

Index

A

Accumulated proportion, 240, 241
Adjusted coefficient of determination, 34, 37, 96, 98
Akaike's information criterion (AIC), 95, 96, 98, 100, 107
AUC, 67
Average linkage, 235

B

Back-fitting, 154
Bagging, 185
Bayesian information criterion (BIC), 95, 96, 98
Boosting, 189
Bootstrap, 85, 86
Bootstrapping, 89
Branch node, 171

C

Centroid linkage, 237
Coefficient of determination, 34, 36, 37, 96
Complete linkage, 234
Confidence interval, 39, 85, 115, 127
Convex, 118
Convex function, 118, 119
Correlation coefficient, 36
Cramér-Rao inequality, 100, 102
Cross-validation, 77, 78, 82, 126, 173, 191

D

Decision tree, 171, 181
Dendrogram, 236, 238, 247

Dual problem, 203, 205, 209

E

Effective degrees of freedom, 147
Efficient estimator, 100, 101
Entropy, 181
Epanechnikov kernel, 150
Error probability, 181
ESS, 35

F

False negative, 67
False positive, 67
False positive rate, 67
Fisher information matrix, 100–102
Fitness, 95, 98

G

Gini index, 181
Gradient boosting, 192

H

Hat matrix, 26
Hierarchical clustering, 233, 236, 237, 247
Hypothesis testing, 30

I

Information criterion, 95
Interior node, 171
Inversion, 237

K

K-nearest neighbor, 65
 Karush–Kuhn–Tucker (KKT), 204
 Kernel, 148, 149
 K-means clustering, 230
 Kullback-Leibler, 103, 105

L

L1 norm, 121
 L1 regularization, 125
 L2 norm, 121
 L2 regularization, 125
 Lasso, 124, 126
 Least squares, 19, 22, 25, 41
 Linear discrimination, 62
 Linear regression, 19, 77, 78, 80, 81, 85, 115
 Local linear regression, 148, 151
 Logistic regression, 54, 60
 LOOCV, 81

M

Margin, 200, 201
 Maximum likelihood, 55, 59, 60
 Maximum posterior probability, 64
 Multiple regression, 23

N

Nadaraya-Watson estimator, 149, 150
 Natural spline, 140, 142, 144, 153, 158
 Newton–Raphson method, 56
 Neyman–Pearson criterion, 67
 Nonnegative definite, 15
 Null hypothesis, 30, 35

O

Overfitting, 79, 80

P

Polynomial kernel, 210
 Positive definite, 15
 Posterior probability, 61
 Power, 67
 Prediction interval, 40
 Primary problem, 203, 204, 209
 Principal component analysis (PCA), 239
 Principle component regression, 246
 Prior probability, 61, 62
 Proportion, 240, 241

Q

Quadprog, 207
 Quadratic discrimination, 62

R

Radical kernel, 213
 Random forest, 187
 Receiver operating characteristic (ROC), 68
 Ridge, 115, 121, 124
 ROC curve, 67
 Root, 171
 RSS, 26–28, 32, 96, 125

S

Schwarz's inequality, 105
 Separable, 200
 Sherman-Morrison-Woodbury, 82
 Significance level, 30
 Simplicity, 95, 98
 Single linkage, 237
 Single regression, 22, 36
 Smoothing spline, 144, 145, 154
 Sparse, 115
 Spline curve, 141
 Spline function, 137, 138
 Spline regression, 138, 140
 Subderivative, 118, 119
 Supervised learning, 227
 Support vector, 200, 201
 Support vector machine, 199, 202, 210, 212

T

Terminal node, 171
 True negative, 67
 True positive, 67
 TSS, 36
 Type I Error, 66
 Type II Error, 66

U

Unbiased, 26
 Unbiased estimate, 102, 105
 Unbiased estimator, 100
 Unsupervised learning, 227

V

Vandermonde's determinant, 5
 Vandermonde's matrix, 135
 Variable selection, 106
 VIF, 38