

**Essential Mathematics for Political  
and Social Research**

Jeff Gill

*Cambridge University Press*

Cambridge, New York

# Contents

<i>List of Tables</i>	<i>page</i> vii
<i>List of Illustrations</i>	viii
<i>List of Examples</i>	xi
<i>Preface</i>	xv
<b>1 The Basics</b>	<b>1</b>
1.1 Introduction	1
1.2 Notation, Notation, Notation	2
1.3 Basic Terms	5
1.3.1 Indexing and Referencing	9
1.3.2 Specific Mathematical Use of Terms	11
1.4 Functions and Equations	14
1.4.1 Applying Functions: the Equation of a Line	23
1.4.2 The Factorial Function	27
1.4.3 The Modulo Function	29
1.5 Polynomial Functions	30
1.6 Logarithms and Exponents	31

1.7	New Terminology	38
1.8	Chapter Appendix: Its All Greek to Me	46
<b>2</b>	<b>Analytic Geometry</b>	<b>48</b>
2.1	Introduction: the Width of a Circle	48
2.2	Radian Measurement and Polar Coordinates	48
2.3	What is Trigonometry?	52
2.3.1	Radian Measures for Trigonometric Functions	56
2.3.2	Conic Sections and Some Analytical Geometry	60
2.4	New Terminology	73
<b>3</b>	<b>Linear Algebra: Vectors, Matrices, and Operations</b>	<b>81</b>
3.1	Motivation	81
3.2	Working with Vectors	82
3.2.1	Vector Norms	90
3.3	So What is the Matrix?	96
3.3.1	Some Special Matrices	97
3.4	Controlling the Matrix	102
3.5	Matrix Transposition	110
3.6	Advanced Topics	112
3.6.1	Special Matrix Forms	112
3.6.2	Vectorization of Matrices	113
3.7	New Terminology	117
<b>4</b>	<b>Linear Algebra Continued: Matrix Structure</b>	<b>124</b>
4.1	Objectives	124
4.2	Space and Time	125
4.3	The Trace and Determinant of a Matrix	132
4.4	Matrix Rank	139
4.5	Matrix Norms	142

4.6	Matrix Inversion	143
4.7	Linear Systems of Equations	150
4.8	Eigen-Analysis of Matrices	152
4.9	Quadratic Forms and Descriptions	160
4.10	New Terminology	162
<b>5</b>	<b>Elementary Scalar Calculus</b>	<b>170</b>
5.1	Limits and Lines	170
5.2	Understanding Rates, Changes, and Derivatives	174
5.3	Derivative Rules for Common Functions	181
5.3.1	Basic Algebraic Rules for Derivatives	182
5.3.2	Derivatives of Logarithms and Exponents	188
5.3.3	L'Hospital's Rule	193
5.3.4	Applications: Rolle's Theorem and the Mean Value Theorem	195
5.4	Understanding Areas, Slices, and Integrals	197
5.4.1	Riemann Integrals	197
5.5	The Fundamental Theorem of Calculus	201
5.5.1	Integrating Polynomials with Antiderivatives	204
5.5.2	Indefinite Integrals	210
5.5.3	Integrals Involving Logarithms and Exponents	211
5.5.4	Integration by Parts	212
5.6	Additional Topics: Calculus of Trigonometric Functions	218
5.6.1	Derivatives of Trigonometric Functions	218
5.6.2	Integrals of Trigonometric Functions	219
5.7	New Terminology	220
<b>6</b>	<b>Additional Topics in Scalar and Vector Calculus</b>	<b>227</b>
6.1	Partial Derivatives	227
6.2	Derivatives and Partial Derivatives of Higher Order	231

6.3	Maxima, Minima, and Root Finding	233
6.3.1	Evaluating Zero-Derivative Points	238
6.3.2	Root Finding with Newton-Raphson	239
6.4	Multidimensional Integrals	242
6.5	Finite and Infinite Series	249
6.5.1	Convergence	251
6.6	The Calculus of Vector and Matrix Forms	259
6.6.1	Vector Function Notation	259
6.6.2	Differentiation and Integration of a Vector Function	261
6.7	New Terminology	265
<b>7</b>	<b>Probability Theory</b>	<b>272</b>
7.1	Introduction and Motivation	272
7.2	Counting Rules and Permutations	273
7.2.1	The Binomial Theorem and Pascal's Triangle	277
7.3	Sets and Operations on Sets	279
7.3.1	General Characteristics of Sets	281
7.3.2	A Special Set: the Empty Set	283
7.3.3	Operations on Sets	284
7.4	The Probability Function	293
7.5	Calculations with Probabilities	296
7.6	Conditional Probability and Bayes Law	299
7.6.1	Simpson's Paradox	303
7.7	Independence	305
7.8	New Terminology	310
<b>8</b>	<b>Random Variables</b>	<b>316</b>
8.1	Levels of Measurement	316
8.2	Distribution Functions	320
8.2.1	Randomness and Variables	323

8.2.2	Probability Mass Functions	325
8.2.3	Bernoulli Trials	326
8.2.4	Binomial Experiments	326
8.2.5	Poisson Counts	331
8.2.6	The Cumulative Distribution Function: Discrete Version	334
8.2.7	Probability Density Functions	337
8.2.8	Exponential and Gamma PDFs	338
8.2.9	Normal PDF	341
8.2.10	The Cumulative Distribution Function: Continuous Version	344
8.2.11	The Uniform Distributions	346
8.3	Expected Value and Variance	349
8.4	Some Handy Properties and Rules	351
8.4.1	Application: Craps in a Casino	352
8.5	Inequalities Based on Expected Values	357
8.6	Moments of a Distribution	359
8.7	Measures of Central Tendency: Mean, Median, and Mode	361
8.8	Measures of Dispersion: Variance, Standard Deviation, and MAD	365
8.9	Correlation and Covariance	367
8.10	New Terminology	371
<b>9</b>	<b>Markov Chains</b>	<b>382</b>
9.1	Defining Stochastic Processes and Markov Chains	382
9.1.1	The Markov Chain Kernel	387
9.1.2	The Stationary Distribution of a Markov Chain	391
9.2	Properties of Markov Chains	395
9.2.1	Homogeneity and Periodicity	395

9.2.2	Irreducibility	398
9.2.3	Recurrence	400
9.2.4	Stationarity and Ergodicity	404
9.2.5	Reversibility	408
9.3	New Terminology	411
	<i>References</i>	420

## List of Tables

7.1	ILLUSTRATION OF SIMPSON'S PARADOX: JOB TRAINING	303
8.1	BINOMIAL OUTCOMES AND PROBABILITIES	327
8.2	PERCENT EMPLOYMENT BY RACE, 1998	362
8.3	MEASURES OF DISPERSION, RACE IN AGENCIES	366
8.4	CORRELATION COEFFICIENT OF FAMILY UNIT VS. SPACE	370

## List of Illustrations

1.1	TWO IDEAL POINTS IN THE SENATE	6
1.2	A RELATION THAT IS NOT A FUNCTION	19
1.3	RELATING $x$ AND $f(x)$	20
1.4	THE CUBE LAW	23
1.5	PARALLEL AND PERPENDICULAR LINES	25
1.6	POVERTY AND READING TEST SCORES	26
1.7	NAUTILUS CHAMBERS	36
2.1	RADIAN MEASUREMENT OF ANGLES	49
2.2	POLAR COORDINATES	50
2.3	A GENERAL TRIGONOMETRIC SETUP	53
2.4	BASIC TRIGONOMETRIC FUNCTION IN RADIANS	57
2.5	VIEWS OF CONFLICT ESCALATION	59
2.6	CHARACTERISTICS OF A PARABOLA	62
2.7	PARABOLIC PRESIDENTIAL POPULARITY	63
2.8	CHARACTERISTICS OF AN ELLIPSE	65
2.9	MULTIDIMENSIONAL ISSUE PREFERENCE	66
2.10	CHARACTERISTICS OF A HYPERBOLA	68

2.11	EXPONENTIAL CURVES	70
2.12	DERIVED HYPERBOLA FORM	71
2.13	GENERAL HYPERBOLA FORM FOR DISCOUNTING	72
2.14	CONFLICT PROBABILITIES	77
2.15	DERIVATION OF THE DISTANCE	79
3.1	THE RIGHT-HAND-RULE ILLUSTRATED	89
4.1	VISUALIZING POINTS IN SPACES	125
4.2	VISUALIZING VECTORS IN SPACES	126
4.3	VECTOR PROJECTION, ADDITION, AND SUBTRACTION	127
4.4	RELATIVE INCOME AND SENIOR POVERTY, EU COUNTRIES	150
5.1	$f(x) = 3 - (x - 2)^2$	171
5.2	$f(x) = 1 + 1/x^2$	172
5.3	DESCRIBING THE RATE OF CHANGE	175
5.4	TANGENT LINES ON $f(x) = 5x - \frac{1}{20}x^2$	178
5.5	RIEMANN INTEGRATION	198
5.6	RIEMANN SUMS FOR $f(x) = x^2$ OVER $[0:1]$	202
5.7	INTEGRATING BY PIECES	206
5.8	PLACING THE MEDIAN VOTER	208
6.1	ILLUSTRATING THE INFLECTION POINT	234
6.2	POWER/ROLE GAP CHANGES INDUCED BY THE INFLECTION POINT	236
6.3	CONCAVE AND CONVEX PORTIONS OF A FUNCTION	237
6.4	ILLUSTRATION OF ITERATED INTEGRATION OF VOLUMES	244
6.5	INTEGRATING $f(x, y) = 2x^3y$ OVER PART OF THE UNIT CIRCLE	246
6.6	IRREGULAR REGION DOUBLE INTEGRAL	247
7.1	THREE SETS	288
7.2	THE MAPPING OF PROBABILITY FUNCTION	295
7.3	THEOREM OF TOTAL PROBABILITY ILLUSTRATED	297
7.4	MARGINAL CONTRIBUTION OF THIRD PARTIES	308

8.1	SELF-IDENTIFIED IDEOLOGY, ANES 2000	322
8.2	EXAMPLE BINOMIAL PROBABILITIES	328
8.3	POISSON PROBABILITIES OF WAR	334
8.4	BINOMIAL CDF PROBABILITIES, $n = 3, p = 0.5$	336
8.5	EXPONENTIAL PDF FORMS	338
8.6	GAMMA PDF FORMS	339
8.7	FITTING GAMMA DISTRIBUTIONS TO INCOME	340
8.8	NORMAL PDF FORMS	342
8.9	FITTING A NORMAL DISTRIBUTION TO LEGISLATIVE PARTICIPATION	343
8.10	PROBIT MODELS FOR PARTISAN VOTE-CHOICE	345

## List of Examples

1.1	Explaining Why People Vote.	3
1.2	Graphing Ideal Points in the Senate.	5
1.3	The “Cube Rule” in Votes to Seats.	21
1.4	Child Poverty and Reading Scores	25
1.5	Coalition Cabinet Formation.	28
2.1	An Expected Utility Model of Conflict Escalation.	56
2.2	Testing for a Circular Form	61
2.3	Presidential Support as a Parabola.	62
2.4	Elliptical Voting Preferences.	66
2.5	Hyperbolic Discounting in Evolutionary Psychology and Behavioral Economics.	68
3.1	Vector Addition	83
3.2	Vector Subtraction	83
3.3	Scalar Multiplication	84
3.4	Scalar Division	84
3.5	Inner Product	86
3.6	Cross Product	88

3.7	Outer Product	90
3.8	Votes in the House of Commons	92
3.9	The Political Economy of Taxation	93
3.10	Marriage Satisfaction	100
3.11	Matrix Addition	102
3.12	Matrix Subtraction	103
3.13	Scalar Multiplication	103
3.14	Scalar Division	103
3.15	LU Matrix Decomposition	107
3.16	Kronecker Product	114
4.1	Linear Transformation of Voter Assessments	129
4.2	Structural Shortcuts	136
4.3	Structural Equation Models	140
4.4	Calculating Regression Parameters	147
4.5	Solving Systems of Equations by Inversion	151
4.6	Basic Eigenanalysis	154
4.7	Analyzing Social Mobility with Eigens	156
4.8	LDU Decomposition	161
5.1	Radical Expression	173
5.2	Polynomial Ratio	173
5.3	Fractions and Exponents	174
5.4	Mixed Polynomials	174
5.5	Using Derivatives to Analyze Legislative Committee Size	179
5.6	Productivity and Interdependence in Work-Groups	186
5.7	Security Tradeoffs for Arms Versus Alliances	190
5.8	Analyzing an Infinite Series for Sociology Data	194
5.9	The Median Voter Theorem	206
5.10	Utility Functions From Voting	216
6.1	The Relationship Between Age and Earnings	229
6.2	Indexing Socio-Economic Status (SES)	230

6.3	Power Cycle Theory	235
6.4	Double Integral with Constant Limits	243
6.5	Revisiting Double Integral with Constant Limits	244
6.6	Double Integral with Dependent Limits	245
6.7	Complex Double Integral	247
6.8	The Nazification of German Sociological Genetics	249
6.9	Measuring Small Group Standing	249
6.10	Consecutive Integers	253
6.11	Telescoping sum	253
6.12	Geometric series	254
6.13	Repeating Values as a Geometric Series	255
6.14	An Equilibrium Point in Simple Games	256
7.1	Survey Sampling	276
7.2	Dice	280
7.3	Countably Finite Set	281
7.4	Multiple Views of Countably Finite	281
7.5	Countably Infinite Set	282
7.6	Uncountably Infinite Set	282
7.7	Single Die Experiment	287
7.8	Approval Voting	291
7.9	Single Coin Flip	294
7.10	Probabilistic Analysis of Supreme Court Decisions	298
7.11	Updating Probability Statements	299
7.12	Conditional Probability with Dice	300
7.13	Analyzing Politics with Decision Trees	307
8.1	Binomial Analysis of Bill Passage	329
8.2	Poisson Counts of Supreme Court Decisions	331
8.3	Modeling Nineteenth Century European Alliances	333
8.4	Poisson Process Model of Wars	333
8.5	Characterizing Income Distributions	339

8.6	Levels of Women Serving in U.S. state legislatures	342
8.7	The Standard Normal CDF: Probit Analysis	344
8.8	Employment by Race in Federal Agencies	363
8.9	Employment by Race in Federal Agencies	366
8.10	An Ethnoarchaeological Study in the South American Tropical Lowlands	369
9.1	Contraception Use in Barbados	385
9.2	Campaign Contributions	389
9.3	Shuffling Cards	391
9.4	Conflict and Cooperation in Rural Andean Communities	401
9.5	Internal Population Migration in Malawi	406

## Preface

This book is intended to serve several needs. First (and perhaps foremost), it is supposed to be an introduction to mathematical principles for incoming social science graduate students. For this reason, there are a large set of examples (eighty-six of them actually) drawn from various literatures including sociology, political science, anthropology, psychology, public policy, and others. Almost no example is produced from “hypothetical data.” This approach is intended not only to motivate specific mathematical principles and practices, but also introduce the *way* that social science researchers use these tools. With this approach the topics presumably retain needed relevance. The design of the book is such that this endeavor can be a semester-long adjunct to another topic like data analysis or it can support the prefresher “math-camp” approach that is increasingly popular. Second, this book can also serve as a single reference work where multiple books would ordinarily be needed. To support this function there is extensive indexing and referencing designed to make it easy to find topics. Also in support of this purpose, there are some topics that may not be suitable for coursework that are deliberately included for this purpose (i.e. calculus on

trigonometric functions and advanced linear algebra topics). Third, the format is purposely made conducive to self-study. Many of us missed or have forgotten various mathematical topics and then find we need them later.

The main purpose of the proposed work is to address an educational deficiency in the social and behavioral sciences. Undergraduate curriculum in the social sciences tends to under-emphasize mathematical topics that are then needed by students at the graduate level. This leads to some discomfort whereby incoming graduate students are often unprepared for and uncomfortable with the necessary mathematical level of research in these fields. As a result, the methodological training of graduate students increasingly often begins with intense “prequel” seminars wherein basic mathematical principles are taught in short (generally week-long) programs just before beginning regular first year course-work. Usually these courses are taught from the instructor’s notes, selected chapters from textbooks, or assembled sets of monographs or books. There is currently no *tailored* book-length work that specifically addresses the mathematical topics of these programs. This work fills this need by providing a comprehensive introduction to the mathematical principles needed by modern research social scientists. The material introduces basic mathematical principles necessary to do analytical work in the social sciences, starting from first principles, but without unnecessary complexity. The core purpose is to present fundamental notions in standard notation and standard language with a clear, unified framework throughout

Although there is an extensive literature on mathematical and statistical methods in the social sciences, there is a dearth of introduction to the underlying *language* used in these works, exacerbating the fact that many students in social science graduate programs enter with an undergraduate education that contains no regularized exposure to the mathematics they will need to succeed in graduate school.

Actually, the book *is itself* a prerequisite, so for obvious reasons the prerequisites to this prerequisite are minimal. The only required material is knowledge of high school algebra and geometry. Most target students will have had very little mathematical training beyond this modest level. Furthermore, the first chapter is sufficiently basic that readers who are comfortable with only arithmetic operations on symbolic quantities will be able to work through the material. No prior knowledge of statistics, probability, or non-scalar representations will be required. The intended emphasis is on conceptual understanding of key principles and the subsequent application through examples and exercises. No proofs or detailed derivations will be provided.

The book has two general divisions reflecting a core component along with associated topics. The first section comprises six chapters, and is focused basic the mathematical tools, matrix algebra, and calculus. The topics are all essential, deterministic, mathematical principles. The primary goal of this section is to establish the mathematical language used in formal theory and mathematical analysis, as practiced in the social sciences. The second section, consisting of three chapters, is designed to give the background required to proceed in standard empirical quantitative analysis courses such as social science statistics, mathematical analysis, and others.

Although structure differs somewhat by chapter, there is a general format followed within each. There is motivation given for the material, followed by a detailed exposition of the concepts. The concepts are illustrated with examples that social scientists care about and can relate to. This last point is not trivial. A great many books in these areas center on engineering and biology examples, and the result is often to reduce reader interest and perceived applicability in the social sciences. Therefore, every example is taken from the social and behavioral sciences. Finally, each chapter has a set of exercises designed to reinforce the primary concepts.

There are different ways to teach from this book. The obvious way is to cover the first six chapters sequentially, although aspects of the first two chapters may be skipped for a suitably prepared audience. Chapter 2 focuses on trigonometry and this may not be needed for some programs. The topics in Chapters 7, 8 and 9 essentially constitute a “pre-statistics” course for social science graduate students. This may or may not be useful for specific purposes. The last chapter on Markov chains addresses a topic that has become increasingly important. This tool is used extensively in both mathematical modeling and Bayesian statistics. In addition, this chapter is a useful way to practice and reinforce the matrix algebra principles covered in Chapters 3 and 4. This book can also be used with in a “just in time” way whereby a course on mathematical modeling or statistics proceeds until certain topics in matrix algebra, calculus, or random variables are needed.

One intended use of this book is through a “math-camp” approach where incoming graduate students are given a pre-semester intensive introduction to the mathematical methods required for their forthcoming study. This is pretty standard in economics and increasingly done in political science, sociology, and other fields. For this purpose, I recommend one of two possible tracks through the material:

### **3 Day Program**

- Chapter 1: The Basics.
- Chapter 3: Linear Algebra: Vectors, Matrices, and Operations.
- Chapter 5: Elementary Scalar Calculus.

### **5 Day Program**

- Chapter 1: The Basics.
- Chapter 3: Linear Algebra: Vectors, Matrices, and Operations.

- Chapter 5: Elementary Scalar Calculus.
- Chapter 7: Probability Theory.
- Chapter 8: Random Variables.

The five day program above is focuses on a pre-statistics curriculum after the introductory mathematics. If this is not appropriate or desired then the continuation chapters on linear algebra and calculus (Chapters 4 and 6) be substituted for the latter two chapters. Conversely, a lighter pre-statistics approach that did not need to focus on theory involving calculus might look the following:

#### **Standard Pre-Calculus Program**

- Chapter 1: The Basics.
- Chapter 3: Linear Algebra: Vectors, Matrices, and Operations.
- Chapter 7: Probability Theory.
- Chapter 8: Random Variables.

This program omits Chapter 5 from the previous listing but sets students up for a standard regression text such as Hanushek and Jackson (1977), Gujarati (1995), Neter, *et al.* (1996), , Fox (1997), or the forthcoming text in this series by Schneider and Jacoby.

Its cliché to say, but this book was not created in a vacuum and numerous people read, commented on, criticized, and even taught from the manuscript. These people include: Attic Access, Mike Alvarez, Maggie Bakhos, Ryan Bakker, Neal Beck, James Fowler, Jason Gainous, Scott Gartner, Hank Heitowit, Bob Huckfeldt, Bob Jackman, Renee Johnson Cindy Kam, Gary King, Michael Martinez, Bill Reed, Les Thiele, and Kevin Wagner. In particular I would like to thank Ed Parsons for his continued assistance and patience in helping get this done. I have also enjoyed the support of various chairs and deans at the University of California–Davis and the ICPSR Summer Program at the University of Michigan.