WIKIPEDIA

Convex optimization

Convex optimization is a subfield of mathematical optimization that studies the problem of minimizing convex functions over convex sets. Many classes of convex optimization problems admit polynomial-time algorithms,[1] whereas mathematical optimization is in general NP-hard.[2][3][4]

Convex optimization has applications in a wide range of disciplines, such as automatic control systems, estimation and signal processing, communications and networks, electronic circuit design,[5] data analysis and modeling, finance, statistics (optimal experimental design),[6] and structural optimization, where the approximation concept has proven to be efficient.[7][8] With recent advancements in computing and optimization algorithms, convex programming is nearly as straightforward as linear programming.[9]

Contents

Definition

A convex optimization problem is an optimization problem in which the objective function is a convex function and the feasible set is a convex set. A function f mapping some subset of $\mathbb{R}^n$ into $\mathbb{R} \cup \{\pm\infty\}$ is convex if its domain is convex and for all $\theta \in [0,1]$ and all x, y in its domain, the following condition holds: $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$. A set S is convex if for all members $x, y \in S$ and all $\theta \in [0,1]$, we have that $\theta x + (1-\theta)y \in S$.

Concretely, a convex optimization problem is the problem of finding some $\mathbf{x}^* \in C$ attaining

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in C\},$$

where the objective function $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex, as is the feasible set C.[10] [11] If such a point exists, it is referred to as an *optimal point* or *solution*; the set of all optimal points is called the *optimal set*. If f is unbounded below over C or the infimum is not attained, then the optimization problem is said to be *unbounded*. Otherwise, if C is the empty set, then the problem is said to be *infeasible*.[12]

Standard form

A convex optimization problem is in *standard form* if it is written as

$$\begin{aligned} &\underset{\mathbf{x}}{\operatorname{minimize}} && f(\mathbf{x}) \\ &\operatorname{subject\ to} && g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \\ &&& h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p, \end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, the function $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are convex, and $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, p$, are affine.[12] This notation describes the problem of finding $\mathbf{x} \in \mathbb{R}^n$ that minimizes $f(\mathbf{x})$ among all $\mathbf{x}$ satisfying $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, m$ and $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, p$. The function f is the objective function of the problem, and the functions g_i and h_i are the constraint functions.

The feasible set C of the optimization problem consists of all points $\mathbf{x} \in \mathcal{D}$ satisfying the constraints. This set is convex because $\mathcal{D}$ is convex, the sublevel sets of convex functions are convex, affine sets are convex, and the intersection of convex sets is convex.[13]

A solution to a convex optimization problem is any point $\mathbf{x} \in C$ attaining $\inf\{f(\mathbf{x}) : \mathbf{x} \in C\}$. In general, a convex optimization problem may have zero, one, or many solutions.

Many optimization problems can be equivalently formulated in this standard form. For example, the problem of maximizing a concave function f can be re-formulated equivalently as the problem of minimizing the convex function $-f$. The problem of maximizing a concave function over a convex set is commonly called a convex optimization problem.

Properties

The following are useful properties of convex optimization problems:[14][12]

- every local minimum is a global minimum;
- the optimal set is convex;
- if the objective function is *strictly* convex, then the problem has at most one optimal point.

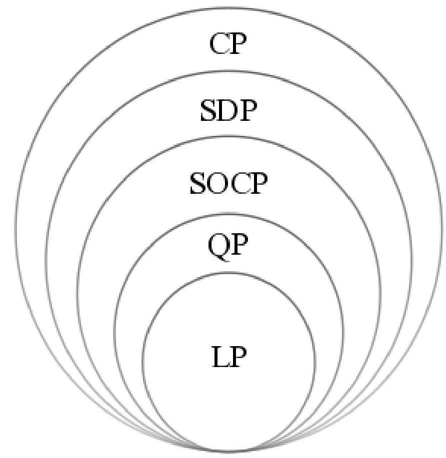
These results are used by the theory of convex minimization along with geometric notions from functional analysis (in Hilbert spaces) such as the Hilbert projection theorem, the separating hyperplane theorem, and Farkas' lemma.

Applications

Ben-Hain and Elishakoff[15] (1990), Elishakoff et al.[16] (1994) applied convex analysis to **model uncertainty**.

The following problem classes are all convex optimization problems, or can be reduced to convex optimization problems via simple transformations:[12][17]

- Least squares
- Linear programming
- Convex quadratic minimization with linear constraints
- Quadratic minimization with convex quadratic constraints
- Conic optimization
- Geometric programming
- Second order cone programming
- Semidefinite programming
- Entropy maximization with appropriate constraints



A hierarchy of convex optimization problems. (LP: linear program, QP: quadratic program, SOCP second-order cone program, SDP: semidefinite program, CP: cone program.)

Convex optimization has practical applications for the following.

- portfolio optimization[18]
- worst-case risk analysis[18]
- optimal advertising[18]
- variations of statistical regression (including regularization and quantile regression)[18]
- model fitting[18] (particularly multiclass classification[19])
- electricity generation optimization[19]
- combinatorial optimization[19]

Lagrange multipliers

Consider a convex minimization problem given in standard form by a cost function $f(x)$ and inequality constraints $g_i(x) \leq 0$ for $1 \leq i \leq m$. Then the domain $\mathcal{X}$ is:

$$\mathcal{X} = \{x \in X | g_1(x), \ldots, g_m(x) \leq 0\}.$$

The Lagrangian function for the problem is

$$L(x, \lambda_0, \lambda_1, \ldots, \lambda_m) = \lambda_0 f(x) + \lambda_1 g_1(x) + \cdots + \lambda_m g_m(x).$$

For each point x in X that minimizes f over X, there exist real numbers $\lambda_0, \lambda_1, \ldots, \lambda_m$, called Lagrange multipliers, that satisfy these conditions simultaneously:

1. x minimizes $L(y, \lambda_0, \lambda_1, \ldots, \lambda_m)$ over all $y \in X$,
2. $\lambda_0, \lambda_1, \ldots, \lambda_m \geq 0$, with at least one $\lambda_k > 0$,
3. $\lambda_1 g_1(x) = \cdots = \lambda_m g_m(x) = 0$ (complementary slackness).

If there exists a "strictly feasible point", that is, a point z satisfying

$g_1(z), \ldots, g_m(z) < 0,$

then the statement above can be strengthened to require that $\lambda_0 = 1$.

Conversely, if some x in X satisfies (1)–(3) for scalars $\lambda_0, \ldots, \lambda_m$ with $\lambda_0 = 1$ then x is certain to minimize f over X.

Algorithms

Unconstrained convex optimization can be easily solved with gradient descent (a special case of steepest descent) or Newton's method, combined with line search for an appropriate step size; these can be mathematically proven to converge quickly, especially the latter method.[20] Convex optimization with linear equality constraints can also be solved using KKT matrix techniques if the objective function is a quadratic function (which generalizes to a variation of Newton's method, which works even if the point of initialization does not satisfy the constraints), but can also generally be solved by eliminating the equality constraints with linear algebra or solving the dual problem.[20] Finally, convex optimization with both linear equality constraints and convex inequality constraints can be solved by applying an unconstrained convex optimization technique to the objective function plus logarithmic barrier terms.[20] (When the starting point is not feasible - that is, satisfying the constraints - this is preceded by so-called *phase I* methods, which either find a feasible point or show that none exist. Phase I methods generally consist of reducing the search in question to yet another convex optimization problem.[20])

Convex optimization problems can also be solved by the following contemporary methods:[21]

- Bundle methods (Wolfe, Lemaréchal, Kiwiel), and
- Subgradient projection methods (Polyak),
- Interior-point methods,[1] which make use of self-concordant barrier functions [22] and self-regular barrier functions.[23]
- Cutting-plane methods
- Ellipsoid method
- Subgradient method
- Dual subgradients and the drift-plus-penalty method

Subgradient methods can be implemented simply and so are widely used.[24] Dual subgradient methods are subgradient methods applied to a dual problem. The drift-plus-penalty method is similar to the dual subgradient method, but takes a time average of the primal variables.

Implementations

Convex optimization and related algorithms have been implemented in the following software programs:

Program	Language	Description	FOSS?	Ref
CVX	MATLAB	Interfaces with SeDuMi and SDPT3 solvers; designed to only express convex optimization problems.	Yes	[25]
CVXMOD	Python	Interfaces with the CVXOPT solver.	Yes	[25]
CVXPY	Python			[26]
CVXR	R		Yes	[27]
YALMIP	MATLAB	Interfaces with CPLEX, GUROBI, MOSEK, SDPT3, SEDUMI, CSDP, SDPA, PENNON solvers; also supports integer and nonlinear optimization, and some nonconvex optimization. Can perform robust optimization with uncertainty in LP/SOCP/SDP constraints.	Yes	[25]
LMI lab	MATLAB	Expresses and solves semidefinite programming problems (called "linear matrix inequalities")	No	[25]
LMIlab translator		Transforms LMI lab problems into SDP problems.	Yes	[25]
xLMI	MATLAB	Similar to LMI lab, but uses the SeDuMi solver.	Yes	[25]
AIMMS		Can do robust optimization on linear programming (with MOSEK to solve second-order cone programming) and mixed integer linear programming. Modeling package for LP + SDP and robust versions.	No	[25]
ROME		Modeling system for robust optimization. Supports distributionally robust optimization and uncertainty sets.	Yes	[25]
GloptiPoly 3	MATLAB, Octave	Modeling system for polynomial optimization.	Yes	[25]
SOSTOOLS		Modeling system for polynomial optimization. Uses SDPT3 and SeDuMi. Requires Symbolic Computation Toolbox.	Yes	[25]
SparsePOP		Modeling system for polynomial optimization. Uses the SDPA or SeDuMi solvers.	Yes	[25]
CPLEX		Supports primal-dual methods for LP + SOCP. Can solve LP, QP, SOCP, and mixed integer linear programming problems.	No	[25]
CSDP	C	Supports primal-dual methods for LP + SDP. Interfaces available for MATLAB, R, and Python. Parallel version available. SDP solver.	Yes	[25]
CVXOPT	Python	Supports primal-dual methods for LP + SOCP + SDP. Uses Nesterov-Todd scaling. Interfaces to MOSEK and DSDP.	Yes	[25]
MOSEK		Supports primal-dual methods for LP + SOCP.	No	[25]
SeDuMi	MATLAB, MEX	Solves LP + SOCP + SDP. Supports primal-dual methods for LP + SOCP + SDP.	Yes	[25]

Program	Language	Description	FOSS?	Ref
SDPA	C++	Solves LP + SDP. Supports primal-dual methods for LP + SDP. Parallelized and extended precision versions are available.	Yes	[25]
SDPT3	MATLAB, MEX	Solves LP + SOCP + SDP. Supports primal-dual methods for LP + SOCP + SDP.	Yes	[25]
ConicBundle		Supports general-purpose codes for LP + SOCP + SDP. Uses a bundle method. Special support for SDP and SOCP constraints.	Yes	[25]
DSDP		Supports general-purpose codes for LP + SDP. Uses a dual interior point method.	Yes	[25]
LOQO		Supports general-purpose codes for SOCP, which it treats as a nonlinear programming problem.	No	[25]
PENNON		Supports general-purpose codes. Uses an augmented Lagrangian method, especially for problems with SDP constraints.	No	[25]
SDPLR		Supports general-purpose codes. Uses low-rank factorization with an augmented Lagrangian method.	Yes	[25]
GAMS		Modeling system for linear, nonlinear, mixed integer linear/nonlinear, and second-order cone programming problems.	No	[25]
Optimization Services		XML standard for encoding optimization problems and solutions.		[25]

Extensions

Extensions of convex optimization include the optimization of biconvex, pseudo-convex, and quasiconvex functions. Extensions of the theory of convex analysis and iterative methods for approximately solving non-convex minimization problems occur in the field of generalized convexity, also known as abstract convex analysis.

See also

- Duality
- Karush–Kuhn–Tucker conditions
- Optimization problem
- Proximal gradient method

Notes

1. Nesterov & Nemirovskii 1994
2. Murty, Katta; Kabadi, Santosh (1987). "Some NP-complete problems in quadratic and nonlinear programming". *Mathematical Programming*. **39** (2): 117–129. doi:10.1007/BF02592948 (https://doi.org/10.1007%2FBF02592948). hdl:2027.42/6740 (https://hdl.handle.net/2027.42%2F6740). S2CID 30500771 (https://api.semanticscholar.org/CorpusID:30500771).
3. Sahni, S. "Computationally related problems," in SIAM Journal on Computing, 3, 262--279, 1974.

4. Quadratic programming with one negative eigenvalue is NP-hard (https://link.springer.com/article/10.1007/BF00120662), Panos M. Pardalos and Stephen A. Vavasis in *Journal of Global Optimization*, Volume 1, Number 1, 1991, pg.15-22.
5. Boyd & Vandenberghe 2004, p. 17
6. Chritensen/Klarbring, chpt. 4.
7. Boyd & Vandenberghe 2004
8. Schmit, L.A.; Fleury, C. 1980: *Structural synthesis by combining approximation concepts and dual methods*. J. Amer. Inst. Aeronaut. Astronaut 18, 1252-1260
9. Boyd & Vandenberghe 2004, p. 8
10. Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1996). *Convex analysis and minimization algorithms: Fundamentals* (https://books.google.com/books?id=Gdl4Jc3RVjcC&q=lemarechal+convex+analysis+and+minimization). p. 291. ISBN 9783540568506.
11. Ben-Tal, Aharon; Nemirovskiĭ, Arkadiĭ Semenovich (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications* (https://books.google.com/books?id=M3MqpEJ3jzQC&q=Lectures+on+Modern+Convex+Optimization:+Analysis,+Algorithms,). pp. 335–336. ISBN 9780898714913.
12. Boyd & Vandenberghe 2004, chpt. 4
13. Boyd & Vandenberghe 2004, chpt. 2
14. Rockafellar, R. Tyrrell (1993). "Lagrange multipliers and optimality" (http://web.williams.edu/Mathematics/sjmiller/public_html/105Sp10/handouts/Rockafellar_LagrangeMultAndOptimality.pdf) (PDF). *SIAM Review*. **35** (2): 183–238. CiteSeerX 10.1.1.161.7209 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.161.7209). doi:10.1137/1035044 (https://doi.org/10.1137%2F1035044).
15. Ben Haim Y. and Elishakoff I., Convex Models of Uncertainty in Applied Mechanics, Elsevier Science Publishers, Amsterdam, 1990
16. I. Elishakoff, I. Lin Y.K. and Zhu L.P., Probabilistic and Convex Modeling of Acoustically Excited Structures, Elsevier Science Publishers, Amsterdam, 1994
17. Agrawal, Akshay; Verschueren, Robin; Diamond, Steven; Boyd, Stephen (2018). "A rewriting system for convex optimization problems" (https://web.stanford.edu/~boyd/papers/pdf/cvxpy_rewriting.pdf) (PDF). *Control and Decision*. **5** (1): 42–60. arXiv:1709.04494 (https://arxiv.org/abs/1709.04494). doi:10.1080/23307706.2017.1397554 (https://doi.org/10.1080%2F23307706.2017.1397554). S2CID 67856259 (https://api.semanticscholar.org/CorpusID:67856259).
18. Boyd, Stephen; Diamond, Stephen; Zhang, Junzi; Agrawal, Akshay. "Convex Optimization Applications" (https://web.stanford.edu/~boyd/papers/pdf/cvx_applications.pdf) (PDF). Archived (https://web.archive.org/web/20151001185038/http://web.stanford.edu/~boyd/papers/pdf/cvx_applications.pdf) (PDF) from the original on 2015-10-01. Retrieved 12 Apr 2021.
19. Malick, Jérôme (2011-09-28). "Convex optimization: applications, formulations, relaxations" (https://www-ljk.imag.fr//membres/Jerome.Malick/Talks/11-INRIA.pdf) (PDF). Archived (https://web.archive.org/web/20210412044738/https://www-ljk.imag.fr//membres/Jerome.Malick/Talks/11-INRIA.pdf) (PDF) from the original on 2021-04-12. Retrieved 12 Apr 2021.
20. Boyd, Stephen; Vandenberghe, Lieven (2004). *Convex Optimization* (https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf) (PDF). Cambridge University Press. ISBN 978-0-521-83378-3. Retrieved 12 Apr 2021.
21. For methods for convex minimization, see the volumes by Hiriart-Urruty and Lemaréchal (bundle) and the textbooks by Ruszczyński, Bertsekas, and Boyd and Vandenberghe (interior point).
22. Nesterov, Yurii; Arkadii, Nemirovskii (1995). *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics. ISBN 978-0898715156.

23. Peng, Jiming; Roos, Cornelis; Terlaky, Tamás (2002). "Self-regular functions and new search directions for linear and semidefinite optimization". *Mathematical Programming*. **93** (1): 129–171. doi:10.1007/s101070200296 (https://doi.org/10.1007%2Fs101070200296). ISSN 0025-5610 (https://www.worldcat.org/issn/0025-5610). S2CID 28882966 (https://api.semanticscholar.org/CorpusID:28882966).
24. Bertsekas
25. Borchers, Brian. "An Overview Of Software For Convex Optimization" (https://web.archive.org/web/20170918180026/http://infohost.nmt.edu/~borchers/presentation.pdf) (PDF). Archived from the original (http://infohost.nmt.edu/~borchers/presentation.pdf) (PDF) on 2017-09-18. Retrieved 12 Apr 2021.
26. "Welcome to CVXPY 1.1 — CVXPY 1.1.11 documentation" (https://www.cvxpy.org/). *www.cvxpy.org*. Retrieved 2021-04-12.
27. "Disciplined Convex Optimiation - CVXR" (https://www.cvxgrp.org/CVXR/). *www.cvxgrp.org*. Retrieved 2021-06-17.

References

- Bertsekas, Dimitri P.; Nedic, Angelia; Ozdaglar, Asuman (2003). *Convex Analysis and Optimization*. Belmont, MA.: Athena Scientific. ISBN 978-1-886529-45-8.
- Bertsekas, Dimitri P. (2009). *Convex Optimization Theory*. Belmont, MA.: Athena Scientific. ISBN 978-1-886529-31-1.
- Bertsekas, Dimitri P. (2015). *Convex Optimization Algorithms*. Belmont, MA.: Athena Scientific. ISBN 978-1-886529-28-1.
- Borwein, Jonathan; Lewis, Adrian (2000). *Convex Analysis and Nonlinear Optimization: Theory and Examples, Second Edition* (https://carma.newcastle.edu.au/resources/jon/Preprints/Books/CaNo2/cano2f.pdf) (PDF). Springer. Retrieved 12 Apr 2021.
- Christensen, Peter W.; Anders Klarbring (2008). *An introduction to structural optimization* (https://books.google.com/books?id=80IeN__MYI8C). Vol. 153. Springer Science & Business Media. ISBN 9781402086663.
- Hiriart-Urruty, Jean-Baptiste, and Lemaréchal, Claude. (2004). *Fundamentals of Convex analysis*. Berlin: Springer.
- Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993). *Convex analysis and minimization algorithms, Volume I: Fundamentals*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Vol. 305. Berlin: Springer-Verlag. pp. xviii+417. ISBN 978-3-540-56850-6. MR 1261420 (https://www.ams.org/mathscinet-getitem?mr=1261420).
- Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993). *Convex analysis and minimization algorithms, Volume II: Advanced theory and bundle methods*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Vol. 306. Berlin: Springer-Verlag. pp. xviii+346. ISBN 978-3-540-56852-0. MR 1295240 (https://www.ams.org/mathscinet-getitem?mr=1295240).
- Kiwiel, Krzysztof C. (1985). *Methods of Descent for Nondifferentiable Optimization* (https://archive.org/details/methodsofdescent0000kiwi). Lecture Notes in Mathematics. New York: Springer-Verlag. ISBN 978-3-540-15642-0.
- Lemaréchal, Claude (2001). "Lagrangian relaxation". In Michael Jünger and Denis Naddef (ed.). *Computational combinatorial optimization: Papers from the Spring School held in Schloß Dagstuhl, May 15–19, 2000*. Lecture Notes in Computer Science. Vol. 2241. Berlin: Springer-Verlag. pp. 112–156. doi:10.1007/3-540-45586-8_4 (https://doi.org/10.1007%2F3-540-45586-8_4).

ISBN 978-3-540-42877-0. MR 1900016 (https://www.ams.org/mathscinet-getitem?mr=1900016). S2CID 9048698 (https://api.semanticscholar.org/CorpusID:9048698).

- Nesterov, Yurii; Nemirovskii, Arkadii (1994). *Interior Point Polynomial Methods in Convex Programming*. SIAM.
- Nesterov, Yurii. (2004). *Introductory Lectures on Convex Optimization (https://books.google.com/books?hl=en&lr=&id=2-ElBQAAQBAJ&oi=fnd&pg=PA1&dq=%22Introductory+Lectures+on+Convex+Optimization%22&ots=wltU7svijv&sig=iknjb0X1jb2uiVAPSn0QPyYGBYg#v=onepage&q=%22Introductory%20Lectures%20on%20Convex%20Optimization%22&f=false)*, Kluwer Academic Publishers
- Rockafellar, R. T. (1970). *Convex analysis*. Princeton: Princeton University Press.
- Ruszczyński, Andrzej (2006). *Nonlinear Optimization*. Princeton University Press.
- Schmit, L.A.; Fleury, C. 1980: *Structural synthesis by combining approximation concepts and dual methods*. J. Amer. Inst. Aeronaut. Astronaut 18, 1252-1260

External links

- EE364a: Convex Optimization I (https://web.stanford.edu/class/ee364a/) and EE364b: Convex Optimization II (https://web.stanford.edu/class/ee364b/), Stanford course homepages
- 6.253: Convex Analysis and Optimization (https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-253-convex-analysis-and-optimization-spring-2012/lecture-notes/), an MIT OCW course homepage
- Brian Borchers, An overview of software for convex optimization (https://web.archive.org/web/20170918180026/http://infohost.nmt.edu/~borchers/presentation.pdf)