

FINANCIAL SIGNAL PROCESSING AND MACHINE LEARNING

FINANCIAL SIGNAL PROCESSING AND MACHINE LEARNING

Edited by

Ali N. Akansu

New Jersey Institute of Technology, USA

Sanjeev R. Kulkarni

Princeton University, USA

Dmitry Malioutov

IBM T.J. Watson Research Center, USA

The logo for IEEE Press, featuring a stylized sunburst or fan-like graphic above the text "IEEE PRESS".

IEEE PRESS

WILEY

This edition first published 2016

© 2016 John Wiley & Sons, Ltd

First Edition published in 2016

Registered office

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

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

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 the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data applied for

ISBN: 9781118745670

A catalogue record for this book is available from the British Library.

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

Contents

List of Contributors	xiii
Preface	xv
1 Overview	1
<i>Ali N. Akansu, Sanjeev R. Kulkarni, and Dmitry Malioutov</i>	
1.1 Introduction	1
1.2 A Bird's-Eye View of Finance	2
1.2.1 Trading and Exchanges	4
1.2.2 Technical Themes in the Book	5
1.3 Overview of the Chapters	6
1.3.1 Chapter 2: "Sparse Markowitz Portfolios" by Christine De Mol	6
1.3.2 Chapter 3: "Mean-Reverting Portfolios: Tradeoffs between Sparsity and Volatility" by Marco Cuturi and Alexandre d'Aspremont	7
1.3.3 Chapter 4: "Temporal Causal Modeling" by Prabhanjan Kambadur, Aurélie C. Lozano, and Ronny Luss	7
1.3.4 Chapter 5: "Explicit Kernel and Sparsity of Eigen Subspace for the AR(1) Process" by Mustafa U. Torun, Onur Yilmaz and Ali N. Akansu	7
1.3.5 Chapter 6: "Approaches to High-Dimensional Covariance and Precision Matrix Estimation" by Jianqing Fan, Yuan Liao, and Han Liu	7
1.3.6 Chapter 7: "Stochastic Volatility: Modeling and Asymptotic Approaches to Option Pricing and Portfolio Selection" by Matthew Lorig and Ronnie Sircar	7
1.3.7 Chapter 8: "Statistical Measures of Dependence for Financial Data" by David S. Matteson, Nicholas A. James, and William B. Nicholson	8
1.3.8 Chapter 9: "Correlated Poisson Processes and Their Applications in Financial Modeling" by Alexander Kreinin	8
1.3.9 Chapter 10: "CVaR Minimizations in Support Vector Machines" by Junya Gotoh and Akiko Takeda	8
1.3.10 Chapter 11: "Regression Models in Risk Management" by Stan Uryasev	8
1.4 Other Topics in Financial Signal Processing and Machine Learning	9
References	9

2	Sparse Markowitz Portfolios	11
	<i>Christine De Mol</i>	
2.1	Markowitz Portfolios	11
2.2	Portfolio Optimization as an Inverse Problem: The Need for Regularization	13
2.3	Sparse Portfolios	15
2.4	Empirical Validation	17
2.5	Variations on the Theme	18
	2.5.1 <i>Portfolio Rebalancing</i>	18
	2.5.2 <i>Portfolio Replication or Index Tracking</i>	19
	2.5.3 <i>Other Penalties and Portfolio Norms</i>	19
2.6	Optimal Forecast Combination	20
	Acknowledgments	21
	References	21
3	Mean-Reverting Portfolios	23
	<i>Marco Cuturi and Alexandre d'Aspremont</i>	
3.1	Introduction	23
	3.1.1 <i>Synthetic Mean-Reverting Baskets</i>	24
	3.1.2 <i>Mean-Reverting Baskets with Sufficient Volatility and Sparsity</i>	24
3.2	Proxies for Mean Reversion	25
	3.2.1 <i>Related Work and Problem Setting</i>	25
	3.2.2 <i>Predictability</i>	26
	3.2.3 <i>Portmanteau Criterion</i>	27
	3.2.4 <i>Crossing Statistics</i>	28
3.3	Optimal Baskets	28
	3.3.1 <i>Minimizing Predictability</i>	29
	3.3.2 <i>Minimizing the Portmanteau Statistic</i>	29
	3.3.3 <i>Minimizing the Crossing Statistic</i>	29
3.4	Semidefinite Relaxations and Sparse Components	30
	3.4.1 <i>A Semidefinite Programming Approach to Basket Estimation</i>	30
	3.4.2 <i>Predictability</i>	30
	3.4.3 <i>Portmanteau</i>	31
	3.4.4 <i>Crossing Stats</i>	31
3.5	Numerical Experiments	32
	3.5.1 <i>Historical Data</i>	32
	3.5.2 <i>Mean-reverting Basket Estimators</i>	33
	3.5.3 <i>Jurek and Yang (2007) Trading Strategy</i>	33
	3.5.4 <i>Transaction Costs</i>	33
	3.5.5 <i>Experimental Setup</i>	36
	3.5.6 <i>Results</i>	36
3.6	Conclusion	39
	References	39

4	Temporal Causal Modeling	41
	<i>Prabhanjan Kambadur, Aurélie C. Lozano, and Ronny Luss</i>	
4.1	Introduction	41
4.2	TCM	46
4.2.1	<i>Granger Causality and Temporal Causal Modeling</i>	46
4.2.2	<i>Grouped Temporal Causal Modeling Method</i>	47
4.2.3	<i>Synthetic Experiments</i>	49
4.3	Causal Strength Modeling	51
4.4	Quantile TCM (Q-TCM)	52
4.4.1	<i>Modifying Group OMP for Quantile Loss</i>	52
4.4.2	<i>Experiments</i>	53
4.5	TCM with Regime Change Identification	55
4.5.1	<i>Model</i>	56
4.5.2	<i>Algorithm</i>	58
4.5.3	<i>Synthetic Experiments</i>	60
4.5.4	<i>Application: Analyzing Stock Returns</i>	62
4.6	Conclusions	63
	References	64
5	Explicit Kernel and Sparsity of Eigen Subspace for the AR(1) Process	67
	<i>Mustafa U. Torun, Onur Yilmaz, and Ali N. Akansu</i>	
5.1	Introduction	67
5.2	Mathematical Definitions	68
5.2.1	<i>Discrete AR(1) Stochastic Signal Model</i>	68
5.2.2	<i>Orthogonal Subspace</i>	69
5.3	Derivation of Explicit KLT Kernel for a Discrete AR(1) Process	72
5.3.1	<i>A Simple Method for Explicit Solution of a Transcendental Equation</i>	73
5.3.2	<i>Continuous Process with Exponential Autocorrelation</i>	74
5.3.3	<i>Eigenanalysis of a Discrete AR(1) Process</i>	76
5.3.4	<i>Fast Derivation of KLT Kernel for an AR(1) Process</i>	79
5.4	Sparsity of Eigen Subspace	82
5.4.1	<i>Overview of Sparsity Methods</i>	83
5.4.2	<i>pdf-Optimized Midtread Quantizer</i>	84
5.4.3	<i>Quantization of Eigen Subspace</i>	86
5.4.4	<i>pdf of Eigenvector</i>	87
5.4.5	<i>Sparse KLT Method</i>	89
5.4.6	<i>Sparsity Performance</i>	91
5.5	Conclusions	97
	References	97

6	Approaches to High-Dimensional Covariance and Precision Matrix Estimations	100
	<i>Jianqing Fan, Yuan Liao, and Han Liu</i>	
6.1	Introduction	100
6.2	Covariance Estimation via Factor Analysis	101
6.2.1	<i>Known Factors</i>	103
6.2.2	<i>Unknown Factors</i>	104
6.2.3	<i>Choosing the Threshold</i>	105
6.2.4	<i>Asymptotic Results</i>	105
6.2.5	<i>A Numerical Illustration</i>	107
6.3	Precision Matrix Estimation and Graphical Models	109
6.3.1	<i>Column-wise Precision Matrix Estimation</i>	110
6.3.2	<i>The Need for Tuning-insensitive Procedures</i>	111
6.3.3	<i>TIGER: A Tuning-insensitive Approach for Optimal Precision Matrix Estimation</i>	112
6.3.4	<i>Computation</i>	114
6.3.5	<i>Theoretical Properties of TIGER</i>	114
6.3.6	<i>Applications to Modeling Stock Returns</i>	115
6.3.7	<i>Applications to Genomic Network</i>	118
6.4	Financial Applications	119
6.4.1	<i>Estimating Risks of Large Portfolios</i>	119
6.4.2	<i>Large Panel Test of Factor Pricing Models</i>	121
6.5	Statistical Inference in Panel Data Models	126
6.5.1	<i>Efficient Estimation in Pure Factor Models</i>	126
6.5.2	<i>Panel Data Model with Interactive Effects</i>	127
6.5.3	<i>Numerical Illustrations</i>	130
6.6	Conclusions	131
	References	131
7	Stochastic Volatility	135
	<i>Matthew Lorig and Ronnie Sircar</i>	
7.1	Introduction	135
7.1.1	<i>Options and Implied Volatility</i>	136
7.1.2	<i>Volatility Modeling</i>	137
7.2	Asymptotic Regimes and Approximations	141
7.2.1	<i>Contract Asymptotics</i>	142
7.2.2	<i>Model Asymptotics</i>	142
7.2.3	<i>Implied Volatility Asymptotics</i>	143
7.2.4	<i>Tractable Models</i>	145
7.2.5	<i>Model Coefficient Polynomial Expansions</i>	146
7.2.6	<i>Small “Vol of Vol” Expansion</i>	152
7.2.7	<i>Separation of Timescales Approach</i>	152
7.2.8	<i>Comparison of the Expansion Schemes</i>	154
7.3	Merton Problem with Stochastic Volatility: Model Coefficient Polynomial Expansions	155

7.3.1	<i>Models and Dynamic Programming Equation</i>	155
7.3.2	<i>Asymptotic Approximation</i>	157
7.3.3	<i>Power Utility</i>	159
7.4	Conclusions	160
	Acknowledgements	160
	References	160
8	Statistical Measures of Dependence for Financial Data	162
	<i>David S. Matteson, Nicholas A. James, and William B. Nicholson</i>	
8.1	Introduction	162
8.2	Robust Measures of Correlation and Autocorrelation	164
8.2.1	<i>Transformations and Rank-Based Methods</i>	166
8.2.2	<i>Inference</i>	169
8.2.3	<i>Misspecification Testing</i>	171
8.3	Multivariate Extensions	174
8.3.1	<i>Multivariate Volatility</i>	175
8.3.2	<i>Multivariate Misspecification Testing</i>	176
8.3.3	<i>Granger Causality</i>	176
8.3.4	<i>Nonlinear Granger Causality</i>	177
8.4	Copulas	179
8.4.1	<i>Fitting Copula Models</i>	180
8.4.2	<i>Parametric Copulas</i>	181
8.4.3	<i>Extending beyond Two Random Variables</i>	183
8.4.4	<i>Software</i>	185
8.5	Types of Dependence	185
8.5.1	<i>Positive and Negative Dependence</i>	185
8.5.2	<i>Tail Dependence</i>	187
	References	188
9	Correlated Poisson Processes and Their Applications in Financial Modeling	191
	<i>Alexander Kreinin</i>	
9.1	Introduction	191
9.2	Poisson Processes and Financial Scenarios	193
9.2.1	<i>Integrated Market–Credit Risk Modeling</i>	193
9.2.2	<i>Market Risk and Derivatives Pricing</i>	194
9.2.3	<i>Operational Risk Modeling</i>	194
9.2.4	<i>Correlation of Operational Events</i>	195
9.3	Common Shock Model and Randomization of Intensities	196
9.3.1	<i>Common Shock Model</i>	196
9.3.2	<i>Randomization of Intensities</i>	196
9.4	Simulation of Poisson Processes	197
9.4.1	<i>Forward Simulation</i>	197
9.4.2	<i>Backward Simulation</i>	200
9.5	Extreme Joint Distribution	207

9.5.1	<i>Reduction to Optimization Problem</i>	207
9.5.2	<i>Monotone Distributions</i>	208
9.5.3	<i>Computation of the Joint Distribution</i>	214
9.5.4	<i>On the Frechet–Hoeffding Theorem</i>	215
9.5.5	<i>Approximation of the Extreme Distributions</i>	217
9.6	Numerical Results	219
9.6.1	<i>Examples of the Support</i>	219
9.6.2	<i>Correlation Boundaries</i>	221
9.7	Backward Simulation of the Poisson–Wiener Process	222
9.8	Concluding Remarks	227
	Acknowledgments	228
Appendix A		229
A.1	Proof of Lemmas 9.2 and 9.3	229
A.1.1	<i>Proof of Lemma 9.2</i>	229
A.1.2	<i>Proof of Lemma 9.3</i>	230
	References	231
10	CVaR Minimizations in Support Vector Machines	233
	<i>Jun-ya Gotoh and Akiko Takeda</i>	
10.1	What Is CVaR?	234
10.1.1	<i>Definition and Interpretations</i>	234
10.1.2	<i>Basic Properties of CVaR</i>	238
10.1.3	<i>Minimization of CVaR</i>	240
10.2	Support Vector Machines	242
10.2.1	<i>Classification</i>	242
10.2.2	<i>Regression</i>	246
10.3	ν -SVMs as CVaR Minimizations	247
10.3.1	<i>ν-SVMs as CVaR Minimizations with Homogeneous Loss</i>	247
10.3.2	<i>ν-SVMs as CVaR Minimizations with Nonhomogeneous Loss</i>	251
10.3.3	<i>Refining the ν-Property</i>	253
10.4	Duality	256
10.4.1	<i>Binary Classification</i>	256
10.4.2	<i>Geometric Interpretation of ν-SVM</i>	257
10.4.3	<i>Geometric Interpretation of the Range of ν for ν-SVC</i>	258
10.4.4	<i>Regression</i>	259
10.4.5	<i>One-class Classification and SVDD</i>	259
10.5	Extensions to Robust Optimization Modelings	259
10.5.1	<i>Distributionally Robust Formulation</i>	259
10.5.2	<i>Measurement-wise Robust Formulation</i>	261
10.6	Literature Review	262
10.6.1	<i>CVaR as a Risk Measure</i>	263
10.6.2	<i>From CVaR Minimization to SVM</i>	263
10.6.3	<i>From SVM to CVaR Minimization</i>	263
10.6.4	<i>Beyond CVaR</i>	263
	References	264

11	Regression Models in Risk Management	266
	<i>Stan Uryasev</i>	
11.1	Introduction	267
11.2	Error and Deviation Measures	268
11.3	Risk Envelopes and Risk Identifiers	271
	<i>11.3.1 Examples of Deviation Measures \mathcal{D}, Corresponding Risk Envelopes \mathcal{Q}, and Sets of Risk Identifiers $\mathcal{QD}(X)$</i>	<i>272</i>
11.4	Error Decomposition in Regression	273
11.5	Least-Squares Linear Regression	275
11.6	Median Regression	277
11.7	Quantile Regression and Mixed Quantile Regression	281
11.8	Special Types of Linear Regression	283
11.9	Robust Regression	284
	References, Further Reading, and Bibliography	287
Index		289

List of Contributors

Ali N. Akansu, New Jersey Institute of Technology, USA

Marco Cuturi, Kyoto University, Japan

Alexandre d'Aspremont, CNRS - Ecole Normale supérieure, France

Christine De Mol, Université Libre de Bruxelles, Belgium

Jianqing Fan, Princeton University, USA

Jun-ya Gotoh, Chuo University, Japan

Nicholas A. James, Cornell University, USA

Prabhanjan Kambadur, Bloomberg L.P., USA

Alexander Kreinin, Risk Analytics, IBM, Canada

Sanjeev R. Kulkarni, Princeton University, USA

Yuan Liao, University of Maryland, USA

Han Liu, Princeton University, USA

Matthew Lorig, University of Washington, USA

Aurélié C. Lozano, IBM T.J. Watson Research Center, USA

Ronny Luss, IBM T.J. Watson Research Center, USA

Dmitry Malioutov, IBM T.J. Watson Research Center, USA

David S. Matteson, Cornell University, USA

William B. Nicholson, Cornell University, USA

Ronnie Sircar, Princeton University, USA

Akiko Takeda, The University of Tokyo, Japan

Mustafa U. Torun, New Jersey Institute of Technology, USA

Stan Uryasev, University of Florida, USA

Onur Yilmaz, New Jersey Institute of Technology, USA

Preface

This edited volume collects and unifies a number of recent advances in the signal-processing and machine-learning literature with significant applications in financial risk and portfolio management. The topics in the volume include characterizing statistical dependence and correlation in high dimensions, constructing effective and robust risk measures, and using these notions of risk in portfolio optimization and rebalancing through the lens of convex optimization. It also presents signal-processing approaches to model return, momentum, and mean reversion, including both theoretical and implementation aspects. Modern finance has become global and highly interconnected. Hence, these topics are of great importance in portfolio management and trading, where the financial industry is forced to deal with large and diverse portfolios in a variety of asset classes. The investment universe now includes tens of thousands of international equities and corporate bonds, and a wide variety of other interest rate and derivative products—often with limited, sparse, and noisy market data.

Using traditional risk measures and return forecasting (such as historical sample covariance and sample means in Markowitz theory) in high-dimensional settings is fraught with peril for portfolio optimization, as widely recognized by practitioners. Tools from high-dimensional statistics, such as factor models, eigen-analysis, and various forms of regularization that are widely used in real-time risk measurement of massive portfolios and for designing a variety of trading strategies including statistical arbitrage, are highlighted in the book. The dramatic improvements in computational power and special-purpose hardware such as field programmable gate arrays (FPGAs) and graphics processing units (GPUs) along with low-latency data communications facilitate the realization of these sophisticated financial algorithms that not long ago were “hard to implement.”

The book covers a number of topics that have been popular recently in machine learning and signal processing to solve problems with large portfolios. In particular, the connections between the portfolio theory and sparse learning and compressed sensing, robust optimization, non-Gaussian data-driven risk measures, graphical models, causal analysis through temporal-causal modeling, and large-scale copula-based approaches are highlighted in the book.

Although some of these techniques already have been used in finance and reported in journals and conferences of different disciplines, this book attempts to give a unified treatment from a common mathematical perspective of high-dimensional statistics and convex optimization. Traditionally, the academic quantitative finance community did not have much overlap with the signal and information-processing communities. However, the fields are seeing more interaction, and this trend is accelerating due to the paradigm in the financial sector which has

embraced state-of-the-art, high-performance computing and signal-processing technologies. Thus, engineers play an important role in this financial ecosystem. The goal of this edited volume is to help to bridge the divide, and to highlight machine learning and signal processing as disciplines that may help drive innovations in quantitative finance and electronic trading, including high-frequency trading.

The reader is assumed to have graduate-level knowledge in linear algebra, probability, and statistics, and an appreciation for the key concepts in optimization. Each chapter provides a list of references for readers who would like to pursue the topic in more depth. The book, complemented with a primer in financial engineering, may serve as the main textbook for a graduate course in financial signal processing.

We would like to thank all the authors who contributed to this volume as well as all of the anonymous reviewers who provided valuable feedback on the chapters in this book. We also gratefully acknowledge the editors and staff at Wiley for their efforts in bringing this project to fruition.