Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

268

P. M. Pardalos
J. B. Rosen

Constrained Global Optimization: Algorithms and Applications

Springer-Verlag

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

268

P.M. Pardalos
J.B. Rosen

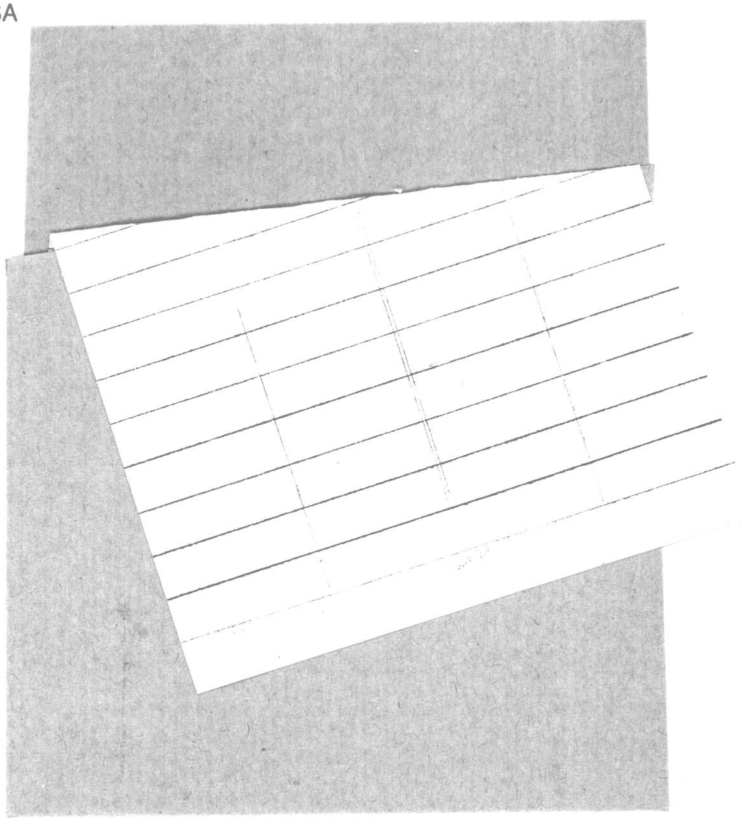
Constrained Global Optimization: Algorithms and Applications

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo

Authors

Panos M. Pardalos
Computer Science Department, The Pennsylvania State University
University Park, PA 16802, USA

J. Ben Rosen
Computer Science Department, University of Minnesota
Minneapolis, MN 55455, USA



CR Subject Classification (1987): G.1.6, G.1.2, F.2.1

ISBN 3-540-18095-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-18095-8 Springer-Verlag New York Berlin Heidelberg

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2145/3140-543210

Lecture Notes in Computer Science

Vol. 167: International Symposium on Programming. Proceedings, 1984. Edited by C. Girault and M. Paul. VI, 262 pages. 1984.

Vol. 168: Methods and Tools for Computer Integrated Manufacturing. Edited by R. Dillmann and U. Rembold. XVI, 528 pages. 1984.

Vol. 169: Ch. Ronse, Feedback Shift Registers. II, 1–2, 145 pages. 1984.

Vol. 170: Seventh International Conference on Automated Deduction. Edited by R.E. Shostak. IV, 508 pages. 1984.

Vol. 171: Logic and Machines: Decision Problems and Complexity. Proceedings, 1983. Edited by E. Börger, G. Hasenjaeger and D. Rödding. VI, 456 pages. 1984.

Vol. 172: Automata, Languages and Programming. Proceedings, 1984. Edited by J. Paredaens. VIII, 527 pages. 1984.

Vol. 173: Semantics of Data Types. Proceedings, 1984. Edited by G. Kahn, D.B. MacQueen and G. Plotkin. VI, 391 pages. 1984.

Vol. 174: EUROSAM 84. Proceedings, 1984. Edited by J. Fitch. XI, 396 pages. 1984.

Vol. 175: A. Thayse, P-Functions and Boolean Matrix Factorization. VII, 248 pages. 1984.

Vol. 176: Mathematical Foundations of Computer Science 1984. Proceedings, 1984. Edited by M.P. Chytil and V. Koubek. XI, 581 pages. 1984.

Vol. 177: Programming Languages and Their Definition. Edited by C.B. Jones. XXXII, 254 pages. 1984.

Vol. 178: Readings on Cognitive Ergonomics – Mind and Computers. Proceedings, 1984. Edited by G.C. van der Veer, M.J. Tauber, T.R.G. Green and P. Gorny. VI, 269 pages. 1984.

Vol. 179: V. Pan, How to Multiply Matrices Faster. XI, 212 pages. 1984.

Vol. 180: Ada Software Tools Interfaces. Proceedings, 1983. Edited by P.J.L. Wallis. III, 164 pages. 1984.

Vol. 181: Foundations of Software Technology and Theoretical Computer Science. Proceedings, 1984. Edited by M. Joseph and R. Shyamasundar. VIII, 468 pages. 1984.

Vol. 182: STACS 85. 2nd Annual Symposium on Theoretical Aspects of Computer Science. Proceedings, 1985. Edited by K. Mehlhorn. VII, 374 pages. 1985.

Vol. 183: The Munich Project CIP. Volume I: The Wide Spectrum Language CIP-L. By the CIP Language Group. XI, 275 pages. 1985.

Vol. 184: Local Area Networks: An Advanced Course. Proceedings, 1983. Edited by D. Hutchison, J. Mariani and D. Shepherd. VIII, 497 pages. 1985.

Vol. 185: Mathematical Foundations of Software Development. Proceedings, 1985. Volume 1: Colloquium on Trees in Algebra and Programming (CAAP'85). Edited by H. Ehrig, C. Floyd, M. Nivat and J. Thatcher. XIV, 418 pages. 1985.

Vol. 186: Formal Methods and Software Development. Proceedings, 1985. Volume 2: Colloquium on Software Engineering (CSE). Edited by H. Ehrig, C. Floyd, M. Nivat and J. Thatcher. XIV, 455 pages. 1985.

Vol. 187: F.S. Chaghaghi, Time Series Package (TSPACK). III, 305 pages. 1985.

Vol. 188: Advances in Petri Nets 1984. Edited by G. Rozenberg with the cooperation of H. Genrich and G. Roucairol. VII, 467 pages. 1985.

Vol. 189: M.S. Sherman, Paragon. XI, 376 pages. 1985.

Vol. 190: M.W. Alford, J.P. Ansart, G. Hommel, L. Lamport, B. Liskov, G.P. Mullery and F.B. Schneider, Distributed Systems. Edited by M. Paul and H.J. Siegert. VI, 573 pages. 1985.

Vol. 191: H. Barringer, A Survey of Verification Techniques for Parallel Programs. VI, 115 pages. 1985.

Vol. 192: Automata on Infinite Words. Proceedings, 1984. Edited by M. Nivat and D. Perrin. V, 216 pages 1985.

Vol. 193: Logics of Programs. Proceedings, 1985. Edited by R. Parikh. VI, 424 pages. 1985.

Vol. 194: Automata, Languages and Programming. Proceedings, 1985. Edited by W. Brauer. IX, 520 pages. 1985.

Vol. 195: H.J. Stüttgen, A Hierarchical Associative Processing System. XII, 273 pages. 1985.

Vol. 196: Advances in Cryptology. Proceedings of CRYPTO '84. Edited by G.R. Blakley and D. Chaum. IX, 491 pages. 1985.

Vol. 197: Seminar on Concurrency. Proceedings, 1984. Edited by S.D. Brookes, A. W. Roscoe and G. Winskel. X, 523 pages. 1985.

Vol. 198: A. Businger, PORTAL Language Description. VIII, 186 pages. 1985.

Vol. 199: Fundamentals of Computation Theory. Proceedings, 1985. Edited by L. Budach. XII, 533 pages. 1985.

Vol. 200: J.L.A. van de Snepscheut, Trace Theory and VLSI Design. VI, 0–140 pages. 1985.

Vol. 201: Functional Programming Languages and Computer Architecture. Proceedings, 1985. Edited by J.-P. Jouannaud. VI, 413 pages. 1985.

Vol. 202: Rewriting Techniques and Applications. Edited by J.-P. Jouannaud. VI, 441 pages. 1985.

Vol. 203: EUROCAL '85. Proceedings Vol. 1, 1985. Edited by B. Buchberger. V, 233 pages. 1985.

Vol. 204: EUROCAL '85. Proceedings Vol. 2, 1985. Edited by B. F. Caviness. XVI, 650 pages. 1985.

Vol. 205: P. Klint, A Study in String Processing Languages. VIII, 165 pages. 1985.

Vol. 206: Foundations of Software Technology and Theoretical Computer Science. Proceedings, 1985. Edited by S.N. Maheshwari. IX, 522 pages. 1985.

Vol. 207: The Analysis of Concurrent Systems. Proceedings, 1983. Edited by B.T. Denvir, W.T. Harwood, M.I. Jackson and M.J. Wray. VII, 398 pages. 1985.

Vol. 208: Computation Theory. Proceedings, 1984. Edited by A. Skowron. VII, 397 pages. 1985.

Vol. 209: Advances in Cryptology. Proceedings of EUROCRYPT '84. Edited by T. Beth, N. Cot and I. Ingemarsson. VII, 491 pages. 1985.

Vol. 210: STACS 86. Proceedings, 1986. Edited by B. Monien and G. Vidal-Naquet. IX, 368 pages. 1986.

Vol. 211: U. Schöning, Complexity and Structure. V, 99 pages. 1986.

Vol. 212: Interval Mathematics 1985. Proceedings, 1985. Edited by K. Nickel. VI, 227 pages. 1986.

Vol. 213: ESOP 86. Proceedings, 1986. Edited by B. Robinet and R. Wilhelm. VI, 374 pages. 1986.

Vol. 214: CAAP '86. 11th Colloquium on Trees in Algebra and Programming. Proceedings, 1986. Edited by P. Franchi-Zannettacci. VI, 306 pages. 1986.

Vol. 215: Mathematical Methods of Specification and Synthesis of Software Systems '85. Proceedings, 1985. Edited by W. Bibel and K.P. Jantke. 245 pages. 1986.

Vol. 216: Ch. Fernstrom, I. Kruzela, B. Svensson, LUCAS Associative Array Processor: Design, Programming and Application Studies. XII, 323 pages. 1986.

Vol. 217: Programs as Data Objects. Proceedings, 1985. Edited by H. Ganzinger and N.D. Jones. X, 324 pages. 1986.

Vol. 218: Advances in Cryptology – CRYPTO '85. Proceedings, 1985. Edited by H.C. Williams. X, 548 pages. 1986.

Vol. 219: Advances in Cryptology – EUROCRYPT '85. Proceedings, 1985. Edited by F. Pichler. IX, 281 pages. 1986.

Vol. 220: RIMS Symposia on Software Science and Engineering II. Proceedings, 1983 and 1984. Edited by E. Goto, K. Araki and T. Yuasa. XI, 323 pages. 1986.

Vol. 221: Logic Programming '85. Proceedings, 1985. Edited by E. Wada. IX, 311 pages. 1986.

Vol. 222: Advances in Petri Nets 1985. Edited by G. Rozenberg. VI, 498 pages. 1986.

Vol. 223: Structure in Complexity Theory. Proceedings, 1986. Edited by A.L. Selman. VI, 401 pages. 1986.

Vol. 224: Current Trends in Concurrency. Overviews and Tutorials. Edited by J.W. de Bakker, W.-P. de Roever and G. Rozenberg. XII, 716 pages. 1986.

Vol. 225: Third International Conference on Logic Programming. Proceedings, 1986. Edited by E. Shapiro. IX, 720 pages. 1986.

Vol. 226: Automata, Languages and Programming. Proceedings, 1986. Edited by L. Kott. IX, 474 pages. 1986.

Vol. 227: VLSI Algorithms and Architectures – AWOC 86. Proceedings, 1986. Edited by F. Makedon, K. Mehlhorn, T. Papatheodorou and P. Spirakis. VIII, 328 pages. 1986.

Vol. 228: Applied Algebra, Algorithmics and Error-Correcting Codes. AAECC-2. Proceedings, 1984. Edited by A. Poli. VI, 265 pages. 1986.

Vol. 229: Algebraic Algorithms and Error-Correcting Codes. AAECC-3. Proceedings, 1985. Edited by J. Calmet. VII, 416 pages. 1986.

Vol. 230: 8th International Conference on Automated Deduction. Proceedings, 1986. Edited by J.H. Siekmann. X, 708 pages. 1986.

Vol. 231: R. Hausser, NEWCAT: Parsing Natural Language Using Left-Associative Grammar. II, 540 pages. 1986.

Vol. 232: Fundamentals of Artificial Intelligence. Edited by W. Bibel and Ph. Jorrand. VII, 313 pages. 1986.

Vol. 233: Mathematical Foundations of Computer Science 1986. Proceedings, 1986. Edited by J. Gruska, B. Rovan and J. Wiedermann. IX, 650 pages. 1986.

Vol. 234: Concepts in User Interfaces: A Reference Model for Command and Response Languages. By Members of IFIP Working Group 2.7. Edited by D. Beech. X, 116 pages. 1986.

Vol. 235: Accurate Scientific Computations. Proceedings, 1985. Edited by W.L. Miranker and R.A. Toupin. XIII, 205 pages. 1986.

Vol. 236: TEX for Scientific Documentation. Proceedings, 1986. Edited by J. Désarménien. VI, 204 pages. 1986.

Vol. 237: CONPAR 86. Proceedings, 1986. Edited by W. Händler, D. Haupt, R. Jeltsch, W. Juling and O. Lange. X, 418 pages. 1986.

Vol. 238: L. Naish, Negation and Control in Prolog. IX, 119 pages. 1986.

Vol. 239: Mathematical Foundations of Programming Semantics. Proceedings, 1985. Edited by A. Melton. VI, 395 pages. 1986.

Vol. 240: Category Theory and Computer Programming. Proceedings, 1985. Edited by D. Pitt, S. Abramsky, A. Poigné and D. Rydeheard. VII, 519 pages. 1986.

Vol. 241: Foundations of Software Technology and Theoretical Computer Science. Proceedings, 1986. Edited by K.V. Nori. XII, 519 pages. 1986.

Vol. 242: Combinators and Functional Programming Languages. Proceedings, 1985. Edited by G. Cousineau, P.-L. Curien and B. Robinet. V, 208 pages. 1986.

Vol. 243: ICDT '86. Proceedings, 1986. Edited by G. Ausiello and P. Atzeni. VI, 444 pages. 1986.

Vol. 244: Advanced Programming Environments. Proceedings, 1986. Edited by R. Conradi, T.M. Didriksen and D.H. Wanvik. VII, 604 pages. 1986

Vol. 245: H.F. de Groote, Lectures on the Complexity of Bilinear Problems. V, 135 pages. 1987.

Vol. 246: Graph-Theoretic Concepts in Computer Science. Proceedings, 1986. Edited by G. Tinhofer and G. Schmidt. VII, 307 pages. 1987.

Vol. 247: STACS 87. Proceedings, 1987. Edited by F.J. Brandenburg, G. Vidal-Naquet and M. Wirsing. X, 484 pages. 1987.

Vol. 248: Networking in Open Systems. Proceedings, 1986. Edited by G. Müller and R.P. Blanc. VI, 441 pages. 1987.

Vol. 249: TAPSOFT '87. Volume 1. Proceedings, 1987. Edited by H. Ehrig, R. Kowalski, G. Levi and U. Montanari. XIV, 289 pages. 1987.

Vol. 250: TAPSOFT '87. Volume 2. Proceedings, 1987. Edited by H. Ehrig, R. Kowalski, G. Levi and U. Montanari. XIV, 336 pages. 1987.

Vol. 251: V. Akman, Unobstructed Shortest Paths in Polyhedral Environments. VII, 103 pages. 1987.

Vol. 252: VDM '87. VDM – A Formal Method at Work. Proceedings, 1987. Edited by D. Bjørner, C.B. Jones, M. Mac an Airchinnigh and E.J. Neuhold. IX, 422 pages. 1987.

Vol. 253: J.D. Becker, I. Eisele (Eds.), WOOPLOT 86. Parallel Processing: Logic, Organization, and Technology. Proceedings, 1986. V, 226 pages. 1987.

Vol. 254: Petri Nets: Central Models and Their Properties. Advances in Petri Nets 1986, Part I. Proceedings, 1986. Edited by W. Brauer, W. Reisig and G. Rozenberg. X, 480 pages. 1987.

Vol. 255: Petri Nets: Applications and Relationships to Other Models of Concurrency. Advances in Petri Nets 1986, Part II. Proceedings, 1986. Edited by W. Brauer, W. Reisig and G. Rozenberg. X, 516 pages. 1987.

Vol. 256: Rewriting Techniques and Applications. Proceedings, 1987. Edited by P. Lescanne. VI, 285 pages. 1987.

Vol. 257: Database Machine Performance: Modeling Methodologies and Evaluation Strategies. Edited by F. Cesarini and S. Salza. X, 250 pages. 1987.

Vol. 258: PARLE, Parallel Architectures and Languages Europe. Volume I. Proceedings, 1987. Edited by J.W. de Bakker, A.J. Nijman and P.C. Treleaven. XII, 480 pages. 1987.

Vol. 259: PARLE, Parallel Architectures and Languages Europe. Volume II. Proceedings, 1987. Edited by J.W. de Bakker, A.J. Nijman and P.C. Treleaven. XII, 464 pages. 1987.

Vol. 260: D.C. Luckham, F.W. von Henke, B. Krieg-Brückner, O. Owe, ANNA, A Language for Annotating Ada Programs. V, 143 pages. 1987.

Vol. 261: J. Ch. Freytag, Translating Relational Queries into Iterative Programs. XI, 131 pages. 1987.

Vol. 262: A. Burns, A.M. Lister, A.J. Wellings, A Review of Ada Tasking. VIII, 141 pages. 1987.

Vol. 263: A.M. Odlyzko (Ed.), Advances in Cryptology – CRYPTO '86. Proceedings. XI, 489 pages. 1987.

Vol. 264: E. Wada (Ed.), Logic Programming '86. Proceedings, 1986. VI, 179 pages. 1987.

Vol. 265: K.P. Jantke (Ed.), Analogical and Inductive Inference. Proceedings, 1986. VI, 227 pages. 1987.

Vol. 266: G. Rozenberg (Ed.), Advances in Petri Nets 1987. VI, 451 pages. 1987.

Vol. 267: Th. Ottmann (Ed.), Automata, Languages and Programming. Proceedings, 1987. X, 565 pages. 1987.

Vol. 268: P.M. Pardalos, J.B. Rosen, Constrained Global Optimization: Algorithms and Applications. VII, 143 pages. 1987.

Preface

Global optimization is concerned with the characterization and computation of global minima or maxima of nonlinear functions. The general constrained global minimization problem has the following form:

Given: $K \subseteq R^n$ *compact set,* $f: K \to R$ *continuous function*

Find: $x^* \in K$, $f^* = f(x^*)$ *such that* $f^* \le f(x)$ *for all* $x \in K$

Such problems are widespread in mathematical modeling of real world systems for a very broad range of applications. Such applications include economies of scale, fixed charges, allocation and location problems, quadratic assignment and a number of other combinatorial optimization problems. More recently it has been shown that certain aspects of VLSI chip design and database problems can be formulated as constrained global optimization problems with a quadratic objective function. Although standard nonlinear programming algorithms will usually obtain a local minimum to the problem, such a local minimum will only be global when certain conditions are satisfied (such as f and K being convex). In general several local minima may exist and the corresponding function values may differ substantially. The problem of designing algorithms that obtain global solutions is very difficult, since in general, there is no local criterion for deciding whether a local solution is global.

Active research during the past two decades has produced a variety of methods for finding constrained global solutions to nonlinear optimization problems. In this monograph we consider deterministic methods which include those concerned with enumerative techniques, cutting planes, branch and bound, bilinear programming, general approximate algorithms and large-scale approaches.

There has been a significant recent increase in research activity on the subject of constrained global optimization and related computational algorithms. This monograph summarizes much of this recent work and contains an extensive list of references to papers on constrained global optimization, deterministic solution methods, and applications. We hope that this work will be valuable for other researchers in global optimization.

We wish to express our appreciation and thanks to Andrew T. Phillips and Nainan Kovoor who carefully read an earlier version of this monograph and suggested a number of valuable improvements. The authors' research described in this monograph was supported in part by the National Science Foundation Grant DCR8405489.

June 1987 P.M. Pardalos, J.B. Rosen

Table of Contents

Chapter 1 Convex sets and functions

Convex sets and functions play a dominant role in optimization and some of their properties are essential in the study of several algorithms. In this introductory chapter we start with summary of some of the most important properties that we are going to use.

1.1 Convex sets

A subset C of the Euclidean space R^n is said to be convex if for every $x_1, x_2 \in C$ and λ real, $0 \le \lambda \le 1$, the point $\lambda x_1 + (1-\lambda) x_2 \in C$.

The geometric interpretation of this definition is clear. For any two points of C, the line segment joining these two points lies entirely in C.

Given the vectors $x_1, \ldots, x_m$ in R^n and real numbers $\lambda_i \ge 0$ with $\sum_{i=1}^{m} \lambda_i = 1$, the vector sum $\lambda_1 x_1 + \cdots + \lambda_m x_m$ is called a convex combination of these points. Some properties of convex sets are summarized in the next theorems.

Theorem 1.1.1: A subset of R^n is convex iff it contains all the convex combinations of its elements.

Proof: Let C be a convex set in R^n. If $x_i \in C$ and $\lambda_i \ge 0$, $i=1, \ldots, k$, such that $\sum_{i=1}^{k} \lambda_i = 1$, prove by induction on k that the convex combination $\sum_{i=1}^{k} \lambda_i x_i \in C$.

Theorem 1.1.2: Let F be a family of convex sets. Then the intersection $\bigcap_{C \in F} C$ is also a convex set.

However, it is easy to see that the union of convex sets need not be convex. Some other algebraic set operations that preserve convexity are defined below.

Theorem 1.1.3: 1) Let C be a convex set in R^n and a a real number. Then the set $aC = \{x: x = ay, y \in C\}$ is also convex.
2) Let C_1, C_2 be convex sets in R^n. Then the set

$$C_1 + C_2 = \{x: x = x_1+x_2, x_1 \in C_1, x_2 \in C_2\}$$

is also convex.

1.2 Linear and affine spaces

For any $x, y \in R^n$ the inner product $x^T y$ is the real number $\sum_{i=1}^{n} x_i y_i$. The Euclidean norm is defined to be $\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$. Other notations and terminology not defined here are the standard ones used in the literature.

A hyperplane H in R^n is a set of the form

$$H = \{x \in R^n : c^T x = b\}.$$

for some vector $c \in R^n$ and $b \in R$. Similarly we define the closed half spaces

$$H_1 = \{x \in R^n : c^T x \geq b\}, \quad H_2 = \{x \in R^n : c^T x \leq b\}.$$

It is very easy to see that H, H_1, H_2 are all convex sets.

A nonempty subset V of R^n is called a (linear) subspace if the following conditions are true:
i) if $x, y \in V$ then $x+y \in V$,
ii) if $x \in V$ and $r \in R$ then $rx \in V$.

Next we define the structure of linear subspaces. Let $S=\{v_1, \ldots, v_m\}$ be a set of vectors in V. We say that S spans V if for every vector $v \in V$, $v = \sum_{i=1}^{m} c_i v_i$, where the c_i's are real numbers. The set S is said to be linearly independent if we cannot find constants $c_1, \ldots, c_m$ not all zero such that $c_1 v_1 + \cdots + c_m v_m = 0$ (otherwise S is called linearly dependent).

If the set S spans V and is linearly independent we call it a basis of V. The dimension of the subspace V, $dim(V)$, is defined to be the number of vectors in some basis S.

To get more insight into the geometric and algebraic nature of a subspace we equivalently define a linear subspace to be the set

$$V = \{x \in R^n : c_{i1}x_1 + \cdots + c_{in}x_n = 0, \ i = 1, \ldots, m\},$$

that is, V is the solution set of the homogeneous system of linear equations $Cx = 0$ where C is the $m \times n$ matrix of the coefficients c_{ij}. Here the dimension of V is equal to $n - rank(C)$ where the $rank$ of the matrix is the maximum number of linearly independent columns (or rows) of the matrix.

An affine subspace A of R^n is a linear subspace V translated by some vector u, that is $A = \{x \in R^n : x = u + v, \ v \in V\}$. Also $dim(A) = dim(V)$. Equivalently we can define

$$A = \{x \in R^n : c_{i1}x_1 + \cdots + c_{in}x_n = b, \ i = 1, \ldots, m\},$$

that is, A is the solution set of the (nonhomogeneous) linear system $Cx = b$.

From the above discussion it is clear that a hyperplane in R^n is an affine subspace of dimension $n-1$.

1.3 Convex hull

Another important concept in convexity is that of forming the smallest convex set containing a given subset S in R^n. The convex hull of S is the set

$$Co(S) = \cap\{C : C \text{ convex in } R^n \text{ and } C \supseteq S\}.$$

The convex hull of a finite set of points is called a convex polytope.

It is clear that a convex polytope is always bounded. A convex polytope that contains all its boundary points is closed. A point x on the boundary of S is called an extreme point (or vertex) if there are no distinct points $x_1, x_2 \in S$ such that $x = \lambda x_1 + (1-\lambda)x_2$, $0 < \lambda < 1$. For example in the plane a triangle has 3 extreme points, and the sphere has all its boundary points as extreme points. The following theorem gives a very important characterization of a certain kind of convex set.

Theorem (Krein-Milman) 1.3.1: A closed, bounded convex set in R^n is the convex hull of its extreme points.

A (convex) polyhedron is the intersection of finitely many half spaces. Using matrix notation we can define a polyhedron to be the set of points $P=\{x \in R^n : Ax \leq b\}$ where A is an $m \times n$ matrix and $b \in R^m$. Polyhedral sets of this form are of central importance in mathematical programming.

1.4 Convex and concave functions

If $c \in R^n$, the linear function $f : R^n \to R$ defined by $f(x) = c^T x$ is known as a linear function on R^n.

Theorem 1.4.1: Let C be a convex polyhedron in R^n. Consider the linear programming problem

$$\min_{x \in C} f(x) = c^T x. \qquad \text{(LP)}$$

If (LP) has a solution then it occurs at some vertex of C.

Proof: Let $v_1, \ldots, v_k$ be the vertices of C, and let v be the vertex such that $f(v) = \min_{1 \leq i \leq k} \{f(v_i)\}$. Since for any $x \in C$, $x = \sum_{i=1}^{k} \lambda_i v_i$, $\lambda_i \geq 0$, and $\sum_{i=1}^{k} \lambda_i = 1$, we have that $f(x) = \sum_{i=1}^{k} \lambda_i f(v_i) \geq \sum_{i=1}^{k} \lambda_i f(v) = f(v)$.

A function $f : C \subseteq R^n \to R$, where C is a convex set, is called convex if

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

for any $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$. The function f is called concave iff $-f$ is convex.

If the function $f(x)$ has continuous second derivatives, then the following conditions give necessary and sufficient conditions for convexity:

a) $f(y) \geq f(x) + \nabla f(x)(y-x)$ for all $x, y \in C$, or

b) The Hessian is positive semidefinite for all $x \in C$.

Theorem 1.4.2: Let $f_i : S \subseteq R^n \rightarrow R$ be convex functions. Then

1) $\sum_{i=1}^{k} \alpha_i f_i(x)$, $\alpha_i \geq 0$, is also convex

2) $\max_{1 \leq i \leq k} (f_i(x))$ is convex

3) $\max_{1 \leq i \leq k} (0, f_i(x))$ is convex.

We are concerned here with the constrained nonlinear minimization problem of the general form

$$\min_{x \in P} f(x) \qquad \text{(NP)}$$

where P is a compact convex set in R^n and $f(x)$ is a continuous function defined on P.

A point $x^* \in P$ is said to be a relative or local minimum point if $f(x^*) \leq f(x)$ for all $\| x - x^* \| \leq \varepsilon$ for some $\varepsilon > 0$. We say that x^* is a global minimum point if $f(x^*) \leq f(x)$ for all $x \in P$.

When $f(x)$ is convex the problem (NP) is referred to as a convex programming problem, and when $f(x)$ is nonconvex we are talking about nonconvex programming. When the objective function $f(x)$ is convex (or more generally quasiconvex) then every local minimum is also global. This is no longer true for nonconvex functions.

Theorem 1.4.3: Suppose P is a convex compact set and $f : P \subseteq R^n \rightarrow R$ is a convex function. Then every local minimum of f over P is also global.

Proof: Let x^* be a local minimum and suppose that there exists another point y such that $f(y) < f(x^*)$. Then on the line $\lambda y + (1-\lambda)x^*$ $(0 < \lambda < 1)$ we have

$$f(\lambda y + (1-\lambda)x^*) \leq \lambda f(y) + (1-\lambda) f(x^*) < f(x^*)$$

contradicting the fact that x^* is a local minimum.

The above theorem makes convex programming a much easier problem to solve than the general nonlinear programming problem. Consider now the case where $f(x)$ is a concave function. In this case we may have many local minima which are not global. However, this problem has the following important property that also characterizes linear programming.

Theorem 1.4.4: Consider the problem

$$\text{global} \min_{x \in P} f(x)$$

where $f(x)$ is a continuous concave function defined on the bounded polyhedron P. Then every global (and local) minimum is attained at some vertex of P.

Proof: Similar to that of Theorem 1.4.1.

Note that since $\min f(x) = -\max(-f(x))$ minimization of a convex function is equivalent to maximization of a concave function(and vice versa). For continuously differentiable functions convexity and local optima are characterized using the gradient and the Hessian matrix of the function (e.g. [LUEN84], [STOE70]). For additional details regarding different convexity results see [MANG69], [GRUN67] and [ROCK70].

1.5 Convex envelopes

An important concept in nonconvex optimization is that of the convex envelope of a function.

Definition 1.5.1: Let $f : S \to R$ be a lower semi-continuous function, where S is a nonempty subset (of its domain) in R^n. Then the convex envelope of $f(x)$ taken over S is a function $F_S(x)$ such that

i) $F_S(x)$ is convex on the convex hull $Co(S)$

ii) $F_S(x) \leq f(x)$ for all $x \in S$

iii) If $h(x)$ is any convex function defined on $Co(S)$ such that $h(x) \le f(x)$ for all $x \in S$, then $h(x) \le F_S(x)$ for all $x \in Co(S)$.

From this definition, the convex envelope of a function is actually the best convex underestimator over S. Convex envelopes were first introduced by Kleibohm [KLEI67], who proved that with each nonconvex optimization problem is associated a convex one whose global solution is the same as that of the original problem. More precisely we have the following:

Theorem 1.5.1: Consider the problem

$$\text{global} \min_{x \in P} f(x)$$

where P is a convex set in R^n, and assume that the global minimum occurs at $x^* \in P$. Let $F(x)$ be the convex envelope of $f(x)$ over P. Then we have

$$\min_{x \in P} f(x) = \min_{x \in P} F(x)$$

and

$$\{y \in P : f(y) = \min_{x \in P} f(x)\} \subseteq \{y \in P : F(y) = \min_{x \in P} F(x)\}.$$

Proof: By definition $F(x) \le f(x)$ for all $x \in P$. Therefore

$$\min_{x \in P} F(x) \le \min_{x \in P} f(x) = f(x^*).$$

The constant function $G(x) = f(x^*) \le f(x)$ is a convex underestimating function. Again by the definition of the convex envelope we have that $F(x) \ge f(x^*)$ for all x and so

$$\min_{x \in P} F(x) \ge f(x^*) = \min_{x \in P} f(x).$$

We prove the second part by contradiction. Let x^* be a global minimum of $f(x)$ over P, and suppose that x^* is not a minimum of $F(x)$ over P. Let y^* be the minimum. Then

$$F(y^*) < F(x^*) \le f(x^*) = f^*.$$

Consider now the function $H(x) = \max(f^*, F(x))$. Then $H(x)$ is convex since the maximum of two convex functions is convex. Now $H(x) \ge F(x)$ for all $x \in P$ and