

Lineare Optimierung in der Bioinformatik

| | | |
|----------------------------|--|-------------|
| Title | Lineare Optimierung in der Bioinformatik | |
| Typ | Vorlesung mit Übungen | |
| Credits | 5 | |
| Lehrform/SWS | 2V + 2Ü | |
| Sprache | Deutsch oder Englisch (wird vom Dozenten zum Vorlesungsbeginn bekannt gegeben) | |
| Modulniveau | Master | |
| Arbeitsaufwand | Präsenzstunden | 60 Stunden |
| | Eigenstudium | 90 Stunden |
| | Gesamtaufwand | 150 Stunden |
| Angestrebte Lernergebnisse | Die Teilnehmer beherrschen die Grundlagen der Optimierung von linearen Funktionen unter linearen Nebenbedingungen und sind in der Lage kombinatorische Optimierungsprobleme wie sie in der Bioinformatik auftreten als (ganzzahlige) lineare Programme zu beschreiben und zu lösen. | |
| Intended Learning Outcomes | Students are familiar with fundamental concepts and algorithmic methods for optimizing a linear function subject to linear constraints. They will learn how to formulate combinatorial optimization problems that occur in Bioinformatics as (integer) linear programs. | |
| Inhalt | <p>Themen des Moduls Lineare Optimierung in der Bioinformatik sind:</p> <ul style="list-style-type: none"> • Geometrische Interpretation von Linearen Programmen: Konvexe Mengen, Polyhedra, Extrempunkte • Die Simplex-Methode: Optimalitätsbedingungen, Herleitung für den nicht-degenerierten Fall, Laufzeit • Dualitätstheorie: Das duale Problem, schwache und starke Dualität, komplementärer Schlupf • Netzwerkflüsse: Problemformulierung, Netzwerksimplex, Cycle Cancelling Algorithmus, max flow - min cut • Ganzzahlige Lineare Programmierung: Starke Formulierungen • Ganzzahlige Lineare Programmierung Methoden: Schnittebenenverfahren, Branch-and-Bound, Lagrange-Relaxierung • Ausgewählte Anwendungen in der Bioinformatik: Vergleich von Einzelzell-Trajektorien, Analyse der Proteindynamik, | |

| | |
|-------------|---|
| | Seitenkettenplatzierung, Netzwerkanalyse |
| Contents | <p>Topics of Linear Optimization in Bioinformatics are:</p> <ul style="list-style-type: none"> • Geometry of linear programming: Convex sets, polyhedra, extreme points • The simplex method: Optimality conditions, development in the non-degenerate case, running time • Duality theory: The dual problem, weak and strong duality, complementary slackness • Network flows: Problem formulation, network simplex algorithm, negative cost cycle algorithm, max flow – min cut • Integer programming: Strong formulations • Integer programming methods: Cutting plane methods, branch and bound, Lagrangean relaxation • Selected Applications in Bioinformatics: Comparison of single-cell trajectories, Analysis of protein dynamics, side chain placement, network analysis |
| Prüfung | <p>Prüfungsleistung (benotet): Klausur (120min)</p> <p>Wiederholungsklausur zu Ende des Semesters. Details werden zu Beginn des Moduls bekannt gegeben.</p> <p>In der Klausur weisen die Studierenden nach, dass sie kombinatorische Optimierungsprobleme in der Bioinformatik als (ganzzahlige) linear Programme modellieren, ihre Eigenschaften analysieren, und den Ablauf der im Modul behandelten Algorithmen erläutern bzw. diese an neue Problemstellungen anpassen können. Konkret werden in der Klausur 4-7 Aufgaben bearbeitet, die eine eigenständige Anwendung der Methoden zur Formulierung und Analyse von (ganzzahligen) linearen Programmen zur Lösung einer anspruchsvollen Problemstellung erfordern, wie der konkrete Entwurf von Variablen und linearen Ungleichungen, die Ermittlung des Gaps einer LP Relaxierung, die Beschreibung des Verhaltens eines im Modul behandelten Algorithmus (z.B. Simplex Methode, max flow), und das Entwerfen oder Anpassen eines Separierungsalgorithmus für eine Problembeschreibung mit exponentiell vielen Ungleichungen.</p> |
| Examination | <p>Examination requirements (graded): written exam (120 min)</p> <p>A makeup exam will be offered at the end of the</p> |

| | |
|-----------------------|--|
| | <p>semester. Details will be announced at the beginning of the module.</p> <p>Within the written exam, students demonstrate that they are able to model combinatorial optimization problems in Bioinformatics as (integer) linear programs, to analyze their properties, to explain the behavior of an algorithm presented in the module, and to adapt them to related problems. The written exam contains 4 to 7 assignments, which require the independent application of methods for the formulation and analysis of (integer) linear programs to solve demanding problems such as the design of variables and linear inequalities, the determination of the integrality gap of a LP relaxation, the description of the behavior of an algorithm presented in the module (e.g., simplex method, max flow) as well as the design or adaptation of a separation algorithm for a problem description involving an exponential number of constraints.</p> |
| Literatur/Literature | <p>Dan Gusfield: Integer Linear Programming in Computational and Systems Biology, Cambridge University Press, 2019.</p> <p>D. Bertsimas, J.N. Tsitsiklis: Introduction to Linear Optimization, Athena Scientific, 1997.</p> <p>L.A. Wolsey: Integer Programming, Wiley-Interscience, 1998.</p> <p>A. Schrijver: Theory of Linear and Integer Programming, Wiley-Interscience, 1998</p> |
| Literature | <p>Dan Gusfield: Integer Linear Programming in Computational and Systems Biology, Cambridge University Press, 2019.</p> <p>D. Bertsimas, J.N. Tsitsiklis: Introduction to Linear Optimization, Athena Scientific, 1997.</p> <p>L.A. Wolsey: Integer Programming, Wiley-Interscience, 1998.</p> <p>A. Schrijver: Theory of Linear and Integer Programming, Wiley-Interscience, 1998</p> |
| Medienformen | Tafelanschrieb, Folienpräsentation |
| Media | blackboard, slides |
| Lehr- und Lernmethode | <p>Vorlesung, Tutorübung, Aufgaben zum Selbststudium.</p> <p>Das Modul besteht aus einer Vorlesung und einer Übung. In den Hausaufgaben, die freiwillig abzugeben sind, wird das Verständnis der Konzepte, Techniken und Algorithmen, die in der Vorlesung vorgestellt werden, anhand konkreter Daten und Beispiele vertieft. Die Studierenden modellieren kombinatorische Optimierungsprobleme aus der Bioinformatik als</p> |

| | |
|-------------------------------|--|
| | <p>(ganzzahlige) lineare Optimierungsprobleme, analysieren deren Eigenschaften mithilfe der in der Vorlesung vorgestellten Methoden und Techniken, und entwickeln bzw. vergleichen Algorithmen zu deren effizienter Lösung. In den Hausaufgaben werden selbständig anspruchsvolle Übungsaufgaben bearbeitet, die ähnlich wie die Klausuraufgaben sind (siehe oben) und daher zur Vorbereitung darauf dienen. In den Übungen werden mögliche Lösungsansätze der Aufgaben zum Selbststudium diskutiert.</p> |
| Teaching and Learning Methods | <p>Lecture, tutorial, assignments for individual study.</p> <p>The module consists of a lecture and exercises. Within the assignments (submission is optional) concepts, techniques and algorithms (presented in the lecture) will be applied to real data and examples to deepen the understanding. The students model combinatorial optimization problems in Bioinformatics as (integer) linear optimization problems, analyze their characteristics using methods and techniques taught in the lectures, and develop or compare algorithms for their efficient solution. The assignments consist of demanding problems similar to the assignments in the written exam (for details see above) and serve as a preparation for the written exam. Within the tutorials possible approaches for solutions of the assignments will be discussed.</p> |
| Turnus | Wintersemester |
| Modulverantwortlicher | Dr. Stefan Canzar |
| Dozenten | Dr. Stefan Canzar |