

Practical Smoothing

The Joys of P-splines

Paul H. C. Eilers
Erasmus University Medical Center

Brian D. Marx
Louisiana State University



CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>Preface</i>	<i>page xi</i>
1 Introduction	1
2 Bases, Penalties, and Likelihoods	6
2.1 Linear and Polynomial Regression	6
2.2 B-splines	9
2.3 Penalized Least Squares	14
2.4 Interpolation and Extrapolation	18
2.5 Derivatives	20
2.6 The Effective Dimension	21
2.7 Standard Errors	23
2.8 Heavy Smoothing and Polynomial Limits	24
2.9 P-splines as a Parametric Model	24
2.10 Whittaker: P-splines without B-splines	26
2.11 Equivalent Kernels	26
2.12 Smoothing of a Non-normal Response	28
2.12.1 Poisson Smoothing	28
2.12.2 Binomial Smoothing	31
2.12.3 GLM Effective Dimension and Standard Errors	32
2.13 Notes and Details	34
3 Optimal Smoothing in Action	36
3.1 Cross-Validation	37
3.2 Akaike's Information Criterion	38
3.3 Density Estimation	40
3.4 Mixed Models	41
3.5 Bayesian P-splines	45
3.6 Dangers of Automatic Smoothing	51

3.7	L- and V-curves	54
3.8	Transformation of the Independent Variable	56
3.9	Notes and Details	58
4	Multidimensional Smoothing	59
4.1	Generalized Additive Models	60
4.2	Varying Coefficient Models	63
4.3	Tensor Product Models	67
4.4	Tensor Product Bases	69
4.5	Two-Dimensional Penalties	71
4.6	Interpolation and Extrapolation	73
4.7	Smoothing on Large Grids	74
4.8	Generalized Two-Dimensional Smoothing	75
4.9	Optimal Two-Dimensional Smoothing	77
4.10	Issues with Isotropic Smoothing	79
4.11	Higher Dimensions	79
4.12	Nested Bases and PS-ANOVA	79
4.13	Notes and Details	83
5	Smoothing of Scale and Shape	86
5.1	Quantile Smoothing	85
5.2	Expectile Smoothing	91
5.3	Models for Shape and Scale Parameters	97
5.4	Baseline Estimation	101
5.5	Notes and Details	102
6	Complex Counts and Composite Links	103
6.1	Histograms with Wide Bins	104
6.2	Histograms and Scale Transformation	107
6.3	Individual Censoring	109
6.4	Latent Mixtures	110
6.5	Notes and Details	112
7	Signal Regression	114
7.1	A Chemical Calibration Problem	115
7.2	Extensions to the Generalized Linear Model	120
7.3	Multidimensional Signal Regression	122
7.4	Further Extensions	126
7.5	Notes and Details	129
8	Special Subjects	131
8.1	The Proper B-spline Basis	132
8.2	Harmonic Smoothing	132

8.3	Circular Smoothing	135
8.4	Signal Separation with Penalties	138
8.5	Double Penalties	141
8.6	Piecewise Constant Smoothing	143
8.7	Spline Constraints	146
8.8	Variable and Adaptive Penalties	152
8.9	Survival Analysis and Mortality Modeling	155
8.10	Notes and Details	158
Appendix A P-splines for the Impatient		159
Appendix B P-splines and Competitors		161
Appendix C Computational Details		168
Appendix D Array Algorithms		174
Appendix E Mixed Model Equations		176
Appendix F Standard Errors in Detail		182
Appendix G The Website		184
<i>References</i>		188
<i>Index</i>		196