英文索引

※掲載ページは各項目の末尾に示し，ローマン体が上巻，イタリック体が下巻を表すものとする．

A

acceptance criterion（受理規準） ***252***, *255*, *259*
activation function（活性化関数） **178**, 212, 227
active constraint（有効制約） **327**, *38*
AdaBoost *374*, ***375***
adaline 194
adaptive rejection sampling（適応的棄却サンプリング） *244*
ADF ➜ assumed density filtering
AIC ➜ Akaike information criterion
Akaike information criterion（赤池情報量規準） **33**, 216
α divergence（α ダイバージェンス） *183*
α recursion（α 再帰） *338*
ancestral sampling（伝承サンプリング） ***77***, *238*, *330*
annular flow（環状流） 297
AR model（AR モデル）
  ➜ autoregressive model
arc（弧） *71*
ARD ➜ automatic relevance determination
ARMA ➜ autoregressive moving average
assumed density filtering（仮定密度フィルタリング） *224*
autoassociative mapping（自己連想写像）*311*
autoassociative network（自己連想ネットワーク） *310*
automatic relevance determination（関連度自動決定） 261, *23*, ***60***, *199*, *299*
autoregressive hidden Markov model（自己回帰隠れマルコフモデル） *350*
autoregressive model（自己回帰モデル） *327*
autoregressive moving average（自己回帰移動平均） *15*

B

back-tracking（バックトラック） ***129***, *348*
backgammon（バックギャモン） 3
backpropagation（逆伝播） 242
bagging（バギング） *373*
basis function（基底関数） **136**, 171, 202, 226
batch training（バッチ訓練） 241
Baum–Welch algorithm（Baum–Welch アルゴリズム） *336*
Bayes factor（ベイズ因子） 160
Bayes' theorem（ベイズの定理） 14
Bayes, Thomas 20
Bayesian analysis（ベイズ解析） v, 9, **20**
  hierarchical *84*
  model averaging *372*
Bayesian information criterion（ベイズ情報量規準） 33, **215**
Bayesian model comparison（ベイズモデル比較） **160**, *187*, *197*
Bayesian network（ベイジアンネットワーク） *71*
Bayesian probability（ベイズ確率） 20
belief propagation（確率伝播） *117*

Bernoulli distribution（ベルヌーイ分布） **66**, 111, 303
  mixture model *160*
Bernoulli, Jacob 67
beta distribution（ベータ分布） **68**, 304
β recursion（β 再帰） *339*
between-class covariance（クラス間共分散） 186
bias（バイアス） **27**, 147
bias parameter（バイアスパラメータ） **136**, 179, 227, *56*
bias-variance trade-off（バイアス–バリアンスのトレードオフ） 146
BIC ➜ Bayesian information criterion
binary entropy function（二値エントロピー関数） *210*
binomial distribution（二項分布） **68**, 304
biological sequence（生物学的配列） *328*
bipartite graph（二部グラフ） *114*
bit（ビット） 49
blind source separation（未知音源分離） *309*
blocked path（遮断された経路） *86*, ***91***, *97*
Boltzmann distribution（ボルツマン分布） *100*
Boltzmann, Ludwig Eduard 52
Boolean logic（ブール論理） 21
boosting（ブースティング） *374*
bootstrap（ブートストラップ） **22**, *373*
bootstrap filter（ブートストラップフィルタ） *365*
box constraint（矩形制約） ***43***, *52*
Box–Muller method（Box–Muller 法） *241*

C

C4.5 *381*
calculus of variations（変分法） 321, *176*
canonical correlation analysis（正準相関分析） *282*
canonical link function（正準連結関数） 211
CART ➜ classification and regression tree
Cauchy distribution（コーシー分布） 310, ***240***, *243*
causality（因果） *78*
CCA ➜ canonical correlation analysis
central difference（中心差分） 248
central limit theorem（中心極限定理） 76
chain graph（連鎖グラフ） *107*
chaining（連鎖） *270*
Chapman–Kolmogorov equations（チャップマン–コルモゴロフの方程式） *111*
child node（子ノード） *72*
Cholesky decomposition（コレスキー分解） *242*
chunking（チャンキング） *45*
circular normal distribution（循環正規分布） ➜ von Mises distribution
classical probability（古典的確率） 20
classification（クラス分類） 2
classification and regression tree *381*
clique（クリーク） *98*
clustering（クラスタリング） 3
clutter problem（雑音データ問題） *225*
co-parent（共同親） ***95***, *206*
code-book vector（符号表ベクトル） *145*
combining models（モデルの結合） 44, ***371***
committee（コミッティ） *373*
complete data set（完全データ集合） *155*
completing the square（平方完成） 84
computational learning theory（計算論的学習理論） *36*, *54*
concave function（凹関数） 55
concentration parameter（集中度パラメータ） **105**, 312
condensation algorithm（コンデンセーションアルゴリズム） *365*
conditional entropy（条件付きエントロピー） 54
conditional expectation（条件付き期待値） 19
conditional independence（条件付き独立） 45, ***84***, *96*
conditional mixture model（条件付き混合モデル） ➜ mixture model
conditional probability（条件付き確率） 13
conjugate prior（共役事前分布） 65, 95, **114**, *205*
convex duality（凸双対性） *208*
convex function（凸関数） **55**, *207*
convolutional neural network（たたみ込みニ

ューラルネットワーク） 270
correlation coefficient（相関係数） *283*
cost function（コスト関数） 40
covariance（共分散） 20
  between-class 186
  within-class 186
covariance matrix（共分散行列）
  diagonal 81
  isotropic 81
  partitioned **83**, *19*
  positive definite *19*
covariant（共変的） *321*
Cox's axioms（コックスの公理） 21
credit assignment（信頼度割り当て） 3
cross-entropy error function（交差エントロピー誤差関数）**205**, 209, 235, *349*, *383*
cross-validation（交差確認） **32**, 160
cumulative distribution function（累積分布関数） 18
curse of dimensionality（次元の呪い）**33**, 36
curve fitting（曲線フィッティング） 4

D

D map（D マップ） ➜ dependency map
d-separation（有向分離） *85*, ***90***, *158*
DAG ➜ directed acyclic graph
DAGSVM *49*
data augmentation algorithm（データ拡大アルゴリズム） *251*
data compression（データ圧縮） *145*
decision boundary（決定境界） **38**, 177
decision region（決定領域） **38**, 177
decision surface（決定面）
  ➜ decision boundary
decision theory（決定理論） 37
decision tree（決定木） *371*, ***380***, *391*
decomposition method（分解法） *45*
degree of freedom（自由度） *275*
degrees-of-freedom parameter（自由度パラメータ） **99**, 312
density estimation（密度推定） 3, **65**
density network（密度ネットワーク） *316*
dependency map（依存性マップ） *106*
descendant node（子孫ノード） *88*
design matrix（計画行列） **139**, *57*
detailed balance（詳細釣り合い条件） *254*
differential entropy（微分エントロピー） 52
digamma function（ディガンマ関数） 306
directed acyclic graph（有向非循環グラフ）*74*
directed cycle（有向閉路） *74*
directed factorization（有向分解） *94*
Dirichlet distribution（ディリクレ分布） **74**, 305
Dirichlet, Lejeune 74
discriminant function（識別関数） 42, 177, **179**
discriminative model（識別モデル） **42**, 202
distortion measure（歪み尺度） *140*
distributive law of multiplication（積の分配則） *110*
DNA *328*
document retrieval（文書検索） *9*
dual representation（双対表現） ***2***, *38*
dual-energy gamma densitometry（2-エネルギーガンマ濃度測定法） 296
dynamic programming（動的計画法） *126*
dynamical system（力学系） *263*

E

E step（E ステップ） ➜ expectation step
early stopping（早期終了） 261
ECM ➜ expectation conditional maximization
edge（辺） *71*
effective number of observations（有効観測数） **70**, 98
effective number of parameters（有効パラメータ数） 9, **169**, 284
elliptical K-means algorithm（楕円 K-means アルゴリズム） *160*
EM ➜ expectation maximization algorithm
emission probability（出力確率） *329*
empirical Bayes（経験ベイズ）
  ➜ evidence approximation
energy function（エネルギー関数） *100*
entropy（エントロピー） 49
  conditional 54
  differential 52
  relative 54

EP ➜ expectation propagation method
ϵ-insensitive error function（ϵ 許容誤差関数） *50*
ϵ-tube（ϵ チューブ） *51*
equality constraint（等式制約） 327
equivalent kernel（等価カーネル） **157**, *12*
erf function（erf 関数） 211
ergodicity（エルゴード性） *255*
error backpropagation（誤差逆伝播） ➜ backpropagation
error function（誤差関数） **5**, 22
error-correcting output code（誤り訂正出力符号） *49*
Euler, Leonhard *178*
Euler–Lagrange equation（オイラー–ラグランジュ方程式） 323
evidence approximation（エビデンス近似） **164**, *57*, *299*
evidence function（エビデンス関数） 160
expectation（期待値） 19
expectation conditional maximization（ECM 法） *170*
expectation maximization algorithm（EM アルゴリズム） 110, ***139***, *156*
  Gaussian mixture *151*
  generalized *170*
  sampling method *250*
expectation propagation method（EP 法） *27*, *181*, ***219***
expectation step（E ステップ） *152*
explaining away（弁明） *90*
exploitation（知識利用） 3
exploration（探査） 3
exponential distribution（指数分布） 306, ***240***
exponential family（指数型分布族） 66, **110**, 201, *204*
extensive variable（外延的変数） *205*

F

face detection（顔検出） 2
face tracking（顔追跡） *66*
factor analysis（因子分析） *302*
  mixture model *313*
factor graph（因子グラフ） *72*, ***113***, *343*
factor loading（因子負荷） *302*
factorial hidden Markov model（階乗隠れマルコフモデル） *352*
factorized distribution（分解された分布）***177***, *189*
feature extraction（特徴抽出） 2
feature map（特徴マップ） 270
feature space（特徴空間） ***1***, *304*
Fisher information matrix（フィッシャー情報量行列） *9*
Fisher kernel（フィッシャーカーネル） *8*
Fisher's linear discriminant（フィッシャーの線形判別） 185
flooding schedule（フラッディングスケジュール） *133*
forward kinematics（順運動学） 275
forward problem（順問題） 274
forward propagation（順伝播） **228**, 244
forward-backward algorithm（フォワード–バックワードアルゴリズム） *336*
fractional belief propagation（部分的確率伝播） *231*
frequentist probability（頻度主義的確率） 20
fuel system（燃料装置） *89*
function interpolation（関数補間） *10*
functional（汎関数） **321**, *176*
  derivative *176*

G

gamma densitometry（ガンマ濃度測定法）296
gamma distribution（ガンマ分布） **306**, *243*
gamma function（ガンマ関数） 69
gating function（ゲート関数） *391*
Gauss, Carl Friedrich 77
Gaussian distribution（ガウス分布） 24, **76**, 307
  conditional **82**, 90
  marginal **85**, 90
  maximum likelihood 91
  mixture 107, 272, 276, ***146***
  sequential estimation 92
  sufficient statistic 91
  wrapped 107

Gaussian kernel（ガウスカーネル） *6*
Gaussian process（ガウス過程） 159, ***14***
Gaussian random field（ガウス確率場） *16*
Gaussian-gamma distribution（ガウス–ガンマ分布） **99**, 308
Gaussian-Wishart distribution（ガウス–ウィシャート分布） **100**, 309, *188*, *192*
GEM ➜ expectation maximization, generalized
generalization（汎化） 2
generalized linear model（一般化線形モデル） **178**, 212
generalized maximum likelihood（一般化最尤推定） ➜ evidence approximation
generative model（生成モデル） **42**, 195, *7*, *77*, *288*, *349*
generative topographic mapping（GTM） *316*
  directional curvature *318*
  magnification factor *318*
geodesic distance（測地線距離） *315*
Gibbs sampling（ギブスサンプリング） *257*
  blocking *261*
Gibbs, Josiah Willard *258*
Gini index（ジニ係数） *383*
global minimum（大域的最小点） 238
gradient descent method（勾配降下法） 241
Gram matrix（グラム行列） *3*
graph-cut algorithm（グラフカットアルゴリズム） *103*
graphical model（グラフィカルモデル） *71*
  bipartite *114*
  directed *71*
  factorization *74*, *98*
  fully connected *73*
  inference *107*
  tree *112*
  treewidth *132*
  triangulated *131*
  undirected *71*
Green's function（グリーン関数） *10*
GTM ➜ generative topographic mapping

H

Hamilton, William Rowan *264*
Hamiltonian dynamics（ハミルトン力学） *263*
Hamiltonian function（ハミルトン関数） *265*
Hammersley–Clifford theorem（Hammersley–Clifford の定理） *100*
handwriting recognition（手書き文字認識） 1, *328*, *331*
handwritten digit（手書き数字） **295**, *282*, *331*
head-to-head path（head-to-head 経路） *88*
head-to-tail path（head-to-tail 経路） *87*
Heaviside step function（Heaviside ステップ関数） 205
Hellinger distance（ヘリンガー距離） *184*
Hessian matrix（ヘッセ行列） 166, 215, 216, 238, **250**
  diagonal approximation 251
  exact evaluation 255
  fast multiplication 256
  finite difference 254
  inverse 253
  outer product approximation 252
heteroscedastic（異分散） **276**, *23*
hidden Markov model（隠れマルコフモデル） *7*, ***328***
  autoregressive *350*
  factorial *352*
  forward-backward algorithm *336*
  input-output *351*
  left-to-right *331*
  maximum likelihood *333*
  scaling factor *345*
  sum-product algorithm *343*
  switching *363*
  variational inference *343*
hidden unit（隠れユニット） 227
hidden variable（隠れ変数） 82, ***76***, *146*, *275*
hierarchical Bayesian model（階層ベイズモデル） *84*
hierarchical mixture of experts（階層的混合エキスパートモデル） *391*
hinge error function（ヒンジ形誤差関数） *47*
Hinton diagram（ヒントン図） *300*
histogram density estimation method（ヒスト

グラム密度推定法） 118
HME ➜ hierarchical mixture of experts
hold-out set（ホールドアウト集合） 11
homogeneous flow（一様流） 297
homogeneous kernel（均一カーネル） *2*
homogeneous Markov chain（均一マルコフ連鎖） ***254***, *326*
Hooke's law（フックの法則） *297*
hybrid Monte Carlo（ハイブリッドモンテカルロ） *263*
hyperparameter（超パラメータ） **69**, 283, *22*, *57*, *84*, *216*
hyperprior（超事前分布） *84*

I

I map（I マップ） ➜ independence map
ICA ➜ independent component analysis
ICM ➜ iterated conditional modes
ID3 *381*
identifiability（識別不可能性） *150*
i.i.d. ➜ independent identically distributed
image de-noising（画像のノイズ除去） *100*
importance sampling（重点サンプリング） *238*, ***246***
importance weight（重要度重み） *247*
improper prior（変則事前分布） **115**, 260, *186*
imputation step（I ステップ） *251*
imputation-posterior algorithm（IP アルゴリズム） *251*
inactive constraint（無効制約） **327**, *38*
incomplete data set（不完全データ集合） *155*
independence map（独立性マップ） *106*
independent component analysis（独立成分分析） *309*
independent factor analysis（独立因子分析） *310*
independent identically distributed（独立同分布） **26**, *92*, *323*
independent variable（独立変数） 17
induced factorization（導出された分解） *199*
inequality constraint（不等式制約） 327
inference（推論） 37, **42**
information criterion（情報量規準） 33
information geometry（情報幾何） *9*
information theory（情報理論） 47
input-output hidden Markov model（input-output 隠れマルコフモデル） *351*
intensive variable（内包的変数） *205*
intrinsic dimensionality（実効次元） *275*
invariance（不変性） 264
inverse gamma distribution（逆ガンマ分布） 98
inverse kinematics（逆運動学） 275
inverse problem（逆問題） 274
inverse Wishart distribution（逆ウィシャート分布） 100
IP algorithm（IP アルゴリズム） ➜ imputation-posterior algorithm
IRLS ➜ iterative reweighted least squares method
Ising model（イジングモデル） *103*
isomap ➜ isometric feature map
isometric feature map（等長特徴写像） *315*
iterated conditional modes（反復条件付きモード） ***103***, *131*
iterative reweighted least squares method（反復再重み付け最小二乗法） **206**, 209, *28*, *65*, *389*

J

Jacobian matrix（ヤコビ行列） **248**, 266
Jensen's inequality（イェンセンの不等式） 56
joint probability（同時確率） 12
junction tree algorithm（ジャンクションツリーアルゴリズム） *105*, ***131***

K

K nearest neighbor（K 近傍法） 122
K-means algorithm（K-means アルゴリズム） ***140***, *159*
K-medoids algorithm（K-medoids アルゴリズム） *143*
Kalman filter（カルマンフィルタ） *15*, ***355***
  extended *363*
Kalman gain matrix（カルマン利得行列） *357*
Kalman smoother（カルマンスムーザ） *355*
Karhunen–Loève transform（Karhunen–Loève 変換） *277*

Karush–Kuhn–Tucker condition（Karush–Kuhn–Tucker 条件） **328**, *39*, *43*, *52*
kernel（カーネル） 121, ***1***, *4*
  Fisher *8*
  Gaussian *6*
  homogeneous *2*
  nonvectorial input *7*
  stationary *2*
kernel density estimator（カーネル密度推定法） **119**, *37*
kernel PCA（カーネル主成分分析） *304*
kernel regression（カーネル回帰） *11*, ***13***
kernel substitution（カーネル置換） *2*
kernel trick（カーネルトリック） *2*
kinetic energy（運動エネルギー） *264*
KKT ➜ Karush–Kuhn–Tucker condition
KL divergence（KL ダイバージェンス） ➜ Kullback–Leibler divergence
kriging（クリギング） ➜ Gaussian process
Kullback–Leibler divergence（カルバック–ライブラーダイバージェンス） **55**, *166*, *182*, *219*

L

Lagrange multiplier（ラグランジュ乗数） 325
Lagrange, Joseph-Louis *39*
Lagrangian（ラグランジュ関数） **326**, *38*, *42*, *51*
laminar flow（層流） 296
Laplace approximation（ラプラス近似） **213**, 217, 282, *27*, *65*
Laplace, Pierre-Simon 24
large margin（大きなマージン） ➜ margin
lasso 143
latent class analysis（潜在クラス分析） *160*
latent trait model（潜在特性モデル） *315*
latent variable（潜在変数） 82, ***76***, *146*, *275*
lattice diagram（格子図） ***129***, *329*, *338*, *347*
LDS ➜ linear dynamical system
leapfrog discretization（リープフロッグ離散化） *266*
learning（学習） 2
learning rate parameter（学習率パラメータ） 241
least-mean-squares algorithm（最小平均二乗アルゴリズム） 142
leave-one-out method（LOO 法） 32
likelihood function（尤度関数） 22
likelihood weighted sampling（尤度重み付きサンプリング） *248*
linear discriminant（線形識別） 179
  Fisher 185
linear dynamical system（線形動的システム） 82, ***353***
  inference *356*
linear independence（線形独立） 314
linear regression（線形回帰） 136
  EM algorithm *164*
  variational *200*
  mixture model *384*
linear smoother（線形平滑器） 157
linear-Gaussian model（線形ガウスモデル） 85, ***82***
linearly separable（線形分離可能） 177
link（リンク） *71*
link function（連結関数） 178, 212
Liouville's Theorem（リューヴィルの定理） *265*
LLE ➜ locally linear embedding
LMS algorithm（LMS アルゴリズム） ➜ least-mean-squares algorithm
local minimum（局所的極小点） 238
local receptive field（局所的受容野） 270
locally linear embedding（局所線形埋め込み） *315*
location parameter（位置パラメータ） 116
log odds（対数オッズ） 196
logic sampling（ロジックサンプリング） *238*
logistic regression（ロジスティック回帰） **204**, *47*
  Bayesian 217, *212*
  mixture model *387*
  multiclass 208
logistic sigmoid function（ロジスティックシグモイド関数） 111, 137, **195**, 204, 219, 227, *210*
logit function（ロジット関数） 196

loopy belief propagation（ループあり確率伝播） *132*
loss function（損失関数） 40
loss matrix（損失行列） 40
lossless data compression（無歪みデータ圧縮） *145*
lossy data compression（歪みのあるデータ圧縮） *145*
lower bound（下限） *198*

M

M step（M ステップ） ➜ maximization step
machine learning（機械学習） v
macrostate（マクロ状態） 50
Mahalanobis distance（マハラノビス距離） 78
manifold（多様体） **37**, 299, *308*, *313*
MAP ➜ maximum posterior
margin（マージン） *35*, ***36***, *216*
  error *44*
  soft *42*
marginal likelihood（周辺尤度） **160**, 164
marginal probability（周辺確率） 13
Markov blanket（マルコフブランケット） ***95***, *97*, *260*
Markov boundary（マルコフ境界） ➜ Markov blanket
Markov chain（マルコフ連鎖） *111*, ***253***
  first order *325*
  homogeneous ***254***, *326*
  second order *326*
Markov chain Monte Carlo（マルコフ連鎖モンテカルロ） *252*
Markov model（マルコフモデル） *324*
  homogeneous *330*
Markov network（マルコフネットワーク） ➜ Markov random field
Markov random field（マルコフ確率場） 82, *71*, ***96***
max-sum algorithm（max-sum アルゴリズム） ***126***, *347*
maximal clique（極大クリーク） *98*
maximal spanning tree（極大全域木） *131*
maximization step（M ステップ） *152*
maximum likelihood（最尤推定） 8, **22**, 26, 113
  Gaussian mixture *149*
  singularity *194*
  type 2 ➜ evidence approximation
maximum margin（最大マージン） ➜ margin
maximum posterior（最大事後確率推定） **30**, *157*
MCMC ➜ Markov chain Monte Carlo
MDN ➜ mixture density network
MDS ➜ multidimensional scaling
mean（平均） 24
mean field approximation（平均場近似） *178*
mean value theorem（平均値の定理） 51
measure theory（測度論） 18
memory-based method（メモリベース法） *1*
message passing（メッセージパッシング） *110*
  pending message *133*
  schedule *133*
  variational *206*
Metropolis algorithm（Metropolis アルゴリズム） *252*
Metropolis–Hastings algorithm（Metropolis–Hastings アルゴリズム） *255*
microstate（ミクロ状態） 50
minimum risk（リスク最小化） 44
Minkowski loss（ミンコフスキー損失） 47
missing at random（ランダム欠損） ***157***, *296*
missing data（欠損データ） *296*
mixing coefficient（混合係数） 109
mixture component（混合要素） 108
mixture density network（混合密度ネットワーク） **274**, *391*
mixture distribution（混合分布） ➜ mixture model
mixture model（混合モデル） 161, ***139***
  conditional 276, ***384***
  linear regression *384*
  logistic regression *387*
  symmetry *197*
mixture of experts（混合エキスパートモデル） *390*

mixture of Gaussians（混合ガウス分布） 107, 272, 276, ***146***
MLP ➜ multilayer perceptron
MNIST data（MNIST データ） 295
model comparison（モデル比較） 5, 31, **160**, *187*, *197*
model evidence（モデルエビデンス） 160
model selection（モデル選択） 161
moment matching（モーメント一致法） ***220***, *224*
momentum variable（運動量変数） *264*
Monte Carlo EM algorithm（モンテカルロ EM アルゴリズム） *250*
Monte Carlo sampling（モンテカルロサンプリング） 23, ***237***
Moore–Penrose pseudo-inverse matrix（ムーア–ペンローズの擬似逆行列） ➜ pseudo-inverse matrix
moralization（モラル化） ***105***, *115*
MRF ➜ Markov random field
multidimensional scaling（多次元尺度構成法） *315*
multilayer perceptron（多層パーセプトロン） 225, **229**
multimodality（多峰性） 275
multinomial distribution（多項分布） 74, 111, **309**
multiplicity（多重度） 50
mutual information（相互情報量） 54, **57**

N

Nadaraya–Watson（Nadaraya–Watson モデル） ➜ kernel regression
naive Bayes model（ナイーブベイズモデル） 45, ***93***
nat（ナット） 50
natural language modelling（自然言語のモデル化） *328*
natural parameter（自然パラメータ） 110
nearest-neighbor method（最近傍法） 122
neural network（ニューラルネットワーク） 225
  convolutional 270
  regularization 258
  relation to Gaussian process *31*
Newton–Raphson method（ニュートン–ラフソン法） **206**, *29*
node（ノード） *71*
noiseless coding theorem（ノイズなし符号化定理） 49
nonidentifiability（識別不可能性） *303*
noninformative prior（無情報事前分布） 23, **115**
nonmetric MDS（非計量多次元尺度構成法） *315*
nonparametric method（ノンパラメトリック法） 66, **117**
normal distribution（正規分布） ➜ Gaussian distribution
normal equation（正規方程式） 139
normal-gamma distribution（正規–ガンマ分布） **99**, 310
normal-Wishart distribution（正規–ウィシャート分布） **100**, 310, *188*, *192*
normalized exponential（正規化指数関数） ➜ softmax function
novelty detection（新規性検出） 43
ν-SVM *44*

O

object recognition（物体認識） *77*
observed variable（観測変数） *75*
Occam factor（Occam 係数） 216
oil flow data（送油データ） 33, **296**, *276*, *284*
Old Faithful data（Old Faithful 間欠泉データ） 107, **299** *193*, *198*
on-line learning（オンライン学習） ➜ sequential learning
1-of-K coding scheme（一対 K 符号化法） *140*
one-versus-one classifier（1 対 1 分類器） **180**, *49*
one-versus-the-rest classifier（1 対他分類器） **180**, *48*
ordered over-relaxation（順序付き過剰緩和） *260*
Ornstein–Uhlenbeck process（オルンシュタイン–ウーレンベック過程） *17*
orthogonal least squares（直交最小二乗法） *12*
outlier（外れ値） 43, **101**, 183, 211

output unit activation（出力ユニット活性）227
over-fitting（過学習） **6**, 146, *150*, *177*
over-relaxation（過剰緩和） *260*

P

PAC 学習（PAC learning） ➜ probably approximately correct learning
PAC-Bayesian framework（PAC–ベイズ理論） *55*
parameter shrinkage（パラメータ縮小推定） 142
parent node（親ノード） *72*
particle filter（粒子フィルタ） *364*
partition function（分配関数） ***99***, *269*
Parzen estimator（Parzen 推定法） ➜ kernel density estimator
Parzen window（Parzen 窓） 121
pattern recognition（パターン認識） v
PCA ➜ principal component analysis
pending message（保留メッセージ） *133*
perceptron（パーセプトロン） 190
  convergence theorem 192
  hardware 194
perceptron criterion（パーセプトロン規準） 191
perfect map（完全マップ） *106*
periodic variable（周期変数） 102
phase space（位相空間） *264*
photon noise（光子ノイズ） 298
plate（プレート） *75*
polynomial curve fitting（多項式曲線フィッティング） 4, *74*
polytree（多重木） *113*
position variable（位置変数） *264*
positive definite covariance（正定値共分散） 79
positive definite matrix（正定値行列） 319
positive semidefinite covariance（半正定値共分散） 79
positive semidefinite matrix（半正定値行列） 319
posterior probability（事後確率） 16
posterior step（P ステップ） *251*
potential energy（ポテンシャルエネルギー） *264*
potential function（ポテンシャル関数） *99*
Power EP method（Power EP 法） *231*
power method（べき乗法） *279*
pre-image（原像） *308*
precision matrix（精度行列） 83
precision parameter（精度パラメータ） 24
predictive distribution（予測分布） **29**, 155
preprocessing（前処理） 2
principal component analysis（主成分分析） ***277***, *288*, *311*
  Bayesian *297*
  EM algorithm *294*
  Gibbs sampling *300*
  mixture distribution *313*
  physical analogy *297*
principal curve（主成分曲線） *314*
principal subspace（主部分空間） *277*
principal surface（主成分曲面） *314*
prior（事前分布）
  conjugate 65, 95, **114**, *205*
  consistent 259
  improper **115**, 261, *186*
  noninformative 23, **115**
prior probability（事前確率） 16
probabilistic graphical model（確率的グラフィカルモデル） ➜ graphical model
probabilistic PCA（確率的主成分分析） *287*
probability（確率） 11
  Bayesian 20
  classical 20
  density 17
  frequentist 20
  mass function 18
  prior 44
  product rule 12, **14**, *71*
  sum rule 12, **14**, *71*
  theory 11
probably approximately correct learning（PAC 学習） *54*
probit function（プロビット関数） **210**, 219
probit regression（プロビット回帰） 209
product rule of probability（確率の乗法定理） 12, **14**, *71*
proposal distribution（提案分布） ***242***, *246*, *252*

protected conjugate gradient method（保護共役勾配法） *45*
protein sequence（タンパク質系列） *328*
pseudo-inverse matrix（擬似逆行列） **140**, 183
pseudo-random number（擬似乱数） *239*

Q

quadratic discriminant（2 次判別関数） 198
quality parameter（品質パラメータ） *62*

R

radial basis function *2*, ***10***
Rauch–Tung–Striebel equations（Rauch–Tung–Striebel 方程式） *355*
regression（回帰） 3
regression function（回帰関数） **46**, 93
regularization（正則化） 9
  Tikhonov 269
regularized least squares method（正則化最小二乗法） 142
reinforcement learning（強化学習） 3
reject option（棄却オプション） **41**, 44
rejection sampling（棄却サンプリング） *242*
relative entropy（相対エントロピー） 54, 55
relevance vector（関連ベクトル） *59*
relevance vector machine（関連ベクトルマシン） 160, ***56***
responsibility（負担率） 110, ***148***, *190*
ridge regression（リッジ回帰） 9
RMS error（RMS 誤差） ➜ root-mean-square error
Robbins–Monro algorithm（Robbins–Monro アルゴリズム） 92
robot arm（ロボットアーム） 274
robustness（頑健性） **101**, 183
root node（根ノード） *112*
root-mean-square error（平均二乗平方根誤差） 7
Rosenblatt, Frank 191
rotation invariance（回転不変性） ***289***, *303*
RTS equation（RTS 方程式） ➜ Rauch–Tung–Striebel equation
running intersection property（連結横断特性） *132*
RVM ➜ relevance vector machine

S

sample mean（サンプル平均） 27
sample variance（サンプル分散） 27
sampling-importance-resampling（SIR） *249*
scale invariance（尺度不変性） 116, **264**
scale parameter（尺度パラメータ） 116
scaling factor（スケーリング係数） *345*
Schur complement matrix（シューア補行列） 85
Schwarz criterion（Schwarz 規準） ➜ Bayesian information criterion
self-organizing map（自己組織化マップ） *317*
sequential data（系列データ） *323*
sequential estimation（逐次推定） 92
sequential gradient descent（逐次的勾配降下法） 141, 241
sequential learning（逐次学習） **70**, 141
sequential minimal optimization（逐次最小問題最適化法） *45*
serial message passing schedule（直列メッセージパッシングスケジュール） *133*
Shannon, Claude 54
shared parameter（共有パラメータ） *80*
shrinkage（縮小推定） 9
sigmoid function（シグモイド関数） ➜ logistic sigmoid function
simplex（単体） 75
single-class support vector machine（単一クラス SVM） *50*
singular value decomposition（特異値分解） 141
sinusoidal data（三角関数データ） 300
SIR ➜ sampling-importance-resampling
skip-layer connection（層を飛び越えた結合） 229
slack variable（スラック変数） *41*
slice sampling（スライスサンプリング） *261*
SMO ➜ sequential minimal optimization
smoother matrix（平滑化行列） 157
smoothing parameter（平滑化パラメータ） 119
soft margin（ソフトマージン） *42*

soft weight sharing（ソフト重み共有） 272
softmax function（ソフトマックス関数）112, **196**, 236, 277, *67*, *211*
SOM ➜ self-organizing map
sparsity（疎性） 143, *57*, ***60***, *299*
sparsity parameter（疎性パラメータ） *62*
spectrogram（スペクトログラム） *323*
speech recognition（音声認識） ***323***, *328*
sphereing（球状化） *284*
spline function（スプライン関数） 136
standard deviation（標準偏差） 24
standardizing（標準化） *142*, ***283***
state space model（状態空間モデル） *327*
  switching *363*
stationary kernel（不変カーネル） *2*
statistical bias（統計的バイアス） ➜ bias
statistical independence（統計的独立性） ➜ independent variable
statistical learning theory（統計的学習理論） ➜ computational learning theory
steepest descent method（最急降下法） 241
Stirling's approximation（スターリングの近似式） 50
stochastic（確率的） 4
stochastic EM（確率的 EM） *251*
stochastic gradient descent（確率的勾配降下法） 141, 241
stochastic process（確率過程） *16*
stratified flow（層状流） 296
Student's t-distribution（スチューデントの t 分布） **100**, 310, *197*
subsampling（部分サンプリング） 270
sufficient statistic（十分統計量） 67, 73, **113**
sum rule of probability（確率の加法定理）12, **14**, *71*
sum-of-squares error（二乗和誤差） **5**, 29, 182, 233, *380*
sum-product algorithm（積和アルゴリズム） *112*, ***116***
  for hidden Markov model *343*
supervised learning（教師あり学習） 2
support vector（サポートベクトル） *40*
support vector machine（サポートベクトルマシン） 225
  for regression *50*
  multiclass *48*
survival of the fittest *365*
SVD ➜ singular value decomposition
SVM ➜ support vector machine
switching hidden Markov model（スイッチング隠れマルコフモデル） *363*
switching state space model（スイッチング状態空間モデル） *363*
synthetic data set（人工データ集合） 300

T

tail-to-tail path（tail-to-tail 経路） *86*
tangent distance（接距離） 268
tangent propagation（接線伝播法） 264, **265**
tapped delay line（タップ付き遅延線） *327*
target vector（目標ベクトル） 2
test set（テスト集合） 2, **32**
threshold parameter（しきい値パラメータ） 179
tied parameter（結合されたパラメータ） *80*
Tikhonov regularization（ティホノフ正則化） 269
time warping（時間軸伸縮） *333*
tomography（断層撮影） 298
training（訓練） 2
training set（訓練集合） 1
transition probability（遷移確率） ***254***, *328*
translation invariance（平行移動不変性）116, **264**
tree-reweighted message passing（再重み付け木メッセージパッシング） *231*
treewidth（木幅） *132*
trellis diagram（トレリス図） ➜ lattice diagram
triangulated graph（三角形分割グラフ） *131*
type 2 maximum likelihood（第二種の最尤推定） ➜ evidence approximation

U

undetermined multiplier（未定乗数） ➜ Lagrange multiplier
undirected graph（無向グラフ） ➜ Markov

random field
uniform distribution（一様分布） 311
uniform sampling（一様サンプリング） *248*
uniqueness（独自性） *302*
unobserved variable（非観測変数） ➜ latent variable
unsupervised learning（教師なし学習） 3
utility function（効用関数） 40

V

validation set（確認用集合） 11, **32**
Vapnik–Chervonenkis dimension（Vapnik–Chervonenkis 次元） *55*
variance（バリアンス） 147
variance（分散） **19**, 24
variational inference（変分推論法） *27*, ***176***, *353*
for Gaussian mixture *187*
for hidden Markov model *343*
local *207*
VC dimension（VC 次元） ➜ Vapnik–Chervonenkis dimension
vector quantization（ベクトル量子化） *145*
vertex（頂点） ➜ node
visualization（視覚化） 3
Viterbi algorithm（Viterbi アルゴリズム） *130*, ***347***
von Mises distribution（フォン・ミーゼス分布） **105**, 312

W

wavelet（ウェーブレット） 137
weak learner（弱学習器） *375*
weight decay（荷重減衰） 9, **142**, 259
weight parameter（重みパラメータ） 227
weight sharing（重み共有） 270
soft 272
weight vector（重みベクトル） 179
weight-space symmetry（重み空間対称性） **232**, 285
weighted least squares（重み付き最小二乗） *386*
well-determined parameter（well-determined パラメータ） 169
whitening（白色化） *9*, ***284***
Wishart distribution（ウィシャート分布） **99**, 312
within-class covariance（クラス内共分散） 186
Woodbury identity（Woodbury の公式） 314
wrapped distribution（巻き込み分布） 107

Y

Yellowstone National Park（イエローストーン国立公園） 107, **299**