

Index

- A-stability, 322, 323
- acquisition function, 254
- Adam, 230
- affine transformation, 23
- agent, 3, 4, 11
- aleatory uncertainty, 11, 71
- analytic, 1
- Arenstorf orbit, 336
- Armijo condition, *see* line search
- Arnoldi process, 135
- atomic operation, 70
- Automated Machine Learning, 276
- average-case, 182
- average-case analysis, 7
- backpack package, *see* software219
- Bayes' theorem, 21
- Bayesian, 8
- Bayesian ODE filters and smoothers, *see* ODE filters and smoothers
- Bayesian Optimisation, 251
- Bayesian quadrature, 72
- belief propagation, 43
- BFGS, 236
- bias, 11
- bifurcation, 311
- boundary value problem, 286, 339
- Brownian motion, 50
- calibration, 12
- CARE, *see* Riccati equation
- Cauchy-Schwarz inequality, 37
- Chain graph, 42
- chaos, 332
- Chapman-Kolmogorov equation, 43, 45
- Chebyshev polynomials, 103
- Cholesky decomposition, 36, 138
- code, *see* software
- companion matrix, 53
- conditional distribution, 21
- conditional entropy, 272
- conjugate gradients, 134, 137, 144, 145, 165, 166
 - probabilistic, 165
- conjugate prior, 55, 56
- continuous time, 48, 296
- continuous-time Riccati equation, *see* Riccati equation
- convergence rate, 201
- convex function, 199
- covariance, 23
- covariance function, *see* kernel
- cubic splines, *see* splines
- curse of dimensionality, 79
- Dahlquist test equation, 323
- DARE, *see* Riccati equation
- data, 21
- decision theory, 4
- Dennis family, 236
- detailed balance, 85
- determinant lemma, 130
- determinantal point process, 80
- DFP, 237
- Dirac delta, 27
- discrete time, 42, 297
- discrete-time algebraic Riccati equation, *see* Riccati equation
- dynamic model, 45, 298
- early stopping, 224
- EKF0, EKF1, *see* ODE filters and smoothers
- EKS0, EKS1, *see* ODE filters and smoothers
- empirical risk minimisation, 37, 204
- emukit package, *see* software
- entropy, 23
- epistemic uncertainty, 11, 70
- equidistant grid, 92
- ergodicity, 85
- error analysis vs error estimation, 92
- error function, 215
- Euler's method, 289
- Eulerian integrals, 56
- evidence, 99
- expected improvement, 259
- expensive evaluations, 256
- exploration-exploitation trade-off, 247
- exponential kernel, *see* Ornstein-Uhlenbeck process
- exponentiated quadratic kernel, *see* Gaussian kernel
- filter, 44
 - Kalman, 46
 - ODE, *see* ODE filters and smoothers
 - optimal, 46
 - particle, 309
- forward-backward algorithm, *see* sum-product algorithm
- Frobenius norm, *see* norm
- function-space view, 28
- Galerkin condition, 143
- gamma distribution, 56
- gamma function, 56
- Gauss-Markov process, 41
- Gauss-inverse-gamma, 56
- Gauss-inverse-Wishart, 59
- Gaussian
 - elimination, 133

396 Index

- Gaussian distribution, 23
 Gaussian kernel, 32, 40
 Gaussian ODE filters and smoothers, *see* ODE filters and smoothers
 Gaussian process, 30, 34
 Gaussian quadrature, 102
 Gaussian vs nonparametric, 314, 336
 general linear model, *see* linear regression
 generative model, 73
 genetic algorithm, 200
 Gittins indices, 255
 GNU C library, 70
 GPU, 1, 8
 gradient descent, 201
 Gram matrix, 36
 graphical model, 43
 Greenstadt's method, 236
- Hamiltonian matrix, 54
 harmonic potential, 52
 heavy ball, *see* momentum
 Hermite interpolation, 291
 Hermite polynomials, 103
 hierarchical inference, 55
 hyperparameter, 55
 Hölder continuity, 95
- importance sampling, 71, 310
 sequential, 311
 inference, 22
 infill function, 254
 information operator, 320
 information theory, 13
 information-based complexity, 7
 informed prior, 74
 initial value problem, *see* IVP
 integrated Brownian motion, *see* integrated Wiener process
 integrated Wiener process, 51, 299
 interpolant, 52
 inverse problems, 340
 Itô integral, 49, 50
 IVP, 286
 regularity, 287
 solver, *see* ODE solver
- Jacobi polynomials, 103
 Jacobian-matrix estimator, 344
- Kalman filter, *see* filter
 Kalman smoother, *see* smoother
 kernel, 31
 kernel herding, 76, 121
 kernel mean embedding, 76
 kernel ridge regression, 30, 37
 knot, *see* node
 knowledge vs assumptions, 95
 Kronecker product, 127
 covariance, 152
 skew-symmetric, 157
 symmetric, 157
 Krylov sequence, 144
- Laguerre polynomials, 103
 Lanczos process, 145
 Laplace approximation, 55
 learning curves, 277
 Legendre polynomials, 29, 103
 likelihood, 21
 line search, 200, 205
 probabilistic, 211
 linear differential equation, 48
 linear Gaussian system, *see* LTI
 linear regression, 27
 linear spline, 88
 linear splines, *see* splines, *see* splines
 linear time-invariant system, *see* LTI, *see* LTI
 LTI, 44, 296
 LU decomposition, 133
- machine learning, 1, 348
 MAP, *see* maximum a posteriori estimate
 marginal, 21
 marginal likelihood, 55
 Markov
 chain, 42
 process, 41
 property, 42
 Markov chain Monte Carlo, *see* MCMC
 matrix exponential, 48
 matrix inversion lemma, 129
 matrix perturbation theory, 160
 Matérn family, 52
 maximum a posteriori estimate, 109, 305
 MCMC, 70, 72, 108, 341
- mean, 23
 measurement matrix, 45
 measurement model, 44, 45
 Mercer kernel, *see* kernel
 message passing, 42, 44
 meta-criterion for Bayesian optimisation, 274
 model, 73
 modified Bessel function, 52
 momentum, 230
 multi-armed bandits, 247
 multi-point expected improvement, 275
 myopic Bayesian optimisation, 255
- natural spline, 88
 Nesterov acceleration, 230
 neural network, 251
 neuroscience, 338
 nilpotent, 48
 node, 102
 nonparametric vs Gaussian, 314, 336
 Nordsieck methods, 327
 norm
 Euclidean, 128
 Frobenius, 128
 matrix, 128
 normal distribution, *see* Gaussian distribution
 numerical algorithm, 70
 numerical stability, 131, 317, 322–324
 Nyström approximation, 138
- observation noise, 45
 ODE, 285
 bifurcation, 311
 flow, 289
 higher-order, 286, 298
 inverse problem, 340
 Lorenz, 332
 periodic, 300
 regularity, 287
 sensitivity, 332
 stiff, 322
 ODE filters and smoothers, 301, 314
 A-stability, 322
 bootstrap particle, 311
 BVPs, 340
 convergence, 317, 318

- EKF0, 318
- global, 319, 321, 329
- IEKS, 321
- local, 318
- MAP, 320
- EK0, EK1, 305
- EKF0, EKF1, 304
- EKS0, EKS1, 305
- equivalences, 326
 - Nordsieck, 327
 - trapezoidal rule, 327
- error estimation, 313
- extended Kalman, 303–305
- Gaussian, 301–303, 336
- gradient estimator, 345
- Hessian estimator, 345
- IEKF, 306
- IEKS, 306
- implementation, 315
- inverse problems, 343
- iterated extended Kalman, 306
- Jacobian-matrix estimator, 344
- Kalman, 309
- MAP estimate, 305
- multi-step, 329
- Nordsieck form, 327
- numerical stability, 317, 322
 - A-stability, 323
 - linear algebra, 324
- particle, 309–311
- re-scaled variables, 324, 328
- square-root, 325
- steady state, 329
- step-size selection, 314
- terminology, 301, 305, 309
- uncertainty calibration
 - asymptotic, 319
 - scaling, 312, 313
- ODE solver, 287, 289
 - BVP, 339
 - classical, 289, 291, 293
 - geometric, 334
 - multi-step, 290, 322, 329
 - Nordsieck, 327
 - probabilistic, 287, 293
 - single-step, 290, 322
 - trapezoidal, 327
- ordinary differential equation, *see* ODE
- Ornstein–Uhlenbeck process, 32, 52, 232, 299
- orthogonal polynomials, 102
- parametric
 - model, 27
 - regression, 27
- partial differential equation, 285, 347
- Pascal triangle, 50
- perturbative solvers, 331, 336, 347
 - additive noise, 332
 - convergence, 333, 335
 - inverse problems, 342
 - random step sizes, 334
- Picard–Lindelöf theorem, 286
- Poincaré, Henri, 5
- positive definite kernel, *see* kernel
- positive definite matrix, 159
- posterior, 21
- Powell symmetric Broyden, 236
- preconditioning, 145, 166
- prediction-error decomposition, 60
- prior, 3, 21, 298
 - conjugate, 55, 56
- ProbNum package, *see* software
- process noise, 45
- product rule, 21
- proposal distribution, 310, 311
- pseudoinverse, 171
- quadrature, 69
 - adaptive, 73
 - Bayesian, 72
 - Gaussian, 102
- quasi-Newton method, 234
- query selection function, 254
- radial basis function, *see* Gaussian kernel
- Radon–Nikodym derivative, 21
- random, 11
- random forest, 251
- Rauch–Tung–Striebel smoother, *see* smoother
- Rayleigh regression, 176
- recommendation strategy for
 - Bayesian optimisation, 255
- regression in ODEs, 289
- reinforcement learning, 247, 276
- rejection sampling, 71
- reproducing kernel Hilbert space, 36–38
- Riccati equation, 53
- Riemannian manifolds, 340
- RKHS, *see* reproducing kernel Hilbert space
- Runge–Kutta methods, 291
- scattered-data interpolation, 320
- Schur complement, 130
- Schur’s product theorem, 33
- secant method, *see* quasi-Newton method
- semi-ring (of kernels), 33
- sigma point, *see* node
- Simpson’s rule, 95
- singular value decomposition, 128
- smoother, 44, 47
 - Kalman, 47
 - ODE, *see* ODE filters and smoothers
 - particle, 311
- Sobolev space, 52, 320
- software
 - backpack, 219
 - emukit, 120, 278
 - ProbNum, 121, 194, 315
- splines
 - cubic, 50, 212
 - linear, 32, 35
 - polynomial, 50
- square exponential kernel, *see* Gaussian kernel
- square-exponential, *see* Gaussian kernel
- SR1, 236
- state vector, 296
- state-space model, 44, 297, 307, 340
- statistical inference, 13
- steady-state, 54
- steepest descent, 201
- step function, 29
- stochastic differential equation, 48
- stochastic filtering problem, 296
- stochastic integral equation, 50
- sufficient statistics, 57
- sum rule, 21
- sum-product algorithm, 43
- surrogate function, 251
- switch function, 29
- symmetric matrix, skew-symmetric matrix, 156
- symmetric positive (semi-) definite,

398 Index

128	uncertainty sampling, 83	weight space, 28
symplectic matrix, 54	uncertainty-awareness, 2, 291, 331, 342, 345	weight-space view, 28
	under-confidence, 95	white noise, 109
time series, 42	Vandermonde matrix, 102	Wiener process, 32, 50
training curves, 277	variable metric, <i>see</i> quasi-Newton method	winner’s curse, 263
transition matrix, 45	vectorisation, 127	Wishart distribution, 59, 149
trigonometric functions, 29		Wolfe conditions, <i>see</i> line search
trust-region, 234		Woodbury Identity, <i>see</i> matrix in-
type-II likelihood, <i>see</i> marginal like- lihood	weak adaptivity, 85	version lemma
		worst-case error, 37