Background results for robust minmax control of linear dynamical systems*

James B. Rawlings, Davide Mannini, and Steven J. Kuntz †
Department of Chemical Engineering
University of California, Santa Barbara

June 25, 2024

The purpose of this note is to summarize the arguments required to derive the results appearing in robust minmax control of linear dynamical systems using a quadratic stage cost. The main results required in robust minmax control are Corollary 19 and Proposition 20. Moreover, the solution to the trust-region problem given in Proposition 15 and Lemma 16 may be of more general interest.

Linear algebra

We assume throughout that the parameters $D \in \mathbb{R}^{n\times n} \geq 0$, $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, d \in \mathbb{R}^n$ or $d \in \mathbb{R}^{n+m}$. Let $A^+ \in \mathbb{R}^{n\times m}$ denote the pseudoinverse of matrix $A \in \mathbb{R}^{m\times n}$. Let $N(A)$ and $R(A)$ denoted the null space and range space of matrix A, respectively. We will also make use of the singular value decomposition (SVD) of A given by

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \end{bmatrix} = U_1 \Sigma_r V_1' \tag{1}$$

and r is the rank of A. The properties of the SVD and the fundamental theorem of linear algebra imply that the orthonormal columns of U_1 and U_2 are bases for $R(A)$ and $N(A')$, respectively, and the orthonormal columns of V_1 and V_2 are bases for $R(A')$ and $N(A)$, respectively.[1] We also have that $A^+ = V_1 \Sigma_r^{-1} U_1'$.

First, we require solutions to linear algebra problems when such solutions exist.

Proposition 1 (Solving linear algebra problems.)**.** *Consider the linear algebra problem*

$$Ax = b$$

1. A solution exists if and only if $b \in R(A)$.

*An earlier version of these results was presented at the short course "Robust Nonlinear Model Predictive Control: Recent Advances in Design and Computation," University of California, Santa Barbara, CA, March 25–28, 2024.

†The authors gratefully acknowledge the financial support of the National Science Foundation (NSF) under Grant Nos. 2027091 and 2138985. jbraw@ucsb.edu, dmannini@ucsb.edu, skuntz@ucsb.edu.

[1]Edge cases: $A = 0$ has $r = 0$ and empty U_1, V_1, Σ_r matrices, so $U = U_2, V = V_2$ and $R(A) = \{0\}, R(A') = \{0\}, N(A) = \mathbb{R}^n, N(A') = \mathbb{R}^m$. At the other extreme, if A is square and invertible, $r = m = n$ and U_2, V_2 are empty so $U = U_1, V = V_1$, and $R(A) = \mathbb{R}^n, R(A') = \mathbb{R}^m, N(A) = \{0\}, N(A') = \{0\}$.

2. For $b \in R(A)$, the solution (set of solutions) is given by[2]

$$x^0 \in A^+ b + N(A) \tag{2}$$

Proof. By definition of range, if $b \notin R(A)$ there is no x such that $Ax = b$, and if $b \in R(A)$, there is an x such that $Ax = b$, which is the same as the existence condition. For $b \in R(A)$, let $z \in \mathbb{R}^n$ denote a value so that $Az = b$, and let q be an arbitrary element in $N(A)$ so $x^0 = A^+ b + q$. To show (2) are solutions, note that

$$Ax^0 = A(A^+ b + q) = AA^+ Az = Az = b$$

where we have used the definition of the null space and one of the pseudoinverse's defining properties, $AA^+A = A$. To show that (2) are all solutions, let x' denote a solution. We then have $A(x' - A^+ b) = b - b = 0$, so $x' - A^+ b \in N(A)$ or $x' \in A^+ b + N(A)$. Since x' is an arbitrary solution, (2) gives all solutions. □

Note that if one is interested in *deriving* (2) rather than establishing that it is correct as we did here, use the two orthogonal coordinate systems provided by the SVD of A, and let $x = V\alpha$, $b = U\beta$, and solve that simpler *decoupled* linear algebra problem for α^0 as a function of β, and convert back to x^0 in terms of b.

If $b \notin R(A)$, x^0 is still well-defined, but $Ax^0 - b = (AA^+ - I)b = -U_2 U_2' b \neq 0$. In this case, the x^0 given in (2) solves $\min_x |Ax - b|$ (least-squares solution), and achieves value $\left|Ax^0 - b\right| = |U_2' b|$.

Positive semidefinite matrices. We say that a matrix $M \in \mathbb{R}^{n \times n}$ is positive semidefinite, denoted $M \geq 0$, if M is symmetric and $x'Mx \geq 0$ for all $x \in \mathbb{R}^n$.

Optimization

We shall appeal without proof to one theorem for existence of solutions to optimization problems, the Weierstrass (extreme value) theorem. It says that a continuous function on a closed and bounded set attains its min and max on the set.[3] As we specialize to the results of interest in this note, next we consider convex, differentiable functions.

Definition 2 (Convex function)**.** A function $V : \mathbb{R}^n \to \mathbb{R}$ is convex if

$$V(\alpha u + (1 - \alpha)v) \leq \alpha V(u) + (1 - \alpha)V(v) \tag{3}$$

for all $u, v \in \mathbb{R}^n$ and $0 \leq \alpha \leq 1$.

If the function V is differentiable, then it is convex if and only if

$$V(v) \geq V(u) + (v - u)' \frac{dV}{du}(u) \tag{4}$$

for all $u, v \in \mathbb{R}^n$. See Boyd and Vandenberghe (2004, pp.69–70) for a proof of this fact.

An immediate consequence of this global lower bound is that u^0 is a minimizer of V if and only if $(dV/du)(u^0) = 0$.

[2] We overload the addition symbol to mean set addition when adding singletons ($A^+ b$) and sets ($N(A)$).

[3] Proofs for the multivariate version required here can be found in Mangasarian (1994, p. 198), Polak (1997, Corollary 5.1.25), Rockafellar and Wets (1998, p. 11), and Rawlings, Mayne, and Diehl (2020, Proposition A.7).

Proposition 3. *A convex, differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ has a minimizer u^0 if and only if $(dV/du)(u^0) = 0$.*

Proof. To establish sufficiency, assume $(dV/du)(u^0) = 0$; (4) then implies $V(v) \geq V(u^0)$ for all $v \in \mathbb{R}^n$, and therefore u^0 is the minimizer of V.

To establish necessity, assume that u^0 is optimal but that, contrary to what is to be proven, $(dV/du)(u^0) \neq 0$, and let $h = -(dV/du)(u^0)$ so that the directional derivative satisfies

$$\lim_{\lambda \to 0^+} \frac{V(u^0 + \lambda h) - V(u^0)}{\lambda} = h' \frac{dV}{du}(u^0) = -\left| \frac{dV}{du}(u^0) \right|^2$$

Given this limit, for every $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$\frac{V(u^0 + \lambda h) - V(u^0)}{\lambda} \leq -\left| \frac{dV}{du}(u^0) \right|^2 + \epsilon$$

for all $0 < \lambda \leq \delta$. Choose $\epsilon = (1/2)\left|(dV/du)(u^0)\right|^2 > 0$, and we have that

$$V(u^0 + \lambda h) \leq V(u^0) - (\lambda/2)\left| \frac{dV}{du}(u^0) \right|^2$$

for $0 < \lambda \leq \delta$. This inequality contradicts the optimality of u^0 and, therefore $(dV/du)(u^0) = 0$, which establishes necessity, and the proposition is proven. □

When considering robust control of linear dynamical systems with quadratic stage cost, quadratic functions play a central role. We have the following result about their convexity.

Proposition 4 (Convex quadratic functions)**.** *The quadratic function $V(u) = (1/2)u'Du + u'd + c$ is convex if and only if $D \geq 0$.*

Proof. We establish that the quadratic term $f(u) \coloneqq u'Du$ is convex by substituting $\alpha u + (1-\alpha)v$ into function f and rearranging the terms

$$f(\alpha u + (1-\alpha)v) - \big(\alpha f(u) + (1-\alpha)f(v)\big) = -\alpha(1-\alpha)(u-v)'D(u-v)$$

Since $-\alpha(1-\alpha) < 0$ for $\alpha \in (0,1)$, we have that the right-hand side is less than or equal to zero for every $u, v \in \mathbb{R}^n$ if and only if $D \geq 0$, verifying (3) for the function f.

It is then straightforward to show that the linear function $u'd$ and the constant function c are both convex by directly verifying (3). It is also straightforward to establish that linear combinations of convex functions are convex by verifying (3), and, therefore the function V is convex if and only if $D \geq 0$. □

The following optimization result for convex quadratic functions is then useful in the ensuing discussion.

Proposition 5 (Minimum of quadratic functions)**.** *Consider the quadratic function $V(\cdot) : \mathbb{R}^n \to \mathbb{R}$ with $D \in \mathbb{R}^{n \times n} \geq 0$*

$$V(u) \coloneqq (1/2)u'Du + u'd$$

1. *A solution to $\min_u V$ exists if and only if $d \in R(D)$.*

2. *For $d \in R(D)$, the minimizer and optimal value function are*

$$u^0 \in -D^+d + N(D) \qquad V^0 = -(1/2)d'D^+d \tag{5}$$

and $(d/du)V(u) = 0$ at u^0.

Proof. The function $V(u)$ is differentiable and convex (Proposition 4) so from Proposition 3, a solution exists if and only if the derivative is zero. Taking the derivative gives $(d/du)V(u) = Du + d$. From Proposition 1, $Du + d = 0$ has a solution if and only if $d \in R(D)$ and the set of all solutions is $d^0 = -D^+d + N(D)$, and evaluating V at the solution gives $V(u^0) = -(1/2)d'D^+d$ establishing (5). □

For maximization problems, we can replace $D \geq 0$ with $D \leq 0$ and min with max.

Partitioned semidefinite matrices. We make extensive use of partitioned matrices

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix}$$

We have the following result for positive semidefinite partitioned matrices (Boyd and Vandenberghe, 2004, p.651).

Proposition 6 (Positive semidefinite partitioned matrices)**.** *The matrix $M \geq 0$ if and only if $M_{11} \geq 0$, $M_{22} - M_{12}'M_{11}^+M_{12} \geq 0$, and $R(M_{12}) \subseteq R(M_{11})$.*

Proof.

1. Forward implication. Define $V(x,y) \coloneqq (1/2)(x,y)'M(x,y)$, and assume $M \geq 0$. Expanding V using the partitioned matrix

$$\begin{aligned} V(x,y) &= (1/2)\begin{bmatrix} x \\ y \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= (1/2)\big(x'M_{11}x + 2x'M_{12}y + y'M_{22}y\big) \\ &\geq 0, \quad \text{for all } (x,y) \end{aligned} \tag{6}$$

   Setting $y = 0$ in (6) implies that $M_{11} \geq 0$. Since $M_{11} \geq 0$, $V(x,y)$ is a differentiable, convex function of x for any y. Therefore $\min_x V(x,y)$ has a solution for every y, and Proposition 5 then implies $M_{12}y \in R(M_{11})$ for every y, which is equivalent to $R(M_{12}) \subseteq R(M_{11})$. Substituting the minimizer over x, $x^0 = -M_{11}^+M_{12}y$ into V gives

$$V(x^0,y) = (1/2)y'(M_{22} - M_{12}'M_{11}^+M_{12})y \tag{7}$$

   and since $V(x,y) \geq 0$ for all (x,y), we have that $M_{22} - M_{12}'M_{11}^+M_{12} \geq 0$, and the forward implication is established.

2. Reverse implication. Assume $M_{11} \geq 0$, $M_{22} - M_{12}'M_{11}^+M_{12} \geq 0$, and $R(M_{12}) \subseteq R(M_{11})$, and we establish that $M \geq 0$. For proof by contradiction, assume there exists an $(\overline{x},\overline{y})$ such that $(\overline{x},\overline{y})'M(\overline{x},\overline{y}) < 0$. By Proposition 5, we know that $\min_x V(x,\overline{y})$ exists since $M_{11} \geq 0$ and $M_{12}\overline{y} \in R(M_{11})$, and it has value $V^0 = V(x^0,\overline{y})$ with $x^0 = -M_{11}^+M_{12}\overline{y}$. Substituting this into V gives $V^0 = (1/2)\overline{y}'(M_{22} - M_{12}'M_{11}^+M_{12})\overline{y} \geq 0$ because matrix $M_{22} - M_{12}'M_{11}^+M_{12} \geq 0$. By optimality of x^0, $V(x,\overline{y}) \geq V^0 \geq 0$ for all x. But that contradicts $V(\overline{x},\overline{y}) < 0$, and we conclude $M \geq 0$, and the proof is complete. □

Note that

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \geq 0 \quad \text{if and only if} \quad \begin{bmatrix} M_{22} & M_{12}' \\ M_{12} & M_{11} \end{bmatrix} \geq 0$$

So we can also conclude that $M \geq 0$ if and only if $M_{11} \geq 0$, $M_{22} \geq 0$, $M_{22} - M_{12}' M_{11}^{+} M_{12} \geq 0$, $M_{11} - M_{12} M_{22}^{+} M_{12}' \geq 0$, $R(M_{12}) \subseteq R(M_{11})$, and $R(M_{12}') \subseteq R(M_{22})$. Note also that given the partitioning in M, we define

$$\begin{aligned} \tilde{M}_{11} &\coloneqq M_{22} - M_{12}' M_{11}^{+} M_{12} \\ \tilde{M}_{22} &\coloneqq M_{11} - M_{12} M_{22}^{+} M_{12}' \end{aligned} \tag{8}$$

and $\tilde{M}_{11}$ is known as the Schur complement of M_{11}, and $\tilde{M}_{22}$ is known as the Schur complement of M_{22}.

Constraints and Lagrangians. Next we require a standard optimization result for using a Lagrangian to reformulate a constrained minimization as an unconstrained min-max problem. The following result will be useful for this purpose. Let $U \subseteq \mathbb{R}^n$ be a nonempty compact set and $V(\cdot) : U \to \mathbb{R}$ be a continuous function on U. Define the Lagrangian function $L(\cdot) : U \times \mathbb{R} \to \mathbb{R}$ as

$$L(u, \lambda) = V(u) - \lambda \rho(u, U) \tag{9}$$

where $\rho(\cdot) : \mathbb{R}^n \times U \to \mathbb{R}_{\geq 0}$ is any convenient continuous indicator function that evaluates to zero if and only if $u \in U$. Denote a minmax problem as

$$\inf_u \sup_\lambda L(u, \lambda)$$

When a solution to this problem exists, we define the optimal value L^* and solution set u^* as

$$L^* = \min_u \max_\lambda L(u, \lambda) \qquad u^* = \arg\min_u \max_\lambda L(u, \lambda)$$

It is convenient in the subsequent development to define the maximizer of the inner problem

$$\overline{\lambda}(u) \coloneqq \arg\max_\lambda L(u, \lambda), \quad u \in U$$

Proposition 7 (Constrained minimization and Lagrangian minmax)**.** *Let $U \subseteq \mathbb{R}^n$ be a nonempty compact set and $V(\cdot) : U \to \mathbb{R}$ be a continuous function on U, and $L(\cdot) : U \times \mathbb{R} \to \mathbb{R}$ be defined as $L(u, \lambda) \coloneqq V(u) - \lambda \rho(u, U)$. Consider the constrained optimization problem*

$$\inf_{u \in U} V(u) \tag{10}$$

and the (unconstrained) Lagrangian minmax problem

$$\inf_u \sup_\lambda L(u, \lambda) \tag{11}$$

1. *Solutions to both problems exist.*
2. *Let V^0 be the solution and u^0 be the set of optimizers of $\min_{u \in U} V(u)$. Let L^* be the solution and u^* the set of optimizers of $\min_u \max_\lambda L(u, \lambda)$. Then*

$$V^0 = L^* \qquad u^0 = u^* \qquad \overline{\lambda}(u^*) = \mathbb{R}$$

Proof. The solution to (10) exists by the Weierstrass theorem. Denote the optimal value V^0 and solution set $u^0 \subseteq U$ which satisfy $V(u^0) = V^0$. We show that a solution to (11) also exists. Consider the inner supremum. From the definitions of functions L and ρ, we conclude

$$\sup_\lambda L(u,\lambda) = \begin{cases} V(u), & u \in U \\ +\infty, & u \notin U \end{cases}$$

Then consider the outer infimum. We have that

$$L^* = \inf_u \sup_\lambda L(u,\lambda) = \inf_{u\in U} V(u) = \min_{u\in U} V(u) = V^0$$

So the solution to (11) exists with value $L^* = V^0$. Taking the argument gives

$$u^* \coloneqq \arg\inf_u \sup_\lambda L(u,\lambda) = \arg\min_{u\in U} V(u) = u^0$$

For the inner problem evaluated at u^*, we note that

$$\sup_\lambda L(u^*,\lambda) = \sup_\lambda V(u^0) = \max_\lambda V^0 = V^0$$

Taking the argument then gives

$$\overline{\lambda}(u^*) = \arg\sup_\lambda L(u^*,\lambda) = \arg\max_\lambda V^0 = \mathbb{R}$$

and the result is established. □

Minmax and Maxmin

More generally, we are interested in a function $V(u,w)$ $V : U \times W \to \mathbb{R}$ and the optimization problems

$$\inf_{u\in U}\sup_{w\in W} V(u,w) \qquad \sup_{w\in W}\inf_{u\in U} V(u,w)$$

We assume in the following that the inf and sup are achieved on the respective sets and replace them with min and max.

Continuous functions. Let's start here. According to Wikipedia, von Neumann's minimax theorem states (von Neumann, 1928)

Theorem 8 (Minimax Theorem)**.** *Let $U \subset \mathbb{R}^m$ and $W \subset \mathbb{R}^n$ be compact convex sets. If $V : U \times W \to \mathbb{R}$ is a continuous function that is convex-concave, i.e., $V(\cdot, w) : U \to \mathbb{R}$ is convex for all $w \in W$, and $V(u,\cdot) : W \to \mathbb{R}$ is concave for all $u \in U$*
Then we have that

$$\min_{u\in U}\max_{w\in W} V(u,w) = \max_{w\in W}\min_{u\in U} V(u,w)$$

Note that existence of min and max is guaranteed by compactness of U, W (closed, bounded). Also note that the following holds for any continuous function V

$$\min_{u\in U}\max_{w\in W} V(u,w) \geq \max_{w\in W}\min_{u\in U} V(u,w)$$

This is often called *weak duality*. It's easy to establish. We are regarding the switching of the order of min and max as a form of duality. (Think of observability and controllability as duals of each other.)

So when this inequality achieves equality, that's often called *strong duality.* So the minimax theorem says that continuous functions that are convex-concave on compact sets satisfy strong duality. When strong duality is not achieved, we refer to the difference as the *duality gap*, which is positive due to weak duality

$$\min_{u\in U}\max_{w\in W} V(u,w) - \max_{w\in W}\min_{u\in U} V(u,w) > 0$$

Saddle Points. In characterizing solutions of these problems, it is useful to define a saddle point of the function $V(u,w)$.

Definition 9 (Saddle point)**.** The point (set) $(u^*,w^*)\subseteq U\times W$ is called a saddle point (set) of $V(\cdot)$ if

$$V(u^*,w)\leq V(u^*,w^*)\leq V(u,w^*) \quad \text{for all } u\in U, w\in W \tag{12}$$

Proposition 10 (Saddle-point theorem)**.** *The point (set) $(u^*,w^*)\subseteq U\times W$ is a saddle point (set) of function $V(\cdot)$ if and only if strong duality holds and (u^*,w^*) is a solution to the two problems*

$$\min_{u\in U}\max_{w\in W} V(u,w) = \max_{w\in W}\min_{u\in U} V(u,w) = V(u^*,w^*) \tag{13}$$

$$u^* = \arg\min_{u\in U}\max_{w\in W} V(u,w) \qquad w^* = \arg\max_{w\in W}\min_{u\in U} V(u,w) \tag{14}$$

In the following development it is convenient to define the solutions to the minimization and maximization problems

$$\overline{V}(u) \coloneqq \max_{w\in W} V(u,w), \quad u\in U \qquad \underline{V}(w) \coloneqq \min_{u\in U} V(u,w), \quad w\in W \tag{15}$$

Note that (14) implies that $\max_{w\in W}\underline{V}(w) = \underline{V}(w^*)$ and $\min_{u\in U}\overline{V}(u) = \overline{V}(u^*)$.

Remark 11. Note that (14) also implies that

$$\max_{w\in W}\min_{u\in U} V(u,w) = \min_{u\in U} V(u,w^*) \qquad \min_{u\in U}\max_{w\in W} V(u,w) = \max_{w\in W} V(u^*,w) \tag{16}$$

To establish this remark, note that

$$\max_{w\in W}\min_{u\in U} V(u,w) = \max_{w\in W}\underline{V}(w) = \underline{V}(w^*) = \min_{u\in U} V(u,w^*)$$

Similarly,

$$\min_{u\in U}\max_{w\in W} V(u,w) = \min_{u\in U}\overline{V}(u) = \overline{V}(u^*) = \max_{w\in W} V(u^*,w)$$

Next we prove Proposition 10

Proof. First we establish that (14) implies (12). Note that by optimality, the first equality in (16), which is a consequence of assuming (14), implies that $V(u^*,w^*)\leq V(u,w^*)$ for all $u\in U$, and the second implies that $V(u^*,w^*)\geq V(u^*,w)$ for all $w\in W$. Taken together, these are (12).

Next we show that (12) implies (14). We know that the following holds by weak duality

$$\max_{w\in W}\min_{u\in U} V(u,w) \leq \min_{u\in U}\max_{w\in W} V(u,w) \tag{17}$$

So we wish to show that the reverse inequality also holds to establish strong duality, i.e., the first equality in (14). To that end note that from (12)

$$V(u^*, w) \leq V(u, w^*) \quad \text{for all } w \in W, u \in U$$

Since this holds for all $w \in W$, it also holds for a maximizer, and therefore

$$\max_{w \in W} V(u^*, w) \leq V(u, w^*) \quad \text{for all } u \in U$$

The left-hand side will not be larger if instead of evaluating at $u = u^* \in U$, we minimize over all $u \in U$, giving

$$\min_{u \in U} \max_{w \in W} V(u, w) \leq V(u, w^*) \quad \text{for all } u \in U$$

Now if this inequality holds for all $u \in U$, it also holds for the minimizer on the right-hand side so that

$$\min_{u \in U} \max_{w \in W} V(u, w) \leq \min_{u \in U} V(u, w^*)$$

We can only increase the value of the right-hand side if instead of evaluating at $w = w^* \in W$, we maximize over all $w \in W$, giving

$$\min_{u \in U} \max_{w \in W} V(u, w) \leq \max_{w \in W} \min_{u \in U} V(u, w)$$

Note that this is the weak duality inequality (17) written in the reverse direction, so combining with weak duality, we have that

$$\min_{u \in U} \max_{w \in W} V(u, w) = \max_{w \in W} \min_{u \in U} V(u, w)$$

and strong duality is established.

We next show that u^* solves the minmax problem. From the defined optimizations in (21) we have that

$$\min_{u \in U} \max_{w \in W} V(u, w) = \min_{u \in U} \overline{V}(u) \tag{18}$$

$$\max_{w \in W} \min_{u \in U} V(u, w) = \max_{w \in W} \underline{V}(w) \tag{19}$$

Next choose an arbitraru $u_1 \in U$ and assume for contradiction that $\overline{V}(u_1) < \overline{V}(u^*)$. From the definition of $\overline{V}$ we then have have that

$$\max_{w \in W} V(u_1, w) < \max_{w \in W} V(u^*, w)$$

Therefore since $w^* \in W$

$$V(u_1, w^*) < \max_{w \in W} V(u^*, w)$$

But from the saddle-point condition, (12), $\max_{w \in W} V(u^*, w) \leq V(u, w^*)$ for all $u \in U$, which contradicts the previous inequality since $u_1 \in U$. Therefore $\overline{V}(u_1) \geq \overline{V}(u^*)$, and since u_1 is an arbitrary element of U, u^* solves the minmax problem (18).

Similarly we can show that w^* solves the maxmin problem (19) by exchanging the variables w and u and the operations max and min. Therefore (w^*, u^*) solves (14), and we have established that (12) implies (14). □

In the following development it is convenient to define the solutions to the inner minimization and maximization problems

$$\underline{u}^0(w) := \arg\min_{u\in U} V(u,w), \quad w \in W \tag{20}$$

$$\overline{w}^0(u) := \arg\max_{w\in W} V(u,w), \quad u \in U \tag{21}$$

Note that these inner solution sets are too "large" in the following sense. Even if we evaluate them at the optimizers of their respective outer problems, we know only that

$$u^* \subseteq \underline{u}^0(w^*) \qquad w^* \subseteq \overline{w}^0(u^*)$$

and these subsets may be strict. So we have to exercise some care when we exploit strong duality and want to extract the optimizer from a dual problem. We shall illustrate this issue in the upcoming results.

Quadratic functions. In control problems, we min and max over possibly unbounded sets, so we need something other than compactness to guarantee existence of solutions. When we have linear dynamic models and quadratic stage cost (LQ), we can use the following results for quadratic functions.

Proposition 12 (Saddle-point theorem for quadratic functions)**.** *Consider the quadratic function* $V(\cdot) : \mathbb{R}^{n+m} \to \mathbb{R}$

$$V(u,w) \coloneqq (1/2)\begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} + \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

with $M_{22} \in \mathbb{R}^{n\times n} \leq 0$, $M_{11} \in \mathbb{R}^{m\times m} \geq 0$, $M_{12} \in \mathbb{R}^{m\times n}, d \in \mathbb{R}^{n+m}$.

1. *A solution to* $\min_u \max_w V$ *exists if and only if* $d \in R(M)$. *Similarly, a solution to* $\max_w \min_u V$ *exists if and only if* $d \in R(M)$.

2. *For* $d \in R(M)$, *strong duality holds so that*

$$\min_u \max_w V(u,w) = \max_w \min_u V(u,w) = V(u^*,w^*)$$

where (u^*, w^*) *are saddle points of the function* V, *satisfying*

$$\begin{bmatrix} u^* \\ w^* \end{bmatrix} \in -M^+ d + N(M) \qquad V(u^*,w^*) = -(1/2)d'M^+d \tag{22}$$

and $dV(u,w)/d(u,w) = 0$ *at* (u^*, w^*).

3. *For* $d \in R(M)$, *let* $\underline{u}^0(w) \coloneqq \min_u V(u,w)$ *and* $\overline{w}^0(u) \coloneqq \max_w V(u,w)$. *The solution sets and saddle points satisfy the following relationships*

$$u^* = \arg\min_u \max_w V(u,w), \quad u^* \subseteq \underline{u}^0(w^*), \tag{23}$$

$$w^* = \arg\max_w \min_u V(u,w), \quad w^* \subseteq \overline{w}^0(u^*). \tag{24}$$

Proof. First we establish that (u^*, w^*) satisfy (22) by analyzing the $\min_u \max_w V$ problem. We assume $d \in R(M)$ and expand $V(\cdot)$ as

$$V(u,w) = (1/2)w'M_{22}w + w'(M_{12}'u + d_2) + (1/2)u'M_{11}u + u'd_1 \tag{25}$$

From Proposition 5, $\max_w V$ exists if and only if $M_{12}'u + d_2 \in R(M_{22})$. This condition is satisfied for some nonempty set of u by the bottom half of $d \in R(M)$. For such u we have the necessary and sufficient condition for the optimum

$$M_{22}\overline{w}^0 + M_{12}'u + d_2 = 0 \tag{26}$$

which defines an implicit function $\overline{w}^0(u)$, and optimal value given by (5)

$$\begin{aligned}\overline{w}^0(u) &= -M_{22}^+(M_{12}'u + d_2) + N(M_{22}) \\ V(u, \overline{w}^0(u)) &= (1/2)u'\tilde{M}_{22}u + u'(d_1 - M_{12}M_{22}^+d_2) - (1/2)d_2M_{22}^+d_2\end{aligned}$$

where $\tilde{M}_{22}$ is the Schur complement of M_{22} defined in (8). Note that $\tilde{M}_{22} \geq 0$ since $M_{11} \geq 0$ and $M_{22} \leq 0$, which implies $M_{22}^+ \leq 0$. However, we cannot simply set the derivative to zero because we require $M_{12}'u + d_2 \in R(M_{22})$ for the existence of $V(u, \overline{w}^0(u)) = \max_w V(u, w)$. To handle this range constraint, we use a linear equality constraint $M_{12}'u + d_2 = M_{22}y$ where y is a slack variable. Under the equality constraint, the problem $\min_u \max_w V$ is equivalent to the following constrained minimization:

$$\min_{u,y} V(u, \overline{w}^0(u)) \qquad \text{subject to} \qquad M_{12}'u + d_2 = M_{22}y.$$

Because $V(u, \overline{w}^0(u))$ is convex and differentiable in (u, y), its minimum subject to the affine constraint $M_{12}'u + d_2 = M_{22}y$ is achieved by the stationary points of the Lagrangian (Boyd and Vandenberghe, 2004, pp. 141–142):

$$L(u, y, \lambda) := V(u, \overline{w}^0(u)) + \lambda'(M_{12}'u + d_2 - M_{22}y).$$

Taking derivatives, we have (u^*, y^*, λ^*) is a stationary point if and only if

$$\tilde{M}_{22}u^* + d_1 - M_{12}M_{22}^+d_2 + M_{12}\lambda^* = 0 \tag{27}$$

$$-M_{22}\lambda^* = 0 \tag{28}$$

$$M_{12}'u^* + d_2 - M_{22}y^* = 0. \tag{29}$$

Substituting (29) into (27), we have

$$M_{11}u^* + M_{12}(\lambda^* - M_{22}^+M_{22}y^*) + d_1 = 0.$$

Next, (28) implies $\lambda^* \in N(M_{22})$, so we can rewrite (29) as

$$M_{12}'u^* + M_{22}(\lambda^* - M_{22}^+M_{22}y^*) + d_2 = 0.$$

With $w^* := \lambda^* - M_{22}^+M_{22}y^*$, we have the system (27) and (29) as

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} u^* \\ w^* \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \tag{30}$$

which has solutions since $d \in R(M)$, and moreover, any such solution gives $y^* = -M_{22}^+M_{22}w^*$ and $\lambda^* = w^* + y^*$ satisfying (27)–(29). In fact, $w^* \subseteq \overline{w}^0(u^*)$ is implied by (28) and (29). Finally, solving (30) and substituting the solution into $V(u, w)$ gives (22).

To solve the $\max_w \min_u V$ problem, take the negative of the objective to obtain $\max_w \min_u V = -\min_w \max_u(-V)$. Therefore the exact same procedure can be used here, and it produces the same solutions and optimal values (22), along with $u^* \subseteq \underline{u}^0(w^*)$.

Finally, note that if $d \notin R(M)$, we have no solution to $\max_w V(u, w)$ for any u, and therefore no solution to $\min_u \max_w V(u, w)$. Similarly we have no solution to $\min_u V(u, w)$ for any w, and therefore no solution to $\max_w \min_u V(u, w)$. We have thus established all the claims of the proposition. □

Applying Proposition 12 to the following example with $M_{11} = M_{22} = 0$ and $M_{12} = 1$

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad V(u, w) = uw + \begin{bmatrix} u \\ w \end{bmatrix}' d$$

gives $(u^*, w^*) = -(d_2, d_1)$, $V(u^*, w^*) = -d_1 d_2$, $\overline{w}^0(u^*) = \mathbb{R}$, $\underline{u}^0(w^*) = \mathbb{R}$. Note that both functions $\overline{w}^0(\cdot)$ and $\underline{u}^0(\cdot)$ are defined at only a single point, u^* and w^*, respectively. So in this degenerate case, these functions are not even differentiable.

Lagrangian functions. The connections between constrained optimization problems via the use of Lagrange multipliers and game theory problems are useful (Rockafellar, 1993).

> For optimization problems of convex type, Lagrange multipliers take on a game-theoretic role that could hardly even have been imagined before the creative insights of von Neumann [32], [33], in applying mathematics to models of social and economic conflict.
>
> –T.A. Rockafellar

Next we are interested in the Lagrangian function $L(\cdot) : \mathbb{R}^{n+m+1} \to \mathbb{R}$

$$\begin{aligned} L(u, w, \lambda) &\coloneqq (1/2) \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} - (1/2)\lambda(w'w - 1) \\ &= (1/2) \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} - \lambda I \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} + \lambda/2 \end{aligned}$$

with $M \in \mathbb{R}^{(n+m)\times(n+m)} \geq 0$, and $M_{11} \in \mathbb{R}^{m\times m}$, $M_{12} \in \mathbb{R}^{m\times n}$, $M_{22} \in \mathbb{R}^{n\times n}$. Note that from Proposition 6, both $M_{11} \geq 0$ and $M_{22} \geq 0$ as well, so that $\max_w$ is not bounded unless λ is large enough to make $M_{22} - \lambda I \leq 0$. The Schur complements of M_{11} and $M_{22} - \lambda I$ are useful for expressing the solution.

$$\begin{aligned} \tilde{M}_{11}(\lambda) &\coloneqq (M_{22} - \lambda I) - M_{12}' M_{11}^{+} M_{12} \\ \tilde{M}_{22}(\lambda) &\coloneqq M_{11} - M_{12}(M_{22} - \lambda I)^{+} M_{12}' \end{aligned}$$

Note that both Schur complements depend on the parameter λ.

Corollary 13 (Minmax and maxmin of a quadratic function with a parameter)**.** *Consider the quadratic function $L(\cdot) : \mathbb{R}^{n+m+1} \to \mathbb{R}$ expressed as*

$$L(u, w, \lambda) \coloneqq (1/2) \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} - \lambda I \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} + \lambda/2 \tag{31}$$

1. *A solution to $\min_u \max_w L$ exists if and only if*

$$\lambda \geq |M_{22}| \tag{32}$$

and the solution (set) and optimal value function are

$$\begin{aligned} w^0(u) &\in -(M_{22} - \lambda I)^{+} M_{12}'\, u + N(M_{22} - \lambda I) \\ u^0 &\in N(\tilde{M}_{22}(\lambda)) \\ L(u^0, w^0(u^0), \lambda) &= \begin{cases} \lambda/2, & \lambda \geq |M_{22}| \\ +\infty, & \lambda < |M_{22}| \end{cases} \end{aligned}$$

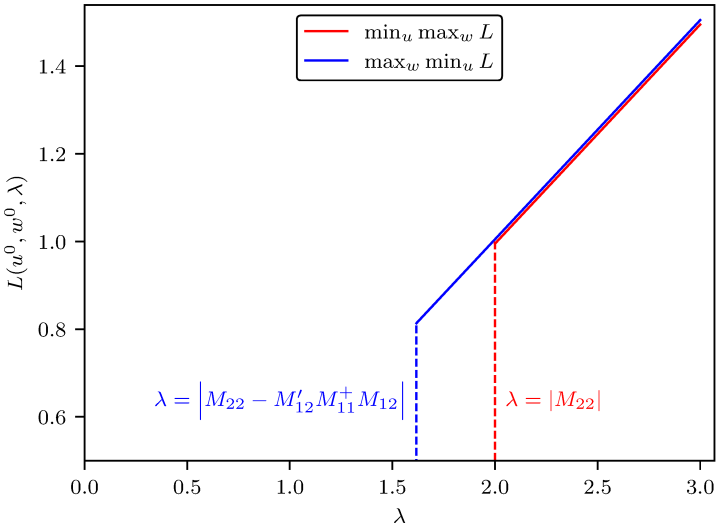

Figure 1: The optimal value function L^0 for $\min_u \max_w L$ and $\max_w \min_u L$ versus parameter λ. Strong duality holds only when $\lambda \geq |M_{22}|$. For $\lambda < |M_{22}|$, $\min_u \max_w L^0 = +\infty$. For $\lambda < \left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right|$, $\max_w \min_u L^0 = +\infty$.

2. *A solution to* $\max_w \min_u L$ *exists if and only if*

$$\lambda \geq \left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right| \tag{33}$$

and the solution (set) and optimal value function are

$$u^0(w) \in -M_{11}^{+}M_{12}\, w + N(M_{11})$$
$$w^0 \in N(\tilde{M}_{11}(\lambda))$$
$$L(u^0(w^0), w^0, \lambda) = \begin{cases} \lambda/2, & \lambda \geq \left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right| \\ +\infty, & \lambda < \left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right| \end{cases}$$

3. *Strong duality holds and the duality gap is zero for* $\lambda \geq |M_{22}|$.

4. *If* $\left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right| < |M_{22}|$, *then there is an unbounded duality gap for* λ *in that interval*

$$\min_u \max_w L - \max_w \min_u L = +\infty, \quad \textit{for } \left|M_{22} - M'_{12}M_{11}^{+}M_{12}\right| \leq \lambda < |M_{22}|$$

Figure 1 illustrates a common outcome for Corollary 13. Note that the proof of Corollary 13 follows the proof of Proposition 14.

Proposition 14 (Minmax and maxmin of a quadratic function with a parameter and a linear term)**.** *Consider the quadratic function* $L(\cdot) : \mathbb{R}^{n+m+1} \to \mathbb{R}$ *expressed as*

$$\begin{aligned} L(u, w, \lambda) &:= \frac{1}{2}\begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} + \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} - \frac{\lambda}{2}(w'w - 1) \\ &= \frac{1}{2}\begin{bmatrix} u \\ w \end{bmatrix}' \underbrace{\begin{bmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} - \lambda I \end{bmatrix}}_{M(\lambda)} \begin{bmatrix} u \\ w \end{bmatrix} + \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \frac{\lambda}{2} \end{aligned}$$

with $M(\lambda = 0) \geq 0$.

1. *A solution to* $\min_u \max_w L$ *exists if and only if*

$$\lambda \geq |M_{22}| \tag{34}$$

and the solution (set) and optimal value function are

$$\begin{aligned} w^0(u,\lambda) &\in -(M_{22} - \lambda I)^+(M_{12}'\, u + d_2) + N(M_{22} - \lambda I) \\ u^0(\lambda) &\in -\tilde{M}_{22}^+(\lambda)(d_1 - M_{12}(M_{22} - \lambda I)^+ d_2) + N(\tilde{M}_{22}(\lambda)) \\ L(u^0, w^0(u^0), \lambda) &= \begin{cases} \frac{\lambda}{2} - \frac{1}{2} d' M^+(\lambda) d, & \lambda \geq |M_{22}| \\ +\infty, & \lambda < |M_{22}| \end{cases} \end{aligned}$$

2. *A solution to* $\max_w \min_u L$ *exists if and only if*

$$\lambda \geq \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right| \tag{35}$$

and the solution (set) and optimal value function are

$$\begin{aligned} u^0(w,\lambda) &\in -M_{11}^+(M_{12}\, w + d_1) + N(M_{11}) \\ w^0(\lambda) &\in -\tilde{M}_{11}^+(\lambda)(d_2 - M_{12}' M_{11}^+ d_1) + N(\tilde{M}_{11}(\lambda)) \\ L(u^0(w^0), w^0, \lambda) &= \begin{cases} \frac{\lambda}{2} - \frac{1}{2} d' M^+(\lambda) d, & \lambda \geq \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right| \\ +\infty, & \lambda < \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right| \end{cases} \end{aligned}$$

3. *Strong duality holds and the duality gap is zero for* $\lambda \geq |M_{22}|$.

4. *If* $\left|M_{22} - M_{12}' M_{11}^+ M_{12}\right| < |M_{22}|$*, then there is an unbounded duality gap for* λ *in that interval*

$$\begin{gathered} \min_u \max_w L - \max_w \min_u L = +\infty, \\ \text{for } \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right| \leq \lambda < |M_{22}| \end{gathered}$$

Proof. Expand L as

$$L(u,w,\lambda) = \frac{1}{2}\big(u' M_{11} u + 2u' M_{12} w + w'(M_{22} - \lambda I) w + 2u' d_1 + 2w' d_2 + \lambda\big)$$

1. First note that $\max_w L$ exists if and only if $M_{22} - \lambda I \leq 0$. Otherwise $\max_w L(w,u,\lambda) = +\infty$. And $M_{22} - \lambda I \leq 0$ if and only if $\lambda \geq |M_{22}|$, which establishes (34). If this condition is satisfied, from Proposition 5 the solution is

$$\begin{aligned} w^0(u) &\in -(M_{22} - \lambda I)^+(M_{12}'\, u + d_2) + N(M_{22} - \lambda I) \\ L(u, w^0(u), \lambda) &= \frac{1}{2}\big(u' \tilde{M}_{22}(\lambda) u + 2u'(d_1 - M_{12}(M_{22} - \lambda I)^+ d_2) - d_2'(M_{22} - \lambda I)^+ d_2 + \lambda\big) \end{aligned}$$

Since $M_{22} - \lambda I \leq 0$, we have that $(M_{22} - \lambda I)^+ \leq 0$ as well and therefore $\tilde{M}_{22}(\lambda) \geq 0$. Therefore, $\min_u L(u, w^0(u))$ exists and the solution from (5) is

$$u^0(\lambda) \in -\tilde{M}_{22}^+(\lambda)(d_1 - M_{12}(M_{22} - \lambda I)^+ d_2) + N(\tilde{M}_{22}(\lambda))$$

substituting u^0 and $w(u^0)$ into L gives

$$L(u^0, w^0(u^0), \lambda) = \begin{cases} \frac{\lambda}{2} - \frac{1}{2} d' M^+(\lambda) d, & \lambda \geq |M_{22}| \\ +\infty, & \lambda < |M_{22}| \end{cases}$$

and this part is established.

2. Since $M_{11} \geq 0$, $\min_u L$ exists from Proposition 5 and we have

$$u^0(w) \in -M_{11}^+(M_{12}\, w + d_1) + N(M_{11})$$
$$L(u^0(w), w, \lambda) = \frac{1}{2}\big(w'\tilde{M}_{11}(\lambda)w + 2w'(d_2 - M_{12}'M_{11}^+ d_1) - d_1'M_{11}^+ d_1 + \lambda\big)$$

Next note that $\max_w L(u^0(w), w, \lambda)$ exists if and only if $\tilde{M}_{11}(\lambda) \leq 0$, or

$$0 \geq (M_{22} - \lambda I) - M_{12}'M_{11}^+ M_{12}$$
$$\lambda I \geq M_{22} - M_{12}'M_{11}^+ M_{12}$$

and the last inequality is satisfied if and only if $\lambda \geq \big|M_{22} - M_{12}'M_{11}^+ M_{12}\big|$. If that does not hold $\max_w L(u^0(w), w, \lambda) = +\infty$, which establishes (34). When this condition is satisfied

$$w^0(\lambda) \in -\tilde{M}_{11}^+(\lambda)(d_2 - M_{12}'M_{11}^+ d_1) + N(\tilde{M}_{11}(\lambda))$$

Substituting these values into L gives

$$L(u^0(w^0), w^0, \lambda) = \begin{cases} \frac{\lambda}{2} - \frac{1}{2}d'M^+(\lambda)d, & \lambda \geq \big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| \\ +\infty, & \lambda < \big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| \end{cases}$$

and this part is established.

3. Recall from Proposition 6 that $M \geq 0$ implies $M_{11} \geq 0$, $M_{22} \geq 0$, and $M_{22} - M_{12}'M_{11}^+ M_{12} \geq 0$. Since $M_{11} \geq 0$, $M_{11}^+ \geq 0$ as well and therefore $M_{12}'M_{11}^+ M_{12} \geq 0$. Therefore $0 \leq M_{22} - M_{12}'M_{11}^+ M_{12} \leq M_{22}$, which implies $\big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| \leq |M_{22}|$. So for $\lambda > |M_{22}|$, both $\min_u \max_w L$ and $\max_w \min_u L$ have value $\frac{\lambda}{2} - \frac{1}{2}d'M^+(\lambda)d$, and strong duality holds.

4. If M is such that $\big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| < |M_{22}|$, then for λ satisfying $\big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| \leq \lambda < |M_{22}|$, we have $\min_u \max_w L = +\infty$ and $\max_w \min_u L = \frac{\lambda}{2} - \frac{1}{2}d'M^+(\lambda)d$, so the duality gap is infinite, which establishes this part. □

Note that setting $d = 0$ in the proof of Proposition 14 establishes Corollary 13.

For strong duality to hold for all λ such that either problem has a bounded solution requires that $|M_{22}| = \big|M_{22} - M_{12}'M_{11}^+ M_{12}\big|$. The following example shows that it is not *necessary* for $M_{12}'M_{11}^+ M_{12} = 0$ for this condition to hold.

$$M_{22} = M_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad M_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = M_{11}^+$$

$$M_{22} - M_{12}'M_{11}^+ M_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

We have that $|M_{22}| = 1$ and $\big|M_{22} - M_{12}'M_{11}^+ M_{12}\big| = 1$, so the norms are equal but $M_{12}'M_{11}^+ M_{12} = M_{11}^+ \neq 0$.

Constrained quadratic optimization

> A mysterious piece of information has been uncovered. In our innocence we thought we were engaged straightforwardly in solving a single problem (P). But we find we've assumed the role of Player 1 in a certain game in which we have an adversary, Player 2, whose interests are diametrically opposed to ours!
>
> –T.A. Rockafellar

We next consider *maximization* of a *convex* function so that a constraint is required for even existence of a solution. We establish the following result.

Proposition 15 (Constrained quadratic optimization)**.** *Define the convex quadratic function, $V(\cdot) : \mathbb{R}^n \to \mathbb{R}$ and compact constraint set W*

$$V(w) \coloneqq (1/2)w'Dw + w'd \qquad W \coloneqq \{w \mid w'w = 1\}$$

with $D \in \mathbb{R}^{n\times n} \geq 0$. Consider the constrained maximization problem

$$\max_{w\in W} V(w) \tag{36}$$

Define the Lagrangian function

$$L(w, \lambda) = V(w) - (1/2)\lambda(w'w - 1)$$

and the (unconstrained) Lagrangian problem

$$\max_{w} \min_{\lambda} L(w, \lambda) \tag{37}$$

and the (unconstrained) dual Lagrangian problem

$$\min_{\lambda} \max_{w} L(w, \lambda) \tag{38}$$

1. *Solutions to all three problems* (36)*,* (37)*, and* (38) *exist for all $D \geq 0$ and $d \in \mathbb{R}^n$ with optimal value*

$$V^0 = L^0 = -(1/2)d'(D - \lambda_P I)^+ d + \lambda_P/2$$

where

$$\lambda_P \coloneqq \text{ the largest real eigenvalue of } P \qquad P \coloneqq \begin{bmatrix} D & I \\ dd' & D \end{bmatrix} \tag{39}$$

2. *Problems* (37) *and* (38) *satisfy strong duality, and the function $L(w, \lambda)$ has saddle points (sets) (w^*, λ^*) given by*

$$w^* = \begin{cases} \Big(-(D - \lambda_P I)^+ d + N(D - \lambda_P I) \Big) \cap W, & \lambda_P = |D| \\ -(D - \lambda_P I)^{-1} d, & \lambda_P > |D| \end{cases}$$
$$\lambda^* = \lambda_P$$

3. *The optimizer of* (36) *is given by $w^0 = w^*$.*

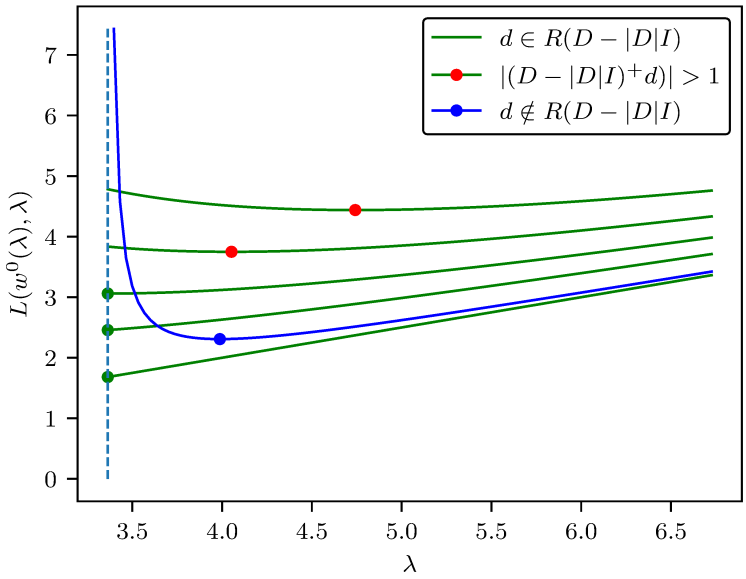

Figure 2: $L(w^0(\lambda), \lambda)$ versus λ for the same D but different d. Green lines: for $d \in R(D - |D|\, I)$, L is bounded for all $\lambda \geq |D|$. Green dots: for $|(D - |D|\, I)^+ d| \leq 1$, the optimum is on the boundary and $\lambda_P = |D|$. Red dots: for $|(D - |D|\, I)^+ d| > 1$, the optimum is in the interior and $\lambda_P > |D|$. Blue line and dot: for $d \notin R(D - |D|\, I)$, L is unbounded at $\lambda = |D|$, and the optimum is in the interior, and $\lambda_P > |D|$.

4. *The optimizer of* (37) *is given by*

$$w^0 = w^* \qquad \underline{\lambda}^0(w^0) = \mathbb{R}$$

5. *The optimizer of* (38) *is given by*

$$\overline{w}^0(\lambda^0) = \begin{cases} -(D - \lambda_P I)^+ d + N(D - \lambda_P I), & \lambda_P = |D| \\ -(D - \lambda_P I)^{-1} d, & \lambda_P > |D| \end{cases}$$
$$\lambda^0 = \lambda_P$$

6. *Additionally* $\lambda_P = |D|$ *if and only if (i)* $d \in R(D - |D|\, I)$ *and (ii)* $|(D - |D|\, I)^+ d| \leq 1$. *If (i) or (ii) do not hold, then* $\lambda_P > |D|$ *and* $\left|(D - \lambda_P I)^{-1} d\right| = 1$.

Figure 2 shows the possible behaviors. The green lines show the case $d \in R(D - |D|\, I)$, for different $|d|$. For $d = 0$ (bottom green line), $\lambda_P = |D|$ and the optimum is on the boundary. Increasing $|d|$ eventually produces a zero derivative at $\lambda = |D|$ (third green line from bottom). Further increasing $|d|$ makes the derivative at $\lambda = |D|$ negative (top two green lines), and $\lambda_P > |D|$ (red dots), and the optimum moves to the interior. The blue line shows the case $d \notin R(D - |D|\, I)$. L is unbounded at $\lambda = |D|$, $\lambda_P > |D|$ (blue dot) and the optimum is again in the interior.

To organize the proof of this proposition, we treat the Lagrangian, dual Lagrangian, and saddle-point problems in separate lemmas, and then combine them. We start with

the dual Lagrangian minmax problem. We shall find that all of the information about λ_P emerges from this problem.

Lemma 16 (Dual Lagrangian of constrained quadratic optimization)**.** *Consider the dual Lagrangian problem*

$$\min_{\lambda} \max_{w} L(w, \lambda) \qquad L(w, \lambda) \coloneqq (1/2) w' D w + w' d - (1/2)\lambda(w'w - 1) \tag{40}$$

with $D \geq 0$. We have the following results.

1. *This problem is equivalent to*

$$\min_{\lambda \geq |D|} \max_{w} L(w, \lambda)$$

2. *The solution exists for all D and d and has optimal value*

$$L^0 = -(1/2) d' (D - \lambda_P I)^+ d + \lambda_P / 2 \tag{41}$$

where λ_P and matrix $P \in \mathbb{R}^{2n \times 2n}$ are defined as

$$\lambda_P \coloneqq \textit{ the largest real eigenvalue of } P \qquad P \coloneqq \begin{bmatrix} D & I \\ dd' & D \end{bmatrix} \tag{42}$$

3. *The optimal λ and $w^0(\lambda)$ are given by*

$$\lambda^0 = \lambda_P \in [|D|, \infty)$$
$$w^0(\lambda^0) = \begin{cases} -(D - |D| I)^+ d + N(D - |D| I), & \lambda_P = |D| \\ -(D - \lambda_P I)^{-1} d, & \lambda_P \in (|D|, \infty) \end{cases}$$

4. *We have that $\lambda_P = |D|$ if and only if (i) $d \in R(D - |D| I)$, and (ii) $|(D - |D| I)^+ d| \leq 1$. Otherwise $\lambda_P > |D|$. If (i) is violated, $L(w^0(\lambda), \lambda) = +\infty$ at $\lambda = |D|$. If (i) holds but (ii) is violated, then $L(w^0(\lambda, \lambda)$ is finite at $\lambda = |D|$, but $(d/d\lambda) L(w^0(\lambda), \lambda) < 0$ at $\lambda = |D|$.*

Proof. To establish statement 1 in the lemma, note that if $\lambda < |D|$, then $D - \lambda I > 0$ and $\max_w L(w, \lambda) = +\infty$. So adding the constraint $\lambda \geq |D|$ to the outer minimization does not alter the solution.

To establish statements 2–4 in the lemma, we make use of the SVD of matrix $D = UMU'$, which we partition as

$$D = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} |D| I_p & \\ & M_2 \end{bmatrix} \begin{bmatrix} U_1' \\ U_2' \end{bmatrix}$$

where p is the multiplicity of the largest eigenvalue of D, $1 \leq p \leq n$. Also denote $y = U'd, y_1 = U_1'd, y_2 = U_2'd$.

We break the problem into two cases.

1. Case $\lambda_P \in (|D|, \infty)$.

   For $\lambda \in (|D|, \infty)$, we have from Proposition 5 that $w^0(\lambda) = -(D - \lambda I)^{-1} d$ and $L(w^0(\lambda), \lambda) = -(1/2) d' (D - \lambda I)^{-1} d + \lambda/2$. $L(w^0(\lambda), \lambda)$ is differentiable, and taking two derivatives gives

$$\frac{dL}{d\lambda} = (1/2)(1 - d'(D - \lambda I)^{-2} d)$$
$$\frac{d^2 L}{d\lambda^2} = -d'(D - \lambda I)^{-3} d$$

Setting the first derivative to zero yields

$$\begin{aligned} 0 &= 1 - d'(D-\lambda I)^{-2}d \\ &= \det(1 - d'(D-\lambda I)^{-2}d) \\ &= \det(I - dd'(D-\lambda I)^{-2}) \\ &= \det(D-\lambda I) - dd'(D-\lambda I)^{-1})\det(D-\lambda I)^{-1} \end{aligned} \tag{43}$$

where we have used the fact that $\det(I+AB) = \det(I+BA)$. Since $\det(D-\lambda I) \neq 0$, we can multiply both sides of the last equality by $\det(D-\lambda I)^2$ to obtain

$$\begin{aligned} 0 &= \det(D-\lambda I) - dd'(D-\lambda I)^{-1})\det(D-\lambda I) \\ &= \det\left(\begin{bmatrix} D-\lambda I & I \\ dd' & D-\lambda I \end{bmatrix}\right) = \det(P-\lambda I) \end{aligned}$$

where we have used the partitioned determinant formula, which is valid since $D-\lambda I$ is nonsingular for $\lambda > |D|$. Therefore the first derivative of L vanishes in the interval $(|D|, \infty)$ if and only if there exists a real-valued eigenvalue of P in this interval. Also, we have from (43) that

$$1 = d'(D-\lambda I)^{-2}d = d'U(M-\lambda I)^{-2}U'd = y'(M-\lambda I)^{-2}y$$

with $y \coloneqq U'd$. So we conclude that $y \neq 0$ for this case. Examining the second derivative, we have

$$\frac{d^2L}{d\lambda^2} = d'(\lambda I - D)^{-3}d = y'(\lambda I - M)^{-3}y$$

Note that since $\lambda > |D|$, $(\lambda I - M)^{-3} > 0$, and since $y \neq 0$, we have that $d^2L/d\lambda^2 > 0$ on $(|D|, \infty)$ and therefore $L(w^0(\lambda), \lambda)$ is strictly convex on this interval. Therefore the minimizer of L is unique and the first derivative is zero at the solution. We also know that there is only one real eigenvalue of P in this interval due to the uniqueness of the optimal solution. Therefore we have established that $\lambda^0 = \lambda_P \in (|D|, \infty)$ is the optimal solution. Substituting this solution into L gives

$$L^0 = L(w^0(\lambda^0), \lambda^0) = -(1/2)d'(D-\lambda_P I)^{-1}d + \lambda_P/2$$

verifying that (41) holds for the first case.

2. Case $\lambda_P \notin (|D|, \infty)$. In this case, we first show that the $\lambda^0 = |D|$. Using the SVD, we have for $\lambda \in (|D|, \infty)$

$$\begin{aligned} L(w^0(\lambda), \lambda) &= -(1/2)d'(D-\lambda I)^{-1}d + \lambda/2 \\ &= -(1/2)\begin{bmatrix} y_1' & y_2' \end{bmatrix}\begin{bmatrix} \frac{1}{|D|-\lambda} I_p & \\ & (M_2 - \lambda I)^{-1} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \lambda/2 \end{aligned}$$

From this expression for L, note that $y_1 = U_1'd$ must be zero for this case, or $\lim_{\lambda \to |D|^+} L(w^0(\lambda), \lambda) = +\infty$, which is a contradiction since $\lim_{\lambda \to +\infty} L(w^0(\lambda), \lambda) = +\infty$ as well, and L is a smooth function on the interval $(|D|, \infty)$, so it must have a minimum on that interval (zero derivative), but by assumption it does not have

a zero derivative on that interval. Note that $y_1 = U_1'd = 0$ is equivalent to $d \in R(D - |D|\, I)$, which can be seen from the SVD of $D - |D|\, I$

$$D - |D|\, I = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} 0 & \\ & M_2 - |D|\, I \end{bmatrix} \begin{bmatrix} U_1' \\ U_2' \end{bmatrix}$$

so the columns of U_1 are a basis for $N(D - |D|\, I)$ and d is orthogonal to the columns of U_1 so $d \in R(D - |D|\, I)$. Substituting $y_1 = 0$, into the expression for $L(w^0(\lambda), \lambda)$ gives

$$L(w^0(\lambda), \lambda) = -(1/2)d'(D - \lambda I)^+ d + \lambda/2, \qquad \lambda \geq |D|\,,\ d \in R(D - |D|\, I) \quad (44)$$

and $L(w^0(\lambda), \lambda)$ is smooth on the interval *including* the left boundary, $[|D|\,, \infty)$, and the optimizer must be on the boundary, $\lambda^0 = |D|$. For this value of λ, the inner maximization over w gives from Proposition 5

$$w^0(\lambda^0) = -(1/2)(D - |D|\, I)^+ d + N(D - |D|\, I)$$

and evaluating L^0 gives

$$\begin{aligned} L(w^0(\lambda^0), \lambda^0) &= -(1/2)y_2'(M_2 - |D|\, I)^{-1} y_2 + |D|\,/2 \\ &= -(1/2)d'(D - |D|\, I)^+ d + |D|\,/2 \end{aligned} \quad (45)$$

verifying (41) for this case.

Taking the derivative of (44) and evaluating at $\lambda = |D|$ gives

$$(d/d\lambda)L(w^0(\lambda), \lambda) = (1/2)(1 - d'((D - |D|\, I)^+)^2 d) = (1/2)(1 - \left|(D - |D|\, I)^+ d\right|^2)$$

which is non-negative if and only if $|(D - |D|\, I)^+ d| \leq 1$. Otherwise the derivative at the boundary is negative and the optimal λ is in the interval $(|D|\,, \infty)$, which is the previous case. Therefore $\lambda^0 = |D|$ if and only if

$$d \in R(D - |D|\, I), \quad \left|(D - |D|\, I)^+ d\right| \leq 1$$

Next we show that $|D|$ is an eigenvalue of P in this case. Factoring $P - \lambda I$ gives

$$\begin{aligned} P - \lambda I &= \begin{bmatrix} U & \\ & U \end{bmatrix} \begin{bmatrix} U'DU - \lambda I & I \\ U'dd'U & U'DU - \lambda I \end{bmatrix} \begin{bmatrix} U' & \\ & U' \end{bmatrix} \\ &= \begin{bmatrix} U & \\ & U \end{bmatrix} \begin{bmatrix} (|D| - \lambda) I_p & & I & \\ & M_2 - \lambda I & & I \\ y_1 y_1' & y_1 y_2' & (|D| - \lambda) I_p & \\ y_2 y_1' & y_2 y_2' & & M_2 - \lambda I \end{bmatrix} \begin{bmatrix} U' & \\ & U' \end{bmatrix} \end{aligned}$$

Since the leading and trailing matrices are inverses of each other, we have a similarity transformation, and the eigenvalues of the inner matrix are the eigenvalues of P. Setting $y_1 = 0$ in the inner matrix and setting $\lambda = |D|$ gives a zero third block row of the inner matrix, and it is singular. Therefore $\lambda = |D|$ is an eigenvalue of P. Since there are no real eigenvalues of P in $(|D|\,, \infty)$, we have that $|D|$ is the *largest* real eigenvalue of P for this case, and we have established that $\lambda^0 = \lambda_P$ also for this case.

Summarizing, we have broken the problem into two cases. In the first case we have shown that $\lambda^0 = \lambda_P > |D|$, and $(d/d\lambda)L(w^0(\lambda), \lambda)$ is zero at $\lambda = \lambda_P$. In this case there is only one real eigenvalue of P in $(|D|, \infty)$.

In the second case, we have that $\lambda^0 = \lambda_P = |D|$, the boundary of the feasible set. We have also shown that $d \in R(D - |D| I)$, and $|(D - |D| I)^+ d| \leq 1$ for this case. If $d \notin R(D - |D| I)$, then $L(w^0(\lambda), \lambda)$ is $+\infty$ at $\lambda = |D|$, which is in the first case. If $d \in R(D - |D| I)$, but $|(D - |D| I)^+ d| > 1$, then $L(w^0(\lambda), \lambda)$ is finite at $\lambda = |D|$, but the derivative is negative, which is again in the first case. Thus we have established statements 2–4 in the lemma and the proof is complete. □

Next we turn to the saddle points.

Lemma 17 (Saddle points of the Lagrangian of constrained quadratic optimization)**.** *The following* (w^*, λ^*) *are saddle points of* $L(w, \lambda) \coloneqq (1/2)w'Dw + w'd - (1/2)\lambda(w'w - 1)$.

$$w^* = \begin{cases} \left((D - |D| I)^+ d + N(D - |D| I) \right) \cap W & \lambda_P = |D| \\ (D - \lambda_P I)^{-1} d, & \lambda_P > |D| \end{cases}$$
$$\lambda^* = \lambda_P$$

Proof. From the definition of a saddle point we need to establish the inequalities

$$L(w, \lambda^*) \leq L(w^*, \lambda^*) \leq L(w^*, \lambda)$$

hold for all $w \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

Taking the second inequality first, we have that $L(w^*, \lambda) = (1/2)(w^*)'Dw^* + (w^*)'d$ for all λ since $(w^*)'w^* = 1$. Therefore $L(w^*, \lambda) = L(w^*, \lambda^*)$ for all λ and the second inequality is established with equality.

Turning to the first inequality, we consider the two cases; (i) $\lambda^* = \lambda_P > |D|$, and (ii) $\lambda^* = \lambda_P = |D|$ and $d \in R(D - |D| I)$. For the first case we know that

$$\max_w L(w, \lambda_P) = -(1/2)d'(D - \lambda_P I)^{-1} d + \lambda_P/2 = L(w^*, \lambda_P) = L(w^*, \lambda^*)$$

where the maximization over w is unconstrained, and the first equality follows from Proposition 5. Since this equality holds for the maximizer, we have that $L(w, \lambda^*) = L(w, \lambda_P) \leq L(w^*, \lambda^*)$ for all $w \in \mathbb{R}^n$, where the first equality comes from the definition of λ^*. Thus the first inequality holds for the first case.

Turning to the second case, we have similarly from Proposition 5

$$\max_w L(w, |D|) = -(1/2)d'(D - |D| I)^+ d + |D|/2 = L(w^*, |D|) = L(w^*, \lambda^*)$$

where again the maximization over w is unconstrained, and we have $L(w, \lambda^*) = L(w, |D|) \leq L(w^*, \lambda^*)$ for all $w \in \mathbb{R}^n$. Thus both cases satisfy the first inequality, both inequalities have been established, and the result is proven. □

Finally, we address the original constrained optimization of the concave quadratic function and its Lagrangian

Lemma 18 (Constrained quadratic optimization and its Lagrangian)**.** *We are given the following: (i) a convex quadratic function and compact constraint set* W *defined in Proposition* 15

$$V(w) \coloneqq (1/2)w'Dw + w'd \qquad W \coloneqq \{w \mid w'w = 1\}$$

with $D \in \mathbb{R}^{n\times n} \geq 0$, (ii) the constrained maximization problem (36)

$$\max_{w\in W} V(w)$$

with Lagrangian function

$$L(w,\lambda) = V(w) - (1/2)\lambda(w'w - 1)$$

and the (unconstrained) Lagrangian problem (37)

$$\max_{w}\min_{\lambda} L(w,\lambda)$$

1. *Solutions to* (36) *and* (37) *exist for all $D \geq 0$ and $d \in \mathbb{R}^n$ and achieve the same optimal value $V^0 = L^0 = -(1/2)d'(D - \lambda_P I)^+ d + \lambda_P/2$.*

2. *The optimizer of* (36)*, denoted w^0, is given by $w^0 = w^*$ where w^* the saddle-point solution set from Lemma* 17*.*

3. *The optimizer of* (37)*, denoted $(w_L^0, \underline{\lambda}(w_L^0))$ is given by $w_L^0 = w^*$ and $\underline{\lambda}(w_L^0) = \mathbb{R}$.*

Proof.

The solution to (36) exists since $V(\cdot)$ is continuous and W is compact. The solution to (37) exists and satisfies strong duality with (38) due to the saddle-point theorem (Proposition 12) and Lemma 17, so we have that

$$L^0 = \max_{w}\min_{\lambda} L = \min_{\lambda}\max_{w} L = -(1/2)d'(D - \lambda_P I)^+ d + \lambda_P/2$$

where the last equality follows by (41). From the saddle-point theorem, we also have that $w_L^0 = w^*$ for the optimizer of the Lagrangian. Since w_L^0 satisfies $(w_L^0)'w_L^0 = 1$, it follows that $\underline{\lambda}^0(w_L^0) = \mathbb{R}$. Finally, we have that value $V^0 = L^0$ and set $w^0 = w_L^0$ by Proposition 7, and the proof is complete. □

The proofs of Lemma 16, Lemma 17, and Lemma 18 have proven Proposition 15.

Discussion of Proposition 15. The basic problem of constrained, nonconvex quadratic optimization has appeared in several fields. In the optimization literature it is known as the "trust-region" problem. Nocedal and Wright discuss the numerical solution of the trust-region problem in the context of nonlinear programming (Nocedal and Wright, 2006, p.69). Boyd and Vandenberghe establish strong dualilty of the Lagrangian and dual Langrangian formulations of the problem (Boyd and Vandenberghe, 2004, Appendix B). The complete solution provided in Proposition 15 appears to be new to this work. The authors would also like to acknowledge Robin Strässer for his work on earlier versions of Proposition 15 (Mannini, Strässer, and Rawlings, 2024).

Constrained minmax and maxmin. To compactly state the results in this section it is convenient to define two functions

$$M(\lambda) \coloneqq \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} - \lambda I \end{bmatrix} \qquad L(\lambda) \coloneqq -(1/2)d'M^+(\lambda)d + \lambda/2$$

and for convenience, we repeat the definition of the Schur complements

$$\tilde{M}_{11}(\lambda) \coloneqq (M_{22} - \lambda I) - M_{12}'M_{11}^+M_{12} \qquad \tilde{M}_{22}(\lambda) \coloneqq M_{11} - M_{12}(M_{22} - \lambda I)^+M_{12}'$$

Corollary 19 (Minmax and maxmin of constrained quadratic functions)**.** *Consider quadratic function* $V(\cdot) : \mathbb{R}^{n+m} \to \mathbb{R}$ *and compact constraint set* W

$$V(u,w) \coloneqq \frac{1}{2}\begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} \qquad W \coloneqq \{w \mid w'w = 1\}$$

with $M \in \mathbb{R}^{(n+m)\times(n+m)} \geq 0$*,* $M_{22} \in \mathbb{R}^{n\times n}$*,* $M_{11} \in \mathbb{R}^{m\times m}$*,* $M_{12} \in \mathbb{R}^{m\times n}$*, and the two constrained optimization problems.*

$$\min_u \max_{w\in W} V(u,w) \tag{46}$$

$$\max_{w\in W} \min_u V(u,w) \tag{47}$$

1. *The solution to* (46) *is*

$$\begin{aligned} V(u^0, w^0(u^0)) &= \frac{\lambda^0}{2} \\ u^0 &\in N(\tilde{M}_{22}(\lambda^0)) \\ w^0(u^0) &\in \Big(-(M_{22} - \lambda^0 I)^+ M_{12}'\, u^0 + N(M_{22} - \lambda^0 I) \Big) \cap W \end{aligned}$$

   with $\lambda^0 = |M_{22}|$*.*

2. *The solution to* (47) *is*

$$\begin{aligned} V(u^0(w^0), w^0) &= \frac{\lambda^0}{2} \\ w^0 &\in \big(N(\tilde{M}_{11}(\lambda^0)) \big) \cap W \\ u^0(w^0) &\in -M_{11}^+ M_{12}\, w^0 + N(M_{11}) \end{aligned}$$

   with $\lambda^0 = \big| M_{22} - M_{12}' M_{11}^+ M_{12} \big|$*.*

The proof of Corollary 19 is given at the end of the proof of Proposition 20.

Next we add the linear term to $V(u,w)$, which may seem harmless but actually precludes a closed-form solution as in Corollary 19. Here a nonlinear optimization over scalar λ remains.

Proposition 20 (Minmax and maxmin of constrained quadratic functions with linear term)**.** *Consider quadratic function* $V(\cdot) : \mathbb{R}^{n+m} \to \mathbb{R}$ *and compact constraint set* W

$$V(u,w) \coloneqq \frac{1}{2}\begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} + \begin{bmatrix} u \\ w \end{bmatrix}' \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \qquad W \coloneqq \{w \mid w'w = 1\}$$

with $M \in \mathbb{R}^{(n+m)\times(n+m)} \geq 0$*, and the two constrained optimization problems.*

$$\min_u \max_{w\in W} V(u,w) \tag{48}$$

$$\max_{w\in W} \min_u V(u,w) \tag{49}$$

1. *The solution to* (48) *is*

$$\begin{aligned} V(u^0, w^0(u^0)) &= -(1/2) d' M^+(\lambda^0) d + \lambda^0/2 \\ u^0 &\in -\tilde{M}_{22}^+(\lambda^0)(d_1 - M_{12}(M_{22} - \lambda^0 I)^+ d_2) + N(\tilde{M}_{22}(\lambda^0)) \\ w^0(u^0) &\in \Big(-(M_{22} - \lambda^0 I)^+ (M_{12}'\, u^0 + d_2) + -N(M_{22} - \lambda^0 I) \Big) \cap W \end{aligned}$$

where λ^0 must be computed from the following nonlinear optimization problem

$$\lambda^0 = \arg \min_{\lambda \geq |M_{22}|} L(\lambda)$$

2. *The solution to* (49) *is*

$$\begin{aligned} V(u^0(w^0), w^0) &= -(1/2) d' M^+(\lambda^0) d + \lambda^0/2 \\ w^0 &\in \left(- \tilde{M}_{11}^+(\lambda^0)(d_2 - M_{12}' M_{11}^+ d_1) + N(\tilde{M}_{11}(\lambda^0)) \right) \cap W \\ u^0(w^0) &\in -M_{11}^+(M_{12}\, w^0 + d_1) + N(M_{11}) \end{aligned}$$

where λ^0 must be computed from the following nonlinear optimization problem

$$\lambda^0 = \arg \min_{\lambda \geq \left| M_{22} - M_{12}' M_{11}^+ M_{12} \right|} L(\lambda)$$

Proof. Expand V as

$$V(u, w) = \frac{1}{2} u' M_{11} u + u' M_{12} w + \frac{1}{2} w' M_{22} w + u' d_1 + w' d_2$$

1. Note that from Proposition 7 the optimization problem

$$\min_u \max_{w \in W} V(u, w)$$

is equivalent to the Lagrangian problem

$$\min_u \max_w \min_\lambda L(u, w, \lambda), \qquad L(u, w, \lambda) := V(u, w) - \frac{\lambda}{2}(w'w - 1)$$

From Proposition 15 strong duality holds for $\max_w \min_\lambda L(u, w, \lambda)$, so the optimization problem is also equivalent to the dual Lagrangian problem

$$\min_\lambda \min_u \max_w L(u, w, \lambda)$$

but when we use the dual to obtain a solution set for the inner $\max_w$ problem, which gives $w^0(u, \lambda)$, that solution set may be too large. We fix this issue subsequently. From Proposition 14 a solution to $\min_u \max_w L(u, w, \lambda)$ exists if and only if $\lambda \geq |M_{22}|$ and the solution (set) and optimal value function are

$$\begin{aligned} w^0(u, \lambda) &\in -(M_{22} - \lambda I)^+ (M_{12}'\, u + d_2) + N(M_{22} - \lambda I) \\ u^0(\lambda) &\in -\tilde{M}_{22}^+(\lambda)(d_1 - M_{12}(M_{22} - \lambda I)^+ d_2) + N(\tilde{M}_{22}(\lambda)) \\ L(u^0(\lambda), w^0(u^0, \lambda), \lambda) &= \begin{cases} L(\lambda), & \lambda \geq |M_{22}| \\ +\infty, & \lambda < |M_{22}| \end{cases} \end{aligned}$$

The remaining optimization, which is equivalent to solving (48), is

$$\min_{\lambda \geq |M_{22}|} L(\lambda) \tag{50}$$

which establishes that

$$\lambda^0 := \arg \min_{\lambda \geq |M_{22}|} L(\lambda)$$

Substituting λ^0 in $u^0(\lambda)$, $w^0(u(\lambda))$, and $L(u^0(\lambda), w^0(u^0, \lambda), \lambda)$ gives

$$\begin{aligned} w^0(u) &\in -(M_{22} - \lambda^0 I)^+(M_{12}'\, u + d_2) + N(M_{22} - \lambda^0 I) \\ u^0(\lambda^0) &\in \tilde{M}_{22}^+(\lambda^0)(d_1 - M_{12}(M_{22} - \lambda^0 I)^+ d_2) + N(\tilde{M}_{22}(\lambda^0)) \\ L(u^0, w^0(u^0), \lambda^0) &= -(1/2) d' M^+(\lambda^0) d + \lambda^0/2 \end{aligned}$$

Finally, restricting the dual solution $w^0(u^0)$ to satisfy the constraint $w \in W$ by intersecting $w^0(u^0)$ with W giving

$$w^0(u^0) \in \Big(-(M_{22} - \lambda^0 I)^+(M_{12}'\, u^0 + d_2) + N(M_{22} - \lambda I) \Big) \cap W$$

Since the constraint is satisfied, $L(u^0, w^0(u^0), \lambda^0) = V(u^0, w^0(u^0))$ giving

$$V(u^0, w^0(u^0)) = -(1/2) d' M^+(\lambda^0) d + \lambda^0/2$$

and part 1 is established.

2. From Proposition 7 the optimization problem

$$\max_{w \in W} \min_u V(u, w)$$

is equivalent to the Lagrangian minmax problem

$$\max_w \min_\lambda \min_u L(u, w, \lambda)$$

Unlike the previous part, before we can use Proposition 15 and invoke strong duality, we must first examine the form of the innermost problem $\min_u L(u, w, \lambda)$. Using Proposition 5 to solve the $\min_u L(u, w, \lambda)$ problem and evaluating at the optimal u gives

$$L(u^0(w, \lambda), w, \lambda) = \frac{1}{2} w' \tilde{M}_{11}(\lambda) w + w'(d_2 - M_{12}' M_{11}^+ d_1) - \frac{1}{2} d_1' M_{11}^+ d_1 + \frac{\lambda}{2}$$

Given the functional form of w and λ above, we see that Proposition 15 indeed applies, and the optimization problem (49) is also equivalent to the dual Lagrangian problem

$$\min_\lambda \max_w \min_u L(u, w, \lambda)$$

Again, using the dual to obtain the solution for the inner $\max_w$ problem, which gives $w^0(\lambda)$, may give a solution set that is too large, and we further restrict that set subsequently. From Proposition 14 a solution to $\max_w \min_u L(u, w, \lambda)$ exists if and only if $\lambda \geq \left| M_{22} - M_{12}' M_{11}^+ M_{12} \right|$ and the solution (set) and optimal value function are

$$\begin{aligned} u^0(w, \lambda) &\in -M_{11}^+(M_{12}\, w + d_1) + N(M_{11}) \\ w^0(\lambda) &\in -\tilde{M}_{11}^+(\lambda)(d_2 - M_{12}' M_{11}^+ d_1) + N(\tilde{M}_{11}(\lambda)) \\ L(u^0(w^0, \lambda), w^0(\lambda), \lambda) &= \begin{cases} L(\lambda), & \lambda \geq \left| M_{22} - M_{12}' M_{11}^+ M_{12} \right| \\ +\infty, & \lambda < \left| M_{22} - M_{12}' M_{11}^+ M_{12} \right| \end{cases} \end{aligned}$$

The remaining optimization, which is equivalent to solving (49), is

$$\min_{\lambda \geq \left| M_{22} - M_{12}' M_{11}^+ M_{12} \right|} L(\lambda) \tag{51}$$

which establishes that

$$\lambda^0 := \arg \min_{\lambda \geq \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right|} L(\lambda)$$

Restricting w^0 by enforcing the constraint $w^0 \in W$ gives

$$w^0(\lambda^0) \in \left(-\tilde{M}_{11}^+(\lambda^0)(d_2 - M_{12}' M_{11}^+ d_1) + N(\tilde{M}_{11}(\lambda^0)) \right) \cap W$$

Since the constraint is satisfied, $L(u^0(w^0), w^0, \lambda) = V(u^0(w^0), w^0)$ giving

$$V(u^0(w^0), w^0) = -(1/2) d' M^+(\lambda^0) d + \lambda^0/2$$

and part 2 is established. □

To prove Corollary 19, note that $d = 0$ so that (50) and (51) can be solved analytically giving $\lambda^0 = |M_{22}|$ for (50) and $\lambda^0 = \left|M_{22} - M_{12}' M_{11}^+ M_{12}\right|$ for (51). Substituting these λ^0 values and $d = 0$ into the statement of Proposition 20 then establishes the results of Corollary 19.

References

S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

O. Mangasarian. *Nonlinear Programming*. SIAM, Philadelphia, PA, 1994.

D. Mannini, R. Strässer, and J. B. Rawlings. Optimal design of disturbance attenuation feedback controllers for linear dynamical systems. In *American Control Conference*, Toronto, CA, July 8–12, 2024.

J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, second edition, 2006.

E. Polak. *Optimization: Algorithms and Consistent Approximations*. Springer Verlag, New York, 1997. ISBN 0-387-94971-2.

J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Design, and Computation*. Nob Hill Publishing, Santa Barbara, CA, 2nd, paperback edition, 2020. 770 pages, ISBN 978-0-9759377-5-4.

R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Rev.*, 35(2):183–238, 1993.

R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer-Verlag, 1998.

J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Math. Ann.*, 100:295–320, 1928. doi: 10.1007/BF01448847.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton and Oxford, 1944.