Nonnegative polynomials, SDP formulations, and primal-dual interior-point methods

Lieven Vandenberghe

Electrical Engineering Department, UCLA

CDC 2003

Outline

- SDP representation of nonnegative (trigonometric) polynomials
- primal-dual interior-point methods for SDP

Outline

- SDP representation of nonnegative (trigonometric) polynomials
- primal-dual interior-point methods for SDP

Nonnegative trigonometric polynomials

$$X(\omega) = x_0 + 2x_1 \cos\omega + \cdots + 2x_n \cos n\omega \geq 0, \quad \omega \in [0, \pi]$$

- an infinite set of linear inequalities in $x \in \mathbf{R}^{n+1}$
- defines a closed convex cone $C_{n+1} = \{x \mid X(\omega) \geq 0\}$

spectral factorization (Riesz-Fejér theorem)

$x \in C_{n+1}$ if and only if there exist $y_0, \ldots, y_n \in \mathbf{R}$ such that

$$X(\omega) = \left| y_0 + y_1 e^{-j\omega} + y_2 e^{-2j\omega} + \cdots + y_n e^{-nj\omega} \right|^2$$

LMI characterization (equality form)

$x \in C_{n+1}$ if and only if there exists

$$Y = \begin{bmatrix} Y_{00} & Y_{10} & \cdots & Y_{n0} \\ Y_{10} & Y_{11} & \cdots & Y_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n0} & Y_{n1} & \cdots & Y_{nn} \end{bmatrix} \succeq 0$$

such that

$$\begin{aligned} x_0 &= Y_{00} + Y_{11} + \cdots + Y_{nn} \\ x_1 &= Y_{10} + Y_{21} + \cdots + Y_{n-1,n} \\ &\vdots \\ x_n &= Y_{n0} \end{aligned}$$

$i.e.$, $x_k = \mathbf{Tr}(E^k Y)$ where $E = \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix}$

proof

- if $x_k = \mathbf{Tr}(E^k Y)$ with $Y \succeq 0$, then

$$X(\omega) = \begin{bmatrix} 1 \\ e^{j\omega} \\ \vdots \\ e^{nj\omega} \end{bmatrix}^H \begin{bmatrix} Y_{00} & Y_{10} & \cdots & Y_{n0} \\ Y_{10} & Y_{11} & \cdots & Y_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n0} & Y_{n1} & \cdots & Y_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ e^{j\omega} \\ \vdots \\ e^{nj\omega} \end{bmatrix} \geq 0$$

- if $X(\omega) \geq 0$, expanding $X(\omega) = |y_0 + y_1 e^{-j\omega} + \cdots + y_n e^{-nj\omega}|^2$ gives

$$\begin{aligned} x_0 &= y_0^2 + y_1^2 + \cdots + y_n^2 \\ x_1 &= y_0 y_1 + y_1 y_2 + \cdots + y_{n-1} y_n \\ &\vdots \\ x_n &= y_0 y_n \end{aligned}$$

$i.e.$, $x_k = y^T E^k y = \mathbf{Tr}(E^k y y^T)$

LMI characterization (inequality form)

$$
\begin{aligned}
x_0 &= Y_{00} + Y_{11} + \cdots + Y_{nn} \\
x_1 &= Y_{10} + Y_{21} + \cdots + Y_{n-1,n} \\
&\vdots \\
x_n &= Y_{n0}
\end{aligned}
$$

if and only if there exists exists $P \in \mathbf{S}^n$ such that

$$
Y(x,P) = \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & x_n \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & x_1 \\ x_n & \cdots & x_1 & x_0 \end{bmatrix}
$$

therefore, $x \in C_{n+1}$ if and only if $Y(x,P) \succeq 0$ for some $P \in \mathbf{S}^n$

Dual cone

$$C_{n+1}^* = \{z \mid z^T x \geq 0 \text{ for all } x \in C_{n+1}\}$$

LMI characterization: $z \in C_{n+1}^*$ if and only if

$$Z = \begin{bmatrix} 2z_0 & z_1 & z_2 & \cdots & z_n \\ z_1 & 2z_0 & z_1 & \cdots & z_{n-1} \\ z_2 & z_1 & 2z_0 & \cdots & z_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_n & z_{n-1} & z_{n-2} & \cdots & 2z_0 \end{bmatrix} \succeq 0$$

proof: $z^T x \geq 0$ for all $x \in C_{n+1}$ if for all y,

$$\sum_{k=0}^{n} z_k (y^T E^k y) = \frac{1}{2} y^T Z y \geq 0$$

Nonnegative real polynomials on $[-1,1]$

change of variables $t = \cos\omega$ maps $[0,\pi]$ to $[-1,1]$ and $X(\omega)$ to

$$Q(t) = X(\cos^{-1} t) = x_0 p_0(t) + 2x_1 p_1(t) + \cdots + 2x_n p_n(t)$$

where $p_k(t) = \cos(k\cos^{-1} t)$ (the kth Chebyshev polynomial)

SOS representation (for $n = 2m$): polynomial $Q(t) \geq 0$ on $[-1,1]$ iff

$$\begin{aligned} Q(t) &= \left. \left| y_0 + y_1 e^{-j\omega} + y_2 e^{-2j\omega} + \cdots + y_n e^{-nj\omega} \right|^2 \right|_{\omega = \cos^{-1} t} \\ &= \left(y_m p_0(t) + (y_{m-1} + y_{m+1}) p_1(t) + \cdots + \cdots + (y_0 + y_n) p_m(t) \right)^2 \\ &\quad + (1 - t^2) \left((y_{m-1} - y_{m+1}) q_0(t) + \cdots + (y_0 - y_n) q_{n-1}(t) \right)^2 \end{aligned}$$

where $q_{k-1}(t) = \sin(k\cos^{-1} t) / \sin(\cos^{-1} t)$ (Cheb. polyn. of 2nd kind)

Extension to arbitrary bounded intervals

$$Q(t) = x_0 p_0(t) + 2x_1 p_1(t) + \cdots + 2x_n p_n(t) \geq 0, \quad t \in [t_1, t_2]$$

$$\Updownarrow$$

$$\exists P \in \mathbf{S}^n : \quad \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & a_n^T x \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_1^T x \\ a_n^T x & \cdots & a_1^T x & a_0^T x \end{bmatrix} \succeq 0$$

$a_i^T x$ are coefficients of

$$Q(t) = (a_0^T x)\tilde{p}_0(t) + (a_1^T x)\tilde{p}_1(t) + \cdots + (a_n^T x)\tilde{p}_n(t)$$

in basis of shifted Chebyshev polynomials

$$\tilde{p}_k(t) = p_k((2t - (t_1 + t_2))/(t_2 - t_1))$$

Example: Magnitude FIR filter design

magnitude constraints

$$L \leq |h_0 + h_1 e^{-j\omega} + \cdots + h_n e^{-nj\omega}| \leq U, \quad \omega \in [\omega_1, \omega_2]$$

- $H(\omega) = \sum_k h_k e^{-kj\omega}$ is frequency response of FIR filter
- not convex in filter coefficients h_i

change of variables $x_k = \sum_{i=0}^{n-k} h_i h_{i+k}$ gives equivalent constraints

$$L^2 \leq x_0 + 2x_1 \cos\omega + \cdots + 2x_n \cos n\omega \leq U^2, \quad \omega \in [\omega_1, \omega_2]$$

$$x_0 + 2x_1 \cos\omega + \cdots + 2x_n \cos n\omega \geq 0, \quad \omega \in [0, \pi]$$

- from x, obtain h by spectral factorization
- convex constraints in x, representable as LMIs

LMI formulation (auxiliary variables $P_1, P_2, P_3 \in \mathbf{S}^n$)

- $X(\omega) \geq 0$, $\omega \in [0, \pi]$:

$$\begin{bmatrix} 0 & 0 \\ 0 & P_1 \end{bmatrix} - \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & x_n \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & x_1 \\ x_n & \cdots & x_1 & x_0 \end{bmatrix} \succeq 0$$

- $L^2 \leq X(\omega) \leq U^2$, $\omega \in [\omega_1, \omega_2]$:

$$\begin{bmatrix} 0 & 0 \\ 0 & P_2 \end{bmatrix} - \begin{bmatrix} P_2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & a_n^T x \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_1^T x \\ a_n^T x & \cdots & a_1^T x & a_0^T x - L^2 \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} 0 & 0 \\ 0 & P_3 \end{bmatrix} - \begin{bmatrix} P_3 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & -a_n^T x \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -a_1^T x \\ -a_n^T x & \cdots & -a_1^T x & U^2 - a_0^T x \end{bmatrix} \succeq 0$$

example: peak-constrained least-squares filter with N bands $[\alpha_k, \beta_k]$

$$\begin{array}{ll} \text{minimize} & \int_{\text{stopbands}} |H(\omega)|^2 \, d\omega \\ \text{subject to} & L_k \leq |H(\omega)| \leq U_k, \quad \omega \in [\alpha_k, \beta_k], \quad k = 1, \ldots, N \end{array}$$

use $X(\omega) = |H(\omega)|^2 = x_0 + 2x_1 \cos\omega + \cdots + 2x_n \cos n\omega$, to get

$$\begin{array}{ll} \text{minimize} & \int_{\text{stopbands}} X(\omega) \, d\omega \\ \text{subject to} & L_k^2 \leq X(\omega) \leq U_k^2, \quad \omega \in [\alpha_k, \beta_k], \quad k = 1, \ldots, N \\ & X(\omega) \geq 0, \quad \omega \in [0, \pi] \end{array}$$

- standard method: discretize constraints and solve an LP
- SDP method: solve SDP with variables x, $2N + 1$ matrices $P_k \in \mathbf{S}^n$

general problem

$$
\begin{array}{ll}
\text{minimize} & q^T x \\
\text{subject to} & \left[\begin{array}{cc} 0 & 0 \\ 0 & P_k \end{array}\right] - \left[\begin{array}{cc} P_k & 0 \\ 0 & 0 \end{array}\right] + \sum_{i=0}^{p} x_i M_{ki} \succeq N_k, \quad k = 1, \ldots, L
\end{array}
$$

- variables $x \in \mathbf{R}^p$, $P_k \in \mathbf{S}^n$
- P_k are auxiliary variables, introduced to formulate semi-infinite inequalities as LMIs
- expensive to solve using general-purpose SDP software

Outline

- SDP representation of nonnegative (trigonometric) polynomials
- primal-dual interior-point methods for SDP

Interior-point SDP methods

early methods (1990-1995)

- projective method, implemented in the Matlab LMI control toolbox
- potential reduction methods, implemented in SDPSOL
- barrier methods

more recent methods

- primal-dual path-following methods
- general-purpose software packages: Sedumi, SDPT3, SDPPACK, SDPA, CSDP, DSDP, Yalmip, . . .

SDP duality

primal SDP

$$\begin{array}{ll} \text{minimize} & \langle c, y \rangle \\ \text{subject to} & \mathcal{A}(y) + S = B, \quad S \succeq 0 \end{array}$$

- variable $x \in \mathcal{V}$, $S \in \mathbf{S}^n$ (slack variable)
- $\mathcal{A}$ is linear mapping from $\mathcal{V}$ to $\mathbf{S}^n$

dual SDP

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

variable $Z \in \mathbf{S}^n$; $\mathcal{A}^{\text{adj}} : \mathbf{S}^n \to \mathcal{V}$ is adjoint of $\mathcal{A}$

optimal values are equal (if primal or dual is strictly feasible)

Primal-dual path-following algorithm

(Tütüncü, Toh, Todd)

select starting point $S \succ 0$, $Z \succ 0$, any y; repeat the following steps

1. *Verify stopping criteria.*
2. *Compute the Nesterov-Todd scaling matrix* R: defined by

$$R^T S^{-1} R = \mathbf{diag}(\lambda)^{-1}, \qquad R^T Z R = \mathbf{diag}(\lambda), \qquad \lambda \in \mathbf{R}^n_{++}$$

3. *Compute affine scaling directions* ΔZ^{a}, ΔS^{a}, Δy^{a}:

$$\begin{aligned} \mathcal{H}(\Delta Z^{\mathsf{a}} S + Z \Delta S^{\mathsf{a}}) &= -\mathbf{diag}(\lambda)^2 \\ \Delta S^{\mathsf{a}} + \mathcal{A}(\Delta y^{\mathsf{a}}) &= -(\mathcal{A}(y) + S - B) \\ \mathcal{A}^{\mathsf{adj}}(\Delta Z^{\mathsf{a}}) &= -(\mathcal{A}^{\mathsf{adj}}(Z) + c) \end{aligned}$$

where $\mathcal{H}(X) = \frac{1}{2}(R^T X R^{-T} + R^{-1} X^T R)$

4. *Compute* *centering-corrector steps* ΔZ^{c}, ΔS^{c}, Δy^{c}:

$$\begin{aligned} \mathcal{H}(\Delta Z^{\mathsf{c}} S + Z \Delta S^{\mathsf{c}}) &= \rho I - \mathcal{H}(\Delta Z^{\mathsf{a}} \Delta S^{\mathsf{a}}) \\ \Delta S^{\mathsf{c}} + \mathcal{A}(\Delta y^{\mathsf{c}}) &= 0 \\ \mathcal{A}^{\mathsf{adj}}(\Delta Z^{\mathsf{c}}) &= 0 \end{aligned}$$

with ρ calculated based on $\mathbf{Tr}(SZ)$, ΔZ^{a}, ΔS^{a}

5. *Update primal and dual iterates:*

$$y := y + \alpha \Delta y, \qquad S := S + \alpha \Delta S, \qquad Z := Z + \beta \Delta Z$$

where $\Delta y = \Delta y^{\mathsf{a}} + \Delta y^{\mathsf{c}}$, $\Delta S = \Delta S^{\mathsf{a}} + \Delta S^{\mathsf{c}}$, $\Delta Z = \Delta Z^{\mathsf{a}} + \Delta Z^{\mathsf{c}}$,

$$\begin{aligned} \alpha &= \min\{1, 0.99 \sup\{\alpha \mid S + \alpha \Delta S \succeq 0\}\} \\ \beta &= \min\{1, 0.99 \sup\{\beta \mid Z + \beta \Delta Z \succeq 0\}\} \end{aligned}$$

Overall complexity

- number of iterations is small (< 30)
- at each iteration, solve two sets of equations ('Newton equations')

$$\begin{aligned} \mathcal{H}(\Delta Z S + Z \Delta S) &= D_1 \\ \Delta S + \mathcal{A}(\Delta y) &= D_2 \\ \mathcal{A}^{\mathsf{adj}}(\Delta Z) &= d \end{aligned}$$

  where
  $$\mathcal{H}(X) = \frac{1}{2}(R^T X R^{-T} + R^{-1} X^T R)$$

  values of R (NT scaling matrix), D_1, D_2, d change at each iteration

- equations for other primal-dual methods are similar (with different R)

General-purpose implementation

- eliminate ΔS from $\mathcal{H}(\Delta Z S + Z \Delta S) = D_1$:

$$\begin{aligned} -W \Delta Z W + \mathcal{A}(\Delta y) &= D \qquad (1) \\ \mathcal{A}^{\mathrm{adj}}(\Delta Z) &= d \end{aligned}$$

  where $W = RR^T$

- eliminate ΔZ from (1):

$$\mathcal{A}^{\mathrm{adj}}(W^{-1}\mathcal{A}(\Delta y)W^{-1}) = d + \mathcal{A}^{\mathrm{adj}}(W^{-1}DW^{-1}) \qquad (2)$$

  a positive definite set of linear equations in Δy, and *usually dense*

total cost: cost of forming the equations (2) plus cost of solving

SDP with structure

$$\begin{array}{ll} \text{minimize} & q^T x \\ \text{subject to} & \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \sum_{i=1}^p x_i M_i \succeq N \end{array}$$

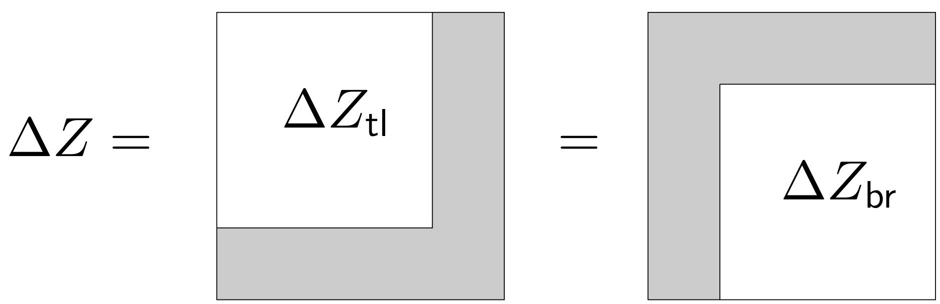
- $p + n(n+1)/2$ variables x, P
- we will assume that $p = O(n)$
- discussion extends to problems with multiple constraints

$$\begin{bmatrix} 0 & 0 \\ 0 & P_k \end{bmatrix} - \begin{bmatrix} P_k & 0 \\ 0 & 0 \end{bmatrix} + \sum_{i=1}^p x_i M_{ki} \succeq N_k, \quad k = 1, \ldots, L$$

Newton equations

$$
\begin{aligned}
W\Delta Z W + \begin{bmatrix} 0 & 0 \\ 0 & \Delta P \end{bmatrix} - \begin{bmatrix} \Delta P & 0 \\ 0 & 0 \end{bmatrix} + \sum_{i=1}^{p} \Delta x_i M_i &= D_1 \\
\Delta Z_{\mathsf{br}} - \Delta Z_{\mathsf{tl}} &= D_2 \\
\mathbf{Tr}(M_i \Delta Z) &= d_i, \quad i = 1, \ldots, p
\end{aligned}
$$

where ΔZ_{tl}, ΔZ_{br} are leading and trailing $n \times n$ submatrices of ΔZ



$W \succ 0$; value changes at each iteration

Standard solution method

$$
\begin{array}{rcl}
W\Delta Z W + \left[\begin{array}{cc} 0 & 0 \\ 0 & \Delta P \end{array}\right] - \left[\begin{array}{cc} \Delta P & 0 \\ 0 & 0 \end{array}\right] + \sum_{i=1}^{p} \Delta x_i M_i & = & D_1 \\
\Delta Z_{\mathsf{br}} - \Delta Z_{\mathsf{tl}} & = & D_2 \\
\mathbf{Tr}(M_i \Delta Z) & = & d_i, \quad i = 1, \ldots, p
\end{array}
$$

- eliminate ΔZ from 1st equation
- solve dense set of equations in Δx, ΔP
- cost: cost of forming reduced equations plus at least $O(n^6)$ for solving

used in general-purpose solvers

Alternative method for solving Newton equations

first equation

$$W\Delta Z W + \begin{bmatrix} 0 & 0 \\ 0 & \Delta P \end{bmatrix} - \begin{bmatrix} \Delta P & 0 \\ 0 & 0 \end{bmatrix} + \sum_{i=1}^{p} \Delta x_i M_i = D_1$$

eliminate ΔP by taking inner product with E^k, $E = \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix}$:

$$\mathbf{Tr}(E^k W \Delta Z W) + \sum_{i=1}^{p} \Delta x_i \, \mathbf{Tr}(E^k M_i) = \mathbf{Tr}(E^k D_1), \quad k = 0, \ldots, n$$

(note: $\mathbf{Tr}(E^k X)$ is sum of elements on kth diagonal of X)

second equation

$$\Delta Z_{\mathsf{br}} - \Delta Z_{\mathsf{tl}} = D_2$$

means

$$\Delta Z = \mathbf{T}(\Delta u) + Z_0 \text{ for some } \Delta u$$

- $\mathbf{T}(\Delta u)$ is symmetric Toeplitz matrix constructed from Δu

$$\begin{aligned} \mathbf{T}(\Delta u) &= \begin{bmatrix} 2\Delta u_0 & \Delta u_1 & \cdots & \Delta u_n \\ \Delta u_1 & 2\Delta u_0 & \cdots & \Delta u_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta u_n & \Delta u_{n-1} & \cdots & 2\Delta u_0 \end{bmatrix} \\ &= \sum_{k=0}^{n} \Delta u_k (E^k + (E^k)^T) \end{aligned}$$

- Z_0 is any solution to $Z_{0,\mathsf{br}} - Z_{0,\mathsf{tl}} = D_2$

reduced Newton equations

$$\begin{array}{rcl}\mathbf{Tr}(E^kW\mathbf{T}(\Delta u)W) + \sum_{i=1}^{p} \Delta x_i\, \mathbf{Tr}(E^kM_i) & = & \mathbf{Tr}(E^kD_1), \quad k=0,\ldots,n \\ \mathbf{Tr}(M_i\mathbf{T}(\Delta u)) & = & d_i, \quad i=1,\ldots,p\end{array}$$

in matrix notation:

$$\left[\begin{array}{cc} H & G \\ G^T & 0 \end{array}\right] \left[\begin{array}{c} \Delta u \\ \Delta x \end{array}\right] = \left[\begin{array}{c} r \\ d \end{array}\right]$$

$$H_{ki} = \mathbf{Tr}\left(E^kW\left(E^i + (E^i)^T\right)W\right), \qquad G_{ki} = \mathbf{Tr}(E^kM_i)$$

total cost: $O(n^3)$ operations for solving, plus

- cost of forming G (can be pre-computed in at most $O(n^3)$, usually less)
- cost of forming H

fast evaluation of H via DFT

$$H_{ki} = \mathbf{Tr}(E^k W E^i W) + \mathbf{Tr}(E^k W (E^i)^T W), \quad i,k = 0,\ldots,n$$

- factor $W = RR^T = \sum_{l=0}^{n} r_l r_l^T$
- take zero-padded (length $\geq 2(n+1)$) DFTs of r_l:

$$V = W_{\mathsf{DFT}} R$$

- evaluate H using Hadamard products:

$$H = W_{\mathsf{DFT}}^H \left((VV^H) \circ (VV^H)^T + (VV^T) \circ (VV^T)^H \right) W_{\mathsf{DFT}}$$

cost: $O(n^3)$

Summary

SDP formulation of a class of problems involving nonnegative polynomials:

- difficult to solve using general-purpose software
  - large number of auxiliary variables ($O(n^2)$)
  - complexity typically $O(n^6)$ per iteration
- custom implementation of primal-dual interior-point method:
  - exploit structure in Newton equations using direct linear algebra
  - cost: $O(n^3)$ per iteration (times 20-30 iterations)

References

- Genin, Hachez, Nesterov, Van Dooren (SIMAX 2003)

  nonnegative generalized matrix polynomials; fast implementations of dual barrier method using generalized Schur algorithm

- Alkire, Vandenberghe (Mathematical Programming 2002)

  trigonometric polynomials; fast implementation of dual barrier method using DFT

- Vandenberghe, Balakrishnan, Wallin, Hansson (this CDC)

  extension to primal-dual methods, SDPs derived from KYP lemma