

## Index

### a

- abduction 301
- accreditation 351
- accuracy 97, 98, 261, 346–348
  - of analysis 98
- adaptive neuro-fuzzy inference system (ANFIS) 330, 331
- adaptive resonance theory (ART net) 258, 265, 330, 331
- adjacency matrix 280, 281, 294
- aggregation, performance characteristics 100, 102
- alignment methods
  - DTW 82
  - left panel, simulated chromatogram 83
- allele 334
- $\alpha$  error 40, 41
- alternating conditional expectations (ACE) 258, 262–264, 266, 267
- alternating least squares (ALS) 169, 172
- alternative 300
- alternative hypothesis *see* hypothesis
- American food and drug administration (FDA) 351
- analog-to-digital converter (ADC) 3, 5, 7
- analysis of variance (ANOVA)
  - definition 44
  - MANOVA (*see* multidimensional analysis of variance (MANOVA))
  - multi dimensional (MANOVA) 49–50
  - multi-way, 46, 49 (*see also* multi-way analysis of variance)
  - one-way, 44–46, 48 (*see* one-way analysis of variance)
  - regression analysis 218
  - sums of squares (SS) 216, 217
  - two-way, 46–49 (*see* two-way analysis of variance)
- analytical databases
  - coding of chemical structures (*see* chemical structures coding)
  - coding of spectra 276, 277
  - examples 273, 274
  - exchange files 275, 276
  - information types 274
  - laboratory information and management systems (LIMS) 285
  - source and library files 275
  - structure 275
  - working and index file 284
- analytical range 97
- analyzers 9, 298
  - discrete 9
- Anthranilic acid 266, 267
- apodization function 74
- area determination 55, 66, 68
- artificial intelligence (AI) 2, 10–12, 11, 292, 297–306
  - definition 297
  - knowledge acquisition and processing 298
  - methods 298
  - modules 297, 298
- artificial neural networks 306, 320, 321 *see also* neural networks
- artificial neuron 306, 307
- assembler language 7
- associative learning law 312
- atomic absorption spectrometry 35, 96
- atomic emission spectroscopy 83, 299
  - inductively coupled plasma (ICP-AES) 83
- autocorrelation 55, 83, 84, 86, 87, 90
  - correlated data 87
  - correlogram 86
  - drift and periodicities 88
  - empirical 87
  - function 86–89
  - point-wise 85

- autocorrelation (*contd.*)
  - stationary process 85
  - time lags 84, 86
  - uncorrelated data 87
- autocovariance 83–85
- automation 1, 3, 8–10
- autoregression 89–90
- autoscaling 138
- axiom 300
- axon 306, 307
  
- b**
- background correction 55, 68, 75, 88
- backpropagation 312, 316, 317, 320, 321, 338
  - networks 316–318
- back-tracking 302, 303
- bagging 204–206
- basic 7, 15–50, 104, 120, 195, 267, 302, 351
- Bayesian classification 191
- Bayesian regulation backpropagation algorithm 320
- Bayes's theorem 191
- Beer's law 240, 243, 258
- $\beta$  error 40, 41
- best linear unbiased estimator (BLUE) 232
- bias 16, 44, 98, 103, 254, 271, 311, 317, 320 *see also*
  - error; systematic
- bidirectional associative memory (BAM) 308–310, 308–312, 344
- binary number system 3–5
- biological fluids 83
- biplot 148, 149
- bit 1, 4–7, 286, 289, 334–338, 340, 342
- blank 96, 97, 330, 346, 348
  - samples 347
- blind source separation problem 165
- blocking 104–105, 108, 110, 124
- blood 9, 16, 83, 174, 175, 184, 203, 351
- bond electron (BE) matrix 281
- bootstrapping 204, 262, 320
- boosting 204–206
- box-and-whisker plots 24, 25, 52
- Box–Behnken design 117
- bulk storage devices 7
- bus 6
- byte 6
  
- c**
- C 7
- Calcium 174, 175
- calibration 213
  - curved 96
  - direct 241–243
  - function 96–97, 214, 346
  - indirect 243–247
  - inverse 245–247, 251
  - K-matrix approach 243–245
  - P-matrix approach 245
- calibration function 96
- canonization 284
- cardinality 328, 329, 344
- centering 137, 138, 210, 236
- central composite design 115, 117
- central limit theorem 21–22
- central processing unit (CPU) 6
- certified reference materials 350
- ceteris paribus principle 102–103
- “change and hope” optimization procedure 125
- chaos 12
- chemical databases 273 *see also* analytical databases
- chemical structure 11, 273–281, 283, 286, 287, 290–292, 295
  - matrix representation 149, 280, 281
- chemical structures coding
  - canonization 284
  - connection table 283, 284
  - fragmentation codes 278–280, 279
  - Markush structures 284
  - matrix representation 280, 281, 283
- chemometrics
  - computer-based methods (*see* computer-based laboratory)
  - data evaluation and interpretation methods 10
  - definition 1
- chemometrics in Analytical Chemistry 1
- Chernoff faces 182–184, 183, 184
- chromatographic response function (CRF) 101
- chromatography 9, 100, 101, 118, 135, 161, 164, 168, 256, 277, 299, 332
- chromosome 334–336, 338
- $\chi^2$  goodness-of-fit test 37, 38
- city-block distance 172
- class 11, 16, 17, 23, 38, 135–137, 184, 186–199, 201, 204, 208, 284, 288, 311, 313, 314, 319, 331, 332
- classification
  - by discriminant analysis 186–193, 198, 199, 267
  - by k-nearest neighbors (kNN) 178, 193
  - by soft independent modeling of class analogies (SIMCA) 184, 195–196
- classification and regression trees (CART) 201, 202, 209, 267
- class width 16, 17
- cluster analysis 11, 135, 136, 140, 147, 148, 172–175, 174, 175, 177–182, 178–180, 184, 201, 287, 288, 311, 319, 332
  - distance measure 172
  - hierarchical 174
  - nonhierarchical 179
  - similarity measure 173

- clustering
    - crisp 179, 180, 290, 322, 323
    - fuzzy 179, 180, 332, 333
    - hierarchical 174, 180
    - non-hierarchical 179–182, 201
  - c-means algorithm 179, 181, 332
  - c-means clustering 181, 182
  - coding
    - chemical structures 278–295
    - spectra 273
  - coefficient
    - of determination 220, 221, 252
    - of variation 23
  - Coiflet 78, 79
  - collinearity 243
  - common factors 150–152, 158
  - communality 151–153, 156, 160
  - competitive learning law(s) 319
    - steps 312
    - supervised 313
    - unsupervised 313
  - compiler 7
  - complement 328
  - compositional rule of inference 329
  - computer application in analytics (COMPANA) 1
  - computer-based analytical chemistry (COBAC) 1
  - computer-based information systems 11–12
  - computer-based laboratory 2–10
    - analog and digital data 3
    - analog-to-digital converters (ADCs) 5
    - binary vs. decimal number system 3, 5
    - central processing units and buses 6
    - computer terminology 6
    - digital-to-analog converters (DACs) 5
    - input/output-systems 7
    - memory 6
    - networking 7, 8
    - programming 7
    - robotics and automation 8, 10
  - computer program 3, 6, 7, 304
  - conditioning 9, 232
  - condition number 210, 232, 233
  - confidence
    - band 224, 225, 262
    - interval 23, 24, 52, 223–231, 234, 271, 320, 347
  - confidence interval 23, 24
  - confounding 105, 114
    - parameter estimations 105
  - confusion matrix 188
  - conjunction 300
  - connection table 283, 284, 290, 291, 295
    - bond atoms and bonds 283
    - connection matrix of atoms 283
    - nonredundant 283
  - continuous/discrete continuous quantities 96
  - control
    - chart 348, 349
    - sample 347
  - convolution 55, 68, 73
  - Cook's distance 250, 253, 254
  - correlation
    - coefficient 82, 84, 85, 140, 153, 220, 221, 288
    - matrix 139, 140, 145, 149, 151, 153, 155, 156, 158, 159, 175, 309, 310
    - – reduced 151–153, 156, 157
    - measure 16, 84, 148
  - correlation matrix 140
  - corticoids 288
  - cortisone 287, 288
  - counting rates 20
  - covariance matrix 62, 63, 83, 139, 140, 144, 151, 188, 192, 193, 195, 223, 224, 262, 266
  - cross correlation 55, 83, 89
  - crossover 334, 335, 337
  - cross-validation 78, 144, 168, 194, 196, 204–206, 209, 236, 238, 247, 253, 255, 268
  - cross-validation principal components 145
- d**
- data
    - analog 3
    - correlated 16, 55, 87, 88
    - digital 3, 5
    - interpretation 10–11
    - missing 137
    - preprocessing 135, 210
    - random 10, 16, 55, 193
    - set characterization 22, 137, 141, 145, 146, 168, 172, 174, 189, 204, 207, 249, 263, 265, 278, 286, 288, 290, 320, 331, 332
    - three-way 168, 169, 172, 256, 257
    - transfer
      - – parallel 6
      - – serial 6
    - two-way 47, 49, 167, 256
  - database
    - analytical 273–297
    - model 285
    - structure 274–275
  - data preprocessing
    - correlation matrix 140
    - mean centering 137
    - missing data 137
    - scaling methods 138
    - variance–covariance matrix 139
  - data reduction 71, 75, 278, 285

- Daubechies 78, 79, 81
- decimal number system 3–5
- decision boundary 184–186, 198, 199, 321
- deconvolution 55, 68, 73, 74
- deduction 298, 300, 301
- degrees of freedom 23, 31, 32, 35, 37, 46, 49, 107, 217, 220, 221, 249, 252, 349, 358–360
- delta rule 312, 316
- dendrite 306, 307
- dendrogram 177, 179
- density function *see* frequency function
- derivation algorithm 10
- descriptive statistics
  - box-and-whisker plots 24, 25
  - confidence interval 23, 24
  - dispersion measure 23
  - error propagation 25, 27
  - geometric mean 22
  - harmonic mean 22
  - median 22
  - quartiles 22, 23
  - random numbers (*see* random number distribution)
  - uncertainty and error 27
- detection limit 97, 330, 346
- diagnostic statistics
  - hat matrix 248
  - influential observations 250, 256
  - inverse calibration model 247
  - outliers 249
  - residuals and prediction error 248
  - SEC 247
- digit
  - least significant 4
  - most significant 4, 75
- digital smoothing and filtering
  - advantages 56
  - filter selection process 58, 62
  - Kalman filter 62, 63
  - moving-average filter 56, 57
  - Savitzky–Golay filter (*see* Savitzky–Golay filter)
  - Simpson rules 67, 68
  - trapezoidal rule 67
- digital-to-analog converter (DAC) 5, 7
- discrete membership functions 326, 327
- discrete quantities 96
- discriminant analysis
  - linear (LDA) 186–192, 207–209
  - quadratic (QDA) 192, 207–209
  - regularized (RDA) 192
- dispersion 15, 18, 23, 98, 141, 151, 347, 348
  - matrix 141, 151 (*see also* variance)
  - measure 23
- dissimilarity measure 288, 289, 295
- distance
  - matrix 174–176, 178
  - measure 172–175, 173, 178, 287, 288, 312, 318
- distribution
  - binomial 21
  - chi-square 360
  - F- 35, 36, 359
  - function 18, 29, 38, 39, 323
  - Gaussian (*see* normal distribution)
  - kurtosis 19, 22, 24, 160
  - moment 19
  - Poisson 20
  - skewness 19
  - Student's 23, 31, 32, 347, 358
- Dixon's Q-test 42, 43
- drift 55, 63, 72, 73, 78, 84, 88, 89, 103
- dynamic range 97
- dynamic time warping (DTW) 82
- e**
- eigenvalue 50, 145, 149, 150, 153–156, 158–160, 165, 187, 188, 210, 232, 245, 247, 369
  - one criterion 144, 145
- eigenvector
  - analysis 141, 153, 369
  - projection 141
- elemental composition 50, 136, 137, 146, 160, 164, 189, 190
- ensemble methods 204
- enzyme 105, 111, 113, 121, 124, 129
  - ceruloplasmin (CP) 120, 121
- equivalence 137, 300
- error
  - first and second kind 40–41
  - of first kind 40, 41
  - integral (*see* distribution function)
  - propagation 15, 25, 27, 65, 66, 100
  - random 16, 36, 44, 47, 90, 98, 226, 227, 230
  - real 159
  - round-off 3, 25, 56, 233
  - of second kind 40, 41
  - systematic 16, 97, 98, 103, 345
- ESPRIT 305
- Euclidean distance 172, 173, 179, 189, 190, 193, 312
- European Union regulation 33
- evolutional component 334
- evolutionary strategies 11, 12
- evolving factor analysis (EVA) 164
- exchange file 275, 276
- exclusive OR 289, 290, 314, 315
- experimental design
  - blocking factors 104, 105
  - Box–Behnken 115, 117, 118, 120, 121, 367
  - central composite 115–117, 366

- ceteris paribus principle 102, 103
  - confounding 105
  - factorial experiments 105
  - fractional factorial 107–109, 115
  - full factorial 106, 107, 111, 114–116, 125
  - Latin square 52, 108, 110
  - lattice 119, 368
  - mixture 118, 119, 368
  - Plackett–Burman 108, 109
  - randomization 103
  - replication 103
  - response surface (*see* three-level)
  - screening (*see* two-level)
  - star 115, 116
  - symmetry 105
  - three-level 114, 115, 117
  - three-level designs (*see* response surface designs)
  - two-level 106, 108, 114–116, 364
  - two-level designs (*see* screening designs)
  - expert systems 4, 7, 11, 297–306, 333
    - aims 304
    - analytical chemistry 305, 306
    - definition 303
    - development tools 304
    - structure 304, 305
  - explained variance principal components 145
  - extended *t*-test and *F*-test 35
  - externally studentized residual 250
- f**
- fact 2, 56, 71, 75, 141, 145, 165, 179, 184, 187, 190, 244, 245, 255, 259, 265, 281, 300, 303
  - factor
    - coding 122–123
    - experimental 15, 133
    - loadings 153
    - rotation methods 159
    - score determination 157
    - score matrix 157
    - significance 94
  - factor analysis (FA)
    - evolving (EVA) 164, 165
    - factorial methods 150
    - parallel (PARAFAC) 168–172, 256
    - vs. singular value decomposition (SVD) 157
  - factor effects 50, 93, 105–107, 110, 111, 113, 114, 119, 120, 123–125
    - estimation 110, 114
    - vs. regression parameters 123, 124
    - relation with regression parameters 234, 248
  - factorial
    - experiments 105, 107
    - methods 78, 140, 141, 144, 168, 311, 369
    - – factor analysis 150
    - – principal component analysis (PCA) 141
    - – singular value decomposition (SVD) 149
  - fast Hadamard transformation (FHT) 71
  - feature 31, 81, 137, 138, 141, 145, 147, 148, 150–153, 156, 157, 159, 160, 163, 164, 172–175, 180, 182–184, 187, 189, 191, 198–202, 204, 256, 276, 286, 288, 289, 295, 311, 318, 334, 345
  - figure-of-merit *see* objective function
  - filter
    - high pass 72
    - Kalman 55, 62, 63
    - low pass 72
    - moving average 56–58
    - polynomial 55, 61, 64, 65, 78
    - recursive 62
    - Savitzky–Golay 57–59, 61, 63, 65
    - width 56–58, 61, 91
  - filtering algorithm 63
  - fixed-size simplex optimization method 127, 128
  - flowing system 9
  - forecasting *see* prediction
  - FORTRAN 7
  - forward directed learning algorithm 316
  - Fourier transformation (FT)
    - conventional spectroscopy 69
    - convolution and deconvolution 73, 74
    - data reduction and background correction 75
    - discrete 69
    - integration 74, 75
    - inverse 70
    - Nyquist frequency 68
    - signal filtering 72
    - time domain 68
  - fractal structures 12
  - fractional factorial designs 107, 110
  - fragmentation codes 278, 280, 290
    - HORD code 280
    - HOSE code 278, 279
    - structure generator 280
  - frame 299, 300, 345
  - frequency
    - aliased 69
    - function 17
  - frequency function 17
  - F*-tests 35
  - full-factorial design 106
  - fuzzy arithmetics 333
  - fuzzy clustering 180, 181
  - fuzzy inferences 333
  - fuzzy inference systems 329
  - fuzzy logic 333
  - fuzzy set 180, 290, 297, 322, 323, 326–330, 333
    - operation 327–329

- fuzzy theory 11, 322–333
  - applications 332, 333
  - cardinality 328, 329
  - complement 328
  - intersection 327
  - relative cardinality 328, 329
  - set theory 322
  - union 327
- F value 35, 36, 45, 46, 48, 49, 214, 220–222, 226, 252
  
- g**
- gas chromatography 135, 256, 277, 332
  - GC/MS 8
- Gaussian distribution 17, 19 *see* normal distribution
- Gaussian elimination 233
- Gauss-Jordan elimination 233
- generic structure *see* Markush structure
- genetic algorithms 12, 297, 334–342
  - applications 337, 338
  - computer heredity 334
  - crossover 335, 337
  - initial population 334
  - mutations 335, 336
  - selection/reproduction 335, 336
  - simulated annealing (SA) 338–340
  - tabu search (TS) 340–342
- geometric mean 22
- global search strategies *see* genetic algorithms
- glucose 16, 83, 88
- good laboratory practice (GLP) 351, 352
- goodness-of-fit test 37, 222, 252
- graph theory 280
- Grubb's test 43, 52
  
- h**
- Haar 78, 79
- Hadamard transformation (HT)
  - data reduction and background correction 75
  - FHT 71
  - Walsh function 70
  - wavelet transformations 70
- hair 136, 137, 145–149, 160, 179, 183–185, 189, 190
- hamming distance 289
- hard disk 7
- hardware 6, 56
- harmonic mean 22
- hat-matrix 254
- Hebb learning *see* learning, associative
- Hebb's law 312
- heredity 334
- Hessian matrix 261
- heteroscedasticity 227, 228
- hexadecimal number 4, 5
- hidden layer 200, 306, 313, 315–317, 320, 321
- hierarchical cluster analysis
  - centroid linkage 179
  - complete linkage 178
  - first reduced matrix 175
  - fourth reduced matrix 177
  - median linkage 179
  - second reduced matrix 176
  - single linkage 178
  - third reduced matrix 176
  - unweighted average linkage 179
  - Ward's method 179
  - weighted average linkage 178
- hierarchically ordered ring description (HORD) code 280 *see also* HORD code
- hierarchically ordered spherical description of environment (HOSE) code 279 *see also* HOSE code
- hierarchical search trees 286, 287
- histogram 17, 18, 227
- hit list 288
- homoscedasticity 228
- HORD code 280
- horn clause 303
- HOSE code 278, 279
- householder reduction 233
- human being 8, 298
- 3-hydroxypropanal 282
- hypotheses testing
  - comparison of mean with true value 30
  - comparison of two means 34, 35
  - comparison of variances 35, 36
  - distribution (*see* hypothetical distribution)
  - error of first and second kind 40, 41
  - mean *vs.* true value 34
  - steps 30
- hypothesis
  - alternative 30–33, 38
  - null 30–35, 38, 40, 42, 43, 45, 50, 233, 234, 249
  - testing 20, 28, 30, 33, 34, 40
- hypothetical distribution
  - $\chi^2$  goodness-of-fit test 37, 38
  - Kolmogorov–Smirnov's test 38, 40
  
- i**
- ICA *see* Independent Component Analysis (ICA)
- identification 285, 320, 333
- identity matrix 63, 154, 192, 363
- image processing 10
- implication 300, 305
- in-breadth searching 302
- Independent Component Analysis (ICA) 165, 195
- in-depth search 302
- indicator function 159

- induction 301
- inferences 10, 15, 16, 300–305, 329–331, 333
  - abduction 301
  - deduction 300, 301
  - engine 304, 305
  - induction 301
  - logical connectives 300
  - rule 301
- influential observation 227, 250, 253, 254, 272
- information and system theory 10
- infra red spectroscopy 213
  - NIR 213, 239, 240, 251, 252, 320
- initial population 334, 336
- integration algorithm 74
- interface
  - IEEE-488 6
  - RS-232 6
- interference noise 73
- interlaboratory comparison 35, 45, 48
- interpolating splines 77
- interpretable factors 150
- interquartile range 22, 23, 25
- intersection 286, 290, 327, 328, 344
- inverse sensitivity 96
- inverted lists 286, 287, 295
  - library search 286, 287
- I/O system 7
  
- j**
- jack-knifed residuals 253 *see also* studentized residual
- Jacobian matrix 261
- JCAMP/DX exchange file 275, 276
- joint committee on atomic and molecular data (JCAMP/DX) 275, 276
  
- k**
- Kalman filter 62, 63
- Karhunen–Loeve expansion 147
- kernel functions 200
- kinetic-enzymatic oxidation 113
- k-matrix approach 244, 245
- k-nearest neighbor classification 193, 195
- k-nearest neighbor (k-NN) method 193
- knowledge
  - acquisition 11, 298
  - base 302, 304, 305
  - processing 11, 297–342
  - representation 299
- knowledge processing and soft computing *see* artificial intelligence (AI); expert systems; fuzzy theory; genetic algorithms; neural networks
- known factors 105
- Kohonen networks 311, 318–320, 318–321
  
- Kolmogoroff–Smirnov's test 38, 40
- kurtosis *see* distribution
  
- l**
- laboratory-information-and-management system (LIMS) 8, 285
- lack-of-fit test 222
- Lagrangian theory 199
- latent variable 143, 210, 235, 264, 265
- Latin square *see* experimental design
- Latin square factorial design 108, 110
- learning
  - algorithm 308, 312, 313, 316, 318, 319, 338
  - associative 312, 314, 319
  - competitive 312, 313, 319, 344
  - law 312, 314, 319, 344
  - paradigm 311
  - supervised 311, 314, 317, 332, 333
  - unsupervised 135, 172, 311, 314, 318, 332, 333
- Leary criterion 101
- legal inference 300
- level attained *see* p-level
- library file 275
- library search 193, 273, 286, 287, 333
  - hierarchical search trees 286, 287
  - inverted lists 286, 287
  - sequential search 286
  - similarity measures for chemical structures 290, 291
  - similarity measures for spectra (*see* similarity measures for spectra)
- library spectrum 287–290
- limited selective method 99
- limit of determination 97, 346, 353
- linear discriminant analysis (LDA) 186, 207
- linearization method 259, 260, 262
- linear learning machine (LLM) 184–187, 198
- linear model 214
- linguistic modifiers 325, 326
- linguistic variables 325, 326, 330
- linkage
  - centroid 179
  - complete 178
  - median 179
  - single 178, 179
  - unweighted average 178
  - weighted average 175, 178–179
- liquid chromatography 118, 161
  - HPLC 103, 119, 129, 161, 257, 258, 266, 267, 276, 288, 306, 320, 338
- liquid handling 9
- LISP language 302–304

- list processing language (LISP) 7, 298, 302–306, 303
- local area networks (LANs) 8
- location parameter 22
- logical connective 5, 289, 300
  
- m**
- Machine code 7
- machine learning 11, 305, 319
- Mahalanobis distance 173, 192, 248
- Manhattan distance 172, 288
- manipulation 8, 9, 11, 55, 285
  - software 275
- MANOVA *see* analysis of variance, multidimensional
- Markush chemical structures 284
- Markush structure 284
- Marquardt algorithm 262
- mass spectrometry 80, 81, 135, 169, 256, 277, 288, 289, 299
  - electrospray-ionization (ESI-MS) 169–171
- matrix inversion 245, 369
- mean
  - arithmetic 19, 22, 24, 52
  - centering 137
  - geometric 22, 24
  - harmonic 22
- measurement 1, 3, 8–10, 15–19, 22, 24–32, 39–45, 47–49, 62, 69, 75, 76, 83, 84, 86, 89, 90, 95–98, 101, 103, 111, 120, 214–218, 222, 229, 234, 256, 273, 277, 345, 347, 349, 366, 367
  - uncertainty 15, 16
- median 22, 24, 52, 179, 180, 231, 348
- membership functions 180, 322–330, 332, 333
  - characteristic function 322
  - discrete 326, 327
  - fuzzy observations 323–325
  - linguistic variables 325, 326
  - truth values 325, 326
- memory
  - random access (RAM) 6, 286
  - read-only (ROM) 6, 7
- method of steepest descent 259, 262
- Mexican hat 78, 79, 318, 319
- Minkowski distance 174
- missing data 137
- mixture designs 118, 119
- mode 3, 22, 23, 151, 158, 170, 256
- model(ing)
  - dynamical system model 62
  - empirical 114, 119, 120, 213, 261, 267, 269
  - linear 107, 120, 214, 221, 259, 311
  - mechanistic 119, 261, 267
  - power 196, 197
  - model adequacy test 107, 215, 216, 218, 221, 222
  - moments of distribution 19
  - Monte Carlo method 261
  - Moore–Penrose-matrix *see* pseudo inverse matrix
  - Morgan's algorithm 284
  - Morlet 78, 79
  - multicomponent analysis 63, 99, 100, 213, 231, 240, 243, 246, 247, 251, 256, 264, 334, 338, 340
  - multicriteria optimization 333
  - multidimensional analysis of variance (MANOVA) 49, 50
  - multilayer perceptrons 316
  - multiple linear regression
    - diagnostic statistics (*see* diagnostic statistics)
    - direct calibration method 241, 243
    - indirect calibration methods 243, 245
    - multiway regression 256
    - OLS (*see* ordinary least squares (OLS) regression)
    - PCR 235, 246
    - PLS 235, 238, 247
    - P-matrix approach 245
    - soft modeling 246
    - spectrometric multicomponent analysis 240
  - multivariate adaptive regression splines (MARS) 258, 265, 267
  - multivariate regression spline model 265
  - multi-way analysis of variance 49, 50
  - multi-way decompositions 167
  - mutation 334, 335, 337
  
- n**
- negation 286, 300
- Nelder and Mead method 93, 126, 127
- network 3, 7–8, 12, 200, 258, 273, 278, 285, 288, 297, 298, 306–309, 311–321, 330, 334, 338
  - worldwide 8
- Network Common Data Format (Netcdf) 278
- neural networks 12, 288, 297, 298, 306–322, 330, 338
  - aggregation 307, 308
  - applications 319, 320
  - architecture 313
  - artificial neuron 306
  - associative learning law 312
  - backpropagation algorithm 316–318
  - bidirectional associative memory (BAM) 308–310
  - competitive learning law 312, 313
  - hidden layers 306
  - Kohonen networks 318–320
  - learning paradigms 311
  - models 311
  - perceptron models 314–316



- transfer functions 308
- weights of synapses 307
- neuron 306–308, 311–313, 318, 319
- Nitrate 32–34
- noise
  - interference 73
  - of signal derivatives 63–66
- nonhierarchical cluster analysis
  - c-means clustering 181, 182
  - fuzzy clustering 180, 181
- nonlinear iterative least squares algorithm (NIPALS) 143, 148, 149, 195, 196, 237
- nonlinear partial least squares (NPLS) 264
- nonlinear PLS 264
- nonlinear regression (NLR)
  - chemical equilibria 258
  - grid search method 261
  - Hessian matrix 261
  - Jacobian matrix 261
  - Marquardt algorithm 262
  - Monte Carlo method 261
  - parameters 259
  - regression diagnostics 262
  - simplex method 261
  - steepest descent 259, 260
- normal distribution 17, 19, 21–23, 28–31, 30, 37–40, 42, 191, 192, 195, 227, 356, 357, 361
- normalization 26, 58, 138, 210, 259
- nuclear magnetic resonance 69, 274, 275, 277, 278, 292, 293, 299
- null hypothesis *see* hypothesis
- N-way PLS (N-PLS) 256, 257
- Nyquist frequency 68
- o**
- object 22, 137, 138, 141, 145, 148, 172, 174–182, 189–191, 193–198, 350
- objective criterion 10, 101, 129, 130, 348, 349
  - aggregation of 175
- objective function 93–102, 259, 261, 312, 333, 335, 337, 339–341
  - accuracy of analysis 98
  - aggregation, performance characteristics 100, 102
  - calibration function 96
  - detection limit and limit of determination 97
  - polyoptimization 102
  - specificity and selectivity 98, 100
  - time, cost and risk 100
- oblimax criterion 160
- one-variable *t*-test
  - calculated value *vs.* tabulated value 31, 32
  - nonparametric tests 31
  - null hypothesis 30
- one-sided *t*-test 32
- p-level 34
- significance level 30
- standard deviation 31
- two-sided *t*-test 31
- one-way analysis of variance 44, 46
- OPLS *see* orthogonal PLS (OPLS)
- optimization
  - poly 102
  - sequential 94, 125–132
  - simultaneous 100
  - systematic 93, 94
- ordinal scaled values 96
- ordinary least squares (OLS) regression
  - definition 232
  - matrix notation 232
  - parameter estimation 232, 233
  - parameter significance 233
  - prediction 234, 235
  - singular value decomposition (SVD) 233
- organization for economic cooperation and development (OECD) 352
- orthogonal PLS (OPLS) 238, 239
- outlier 15, 22, 24, 25, 41–43, 52, 179, 186, 190, 196, 227, 231, 247, 249, 250, 253, 254, 272
- outlier testing
  - Dixon's *Q*-test 42, 43
  - Grubbs's test 43
- overfitting 168, 203, 233
- p**
- parallel factor analysis 168
- parameter estimation 105, 121, 123, 136, 215, 223, 231–235, 261, 262, 266, 320, 334
- Pareto optimality 102
- Pareto scaling 138
- partial least squares (PLS) 143, 184, 231, 233, 235, 236, 238, 239, 245–247, 254–257, 264
- partial least squares regression (PLS) 235, 238, 247
- partition 16, 44, 105, 179–182, 201, 202, 267
  - pattern 179–182, 201, 202
- pattern recognition 10, 11, 41, 135–211
- pattern recognition and classification *see* cluster analysis; supervised learning methods; unsupervised learning methods
- PCR *see* regression, principal component
- peak separation after Kaiser 101
- peak shape
  - Gaussian 61, 74
  - Lorentz 58, 61, 65
- Pearson distance 173
- percentile 23
- perceptron 200, 314, 316, 317, 320, 344

- performance characteristic 95, 96, 100, 101, 285, 346
  - see also* objective function
- periodicity 84, 86
- pharmaceutical industry 118, 240
- Phenol 29
- pH monitoring 83
- Phosgene 280, 281, 283
- Phosphate 174, 175
- photomultiplier 3, 20, 73
- p-level 34, 49, 50, 252
- Poisson distribution 20, 21
- polycyclic aromatic hydrocarbons (PAH) 42, 161
- polyoptimization 102
- Potassium 45, 46, 49
- Powell's method 129
- p-Phenylenediamine (PPD) 111, 113, 114, 120–123, 129–132
- precision 10, 15, 16, 52, 93, 94, 98, 105, 121, 233, 237, 243, 346, 347
- prediction 2, 63, 84, 90, 164, 203, 204, 215, 224–226, 229, 234–236, 238, 239, 246–249, 252–255, 257, 263, 264, 268, 269, 320, 321, 338
  - error 235, 247–249, 255, 268
- principal component 141, 143, 146–149, 195, 196, 210, 231, 235, 237, 239, 243, 246–248, 254, 255, 288
  - analysis (PCA) 139, 141, 143, 144, 146, 147, 150, 151, 153, 156–158, 165–167, 170, 210
- principal component analysis (PCA) 141, 153, 165, 167
- principal component regression (PCR) 235, 246
- principal factor analysis 156
- probability 16–18, 20, 21, 23, 27–30, 36, 191, 192, 222, 234, 322, 323, 335–339, 356
- probability density function *see* frequency function
- probability function of Poisson distribution 20, 338
- process
  - random 88
  - stationary 85, 87
- productivity 10
- programming 3, 7, 9, 298, 302–304
- programming in logic (PROLOG) 7, 298, 303–306
- PROLOG language 303–305
- Protein 80, 81, 169–171, 213, 239, 240, 251–253, 278
  - Apomyoglobin 169
- pseudo-inverse matrix 235
  
- q**
- Q analysis 151, 158
- Q and R analysis 151
- quadratic discriminant analysis (QDA) 209
- quality assurance
  - external 350
  - internal 347
- quality control 10, 333, 345–349
- quality, definition 345
- quantile 22, 23, 31, 32, 347, 348
- quartiles 22–25, 23
- quartimax criterion 160
  
- r**
- R analysis 151, 154, 158
- random access memory (RAM) 6 *see also* memory, random access
- random errors 16, 98
- randomization 103, 104, 133
- random number distribution
  - central limit theorem 21, 22
  - classes 16
  - frequency distribution 17
  - Gaussian distribution 17, 19
  - moments of distribution 19
  - Poisson distribution 20, 21
  - spectrophotometric measurements 16, 17
- range
  - analytical 73, 96, 97, 346, 347
  - dynamic 97, 346, 347
  - scaling 138
- rank 22, 96, 137, 141, 154, 159, 235, 248, 253, 366
- rank annihilation 141
- read-only memory (ROM) 6 *see* memory; read-only (ROM)
- recovery function 252, 262, 268, 269, 346
- recursive filters 62
- reduced correlation matrix 151
- regression
  - linear 78, 87, 90, 161, 163, 187, 213, 214, 216, 217, 220, 227, 228, 231, 258, 260, 262, 266, 269, 271, 292, 340, 346
  - multiple linear 161, 163, 213, 231–257, 292
  - nonlinear (NLR) 258, 259, 264, 266, 267
  - nonparametric 267
  - ordinary least squares (OLS) 231–233, 235, 243, 245, 252, 253
  - partial least squares 235, 238, 247
  - principal component (PCR) 231, 233, 235, 245, 246, 254, 255
  - robust 228, 231
  - tree-based regression 268, 269
  - univariate linear 214–231
  - weighted 227–230
- regression diagnostic 213, 227, 231, 247, 253, 254, 262
- regression spline *see* spline
- regression trees 268, 269

- relative cardinality 328, 329
- replication 103, 105, 115, 117, 118, 218
- residual
  - analysis 226–228
  - studentized 249–251
- resolution
  - analytical 99
  - chromatographic 100
  - visual 56, 64, 65
- response 49, 50, 93–95, 101, 102, 106–127, 129–132, 186, 201, 204, 213, 234, 243, 257, 258, 267–269, 335, 336
- response surface designs
  - blocking of experiments 124, 125
  - Box–Behnken design 117
  - central composite design 115, 117
  - factor effects *vs.* regression parameters 123, 124
  - mixture designs 118, 119
  - response surface methods 119, 123
- response surface methods (RSM) 93, 95, 119–121, 123, 213
- risk 23, 28, 30, 31, 35, 40, 42, 50, 100, 191, 214, 222, 234, 249, 358, 359
- robot
  - geometry 8, 9
  - oblique 159
  - orthogonal 159
  - target 160
- robotics and automation 8, 10
- robustness 93, 231, 250, 346, 347, 353
- robust regression 231
- rotation of factors 159, 160
- round-robin test 48, 347, 350
- rule 2, 3, 9, 15, 30, 58, 67, 68, 71, 114, 143, 169, 191, 198, 277, 284, 298–305, 312, 316, 319, 329, 330, 337
- rule-based fuzzy systems 329, 332
- S**
- sample 5, 15, 82, 99, 135, 215, 274, 306, 346
- Savitzky–Golay filter
  - filter coefficients 58
  - signal derivatives 63, 66
- scaling 78, 80, 96, 106, 137–140, 148, 151, 173, 237, 250, 345
  - auto 138, 139, 210
  - methods 138
  - range 137, 138, 210
- score 141–144, 146–151, 156–158, 162, 189, 192, 196, 197, 210, 236–239, 247, 248, 264
- screening designs
  - factor effects estimation 110, 114
  - fractional factorial designs 107, 110
  - full-factorial design 106
- scree test 144, 145, 158
- scree test principal components 145
- search methods 95, 259, 261, 273
- search problems *see* genetic algorithms
- search strategies 273, 286, 302, 334–342
- seed 172
- selection 58, 94, 104, 227, 286, 288, 290, 299, 306, 313, 317, 334–336, 338, 340, 342
- selective analytical method 98
- selectivity 94, 98–102, 129, 240, 350
- self-organizing feature maps 318 *see* Kohonen network; Kohonen networks
- self-organizing nets *see* Kohonen networks
- semantic nets 299, 300
- sensitivity 93, 94, 96, 97, 99, 104
- separation 2, 9, 100, 101, 129, 165, 186, 189, 193, 198, 200, 257, 259, 277, 278, 306, 314, 315
- sequential optimization 125, 132
- sequential search strategy 286
- sequential strategies 95
- shell 7, 304, 306
- shot noise 72 *see* noise, white
- signal
  - derivative 63, 64
  - filtering 72, 73, 75
  - Gaussian 165
  - integration 66–68
  - Lorentz 76
  - processing 10, 55–90
  - smoothing 75, 76
- signal-to-background ratio 100, 101
- signal-to-noise ratio 72, 93, 100
- significance
  - level 30, 32, 34, 35, 37, 39, 42, 43, 48, 49, 52, 53, 222–224, 226, 233, 234, 252, 347, 348, 361
  - of parameters 233–234
- significant factors 158, 159
- similarity measure 135, 172–174, 286–289
- similarity measures for spectra
  - correlation measures 288
  - dissimilarity measures 288
  - grouping and feature selection 288, 289
  - logical operations 289, 290
- simplex
  - fixed-size 127, 128
  - variable-size 128, 129
- simplex method 10, 93, 125–132, 261, 339
- simplex optimization methods 125, 132
- Simpson rules 67, 68
- simulated annealing (SA) 338–340
- simultaneous strategies 94
- singular value decomposition (SVD) 141, 143, 148, 149, 153, 157, 158, 233, 235, 245, 246, 257

- skewness 19, 22, 24
  - smoothing algorithm 263
  - smoothing splines 78
  - soft computing 11, 297–342
  - soft independent modeling of class analogies (SIMCA) 195
  - soft modeling 244–246
  - software 2, 6, 9, 34, 39, 49, 50, 53, 56, 108, 146, 156, 247, 284, 369
  - soil 42, 83, 210
  - source file 275
  - specific analytical method 99
  - specificity 98
  - spectra interpretation 11, 306
  - spectra simulation 292, 294
  - spline
    - adaptive 77
    - function 75–78, 265
    - interpolating 77
    - regression 258, 265
    - smoothing 55, 78
  - spline functions
    - definition 76
    - interpolating splines 77
    - smoothing 78
  - standard deviation
    - of the procedure 346
    - relative 20, 23, 24
  - standard error
    - of calibration (SEC) 247, 253
    - of prediction (SEP) 229, 236, 238, 247, 248, 253
  - standard normal distribution 28, 31 *see* normal distribution
  - standard operating procedure (SOP) 347, 348
  - standard solution 271, 346, 347
  - Staphylococcus nuclease (SNase) 81
  - star and sun-ray plots 182
  - star design *see* experimental design
  - star plot 182–183
  - statistical testing
    - hypotheses (*see* hypotheses testing)
    - normal distribution 28, 30
    - outlier (*see* outlier testing)
  - statistical tests 28, 30, 32, 34, 41, 94, 104, 216
  - statistics
    - descriptive 16–28, 87
    - inference 10, 15, 16
  - steel 35, 48, 49, 285, 299, 306
  - storage
    - bulk 7, 275, 276
    - optical 3, 20, 50, 167, 314
  - straight-line model *see* regression, univariate linear
    - ANOVA 216, 218
    - coefficient of determination 220, 221
    - confidence bands 224, 226
    - confidence intervals 223, 224
    - correlation coefficient 220, 221
    - lack-of-fit test 221, 222
    - parameter estimations 214, 215
    - prediction 215
    - residual analysis 226, 227
    - residual variance 215
    - robust regression 231
    - standard deviation 215
    - standard generalization 216
    - weighted regression 228, 230
  - structure activity relationships 2, 294, 342
  - structure generator 280
  - studentized residual 249–251
  - Student's t-value 31, 33, 112, 249
  - substructure search 284, 291
  - Sulfate 84, 203
  - sum of residuals 220
  - sum of squares
    - mean (MSS) 220–224, 252
    - predictive residual (PRESS) 249, 253–255, 267
  - Sun ray plot 182, 183
  - supervised competitive learning 313
  - supervised learning 311, 332
  - supervised learning methods
    - discriminant analysis 186
    - k-nearest neighbor method (k-NN) 193
    - linear learning machine (LMM) 184
    - SIMCA 195
    - tree-based classification 201
    - support vector machines 198
  - supervision 9
  - support vector machines (SVM) 184, 198, 200, 207–209
  - symbolic knowledge processing 299, 329, 344
    - inferencing 300
    - knowledge representation 299
    - programming 302, 303
    - search strategies 302
  - symbolic programming 298, 302–304
  - symmetry 105, 114
  - Symmlet 79
  - systematic errors 16, 98
  - systematic optimizations 94, 95
- t**
- Tabu search (TS) 340–342
  - tail area 27, 28
  - target-transformation factor analysis (TTFA) 150, 160
  - Taylor expansion 261
  - ternary mixture 119

- test
    - chi-square 360
    - distribution 15, 20, 30, 31, 36–43
    - for distributions 36
    - F-32
    - Kolmogorov–Smirnov 36, 38, 39, 361
    - nonparametric 31, 193, 213, 258, 266, 267
    - outlier 41–43
    - t-, extended 35, 36
    - t-, general 30, 35
    - t-, one-sided 31–34
    - t-, one-variable 30
    - t-, two-sided 15, 31, 33, 34, 36
    - t-, two-variable 34
    - Wilk's 50
  - textile industry 118
  - time-series analysis 16, 55–90
    - autocorrelation (*see* autocorrelation)
    - autocovariance 85
    - autoregression 89, 90
    - cross-correlation 89
  - Titanium 35, 36
  - traceability 350
  - transfer function 306, 308, 320
  - transformation
    - Fourier 10, 68–70, 73, 74, 210, 275
    - – discrete 78
    - Hadamard 10, 55, 68, 70, 71
    - inverse 70
    - wavelet discrete (DWT) 80, 81
  - trapezoidal rule 67
  - tree-based classification
    - CART 201, 202, 209
    - classification 201, 203
    - ensemble methods 204, 207
    - recursive binary partitioning 201
  - trueness 10, 16, 93, 98, 346, 347, 350
  - truth value 5, 16, 30, 31, 98, 235, 303, 325, 327, 350
  - Tucker models 168, 170, 172, 256
  - two-variable *t*-tests 34
  - two-way analysis of variance 46, 49
- u**
- uncertainty 15, 16, 25–28, 63, 323, 324 *see also*
    - measurement uncertainty
    - vs. error 27
    - error propagation 25
    - measurement uncertainty 15
    - random and systematic errors 16
    - sources 25
    - vagueness 16
    - variability 15
  - uncontrolled factors 104, 105
  - undirected graphs 281
  - univariate linear regression
    - straight-line model 214 (*see also* straight-line model)
  - unknown factors 104
  - unselective/partially selective method 99
  - unsupervised competitive learning 313
  - unsupervised learning 311
  - unsupervised learning methods
    - cluster analysis 172
    - factorial methods 141
    - graphical methods 182
  - urine 88
  - UV spectroscopy 161
- v**
- vagueness 15, 16
  - validation 10, 78, 144, 145, 331, 332, 346, 351
  - variability 15, 16, 332
  - variable-size simplex optimization method 128, 132
  - variance
    - explained 144, 145
    - meaningful 197
    - residual 145, 196–198, 215, 223
  - variance–covariance matrix 139 *see* covariance matrix
  - variate 20, 28, 357
  - standard normal 28, 31, 357
  - varimax criterion 159, 160
- w**
- Walsh function 70, 75
  - Walter Shewhart 348
  - Ward's method 179
  - warping 82
  - water 29, 31, 32, 34, 45, 49, 52, 83, 119, 213, 240, 288, 325
  - wavelet analysis 78, 82
  - weighing 9, 256, 257
  - weight 42, 67, 101, 118, 185, 186, 198, 199, 204, 228, 229, 236, 306, 310, 312, 313, 317, 319, 320, 344
  - weighted regression 228, 230
  - Welch test 35
  - Wilcoxon test 31
  - Word 2, 5, 6, 31, 87, 141, 222
  - working range 97 *see* range, analytical
- x**
- X-ray analysis 20
  - X-ray diffraction spectrometry 275
  - X-ray fluorescence analysis 101, 299, 322, 323, 333, 338















