Semidefinite Programming Duality
Implications for System Theory and Computation

Venkataramanan (Ragu) Balakrishnan
School of ECE, Purdue University

6 July, 2004

MTNS Symposium, Leuven

Semidefinite Programming Duality
Implications for System Theory and Computation

Venkataramanan (Ragu) Balakrishnan
School of ECE, Purdue University

6 July, 2004

MTNS Symposium, Leuven

Joint work with Lieven Vandenberghe

Semidefinite Programming Duality
Implications for System Theory and Computation

Venkataramanan (Ragu) Balakrishnan
School of ECE, Purdue University

6 July, 2004

MTNS Symposium, Leuven

Joint work with Lieven Vandenberghe and Tae Roh, UCLA
Anders Hansson and Ragnar Wallin, Linköping University

Basic ideas

- Many engineering problems yield LMIs and SDPs

Basic ideas

- Many engineering problems yield LMIs and SDPs

- LMIs are convex constraints, SDPs are convex optimization problems

Basic ideas

- Many engineering problems yield LMIs and SDPs

- LMIs are convex constraints, SDPs are convex optimization problems

- Rich duality theory in convex optimization

Basic ideas

- Many engineering problems yield LMIs and SDPs
- LMIs are convex constraints, SDPs are convex optimization problems
- Rich duality theory in convex optimization
- Implications of convex duality on the underlying engineering problem:

Basic ideas

- Many engineering problems yield LMIs and SDPs
- LMIs are convex constraints, SDPs are convex optimization problems
- Rich duality theory in convex optimization
- Implications of convex duality on the underlying engineering problem:
  - ⋆ theoretical
  - ⋆ for computation

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
- Computational implications of SDP duality

LMIs and Semidefinite Programming

- $\mathcal{V}$ is a finite-dimensional Hilbert space, $\mathcal{S}$ is a subspace of Hermitian matrices, $\mathcal{F} : \mathcal{V} \rightarrow \mathcal{S}$ is a linear mapping, $F_0 \in \mathcal{S}$
- SDP is an optimization of the form:

$$\begin{array}{ll} \texttt{minimize:} & \langle c, x \rangle_{\mathcal{V}} \\ \texttt{subject to:} & \mathcal{F}(x) + F_0 \succeq 0 \end{array}$$

LMIs and Semidefinite Programming

- $\mathcal{V}$ is a finite-dimensional Hilbert space, $\mathcal{S}$ is a subspace of Hermitian matrices, $\mathcal{F} : \mathcal{V} \to \mathcal{S}$ is a linear mapping, $F_0 \in \mathcal{S}$
- SDP is an optimization of the form:

$$\begin{array}{ll} \texttt{minimize:} & \langle c, x \rangle_{\mathcal{V}} \\ \texttt{subject to:} & \mathcal{F}(x) + F_0 \succeq 0 \end{array}$$

- Inequality $\mathcal{F}(x) + F_0 \succeq 0$ is an LMI
- $F \succeq 0$ means F is positive semidefinite, that is $u^T F u \geq 0$ for all vectors u
- LMIs are nonlinear, but *convex* constraints:
  If $F(x) \succeq 0$ and $F(y) \succeq 0$, then

$$F\left(\lambda x + (1-\lambda)y\right) = \lambda F(x) + (1-\lambda)F(y) \succeq 0 \text{ for all } \lambda \in [0, 1]$$

- Can have linear equality constraints on x

SDP vs. LP

SDP		LP	
minimize	$\langle c, x\rangle$	minimize	$\langle c, x\rangle$
subject to	$\mathcal{F}(x) + F_0 \succeq 0$	subject to	$\langle a_i, x\rangle \le b_i, i = 1, \ldots, N$

- Same linear objective
- Linear matrix inequality constraint instead of linear scalar inequalities

SDP applications

- Systems and control (quite well-known)

SDP applications

- Systems and control (quite well-known)

- Signal processing

SDP applications

- Systems and control (quite well-known)
- Signal processing
- Nonconvex optimization

SDP applications

- Systems and control (quite well-known)
- Signal processing
- Nonconvex optimization
- ... many others

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
  - ⋆ Theorems of alternatives, and the KYP Lemma
  - ⋆ Primal and Dual SDPs, and new bounds on the $\mathbf{H}_\infty$ norm
- Computational implications of SDP duality

A theorem of alternatives for LMIs

Exactly one of the following statements is true

1. $\mathcal{F}(x) + F_0 \succ 0$ is feasible

2. There exists $Z \in \mathcal{S}$ s.t. $Z \succeq\neq 0$, $\mathcal{F}^{\mathrm{adj}}(Z) = 0$, $\langle F_0, Z \rangle_{\mathcal{S}} \leq 0$

$\mathcal{F}^{\mathrm{adj}}(\cdot)$ denotes adjoint map, i.e., $\forall\, x \in \mathcal{V}$, $Z \in \mathcal{S}$, $\langle \mathcal{F}(x), Z \rangle_{\mathcal{S}} = \langle x, \mathcal{F}^{\mathrm{adj}}(Z) \rangle_{\mathcal{V}}$

A theorem of alternatives for LMIs

Exactly one of the following statements is true

1. $\mathcal{F}(x) + F_0 \succ 0$ is feasible

2. There exists $Z \in \mathcal{S}$ s.t. $Z \succneqq 0$, $\mathcal{F}^{\mathrm{adj}}(Z) = 0$, $\langle F_0, Z \rangle_{\mathcal{S}} \leq 0$

Idea of proof: Easy part, contradiction if (1) and (2) both hold:

$$0 < \langle Z, \mathcal{F}(x) + F_0 \rangle_{\mathcal{S}} = \langle \mathcal{F}^{\mathrm{adj}}(Z), x \rangle_{\mathcal{V}} + \langle Z, F_0 \rangle_{\mathcal{S}} \leq 0$$

Difficult part: If (1) is false, (2) is true. Follows from separating hyperplane theorem for convex sets

A theorem of alternatives for LMIs

Exactly one of the following statements is true

1. $\mathcal{F}(x) + F_0 \succ 0$ is feasible

2. There exists $Z \in \mathcal{S}$ s.t. $Z \succneqq 0$, $\mathcal{F}^{\mathrm{adj}}(Z) = 0$, $\langle F_0, Z \rangle_{\mathcal{S}} \leq 0$

Idea of proof: Easy part, contradiction if (1) and (2) both hold:

$$0 < \langle Z, \mathcal{F}(x) + F_0 \rangle_{\mathcal{S}} = \langle \mathcal{F}^{\mathrm{adj}}(Z), x \rangle_{\mathcal{V}} + \langle Z, F_0 \rangle_{\mathcal{S}} \leq 0$$

Difficult part: If (1) is false, (2) is true. Follows from separating hyperplane theorem for convex sets

- Variants available for nonstrict inequalities such as $\mathcal{F}(x) + F_0 \succneqq 0$ and $\mathcal{F}(x) + F_0 \succeq 0$, and with additional linear equality constraints $\mathcal{G}(x) = 0$
- Typically get *weak* alternatives, need additional conditions (constraint qualifications) to make them strong

Application: Frequency-domain inequalities

Consider frequency domain inequality

$$\sigma_{\max}(H(j\omega)) < 1, \quad \text{where } H(s) = C(sI - A)^{-1}B$$

Application: Frequency-domain inequalities

Consider frequency domain inequality

$$\sigma_{\max}(H(j\omega)) < 1, \quad \text{where } H(s) = C(sI - A)^{-1}B$$

This, and similar inequalities ubiquitous in:

- Robust control
- Digital filter design
- Linear system analysis and design

Kalman-Yakubovich-Popov Lemma

$$\sigma_{\max}\left(H(j\omega)\right) < 1, \quad \text{for all } \omega$$

iff LMI is feasible:

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - \begin{bmatrix} -C^TC & 0 \\ 0 & I \end{bmatrix} \prec 0$$

Kalman-Yakubovich-Popov Lemma

$$\sigma_{\max}\left(H(j\omega)\right) < 1, \quad \text{for all } \omega$$

iff LMI is feasible:

$$\left[\begin{array}{cc} A^*P + PA & PB \\ B^*P & 0 \end{array}\right] - \left[\begin{array}{cc} -C^TC & 0 \\ 0 & I \end{array}\right] \prec 0$$

- Semi-infinite frequency domain inequality is exactly equivalent to LMI (no sampling)
- P serves as an auxiliary variable
- Of enormous importance for computation

Proof of KYP Lemma

Suppose LMI

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - \begin{bmatrix} -C^TC & 0 \\ 0 & I \end{bmatrix} \prec 0$$

is feasible

Then

$$0 \succ \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} A^*P + PA - C^TC & PB \\ B^*P & I \end{bmatrix} \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}$$
$$= I - H(j\omega)^* H(j\omega)$$

Proof of KYP Lemma: Converse

Proof: Using theorem of alternatives, infeasibility of LMI in P equivalent to

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \succneqq 0,$$
$$Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \mathbf{Tr}(Z_{22} - CZ_{11}C^T) \leq 0$$

Factor

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} \begin{bmatrix} U^* & V^* \end{bmatrix}$$

with U full rank. Then,

$$AUU^* + BVU^* = USU^*, \quad \mathbf{Tr}(V^*V - U^*C^TCU) \leq 0$$

with S skew-symmetric

$$AUU^* + BVU^* = USU^*, \quad \textbf{Tr}(V^*V - U^*C^TCU) \leq 0$$

As S is skew-symmetric, it has e.v.d. $S = \sum_i j\omega_i q_i q_i^*$, with $Q^*Q = I$

$$AUU^* + BVU^* = USU^*, \quad \textbf{Tr}(V^*V - U^*C^TCU) \leq 0$$

As S is skew-symmetric, it has e.v.d. $S = \sum_i j\omega_i q_i q_i^*$, with $Q^*Q = I$

From $\textbf{Tr}(V^*V - U^*C^TCU) \leq 0$, we have

$$\textbf{Tr}(V^*V - U^*C^TCU)QQ^* = \sum q_i^*(V^*V - U^*C^TCU)q_i \leq 0$$

so that

$$q_i^*(V^*V - U^*C^TCU)q_i \leq 0 \quad \text{for some } i$$

$$AUU^* + BVU^* = USU^*, \quad \textbf{Tr}(V^*V - U^*C^TCU) \le 0$$

As S is skew-symmetric, it has e.v.d. $S = \sum_i j\omega_i q_i q_i^*$, with $Q^*Q = I$

From $\textbf{Tr}(V^*V - U^*C^TCU) \le 0$, we have

$$\textbf{Tr}(V^*V - U^*C^TCU)QQ^* = \sum q_i^*(V^*V - U^*C^TCU)q_i \le 0$$

so that

$$q_i^*(V^*V - U^*C^TCU)q_i \le 0 \quad \text{for some } i$$

From $AUU^* + BVU^* = USU^*$, we have

$$AUq_i + BVq_i = j\omega_i Uq_i$$

Thus

$$q_i^* V^T(-B^T(-j\omega_i I - A^T)^{-1}C^T C(j\omega_i I - A)^{-1}B + I)Vq_i \le 0$$

or

$$q_i^* V^T(I - H(j\omega_i)^* H(j\omega_i))Vq_i \le 0,$$

i.e., frequency domain condition violated

Thus

$$q_i^* V^T(-B^T(-j\omega_i I - A^T)^{-1}C^T C(j\omega_i I - A)^{-1}B + I)Vq_i \leq 0$$

or

$$q_i^* V^T(I - H(j\omega_i)^* H(j\omega_i))Vq_i \leq 0,$$

i.e., frequency domain condition violated

- Usual proof of KYP Lemma based on optimal control theory
- This proof based only on SDP duality (no system theory)

Other results

- LMI "$A^*P + PA \prec 0$" is infeasible iff A has a pure imaginary eigenvalue
- "$P \succ 0$, $A^*P + PA \prec 0$" is infeasible iff $\lambda_i(A) \geq 0$ for some i
- "$A^*P + PA \precneqq 0$" is infeasible iff A is similar to a purely imaginary diagonal matrix
- "$A^*P + PA \preceq 0$, $P \succneqq 0$" is infeasible iff $\lambda_i(A) \geq 0$ for all i
- "$A^*P + PA \precneqq 0$, $PB = 0$" is infeasible iff all uncontrollable modes of (A, B) are nondefective and correspond to imaginary eigenvalues
- "$P \succneqq 0$, $A^*P + PA \preceq 0$, $PB = 0$" is infeasible iff all uncontrollable modes of (A, B) correspond to eigenvalues with positive real part
- "$P \neq 0$, $A^*P + PA \preceq 0$, $PB = 0$" is infeasible iff (A, B) is controllable

Other results

- LMI "$A^*P + PA \prec 0$" is infeasible iff A has a pure imaginary eigenvalue
- "$P \succ 0$, $A^*P + PA \prec 0$" is infeasible iff $\lambda_i(A) \geq 0$ for some i
- "$A^*P + PA \precneqq 0$" is infeasible iff A is similar to a purely imaginary diagonal matrix
- "$A^*P + PA \preceq 0$, $P \succneqq 0$" is infeasible iff $\lambda_i(A) \geq 0$ for all i
- "$A^*P + PA \precneqq 0$, $PB = 0$" is infeasible iff all uncontrollable modes of (A, B) are nondefective and correspond to imaginary eigenvalues
- "$P \succneqq 0$, $A^*P + PA \preceq 0$, $PB = 0$" is infeasible iff all uncontrollable modes of (A, B) correspond to eigenvalues with positive real part
- "$P \neq 0$, $A^*P + PA \preceq 0$, $PB = 0$" is infeasible iff (A, B) is controllable

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
  - ⋆ Theorems of alternatives, and the KYP Lemma
  - ⋆ Primal and Dual SDPs, and new bounds on the $\mathbf{H}_\infty$ norm
- Computational implications of SDP duality

SDP duality

Rewrite SDP as

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & F_0 + \mathcal{F}(x) - S = 0 \\ & S \succeq 0 \end{array}$$

SDP duality

Primal SDP minimize $\langle c, x \rangle$
subject to $F_0 + \mathcal{F}(x) - S = 0$
$S \succeq 0$

Dual SDP maximize $-\langle F_0, Z \rangle$
subject to $Z \succeq 0$
$\mathcal{F}^{\mathrm{adj}}(Z) = c$

SDP duality

Primal SDP minimize $\langle c, x\rangle$
subject to $F_0 + \mathcal{F}(x) - S = 0$
$S \succeq 0$

Dual SDP maximize $-\langle F_0, Z\rangle$
subject to $Z \succeq 0$
$\mathcal{F}^{\mathrm{adj}}(Z) = c$

- If Z is dual feasible, then $-\langle F_0, Z\rangle \leq p^*$
- If x is primal feasible, then $\langle c, x\rangle \geq d^*$
- Under mild conditions, $p^* = d^*$
- At optimum, $S_{\mathrm{opt}} Z_{\mathrm{opt}} = 0$

Application of duality: Bounds on $\mathbf{H}_\infty$ norm

Stable LTI system

$$\dot{x} = Ax + Bu, \quad x(0) = 0, \qquad y = Cx$$

- Transfer function $H(s) = C(sI - A)^{-1}B$
- $\mathbf{H}_\infty$ norm of H defined as

$$\|H\|_\infty = \sup_{\Re s > 0} \sigma_{\max}(H(s))$$

- $\|H\|_\infty^2$ equals maximum energy gain

$$\|H\|_\infty^2 = \max_u \frac{\int y^T y}{\int u^T u}$$

$\|H\|_\infty$ computation as an SDP

$$\begin{array}{ll} \texttt{minimize:} & \beta \\ \texttt{subject to:} & \left[\begin{array}{cc} A^*P + PA + C^*C & PB \\ B^*P & -\beta I \end{array}\right] \preceq 0 \end{array}$$

($\|H\|_\infty^2 = \beta_{\mathrm{opt}}$)

Dual problem

$$\begin{array}{ll} \texttt{maximize:} & \mathbf{Tr}\, C Z_{11} C^* \\ \texttt{subject to:} & Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \\ & \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{array}\right] \succeq 0, \quad \mathbf{Tr}\, Z_{22} = 1 \end{array}$$

New upper bounds on $\|H\|_\infty$

Recall primal problem:

$$\begin{array}{ll} \texttt{minimize:} & \beta \\ \texttt{subject to:} & \begin{bmatrix} A^*P + PA + C^*C & PB \\ B^*P & -\beta I \end{bmatrix} \preceq 0 \end{array}$$

A primal feasible point is

$$P = 2W_\text{o}, \quad \beta = 4\lambda_\text{max}(W_\text{o}BB^*W_\text{o}, C^*C)$$

where W_o is observability Gramian, obtained by solving $W_\text{o}A + A^*W_\text{o} + C^*C = 0$

Thus, new upper bound on $\|H\|_\infty$ is given by

$$2\sqrt{\lambda_\text{max}(W_\text{o}BB^*W_\text{o}, C^*C)}$$

New lower bounds on $\|H\|_\infty$

Recall dual problem

$$
\begin{array}{ll}
\texttt{maximize:} & \mathbf{Tr}\, C Z_{11} C^* \\
\texttt{subject to:} & Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \\
& \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{array}\right] \succeq 0, \quad \mathbf{Tr}\, Z_{22} = 1
\end{array}
$$

A dual feasible point is

$$
Z_{11} = W_{\rm c}/\alpha, \quad Z_{12} = B/(2\alpha), \quad Z_{22} = B^* W_{\rm c}^{-1} B/(4\alpha),
$$

where $\alpha = \mathbf{Tr}(B^* W_{\rm c}^{-1} B/4)$

Thus new lower bound is

$$
2\sqrt{\mathbf{Tr}\, C W_{\rm c} C^* / (\mathbf{Tr}\, B^* W_{\rm c}^{-1} B)}
$$

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
- Computational implications of SDP duality
  - ⋆ Primal-dual algorithms
  - ⋆ KYP SDPs
  - ⋆ Fast primal-dual algorithms for KYP SDPs

Primal and dual problems

Primal SDP minimize $\langle c, x\rangle$

subject to $F_0 + \mathcal{F}(x) - S = 0$

$S \succeq 0$

Dual SDP maximize $-\langle F_0, Z\rangle$

subject to $Z \succeq 0$

$\mathcal{F}^{\mathrm{adj}}(Z) = c$

(At optimum, $S_{\mathrm{opt}} Z_{\mathrm{opt}} = 0$)

Primal-dual algorithms

Solve primal and dual problem together:

$$
\begin{array}{ll}
\text{minimize} & c^T x + \textbf{Tr} F_0 Z \quad (= \langle S, Z \rangle) \\
\text{subject to} & F_0 + \mathcal{F}(x) - S = 0 \\
& S \succeq 0, \ Z \succeq 0 \\
& \mathcal{F}^{\mathrm{adj}}(Z) = c
\end{array}
$$

Primal-dual algorithms

Solve primal and dual problem together:

$$\begin{array}{ll} \text{minimize} & c^T x + \mathbf{Tr} F_0 Z \quad (= \langle S, Z \rangle) \\ \text{subject to} & F_0 + \mathcal{F}(x) - S = 0 \\ & S \succeq 0,\ Z \succeq 0 \\ & \mathcal{F}^{\mathrm{adj}}(Z) = c \end{array}$$

- At every iteration, we have upper and lower bounds, thus guaranteed accuracy
- Early termination possible
- Other advantages at algorithmic level

Primal-dual algorithm outline

For simplicity, suppose we have a feasible point, i.e., x, Z and S s.t.

$$F_0 + \mathcal{F}(x) - S = 0$$
$$S \succeq 0, Z \succeq 0$$
$$\mathcal{F}^{\mathrm{adj}}(Z) = c$$

(More general case, with infeasible starting points, essentially the same)

Primal-dual algorithm outline

For simplicity, suppose we have a feasible point, i.e., x, Z and S s.t.

$$\begin{gathered} F_0 + \mathcal{F}(x) - S = 0 \\ S \succeq 0, Z \succeq 0 \\ \mathcal{F}^{\mathrm{adj}}(Z) = c \end{gathered}$$

At each iteration:

- Compute product SZ. If it is "small", stop
- Otherwise, take steps ΔS, ΔZ, and Δx such that

$$\left.\begin{array}{l} \mathcal{F}(\Delta x) - \Delta S = 0 \\ \mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0 \\ S + \Delta S \succeq 0, \; Z + \Delta Z \succeq 0 \end{array}\right\} \quad \text{(maintain feasibility)}$$

Primal-dual algorithm outline

For simplicity, suppose we have a feasible point, i.e., x, Z and S s.t.

$$\begin{gathered} F_0 + \mathcal{F}(x) - S = 0 \\ S \succeq 0, Z \succeq 0 \\ \mathcal{F}^{\mathrm{adj}}(Z) = c \end{gathered}$$

At each iteration:

- Compute product SZ. If it is "small", stop
- Otherwise, take steps ΔS, ΔZ, and Δx such that

$$\left.\begin{array}{l} \mathcal{F}(\Delta x) - \Delta S = 0 \\ \mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0 \\ S + \Delta S \succeq 0,\ Z + \Delta Z \succeq 0 \end{array}\right\} \quad \text{(maintain feasibility)}$$

$(S + \Delta S)(Z + \Delta Z)$ is made "smaller" (address objective)

Solving search equations

1. $\mathcal{F}(\Delta x) - \Delta S = 0$
2. $\mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0$
3. $(S + \Delta S)(Z + \Delta Z)$ is made “smaller”
4. $S + \Delta S \succeq 0,\ Z + \Delta Z \succeq 0$

Solving search equations

1. $\mathcal{F}(\Delta x) - \Delta S = 0$
2. $\mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0$
3. $(S + \Delta S)(Z + \Delta Z)$ is made “smaller”
4. $S + \Delta S \succeq 0, \; Z + \Delta Z \succeq 0$

(1), (2) linear equations

Solving search equations

1. $\mathcal{F}(\Delta x) - \Delta S = 0$
2. $\mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0$
3. $(S + \Delta S)(Z + \Delta Z)$ is made “smaller”
4. $S + \Delta S \succeq 0,\ Z + \Delta Z \succeq 0$

(1), (2) linear equations

(3) accomplished via Newton step, another linear equation

Solving search equations

1. $\mathcal{F}(\Delta x) - \Delta S = 0$
2. $\mathcal{F}^{\mathrm{adj}}(\Delta Z) = 0$
3. $(S + \Delta S)(Z + \Delta Z)$ is made “smaller”
4. $S + \Delta S \succeq 0,\ Z + \Delta Z \succeq 0$

(1), (2) linear equations

(3) accomplished via Newton step, another linear equation

Solution strategy:

- First, eliminate ΔS from the linear equations
- Next eliminate ΔZ
- Solve a dense linear system in variable Δx
- Reconstruct ΔZ and ΔS
- $S + \Delta S \succeq 0,\ Z + \Delta Z \succeq 0$ ensured using line search

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
- Computational implications of SDP duality
  - ⋆ SDP duality
  - ⋆ Primal-dual algorithms
  - ⋆ KYP SDPs
  - ⋆ Fast primal-dual algorithms for KYP SDPs

KYP SDPs

Owing to importance of FDIs and KYP LMIs in engineering, focus on

$$\begin{array}{ll} \text{minimize} & c^T x + \mathbf{Tr}(CP) \\ \text{subject to} & \left[\begin{array}{cc} A^T P + PA & PB \\ B^T P & 0 \end{array}\right] + \sum_{i=1}^{p} x_i M_i \succeq N \end{array}$$

where $c \in \mathbf{R}^p, C \in \mathbf{S}^n, A \in \mathbf{R}^{n \times n}, B \in \mathbf{R}^{n \times 1}, M_i \in \mathbf{S}^{n+1}, N \in \mathbf{S}^{n+1}$

KYP SDPs

Owing to importance of FDIs and KYP LMIs in engineering, focus on

$$\begin{array}{ll} \text{minimize} & c^T x + \mathbf{Tr}(CP) \\ \text{subject to} & \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} + \sum_{i=1}^{p} x_i M_i \succeq N \end{array}$$

where $c \in \mathbf{R}^p, C \in \mathbf{S}^n, A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times 1}, M_i \in \mathbf{S}^{n+1}, N \in \mathbf{S}^{n+1}$

(Extension to multiple LMIs in multiple variables straightforward)

$$\begin{array}{ll} \text{minimize} & c^T x + \sum_{k=1}^{K} \mathbf{Tr}(C_k P_k) \\ \text{subject to} & \begin{bmatrix} A_k^T P_k + P_k A_k & P_k B_k \\ B_k^T P_k & 0 \end{bmatrix} + \sum_{i=1}^{p} x_i M_{ki} \succeq N_k, \quad k = 1, \ldots, K \end{array}$$

KYP SDPs

Owing to importance of FDIs and KYP LMIs in engineering, focus on

$$\begin{array}{ll} \text{minimize} & c^T x + \textbf{Tr}(CP) \\ \text{subject to} & \left[\begin{array}{cc} A^T P + PA & PB \\ B^T P & 0 \end{array}\right] + \sum_{i=1}^p x_i M_i \succeq N \end{array}$$

where $c \in \mathbf{R}^p, C \in \mathbf{S}^n, A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times 1}, M_i \in \mathbf{S}^{n+1}, N \in \mathbf{S}^{n+1}$

Define

$$\mathcal{K}(P) = \left[\begin{array}{cc} A^T P + PA & PB \\ B^T P & 0 \end{array}\right], \qquad \mathcal{M}(x) = \sum_{i=1}^p x_i M_i$$

Primal and Dual KYP SDPs

Primal SDP

minimize $c^T x + \mathbf{Tr}(CP)$

subject to $\mathcal{K}(P) + \mathcal{M}(x) \succeq N$

Primal and Dual KYP SDPs

Primal SDP

minimize $c^T x + \textbf{Tr}(CP)$

subject to $\mathcal{K}(P) + \mathcal{M}(x) \succeq N$

Dual SDP

maximize $-\textbf{Tr}(NZ)$

subject to $\mathcal{K}^{\mathrm{adj}}(Z) = C, \quad \mathcal{M}^{\mathrm{adj}}(Z) = c, \quad Z \succeq 0$

Primal and Dual KYP SDPs

Primal SDP

$$\begin{array}{ll} \text{minimize} & c^T x + \mathbf{Tr}(CP) \\ \text{subject to} & \mathcal{K}(P) + \mathcal{M}(x) \succeq N \end{array}$$

Dual SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{Tr}(NZ) \\ \text{subject to} & \mathcal{K}^{\text{adj}}(Z) = C, \quad \mathcal{M}^{\text{adj}}(Z) = c, \quad Z \succeq 0 \end{array}$$

where, partitioning

$$Z = \left[\begin{array}{cc} Z_{11} & \tilde{z} \\ \tilde{z}^T & 2z_{n+1} \end{array} \right],$$

and with $z = [\tilde{z}^T, z_{n+1}]^T$,

$$\mathcal{K}^{\text{adj}}(Z) = AZ_{11} + Z_{11}A^T + \tilde{z}B^T + B\tilde{z}^T, \quad \mathcal{M}^{\text{adj}}(Z) = \{\mathbf{Tr} M_i Z\}$$

Search equations for KYP SDPs

$$\begin{aligned} W\Delta Z W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{K}^{\mathrm{adj}}(\Delta Z) &= 0 \\ \mathcal{M}^{\mathrm{adj}}(\Delta Z) &= 0 \end{aligned}$$

$W \succ 0$; values of W, D change at each iteration

Search equations for KYP SDPs

$$\begin{aligned} W\Delta Z W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{K}^{\mathrm{adj}}(\Delta Z) &= 0 \\ \mathcal{M}^{\mathrm{adj}}(\Delta Z) &= 0 \end{aligned}$$

$W \succ 0$; values of W, D change at each iteration

General-purpose solvers eliminate ΔZ from first equation:

$$\begin{aligned} \mathcal{K}^{\mathrm{adj}}(W^{-1}(\mathcal{K}(\Delta P) + \mathcal{M}(\Delta x))W^{-1}) &= \mathcal{K}^{\mathrm{adj}}(W^{-1}DW^{-1}) \\ \mathcal{M}^{\mathrm{adj}}(W^{-1}(\mathcal{K}(\Delta P) + \mathcal{M}(\Delta x))W^{-1}) &= \mathcal{M}^{\mathrm{adj}}(W^{-1}DW^{-1}) \end{aligned}$$

A dense set of linear equations in ΔP, Δx

Cost: At least $O(n^6)$ (assuming $p = O(n)$). Prohibitive, as KYP SDPs tend to be of very large scale

Outline

- Introduction to SDP
- Theoretical implications of SDP duality
- Computational implications of SDP duality
  - ⋆ SDP duality
  - ⋆ Primal-dual algorithms
  - ⋆ KYP SDPs
  - ⋆ Fast primal-dual algorithms for KYP SDPs

Alternative method: Step 1, eliminate ΔZ_{11}

$$\begin{aligned} W\Delta Z W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{K}^{\mathrm{adj}}(\Delta Z) &= 0 \\ \mathcal{M}^{\mathrm{adj}}(\Delta Z) &= 0 \end{aligned}$$

Alternative method: Step 1, eliminate ΔZ_{11}

$$\begin{aligned} W\Delta Z W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ A\Delta Z_{11} + \Delta Z_{11} A^T + \Delta\tilde{z} B^T + B\Delta\tilde{z}^T &= 0 \\ \mathcal{M}^{\text{adj}}(\Delta Z) &= 0 \end{aligned}$$

Use second equation to express ΔZ_{11} in terms of $\Delta\tilde{z}$:

$$\Delta Z_{11} = \sum_{i=1}^n \tilde{\Delta} z_i X_i, \quad \text{where } AX_i + X_i A^T + Be_i^T + e_i B^T = 0$$

$$\text{Thus} \quad \Delta Z = \mathcal{L}(\tilde{\Delta} z) = \begin{bmatrix} \sum_{i=1}^n \Delta z_i X_i & \Delta\tilde{z} \\ \Delta\tilde{z}^T & 2\Delta z_{n+1} \end{bmatrix}$$

Alternative method: Step 1, eliminate ΔZ_{11}

$$\begin{aligned} W\Delta Z W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ A\Delta Z_{11} + \Delta Z_{11}A^T + \Delta\tilde{z}B^T + B\Delta\tilde{z}^T &= 0 \\ \mathcal{M}^{\mathrm{adj}}(\Delta Z) &= 0 \end{aligned}$$

Use second equation to express ΔZ_{11} in terms of $\Delta\tilde{z}$:

$$\Delta Z_{11} = \sum_{i=1}^n \tilde{\Delta} z_i X_i, \quad \text{where } AX_i + X_iA^T + Be_i^T + e_iB^T = 0$$

$$\text{Thus} \quad \Delta Z = \mathcal{L}(\tilde{\Delta} z) = \begin{bmatrix} \sum_{i=1}^n \Delta z_i X_i & \Delta\tilde{z} \\ \Delta\tilde{z}^T & 2\Delta z_{n+1} \end{bmatrix}$$

Substituting in first and third equations gives

$$\begin{aligned} W\mathcal{L}(\tilde{\Delta} z)W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{M}^{\mathrm{adj}}(\mathcal{L}(\tilde{\Delta} z)) &= 0 \end{aligned}$$

Alternative method: Step 2, eliminate ΔP

$$\begin{aligned} W\mathcal{L}(\tilde{\Delta} z)W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{M}^{\mathrm{adj}}(\mathcal{L}(\tilde{\Delta} z)) &= 0 \end{aligned}$$

Alternative method: Step 2, eliminate ΔP

$$\begin{aligned} W\mathcal{L}(\tilde{\Delta} z)W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{M}^{\mathrm{adj}}(\mathcal{L}(\tilde{\Delta} z)) &= 0 \end{aligned}$$

$G = \mathcal{K}(\Delta P)$ for some $\Delta P \quad \Longleftrightarrow \quad \mathcal{L}^{\mathrm{adj}}(G) = 0$

Alternative method: Step 2, eliminate ΔP

$$\begin{aligned} W\mathcal{L}(\tilde{\Delta} z)W + \mathcal{K}(\Delta P) + \mathcal{M}(\Delta x) &= D \\ \mathcal{M}^{\text{adj}}(\mathcal{L}(\tilde{\Delta} z)) &= 0 \end{aligned}$$

$$G = \mathcal{K}(\Delta P) \text{ for some } \Delta P \quad \Longleftrightarrow \quad \mathcal{L}^{\text{adj}}(G) = 0$$

Use to eliminate ΔP:

$$\begin{aligned} \mathcal{L}^{\text{adj}}(W\mathcal{L}(\tilde{\Delta} z)W) + \mathcal{L}^{\text{adj}}(\mathcal{M}(\Delta x)) &= \mathcal{L}^{\text{adj}}(D) \\ \mathcal{M}^{\text{adj}}(\mathcal{L}(\tilde{\Delta} z)) &= 0 \end{aligned}$$

$n+p+1$ linear equations in $n+p+1$ variables $\tilde{\Delta} z$, Δx

Alternative method: Summary

Reduced search equations of the form

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Delta} z \\ \Delta x \end{bmatrix} = \begin{bmatrix} q_1 \\ 0 \end{bmatrix}$$

- Cost of solving is $O(n^3)$ operations (if we assume $p = O(n)$)
- From $\tilde{\Delta} z$, Δx, can find ΔZ, ΔP in $O(n^3)$ operations
- Need to pre-compute X_is ($O(n^4)$)
- P_{12} is independent of current iterates and can be pre-computed, in $O(n^4)$
- Constructing P_{11} requires constructing terms such as **Tr**$(X_i W_{11} X_j W_{11})$ and $W_{11} X_i W_{12}$ (also $O(n^4)$)
- **Overall cost dominated by** $O(n^4)$

Numerical example

$n = p$	KYP IPM		SeDuMi (primal)
	prep. time	time/iter.	time/iter.
25	0.1	0.04	0.3
50	1.1	0.3	8.1
100	21.4	3.3	307.1
200	390.7	30.9	

- CPU time in seconds on 2.4GHz PIV with 1GB of memory
- KYP-IPM: Matlab implementation of alternative method
- SeDuMi (primal): SeDuMi version 1.05 applied to primal problem
- Prep. time is time to compute matrices X_i
- #iterations in both methods is comparable (7–15)

Further reduction in computation

Use factorization of A to compute terms such as $\mathbf{Tr}(X_i W_{11} X_j W_{11})$ without computing X_i, i.e., without explicitly solving

$$AX_i + X_i A^T + B e_i^T + e_i B^T = 0, \quad i = 1, \ldots, n$$

- Advantages: no need to store matrices X_i, faster construction of reduced search equations
- Possible factorizations: eigenvalue decomposition, companion form, . . .
- For example, if A has distinct eigenvalues $A = V\,\mathbf{diag}(\lambda) V^{-1}$, easy to write down search equations in $O(n^3)$, in terms of V and λ

Existence of distinct stable eigenvalues

- If (A, B) is controllable, can arbitrarily assign eigenvalues of $A + BK$ by choosing K
- Choose $T = \begin{bmatrix} I & 0 \\ K & I \end{bmatrix}$, and replace LMI by equivalent LMI

$$T^T \left(\begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} + \sum_{i=1}^{N} x_i M_i \right) T \succeq T^T N T$$

$$\begin{bmatrix} (A + BK)^T P + P(A + BK) & PB \\ B^T P & 0 \end{bmatrix} + \sum_{i=1}^{N} x_i (T^T M_i T) \succeq T^T N T$$

Conclusion: Can assume without loss of generality that A is stable with distinct eigenvalues

Numerical example

Five randomly generated problems with $p = 50$, $n = 100, \ldots, 500$

n	KYP IPM (fast)		KYP IPM		SeDuMi (primal)	
	prep. time	time/iter	prep. time	time/iter	prep. time	time/iter
100	1.0	0.6	21.4	3.3	0	324.7
200	8.3	4.7	390.7	30.9		
300	28.1	16.7				
400	62.3	36.2				
500	122.0	70.3				

- KYP-IPM (fast) uses eigenvalue decomposition of A to construct reduced search equations
- Preprocessing time and time/iteration grow as $O(n^3)$

Conclusions

SDP duality theory has implications for:

- System theory: New proofs, results
- Fast implementation of interior-point methods (cost $O(n^4)$ or $O(n^3)$, as compared to $O(n^6)$ with standard methods)