

Index

- L_∞ norm, 188
- L_p norm, 188
- CAT(0), 267
- CAT(0) space, 267
- CAT(κ) space, 268
- ϵ -net, 35
- σ -algebra, 183
- p -Wasserstein distance, 193
- adult T-cell lymphoma or ATL, 375
- ancestral recombinant graphs (ARGs), 342, 462, 464
- aneuploidy, 358, 366, 370
- animalia paradoxa, 274
- ARMA models, 430
- Arsuaga, J., 427, 435
- asynchronous differentiation processes, 405
- atlas, 114
- ATRX, 370
- attaching map, 58
- autoregressive models (AR), 430
- Ayasdi, 446
- B-cell lymphomas, 371
- Baas, N., 435
- bacteria, 315
 - Caulobacter crescentus*, 419
 - conjugation, 317
 - E. coli*, 318
 - H. pylori*, 319
 - horizontal gene transfer, 316
 - K. pneumoniae*, 319
 - S. enterica*, 319
 - Staphylococcus aureus*, 322
 - transduction, 317
 - transformation, 317
- ball, 32
 - closed, 32
 - open, 32
- Baltimore, D., 360
- Bandelt-Dress decomposition, 465
- Barbacid, 361
- barcode, 136
 - confidence set, 208
 - metric, 146, 147
- barcode ensemble, 342, 346
- BCR-ABL, 371
- bijective, 26
- Billera-Holmes-Vogtmann space of trees, 264
- Bishop, M., 360
- blastocyst, 408
- bottleneck distance, 146
- boundary, 93
- boundary map, 104
- Boveri, T., 358
- BRAF*, 394
- brain functional networks, 437
- branch and bound, 457
- Bray-Curtis dissimilarity, 30
- Bryant, D., 465
- Buneman, P., 283, 339
- Buneman graph, 337, 339
- Burkitt's lymphomas, 374
- cancer, 356
- Cancer Cell Line Encyclopedia (CCLE), 394
- Cancer Genome Atlas (TCGA), 370, 380, 381, 387, 391, 398
- Cantor, 28
- Cartan-Hadamard theorem, 268
- Cartesian product
 - groups, 70
 - sets, 26
 - topological spaces, 41
- category, 81
 - of abstract simplicial complexes, 98
 - of chain complexes, 109
 - simplicial complexes, 96
- CDKN2A*, 434

496

cell attachment, 58
cell cycle, 424
centroid, 408
chain complex, 109
 homology of, 110
 of abelian groups, 112
 of vector spaces, 109
chains, 103
chart, 114
Čech complex, 124
Čech nerve
 geometric, 124
chromatin, 412
chromosomal gains, 366, 370, 434
chromosomal loss, 366, 370, 434
chromosomal territories, 412
chromosome conformation capture, 413
chronic myelogenous leukemia, 371
CIC, 385
circadian rhythm, 424
classification of surfaces, 117
closed set, 38
clustering
 relationship to path components, 48
 single-linkage, 48, 137
coalescence, 286, 324, 327, 450
Cohen-Steine, D, 424
commutative diagram, 83
compact
 closed and bounded, 43
complex
 median, 337, 339
 Vietoris-Rips, 284
compressibility, 431
congruence class, 65
contact maps, 414
continuous, 37, 43
convolution, 188
copy number variation, 370, 434
correlograms, 430
cosets, 68
cover
 open, 42
cover of a space, 124
Crowford, L., 361
CW complex, 58
 quotient, 60
 subcomplex, 60
cyclic group, 65

Darwin, C., 274, 351
Darwin's finches, 351
deformation retraction, 52
dengue virus, 315
dense subset, 35
diameter, 36
diffusion distance, 251
Dionysus, 445

Index

Dipha, 445
discrete harmonic analysis, 250
disease space, 440
dispersion, 408
distance distribution, 218
distance to a measure, 204
distributional persistent homology, 199
Drosophila melanogaster, 342
Dulbecco, R., 360

edge
 external, 262
 internal, 262
EGFR, 370, 385, 403
eigenvalue, 238
eigenvector, 238
Eirene, 445
embryos, 408
endoderm, 408
enhancer-promoter interactions, 415
epiblast, 408
equivalence relation, 29
 homeomorphism of spaces, 47
 homotopy equivalence, 51
 homotopy of maps, 50
 isomorphism of sets, 29
 weak homotopy equivalence, 74
Euclidean space, 31
Euler characteristic, 101
evolution, 273
 clonal, 277
 horizontal, 277
 reticulate, 277
 vertical, 277
exome, 365

face, 94
Felsenstein, J., 459
FGFR-TACC gene fusions, 371
field, 75
filtered simplicial complex
 tame, 136
filtered system, 135
 of abelian groups, 135
 of simplicial complexes, 135
 of vector spaces, 135
filtered vector space
 finite type, 137
finitely generated, 69
Fisher, R., 447
Fitch algorithm, 456
four gamete test, 337, 339
four point condition, 262, 283
Fourier transform, 428
Fréchet mean, 190
Fréchet variance, 190
fractal globule, 412

functional MRI, 437	and persistent homology, 137
functor, 90	dendrogram, 138
functoriality	homologous recombination, 328
of path components, 46	homology group, 105, 110
fundamental group, 53	algorithm for computing, 108
	homotopy, 49
gene expression, 376	homotopy class, 52
gene fusions, 371	homotopy equivalence, 50
gene trees, 333	homotopy group, 53
generator and relations, 69	homotopy type, 51
geometric random complex, 227	Human Immunodeficiency Virus, 303
geometric random graph, 227	Env, 304
geometric realization, 95	Gag, 304
geometric sampling, 189	Pol, 304
glioblastoma, 371, 377, 384, 403, 438	Recombination, 307
classical, 404	hippocampus, 436
mesenchymal, 404	
multicentric glioblastoma, 403	<i>IDH1</i> , 385
Multifocal, 438	imaging, 424
neural, 404	imatinib, 371
proneural, 404	indels, 366
graph, 31	inf, 36
geometric random, 227	infectious diseases, 440
Laplacian, 246	infimum, 36
neighborhood, 48	influenza
path components, 48	2009 pandemic, 293
weighted, 31	avian, 287
Gromov, M., 339	hemagglutinin, 289
Gromov-Hausdorff distance, 30, 144	neuraminidase, 289
Gromov-Prohorov distance, 194	pandemic, 289
Gromov-Wasserstein distance, 194	reassortment, 289, 291, 293
group, 62	seasonal, 288
abelian, 64	Spanish influenza, 291
cyclic, 65	influenza A, 287
free, 69	initial object, 83
infinite cyclic, 66	injective, 26
Lie, 118	interior, 93
quotient, 68	inverse function, 27
rank of, 71	Isomap, 242
subgroup, 67	isomorphism, 27
group homomorphism, 66	categorical, 82
image of, 66	of abstract simplicial complexes, 98
isomorphism, 67	of groups, 67
kernel of, 66	of sets, 27
Gudhi, 445	of simplicial complexes, 96
	of vector spaces, 80
Haldane, S., 447	isomorphism class, 29
Hamming distance, 31, 460	
Harer, J., 424	Javaplex, 445
Hausdorff distance, 142	Jonckheere trend test, 429
hepatitis C virus, 315	Jonckheere-Terpstra-Kendall (JTK) test, 429
hepatocellular carcinomas, 375	Jukes-Cantor distance, 460
Hera, 445	
Hessian eigenmaps, 247	Kaposi sarcomas, 375
heterochromatin, 412	KeplerMapper, 446
Hi-C, 413	kernel methods, 248
hierarchical clustering	kernel PCA, 248

498

Kimura diffusion model, 450
 Kimura, M., 447
 Kolmogorov, A., 447
 Kullback-Leibler divergence, 30, 253

landmark embedding distribution, 219
 landmark point, 159
 Lane, D., 361
 Laplacian eigenmaps, 246
 leaf, 262
 Levine, A., 361
 Li-Fraumeni families, 363
 Lie group, 118
 linear transformation, 79
 image, 79
 kernel, 79
 linearly independent, 77
 Linnaeus, C., 273
 LLE, 244
 local linear embedding, 244
 local tangent space alignment, 252
 Lomb-Scargle periodogram, 428
 loss of heterozygosity, 366, 370

MacPherson, R., 415
 malaria, 424, 427
 manifold
 Riemannian, 117
 smooth, 115
 topological, 115
 manifold charting, 252
 manifold learning
 LLE, 244
 local tangent space alignment, 252
 manifold charting, 252
 maximum variance unfolding, 252
 map, 26
 bijective, 26
 continuous (metric spaces), 37
 continuous (topological spaces), 43
 homeomorphism, 46
 homotopic, 49
 homotopy, 49
 homotopy classes of, 52
 homotopy equivalence, 50
 injective, 26
 isomorphism, 27
 Lipschitz, 38
 of sets, 26
 quasi-isomorphism of chain complexes, 110
 quotient, 56
 simplicial, 96
 surjective, 26
 Mapper, 404, 406
 packages for computing, 445
 maximum variance unfolding, 252
 measurable function, 184

Index

measurable isomorphism, 184
 measurable space, 184
 measure, 184
 counting, 189
 support of, 186
 measure space, 184
 median graph, 337
 median networks, 463
 meiotic recombination, 328
 melanomas, 385
 Merkel cell carcinomas, 375
 metric, 30
 bottleneck, 146
 diffusion, 251
 earth-mover, 193
 graph, 31
 Gromov-Hausdorff, 30, 144
 Gromov-Prohorov, 194
 Gromov-Wasserstein, 194
 Hamming, 31, 460
 Hausdorff, 142
 Jukes-Cantor, 460
 Pearson correlation, 32
 Prohorov, 192
 Tajima-Nei, 461
 tree, 262
 Wasserstein, 147, 193
 metric measure space, 186
 metric space, 30
 complete, 35
 curvature of, 267
 finite, 123
 geodesic, 267
 length of path, 266
 length space, 267
 non-positively curved, 268
 of phylogenetic trees, 264
 separable, 35
 minimal cycles, 415
 modular arithmetic, 65
 Moran model, 449
 morphism, 81
 Morse function, 119
 Morse theory, 118
 morula, 408
 most recent common ancestor, 452
 Moulton, V., 465
 moving average process, 430
 MRI, 438
 Muller, H. J., 358
 multifiltered complex, 154
 multifiltered vector space, 154
 mutations
 germline, 363
 somatic, 363
 MYC, 371, 376

nasopharyngeal carcinomas, 374

- natural transformation, 91
- Needleman-Wunsch algorithm, 455
- neighbor-joining, 262
- neighbor-net, 465
- neighborhood, 41
- neighborhood graph, 48
- nerve of a cover, 124
- networks, 423
- neuronal activity, 436
- neuroscience, 436
- NMF, 241
- non-coding RNA, 376
- non-negative matrix factorization, 241
- Nordling, C. O., 356
- NOTCH1*, 385
- Nowell, P., 363, 371
- nucleosomes, 412
- object
 - initial, 83
 - terminal, 83
- oncogene, 360
- oncovirus, 360, 374
- open set, 38
- p53, 361
- partition, 57
- path, 44
- path components, 45
- path-connected, 45
- PDGFRA*, 370, 410
- Pearson correlation distance, 32
- Perea, J., 424
- periodogram, 428
- Perseus, 445
- persistence diagram, 137
- persistence landscape, 222
- persistence landscape toolbox, 445
- persistent homology, 137, 415
 - and hierarchical clustering, 137
 - barcode, 136
 - distributional, 199
 - packages for computing, 445
- persistent homology transform, 157
- phylogenetic networks, 462
- phylogenetic tree, 262, 274
- phylogenetics
 - agglomerative methods, 461
 - Bayesian methods, 459
 - Distance methods, 460
 - Fitch-Margoliash method, 461
 - likelihood methods, 458
 - Markov chain Monte Carlo (MCMC), 459
 - nearest neighbor interchange (NNI), 458
 - neighbor-joining (NJ), 461
 - parsimony, 456
 - pruning and regrafting (SPR), 458
 - tree bisection and reconnection (TBR), 458
- PIK3CA*, 438
- Plasmodium chabaudi*, 440
- Plasmodium falciparum*, 427
- plectoneme, 417
- point mutations, 366
- polymer, 416
- population genetics, 447
- probability density, 186
- probability measure, 186
- product measure, 187
- Prohorov metric, 192
- projective tree moduli space, 377
- PTEN*, 370, 385
- pushforward measure, 188
- pushout, 85
- Python Mapper, 445
- quasi-isomorphism, 110
- quotient group, 68
- quotient map, 56
- quotient space, 57
- quotient topology, 56
- rank, 71
- RAS*, 361
- RBI*, 370
- reciprocal translocation, 371
- recombination hotspots, 331
- Reeb graph, 119
- Reeb space, 119
 - categorical, 120
- reference curve comparison, 429
- retrovirus, 304
- Riemannian manifold, 117
 - curvature, 117
 - volume form, 117
- ring, 75
- Rips zigzag, 153
- Ripser, 445
- Rivet, 445
- Rous, P., 360
- Rubin, H., 360
- Russell's paradox, 81
- sampling
 - geometric, 189
- scalar multiplication, 76
- sequence, 33
 - Cauchy, 33
 - converging (metric spaces), 33
 - converging (topological spaces), 42
- sequence alignment, 454
- sets, 25
 - Cartesian product, 26
 - closed, 38
 - countable, 28

500

finite, 25
infinite, 25
isomorphic, 27
open, 38
subset, 26
uncountable, 28
simplex, 93
simplicial complex
 abstract, 97
 geometric, 94
 geometric random, 227
 geometric realization of, 95
 multifiltered, 154
 orientation of vertices, 103
simplicial map, 96
single cell, 399
 amplification, 400
 isolation, 400
 technologies, 400
single-linkage clustering, 137
sliding window, 433
Smith normal form, 108
Smith-Waterman algorithm, 455
smoothed Euler characteristic curve, 230
smoothed Euler characteristic transform, 231
SNE, 253
 t-SNE, 254
space, 30
 metric, 30
 quotient, 57
 topological, 38
span, 77
species trees, 333
split decomposition, 465
split networks, 463
stochastic neighbor embedding, 253
stratified space, 118
subcategory, 82
subgroup, 67
subset, 26
 dense, 35
sup, 36
supercoiling, 417
support, 186
supremum, 36
surjective, 26

Tajima-Nei distance, 461
tangent space, 116
taxonomy, 273
TDA Mapper, 446
TDA R package, 445
TDGF1, 410
Temin, H., 360
temozolomide (TMZ), 377
terminal object, 83
time series, 423
topoisomerases, 418

Index

topological ARGs, 342, 344
topological associated domains, 412, 420
topological data analysis, 415
topological obstruction to phylogeny, 285
topological space, 38
 Cartesian product, 41
 compact, 43
 Hausdorff, 40
 homeomorphic, 46
 homotopic, 49
 homotopy equivalent, 50
 path-connected, 45
 second countable, 40
topology, 39
 base for, 39
 discrete, 39
 indiscrete, 39
 metric, 39
 product, 41
 quotient, 56
 subspace, 41
torsion, 71
totally bounded, 36
TP53, 370, 385, 434
transcription factories, 415
transcription factors, 376
triangle inequality, 30
trophectoderm, 408
tumor suppressors, 361
tumors, 356

unit, 75

Varmus, H., 360
vector space, 76
 basis, 77
 dimension, 79
 inner product, 116
 multifiltered, 154
 norm, 117
 subspace, 77
vertices, 95
Vietoris-Rips complex, 126
virus
 Epstein-Barr virus (HHV4), 374
 hepatitis B virus, 375
 hepatitis C virus, 375
 human Kaposi-associated herpesvirus (HHV8), 375
 human papilloma virus (HPV), 375
 human T-cell lymphotropic virus type 1, 375
 Merkel cell carcinoma virus (MCCV), 375
 Rous sarcoma virus, 360
 SV40, 361
viSNE, 256
visual cortex, 436
volume measure, 187

Wasserstein distance, 147	Wright-Fisher model, 448
Watterson estimator, 452	
wavelets, 429	
weak convergence, 191	xenografts, 405
weak equivalence, 74	
weak homotopy equivalence, 74	Yamagiwa, K., 357
weak homotopy type, 74	
weak witness, 159	
Weinberg, R., 361	zigzag diagram, 149
West Nile virus, 315	zigzag module, 149
Wigler, M., 361	interval, 151
witness complex, 160	zigzag persistence, 310
Woese, C., 277	zigzag submodule, 150
Wright, S., 447	zona pellucida, 408