decomposed into singular values, and the smaller singular values are selected according to the criterion $\sum_{i=1}^r (1/\sigma_i) / \sum_{i=1}^n (1/\sigma_i) \leq 5\%$ (in the singular value matrix, $\sigma_1 > \sigma_1 > \dots > \sigma_r > \dots > \sigma_n$). Second, the relatively smaller singular values are modified by the biased parameter to suppress the magnification of the estimated variance so as to effectively reduce the variance of parameter estimation, reduce the introduction of deviation and obtain more reliable parameter estimation. The results of the numerical experiments show that the improved singular value modification restriction method can not only overcome the effect of the ill-posed normal matrix on the parameter estimation solution but also correctly separate the systematic errors and improve the accuracy of semiparametric regression model calculation results.

1. Introduction

The compensated least squares regularization criterion $V_{re}^T P V_{re} + \alpha \hat{s}^T R \hat{s} = \min$ [1–6] has been widely used as a good estimator since the establishment of the semiparametric regression model. However, in practice it has gradually become clear that the properties of the compensated least squares estimator deteriorate significantly, the solution for the undetermined parameters is extremely unstable, and the effect of separating systematic errors becomes worse, all of which are due to the coefficient matrix containing linear correlation column vectors. To deal with this problem, researchers introduced ridge estimation into the semiparametric model and deduced the semiparametric general penalized least squares (GPLS) estimation method with the criterion $V_{re}^T P V_{re} + \alpha \hat{s}^T R \hat{s} + k \hat{X}^T Q \hat{X} = \min$ [7, 8] and generalized penalized least squares by adding two smoothing parameters (GPLSBASP estimation method) with the criterion $\alpha V_{re}^T P V_{re} + (1 - \alpha) \hat{s}^T R \hat{s} + k \hat{X}^T Q \hat{X} = \min$ [9]. These theories and methods first use the L-curve

method or generalized cross-validation method to determine the regularization parameter α and then use ridge trace method or generalized cross-validation method to determine ridge regularization parameter k and determine R and Q as the regularization unit matrix. Because the above estimation contains the regularization matrix in the form of a unit matrix, the analysis of variance and deviation shows that the singular values of the coefficient matrix are modified by the ridge regularization parameter k . Although the ill-posed problem of the normal matrix is improved, the reliability of parameter estimation is reduced.

In dealing with ill-posed problems, more and more attention has been paid to the direct solution algorithm for the ill-posed observation equation based on singular value decomposition (SVD) [10, 11]. After SVD, relatively large singular values and corresponding left eigenvectors represent the more reliable part of the model parameters, while small singular values and corresponding right eigenvectors represent the unreliable part. The singular value correction method suppresses the unreliable components. It not only corrects the

smaller singular values but also corrects the relatively larger singular values, and this distorts the determined components in the model and increases the deviation. To solve this problem, scholars have proposed and developed many and varied effective numerical methods. Based on the spectral decomposition of the coefficient matrix of the linear model, the difference between Tikhonov and Tsvd regularization methods is explained theoretically in the paper by Wu [12]. In the two-step solution proposed in the paper by Wang [13], the essence of selecting a suitable regularization matrix can also be summed up in the better modification of singular values. The paper distinguishes between relatively large and relatively small singular values of the coefficient matrix in the linear model [14–18]. At the same time, it proposes an improved singular value correction method based on the characteristics of the truncated singular value method.

This paper synthesizes the characteristics of the truncated singular value method and the singular value correction method, draws lessons from the idea of dividing all singular values into two parts and modifying them separately, and proposes a solution method for the ill-posed semiparametric regression model based on the limitations of singular value correction.

2. Regularized Solutions of Ill-Posed Problems

2.1. Compensated Least Squares Estimation of the Semiparametric Model. Semiparametric model is

$$L = BX + s + \Delta \quad (1)$$

where coefficient B is an $m \times n$ -dimensional vector, estimated parameters X and s are n -dimensional vectors, observation vector L is an m -dimensional vector (i.e., L is an equal precision observation vector), and Δ is an m -dimensional error vector.

From the semiparametric regression model $L = BX + s + \Delta$, the following error equation can be obtained:

$$V_{re} = B\hat{X} + \hat{s} - L \quad (2)$$

Under the criterion of compensated least squares estimation, according to Lagrange function method for solving conditional extremum, the constructed extremum function is as follows:

$$\phi = V_{re}^T PV_{re} + \alpha \hat{s}^T R \hat{s} + 2K^T (B\hat{X} + \hat{s} + L - V_{re}) \quad (3)$$

where K is a Lagrange constant, α is the regularization parameter, and R is the regularization matrix.

The following can be derived from (3) and $\partial\phi/\partial V_{re} = 0$, $\partial\phi/\partial \hat{s} = 0$, and $\partial\phi/\partial \hat{X} = 0$:

$$\begin{aligned} K &= PV_{re} \\ K &= -\alpha R \hat{s} \end{aligned} \quad (4)$$

$$B^T K = 0$$

This gives the following normal equation:

$$\begin{bmatrix} B^T PB & B^T P \\ PB & P + \alpha R \end{bmatrix} \begin{bmatrix} \hat{X} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} B^T PL \\ PL \end{bmatrix} \quad (5)$$

When $B^T PB$ is reversible, the regularization matrix R can be determined by the natural spline smoothing function method, the time series method, the prior variance, or the prior variance-covariance method. The regularization parameter α can be determined by the L-curve method or generalized cross-validation method, and the following parametric and nonparametric estimation solution can be obtained:

$$\begin{aligned} \hat{X}_{(\alpha)} &= N^{-1} B^T P (L - \hat{s}_{(\alpha)}) \\ \hat{s}_{(\alpha)} &= (M + \alpha R)^{-1} ML \end{aligned} \quad (6)$$

where $N = B^T PB$ and $M = P - PBN^{-1}B^T P$. The variance and deviation are

$$\begin{aligned} D(\hat{X}_{\alpha}) &= \sigma_0^2 \alpha^2 N^{-1} B^T P (\alpha R + M)^{-1} \\ &\quad \cdot RQR (\alpha R + M)^{-1} (N^{-1} B^T P)^T \end{aligned} \quad (7)$$

$$\text{bias}(\hat{X}) = E(\hat{X}) - X = \alpha N^{-1} B^T P (\alpha R + M)^{-1} R s \quad (8)$$

2.2. Generalized Compensated Least Squares Estimation in the Semiparametric Model. When $B^T PB$ is ill-posed or irreversible, the criterion for the semiparametric generalized compensated least squares estimation is defined as

$$V_{re}^T PV_{re} + \alpha \hat{s}^T R \hat{s} + k \hat{X}^T Q \hat{X} = \min \quad (9)$$

where α is called regularization parameter, k is called ridge regularization parameter (ridge regularization parameter are called in the regularization method for ill-posed problems, and biased parameter are called in the singular value decomposition method for ill-posed problems.) that plays a balancing role in the process of minimization for V_{re} , s , and X . The regularization parameter α is first determined by the L-curve method or generalized cross-validation method, and then the parameter k is determined by the ridge trace method or generalized cross validation method and Q is a given regularization matrix, usually a unit array. The regularized matrix R can be determined by the natural spline smoothing function method, the time series method, the prior variance, or the prior variance-covariance method. The criterion function is constructed by the Lagrangian extremum method as follows:

$$\begin{aligned} \phi &= V_{re}^T PV_{re} + \alpha \hat{s}^T R \hat{s} + k \hat{X}^T Q \hat{X} \\ &\quad + 2K^T (B\hat{X} + \hat{s} - L - V_{re}) \end{aligned} \quad (10)$$

The following normal equation can be derived from (10) and $\partial\phi/\partial V_{re} = 0$, $\partial\phi/\partial \hat{s} = 0$, and $\partial\phi/\partial \hat{X} = 0$:

$$\begin{bmatrix} B^T PB + kQ & B^T P \\ PB & P + \alpha R \end{bmatrix} \begin{bmatrix} \hat{X} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} B^T PL \\ PL \end{bmatrix} \quad (11)$$

This gives the following:

$$\begin{aligned}\widehat{X}_{(k,\alpha)} &= N_{(k,\alpha)}^{-1} B^T P (L - \widehat{s}_{(k,\alpha)}) \\ \widehat{s}_{(k,\alpha)} &= (P + \alpha R - P B N_{(k,\alpha)}^{-1} B^T P)^{-1} \\ &\quad \cdot (P - P B N_{(k,\alpha)}^{-1} B^T P) L\end{aligned}\quad (12)$$

where $N_{(k,\alpha)} = B^T P B + kQ$ and $M_{(k,\alpha)} = P - P B N_{(k,\alpha)}^{-1} B^T P$.

The variance and deviation are

$$D(\widehat{X}_{(k,\alpha)}) = \sigma_0^2 \alpha^2 N_{(k,\alpha)}^{-1} B^T P (\alpha R + M_{(k,\alpha)})^{-1} \cdot RQR (\alpha R + M_{(k,\alpha)})^{-1} (N_{(k,\alpha)}^{-1} B^T P)^T \quad (13)$$

$$\begin{aligned}bias(\widehat{X}_{(k,\alpha)}) &= E(\widehat{X}_{(k,\alpha)}) - X \\ &= \alpha N_{(k,\alpha)}^{-1} B^T P (\alpha R + M_{(k,\alpha)})^{-1} R s.\end{aligned}\quad (14)$$

When the matrix $B^T P B$ is ill-posed or singular, the semi-parametric compensated least squares method will fail. The generalized compensated least squares estimation method can overcome the ill-posed problem of the normal matrix and obtain the parameter estimation solution.

3. Singular Value Decomposition Solutions of Ill-Posed Problems

For the semiparametric model, the parametric estimation solution, variance, and deviation of the estimation are shown by (6) to (8). The weight matrix P , which can be united, is set as the unit matrix to facilitate derivation, and the coefficient matrix B is decomposed to a singular value:

$$B_{m \times n} = V_{m \times m} D_{m \times n} U_{n \times n}^T \quad (15)$$

$$D = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \sigma_n \end{bmatrix} D^+ = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \frac{1}{\sigma_n} \end{bmatrix}. \quad (16)$$

Whereas V and U^T are orthogonal matrices that are the left and right singular value vectors of B , D is a singular value matrix, and B^+ , which is the singular value decomposition equation for the Moore-Penrose generalized inverse of B , is obtained as follows:

$$B_{n \times m}^+ = U_{n \times n} (D_{n \times m}^+)^T V_{m \times m}^T \quad (17)$$

The singular value decomposition expressions for the parametric and nonparametric estimates in the semiparametric models are

$$\begin{aligned}\widehat{X}_{(\alpha)} &= U (D^+)^T V^T (L - \widehat{s}_{(\alpha)}) \\ \widehat{s}_{(\alpha)} &= (M_{(\alpha)} + \alpha R)^{-1} M_{(\alpha)} L\end{aligned}\quad (18)$$

The variance and deviation of the parameter estimation are

$$D(\widehat{X}_{(\alpha)}) = \sigma_0^2 \alpha^2 U (D^+)^T V^T (\alpha R + M_{(\alpha)})^{-1} \cdot RQR (\alpha R + M_{(\alpha)})^{-1} (U (D^+)^T V^T) \quad (19)$$

$$\begin{aligned}bias(\widehat{X}_{(\alpha)}) &= E(\widehat{X}_{(\alpha)}) - X = \alpha U (D^+)^T \\ &\quad \cdot V^T (\alpha R + M_{(\alpha)})^{-1} R s.\end{aligned}\quad (20)$$

In (18), the L-curve method or generalized cross-validation method is used to determine the regularization parameter α , and the regularization matrix R can be determined by the natural spline smoothing function method or the time series method according to the actual situation, where $M_{(\alpha)} = P - P B U (D^+)^T V^T$, $\sigma_1 > \sigma_2 > \dots > \sigma_r > \dots > \sigma_n$ is the singular value of the design matrix [19, 20]. If the normal equation is ill-posed, then σ_n is a smaller value close to zero. It can be seen from (18) and (19) that smaller singular values will have a serious impact on the solution of the parametric estimation, and the estimation variance will be greatly amplified by smaller singular values, which will lead to unreliable compensation for least squares estimation and inaccurate estimation of the parameters.

The real cause of normal matrix morbidity is that some singular values of B are close to zero. To overcome the unwanted influence of the ill-posed problem of design matrix B on parameter estimation, we should start by improving the degree of non-zero singular values approaching zero. That is, by modifying the singular value of the coefficient matrix, the ill-conditioning of the coefficient matrix can be improved, and the stability and accuracy of parametric estimation can be improved. For this reason, an improved form of (16) is proposed, from which we can obtain

$$D_{(k,\alpha)}^+ = \begin{bmatrix} \frac{1}{\sigma_1 + k} & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \frac{1}{\sigma_n + k} \end{bmatrix} \quad (21)$$

The singular value decomposition expressions of the parametric and nonparametric estimation solutions of the semiparametric model based on singular value modification are as follows:

$$\begin{aligned}\widehat{X}_{(k,\alpha)} &= \mathbf{U}(\mathbf{D}_{(k,\alpha)}^+)^T \mathbf{V}^T (\mathbf{L} - \widehat{\mathbf{s}}) \\ \widehat{\mathbf{s}}_{(k,\alpha)} &= (\mathbf{M}_{(k,\alpha)} + \alpha \mathbf{R})^{-1} \mathbf{M}_{(k,\alpha)} \mathbf{L}\end{aligned}\quad (22)$$

The variance and deviation of the parametric estimation are

$$\begin{aligned}\mathbf{D}(\widehat{X}_{(k,\alpha)})_{all} &= \sigma_0^2 \alpha^2 \mathbf{U}(\mathbf{D}_{(k,\alpha)}^+)^T \mathbf{V}^T (\alpha \mathbf{R} + \mathbf{M}_{(k,\alpha)})^{-1} \\ &\cdot \mathbf{RQR}(\alpha \mathbf{R} + \mathbf{M}_{(k,\alpha)})^{-1} (\mathbf{U}(\mathbf{D}_{(k,\alpha)}^+)^T \mathbf{V}^T)^T\end{aligned}\quad (23)$$

$$\begin{aligned}bias(\widehat{X}_{(k,\alpha)})_{all} &= E(\widehat{X}) - X = \alpha \mathbf{U}(\mathbf{D}_{(k,\alpha)}^+)^T \\ &\cdot \mathbf{V}^T (\alpha \mathbf{R} + \mathbf{M}_{(k,\alpha)})^{-1} \mathbf{R} \mathbf{s} + \sum_{i=1}^n \frac{k}{\sigma_i + k} v_i^T X v_i\end{aligned}\quad (24)$$

In (22), the method of determining the regularization parameter α is L-curve method or generalized cross-validation method, the biased parameter k is determined by $k = n\sigma_0^2/\sigma_1 \widehat{x}_{pls}^T \widehat{x}_{pls}$, and the regularization matrix $\mathbf{R}$ can be determined by the natural spline smoothing function method or the time series method according to the actual situation, where $\mathbf{M}_{(k,\alpha)} = \mathbf{P} - \mathbf{PBU}(\mathbf{D}_{(k,\alpha)}^+)^T \mathbf{V}^T$. Thus, compared with (23) and (19), (24), and (20), the solution of ill-conditioned semiparametric regression model based on singular value modification reduces the variance of the estimator and the introduction of deviation and improves the stability of the solution by modifying the singular value of the coefficient matrix. Since the regularization matrix is a unit matrix, the biased parameters modify each eigenvalue, and the degree of correction is the biased parameter k .

4. Solution of Ill-Posed Semiparametric Regression Model Based on Singular Value Modification Restriction

Equation (21) modifies the smaller singular value to suppress its amplification effect on the parameter estimation solution. At the same time, it also modifies the larger singular value, which distorts the stable components of the model. An improved singular value decomposition method is designed for this purpose. The main idea is as follows. First, the coefficient matrix is decomposed into singular values, and the smaller singular values are selected according to the criterion $\sum_{i=1}^r (1/\sigma_i) / \sum_{i=1}^n (1/\sigma_i) \leq 5\%$ (singular value matrix, $\sigma_1 > \sigma_1 > \dots > \sigma_r > \dots > \sigma_n$). Then, the smaller singular values are modified by the biased parameter, which not only restrains the magnification effect of the smaller singular values on the parameter estimation solution but also avoids the effect of the larger singular values being modified.

For the modified restricted singular value method, set r as the truncation parameter and select it according to the

decision condition $\sum_{i=1}^r (1/\sigma_i) / \sum_{i=1}^n (1/\sigma_i) \leq 5\%$. If k is a biased parameter, the corrected singular value is

$$\sigma = \begin{cases} \sigma_i, & 1 \leq i \leq r \\ \sigma_i + k, & r+1 < i \leq n \end{cases}\quad (25)$$

$$k = \frac{n\sigma_0^2}{\sigma_1 \widehat{x}_{pls}^T \widehat{x}_{pls}}$$

$$\widetilde{\mathbf{D}}_{(k,\alpha)}^+ = \begin{bmatrix} \frac{1}{\sigma_1} & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \dots & \dots & \frac{1}{\sigma_{r+1} + k} & \dots \\ 0 & \dots & \dots & \frac{1}{\sigma_n + k} \end{bmatrix}\quad (26)$$

Then the singular value decomposition expressions of parametric and the nonparametric estimation solutions of the semiparametric models under the limitations of singular value modification can be written as follows:

$$\widehat{X} = \left(\mathbf{U}(\widetilde{\mathbf{D}}_{(k,\alpha)}^+)^T \mathbf{V}^T \right) (\mathbf{L} - \widehat{\mathbf{s}})\quad (27)$$

$$\widehat{\mathbf{s}} = (\widetilde{\mathbf{M}}_{(\alpha)} + \alpha \mathbf{R})^{-1} \widetilde{\mathbf{M}}_{(\alpha)} \mathbf{L}$$

Its variance and deviation are

$$\begin{aligned}\mathbf{D}(\widehat{X}_{(k,\alpha)})_{part} &= \sigma_0^2 \alpha^2 \mathbf{U}(\widetilde{\mathbf{D}}_{(k,\alpha)}^+)^T \\ &\cdot \mathbf{V}^T (\alpha \mathbf{R} + \widetilde{\mathbf{M}}_{(\alpha)})^{-1} \mathbf{RQR}(\alpha \mathbf{R} + \widetilde{\mathbf{M}}_{(\alpha)})^{-1} \\ &\cdot \left(\mathbf{U}(\widetilde{\mathbf{D}}_{(k,\alpha)}^+)^T \mathbf{V}^T \right)^T\end{aligned}\quad (28)$$

$$\begin{aligned}bias(\widehat{X}_{(k,\alpha)})_{part} &= E(\widehat{X}) - X = \alpha \mathbf{U}(\widetilde{\mathbf{D}}_{(k,\alpha)}^+)^T \\ &\cdot \mathbf{V}^T (\alpha \mathbf{R} + \widetilde{\mathbf{M}}_{(\alpha)})^{-1} \mathbf{R} \mathbf{s} + \sum_{i=r+1}^n \frac{k}{\sigma_i + k} v_i^T X v_i.\end{aligned}\quad (29)$$

In (28) and (29), the method of determining the regularization parameter is L-curve method or Generalized cross validation method; and the regularization matrix $\mathbf{R}$ can be determined by the natural spline smoothing function method or the time series method according to the actual situation, where $\widetilde{\mathbf{M}}_{(\alpha)} = \mathbf{P} - \mathbf{PBU}(\widetilde{\mathbf{D}}_{(\alpha)}^+)^T \mathbf{V}^T$. Equations (19), (23), and (28) show that the influence of morbidity on the estimated variance is mainly reflected in the magnification of variance by smaller singular values. The larger singular values have no adverse effect on the difference. Because the information on singular value decomposition is insufficient, the smaller singular values are mainly reflected in the smaller singular values. Therefore, the solution method for the ill-posed semiparametric regression model based on the limitations of singular value correction is applied. According to the criterion $\sum_{i=1}^r (1/\sigma_i) / \sum_{i=1}^n (1/\sigma_i) \leq 5\%$, the smaller singular

TABLE 1: Parameter valuation comparison of various methods (normal matrix is moderately ill-posed).

	Truth value	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5
$\hat{\mathbf{X}}$	2	2.6274	1.9896	2.5172	1.9035	2.0321
	3	2.7828	2.8825	2.8352	2.9255	2.9769
alpha1	-	1.3629	1.6184	1.4088	1.3629	1.3928
k	-	-	0.013	-	0.0445	0.008
alpha2	-	-	-	-	1.377	3.211
ΔX	-	0.4408	0.0120	0.2947	0.0109	0.0049
Δs	-	83.6400	14.8360	59.2897	13.8154	12.1854
$\hat{\sigma}_0$	-	0.6395	0.6282	0.6345	0.6043	0.6069

Note. $\Delta X = \|\mathbf{X} - \hat{\mathbf{X}}\|$ and $\Delta s = \|\mathbf{s} - \hat{\mathbf{s}}\|$.

values are selectively corrected and the larger singular values are retained. Compared with (28) and (23), (29) and (24), the method proposed in this paper, can not only reduce the variance efficiently but also reduce the damage to information and the introduction of deviation.

5. Comparisons and Analysis of Simulation Examples and Results

Numeric experiments. The data for Experiment 1 come from the paper cited in [9]. The validity of this method is analyzed when the matrix is moderately ill-posed. The data for Experiment 2 are adapted from the data for Experiment 1. The validity of this method is analyzed when the matrix is seriously ill-posed. The data for Experiment 3 are taken from the analysis of the Generalized Measurement Adjustment sample; that is, the epoch interval time is 2 seconds, with five observation satellites and four epochs to solve the ambiguity.

Experiment 1. Suppose there is a linear model $\mathbf{L}_1 = \mathbf{B}\mathbf{X}$, and $\tilde{\mathbf{x}} = [2 \ 3]^T$, where $\mathbf{B}(\mathbf{b}_{i,j})$ is a matrix of dimension 100×2 and $(\mathbf{b}_{i,1}) = i/20$, $(\mathbf{b}_{i,2}) = (1/20)^2$, $i = 1, 2, \dots, 100$. Suppose the system error is $\mathbf{L}_2 = [l_1 \ l_2 \ \dots \ l_{100}]^T$, where $l_i = 10 \sin(2(i-1)\pi/100)$, $i = 1, 2, \dots, 100$. The random error Δ is a column vector composed of 100 random numbers that obey $N(0, 1)$. The observed value is $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 + \Delta$.

In the semiparametric regression model, the regularization matrix is determined in the form of the sum of the squares of the difference between two adjacent observations. In this calculation example, $\text{rank}(\mathbf{R}) = 99 < 100$ and must satisfy $\text{rank}(\mathbf{B}) = 2 = t$. According to the semiparametric compensation least squares solution, there is a unique solution to calculate the condition number of the normal matrix $\mathbf{N}$, $\text{cond}(\mathbf{N}) = 273.4627$. To illustrate the influence of small singular values on parametric estimation, the following five computational schemes are designed.

Scheme 1. PLS estimation, solution by Formula (6), the regularization parameter is determined by L-curve method;

Scheme 2. GPLS estimation, solution by Formula (12), the regularization parameter is determined by L-curve method

and the ridge regularization parameter k is determined by generalized cross validation method;

Scheme 3. Direct solution of ill-posed semiparametric regression model based on singular value decomposition, solution by Formula (18), the regularization parameter is determined by L-curve method;

Scheme 4. Solution of ill-posed semiparametric regression model based on singular value modification, solution by Formula (22), Firstly, the regularization parameter 1 is determined by L-curve method, and then the biased parameter k is obtained by the solution of parameter estimation. Secondly, the biased parameter k is used to correct all singular values. Finally, the regularization parameter 2 is determined by L-curve method;

Scheme 5. Using the estimation method proposed in this paper, solution by Formula (27), Firstly, the regularization parameter 1 is determined by L-curve method, and then the biased parameter k is obtained by the solution of parameter estimation. Secondly, the biased parameter k is used to correct the smaller singular values. Finally, the regularization parameter 2 is determined by L-curve method;

The differences of the parameter estimation $\hat{\mathbf{X}}$, norms of $\mathbf{X}$ and $\hat{\mathbf{X}}$, norms of $\mathbf{s}$ and $\hat{\mathbf{s}}$, and unit weight errors calculated by the five schemes are shown in Table 1.

Experiment 2. There is no obvious difference between the values of the estimated solutions of the undetermined parameters in Table 1. This is because the ill-posed problem of the normal matrix in this numerical experiment is moderate and not very serious. The method of solving the ill-posed semiparametric regression model based on the modification restriction of singular values only shows a weak advantage. For this reason, the example is transformed and the true value of the parameter is taken as follows: $[1 \ 1 \ 1 \ 1]^T$. According to the condition of the Hilbert matrix, a coefficient matrix with dimension 100×4 is obtained and the remaining assumptions remain unchanged. The condition number of the normal matrix is obtained as $\text{cond}(\mathbf{N}) = 2.4864 \times 10^6$ and the morbidity is serious. Similarly, the above five schemes are designed to calculate the numerical examples. The relevant calculation results are listed in Table 2.

Experiment 3. The coefficient matrix of the error equation is

$$\mathbf{B} = \begin{bmatrix} 0.2727 & 1.5127 & -0.5903 & 0.1903 & -0.1903 & 0 & 0 \\ 0.3030 & -1.7179 & 0.1070 & 0 & 0.1903 & -0.1903 & 0 \\ -0.5994 & 1.1626 & -0.3181 & 0 & 0 & 0.1903 & -0.1903 \\ 0.1414 & -0.6323 & -0.0527 & 0 & 0 & 0 & 0.1903 \\ 0.2727 & 1.5126 & -0.5903 & 0.1903 & -0.1903 & 0 & 0 \\ 0.3031 & -1.7179 & 0.1074 & 0 & 0.1903 & -0.1903 & 0 \\ -0.5996 & 1.1625 & -0.3182 & 0 & 0 & 0.1903 & -0.1903 \\ 0.1419 & -0.6322 & -0.0530 & 0 & 0 & 0 & 0.1903 \\ 0.2727 & 1.5125 & -0.5902 & 0.1903 & -0.1903 & 0 & 0 \\ 0.3032 & -1.7179 & 0.1079 & 0 & 0.1903 & -0.1903 & 0 \\ -0.5998 & 1.1625 & -0.3183 & 0 & 0 & 0.1903 & -0.1903 \\ 0.1423 & -0.6322 & -0.0533 & 0 & 0 & 0 & 0.1903 \\ 0.2727 & 1.5124 & -0.5902 & 0.1903 & -0.1903 & 0 & 0 \\ 0.3033 & -1.7178 & 0.1083 & 0 & 0.1903 & -0.1903 & 0 \\ -0.6001 & 1.1624 & -0.3184 & 0 & 0 & 0.1903 & -0.1903 \\ 0.1427 & 0.6321 & -0.0536 & 0 & 0 & 0 & 0.1903 \end{bmatrix}. \quad (30)$$

The true value of the parameter is as follows:

$$= [4.2 \quad -2.1 \quad 3.5 \quad 12 \quad 13 \quad 2 \quad 12] \quad (31)$$

$$\mathbf{X}^T = [\delta_x \quad \delta_y \quad \delta_z \quad N_1 \quad N_2 \quad N_3 \quad N_4]$$

where N_i is the ambiguity parameter and δ_i is the coordinate parameter. The observed values are as follows:

$$\mathbf{L} = [-4.8586 \quad 9.2509 \quad -10.6390 \quad 12.2032 \quad -4.8583 \quad 9.2527 \quad -10.6402 \quad 12.2039 \quad -4.8581 \quad 9.2545 \quad -10.6414 \quad 12.2045 \quad -4.8577 \quad 9.2563 \quad -10.6426 \quad 12.2052]. \quad (32)$$

The conditional number of the normal matrix is 3.2127×10^{10} . There are serious ill-posed problems. Although some systematic errors of observations have been corrected in the calculation process of baseline calculation software, there are still some residual systematic errors in baseline calculation, which may lead to excessive residual errors. Therefore, the above five schemes are also used for calculation. The differences of the parameter estimation $\hat{\mathbf{X}}$, norms of $\mathbf{X}$ and $\hat{\mathbf{X}}$, and unit weight errors calculated by the five schemes are shown in Table 3.

The following conclusions can be drawn from Tables 2-3.

(1) When the normal matrix $\mathbf{B}^T \mathbf{P} \mathbf{B}$ is seriously ill-posed, the deviation between the solution and the true value of scheme 1 and scheme 3 is very large, and the effect of separating the system error becomes worse. The estimation method is not suitable, and the deviation varies with the degree of the ill-posed normal matrix. The more serious

the ill-posed normal matrix is, the greater the deviation is.

(2) Both scheme 2 and scheme 4 can effectively improve the estimation accuracy of scheme 1 when the normal matrix is ill-posed. However, when the normal matrix is seriously ill-posed, the deviation between the estimated value and the true value is still large. Then scheme 2 and scheme 4 can improve the solution effect of the ill-posed normal matrix. However, when the normal matrix is seriously ill-posed, the solution effect is not ideal.

(3) When the normal matrix is ill-posed and scheme 5 is adopted, the result of calculation is close to the simulated true value. The less ill-posed it is, the smaller the deviation is and the closer the estimation is to the true value. The accuracy of the parameter estimation solution obtained by this method is better than by the other two methods, and it is more effective in reducing the variance and deviation.

TABLE 2: Parameter valuation comparison of various methods (normal matrix is seriously ill-posed).

	Truth value	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5
$\widehat{X}$	1	41.1267	-0.0356	-1.3633	-0.5574	-0.5368
	1	-221.1655	-0.0261	-0.4572	0.1660	0.0601
	1	317.8136	-0.0195	4.1109	0.6712	0.6730
	1	-167.8547	1.97117	-2.1045	0.1096	0.1395
Alpha	-	1.5046	1.4699	1.571	1.4699	1.46
k	-	-	0.0119	-	0.6403	0.6403
alpha1	-	-	-	-	1.9778	1.409
Δx	-	3.37046×10^5	4.1956	27.0243	4.1219	4.0926
Δs	-	388.9770	37.5001	39.8220	33.8128	33.5758
$\widehat{\sigma}_0$	-	1.5046	1.0222	1.2490	1.017	1.015

Note. $\Delta X = \|X - \widehat{X}\|$ and $\Delta s = \|s - \widehat{s}\|$.

TABLE 3: Parameter valuation comparison of various methods.

	Truth value	Scheme 1	Scheme 2	Scheme 3 (unsolvable)	Scheme 4	Scheme 5
$\widehat{X}$	4.2	4.0524×10^4	2.7444	-	-1.5617	-1.5682
	-2.1	-2.0009×10^4	-5.9468	-	1.7846	1.7738
	3.5	3.0001×10^4	0.8886	-	3.1834	3.1872
	12	1.0089×10^4	-0.0089	-	2.3527	2.3597
	13	0.7585×10^4	-0.2128	-	6.4307	6.4556
	8	0.5051×10^4	-0.5323	-	-6.2507	-6.2715
	12	0.2567×10^4	2.7444	-	5.0245	5.0445
alpha	-	3.59	3.591	-	3.591	3.59
k	-	-	0.46	-	0.371	0.3712
alpha1	-	-	-	-	0.005	0.006
Δx	-	1.9083×10^8	520.6240	-	436.4224	436.1954
$\widehat{\sigma}_0$	-	-	10.8041	-	10.2609	10.2396

Note. (1) $\Delta X = \|X - \widehat{X}\|$ and (2) because the singular value is 0, scheme 3 has no solution.

6. Conclusion

This paper proposes a method for solving the ill-posed semi-parametric regression model based on the limitations of singular value correction. The method combines the characteristics of the truncated singular value method and the singular value correction method, which overcomes the damage caused by the ill-posed matrix and weakens the influence of biased estimation deviation.

The simulation example shows that the algorithm improves the quality of the solution significantly from the aspect of the mean squared error. It is easy to use in solving practical problems and has good applicability.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the project of Shandong Province Higher Educational Science and Technology Program, grant numbers J18KA195. This research was also partially funded by the Natural Science Foundation of Shandong Province; grant number ZR2019PD016 and the Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents.

References

- [1] P. J. Green, "Penalized likelihood for general semi-parametric regression models," *International Statistical Review / Revue Internationale de Statistique*, vol. 55, no. 3, p. 245, 1987.
- [2] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models*, vol. 58 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London, 1994.
- [3] H. Hu, *Estimation Methods and Their Applications in Semiparametric Model*, Wuhan University, 2004.
- [4] X. Pan, *The Estimation Theory and Application Research in Semi-Parametric Model*, Wuhan University, 2005.

- [5] S. Ding, *Survey Data Modeling and Semiparametric Estimating*, Wuhan University, 2005.
- [6] X. J. Tao, *Study on the Adjustment Criterion of the Semi-Parametric Adjustment Model*, Central South University, 2011.
- [7] S. Ding and Z. Sun, "General compensated least squares and its statistical properties," *Journal of Geodesy and Geodynamics*, vol. 26, no. 4, pp. 22–26, 2006.
- [8] Y. Liu, *Theory and Application in Geodesy of Semi-Parametric Model*, Information Engineering University, 2008.
- [9] J. Zhang, Z. X. Du, X. Y. Zhang, and N. Du, "The method of universal penalized least squares by adding smoothing parameters both to the residual and systematic parts," *Hydrographic Surveying and Charting*, vol. 33, no. 2, pp. 20–23, 2013.
- [10] W. Xu, Y. Chen, and W. You, "Application of SVD method in ill-posed problem," *Bulletin of Surveying and Mapping*, no. 1, pp. 62–63, 2016.
- [11] J. F. Guo, *Diagnostics Processing of Ill-Posed in Surveying Adjustment System*, Information Engineering University, 2002.
- [12] T. Wu, K. Deng, M. Huang, and Y. Ouyang, "An improved singular values decomposition method for ill-posed problem," *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, vol. 36, no. 8, pp. 900–903, 2011.
- [13] Z. Wang, J. Ou, and L. Liu, "A method for resolving Ill-conditioned problems - Two-step solution," *Geomatics and Information Science of Wuhan University*, vol. 30, no. 9, pp. 821–824, 2005.
- [14] X. Zeng, D. Liu, X. Li, J. Su, D. Chen, and W. Qi, "An improved singular value modification method for ill-posed problems," *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, vol. 40, no. 10, pp. 1349–1353, 2015.
- [15] D. Lin, J. Zhu, Y. Song, and Y. He, "Construction method of regularization by singular value decomposition of design matrix," *Cehui Xuebao/Acta Geodaetica et Cartographica Sinica*, vol. 45, no. 8, pp. 883–889, 2016.
- [16] D. Lin and J. Zhu, "Improved ridge estimation with singular value correction constraints," *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, vol. 42, no. 12, pp. 1834–1839, 2017.
- [17] Y. Shen, P. Xu, and B. Li, "Bias-corrected regularized solution to inverse ill-posed models," *Journal of Geodesy*, vol. 86, no. 8, pp. 597–608, 2012.
- [18] S. Sun, *Research on Morbid Problems of gps Baseline Solution*, Northeastern University, 2014.
- [19] G. H. Golub and C. F. Van Loan, *Matrix Computations*, vol. 3, Johns Hopkins University Press, Baltimore, MD, USA, 1983.
- [20] S. L. Ye and J. J. Zhu, "Application of singular value decomposition and generalized ridge estimation in surveying," *The Chinese Journal of Nonferrous Metals*, vol. 8, no. 1, pp. 160–164, 1998.

Submit your manuscripts at
www.hindawi.com