

Nonparametric and Semiparametric Econometrics  
Lecture Notes for Econ 221

Yixiao Sun  
Department of Economics,  
University of California, San Diego

Winter 2007

# Contents

<b>Preface</b>	<b>ix</b>
<b>1 Kernel Smoothing: Density Estimation</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Kernel Estimator . . . . .	3
1.3 Bias and Variance . . . . .	6
1.4 Optimal Bandwidth Choice: Plug-in Approach . . . . .	10
1.5 Optimal Bandwidth Choice: Cross Validation . . . . .	12
1.6 Optimal Kernel . . . . .	15
1.7 Bias Reduction: High Order Kernels . . . . .	17
1.7.1 High-order Kernels . . . . .	17
1.7.2 Flat-top Kernels . . . . .	18
1.8 Asymptotic Normality . . . . .	21
1.9 Uniform Consistency . . . . .	25
1.10 Multivariate Density Estimation . . . . .	28
1.11 Conditional Density Estimation . . . . .	29
1.12 References . . . . .	31
<b>2 Kernel Smoothing: Regression Estimation</b>	<b>32</b>
2.1 Introduction . . . . .	32
2.2 Kernel Estimators: Local Smoothing . . . . .	33
2.2.1 Nadaraya-Watson Estimator . . . . .	33
2.2.2 k-Nearest Neighbor Estimators . . . . .	34
2.2.3 Local Polynomial Estimators . . . . .	35
2.2.4 Robust Smoothing . . . . .	36
2.3 Asymptotic Properties: Local Constant Estimator . . . . .	36
2.3.1 Consistency . . . . .	36
2.3.2 Asymptotic Normality . . . . .	39
2.4 Bandwidth Selection . . . . .	43

2.4.1	MSE-optimal Bandwidth . . . . .	43
2.4.2	Plug-in Implementation . . . . .	44
2.4.3	Cross Validation . . . . .	44
2.4.4	Computation and GCV . . . . .	51
2.5	Model Selection Perspective . . . . .	52
2.5.1	Mallows' $C_p$ Criterion . . . . .	52
2.5.2	A Theoretical Development of AIC (Digression) . . . . .	55
2.5.3	A Shrinkage Interpretation . . . . .	58
2.6	Uniform Consistency . . . . .	59
2.7	Uniform Confidence Intervals (Optional) . . . . .	60
2.8	Asymptotic Properties of Local Linear Estimator . . . . .	62
2.8.1	Asymptotic Variance of Local Linear Estimators . . . . .	62
2.8.2	Asymptotic Bias of Local Linear Estimators . . . . .	64
2.9	References . . . . .	66
<b>3</b>	<b>Sieve Estimation</b> . . . . .	<b>64</b>
3.1	Examples and Motivations . . . . .	64
3.2	Sieve Spaces . . . . .	66
3.2.1	Hölder Class and Finite Dimensional Linear Sieves . . . . .	66
3.2.2	$L_p$ and Finite Dimensional Linear Sieves . . . . .	67
3.2.3	Other Smoothness Classes and Finite Dimensional Nonlinear Sieves. . . . .	68
3.2.4	Infinite Dimensional Sieves . . . . .	68
3.3	Conditional Moment Estimation . . . . .	69
3.3.1	Smoothing Spline: Infinite Dimensional Sieves . . . . .	69
3.3.2	Series Estimators: Finite Dimensional Sieves . . . . .	77
3.4	Density Estimation . . . . .	81
3.4.1	Penalized Likelihood Estimation: Infinite Dimensional Sieves . . . . .	81
3.4.2	Series Estimation: Finite Dimensional Sieves . . . . .	81
3.5	The General Sieve Extremum Estimation . . . . .	82
3.5.1	Consistency of Sieve Extremum Estimators . . . . .	83
3.5.2	Convergence Rates of Sieve M-estimators . . . . .	84
3.6	Bibliographical Remarks . . . . .	89
3.7	References . . . . .	90
<b>4</b>	<b>Examples of Semiparametric Models</b> . . . . .	<b>85</b>
4.1	Introduction . . . . .	85
4.2	Regression Models . . . . .	85
4.2.1	Partially Linear Model . . . . .	86
4.2.2	The Single Index Model . . . . .	87
4.3	Selectivity Models . . . . .	88

4.4	Efficient and Adaptive Estimation . . . . .	91
4.4.1	Efficient Estimation in the Presence of Heteroskedasticity . . . . .	91
4.4.2	Adaptive Estimation . . . . .	92
4.5	Bibliographical Remarks . . . . .	93
4.6	References . . . . .	93
<b>5</b>	<b>Case Study: Partially Linear Model</b>	<b>95</b>
5.1	Introduction . . . . .	95
5.2	A Semiparametric Estimator . . . . .	96
5.3	Asymptotics via Stochastic Equicontinuity . . . . .	98
5.4	References . . . . .	104
<b>6</b>	<b>Semiparametric Methods: A Unified Framework</b>	<b>106</b>
6.1	The Framework . . . . .	106
6.2	Limit Theorem . . . . .	107
6.2.1	Preliminaries . . . . .	107
6.2.2	Smoothness Assumptions . . . . .	109
6.2.3	Stochastic Approximations . . . . .	110
6.2.4	Consistency . . . . .	115
6.2.5	Root-n Consistency . . . . .	116
6.2.6	Asymptotic Normality . . . . .	117
6.3	Some Details . . . . .	120
6.3.1	Orthogonality Conditions . . . . .	120
6.3.2	Asymptotic Normality of Objective Function . . . . .	122
6.4	Case Studies . . . . .	125
6.4.1	Partial Linear Model . . . . .	125
6.4.2	Single Index Model . . . . .	126
6.4.3	Sample Selection Model . . . . .	128
6.4.4	Nonparametric WLS . . . . .	130
6.5	Non-differentiable Objective Function . . . . .	130
6.5.1	Consistency . . . . .	130
6.5.2	Asymptotic Normality . . . . .	132
6.6	References . . . . .	136
<b>7</b>	<b>Empirical Applications of Semiparametric Models</b>	<b>137</b>
7.1	Structural Estimation of Production Functions . . . . .	137
7.1.1	Model . . . . .	137
7.1.2	Estimation . . . . .	137
7.2	More Examples . . . . .	138
7.3	References . . . . .	138

<b>8 Semiparametric Efficiency Bounds</b>	<b>139</b>
8.1 Hilbert Space for Random Vectors . . . . .	139
8.2 RAL Estimators in Parametric Models . . . . .	141
8.2.1 Cramer-Rao Bound . . . . .	141
8.2.2 Hodges Estimator . . . . .	142
8.2.3 RAL Estimators . . . . .	143
8.3 Properties of IFs for RAL Estimators . . . . .	146
8.3.1 Characterization of IFs . . . . .	146
8.3.2 Efficient Influence Function . . . . .	149
8.3.3 Summary of the Results . . . . .	151
8.4 Semiparametric Efficiency Bounds . . . . .	152
8.4.1 Definition . . . . .	152
8.4.2 Characterization of Semiparametric Efficiency Bounds . . . . .	154
8.4.3 Gradient Function . . . . .	157
8.4.4 Tangent Space for Nonparametric Models . . . . .	160
8.4.5 Example: estimating the mean functional . . . . .	160
8.4.6 Example: estimating the mean functional using the gradient function .	162
8.5 Semiparametric Conditional Moment Model . . . . .	163
8.5.1 The model . . . . .	163
8.5.2 Nuisance Parameter Tangent Space . . . . .	163
8.5.3 Computing the Projection . . . . .	166
8.5.4 Optimal Instruments . . . . .	167
8.6 Semiparametric Partially Linear Model . . . . .	168
8.7 References . . . . .	170