

WIKIPEDIA

Duality (optimization)

In mathematical optimization theory, **duality** or the **duality principle** is the principle that optimization problems may be viewed from either of two perspectives, the **primal problem** or the **dual problem**. The solution to the dual problem provides a lower bound to the solution of the primal (minimization) problem.^[1] However in general the optimal values of the primal and dual problems need not be equal. Their difference is called the duality gap. For convex optimization problems, the duality gap is zero under a constraint qualification condition.

Contents

Dual problem

Duality gap

Linear case

Relationship between the primal problem and the dual problem

Nonlinear case

The strong Lagrangian principle: Lagrange duality

Convex problems

History

See also

Notes

References

Books

Articles

Dual problem

Usually the term "dual problem" refers to the *Lagrangian dual problem* but other dual problems are used – for example, the Wolfe dual problem and the Fenchel dual problem. The Lagrangian dual problem is obtained by forming the Lagrangian of a minimization problem by using nonnegative Lagrange multipliers to add the constraints to the objective function, and then solving for the primal variable values that minimize the original objective function. This solution gives the primal variables as functions of the Lagrange multipliers, which are called dual variables, so that the new problem is to maximize the objective function with respect to the dual variables under the derived constraints on the dual variables (including at least the nonnegativity constraints).

In general given two dual pairs of separated locally convex spaces (X, X^*) and (Y, Y^*) and the function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$, we can define the primal problem as finding \hat{x} such that $f(\hat{x}) = \inf_{x \in X} f(x)$. In other words, if \hat{x} exists, $f(\hat{x})$ is the minimum of the function f and the infimum (greatest lower bound) of the function is attained.

If there are constraint conditions, these can be built into the function f by letting $\tilde{f} = f + I_{\text{constraints}}$ where $I_{\text{constraints}}$ is a suitable function on X that has a minimum 0 on the constraints, and for which one can prove that $\inf_{x \in X} \tilde{f}(x) = \inf_{x \text{ constrained}} f(x)$. The latter condition is trivially, but not always conveniently, satisfied for the characteristic function (i.e. $I_{\text{constraints}}(x) = 0$ for x satisfying the constraints and $I_{\text{constraints}}(x) = \infty$ otherwise). Then extend \tilde{f} to a perturbation function $F : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$ such that $F(x, 0) = \tilde{f}(x)$.^[2]

The duality gap is the difference of the right and left hand sides of the inequality

$$\sup_{y^* \in Y^*} -F^*(0, y^*) \leq \inf_{x \in X} F(x, 0),$$

where F^* is the convex conjugate in both variables and **sup** denotes the supremum (least upper bound).^{[2][3][4]}

Duality gap

The duality gap is the difference between the values of any primal solutions and any dual solutions. If d^* is the optimal dual value and p^* is the optimal primal value, then the duality gap is equal to $p^* - d^*$. This value is always greater than or equal to 0. The duality gap is zero if and only if strong duality holds. Otherwise the gap is strictly positive and weak duality holds.^[5]

In computational optimization, another "duality gap" is often reported, which is the difference in value between any dual solution and the value of a feasible but suboptimal iterate for the primal problem. This alternative "duality gap" quantifies the discrepancy between the value of a current feasible but suboptimal iterate for the primal problem and the value of the dual problem; the value of the dual problem is, under regularity conditions, equal to the value of the *convex relaxation* of the primal problem: The convex relaxation is the problem arising replacing a non-convex feasible set with its closed convex hull and with replacing a non-convex function with its convex closure, that is the function that has the epigraph that is the closed convex hull of the original primal objective function.^{[6][7][8][9][10][11][12][13][14][15][16]}

Linear case

Linear programming problems are optimization problems in which the objective function and the constraints are all linear. In the primal problem, the objective function is a linear combination of n variables. There are m constraints, each of which places an upper bound on a linear combination of the n variables. The goal is to maximize the value of the objective function subject to the constraints. A *solution* is a vector (a list) of n values that achieves the maximum value for the objective function.

In the dual problem, the objective function is a linear combination of the m values that are the limits in the m constraints from the primal problem. There are n dual constraints, each of which places a lower bound on a linear combination of m dual variables.

Relationship between the primal problem and the dual problem

In the linear case, in the primal problem, from each sub-optimal point that satisfies all the constraints, there is a direction or subspace of directions to move that increases the objective function. Moving in any such direction is said to remove slack between the candidate solution and one or more constraints. An *infeasible* value of the candidate solution is one that exceeds one or more of the constraints.

In the dual problem, the dual vector multiplies the constraints that determine the positions of the constraints in the primal. Varying the dual vector in the dual problem is equivalent to revising the upper bounds in the primal problem. The lowest upper bound is sought. That is, the dual vector is minimized in order to remove slack between the candidate positions of the constraints and the actual optimum. An infeasible value of the dual vector is one that is too low. It sets the candidate positions of one or more of the constraints in a position that excludes the actual optimum.

This intuition is made formal by the equations in Linear programming: Duality.

Nonlinear case

In nonlinear programming, the constraints are not necessarily linear. Nonetheless, many of the same principles apply.

To ensure that the global maximum of a non-linear problem can be identified easily, the problem formulation often requires that the functions be convex and have compact lower level sets.

This is the significance of the Karush–Kuhn–Tucker conditions. They provide necessary conditions for identifying local optima of non-linear programming problems. There are additional conditions (constraint qualifications) that are necessary so that it will be possible to define the direction to an *optimal* solution. An optimal solution is one that is a local optimum, but possibly not a global optimum.

The strong Lagrangian principle: Lagrange duality

Given a nonlinear programming problem in standard form

$$\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } f_i(x) \leq 0, \ i \in \{1, \dots, m\} \\ &\quad h_i(x) = 0, \ i \in \{1, \dots, p\} \end{aligned}$$

with the domain $\mathcal{D} \subset \mathbb{R}^n$ having non-empty interior, the *Lagrangian function* $\Lambda : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ is defined as

$$\Lambda(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x).$$

The vectors λ and ν are called the *dual variables* or *Lagrange multiplier vectors* associated with the problem. The *Lagrange dual function* $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ is defined as

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \Lambda(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right).$$

The dual function g is concave, even when the initial problem is not convex, because it is a point-wise infimum of affine functions. The dual function yields lower bounds on the optimal value p^* of the initial problem; for any $\lambda \geq 0$ and any ν we have $g(\lambda, \nu) \leq p^*$.

If a constraint qualification such as Slater's condition holds and the original problem is convex, then we have strong duality, i.e. $d^* = \max_{\lambda \geq 0, \nu} g(\lambda, \nu) = \inf f_0 = p^*$.

Convex problems

For a convex minimization problem with inequality constraints,

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

the Lagrangian dual problem is

$$\begin{array}{ll} \underset{u}{\text{maximize}} & \inf_x \left(f(x) + \sum_{j=1}^m u_j g_j(x) \right) \\ \text{subject to} & u_i \geq 0, \quad i = 1, \dots, m \end{array}$$

where the objective function is the Lagrange dual function. Provided that the functions f and g_1, \dots, g_m are continuously differentiable, the infimum occurs where the gradient is equal to zero. The problem

$$\begin{array}{ll} \underset{x, u}{\text{maximize}} & f(x) + \sum_{j=1}^m u_j g_j(x) \\ \text{subject to} & \nabla f(x) + \sum_{j=1}^m u_j \nabla g_j(x) = 0 \\ & u_i \geq 0, \quad i = 1, \dots, m \end{array}$$

is called the Wolfe dual problem. This problem may be difficult to deal with computationally, because the objective function is not concave in the joint variables (u, x) . Also, the equality constraint

$\nabla f(x) + \sum_{j=1}^m u_j \nabla g_j(x)$ is nonlinear in general, so the Wolfe dual problem is typically a nonconvex

optimization problem. In any case, weak duality holds.^[17]

History

According to George Dantzig, the duality theorem for linear optimization was conjectured by John von Neumann immediately after Dantzig presented the linear programming problem. Von Neumann noted that he was using information from his game theory, and conjectured that two person zero sum matrix game was equivalent to linear programming. Rigorous proofs were first published in 1948 by Albert W. Tucker and his group. (Dantzig's foreword to Nering and Tucker, 1993)

See also

- [Convex duality](#)
- [Duality](#)
- [Relaxation \(approximation\)](#)

Notes

1. Boyd, Stephen P.; Vandenberghe, Lieven (2004). *Convex Optimization* (https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf#page=230) (pdf). Cambridge University Press. p. 216. ISBN 978-0-521-83378-3. Retrieved October 15, 2011.
2. Boţ, Radu Ioan; Wanka, Gert; Grad, Sorin-Mihai (2009). *Duality in Vector Optimization*. Springer. ISBN 978-3-642-02885-4.
3. Csetnek, Ernő Robert (2010). *Overcoming the failure of the classical generalized interior-point regularity conditions in convex optimization. Applications of the duality theory to enlargements of maximal monotone operators*. Logos Verlag Berlin GmbH. ISBN 978-3-8325-2503-3.
4. Zălinescu, Constantin (2002). *Convex analysis in general vector spaces* (https://archive.org/details/s/convexanalysisge00zali_934). River Edge, NJ: World Scientific Publishing Co., Inc. pp. 106 (https://archive.org/details/s/convexanalysisge00zali_934/page/n126)–113. ISBN 981-238-067-1. MR 1921556 (<https://www.ams.org/mathscinet-getitem?mr=1921556>).
5. Borwein, Jonathan; Zhu, Qiji (2005). *Techniques of Variational Analysis*. Springer. ISBN 978-1-4419-2026-3.
6. Ahuja, Ravindra K.; Magnanti, Thomas L.; Orlin, James B. (1993). *Network Flows: Theory, Algorithms and Applications*. Prentice Hall. ISBN 0-13-617549-X.
7. Bertsekas, Dimitri; Nedic, Angelia; Ozdaglar, Asuman (2003). *Convex Analysis and Optimization*. Athena Scientific. ISBN 1-886529-45-0.
8. Bertsekas, Dimitri P. (1999). *Nonlinear Programming* (2nd ed.). Athena Scientific. ISBN 1-886529-00-0.
9. Bertsekas, Dimitri P. (2009). *Convex Optimization Theory*. Athena Scientific. ISBN 978-1-886529-31-1.
10. Bonnans, J. Frédéric; Gilbert, J. Charles; Lemaréchal, Claude; Sagastizábal, Claudia A. (2006). *Numerical optimization: Theoretical and practical aspects* (<https://www.springer.com/mathematics/applications/book/978-3-540-35445-1>). Universitext (Second revised ed. of translation of 1997 French ed.). Berlin: Springer-Verlag. pp. xiv+490. doi:10.1007/978-3-540-35447-5 (<https://doi.org/10.1007%2F978-3-540-35447-5>). ISBN 3-540-35445-X. MR 2265882 (<https://www.ams.org/mathscinet-getitem?mr=2265882>).
11. 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 3-540-56850-6. MR 1261420 (<https://www.ams.org/mathscinet-getitem?mr=1261420>).
12. Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993). "14 Duality for Practitioners". *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 3-540-56852-2. MR 1295240 (<https://www.ams.org/mathscinet-getitem?mr=1295240>).

13. Lasdon, Leon S. (2002) [Reprint of the 1970 Macmillan]. *Optimization theory for large systems*. Mineola, New York: Dover Publications, Inc. pp. xiii+523. ISBN 978-0-486-41999-2. MR 1888251 (<https://www.ams.org/mathscinet-getitem?mr=1888251>).
14. Lemaréchal, Claude (2001). "Lagrangian relaxation". In Jünger, Michael; Naddef, Denis (eds.). *Computational combinatorial optimization: Papers from the Spring School held in Schloß Dagstuhl, May 15–19, 2000*. Lecture Notes in Computer Science (LNCS). 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 3-540-42877-1. MR 1900016 (<https://www.ams.org/mathscinet-getitem?mr=1900016>).
15. Minoux, Michel (1986). *Mathematical programming: Theory and algorithms*. Egon Balas (forward); Steven Vajda (trans) from the (1983 Paris: Dunod) French. Chichester: A Wiley-Interscience Publication. John Wiley & Sons, Ltd. pp. xxviii+489. ISBN 0-471-90170-9. MR 0868279 (<https://www.ams.org/mathscinet-getitem?mr=0868279>). (2008 Second ed., in French: *Programmation mathématique : Théorie et algorithmes*, Éditions Tec & Doc, Paris, 2008. xxx+711 pp.).
16. Shapiro, Jeremy F. (1979). *Mathematical programming: Structures and algorithms* (<https://archive.org/details/mathematicalprog0000shap/page/>). New York: Wiley-Interscience [John Wiley & Sons]. pp. xvi+388 (<https://archive.org/details/mathematicalprog0000shap/page/>). ISBN 0-471-77886-9. MR 0544669 (<https://www.ams.org/mathscinet-getitem?mr=0544669>).
17. Geoffrion, Arthur M. (1971). "Duality in Nonlinear Programming: A Simplified Applications-Oriented Development". *SIAM Review*. **13** (1): 1–37. doi:10.1137/1013001 (<https://doi.org/10.1137%2F1013001>). JSTOR 2028848 (<https://www.jstor.org/stable/2028848>).

References

Books

- Ahuja, Ravindra K.; Magnanti, Thomas L.; Orlin, James B. (1993). *Network Flows: Theory, Algorithms and Applications*. Prentice Hall. ISBN 0-13-617549-X.
- Bertsekas, Dimitri; Nedic, Angelia; Ozdaglar, Asuman (2003). *Convex Analysis and Optimization*. Athena Scientific. ISBN 1-886529-45-0.
- Bertsekas, Dimitri P. (1999). *Nonlinear Programming* (2nd ed.). Athena Scientific. ISBN 1-886529-00-0.
- Bertsekas, Dimitri P. (2009). *Convex Optimization Theory*. Athena Scientific. ISBN 978-1-886529-31-1.
- Bonnans, J. Frédéric; Gilbert, J. Charles; Lemaréchal, Claude; Sagastizábal, Claudia A. (2006). *Numerical optimization: Theoretical and practical aspects* (<https://www.springer.com/mathematics/applications/book/978-3-540-35445-1>). Universitext (Second revised ed. of translation of 1997 French ed.). Berlin: Springer-Verlag. pp. xiv+490. doi:10.1007/978-3-540-35447-5 (<https://doi.org/10.1007%2F978-3-540-35447-5>). ISBN 3-540-35445-X. MR 2265882 (<https://www.ams.org/mathscinet-getitem?mr=2265882>).
- Cook, William J.; Cunningham, William H.; Pulleyblank, William R.; Schrijver, Alexander (November 12, 1997). *Combinatorial Optimization* (1st ed.). John Wiley & Sons. ISBN 0-471-55894-X.
- Dantzig, George B. (1963). *Linear Programming and Extensions* (<https://archive.org/details/linearpogrammin00dant>). Princeton, NJ: Princeton University Press.
- 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 3-540-56850-6. MR 1261420 (<https://www.ams.org/mathscinet-getitem?mr=1261420>).
- Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993). "14 Duality for Practitioners". *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 3-540-56852-2. MR 1295240 (<https://www.ams.org/mathscinet-getitem?mr=1295240>).
 - Lasdon, Leon S. (2002) [Reprint of the 1970 Macmillan]. *Optimization theory for large systems*. Mineola, New York: Dover Publications, Inc. pp. xiii+523. ISBN 978-0-486-41999-2. MR 1888251 (<https://www.ams.org/mathscinet-getitem?mr=1888251>).
 - Lawler, Eugene (2001). "4.5. Combinatorial Implications of Max-Flow Min-Cut Theorem, 4.6. Linear Programming Interpretation of Max-Flow Min-Cut Theorem". *Combinatorial Optimization: Networks and Matroids*. Dover. pp. 117–120. ISBN 0-486-41453-1.
 - Lemaréchal, Claude (2001). "Lagrangian relaxation". In Jünger, Michael; Naddef, Denis (eds.). *Computational combinatorial optimization: Papers from the Spring School held in Schloß Dagstuhl, May 15–19, 2000*. Lecture Notes in Computer Science (LNCS). Vol. 2241. Berlin: Springer-Verlag. pp. 112–156. doi:10.1007/3-540-45586-8_4 (https://doi.org/10.1007/3-540-45586-8_4). ISBN 3-540-42877-1. MR 1900016 (<https://www.ams.org/mathscinet-getitem?mr=1900016>).
 - Minoux, Michel (1986). *Mathematical programming: Theory and algorithms*. Egon Balas (forward); Steven Vajda (trans) from the (1983 Paris: Dunod) French. Chichester: A Wiley-Interscience Publication. John Wiley & Sons, Ltd. pp. xxviii+489. ISBN 0-471-90170-9. MR 0868279 (<https://www.ams.org/mathscinet-getitem?mr=0868279>). (2008 Second ed., in French: *Programmation mathématique : Théorie et algorithmes*, Éditions Tec & Doc, Paris, 2008. xxx+711 pp.).
 - Nering, Evar D.; Tucker, Albert W. (1993). *Linear Programming and Related Problems* (<https://archive.org/details/linearprogramsre0000neri>). Boston, MA: Academic Press. ISBN 978-0-12-515440-6.
 - Papadimitriou, Christos H.; Steiglitz, Kenneth (July 1998). *Combinatorial Optimization : Algorithms and Complexity* (Unabridged ed.). Dover. ISBN 0-486-40258-4.
 - Ruszczyński, Andrzej (2006). *Nonlinear Optimization*. Princeton, NJ: Princeton University Press. pp. xii+454. ISBN 978-0691119151. MR 2199043 (<https://www.ams.org/mathscinet-getitem?mr=2199043>).

Articles

- Everett, Hugh, III (1963). "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources" (<https://web.archive.org/web/20110724151508/http://or.journal.informs.org/cgi/reprint/11/3/399>). *Operations Research*. **11** (3): 399–417. doi:10.1287/opre.11.3.399 (<https://doi.org/10.1287/opre.11.3.399>). JSTOR 168028 (<https://www.jstor.org/stable/168028>). MR 0152360 (<https://www.ams.org/mathscinet-getitem?mr=0152360>). Archived from the original (<http://or.journal.informs.org/cgi/reprint/11/3/399>) on 2011-07-24.
- Kiwiel, Krzysztof C.; Larsson, Torbjörn; Lindberg, P. O. (August 2007). "Lagrangian relaxation via ballstep subgradient methods" (<http://mor.journal.informs.org/cgi/content/abstract/32/3/669>). *Mathematics of Operations Research*. **32** (3): 669–686. doi:10.1287/moor.1070.0261 (<https://doi.org/10.1287/moor.1070.0261>). MR 2348241 (<https://www.ams.org/mathscinet-getitem?mr=2348241>).
- Duality in Linear Programming (<http://www.civilized.com/files/duality.pdf>) Gary D. Knott

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Duality_\(optimization\)&oldid=1071877415](https://en.wikipedia.org/w/index.php?title=Duality_(optimization)&oldid=1071877415)"

This page was last edited on 14 February 2022, at 20:29 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.