

# Contents

Preface	vii
<b>1 An introduction using classical examples</b>	<b>1</b>
1.1 Numerical differentiation. First look at the problem of regularization.	
The balancing principle	1
1.1.1 Finite-difference formulae	1
1.1.2 Finite-difference formulae for nonexact data. A priori choice of the stepsize	3
1.1.3 A posteriori choice of the stepsize	6
1.1.4 Numerical illustration	9
1.1.5 The balancing principle in a general framework	10
1.2 Stable summation of orthogonal series with noisy coefficients.	
Deterministic and stochastic noise models. Description of smoothness properties	12
1.2.1 Summation methods	13
1.2.2 Deterministic noise model	14
1.2.3 Stochastic noise model	15
1.2.4 Smoothness associated with a basis	18
1.2.5 Approximation and stability properties of $\lambda$ -methods	19
1.2.6 Error bounds	21
1.3 The elliptic Cauchy problem and regularization by discretization	25
1.3.1 Natural linearization of the elliptic Cauchy problem	27
1.3.2 Regularization by discretization	36
1.3.3 Application in detecting corrosion	39
<b>2 Basics of single parameter regularization schemes</b>	<b>47</b>
2.1 Simple example for motivation	47
2.2 Essentially ill-posed linear operator equations. Least-squares solution. General view on regularization	49
2.3 Smoothness in the context of the problem. Benchmark accuracy levels for deterministic and stochastic data noise models	65
2.3.1 The best possible accuracy for the deterministic noise model	68

2.3.2	The best possible accuracy for the Gaussian white noise model	73
2.4	Optimal order and the saturation of regularization methods in Hilbert spaces	80
2.5	Changing the penalty term for variance reduction. Regularization in Hilbert scales	90
2.6	Estimation of linear functionals from indirect noisy observations	101
2.7	Regularization by finite-dimensional approximation	113
2.8	Model selection based on indirect observation in Gaussian white noise	124
2.8.1	Linear models given by least-squares methods	127
2.8.2	Operator monotone functions	131
2.8.3	The problem of model selection (continuation)	137
2.9	A warning example: an operator equation formulation is not always adequate (numerical differentiation revisited)	141
2.9.1	Numerical differentiation in variable Hilbert scales associated with designs	143
2.9.2	Error bounds in $L_2$	147
2.9.3	Adaptation to the unknown bound of the approximation error	150
2.9.4	Numerical differentiation in the space of continuous functions	151
2.9.5	Relation to the Savitzky–Golay method. Numerical examples	155
<b>3</b>	<b>Multiparameter regularization</b>	<b>163</b>
3.1	When do we really need multiparameter regularization?	163
3.2	Multiparameter discrepancy principle	165
3.2.1	Model function based on the multiparameter discrepancy principle	168
3.2.2	A use of the model function to approximate one set of parameters satisfying the discrepancy principle	170
3.2.3	Properties of the model function approximation	172
3.2.4	Discrepancy curve and the convergence analysis	173
3.2.5	Heuristic algorithm for the model function approximation of the multiparameter discrepancy principle	174
3.2.6	Generalization in the case of more than two regularization parameters	175
3.3	Numerical realization and testing	177
3.3.1	Numerical examples and comparison	177

3.3.2	Two-parameter discrepancy curve . . . . .	182
3.3.3	A numerical check of Proposition 3.1 and use of a discrepancy curve . . . . .	184
3.3.4	Experiments with three-parameter regularization . . . . .	187
3.4	Two-parameter regularization with one negative parameter for problems with noisy operators and right-hand side . . . . .	189
3.4.1	Computational aspects for regularized total least squares . . . . .	191
3.4.2	Computational aspects for dual regularized total least squares . . . . .	192
3.4.3	Error bounds in the case $B = I$ . . . . .	193
3.4.4	Error bounds for $B \neq I$ . . . . .	195
3.4.5	Numerical illustrations. Model function approximation in dual regularized total least squares . . . . .	197
<b>4</b>	<b>Regularization algorithms in learning theory</b>	<b>203</b>
4.1	Supervised learning problem as an operator equation in a reproducing kernel Hilbert space (RKHS) . . . . .	203
4.1.1	Reproducing kernel Hilbert spaces and related operators . . . . .	205
4.1.2	A priori assumption on the problem: general source conditions . . . . .	207
4.2	Kernel independent learning rates . . . . .	209
4.2.1	Regularization for binary classification: risk bounds and Bayes consistency . . . . .	217
4.3	Adaptive kernel methods using the balancing principle . . . . .	218
4.3.1	Adaptive learning when the error measure is known . . . . .	220
4.3.2	Adaptive learning when the error measure is unknown . . . . .	223
4.3.3	Proofs of Propositions 4.6 and 4.7 . . . . .	225
4.3.4	Numerical experiments. Quasibalancing principle . . . . .	231
4.4	Kernel adaptive regularization with application to blood glucose reading . . . . .	235
4.4.1	Reading the blood glucose level from subcutaneous electric current measurements . . . . .	242
4.5	Multiparameter regularization in learning theory . . . . .	249
<b>5</b>	<b>Meta-learning approach to regularization – case study: blood glucose prediction</b>	<b>255</b>
5.1	A brief introduction to meta-learning and blood glucose prediction . . . . .	255
5.2	A traditional learning theory approach: issues and concerns . . . . .	259
5.3	Meta-learning approach to choosing a kernel and a regularization parameter . . . . .	261
5.3.1	Optimization operation . . . . .	263

5.3.2	Heuristic operation .....	267
5.3.3	Learning at metalevel .....	267
5.4	Case-study: blood glucose prediction .....	269
Bibliography		277
Index		289