

Contents

Chapter 1	Introduction	1
1.1	Flow diagram of procedure used,	1
1.2	Role of computer,	1
1.3	Sequence of subjects discussed,	3
Chapter 2	Assumptions and Methods of Fitting Equations	5
2.1	Assumptions,	5
2.2	Methods of fitting equations,	6
2.3	Least squares,	6
2.4	Linear least-squares estimation,	9
2.5	Nonlinear least-squares estimation,	9
Appendix 2A	Linear least-squares estimates—one independent variable,	10
2A.1	Assumptions,	10
2A.2	Basic idea and derivation,	10
2A.3	Confidence regions,	12
Appendix 2B	Linear least-squares estimates—general case,	13
2B.1	Estimates of coefficients,	13
2B.2	Variance and standard errors of coefficients,	15
2B.3	Computer computations,	16
2B.4	Partitioning sums of squares,	16
2B.5	Multiple correlation coefficient squared, R_y^2 ,	17
2B.6	F -value,	17
2B.7	Bias and random error,	17
2B.8	Residual mean square,	17
2B.9	Residual root mean square,	18

Chapter 3	One Independent Variable	19
3.1	Plotting data and selecting form of equation, 19	
3.2	Plots of linearizable equations, 19	
3.3	Plots of nonlinearizable equations, 22	
3.4	Statistical independence and clusters of dependent variable, 24	
3.5	Allocation of data points, 25	
3.6	Outliers, 25	
3.7	Use of computer program, 26	
3.8	Study of residuals, 27	
3.9	Dealing with error in the independent variable, 32	
3.10	Example of fitting a straight line to data—one independent variable, 32	
Appendix 3A	Cumulative distribution plots of random normal deviates, 33	
Appendix 3B	Example of fitting a straight line to data—one independent variable, 33	
Chapter 4	Two or More Independent Variables	50
4.1	Inadequacies of x - y plots with two or more independent variables, 50	
4.2	Equation forms and transformations, 53	
4.3	Variances of estimated coefficients, b_i , and fitted values, Y_j , 55	
4.4	Use of indicator variables for discontinuous or qualitative classifications, 56	
	One qualitative factor at two levels	
	One qualitative two-level factor and one quantitative (continuous) factor	
	Discrete factors at more than two levels	
	Discrete factors interacting with continuous factors	
4.5	Allocation of data in factor space, 57	
	Linear dependences among the x_i	
	The outermost points in data space	
	Nested data	
Chapter 5	Fitting an Equation in Three Independent Variables	60
5.1	Introduction, 60	
5.2	First trials, 60	
5.3	Possible causes of disturbance, 61	

5.4	Outlier or logged response or squared independent variable, 65	
5.5	Random error estimated from near neighbors, 72	
5.6	Independence of observations, 73	
5.7	Systematic examination of alternatives discovered sequentially, 75	
5.8	Remote points in factor space, 77	
5.9	Equation using "lined-out" data, 77	
5.10	Conclusions on stack loss problem, 81	
5.11	General conclusions, 82	
Chapter 6	Selection of Independent Variables	83
6.1	Introduction, 83	
	Assumptions	
	Obvious imperfections	
	2^K possible equations from K candidate variables	
	On stepwise regression	
	F -test	
	All 2^K equations and fractions	
6.2	Total squared error as a criterion for goodness of fit— C_p , 86	
	Definition	
	Derivation of the C_p statistic	
	Mallows' graphical method of comparing fitted equations	
6.3	A four-variable example, 89	
6.4	Disposition of data in factor space, 91	
6.5	A six-variable example, 95	
	a. Search for all 2^K equations	
	b. $t_{K,r}$ -directed search	
	c. Fractional replicates	
Appendix 6A	Computer printouts of four-variable example, 104	
Appendix 6B	Computer printouts of six-variable example, 104	
Appendix 6C	Fractional replication for 2^K equations, 104	
Chapter 7	Some Consequences of the Disposition of the Data Points	121
7.1	Introduction, 121	
7.2	Description of example with ten independent variables, 122	
	First steps	

- 7.3 Interior analysis I. “Component effects” table, 123
- 7.4 Interior analysis II. Component and component-plus-residual plots, 124
- 7.5 Search of all 2^K possible subset equations, 126
- 7.6 $t_{K,i}$ -directed search, 126
- 7.7 Test for an outlier, 127
- 7.8 Interior analysis III. Weighted-squared-standardized-distance to look for far-out points of influential variables, 129
- 7.9 Interior analysis IV. Variance ratio to look for far-out points of uninfluential variables, 131
- 7.10 Interior analysis V. Error estimation from near neighbors, 133
- 7.11 Cross verification, 136
- 7.12 Other examples of error estimation from near neighbors, 137
 - Six-variable example of Chapter 6
 - Stack loss example of Chapter 5
- 7.13 Additional examples of using component and component-plus-residual plots, 139
 - Example of the use of plots in choosing the form of equation
 - Distribution of observations over each independent variable
 - Four-variable example—Distribution of independent variables
 - Eleven-variable example—Inner and outer observations
- 7.14 Other conditions identified by component and component-plus-residual plots, 145
- 7.15 Summary, 146
- Appendix 7A Computer printouts of ten-variable example, 150
- Appendix 7B Computer printout of six-variable example, 199
- Appendix 7C Computer printouts of stack loss example, 200
- Appendix 7D Computer printouts of four-variable octane example, 207
- Appendix 7E Computer printouts of eleven-variable example, 218
- Appendix 7F Critical values for studentized residual to test for single outlier, 232

Chapter 8 Selection of Variables in Nested Data 234

- 8.1 Background of example, 234
- 8.2 Fitting equation, 236
- 8.3 Recognition of nested data, 236
- 8.4 Data identification and entry, 236
- 8.5 Fitting equation to nested data, 236
 - Use of indicator (dummy) variables
 - Test for significance of added variables, “individual” versus “common” slopes
- 8.6 Fitting equation among sets of nested data, 240
- 8.7 Selecting variables based on total error, recognizing bias and random error, 243
- 8.8 Properties of final equation, 244
- 8.9 Comparison of equations, 247
- Appendix 8A Data preparation and computer printouts of example, 248

Chapter 9 Nonlinear Least Squares, a Complex Example 267

- 9.1 Introduction, 267
- 9.2 Background of example, 268
- 9.3 Replicates, 271
- 9.4 Potential equations, 273
- 9.5 Nonlinear fit—observations of individual cements versus time, 276
- 9.6 Fit with indicator variables, 277
- 9.7 Fit with composition, 280
- 9.8 Possible nesting within cements, 293
- 9.9 Comparison with previous linear least-square fits, 295
- 9.10 Fit using composition in mol percents, 296
- 9.11 Conclusions, 299
- Appendix 9A Computer printout of nonlinear example, 300

Glossary 338

- Conventions, 338
- Symbols, 338
- Computer terms, 343

LINWOOD User's Manual	354
LINWOOD, a computer linear least-squares curve-fitting program	
Abstract, 354	
General information, 355	
Input, 356	
Control card entry, 356	
Standard data card entry, 366	
Problems using transformation option, 367	
Order of cards, 374	
Sample problem, 376	
Example of weighted observations, 381	
Example to measure the precision of calculations, 396	
Example of using plots in choosing the form of equation, 406	
Execution time, 415	
Summary of control cards, 418	
Summary of order of cards, 419	
NONLINWOOD User's Manual	420
NONLINWOOD, a computer nonlinear least-squares curve-fitting program	
Abstract, 420	
General information, 421	
Input, 422	
Control card, 422	
Format card, 424	
Starting values of coefficients, 424	
Data cards, 424	
Order of cards, 425	
Control and data card entry forms, 425	
Equation subroutine, 425	
Example of transformations, 427	
Example 1, 2 coefficients, 2 variables, 428	
Example 2, 43 coefficients, 16 variables, 442	
Example 3, 19 coefficients, 7 variables, 444	
Summary of control cards, 446	
Example of subroutine model, 447	
Summary of order of cards, 447	
Bibliography	449
Index	453