mathematics

Article

Application of LADMM and As-LADMM for a High-Dimensional Partially Linear Model

Aifen Feng *, Xiaogai Chang, Jingya Fan and Zhengfen Jin

School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, China
* Correspondence: faf@haust.edu.cn

Abstract: This paper mainly studies the application of the linearized alternating direction method of multiplier (LADMM) and the accelerated symmetric linearized alternating direction method of multipliers (As-LADMM) for high dimensional partially linear models. First, we construct a l_1-penalty for the least squares estimation of partially linear models under constrained contours. Next, we design the LADMM algorithm to solve the model, in which the linearization technique is introduced to linearize one of the subproblems to obtain an approximate solution. Furthermore, we add the appropriate acceleration techniques to form the As-LADMM algorithm and to solve the model. Then numerical simulations are conducted to compare and analyze the effectiveness of the algorithms. It indicates that the As-LADMM algorithm is better than the LADMM algorithm from the view of the mean squared error, the number of iterations and the running time of the algorithm. Finally, we apply them to the practical problem of predicting Boston housing price data analysis. This indicates that the loss between the predicted and actual values is relatively small, and the As-LADMM algorithm has a good prediction effect.

Keywords: partially linear model; l_1-penalty estimation; LADMM; As-LADMM

MSC: 90C25; 90C30; 62J05; 90C06

check for updates

Citation: Feng, A.; Chang, X.; Fan, J.; Jin, Z. Application of LADMM and As-LADMM for a High-Dimensional Partially Linear Model. *Mathematics* **2023**, *11*, 4220. https://doi.org/10.3390/math11194220

Academic Editor: Ionut Florescu

Received: 11 September 2023
Revised: 3 October 2023
Accepted: 7 October 2023
Published: 9 October 2023

1. Introduction

With the development of information and intelligence in the era of big data, the analysis of high-dimensional data has become an important research topic [1,2]. The relationships between variables of high-dimensional data are diverse and complex with the partially linear model being one of the most important relationships among them [3,4], and some research results have been achieved [5]. Various methods have been proposed for variable selection and estimation in high-dimensional partially linear models, such as the SCAD-penalized method [6], and the selection method via the lasso [7,8]. Ma et al. [9] studied the properties of Lasso in high-dimensional partially linear models. The selection method via profile and restricted profile estimation method [10–12]. Lian et al. [13,14], Guo et al. [15], and Wu et al. [16] have all conducted research on variable selection in partially linear additive models, and the other regression and variable selection methods [17–19], such as quantile regression and spline estimation.

This paper considers the following partially linear model (PLM)

$$Y = X^{\mathrm{T}}\beta + B^{\mathrm{T}}\gamma + \varepsilon \tag{1}$$

where $Y \in R$ is a response variable, $X = (X_1, \ldots, X_p)^{\mathrm{T}} \in R^p$ and $Z \in R$ is explanatory variable. $\beta = (\beta_1, \ldots, \beta_p)^{\mathrm{T}} \in R^p$ is parameter vector, $B = B(Z) = (B_1(Z), \ldots, B_{m_n}(Z))^T$ is a set of B-spline basis functions of order r, $\gamma = (\gamma_1, \ldots, \gamma_{m_n})^T$ is the spline coefficient vector, ε is a random error. The parametric part of the model uses a linear model, and the nonparametric part uses the B-spline basis function [20] method to estimate the unknown function, which combines the advantages of the interpretability of the linear model and the flexibility of the nonparametric model.

In [11], restricted profile estimation was proposed for partially linear models with large-dimensional covariates, and solved by the Lagrange multiplier method [21]. In [12], the alternating-direction method of multipliers (ADMM) to solve the model by constructing an augmented Lagrangian function. The ADMM algorithm was studied in [22,23].

Now, we will further investigate this partially linear model. In practically estimating β and γ, it is possible that the model is more complex and there is overfitting, or all training sets can fit well to obtain results, but it does not have generalization ability. Specifically, as the sparsity of the covariates in the parametric part, l_1-norm regularization term can be added to increase the generalization ability of the model. Therefore, we will study the l_1-penalty of partially linear models. We mainly consider the linearized alternating direction method of multiplier(LADMM) and the accelerated symmetric linearized alternating direction method of multipliers (As-LADMM) algorithm when solving a partially linear model.

The linearized alternating direction method of multiplier (LADMM) is studied by [24], and the linearization technique is introduced to linearize one of the subproblems to obtain an approximate solution. As appropriate acceleration techniques are added to the optimization algorithm, the rate of convergence of the algorithm can be effectively improved, such as the Nesterov acceleration technique [25,26]. An accelerated linearized alternating direction method of multipliers (AADMM) was proposed in [27,28], which combines multi-step acceleration schemes into linearized ADMM, and demonstrated that AADMM has a better convergence rate than ADMM. A symmetric ADMM (s-ADMM) is proposed in [29], which is an easy-to-implement strategy for accelerating ADMM. This strategy can be immediately applied to various practical examples. In [30], an inexact accelerated random alternating direction multiplier (AS-ADMM) scheme for separable convex optimization of linearly constrained structures was proposed.

Considering the sparsity of the covariates in the parametric part, by reducing the complexity of the model and avoiding the problem of overfitting, we study the l_1-penalty of partially linear models. Since the subproblems generated by ADMM must have analytical expression in each iteration process, but not all subproblems have analytical expression, the approximate solution of the subproblems is obtained by using the linearization method, LADMM and As-LDMM for solving the model.

This paper is organized as follows. In Section 2, we construct l_1-penalty estimation of the high-dimensional partially linear model. In Section 3, we employ LADMM to solve the l_1-penalty model of L1PLM. In Section 4, we apply the As-LADMM algorithm for l_1-Penalty estimation in a high-dimensional partially linear model. In Section 5, some numerical illustrations are reported. Finally, we apply them to the practical problem.

2. l_1-Penalty Estimation for High-Dimensional Partially Linear Model

Supposing $(Y_1; X_1^T; T_1), \cdots, (Y_n; X_n^T; T_n)$ is an independent homogeneous sample of the model, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_\mathrm{n})^\mathrm{T}$, and then model (1) can be written as

$$Y_i = X_i{}^\mathrm{T}\beta + B(Z_i)^\mathrm{T}\gamma + \varepsilon_i. \tag{2}$$

In order to estimate β and γ, Wang [11] studied restricted profile least squares estimation by using the Lagrange multiplier method for the following optimal problem

$$\begin{aligned} &\min_{\beta,\gamma} \tfrac{1}{2}||Y - X^\mathrm{T}\beta - B(Z)^\mathrm{T}\gamma||^2, \\ &s.t. R\beta = d. \end{aligned} \tag{3}$$

where R is a given $k \times p$ matrix whose rank is k, and d is a known k-dimensional vector. The augmented Lagrange function was constructed to transform constrained optimization problems into unconstrained optimization problems, the ADMM algorithm was applied for this high-dimensional partially linear model in [12].

In practice, when the covariates are sparse in the parametric part of partially linear models, we can add a regularization term to partially linear models. The l_1-norm regularization term can be added to increase the generalization ability of the model. On the one hand, l_1-norm regularization can obtain sparse solutions. That is to say, many dimensions of the parameter vector have values of zero. The existence of a sparse solution discards features that do not affect the results or have weak effects in the sample, and it can effectively simplify the model. On the other hand, the obtained parameter vector is not unique, and it is more likely that the absolute value of a single dimension within the parameter vector is particularly large. In this case, weak changes in the characteristics of a certain dimension can cause pathological changes in the results, so a l_1-penalty of parameter vector has been added to the objective function, which can effectively prevent this situation. Therefore, due to the sparsity of the covariates in the parametric part, by introducing the l_1-penalty, the resulting model can not only cater to the training set but also be simple and effective. Therefore, we study the l_1-penalty of partially linear models.

The objective functions of estimating β and γ using l_1-penalty least squares method is

$$\min_{\beta,\gamma} \frac{1}{2}||Y - X^{\mathrm{T}}\beta - B(Z)^{\mathrm{T}}\gamma||^2 + \theta||\beta||_1 \tag{4}$$

Therefore, we study the following optimization problem of estimating β and γ, denoted L1PLM:

$$\begin{aligned} &\min_{\beta,\gamma} \tfrac{1}{2}||Y - X^{\mathrm{T}}\beta - B(Z)^{\mathrm{T}}\gamma||^2 + \theta||\beta||_1 \\ &s.t. R\beta = d \end{aligned}$$

3. LADMM Algorithms of l_1-Penalty Estimation for High-Dimensional Partially Linear Model

In this section, we apply LADMM to solve the l_1-penalty model of L1PLM and provide an algorithm framework for solving the problem.

3.1. Solution of l_1-Penalty Estimation for High-Dimensional Partially Linear Model Using LADMM

For the optimization problem L1PLM, by using the augmented Lagrange multiplier method, the constrained programming problem is transformed into an unconstrained programming problem, and the augmented Lagrange function is

$$\min Q_\rho(\beta,\gamma,\lambda) = \frac{1}{2}\|Y - X\beta - B\gamma\|^2 + \theta||\beta||_1 + \langle \lambda, R\beta - d\rangle + \frac{\rho}{2}\|R\beta - d\|^2. \tag{5}$$

Using the classical alternating direction method of multiplier (ADMM), its n-step iteration starts from the given (β^n, λ^n) and iterates to obtain a new iteration point $(\gamma^{n+1}, \beta^{n+1}, \lambda^{n+1})$ via the following scheme

$$\begin{cases} \gamma^{n+1} = \arg\min\limits_{\gamma} Q_\rho(\beta^n, \gamma, \lambda^n) \\ \beta^{n+1} = \arg\min\limits_{\beta} Q_\rho(\beta, \gamma^{n+1}, \lambda^n) \\ \lambda^{n+1} = \lambda^n + \rho(R\beta^{n+1} - d) \end{cases} \tag{6}$$

Now, let us solve these subproblems.

Firstly, for the solution of the subproblem, the problem can be written as follows

$$\gamma^{n+1} = \arg\min_{\gamma} \{\frac{1}{2}\|Y - X\beta^n - B\gamma\|^2\}.$$

Since β^n can be given, taking the partial derivative of γ

$$\frac{\partial Q_\rho(\beta,\gamma,\lambda)}{\partial\gamma} = -B(Y - X\beta^n - B\gamma) = 0$$

The analytical solution of γ is

$$\gamma^{n+1} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}(Y - X\beta^n) \tag{7}$$

Secondly, for the solution of β-subproblem, by substituting γ into Equation (5), we can write the problem as follows

$$\begin{aligned}\beta^{n+1} &= \arg\min_{\beta} \{\frac{1}{2}\left\|Y - X\beta - B\gamma^{n+1}\right\|^2 + \theta||\beta||_1 + \langle\lambda^n, R\beta - d\rangle + \frac{\rho}{2}\|R\beta - d\|^2\} \\ &= \arg\min_{\beta} \{\theta||\beta||_1 + \frac{1}{2}\left\|Y - X\beta - B\gamma^{n+1}\right\|^2 + \frac{\rho}{2}\left\|R\beta - d + \frac{\lambda^n}{\rho}\right\|^2\} \\ &= \arg\min_{\beta} \{\theta||\beta||_1 + \frac{1}{2}\left\|\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y}\right\|^2\}\end{aligned} \tag{8}$$

where $\hat{X} = (X^{\mathrm{T}}, \sqrt{\rho}R)^{\mathrm{T}}$, $\hat{B} = (B^{\mathrm{T}}, 0)$.

Since $\hat{X}$ may be a non-positive-definite matrix, there is no closed-form solution for this subproblem. We can use the LADMM method by the linearized quadratic term as follows.

$$\frac{1}{2}\left\|\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y}\right\|^2 \cong (\hat{X}^{\mathrm{T}}(\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y}))^{\mathrm{T}}(\beta - \beta^n) + \frac{\nu}{2}||\beta - \beta^n||^2 \tag{9}$$

Therefore, the solved subproblem is equivalent to

$$\begin{aligned}\beta^{n+1} &= \arg\min_{\beta} \{\theta||\beta||_1 + (\hat{X}^{\mathrm{T}}(\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y}))^{\mathrm{T}}(\beta - \beta^n) + \frac{\nu}{2}||\beta - \beta^n||^2\} \\ &= \arg\min_{\beta} \{\theta||\beta||_1 + \frac{\nu}{2}||\beta - \beta^n + \frac{\hat{X}^{\mathrm{T}}(\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y})}{\nu}||^2\}\end{aligned} \tag{10}$$

The closed-form solution of this subproblem can be obtained by using the method of soft threshold

$$\beta^{n+1} = \mathrm{shrink}_{1,2}\{\beta^n - \frac{\hat{X}^{\mathrm{T}}(\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y})}{\nu}, \frac{\theta}{\nu}\} = \tilde{u}^n - P_C(\tilde{u}^n). \tag{11}$$

where $\tilde{u}^n = \beta^n - \frac{\hat{X}^{\mathrm{T}}(\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y})}{\nu}, C = [-\frac{\theta}{\nu}, \frac{\theta}{\nu}]$.

By solving the γ, β, λ subproblem separately, we obtained $\gamma^{n+1}, \beta^{n+1}, \lambda^{n+1}$ as follows

$$\begin{cases} \gamma^{n+1} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}(Y - X\beta^n) \\ \beta^{n+1} = \tilde{u}^n - P_C(\tilde{u}^n) \\ \lambda^{n+1} = \lambda^n + \rho(R\beta^{n+1} - d) \end{cases} \tag{12}$$

3.2. LADMM Algorithm Design for l_1-Penalty Estimation High-Dimensional Partially Linear Model

Based on the characteristics of the solution of l_1-penalty estimation for the high-dimensional partially linear model, the algorithm scheme using LADMM is as follows (Algorithm 1).

Algorithm 1 Iterative Scheme of LADMM for LIPLM

Step 1: Input X, Y, B. Given the initial variables $(\beta^0, \gamma^0, \lambda^0)$.
Choose penalty parametric $\rho > 0, \theta > 0$. Let $n = 1$ be iteration;
Step 2: Update γ^{n+1}, β^{n+1}, λ^{n+1} by Equation (12);
Step 3: If the algorithm does not meet the termination criteria at N-th iteration, let $n = n + 1$ go to Step 2; otherwise, go to the next step;
Step4: Output $(\beta^N, \gamma^N, \lambda^N)$ is the approximate solution of $(\hat{\beta}, \hat{\gamma}, \hat{\lambda})$.

We can prove that the LADMM algorithm should converge to the optimal solution under certain conditions, see references [28,29].

4. As-LADMM Algorithm for l_1-Penalty Estimation in High-Dimensional Partially Linear Model

In this section, we apply As-LADMM to solve the l_1-penalty model of L1PLM and provide an algorithm framework for solving the problem. He B. el studied the symmetric version of ADMM with larger step sizes and provided an easily implementable strategy to accelerate the ADMM numerically that can be immediately applied to a variety of applications [29]. We use the symmetric version of ADMM for solving the l_1-penalty model of L1PLM.

4.1. The Solution of l_1-Penalty Estimation for High-Dimensional Partially Linear Model by As-LADMM

In order to solve the optimization problem L1PLM using As-LADMM, the augmented Lagrange function is constructed

$$\min \phi_\rho(\beta,\gamma,\lambda) = \frac{1}{2}\|Y - X\beta - B\gamma\|^2 + \theta||\beta||_1 + \langle \lambda, R\beta - d\rangle + \frac{\rho}{2\eta}\|R\beta - d\|^2 \tag{13}$$

For a given (β^n, λ^n), we obtain $(\gamma^{n+1}, \beta^{n+1}, \lambda^{n+1})$ by the following iteration scheme

$$\begin{cases} \nu^n = \beta^n + \frac{\eta^n(1-\eta^{n-1})}{\eta^{n-1}}(\beta^n - \beta^{n-1}) \\ \gamma^{n+1} = \arg\min\limits_{\gamma} \phi_\rho(\nu^n, \gamma, \lambda^n) \\ \lambda^{n+\frac{1}{2}} = \lambda^n + \rho\tau(R\nu^{n+1} - d) \\ \beta^{n+1} = \arg\min\limits_{\beta} \phi_\rho(\beta, \gamma^{n+1}, \lambda^{n+\frac{1}{2}}) \\ \lambda^{n+1} = \lambda^{n+\frac{1}{2}} + \rho\tau(R\beta^{n+1} - d) \end{cases} \tag{14}$$

Now, let us solve these subproblems.
Firstly, for the solution of the γ-subproblem, the problem can be written as follows

$$\gamma^{n+1} = \arg\min_{\gamma} \{\frac{1}{2}\|Y - X\nu^n - B\gamma\|^2\}$$

Taking the partial derivative of γ

$$\frac{\partial\phi_\rho(\beta,\gamma,\lambda)}{\partial\gamma} = -B(Y - X\nu^n - B\gamma) = 0$$

The analytical expression of γ is

$$\gamma^{n+1} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}(Y - X\nu^n) \tag{15}$$

Secondly, for the solution of β-subproblems, the problem can be written as follows

$$\begin{aligned}
\beta^{n+1} &= \arg\min_{\beta}\{\frac{1}{2}\left\|Y - X\beta - B\gamma^{n+1}\right\|^2 + \theta||\beta||_1 + \left\langle \lambda^{n+\frac{1}{2}}, R\beta - d\right\rangle + \frac{\rho}{2\eta^n}\|R\beta - d\|^2\} \\
&= \arg\min_{\beta}\{\theta||\beta||_1 + \frac{1}{2}\left\|Y - X\beta - B\gamma^{n+1}\right\|^2 + \frac{\rho}{2\eta^n}\left\|R\beta - d + \frac{\lambda^{n+\frac{1}{2}}\eta^n}{\rho}\right\|^2\} \\
&= \arg\min_{\beta}\{\theta||\beta||_1 + \frac{1}{2}\left\|\widehat{X}\beta + \widehat{B}\gamma^{n+1} - \widehat{Y}\right\|^2\}.
\end{aligned}$$

where $\widehat{X} = (X^{\mathrm{T}}, \sqrt{\frac{\rho}{\eta^n}}R)^{\mathrm{T}}, \widehat{Y} = (Y^{\mathrm{T}}, \sqrt{\frac{\rho}{\eta^n}}(d - \frac{\lambda^{n+\frac{1}{2}}\eta^n}{\rho}))^{\mathrm{T}}, \widehat{B} = (B^{\mathrm{T}}, 0)$.

Since $\hat{X}$ may be a non-positive-definite matrix, there is no closed-form solution to this subproblem. The quadratic term $\frac{1}{2}\|\hat{X}\beta + \hat{B}\gamma^{n+1} - \hat{Y}\|^2$ can be replaced by linearized. LADMM method can be used. Therefore, the solved subproblem is equivalent to

$$\begin{aligned}
\beta^{n+1} &= \arg\min_{\beta}\{\theta||\beta||_1 + (\widehat{X}^{\mathrm{T}}(\widehat{X}\beta + \widehat{B}\gamma^{n+1} - \widehat{Y}))^{\mathrm{T}}(\beta - \beta^n) + \frac{\nu}{2}||\beta - \beta^n||^2\} \\
&= \arg\min_{\beta}\{\theta||\beta||_1 + \frac{\nu}{2}||\beta - \beta^n + \frac{\widehat{X}^{\mathrm{T}}(\widehat{X}\beta + \widehat{B}\gamma^{n+1} - \widehat{Y})}{\nu}||^2\}.
\end{aligned}$$

The closed-form solution of this subproblem can be obtained by using the method of soft threshold

$$\beta^{n+1} = \mathrm{shrink}_{1,2}\{\beta^n - \frac{\widehat{X}^{\mathrm{T}}(\widehat{X}\beta + \widehat{B}\gamma^{n+1} - \widehat{Y})}{\nu}, \frac{\theta}{\nu}\} = \tilde{v}^n - P_C(\tilde{v}^n) \tag{16}$$

where

$$\tilde{v}^n = \beta^n - \frac{\widehat{X}^{\mathrm{T}}(\widehat{X}\beta + \widehat{B}\gamma^{n+1} - \widehat{Y})}{\nu}, C = [-\frac{\theta}{\nu}, \frac{\theta}{\nu}].$$

Overall, by solving the γ, β, λ subproblems separately, the solutions $\gamma^{n+1}, \beta^{n+1}, \lambda^{n+1}$ are obtained. Therefore, its algorithm iteration framework is

$$\begin{cases}
\gamma^{n+1} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}(Y - X\nu^n), \\
\beta^{n+1} = \tilde{v}^n - P_C(\tilde{v}^n), \\
\lambda^{n+1} = \lambda^n + \rho\tau(R\nu^{n+1} - d) + \rho\tau(R\beta^{n+1} - d).
\end{cases} \tag{17}$$

4.2. As-LADMM Algorithm Design for l_1-Penalty Estimation in High-Dimensional Partial Linear Model

Based on the idea of the As-LADMM solution for l_1-penalty estimation in a high-dimensional partially linear model, we design an algorithm framework as follows (Algorithm 2).

Algorithm 2 Iterative Scheme of As-LADMM for LIPLM

Step 1: Input X, Y, B and tol. Given the initial variables $(\beta^0, \gamma^0, \lambda^0)$,
choose $\rho > 0, \theta^0 = -1, \theta^{-1} = \frac{1}{\tau}, 0.5 < \tau < 1$. Let $n = 1$;
Step 2: Update γ^{n+1}, β^{n+1}, λ^{n+1} by Equation (17);
Step 3: If the algorithm meets the termination criteria at N-th iteration, go to the next step.
Otherwise, let $n = n + 1$ go to Step 2; ;
Step 4: Output $(\beta^N, \gamma^N, \lambda^N)$. It is the approximate solution of $(\hat{\beta}, \hat{\gamma}, \hat{\lambda})$.

We can similarly prove that this algorithm should converge to the optimal solution under certain conditions, see references [28,29].

5. Numerical Simulation

5.1. Parameter Setting

Numerical simulation is performed for high-dimensional partial linear model estimation, with a sample size of n generated by the model. Here, $\varepsilon \sim N(0,\sigma^2)$, data set X follows the p-dimensional multivariate normal distribution, $X \sim N(0,\Sigma)$, $\Sigma = 0.5^{|j-k|}$, j and k is the jth and kth components of the covariance, respectively. $Z \sim U(0,1)$, $g(z) = 3\cos(2\pi z)$, $\beta = (1,2,0.5,-1,0,\ldots,0)^{\mathrm{T}}$ and $\beta_5 = \ldots = \beta_p = 0$. Parameter $\tau = 0.95$, the smoothness function is estimated by cubic spline interpolation, and the cubic B-spline basis function is used for numerical simulation. The results show that the effect of the smoothness is good.

5.2. Simulation Results

The simulation effect is expressed by the mean squared error (mse), objective value (obj), iteration times (iter) and running time (time) of the algorithm, where the sample size is taken as n = 100, 200, p = 109, 209, 409, 509, 1009.

$$mse = \|\hat{\beta} - \beta\|^2 \quad obj = \frac{1}{2}||Y - X\hat{\beta} - B\gamma||^2 + \theta||\hat{\beta}||_1$$

By determining the value of sample size and dimensionality from small to large, we study the effectiveness of simulation in high dimensional ($p >> n$) situations.

Based on the above parameters settings, the specific results are obtained by using the LADMM algorithm are shown in Table 1; The specific results obtained using the As-LDMM algorithm are shown in Table 2.

Table 1. Simulation results of LADMM in different parameters set.

n	p	σ	mse	obj	Iter	Time
100	109	0.5	0.0028	0.9912	189	0.023358
	209		0.0026	0.4000	71	0.017601
	509		0.0022	0.1675	24	0.010753
	1009		0.0021	0.1444	18	0.011368
200	209		0.0031	1.2943	262	0.032774
	409		0.0026	0.3700	46	0.014811
	509		0.0026	0.3205	34	0.011893
	1009		0.0024	0.2101	22	0.011691
100	109	1	0.0030	1.1306	230	0.026888
	209		0.0028	0.4549	76	0.015361
	509		0.0020	0.2015	25	0.010161
	1009		0.0020	0.1735	18	0.010081
200	209		0.0039	1.6863	329	0.041064
	409		0.0027	0.4233	50	0.014497
	509		0.0026	0.4001	40	0.012453
	1009		0.0025	0.2704	23	0.012234
100	109	2	0.0040	1.6747	196	0.028308
	209		0.0033	0.5872	86	0.014744
	509		0.0023	0.3052	26	0.013412
	1009		0.0024	0.2498	20	0.010456
200	209		0.0062	2.8058	372	0.044700
	409		0.0031	0.5956	56	0.015718
	509		0.0030	0.5985	46	0.012268
	1009		0.0027	0.4011	26	0.012453

According to the results of Tables 1 and 2, in the high-dimensional case, for a fixed σ, the mean squared error of parameter estimation for these algorithm decreases with the increase in dimension p, and the mean squared error of the As-LADMM algorithm is slightly lower than that of the LADMM algorithm. It indicates that the As-LADMM algorithm is better than the LADMM algorithm.

Compared to the LADMM algorithm, the As-LADMM algorithm performs accelerated symmetry transformation to improve the performance of the algorithm. For a fixed value of p, the mean squared error, the objective value, the number of iteration and the running time of the algorithm increases with the increase in σ. However, the As-LDMM algorithm performs better than the LADMM algorithm.

We draw a comparison line of mean squared error and objective value under different variances with the sample size of 100 and 200, specifically comparing and expressing the effectiveness of the As-LDMM algorithm with Figures 1–12.

Table 2. Simulation results of As-LADMM in different parameters set.

n	p	σ	mse	obj	Iter	Time
100	109	0.5	0.0028	0.9879	189	0.020590
	209		0.0026	0.3958	72	0.014416
	509		0.0019	0.1626	26	0.014585
	1009		0.0019	0.1435	17	0.011211
200	209		0.0031	1.2944	262	0.035734
	409		0.0026	0.3708	46	0.018151
	509		0.0026	0.3216	34	0.015442
	1009		0.0023	0.2097	19	0.012723
100	109	1	0.0030	1.1288	230	0.023713
	209		0.0028	0.4502	77	0.014860
	509		0.0020	0.2013	24	0.013338
	1009		0.0017	0.1718	16	0.010901
200	209		0.0039	1.6853	329	0.037359
	409		0.0027	0.4133	50	0.015312
	509		0.0027	0.3881	40	0.015518
	1009		0.0024	0.2635	25	0.012891
100	109	2	0.0040	1.6727	296	0.026368
	209		0.0032	0.5740	86	0.015204
	509		0.0023	0.2957	26	0.012608
	1009		0.0021	0.2470	19	0.011138
200	209		0.0062	2.8041	372	0.042947
	409		0.0032	0.5755	56	0.016475
	509		0.0032	0.5474	53	0.015811
	1009		0.0022	0.3819	30	0.013330

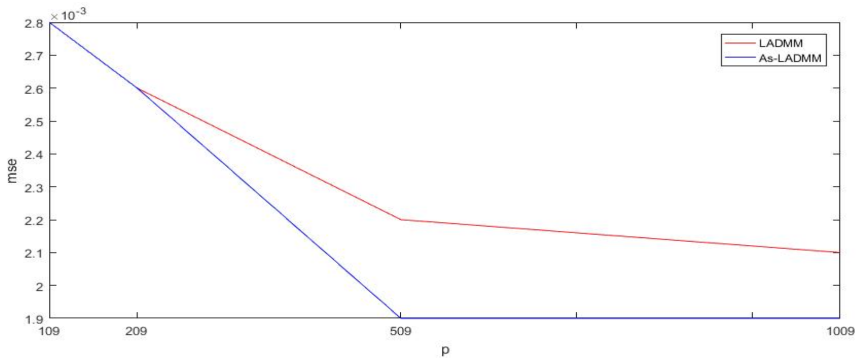

Figure 1. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 0.5$.

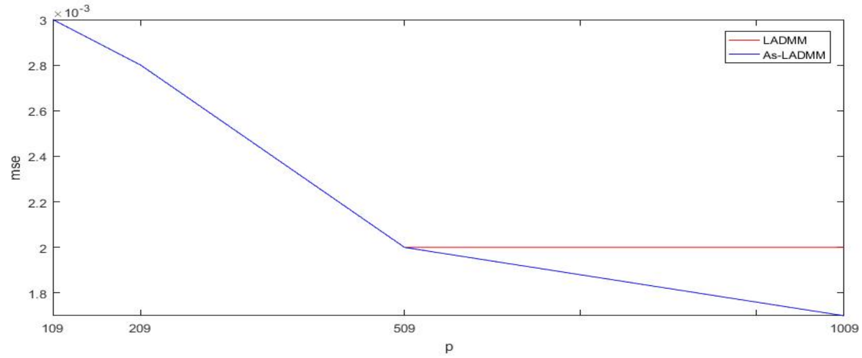

Figure 2. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 1$.

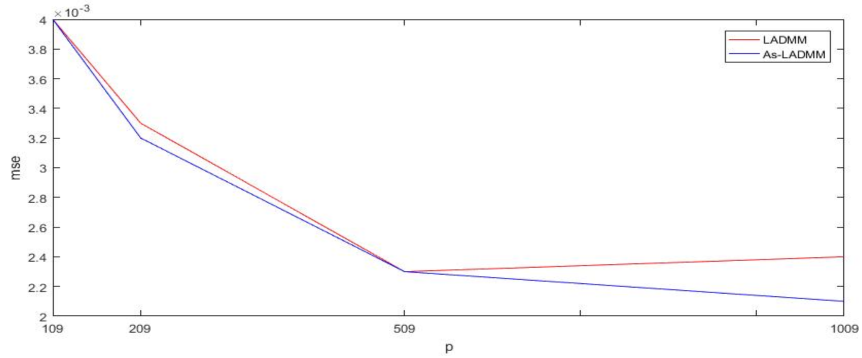

Figure 3. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 2$.

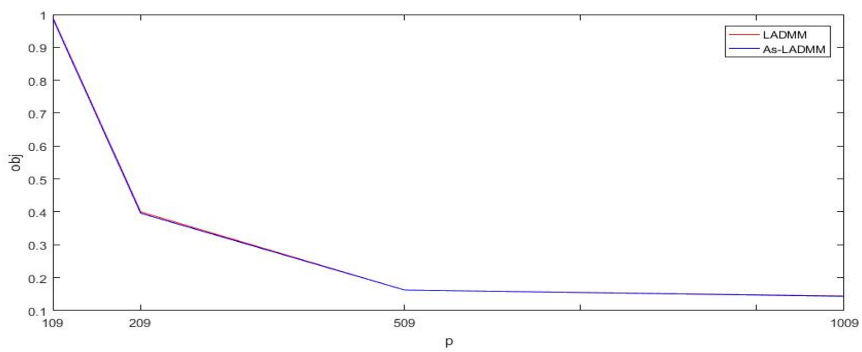

Figure 4. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 0.5$.

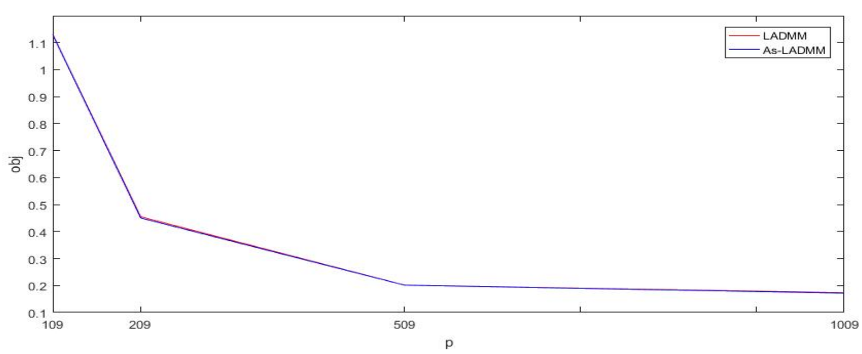

Figure 5. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 1$.

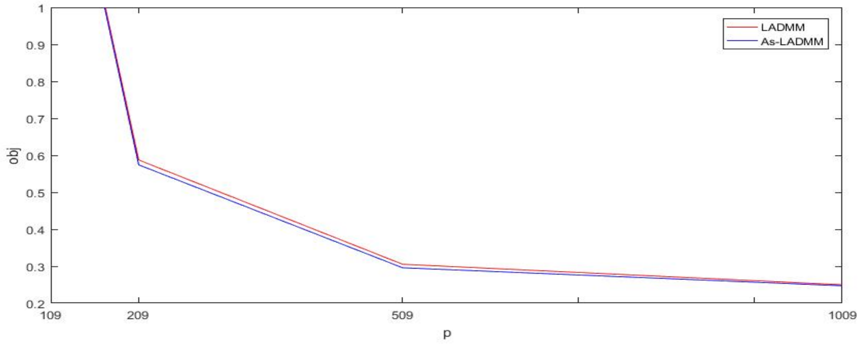

Figure 6. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 100$, $\sigma = 2$.

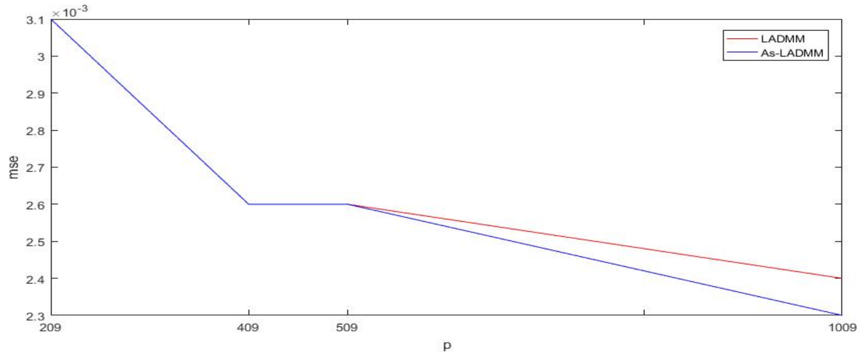

Figure 7. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 0.5$.

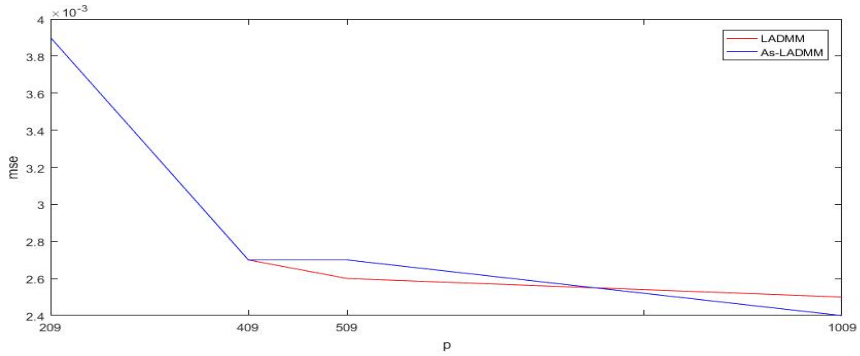

Figure 8. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 1$.

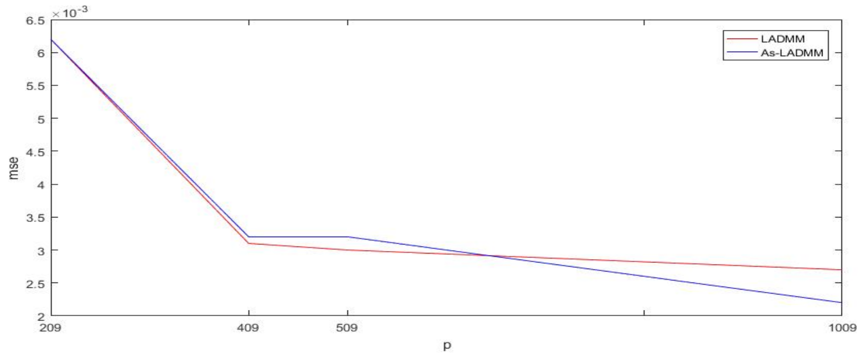

Figure 9. Comparison of mean squared error line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 2$.

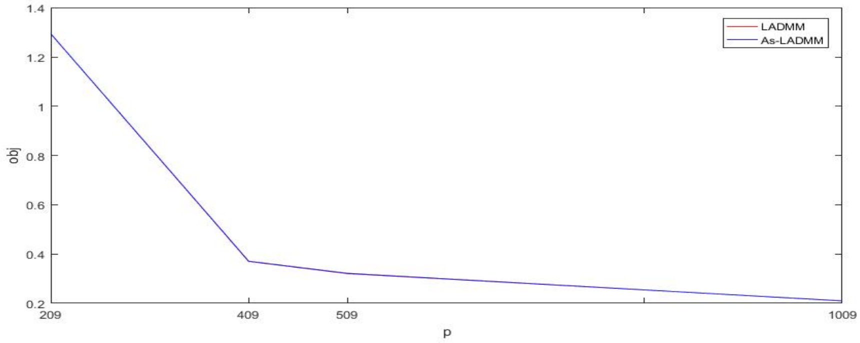

Figure 10. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 0.5$.

It can be seen from these figures that the mean squared error of the LADMM algorithm and the As-LADMM algorithm is very small and nearly zero, but the mean squared error of the As-LADMM algorithm is smaller than that. The objective values of the LADMM algorithm and the As-LDMM algorithm are both small, while the objective values of the As-LDMM algorithm are smaller. Therefore, it indicates that the As-LADMM algorithm has better performance and is suitable for solving high-dimensional partially linear models.

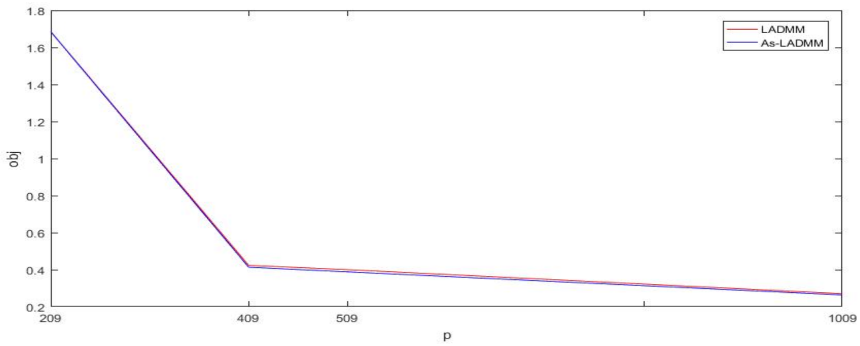

Figure 11. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 1$.

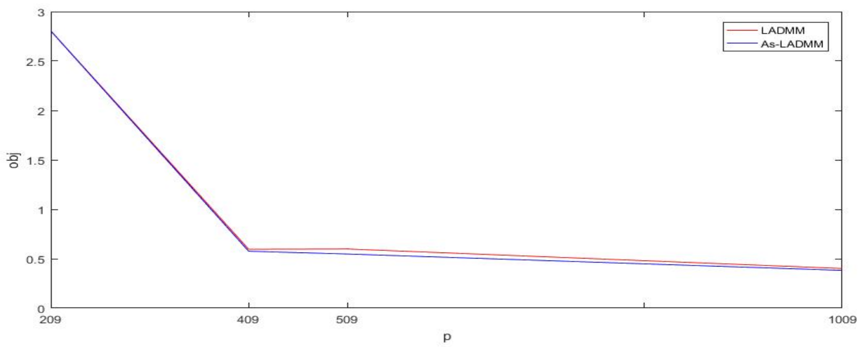

Figure 12. Comparison of objective value line of LADMM algorithm and As-LADMM algorithm under $n = 200$, $\sigma = 2$.

6. Application: Boston Housing Price Data Analysis

In order to verify the application of As-LADMM in high-dimensional data, we take Boston house price data as an example to analyze. The Boston home price data were information about home prices in Boston, Massachusetts, collected by the U. S. Census Bureau in the 1970 U. S. Census. It is obtained from the StatLib archive (http://lib.stat.cmu.edu/datasets/boston, accessed on 6 March 2023). The data are more representative of the actual situation [31,32]. The data set is composed of 13 input variables and one output variable that is the median value of owner-occupied homes in $1000's (MEDV). The input variables include per capita crime rate by town (CRIM), the proportion of residential land zoned for lots over 25,000 sq.ft. (ZN), the proportion of non-retail business acres per town (INDUS), nitric oxides concentration (parts per 10 million) (NOX), the average number of rooms per dwelling (RM), the proportion of owner-occupied units built prior to 1940 (AGE), weighted distances to five Boston employment centers (DIS), index of accessibility to radial highways(RAD), and full-value property-tax rate per $10,000 (TAX). In addition, PTRATIO is the pupil–teacher ratio by town; LSTAT is the lower status of the population. For thirteen independent variables, the PTRATIO is not necessarily linear with the proportion of MEDV. Therefore, the importance of other independent variables to MEDV is mainly considered. The model is as follows

$$Y_i = \sum_{j}^{12} X_{ij}\beta_j + m(U_i) + \varepsilon_i$$

where Y_i is MEDV of the i-th sample , and U_i is PTRATIO of the i-th sample, X_{ij} is the j-th variable of the i-th sample, and $\varepsilon_i \sim N(0, \sigma^2)$.

During the experiment, we selected 75% of the training samples and 25% of the test samples. The predictive value of MEDV in the prediction sample is expressed in y, with median absolute error (MAE) and standard error (SE) to evaluate the predictive ability of the model. The MAE can reduce the impact of outliers. The SE reflects the degree to which the sample deviates from the average value, and the smaller the value, the more reliable the method. The calculation method is

$$MAE = medain\{|y_1 - \hat{y}_1|, |y_2 - \hat{y}_2|, \cdots, |y_n - \hat{y}_n|\}.$$

$$SE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}.$$

The prediction ability of the As-LADMM algorithm was compared with LADMM, as shown in Table 3.

Table 3. MAE and SE effects of housing price forecasts for owner-occupied housing.

Variable	LADMM		AS-LADMM	
	MAE	**SE**	**MAE**	**SE**
CRIM	1.4785	0.0748	0.972	0.0504
ZN	1.4809	0.0752	0.9717	0.0549
INDUS	1.48	0.0782	0.9712	0.0544
CHAS	1.4474	0.0748	0.9726	0.0563
NOX	1.5046	0.0787	0.9612	0.0535
RM	1.5063	0.0776	0.977	0.0549
AGE	1.4715	0.0775	0.9649	0.0466
DIS	1.45596	0.0748	0.97	0.0523
RAD	1.4776	0.0746	0.9708	0.0538
TAX	1.4552	0.0747	0.9712	0.0524
B	1.4848	0.0752	0.9727	0.0513
LSTAT	1.2362	0.0598	0.8972	0.0323

From the results in Table 3, it can be seen that the values of MAE predicted using the As-LADMM algorithm are basically lower than those predicted by the LADMM algorithm. This indicates that the loss between the predicted and actual values of MEDV results is relatively small, and the As-LADMM algorithm has a good prediction effect. The values of SE are all below 0.09, and the corresponding As-LADMM algorithm predicts smaller values, indicating that the As-LADMM algorithm has a more reliable prediction ability than the LADMM algorithm. The overall performance is very good, so the parameter estimation method in this article is relatively effective.

7. Conclusions

In this paper, we mainly studied the application of LADMM and As-LADMM for high-dimensional partially linear models. As the sparsity of the covariates in the parametric part, we added to l_1-norm regularization term to estimate the parametric and increase the generalization ability of the model. We constructed the augmented Lagrange function to transform the constrained optimization problems into unconstrained optimization problems and solved the model using LADMM and As-LADMM. Through numerical simulation, we compared and analyzed the superiority of the designed algorithm. From the view of the mean squared error, the number of iterations and the running time of the algorithm, the As-LADMM algorithm is better than the LADMM algorithm. Finally, the two algorithms were applied to Boston housing price data, and the comparison showed the effectiveness of As-LADMM as well.

Author Contributions: Conceptualization, A.F. and X.C.; methodology, X.C. and A.F.; validation, X.C.; formal analysis, A.F. and Z.J.; data curation, X.C. and J.F.; writing—original draft preparation, X.C.; writing—review and editing, A.F., X.C. and J.F.; supervision, A.F. and Z.J. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded in part by the National Natural Science Foundation of China (Grant No. 12101195, 12071112).

Data Availability Statement: Not applicable.

Acknowledgments: The research was funded in part by the Provincial first-class undergraduate curriculum project of mathematical models. We sincerely thank the anonymous reviewers for their insightful comments, which greatly improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.; Zhang, S.; Ma, S.; Zhang, Q. Tests for regression coefficients in high dimensional partially linear models. *Stat. Probab. Lett.* **2020**, *163*, 108772. [CrossRef] [PubMed]
2. Zhao, F.; Lin, N.; Zhang, B. A new test for high-dimensional regression coefficients in partially linear models. *Can. J. Stat.* **2023**, *51*, 5–18. [CrossRef]
3. Engle, R.F.; Granger, C.W.; Rice, J.; Weiss, A. Semiparametric estimates of the relation between weather and electricity sales. *J. Am. Stat. Assoc.* **1986**, *81*, 310–320. [CrossRef]
4. Heckman, N.E. Spline smoothing in a partly linear model. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1986**, *48*, 244–248. [CrossRef]
5. Härdle, W.; Liang, H.; Gao, J. *Partially Linear Models*; Springer Science & Business Media: Berlin/Heidelberg, Germany , 2000.
6. Xie, H.; Huang, J. SCAD-penalized regression in high-dimensional partially linear models. *Ann. Stat.* **2009**, *37*, 673–696. [CrossRef]
7. Tibshirani, R. Regression shrinkage and selection via the lasso: A retrospective. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2011**, *73*, 268–288. [CrossRef]
8. He, K.; Wang, Y.; Zhou, X.; Xu, H.; Huang, C. An improved variable selection procedure for adaptive Lasso in high-dimensional survival analysis. *Lifetime Data Anal.* **2019**, *25*, 569–585. [CrossRef]
9. Ma, C.; Huang, J. Asymptotic properties of lasso in high-dimensional partially linear models. *Sci. China Math.* **2016**, *59*, 769–788. [CrossRef]
10. Chuanhua, W.; Xizhi, W. Profile Lagrange multiplier test for partially linear varying-coefficient regression models. *J. Syst. Sci. Math. Sci.* **2008**, *28*, 416.
11. Wang, X.; Zhao, S.; Wang, M. Restricted profile estimation for partially linear models with large-dimensional covariates. *Stat. Probab. Lett.* **2017**, *128*, 71–76. [CrossRef]
12. Feng, A.; Chang, X.; Shang, Y.; Fan, J. Application of the ADMM Algorithm for a High-Dimensional Partially Linear Model. *Mathematics* **2022**, *10*, 4767. [CrossRef]
13. Lian, H. Variable selection in high-dimensional partly linear additive models. *J. Nonparametr. Stat.* **2012**, *24*, 825–839. [CrossRef]
14. Lian, H.; Liang, H.; Ruppert, D. Separation of covariates into nonparametric and parametric parts in high-dimensional partially linear additive models. *Stat. Sin.* **2015**, *25*, 591–607.

15. Guo, J.; Tang, M.; Tian, M.; Zhu, K. Variable selection in high-dimensional partially linear additive models for composite quantile regression. *Comput. Stat. Data Anal.* **2013**, *65*, 56–67. [CrossRef]
16. Wu, Q.; Zhao, H.; Zhu, L.; Sun, J. Variable selection for high-dimensional partly linear additive Cox model with application to Alzheimer's disease. *Stat. Med.* **2020**, *39*, 3120–3134. [CrossRef]
17. Xu, H.X.; Chen, Z.L.; Wang, J.F.; Fan, G.L. Quantile regression and variable selection for partially linear model with randomly truncated data. *Stat. Pap.* **2019**, *60*, 1137–1160. [CrossRef]
18. Chen, W. Polynomial-based smoothing estimation for a semiparametric accelerated failure time partial linear model. *Open Access Libr. J.* **2020**, *7*, 1. [CrossRef]
19. Auerbach, E. Identification and estimation of a partially linear regression model using network data. *Econometrica* **2022**, *90*, 347–365. [CrossRef]
20. Liu, Y.; Yin, J. Spline estimation of partially linear regression models for time series with correlated errors. *Commun. Stat.-Simul. Comput.* **2021**, 1–15. [CrossRef]
21. Powell, M.J. A method for nonlinear constraints in minimization problems. In *Optimization*; Academic Press: New York, NY, USA, 1969; pp. 283–298.
22. Wahlberg, B.; Boyd, S.; Annergren, M.; Wang, Y. An ADMM algorithm for a class of total variation regularized estimation problems. *IFAC Proc. Vol.* **2012**, *45*, 83–88. [CrossRef]
23. Glowinski, R.; Song, Y.; Yuan, X.; Yue, H. Application of the alternating direction method of multipliers to control constrained parabolic optimal control problems and beyond. *Ann. Appl. Math.* **2022**, *38*, 115–158. [CrossRef]
24. Li, X.; Mo, L.; Yuan, X.; Zhang, J. Linearized alternating direction method of multipliers for sparse group and fused LASSO models. *Comput. Stat. Data Anal.* **2014**, *79*, 203–221. [CrossRef]
25. Nesterov, Y. *Introductory Lectures on Convex Optimization: A Basic Course*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2003; Volume 87.
26. Nesterov, Y. Smooth minimization of non-smooth functions. *Math. Program.* **2005**, *103*, 127–152. [CrossRef]
27. Lin, Z.; Li, H.; Fang, C. *Accelerated Optimization for Machine Learning*; Springer: Singapore, 2020.
28. Ouyang, Y.; Chen, Y.; Lan, G.; Pasiliao, E., Jr. An accelerated linearized alternating direction method of multipliers. *SIAM J. Imaging Sci.* **2015**, *8*, 644–681. [CrossRef]
29. He, B.; Ma, F.; Yuan, X. Convergence study on the symmetric version of ADMM with larger step sizes. *SIAM J. Imaging Sci.* **2016**, *9*, 1467–1501. [CrossRef]
30. Bai, J.; Hager, W.W.; Zhang, H. An inexact accelerated stochastic ADMM for separable convex optimization. *Comput. Optim. Appl.* **2022**, *81*, 479–518. [CrossRef]
31. Berndt, E.R. *The Practice of Econometrics: Classic and Contemporary*; Addison-Wesley Pub. Co.: San Francisco, CA, USA, 1991.
32. Harrison, D., Jr.; Rubinfeld, D.L. Hedonic housing prices and the demand for clean air. *J. Environ. Econ. Manag.* **1978**, *5*, 81–102. [CrossRef]