Chapter 1
Optimization and Control for Systems in the Big Data Era: An Introduction

Tsan-Ming Choi, Jianjun Gao, James H. Lambert, Chi-Kong Ng, and Jun Wang

Abstract The big data era is characterized by the presence of many Vs in terms of data and data usage. In this introductory chapter, we first discuss some challenges in optimization and control for systems in the presence of massive amount of data. We then introduce the papers featured in this book.

Keywords Introduction • Big data • Optimization • Control

T.-M. Choi (✉)
Institute of Textiles and Clothing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
e-mail: jason.choi@polyu.edu.hk

J. Gao
School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai, People's Republic of China
e-mail: gao.jianjun@shufe.edu.cn

J.H. Lambert
Department of Systems and Information Engineering, University of Virginia, Charlottesville, VA, USA
e-mail: lambert@virginia.edu

C.-K. Ng
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
e-mail: ckng@se.cuhk.edu.hk

J. Wang
Department of Management Science and Engineering, Business School, Qingdao University, Shandong, People's Republic of China
e-mail: jwang@qdu.edu.cn

T.-M. Choi et al. (eds.), *Optimization and Control for Systems in the Big-Data Era*, International Series in Operations Research & Management Science 252,
DOI 10.1007/978-3-319-53518-0_1

1.1 Optimization and Control in Big Data Era

We are now in the big data era. The popularity of social media, mobile devices, cloud storage and application services, etc., all means that the amount of data which can be collected and used by organizations and companies is increasing everyday (Chan et al. 2016; Wang et al. 2016).

Traditionally, the term "big data" is associated with many Vs (Choi et al. 2016), such as volume (the amount of data), velocity (speed of data collection and processing), variety (the structured and unstructured data; complex data), veracity (the data accuracy and uncertainty) and value (the value associated with data). Undoubtedly, the presence of a massive amount of data means that we have to rethink about the strengths and weaknesses of the existing optimization and control methods.

In the recent literature, big data related optimization and control problems have been examined. For instance, Facchinei and Scutari (2015) propose a decomposition framework for achieving parallel optimization for a class of nonconvex problems with a massive amount of data. The authors demonstrate that their proposed method outperforms other existing methods. Daneshmand et al. (2015) develop a novel hybrid and parallel decomposition scheme for solving convex and nonconvex big data optimization problems. For very big problems, the authors show that their proposed decomposition scheme works well compared to other random or deterministic schemes. Bhattacharya et al. (2016) explore how an evolutionary optimization based algorithm can handle optimization problems with a high volume dataset. The authors claim that they have successfully applied the proposed algorithm in real world financial portfolio management. Richtarik and Takac (2016) explore how parallel randomized block coordinate descent methods can be used for developing a big data optimization algorithm. The authors show that their proposed algorithm can solve a class of large scale problems efficiently. Most recently, Boone et al. (2016) develop a framework for exploring service parts performance optimization problems with big data. They propose how, where and why big data applications can be applied in their proposed framework. For more recent developments of big data optimization, refer to the books by Emrouznejad (2016) and Japkowicz and Stefanowski (2016).

In this introductory chapter, we briefly review the papers featured in this book. According to the sectioning of the book, we present the papers in four sections: Sect. 1.2 reports the papers in the "Reviews on Theories" section, Sect. 1.3 examines the papers in the "Reviews on Applications" section, Sect. 1.4 introduces the "Financial Optimization Analysis" papers and Sect. 1.5 describes the papers in "Operations Analysis".

1.2 Reviews on Theories

For the review on optimization and control theories, this book features three papers. First, Fu examines the dual control theory in the big data era. The author first presents an overview of dual control and its probable applications. She explores different dual and non-dual controllers and highlights their complexity and limitations. In particular, she focuses on a class of discrete-time LQG problems with unknown parameters. She shows the optimal dual control, active open-loop feedback control via variance minimization (Li et al. 2002) and optimal nominal dual control for this class of problems. Finally, the author also discusses the probable usage of dual control in economic systems, information retrieval as well as mechanical engineering.

Shi and Cui review the time inconsistency and self control optimization problems, which are commonly found in financial optimization and conflict decision making (Cui et al. 2012). The authors examine different approaches which can effectively deal with the time inconsistency challenge in decision making. Moreover, they report the recent progress in the area and mention the challenges of time inconsistency optimization in the big data era.

Wu and Jiang examine the recent developments in the quadratic convex reformulation method. In fact, the quadratic convex reformulation method is commonly used to derive efficient equivalent reformulations for mixed-integer quadratically constrained quadratic optimization problems. The authors comment that even though the proposed problems can be solved by a standard mixed-integer quadratic solver using the branch-and-bound method, the solver's performance is far from satisfactory. They thus argue that the quadratic convex reformulation approach provides a systematic way to solve the above-mentioned challenging optimization problems. The authors also review some recent extensions of the quadratic convex reformulation method for problems such as the challenging semi-continuous quadratic optimization problems.

1.3 Reviews on Applications

For the review on applications of optimization and control methods in the big data era, there are four related papers in this book.

Pei and Zhu explore measurements of financial contagion. The authors propose that the financial contagion is a timely issue which is closely related to financial systemic risk. In their paper, they first clarify and summarize various critical concepts and measurements of financial contagion and then highlight their common features and differences. Since the structural break is known to be especially crucial, the authors review and discuss the respective financial contagion measurements. The authors conclude that the big data technology may be helpful for advancing risk management relevant to financial contagion in, e.g., information acquisition as well as model specification.

Dynamic portfolio optimization is a critical application area of advanced optimization models and methods. Chiu reports the summary of recent advances in the optimal asset-liability management system, which employs the dynamic portfolio optimization technique, in a continuous-time domain. From the stochastic optimal control perspective, the author derives a new asset-liability management solution for the case with insurers possessing a constant absolute risk averse utility function and Poisson-type insurance liabilities.

Cryptography is a very interesting area for secret communication. Traditionally, cryptography focuses on confidentiality, integrity and authentication of information. Lu reports in her paper a review of modern cryptography. The author chooses two important systems in cryptography, namely the Merkle–Hellman knapsack cryptography system and the subset-sum problem based cryptography system in the review. She examines the encryption and decryption processes of the above two systems.

In supply risk analysis, Li et al. conduct a systematic review on supply risk modeling. From the operations management perspective, the authors present various recent developments in supply risk modeling which include vertical supply chain interaction, horizontal supply chain competition and network competition problems. The authors present analytical models for each scenario. They also discuss future research directions in the big data era.

1.4 Financial Optimization Analysis

In addition to reviews on optimization and control theories and applications, this book also features many technical papers with novel insights and new findings. In the area of financial optimization, there are four related papers.

First, Li et al. develop a parameterized method for achieving optimal multi-period mean-variance portfolio selection with liability considerations. The authors note that the financial market generates a massive amount of data which are related to portfolio management. To effectively select the optimal portfolio requires very careful planning and the use of an efficient method. The authors hence propose a new method to help derive the analytical optimal portfolio strategies and efficient frontiers accurately. The authors demonstrate the applicability of their work by presenting a numerical example.

Gao and Wu explore the data driven mean-CVaR (DDMC) portfolio optimization problem. The authors consider the case when the out of sample performance of the DDMC portfolio optimization model is unstable, which occurs in practice owing to the availability of historical data. To deal with this challenge, the authors propose a novel method by adding a penalty on the sparsity of the portfolio weight and combine the variance term in the DDMC model. The authors run a numerical analysis and confirm that the out-of-sample performance fragility is being well mitigated using the new method.

Options are critical in the financial market. Liang conducts an analysis on the multi-stage optioned portfolio selection problem. The author introduces the mean-variance models first and then establishes the target tracking model for the optioned portfolio selection problem in both dynamic and static settings. She proposes two different solution schemes, namely the stochastic programming approach with optimality condition scheme, and the stochastic control with dynamic programming scheme. She also reveals the closed form relationship between the mean-variance model and the target tracking model.

Yi studies the multi-period portfolio selection problem with a stochastic investment horizon. She argues that the problem is commonly seen in the real world as an investor may suddenly terminate an investment (which leads to the stochastic investment horizon) owing to various factors. The author finds that the formulated problem is non-separable in the sense of dynamic programming. She thus employs an embedding technique to derive the optimal policy.

1.5 Operations Analysis

Finally, in this book, two papers on operations analysis are featured. First, Wang et al. present a new model and a novel method to solve the order selection problems in flow-shop operations. The authors notice that traditional order selection models separate the production scheduling process and the order selection problem, and the performance of order selection solely depends on production scheduling. The authors hence propose a new model which simultaneously considers both production scheduling and order selection. By computational experiments, the authors demonstrate that the proposed new model is much better than the traditional ones.

Quick response is a well-established strategy in supply chain management (Choi et al. 2003, 2004, 2006). Under quick response, companies can better utilize market information to improve their forecasts. In the big data era, it is easier for companies to collect and use a large amount of data. Motivated by the presence of massive amount of data and the importance of quick response in supply chain operations, Choi explores quick response in fashion supply chains. He focuses his attention on how the number of observations affects the expected values of quick response for the centralized supply chain system as well as the individual supply chain agents in the decentralized setting. He proves that quick response is a beneficial practice for the centralized supply chain system as well as the fashion retailer, and the benefit is increasing in the number of observations. If the number of observations goes to infinity, resembling the presence of "big data", the expected values of quick response will go to the steady states and the analytical expressions are found. However, there exist cases in which the fashion supplier suffers from a loss after adopting quick response. As such, the author derives a wholesale pricing markdown contract to achieve win–win coordination, which means both the supplier and the fashion retailer are better off (i.e. "win–win") and the supply chain is simultaneously globally optimized (i.e. "coordinated").

References

M. Bhattacharya, R. Islam, J. Abawajy, Evolutionary optimization: a big data perspective. J. Network Comput. Appl. **59**, 416–426 (2016)

C.A. Boone, B.T. Hazen, J.B. Skipper, R.E. Overstreet. A framework for investigating optimization of service parts performance with big data. Ann. Oper. Res. (2016), in press

H.K. Chan, T.M. Choi, X. Yue, Big data analytics: risk and operations management for industrial applications. IEEE Trans. Ind. Inf. **12**, 1214–1218 (2016)

T.M. Choi, D. Li, H. Yan, Optimal two-stage ordering policy with Bayesian information updating. J. Oper. Res. Soc. **54**(8), 846–859 (2003)

T.M. Choi, D. Li, H. Yan, Optimal single ordering policy with multiple delivery modes and Bayesian information updates. Comput. Oper. Res. **31**, 1965–1984 (2004)

T.M. Choi, D. Li, H. Yan, Quick response policy with Bayesian information updates. Eur. J. Oper. Res. **170**, 788–808 (2006)

T.M. Choi, H.K. Chan, X. Yue, Recent development in big data analytics for business operations and risk management. IEEE Trans. Cybern. **47**, 81–92 (2016). doi:10.1109/TCYB.2015.2507599

X.Y. Cui, D. Li, S.Y. Wang, S.S. Zhu, Better than dynamic mean-variance: Time inconsistency and free cash flow stream. Math. Financ. **22**(2), 346–378 (2012)

A. Daneshmand, F. Facchinei, V. Kungurtsev, G. Scutari, Hybrid random/deterministic parallel algorithms for convex and nonconvex big data optimization. IEEE Trans. Signal Process. **63**(15), 3914–3929 (2015)

A. Emrouznejad (ed.), *Big Data Optimization: Recent Developments and Challenges*, Series on Studies in Big Data (Springer, Switzerland, 2016)

F. Facchinei, G. Scutari, Parallel selective algorithms for nonconvex big data optimization. IEEE Trans. Signal Process. **63**(7), 1874–1889 (2015)

N. Japkowicz, J. Stefanowski, *Big Data Analysis: New Algorithms for a New Society*, Series on Studies in Big Data (Springer, Switzerland, 2016)

D. Li, F.C. Qian, P.L. Fu, Variance minimization approach for a class of dual control problems. IEEE Trans. Autom. Control **47**(12), 2010–2020 (2002)

P. Richtarik, M. Takac, Parallel coordinate descent methods for big data optimization. Math. Program. Ser. A **156**, 433–484 (2016)

X. Wang, Y. He, Learning from uncertainty for big data. IEEE Syst. Man Cybernet. Mag. **2**(2), 26–32 (2016)