# EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
# Lecture 9

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

May 1, 2012

Caltech

# Outline

# Monomials, Posynomials, and Geometric Programs

**Monomial Functions:**
A function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(\mathbf{x}) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \, , \; \mathrm{dom}(f) = \mathbb{R}^n_{++} \, ,$$

where $c > 0$ and $a_i \in \mathbb{R}$ for $i = 1, \ldots, n$ is called a *monomial function* or *monomial*. For example, $f(x_1, x_2) = 1.7 x_1^{-0.3} x_2^{\sqrt{\pi}}$ is a monomial of two variables.

**Posynomial Functions:**
A sum of monomials, i.e., a function of the form

$$f(\mathbf{x}) = \sum_{k=1}^{K} c_k x_1^{a_{1,k}} x_2^{a_{2,k}} \cdots x_n^{a_{n,k}} \, , \; \mathrm{dom}(f) = \mathbb{R}^n_{++} \, ,$$

is called a *posynomial function* (with $K$ terms) or *posynomial*. For example, $f(x_1, x_2) = 1.2 x_1^{-3.7} x_2^{e} + \pi^2 x_1^5 x_2^{-1.9}$ is a posynomial.

**Geometric Program (GP):**
An optimization problem of the form

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \leq 1 \, , \; i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 1 \, , \; i = 1, \ldots, p \end{array} \; ,$$

where $f_0, f_1, \ldots, f_m$ are *posynomials* and $h_1, \ldots, h_p$ are *monomials* is called a *geometric program* (GP). Here, the domain of the problem is $\mathcal{D} = \mathbb{R}^n_{++}$ and the constraint $\mathbf{x} \succ \mathbf{0}$ is implicit. Note that this form of the GP as stated here is *not* a convex optimization problem.

# Geometric Program in Convex Form

With a change of variables and a transformation of the objective and constraint functions, a GP can be converted to a convex optimization problem. For this, we change variables to $y_i = \log x_i$ (so that $x_i = e^{y_i}$) and take the logarithm of the objective and constraint functions.

- **Monomial transformation:**

$$f(\mathbf{x}) = c x_1^{a_1} \cdots x_n^{a_n} \to \log f(e^{y_1}, \ldots, e^{y_n}) = \mathbf{a}^T \mathbf{y} + b \,,$$

  where $b = \log c$.

- **Posynomial transformation:**

$$f(\mathbf{x}) = \sum_{k=1}^{K} c_k x_1^{a_{1,k}} \cdots x_n^{a_{n,k}} \to \log f(e^{y_1}, \ldots, e^{y_n}) = \log \left( \sum_{k=1}^{K} e^{\mathbf{a}_k^T \mathbf{y} + b_k} \right) \,,$$

  where $b_k = \log c_k$ for $k = 1, \ldots, K$.

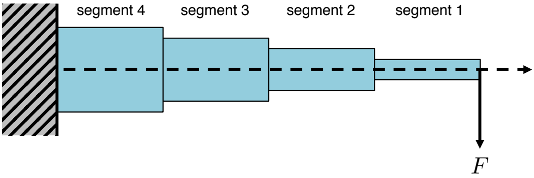
- **GP transformation:**
  The geometric program transforms to the following convex form.

$$\begin{array}{ll} \text{minimize} & \widetilde{f}_0(\mathbf{y}) = \log\left(\sum_{k=1}^{K_0} e^{\mathbf{a}_{0,k}^T \mathbf{y} + b_{0,k}}\right) \\ \text{subject to} & \widetilde{f}_i(\mathbf{y}) = \log\left(\sum_{k=1}^{K_i} e^{\mathbf{a}_{i,k}^T \mathbf{y} + b_{i,k}}\right) \le 0 \,, \; i = 1, \ldots, m \\ & \widetilde{h}_i(\mathbf{y}) = \mathbf{g}_i^T \mathbf{y} + d_i = 0 \,, \; i = 1, \ldots, p \end{array} ,$$

  Here, the functions $\widetilde{f}_i$ are convex and $\widetilde{h}_i$ are affine, as desired. We refer to this as a *geometric program in convex form*, as opposed the original GP formulation, which we refer to as a *geometric program in posynomial form*.

# Design of a Cantilever Beam

It can be shown that the design of a *cantilever*, which is a beam anchored only at one end, can be posed as a GP.




- The beam consists of $N$ segments with unit lengths and rectangular cross-sections of size $w_i \times h_i$ for $i = 1, \ldots, N$.
- A given vertical force $F$ is applied at the right end of the beam.

**Design Problem:**

| | |
|---|---|
| minimize | total weight |
| subject to | upper & lower bounds on $w_i$ and $h_i$ |
| | upper & lower bounds on aspect ratios $w_i/h_i$ |
| | upper bound on stress in each segment |
| | upper bound on vertical deflection at the end of the beam |

,

with variables $w_i, h_i$ for $i = 1, \ldots, N$.

Under the assumption that the deflections are small and that the material is linearly elastic, we can pose this design problem as a GP.

# Design of a Cantilever Beam (Continued)

**Objective Function:**

- The total weight is propotional to the total volume, which is given by

$$w_1 h_1 + \cdots + w_N h_N \,,$$

assuming unit lengths for all the segments. Note that this is a posynomial in $\mathbf{w}, \mathbf{h}$.

**Constraint Functions:**

- The aspect ratio $w_i/h_i$ and inverse aspect ratio $h_i/w_i$ are monomials in $\mathbf{w}, \mathbf{h}$.
- The maximum stress in the $i$-th segment is given by $6iF/\left(w_i h_i^2\right)$, which is a monomial in $\mathbf{w}, \mathbf{h}$.
- The vertical deflection $y_i$ and slope $v_i$ of the central axis at the right end of the $i$-th segment is given recursively as

$$\begin{aligned} v_i &= 12\left(i - 1/2\right) \frac{F}{E w_i h_i^3} + v_{i+1} \\ y_i &= 6\left(i - 1/3\right) \frac{F}{E w_i h_i^3} + v_{i+1} + y_{i+1} \end{aligned} \,,$$

for $i = N, N-1, \ldots, 1$, with the boundary conditions $v_{N+1} = y_{N+1} = 0$. Here, $E > 0$ is a constant which is a measure of the stiffness of the elastic material used to make the beam known as *Young's modulus*. Note that $v_i$ and $y_i$ are posynomial functions of $\mathbf{w}, \mathbf{h}$.

# Design of a Cantilever Beam (Continued)

**Design Problem:**

$$\begin{array}{ll} \text{minimize} & w_1 h_1 + \cdots + w_N h_N \\ \text{subject to} & w_{\min} \leq w_i \leq w_{\max}\,,\ h_{\min} \leq h_i \leq h_{\max}\,,\ i = 1, \ldots, N \\ & S_{\min} \leq w_i / h_i \leq S_{\max}\,,\ i = 1, \ldots, N \\ & 6iF / \left(w_i h_i^2\right) \leq \sigma_{\max}\,,\ i = 1, \ldots, N \\ & y_1 \leq y_{\max} \end{array} .$$

Here, $w_{\min}$, $w_{\max}$, $h_{\min}$ and $h_{\max}$ denote the minimum and maximum allowable width and height of each segment, respectively, $S_{\min}$ and $S_{\max}$ denote the minimum and maximum allowable aspect ratio for each segment, respectively, $\sigma_{\max}$ denotes the maximum allowable stress in each segment, and $y_{\max}$ denotes the maximum allowable vertical deflection of the end of the beam.

**Formulation as a GP:**

$$\begin{array}{ll} \text{minimize} & w_1 h_1 + \cdots + w_N h_N \\ \text{subject to} & w_{\max}^{-1} w_i \leq 1\,,\ w_{\min} w_i^{-1} \leq 1\,,\ h_{\max}^{-1} h_i \leq 1\,,\ h_{\min} h_i^{-1} \leq 1\,,\ i = 1, \ldots, N \\ & S_{\max}^{-1} w_i h_i^{-1} \leq 1\,,\ S_{\min} w_i^{-1} h_i \leq 1\,,\ i = 1, \ldots, N \\ & 6iF \sigma_{\max}^{-1} w_i^{-1} h_i^{-2} \leq 1\,,\ i = 1, \ldots, N \\ & y_{\max}^{-1} y_1 \leq 1 \end{array} .$$

# Spectral Radius Minimization of a Nonnegative Matrix

If a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is *elementwise nonnegative* (i.e., $A_{\ell,m} \geq 0$ for $\ell, m = 1, \ldots, n$) and is *irreducible* (meaning that $(\mathbf{I} + \mathbf{A})^{n-1}$ is elementwise positive), then by the *Perron-Frobenius theorem*, $\mathbf{A}$ has a positive real eigenvalue $\lambda_{\text{pf}}(\mathbf{A})$, called the *Perron-Frobenius eigenvalue*, which is equal to its spectral radius, i.e., $\max_i \{|\lambda_i(\mathbf{A})|\}$.

**Properties of Perron-Frobenius Eigenvalue $\lambda_{\text{pf}}(\mathbf{A})$:**

- It determines the asymptotic growth or decay rate of $\mathbf{A}^k$: $\mathbf{A}^k \simeq \lambda_{\text{pf}}^k$ as $k \to \infty$.
- It has an alternate characterization given by

$$\lambda_{\text{pf}}(\mathbf{A}) = \inf \{\lambda : \mathbf{A}\mathbf{v} \preceq \lambda \mathbf{v} \text{ for some } \mathbf{v} \succ \mathbf{0}\} \ .$$

**Minimizing Spectral Radius of Matrix of Posynomials:**

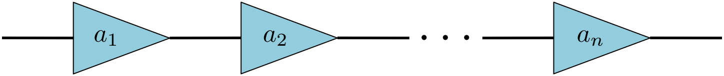
- The goal is to minimize $\lambda_{\text{pf}}(\mathbf{A}(\mathbf{x}))$, where the elements $A_{\ell,m}(\mathbf{x})$ are posynomials in $\mathbf{x}$, subject to possible posynomial inequalities on $\mathbf{x}$.
- Based on the above characterization of the Perron-Frobenius eigenvalue, this leads to the equivalent GP given below.

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \sum_{m=1}^{n} A_{\ell,m}(\mathbf{x})\, v_m / (\lambda v_\ell) \leq 1 \,, \ \ell = 1, \ldots, n \quad , \\ & f_i(\mathbf{x}) \leq 1 \,, \ i = 1, \ldots, p \end{array}$$

where the optimization variables are $\lambda$, $\mathbf{v}$, and $\mathbf{x}$.

# Design of Cascaded Amplifier Gains

The optimal selection of gain values to use for a *cascaded amplifier system* can be posed as a GP, as we show.




- The amplifier gains $a_1, \ldots, a_n > 0$ are chosen such that the overall gain remains fixed at some prescribed level.
- Individual amplifier saturation effects limit the maximum allowable output level.
- Noise and amplifier overload effects will limit the dynamic range of the system.

**Design Problem:**

| | |
|---|---|
| maximize | dynamic range |
| subject to | upper bounds on individual amplifier output levels |
| | upper bounds on individual amplifier gains $a_i$ |
| | fixed overall gain of cascaded system |

,

with variables $a_i$ for $i = 1, \ldots, n$.

Using practical modeling assumptions, we can pose this design problem as a GP.

# Design of Cascaded Amplifier Gains (Continued)

**Objective Function:**

- The *dynamic range* $D$ is given by $D = S_{\max}/N_{\text{out}}$, where $S_{\max}$ is the *maximum output signal level* and $N_{\text{out}}$ is the *output noise level*.
- The signal level at the output of the $i$-th amplifier, $S_i$, is given recursively as
$$S_0 = S_{\text{in}}\,,\ S_i = a_i S_{i-1}\,,\ i = 1, \ldots, n\,,$$
where $S_{\text{in}}$ is the *input signal level*.
- The root-mean-square (RMS) noise level at the output of the $i$-th amplifier, denoted $N_i$, is given recursively as
$$N_0 = 0\,,\ N_i = a_i \sqrt{N_{i-1}^2 + \eta_i^2}\,,$$
where $\eta_i$ is the *input-referred* RMS noise level of the $i$-th amplifier. We have $N_{\text{out}} = N_n$, i.e., the noise level of the last amplifier.

**Constraint Functions:**

- Each amplifier has a maximum allowable output level $M_i > 0$. Thus, we require $S_i \leq M_i$.
- Each amplifier has a maximum gain value given by $A_i^{\max} > 0$. Hence, we require $a_i \leq A_i^{\max}$.
- The overall gain of the system $a_1 a_2 \cdots a_n$ must be fixed at some level given by $A^{\text{tot}}$. So, we require $a_1 \cdots a_n = A^{\text{tot}}$.

# Design of Cascaded Amplifier Gains (Continued)

Note that for the signal levels, we have

$$S_i = S_{\text{in}}\prod_{m=1}^{i} a_m \le M_i \,,\; i = 1,\ldots,n \Longrightarrow S_{\text{in,max}} = \min_{i=1,\ldots,n} \frac{M_i}{\prod_{m=1}^{i} a_m} \,.$$

Since $S_n = A^{\text{tot}} S_{\text{in}}$, we have

$$S_{\text{max}} = A^{\text{tot}} S_{\text{in,max}} = \min_{i=1,\ldots,n} M_i \prod_{m=i+1}^{n} a_m \,.$$

Regarding noise levels, we have

$$N_0^2 = 0 \,,\; N_i^2 = a_i^2 \left(N_{i-1}^2 + \eta_i^2\right) \,,\; i = 1,\ldots,n \Longrightarrow N_{\text{out}}^2 = \sum_{\ell=1}^{n} \prod_{k=\ell}^{n} a_k^2 \eta_\ell^2 \,.$$

Therefore, the dynamic range is given by

$$D = \frac{\min_i M_i \prod_{m=i+1}^{n} a_m}{\left(\sum_{\ell=1}^{n} \prod_{k=\ell}^{n} a_k^2 \eta_\ell^2\right)^{1/2}} \,.$$

To maximize $D$, we can equivalently minimize $1/D^2$. Using an epigraph form of the problem, we obtain the following, which is essentially a GP.

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \sum_{\ell=1}^{n} \prod_{k=\ell}^{n} a_k^2 \eta_\ell^2 \le t M_i^2 \prod_{m=i+1}^{n} a_m^2 \,,\; i = 1,\ldots,n \\ & a_i \le A_i^{\max} \,,\; i = 1,\ldots,n \\ & \prod_{k=1}^{n} a_k = A^{\text{tot}} \end{array} \,.$$

# Problems with Generalized Inequality Constraints

One useful generalization of the standard form convex optimization problem is to allow the inequality constraint functions to be vector-valued and use generalized inequalities in the constraints. This yields the following generalization of the convex optimization problem.

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & \mathbf{f}_i(\mathbf{x}) \preceq_{\mathcal{K}_i} \mathbf{0} \,, \; i = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} \quad .$$

Here, $f_0 : \mathbb{R}^n \to \mathbb{R}$ is convex, while $\mathbf{f}_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ is $\mathcal{K}_i$-convex with respect to the proper cone $\mathcal{K}_i$ for $i = 1, \ldots, m$. Also, $\mathbf{A} \in \mathbb{R}^{p \times n}$ and $\mathbf{b} \in \mathbb{R}^p$.
This generalization has many of the same properties as the standard convex problem including the following.

- The feasible set, any sublevel set, and the optimal set are convex.
- Any locally optimal point for the problem is also globally optimal.
- The optimality condition for differentiable $f_0$ holds without any change.

**Conic Form Problem:**
The special case with an affine objective and constraints is referred to as a *conic form problem* and is given by the following.

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{F}\mathbf{x} + \mathbf{g} \preceq_{\mathcal{K}} \mathbf{0} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} \quad .$$

This extends linear programming (the case in which $\mathcal{K} = \mathbb{R}_+^m$) to nonpolyhedral cones.

# Semidefinite Programs

When $\mathcal{K} = \mathbb{S}_{+}^{k}$, the associated conic form problem is referred to as a *semidefinite program* (SDP). This has the following form.

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & x_1 \mathbf{F}_1 + x_2 \mathbf{F}_2 + \cdots + x_n \mathbf{F}_n + \mathbf{G} \preceq \mathbf{0} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

Here, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{F}_i \in \mathbb{S}^k$ for $i = 1, \ldots, n$, $\mathbf{G} \in \mathbb{S}^k$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, and $\mathbf{b} \in \mathbb{R}^p$. The inequality constraint is referred to as a *linear matrix inequality* (LMI).

**Multiple LMIs and Linear Inequalities:**
Consider a problem with a linear objective, several LMI constraints, along with linear inequality and equality constraints, of the form:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{F}^{(i)}(\mathbf{x}) = x_1 \mathbf{F}_1^{(i)} + x_2 \mathbf{F}_2^{(i)} + \cdots + x_n \mathbf{F}_n^{(i)} + \mathbf{G}^{(i)} \preceq \mathbf{0} \, , \; i = 1, \ldots, K \\ & \mathbf{G}\mathbf{x} \preceq \mathbf{h} \, , \; \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

All inequality constraints can be encompassed into one LMI and the problem can be recast as the SDP:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathrm{diag}\Big(\mathbf{G}\mathbf{x} - \mathbf{h}, \mathbf{F}^{(1)}(\mathbf{x}), \ldots, \mathbf{F}^{(K)}(\mathbf{x})\Big) \preceq \mathbf{0} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

# LPs and SOCPs as SDPs

## LP and Equivalent SDP:

LP: $$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array}$$

SDP: $$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathrm{diag}(\mathbf{A}\mathbf{x} - \mathbf{b}) \preceq \mathbf{0} \end{array}$$

Note the different interpretation of the generalized inequality $\preceq$. For the LP, $\preceq$ denotes componentwise inequality, whereas for the SDP, $\preceq$ denotes the usual matrix inequality.

## SOCP and Equivalent SDP:

SOCP: $$\begin{array}{ll} \text{minimize} & \mathbf{f}^T\mathbf{x} \\ \text{subject to} & ||\mathbf{A}_i\mathbf{x} + \mathbf{b}_i||_2 \leq \mathbf{c}_i^T\mathbf{x} + d_i \,, \; i = 1, \ldots, m \end{array}$$

SDP: $$\begin{array}{ll} \text{minimize} & \mathbf{f}^T\mathbf{x} \\ \text{subject to} & \begin{bmatrix} \left(\mathbf{c}_i^T\mathbf{x} + d_i\right)\mathbf{I} & (\mathbf{A}_i\mathbf{x} + \mathbf{b}_i) \\ (\mathbf{A}_i\mathbf{x} + \mathbf{b}_i)^T & \left(\mathbf{c}_i^T\mathbf{x} + d_i\right) \end{bmatrix} \succeq \mathbf{0} \,, \; i = 1, \ldots, m \end{array}$$

The SDP formulation of the SOCP arises through the use of a Schur complement.

# Eigenvalue / Matrix Norm Minimization

**Eigenvalue Minimization:**

$$\text{minimize} \quad \lambda_{\max}(\mathbf{A}(\mathbf{x})) \quad ,$$

where $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + x_1\mathbf{A}_1 + \cdots + x_n\mathbf{A}_n$ (with given $\mathbf{A}_i \in \mathbb{S}^k$).

*Equivalent SDP:*

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathbf{A}(\mathbf{x}) \preceq t\mathbf{I} \end{array} .$$

- Variables are $\mathbf{x} \in \mathbb{R}^n$, $t \in \mathbb{R}$.
- Constraint follows from

$$\lambda_{\max}(\mathbf{A}) \le t \iff \mathbf{A} \preceq t\mathbf{I} .$$

**Matrix Norm Minimization:**

$$\text{minimize} \quad \left(\lambda_{\max}\left(\mathbf{A}(\mathbf{x})^T\mathbf{A}(\mathbf{x})\right)\right)^{1/2} \quad ,$$

where $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + x_1\mathbf{A}_1 + \cdots + x_n\mathbf{A}_n$ (with given $\mathbf{A}_i \in \mathbb{R}^{p\times q}$).

*Equivalent SDP:*

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} t\mathbf{I} & \mathbf{A}(\mathbf{x}) \\ \mathbf{A}(\mathbf{x})^T & t\mathbf{I} \end{bmatrix} \succeq \mathbf{0} \end{array} .$$

- Variables are $\mathbf{x} \in \mathbb{R}^n$, $t \in \mathbb{R}$.
- Constraint follows from

$$||\mathbf{A}||_2 \le t \iff \mathbf{A}^T\mathbf{A} \preceq t^2\mathbf{I}\, ,\ t \ge 0 \iff \begin{bmatrix} t\mathbf{I} & \mathbf{A} \\ \mathbf{A}^T & t\mathbf{I} \end{bmatrix} \succeq \mathbf{0} .$$

# Condition Number Minimization

**Optimization Problem Statement:**

$$\text{minimize} \quad \kappa(\mathbf{A}(\mathbf{x})) \triangleq \frac{\lambda_{\max}(\mathbf{A}(\mathbf{x}))}{\lambda_{\min}(\mathbf{A}(\mathbf{x}))} ,$$

where $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + x_1\mathbf{A}_1 + \cdots + x_n\mathbf{A}_n$ (with given $\mathbf{A}_i \in \mathbb{S}^m$). The domain is $\{\mathbf{x} : \mathbf{A}(\mathbf{x}) \succ \mathbf{0}\}$.

**Equivalent Form:**

$$\begin{array}{ll} \text{minimize} & \lambda/\gamma \\ \text{subject to} & \gamma\mathbf{I} \preceq \mathbf{A}(\mathbf{x}) \preceq \lambda\mathbf{I} \end{array} ,$$

where the domain of the objective is $\{(\gamma, \lambda) : \gamma > 0\}$.

**Quasiconvex Formulation:**

$$\begin{array}{ll} \text{find} & \gamma, \lambda, \mathbf{x} \\ \text{subject to} & \lambda \leq \alpha\gamma\,,\ \gamma\mathbf{I} \preceq \mathbf{A}(\mathbf{x}) \preceq \lambda\mathbf{I}\,,\ \gamma > 0 \end{array} ,$$

**SDP Formulation:**

With the change of variables $s = 1/\gamma$, $t = \lambda/\gamma$, and $\mathbf{y} = \mathbf{x}/\gamma$, we get

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathbf{I} \preceq s\mathbf{A}_0 + y_1\mathbf{A}_1 + \cdots + y_n\mathbf{A}_n \preceq t\mathbf{I} \\ & s \geq 0 \end{array} .$$