
Index

- Absorbing state 12
- Accept-reject algorithm 166–170, 173
 - in sequential Monte Carlo 224, 261
- Acceptance probability
 - in accept-reject 169
 - in Metropolis-Hastings 171
- Acceptance ratio
 - in Metropolis-Hastings 171
 - in reversible jump MCMC 492
- Accessible set 523
- AEP *see* Asymptotic equipartition property
- Asymptotic equipartition property
 - see* Shannon-McMillan-Breiman theorem, 574
- Asymptotically tight *see* Bounded in probability
- Atom 524
- Auxiliary variable 260
 - in sequential Monte Carlo 256–264
- Averaging
 - in MCEM 407, 429
 - in SAEM 416
 - in stochastic approximation 414, 433
- Backward smoothing
 - decomposition 70
 - kernels 70–71, 125, 130
- Bahadur efficiency 565
- Balance equations
 - detailed 41
 - global 41
 - local 41
- Baum-Welch *see* Forward-backward
- Bayes
 - formula 71
 - operator 102
 - rule 64, 157
 - theorem 172
- Bayesian
 - decision procedure 472
 - estimation 360, 471
 - model 71, 472
 - network *see* Graphical model
 - posterior *see* Posterior
 - prior *see* Prior
- Bayesian information criterion 566, 569, 574
- BCJR algorithm 74
- Bearings-only tracking 24
- Bennett inequality 590
- Bernoulli-Gaussian model 197
- BIC *see* Bayesian information criterion
- Binary deconvolution model 375
 - estimation using EM 376
 - estimation using quasi-Newton 376
 - estimation using SAME 506
- Binary symmetric channel 7, 8
- Bootstrap filter 238, 254–256, 259
- Bounded in probability 334
- Bryson-Frazier *see* Smoothing
- Burn-in 399, 498
- Canonical space 38
- Capture-recapture model 12, 485
- Cauchy sequence 606

- CGLSSM *see* State-space model
- Chapman-Kolmogorov equations 36
- Coding probability 571, 574
 - mixture 573
 - normalized maximum likelihood 572
 - universal 572
- Communicating states 513
- Companion matrix 17, 30
- Computable bounds 186
- Conditional likelihood function 218
 - log-concave 225
- Contrast function 442
- Coordinate process 38
- Coupling
 - inequality 543
 - of Markov chains 543–545
 - set 544
- Critical region 570

- Darroch model 13
- Data augmentation 482
- Dirichlet distribution 476, 573
- Disturbance noise 127
- Dobrushin coefficient 96
- Doebelin condition 97
 - for hidden Markov model 561
- Drift conditions
 - for hidden Markov model 562
 - for Markov chain 537–541, 548–552
 - Foster-Lyapunov 549

- ECM *see* Expectation-maximization
- Effective sample size 235
- Efficiency 580
 - Bahadur 581
 - Pitman 580
- Efficient score test 467
- EKF *see* Kalman, extended filter
- EM *see* Expectation-maximization
- Equivalent parameters 451
- Error
 - exponent 581
 - overestimation 568
 - underestimation 568
- Exchangeable distribution 478
- Expectation-maximization 349–353
 - convergence of 389–394
 - ECM 394
 - for MAP estimation 360
 - for missing data models 359
 - in exponential family 352
 - intermediate quantity of 349
 - SAGE 395
- Exponential family 352
 - natural parameterization 473
 - of the Normal 150
- Exponential forgetting *see* Forgetting

- Filtered space 37
- Filtering 54
- Filtration 37
 - natural 38
- Fisher identity 354, 362, 458
- Forgetting 100–120
 - exponential 109, 446
 - of time-reversed chain 461
 - strong mixing condition 105, 108
 - uniform 100, 105–110
- Forward smoothing
 - decomposition 66
 - kernels 66, 101, 327
- Forward-backward 56–66
 - α *see* forward variable
 - β *see* backward variable
 - backward variable 57
 - Baum-Welch denomination 74
 - decomposition 57
 - forward variable 57
 - in finite state space HMM 123–124
 - in state-space model 154
 - scaling 62, 75

- Gaussian linear model 128, 150
- Generalized likelihood ratio test *see* Likelihood ratio test
- Gibbs sampler 180–182
 - in CGLSSM 194
 - in hidden Markov model 481–486
 - random scan 182
 - sweep of 181, 401, 484
 - systematic scan 182
- Gilbert-Elliott channel 6
- Global sampling *see* Resampling, global
- Global updating *see* Updating of hidden chain
- Gram-Schmidt orthogonalization 136
- Graphical model 1, 4

- Growth model
 - comparison of SIS kernels 230–231
 - performance of bootstrap filter 240–242
- Hahn-Jordan decomposition 91
- Harris recurrent chain *see* Markov chain, Harris recurrent
- Harris recurrent set 533
- Hidden Markov model 1–6, 42–44
 - aperiodic 559
 - discrete 43
 - ergodic 33
 - finite 6–13
 - fully dominated 43
 - hierarchical 46–47
 - in biology 10
 - in ion channel modelling 13
 - in speech recognition 14
 - left-to-right 33
 - likelihood 53
 - log-likelihood 53
 - normal *see* Normal hidden Markov model
 - partially dominated 43
 - phi-irreducible 559
 - positive 560
 - recurrent 560
 - transient 560
 - with finite state space 121–127
- Hilbert space 618
- Hitting time 513, 521
- HMM *see* Hidden Markov model
- Hoeffding inequality 292
- Homogeneous *see* Markov chain
- HPD (highest posterior density) region 240
- Hybrid MCMC algorithms 179
- Hyperparameter *see* Prior
- Hypothesis testing
 - composite 565, 567, 569, 581
 - simple 570
- Ideal codeword length 571
- Identifiability 450–457, 468, 478, 565, 568
 - in Gaussian linear state-space model 384
 - of finite mixtures 454
 - of mixtures 454–455
- Implicit conditioning convention 58
- Importance kernel *see* Instrumental kernel
- Importance sampling 173, 210–211, 287–295
 - self-normalized 211, 293–295
 - asymptotic normality 293
 - consistency 293
 - deviation bound 294
 - sequential *see* Sequential Monte Carlo
 - unnormalized 210, 287–292
 - asymptotic normality 288
 - consistency 288
 - deviation bound 292
- Importance weights 173
 - normalized 211
 - coefficient of variation of 235
 - Shannon entropy of 235
- Incremental weight 216
- Information divergence rate 574
- Information matrix 464
 - observed 442
 - convergence of 465
- Information parameterization 148–150
- Initial distribution 38
- Innovation sequence 136
- Instrumental distribution 210
- Instrumental kernel 215
 - choice of 218
 - optimal 220–224
 - local approximation of 225–231
 - prior kernel 218
- Integrated autocorrelation time 192
- Invariant measure 517, 534
 - sub-invariant measure 534
- Inversion method 242
- Irreducibility measure
 - maximal 522
 - of hidden Markov model 557
 - of Markov chain 521
- Jacobian 487, 492, 495–496
- Kalman
 - extended filter 228
 - filter 141–143
 - gain 142

- filtering with non-zero means 143
 - predictor 137–140
 - gain 138
 - unscented filter 228
- Kernel *see* Transition
- Kraft-McMillan inequality 571
- Krichevsky-Trofimov mixture 573
- Kullback-Leibler divergence 350

- Label switching 479
- Lagrange multiplier test 467
- Large deviations 584
- Latent variable model 2
- Law of iterated logarithm 571
- Level 570
 - asymptotic 570
- Likelihood 53, 359, 443–445
 - conditional 65, 66, 444
 - in state-space model 140
- Likelihood ratio test 466–468
 - generalized 467, 565, 570, 574, 584
- Linear prediction 131–137
- Local asymptotic normality 443
- Local updating *see* Updating of hidden chain
- Log-likelihood *see* Likelihood
- Log-normal distribution 487
- Louis identity 354
- Lyapunov function 421
 - differential 430

- MAP *see* Maximum *a posteriori*
- Marcinkiewicz-Zygmund inequality 292
- Markov chain
 - aperiodic 520, 542
 - canonical version 39
 - central limit theorem 555, 556
 - ergodic theorem 520, 542
 - geometrically ergodic 548
 - Harris recurrent 533
 - homogeneous 2
 - irreducible 514
 - law of large numbers 553
 - non-homogeneous 40, 163
 - null 519, 534
 - on countable space 513–520
 - on general space 520–556
 - phi-irreducible 521
 - positive 534
 - positive recurrent 519
 - recurrent 517
 - reverse 40
 - reversible 41
 - solidarity property 516
 - strongly aperiodic 542
 - transient 517
- Markov chain Monte Carlo 170–186
- Markov jump system *see* Markov-switching model
- Markov property 39
 - strong 40
- Markov-switching model 5
 - maximum likelihood estimation 469
 - smoothing 86
- Matrix inversion lemma 149, 152
- Maximum *a posteriori* 360, 473, 501–510
 - state estimation 125, 208
- Maximum likelihood estimator 360, 441
 - asymptotic normality 443, 465
 - asymptotics 442–443
 - consistency 442, 446–450, 465
 - convergence in quotient topology 450
 - efficiency 443
- Maximum marginal posterior estimator 472
 - in CGLSSM 208
- MCEM *see* Monte Carlo EM
- MCMC *see* Markov chain Monte Carlo
- MDL *see* Minimum description length
- Mean field in stochastic approximation 430
- Mean square
 - convergence 618
 - error 620
 - prediction 620
- Measurable
 - function 605
 - set 605
 - space 605
- Measure
 - positive 605
 - probability 605
- MEM algorithm *see* SAME algorithm

- Metropolis-Hastings algorithm 171
 - one-at-a-time 188
 - geometric ergodicity 549
 - independent 173
 - phi-irreducibility 523
 - random walk 176
- Minimum description length 573
- Missing information principle 465
- Mixing distribution 454
- Mixture density 454
- Mixture Kalman filter 275
- ML, MLE *see* Maximum likelihood estimator
- Model averaging 489
- Moderate deviations 568, 584
- Monte Carlo
 - estimate 162
 - integration 162
- Monte Carlo EM 398–399
 - analysis of 419–429
 - averaging in 407
 - in hidden Markov model 399
 - rate of convergence 426–429
 - simulation schedule 403–408
 - with importance sampling 402
 - with sequential Monte Carlo 402
- Monte Carlo steepest ascent 408

- Neyman-Pearson lemma 570
- NML *see* Coding probability
- Noisy AR(1) model
 - SIS with optimal kernel 222–224
 - SIS with prior kernel 218–220
- Non-deterministic process 136
- Normal hidden Markov model 13–15
 - Gibbs sampling 483
 - identifiability 456
 - likelihood ratio testing in 467
 - Metropolis-Hastings sampling 486
 - prior for 477
 - reversible jump MCMC 493
 - SAME algorithm 504
- Normalizing constant 211
 - in accept-reject 169
 - in Metropolis-Hastings 172–173
- Occupation time
 - of set 521
 - of state 514
- Optional sampling 590
- Order 565
 - estimator
 - BIC 587
 - MDL 576
 - PML 577
 - identification 565
 - Markov 566, 567, 569, 587
 - of hidden Markov model 566, 567
- Oscillation semi-norm 92
 - essential 292

- Particle filter 209, 237
- Penalized maximum likelihood 565, 568, 574
- Perfect sampling 186
- Period
 - of irreducible Markov chain 520
 - of phi-irreducible HMM 559
 - of phi-irreducible Markov chain 542
 - of state in Markov chain 520
- PML *see* Penalized maximum likelihood
- Polish space 606
- Posterior 65, 71, 360, 472
- Power 570
 - function 570
- Precision matrix 149
- Prediction 54
- Prior 65, 71, 360
 - conjugate 473
 - diffuse 148
 - Dirichlet 573
 - distribution 471
 - flat 151, 475
 - for hidden Markov model 475–478
 - hyper- 474
 - hyperparameter 473
 - improper 151, 474
 - non-informative 472, 474
 - regularization 360
 - selection 473
 - subjective 472
- Probability space 606
 - filtered 37
- Projection theorem 619
- Proper set 299
- Properly weighted sample 268

- Radon-Nikodym derivative 210
- Rao test 467
- Rao-Blackwellization 182
- Rauch-Tung-Striebel *see* Smoothing
- Rayleigh-fading channel 18
- Recurrent
 - set 524
 - state 514
- Recursive estimation 374
- Regeneration time 529
- Regret 572
- Regularization 360
- Reprojection 420
- Resampling
 - asymptotic normality 306
 - consistency 303
 - global 267
 - in SMC 236–242
 - multinomial 212–213
 - alternatives to 244–250
 - implementation of 242–244
 - optimal 267–273
 - remainder *see* residual
 - residual 245–246
 - stratified 246–248
 - systematic 248–250
 - unbiased 244, 268
- Resolvent kernel *see* Transition
- Return time 513, 521
- Reversibility 41
 - in Gibbs sampler 181
 - of Metropolis-Hastings 171
 - of reversible jump MCMC 491
- Reversible jump MCMC 488, 490
 - acceptance ratio 492
 - birth move 493
 - combine move 493–495
 - death move 493
 - merge move 493
 - split move 493–495
- Riccati equation 139
 - algebraic 141
- Robbins-Monro *see* Stochastic approximation
- RTS *see* Smoothing
- SAEM *see* Stochastic approximation
EM
- SAGE *see* Expectation-maximization
- SAME algorithm 503
 - for normal HMM 504
 - in binary deconvolution model 506
- Sample impoverishment *see* Weight degeneracy
- Sampling importance resampling
 - 211–214, 295–310
 - asymptotic normality 307
 - consistency 307
 - deviation bound 308
 - estimator 213
 - mean squared error of 214
 - unbiasedness 213
- Score function 457
 - asymptotic normality 457–464
- SEM *see* Stochastic EM
- Sensitivity equations 363–367
- Sequential Monte Carlo 209, 214–231
 - i.i.d. sampling 253, 324
 - analysis of 324–333
 - asymptotic normality 325
 - asymptotic variance 326
 - consistency 325
 - deviation bound 328, 330
 - for smoothing functionals 278–286
 - implementation in HMM 214–218
 - mutation step 311–315
 - asymptotic normality 313
 - consistency 312
 - mutation/selection 255, 316
 - analysis of 319
 - asymptotic normality 319
 - consistency 319
 - optimal kernel 322
 - prior kernel 322
 - selection/mutation 253, 255, 316
 - analysis of 320
 - asymptotic normality 320
 - consistency 320
- SISR 322
 - analysis of 321–324
 - asymptotical normality 323
 - consistency 323
 - with resampling 231–242
- Shannon-McMillan-Breiman theorem
 - 61, 568, 574, 575
- Shift operator 39
- Sieve 577
- Simulated annealing 502

- cooling schedule 502
- SIR *see* Sampling importance resampling
- SIS *see* Importance sampling
- SISR *see* Sequential Monte Carlo
- Slice sampler 183
- Small set
 - existence 527
 - of hidden Markov model 559
 - of Markov chain 526
- SMC *see* Sequential Monte Carlo
- Smoothing 51, 54
 - Bryson-Frazier 143
 - disturbance 143–147
 - fixed-interval 51, 59–76
 - fixed-point 78–79
 - forward-backward 59
 - functional 278
 - in CGLSSM 157–159
 - in hierarchical HMM 87–89
 - in Markov-switching model 86
 - Rauch-Tung-Striebel 66, 131
 - recursive 79–85
 - smoothing functional 80
 - two-filter formula 76, 148–155
 - with Markovian decomposition
 - backward 70, 124, 131
 - forward 66
- Source coding 565
- Splitting construction 529–530
 - split chain 529
- Stability in stochastic algorithms 420
- State space 38
- State-space model 3
 - conditionally Gaussian linear 17–23, 46, 194–208, 274–278
 - Gaussian linear 15–17, 127–155
- Stationary distribution
 - of hidden Markov model 560
 - of Markov chain 517
- Stein’s lemma 581, 584
- Stochastic approximation 411
 - analysis of 429–433
 - gradient algorithm 412
 - rate of convergence 432–433
 - Robbins-Monro form 412
- Stochastic approximation EM 414
 - convergence of 433–435
- Stochastic EM 416
- Stochastic process 37
 - adapted 38
 - stationary 41
- Stochastic volatility model 25–29
 - approximation of optimal kernel 227–228
 - EM algorithm 399
 - identifiability 456
 - one-at-a-time sampling 188–194
 - performance of SISR 239–240
 - single site sampling 183–185
 - smoothing with SMC 281
 - weight degeneracy 234–236
- Stopping time 39
- Strong mixing condition 105, 108
- Subspace methods 384
- Sufficient statistic 352
- Sweep *see* Gibbs sampler
- Tangent filter 366
- Target distribution 171
- Tight *see* Bounded in probability
- Total variation distance 91, 93
 - V-total variation 544
- Transient
 - set (uniformly) 524
 - state 514
- Transition
 - density function 35
 - kernel 35
 - Markov 35
 - resolvent 522
 - reverse 37
 - unnormalized 35
 - matrix 35
- Triangular array 297
 - central limit theorems 338–342
 - conditionally independent 298
 - conditionally i.i.d. 298
 - laws of large numbers 333–338
- Two-filter formula *see* Smoothing
- UKF *see* Kalman, unscented filter
- Uniform spacings 243
- Universal coding 565, 567, 571
- Updating of hidden chain
 - global 481
 - local 482

- V-total variation distance *see* Total variation distance
- Variable dimension model 488
- Viterbi algorithm 125
- Wald test 467
- Weight degeneracy 209, 231–236
- Weighted sample 298
 - asymptotic normality 299, 304
 - consistency 298, 301
- Weighting and resampling algorithm 301
- Well-log data model 21–22
 - with Gibbs sampler 204
 - with mixture Kalman filter 276