Cichocki, A., Zdunek, R., Phan, A. H., and Amari, S. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley & Sons Ltd., Chichester West Sussex, United Kingdom, 2009.

Coifman, R. R. and Lafon, S. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, July 2006.

Colombo, A., Cusano, C., and Schettini, R. 3D face detection using curvature analysis. *Pattern Recognition*, 39(3):444–455, March 2006.

Comon, P. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.

Comon, P. and Mourrain, B. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53:93–108, 1996.

Comon, P., Luciani, X., and De Almeida, A.L.F. Tensor decompositions, alternating least squares and other tales. *Journal of Chemometrics*, 23(7-8):393–405, 2009.

Cootes, T. F., Edwards, G. J., and Taylor, C. J. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6): 681–685, June 2001.

Cristianini, N. and Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press, United Kingdom, 2000.

Cunado, D., Nixon, M. S., and Carter, J. N. Automatic extraction and description of human gait models for recognition purposes. *Computer Vision and Image Understanding*, 90(1):1–41, January 2003.

Dai, G. and Yeung, D. Y. Tensor embedding methods. In *Proc. Twenty-First National Conference on Artificial Intelligence*, pp. 330–335, July 2006.

Dalal, N. and Triggs, B. Histograms of oriented gradients for human detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pp. 886–893, 2005.

Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961–1005, 1990.

De la Torre, F. A least-squares framework for component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (34):1041–1055, 2012.

De Lathauwer, L. *Signal Processing Based on Multilinear Algebra*. PhD thesis, Katholieke Universiteit Leuven, 1997. URL `ftp://ftp.esat.kuleuven.ac.be/sista/delathauwer/reports/PHD.pdf`.

He, X., Yan, S., Hu, Y., Niyogi, P., and Zhang, H. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):328–340, March 2005b.

Hillis, D. M. and Bull, J. J. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, 42(2): 182–192, 1993.

Hinton, G.E. and Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

Hinton, G.E., Osindero, S., and Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554, 2006.

Hitchcock, F. L. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematical Physics*, 6(1):164–189, 1927a.

Hitchcock, F. L. Multiple invariants and generalized rank of a p-way matrix or tensor. *Journal of Mathematical Physics*, 7(1):39–79, 1927b.

Ho, T. K. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8): 832–844, August 1998.

Höskuldsson, A. PLS regression methods. *Journal of Chemometrics*, 2(3): 211–228, 1988.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441, 1933.

Hotelling, H. Relations between two sets of variables. *Biometrika*, 28(3/4): 312–377, 1936.

Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., Hill, D. P, Kania, R., Schaeffer, M., St. Pierre, S., et al. Big data: The future of biocuration. *Nature*, 455(7209):47–50, 2008.

Hsu, R.-L., Abdel-Mottaleb, M., and Jain, A. K. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):696–706, 2002.

Hua, G., Viola, P. A., and Drucker, S. M. Face recognition using discriminatively trained orthogonal rank one tensor projections. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2007.

Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, May 1999.

Index

N-way PLS, 184
Q-based method, 111, 119
ℓ_1 norm, 42, 209
ℓ_1 regularization, *see* Lasso
ℓ_2 norm, 42, 209
ℓ_p norm, 209
n-mode, 50
2D-CCA, *see* two-dimensional CCA
2DLDA, *see* two-dimensional LDA
2DPCA, *see* two-dimensional PCA

active learning, 72, 83
adaptive boosting (AdaBoost), 44, 132
aggregation, 43, 83, 153
alternating least squares, 79, 90, 104
    convergence, 104
alternating partial projections, 90, 109, 115, 122, 126, 143, 146, 148, 151, 155, 169, 174, 179
architecture
    ICA, 26
    multilinear CCA, 177
    multilinear modewise ICA, 170
assembled matrix distance, 57
audio processing, 85, 202

bagging, 43
base learner, 43, 132
basis, 10, 207
Bayesian tensor analysis, 138
between-class scatter, 28
    matrix, 67
    scalar, 68
    tensor, 66
bias, 40
bias-variance trade-off, 41
big data, 1, 74, 84, 85
bioinformatics, 85, 202
blind source separation, 24, 167, 172
block matrix, 9
boosting, 43, 131
bootstrap, 43
bootstrap aggregating, *see* bagging

canonical correlation, 31
    tensor, 177, 178, 180
canonical correlation analysis, 7, 20, 31, 46
    multilinear extension, 172
canonical decomposition, 57, 79
canonical variate, 31
    tensor, 177, 178
centering, 21, 26, 75, 125, 154
characteristic equation, 211
class, 27
class discrimination power, 100
class mean
    matrix, 67
    tensor, 66
cocktail-party problem, 24
composite learner, 43
computational complexity, 102
computational thinking, 9
computer vision, 84
concurrent subspaces analysis, 75, 136
constrained optimization, 29, 180, 215
convergence, 97
coordinate matrix, 180
coordinate vector, 21, 36, 124, 125, 153, 154, 157, 177
core tensor, 58, 59, 186

correct recognition rate, 193
correlation, 31, 214
    sample, 179
    sample Pearson, 124
covariance, 214
    matrix, 31, 40
    maximization, 35
CP decomposition, *see* canonical decomposition and parallel factors
cross-correlation, 35
cross-covariance
    matrix, 31
    mode-n, 185
cross-validation, 42
CSA, *see* concurrent subspaces analysis
cubical tensor, 52
cumulant, 26, 166
curse of dimensionality, 5

data mining, 85, 203
decision tree, 43
deep learning, 83
deflation, 37, 38, 76, 121, 185
derivative
    determinant, 217
    matrix, 216, 217
    rules, 215
    trace, 216
    vector, 216
determinant, 210
diagonal tensor, 53
dimensionality reduction, 5, 19, 26
directional tensor ICA, 166
discriminant analysis with tensor representation, 77, 145
distance measure, 56
dot product, 207

EEG, *see* electroencephalography
eigendecomposition, 211
eigentensor, 95
eigenvalue, 211
    problem, 211
eigenvector, 211
electroencephalography, 1, 2, 85, 200
elementary multilinear projection, 62
ensemble learning, 43, 131, 160
entropy, 26
entry, 205
environmental sensor monitoring, 2, 3
expected value, 214

face
    2D, 1, 2, 192
    3D, 2, 3
    Gabor filter, 5
    recognition, 84, 191
factor analysis, 79, 80, 108
feature extraction, 6
feature fusion, 3
feature selection, 99, 171
    supervised, 100
    unsupervised, 101
fiber, 51
filter bank, 11
Fisher discriminant analysis, 27
Fisher's discrimination criterion, 153
Fisherface, 30
flattening, 53
fMRI, *see* functional magnetic resonance imaging
Frobenius norm, 55, 209
full projection, 93, 111, 116
    truncation, 93, 95, 96, 116
functional magnetic resonance imaging, 3, 4, 85, 201
fusion, 3

Gabor filter, 4, 73
gait, 2, 3, 197
    recognition, 196
gallery, 190
Gaussian processes, 82
general tensor discriminant analysis, 77, 147
generalization, 40, 160
generalized eigenvalue, 212

problem, 212
generalized eigenvector, 212
generalized low rank approximation of matrices, 75, 136
generalized PCA, 75, 108
generative model, 25
GLRAM, *see* generalized low rank approximation of matrices
graph-embedding, 82, 162
graphical abstract, 14

higher-order
PLS, 185
SVD, 58, 79
tensors, 2, 79
hyperspectral data, 2, 3
hypothesis, 44, 45, 132, 133

ill-conditioned, 35, 40, 41
implementation, 102
incremental learning, 81
incremental tensor analysis, 138
independent, 20, 23, 25
independent component, 24
independent component analysis, 20, 24, 46
architectures, 26
multilinear extension, 166
indicator
matrix, 38
vector, 38
initialization, 92, 110, 116, 144, 160, 161, 169, 176, 183
inner product, 55, 207
invertible, 206

joint probability mass function, 214

kerned-based learning, 82
Kronecker delta, 125
Kronecker product, 207
Kullback–Leibler divergence, 25
kurtosis, 26

label, 27
Lagrange multiplier, 22, 215
Lagrangian function, 215
Lasso, 42
latent
factor, 185
space, 185
variable, 6, 25, 36, 83, 139, 168, 184, 185
vector, 186
learning paradigm, 72
least absolute shrinkage and selection operator, *see* Lasso
left inverse, 169
linear algebra, 205
linear discriminant analysis, 7, 20, 27, 46
Fisher discriminant analysis, 27
multilinear extension, 77, 141
linear projection, 59
linear subspace learning, 7, 19
loading
matrix, 36, 186
vector, 36, 187

manifold learning, 82
margin, 44
MATLAB implementation tips, 102
matricization, 53
matrix
condition number, 213
diagonal, 206
identity, 206
ill-conditioned, 213
inverse, 206
left inverse, 206
nonnegative, 205
nonsingular, 211
orthogonal, 207
positive definite, 208
product, 207
pseudoinverse, 206
rectangular diagonal, 206
singular, 211
symmetric, 205
trace, 209
matrix calculus, 215

maximum likelihood estimator, 179
mean, 214
    matrix, 67
    tensor, 65
memory requirement, 9, 101
mislabel, 132
mixing
    matrix, 25, 170
    tensor, 168
mixture, 24
mobile
    computing, 74
    Internet, 84
    vision, 85
mode, 2, 50
mode-n, 50, 51
    eigenvalue, 94
    between-class scatter matrix, 66
    eigenvalue, 95
    product, 54, 55
    slice, 51
    source matrix, 168
    total scatter matrix, 65
    vector, 51
    within-class scatter matrix, 66
model complexity, 41
model selection, 42
multidimensional array, 1, 2
multilinear algebra, 50, 78
multilinear CCA, 172, 176
multilinear ICA, 166
multilinear mapping, 9
multilinear mixing model, 167
multilinear modewise ICA, 167
multilinear PCA, 75, 113
    nonnegative, 75
    robust, 75
    uncorrelated, 76
multilinear projection, 49, 59, 72
    partial, 83, 90
    relationships, 63
multilinear subspace learning, 1
    application, 84, 189
    constraint, 73
    framework, 72
    future direction, 81
    graph-embedding, 82
    history, 78
    incremental, 81
    kernel-based, 82
    model, 73
    optimization, 83
    probabilistic, 81
multiple discriminant analysis, 27
multiway analysis, 79
multiway array, 2
multiway CCA, 173
mutual information, 23, 25

negentropy, 26
network traffic data, 3, 4, 203
neurotechnology, 85
non-Gaussianity, 24, 26
nonlinear iterative partial least squares, 36
nonnegative
    matrix factorization, 79
    MPCA, 137
    tensor factorization, 80
nonnegative tensor factorization, 86
nonsingular, 206
normalization, 21, 22, 36

Occam's razor, 41
offline learning, 81
online learning, 81
open-source software, 228
optimality criterion, 73
optimization
    constrained, 215
order, 2, 50
orthogonal matrix, 207
orthonormal vector, 23, 207
outer product, 50, 51, 207
overfitting, 40

paired data, 31, 36, 176, 184
parallel factors, 57, 79
partial least squares, 20, 35, 47
    multilinear extension, 184
partial multilinear projection, 90

pattern recognition
  system, 190
Pearson correlation
  sample, 179
penalty, 42
PLS1, 38
poorly conditioned, 35
population, 19
power method, 37, 213
predictor, 37, 184, 185
preprocessing, 21, 26, 219
principal component, 20
principal component analysis, 6, 20, 46
  multilinear extension, 74, 107
probability mass function, 214
probe, 190
projection matrix, 19, 21, 28
projection order, 96, 160
pseudo-identity matrix, 93

random forest, 43
random initialization, 92, 93
rank-1 identification rate, 193
rank-one approximation, 36
rank-one tensor, 51, 58
Rayleigh quotient
  generalized, 39
RDA, *see* regularized discriminant analysis
real-world application, 14, 87, 189
reconstruction error
  minimization, 112
regression, 37, 184
  PLS, 37
regularization, 30, 35, 40, 135, 153, 156, 169
  covariance matrix, 41
  model complexity, 41
regularized discriminant analysis, 41
reinforcement learning, 83
reshape, 7, 19
residual matrix, 36
residue calculation, 76, 121
response, 37, 185
robust MPCA, 137

sample, 19
sample mean, 21
scalar product, 55, 207
scatter difference, 147, 151, 152
scatter matrix, 40
  between-class, 28
  total, 20
  within-class, 28
scatter measure, 64, 73
scatter ratio, 29, 143, 145, 146, 156
score
  matrix, 36, 185
  vector, 36, 185
semi-supervised learning, 72, 83
separating matrix, 25, 170
sequential mode truncation, 119
singular value, 213
  decomposition, 212
singular vector
  left, 213
  right, 213
slice, 51
small sample size, 5, 8, 10, 30, 41, 64, 153, 156
social network, 2, 3
source, 25
  matrix, 167, 171
sparse representation, 10
sparse tensor, 227
spectrogram, 1, 2, 202, 203
speech processing, 85
standard deviation, 214
strong learner, 43
structured representation, 11
subspace, 5, 19
  dimension determination, 91, 111, 118
  learning, 6
supervised learning, 20, 27, 72
support tensor machine, 82
support vector machine, 82, 100, 201

tensor, 1, 2, 50

approximation, 75, 79
decomposition, 78
factor analysis, 80
higher-order, 50
index of, 51
measure, 64
mode of, 2, 50
order of, 2, 50, 73
reconstruction, 75
reorganization, 83, 85
unfolding of, 53
tensor CCA, 173
tensor data, 73
fourth-order, 2, 4
second-order, 1, 2
synthetic, 97, 171
third-order, 2, 3, 5
tensor field, 2, 50
tensor product, 50, 207
tensor rank-one decomposition, 76, 120
tensor rank-one discriminant analysis, 77, 150
tensor space, 50, 61, 72, 108, 114
tensor subspace, 59
analysis, 163
tensor toolbox, 227
tensor-based computation, 1
tensor-to-tensor projection, 61
tensor-to-vector projection, 61
TensorFace, 80
termination, 97
test error, 40
test set, 42
three-way factor analysis, 90
total scatter, 28
matrix, 67
scalar, 67
tensor, 65, 94
training error, 40, 41
training set, 42
transfer learning, 83
transpose, 205
tri-linear PLS1, 185
TSA, *see* tensor subspace analysis
Tucker decomposition, 58, 79
two-dimensional
CCA, 172, 173
LDA, 77, 142
PCA, 74, 135

uncorrelated, 20, 23, 33, 176, 214
uncorrelated multilinear discriminant analysis, 78, 153
aggregation, 78, 160
regularized, 78, 154
uncorrelated multilinear PCA, 124
unfolding, 53
uniform initialization, 93
unmixing matrix, 25
unsupervised learning, 20, 72

validation set, 42
variance, 40, 68, 214
variation, 20, 21
vector
linearly dependent, 206
linearly independent, 206
normalized, 207
orthogonal, 207
orthonormal, 23, 207
unit, 207
vector space, 50, 207
basis, 207
span, 207
vector-to-vector projection, 60
vectorization, 7, 19, 53
video, 2, 3, 198
volume measure, 57

weak learner, 43, 132
web graph, 2, 3
weight parameter, 24
weight vector, 36
whitening, 26
within-class scatter, 28
matrix, 67
scalar, 68
tensor, 66

zero-correlation, 22, 32, 35