Bayes Linear Statistics

Theory and Methods

Michael Goldstein and David Wooff
Durham University, UK

Contents

Preface xvii

1 The Bayes linear approach 1
- 1.1 Combining beliefs with data . . . 2
- 1.2 The Bayesian approach . . . 3
- 1.3 Features of the Bayes linear approach . . . 6
- 1.4 Example . . . 7
  - 1.4.1 Expectation, variance, and standardization . . . 8
  - 1.4.2 Prior inputs . . . 8
  - 1.4.3 Adjusted expectations . . . 9
  - 1.4.4 Adjusted versions . . . 10
  - 1.4.5 Adjusted variances . . . 10
  - 1.4.6 Checking data inputs . . . 11
  - 1.4.7 Observed adjusted expectations . . . 12
  - 1.4.8 Diagnostics for adjusted beliefs . . . 12
  - 1.4.9 Further diagnostics for the adjusted versions . . . 12
  - 1.4.10 Summary of basic adjustment . . . 13
  - 1.4.11 Diagnostics for collections . . . 14
  - 1.4.12 Exploring collections of beliefs via canonical structure . 17
  - 1.4.13 Modifying the original specifications . . . 19
  - 1.4.14 Repeating the analysis for the revised model . . . 20
  - 1.4.15 Global analysis of collections of observations . . . 22
  - 1.4.16 Partial adjustments . . . 24
  - 1.4.17 Partial diagnostics . . . 27
  - 1.4.18 Summary . . . 29
- 1.5 Overview . . . 30

2 Expectation 33
- 2.1 Expectation as a primitive . . . 33
- 2.2 Discussion: expectation as a primitive . . . 35
- 2.3 Quantifying collections of uncertainties . . . 37
- 2.4 Specifying prior beliefs . . . 39
  - 2.4.1 Example: oral glucose tolerance test . . . 39

2.5 Qualitative and quantitative prior specification 41
2.6 Example: qualitative representation of uncertainty 42
2.6.1 Identifying the quantities of interest 42
2.6.2 Identifying relevant prior information 42
2.6.3 Sources of variation 43
2.6.4 Representing population variation 44
2.6.5 The qualitative representation 44
2.6.6 Graphical models 46
2.7 Example: quantifying uncertainty 48
2.7.1 Prior expectations 48
2.7.2 Prior variances 49
2.7.3 Prior covariances 51
2.7.4 Summary of belief specifications 52
2.8 Discussion: on the various methods for assigning expectations . . 52

3 Adjusting beliefs **55**
3.1 Adjusted expectation 55
3.2 Properties of adjusted expectation 56
3.3 Adjusted variance 57
3.4 Interpretations of belief adjustment 58
3.5 Foundational issues concerning belief adjustment 60
3.6 Example: one-dimensional problem 63
3.7 Collections of adjusted beliefs 64
3.8 Examples 65
3.8.1 Algebraic example 65
3.8.2 Oral glucose tolerance test 69
3.8.3 Many oral glucose tolerance tests 73
3.9 Canonical analysis for a belief adjustment 75
3.9.1 Canonical directions for the adjustment 75
3.9.2 The resolution transform 77
3.9.3 Partitioning the resolution 79
3.9.4 The reverse adjustment 80
3.9.5 Minimal linear sufficiency 81
3.9.6 The adjusted belief transform matrix 81
3.10 The geometric interpretation of belief adjustment 82
3.11 Examples 84
3.11.1 Simple one-dimensional problem 84
3.11.2 Algebraic example 84
3.11.3 Oral glucose tolerance test 90
3.12 Further reading 93

4 The observed adjustment **95**
4.1 Discrepancy 95
4.1.1 Discrepancy for a collection 96

4.1.2 Evaluating discrepancy over a basis 98
4.1.3 Discrepancy for quantities with variance zero 98
4.2 Properties of discrepancy measures 98
4.2.1 Evaluating the discrepancy vector over a basis 100
4.3 Examples 100
4.3.1 Simple one-dimensional problem 100
4.3.2 Detecting degeneracy 100
4.3.3 Oral glucose tolerance test 102
4.4 The observed adjustment 104
4.4.1 Adjustment discrepancy 104
4.4.2 Adjustment discrepancy for a collection 105
4.4.3 Maximal discrepancy 106
4.4.4 Construction over a basis 107
4.4.5 Partitioning the discrepancy 108
4.5 Examples 108
4.5.1 Simple one-dimensional problem 108
4.5.2 Oral glucose tolerance test 109
4.6 The size of an adjustment 113
4.6.1 The size of an adjustment for a collection 113
4.7 The bearing for an adjustment 114
4.7.1 Construction via a basis 115
4.7.2 Representing discrepancy vectors as bearings 115
4.8 Joint bearings 116
4.9 Size diagnostics 116
4.10 Geometric interpretation 117
4.11 Linear likelihood 118
4.12 Examples 119
4.12.1 Algebraic example 119
4.12.2 Oral glucose tolerance test 120

5 Partial Bayes linear analysis **125**
5.1 Partial adjustment 125
5.2 Partial variance 127
5.3 Partial resolution transforms 128
5.4 Relative belief adjustment 129
5.5 Example: oral glucose tolerance test 130
5.5.1 Performing an initial adjustment 131
5.5.2 Partial resolved variances 132
5.5.3 Partial canonical directions 132
5.5.4 Deducing changes for other linear combinations 133
5.5.5 Relative belief adjustment 133
5.5.6 Withdrawing quantities from the adjustment 134
5.6 Partial bearings 135
5.7 Partial data size 137

5.8 Bearing and size for a relative adjustment 137
5.9 Path correlation 138
5.10 Example: oral glucose tolerance test 139
5.10.1 The initial observed adjustment 139
5.10.2 Observed partial expectations 140
5.10.3 The size of the partial adjustment 141
5.10.4 The bearing for the partial adjustment 142
5.10.5 The path correlation for the partial adjustment 143
5.11 Sequential adjustment 144
5.11.1 The data trajectory 144
5.12 The canonical trajectory 145
5.13 Detection of systematic bias 146
5.14 Examples 147
5.14.1 Anscombe data sets 147
5.14.2 Regression with correlated responses 153
5.15 Bayes linear sufficiency and belief separation 166
5.16 Properties of generalized conditional independence 168
5.17 Properties of belief separation 169
5.18 Example: regression with correlated responses 172
5.18.1 Exploiting separation 172
5.18.2 Heart of the transform 173
5.19 Further reading 176

6 Exchangeable beliefs **177**
6.1 Exchangeability 177
6.2 Coin tossing 180
6.3 Exchangeable belief structures 183
6.4 The representation theorem 185
6.5 Finite exchangeability 188
6.6 Example: oral glucose tolerance test 189
6.7 Example: analysing exchangeable regressions 191
6.7.1 Introduction 191
6.7.2 Error structure and specifications 192
6.7.3 Regression coefficient specifications 193
6.7.4 Structural implications 194
6.8 Adjusting exchangeable beliefs 194
6.9 Predictive sufficiency for exchangeable models 195
6.10 Bayes linear sufficiency for sample means 196
6.11 Belief adjustment for scalar exchangeable quantities 197
6.12 Canonical structure for an exchangeable adjustment 198
6.12.1 Standard form for the adjustment 200
6.12.2 Further properties of exchangeable adjustments 201
6.13 Algebraic example 202
6.13.1 Representation 203

6.13.2 Coherence 203
6.13.3 Bayes linear sufficiency 204
6.14 Example: adjusting exchangeable regressions 205
6.14.1 Bayes linear sufficiency 205
6.14.2 Adjustment 206
6.14.3 Resolution transforms 208
6.14.4 Resolution partition for exchangeable cases 210
6.14.5 Data diagnostics 211
6.14.6 Sample size choice 212
6.14.7 Adjustment for an equivalent linear space 214
6.14.8 Data diagnostics for an equivalent linear space 214
6.14.9 Compatibility of data sources 215
6.15 Predictive adjustment 218
6.16 Example: oral glucose tolerance test 220
6.16.1 Context of exchangeability 220
6.16.2 Mean component adjustment 220
6.16.3 Variance reduction for a predictive adjustment 221
6.16.4 Observed exchangeable adjustments 223
6.16.5 Path diagnostics 226
6.17 Example: predictive analysis for exchangeable regressions 228
6.17.1 Choice of canonical directions 230
6.18 Further reading 231

7 Co-exchangeable beliefs **233**
7.1 Respecting exchangeability 233
7.2 Adjustments respecting exchangeability 234
7.3 Example: simple algebraic problem 235
7.3.1 Coherence 236
7.3.2 Resolution transform 236
7.4 Co-exchangeable adjustments 238
7.5 Example: analysing further exchangeable regressions 240
7.5.1 The resolution envelope 243
7.6 Example: exchangeability in a population dynamics experiment . 244
7.6.1 Model 244
7.6.2 Specifications 248
7.6.3 Issues 251
7.6.4 Analysis 251

8 Learning about population variances **265**
8.1 Assessing a population variance with known population mean . . 265
8.2 Assessing a population variance with unknown population mean 266
8.3 Choice of prior values 268
8.4 Example: oral glucose tolerance test 271

8.5 Adjusting the population residual variance in multiple linear regression: uncorrelated errors . . . 273
8.5.1 Sample information . . . 274
8.5.2 Choice of prior values . . . 276
8.6 Example: Anscombe data sets . . . 276
8.7 Adjusting the population residual variance in multiple linear regression: correlated errors . . . 277
8.8 Example: regression with correlated responses . . . 278
8.9 Example: analysing exchangeable regressions . . . 280
8.10 Adjusting a collection of population variances and covariances . . . 282
8.11 Direct adjustment for a population variance matrix . . . 283
8.12 Example: regression with correlated responses . . . 284
8.13 Separating direct adjustment for population variances and for correlation structure . . . 285
8.13.1 Assessing the equivalent sample size . . . 286
8.14 Example: oral glucose tolerance test . . . 287
8.15 Two-stage Bayes linear analysis . . . 288
8.16 Example: oral glucose tolerance test . . . 290
8.17 Example: analysing exchangeable regressions . . . 290
8.18 Further reading . . . 292

9 Belief comparison **293**
9.1 Comparing variance specifications . . . 294
9.1.1 Rank-degenerate case . . . 296
9.1.2 Comparison of orthogonal subspaces . . . 298
9.2 Example: variance comparison . . . 298
9.2.1 Canonical structure for the comparison . . . 299
9.2.2 Consistency checks . . . 301
9.2.3 Comparisons for further constructed quantities . . . 301
9.2.4 Construction of specifications . . . 302
9.3 Comparing many variance specifications . . . 302
9.4 Example: comparing some simple nested hypotheses . . . 304
9.5 General belief transforms . . . 306
9.5.1 General belief transforms . . . 306
9.5.2 Properties of general belief transforms . . . 307
9.5.3 Adjusted belief transforms as general belief transforms . 309
9.5.4 Example: adjustment of exchangeable structures . . . 310
9.5.5 Example: analysing exchangeable regressions . . . 311
9.6 Comparing expectations and variances . . . 312
9.7 Geometric interpretation . . . 314
9.8 Residual forms for mean and variance comparisons . . . 315
9.8.1 Rank-degenerate case . . . 317

9.9 The observed comparison 318
9.9.1 Combined directions 319
9.10 Example: mean and variance comparison 320
9.10.1 The observed comparison 323
9.11 Graphical comparison of specifications 324
9.11.1 Belief comparison diagram 325
9.11.2 The observed comparison 327
9.11.3 Combining information 329
9.11.4 Residual belief comparison diagrams 329
9.12 Example: exchangeable regressions 331
9.12.1 Basic canonical analysis 332
9.12.2 Mean and residual comparisons 333
9.13 Comparisons for exchangeable structures 337
9.13.1 The observed comparison 338
9.13.2 Example: exchangeable regressions 340
9.14 Example: fly population dynamics 342
9.14.1 Differences for the mean part of the average 343
9.14.2 Differences for the residual part of the average 343
9.14.3 Differences for the residual part of the average 344
9.15 Assessing robustness of specifications 346
9.15.1 Sensitivity analyses for expectations 347
9.15.2 Example: robustness analysis for exchangeable regressions 349
9.15.3 Sensitivity analyses for variances 350
9.15.4 Example: robustness analysis for variance specifications . 351
9.16 Further reading 353

10 Bayes linear graphical models **355**
10.1 Directed graphical models 356
10.1.1 Construction via statistical models 358
10.2 Operations on directed graphs 358
10.3 Quantifying a directed graphical model 361
10.4 Undirected graphs 362
10.4.1 Node removal via the moral graph 364
10.5 Example 364
10.5.1 Plates for duplicated structures 367
10.5.2 Reading properties from the diagram 367
10.5.3 Alternative diagrams 368
10.5.4 Diagrams for inference and prediction 370
10.6 Displaying the flow of information 372
10.6.1 Node shading 373
10.6.2 Arc labelling 374
10.6.3 Tracking information as it is received 376
10.6.4 Example 377

10.7 Displaying diagnostic information . . . 383
10.7.1 Node diagnostics . . . 385
10.7.2 Arc diagnostics . . . 387
10.7.3 Showing implications across all nodes . . . 388
10.7.4 Interpreting diagnostic warnings . . . 388
10.7.5 Example: inference and prediction . . . 389
10.8 Local computation: directed trees . . . 395
10.8.1 Propagation . . . 397
10.8.2 Example . . . 398
10.9 Junction trees . . . 399
10.10 Sequential local computation on the junction tree . . . 400
10.11 Example: correlated regressions . . . 402
10.12 Example: problems of prediction in a large brewery . . . 402
10.12.1 Problem summary . . . 402
10.12.2 Identifying the quantities of interest . . . 403
10.12.3 Modelling . . . 404
10.12.4 Initialization values and specifications . . . 406
10.12.5 Examining the generated model . . . 412
10.12.6 Basic adjustment . . . 414
10.12.7 Exploration via graphical models . . . 416
10.13 Local computation for global adjustment of the junction tree . . . 424
10.13.1 Merging separate adjustments . . . 425
10.13.2 The global adjustment algorithm . . . 427
10.13.3 Absorption of evidence . . . 427
10.14 Further reading . . . 429

11 Matrix algebra for implementing the theory 431
11.1 Basic definitions . . . 431
11.2 Covariance matrices and quadratic forms . . . 431
11.3 Generalized inverses . . . 432
11.3.1 Basic properties . . . 432
11.3.2 Computing the Moore–Penrose inverse . . . 432
11.3.3 Other properties of generalized inverses . . . 433
11.4 Multiplication laws . . . 434
11.5 Range and null space of a matrix . . . 435
11.6 Rank conditions . . . 436
11.7 Partitioned matrices . . . 436
11.7.1 Definiteness for a partitioned real symmetric matrix . . . 436
11.7.2 Generalized inverses for partitioned non-negative definite matrices . . . 437
11.8 Solving linear equations . . . 438
11.9 Eigensolutions to related matrices . . . 439

11.10 Maximizing a ratio of quadratic forms 440
11.11 The generalized eigenvalue problem 441
11.11.1 Introduction 441
11.11.2 The QZ algorithm 442
11.11.3 An alternative algorithm 442
11.11.4 An algorithm for $B - A$ non-negative definite 444
11.12 Direct products of matrices 447
11.12.1 The Helmert matrix 447
11.12.2 Direct products 448

12 Implementing Bayes linear statistics **451**
12.1 Introduction 451
12.2 Coherence of belief specifications 452
12.2.1 Coherence for a single collection 452
12.2.2 Coherence for two collections 452
12.2.3 Coherence for three collections 453
12.3 Consistency of data with beliefs 455
12.3.1 Consistency for a single collection 455
12.3.2 Consistency for a partitioned collection 456
12.4 Adjusted expectation 457
12.5 Adjusted and resolved variance 458
12.6 The resolved variance matrix 459
12.7 Matrix representations of the resolution transform 460
12.7.1 The symmetrized resolution transform matrix 461
12.7.2 The transform for the reverse adjustment 463
12.7.3 Inverses for the resolved variance matrix 464
12.7.4 Canonical quantities 465
12.7.5 Coherence via the resolution transform matrix 466
12.8 Assessing discrepant data 467
12.9 Consistency of observed adjustments 468
12.9.1 Partitioning the discrepancy 469
12.10 The bearing and size of adjustment 472
12.11 Partial adjustments 473
12.11.1 Partial and relative adjustment transforms 475
12.11.2 Calculating the partial bearing 475
12.12 Exchangeable adjustments 476
12.12.1 Notation 476
12.12.2 Coherence requirements for exchangeable adjustments 477
12.12.3 Data consistency 477
12.12.4 Pure exchangeable adjustments 477
12.12.5 General exchangeable adjustments 481

12.13 Implementing comparisons of belief 483
12.13.1 Expectation comparisons 483
12.13.2 Comparison of exchangeable beliefs 483

A Notation **487**

B Index of examples **491**

C Software for Bayes linear computation **495**
C.1 [B/D] 495
C.2 BAYES-LIN 495

References **497**

Index **503**