DIGITAL SIGNAL AND IMAGE PROCESSING SERIES

Multi-factor Models and Signal Processing Techniques

Application to Quantitative Finance

Serge Darolles, Patrick Duvaut
Emmanuelle Jay

ISTE

WILEY

Multi-factor Models and Signal Processing Techniques

Application to Quantitative Finance

常州大学图书馆
藏书章

Serge Darolles
Patrick Duvaut
Emmanuelle Jay

ISTE

WILEY

First published 2013 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

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

www.wiley.com

© ISTE Ltd 2013
The rights of Serge Darolles, Patrick Duvaut and Emmanuelle Jay to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Library of Congress Control Number: 2013941318

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN: 978-1-84821-419-4

MIX
Paper from responsible sources
FSC® C013604

Printed and bound in Great Britain by CPI Group (UK) Ltd., Croydon, Surrey CR0 4YY

Multi-factor Models and Signal Processing Techniques

Foreword

With the irruption of electronic markets and the widespread use of quantitative models for trading, order execution and risk management, financial markets have evolved from the low-tech environment where they started from a few decades ago to the one where market participants routinely employ advanced data processing techniques to analyze large data flows in real time and feed the results into trading algorithms and risk management systems. Operational constraints arising from the sheer volume of data to be analyzed have created the need for intelligent data processing and modeling methods and parsimonious models whose estimation and simulation is feasible in the high-dimensional settings represented by the large number of risk factors involved in risk decisions.

In this context, advanced tools and methods from signal processing – the branch of science and engineering that deals with the modeling and analysis of complex streams of data – are increasingly used to analyze financial data and estimate the statistical models used by portfolio managers and risk managers.

In contrast to the traditional toolkit of portfolio theory, which is mostly based on *static* analysis of portfolio loss

此为试读，需要完整PDF请访问：www.ertongbook.com

distributions, the use of models based on stochastic process allows risk managers a dynamic, and more realistic, view of portfolio risk which is dynamically updated as new information is revealed. However, the use of such dynamic models comes at a price: we need more advanced tools to estimate, simulate and analyze portfolio risk in these models. Tools that have long been used in engineering, such as the Kalman filter, but may be unfamiliar to some risk management professionals, turn out to be relevant when moving beyond the static setting.

Simultaneously, the use of high-dimensional models entail new estimation problems: in many asset allocation problems, one has a few years (so, a few hundreds of daily observations) to estimate the parameters describing the evolution of hundreds, if not thousands, of risk factors. In this context, traditional estimation methods, such as principal component analysis and basic covariance analysis, fail to be consistent and yield unexpected (and biased!) results ... Tools such as random matrix theory, initially developed in nuclear physics to model the energy levels of complex nuclei, have been recently used to explore such estimation biases and propose methods for overcoming them.

Leveraging on their double profile of quantitative risk professionals with a strong research and technical background, Serge Darolles, Patrick Duvaut and Emmanuelle Jay have done an excellent job in presenting a nice selection of advanced signal processing tools and statistical models that address some of these points in a self-contained manner, at a level accessible to students and professionals involved with quantitative risk management.

By drawing on their professional experience in quantitative modeling to achieve the right level of technical sophistication, Serge, Patrick and Emmanuelle have succeeded in covering a variety of relevant concepts without sacrificing the technical details necessary for a serious implementation: multi-factor models, advanced estimation methods based on random matrix theory, filtering algorithms and regularization methods for identification of latent risk factors. The use of these tools in risk management and asset allocation is nicely illustrated through examples using financial data.

I trust that the hard work done by the authors will benefit a wide range of readers across academic and professional circles and act as an incentive for readers to delve into the vast research literature on statistical modeling in finance. *Bonne lecture!*

Rama CONT
London
June 2013

Introduction

I.1. Digital society and new paradigms in fund management

The massive digitization of all our activities from entertainment to all public and economical domains marks the era of the digital society. This latter generates an overwhelming amount of data, measured in zetta (10^{21}) bytes.

This phenomenon called *Big Data* impacts the finance sphere as well. The so-called *five Vs* need to be handled by fund management companies: volume, velocity, variety, visualization and value. Moreover, the 2008 mortgage crisis and its macroeconomic effects had two major consequences that make the situation even more difficult for fund managers. First, these large amounts of data are highly *hectic* (non-stationary), even more unpredictable and noisy. Second, the fund management paradigm has been shifted from an "abstruse gain race" at any risk to being able to control the risks (draw down) while securing a minimum gain, with a maximum transparency.

One way to address the above mentioned challenging issues is to provide fund managers with appropriate and "user friendly" QUANT(ITATIVE) tools that are market driven. This is where Factor Models, Statistical Signal Processing (SSP) and Multi-factor Models and Signal Processing Techniques come into play.

I.2. Statistical signal processing and factor models in finance

The linear factor model is nowadays a benchmark in portfolio management theory and in the understanding of asset returns, even in the hedge fund industry where returns may have high nonlinearities. The idea is to relate a large number and variety of returns to a few relevant *factors*, up to Gaussian errors that cannot be described solely by the factors. The factors thus drive and parsimoniously explain the *spatial correlation between* assets, while the errors are *idiosyncratic to each* asset.

Factors need to be chosen and/or identified first, then the parameters of the linear model that connect them to the assets, such as the *alphas* (gains) and the *betas* (assets exposures) have to be estimated and tracked. Based on a linear regression fed by market data, linear factor models are thus suitable for all the requirements listed above of friendliness to the users, transparency and as being market driven.

Statistical signal processing (SSP) methods allow us to access both the factors and the model parameters, in a very noisy, highly unpredictable and hectic environment. SSP methods have proved their ergonomy, efficiency and practicality in highly hostile and changing conditions in the past few decades when applied to medical diagnostics, non-destructive testing of nuclear plants, aircrafts, cars and in boosting both the performance and the robustness of mobile phones, with some top achievements in the 4G Long Term Evolution (LTE) standards.

The aim of this book is to share the same when dealing with fund management and more specifically when facing the identification of linear factor models. While application of SSP to finance has been ongoing for a few years, it is still a big challenge due to the difficulty of relying on joint expertise in fund management, statistical signal processing and being at the crossroad of Academia and Business, to make the recent results of research usable for fund managers.

I.3. Multiexpertise team of authors

The three authors' complementary expertise and experience meet the challenging above mentioned conditions. First, the three of them have both academia and business experience. Second, two of them are experienced fund managers, while the third was an executive of a US company involved in asset capitalization. Moreover, the three authors have been working together many years, as partners of the same Company QAMLab©, involved in quantitative solutions for fund management. It is worth mentioning that the authors' expertise range from mathematical finance to SSP.

This book is thus the outcome of such a unique pluridisciplinary and market-oriented expertise.

I.4. Book positioning and targeted readers

The book targets three worldwide audiences. First, graduate students of Business, Finance, Management Masters, including, of course, MBAs who would like to understand how SSP methods can leverage linear factor models in the complicated and hectic digital asset management activities of the emerging digital society. Second, SSP grad students who would like to ramp up in quantitative finance following familiar tracks of least squares, Kalman filtering (KF) and more advanced tools based on most recent regularized Kalman approaches. Third, qualitative fund managers who are eager to add to their daily management softwares and practices, several standalone and/or end-to-end SSP methods in order to enhance and boost their own management style, in a very practical, transparent and market-oriented manner.

The book material is unique since it presents QUANT tools as practical answers to market driven issues faced by fund managers. First in the associated chapters, the SSP material is introduced from the users' stand-point, answering the questions, why do we need these tools, how do they work, how do we use them, how can they be tuned to the real data? The derivations of the results are provided in appendices and can be skipped for user-only readers. On the other hand, SSP skillful audience can access quickly this advanced material. Second, numerous examples and real market data are used through the whole book to better understand the SSP methods and their pluses when dealing with linear factor models. Finally, highlights emphasize the most relevant contents at the end of each chapter, to allow quick access.

It is important to note that the book contents were already taught by all the authors, both in finance and Information Science Masters. Feedback from the students has thus been a valuable input to improve the clarity and the depth of the writing.

I.5. Book organization

The book is organized into four chapters.

Chapter 1 introduces the factor model foundations of the whole book, the equations, the notations, the main concepts, the origin, the applications, etc. While most of it can be

此为试读,需要完整PDF请访问: www.ertongbook.com

skipped by readers familiar with fund management, it is nevertheless valuable to check it to master the model that will be used in the remainder of the book.

Chapter 2 presents qualitative and quantitative ways to choose the factors. While qualitative approaches leverage pure markets know-how, quantitative methods are based on the so-called *eigenfactors*, derived from the eigen-decomposition of the returns covariance matrix. Statistical methods to select the "optimum" number of factors are also detailed. Even SSP readers can enrich their knowledge by grasping the connection between eigen-methods and the determination of the appropriate number of factors.

Least squares and Kalman algorithms that aim to estimate the alphas and the betas of the factor model are the main focus of Chapter 3. A geometrical presentation of both methods enables us to understand their objectives, usage, performance, and limitations in a practical way, especially to the attention of readers from finance and business backgrounds. The whole derivation of KF is provided in Chapter 3 appendix, still along a geometrical fashion. These two methods are basic tools in SSP.

Chapter 4 contents are the most advanced from SSP standpoint and share results at the edge (at the time the book is written) of relevant research about KF regularization to handle some impulsive and non-Gaussian markets behaviors. Spectacular improvement with respect to traditional KF is obtained and illustrated on market data.

Notations and Acronyms

a, **b**, etc.	Column vectors
A, **B**, etc.	Matrices
$\mathbf{A}'$	Transpose of matrix **A**
$\mathbf{A}^H$	Transpose conjugate of matrix **A**
$\mathbf{A}^{-1}$	Inverse of matrix **A**
$\mathrm{tr}(\mathbf{A})$	Trace of matrix **A**
$\nabla_{\mathbf{x}}(.)$	Gradient vector-operator
$\mathbf{0}_N$	$N \times N$ null matrix
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{1}_N$	N-dimensional vector of 1s
a, b, etc.	Scalars
t	Time
T	Total number of time observations
$t = 1, \cdots, T$	One-step vector of time going from 1 to T by step of 1
N	Number of assets
$s_{i,t}$	Price of asset i at time t
r_t	Return of an asset at time t
r_f	Return of the risk-free asset
$r_{i,t}$	Return of asset i at time t
$r_{m,t}$	Return of the market m at time t
$\mathbf{r}_t$	N-dimensional vector of returns at time t
$\mathbf{r}_k$	T-dimensional vector of returns for asset k
$\mathbf{R} = \{\mathbf{r}_k\}_{k=1}^N$	Matrix of returns for the N observed assets

$\mathbf{F} = \{\mathbf{f}_k\}_{k=1}^{K}$	Matrix of the K factor returns included in factor models
$\mathbf{f}_k$	kth column of $\mathbf{F}$
$\mathbf{f}_t$	tth row of $\mathbf{F}$ and K-dimensional vector of the factor values at t
$\mathbf{G} = \left[\mathbf{1}\ \mathbf{F}\right]$	Matrix $\mathbf{F}$ augmented by a vector of 1s having the same number of rows
$\mathbf{g}_k$	kth column of $\mathbf{G}$
$\mathbf{g}_t$	tth row of $\mathbf{G}$ and $K+1$-dimensional vector of values at t
α	"Alpha" or the specific performance of an asset over the market performance
β	"Beta" or the exposure of an asset to a specific factor
$\boldsymbol{\beta}$ or b	"Betas" or the vector of asset exposures to the K selected factors
$\boldsymbol{\theta} = \left[\alpha\ \boldsymbol{\beta}^T\right]^T$	$K+1$-dimensional vector of the unknown parameters of a multi-factor model
$\lVert . \rVert_q$	l^q-norm
$\mathbb{1}(.)$	Indicator function
$\hat{\mathbf{a}}$	Estimated value of a
i.i.d.	Independent and identically distributed
$\sim$	Distributed according to
$\perp\!\!\!\perp$	Statistically independent
$\mathbb{E}(.)$	Statistical expectation
$V(.)$ or $Var(.)$	Variance
$Cov(.)$	Covariance
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Multivariate Gaussian distribution with a mean vector $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$
AIC	Akaike information criterion
APT	Arbitrage pricing theory
BIC	Bayes information criterion
CAPM	Capital asset pricing model

FP	Fixedpoint
KF	Kalman filter
LSE	Least squares estimate
MDL	Minimum description length
ML	Maximum likelihood
MMSE	Minimum mean square error
MSE	Mean square error
MUSIC	Multiple signal characterization
NAV	Net asset value
OLS	Ordinary least squares
PCA	Principal component analysis
QAMLAB©	Quantitative Asset Management Laboratory
rgKF	Regularized Kalman filter
RKF	Robust Kalman filter
RMSE	Root mean square error
RMT	Random matrix theory
SCM	Sample covariance matrix
SSP	Statistical signal processing
SW-OLS	Sliding window ordinary least squares

Table of Contents

此为试读，需要完整PDF请访问：www.ertongbook.com