



A PRIMER TO MATHEMATICAL OPTIMIZATION

PROF. DEBDAS GHOSH

Department of Mathematical Sciences
IIT (BHU), Varanasi

PRE-REQUISITES : Calculus, Linear Algebra, Coordinate Geometry

INTENDED AUDIENCE : Third Year Undergraduates of Mathematics / Computer Science /Electrical / Mechanical Engineering

INDUSTRY SUPPORT :Control, Machine Learning

COURSE OUTLINE :

In this era of Machine Learning and Data Science, students often crave for learning the basic tools of optimization theory because machine learning ideas essentially exploit the power of Numerical Linear Algebra, Optimization, and Statistics. The primary aim of this course is to hand over a complete readymade package for beginners in mathematical optimization. 'Complete' in the sense of its mathematical orientation, geometrical explanation, problem-solution sheets, and tutorial sheets. After attending this course, students will get to know the standard methods, and basic and modern results in optimization. The concepts will be explained not only with mathematical rigor but also with geometrical essence so that students feel optimization is fun. The course covers mathematical foundation for optimization and basic techniques of unconstrained and constrained optimization problems.

ABOUT INSTRUCTOR :

Prof. Debdas Ghosh is currently an Assistant Professor in the Department of Mathematical Sciences, Indian Institute of Technology (BHU) Varanasi, India. In 2014, he obtained Ph.D. degree from IIT Kharagpur in Mathematics. His research interest broadly includes Optimization Theory and Applications. He is a recipient of Prof. J. C. Bose Memorial Gold Medal from IIT Kharagpur (2009). He has been awarded Outstanding Potential for Excellence in Research and Academics Award (2014) from his earlier employer, BITS-Pilani. So far, he has published forty international journal papers and fifteen conference papers. He has written a research monograph entitled An Introduction to Analytical Fuzzy Plane Geometry which is published by Springer in the book series of Studies in Fuzziness and Soft Computing. He has edited two conference proceedings volumes on Mathematics and Computing which are published by Springer.

COURSE PLAN :

Week 1: Lecture 1.Introduction and history on the development of optimization theory

Section I: Tools from Linear Algebra

Lecture 2.Vector space, basis and dimension, subspace, inner products and norms, matrix norms, projection,eigen values and eigen vectors

Lecture 3.Definiteness of matrices: eigen value criterion and Sylvester criterion, quadratic forms andquadratic functions

Section II: Tools from Calculus of Several Variables

Lecture 4.Sets in \mathbb{R}^n : balls, neighborhood, interior point, open sets, closed sets, bounded sets, compact sets,level sets, epigraphs

Lecture 5.Supremum and infimum of a set, sequence, subsequence, limsup and liminf

Week 2: Lecture 6.Order of convergence of a sequence

Lecture 7.Limit, continuity, uniform continuity and Lipschitz continuity and their geometries

Lecture 8.Partial derivative, gradient, Hessian, directional derivative

Lecture 9.Differentiability, hierarchy of 'PDs, DD and differentiability'

Lecture 10.Little oh and big Oh notations, Taylor's theorem

Week 3: Section III: Convex Sets and Convex Functions

Lecture 11.Convex sets, examples of commonly occurring convex sets, convexity preserving operations

Lecture 12.Convexity preserving operations continued, separation result (a point and a closed convex set)Lecture 13.Theorems of alternatives: Farkas' theorem

Lecture 14.Theorems of alternatives: Gordan's theorem and other results

Lecture 15.Convex functions, their continuity, Lipschitz continuity and directional derivative

Week 4: Lecture 16.Zeroth order and first order characterizations of convexity

Lecture 17.Second order characterization of convexity, convexity preserving operations

Section IV: Unconstrained Optimization Theory

Lecture 18.Local optima and its variants, saddle point, coercive functions, existence of optima: Weierstrass theorem

Lecture 19.Existence of optima: Theorem for coercive functions. Descent direction and its first order and second order characterizations

Lecture 20.First order and second order necessary and sufficient conditions for minima

Week 5: Section V: Unconstrained Optimization Methods

Lecture 21.General structure of an optimization algorithm, global convergence, descent property, quadratic termination property.

Introduction to line search: exact and inexact line searches

Lecture 22.Inexact line searches: Armijo-Goldstein, Wolfe-Powell, Armijo-backtracking

Lecture 23.On global convergence of gradient descent methods using line search methods: for exact line search

Lecture 24.On global convergence of gradient descent methods using line search methods: for inexact line searches

Lecture 25.Trust region methods: Cauchy point, dog leg and double dog leg methods

Week 6: Lecture 26.Where do general descent methods converge? —almost always to a local minimizer. Capture theorem.

Lecture 27.Scaling of variables

Lecture 28.Practical stopping criteria

Lecture 29.Steepest descent method, descent property, zigzagging, scaling effect, Barzilai and Borwin gradient method

Lecture 30.Global convergence of steepest descent method and order of convergence. Scaling effect and dimension independency.

Week 7: Lecture 31.Newton method and its local convergence.

Lecture 32.Dimension dependency of Newton method. Scaling effect.

Lecture 33.Levenberg-Marquardt and other modified Newton methods

Lecture 34.Quasi-Newton methods - quasi-Newton equation, general quasi-Newton algorithm, variable metric method, general comparison of Newton and quasi-Newton methods

Lecture 35.Global convergence of quasi-Newton methods for uniformly convex functions

Week 8: Lecture 36.Superlinear convergence of quasi-Newton methods

Lecture 37.Basic conjugate direction method

Lecture 38.Convergence rate of the basic CG method

Section VI. Constraint Optimization Theory

Lecture 39.A revisit to Lagrange multipliers method

Lecture 40.Cones for constraint optimization: cone, dual cone, cone of feasible directions, linearizing cone, cone of descent directions

Week 9: Lecture 41.Cones for constraint optimization: tangent cone. Geometric optimality conditions.

Lecture 42.First order FJ and KKT necessary optimality conditions

Lecture 43.First order KKT sufficient optimality condition

Lecture 44.Second order KKT necessary and sufficient optimality conditions

Lecture 45.Constraint qualification

Week 10: Section VII: Lagrangian Duality Theory

Lecture 46.Lagrangian duality: how does one discover duality? Geometric interpretation of duality.

Lecture 47.Results on dual function: dual problem is a convex optimization problem, differentiability of the dual function

Lecture 48.Weak duality theorem. Duality gap. Equal primal and dual objective values imply optimal.

Lecture 49.Strong duality theorem. Convex optimization problem and SCQ implies strong duality.

Lecture 50.Saddle point optimality and absence of duality gap. Relationship between saddle point criteria and KKT conditions

Week 11: Section VIII: Linearly Constrained Nonlinear Optimization Algorithms

Lecture 51.Quadratic programming methods - Active set method

Lecture 52.Quadratic programming method - Interior point method

Lecture 53.Projective gradient algorithm

Lecture 54.Reduced gradient algorithm

Section IX: Nonlinearly Constrained Nonlinear Optimization Algorithms

Lecture 55.Penalty method - Exterior penalty method

Lecture 56.Penalty method - Interior penalty method

Week 12: Lecture 57.Penalty method - Augmented Lagrangian method

Lecture 58.Sequential quadratic programming - Lagrange-Newton method)

Lecture 59.Sequential quadratic programming - Reduced Hessian matrix method

Lecture 60.Trust region method

Lecture 61.Null space method