Introduction to Semidefinite Programming (SDP)

Robert M. Freund and Brian Anthony
with assistance from David Craft

May 4-6, 2004

1 Outline

SLIDE 1

- Alternate View of Linear Programming
- Facts about Symmetric and Semidefinite Matrices
- SDP
- SDP Duality
- Examples of SDP
  - Combinatorial Optimization: MAXCUT
  - Convex Optimization: Quadratic Constraints, Eigenvalue Problems, $\log\det(X)$ problems
- Interior-Point Methods for SDP
- Application: Truss Vibration Dynamics via SDP

2 Linear Programming

2.1 Alternative Perspective

SLIDE 2

$$\begin{array}{llll} LP: & \text{minimize} & c \cdot x & \\ & \text{s.t.} & a_i \cdot x = b_i, & i = 1, \ldots, m \\ & & x \in \Re^n_+. & \end{array}$$

"$c \cdot x$" means the linear function "$\sum_{j=1}^n c_j x_j$"
$\Re^n_+ := \{x \in \Re^n \mid x \geq 0\}$ is the nonnegative orthant.
$\Re^n_+$ is a *convex cone.*
K is convex cone if $x, w \in K$ and $\alpha, \beta \geq 0 \quad \Rightarrow \quad \alpha x + \beta w \in K$.

SLIDE 3

$$\begin{array}{llll} LP: & \text{minimize} & c \cdot x & \\ & \text{s.t.} & a_i \cdot x = b_i, & i = 1, \ldots, m \\ & & x \in \Re^n_+. & \end{array}$$

"Minimize the linear function $c \cdot x$, subject to the condition that x must solve m given equations $a_i \cdot x = b_i, i = 1, \ldots, m$, and that x must lie in the convex cone $K = \Re^n_+$."

2.1.1 LP Dual Problem

SLIDE 4

$$\begin{array}{lll} LD: & \text{maximize} & \displaystyle\sum_{i=1}^m y_i b_i \\ & \text{s.t.} & \displaystyle\sum_{i=1}^m y_i a_i + s = c \\ & & s \in \Re^n_+. \end{array}$$

For feasible solutions x of LP and (y,s) of LD, the duality gap is simply

$$c \cdot x - \sum_{i=1}^{m} y_i b_i = \left(c - \sum_{i=1}^{m} y_i a_i \right) \cdot x = s \cdot x \geq 0$$

SLIDE 5

If LP and LD are feasible, then there exists x^* and (y^*, s^*) feasible for the primal and dual, respectively, for which

$$c \cdot x^* - \sum_{i=1}^{m} y_i^* b_i = s^* \cdot x^* = 0$$

3 Facts about the Semidefinite Cone

SLIDE 6

If X is an $n \times n$ matrix, then X is a symmetric positive semidefinite (SPSD) matrix if $X = X^T$ and

$$v^T X v \geq 0 \quad \text{for any} \quad v \in \Re^n$$

If X is an $n \times n$ matrix, then X is a symmetric positive definite (SPD) matrix if $X = X^T$ and

$$v^T X v > 0 \quad \text{for any} \quad v \in \Re^n, v \neq 0$$

4 Facts about the Semidefinite Cone

SLIDE 7

S^n denotes the set of symmetric $n \times n$ matrices
S^n_+ denotes the set of (SPSD) $n \times n$ matrices.
S^n_{++} denotes the set of (SPD) $n \times n$ matrices. Let $X, Y \in S^n$. SLIDE 8
"$X \succeq 0$" denotes that X is SPSD
"$X \succeq Y$" denotes that $X - Y \succeq 0$
"$X \succ 0$" to denote that X is SPD, etc.
Remark: $S^n_+ = \{X \in S^n \mid X \succeq 0\}$ is a convex cone.

5 Facts about Eigenvalues and Eigenvectors

SLIDE 9

If M is a square $n \times n$ matrix, then λ is an eigenvalue of M with corresponding eigenvector q if

$$Mq = \lambda q \quad \text{and} \quad q \neq 0 \ .$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ enumerate the eigenvalues of M.

6 Facts about Eigenvalues and Eigenvectors

SLIDE 10

The corresponding eigenvectors $q^1, q^2, \ldots, q^n$ of M can be chosen so that they are orthonormal, namely

$$\left(q^i\right)^T \left(q^j\right) = 0 \text{ for } i \neq j, \quad \text{and} \quad \left(q^i\right)^T \left(q^i\right) = 1$$

Define:

$$Q := \begin{bmatrix} q^1 & q^2 & \cdots & q^n \end{bmatrix}$$

Then Q is an *orthonormal* matrix:

$$Q^T Q = I, \quad \text{equivalently} \quad Q^T = Q^{-1}$$

SLIDE 11

$\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of M
$q^1, q^2, \ldots, q^n$ are the corresponding orthonormal eigenvectors of M

$$Q := \begin{bmatrix} q^1 & q^2 & \cdots & q^n \end{bmatrix}$$

$$Q^T Q = I, \quad \text{equivalently} \quad Q^T = Q^{-1}$$

Define D:

$$D := \begin{pmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Property: $M = QDQ^T$. SLIDE 12

The decomposition of M into $M = QDQ^T$ is called its *eigendecomposition.*

7 Facts about Symmetric Matrices

SLIDE 13

- If $X \in S^n$, then $X = QDQ^T$ for some orthonormal matrix Q and some diagonal matrix D. The columns of Q form a set of n orthogonal eigenvectors of X, whose eigenvalues are the corresponding entries of the diagonal matrix D.
- $X \succeq 0$ if and only if $X = QDQ^T$ where the eigenvalues (i.e., the diagonal entries of D) are all nonnegative.
- $X \succ 0$ if and only if $X = QDQ^T$ where the eigenvalues (i.e., the diagonal entries of D) are all positive.

SLIDE 14

- If M is symmetric, then

$$\det(M) = \prod_{j=1}^{n} \lambda_j$$

SLIDE 15

- Consider the matrix M defined as follows:

$$M = \begin{pmatrix} P & v \\ v^T & d \end{pmatrix},$$

  where $P \succ 0$, v is a vector, and d is a scalar. Then $M \succeq 0$ if and only if $d - v^T P^{-1} v \geq 0$.
- For a given column vector a, the matrix $X := aa^T$ is SPSD, i.e., $X = aa^T \succeq 0$.
- If $M \succeq 0$, then there is a matrix N for which $M = N^T N$. To see this, simply take $N = D^{\frac{1}{2}} Q^T$.

8 SDP

8.1 Semidefinite Programming

8.1.1 Think about X

SLIDE 16

Let $X \in S^n$. Think of X as:

- a matrix
- an array of n^2 components of the form $(x_{11}, \ldots, x_{nn})$
- an object (a vector) in the space S^n.

All three different equivalent ways of looking at X will be useful.

8.1.2 Linear Function of X

SLIDE 17

Let $X \in S^n$. What will a linear function of X look like?
If $C(X)$ is a linear function of X, then $C(X)$ can be written as $C \bullet X$, where

$$C \bullet X := \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij}.$$

There is no loss of generality in assuming that the matrix C is also symmetric.

8.1.3 Definition of SDP

SLIDE 18

$$\begin{array}{lll} SDP: & \text{minimize} & C \bullet X \\ & \text{s.t.} & A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ & & X \succeq 0, \end{array}$$

"$X \succeq 0$" is the same as "$X \in S^n_+$"
The data for SDP consists of the symmetric matrix C (which is the data for the objective function) and the m symmetric matrices $A_1, \ldots, A_m$, and the $m-$vector b, which form the m linear equations.

8.1.4 Example

SLIDE 19

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix}, \; b = \begin{pmatrix} 11 \\ 19 \end{pmatrix}, \; \text{and} \; C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix},$$

The variable X will be the 3×3 symmetric matrix:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix},$$

$$\begin{array}{llll} SDP: & \text{minimize} & x_{11}+4x_{12}+6x_{13}+9x_{22}+0x_{23}+7x_{33} & \\ & \text{s.t.} & x_{11}+0x_{12}+2x_{13}+3x_{22}+14x_{23}+5x_{33} & = 11 \\ & & 0x_{11}+4x_{12}+16x_{13}+6x_{22}+0x_{23}+4x_{33} & = 19 \end{array}$$

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \succeq 0.$$

SLIDE 20

$$\begin{array}{llll} SDP: & \text{minimize} & x_{11}+4x_{12}+6x_{13}+9x_{22}+0x_{23}+7x_{33} & \\ & \text{s.t.} & x_{11}+0x_{12}+2x_{13}+3x_{22}+14x_{23}+5x_{33} & = 11 \\ & & 0x_{11}+4x_{12}+16x_{13}+6x_{22}+0x_{23}+4x_{33} & = 19 \end{array}$$

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \succeq 0.$$

It may be helpful to think of "$X \succeq 0$" as stating that each of the n *eigenvalues* of X must be nonnegative.

8.1.5 $LP \subset SDP$

SLIDE 21

$$\begin{array}{lll} LP: & \text{minimize} & c \cdot x \\ & \text{s.t.} & a_i \cdot x = b_i, \quad i = 1, \ldots, m \\ & & x \in \Re^n_+. \end{array}$$

Define:

$$A_i = \begin{pmatrix} a_{i1} & 0 & \ldots & 0 \\ 0 & a_{i2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{in} \end{pmatrix}, \quad i = 1, \ldots, m, \quad \text{and} \quad C = \begin{pmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_n \end{pmatrix}.$$

$$\begin{array}{lll} SDP: & \text{minimize} & C \bullet X \\ & \text{s.t.} & A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ & & X_{ij} = 0, \quad i = 1, \ldots, n, \quad j = i+1, \ldots, n, \\ & & X = \begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_n \end{pmatrix} \succeq 0, \end{array}$$

9 SDP Duality

SLIDE 22

$$\begin{array}{lll} SDD: & \text{maximize} & \sum_{i=1}^{m} y_i b_i \\ & \text{s.t.} & \sum_{i=1}^{m} y_i A_i + S = C \\ & & S \succeq 0. \end{array}$$

Notice

$$S = C - \sum_{i=1}^{m} y_i A_i \succeq 0$$

10 SDP Duality

Slide 23

and so equivalently:

$$\begin{array}{lll} SDD: & \text{maximize} & \sum_{i=1}^{m} y_i b_i \\ & \text{s.t.} & C - \sum_{i=1}^{m} y_i A_i \succeq 0 \end{array}$$

10.1 Example

Slide 24

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix}, \; b = \begin{pmatrix} 11 \\ 19 \end{pmatrix}, \quad \text{and } C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix},$$

$$\begin{array}{lll} SDD: & \text{maximize} & 11y_1 + 19y_2 \\ & \text{s.t.} & y_1 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix} + S = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix} \\ & & S \succeq 0 \end{array}$$

Slide 25

$$\begin{array}{lll} SDD: & \text{maximize} & 11y_1 + 19y_2 \\ & \text{s.t.} & y_1 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix} + S = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix} \\ & & S \succeq 0 \end{array}$$

is the same as:

$$\begin{array}{lll} SDD: & \text{maximize} & 11y_1 + 19y_2 \\ & \text{s.t.} & \begin{pmatrix} 1 - 1y_1 - 0y_2 & 2 - 0y_1 - 2y_2 & 3 - 1y_1 - 8y_2 \\ 2 - 0y_1 - 2y_2 & 9 - 3y_1 - 6y_2 & 0 - 7y_1 - 0y_2 \\ 3 - 1y_1 - 8y_2 & 0 - 7y_1 - 0y_2 & 7 - 5y_1 - 4y_2 \end{pmatrix} \succeq 0. \end{array}$$

10.2 Weak Duality

Slide 26

Weak Duality Theorem: Given a feasible solution X of SDP and a feasible solution (y, S) of SDD, the duality gap is

$$C \bullet X - \sum_{i=1}^{m} y_i b_i = S \bullet X \geq 0 \ .$$

If

$$C \bullet X - \sum_{i=1}^{m} y_i b_i = 0 \ ,$$

then X and (y, S) are each optimal solutions to SDP and SDD, respectively, and furthermore, $SX = 0$.

10.3 Strong Duality

SLIDE 27

Strong Duality Theorem: Let z_P^* and z_D^* denote the optimal objective function values of SDP and SDD, respectively. Suppose that there exists a feasible solution $\hat{X}$ of SDP such that $\hat{X} \succ 0$, and that there exists a feasible solution $(\hat{y}, \hat{S})$ of SDD such that $\hat{S} \succ 0$. Then both SDP and SDD attain their optimal values, and

$$z_P^* = z_D^* \ .$$

11 Some Important Weaknesses of SDP

SLIDE 28

- There may be a finite or infinite duality gap.
- The primal and/or dual may or may not attain their optima.
- Both programs will attain their common optimal value if both programs have feasible solutions that are SPD.
- There is no finite algorithm for solving SDP.
- There is a simplex algorithm, but it is not a finite algorithm. There is no direct analog of a "basic feasible solution" for SDP.

12 SDP in Combinatorial Optimization

12.0.1 The MAX CUT Problem

SLIDE 29

G is an undirected graph with nodes $N = \{1, \ldots, n\}$ and edge set E.
Let $w_{ij} = w_{ji}$ be the weight on edge (i, j), for $(i, j) \in E$.
We assume that $w_{ij} \geq 0$ for all $(i, j) \in E$.
The MAX CUT problem is to determine a subset S of the nodes N for which the sum of the weights of the edges that cross from S to its complement $\bar{S}$ is maximized ($\bar{S} := N \setminus S$).

12.0.2 Formulations

SLIDE 30

The MAX CUT problem is to determine a subset S of the nodes N for which the sum of the weights w_{ij} of the edges that cross from S to its complement $\bar{S}$ is maximized ($\bar{S} := N \setminus S$).
Let $x_j = 1$ for $j \in S$ and $x_j = -1$ for $j \in \bar{S}$.

$$\begin{array}{lll} MAXCUT: & \text{maximize}_x & \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_i x_j) \\ & & \\ & \text{s.t.} & x_j \in \{-1, 1\}, \quad j = 1, \ldots, n. \end{array}$$

SLIDE 31

$$\begin{array}{lll} MAXCUT: & \text{maximize}_x & \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_i x_j) \\ & & \\ & \text{s.t.} & x_j \in \{-1, 1\}, \quad j = 1, \ldots, n. \end{array}$$

Let

$$Y = xx^T \ .$$

Then

$$Y_{ij} = x_i x_j \quad i = 1, \ldots, n, \quad j = 1, \ldots, n.$$

SLIDE 32

Also let W be the matrix whose $(i,j)^{\text{th}}$ element is w_{ij} for $i = 1, \ldots, n$ and $j = 1, \ldots, n$. Then

$$\begin{array}{llll} MAXCUT: & \text{maximize}_{Y,x} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y \\ & \text{s.t.} & x_j \in \{-1, 1\}, \quad j = 1, \ldots, n \\ & & Y = xx^T. \end{array}$$

SLIDE 33

$$\begin{array}{llll} MAXCUT: & \text{maximize}_{Y,x} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y \\ & \text{s.t.} & x_j \in \{-1, 1\}, \quad j = 1, \ldots, n \\ & & Y = xx^T. \end{array}$$

SLIDE 34

The first set of constraints are equivalent to $Y_{jj} = 1, j = 1, \ldots, n$.

$$\begin{array}{llll} MAXCUT: & \text{maximize}_{Y,x} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y \\ & \text{s.t.} & Y_{jj} = 1, \quad j = 1, \ldots, n \\ & & Y = xx^T. \end{array}$$

SLIDE 35

$$\begin{array}{llll} MAXCUT: & \text{maximize}_{Y,x} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y \\ & \text{s.t.} & Y_{jj} = 1, \quad j = 1, \ldots, n \\ & & Y = xx^T. \end{array}$$

SLIDE 36

Notice that the matrix $Y = xx^T$ is a rank-1 SPSD matrix.
We *relax* this condition by removing the rank-1 restriction:

$$\begin{array}{llll} RELAX: & \text{maximize}_{Y} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y \\ & \text{s.t.} & Y_{jj} = 1, \quad j = 1, \ldots, n \\ & & Y \succeq 0. \end{array}$$

It is therefore easy to see that RELAX provides an upper bound on MAXCUT, i.e.,

$$MAXCUT \le RELAX.$$

SLIDE 37

$$RELAX: \quad \text{maximize}_Y \quad \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y$$

$$\text{s.t.} \quad Y_{jj} = 1, \quad j = 1, \ldots, n$$

$$Y \succeq 0.$$

As it turns out, one can also prove without too much effort that:

$$0.87856 \; RELAX \le MAXCUT \le RELAX.$$

SLIDE 38

This is an impressive result, in that it states that the value of the semidefinite relaxation is guaranteed to be no more than 12.2% higher than the value of NP-hard problem MAX CUT.

13 SDP for Convex QCQP

SLIDE 39

A convex quadratically constrained quadratic program (QCQP) is a problem of the form:

$$\begin{array}{lll} QCQP: & \underset{x}{\text{minimize}} & x^T Q_0 x + q_0^T x + c_0 \\ & \text{s.t.} & x^T Q_i x + q_i^T x + c_i \le 0 \quad, i = 1, \ldots, m, \end{array}$$

where the $Q_0 \succeq 0$ and $Q_i \succeq 0$, $\; i = 1, \ldots, m$. This is the same as:

$$\begin{array}{lll} QCQP: & \underset{x,\theta}{\text{minimize}} & \theta \\ & \text{s.t.} & x^T Q_0 x + q_0^T x + c_0 - \theta \le 0 \\ & & x^T Q_i x + q_i^T x + c_i \le 0 \quad, i = 1, \ldots, m. \end{array}$$

SLIDE 40

$$\begin{array}{lll} QCQP: & \underset{x,\theta}{\text{minimize}} & \theta \\ & \text{s.t.} & x^T Q_0 x + q_0^T x + c_0 - \theta \le 0 \\ & & x^T Q_i x + q_i^T x + c_i \le 0 \quad, i = 1, \ldots, m. \end{array}$$

Factor each Q_i into

$$Q_i = M_i^T M_i$$

and note the equivalence:

$$\begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - q_i^T x \end{pmatrix} \succeq 0 \quad \Longleftrightarrow \quad x^T Q_i x + q_i^T x + c_i \le 0.$$

SLIDE 41

$$QCQP: \quad \underset{x,\theta}{\text{minimize}} \quad \theta$$
$$\text{s.t.} \quad x^T Q_0 x + q_0^T x + c_0 - \theta \leq 0$$
$$x^T Q_i x + q_i^T x + c_i \leq 0 \quad , i = 1, \ldots, m.$$

Re-write $QCQP$ as:

$$QCQP: \quad \underset{x,\theta}{\text{minimize}} \quad \theta$$
$$\text{s.t.} \quad \begin{pmatrix} I & M_0 x \\ x^T M_0^T & -c_0 - q_0^T x + \theta \end{pmatrix} \succeq 0$$
$$\begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - q_i^T x \end{pmatrix} \succeq 0 \quad , i = 1, \ldots, m.$$

14 SDP for SOCP

14.1 Second-Order Cone Optimization

SLIDE 42

Second-order cone optimization:

$$\text{SOCP}: \quad \min_x \quad c^T x$$
$$\text{s.t.} \quad Ax = b$$
$$\|Q_i x + d_i\| \leq \left(g_i^T x + h_i\right) \quad , \quad i = 1, \ldots, k \ .$$

Recall $\|v\| := \sqrt{v^T v}$

SLIDE 43

$$\text{SOCP}: \quad \min_x \quad c^T x$$
$$\text{s.t.} \quad Ax = b$$
$$\|Q_i x + d_i\| \leq \left(g_i^T x + h_i\right) \quad , \quad i = 1, \ldots, k \ .$$

Property:

$$\|Qx + d\| \leq \left(g^T x + h\right) \iff \begin{pmatrix} (g^T x + h)I & (Qx + d) \\ (Qx + d)^T & g^T x + h \end{pmatrix} \succeq 0 \ .$$

This property is a direct consequence of the fact that

$$M = \begin{pmatrix} P & v \\ v^T & d \end{pmatrix} \succeq 0 \iff d - v^T P^{-1} v \geq 0 \ .$$

SLIDE 44

$$\text{SOCP}: \quad \min_x \quad c^T x$$
$$\text{s.t.} \quad Ax = b$$
$$\|Q_i x + d_i\| \leq \left(g_i^T x + h_i\right) \quad , \quad i = 1, \ldots, k \ .$$

Re-write as:

$$\text{SDPSOCP}: \quad \min_x \quad c^T x$$
$$\text{s.t.} \quad Ax = b$$
$$\begin{pmatrix} (g_i^T x + h_i)I & (Q_i x + d_i) \\ (Q_i x + d_i)^T & g_i^T x + h_i \end{pmatrix} \succeq 0 \ , \quad i = 1, \ldots, k \ .$$

15 Eigenvalue Optimization

Slide 45

We are given symmetric matrices B and $A_i, i = 1, \ldots, k$
Choose weights $w_1, \ldots, w_k$ to create a new matrix S:

$$S := B - \sum_{i=1}^{k} w_i A_i \ .$$

There might be restrictions on the weights $Gw \leq d$.
The typical goal is for S is to have some nice property such as:

- $\lambda_{\min}(S)$ is maximized
- $\lambda_{\max}(S)$ is minimized
- $\lambda_{\max}(S) - \lambda_{\min}(S)$ is minimized

15.1 Some Useful Relationships

Slide 46

Property: $M \succeq tI$ if and only if $\lambda_{\min}(M) \geq t$.

Proof: $M = QDQ^T$. Define

$$R = M - tI = QDQ^T - tI = Q(D - tI)Q^T \ .$$

$$M \succeq tI \iff R \succeq 0 \iff D - tI \succeq 0 \iff \lambda_{\min}(M) \geq t \ .$$

q.e.d.

Property: $M \preceq tI$ if and only if $\lambda_{\max}(M) \leq t$.

15.2 Design Problem

Slide 47

Consider the design problem:

$$\begin{array}{lll} EOP: & \underset{w,S}{\text{minimize}} & \lambda_{\max}(S) - \lambda_{\min}(S) \\ & \text{s.t.} & S = B - \sum_{i=1}^{k} w_i A_i \\ & & Gw \leq d \ . \end{array}$$

Slide 48

$$\begin{array}{llll} EOP: & \underset{w,S}{\text{minimize}} & \lambda_{\max}(S) - \lambda_{\min}(S) \\ & \text{s.t.} & S = B - \sum_{i=1}^{k} w_i A_i \\ & & Gw \le d\ . \end{array}$$

This is equivalent to:

$$\begin{array}{llll} EOP: & \underset{w,S,\mu,\lambda}{\text{minimize}} & \mu - \lambda \\ & \text{s.t.} & S = B - \sum_{i=1}^{k} w_i A_i \\ & & Gw \le d \\ & \lambda I \preceq S \preceq \mu I. & \end{array}$$

16 The Logarithmic Barrier Function for SPD Matrices

SLIDE 49

Let $X \succeq 0$, equivalently $X \in S^n_+$.
X will have n nonnegative eigenvalues, say $\lambda_1(X), \ldots, \lambda_n(X) \ge 0$ (possibly counting multiplicities).

$$\begin{aligned} \partial S^n_+ = \{X \in S^n \mid \lambda_j(X) \ge 0, j = 1, \ldots, n, \\ \text{and } \lambda_j(X) = 0 \text{ for some } j \in \{1, \ldots, n\}\}. \end{aligned}$$

SLIDE 50

$$\begin{aligned} \partial S^n_+ = \{X \in S^n \mid \lambda_j(X) \ge 0, j = 1, \ldots, n, \\ \text{and } \lambda_j(X) = 0 \text{ for some } j \in \{1, \ldots, n\}\}. \end{aligned}$$

A natural barrier function is:

$$B(X) := -\sum_{j=1}^{n} \ln(\lambda_i(X)) = -\ln\left(\prod_{j=1}^{n} \lambda_i(X)\right) = -\ln(\det(X)).$$

This function is called the log-determinant function or the logarithmic barrier function for the semidefinite cone.

SLIDE 51

$$B(X) := -\sum_{j=1}^{n} \ln(\lambda_i(X)) = -\ln\left(\prod_{j=1}^{n} \lambda_i(X)\right) = -\ln(\det(X)).$$

Quadratic Taylor expansion at $X = \bar{X}$:

$$B(\bar{X} + \alpha D) \approx B(\bar{X}) + \alpha \bar{X}^{-1} \bullet D + \frac{1}{2}\alpha^2 \left(\bar{X}^{-\frac{1}{2}} D \bar{X}^{-\frac{1}{2}}\right) \bullet \left(\bar{X}^{-\frac{1}{2}} D \bar{X}^{-\frac{1}{2}}\right)\ .$$

$B(X)$ has the same remarkable properties in the context of interior-point methods for SDP as the barrier function $-\sum_{j=1}^{n} \ln(x_j)$ does in the context of linear optimization.

17 The SDP Analytic Center Problem

Slide 52

Given a system:

$$\sum_{i=1}^{m} y_i A_i \preceq C \ ,$$

the *analytic center* is the solution $(\hat{y}, \hat{S})$ of:

$$\begin{array}{lll} \text{(ACP:)} & \text{maximize}_{y,S} & \prod_{i=1}^{n} \lambda_i(S) \\ & \text{s.t.} & \sum_{i=1}^{m} y_i A_i + S = C \\ & & S \succeq 0 \ . \end{array}$$

Slide 53

$$\begin{array}{lll} \text{(ACP:)} & \text{maximize}_{y,S} & \prod_{i=1}^{n} \lambda_i(S) \\ & \text{s.t.} & \sum_{i=1}^{m} y_i A_i + S = C \\ & & S \succeq 0 \ . \end{array}$$

This is the same as:

$$\begin{array}{lll} \text{(ACP:)} & \text{minimize}_{y,S} & -\ln\det(S) \\ & \text{s.t.} & \sum_{i=1}^{m} y_i A_i + S = C \\ & & S \succ 0 \ . \end{array}$$

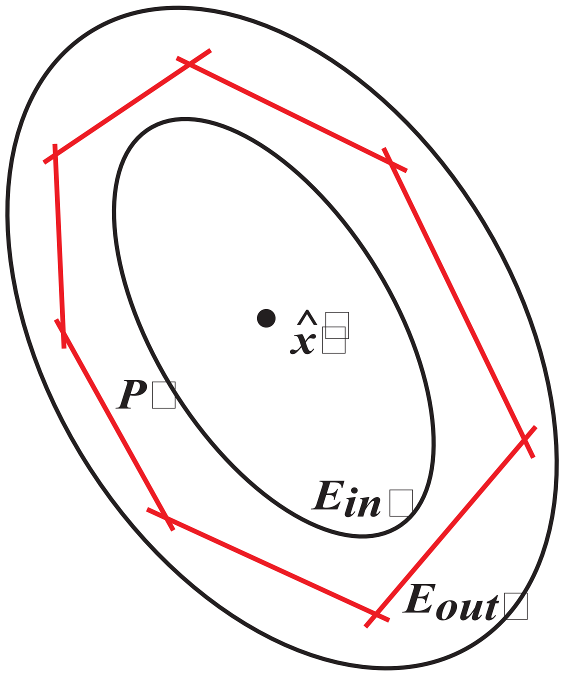
Slide 54

$$\begin{array}{lll} \text{(ACP:)} & \text{minimize}_{y,S} & -\ln\det(S) \\ & \text{s.t.} & \sum_{i=1}^{m} y_i A_i + S = C \\ & & S \succ 0 \ . \end{array}$$

Let $(\hat{y}, \hat{S})$ be the analytic center.
There are easy-to-construct ellipsoids E_{IN} and E_{OUT}, both centered at $\hat{y}$ and where E_{OUT} is a scaled version of E_{IN} with scale factor n, with the property that:

$$E_{\text{IN}} \subset P \subset E_{\text{OUT}}$$

Slide 55



18 Minimum Volume Circumscription

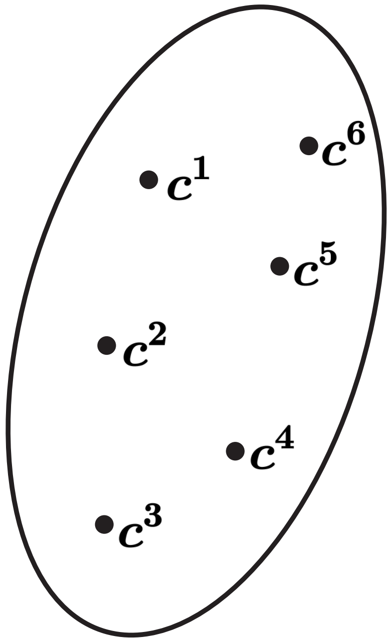
SLIDE 56

$R \succ 0$ and $z \in \Re^n$ define an ellipsoid in $\Re^n$:

$$E_{R,z} := \{y \mid (y-z)^T R(y-z) \leq 1\}.$$

The volume of $E_{R,z}$ is proportional to $\sqrt{\det(R^{-1})}$. SLIDE 57

Given k points $c_1, \ldots, c_k$, we would like to find an ellipsoid circumscribing $c_1, \ldots, c_k$ that has minimum volume:



Slide 58

$$MCP: \quad \underset{R,z}{\text{minimize}} \quad \text{vol}\ (E_{R,z})$$
$$\text{s.t.} \quad c_i \in E_{R,z}, \quad i = 1, \ldots, k$$

which is equivalent to:

$$MCP: \quad \underset{R,z}{\text{minimize}} \quad -\ln(\det(R))$$
$$\text{s.t.} \quad (c_i - z)^T R (c_i - z) \le 1, \quad i = 1, \ldots, k$$
$$R \succ 0$$

Slide 59

$$MCP: \quad \underset{R,z}{\text{minimize}} \quad -\ln(\det(R))$$
$$\text{s.t.} \quad (c_i - z)^T R (c_i - z) \le 1, \quad i = 1, \ldots, k$$
$$R \succ 0$$

Factor $R = M^2$ where $M \succ 0$ (that is, M is a square root of R):

$$MCP: \quad \underset{M,z}{\text{minimize}} \quad -\ln(\det(M^2))$$
$$\text{s.t.} \quad (c_i - z)^T M^T M (c_i - z) \le 1, \quad i = 1, \ldots, k,$$
$$M \succ 0$$

Slide 60

$$\begin{array}{llll} MCP: & \underset{M,z}{\text{minimize}} & -\ln(\det(M^2)) & \\ & \text{s.t.} & (c_i - z)^T M^T M(c_i - z) \leq 1, & i = 1, \dots, k, \\ & & M \succ 0. & \end{array}$$

Notice the equivalence:

$$\begin{pmatrix} I & Mc_i - Mz \\ (Mc_i - Mz)^T & 1 \end{pmatrix} \succeq 0 \quad \Longleftrightarrow \quad (c_i - z)^T M^T M(c_i - z) \leq 1$$

Re-write MCP:

$$\begin{array}{lll} MCP: & \underset{M,z}{\text{minimize}} & -2\ln(\det(M)) \\ & \text{s.t.} & \begin{pmatrix} I & Mc_i - Mz \\ (Mc_i - Mz)^T & 1 \end{pmatrix} \succeq 0, \quad i = 1, \dots, k, \\ & & M \succ 0. \end{array}$$

Slide 61

$$\begin{array}{lll} MCP: & \underset{M,z}{\text{minimize}} & -2\ln(\det(M)) \\ & \text{s.t.} & \begin{pmatrix} I & Mc_i - Mz \\ (Mc_i - Mz)^T & 1 \end{pmatrix} \succeq 0, \quad i = 1, \dots, k, \\ & & M \succ 0. \end{array}$$

Substitute $y = Mz$:

$$\begin{array}{lll} MCP: & \underset{M,y}{\text{minimize}} & -2\ln(\det(M)) \\ & \text{s.t.} & \begin{pmatrix} I & Mc_i - y \\ (Mc_i - y)^T & 1 \end{pmatrix} \succeq 0, \quad i = 1, \dots, k, \\ & & M \succ 0. \end{array}$$

Slide 62

$$\begin{array}{lll} MCP: & \underset{M,y}{\text{minimize}} & -2\ln(\det(M)) \\ & \text{s.t.} & \begin{pmatrix} I & Mc_i - y \\ (Mc_i - y)^T & 1 \end{pmatrix} \succeq 0, \quad i = 1, \dots, k, \\ & & M \succ 0. \end{array}$$

This problem is very easy to solve.
Recover the original solution R, z by computing:

$$R = M^2 \quad \text{and} \quad z = M^{-1}y.$$

19 SDP in Control Theory

Slide 63

A variety of control and system problems can be cast and solved as instances of SDP. This topic is beyond the scope of this lecturer's expertise.

20 Interior-point Methods for SDP

20.1 Primal and Dual SDP

SLIDE 64

$$\begin{array}{lll} SDP: & \text{minimize} & C \bullet X \\ & \text{s.t.} & A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ & & X \succeq 0 \end{array}$$

and

$$\begin{array}{lll} SDD: & \text{maximize} & \sum_{i=1}^m y_i b_i \\ & \text{s.t.} & \sum_{i=1}^m y_i A_i + S = C \\ & & S \succeq 0 \ . \end{array}$$

If X and (y, S) are feasible for the primal and the dual, the duality gap is:

$$C \bullet X - \sum_{i=1}^m y_i b_i = S \bullet X \geq 0 \ .$$

Also,

$$S \bullet X = 0 \iff SX = 0 \ .$$

SLIDE 65

$$B(X) = -\sum_{j=1}^n \ln(\lambda_i(X)) = -\ln\left(\prod_{j=1}^n \lambda_i(X)\right) = -\ln(\det(X)) \ .$$

Consider:

$$\begin{array}{lll} BSDP(\mu): & \text{minimize} & C \bullet X - \mu \ln(\det(X)) \\ & \text{s.t.} & A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ & & X \succ 0. \end{array}$$

Let $f_\mu(X)$ denote the objective function of $BSDP(\mu)$. Then:

$$-\nabla f_\mu(X) = C - \mu X^{-1}$$

SLIDE 66

$$\begin{array}{lll} BSDP(\mu): & \text{minimize} & C \bullet X - \mu \ln(\det(X)) \\ & \text{s.t.} & A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ & & X \succ 0. \end{array}$$

$\nabla f_\mu(X) = C - \mu X^{-1}$
Karush-Kuhn-Tucker conditions for $BSDP(\mu)$ are:

$$\begin{cases} A_i \bullet X = b_i \quad, i = 1, \ldots, m, \\ X \succ 0, \\ C - \mu X^{-1} = \sum_{i=1}^m y_i A_i. \end{cases}$$

SLIDE 67

$$\begin{cases} A_i \bullet X = b_i \quad , i = 1, \ldots, m, \\ X \succ 0, \\ C - \mu X^{-1} = \sum_{i=1}^{m} y_i A_i. \end{cases}$$

Define

$$S = \mu X^{-1} \ ,$$

which implies

$$XS = \mu I \ ,$$

and rewrite KKT conditions as: SLIDE 68

$$\begin{cases} A_i \bullet X = b_i \quad , i = 1, \ldots, m, \quad X \succ 0 \\ \sum_{i=1}^{m} y_i A_i + S = C \\ XS = \mu I. \end{cases}$$

SLIDE 69

$$\begin{cases} A_i \bullet X = b_i \quad , i = 1, \ldots, m, \quad X \succ 0 \\ \sum_{i=1}^{m} y_i A_i + S = C \\ XS = \mu I. \end{cases}$$

If (X, y, S) is a solution of this system, then X is feasible for SDP, (y, S) is feasible for SDD, and the resulting duality gap is

$$S \bullet X = \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} X_{ij} = \sum_{j=1}^{n} (SX)_{jj} = \sum_{j=1}^{n} (\mu I)_{jj} = n\mu.$$

SLIDE 70

$$\begin{cases} A_i \bullet X = b_i \quad , i = 1, \ldots, m, \quad X \succ 0 \\ \sum_{i=1}^{m} y_i A_i + S = C \\ XS = \mu I. \end{cases}$$

If (X, y, S) is a solution of this system, then X is feasible for SDP, (y, S) is feasible for SDD, the duality gap is

$$S \bullet X = n\mu.$$

SLIDE 71

This suggests that we try solving $BSDP(\mu)$ for a variety of values of μ as $\mu \to 0$. Interior-point methods for SDP are very similar to those for linear optimization, in that they use Newton's method to solve the KKT system as $\mu \to 0$.

21 Website for SDP

SLIDE 72

A good website for semidefinite programming is:

`http://www-user.tu-chemnitz.de/ helmberg/semidef.html`.

22 Optimization of Truss Vibration

22.1 Motivation

SLIDE 73

- The design and analysis of trusses are found in a wide variety of scientific applications including engineering mechanics, structural engineering, MEMS, and biomedical engineering.
- As finite approximations to solid structures, a truss is the fundamental concept of Finite Element Analysis.
- The truss problem also arises quite obviously and naturally in the design of scaffolding-based structures such as bridges, the Eiffel tower, and the skeletons for tall buildings.

SLIDE 74

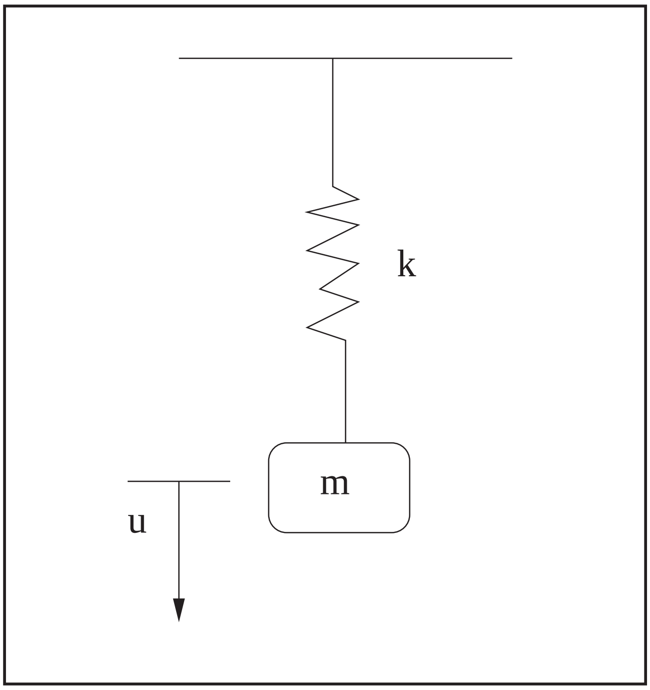
- Using semidefinite programming (SDP) and the interior-point software SDPT3, we will explore an elegant and powerful technique for optimizing truss vibration dynamics.
- The problem we consider here is designing a truss such that the lowest frequency Ω at which it vibrates is above a given lower bound $\bar{\Omega}$.
- November 7, 1940, Tacoma Narrows Bridge in Tacoma, Washington

22.2 The Dynamics Model

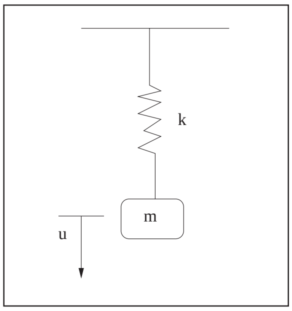
SLIDE 75

Newton's Second Law of Motion:

$$F = m \times a \ .$$



SLIDE 76



If the mass is pulled down, the displacement u produces a force in the spring tending to move the mass back to its equilibrium point (where $u = 0$).
The displacement u causes an upward force $k \times u$, where k is the spring constant.
We obtain from $F = m \times a$ that:

$$-ku(t) = m\ddot{u}(t)$$

SLIDE 77

Law of Motion:

$$-ku(t) = m\ddot{u}(t)$$

Solution:

$$u(t) = \sin\left(\sqrt{\frac{k}{m}}\; t\right)$$

Frequency of vibration:

$$\omega = \sqrt{\frac{k}{m}}\;.$$

22.2.1 Apply to Truss Structure

SLIDE 78

Law of Motion:

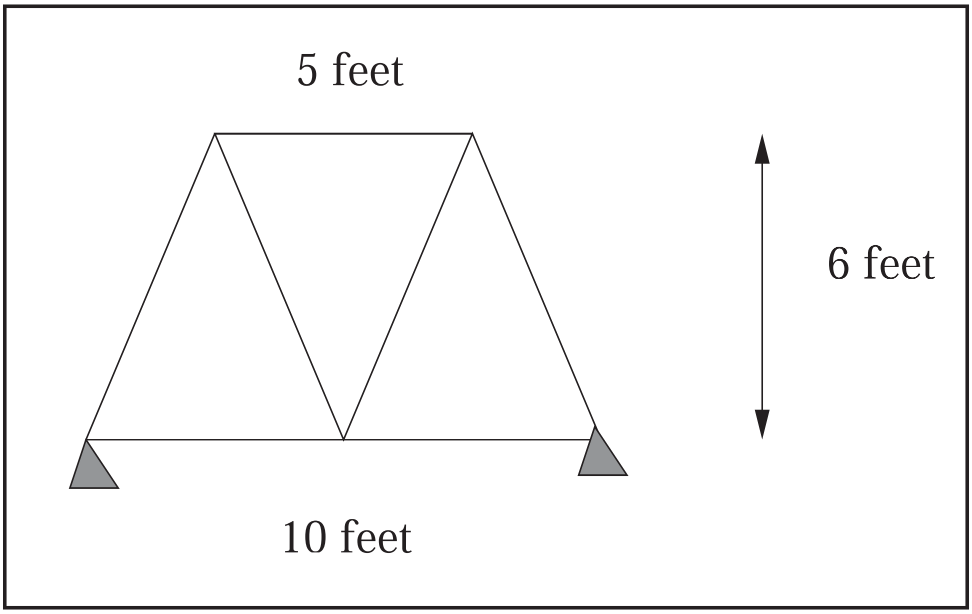
$$-ku(t) = m\ddot{u}(t)$$

Solution:

$$u(t) = \sin\left(\sqrt{\frac{k}{m}}\; t\right)$$

$$\omega = \sqrt{\frac{k}{m}}$$

For truss structure, we need multidimensional analogs for $k, u(t)$, and m. SLIDE 79
A simple truss.
Each bar has both stiffness and mass that depend on material properties and the bar's cross-sectional area.



22.2.2 Analog of k

SLIDE 80

The spring constant k extends to the stiffness matrix of a truss.
We used G to denote the stiffness matrix.
Here we will use K.

$$K = G = AB^{-1}A^T$$

Each column of A, denoted as a_i, is the projection of bar i onto the degrees of freedom of the nodes that bar i meets.

$$B = \begin{pmatrix} \frac{L_1^2}{E_1 t_1} & & 0 \\ & \ddots & \\ 0 & & \frac{L_m^2}{E_m t_m} \end{pmatrix} \quad , \quad B^{-1} = \begin{pmatrix} \frac{E_1 t_1}{L_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{E_m t_m}{L_m^2} \end{pmatrix} .$$

22.2.3 Analog of m

SLIDE 81

Instead of a single displacement scalar $u(t)$, we have N degrees of freedom, and the vector

$$u(t) = (u_1(t), \ldots, u_N(t))$$

is the vector of displacements.
The mass m extends to a mass matrix M

22.2.4 Laws of Motion

SLIDE 82

$$-ku(t) = m\ddot{u}(t)$$

becomes:

$$-Ku(t) = M\ddot{u}(t)$$

Both K and M are SPD matrices, and are easily computed once the truss geometry and the nodal constraints are specified. SLIDE 83

$$-Ku(t) = M\ddot{u}(t)$$

The truss structure vibration involves sine functions with frequencies

$$\omega_i = \sqrt{\lambda_i}$$

where

$$\lambda_1, \dots, \lambda_N$$

are the eigenvalues of

$$M^{-1}K$$

The *threshold frequency* Ω of the truss is the lowest frequency $\omega_i, i = 1, \dots, N$, or equivalently, the square root of the smallest eigenvalue of $M^{-1}K$. SLIDE 84

$$-Ku(t) = M\ddot{u}(t)$$

The *threshold frequency* Ω of the truss is the square root of the smallest eigenvalue of $M^{-1}K$.
Lower bound constraint on the threshold frequency

$$\Omega \geq \bar{\Omega}$$

Property:

$$\Omega \geq \bar{\Omega} \iff K - \bar{\Omega}^2 M \succeq 0 \ .$$

22.3 Truss Vibration Design

SLIDE 85

We wrote the stiffness matrix as a linear function of the volumes t_i of the bars i:

$$K = \sum_{i=1}^{m} t_i \frac{E_i}{L_i^2} (a_i)(a_i)^T \ ,$$

L_i is the length of bar i
E_i is the Young's modulus of bar i
t_i is the volume of bar i.

22.4 Truss Vibration Design

SLIDE 86

Here we use y_i to represent the area of bar i ($y_i = \frac{t_i}{L_i}$)

$$K = K(y) = \sum_{i=1}^{m} \left[\frac{E_i}{L_i} (a_i)(a_i)^T \right] y_i = \sum_{i=1}^{m} K_i y_i$$

where

$$K_i = \left[\frac{E_i}{L_i} (a_i)(a_i)^T \right] \ , \ i = 1, \dots, m$$

SLIDE 87

There are matrices $M_1, \ldots, M_m$ for which we can write the mass matrix as a linear function of the areas $y_1, \ldots, y_m$:

$$M = M(y) = \sum_{i=1}^{m} M_i y_i$$

In truss vibration design, we seek to design a truss of minimum weight whose threshold frequency Ω is at least a pre-specified value $\bar{\Omega}$. SLIDE 88

$$\begin{array}{lll} TSDP: & \text{minimize} & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & & \sum_{i=1}^{m} (K_i - \bar{\Omega}^2 M_i) y_i \succeq 0 \\ & & l_i \leq y_i \leq u_i \ , \ i = 1, \ldots, m \ . \end{array}$$

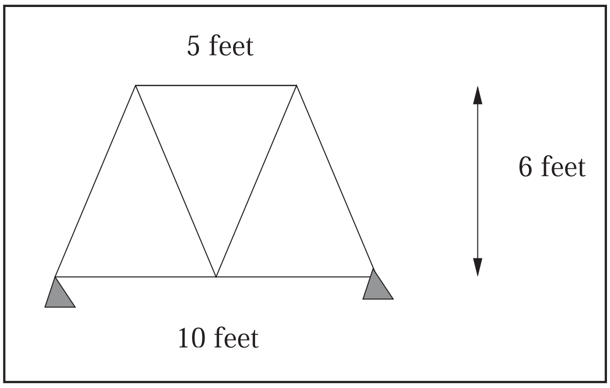
The decision variables are $y_1, \ldots, y_m$ SLIDE 89
l_i, u_i are bounds on the area y_i of bar i (perhaps from the output of the static truss design model)
b_i is the length of bar i times the material density of bar i SLIDE 90

$$\begin{array}{lll} TSDP: & \text{minimize}_{\text{y}} & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & & \sum_{i=1}^{m} (K_i - \bar{\Omega}^2 M_i) y_i \succeq 0 \\ & & l_i \leq y_i \leq u_i \ , \ i = 1, \ldots, m \ . \end{array}$$

22.5 Computational Example

SLIDE 91

$$\begin{array}{lll} TSDP: & \text{minimize}_{\text{y}} & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & & \sum_{i=1}^{m} (K_i - \bar{\Omega}^2 M_i) y_i \succeq 0 \\ & & l_i \leq y_i \leq u_i \ , \ i = 1, \ldots, m \ . \end{array}$$



SLIDE 92

$$\begin{array}{lll} TSDP: & \text{minimize}_{\text{y}} & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & & \sum_{i=1}^{m} (K_i - \bar{\Omega}^2 M_i) y_i \succeq 0 \\ & & l_i \leq y_i \leq u_i \ , \ i = 1, \ldots, m \ . \end{array}$$

- $l_i = 5.0$ square inches for all bars i
- $u_i = 8.0$ square inches for all bars i
- mass density for steel, which is ρ =0.736e−03
- Young's modulus for steel, which is 3.0e+07 pounds per square inch
- $\bar{\Omega} = 220$Hz

22.5.1 SDPT3

SLIDE 93

SDPT3 is the semidefinite programming software developed by "T3":

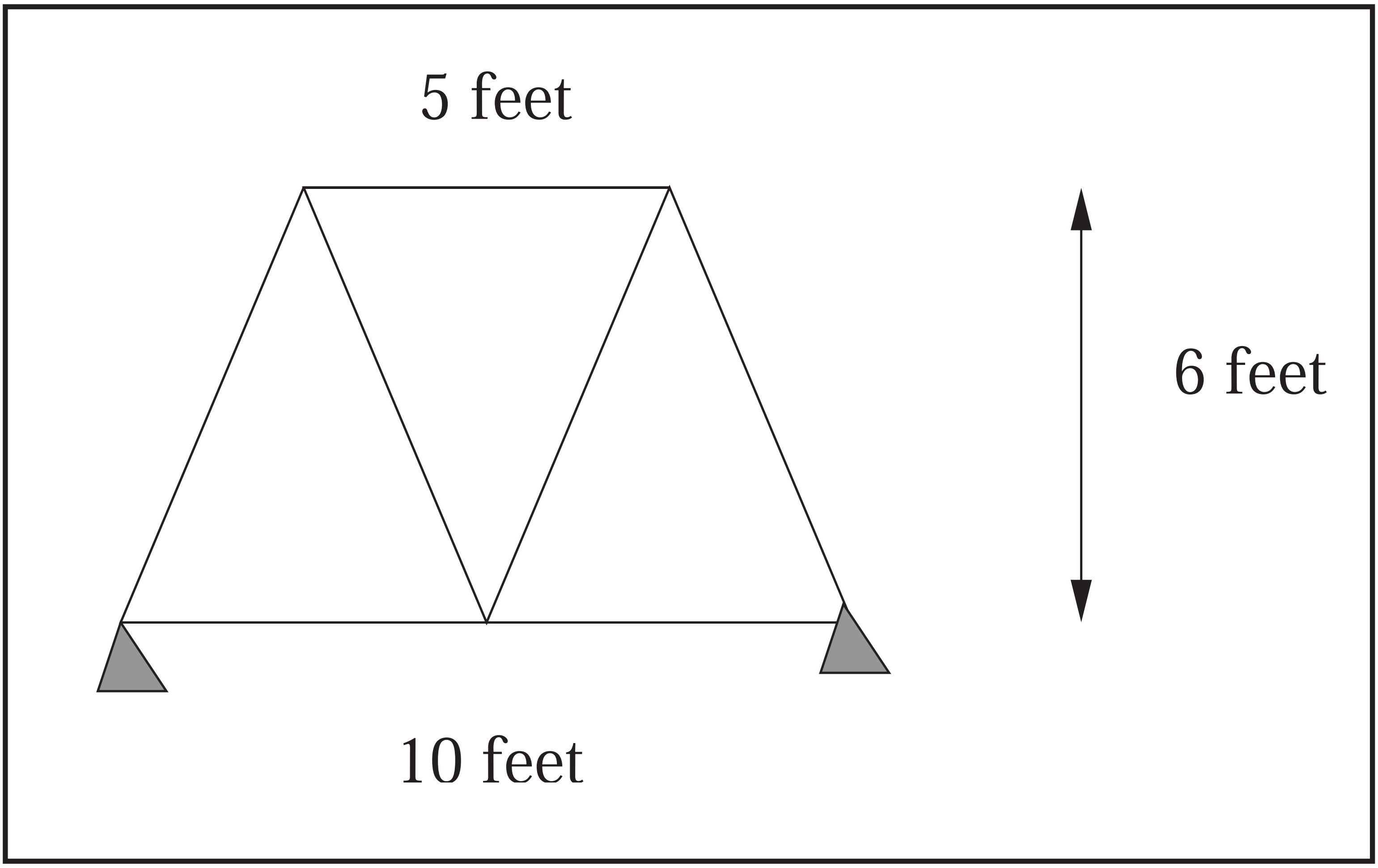
- Kim Chuan Toh of National University of Singapore
- Reha Tütünçu of Carnegie Mellon University
- Michael Todd of Cornell University

SLIDE 94

Statistics for TSDP problem run using SDPT3

Linear Inequalities	14
Semidefinite block size	6×6
CPU time (seconds):	0.8
IPM Iterations:	15
Optimal Solution	
Bar 1 area (square inches)	8.0000
Bar 2 area (square inches)	8.0000
Bar 3 area (square inches)	7.1797
Bar 4 area (square inches)	6.9411
Bar 5 area (square inches)	5.0000
Bar 6 area (square inches)	6.9411
Bar 7 area (square inches)	7.1797

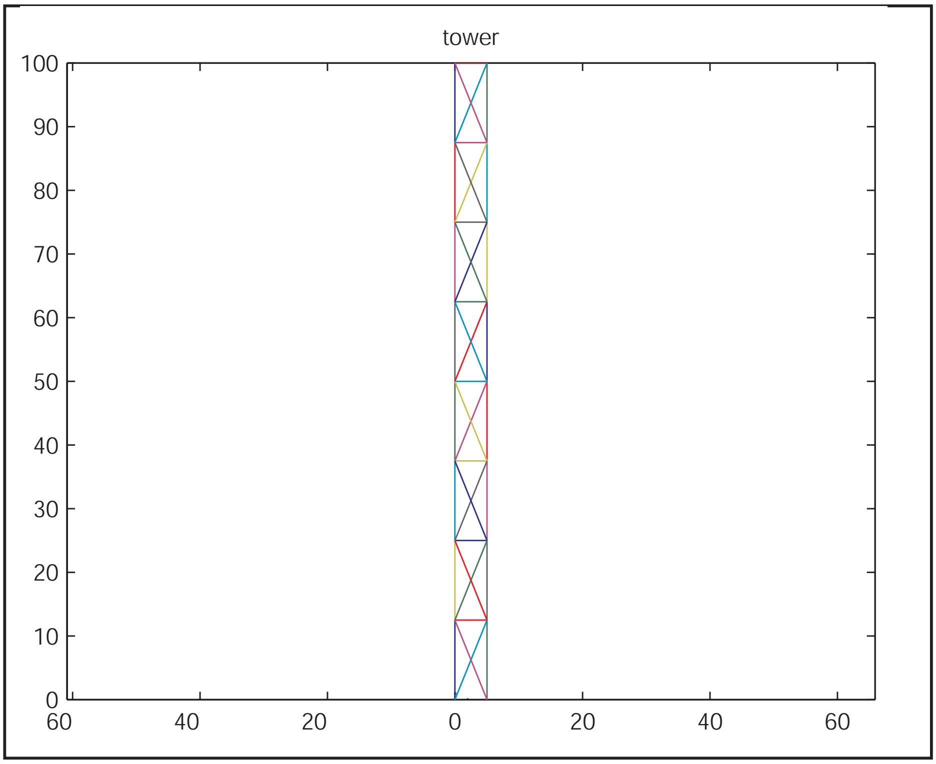
SLIDE 95



22.6 More Computation

Slide 96

A truss tower used for computational experiments. This version of the tower has 40 bars and 32 degrees of freedom.



SLIDE 97

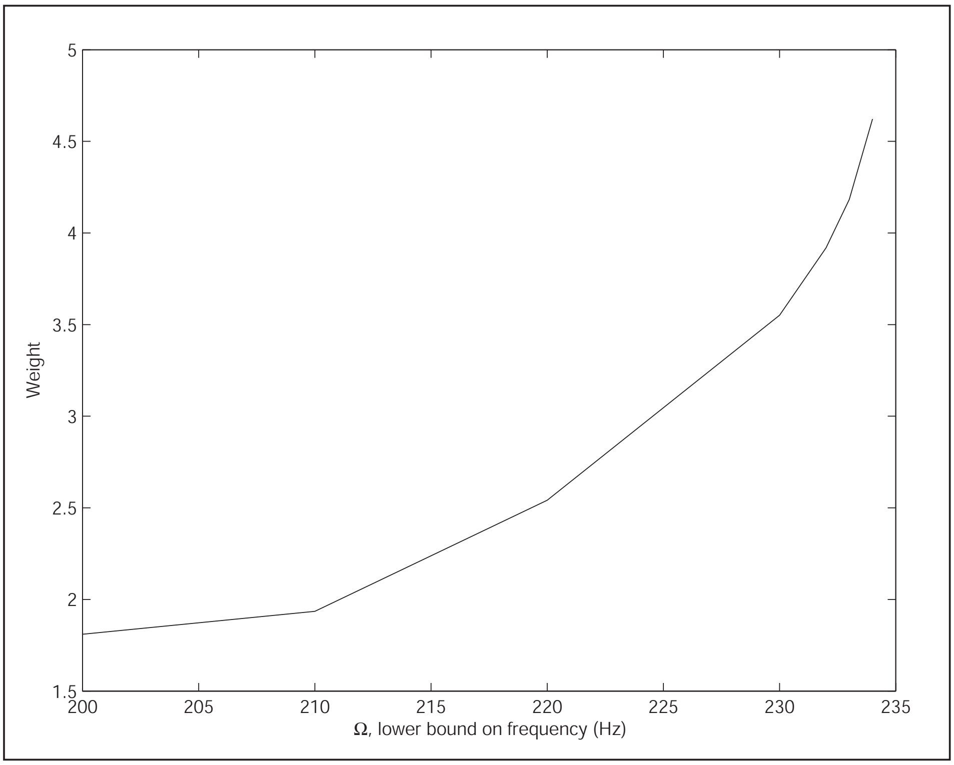
Computational results using SDPT3 for truss frequency optimization.

Semidefinite Block	Linear Inequalities	Scalar Variables	IPM Iterations	CPU time (sec)
12 × 12	30	15	17	1.17
20 × 20	50	25	20	1.49
32 × 32	80	40	21	1.88
48 × 48	120	60	20	2.73
60 × 60	150	75	20	3.76
80 × 80	200	100	23	5.34
120 × 120	300	150	23	9.46

22.6.1 Frontier Solutions

SLIDE 98

Lower bound on Threshold Frequency Ω versus Weight of Structure



Weight
5
4.5
4
3.5
3
2.5
2
1.5
200
205
210
215
220
225
230
235
Ω, lower bound on frequency (Hz)