Connections Between Duality in Control Theory and Convex Optimization

V. Balakrishnan[1] and L. Vandenberghe[2]

Abstract

Several important problems in control theory can be reformulated as convex optimization problems. From duality theory in convex optimization, dual problems can be derived for these convex optimization problems. These dual problems can in turn be reinterpreted in control or system theoretic terms, often yielding new results or new proofs for existing results from control theory. Moreover, the most efficient algorithms for convex optimization solve the primal and dual problems simultaneously. Insight into the system-theoretic meaning of the dual problem can therefore be very helpful in developing efficient algorithms. We demonstrate these observations with some examples.

1. Introduction

Over the past few years, convex optimization has come to be recognized as a valuable tool for control system analysis and design via numerical methods. Convex optimization problems enjoy a number of advantages over more general optimization problems: Every stationary point is also a global minimizer; they can be solved in polynomial-time; we can immediately write down necessary and sufficient optimality conditions; and there is a well-developed duality theory.

From a practical standpoint, there are effective and powerful algorithms for the solution of convex optimization problems, that is, algorithms that rapidly compute the global optimum, with non-heuristic stopping criteria. These algorithms range from simple descent-type or quasi-Newton methods for smooth problems to sophisticated cutting-plane or interior-point methods for non-smooth problems. A comprehensive literature is available on algorithms for convex programming; see for example, [1] and [2]; see also [3].

[1]School of Electrical Engineering, Purdue University, West Lafayette, IN 47907-1285, `ragu@ecn.purdue.edu`.

[2]K. U. Leuven, Electrical Engineering Department, K. Mercierlaan 94, 3001 Leuven, Belgium, `Lieven.Vandenberghe@esat.kuleuven.ac.be`. Research supported by the NFWO, and IUAP projects 17 and 50.

A typical application of convex optimization to a problem from control theory proceeds as follows: The control problem is reformulated (or in many cases, approximately reformulated) as a convex optimization problem, which is then solved using convex programming methods. Once the control problem is reformulated into a convex optimization problem, then it is straightforward to write down the *dual* convex optimization problem. This dual problem can be often be reinterpreted in control-theoretic terms, yielding new insight. The control-theoretic interpretation of the dual problem in turn helps in the efficient (numerical) implementation of primal-dual algorithms, which are among the most efficient techniques known for solving convex optimization problems. In this paper, we illustrate each of these points. First, we examine the standard LQR problem from control theory, and show how convex duality provides insight into its solution. We then discuss the implementation of primal-dual algorithms for another important control problem, namely the Linear Quadratic Regulator problem for linear time-varying systems with state-space parameters that lie in a polytope.

2. Convex programming duality

In the sequel, we will be concerned with convex optimization problems involving *linear matrix inequalities* or LMIs. These optimization problems have the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F(x) > 0 \end{array} \qquad (1)$$

where

$$F(x) \triangleq F_0 + x_1 F_1 + \cdots + x_m F_m.$$

The problem data are the vector $c = \mathbb{R}^m$ and $m+1$ symmetric matrices $F_0, F_1, \ldots, F_m \in \mathbb{R}^{n\times n}$. The inequality $F(x) > 0$ means that $F(x)$ is positive definite. We call problem (1) a *semidefinite program.*[3] Semidefinite programs can be solved efficiently using recently developed interior-point methods (see [2, 4]). The book [5] lists a large number of problems in

[3]Strictly speaking, the term "semidefinite program" refers to problem (1) with the constraint $F(x) \geq 0$ instead of $F(x) > 0$.

control and system theory that can be reduced to a semidefinite program.

As a consequence of convexity, we have a complete duality theory for semidefinite programs. For the special form of problem (1) duality reduces to the following. With every problem (1) we associate a *dual problem*

$$\begin{array}{ll} \text{maximize} & -\mathbf{Tr}\, F_0 Z \\ \text{subject to} & Z > 0 \\ & \mathbf{Tr}\, F_i Z = c_i,\ i = 1, \ldots, m. \end{array} \tag{2}$$

Here $\mathbf{Tr}\, X$ denotes the trace of a matrix X. The variable in (2) is the matrix $Z = Z^T \in \mathbb{R}^{n\times n}$. We have the following properties.

- If a matrix Z is dual feasible, i.e., $Z > 0$ and $\mathbf{Tr}\, F_i Z = c_i$, $i = 1, \ldots, m$, then the dual objective $-\mathbf{Tr}\, F_0 Z$ is a lower bound for the optimal value of (1):
$$-\mathbf{Tr}\, F_0 Z \le \inf \{c^T x | F(x) > 0\}.$$
- If an $x \in \mathbb{R}^m$ is primal feasible, i.e., $F(x) > 0$, then the primal objective $c^T x$ is an upper bound for the optimal value of (2):
$$c^T x \ge \sup \left\{ -\mathbf{Tr}\, F_0 Z \;\middle|\; \begin{array}{l} Z = Z^T > 0, \\ \mathbf{Tr}\, F_i Z = c_i, \\ i = 1, \ldots, m \end{array} \right\}.$$
- Under mild conditions, the optimal values of the primal problem (1) and its dual (2) are equal.

3. Primal-dual algorithms

Primal-dual algorithms are a class of iterative numerical algorithms for solving semidefinite programs. These algorithms solve problems (1) and (2) simultaneously; as they proceed, they generate a sequence of primal and dual feasible points $x^{(k)}$ and $Z^{(k)}$ ($k = 0, 1, \ldots$ denotes iteration number). This means that for every k, we have an upper bound $c^T x^{(k)}$ and a lower bound $-\mathbf{Tr}\, F_0 Z^{(k)}$ on the optimal value of problem (1).

General primal-dual interior-point methods that solve semidefinite programs are often more efficient than methods that work on the primal problem only. Their worst-case complexity is typically lower, and they are often faster in practice as well.

An important class of interior-point methods is based on the primal-dual potential function

$$\phi(x, Z) = (n + \nu\sqrt{n}) \log(c^T x + \mathbf{Tr}\, F_0 Z) - \log\det F(x) Z.$$

($\nu \ge 1$ is fixed.) If a method decreases this function by at least a fixed amount, independent of the problem size, in every iteration, then it can be shown that the number of iterations grows at most as $O(\sqrt{n})$ with the problem size. In practice the number of iterations appears to grow slower with n. Moreover the amount of work per iteration can be reduced considerably by taking advantage of the structure in the equations (see [6]).

An outline of a potential-reduction method due to Nesterov and Todd [7]—this is the algorithm used for solving the semidefinite programs that occur in this paper—is as follows. The method starts at strictly feasible x and Z. Each iteration consists of the following steps.

1. Compute a matrix R that simultaneously diagonalizes $F(x)^{-1}$ and Z:
$$R^T F(x)^{-1} R = \Lambda^{-1/2},\ R^T Z R = \Lambda^{1/2}.$$
The matrix Λ is diagonal, with as diagonal elements the eigenvalues of $F(x)Z$.
2. Compute $\delta x \in \mathbb{R}^m$ and $\delta Z = \delta Z^T \in \mathbb{R}^{n\times n}$ from
$$RR^T \delta Z RR^T + \textstyle\sum_{i=1}^m \delta x_i F_i = -\rho F(x) + Z^{-1}$$
$$\mathbf{Tr}\, F_j \delta Z = 0,\ \ j = 1, \ldots, m$$
with $\rho = (n + \nu\sqrt{n})/(c^T x + \mathbf{Tr}\, F_0 Z)$.
3. Find $p, q \in \mathbb{R}$ that minimize $\phi(x + p\delta x, Z + q\delta Z)$ and update $x := x + p\delta x$ and $Z := Z + q\delta Z$.

For details, we refer the reader to [7]; see also [4].

4. Convex duality and control theory

We first consider the standard Linear Quadratic Regulator problem, and show how convex duality described in §2 can be used to reinterpret the standard LQR solution. We then consider a multi-model (or "robust") version of the LQR problem, and describe an application of the primal-dual algorithm of §3 for computing bounds for this problelm.

4.1. The Linear Quadratic regulator

Consider the following optimal control problem: For the system

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \tag{3}$$

find u that minimizes

$$J = \int_0^\infty \left(x(t)^T Q x(t) + u(t)^T R u(t) \right)\, dt, \tag{4}$$

with $Q \ge 0$ and $R > 0$, subject to $\lim_{t\to\infty} x(t) = 0$. We assume the pair (A, B) is controllable. Let $J_{\rm opt}$ denote the minimum value.

Lower bound via quadratic functions
We can write down a lower bound for $J_{\rm opt}$ using quadratic functions; the following is essentially from [8, Theorem 2].

Suppose the quadratic function $\psi^T P\psi$ with $P > 0$ satisfies

$$\frac{d}{dt}x(t)^T Px(t) > -\left(x(t)^T Qx(t) + u(t)^T Ru(t)\right), \quad (5)$$

for all $t \geq 0$, and for all x and u satisfying $\dot{x} = Ax + Bu$, $x(T) = 0$. Then, integrating both sides from 0 to T, we get

$$x_0^T Px_0 < \int_0^T \left(x(t)^T Qx(t) + u(t)^T Ru(t)\right)\, dt,$$

or we have a lower bound for $J_{\rm opt}$.

Condition (5) holds for *all* x and u (not necessarily those that steer state to zero) if the Linear Matrix Inequality

$$\begin{bmatrix} A^T P + PA + Q & PB \\ B^T P & R \end{bmatrix} > 0 \quad (6)$$

is satisfied. Thus, the problem of computing the best lower bound using quadratic functions is

$$\begin{array}{ll} \text{maximize:} & x_0^T Px_0 \\ \text{subject to:} & P > 0, \ \ (6) \end{array} \quad (7)$$

The optimization variable in problem (7) is the symmetric matrix P.

Upper bound with state-feedback
Consider system (3) with a constant, linear state-feedback $u = Kx$ that stabilizes the system:

$$\dot{x} = (A + BK)x, \quad x(0) = x_0, \quad (8)$$

with $A + BK$ stable. Then the LQR objective J reduces to

$$J_K = \int_0^\infty x(t)^T \left(Q + K^T RK\right) x(t)\, dt.$$

Clearly, for every K, J_K yields an upper bound on the optimum LQR objective $J_{\rm opt}$. From standard results in control theory, J_K can be evaluated as

$$\mathbf{Tr}\, Z(Q + K^T RK),$$

where Z satisfies

$$(A + BK)Z + Z(A + BK)^T + x_0 x_0^T = 0,$$

with $A + BK$ stable.

It will be useful for us to rewrite this expression for J_K as

$$\inf_{Z>0} \mathbf{Tr}\, Z(Q + K^T RK),$$

where Z satisfies

$$(A + BK)Z + Z(A + BK)^T + x_0 x_0^T < 0. \quad (9)$$

Thus, the best upper bound on $J_{\rm opt}$, achievable using state-feedback control, is given by the optimization problem with the optimization variables Z and K:

$$\begin{array}{ll} \text{minimize:} & \mathbf{Tr}\, Z(Q + K^T RK) \\ \text{subject to:} & Z > 0, \ \ (9) \end{array} \quad (10)$$

Duality
We observe the following:

Problems (7) and (10) are duals of each other.

Proof: The proof is by direct verification. For problem (7), the dual problem is given by

$$\text{minimize} \ \ \mathbf{Tr} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} Z, \quad \text{over } Z,$$

subject to

$$Z = Z^T = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^T & Z_{22} \end{bmatrix} > 0$$

$$[A \quad B]\, Z \begin{bmatrix} I \\ 0 \end{bmatrix} + [I \quad 0]\, Z \begin{bmatrix} A^T \\ B^T \end{bmatrix} + x_0 x_0^T < 0.$$

With the change of variables $K = Z_{12}^T Z_{11}^{-1}$, and some standard arguments, we get the equivalent problem with variables Z_{11} and K:

$$\begin{array}{ll} \text{minimize:} & \mathbf{Tr}(Q + K^T RK)Z_{11}, \\ \text{subject to:} & Z_{11} > 0, \\ & (A + BK)Z_{11} + Z_{11}(A + BK)^T \\ & \qquad +x_0 x_0^T < 0, \end{array} \quad (11)$$

which is the same as problem (10).

□

Note that this shows that the optimal solution to the LQR problem is a linear state-feedback.

4.2. The multi-model LQR problem
Let us now consider a mutlti-model version of the LQR problem. We consider the multi-model or polytopic system (see [5])

$$\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t), \quad x(0) = x_0, \quad (12)$$

where for every time t,

$$[A(t) \ \ B(t)] \in \Omega \triangleq \mathbf{Co}\left\{[A_1 \ \ B_1], \ldots, [A_L \ \ B_L]\right\}. \quad (13)$$

Our objective now is to find u that minimizes

$$\sup_{A(\cdot),B(\cdot)\in\Omega} \int_0^\infty \left(x(t)^T Qx(t) + u(t)^T Ru(t)\right)\, dt,$$

with $Q \geq 0$ and $R > 0$, subject to $\lim_{t\to\infty} x(t) = 0$. Let $J_{\rm opt}$ denote the minimum value.

Lower bound via quadratic functions

We now repeat the steps of the previous section to write down a lower bound for $J_{\rm opt}$ using quadratic functions. Suppose the quadratic function $\psi^T P\psi$ with $P>0$ satisfies

$$\frac{d}{dt}x(t)^T Px(t) > -\left(x(t)^T Qx(t) + u(t)^T Ru(t)\right), \tag{14}$$

for all $t\geq 0$, and for all x and u satisfying (12) with $x(T)=0$. Then, integrating both sides from 0 to T, we get

$$x_0^T Px_0 < \int_0^T \left(x(t)^T Qx(t) + u(t)^T Ru(t)\right)\, dt,$$

or we have a lower bound for $J_{\rm opt}$.

Condition (14) holds for all x and u if the inequality

$$\left[\begin{array}{cc} A(t)^T P + PA(t) + Q & PB(t) \\ B(t)^T P & R \end{array}\right] > 0$$

holds for all $t\geq 0$, which in turn is equivalent to

$$\left[\begin{array}{cc} A_i^T P + PA_i + Q & PB_i \\ B_i^T P & R \end{array}\right] > 0,\ i=1,\ldots,L. \tag{15}$$

Thus, the problem of computing the best lower bound via quadratic functions is

$$\begin{array}{ll} \text{maximize:} & x_0^T Px_0 \\ \text{subject to:} & P>0,\ \ (15) \end{array} \tag{16}$$

The optimization variable in problem (16) is the symmetric matrix P.

The dual of problem (16) is

$$\text{minimize}\ \sum_{i=1}^L \mathbf{Tr} \left[\begin{array}{cc} Q & 0 \\ 0 & R \end{array}\right] Z_i,$$

over Z_i, $i=1,\ldots,L$, subject to

$$Z_i = Z_i^T = \left[\begin{array}{cc} Z_{i,11} & Z_{i,12} \\ Z_{i,12}^T & Z_{i,22} \end{array}\right] > 0$$

$$\sum_{i=1}^L \left([A_i \quad B_i]\, Z_i \left[\begin{array}{c} I \\ 0 \end{array}\right] + [I \quad 0]\, Z_i \left[\begin{array}{c} A_i^T \\ B_i^T \end{array}\right]\right) + x_0x_0^T < 0.$$

With the change of variables $K_i = Z_{i,12}^T Z_{i,11}^{-1}$, and some standard arguments, we get the equivalent problem with variables $Z_{i,11}$ and K_i:

$$\begin{array}{ll} \text{minimize:} & \mathbf{Tr} \sum_{i=1}^L \left((Q+K_i^T RK_i)Z_{i,11}\right), \\ \text{subject to:} & Z_{i,11} > 0, \\ & \sum_{i=1}^L \left((A_i+B_iK_i)Z_{i,11} \right. \\ & \left. +Z_{i,11}(A_i+B_iK_i)^T\right) + x_0x_0^T < 0. \end{array} \tag{17}$$

We are not aware of a nice control-theoretic interpretation of the dual problem at the time of writing of this paper.

It is easy to calculate feasible points for (17) by solving an LQR problem (recall that this is important for the application of the primal-dual algorithm of §3). Select any system $[\bar{A},\bar{B}]$ from the convex hull (13), i.e., choose

$$\bar{A} = \sum_{i=1}^L \lambda_i A_i, \quad \bar{B} = \sum_{i=1}^L \lambda_i B_i$$

for some $\lambda_i\geq 0$, $i=1,\ldots,L$, $\sum_{i=1}^L \lambda_i = 1$. By solving the LQR problem with $\bar{A}$, $\bar{B}$, we obtain matrices $\bar{Z}_{11}>0$ and $\bar{K}$ that satisfy

$$(\bar{A}+\bar{B}\bar{K})\bar{Z}_{11} + \bar{Z}_{11}(\bar{A}+\bar{B}\bar{K})^T + x_0x_0^T < 0. \tag{18}$$

From this it is clear that $Z_{i,11} = \lambda_i\bar{Z}_{11}$, $K_i=\bar{K}$, are feasible solutions in (17). Those dual solutions can be used as starting points for a primal-dual algorithm.

Upper bound with state-feedback

Restricting u to be a constant, linear state-feedback yields an upper bound on $J_{\rm opt}$. With $u=Kx$, the equations governing system (12) are

$$\frac{d}{dt}x(t) = (A(t)+B(t)K)\,x(t), \quad x(0)=x_0, \tag{19}$$

with the matrices A and B satisfying (13). Then the LQR objective J reduces to

$$J_K = \sup_{A(\cdot),B(\cdot)} \int_0^\infty x(t)^T \left(Q+K^T RK\right) x(t)\, dt.$$

Once again, for every K, J_K yields an upper bound on the optimum LQR objective $J_{\rm opt}$. Unlike with the LQR problem however, J_K is not easy to compute. We therefore present a simple upper bound for J_K using quadratic functions.

Suppose the quadratic function $\psi^T P\psi$ with $P>0$ satisfies

$$\frac{d}{dt}x(t)^T Px(t) < -x(t)^T \left(Q+K^T RK\right) x(t), \tag{20}$$

for all $t\geq 0$, and for all x and u satisfying (12) with $x(T)=0$. Then, integrating both sides from 0 to T, we get

$$x_0^T Px_0 > \int_0^T x(t)^T \left(Q+K^T RK\right) x(t)\, dt,$$

or we have an upper bound for $J_{\rm opt}$.

Condition (20) holds for *all* x and u (not necessarily those that steer state to zero) if the inequality

$$(A(t)+B(t)K)^T P + P(A(t)+B(t)K) + Q + K^T RK < 0$$

holds for all $t \geq 0$, which in turn is equivalent to

$$P^{-1}(A_i + B_i K)^T + (A_i + B_i K) P^{-1} + P^{-1}(Q + K^T R K) P^{-1} < 0, \; i = 1, \ldots, L.$$

With the change of variables $W = P^{-1}$ and $Y = KP^{-1}$, we get the matrix inequality (which can be written as an LMI using Schur complements)

$$\begin{array}{l} WA_i^T + A_i W + B_i Y + Y^T B_i^T \\ WQW + Y^T R Y < 0, \; i = 1, \ldots, L. \end{array} \quad (21)$$

Thus the best upper bound on J_{opt} using constant state-feedback and quadratic functions can be obtained by solving the semidefinite program with variables $W = W^T$ and Y:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, x_0^T W^{-1} x_0 \\ \text{subject to} & W > 0, \;\; (21) \end{array} \quad (22)$$

The dual problem is

$$\begin{array}{l} \text{maximize} \quad -\sum_{i=1}^{L} (\mathbf{Tr}\, Z_{i,22} + \mathbf{Tr}\, Z_{i,33}) - 2z^T x_0, \\ \text{subject to} : Z_i = Z_i^T = \left[\begin{array}{ccc} Z_{i,11} & Z_{i,12} & Z_{i,13} \\ Z_{i,12}^T & Z_{i,22} & Z_{i,23} \\ Z_{i,13}^T & Z_{i,23}^T & Z_{i,33} \end{array}\right] > 0 \\ \sum_{i=1}^{L} \left(Z_{i,11} A_i + A_i^T Z_{i,11} + Z_{i,12} Q^{1/2} + Q^{1/2} Z_{i,12}^T\right) > zz^T \\ \sum_{i=1}^{L} \left(B_i^T Z_{i,11} + R^{1/2} Z_{i,13}^T\right) = 0 \end{array} \quad (23)$$

where the variables are the L matrices Z_i and the vector z.

As with the lower bound, it is possible to obtain a dual feasible solution by solving an LQR problem. We omit the details here.

A Numerical Example

Figure 1 shows the results of a numerical example. The data are five matrices $A_i \in \mathbb{R}^{5\times5}$ and five matrices $B_i \in \mathbb{R}^{5\times3}$. The figures shows the objective values of the four semidefinite programs that we discussed above.

5. Conclusion

We have considered an optimal control problem with a quadratic objective, and have shown how we may obtain useful bounds for the optimal value using LMI-based convex optimization. Convex duality can be used to rederive the well-known LQR solution; control theory duality can be used to devise efficient primal-dual convex optimization algorithms.

The results presented herein are preliminary; it would be interesting to derive control-theoretic interpretations of the many primal-dual convex optimization problems presented here. Also of interest would be a careful study of the numerical advantages gained by using primal-dual algorithms over primal-only solvers. The primal-dual method outlined in this paper can be used to efficiently solve a large class of semidefinite programs from system and control theory, e.g., most of the ones presented in the book [5].

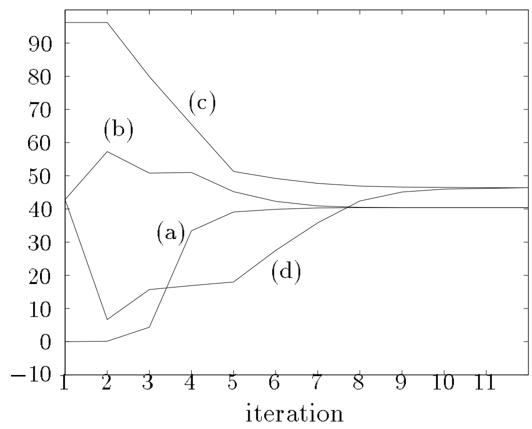

Figure 1: Upper and lower bounds versus iteration number. Curve (a) shows the lower bound (16) during execution of the primal-dual algorithm. Curve (b) shows the value of the associated dual problem (17). Curve (c) shows the upper bound (22) during execution of the primal-dual algorithm. Curve (d) shows the value of the associated dual problem (23).

References

[1] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*, volume 306 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 1993.

[2] Yu. Nesterov and A. Nemirovsky. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[3] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance*. Prentice-Hall, 1991.

[4] L. Vandenberghe and S. Boyd. Semidefinite programming. Submitted to *SIAM Review*, July 1994.

[5] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, June 1994.

[6] L. Vandenberghe and S. Boyd. Primal-dual potential reduction method for problems involving matrix inequalities. To be published in *Math. Programming*, 1995.

[7] Yu. E. Nesterov and M. J. Todd. Self-scaled cones and interior-point methods in nonlinear programming. Technical Report 1091, Cornell University, April 1994.

[8] J. C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Trans. Aut. Control*, AC-16(6):621–634, December 1971.