QUANTITATIVE TRADING

Algorithms, Analytics, Data, Models, Optimization

Xin Guo
University of California, Berkeley, USA

Tze Leung Lai
Stanford University, California, USA

Howard Shek
Tower Research Capital, New York City, New York, USA

Samuel Po-Shing Wong
5Lattice Securities Limited, Hong Kong, China

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

Contents

Preface xiii

List of Figures xvii

List of Tables xxi

1 Introduction 1

1.1 Evolution of trading infrastructure 1
1.2 Quantitative strategies and time-scales 5
1.3 Statistical arbitrage and debates about EMH 6
1.4 Quantitative funds, mutual funds, hedge funds 8
1.5 Data, analytics, models, optimization, algorithms 10
1.6 Interdisciplinary nature of the subject and how the book can be used 11
1.7 Supplements and problems 13

2 Statistical Models and Methods for Quantitative Trading 17

2.1 Stylized facts on stock price data 18
2.1.1 Time series of low-frequency returns 18
2.1.2 Discrete price changes in high-frequency data 18
2.2 Brownian motion models for speculative prices 22
2.3 MPT as a "walking shoe" down Wall Street 22
2.4 Statistical underpinnings of MPT 24
2.4.1 Multifactor pricing models 24
2.4.2 Bayes, shrinkage, and Black-Litterman estimators . . 25
2.4.3 Bootstrapping and the resampled frontier 26
2.5 A new approach incorporating parameter uncertainty 27
2.5.1 Solution of the optimization problem 27
2.5.2 Computation of the optimal weight vector 28
2.5.3 Bootstrap estimate of performance and NPEB 29
2.6 From random walks to martingales that match stylized facts 30
2.6.1 From Gaussian to Paretian random walks 31
2.6.2 Random walks with optional sampling times 32
2.6.3 From random walks to ARIMA, GARCH 35
2.7 Neo-MPT involving martingale regression models 37

2.7.1 Incorporating time series effects in NPEB 38
2.7.2 Optimizing information ratios along efficient frontier . 38
2.7.3 An empirical study of neo-MPT 39
2.8 Statistical arbitrage and strategies beyond EMH 41
2.8.1 Technical rules and the statistical background 41
2.8.2 Time series, momentum, and pairs trading strategies . 43
2.8.3 Contrarian strategies, behavioral finance, and investors' cognitive biases 44
2.8.4 From value investing to global macro strategies 44
2.8.5 In-sample and out-of-sample evaluation 45
2.9 Supplements and problems 46

3 Active Portfolio Management and Investment Strategies 61

3.1 Active alpha and beta in portfolio management 62
3.1.1 Sources of alpha 63
3.1.2 Exotic beta beyond active alpha 63
3.1.3 A new approach to active portfolio optimization . . . 64
3.2 Transaction costs, and long-short constraints 67
3.2.1 Cost of transactions and its components 67
3.2.2 Long-short and other portfolio constraints 68
3.3 Multiperiod portfolio management 69
3.3.1 The Samuelson-Merton theory 69
3.3.2 Incorporating transaction costs into Merton's problem 72
3.3.3 Multiperiod capital growth and volatility pumping . . 73
3.3.4 Multiperiod mean-variance portfolio rebalancing . . . 74
3.3.5 Dynamic mean-variance portfolio optimization 75
3.3.6 Dynamic portfolio selection 76
3.4 Supplementary notes and comments 78
3.5 Exercises 101

4 Econometrics of Transactions in Electronic Platforms 103

4.1 Transactions and transactions data 104
4.2 Models for high-frequency data 104
4.2.1 Roll's model of bid-ask bounce 105
4.2.2 Market microstructure model with additive noise . . . 106
4.3 Estimation of integrated variance of X_t 107
4.3.1 Sparse sampling methods 108
4.3.2 Averaging method over subsamples 109
4.3.3 Method of two time-scales 109
4.3.4 Method of kernel smoothing: Realized kernels 110
4.3.5 Method of pre-averaging 111
4.3.6 From MLE of volatility parameter to QMLE of $[X]_T$. 112
4.4 Estimation of covariation of multiple assets 113

4.4.1 Asynchronicity and the Epps effect 113
4.4.2 Synchronization procedures 114
4.4.3 QMLE for covariance and correlation estimation . . . 115
4.4.4 Multivariate realized kernels and two-scale estimators 116
4.5 Fourier methods 118
4.5.1 Fourier estimator of $[X]_T$ and spot volatility 118
4.5.2 Statistical properties of Fourier estimators 120
4.5.3 Fourier estimators of spot co-volatilities 121
4.6 Other econometric models involving TAQ 122
4.6.1 ACD models of inter-transaction durations 123
4.6.2 Self-exciting point process models 124
4.6.3 Decomposition of D_i and generalized linear models . . 125
4.6.4 McCulloch and Tsay's decomposition 126
4.6.5 Joint modeling of point process and its marks 127
4.6.6 Realized GARCH and other predictive models 128
4.6.7 Jumps in efficient price process and power variation . 130
4.7 Supplementary notes and comments 132
4.8 Exercises 139

5 Limit Order Book: Data Analytics and Dynamic Models 143

5.1 From market data to limit order book (LOB) 144
5.2 Stylized facts of LOB data 145
5.2.1 Book price adjustment 145
5.2.2 Volume imbalance and other indicators 148
5.3 Fitting a multivariate point process to LOB data 151
5.3.1 Marketable orders as a multivariate point process . . . 151
5.3.2 Empirical illustration 153
5.4 LOB data analytics via machine learning 157
5.5 Queueing models of LOB dynamics 159
5.5.1 Diffusion limits of the level-1 reduced-form model . . . 160
5.5.2 Fluid limit of order positions 163
5.5.3 LOB-based queue-reactive model 166
5.6 Supplements and problems 169

6 Optimal Execution and Placement 183

6.1 Optimal execution with a single asset 184
6.1.1 Dynamic programming solution of problem (6.2) . . . 185
6.1.2 Continuous-time models and calculus of variations . . 187
6.1.3 Myth: Optimality of deterministic strategies 189
6.2 Multiplicative price impact model 190
6.2.1 The model and stochastic control problem 190
6.2.2 HJB equation for the finite-horizon case 191
6.2.3 Infinite-horizon case $T = \infty$ 193

6.2.4 Price manipulation and transient price impact 196
6.3 Optimal execution using the LOB shape 196
6.3.1 Cost minimization 199
6.3.2 Optimal strategy for Model 1 202
6.3.3 Optimal strategy for Model 2 203
6.3.4 Closed-form solution for block-shaped LOBs 204
6.4 Optimal execution for portfolios 204
6.5 Optimal placement 207
6.5.1 Markov random walk model with mean reversion . . . 208
6.5.2 Continuous-time Markov chain model 211
6.6 Supplements and problems 215

7 Market Making and Smart Order Routing 221

7.1 Ho and Stoll's model and the Avellanedo-Stoikov policy . . . 222
7.2 Solution to the HJB equation and subsequent extensions . . 223
7.3 Impulse control involving limit and market orders 225
7.3.1 Impulse control for the market maker 225
7.3.2 Control formulation 226
7.4 Smart order routing and dark pools 228
7.5 Optimal order splitting among exchanges in SOR 230
7.5.1 The cost function and optimization problem 231
7.5.2 Optimal order placement across K exchanges 232
7.5.3 A stochastic approximation method 233
7.6 Censored exploration-exploitation for dark pools 234
7.6.1 The SOR problem and a greedy algorithm 234
7.6.2 Modified Kaplan-Meier estimate $\hat{T}_i$ 235
7.6.3 Exploration, exploitation, and optimal allocation . . . 236
7.7 Stochastic Lagrangian optimization in dark pools 237
7.7.1 Lagrangian approach via stochastic approximation . . 238
7.7.2 Convergence of Lagrangian recursion to optimizer . . 240
7.8 Supplementary notes and comments 241
7.9 Exercises 248

8 Informatics, Regulation and Risk Management 251

8.1 Some quantitative strategies 253
8.2 Exchange infrastructure 255
8.2.1 Order gateway 258
8.2.2 Matching engine 258
8.2.3 Market data dissemination 259
8.2.4 Order fee structure 260
8.2.5 Colocation service 262
8.2.6 Clearing and settlement 263
8.3 Strategy informatics and infrastructure 264

8.3.1 Market data handling 264
8.3.2 Alpha engine 265
8.3.3 Order management 266
8.3.4 Order type and order qualifier 266
8.4 Exchange rules and regulations 269
8.4.1 SIP and Reg NMS 269
8.4.2 Regulation SHO 272
8.4.3 Other exchange-specific rules 273
8.4.4 Circuit breaker 274
8.4.5 Market manipulation 274
8.5 Risk management 274
8.5.1 Operational risk 275
8.5.2 Strategy risk 277
8.6 Supplementary notes and comments 279
8.7 Exercises 289

A Martingale Theory **295**

A.1 Discrete-time martingales 295
A.2 Continuous-time martingales 298

B Markov Chain and Related Topics **303**

B.1 Generator $\boldsymbol{Q}$ of CTMC 303
B.2 Potential theory for Markov chains 304
B.3 Markov decision theory 304

C Doubly Stochastic Self-Exciting Point Processes **307**

C.1 Martingale theory and compensators of multivariate counting processes 307
C.2 Doubly stochastic point process models 308
C.3 Likelihood inference in point process models 309
C.4 Simulation of doubly stochastic SEPP 312

D Weak Convergence and Limit Theorems **315**

D.1 Donsker's theorem and its extensions 316
D.2 Queuing system and limit theorems 317

Bibliography **319**

Index **349**