

Index

- , 7
 $\|\cdot\|_{\mathcal{C}^s(X)}$, 18
 $\|\cdot\|_{\text{Lip}(s)}$, 73
 $\|\cdot\|_{\text{Lip}^*(s, \mathcal{C}(X))}$, 74
 $\|\cdot\|_{\text{Lip}^*(s, \mathcal{L}^p(X))}$, 74
 $\|\cdot\|_s$, 21
 $\|\cdot\|_{H^s(\mathbb{R}^n)}$, 75
 $\mathcal{A}(f_\rho, R)$, 54
 $\mathcal{A}(\mathcal{H})$, 12
 B_R , 76
 \mathbf{C}_K , 22
 $\mathcal{C}(X)$, 9
 $\mathcal{C}^s(X)$, 18
 $\mathcal{C}^\infty(X)$, 18
 $\mathcal{D}(\gamma)$, 136
 $\mathcal{D}(\gamma, \phi)$, 196
 $\text{Diam}(X)$, 72
 $\mathcal{D}_{\mu\rho}$, 110
 Δ_i^r , 74
 \mathcal{E} , 6
 \mathcal{E}_γ , 134
 $\mathcal{E}_{\mathcal{H}}$, 11
 \mathcal{E}^ϕ , 162
 \mathcal{E}_γ^ϕ , 162
 $\mathcal{E}_{z,\gamma}^\phi$, 162
 \mathcal{E}_z^ϕ , 162
 \mathcal{E}_z , 8
 $\mathcal{E}_{z,\gamma}$, 134
 $\epsilon_K(\mathbf{x})$, 112
 η_x , 174
 f_c , 159
 f_γ , 134
 $f_{\mathcal{H}}$, 9
 f_ρ , 3, 6
 f_Y , 8
 f_z , 9
 f_z^0 , 161
 $f_{z,\gamma}$, 134
 $f_{z,\gamma}^\phi$, 164
 \mathcal{H}_K , 23
 $\mathcal{H}_{K,z}$, 34, 163
 $H^s(\mathbb{R}^n)$, 75
 $H^s(X)$, 21
 K_ρ , 159
 κ_ρ , 159
 K_X , 22
 $K[\mathbf{x}]$, 22
 $\ell^2(X_N)$, 90
 $\ell^\infty(X_N)$, 90
 $\text{Lip}(s)$, 73
 $\text{Lip}^*(s, \mathcal{C}(X))$, 74
 $\text{Lip}^*(s, \mathcal{L}^p(X))$, 74
 L_K , 56
 $\mathcal{L}_v^p(X)$, 18
 $\mathcal{L}_v^\infty(X)$, 19
 L_z , 10
 $\mathcal{M}(S, \eta)$, 101
 $\mathcal{N}(S, \eta)$, 37
 $\mathcal{O}(n)$, 30
 ϕ_0 , 162
 $\phi_{\mathbf{h}}$, 165
 $\phi_{\mathbf{s}}$, 162
 $\Pi_s(\mathbb{R})$, 84
 $\mathcal{R}(f)$, 5
 ρ , 5
 ρ_X , 6
 $\rho(y|x)$, 6
 sgn , 159
 σ_ρ^2 , 6
 Υ_k , 87

- X , 5
 $X_{r,t}$, 74
 Y , 5
 Z , 5
 Z^m , 8
- Bayes rule, 160
 Bennett's inequality, 38
 Bernstein's inequality, 40, 42
 best fit, 2
 bias–variance problem, 13, 127
 bounded linear map, 21
 box spline, 28
- Chebyshev's inequality, 38
 classification algorithms, 160
 classifier
 - binary, 157
 - regularized, 164
 compact linear map, 21
 confidence, 42
 constraints, 33
 convex function, 33
 convex programming, 33
 convex set, 33
 convolution, 19
 covering number, 37
- defect, 10
 distortion, 110
 divided difference, 74
 domination of measures, 110
- efficient algorithm, 33
 ε -net, 78
 ERM, 50
 error
 - approximation, 12
 - approximation (associated with ψ), 70
 - empirical, 8
 - empirical (associated with ϕ), 162
 - empirical (associated with ψ), 51
 - excess generalization, 12
 - excess misclassification, 188
 - generalization, 5
 - generalization (associated with ϕ), 162
 - generalization (associated with ψ), 51
 - in \mathcal{H} , 11
 - local, 6, 161
 - misclassification, 157
 - regularized, 134
 - regularized (associated with ϕ), 162
 - regularized empirical, 134
 - regularized empirical (associated with ϕ), 162
 expected value, 5
- feasible points, 33
 feasible set, 33
 feature map, 70
 Fourier coefficients, 55
 Fourier transform, 19
 - nonnegative, 26
 - positive, 26
 full measure, 10
 function
 - completely monotonic, 21
 - even, 26
 - measurable, 19
 generalized Bennett's inequality, 40, 42
 Gramian, 22
- Hoeffding's inequality, 40, 42
 homogeneous polynomials, 17, 29
 hypothesis space, 9
 - convex, 46
 interpolation space, 63
- \mathbb{K} -functional, 63
 kernel, 56
 - box spline, 28
 - dot product, 24
 - Mercer, 22
 - spline, 27
 - translation invariant, 26
 - universal, 212
 Lagrange interpolation polynomials, 84
 Lagrange multiplier, 151
 least squares, 1, 2
 left derivative, 34
 linear programming, 33
 localizing function, 190
 loss
 - classifying function, 187
 - ε -insensitive, 50
 - function, 161
 - function, regression, 50
 - hinge, 165

Cambridge University Press

978-0-521-86559-3 - Learning Theory: An Approximation Theory Viewpoint

Felipe Cucker and Ding-Xuan Zhou

Index

[More information](#)

224

Index

- least squares, 50, 162
- misclassification, 162
- q -norm, 187
- margin, 167
 - hard, 171
 - maximal, 168
 - of the sample, 168
 - soft, 171
- Markov's inequality, 38
- measure
 - finite, 19
 - marginal, 6
 - nondegenerate, 19
 - strictly separable by \mathcal{H}_K , 173
 - support of, 19
 - weakly separable by \mathcal{H}_K , 158, 182
- Mercer kernel, *see* kernel
- metric entropy, 52
- model selection, 127
- multinomial coefficients, 24
- net, 78
- nodal functions, 102
- objective function, 33
- offset, 172
- operator
 - positive, 55
 - self-adjoint, 55
 - strictly positive, 55
- optimal hyperplane, 168
- orthogonal group, 30
- orthogonal invariance, 30
- orthonormal basis, 55
- orthonormal complete system, 55
- orthonormal system, 55
- packing number, 101
- power function, 125
- problems
 - classification, 15
 - regression, 15
- programming
 - convex, 34
 - convex quadratic, 34
 - general nonlinear, 33
 - second-order cone, 34
- projection operator, 195
- radial basis functions, 28
- regression function, 3, 6
- regularization
 - parameter, 135, 164
 - scheme, 135, 164
- regularizing function, 195
- reproducing kernel Hilbert space, 24
- reproducing property, 24
- right derivative, 34
- RKHS, *see* reproducing kernel Hilbert space
- sample, 8
 - separable, 166
 - separable by a hyperplane, 166
- sample error, regularized, 136
- separating hyperplane, 167
- separation
 - exponent, 182
 - triple, 182
- Sobolev embedding theorem, 21
- Sobolev space, 21
 - fractional, 35
- spherical coordinates, 98
- support vector machine, 165
- support vectors, 169
- SVM, *see* support vector machine
- target function, 9, 51, 134
 - empirical, 9, 51, 134
- uniform Glivenko–Cantelli, 52
- variance, 5
- variancing power, 198
- Veronese embedding, 31
- Veronese variety, 31, 36
- weak compactness, 20
- weak convergence, 20
- Weyl inner product, 30
- Zygmund class, 75