Michael Zibulevsky

Search this site

HOME

FILES

MATLAB CODE

OPTIMIZATION COURSE

SLIDES

SITEMAP

Optimization Course

Department of Computer Science, Technion

Lecturer: Dr. Michael Zibulevsky

mzibul@gmail.com, https://sites.google.com/site/michaelzibulevsky/

***FLIPPED CLASSROOM:* Video-lecture based course** CS-236330

Introduction to Optimization

and Deep Learning

3 academic points

Spring semester 2016, Tuesday, 16:30 - 18:30, Room Taub 337

YouTube video lectures here

Snapshots (slides)

Discussion Group: Optimization Technion 2016

We use a FLIPPED CLASSROOM approach:

Students watch video-lecture at home and discuss the lecture and solve problems *in class*.

Course requirements:

- Participation in class discussions;
- 6 homework exercises, partially in MATLAB (30%)
- Final exam (70%)

Announcements

-- This year (Spring 2016) we are going to include several lectures on Deep Learning instead of Conic Programming, see e.g.

- Intro to Neural Networks
- Convolutional Neural Networks in 10 minutes

-- Dear students, please *watch and take notes* of two video lectures: *Lecture 1a* and   *Lecture 1b* towards the first class meeting. The link is here: **YouTube video lectures**

-- All new students should join our Discussion Group Optimization Technion 2016

Course Outline

The field of optimization is concerned with the study of maximization and minimization of mathematical functions of one or many variables. Very often the variables are subject to side conditions or constraints (so-called constrained optimization problems). By great utility in such diverse areas as applied science, engineering, economics, finance, medicine, and statistics, optimization holds an important place in the practical world and the scientific world. Indeed, as far back as the Eighteenth Century, the famous mathematician Leonhard Euler proclaimed that . . . *nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear.*

Very often it is impossible to find analytical solution to unconstrained or constrained optimization problem, however there are efficient numerical algorithms which approach solution iteratively. *Optimization algorithms* constitute the central part of our course. We will also learn how to check, whether optimal solution has been achieved (optimality conditions for constrained and unconstrained problems) and duality theory (Lagrange multiplier representation).

In the final part of the course we will learn a modern topic: Conic optimization, which includes Semidefinite Programming - optimization over set of positive-semidefinite matrices. This field has great importance in modern science and engineering.

Lecture 1a, Introduction; Examples of unconstrained and constrained optimization problems:
Lecture 1b, Linear algebra refresh:
Lecture 2-3: Derivatives of multivariate functions: Gradient and Hessian

Lecture 4-5: Convex sets and functions
Lecture 6. Local and global minimum. Sufficient and necessary unconstrained optimality conditions
Lecture 7 (in Hebrew). Optimality conditions; iterative methods of one-dimensional optimization (line search).
Lecture 8 Iterative methods of multivariate unconstrained optimization
Lecture 9 More on Newton method
Lecture 10 Method of Conjugate Gradients 1
Lecture 11 Method of Conjugate Gradients 2
Lecture 12 Sequential subspace optimization (SESOP) method and Quasi-Newton BFGS
Lecture 13. Summary of unconstrained optimization. Optimization with constraints
Lecture 14 Lagrange multipliers and penalty function method. Augmented Lagrangian
Lecture 15 Minimax theorem, game theory and Lagrange duality

Lecture 16 Conic programming 1 (in particular, semidefinite programming)
Lecture 16b-17 Conic programming 2
Homework on analytical and numerical computation of gradient and Hessian
Penalty method for semidefinite programming and homework on linear matrix approximation

Staff

	Lecturer	
Name	Michael Zibulevsky	
Telephone	0547-847738	

Literature

Basic References

- Jorge Nocedal, Stephen J. Wright, "Numerical Optimization" (1999)
- D. Bertsekas, "Nonlinear Programming", Second Edition (1999)
- S. Boyd, L. Vandenberghe , "Convex Optimization" (Dec. 2003). Click here for 2 or 4 pages per printed page.
- Old Tutorial book by my former assistant Dori Peleg: Tutorial book (2005)

Additional References

- Gilbert Strang, Video-lectures on Linear Algebra, MIT, local link here , http://ocw.mit.edu/OcwWeb/Mathematics/18-06Spring-2005/CourseHome/index.htm , http://videolectures.net/gilbert_strang/
- Gilbert Strang, Video-lectures on Computational Methods in Science and Engineering I - Fall 2007
- Dori Peleg, Hebrew Matlab booklet

GUI

Matlab Code	Status	Tutorial	Status
Unconstrained GUI Code	Available	Unconstrained GUI tutorial	Available
Constrained GUI Code	Available	Constrained GUI tutorial	Available

Useful Links

Yalmip 3

Arkadi Nemirovski, Lectures on modern optimization

Freemat - a free Matlab-like coding environment

Numerical Recipes - Books Online

=====================================

Contents of the video lectures

Lecture 1a, Introduction; Examples of unconstrained and constrained optimization problems:

Linear regression
Function approximation with feed-forward neural

network

Resource assignment with Linear Programming

Lecture 1b, Linear algebra refresh:

Linear and affine subspace;

Lp norm of a vector
Matrix norms
Inner product of matrices
Eigenvalue decomposition
Matrix polynomials and functions
Positive (semi) definite matrices

Lecture 2-3: Derivatives of multivariate functions: Gradient and Hessian

Total differential and gradient
Level sets and directional derivatives
Hessian matrix
Second directional derivative
Example: Gradient and Hessian of linear and quadratic function
Taylor expansion of multivariate functions
Gradient of a function of a matrix
Example: Gradient of a neural network

Lecture 4-5: Convex sets and functions

Definition of convex set and function.

Properties of convex sets
Properties of convex functions
Extended value functions
Epigraph
Convex combination and convex hull
Jensen inequality
Gradient inequality. Hessian of a convex function is positive (semi) definite

Lecture 6. Local and global minimum. Sufficient and necessary unconstrained optimality conditions

Lecture 7 (in Hebrew). Optimality conditions; iterative methods of one-dimensional optimization (line search).

Optimality conditions

One-dimensional optimization, Bisection method
Golden section method
Quadratic interpolation method
Cubic interpolation method

Note: First part of Lecture 7 (optimality conditions) repeats in Hebrew the second part of Lecture 6.

Lecture 8 Iterative methods of multivariate unconstrained optimization

General line search method
Choice of step size: Exact optimization, Backtracking, Armijo stopping rule
Steepest descent (gradient descent)
Newton method

Lecture 9 More on Newton method

Newton method for nonlinear equations
Modified Newton method: enforcing descent direction
Solving symmetric system of equations via Cholesky factorization
Least-squares problem: Gauss-Newton and *Levenberg–Marquardt methods*

Lecture 10 Method of Conjugate Gradients 1

Motivation
Scalar product, definition and examples
Gram-Schmidt orthogonalization
Q-conjugate (Q-orthogonal) directions
Minimization of a quadratic function using conjugate directions

Lecture 11 Method of Conjugate Gradients 2

Derivation of the method of Conjugate Gradients
Summary of CG method
Convergence rate of CG
Preconditioning
Truncated Newton method
Computational burden to compute function, gradient and Hessian-vector product is about the same

Lecture 12 Sequential subspace optimization (SESOP) method and Quasi-Newton BFGS

SESOP method
Fast optimization over subspace
Quasi-Newton methods
How to approximate Hessian
Approximation of inverse Hessian, Sherman-Morrison formula
Broyden family Quasi-Newton methods, DFP, BFGS
Initialization and convergence properties

Lecture 13. Summary of unconstrained optimization. Optimization with constraints

Summary of unconstrained optimization
Constrained optimization, optimality conditions

Karush - Kuhn - Tucker (KKT) first order optimality conditions
Penalty function method

Lecture 14 Lagrange multipliers and penalty function method. Augmented Lagrangian

Lagrange multipliers via penalty function method
Penalty for equality constraints
Barrier method
Augmented Lagrangian method (or method of multipliers)
Augmented Lagrangian for equality constraints
Optimal solution in one step, when multipliers are known

Lecture 15 Minimax theorem, game theory and Lagrange duality

Introduction
Minimax theorem
Game interpretation of Minimax
Saddle point theorem
Minimax of Lagrangian;
Dual problem and weak duality
Strong duality
Slayter condition for strong duality
Examples of dual problems:
    Quadratic program
    Linear program

Lecture 16 Conic programming 1

Introduction
Examples of cones - R^n+, Lorenz cone, cone of positive semidefinite matrices
Semidefinite programming (SDP)
Dual cone , Dual conic problem, weak duality
Strong conic duality and complementarity slackness
Example: Dual SDP problem
Minimax problem
Chebyshev approximation

Lecture 16b-17 Conic programming 2

Continue with Chebyshev approximation, which starts at
Complex Chebyshev approximation
Conversion of different problems to SDP
Lemma of Schur complement
Minimize maximal eigenvalue of symmetric matrix
Linear matrix approximation
Expression of Linear program via SDP
Conic quadratic program via SDP
Barrier method for solution of conic programming problem
Examples of barriers
Matrix functions
Gradient of trace of matrix function
Gradient of the barrier log det A (x)

Homework on analytical and numerical computation of gradient and Hessian

Penalty method for semidefinite programming and homework on linear matrix approximation

==

Contents of the video lectures (time pointers included)

Lecture 1a, Introduction; Examples of unconstrained and constrained optimization problems:

Linear regression (slides 10:08, 11:56)
Function approximation with feed-forward neural network 12:15 (slides 16:15, 24:44, 27.21)
Resource assignment with Linear programming 27:40 (slides 33:02, 35:42)

Lecture 1b, Linear algebra refresh:

Linear and affine subspace; 0:0 (slides 05:39, 07:07

Lp norm of a vector 7:24 (slides 9:44 13:45
Matrix norms - 14:52 (slides 18:29)
Inner product of matrices - 19:30 (slides 22:40)
Eigenvalue decomposition -23:15 (slides 29:20, 31:47)
Matrix polynomials and functions 31:45 (35:12, 40:25)
Positive (semi) definite matrices - 41:17 (slides 45:41)

Lecture 2-3: Derivatives of multivariate functions: Gradient and Hessian

Total differential and gradient - 0:0 (slides 5:14, 13:16)
Level sets and directional derivatives - 13:17 (slides 23:07)
Hessian matrix - 23:20 (slides 28:50, 33:54)
Second directional derivative - 34:45 (slides 35:55)
Example: Gradient and Hessian of linear and quadratic

function -39:04 (slides 39:04)
Taylor expansion of multivariate functions 46:30 (slides 52:56)
Gradient of a function of a matrix - 53:23 (slides 58:54)
Example: Gradient of a neural network 58:56 (slides 1:09:43)

Homework - 1:09:45

Lecture 4-5: Convex sets and functions

Definition of convex set and function. Properties of convex sets - 0:0 (slides 03:05, 9:15, 9:53, 19:10)
Properties of convex functions - 19:10 (slides 20:52, 26:15)
Extended value functions - 26:41 (slides 31:22)
Epigraph - 31:31 (slides 34:34)
Convex combination and convex hull 34:38 (slides 37:18)
Jensen inequality 39:02 (slides 43:24)
Gradient inequality. Hessian of a convex function is positive (semi) definite 43:28 (slides 47:47)

Lecture 6. Local and global minimum. Sufficient and necessary unconstrained optimality conditions

(slides 2:10, 7:54, 12:24, 17:34, 21:55)

Lecture 7 (in Hebrew). Optimality conditions; iterative methods of one-dimensional optimization (line search).

Optimality conditions - 0:0

One-dimensional optimization, Bisection method 13:13 (slides 23:13)
Golden section method 23:17 (slides 32:02)
Quadratic interpolation method 32:04 (slides 36:39, 43:22, 47:00)
Cubic interpolation method 47:04 (slides 51:07, 57:35)

Note: First part of Lecture 7 (optimality conditions) repeats in Hebrew the second part of Lecture 6.

Lecture 8 Iterative methods of multivariate unconstrained optimization

General line search method 0:0 (slides 05:40)
Choice of step size: Exact optimization, Backtracking, Armijo stopping rule 5:53 (slides 17:05, 20:44)
Steepest descent (gradient descent) 20:55 (slides 36:07)
Newton method 36:22 (slides 45:18, 56:04)

End - 56:15 (after this time - garbage from previous lecture)

Lecture 9 More on Newton method

Newton method for nonlinear equations 0:0 (slides 5:28, 8:24)
Modified Newton method: enforcing descent direction 8:26 (slides 15:54)
Solving symmetric system of equations via Cholesky factorization 15:58 (slides 22:35, 28:39)
Least-squares problem: Gauss-Newton and *Levenberg–Marquardt methods* 28:44 (slides 33:39, 43:02, 52:27, 54:18)

Lecture 10 Method of Conjugate Gradients 1

Motivation 0:0
Scalar product, definition 4:47 (slide on 8:53), and examples 8:56 (slides 13:00)
Gram-Schmidt orthogonalization 13:07 (slides 21:13)
Q-conjugate (Q-orthogonal) directions 21:18 (slides 26:24)
Minimization of a quadratic function using conjugate directions 26:30 (slides 36:14 or 38:13)
Expanding manifold property 38:10 (slides 48:53 and 50:58)

Lecture 11 Method of Conjugate Gradients 2

Derivation of the method of Conjugate Gradients 0:0 (slides 5:34, 12:11, 19:32)
Summary of CG method 19:33 (slides 27:06)
Convergence rate of CG 27:08 (slides 32:57)
Preconditioning 32:58 (slides 44:03)
Truncated Newton method 44:04 (slide 50:20)
Computational burden to compute function, gradient and Hessian-vector product is about the same 50:20 (slides 52:36, 57:14)

Lecture 12 Sequential subspace optimization (SESOP) method and Quasi-Newton BFGS

SESOP method 0:42 (slides 7:35)
Fast optimization over subspace 7:37 (slides 15:36)
Quasi-Newton methods 15:37 (slides 22:42)
How to approximate Hessian 22:43 (slide 32:24)
Approximation of inverse Hessian, Sherman-Morrison formula 32:25 (slide 37:43)
Broyden family Quasi-Newton methods, DFP, BFGS 37:44 (slides 41:06, 44:10, 47:01)
Initialization and convergence properties 47:02 (slide 53:26)

Lecture 13. Summary of unconstrained optimization. Optimization with constraints

Summary of unconstrained optimization 0:0 (slide 4:59, one-dimensional methods)
Multivariate methods 05:03 (slides 13:47, 15:41)

Constrained optimization, optimality conditions 15:44 (slides 19:02, 26:32; 30:02)
Karush - Kuhn - Tucker (KKT) first order optimality conditions 30:06 (slide 35:03)
Penalty function method 35:05 (slides 48:36, 54:44, 57:29)

Lecture 14 Lagrange multipliers and penalty function method. Augmented Lagrangian

Lagrange multipliers via penalty function method 0:0 (slides 5:54, 16:30)
Penalty for equality constraints 16:36 (slides penalty 21:08; optimality conditions 24:39)
Barrier method 24:44 (slide 32:24)
Augmented Lagrangian method (or method of multipliers) 32:28 (slides 41:55, 48:40, 51:08)
Augmented Lagrangian for equality constraints 51:11 (slides 53:44)
Optimal solution in one step, when multipliers are known 53:55 (slides 01:02:13)

Lecture 15 Minimax theorem, game theory and Lagrange duality

Introduction 0:0
Minimax theorem 1:19 (slides 6:16)
Game interpretation of minimax 6:20 (slides 12:07)
Saddle point theorem 12:12 (slides 19:03)
Minimax of Lagrangian; weak duality 19:08 (slides 29:07)
Dual problem and weak duality 29:19 (slides 30:39)
Strong duality 30:43 (slides 39:14, 43:03)
Slayter condition for strong duality 43:10 (slides 44:33)
Examples of dual problems:
 Quadratic program 44:38 (slides 49:26)
 Linear program 49:40 (slides 57:48)
Garbage 58:36 -- till end (01:02:24)

Lecture 16 Conic programming 1

Introduction 0:0 (slides 13:43)
Examples of cones - R^n+, Lorenz cone, cone of positive semidefinite matrices13:50 (slides 24:45)
Semidefinite programming (SDP) 24:49 (slides 28:40)
Dual cone 28:43 (slides 33:46, 39:33)
Dual conic problem, weak duality 39:37 (slides 47:26)
Strong conic duality and complementarity slackness 47:34 (slides 56:52)
Example: Dual SDP problem 57:02 (slides 1:05:45)
Minimax problem 1:05:52 (slides 1:13:17, 1:16:12)
Chebyshev approximation 1:16:14 (slides in the next video, 2:10)

Lecture 16b-17 Conic programming 2

Continue with Chebyshev approximation, which starts at 1:16:14 of the previous video (slides 2:10)
Complex Chebyshev approximation 02:16 (slides 15:25, 18:40)

Conversion of different problems to SDP 18:51
Lemma of Schur complement 19:21 (slides 25:14)
Minimize maximal eigenvalue of symmetric matrix 25:18 (slides 31:33)
Linear matrix approximation 31:36 (slides 43:40)
Expression of Linear program via SDP 43:43 (slides 45:39)
Conic quadratic program via SDP 45:40 (slides 49:57)
Barrier method for solution of conic programming problem 50:06 (slides 57:26)
Examples of barriers 57:29 (slides 1:03:12)
Matrix functions 1:03:16 (slides 1:14:53, 1:17:05)
Gradient of trace of matrix function 1:17:09 (slides 1:25:24)
Gradient of log det A 1:25:28 (slides 1:28:28)
Gradient of log det barrier 1:28:31 (slides 1:31:30)

Homework on analytical and numerical computation of gradient and Hessian (file opt_hw_Grad_Hess_compr.avi)

Slides 5:42 6:39 10:12 15:55

Penalty method for semidefinite programming and homework on linear matrix approximation
(file SDP_penalty_method_compress.avi)

slides 3:20 3:26 10:46 11:09 18:39

mzib@cs.technion.ac.il

SUBPAGES (1): SLIDES

תגובות

אין לך הרשאה להוסיף תגובות.