

Use R!

Deepayan Sarkar

Lattice

**Multivariate Data Visualization
with R**

Contents

Preface	vii
1 Introduction	1
1.1 Multipanel conditioning	2
1.1.1 A histogram for every group	2
1.1.2 The Trellis call	3
1.1.3 Kernel density plots	4
1.2 Superposition	5
1.3 The “ <i>trellis</i> ” object	6
1.3.1 The missing Trellis display	7
1.3.2 Arranging multiple Trellis plots	7
1.4 Looking ahead	7

Part I Basics

2 A Technical Overview of lattice	13
2.1 Basic usage	13
2.1.1 The Trellis formula	13
2.1.2 The data argument	14
2.1.3 Conditioning	14
2.1.4 Shingles	15
2.2 Dimension and physical layout	16
2.2.1 Aspect ratio	19
2.2.2 Layout	20
2.2.3 Fine-tuning the layout: between and skip	24
2.3 Grouped displays	24
2.4 Annotation: Captions, labels, and legends	26
2.4.1 More on legends	26
2.5 Graphing the data	28
2.5.1 Scales and axes	28

2.5.2	The panel function	30
2.5.3	The panel function demystified	31
2.6	Return value	33
3	Visualizing Univariate Distributions	35
3.1	Density Plot	35
3.2	Large datasets	37
3.3	Histograms	39
3.4	Normal Q–Q plots	40
3.4.1	Normality and the Box–Cox transformation	42
3.4.2	Other theoretical Q–Q plots	43
3.5	The empirical CDF	44
3.6	Two-sample Q–Q plots	44
3.7	Box-and-whisker plots	47
3.7.1	Violin plots	47
3.8	Strip plots	50
3.9	Coercion rules	52
3.10	Discrete distributions	53
3.11	A note on the formula interface	54
4	Displaying Multiway Tables	55
4.1	Cleveland dot plot	55
4.2	Bar chart	57
4.2.1	Manipulating order	61
4.2.2	Bar charts and discrete distributions	63
4.3	Visualizing categorical data	65
5	Scatter Plots and Extensions	67
5.1	The standard scatter plot	67
5.2	Advanced indexing using <code>subscripts</code>	71
5.3	Variants using the <code>type</code> argument	75
5.3.1	Superposition and <code>type</code>	79
5.4	Scatter-plot variants for large data	82
5.5	Scatter-plot matrix	84
5.5.1	Interacting with scatter-plot matrices	86
5.6	Parallel coordinates plot	87
6	Trivariate Displays	91
6.1	Three-dimensional scatter plots	91
6.1.1	Dynamic manipulation versus stereo viewing	95
6.1.2	Variants and panel functions	96
6.2	Surfaces and two-way tables	98
6.2.1	Data preparation	99
6.2.2	Visualizing surfaces	102
6.2.3	Visualizing discrete array data	105

6.3	Theoretical surfaces	110
6.3.1	Parameterized surfaces	111
6.4	Choosing a palette for false-color plots	113

Part II Finer Control

7	Graphical Parameters and Other Settings	119
7.1	The parameter system	119
7.1.1	Themes	120
7.1.2	Devices	120
7.1.3	Initializing a graphics device	121
7.1.4	Reading and modifying a theme	122
7.1.5	Usage and alternative forms	125
7.1.6	The <code>par.settings</code> argument	125
7.2	Available graphical parameters	126
7.2.1	Nonstandard settings	129
7.3	Non-graphical options	131
7.3.1	Argument defaults	131
7.4	Making customizations persistent	131
8	Plot Coordinates and Axis Annotation	133
8.1	Packets and the <code>prepanel</code> function	133
8.2	The <code>scales</code> argument	134
8.2.1	Relation	134
8.2.2	Axis annotation: Ticks and labels	135
8.2.3	Defaults	138
8.2.4	Three-dimensional displays: <code>cloud()</code> and <code>wireframe()</code>	139
8.3	Limits and aspect ratio	140
8.3.1	The <code>prepanel</code> function revisited	140
8.3.2	Explicit specification of limits	141
8.3.3	Choosing aspect ratio by banking	143
8.4	Scale components and the <code>axis</code> function	144
8.4.1	Components	144
8.4.2	Axis	148
9	Labels and Legends	151
9.1	Labels	151
9.2	Legends	152
9.2.1	Legends as <code>grid</code> graphical objects	152
9.2.2	The <code>colorkey</code> argument	155
9.2.3	The <code>key</code> argument	156
9.2.4	The problem with settings, and the <code>auto.key</code> argument	158
9.2.5	Dropping unused levels from <code>groups</code>	159
9.2.6	A more complicated example	159

9.2.7	Further control: The <code>legend</code> argument	161
9.3	Page annotation	162
10	Data Manipulation and Related Topics	165
10.1	Nonstandard evaluation	165
10.2	The extended formula interface	166
10.3	Combining data sources with <code>make.groups()</code>	170
10.4	Subsetting	173
10.4.1	Dropping of factor levels	176
10.5	Shingles and related utilities	177
10.5.1	Coercion to factors and shingles	182
10.5.2	Using shingles for axis breaks	183
10.5.3	Cut-and-stack plots	184
10.6	Ordering levels of categorical variables	187
10.7	Controlling the appearance of strips	193
10.8	An Example Revisited	198
11	Manipulating the “<i>trellis</i>” Object	201
11.1	Methods for “ <i>trellis</i> ” objects	201
11.2	The <code>plot()</code> , <code>print()</code> , and <code>summary()</code> methods	202
11.3	The <code>update()</code> method and <code>trellis.last.object()</code>	206
11.4	Tukey mean–difference plot	208
11.5	Specialized manipulations	210
11.6	Manipulating the display	211
12	Interacting with Trellis Displays	215
12.1	The traditional graphics model	215
12.1.1	Interaction	216
12.2	Viewports, <code>trellis.vpname()</code> , and <code>trellis.focus()</code>	216
12.3	Interactive additions	217
12.4	Other uses	223

Part III Extending Trellis Displays

13	Advanced Panel Functions	229
13.1	Preliminaries	229
13.1.1	Building blocks for panel functions	229
13.1.2	Accessor functions	231
13.1.3	Arguments	232
13.2	A toy example: Hypotrochoids and hypocycloids	232
13.3	Some more examples	235
13.3.1	An alternative density estimate	235
13.3.2	A modified box-and-whisker plot	237
13.3.3	Corrgrams as customized level plots	238

13.4	Three-dimensional projections	241
13.5	Maps	242
13.5.1	A simple projection scheme	244
13.5.2	Maps with conditioning	245
14	New Trellis Displays	247
14.1	$S3$ methods	248
14.2	$S4$ methods	249
14.3	New functions	251
14.3.1	A complete example: Multipanel pie charts	252
	References	255
	Index	259