

## MASTER'S THESIS

### Variable selection for high dimensional transformation model

Lee, Wai Hong

*Date of Award:*  
2010

[Link to publication](#)

#### **General rights**

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# Variable Selection for High Dimensional Transformation Model

**LEE Wai Hong**

A thesis submitted in partial fulfillment of the requirements

for the degree of

Master of Philosophy

Principal Supervisor: Dr. PENG Heng

Hong Kong Baptist University

August 2010

# Abstract

Transformation Regression Model is an important semiparametric regression model. At the initial stage of the regression model building, a lot of variables are available which may result in a very large model, or make model misspecification. Hence Variable selection is a necessary process in practice. However most traditional variable selection methods are not available for the transformation model because of the unknown transformation function and some constraint identified conditions of the model. On the other hand, when the number of variables in the model is larger than the sample size, such as LASSO, SCAD, those penalized methods developed in recently are still suffered in computational burden, and difficult to be applied for the transformation model directly. This research aims to employ the general idea of SIS proposed by Fan and Lv (2008) and rank correlation to propose a new screening method for the ultrahigh dimensional transformation model. By such screening method, the dimension of the model can be reduced, then those penalized methods can be applied to select variables for the transformation model. Hence the two step variable selection procedure for the ultra-high dimension transformation model is proposed, that is applying a screener at first and followed by a non-parametric penalized estimation. Basic screening properties of the proposed method has been investigated and supported by the numerical study.

# Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	vii
List of Tables	viii
Chapter 1 Introduction	1
1.1 Transformation Model . . . . .	3
1.2 Variable Selection . . . . .	7
1.2.1 Ridge Regression . . . . .	9
1.2.2 Bridge Regression . . . . .	9
1.2.3 LASSO . . . . .	10
1.2.4 Elastic Net . . . . .	10
1.2.5 Nonconcave Penalized Likelihood Estimation . . . . .	11
1.3 The Issues of High Dimensional . . . . .	12

1.3.1	Curse of Dimensionality . . . . .	13
1.3.2	Dimensional Reduction . . . . .	13
1.4	Thesis Organization . . . . .	15
Chapter 2	Methodology	16
2.1	Penalized Method for Linear Model . . . . .	16
2.1.1	Penalty Function . . . . .	17
2.1.2	Selection Criteria . . . . .	22
2.1.3	A Unified Framework . . . . .	25
2.2	Variable Selection for Transformation Model . . . . .	28
2.2.1	Maximum Rank Correlation (MRC) . . . . .	28
2.2.2	Smoothed MRC . . . . .	29
2.2.3	Penalized Maximum Smoothing Rank Correlation . . . . .	29
2.3	Dimensional Reduction . . . . .	31
2.3.1	Sure Independence Screening (SIS) . . . . .	31
2.3.2	Nonparametric Rank Correlation . . . . .	32
2.4	RSIS-PMSRC . . . . .	35
Chapter 3	Screening Properties of RSIS for transformation model	37
3.1	U-Statistics . . . . .	39
3.2	An Exponential Bound for $\hat{\tau}$ . . . . .	40
3.3	Sure Screening Properties of RSIS . . . . .	41
3.4	Controlling False Selection Rates . . . . .	44
Chapter 4	Simulation Studies	46

4.1	Screening Property of Screeners . . . . .	46
4.2	The Performance of PMSRC . . . . .	50
4.3	The Performance RSIS-PMSRC . . . . .	50
Chapter 5 Conclusion		54
Appendix		56
Curriculum Vitae		62