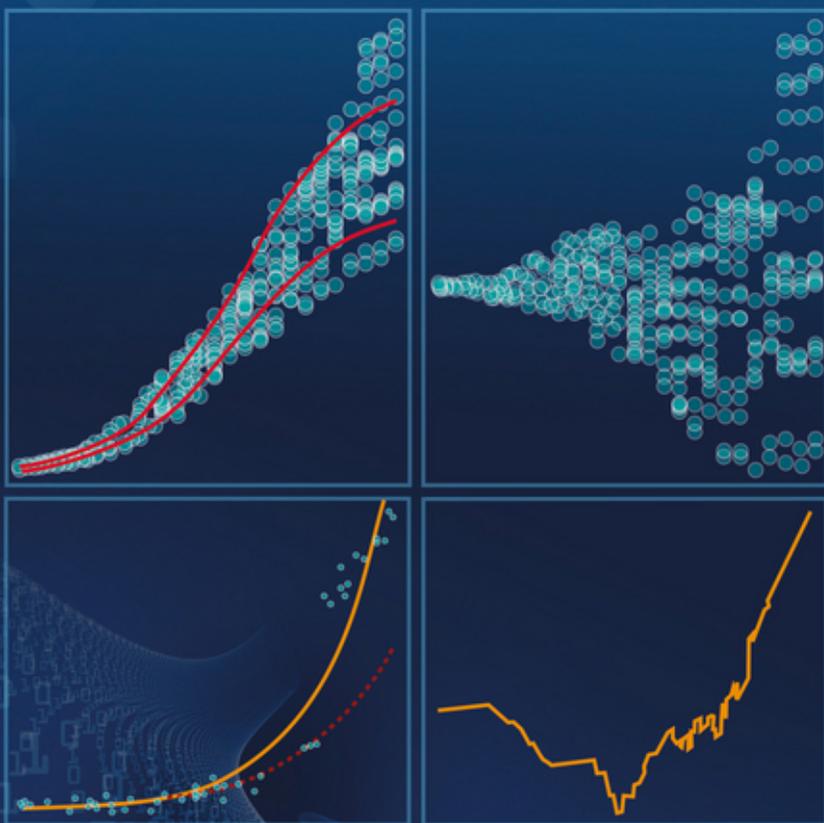


# Robust Nonlinear Regression

with Applications using R



Hossein Riazoshams  
Habshah Midi  
Gebrenegus Ghilagaber

WILEY

## **Robust Nonlinear Regression**



# **Robust Nonlinear Regression: with Applications using R**

*Hossein Riazoshams*

*Lamerd Islamic Azad University, Iran*

*Stockholm University, Sweden*

*University of Putra, Malaysia*

*Habshah Midi*

*University of Putra, Malaysia*

*Gebrenergus Ghilagaber*

*Stockholm University, Sweden*

**WILEY**

This edition first published 2019  
© 2019 John Wiley & Sons Ltd

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Hossein Riazoshams, Habshah Midi and Gebrenegus Ghilagaber to be identified as the authors of this work has been asserted in accordance with law.

*Registered Offices*

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

*Editorial Office*

9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at [www.wiley.com](http://www.wiley.com).

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

*Limit of Liability/Disclaimer of Warranty*

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

*Library of Congress Cataloging-in-Publication Data*

Names: Riazoshams, Hossein, 1971– author. | Midi, Habshah, author. | Ghilagaber, Gebrenegus, author.

Title: Robust nonlinear regression: with applications using R / Hossein Riazoshams, Habshah Midi, Gebrenegus Ghilagaber.

Description: Hoboken, NJ : John Wiley & Sons, 2018. | Includes bibliographical references and index. |

Identifiers: LCCN 2017057347 (print) | LCCN 2018005931 (ebook) | ISBN 9781119010456 (pdf) | ISBN 9781119010449 (epub) | ISBN 9781118738061 (cloth)

Subjects: LCSH: Regression analysis. | Nonlinear theories. | R (Computer program language)

Classification: LCC QA278.2 (ebook) | LCC QA278.2 .R48 2018 (print) | DDC 519.5/36–dc23

LC record available at <https://lccn.loc.gov/2017057347>

Cover Design: Wiley

Cover Image: © Wavebreakmedia Ltd/Getty Images; © Courtesy of Hossein Riazoshams

Set in 10/12pt WarnockPro by SPi Global, Chennai, India

*To my wife Benchamat Hanchana, from Hossein*



## Contents

<b>Preface</b>	<i>xi</i>
<b>Acknowledgements</b>	<i>xiii</i>
<b>About the Companion Website</b>	<i>xv</i>

### Part One Theories 1

<b>1</b>	<b>Robust Statistics and its Application in Linear Regression</b>	<b>3</b>
1.1	Robust Aspects of Data	3
1.2	Robust Statistics and the Mechanism for Producing Outliers	4
1.3	Location and Scale Parameters	5
1.3.1	Location Parameter	5
1.3.2	Scale Parameters	9
1.3.3	Location and Dispersion Models	10
1.3.4	Numerical Computation of M-estimates	11
1.4	Redescending M-estimates	13
1.5	Breakdown Point	13
1.6	Linear Regression	16
1.7	The Robust Approach in Linear Regression	19
1.8	S-estimator	23
1.9	Least Absolute and Quantile Estimates	25
1.10	Outlier Detection in Linear Regression	27
1.10.1	Studentized and Deleted Studentized Residuals	27
1.10.2	Hadi Potential	28
1.10.3	Elliptic Norm (Cook Distance)	28
1.10.4	Difference in Fits	29
1.10.5	Atkinson's Distance	29
1.10.6	DFBETAS	29

<b>2</b>	<b>Nonlinear Models: Concepts and Parameter Estimation</b>	<b>31</b>
2.1	Introduction	31
2.2	Basic Concepts	32
2.3	Parameter Estimations	34
2.3.1	Maximum Likelihood Estimators	34
2.3.2	The Ordinary Least Squares Method	36
2.3.3	Generalized Least Squares Estimate	38
2.4	A Nonlinear Model Example	39
<b>3</b>	<b>Robust Estimators in Nonlinear Regression</b>	<b>41</b>
3.1	Outliers in Nonlinear Regression	41
3.2	Breakdown Point in Nonlinear Regression	43
3.3	Parameter Estimation	44
3.4	Least Absolute and Quantile Estimates	44
3.5	Quantile Regression	45
3.6	Least Median of Squares	45
3.7	Least Trimmed Squares	47
3.8	Least Trimmed Differences	48
3.9	S-estimator	49
3.10	$\tau$ -estimator	50
3.11	MM-estimate	50
3.12	Environmental Data Examples	54
3.13	Nonlinear Models	55
3.14	Carbon Dioxide Data	61
3.15	Conclusion	64
<b>4</b>	<b>Heteroscedastic Variance</b>	<b>67</b>
4.1	Definitions and Notations	69
4.2	Weighted Regression for the Nonparametric Variance Model	69
4.3	Maximum Likelihood Estimates	71
4.4	Variance Modeling and Estimation	72
4.5	Robust Multistage Estimate	74
4.6	Least Squares Estimate of Variance Parameters	75
4.7	Robust Least Squares Estimate of the Structural Variance Parameter	78
4.8	Weighted M-estimate	79
4.9	Chicken-growth Data Example	80
4.10	Toxicology Data Example	85
4.11	Evaluation and Comparison of Methods	87
<b>5</b>	<b>Autocorrelated Errors</b>	<b>89</b>
5.1	Introduction	89
5.2	Nonlinear Autocorrelated Model	90

5.3	The Classic Two-stage Estimator	91
5.4	Robust Two-stage Estimator	92
5.5	Economic Data	93
5.6	ARIMA(1,0,1)(0,0,1) <sup>7</sup> Autocorrelation Function	103
<b>6</b>	<b>Outlier Detection in Nonlinear Regression</b>	<b>107</b>
6.1	Introduction	107
6.2	Estimation Methods	108
6.3	Point Influences	109
6.3.1	Tangential Plan Leverage	110
6.3.2	Jacobian Leverage	111
6.3.3	Generalized and Jacobian Leverages for M-estimator	112
6.4	Outlier Detection Measures	115
6.4.1	Studentized and Deletion Studentized Residuals	116
6.4.2	Hadi's Potential	117
6.4.3	Elliptic Norm (Cook Distance)	117
6.4.4	Difference in Fits	118
6.4.5	Atkinson's Distance	118
6.4.6	DFBETAS	118
6.4.7	Measures Based on Jacobian and MM-estimators	119
6.4.8	Robust Jacobian Leverage and Local Influences	119
6.4.9	Overview	121
6.5	Simulation Study	122
6.6	Numerical Example	128
6.7	Variance Heteroscedasticity	134
6.7.1	Heteroscedastic Variance Studentized Residual	136
6.7.2	Simulation Study, Heteroscedastic Variance	140
6.8	Conclusion	141

## Part Two Computations 143

<b>7</b>	<b>Optimization</b>	<b>145</b>
7.1	Optimization Overview	145
7.2	Iterative Methods	146
7.3	Wolfe Condition	148
7.4	Convergence Criteria	149
7.5	Mixed Algorithm	150
7.6	Robust M-estimator	150
7.7	The Generalized M-estimator	151
7.8	Some Mathematical Notation	151
7.9	Genetic Algorithm	152

<b>8</b>	<b>nlr Package</b>	<b>153</b>
8.1	Overview	153
8.2	nl.form Object	154
8.2.1	selfStart Initial Values	159
8.3	Model Fit by nlr	161
8.3.1	Output Objects, nl.fitt	164
8.3.2	Output Objects, nl.fitt.gn	167
8.3.3	Output Objects, nl.fitt.rob	169
8.3.4	Output Objects, nl.fitt.rgn	169
8.4	nlr.control	170
8.5	Fault Object	172
8.6	Ordinary Least Squares	172
8.7	Robust Estimators	175
8.8	Heteroscedastic Variance Case	179
8.8.1	Chicken-growth Data Example	179
8.8.2	National Toxicology Study Program Data	183
8.9	Autocorrelated Errors	184
8.10	Outlier Detection	193
8.11	Initial Values and Self-start	201
<b>9</b>	<b>Robust Nonlinear Regression in R</b>	<b>207</b>
9.1	Lakes Data Examples	207
9.2	Simulated Data Examples	211
<b>A</b>	<b>nlr Database</b>	<b>215</b>
A.1	Data Set used in the Book	215
A.1.1	Chicken-growth Data	216
A.1.2	Environmental Data	216
A.1.3	Lakes Data	218
A.1.4	Economic Data	221
A.1.5	National Toxicology Program (NTP) Data	223
A.1.6	Cow Milk Data	223
A.1.7	Simulated Outliers	225
A.1.8	Artificially Contaminated Data	227
A.2	Nonlinear Regression Models	227
A.3	Robust Loss Functions Data Bases	229
A.4	Heterogeneous Variance Models	229
	<b>References</b>	<b>233</b>
	<b>Index</b>	<b>239</b>

## Preface

This book is the result of the first author's research, between 2004 and 2016, in the robust nonlinear regression area, when he was affiliated with the institutions listed. The lack of computer programs together with mathematical development in this area encouraged us to write this book and provide an R-package called `nlr` for which a guide is provided in this book. The book concentrates more on applications and thus practical examples are presented.

Robust statistics describes the methods used when the classical assumptions of statistics do not hold. It is mostly applied when a data set includes outliers that lead to violation of the classical assumptions.

The book is divided into two parts. In Part 1, the mathematical theories of robust nonlinear regression are discussed and parameter estimation for heteroscedastic error variances, autocorrelated errors, and several methods for outlier detection are presented. Part 2 presents numerical methods and R-tools for nonlinear regression using robust methods.

In Chapter 1, the basic theories of robust statistics are discussed. Robust approaches to linear regression and outlier detection are presented. These mathematical concepts of robust statistics and linear regression are then extended to nonlinear regression in the rest of the book. Since the book is about nonlinear regression, the proofs of theorems related to robust linear regression are omitted.

Chapter 2 presents the concepts of nonlinear regression and discusses the theory behind several methods of parameter estimation in this area. The robust forms of these methods are outlined in Chapter 3. Chapter 2 presents the generalized least square estimate, which will be used for non-classical situations.

Chapter 3 discusses the concepts of robust statistics, such as robustness and breakdown points, in the context of nonlinear regression. It also presents several robust parameter estimation techniques.

Chapter 4 develops the robust methods for a null condition when the error variances are not homogeneous. Different kinds of outlier are defined and their effects are discussed. Parameter estimation for nonlinear function models and variance function models are presented.

Another null condition, when the errors are autocorrelated, is discussed in Chapter 5. Robust and classical methods for estimating the nonlinear function model and the autocorrelation structure of the error are presented. The effect of different kinds of outlier are explained, and appropriate methods for identifying the correlation structure of errors in the presence of outliers are studied.

Chapter 6 explains the methods for identifying atypical points. The outlier detection methods that are developed in this chapter are based mainly on statistical measures that use robust estimators of the parameters of the nonlinear function model.

In Chapter 7, optimization methods are discussed. These techniques are then modified to solve the minimization problems found in robust nonlinear regressions. They will then used to solve the mathematical problems discussed in Part 1 of the book and their implementation in a new R package called `n1r` is then covered in Chapter 8.

Chapter 8 is a guide to the R package implemented for this book. It covers object definition for a nonlinear function model, parameter estimation, and outlier detection for several model assumption situations discussed in the Part 1. This chapter shows how to fit nonlinear models to real-life and simulated data.

In Chapter 9, another R packages for robust nonlinear regression are presented and compared to `n1r`. Appendix A presents and describes the databases embedded in `n1r`, and the nonlinear models and functions available.

At the time of writing, the `n1r` package is complete, and is available at The Comprehensive R Archive Network (CRAN-project) at <https://cran.r-project.org/package=n1r>.

Because of the large number of figures and programs involved, there are many examples that could not be included in the book. Materials, programs, further examples, and a forum to share and discuss program bugs are all provided at the author's website at <http://www.riazoshams.com/n1r> and at the book's page on the Wiley website.

Response Manager, Shabdiz Music School of Iran,  
Full time faculty member of Islamic Azad University of Lamerd, Iran,  
Department of Statistics, Stockholm University, Sweden,  
Institute for Mathematical, Research University of Putra, Malaysia  
November 2017

*Hossein Riazoshams*

## Acknowledgements

I would like to thank the people and organizations who have helped me in all stages of my research that has culminated in this book. Firstly I would like to express my appreciation to Mohsen Ghodousi Zadeh and Hamid Koohbor for helping me in collecting data for the first time in 2005. This led me to a program of research in nonlinear modeling.

I would like to recognize the Department of Statistics at Stockholm University, Sweden, for financial support while writing most of this book during my stay as a post-doctoral researcher in 2012–2014.

A special note of appreciation is also due to the Islamic Azad University of Abadeh and Lamerd for financial support in connection with collecting some materials for this book.

I would like to note my appreciation for the Institute for Mathematical Research of University Putra Malaysia for financial support during my PhD in 2007–2010 and afterwards.

I owe my gratitude to the John Wiley editing team, specially Shyamala and others for their great editing process during the preparation of the book.

Last but by no means least, I would like to thank my wife, Benchamat Hanchan, for her great patience with the financial and physical adversity that we experienced during this research.

November 2017

*Hossein Riazoshams*

