

Index

- acute angle, 58
- addition
 - audio, 14
 - function, 159
 - matrix, 116
 - vector, 11
- adjacency matrix, 112, 133, 186
- advertising, 125, 234, 341
- affine
 - approximation, 35
 - combination, 17
 - function, 32, 149
 - versus linear, 33
- Affleck, Ben, 84
- age group, 337
- algorithm
 - augmented Lagrangian, 422
 - back substitution, 207
 - computing matrix inverse, 209
 - constrained least squares, 347
 - forward substitution, 207
 - Gauss–Newton, 386
 - Gram–Schmidt, 97, 190
 - k -means, 74
 - least norm, 351
 - least squares, 231
 - Levenberg–Marquardt, 386
 - modified Gram–Schmidt, 102
 - Newton, 388
 - penalty, 421
 - QR factorization, 190
 - solving linear equations, 208
- aligned vectors, 58
- α (alpha), 251
- angle, 56
 - acute, 58
 - document, 58
 - obtuse, 58
 - orthogonal, 58
- annualized return and risk, 359
- anti-aligned vectors, 58
- approximation
 - affine, 35
 - least squares, 226
 - Taylor, 35
- AR model, 28, 164, 259, 280, 283
- argmax, 300
- argument of function, 29
- asset
 - allocation, 357
 - alpha and beta, 251
 - return, 8
 - return matrix, 110
 - risk-free, 358
- attribute vector, 10
- audio
 - addition, 14
 - mixing, 18, 121
- augmented Lagrangian algorithm, 422
- auto-regressive model, *see* AR model
- average, 20
- avg (average), 20
- back substitution, 207
- back-test, 127
- backslash notation, 209, 221, 232
- balancing chemical reactions, 154, 211
- basis, 91
 - dual, 205
 - functions, 246
 - orthonormal, 96
- β (beta), 251
- bi-criterion least squares, 311
- bi-linear interpolation, 162
- big-times-small-squared rule, 333, 442
- bill of materials, 12
- birth rate, 165, 219
- bit, 22
- block
 - matrix, 109, 179
 - vector, 4
- Boeing 747, 379
- Boole, George, 10
- Boolean
 - classification, 285
 - features, 38, 281
 - least squares, 435
 - vector, 10, 26, 87
- Bowie, David, 84
- byte, 22, 122
- calculus, 35, 228, 344, 382, 443

- cash flow, 27, 125
 - discounted, 22
 - net present value, 22
 - replication, 18, 94
 - vector, 8, 93
- categorical feature, 270
- Cauchy, Augustin-Louis, 57
- Cauchy–Schwarz inequality, 56, 68
- centroid, 74
- chain graph, 136, 317
- chain rule, 184, 444, 447
- channel equalization, 146
- Chebyshev inequality, 47, 54, 64, 305
- Chebyshev, Pafnuty, 47
- chemical
 - equilibrium, 384
 - reaction balance, 154
- circular difference matrix, 319
- circulation, 134
- classification, 285
 - Boolean, 285
 - handwritten digits, 290, 404
 - iris flower, 289
 - multi-class, 297
- classifier
 - least squares, 288
 - one-versus-others, 299
- closed-loop, 186
- cluster centroid, 74
- clustering, 69
 - digits, 79
 - objective, 72
 - optimal, 73
- co-occurrence, 20
- coefficients
 - linear equations, 152
 - matrix, 107
 - vector, 3
- colon notation, 4
- color vector, 6
- column-major, 159
- communication channel, 138
- compartmental system, 174
- completing the square, 242
- complexity, 22
 - k -means algorithm, 79
 - Gram–Schmidt algorithm, 102
 - matrix-matrix multiply, 182
 - matrix-vector multiplication, 123
 - vector operations, 24
- compliance matrix, 150
- computer representation
 - matrix, 122
 - vector, 22
- confusion matrix, 287
- conservation of mass, 156
- constrained least squares, 339
 - solution, 344
 - sparse, 349
- constrained optimization, 448
 - KKT conditions, 449
- contingency table, 111
- control, 314
 - closed-loop, 186
 - linear quadratic, 366
 - nonlinear, 425
 - state feedback, 185
- controllability matrix, 195
- convolution, 136
- correlation coefficient, 60, 251
- covariance matrix, 193
- cross product, 159
- cross-validation, 264
 - efficient, 284
- currency exchange rate, 26, 125
- customer purchase
 - matrix, 111
 - vector, 10
- cycle, 145, 195
- data fitting, 245
- data matrix, 112, 116
- de-meaned vector, 52
- de-meaning, 149
- de-trended, 252
- de-tuning, 325
- death rate, 165, 219
- decision threshold, 294
- deformation, 150
- demand, 150
 - elasticity matrix, 150
 - shaping, 315
- dependent variable, 38
- dependent vectors, 89
- derivative, 35, 443
 - chain rule, 184, 444, 447
 - partial, 444
- diag**, 114
- diagonal matrix, 114
- diet, 160
- difference matrix, 119, 317
- difference of vectors, 11
- difference vector, 26
- diffusion, 155
- digits, 79
- dilation, 129
- dimension
 - matrix, 107
 - vector, 3
- directed graph, 112, 132, 186
- Dirichlet energy, 66, 135, 144, 145, 241, 317, 322, 324
- Dirichlet, Peter Gustav Lejeune, 66

- discount factor, 368
- discounted cash flow, 22
- discretization, 170
- disease dynamics, 168
- displacement, 12
- distance, 48
 - spherical, 58
- distributive property, 16, 19, 121, 127
- document
 - angle, 58
 - dissimilarity, 50
 - scoring, 121
 - topic discovery, 82
 - word count, 9
- document-term matrix, 116
- dot product, 19
- down-sampling, 131, 144
- dual basis, 205
- dynamics
 - epidemic, 168
 - matrix, 163
 - supply chain, 171
- edge, 112
- EHR, 65
- elastic deformation, 150
- elasticity, 150, 315
 - matrix, 336, 394
- electronic health record, *see* EHR
- energy use patterns, 71
- epidemic dynamics, 168
- equality
 - matrices, 107
 - vectors, 3
- equalization, 146, 240, 318
- equations
 - homogeneous, 153
 - KKT, 345
 - nonlinear, 381
 - normal, 229
- equilibrium, 162
 - chemical, 384
 - linear dynamical system, 174
 - mechanical, 384
 - Nash, 385
 - prices, 384
- error rate, 287
- Euclidean
 - distance, 48
 - norm, 45
- Euler, Leonhard, 170
- exogenous flow, 134
- expansion in a basis, 92
- expected value, 21
- exponential weighting, 368
- factor-solve method, 208
- false alarm rate, 287
- Fast Fourier Transform, *see* FFT
- feature
 - categorical, 270
 - distance, 50
 - engineering, 269, 293, 330
 - Likert, 270
 - matrix, 112, 152
 - neural network, 273
 - random, 273, 293, 406, 409
 - standardized, 269
 - TFIDF, 273
 - vector, 10, 245
 - winsorized, 269
- FFT, 140
- Fibonacci sequence, 175
- Fibonacci, Leonardo of Pisa, 175
- Fisher, Ronald, 289
- floating point
 - number, 22, 102
 - operation, *see* flop
 - round-off error, 23
- flop, 23
- flow conservation, 133, 156
 - with sources, 134
- forgetting factor, 368
- forward substitution, 207
- Fourier
 - approximation, 283
 - transform, 140
- Fourier, Jean-Baptiste, 140
- friend relation, 116
- Frobenius norm, 118
- Frobenius, Ferdinand Georg, 118
- function
 - affine, 32, 149
 - argument, 29
 - basis, 246
 - composition, 183
 - inner product, 30
 - linear, 30, 147
 - notation, 29
 - objective, 226, 419
 - rational, 160, 218, 282
 - reversal, 148
 - running sum, 149
 - sigmoid, 390, 413
 - sum, 159
- Galton, Sir Francis, 279
- Game of Thrones, 84
- Gauss, Carl Friedrich, 102, 161, 207, 225, 386
- Gauss–Newton algorithm, 386
- generalization, 260
- generalized additive model, 271
- gigabyte, 23

- gigaflop, 23
- global positioning system, *see* GPS
- gone bust, 358
- GPS, 373, 386
- gradient, 228, 445
- Gram matrix, 181, 214, 229, 250, 318, 332, 378
- Gram, Jørgen Pedersen, 97
- Gram–Schmidt algorithm, 97, 190
 - complexity, 102
 - modified, 102
- graph, 112, 132, 186
 - chain, 136
 - circle, 145
 - cycle, 145, 195
 - social network, 116
 - tree, 145
- grayscale, 9
- group representative, 72
- handwritten digits, 79, 290
- heat flow, 155
- hedging, 62
- Hestenes, Magnus, 422
- histogram vector, 9, 50
- homogeneous equations, 153
- house price regression, 39, 258, 265, 274
- identity matrix, 113
- illumination, 234
- image
 - matrix, 110
 - vector, 9
- impulse response, 138
- imputing missing entries, 86
- incidence matrix, 132, 171
- independence-dimension inequality, 91
- independent vectors, 89
- index
 - column, 107
 - range, 4
 - row, 107
 - vector, 3
- inequality
 - Cauchy–Schwarz, 56, 68
 - Chebyshev, 47, 54
 - independence-dimension, 91
 - triangle, 46, 49, 57
- inner product, 19, 178
 - function, 30
 - matrices, 192
- input, 164
- input-output
 - matrix, 157
 - system, 140, 280, 314
- intercept, 38
- interpolation, 144, 154, 160, 162, 210, 218, 354
- inverse
 - left, 199
 - matrix, 202
 - Moore–Penrose, 215
 - pseudo, 214, 337
 - right, 201
- inversion, 316
 - Tikhonov, 317
- invertible matrix, 202
- iris flower classification, 289, 301
- iterative method for least squares, 241
- Jacobi, Carl Gustav Jacob, 151
- Jacobian, 151, 446
- k -means
 - algorithm, 74
 - complexity, 79
 - features, 273
- Kalman, Rudolph, 374
- Karush, William, 345
- Karush–Kuhn–Tucker, *see* KKT
- Kirchhoff’s current law, 156
- Kirchhoff, Gustav, 156
- KKT
 - conditions, 345, 449
 - matrix, 345
- Kuhn, Harold, 345
- Kyoto prize, 374
- label, 38
- Lagrange
 - multipliers, 344, 448
 - polynomial, 211
- Lagrange, Joseph-Louis, 211
- Lambert function, 412
- Lambert, Johann Heinrich, 412
- Laplace, Pierre-Simon, 192
- Laplacian matrix, 192
- Laplacian regularization, 135, 317, 324
- least squares, 225
 - bi-criterion, 311
 - Boolean, 435
 - classifier, 288
 - data fitting, 245
 - iterative method, 241
 - multi-objective, 309
 - nonlinear, 381
 - recursive, 242
 - residual, 225
 - solution method, 231
 - sparse, 232
- LeCun, Yann, 79
- left inverse, 199
- Legendre, Adrien-Marie, 225

- Leonardo of Pisa, 175
- Leontief input-output model, 157, 174
- Leontief, Wassily, 157
- Levenberg, Kenneth, 391
- Levenberg–Marquardt algorithm, 386
- leverage, 358
- Likert scale, 71, 270, 305
- Likert, Rensis, 71
- line, 18, 65, 365
 - segment, 18
- linear
 - combination, 17
 - dynamical system, 163
 - equations, 147, 152
 - function, 30, 147
 - least squares problem, 226
 - quadratic control, 366
 - sparse equations, 210
 - versus affine, 33
- linear dynamical system, 163
 - closed-loop, 186
 - state feedback, 185
- linearity, 147
- linearly independent
 - row vectors, 115
 - vectors, 89
- link, 133
- Lloyd, Stuart, 74
- loan, 8, 93
- location vector, 6
- logarithmic spacing, 314
- logistic regression, 288
- long-only portfolio, 358
- look-ahead, 266
- loss function, 402
- loss leader, 26
- lower triangular matrix, 114
- market
 - clearing, 14
 - return, 251
 - segmentation, 70
- Markov model, 164, 175
- Markov, Andrey, 164
- Markowitz, Harry, 357
- Marquardt, Donald, 391
- mass, 169
- matrix, 107
 - addition, 116
 - adjacency, 112, 133, 186
 - asset return, 110
 - block, 109, 179
 - cancellation, 217
 - circular difference, 319
 - coefficients, 107
 - compliance, 150
 - computer representation, 122
 - confusion, 287
 - controllability, 195
 - covariance, 193
 - data, 112, 116
 - demand elasticity, 150
 - diagonal, 114
 - difference, 119, 317
 - dimensions, 107
 - document-term, 116
 - dynamics, 163
 - elasticity, 336, 394
 - elements, 107
 - equality, 107
 - feature, 152
 - Gram, 181, 214, 229, 250, 318, 332, 378
 - graph, 112
 - identity, 113
 - image, 110
 - incidence, 132, 171
 - inner product, 192
 - inverse, 199, 202, 209
 - invertible, 202
 - Jacobian, 151, 446
 - KKT, 345
 - Laplacian, 192
 - least squares, 233
 - left inverse, 199
 - Leontief input-output, 157
 - lower triangular, 114
 - multiplication, 177
 - negative power, 205
 - nonsingular, 202
 - norm, 117
 - orthogonal, 189, 204
 - permutation, 132, 197
 - population dynamics, 219
 - power, 186
 - projection, 240
 - pseudo-inverse, 214, 229
 - relation, 112
 - resistance, 157
 - return, 110
 - reverser, 131, 148
 - rotation, 129, 191
 - running sum, 120
 - second difference, 183
 - singular, 202
 - sparse, 114
 - square, 108
 - squareroot, 186, 194
 - stacked, 109
 - state feedback gain, 185
 - subtraction, 116
 - sum, 116
 - symmetric, 116

- tall, 108
- Toeplitz, 138, 280, 316
- trace, 192
- transpose, 115
- triangular, 114, 206
- triple product, 182
- upper triangular, 114
- Vandermonde, 121, 127, 154, 210, 256
- vector multiplication, 118
- wide, 108
- zero, 113
- matrix-vector product, 147
- mean, 20, 21
- mean return, 54
- mechanical equilibrium, 384
- minimum mean square error, *see* MMSE
- missing entries, 86
- mixing audio, 18
- mixture of vectors, 17
- MMSE, 247
- MNIST, 79, 290, 404
- model
 - nonlinear, 386, 399
 - over-fit, 261
 - parameter, 246
 - stratified, 272, 336
 - validation, 260
- modified Gram–Schmidt algorithm, 102
- monochrome image, 9
- Moore’s law, 280
- Moore, Eliakim, 215
- Moore, Gordon, 280
- Moore–Penrose inverse, 215
- motion, 169
- moving average, 138
- μ (mu), 20, 53
- multi-class classification, 297
- multi-objective least squares, 309
- multiplication
 - matrix-matrix, 177
 - matrix-vector, 118
 - scalar-matrix, 117
 - scalar-vector, 15
 - sparse matrix, 182
- Nash equilibrium, 385
- Nash, John Forbes Jr., 385
- navigation, 373
- nearest neighbor, 50, 63, 65, 66, 73, 306
- net present value, *see* NPV
- Netflix, 284
- network, 133
- neural network, 273, 413
- Newton algorithm, 388
- Newton’s law of motion, 42, 169, 343
- Newton, Isaac, 42, 386
- nnz** (number of nonzeros), 6, 114
- Nobel prize
 - Leontief, 158
 - Markowitz, 357
 - Nash, 385
- node, 112
- nonlinear
 - control, 425
 - equations, 381
 - least squares, 381
 - model fitting, 386, 399
- nonnegative vector, 27
- nonsingular matrix, 202
- norm, 45
 - Euclidean, 45
 - Frobenius, 118
 - matrix, 117
 - weighted, 68
- normal equations, 229
- notation
 - function, 29
 - overloading, 5
- NPV, 22, 94, 103
- number
 - floating point, 22
 - of nonzeros, 114
- nutrients, 160, 352
- objective
 - clustering, 72
 - function, 226, 419
- observations, 245
- obtuse angle, 58
- occurrence vector, 10
- offset, 38
- one-hot encoding, 270
- one-versus-others classifier, 299
- ones vector, 5
- open-loop, 368
- optimal clustering, 73
- optimal trade-off curve, 311
- optimality condition
 - least squares, 229
 - nonlinear least squares, 382
- optimization, 447
 - constrained, 448
- order, 24
- orthogonal
 - distance regression, 400
 - matrix, 189, 204
 - vectors, 58
- orthogonality principle, 231
- orthonormal
 - basis, 96
 - expansion, 96
 - row vectors, 115
 - vectors, 95
- out-of-sample validation, 261

- outcome, 245
- outer product, 178
- over-determined, 153, 382
- over-fit, 261
- overloading, 5

- parallelogram law, 64
- parameter
 - model, 246
 - regularization, 328
- Pareto optimal, 311, 360
- Pareto, Vilfredo, 311
- partial derivative, 35, 444
- path, 133, 186
- penalty algorithm, 421
- Penrose, Roger, 215
- permutation matrix, 132, 197
- pharmaco-kinetics, 174
- phugoid mode, 379
- piecewise-linear fit, 256
- pixel, 9
- polynomial
 - evaluation, 21, 120
 - fit, 255
 - interpolation, 154, 160, 210
 - Lagrange, 211
- population dynamics, 164, 188
- portfolio
 - gone bust, 358
 - leverage, 358
 - long-only, 358
 - optimization, 357
 - return, 22, 120, 358
 - risk, 359
 - sector exposure, 161
 - trading, 14
 - value, 22
 - vector, 7
 - weights, 357
- potential, 135, 156
- Powell, Michael, 422
- power of matrix, 186
- precision, 287
- prediction error, 50, 152, 246
- price
 - elasticity, 150, 336
 - equilibrium, 384
 - vector, 21
- probability, 21
- product
 - block matrix, 179
 - cross, 159
 - dot, 19
 - inner, 19, 178
 - matrix-matrix, 177
 - matrix-vector, 147
 - outer, 178
 - projection, 65, 129, 144, 240
 - proportions, 7
 - pseudo-inverse, 214, 229, 337
 - push-through identity, 218, 333
 - Pythagoras of Samos, 60

- QR factorization, 189, 206, 231, 348, 351
- quadrature, 161, 220

- random features, 273, 293, 406, 409
- Raphson, Joseph, 388
- rational function, 160, 218, 282
- recall rate, 287
- receiver operating characteristic, *see* ROC
- recommendation engine, 85
- recursive least squares, 242
- regression, 151, 257
 - house price, 39, 258
 - logistic, 288
 - model, 38
 - to the mean, 279
- regressors, 38
- regularization, 364
 - parameter, 328
 - path, 328, 332
 - terms, 314
- relation, 112
 - friend, 116
- residual, 225, 381, 419
- residual sum of squares, *see* RSS
- resistance matrix, 157
- return, 8, 54
 - annualized, 359
 - matrix, 110
 - vector, 22
- reversal function, 148
- reverser matrix, 131, 148
- RGB, 6
- Richardson, Lewis, 241
- ridge regression, 325
- right inverse, 201
- right-hand side, 152
- risk, 54, 359
- risk-free asset, 358
- RMS, 46
 - deviation, 48
 - prediction error, 50
- rms** (root-mean-square), 46
- ROC, 294
- root-mean-square, *see* RMS
- rotation, 129, 191
- round-off error, 23, 102
- row vector, 108
 - linearly independent, 115
- running sum, 120, 149

- samples, 245

- sampling interval, 170
- scalar, 3
- scalar-matrix multiplication, 117
- scalar-vector multiplication, 15
- scaling, 129
- Schmidt, Erhard, 97
- Schwarz, Hermann, 57
- score, 21
- seasonal component, 255
- seasonally adjusted time series, 255
- second difference matrix, 183
- sector exposure, 27, 161, 352
- segment, 18
- sensitivity, 287
- shaping demand, 315
- short position, 7, 22
- shrinkage, 325
- σ (sigma), 53
- sigmoid function, 390, 413
- sign function, 289
- signal, 7
 - flow graph, 413
- Simpson's rule, 161
- Simpson, Thomas, 161
- singular matrix, 202
- sink, 134
- skewed classifier, 294
- slice, 4, 131
- social network graph, 116
- source, 134
- sparse
 - constrained least squares, 349
 - least squares, 232
 - linear equations, 210, 350
 - matrix, 114
 - matrix multiplication, 182
 - QR factorization, 190
 - vector, 6, 24
- specificity, 287
- spherical distance, 58
- spline, 341
- square
 - matrix, 108
 - system of equations, 153, 382
- squareroot of matrix, 194
- stacked
 - matrix, 109
 - vector, 4
- standard deviation, 52, 248
- standardization, 56
- standardized features, 269
- state, 163
- state feedback control, 185, 335, 371
- std** (standard deviation), 52
- steganography, 354
- Steinhaus, Hugo, 74
- stemming, 10, 82
- stoichiometry, 162
- stop words, 10
- straight-line fit, 249
- stratified model, 272, 336
- subadditivity, 46
- submatrix, 109
- subset vector, 10
- subtraction
 - matrix, 116
 - vector, 11
- subvector, 4
- sum
 - linear function, 159
 - matrix, 116
 - of squares, 20, 45, 247
 - vector, 11
- superposition, 30, 147
- supply chain dynamics, 171
- support vector machine, 288
- survey response, 71
- symmetric matrix, 116
- tall matrix, 108
- Taylor approximation, 35, 64, 151, 185, 387, 443
- Taylor, Brook, 36
- term frequency inverse document frequency, *see* TFIDF
- test data set, 261
- TFIDF, 273
- thermal resistance, 157
- Tikhonov, Andrey, 317
- time series
 - auto-regressive model, 259
 - de-trended, 252
 - prediction validation, 266
 - seasonally-adjusted, 255
 - smoothing, 138
 - vector, 7
- time-invariant, 163
- Toeplitz matrix, 138, 280, 316
- Toeplitz, Otto, 138
- topic discovery, 70, 82
- trace, 192
- tracking, 368
- trade list, 14
- trade-off curve, 311
- training data set, 261
- trajectory, 163
- transpose, 115
- tree, 145
- trend line, 252
- triangle inequality, 46, 49, 57, 118
- triangular matrix, 114, 206
- trim conditions, 379
- true negative rate, 287

- true positive rate, 287
- Tucker, Albert, 345
- uncorrelated, 60
- under-determined, 153, 382
- unit vector, 5
- units for vector entries, 51, 63
- up-conversion, 144
- upper triangular matrix, 114
- validation, 260, 314
 - classification, 288
 - limitations, 268
 - set, 261
 - time series prediction, 266
- Vandermonde matrix, 121, 127, 154, 210, 256
- Vandermonde, Alexandre-Théophile, 121
- variable, 225
- vector, 3
 - addition, 11
 - affine combination, 17
 - aligned, 58
 - angle, 56
 - anti-aligned, 58
 - AR model, 164, 283
 - basis, 91
 - block, 4
 - Boolean, 10, 26, 87
 - cash flow, 8, 93
 - clustering, 69
 - coefficients, 3
 - color, 6
 - components, 3
 - computer representation, 22
 - correlation coefficient, 60
 - customer purchase, 10
 - de-meaned, 52
 - dependent, 89
 - difference, 26
 - dimension, 3
 - distance, 48
 - entries, 3
 - equality, 3
 - feature, 10, 21, 245
 - histogram, 9
 - image, 9
 - independence, 89
 - inner product, 19
 - large, 45
 - linear combination, 17
 - linear dependence, 89
 - linear independence, 89
 - location, 6
 - matrix multiplication, 118
 - missing entry, 85
 - mixture, 17
 - nonnegative, 27
 - occurrence, 10
 - ones, 5
 - orthogonal, 58
 - orthonormal, 95
 - outer product, 178
 - portfolio, 7
 - price, 21
 - probability, 21
 - proportions, 7
 - quantities, 7
 - return, 22
 - RMS deviation, 48
 - RMS value, 46
 - row, 108
 - slice, 4
 - small, 45
 - sparse, 6, 24
 - stacked, 4
 - standardization, 56
 - subset, 10
 - sum, 11
 - time series, 7
 - unit, 5
 - units for entries, 51, 63
 - weight, 21, 38
 - word count, 9, 87
 - zero, 5
- vertex, 112
- video, 9
- warm start, 393
- way-point constraint, 371
- weather zones, 71
- weight vector, 38
- weighted
 - average, 17, 334
 - Gram matrix, 334
 - norm, 68
 - sum, 30
 - sum of squares, 310
- wide matrix, 108
- Wikipedia, 51, 82
- Wilkinson, James H., 114
- Winsor, Charles P., 270
- winsorized feature, 269
- word count
 - TFIDF, 273
 - vector, 9, 50, 87
- z-score, 56, 67, 269
- zero
 - matrix, 113
 - vector, 5
- ZIP code, 71, 274