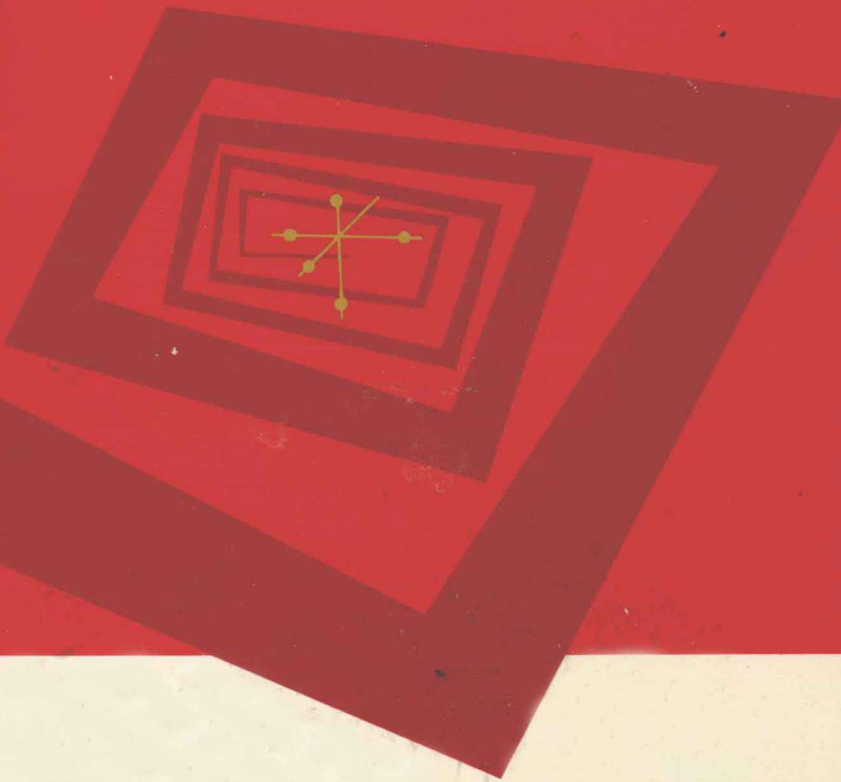
Data Reduction and Error Analysis for the Physical Sciences



Philip R. Bevington

Associate Professor of Physics · Case Western Reserve University

PHILIP R. BEVINGTON
Associate Professor of Physics
Case Western Reserve University

DATA REDUCTION AND ERROR ANALYSIS FOR THE PHYSICAL SCIENCES

McGRAW-HILL BOOK COMPANY
New York San Francisco St. Louis
Toronto London Sydney

DATA REDUCTION AND ERROR ANALYSIS FOR THE PHYSICAL SCIENCES

Printed in the United States of America.

Library of Congress catalog card number: 69–16942

20 21 BKP BKP 8 7 6 5

TO JOAN

PREFACE

The purpose of this text is to provide an introduction to the techniques of data reduction and error analysis commonly employed by individuals doing research in the physical sciences and to present them in sufficient detail and breadth to make them useful for students throughout their undergraduate and graduate studies. The presentation is developed from a practical point of view, including enough derivation to justify the results, but emphasizing the methods more than the theory.

The level of primary concern is that of junior and senior undergraduate laboratory where a thorough study of these techniques is most appropriate. The treatment is intended to be comprehensive enough to be suitable for use by graduate students in experimental research who would benefit from the generalized methods for linear and non-linear least-squares fitting and from the summaries of definitions and techniques.

At the same time, the introduction to the material is made self-supporting in that no prior knowledge of the methods of statistical evaluation is assumed; the material of each section is

developed from first principles. A discussion of differential calculus and manipulation of matrices and determinants is included in the appendixes to supplement their use in the text.

The emphasis, however, is toward the application of more general techniques than are usually presented in undergraduate laboratories. With the proliferation of computers and their use in research laboratories, it is important that sophisticated concepts of data reduction be introduced. Computer routines written explicitly for each section are used throughout to illustrate in a practical way both the concepts and the procedures discussed.

The first five chapters introduce the concepts of errors, uncertainties, probability distributions, and methods of optimizing the estimates of parameters characterizing observations of a single variable. Chapters 6 to 9 deal with the problem of fitting, analytically, complex functions to observations of more than one variable, including estimates of the resulting uncertainties and tests for optimizing the functional form of the fit. The last third of the text contains a description of techniques for searching for the best fit to data with arbitrary functions. Techniques for manipulating data or extracting information without fitting are also discussed.

Computer programs The primary purpose of the computer routines is to clarify the presentation, but they are meant to be usable for calculations as well. They are written as subroutines and function subprograms to provide flexibility for the user in applying them to his own data. The format of the routines is similar to that of the IBM Scientific Subroutine Package for the IBM 360 computer system, including considerable commentary to define the parameters used and to describe the flow of the routine.

The routines are written in Fortran IV, but they are compatible with Fortran II except for the use of double precision and missing suffixes in the names of library functions, which are noted in the program descriptions. The sizes of most of the arrays are to be specified by the user and they are dimensioned in the routines with a size of 1. For most versions of Fortran, the dimen-

sion size in a subprogram is a dummy argument for one-dimensional arrays and need not correspond to the actual size used. Two-dimensional arrays are always dimensioned explicitly; these arrays and those which are wholly contained within the routines are assigned dimensions such that the routines can handle up to 10 terms of a fitting function and up to 100 data points. These dimensions may be increased for larger input dimensions or decreased for more efficient use of memory. All input and output variables are specified as arguments of the calling statement.

These routines have been debugged in both Fortran IV and Fortran II versions on small and large computers. They are intended to be usable operating routines and are reasonably efficient. Their most important function, however, is to serve as a framework on which to build and modify routines to serve the specific needs of the user.

Acknowledgments I am deeply indebted to members of the Stanford University Nuclear Physics Group and especially to Drs. A. S. Anderson, T. R. Fisher, D. W. Heikkinen, R. E. Pixley, F. Riess, and G. D. Sprouse for many helpful comments, suggestions, and techniques. The use of the SCANS (Stanford Computers for the Analysis of Nuclear Structure) PDP-7 and PDP-9 computers for the development of computer routines and the assistance of Wylbur Wryght of the Stanford Computation Center are gratefully acknowledged. I have benefitted from comments by Prof. Jay Orear and especially from many invaluable suggestions and criticisms of Prof. Hugh D. Young.

I wish also to express my appreciation to Mrs. Pat Johnson for her care and diligence in typing and to Mrs. H. Mae Sprouse for her artistic work on the illustrations. Special appreciation is due to my wife, Joan, and our children, Ann and Mark, for enduring so patiently the division of my attention during the writing of this book.

Philip R. Bevington

TABLE OF CONTENTS

APPENDIX C

CHAPTER ONE

SYSTEMATIC AND RANDOM ERRORS

1-1 ERRORS

It is a well-established rule of scientific investigation that the first time an experiment is performed the results bear all too little resemblance to the "truth" being sought. As the experiment is repeated, with successive refinements of technique and method, the results gradually and asymptotically approach what we may accept with some confidence to be a reliable description of events. Some investigators have gone so far as to assert that nature is loath to give up her secrets without a considerable expenditure of effort on our part, and that it is a fundamental fact of life that first steps in experimentation are bound to fail. Whatever the reason, it is certainly true that for all physical experiments, errors

and uncertainties exist which must be reduced by improvements in techniques and ideas and then estimated to establish the validity of results.

Error is defined by Webster as "the difference between a calculated or observed value and the true value." Usually, of course, we do not know what the "true" value is or there would be no reason for performing the experiment. But we often do know approximately what it should be, either from earlier experimentation along the same line or from other theoretical or experimental approaches. Such approximations can yield an indication of whether our result is of the right order of magnitude, but we also need some systematic way to determine from the data themselves how much confidence we can have in our experimental results.

There is one class of errors which we can deal with immediately: that which originates from mistakes or blunders in computation or measurement. Fortunately, these sources of error are usually apparent either as obviously incorrect data points or as results which are not reasonably close to the expected results. They are classified as *illegitimate errors* which can be corrected by performing the erroneous operations again correctly.

Systematic errors There is another class of errors which is not so easy to detect and for which statistical analysis is not generally useful. This is the class of *systematic errors*, such as those which result reproducibly from faulty calibration of equipment or from bias on the part of the observer. These errors must be estimated from an analysis of the experimental conditions and techniques. In some cases, corrections can be made to the data to compensate for systematic errors where the type and extent of the error is known. In other cases, the uncertainties resulting from these errors must be estimated and combined with uncertainties from statistical fluctuations.

EXAMPLE 1-1 A student measures a table top with a steel meter stick and finds that the average of his measurements yields a result of 1.982 m for the length of the table. He subsequently learns that the meter stick was calibrated at 25°C and has an

expansion coefficient of .0005/°C. Since his measurements were made at a temperature of 20°C, he multiplies his results by $1 - 5(.0005) = 0.9975$ so that his new determination of the length is 1.977 m.

When the same student repeats the experiment, he discovers that his technique for reading the meter stick was faulty in that he did not always read the divisions from directly above. By experimentation he determines that this consistently results in a reading which is 2 mm too short. With this correction, his final result is 1.979 m.

Accuracy vs. precision There is considerable confusion among students as to the meaning of and difference between the terms accuracy and precision. To add to this confusion, Webster defines them equally. In scientific investigation, however, they are assigned distinctly different meanings which must be kept separate.

The *accuracy* of an experiment is a measure of how close the result of the experiment comes to the true value. Therefore, it is a measure of the correctness of the result. The *precision* of an experiment is a measure of how exactly the result is determined, without reference to what that result means. It is also a measure of how reproducible the result is. The *absolute precision* indicates the magnitude of the uncertainty in the result in the same units as the result. The *relative precision* indicates the uncertainty in terms of a fraction of the value of the result.

For example, in the experiment of Example 1-1, the first result was given with a fairly high precision. The table top was found to be 1.982 m long, indicating an absolute precision on the order of 1 mm and a relative precision on the order of $\frac{1}{2000}$. The corrections to this result were meant to improve the accuracy by compensating for known deviations of the first result from the best estimate possible. These corrections did not improve the precision at all, but did, in fact, worsen the estimated precision because the corrections were themselves only estimates of the exact corrections.

It is obvious that we must consider accuracy and precision

simultaneously for any experiment. It would be a waste of time and energy to determine a result with a high precision if we knew the result would be highly inaccurate. Conversely, a result cannot be considered to be extremely accurate if the precision is low. For example, if the length of the table is quoted as 2. m, the answer is undoubtedly accurate but the amount of information available is limited by the fact that such a precision only specifies the length to be between 1.5 and 2.5 m long. Similarly, if the length of the table is known to be exactly 2.000 m, there would be no point in improving the experimental precision so long as the inaccuracy remains about 20 mm.

Significant figures and round-off The precision of an experimental result is implied by the way in which the result is written, though it should generally be quoted specifically as well. To indicate the precision, we write a number with as many digits as are significant. The number of *significant figures* in a result is defined as follows:

1. The leftmost nonzero digit is the most significant digit.
2. If there is no decimal point, the rightmost nonzero digit is the least significant digit.
3. If there is a decimal point, the rightmost digit is the least significant digit, even if it is a 0.
4. All digits between the least and most significant digits are counted as significant digits.

For example, the following numbers each have four significant digits: 1,234; 123,400; 123.4; 1,001, 1,000., 10.10, 0.0001010, 100.0. If there is no decimal point, there are ambiguities when the rightmost digit is a 0. For example, the number 1,010 is considered to have only three significant digits even though the last digit might be physically significant. To avoid this ambiguity, it is better to supply decimal points or write such numbers in exponent form as an argument in decimal notation times the appropriate power of 10. Thus, our example of 1,010 would be written as 1,010. or 1.010×10^3 if all four digits are significant.

When quoting results of an experiment, the number of signif-

icant figures given should be approximately one more than that dictated by the experimental precision. The reason for including the extra digit is that in computation one significant figure is sometimes lost. Errors introduced by insufficient precision in calculations are classified as illegitimate. If an extra digit is specified for all numbers used in the computation, the original precision will be retained to a greater extent. For example, in the experiment of Example 1-1, if the absolute precision of the result is 10 mm, the third figure is known with an uncertainty of ± 1 and the fourth figure is not really known at all. We would be barely justified in specifying four figures for computation. If the precision is 2 mm, the third digit is known quite well and the fourth figure is known approximately. We are justified in quoting four figures, but probably not justified in quoting five figures since we cannot even have much confidence in the value of the fourth figure.

When insignificant digits are dropped from a number, the last digit retained should be rounded off for the best accuracy. To round off a number to a smaller number of significant digits than are specified originally, truncate the number to the desired number of significant digits and treat the excess digits as a decimal fraction. Then

1. If the fraction is greater than $\frac{1}{2}$, increment the least significant digit.
2. If the fraction is less than $\frac{1}{2}$, do not increment.
3. If the fraction equals $\frac{1}{2}$, increment the least significant digit only if it is odd.

In this manner, the value of the final result is always within half the least significant digit of the original number. The reason for rule (3) is that in many cases the fraction equals either 0 or $\frac{1}{2}$ and consistently incrementing the least significant digit for a fraction of $\frac{1}{2}$ would lead to a systematic error. For example, 1.235 and 1.245 both become 1.24 when rounded off to three significant figures, but 1.2451 becomes 1.25.

Random errors The *accuracy* of an experiment, as we have defined it, is generally dependent on how well we can control