

---

## Index

- absolute moment, 5, 21
- adjacency matrix, 61
- admissible sequence, 206, 207
- anisotropic random vectors, 40, 134, 136, 257
- approximate isometry, 73, 74, 91, 111
- approximate projection, 75
  
- Bennett's inequality, 36
- Bernoulli distribution, 10, 12
  - symmetric, 13, 25, 46, 62, 136
- Bernstein's inequality, 33, 34, 130
  - for matrices, 113, 119, 120
- binomial
  - coefficients, 4
  - distribution, 11
- bounded differences inequality, 36
- Brownian motion, 148–150
  
- canonical metric, 149, 160
- Caratheodory's theorem, 1, 2
- Cauchy–Schwarz inequality, 6
- centering, 28, 32, 103
- central limit theorem
  - Berry–Esseen, 13
  - de Moivre–Laplace, 10
  - Lindeberg–Lévy, 9
  - projective, 54
- chaining, 178
- chaos, 127
- Chebyshev's inequality, 8
- Chernoff's inequality, 17, 18, 36
- Chevet's inequality, 212, 258
- clustering, 95
- community detection, 87
- concentration
  - for anisotropic random vectors, 134
  - Gaussian, 104
  - of the norm, 39, 134
  - on  $SO(n)$ , 107
  - on a Riemannian manifold, 106
  - on the ball, 108
  - on the cube, 105, 108
  - on the Grassmannian, 107
  - on the sphere, 99, 103
  - on the symmetric group, 106
  - Talagrand's inequality, 110
- contraction principle, 143, 156
  - Talagrand's, 145
- convex
  - body, 50
  - combination, 1
  - hull, 1, 163
  - program, 59
- coordinate distribution, 49, 52
- Coupon collector's problem, 120
- Courant–Fisher's min–max theorem, *see* min–max theorem
- covariance, 6, 41, 93, 94, 122, 224
  - estimation, 94, 122, 223
  - of a random process, 149
- covering number, 4, 75, 77, 78, 160, 162, 193
- Cramér–Wold theorem, 266
- cross-polytope, 165
  
- Davis–Kahan theorem, 89, 96
- de Moivre–Laplace theorem, 10
- decoding map, 81
- decoupling, 127, 128
- degree of a vertex, 19
- diameter, 2, 80, 163, 227
- dimension reduction, 110
- discrepancy, 199
- distance to a subspace, 136
- Dudley's inequality, 176–178, 181, 183, 206, 210
- Dvoretzky–Milman theorem, 259, 260
  
- Eckart–Young–Mirsky theorem, 73
- embedding, 258
- empirical
  - distribution function, 198
  - measure, 188
  - method, 1, 2
  - process, 183, 186, 196
  - risk, 202
- encoding map, 81
- entropy function, 97

282

$\varepsilon$ -net, *see* net  
 $\varepsilon$ -separated set, 76  
 Erdős–Rényi model, 19, 87  
 error correcting code, 81  
 escape theorem, 229, 230, 241, 242, 244  
 exact recovery, 241  
 excess risk, 202  
 expectation, 5  
 exponential distribution, 32

feature map, 68  
 frame, 49, 53  
   tight, 49  
 Frobenius norm, 72, 169  
 functions of matrices, *see* matrix calculus

$\gamma_2$ -functional, 207  
 Garnaev–Gluskin theorem, 239  
 Gaussian  
   complexity, 170, 216, 255  
   distribution, 9  
   integration by parts, 152, 153  
   interpolation, 152  
   measure, 104  
   mixture model, 95, 96  
   orthogonal ensemble, 159  
   process, 150  
     canonical, 161, 162  
     width, 162, 211  
 generic chaining, 206, 208, 209  
 Gilbert–Varshamov bound, 97  
 Glivenko–Cantelli  
   class, 199  
   theorem, 198  
 Golden–Thompson inequality, 116  
 Gordon’s inequality, 157, 159  
 gram matrix, 266  
 graph, 61  
   simple, 61  
 Grassmannian, 107  
 Grothendieck’s identity, 63  
 Grothendieck’s inequality, 55, 67

Haar measure, 107, 108  
 Hamming  
   bound, 97  
   cube, 79, 82, 105, 269  
   distance, 79, 105, 106  
 Hanson–Wright inequality, 130, 131, 134, 135  
 Hermitization trick, 140  
 Hessian, 109  
 Hilbert–Schmidt norm, *see* Frobenius norm  
 Hoeffding’s inequality, 14, 16, 27  
   for matrices, 119  
   general, 27  
 Hölder’s inequality, 6  
 hypothesis space, 202, 204

Index

increments of a random process, 149, 177, 217, 255  
 independent copy of a random variable, 128  
 indicator random variables, 12  
 integer optimization problem, 59  
 integral identity, 7  
 intrinsic dimension, 123, 175  
 isoperimetric inequality, 100, 105  
 isotropic random vectors, 43

Jensen’s inequality, 6, 116  
 Johnson–Lindenstrauss lemma, 110, 111, 170, 225,  
   226, 258

Kantorovich–Rubinstein duality theorem, 188  
 kernel, 65, 68  
   Gaussian, 68  
   polynomial, 68  
 Khintchine’s inequality, 27  
   for matrices, 120

Lasso, 247, 248, 251, 252  
 law of large numbers, 2, 9, 35, 93, 184  
   uniform, 185, 188  
 Lieb’s inequality, 116  
 linear regression, *see* regression  
 Lipschitz  
   function, 98  
   norm, 98  
 low-rank approximation, 73  
 $L^p$  norm, 5  
 $L_{\psi_1}$  norm, 31  
 $L_{\psi_2}$  norm, 24

$M^*$  bound, 227, 229, 234, 239  
 majority decoding, 81  
 majorizing measure theorem, 210  
 Markov’s inequality, 8  
 matrix  
   Bernstein inequality, *see* Bernstein’s inequality  
     for matrices  
   deviation inequality, 216, 255  
   Khintchine’s inequality, 119, 126  
 matrix calculus, 114  
 matrix completion, 140  
 matrix recovery, 239, 241  
 maximum cut, 61  
 McDiarmid’s inequality, *see* bounded differences  
   inequality  
 mean width, *see* spherical width, 165  
 measurements, 232  
 median, 102  
 metric entropy, 79, 80, 160, 176, 177, 187, 194  
 min–max theorem, 71  
 Minkowski’s inequality, 6  
 Minkowski sum, 77  
 moment, 5, 21  
 moment generating function, 5, 14, 22, 25  
 Monte–Carlo method, 184

- net, 75, 83, 84
- network, 19, 87
- non-commutative Bernstein inequality, *see* Bernstein's inequality for matrices
- non-commutative Khintchine inequalities, *see* matrix Khintchine inequalities
- normal distribution, 9, 12, 21, 46, 47, 51
- nuclear norm, 240
  
- operator norm, 71, 83, 84
- ordinary least squares, 248
- Orlicz
  - norm, 33
  - space, 32, 33
  
- packing number, 76
- Pajor's lemma, 191
- perturbation theory, 89
- Poisson
  - distribution, 10, 18, 32
  - limit theorem, 10
- polarization identity, 266
- positive-homogeneous function, 254
- principal component analysis, 42, 93, 95
- probabilistic method, 4, 194
- push-forward measure, 108
  
- Rademacher distribution, 14
- radius, 2, 212
- random
  - field, 148
  - graph, 19, 37
  - matrix
    - norm, 85, 139, 157, 159, 212
    - singular values, 91, 159, 222
  - process, 147
  - projection, 110, 111, 170, 171, 222, 247, 259
  - sections, 227
  - walk, 148
- randomized rounding, 63
- rate of an error correcting code, 83
- regression, 233
- regular graph, 19
- reproducing kernel Hilbert space, 68
- restricted isometry, 245, 246
- Riemannian manifold, 106
- RIP, *see* restricted isometry
- risk, 201, 205
- rotation invariance, 26, 46, 111
  
- sample
  - covariance, 93
  - mean, 9
- Sauer–Shelah lemma, 193
- second moment matrix, 41, 93, 122
- selectors, 129, 140
- semidefinite
  - program, 59
  - relaxation, 59, 62
- shatter, 189
- signal, 232
- singular
  - value decomposition, 70, 131
  - values, 70
    - of random matrices, 91, 222
  - vectors, 70
- Slepian's inequality, 151, 154–157
- small ball probabilities, 16, 41
- sparse recovery, 232, 237, 242
- special orthogonal group, 107
- spectral
  - clustering, 90, 91, 95, 96, 121
  - decomposition, 42, 114
  - norm, *see* operator norm
- spherical
  - distribution, 45, 48, 53, 164
  - width, 164, 165, 171
- stable
  - dimension, 167, 168, 173, 174, 226, 228, 261
  - rank, 124, 169, 224, 263
- standard
  - deviation, 6
  - score, 43
- statistical learning theory, 200
- stochastic
  - block model, 87, 91
  - domination, 151
  - process, *see* random process
- sub-exponential
  - distribution, 28, 29, 31
  - norm, 31
- sub-gaussian
  - distribution, 22, 24, 26, 33, 51
  - increments, 177
  - norm, 24
  - projection, 223
- subadditive function, 254
- Sudakov's minoration inequality, 160, 182
- Sudakov–Fernique inequality, 156–158, 161, 211, 269
- support function, 255, 260
- symmetric
  - Bernoulli distribution, *see* Bernoulli distribution, symmetric distributions, 136
  - group, 106
- symmetrization, 136–138, 144, 149
  - for empirical processes, 197
  
- tails, 7
  - normal, 12
  - Poisson, 18
- Talagrand's
  - comparison inequality, 210, 211
  - concentration inequality, 110

284

contraction principle, 145, 156  
tangent cone, 242  
target function, 200  
tensor, 65  
trace  
  inequalities, 115, 116  
  norm, *see* nuclear norm  
training data, 200  
transportation cost, 188  
truncation, 13, 57  
  
union bound, 20

*Index*

variance, 5  
VC dimension, 189, 194, 196  
  
Wasserstein's  
  distance, 188  
  law of large numbers, 185  
Weyl's inequality, 89  
Wiener process, 148  
  
Young's inequality, 31  
  
zero-one law, 102