

PREFACE

The objective of this book is to provide a tutorial on the computational tools that use mathematical optimization concepts and representations for the curation, analysis and redesign of metabolic networks. Emphasis is placed on explaining the optimization formulation fundamentals and relevant algorithmic solution techniques. This book does not aim to serve as a comprehensive optimization text as it provides only a condensed treatment of different classes of optimization problems. The main goal here is to use the language and tools of mathematical programming to describe and solve frequently occurring problems in the analysis and redesign of metabolic networks. The interested reader is thus encouraged to refer to dedicated optimization textbooks for a more thorough treatment of different classes of mathematical programming problems. Similarly, the description of metabolism in the book is limited, requiring frequent referral to relevant biochemistry and molecular biology resources. The reader is encouraged to consult the original journal publications for all described techniques. Nevertheless, the treatment presented here has benefited from many years of application and customization on a variety of projects. As a consequence, the day-to-day application, implementation and integration of these techniques may differ from their original exposition in the journal publications. Care is exercised to ensure that the metabolic network descriptions, nomenclature and assumptions are consistent throughout. Many important optimization-based techniques in metabolic networks are absent from this treatment due to space limitations, lack of working experience with them or difficulty in integrating them with the rest of the material. Course notes for a graduate elective on "Optimization in Biological Networks" taught at Pennsylvania State University – University Park formed the basis of the techniques described here.

This book can be used by itself or in combination with other books to support coursework on special topics in network analyses of biological and in particular metabolic networks. It can be used to introduce students with knowledge of metabolism to formal mathematical treatments of core computational tasks in metabolic networks or alternatively to expose students with a mathematical programming background to metabolism. The hope is that the book will serve as a starting point for the students for more in-depth investigations of relevant techniques and concepts found in the cited literature.

Chapter 1 uses a regulatory network inference problem to introduce modeling using mathematical programming formulations. Key concepts such as sets, parameters, variables, constraints and optimization formulations are introduced. This is followed by a formal introduction and definition of fundamental concepts in optimization that establish criteria for the existence of a local or global optimal solution. Chapter 2 introduces linear programming (LP), which underpins many analysis techniques in metabolic networks, and linear duality theory, which is used for solving bilevel optimization problems described in Chapter 8. Chapter 3 provides an introduction to metabolic network modeling and flux balance analysis (FBA) in particular. A number of core FBA analysis techniques that rely on LP are described here. Chapter 4 unveils the concept of discrete or binary variables alongside continuous variables for modeling a variety of tasks in metabolic networks. Both representations and solution techniques of mixed-integer linear programming (MILP) problems are highlighted. Chapter 5 describes how reaction free energy of change considerations can be used to restrict reaction directionality and eliminate thermodynamically infeasible cycles in metabolic networks. Chapter 6 addresses the task of resolving metabolic network gaps and growth prediction inconsistencies in metabolic networks using computational techniques relying on MILP problem formulations. Chapter 7 reviews optimization-based approaches for finding paths that link a source to a target metabolite. Chapter 8 addresses the use of bilevel programming for describing and solving metabolic network redesign problems to maximize the yield of a target product in microbial strain design. Chapter 9 introduces basic concepts, optimality criteria, and solution techniques for both unconstrained and constrained nonlinear programming problems (NLPs). Chapter 10 provides examples of how NLP underpins metabolic network analysis problems such as integration of kinetic expressions into genome-scale metabolic models and metabolic flux elucidation using carbon-labeled isotopes in metabolic flux analysis (MFA). Finally, Chapter 11 lays the foundation for the description and solution of mixed-integer NLP (MINLP) problems and highlights how the nonlinearities of kinetic expressions can be integrated in the strain design bilevel optimization formulations introduced in Chapter 8. A number of example problems are addressed in the text and suggested problems are provided at the end of every chapter. The source GAMS code for all examples in the book are available through a dedicated website for this book (www.maranasgroup.com). Alternative representations using MATLAB, LINGO, etc. can also be derived in a straightforward manner. In addition, Appendix A provides a short guide to GAMS.

The book is intended as a textbook for a graduate or advanced undergraduate elective in the use of optimization in metabolic networks. Topics covered in the book start with a formal treatment of the relevant optimization problem class followed by

application in the context of metabolic network analysis. The class of optimization problems becomes progressively more complex starting with LP and MILP and concluding with NLP and MINLP problems. Many of the proofs of convergence and optimality theorems are geared toward the students with a more in-depth interest in optimization and can be skipped without loss of teaching material congruity. Hands-on implementation of all the discussed methods is encouraged to reinforce the introduced concepts.

We would like to acknowledge the people in the C. Maranas group for providing feedback and tirelessly helping with the preparation of the book. In particular, we would like to acknowledge the contribution of Ali Khodayari for assembling the initial set of lecture notes from the related graduate course at Penn State and for helping to create and improve figures; Anupam Chowdhury for performing simulations for examples of Chapters 3, 4, 7, 8, 10, and 11 and for preparing the related figures and computer codes; Chiam Yu Ng and Margaret Simons for helping to create figures and drafting the initial versions of the chapters; and Sarat Ram Gopalakrishnan for helpful feedback on section 10.3 of Chapter 10. Finally, we would like to express our gratitude to Margaret Simons for designing the cover image and for carefully editing the proofs and making numerous suggestions.

Costas D. Maranas
Ali R. Zomorodi