

Index

a

accelerated degradation tests (ADTs) 92, 112
 ad hoc techniques 146
 adaptive direction sampling (ADS) 69
 adjacency matrix 37, 40, 269
 agglomerative hierarchical clustering
 – advantage 276
 – average linkage 276
 – cluster validity 276
 – complete linkage 276
 – cutree function 276
 – disadvantage 278
 – similarity/dissimilarity 274
 – single linkage 276
 Akaike information criterion (AIC) 91, 92, 253
 alternating direction method of multipliers (ADMM) algorithm 235, 240
 anomaly detection
 – adjacency matrix 37
 – chi-squared test 54
 – data analysis 55
 – degree of a vertex 37
 – dynamic graph model 43
 – eigenvectors 52
 – frequent subgraph technique 38, 39
 – implementation in R 50
 – matrix representation of graph 37
 – modularity matrix 52
 – Monte Carlo simulation 51
 – neighborhood-based techniques 37
 – notation 36
 – performance 54
 – random graph 39, 41
 – residuals matrix 54
 – spectral characteristics 54
 – spectrum 52
 – subgraph *see* spectral subgraph detection

Arabidopsis

– *comp.2.cc.fdr* function 25
 – eigen-molecule/eigengene calculation 21
 – flavonoid-deficient and wild-type 26
 – module network visualization 23
 – sinapate and aromatic metabolites 28
 AraMetLeaves (data matrix) 27
 assay networks 157

b

Bayesian inference

– advantages and effectiveness 66
 – ERGM 65, 68
 – GoF analysis *see* predictive goodness-of-fit (GoF) diagnostics
 – posterior distribution 66
 – probabilistic 66
 – R-based software tools 67

Bayesian information criterion (BIC) 91, 230, 232, 249, 253, 254

bc3net gene regulatory network inference

– aggregation of ensembles 294
 – binomial test 294
 – bootstrap ensemble 294
 – c3net advantage 294
 – *getgcc()* function 295
 – global visualization 295
 – igraph R package 294
 – layout function 295
 – mutual information 294
 – NetBioV 296
 – sparse network 294

Bethel algorithm 127

bilinear latent model setting 66

binary hypothesis test 45

Bioconductor 291

- biological network alignment
 - applications 174
 - biomolecules 176
 - challenges 182
 - characteristics 181
 - classic graph matching problem 175
 - classification 175
 - CRF based algorithm *see* conditional random fields (CRF)
 - directed/undirected graph 176
 - MNA 176
 - example of 180
 - many-to-many alignment 180
 - one-to-one alignment 180
 - vs. PNA 182
 - score function 179
 - models and algorithms 180
 - NQ 175
 - classic network querying 178
 - network schemas 179
 - vs. PNA 180
 - schema network querying 178
 - PNA 175
 - definition of 176
 - example of 177
 - feature functions 176
 - gaps 177
 - global network alignment algorithm 178
 - local network alignment algorithm 178
 - many-to-many alignment 178
 - vs. MNA 182
 - vs. NQ 180
 - one-to-one alignment 178
 - optimum mapping 176
 - PPI networks 174
 - protein interaction 174
 - vs. sequence alignment 173, 174
 - weighted/unweighted graph 176
 - biological network visualization
 - algorithms 311, 312
 - cellular system 307
 - discovery driven-data analyses 310
 - gene networks 308
 - graph editing and visualization platform 326
 - hairball approach 325
 - Kamada–Kawai algorithm 311
 - meta information 326
 - metabolic networks 308
 - multidimensional approach 311
 - NetBioV
 - color schemes, node labeling 318
 - global layout style 320
 - global network layouts 313
 - global, modular and layered layouts 313
 - information flow 318
 - interface to R 319
 - layered layout style 323
 - layered network (multiroot) layout 317
 - layout architectures 313, 315
 - library and data loading 320
 - modular and hierarchical structure 313
 - modular layout style 322
 - modular network layout 316
 - number of edges 319
 - number of vertices 319
 - vs. other popular packages 313, 314
 - PPI interactions, Arabidopsis Thaliana 313, 316, 317
 - spiral view 318, 329
 - network layout algorithms 326
 - network structure 308
 - nodes, edges and paths 310
 - public databases 308
 - random layout 311
 - reductionist approach 307
 - statistical data analysis 310
 - structural characteristics 308
 - subnetworks 308
 - system-based approach 307
 - topological diversity and dynamical complexity 308
 - block-coordinate descent algorithm 223
 - breadth-first sampling (BF) 146
- c**
- Cartesian products 42
 - categorical parameters 52
 - cell proliferation 21
 - chi-squared test 54
 - Chinese restaurant process 42
 - cluster analysis, social networks
 - community structure 267
 - conventional vs. network data 269
 - grouping nodes 269
 - grouping set of entities 268
 - modern methods 269
 - R packages
 - agglomerative hierarchical clustering 274
 - cclust 270
 - clusterSim 270
 - clv 270
 - clValid 270
 - creating social and task subgraphs 272
 - data coercing 271
 - data formatting 271
 - data loading 271

- edge betweenness clustering algorithm 279
 - fast Greedy modularity optimization algorithm 281
 - igraph 270
 - modMax 270
 - NbClust 270
 - NetData 270
 - sna 270
 - statnet 270
 - walktrap algorithm 283
 - social network data
 - dichotomous variable 268
 - graph theory 268
 - matrix algebra 269
 - traditional methods 269
 - unsupervised learning 268
 - coefficient solution path 217
 - coexpression networks 2
 - *comp.2.cc.fdr* function 25
 - Comprehensive R Archive Network (CRAN) 67, 291
 - computational capability 50
 - conditional random fields (CRF)
 - CNetA, pairwise network alignment
 - asymmetric results 187
 - evaluation 189
 - vs. IsoRank method 188
 - iterative bi-directional mapping strategy 187
 - many-to-one mapping 186
 - simulated data 188
 - CNetMA, multiple network alignment
 - examples 190
 - Grmlin 189
 - IsoRank 190
 - iterative network-structure update strategy 191
 - large scale multiple network alignment 192
 - running time 192
 - CNetQ, network querying
 - biomolecular network 183
 - conditional probability 184
 - feature function 184, 185
 - gap penalty 185
 - GraphMatch 186
 - labeling problem 183
 - metabolic pathway 186
 - network simplification 186
 - tree-like pathway 186
 - undirected protein metabolism 186
 - Viterbi algorithm 185
 - network alignments with R
 - arguments 194
 - Corbi package 193
 - CRF package 193
 - input file format 194
 - output file format 194
 - connection probability 42
 - connectivity pattern of graph 64
 - corrected Akaike information criterion (AICc) 91
 - crawling 146
 - cumulative distribution function (CDF) 86, 88, 90, 93, 105
 - cyclic coordinate descent (CCD) method 224, 246, 247
 - cyclic coordinate minimization (CCM) algorithm 246, 247
 - cytoscape 15, 23
- d**
- data filtering 9
 - degradation models 83, 84
 - advantage of 85
 - identifiability of 86
 - selection 96
 - degradation paths 83, 86
 - DiffCorr package 4
 - *comp.2.cc.fdr* 6
 - Fisher's z-test method 5
 - *get.eigen.molecule.graph* 6
 - installation 6
 - pattern change identification 4
 - *plot.DiffCorr.group* 6
 - R package DCGL 4
 - differential correlation networks
 - arabidopsis 21, 26
 - construction, transcriptome dataset *see* transcriptome dataset
 - DEGs 2
 - DiffCorr package 4
 - graph clustering 2
 - omics and system biology 2
 - pattern visualization 30
 - differentially expressed genes (DEGs) 2
 - digraph 64
 - directed graph 203
 - disease networks 155
 - diseaseome 155
 - divisive algorithms 274
 - drug-target networks
 - chemogenomic features 155
 - drug repositioning methodologies 155
 - DrugComboRanker package 155
 - FacPad method 155
 - FDA approved drugs 155

drug-target networks (*contd.*)

- follow on drugs 154
- MANTRA 155
- MetaCore and MetaDrug suite 154
- pathway enrichment analyses 154
- pathway reconstruction 154
- protein pockets 155
- protein–protein interaction network 154
- repositioning and activity predictions 155
- TIMMA-R package 155
- topological properties 154

dynamic networks 43, 44

e

edge attributes 45

edge betweenness clustering algorithm 279

edge features 176

edge probability 40, 47

egonet 37

equiangularity condition 216, 220

Euclidean distance (ED) 66

exclusion condition 216

exclusion transition value 216

exponential random graph models (ERGMs) 65

- acceptance rate 70
- ADS 69
- approximate exchange algorithm 69
- iterations, network simulation 69
- Metropolis–Hastings algorithm 68
- overall posterior density estimation 70
- parameters 70
- posterior distribution 68
- posterior mean 70
- posterior sd 70
- two new specification statistics 69

f

factorial dynamic Gaussian graphical model (fdGGM) 243, 244, 250, 251

false discovery rate (FDR) 5

fast Greedy modularity optimization algorithm 16, 281

Fisher's *z*-transformation of coefficient 5

Fruchterman–Reingold algorithm 72, 321, 324

fused graphical lasso 234, 235, 239, 250

g

gamma degradation-based process 84

- lifetime information 88
- log-likelihood function 89

Gaussian Markov random field (GMRF) 206, 207

- goodness-of-fit 253
- gradient and coefficient solution path 217
- graphical lasso estimator 215, 223
- *K*-fold cross-validation 255
- neighborhood selection 207
- osteolytic lesions data set 210
- simone 209
- stability selection 254

Gaussian process 83

gene co-expression analysis 2

gene ontology (GO) 16

gene pair enrichment analysis (GPEA) 301

gene regulatory networks (GRN) 176

- bc3net inference
 - – advantage 294
 - – aggregation of ensembles 294
 - – binomial test 294
 - – bootstrap ensemble 294
 - – *getgcc()* function 295
 - – global visualization 295
 - – igraph R package 294
 - – layout function 295
 - – mutual information 294
 - – NetBioV 296
 - – Pearson correlation coefficient 294
 - – sparse network 294
- data preprocessing
 - – affymetrix microarray 292
 - – entrez gene identifier 293
 - – gene symbols 292
 - – GSE matrix file 292
 - – probe sets 292
- functional analysis, Gene Ontology gene sets 297
- functional enrichment analysis
 - – Achilles heel 301
 - – B-cell network 300
 - – endoplasmatic stress 301
 - – GPEA 301
 - – GSEA 300
 - – hub gene subnetwork 300
 - – hypergeometric test 300
 - – inferred module 300
 - – mitotic cell cycle 302
- gene identifiers 290
- gene symbols 290
- genome-wide gene level data 290
- inference methods 290
- molecular mechanisms 289
- multiple myeloma 291
- pathway and gene set collections 298

- R packages installation 291
- gene set enrichment analysis (GSEA) 300
- generalized linear model (GLM) 42
- get.eigen.molecule* function 21
- global Markov property 204
- GOstats 17
- graph clustering 2
 - fastgreedy.community function 16
 - fidelity, GO enrichment analysis 16
 - hierarchical clustering 16
- graph theory
 - advantages of 202
 - definition 202
 - directed edge 202
 - directed graph 203
 - undirected edge 202
 - undirected graph 203
 - vertex-set 202
- graphical lasso estimator 215
 - coefficient solution path 217
 - computational aspects of 223
 - definition 215
 - equiangularity condition 216
 - exclusion condition 216
 - exclusion transition value 216
 - faster computation, exact covariance thresholding 225
 - gradient path 217
 - inclusion condition 216
 - inclusion transition value 216
 - joint graphical lasso 233, 239, 241
 - KKT conditions 215
 - least angle regression method 216
 - lung cancer microarray data 227
 - structured graphical lasso 243, 248
- graphical user interface (GUI) 84, 122
- group graphical lasso 235, 238, 239
- guide-gene approach 2

h

- Hannan–Quinn criterion (HQC) 91
- Hansen–Hurvitz estimator 146, 148
- HGNC gene symbols 299
- hierarchical clustering techniques
 - agglomerative algorithms 274
 - divisive algorithms 274
- high-dimensional GMRFs, sparse inference *see* Gaussian Markov random field (GMRF)

i

- iDEMO
 - ADTs 112
 - basic information 95
 - computational details 101

- data input format 93
- degradation model selection 96
- fatigue data 106
- functionality 93
- import data 94
- laser data 102
- monotone 95
- non-monotone 95
- odds and ends 101
- optimization algorithms 96
- single degradation model analysis 96
- inclusion condition 216
- inferential framework 41
- inverse Gaussian degradation-based process
 - degradation data analyses 89
 - lifetime distribution 90
 - log-likelihood function 91
- Ising model 205

j

- joint graphical lasso 233, 239
 - definition 234
 - lung cancer microarray data 241

k

- Karush–Kuhn–Tucker (KKT) conditions 215
- Kozak's algorithm 129

l

- Lanczosbidiagonalization algorithm 50
- laser data 102
- latent position cluster models (LPCMs) 66, 73
- latent space models (LSMs) 42, 65
 - Fruchterman–Reingold method 72
 - function and argument 72
 - Gaussian mixture model structure 73
 - latent positions 73, 74
 - LPCMs 73, 76
 - MCMC simulations 75
 - parameter estimation 71
 - posterior probability 72
 - squared Euclidean distance 72
 - variational maximisation algorithm 72
 - visualisation and prediction properties 71
- latent variable model 41
- least angle regression method 216
- Life Satisfaction Index 147, 148
- light-emitting diodes (LEDs) 112, 116, 119
- linear degradation models 85
- linear regression model 216
- link-trace sampling 146
- local Markov property 204

- log-likelihood function
 - gamma degradation-based process 89
 - inverse Gaussian degradation-based process 91
 - Wiener degradation-based process 87
- lung cancer microarray data 227, 241

m

- Markov chain Monte Carlo (MCMC)
 - algorithms 67
- Markov clustering 15
- Markov random field (MRF)
 - definition 203
 - factorization 204
 - global Markov property 204
 - GMRF 206
 - Ising model 205
 - local Markov property 204
 - pairwise Markov property 204
 - partition function 204
 - potential functions 204
- matrix decomposition 51
- matrix representation of graph 43
- maximum-likelihood estimation 42, 43
- metabolic networks 176
- metabolome dataset 27
- methionine overaccumulation 1 (mto 1) 27
- Metropolis–Hastings Random Walk sampling (MHRW) 146
- minimum description length (MDL) heuristic 38
- minimum spanning tree (MST) algorithm 315
- Mode of Action by NeTwoRk Analysis (MANTRA) 155
- multiple myeloma
 - antibodies 291
 - clonal heterogeneity 291
 - plasma cells 291
- multiple network alignment (MNA) 176
 - CNetMA
 - – examples 190
 - – Grmlin 189
 - – IsoRank 190
 - – iterative network-structure update strategy 191
 - – large scale multiple network alignment 192
 - – running time 192
 - example of 180
 - many-to-many alignment 180
 - one-to-one alignment 180

- vs. PNA 182
- score function 179

n

- neighborhood selection 207
- Neisseria meningitis data set 250
- nested network file format (NNF) 23
- NetBioV
 - color schemes, node labeling 318
 - global layout style 320
 - global network layouts 313
 - global, modular and layered layouts 313
 - information flow 318
 - interface to R 319
 - layered layout style 323
 - layered network (multiroot) layout 317
 - layout architectures 313, 315
 - library and data loading 320
 - modular and hierarchical structure 313
 - modular layout style 322
 - modular network layout 316
 - number of edges 319
 - number of vertices 319
 - vs. other popular packages 313, 314
 - PPI interactions, Arabidopsis Thaliana 313, 316, 317
 - spiral view 318, 329
- network querying (NQ) 175
 - classic network querying 178
 - CNetQ
 - – biomolecular network 183
 - – conditional probability 184
 - – feature function 184, 185
 - – gap penalty 185
 - – labeling problem 183
 - – metabolic pathway 186
 - – network simplification 186
 - – tree-liked pathway 186
 - – undirected protein metabolism 186
 - Viterbi algorithm 185
 - network schemas 179
 - vs. PNA 180
 - schema network querying 178
- network similarity graphs (NSG's) 156
- network topology 36
- Newman's modularity matrix 47
- next-generation sequencers (NGS) 1
- node features 176
- nonlinear degradation models 85
- NP-hard problem 40

o

- omics 1
- one-target, one-drug approach 151
- osteolytic lesions 210

p

- pairwise Markov property 204
 - pairwise Markov random field (pMRF) 205
 - pairwise network alignment (PNA) 175
 - CNetA
 - – asymmetric results 187
 - – evaluation 189
 - – vs. IsoRank method 188
 - – iterative bi-directional mapping strategy 187
 - – many-to-one mapping 186
 - – simulated data 188
 - definition of 176
 - example of 177
 - feature functions 176
 - gaps 177
 - global network alignment algorithm 178
 - local network alignment algorithm 178
 - many-to-many alignment 178
 - vs. MNA 182
 - vs. NQ 180
 - one-to-one alignment 178
 - optimum mapping 176
 - PathwayCommons 299
 - Pearson correlation coefficient 2
 - penalty parameter 236
 - Potts model 206
 - PPI networks 176, 182
 - predictive goodness-of-fit (GoF) diagnostics
 - LPCMs 78, 79
 - LSM 78
 - network statistics 76, 77
 - observed network data 76
 - principal component analysis (PCA) 21
 - probabilistic graphical models
 - Bayesian network 203
 - definition 203
 - Markov random field 204
 - probability density function (PDF) 86
 - probability matrix 48
 - protein–protein interaction (PPI) networks 174
 - Pubchem Bioassay meta data and text mining 157
- q**
- quality characteristics (QCs) 102, 106, 112

r

- random graph techniques
 - configuration model 40
 - edge probabilities 40
 - GLM 42
 - maximum likelihood estimation 43
 - models with attributes 41
 - modularity matrix 40
 - NP-hard problem 40
 - observed graph 39
 - planted clique and partition problems 39
- random walk sampling (RW) 146
- re-weighted random walk (RWRW) 146, 148
- real-valued attributes 45
- residual analysis
 - arbitrary parameterization 50
 - assignment matrix 48
 - categorical vertex pair attributes 48
 - coefficients estimation 47
 - edge probability 47
 - expected value 47
 - logistic regression model 47
 - matrix estimation 49
 - matrix–vector multiplication 48
 - parameter approximation scheme 49
 - probability matrix 48
 - residuals matrix 48
 - standard modularity matrix 50
- RNA methylation 21
- robust multichip average (RMA) 9

s

- SamplingStrata
 - general setting
 - – atomic strata 128
 - – auxiliary information 128
 - – Bethel algorithm 127
 - – coefficient of variation 127
 - – cost function 126
 - – datasets 129, 130
 - – genetic algorithm 128, 129
 - – Horvitz–Thompson estimator variance 126
 - – Kozak’s algorithm 129
 - – Lavallée and Hidioglou method 128
 - – multivariate and multi-domain case 129
 - – number of possible alternative stratifications 128
 - – optimisation of stratification 126
 - – population stratification 129
 - – sampling units allocation 129
 - – stratification validity 128
 - – target variables 126, 127
 - – univariate case 129

SamplingStrata (*contd.*)

- open source system 125
- optimisation procedure
 - – example 132
 - – formalising 131
 - – genetic algorithm 130
 - – stratification steps 130
- scaffold networks 158
- scaffold-document networks
 - bioactivity data 162
 - bipartite network 160
 - ChEMBL 159, 160
 - document nodes 162
 - fragment nodes 161
 - imidazonaphthalimides 160
 - limited to three iterations 161
 - MeSH headings 161
- sequence alignment methods
 - vs. biological network alignment 173, 174
 - global and local alignments 173
 - pairwise and multiple sequence alignments 173
- signature identification 43
- single degradation model analysis 96
 - lifetime information 98
 - parameter estimation 97
- singular value decomposition 46
- small molecules
 - assay networks 157
 - disease networks 155
 - drug discovery, networks role in 152
 - drug-target networks
 - – chemogenomic features 155
 - – drug repositioning methodologies 155
 - – DrugComboRanker package 155
 - – FacPad method 155
 - – FDA approved drugs 155
 - – follow on drugs 154
 - – MANTRA 155
 - – MetaCore and MetaDrug suite 154
 - – pathway enrichment analyses 154
 - – pathway reconstruction 154
 - – protein pockets 155
 - – protein–protein interaction network 154
 - – repositioning and activity predictions 155
 - – TIMMA-R package 155
 - – topological properties 154
 - network analyses
 - – graph algorithms 154
 - – protein targets 154
 - – R packages 153
 - polypharmacological behavior 151
 - putative primary target 151
 - R, drug discovery
 - – annotated network 164
 - – bioconductor packages 162
 - – cheminformatics 162
 - – computational representations 162
 - – cutting edge statistical techniques 162
 - – cytoskeletal proteins 164
 - – hashed path fingerprints 163
 - – machine learning methods 162
 - – pairwise Tanimoto similarity matrix 163, 164
 - – protein–protein interactions 164
 - – rcdk package 163
 - – resultant similarity matrix 163
 - – SMILES strings 163
 - – statistical modeling 162
 - SAR networks 156
 - scaffold networks 158
 - scaffold-document networks
 - – bioactivity data 162
 - – bipartite network 160
 - – ChEMBL 159, 160
 - – document nodes 162
 - – fragment nodes 161
 - – imidazonaphthalimides 160
 - – limited to three iterations 161
 - – MeSH headings 161
- social network analysis
 - Bayesian approach *see* Bayesian inference
 - complexity 63
 - Dolphin network dataset 67
 - Dolphin undirected network graph 68
 - network statistics *see* statistical model
 - observed network 63
 - random graphs 64
- social networks 44
- social space 269
- spectral norm, integrated subgraph 46
- spectral subgraph detection
 - analysis, attribute graph *see* residual analysis
 - filter optimization 46
- statistical model
 - Bernoulli random graph model 65
 - ERGM 65
 - latent space models 65
 - LSMs 65
 - Markov random graph model 65
 - stochastic block models 65
- stochastic block model 41, 65
- stochastic degradation-based process
 - choice of $\Lambda(t)$ 91
 - gamma degradation-based process 88

- inverse Gaussian degradation-based process 89
- model selection criteria 91
- threshold degradation 92
- Wiener degradation-based process 85
- stratified sampling
 - auxiliary information 139
 - Italian municipalities, case study
 - – atomic strata 141
 - – auxiliary information 141
 - – Bethel algorithm 141
 - – betweenness 144
 - – centrality measures 139, 144, 145
 - – correlation between centrality measures and target variables 140
 - – data frame structure 140
 - – eigenvector 144
 - – *k*-means clustering method 144
 - – maximum expected coefficient of variation 141
 - – sample size 143, 145
 - – target estimates 139
 - – weighted out degree 142
 - massive networks
 - – BF 146
 - – domains of interest 146
 - – Facebook social network 147
 - – MHRW 146
 - – online networks 145
 - – probabilistic sampling 146
 - – random walk sampling 146
 - – RWRW 146, 148
 - – S-WRS 148
 - – S-WRW sampling methods 147
 - network data applications 125
 - random sampling 125
 - sampling error 125
 - sampling frame 139
 - SamplingStrata 125
 - strata 125
- stratified sampling via weighted random walks (S-WRW) sampling methods 147
- stratified weighted random sampling (S-WRS) 148
- stratified weighted random walk technique 126
- structure activity landscape index (SALI) 156, 157
- structure–activity relationship (SAR) networks 156
- structured graphical lasso 243, 248
 - cyclic coordinate algorithms 246
 - fdGGM 243, 244
- Subdue 38
- t**
 - 3-way tensor 46
 - time series 44
 - traditional graph theory 36
 - transcriptome dataset
 - CEL file 9
 - correlation matrices 12
 - cytoscape 15
 - data filtering 9
 - filter function 11
 - gene ontology enrichment analysis 17
 - GEOquery package 9
 - graph clustering 15
 - igraph package 13, 15
 - NCBI GEO 8
 - probesets identification 12
 - pseudo-color heatmap visualization 12
 - transition value 216
- u**
 - undirected graph 203
- v**
 - vertex pair attributes 45
 - Viterbi algorithm 185
 - Volterra integral equation 86
- w**
 - Walktrap algorithm 283
 - Ward's method 283
 - Wiener degradation-based process 84, 85
 - lifetime information 86
 - log-likelihood function 87

