

The R Series

Spatial Data Science

With Applications in R



Edzer Pebesma
Roger Bivand



CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

bf67
4050-
สำนักหอสมุด มหาวิทยาลัยเชียงใหม่

บ 167055/9
อ 1257440
จ 226858/9

Spatial Data Science

With Applications in R

Edzer Pebesma
Roger Bivand



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an informa business
A CHAPMAN & HALL BOOK

Table of contents

Preface	xi
I Spatial Data	1
1 Getting Started	5
1.1 A first map	5
1.2 Coordinate reference systems	7
1.3 Raster and vector data	8
1.4 Raster types	10
1.5 Time series, arrays, data cubes	11
1.6 Support	12
1.7 Spatial data science software	13
1.7.1 GDAL	14
1.7.2 PROJ	14
1.7.3 GEOS and s2geometry	14
1.7.4 NetCDF, udunits2, liblwgeom	14
1.8 Exercises	15
2 Coordinates	17
2.1 Quantities, units, datum	17
2.2 Ellipsoidal coordinates	18
2.2.1 Spherical or ellipsoidal coordinates	19
2.2.2 Projected coordinates, distances	21
2.2.3 Bounded and unbounded spaces	22
2.3 Coordinate reference systems	22
2.4 PROJ and mapping accuracy	23
2.5 WKT-2	26
2.6 Exercises	27
3 Geometries	29
3.1 Simple feature geometries	29
3.1.1 The big seven	29
3.1.2 Simple and valid geometries, ring direction	31
3.1.3 Z and M coordinates	31
3.1.4 Empty geometries	32
3.1.5 Ten further geometry types	32
3.1.6 Text and binary encodings	33

3.2	Operations on geometries	34
3.2.1	Unary predicates	34
3.2.2	Binary predicates and DE-9IM	34
3.2.3	Unary measures	36
3.2.4	Binary measures	37
3.2.5	Unary transformers	37
3.2.6	Binary transformers	38
3.2.7	N-ary transformers	38
3.3	Precision	39
3.4	Coverages: tessellations and rasters	40
3.4.1	Topological models	40
3.4.2	Raster tessellations	41
3.5	Networks	42
3.6	Exercises	42
4	Spherical Geometries	45
4.1	Straight lines	45
4.2	Ring direction and full polygon	45
4.3	Bounding box, rectangle, and cap	46
4.4	Validity on the sphere	47
4.5	Exercises	47
5	Attributes and Support	49
5.1	Attribute-geometry relationships and support	50
5.2	Aggregating and summarising	52
5.3	Area-weighted interpolation	54
5.3.1	Spatially extensive and intensive variables	54
5.3.2	Dasymetric mapping	55
5.3.3	Support in file formats	55
5.4	Up- and Downscaling	56
5.5	Exercises	57
6	Data Cubes	59
6.1	A four-dimensional data cube	59
6.2	Dimensions, attributes, and support	60
6.2.1	Regular dimensions, GDAL's geotransform	62
6.2.2	Support along cube dimensions	62
6.3	Operations on data cubes	63
6.3.1	Slicing a cube: filter	63
6.3.2	Applying functions to dimensions	64
6.3.3	Reducing dimensions	64
6.4	Aggregating raster to vector cubes	65
6.5	Switching dimension with attributes	67
6.6	Other dynamic spatial data	67
6.7	Exercises	68

II	R for Spatial Data Science	71
7	Introduction to sf and stars	75
7.1	Package sf	75
7.1.1	Creation	76
7.1.2	Reading and writing	77
7.1.3	Subsetting	78
7.1.4	Binary predicates	78
7.1.5	tidyverse	80
7.2	Spatial joins	81
7.2.1	Sampling, gridding, interpolating	82
7.3	Ellipsoidal coordinates	82
7.4	Package stars	84
7.4.1	Reading and writing raster data	85
7.4.2	Subsetting stars data cubes	86
7.4.3	Cropping	88
7.4.4	Redimensioning and combining stars objects	89
7.4.5	Extracting point samples, aggregating	91
7.4.6	Predictive models	92
7.4.7	Plotting raster data	93
7.4.8	Analysing raster data	93
7.4.9	Curvilinear rasters	98
7.4.10	GDAL utils	98
7.5	Vector data cube examples	99
7.5.1	Example: aggregating air quality time series	99
7.5.2	Example: Bristol origin-destination data cube	102
7.5.3	Tidy array data	108
7.5.4	File formats for vector data cubes	109
7.6	Raster-to-vector, vector-to-raster	109
7.6.1	Vector-to-raster	109
7.7	Coordinate transformations and conversions	110
7.7.1	st_crs	110
7.7.2	st_transform , sf_project	111
7.7.3	sf_proj_info	112
7.7.4	Datum grids, proj.db, cdn.proj.org, local cache	112
7.7.5	Transformation pipelines	113
7.7.6	Axis order and direction	115
7.8	Transforming and warping rasters	116
7.9	Exercises	117
8	Plotting spatial data	119
8.1	Every plot is a projection	119
8.1.1	What is a good projection for my data?	120
8.2	Plotting points, lines, polygons, grid cells	121
8.2.1	Colours	121

8.2.2	Colour breaks: <code>classInt</code>	122
8.2.3	Graticule and other navigation aids	123
8.3	Base plot	123
8.3.1	Adding to plots with legends	123
8.3.2	Projections in base plots	125
8.3.3	Colours and colour breaks	125
8.4	Maps with <code>ggplot2</code>	125
8.5	Maps with <code>tmap</code>	127
8.6	Interactive maps: <code>leaflet</code> , <code>mapview</code> , <code>tmap</code>	128
8.7	Exercises	129
9	Large data and cloud native	131
9.1	Vector data: <code>sf</code>	132
9.1.1	Reading from local disk	132
9.1.2	Reading from databases, <code>dbplyr</code>	133
9.1.3	Reading from online resources or web services	134
9.1.4	APIs, <code>OpenStreetMap</code>	134
9.1.5	<code>GeoParquet</code> and <code>GeoArrow</code>	135
9.2	Raster data: <code>stars</code>	136
9.2.1	<code>stars</code> proxy objects	137
9.2.2	Operations on proxy objects	139
9.2.3	Remote raster resources	139
9.3	Very large data cubes	140
9.3.1	Finding and processing assets	140
9.3.2	Cloud native storage: <code>Zarr</code>	140
9.3.3	APIs for data: <code>GEE</code> , <code>openEO</code>	141
9.4	Exercises	142
III	Models for Spatial Data	143
10	Statistical modelling of spatial data	147
10.1	Mapping with non-spatial regression and ML models	148
10.2	Support and statistical modelling	149
10.3	Time in predictive models	150
10.4	Design-based and model-based inference	151
10.5	Predictive models with coordinates	152
10.6	Exercises	153
11	Point Pattern Analysis	155
11.1	Observation window	156
11.2	Coordinate reference systems	160
11.3	Marked point patterns, points on linear networks	161
11.4	Spatial sampling and simulating a point process	163
11.5	Simulating points on the sphere	164
11.6	Exercises	164

12 Spatial Interpolation	165
12.1 A first dataset	165
12.2 Sample variogram	167
12.3 Fitting variogram models	169
12.4 Kriging interpolation	171
12.5 Areal means: block kriging	171
12.6 Conditional simulation	173
12.7 Trend models	174
12.7.1 A population grid	175
12.8 Exercises	179
13 Multivariate and Spatiotemporal Geostatistics	181
13.1 Preparing the air quality dataset	181
13.2 Multivariable geostatistics	183
13.3 Spatiotemporal geostatistics	184
13.3.1 A spatiotemporal variogram model	184
13.3.2 Irregular space time data	189
13.4 Exercises	190
14 Proximity and Areal Data	191
14.1 Representing proximity in <code>spdep</code>	192
14.2 Contiguous neighbours	194
14.3 Graph-based neighbours	197
14.4 Distance-based neighbours	199
14.5 Weights specification	204
14.6 Higher order neighbours	206
14.7 Exercises	208
15 Measures of Spatial Autocorrelation	209
15.1 Measures and process misspecification	209
15.2 Global measures	211
15.2.1 Join-count tests for categorical data	212
15.2.2 Moran's I	214
15.3 Local measures	216
15.3.1 Local Moran's I_i	218
15.3.2 Local Getis-Ord G_i	225
15.3.3 Local Geary's C_i	226
15.3.4 The <code>rgeoda</code> package	229
15.4 Exercises	231
16 Spatial Regression	233
16.1 Markov random field and multilevel models	233
16.1.1 Boston house value dataset	235
16.2 Multilevel models of the Boston dataset	237
16.2.1 IID random effects with <code>lme4</code>	238
16.2.2 IID and CAR random effects with <code>hglm</code>	238

16.2.3	IID and ICAR random effects with R2BayesX	239
16.2.4	IID, ICAR and Leroux random effects with INLA	240
16.2.5	ICAR random effects with mgcv::gam()	241
16.2.6	Upper-level random effects: summary	241
16.3	Exercises	243
17	Spatial Econometrics Models	245
17.1	Spatial econometric models: definitions	245
17.2	Maximum likelihood estimation in spatialreg	248
17.2.1	Boston house value dataset examples	249
17.3	Impacts	252
17.4	Predictions	254
17.5	Exercises	256
A	Older R Spatial Packages	259
A.1	Retiring rgdal and rgeos	259
A.2	Links and differences between sf and sp	259
A.3	Migration code and packages	260
A.4	Package raster and terra	260
B	R Basics	263
B.1	Pipes	263
B.2	Data structures	264
B.2.1	Homogeneous vectors	264
B.2.2	Heterogeneous vectors: list	265
B.2.3	NULL and removing list elements	267
B.2.4	Attributes	267
B.2.5	The names attributes	270
B.2.6	Using structure	271
B.3	Dissecting a MULTIPOLYGON	272
	References	277
	Index	291
	Index of functions	299