

# FINANCIAL SIGNAL PROCESSING AND MACHINE LEARNING

EDITED BY

ALI N. AKANSU SANJEEV R. KULKARNI DMITRY M. MALIOUTOV



 $\overline{\mathsf{WILEY}}$ 

## 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



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. Printed and bound in Singapore by Markono Print Media Pte Ltd

1 2016

## FINANCIAL SIGNAL PROCESSING AND MACHINE LEARNING

此为试读,需要完整PDF请访问: www.ertongbook.com

#### 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

Aurélie 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

xiv List of Contributors

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

xvi Preface

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.

List	ist of Contributors		xiii
Pref	ace		xv
1	Overvi		1
	Ali N. A	Akansu, Sanjeev R. Kulkarni, and Dmitry Malioutov	
1.1	Introdu		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	Overvie	ew 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		opics in Financial Signal Processing and Machine Learning	9
	Referen	ices	9

vi Contents

2	- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	e Markowitz Portfolios ine De Mol	11
2.1	Marko	witz Portfolios	11
2.2	Portfol	lio Optimization as an Inverse Problem: The Need for Regularization	13
2.3		Portfolios	15
2.4		cal Validation	17
2.5	Variations on the Theme		18
	2.5.1	Portfolio Rebalancing	18
	2.5.2	•	19
	2.5.3	Other Penalties and Portfolio Norms	19
2.6		Optimal Forecast Combination	
	_	wlegments	20 21
	Refere	-	21
3	Mean-Reverting Portfolios		23
	Marco	Cuturi and Alexandre d'Aspremont	
3.1	Introduction		
	3.1.1	Synthetic Mean-Reverting Baskets	23 24
	3.1.2	Mean-Reverting Baskets with Sufficient Volatility and Sparsity	24
3.2	Proxies for Mean Reversion		25
	3.2.1		25
	3.2.2	8	26
	3.2.3		27
	3.2.4	Crossing Statistics	28
3.3	Optimal Baskets		28
	3.3.1	Minimizing Predictability	29
	3.3.2		29
	3.3.3	Minimizing the Crossing Statistic	29
3.4	Semidefinite Relaxations and Sparse Components		30
	3.4.1	A Semidefinite Programming Approach to Basket Estimation	30
	3.4.2		30
	3.4.3	•	31
	3.4.4	Crossing Stats	31
3.5	Numerical Experiments		32
	3.5.1	Historical Data	32
	3.5.2	Mean-reverting Basket Estimators	33
	3.5.3	Jurek and Yang (2007) Trading Strategy	33
	3.5.4	Transaction Costs	33
	3.5.5	Experimental Setup	36
	3.5.6	Results	36
3.6	Conclusion		39
	References		39

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

6		aches to High-Dimensional Covariance and Precision x Estimations	100
	Jianqii	ng Fan, Yuan Liao, and Han Liu	
6.1	Introdu	action	100
6.2	Covari	ance Estimation via Factor Analysis	101
	6.2.1	Known Factors	103
	6.2.2	Unknown Factors	104
	6.2.3		105
	6.2.4	Asymptotic Results	105
	6.2.5	A Numerical Illustration	107
6.3	Precision Matrix Estimation and Graphical Models		
0.0	6.3.1 Column-wise Precision Matrix Estimation		
	6.3.2	The Need for Tuning-insensitive Procedures	110 111
	6.3.3	TIGER: A Tuning-insensitive Approach for Optimal Precision	
	0.5.5	Matrix Estimation	112
	6.3.4	Computation	114
	6.3.5	Theoretical Properties of TIGER	114
	6.3.6	Applications to Modeling Stock Returns	115
	6.3.7	Applications to Genomic Network	118
6.4		ial Applications	119
0.1	6.4.1	Estimating Risks of Large Portfolios	119
	6.4.2	Large Panel Test of Factor Pricing Models	121
6.5		cal Inference in Panel Data Models	126
0.5	6.5.1	Efficient Estimation in Pure Factor Models	126
	6.5.2	Panel Data Model with Interactive Effects	127
	6.5.3	Numerical Illustrations	130
6.6	Conclu		131
0.0	References		131
	Refere	nices	131
7	Stochastic Volatility		135
		w Lorig and Ronnie Sircar	
7.1	Introdu	-	135
7.1	7.1.1	Options and Implied Volatility	136
	7.1.2	Volatility Modeling	137
7.2		ototic Regimes and Approximations	141
1.2	7.2.1	Contract Asymptotics	142
	7.2.2	Model Asymptotics	142
	7.2.3	Implied Volatility Asymptotics	143
	7.2.4	Tractable Models	145
	7.2.5	Model Coefficient Polynomial Expansions	146
	7.2.6	Small "Vol of Vol" Expansion	152
	7.2.7	Separation of Timescales Approach	152
	7.2.8	Comparison of the Expansion Schemes	154
7.3		Problem with Stochastic Volatility: Model Coefficient Polynomial	154
	Expansions		

	7.3.1	Models and Dynamic Programming Equation	155
	7.3.2		157
	7.3.3	Power Utility	159
7.4	Concli	·	160
		wledgements	160
	Refere		160
8	Statis	tical Massaves of Danandanas for Einancial Data	162
0		tical Measures of Dependence for Financial Data S. Matteson, Nicholas A. James, and William B. Nicholson	102
8.1	Introd	uction	162
8.2	Robus	t Measures of Correlation and Autocorrelation	164
	8.2.1	Transformations and Rank-Based Methods	166
	8.2.2	Inference	169
	8.2.3	Misspecification Testing	171
8.3	Multiv	rariate Extensions	174
	8.3.1	Multivariate Volatility	175
	8.3.2	Multivariate Misspecification Testing	176
	8.3.3	Granger Causality	176
	8.3.4	Nonlinear Granger Causality	177
8.4	Copulas		
	8.4.1	Fitting Copula Models	180
	8.4.2	Parametric Copulas	181
	8.4.3	Extending beyond Two Random Variables	183
	8.4.4	Software	185
8.5	Types	of Dependence	185
	8.5.1	Positive and Negative Dependence	185
	8.5.2	Tail Dependence	187
	Refere	nces	188
9	Corre	lated Poisson Processes and Their Applications in Financial	
	Model		191
		der Kreinin	
9.1	Introdu	action	191
9.2	Poisso	n Processes and Financial Scenarios	193
	9.2.1	Integrated Market-Credit Risk Modeling	193
	9.2.2	Market Risk and Derivatives Pricing	194
	9.2.3	Operational Risk Modeling	194
	9.2.4	Correlation of Operational Events	195
9.3	Common Shock Model and Randomization of Intensities		
	9.3.1	Common Shock Model	196
	9.3.2	Randomization of Intensities	196
9.4		tion of Poisson Processes	197
	9.4.1	Forward Simulation	197
	9.4.2	Backward Simulation	200
9.5	Extrem	ne Joint Distribution	207

X Contents

	9.5.1	Reduction to Optimization Problem	207
	9.5.2	Monotone Distributions	208
	9.5.3	Computation of the Joint Distribution	214
	9.5.4	On the Frechet–Hoeffding Theorem	215
	9.5.5	Approximation of the Extreme Distributions	217
9.6	Numeri	ical Results	219
	9.6.1	Examples of the Support	219
	9.6.2	Correlation Boundaries	221
9.7	Backwa	ard Simulation of the Poisson-Wiener Process	222
9.8	Conclu	ding Remarks	227
	Acknow	wledgments	228
Appe	ndix A		229
A.1	Proof o	f Lemmas 9.2 and 9.3	229
	A.1.1	Proof of Lemma 9.2	229
	A.1.2	Proof of Lemma 9.3	230
	Referer	nces	231
10	CVaR 1	Minimizations in Support Vector Machines	233
		Gotoh and Akiko Takeda	
10.1		s CVaR?	234
10.1		Definition and Interpretations	234
	10.1.2	Basic Properties of CVaR	238
	10.1.3	Minimization of CVaR	240
10.2		t Vector Machines	242
10.2	10.2.1	Classification	242
	10.2.2	Regression	246
10.3		s as CVaR Minimizations	247
10.0		v-SVMs as CVaR Minimizations with Homogeneous Loss	247
		v-SVMs as CVaR Minimizations with Nonhomogeneous Loss	251
	10.3.3	Refining the v-Property	253
10.4	Duality		256
1011	10.4.1	Binary Classification	256
	10.4.2	Geometric Interpretation of v-SVM	257
	10.4.3	Geometric Interpretation of the Range of v for v-SVC	258
	10.4.4	Regression	259
	10.4.5	One-class Classification and SVDD	259
10.5	Extensions to Robust Optimization Modelings		
	10.5.1	Distributionally Robust Formulation	259
	10.5.2	Measurement-wise Robust Formulation	261
10.6	Literatu	ire Review	262
	10.6.1	CVaR as a Risk Measure	263
	10.6.2	From CVaR Minimization to SVM	263
	10.6.3	From SVM to CVaR Minimization	263
	10.6.4	Beyond CVaR	263
	Referen		264

11	Regression Models in Risk Management Stan Uryasev	266
11.1	Introduction	267
11.2	Error and Deviation Measures	268
11.3	Risk Envelopes and Risk Identifiers	271
	11.3.1 Examples of Deviation Measures D, Corresponding Risk Envelopes	
	Q, and Sets of Risk Identifiers $QD(X)$	272
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	
	References, Further Reading, and Bibliography	287
Index		289