# Preface

The method of least squares has been developed by Carl-Friedrich Gauß in 1795 and has been published 1809 in volume two of his work on celestial mechanics *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium.* Although it is a rather old technique, it is used still today in many practical applications and has lost nothing in its importance, but has been developed further in several directions.

The idea of this method is to determine the behaviour of systems for which just a limited number of samples, i.e. pairs of input and output, are given. If a certain model function describing the system is presumed, then the parameters of this function are searched which explain the system best in terms of minimisation of the quadratic error. In case of simple systems, applying the method of least squares reduces to, for instance, drawing an approximating curve through a scatter plot of data points.

The early fields of application included, for instance, astronomy (modelling of celestial bodies' movement) and geodesy with the purpose of alleviating measurement errors. Nowadays, this method of approximating data points is probably applied in all sciences, especially those dealing traditionally with statistics as, for instance, physics, biology, economy, or psychology, but also in more recent fields as computer vision. There is basically no difference between experimental data and computer generated data with respect to the data analysis via the method of least squares.

The principle of least squares is generally straightforward, however, its computational intricacy is heavily dependent on the character and complexity of the model function which should describe the behaviour of the system under investigation. So, from the introduction of this data-fitting method in the 19th century until now many numerical techniques for proper and elegant treatment have been developed. In addition, mathematical tools for the evaluation of the fitting results have had to be derived.

The intention of this textbook is to introduce the topic in a more comprehensible manner enabling the reader to solve his/her own specific data-fitting problem. It describes the method of least squares and the treatment of uncertainty in a recipe-like way. Emphasis is put on the unified and generic approach from which the solutions for particular applications can be derived. The book fills the gap between those explaining only very simple data-fitting techniques and those looking at the topic at very high mathematical level. The text discusses many real-life and simulated examples and is accompanied with working source code in the programming language C, which is provided on the publishers website (see 'OnlinePlus' area).

The book addresses mainly engineers and software programmers or corresponding undergraduates, which are not primarily interested in proofs and derivations of the formulae, but want to quickly become acquainted with the topic and its pitfalls in order to solve a particular fitting problem. A basic mathematical background, especially a familiarity with matrices and partial derivatives as well as some knowledge about variances and standard deviations, is beneficial. Readers, not yet familiar with the method of least squares, are recommended to start with the first chapter followed by the others in sequential order. Every chapter finishes with a list of test questions, enabling a cross-check of whether the subject has been understood.

The material is organised in two parts. The generic basics of data fitting by the Gaussian method of least squares are introduced in the first part. Several examples of linear systems as well as of nonlinear systems are discussed thoroughly in order to enable the reader to solve similar problems.
Chapter 1 explains the general idea of data fitting and defines the terminology and the notation. Chapter 2 presents all principal formulae for the application of the least-square method. In Chapter 3, methods for the estimation of weights are proposed. This is essential, for instance, when uncertainties of recorded data points are not known in advance or the data contain outliers. A clustering approach is discussed, which separates the possible outliers from the bulk of useful data points. An assessment of the results of the data fitting is described in Chapter 4.

The second part of the book is mostly dedicated to the underlying theory and also some numerical methods are discussed in more detail.
It starts with Chapter 5 which is about basic linear algebra and algorithms for matrix inversion. Chapter 6 explains the idea of least squares introducing the principle of maximum likelihood. On this basis, different techniques for the solution of linear and nonlinear fitting problems (i.e. optimisation) are discussed. In Chapter 7 some methods and tools are collected which are helpful in connection with data fitting.

The appendix contains two chapters. The first is about investigations regarding a cluster-based outlier-detection method. The second chapter describes a data-fitting software and discusses several issues of implementing the least-squares method. It concludes with a detailed performance tests of the software.

Leipzig, July 2010 Tilo Strutz