Application Median

Application to Quantile Regression

Conclusion

Future Research Robust LASSO and Efficient Quantile Regression through Regularization of Case-Specific Parameters

Yoonsuh Jung

Department of Statistics The Ohio State University

Jan 21, 2010

Introduction

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

- 2) Methodology
- 3 Application to LASSO
 - Application to Median Regression
- 5 Application to Quantile Regression

・ロト ・聞ト ・ヨト ・ヨト

- 6 Conclusion
- 7 Future Research

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

1 Introduction

2 Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

6 Conclusion

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

Introduction

- 2 Methodology
- 3 Application to LASSO
 - Application to Median Regression
 - 5 Application to Quantile Regression

・ロッ ・雪 ・ ・ ヨ ・ ・ ヨ ・

-

- 6 Conclusion
- 7 Future Research

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

Introduction

- 2 Methodology
- 3 Application to LASSO
- 4 Application to Median Regression
 - Application to Quantile Regression

- 6 Conclusion
- 7 Future Research

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

Introduction

- 2 Methodology
- 3 Application to LASSO
- 4 Application to Median Regression
- 5 Application to Quantile Regression

★ロト ★課 ト ★注 ト ★注 ト 一注

- 6 Conclusion
- 7 Future Research

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

Introduction

- 2 Methodology
- 3 Application to LASSO
- 4 Application to Median Regression
- 5 Application to Quantile Regression

★ロト ★課 ト ★注 ト ★注 ト 一注

- 6 Conclusion
 - 7 Future Research

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

Introduction

- 2 Methodology
- 3 Application to LASSO
- 4 Application to Median Regression
- 5 Application to Quantile Regression

◆□▶ ◆圖▶ ◆臣▶ ◆臣▶ 三臣

- 6 Conclusion
- 🕖 Future Research

Introduction: Brief Overview

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

3/43

Motivation

• To reduce the effect of outlying or influential cases

2 Method

• Consider a modeling procedure under the frame of loss function minimization

イロト 不得 トイヨト イヨト

• Modify the current modeling procedure by adding case-specific parameters

③ Conjectured results

- Decrease the potential impacts of outliers
- Attain robustness and/or efficiency

Introduction: Brief Overview

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research

Motivation

- To reduce the effect of outlying or influential cases
- 2 Method
 - Consider a modeling procedure under the frame of loss function minimization

• Modify the current modeling procedure by adding case-specific parameters

3 Conjectured results

- Decrease the potential impacts of outliers
- Attain robustness and/or efficiency

Introduction: Brief Overview

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research

Motivation

- To reduce the effect of outlying or influential cases
- 2 Method
 - Consider a modeling procedure under the frame of loss function minimization
 - Modify the current modeling procedure by adding case-specific parameters
- **3** Conjectured results
 - Decrease the potential impacts of outliers
 - Attain robustness and/or efficiency

Introduction: Case-specific Parameter and Regularization

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

① Case-specific parameter in the linear model

$$y_i = \beta_0 + x_i^\top \beta + \gamma_i + \epsilon_i$$

Or in matrix notation,

$$Y = X\beta + I\gamma + \epsilon$$
$$= (X I) \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \epsilon$$

Dimension: Ill posed problem

• Regularization method is needed

Introduction: Case-specific Parameter and Regularization

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research ① Case-specific parameter in the linear model

$$y_i = \beta_0 + x_i^\top \beta + \gamma_i + \epsilon_i$$

Or in matrix notation,

$$Y = X\beta + I\gamma + \epsilon$$
$$= (X I) \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \epsilon$$

Dimension: Ill posed problem
 Regularization method is needed

Introduction: Case-specific Parameter and Regularization

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research ① Case-specific parameter in the linear model

$$y_i = \beta_0 + x_i^\top \beta + \gamma_i + \epsilon_i$$

Or in matrix notation,

$$Y = X\beta + I\gamma + \epsilon$$
$$= (X I) \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \epsilon$$

(日)、(四)、(E)、(E)、(E)

2 Dimension: Ill posed problem

• Regularization method is needed

1 Standard regularization of location model

$$L(\beta) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta) + \lambda_{\beta} J_1(\beta)$$

2 Add case-specific parameters

$$L(\beta,\gamma) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\beta} J_1(\beta) + \lambda_{\gamma} J_2(\gamma)$$

3) With \hat{eta} , find the minimizer, $\hat{\gamma}$, of

 $L(\hat{\beta},\gamma) = \sum_{i=1}^{n} g(r_i - \gamma_i) + \lambda_{\gamma} J_2(\gamma), \text{ where } r_i = y_i - x_i^{\top} \hat{\beta}$

(a) With $\hat{\gamma}$, iterate step 2 and 3 until convergence.

Outline

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

1 Standard regularization of location model

$$L(\beta) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta) + \lambda_{\beta} J_1(\beta)$$

2 Add case-specific parameters

$$L(\beta,\gamma) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\beta} J_1(\beta) + \frac{\lambda_{\gamma} J_2(\gamma)}{\lambda_{\gamma} J_2(\gamma)}$$

3) With \hat{eta} , find the minimizer, $\hat{\gamma}$, of

 $L(\hat{\beta},\gamma) = \sum_{i=1}^{n} g(r_i - \gamma_i) + \lambda_{\gamma} J_2(\gamma), \text{ where } r_i = y_i - x_i^{\top} \hat{\beta}$

(a) With $\hat{\gamma}$, iterate step 2 and 3 until convergence.

Outline

Introduction

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

1 Standard regularization of location model

$$L(\beta) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta) + \lambda_{\beta} J_1(\beta)$$

2 Add case-specific parameters

$$L(\beta,\gamma) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\beta} J_1(\beta) + \frac{\lambda_{\gamma} J_2(\gamma)}{\lambda_{\gamma} J_2(\gamma)}$$

Solution With $\hat{\beta}$, find the minimizer, $\hat{\gamma}$, of

$$L(\hat{eta},\gamma) = \sum_{i=1}^{n} g(r_i - \gamma_i) + \lambda_{\gamma} J_2(\gamma), \text{ where } r_i = y_i - x_i^{\top} \hat{eta}$$

(d) With $\hat{\gamma}$, iterate step 2 and 3 until convergence.

Outline

Introduction

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

1 Standard regularization of location model

$$L(\beta) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta) + \lambda_{\beta} J_1(\beta)$$

2 Add case-specific parameters

$$L(\beta,\gamma) = \sum_{i=1}^{n} g(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\beta} J_1(\beta) + \frac{\lambda_{\gamma} J_2(\gamma)}{\lambda_{\gamma} J_2(\gamma)}$$

(3) With $\hat{\beta}$, find the minimizer, $\hat{\gamma}$, of

 $L(\hat{\beta},\gamma) = \sum_{i=1}^{''} g(r_i - \gamma_i) + \lambda_{\gamma} J_2(\gamma), \text{ where } r_i = y_i - x_i^{\top} \hat{\beta}$

• With $\hat{\gamma}$, iterate step 2 and 3 until convergence.

Outline

Introduction

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

LASSO: Modification Procedure I

Standard LASSO

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

$$L(\beta) = \frac{1}{2} (Y - X\beta)^{\top} (Y - X\beta) + \lambda_{\beta} \sum_{j=1}^{p} |\beta_j|$$

2 Robust LASSO

$$L(\beta,\gamma) = \frac{1}{2} \{Y - (X\beta + \gamma)\}^{\top} \{Y - (X\beta + \gamma)\}$$
$$+ \lambda_{\beta} \sum_{j=1}^{p} |\beta_{j}| + \lambda_{\gamma} \sum_{i=1}^{n} |\gamma_{i}|$$

Case-specific parameters and extra penalty are included

LASSO: Modification Procedure I

1 Standard LASSO

$$L(eta) = rac{1}{2}(Y - Xeta)^{ op}(Y - Xeta) + \lambda_eta \sum_{j=1}^p |eta_j|$$

Robust LASSO 2

$$L(\beta,\gamma) = \frac{1}{2} \{Y - (X\beta + \gamma)\}^{\top} \{Y - (X\beta + \gamma)\} + \lambda_{\beta} \sum_{j=1}^{p} |\beta_{j}| + \lambda_{\gamma} \sum_{i=1}^{n} |\gamma_{i}|$$

Case-specific parameters and extra penalty are included

Application to LASSO

LASSO: Modification Procedure II

 $0 With \hat{\beta},$

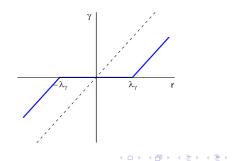
- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

$$L(\hat{\beta},\gamma) = \frac{1}{2}(r-\gamma)^{\top}(r-\gamma) + \lambda_{\hat{\beta}} \sum_{j=1}^{p} |\hat{\beta}_{j}| + \lambda_{\gamma} \sum_{i=1}^{n} |\gamma_{i}|$$

Minimizer is
$$\hat{\gamma} = sgn(r)(|r| - \lambda_{\gamma})_+$$



э

LASSO: Modification Procedure III

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

- Now with $\hat{\gamma}, L(\beta, \hat{\gamma})$ for ith observation becomes $\begin{cases} \frac{1}{2}(y_i - x_i^{\top}\beta)^2 + \lambda_{\beta} \sum_{j=1}^{p} |\beta_j| & \text{for}|y_i - x_i^{\top}\beta| < \lambda_{\gamma} \\ \frac{1}{2}\lambda_{\gamma}^2 + \lambda_{\gamma}(|y_i - x_i^{\top}\beta| - \lambda_{\gamma}) + \lambda_{\beta} \sum_{j=1}^{p} |\beta_j| & \text{otherwise} \end{cases}$
- ② Coincide with Huberized LASSO (Rosset & Zhu, 2004)
 ③ Conjecture : Achieve some Robustness

LASSO: Modification Procedure III

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research

• Now with $\hat{\gamma}, \mathcal{L}(\beta, \hat{\gamma})$ for i^{th} observation becomes

$$\begin{cases} \frac{1}{2}(y_{i} - x_{i}^{\top}\beta)^{2} + \lambda_{\beta}\sum_{j=1}^{p}|\beta_{j}| & \text{for}|y_{i} - x_{i}^{\top}\beta| < \lambda_{\gamma} \\ \frac{1}{2}\lambda_{\gamma}^{2} + \lambda_{\gamma}(|y_{i} - x_{i}^{\top}\beta| - \lambda_{\gamma}) + \lambda_{\beta}\sum_{j=1}^{p}|\beta_{j}| & \text{otherwise} \end{cases}$$

Coincide with Huberized LASSO (Rosset & Zhu, 2004) Conjecture : Achieve some Robustness

LASSO: Modification Procedure III

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research • Now with $\hat{\gamma}, \mathcal{L}(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} \frac{1}{2}(y_i - x_i^{\top}\beta)^2 + \lambda_{\beta}\sum_{j=1}^{p}|\beta_j| & \text{for}|y_i - x_i^{\top}\beta| < \lambda_{\gamma} \\ \frac{1}{2}\lambda_{\gamma}^2 + \lambda_{\gamma}(|y_i - x_i^{\top}\beta| - \lambda_{\gamma}) + \lambda_{\beta}\sum_{j=1}^{p}|\beta_j| & \text{otherwise} \end{cases}$$

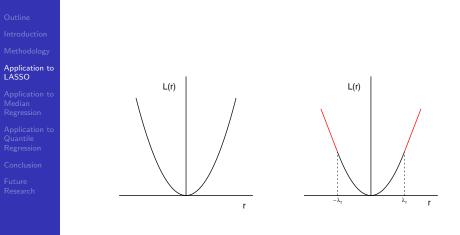
2 Coincide with Huberized LASSO (Rosset & Zhu, 2004)
3 Conjecture : Achieve some Robustness

Graphical Summary



▲口 ▶ ▲圖 ▶ ★ 臣 ▶ ★ 臣 ▶ →

э



Application to Language Data (Baayen, 2007)

Outline

Introduction

Methodology

Application to LASSO

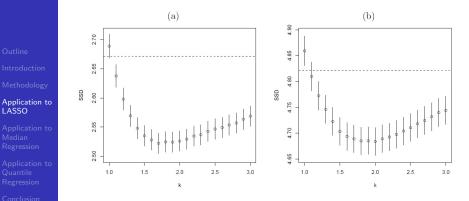
Application to Median Regression

Application to Quantile Regression

Conclusion

- ① Total sample size: about 4500 cases
- Variables: (small) 20 variables (large) 9 additional variables
- Simulated sample size: 400 cases
- 5000 replicates
- Calculate sum of squared deviations (SSD) from Baayens fitted values

Application to Language Data (Baayen, 2007)



Sum of squared deviations (SSD) from Baayens fits in the simulation study. The horizontal line is the mean SSD for the LASSO while the points represent the mean of SSDs for the robust LASSO. The vertical lines give approximate 95% confidence intervals for the mean SSDs. Panel (a) presents results for the small set of covariates and panel (b) presents results for the large set of covariates.

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

• λ_{β} is chosen by Robust Cp (Ronchetti and Staudte, 1994)

2) Set
$$\lambda_{\gamma} = k \cdot \sigma$$

3 Estimate σ with robust $\hat{\sigma}$ such as MAD

) (Huber; 1981) suggests $k \in [1,2]$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research λ_β is chosen by Robust Cp (Ronchetti and Staudte, 1994)
 Set λ_γ = k · σ

3 Estimate σ with robust $\hat{\sigma}$ such as MAD

④ (Huber; 1981) suggests $k \in [1,2]$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research λ_β is chosen by Robust Cp (Ronchetti and Staudte, 1994)
Set λ_γ = k · σ

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

③ Estimate σ with robust $\hat{\sigma}$ such as MAD

④ (Huber; 1981) suggests $k \in [1,2]$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • λ_{β} is chosen by Robust Cp (Ronchetti and Staudte, 1994)

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

2 Set
$$\lambda_{\gamma} = k \cdot \sigma$$

(

- $\textbf{S} \quad \text{Estimate } \sigma \text{ with robust } \hat{\sigma} \text{ such as MAD}$
- (Huber; 1981) suggests $k \in [1, 2]$

- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

() Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

2 Generated $x = (x_1, ..., x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.

Interest of the second seco

- a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
- b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
- c dense: $\beta_j = 0.85$ for all $j = 1, \ldots, 8$
- **5** Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma (\hat{\beta}^i \beta)$

- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

() Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

- **2** Generated $x = (x_1, \ldots, x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.
- 3 Three scenarios were considered
 - a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
 - b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
 - c dense: $\beta_j = 0.85$ for all $j = 1, \ldots, 8$
- **5** Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma (\hat{\beta}^i \beta)$

- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

() Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

2 Generated $x = (x_1, \ldots, x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.

O Three scenarios were considered

- a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
- b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
- c dense: $\beta_j = 0.85$ for all $j = 1, \ldots, 8$

 $\bullet \ \epsilon_i \sim N(0, 3^2)$

- **5** Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma (\hat{\beta}^i \beta)$

- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

(Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

2 Generated $x = (x_1, \ldots, x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.

O Three scenarios were considered

- a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
- b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
- c dense: $\beta_j = 0.85$ for all $j = 1, \dots, 8$

(4) $\epsilon_i \sim N(0, 3^2)$

- 5 Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma (\hat{\beta}^i \beta)$

- Outline
- Introduction
- Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

(Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

- **2** Generated $x = (x_1, \ldots, x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.
- 3 Three scenarios were considered
 - a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
 - b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
 - c dense: $\beta_j = 0.85$ for all $j = 1, \dots, 8$
- $\bullet_i \sim N(0, 3^2)$
- **5** Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma (\hat{\beta}^i \beta)$

Simulation Setting (Tibshirani, 1996)

- Outline
- Introduction
- Methodology

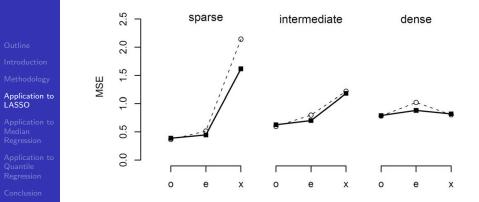
Application to LASSO

- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

(Standard linear model $y = x^{\top}\beta + \epsilon$ was assumed

- **2** Generated $x = (x_1, \ldots, x_8)^{\top}$ from $N(0, \Sigma)$, where $\rho_{i,j} = (0.5)^{|i-j|}$.
- O Three scenarios were considered
 - a Sparse: $\beta = (5, 0, 0, 0, 0, 0, 0, 0)$
 - b Intermediate: $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)$
 - c dense: $\beta_j = 0.85$ for all $j = 1, \dots, 8$
- $\bullet_i \sim N(0, 3^2)$
- **5** Contamination
 - No Contamination
 - first 5% ϵ_i were tripled
 - first 5% of x_1 were tripled
- $MSE = 100^{-1} \sum_{i=1}^{100} (\hat{\beta}^i \beta)^\top \Sigma(\hat{\beta}^i \beta)$

Simulation Results



In each scenario, o, e, and x indicate clean data, data with contaminated measurement errors, and data with mismeasured first covariate. The dotted lines are for LASSO while the solid lines are for Robust LASSO. The points are the average MSE from 100 data sets

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

1 Median Regression

$$L_{\lambda}(\beta) = \sum_{i=1}^{n} |y_i - x_i^{\top}\beta|$$

2 Modified Median Regression

$$L(\beta,\gamma) = \sum_{i=1}^{n} |y_i - x_i^{\top}\beta - \gamma_i| + \frac{\lambda_{\gamma}}{2} \sum_{i=1}^{n} \gamma_i^2$$

Case-specific parameters and extra penalty are included

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

1 Median Regression

$$L_{\lambda}(\beta) = \sum_{i=1}^{n} |y_i - x_i^{\top}\beta|$$

2 Modified Median Regression

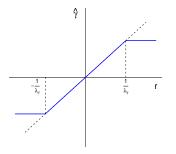
$$L(\beta,\gamma) = \sum_{i=1}^{n} |y_i - x_i^{\top}\beta - \gamma_i| + \frac{\lambda_{\gamma}}{2} \sum_{i=1}^{n} \gamma_i^2$$

Case-specific parameters and extra penalty are included

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨー

- Outline
- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

• With $\hat{\beta}$, $L(\hat{\beta}, \gamma) = \sum_{i=1}^{n} |r_i - \gamma_i| + \frac{\lambda_{\gamma}}{2} \sum_{i=1}^{n} \gamma_i^2$ • Minimizer, $\hat{\gamma} = sgn(r) \frac{1}{\lambda_{\gamma}} I(|r| > \frac{1}{\lambda_{\gamma}}) + rI(|r| \le \frac{1}{\lambda_{\gamma}})$



Outline Introduction

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusior

Future Research **1** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

- 2 Again, Huber's loss function
- 3 Quadratic adjustment of the V shape
- Conjecture : Achieve some Efficiency
- **(5)** Details of bending constant, λ_{γ} , come later
 - **1** Natural extension to Quantile Regression

Outline Introduction

Application to

Application to Median Regression

Application to Quantile Regression

Conclusior

Future Research **1** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

2 Again, Huber's loss function

3 Quadratic adjustment of the V shape

- ④ Conjecture : Achieve some Efficiency
- **(5)** Details of bending constant, λ_{γ} , come later
 - **()** Natural extension to Quantile Regression

Outline Introduction

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **()** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

- **2** Again, Huber's loss function
- **3** Quadratic adjustment of the V shape
- Conjecture : Achieve some Efficiency
- **5** Details of bending constant, λ_{γ} , come later
- **1** Natural extension to Quantile Regression

Outline Introduction

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusior

Future Research **1** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

- **2** Again, Huber's loss function
- **③** Quadratic adjustment of the V shape
- Conjecture : Achieve some Efficiency
- 5 Details of bending constant, λ_{γ} , come later
- Natural extension to Quantile Regression

Outline Introduction

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **1** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

- **2** Again, Huber's loss function
- **③** Quadratic adjustment of the V shape
- Conjecture : Achieve some Efficiency
- **(5)** Details of bending constant, λ_{γ} , come later

Natural extension to Quantile Regression

Outline Introduction

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusior

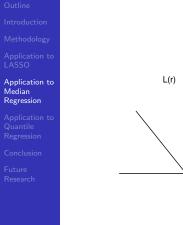
Future Research **1** With $\hat{\gamma}$, $L(\beta, \hat{\gamma})$ for ith observation becomes

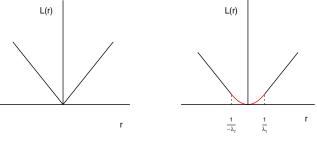
$$\begin{cases} |y_i - x_i^\top \beta| - \frac{1}{2\lambda_{\gamma}} & \text{for} |y_i - x_i^\top \beta| > \frac{1}{\lambda_{\gamma}} \\ \frac{\lambda_{\gamma}}{2} (y_i - x_i^\top \beta)^2 & \text{for} |y_i - x_i^\top \beta| \le \frac{1}{\lambda_{\gamma}} \end{cases}$$

- **2** Again, Huber's loss function
- **③** Quadratic adjustment of the V shape
- Conjecture : Achieve some Efficiency
- **(5)** Details of bending constant, λ_{γ} , come later
- **(6)** Natural extension to Quantile Regression

Graphical Summary

Loss Function: Median Regression and Efficient Median Regression





▲口 ▶ ▲圖 ▶ ★ 臣 ▶ ★ 臣 ▶ →

э

Quantile Regression: Introduction

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • Check loss function for estimating q^{th} regression quantile, 0 < q < 1 (Koenker& Bassett, 1978)

$$ho(u) = \left\{ egin{array}{cc} qu & ext{ for } u \geq 0 \\ (q-1)u & ext{ for } u < 0. \end{array}
ight.$$

2 Finding the minimizer is equivalent to finding the zero of its derivative,

$$\psi(u) = \left\{ egin{array}{cc} q & ext{for } u \geq 0 \ (q-1) & ext{for } u < 0. \end{array}
ight.$$

Quantile Regression: Introduction

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • Check loss function for estimating qth regression quantile, 0 < q < 1 (Koenker& Bassett, 1978) $\rho(u) = \begin{cases} qu & \text{for } u \ge 0\\ (q-1)u & \text{for } u < 0. \end{cases}$

Finding the minimizer is equivalent to finding the zero of its derivative,

$$\psi(u) = \left\{ egin{array}{cc} q & ext{for } u \geq 0 \ (q-1) & ext{for } u < 0. \end{array}
ight.$$

(日) (日) (日) (日) (日) (日) (日) (日) (日)

1 Modified quantile regression

$$L(\beta,\gamma) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\gamma} J_2(\gamma)$$

2 Consider asymmetric $J_2(\gamma)$

$$J_2(\gamma) = \frac{q}{1-q}\gamma^2 I(\gamma \ge 0) + \frac{1-q}{q}\gamma^2 I(\gamma < 0)$$

3 With $\hat{\beta}$, the minimizer of $L(\hat{\beta}, \gamma)$ is $\hat{\gamma} = -\frac{q}{2\lambda_{\gamma}}I(r < -\frac{q}{2\lambda_{\gamma}}) + rI(-\frac{q}{2\lambda_{\gamma}} \leq r < \frac{1-q}{2\lambda_{\gamma}}) + \frac{1-q}{2\lambda_{\gamma}}I(r \geq \frac{1-q}{2\lambda_{\gamma}})$

20/43

Application to Quantile Regression

1 Modified quantile regression

$$L(\beta,\gamma) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\gamma} J_2(\gamma)$$

2 Consider asymmetric $J_2(\gamma)$

$$J_2(\gamma) = \frac{q}{1-q}\gamma^2 I(\gamma \ge 0) + \frac{1-q}{q}\gamma^2 I(\gamma < 0)$$

3 With $\hat{\beta}$, the minimizer of $L(\hat{\beta}, \gamma)$ is $\hat{\gamma} = -\frac{q}{2\lambda\gamma}I(r < -\frac{q}{2\lambda\gamma}) + rI(-\frac{q}{2\lambda\gamma} \leq r < \frac{1-q}{2\lambda\gamma}) + \frac{1-q}{2\lambda\gamma}I(r \geq \frac{1-q}{2\lambda\gamma})$

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

1 Modified quantile regression

$$L(\beta, \gamma) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\gamma} J_2(\gamma)$$

2 Consider asymmetric $J_2(\gamma)$

$$J_2(\gamma) = rac{q}{1-q} \gamma^2 I(\gamma \ge 0) + rac{1-q}{q} \gamma^2 I(\gamma < 0)$$

3 With $\hat{\beta}$, the minimizer of $L(\hat{\beta}, \gamma)$ is

 $\hat{\gamma} = -\frac{q}{2\lambda_{\gamma}} l(r < -\frac{q}{2\lambda_{\gamma}}) + rl(-\frac{q}{2\lambda_{\gamma}} \leq r < \frac{1-q}{2\lambda_{\gamma}}) + \frac{1-q}{2\lambda_{\gamma}} l(r \geq \frac{1-q}{2\lambda_{\gamma}})$

Application to Quantile Regression

1 Modified quantile regression

$$L(\beta, \gamma) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \gamma_i) + \lambda_{\gamma} J_2(\gamma)$$

2 Consider asymmetric $J_2(\gamma)$

$$J_2(\gamma) = rac{q}{1-q} \gamma^2 I(\gamma \ge 0) + rac{1-q}{q} \gamma^2 I(\gamma < 0)$$

3 With $\hat{\beta}$, the minimizer of $L(\hat{\beta}, \gamma)$ is $\hat{\gamma} = -\frac{q}{2\lambda\gamma}I(r < -\frac{q}{2\lambda\gamma}) + rI(-\frac{q}{2\lambda\gamma} \leq r < \frac{1-q}{2\lambda\gamma}) + \frac{1-q}{2\lambda\gamma}I(r \geq \frac{1-q}{2\lambda\gamma})$

Application to Quantile Regression

() With $\hat{\gamma}$

Outline Introductio

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research $L(\beta, \hat{\gamma}) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \hat{\gamma}_i) + \lambda_{\gamma} J_2(\hat{\gamma})$ $= \sum_{i=1}^{n} \rho^M(y_i - x_i^{\top}\beta)$

(2) $\rho^{M}(u)$ is given by

$$\rho_{\gamma}^{M}(u) = \begin{cases} qu - \frac{q(1-q)}{4\lambda_{\gamma}} & \text{if } \frac{1-q}{2\lambda_{\gamma}} \le u\\ \lambda_{\gamma} \frac{q}{1-q} u^{2} & \text{if } 0 \le u < \frac{1-q}{2\lambda_{\gamma}}\\ \lambda_{\gamma} \frac{1-q}{q} u^{2} & \text{if } -\frac{q}{2\lambda_{\gamma}} \le u < 0\\ (q-1)u - \frac{q(1-q)}{4\lambda_{\gamma}} & \text{if } u < -\frac{q}{2\lambda_{\gamma}} \end{cases}$$

21/43

$\textcircled{0} \text{ With } \hat{\gamma}$

Outline Introductio

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

$$L(\beta, \hat{\gamma}) = \sum_{i=1}^{n} \rho(y_i - x_i^{\top}\beta - \hat{\gamma}_i) + \lambda_{\gamma} J_2(\hat{\gamma})$$
$$= \sum_{i=1}^{n} \rho^M(y_i - x_i^{\top}\beta)$$

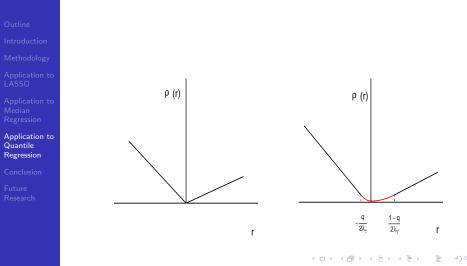
2 $\rho^{M}(u)$ is given by

$$\rho_{\gamma}^{M}(u) = \begin{cases} qu - \frac{q(1-q)}{4\lambda_{\gamma}} & \text{if } \frac{1-q}{2\lambda_{\gamma}} \leq u\\ \lambda_{\gamma} \frac{q}{1-q} u^{2} & \text{if } 0 \leq u < \frac{1-q}{2\lambda_{\gamma}}\\ \lambda_{\gamma} \frac{1-q}{q} u^{2} & \text{if } -\frac{q}{2\lambda_{\gamma}} \leq u < 0\\ (q-1)u - \frac{q(1-q)}{4\lambda_{\gamma}} & \text{if } u < -\frac{q}{2\lambda_{\gamma}} \end{cases}$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ● ● ●

Graphical Summary

Loss Function: Standard QR and Efficient QR



22/43

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • $\psi^{M}(u)$ is employed for computing purpose. rlm(MASS) procedure is used in R.

$$\psi^{M}(u) = \begin{cases} q & \text{if } \frac{1-q}{2\lambda_{\gamma}} \leq u \\ \frac{2\lambda_{\gamma} \frac{q}{1-q}u}{\frac{q}{q}} & \text{if } 0 \leq u < \frac{1-q}{2\lambda_{\gamma}} \\ \frac{2\lambda_{\gamma} \frac{1-q}{q}u}{\frac{q}{q}} & \text{if } -\frac{q}{2\lambda_{\gamma}} \leq u < 0 \\ (q-1) & \text{if } u < -\frac{q}{2\lambda_{\gamma}} \end{cases}$$

2 Recall the $\psi(u)$ for standard QR $\psi(u) = \begin{cases} q & \text{for } u \ge 0 \\ (q-1) & \text{for } u < 0 \end{cases}$

23/43

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • $\psi^{M}(u)$ is employed for computing purpose. rlm(MASS) procedure is used in R.

$$\psi^{M}(u) = \begin{cases} q & \text{if } \frac{1-q}{2\lambda_{\gamma}} \leq u \\ \frac{2\lambda_{\gamma}\frac{q}{1-q}u}{\frac{1-q}{q}} & \text{if } 0 \leq u < \frac{1-q}{2\lambda_{\gamma}} \\ \frac{2\lambda_{\gamma}\frac{1-q}{q}u}{\frac{1-q}{q}} & \text{if } -\frac{q}{2\lambda_{\gamma}} \leq u < 0 \\ (q-1) & \text{if } u < -\frac{q}{2\lambda_{\gamma}} \end{cases}$$

2 Recall the $\psi(u)$ for standard QR $\psi(u) = \begin{cases} q & \text{for } u \ge 0\\ (q-1) & \text{for } u < 0. \end{cases}$

Graphical Summary

ψ Function: Standard QR and Modified QR

Outline

.....

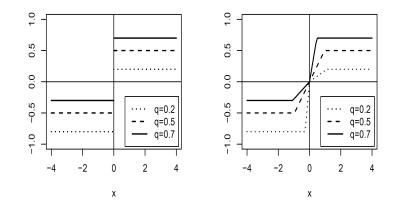
Application to LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

Future Research



Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research

Through simulation search for c which provides 'minimum' MSE.

Error distributions investigated are

- Standard Normal
- T distributions
- Gamma distributions
- Log-normal distribution
- Exponential distribution
- Sample sizes: 10², 10^{2.5}, 10³, 10^{3.5}, 10⁴
- 🕘 quantiles: 0.1, 0.2, ... ,0.9
- **5** The length of interval (of adjustment) is $\frac{1}{2\lambda_{\alpha}} = \frac{\hat{\sigma}}{2c \cdot n^{\alpha}}$

(日) (四) (王) (日) (日) (日)

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Through simulation search for c which provides 'minimum' MSE.

② Error distributions investigated are

- Standard Normal
- T distributions
- Gamma distributions
- Log-normal distribution
- Exponential distribution

3 Sample sizes: 10^2 , $10^{2.5}$, 10^3 , $10^{3.5}$, 10^4

- 🕘 quantiles: 0.1, 0.2, ... ,0.9
- **5** The length of interval (of adjustment) is $\frac{1}{2\lambda_n} = \frac{\hat{\sigma}}{2c_n n^{\alpha}}$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Through simulation search for c which provides 'minimum' MSE.

② Error distributions investigated are

- Standard Normal
- T distributions
- Gamma distributions
- Log-normal distribution
- Exponential distribution
- \bigcirc Sample sizes: $10^2, 10^{2.5}, 10^3, 10^{3.5}, 10^4$
 - 🕘 quantiles: 0.1, 0.2, ... ,0.9
- **5** The length of interval (of adjustment) is $\frac{1}{2\lambda_n} = \frac{\hat{\sigma}}{2c_n \sigma}$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Through simulation search for c which provides 'minimum' MSE.

② Error distributions investigated are

- Standard Normal
- T distributions
- Gamma distributions
- Log-normal distribution
- Exponential distribution
- **③** Sample sizes: 10^2 , $10^{2.5}$, 10^3 , $10^{3.5}$, 10^4
- quantiles: 0.1, 0.2, ..., 0.9

5 The length of interval (of adjustment) is $\frac{1}{2\lambda_{rr}} = \frac{\hat{\sigma}}{2c_{r}n^{0}}$

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

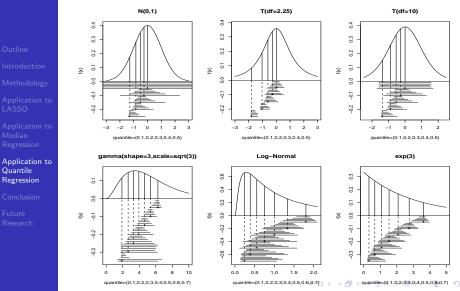
Future Research

25/43

Through simulation search for c which provides 'minimum' MSE.

- ② Error distributions investigated are
 - Standard Normal
 - T distributions
 - Gamma distributions
 - Log-normal distribution
 - Exponential distribution
- **③** Sample sizes: 10^2 , $10^{2.5}$, 10^3 , $10^{3.5}$, 10^4
- quantiles: 0.1, 0.2, ..., 0.9
- **5** The length of interval (of adjustment) is $\frac{1}{2\lambda_{\gamma}} = \frac{\hat{\sigma}}{2c \cdot n^{\alpha}}$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの



Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research • Again, the length of interval adjusted is $\frac{1}{2\lambda_{\gamma}} = \frac{\hat{\sigma}}{2c \cdot n^{\alpha}}$ • Given n, α , rule for c is,

$$\hat{c} = \left\{ \begin{array}{ll} e^{-2.118 - 1.097q} & \text{for } q < 0.5 \\ e^{-2.118 - 1.097(1-q)} & \text{for } q \ge 0.5 \end{array} \right.$$

which is developed from exponential error distribution.

Solution For computation : Embed the rule in the ψ^M(u) then use rlm(MASS) in R.

(日) (四) (王) (日) (日) (日)

Prediction ability?

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Again, the length of interval adjusted is ¹/_{2λγ} = ^ô/_{2c·n^α}
 Given n, α, rule for c is,

$$\hat{c} = \left\{ egin{array}{c} e^{-2.118 - 1.097 q} & ext{for } q < 0.5 \ e^{-2.118 - 1.097 (1 - q)} & ext{for } q \geq 0.5 \end{array}
ight.$$

which is developed from exponential error distribution.

Sor computation : Embed the rule in the ψ^M(u) then use rlm(MASS) in R.

◆□ → ◆圖 → ◆園 → ◆園 → □ ■ □

Prediction ability?

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Again, the length of interval adjusted is ¹/_{2λγ} = [∂]/_{2c·n^α}
 Given n, α, rule for c is,

$$\hat{c} = \left\{ egin{array}{c} e^{-2.118 - 1.097 q} & ext{for } q < 0.5 \ e^{-2.118 - 1.097 (1 - q)} & ext{for } q \geq 0.5 \end{array}
ight.$$

which is developed from exponential error distribution.

So For computation : Embed the rule in the $\psi^{M}(u)$ then use rlm(MASS) in R.

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 少へ⊙

Prediction ability?

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Again, the length of interval adjusted is ¹/_{2λγ} = ^ô/_{2c·n^α}
 Given n, α, rule for c is,

$$\hat{c} = \left\{ egin{array}{c} e^{-2.118 - 1.097 q} & ext{for } q < 0.5 \ e^{-2.118 - 1.097 (1 - q)} & ext{for } q \ge 0.5 \end{array}
ight.$$

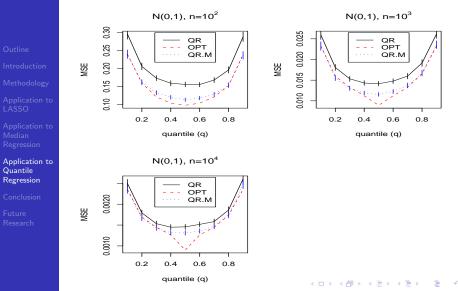
which is developed from exponential error distribution.

So For computation : Embed the rule in the $\psi^{M}(u)$ then use rlm(MASS) in R.

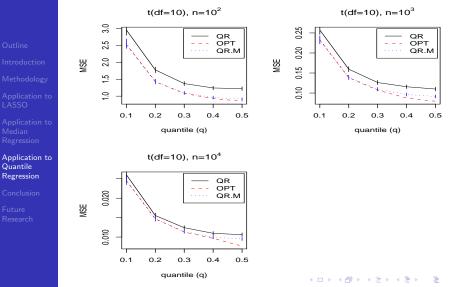
◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 少へ⊙

O Prediction ability?

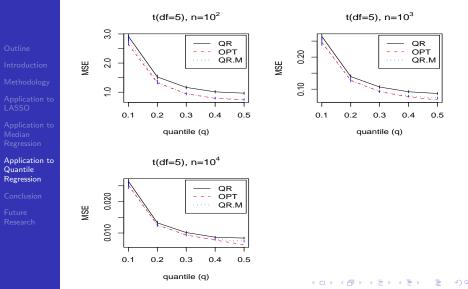
Quantile Regression: Prediction with the Rule on N(0,1) error distribution



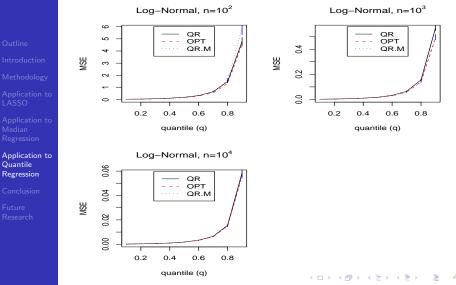
Quantile Regression: Prediction with the Rule on t(df=10) error distribution



Quantile Regression: Prediction with the Rule on t(df=5) error distribution

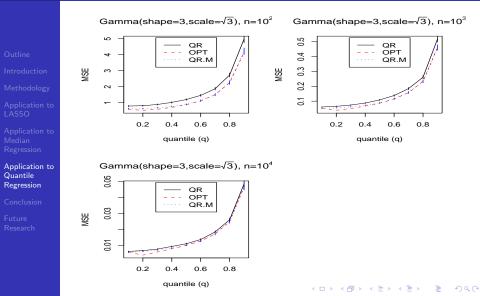


Quantile Regression: Prediction with the Rule on log-normal error distribution



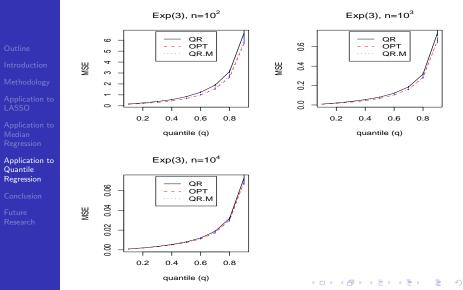
31/43

Quantile Regression: Prediction with the Rule on gamma error distribution

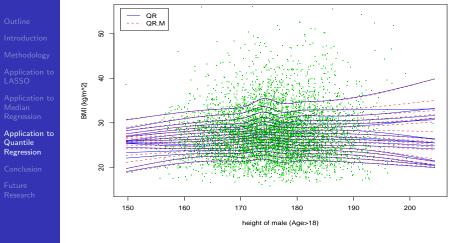


32/43

Quantile Regression: Prediction with the Rule on exponential error distribution

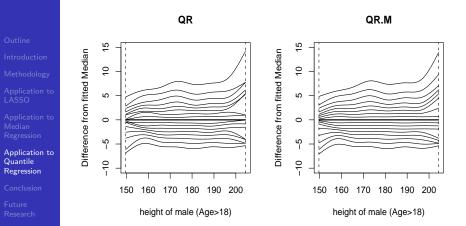


Application to NHANES Data I



Body Mass Index vs height from 5938 U.S. male (age>18)

Application to NHANES Data II



Difference between fitted median line and the other fitted quantiles.

Assumptions

- There exist continuous and differentiable densities, $f_i(\xi)$ uniformly bounded away from 0 and ∞ at ξ_i , i=1,2,....
- 2 $\lim_{n\to\infty} n^{-1} \sum x_i x_i^{\top} = D_0$, where D_0 is positive definite
- 3 $\lim_{n\to\infty} n^{-1} \sum f_i(\xi_i) x_i x_i^{\top} = D_1$, where D_1 is positive definite

Definition: Standard QR Estimator

$$\hat{\beta}_q = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho(y_i - x_i^\top \beta)$$
$$\sqrt{n}(\hat{\beta}_q - \beta) \to N(0, q(1-q)D_1^{-1}D_0D_1^{-1})$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

Outline

Introduction

Methodology

Application t LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Assumptions

- There exist continuous and differentiable densities, $f_i(\xi)$ uniformly bounded away from 0 and ∞ at ξ_i , i=1,2,....
- 2 $\lim_{n\to\infty} n^{-1} \sum x_i x_i^{\top} = D_0$, where D_0 is positive definite
- 3 $\lim_{n\to\infty} n^{-1} \sum f_i(\xi_i) x_i x_i^{\top} = D_1$, where D_1 is positive definite

Definition: Standard QR Estimator

$$\hat{\beta}_q = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho(y_i - x_i^\top \beta)$$
$$\bar{n}(\hat{\beta}_q - \beta) \to N(0, q(1-q)D_1^{-1}D_0D_1^{-1})$$

Outline

Introduction

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

Assumptions

٦

- There exist continuous and differentiable densities, $f_i(\xi)$ uniformly bounded away from 0 and ∞ at ξ_i , i=1,2,....
- 2 $\lim_{n\to\infty} n^{-1} \sum x_i x_i^{\top} = D_0$, where D_0 is positive definite
- 3 $\lim_{n\to\infty} n^{-1} \sum f_i(\xi_i) x_i x_i^{\top} = D_1$, where D_1 is positive definite

Definition: Standard QR Estimator

$$\hat{\beta}_{q} = \arg\min_{\beta \in \mathbb{R}^{p}} \sum_{i=1}^{n} \rho(y_{i} - x_{i}^{\top}\beta)$$

$$\sqrt{n}(\hat{\beta}_{q} - \beta) \rightarrow N(0, q(1-q)D_{1}^{-1}D_{0}D_{1}^{-1})$$

・ロット (雪) (日) (日) (日)

Outline

Introduction

Methodology

Application t LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

1 Modified QR Estimator

$$\hat{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho^M (y_i - x_i^\top \beta)$$

 $\sqrt{n}(\hat{eta}_{q,M}-eta) o N(0,q(1-q)D_1^{-1}D_0D_1^{-1}), ext{if } lpha > 1/3$

2 Modified QR Estimator under Location-Scale Family Model : y_i = x_i^Tβ + (x_i^Tτ)ε_i, where ε_i are iid

$$\check{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\gamma}^M((y_i - x_i^\top \beta) / x_i^\top \tau)$$

With \sqrt{n} -consistent estimator of τ (Koenker,Zhao: 1994) $\sqrt{n}(\check{\beta}_{q,M} - \beta) \rightarrow N(0, q(1-q)D_2^{-1}), \text{ if } \alpha > 1/3$

- Introduction
- Methodology
- Application to LASSO
- Application t Median Regression
- Application to Quantile Regression

Conclusion

1 Modified QR Estimator

$$\hat{eta}_{q,M} = \arg\min_{eta \in \mathbb{R}^p} \sum_{i=1}^n
ho^M (y_i - x_i^\top eta)$$

 $\sqrt{n}(\hat{eta}_{q,M}-eta) o N(0,q(1-q)D_1^{-1}D_0D_1^{-1}), ext{if } lpha > 1/3$

2 Modified QR Estimator under Location-Scale Family Model : y_i = x_i^Tβ + (x_i^Tτ)ε_i, where ε_i are iid

$$\check{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\gamma}^M((y_i - x_i^\top \beta) / x_i^\top \tau)$$

With \sqrt{n} -consistent estimator of τ (Koenker,Zhao: 1994) $\sqrt{n}(\check{\beta}_{q,M} - \beta) \rightarrow N(0, q(1-q)D_2^{-1}), \text{ if } \alpha > 1/3$

Methodology

Application to LASSO

Application t Median Regression

Application to Quantile Regression

Conclusion

1 Modified QR Estimator

$$\hat{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho^M (y_i - x_i^\top \beta)$$

 $\sqrt{n}(\hat{\beta}_{q,M}-\beta) \to N(0,q(1-q)D_1^{-1}D_0D_1^{-1}), \text{if } \alpha > 1/3$

2 Modified QR Estimator under Location-Scale Family Model : $y_i = x_i^{\top}\beta + (x_i^{\top}\tau)\epsilon_i$, where ϵ_i are iid

$$\check{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\gamma}^M((y_i - x_i^\top \beta) / x_i^\top \tau)$$

With \sqrt{n} -consistent estimator of τ (Koenker,Zhao: 1994) $\sqrt{n}(\check{\beta}_{q,M} - \beta) \rightarrow N(0, q(1-q)D_2^{-1}), \text{ if } \alpha > 1/3$

- Outline
- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression

Conclusion

1 Modified QR Estimator

$$\hat{eta}_{q,M} = rg\min_{eta \in \mathbb{R}^p} \sum_{i=1}^n
ho^M(y_i - x_i^ op eta)$$

 $\sqrt{n}(\hat{eta}_{q,M}-eta) o N(0,q(1-q)D_1^{-1}D_0D_1^{-1}), ext{if } lpha > 1/3$

2 Modified QR Estimator under Location-Scale Family Model : $y_i = x_i^{\top}\beta + (x_i^{\top}\tau)\epsilon_i$, where ϵ_i are iid

$$\check{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\gamma}^M((y_i - x_i^\top \beta) / x_i^\top \tau)$$

With \sqrt{n} -consistent estimator of τ (Koenker,Zhao: 1994) $\sqrt{n}(\check{\beta}_{q,M} - \beta) \rightarrow N(0, q(1-q)D_2^{-1}), \text{ if } \alpha > 1/3$

- Outline
- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression

Conclusion

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **1** Modified QR Estimator under Heterogeneous data

$$\widetilde{eta}_{q,M} = \arg\min_{eta \in \mathbb{R}^p} \sum_{i=1}^n f_i(\xi_i)
ho_{\gamma}^M(y_i - x_i^{ op} eta)$$

 $\sqrt{n}(\widetilde{eta}_{q,M}-eta)
ightarrow N(0,q(1-q)D_2^{-1}), ext{if } lpha>1/3$

2 Remark: D₁⁻¹D₀D₁⁻¹ > D₂⁻¹
3 Performance of \$\tilde{\beta}_{g,M}\$ in finite sample size?

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **1** Modified QR Estimator under Heterogeneous data

$$ilde{eta}_{q,M} = \arg\min_{eta \in \mathbb{R}^p} \sum_{i=1}^n f_i(\xi_i)
ho_{\gamma}^M(y_i - x_i^{ op} eta)$$

 $\sqrt{n}(\widetilde{eta}_{q,M}-eta)
ightarrow N(0,q(1-q)D_2^{-1}),$ if lpha>1/3

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨー

2 Remark: D₁⁻¹D₀D₁⁻¹ > D₂⁻¹
3 Performance of \$\tilde{\beta}_{g,M}\$ in finite sample size?

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **1** Modified QR Estimator under Heterogeneous data

$$\tilde{\beta}_{q,M} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n f_i(\xi_i) \rho_{\gamma}^M(y_i - x_i^\top \beta)$$

 $\sqrt{n}(\widetilde{eta}_{q,M}-eta)
ightarrow N(0,q(1-q)D_2^{-1}),$ if lpha>1/3

Remark: D₁⁻¹D₀D₁⁻¹ > D₂⁻¹
Performance of β_{a,M} in finite sample size?

Outline

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research **1** Modified QR Estimator under Heterogeneous data

$$ilde{eta}_{q,M} = rg\min_{eta \in \mathbb{R}^p} \sum_{i=1}^n f_i(\xi_i)
ho_{\gamma}^M(y_i - x_i^{ op}eta)$$

 $\sqrt{n}(\widetilde{eta}_{q,M}-eta)
ightarrow N(0,q(1-q)D_2^{-1}),$ if lpha>1/3

2 Remark: D₁⁻¹D₀D₁⁻¹ > D₂⁻¹ **3** Performance of β̃_{g,M} in finite sample size?

Quantile Regression: Simulation with Heterogeneous Data

- Outline
- Introduction
- Methodology
- Application t LASSO
- Application to Median Regression
- Application to Quantile Regression
- Conclusion
- Future Research

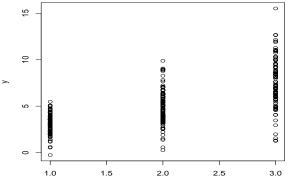
- Consider linear model $y_i = \beta_0 + \beta_1 x_i + x_i \epsilon_i$, ϵ_i 's are *iid* standard normal
- **2** $(\beta_0, \beta_1)^{\top} = (1, 2)^{\top}$
- **③** x is consist of three points; $x \in \{1, 2, 3\}$
- sample size n=300 and n=900 were made and evenly distributed at each design point (200 replicates)
- Standard QR, Modified QR(QR.M), Weighted QR (WQR), and Weighted QR.M (WQR.M) are compared





Conclusion

Future Research



х

▲ロト ▲母 ト ▲ ヨ ト ▲ ヨ - 一 の へ ()

Outline

Introduction

Methodology

Application to LASSO

Application to Median Regression

Application to Quantile Regression

Conclusion

Future Research Table: mean of *MSE* and its standard error in (·) from 200 replications with n=300 and n=900 at q^{th} quantile, multiplied by 1000.

	q=.1	q=.2	q=.3	q=.4	q=.5
			n=300		
QR	92.31(8.85)	66.68(5.47)	56.40(4.80)	44.77(3.57)	42.53(3.33)
QR.M	92.91(8.41)	62.37(5.19)	48.38(4.06)	37.48(2.94)	33.47(2.64)
WQR	85.49(8.32)	62.10(5.53)	54.93(4.97)	42.82(3.46)	42.32(3.48)
WQR.M	84.44(8.03)	57.93(5.19)	44.84(3.92)	35.55(2.89)	31.21(2.48)
			n=900		
RQ	28.93(2.44)	22.10(2.04)	18.06(1.56)	15.50(1.43)	14.41(1.25)
RQ.M	28.89(2.40)	21.09(1.95)	15.90(1.35)	13.13(1.14)	12.22(1.07)
WQR	27.77(2.31)	21.44(1.91)	17.73(1.55)	15.20(1.44)	14.42(1.21)
WRQ.M	28.28(2.59)	20.09(1.82)	15.30(1.28)	12.58(1.10)	11.54(1.00)

Concluding Remarks

Outline

- Introduction
- Methodology
- Application to LASSO
- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research

- 1. A new approach to treat cases
- 2. Regularization of case-specific parameter increases
 - Robustness in LASSO
 - **@** Efficiency and Robustness in Quantile Regression

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

3. Broadly applicable

Future Research

Outline

Introduction

Methodology

Application to LASSO

- Application to Median Regression
- Application to Quantile Regression

Conclusion

Future Research

Classification

- Logistic Regression
- Support Vector Machine

Oross Validation

• Find more accurate loss function

Thank You!