Surveys of Modern Mathematics
Volume III

Application of Elementary Differential Geometry to Influence Analysis

Yat-Sun Poon
Wai-Yin Poon

International Press
www.intlpress.com

高等教育出版社
HIGHER EDUCATION PRESS

Surveys of Modern Mathematics, Volume III
Application of Elementary Differential Geometry to Influence Analysis

Yat-Sun Poon
Department of Mathematics, University of California at Riverside
Wai-Yin Poon
Department of Statistics, The Chinese University of Hong Kong

2010 Mathematics Subject Classification. 53-01, 53A07, 53A30, 53A45, 53B20, 62J05, 97K70.

ISBN: 978-1-57146-252-7

Printed in the United States of America.

16 15 14 13 12   1 2 3 4 5 6 7 8 9

SURVEYS OF MODERN MATHEMATICS

Dedicated to our father Yan Ding Poon
for his 88th birthday

Contents

Part III Local Influence Analysis

Preface

The two authors of this book have different backgrounds within the mathematical sciences. After fifteen years of collaboration, they find it benefical to introduce their research areas jointly to students in the early stages of their research careers. Mathematics students will learn that mathematical concepts have an immediate impact on real life situations. Statistics students will appreciate that abstract mathematical concepts are accessible, relevant and valuable. In fact, an early draft of this book was used by one of the authors as lecture notes to run one year-long undergraduate research seminar for a mixture of mathematics and statistics undergraduate students in their senior years.

We hope that this book will encourage young scientists to develop an appreciation of inter-disciplinary research at a time when their research career is just beginning to emerge.

Differential geometry is a broad mathematical subject. Its global aspects are often presented only to mathematics majors or graduate students. However, its local aspects are the foundation of global issues and can be made accessible to all students with rigorous, multi-variable calculus training. To study local geometry, one studies graphs. It is a generalization of the freshman discussion on the concavity of a real-valued function of a single variable. By studying the geometry of graphs, one can proceed to learn global differential geometry as a mathematician.

One may also begin to apply the geometry of a graph to analyze functions arising from concrete problems, such as through statistics. Many statistical analyses involve a hypothesized model. Once a model is specified, the data collected are used to estimate the parameters that characterize the model. Further inference is affected by the hypothesized model and the data collected. Therefore, one must assess the influence based on the perturbation of various aspects of the model inputs. Perturbations can be represented by a set of perturbation parameters, and the function of these perturbation parameters becomes a mathematical object of interest. Therefore, one can apply geometric concepts to study this function and hence, deduce information on the perturbation.

Working through this process prompts theoretical, practical and technical issues. In particular, we must develop measures for the influence of individual perturbation parameters. We must also develop measures for the joint influence of any two, and then any groups of perturbation parameters. From single to multiple-parameter measures, the process is not a straightforward generalization of lower dimensional geometric problems to higher dimensional problems, because direct geometric generalization may not produce appropriate measures capable of serving the purposes of a statistical analysis. Instead, it is necessary to develop a set of meaningful measures that can constitute a well-structured system that enables further pursuit of relationships among the measures as well as the development of practical tools that facilitate interpretation and data analysis. Typical examples of such practical tools include the establishment of benchmarks for judging and interpreting measures and the search for alternative measures that reduce the computational burden. All such issues must be addressed using geometric techniques in the light of statistical considerations.

In Part I of this book, to fix conventions we recall basics of linear algebra, multi-variable calculus and Euclidean geometry in Chapters 1 and 2. In Chapter 3, we introduce the concept of normal sections, first fundamental forms and second fundamental forms. In Chapter 4, we introduce normal curvature and sectional curvatures. In Chapter 5, we study conformal transformations. This finishes our mathematical preparation.

In Part II, we first review some elementary statistics topics. In Chapters 6 and 7, we introduce basic concepts in relation to univariate distribution for discrete and continuous random variables, including the maximum likelihood estimation method. After generalizing these concepts to the bivariate and multivariate distributions in Chapter 8, we introduce simple linear regression in Chapter 9. Linear regression is the most popular statistical model, and is used as the key example in Part III of this book. To prepare for the illustration, some well-known topics in linear regression are discussed in Chapter 10.

In Part III, we apply the geometric concepts developed in Part I to the statistical issues and models articulated in Part II. The goal is to develop various measures generated by a local perturbation to assess the influence of the perturbation of model inputs. In Chapter 11, we develop the concept of likelihood displacement function. In Chapter 12, we apply the tools developed in Chapters 3 and 4 to analyze the likelihood displacement function. In the process of this development, different or mingled perspectives must be clarified. We use the regression model with some of the most well-known perspectives described in Chapter 10 for illustration. The relations among various measures that are generated in the process must also be analyzed. This is done in Chapter 13. Finally, in Chapter 14, we analyze the modification of perturbations using Chapter 5 as the theoretical foundation.

As we develop our presentation, we often encounter tedious but necessary computations. For the sake of completeness, we have placed computations in the Appendices.

Graduate students in mathematics may choose to begin reading this book starting with Part II. Likewise, graduate students in statistics may skip Part II. However, undergraduate students from both disciplines will benefit from reading this book in its entirety.

Yat-Sun Poon
Wai-Yin Poon
May, 2012