Preface

Measurements through quantitative experiments are one of the most fundamental tasks in all areas of science and technology. Astronomers analyze data from asteroid sightings to predict orbits. Computer scientists develop models for recognizing spam mail. Physicists measure properties of materials at low temperatures to understand superconductivity. Materials engineers study the reaction of materials to varying load levels to develop methods for prediction of failure. Chemical engineers consider reactions as functions of temperature and pressure. The list is endless. From the very small-scale work on DNA to the huge-scale study of black holes, quantitative experiments are performed and the data must be analyzed.

Probably the most popular method of analysis of the data associated with quantitative experiments is least squares. It has been said that the method of least squares was to statistics what calculus was to mathematics. Although the method is hardly mentioned in most engineering and science undergraduate curricula, many graduate students end up using the method to analyze the data gathered as part of their research. There is not a lot of available literature on the subject. Very few books deal with least squares at the level of detail that the subject deserves. Many books on statistics include a chapter on least squares but the treatment is usually limited to the simplest cases of linear least squares. The purpose of this book is to fill the gaps and include the type of information helpful to scientists and engineers interested in applying the method in their own special fields.

The purpose of many engineering and scientific experiments is to determine parameters based upon a mathematical model related to the phenomenon under observation. Even if the data is analyzed using least squares, the full power of the method is often overlooked. For example, the data can be weighted based upon the estimated errors associated with the data. Results from previous experiments or calculations can be combined with the least squares analysis to obtain improved estimate of the model parameters. In addition, the results can be used for predicting values of the dependent variable or variables and the associated uncertainties of the predictions as functions of the independent variables.

Chapter 1 includes a review of the basic statistical concepts that are used throughout the book. The method of least squares is developed in Chapter

2. The treatment includes development of mathematical models using both linear and nonlinear least squares. In Chapter 3 evaluation of models is considered. This chapter includes methods for measuring the "goodness of fit" of a model and methods for comparing different models. The subject of candidate predictors is discussed in Chapter 4. Often there are a number of candidate predictors and the task of the analyst is to try to extract a model using subspaces of the full candidate predictor space. In Chapter 5 attention is turned towards designing experiments that will eventually be analyzed using least squares. The subject considered in Chapter 6 is nonlinear least squares software. Kernel regression is introduced in Chapter 7. Kernel regression is a nonparametric modeling technique that utilizes local least squares estimates.

Although general purpose least squares software is available, the subject of least squares is simple enough so that many users of the method prefer to write their own routines. Often, the least squares analysis is a part of a larger program and it is useful to imbed it within the framework of the larger program. Throughout the book very simple examples are included so that the reader can test his or her own understanding of the subject. These examples are particularly useful for testing computer routines.

I would like to thank David Aronson for the many discussions that we have had over the years regarding the subject of data modeling. My experience in applying least squares to real problems is based upon software that I developed with Ronen Kimche and Victor Leikehman and I would like to thank them for their advice and help. Thanks to Richard Green for introducing me to the first English translation of Gauss's *Theoria Motus* in which Gauss developed the foundations of the method of least squares. I have been teaching a graduate course on analysis and design of experiments and as a result have had many useful discussions with our students throughout the years. When I decided to write this book two years ago, I asked each student taking the course to critically review a section in each chapter that had been written up to that point. Over 20 students in the spring of 2004 and over 20 students in the spring of 2005 submitted reviews that included many useful comments and ideas. A number of typos and errors were located as a result of their efforts and I really appreciated their help. I would also like to thank Donna Bossin for her help in editing the manuscript and teaching me the some of the cryptic subtleties of WORD.

John R. Wolberg
Haifa, Israel July, 2005