

- Noise-contrastive estimation, 616
- Non-parametric model, 113
- Norm, xiv, 38
- Normal distribution, 62, 63, 124
- Normal equations, 108, 108, 111, 232
- Normalized initialization, 301
- Numerical differentiation, *see* finite differences

- Object detection, 449
- Object recognition, 449
- Objective function, 81
- OMP- k , *see* orthogonal matching pursuit
- One-shot learning, 534
- Operation, 202
- Optimization, 79, 81
- Orthodox statistics, *see* frequentist statistics
- Orthogonal matching pursuit, 26, 252
- Orthogonal matrix, 41
- Orthogonality, 40
- Output layer, 166

- Parallel distributed processing, 17
- Parameter initialization, 298, 403
- Parameter sharing, 249, 332, 370, 372, 386
- Parameter tying, *see* Parameter sharing
- Parametric model, 113
- Parametric ReLU, 191
- Partial derivative, 83
- Partition function, 564, 601, 663
- PCA, *see* principal components analysis
- PCD, *see* stochastic maximum likelihood
- Perceptron, 15, 26
- Persistent contrastive divergence, *see* stochastic maximum likelihood
- Perturbation analysis, *see* reparametrization trick
- Point estimator, 121
- Policy, 476
- Pooling, 327, 677
- Positive definite, 88
- Positive phase, 466, 602, 604, 650, 662
- Precision, 420
- Precision (of a normal distribution), 62, 64
- Predictive sparse decomposition, 519

- Preprocessing, 450
- Pretraining, 320, 524
- Primary visual cortex, 362
- Principal components analysis, 47, 145, 146, 486, 626
- Prior probability distribution, 134
- Probabilistic max pooling, 677
- Probabilistic PCA, 486, 487, 627
- Probability density function, 57
- Probability distribution, 55
- Probability mass function, 55
- Probability mass function estimation, 102
- Product of experts, 566
- Product rule of probability, *see* chain rule of probability
- PSD, *see* predictive sparse decomposition
- Pseudolikelihood, 611

- Quadrature pair, 366
- Quasi-Newton methods, 314

- Radial basis function, 195
- Random search, 431
- Random variable, 55
- Ratio matching, 614
- RBF, 195
- RBM, *see* restricted Boltzmann machine
- Recall, 420
- Receptive field, 334
- Recommender Systems, 474
- Rectified linear unit, 170, 191, 422, 503
- Recurrent network, 26
- Recurrent neural network, 375
- Regression, 99
- Regularization, 119, 119, 176, 226, 427
- Regularizer, 118
- REINFORCE, 683
- Reinforcement learning, 24, 105, 476, 683
- Relational database, 479
- Relations, 478
- Reparametrization trick, 682
- Representation learning, 3
- Representational capacity, 113
- Restricted Boltzmann machine, 353, 456, 475, 583, 626, 650, 651, 666, 670,