

Semidefinite optimization with MOSEK

Joachim Dahl

MOSEK ApS

INFORMS annual meeting
Minneapolis, October 5th, 2013

Conic
optimization

Convex cones

Semidefinite
matrices

Linear cone
problems

Conic
modeling

Simple cones

Nearest
correlation

Linear matrix
inequalities

Eigenvalue
optimization

Combinatorial
relaxations

Sum-of-squares
relaxations

Nonnegative
polynomials

Conclusions

Conic optimization

Convex cones

Semidefinite matrices

Linear cone problems

Conic modeling

Simple cones

Nearest correlation

Linear matrix inequalities

Eigenvalue optimization

Combinatorial relaxations

Sum-of-squares relaxations

Nonnegative polynomials

Conclusions

What is a convex cone?

S is a *convex cone* if

$$x \in S \iff \alpha \cdot x \in S, \forall \alpha \geq 0$$

Simple examples

- Nonnegative orthant, $x \geq 0$.
- Quadratic cone,

$$Q^n = \left\{ x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + \dots + x_n^2} \right\}.$$

also known as *second-order* or *Lorentz cone*.

What is a convex cone?

S is a *convex cone* if

$$x \in S \iff \alpha \cdot x \in S, \forall \alpha \geq 0$$

Simple examples

- Nonnegative orthant, $x \geq 0$.
- Quadratic cone,

$$Q^n = \left\{ x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + \dots + x_n^2} \right\}.$$

also known as *second-order* or *Lorentz cone*.

Semidefinite matrices

A symmetric matrix $X \in \mathcal{S}^n$ is positive semidefinite iff

- all eigenvalues are nonnegative.
- it can be factored as $X = VV^T$.
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.

Cone of semidefinite matrices

$$\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0, \forall z \in \mathbb{R}^n \right\}.$$

Notation: $X \succeq Y \iff (X - Y) \in \mathcal{S}_+^n$.

Matrix inner-products and norms

$$\langle A, B \rangle := \text{trace}(A^T B) = \sum_{ij} A_{ij} B_{ij}$$

$$\|A\|_F^2 := \langle A, A \rangle = \sum_{ij} A_{ij}^2.$$

Semidefinite matrices

A symmetric matrix $X \in \mathcal{S}^n$ is positive semidefinite iff

- all eigenvalues are nonnegative.
- it can be factored as $X = VV^T$.
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.

Cone of semidefinite matrices

$$\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0, \forall z \in \mathbb{R}^n \right\}.$$

Notation: $X \succeq Y \iff (X - Y) \in \mathcal{S}_+^n$.

Matrix inner-products and norms

$$\langle A, B \rangle := \text{trace}(A^T B) = \sum_{ij} A_{ij} B_{ij}$$

$$\|A\|_F^2 := \langle A, A \rangle = \sum_{ij} A_{ij}^2.$$

Semidefinite matrices

A symmetric matrix $X \in \mathcal{S}^n$ is positive semidefinite iff

- all eigenvalues are nonnegative.
- it can be factored as $X = VV^T$.
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.

Cone of semidefinite matrices

$$\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0, \forall z \in \mathbb{R}^n \right\}.$$

Notation: $X \succeq Y \iff (X - Y) \in \mathcal{S}_+^n$.

Matrix inner-products and norms

$$\langle A, B \rangle := \text{trace}(A^T B) = \sum_{ij} A_{ij} B_{ij}$$

$$\|A\|_F^2 := \langle A, A \rangle = \sum_{ij} A_{ij}^2.$$

Semidefinite matrices

A symmetric matrix $X \in \mathcal{S}^n$ is positive semidefinite iff

- all eigenvalues are nonnegative.
- it can be factored as $X = VV^T$.
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.

Cone of semidefinite matrices

$$\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0, \forall z \in \mathbb{R}^n \right\}.$$

Notation: $X \succeq Y \iff (X - Y) \in \mathcal{S}_+^n$.

Matrix inner-products and norms

$$\langle A, B \rangle := \text{trace}(A^T B) = \sum_{ij} A_{ij} B_{ij}$$

$$\|A\|_F^2 := \langle A, A \rangle = \sum_{ij} A_{ij}^2.$$

Semidefinite matrices

A symmetric matrix $X \in \mathcal{S}^n$ is positive semidefinite iff

- all eigenvalues are nonnegative.
- it can be factored as $X = VV^T$.
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.

Cone of semidefinite matrices

$$\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0, \forall z \in \mathbb{R}^n \right\}.$$

Notation: $X \succeq Y \iff (X - Y) \in \mathcal{S}_+^n$.

Matrix inner-products and norms

$$\langle A, B \rangle := \text{trace}(A^T B) = \sum_{ij} A_{ij} B_{ij}$$

$$\|A\|_F^2 := \langle A, A \rangle = \sum_{ij} A_{ij}^2.$$

```
>> A=[2,1,1; 1,2,1; 1,1,1]
```

```
A =
```

```
    2    1    1
    1    2    1
    1    1    1
```

```
>> [U, D]=eig(A) % A = U*D*U'
```

```
U =
```

```
    0.3251    0.7071    0.6280
    0.3251   -0.7071    0.6280
   -0.8881         0    0.4597
```

```
D =
```

```
    0.2679         0         0
         0    1.0000         0
         0         0    3.7321
```

```
>> V=U*sqrt(D)
```

```
V =
```

```
    0.1683    0.7071    1.2131
    0.1683   -0.7071    1.2131
   -0.4597         0    0.8881
```

```
>> V*V' % V is a factor of A
```

```
ans =
```

```
    2.0000    1.0000    1.0000
    1.0000    2.0000    1.0000
    1.0000    1.0000    1.0000
```

```
>> A(:)'\*A(:) % squared Frobenius norm
```

```
ans =
```

```
15.0000
```

```
>> sum(diag(D).^2)
```

```
ans =
```

```
15.0000
```

Linear cone problems

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& x \in \mathcal{C} \end{aligned}$$

$$\begin{aligned} &\text{maximize} && b^T y \\ &\text{subject to} && c - A^T y = s \\ &&& s \in \mathcal{C} \end{aligned}$$

where $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \times \dots \times \mathcal{C}_p$ is a product of cones.

Admissible cones

- Nonnegative orthant $x \geq 0$.
- Quadratic cone Q^n .
- Rotated quadratic cone,

$$Q_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \dots + x_n^2, x_1, x_2 \geq 0\}.$$

- Semidefinite cone \mathcal{S}_+^n .

Linear cone problems

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{C} \end{array}$$

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \in \mathcal{C} \end{array}$$

where $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \times \dots \times \mathcal{C}_p$ is a product of cones.

Admissible cones

- Nonnegative orthant $x \geq 0$.
- Quadratic cone Q^n .
- Rotated quadratic cone,

$$Q_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \dots + x_n^2, x_1, x_2 \geq 0\}.$$

- Semidefinite cone \mathcal{S}_+^n .

Linear cone problems

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b \\ & && x \in \mathcal{C} \end{aligned}$$

$$\begin{aligned} & \text{maximize} && b^T y \\ & \text{subject to} && c - A^T y = s \\ & && s \in \mathcal{C} \end{aligned}$$

where $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \times \dots \times \mathcal{C}_p$ is a product of cones.

Admissible cones

- Nonnegative orthant $x \geq 0$.
- Quadratic cone Q^n .
- Rotated quadratic cone,

$$Q_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \dots + x_n^2, x_1, x_2 \geq 0\}.$$

- Semidefinite cone S_+^n .

Example problem:

$$\begin{aligned}
 &\text{minimize} && \left\langle \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, X \right\rangle + z_1 \\
 &\text{subject to} && \left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle + z_1 &= 1 \\
 &&& \left\langle \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, X \right\rangle + z_2 + z_3 &= 1/2 \\
 &&& (z_1, z_2, z_3) \in \mathcal{Q}^3, X \in \mathcal{S}_+^3
 \end{aligned}$$

A standard linear cone problem with

$$x = (z_1 \quad z_2 \quad z_3 \quad X_{11} \quad X_{21} \quad X_{31} \quad X_{12} \quad X_{22} \quad X_{23} \quad X_{13} \quad X_{23} \quad X_{33})$$

$$c = (1 \quad 0 \quad 0 \quad 2 \quad 1 \quad 0 \quad 1 \quad 2 \quad 1 \quad 0 \quad 1 \quad 2)^T$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$b = (1 \quad 1/2)^T$$

Semidefinite optimization

Joachim Dahl

Conic optimization

Convex cones
Semidefinite matrices

Linear cone problems

Conic modeling

Simple cones

Nearest correlation

Linear matrix inequalities

Eigenvalue optimization

Combinatorial relaxations

Sum-of-squares relaxations

Nonnegative polynomials

Conclusions

Example problem:

$$\begin{aligned}
 &\text{minimize} && \left\langle \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, X \right\rangle + z_1 \\
 &\text{subject to} && \left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle + z_1 &= 1 \\
 &&& \left\langle \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, X \right\rangle + z_2 + z_3 &= 1/2 \\
 &&& (z_1, z_2, z_3) \in \mathcal{Q}^3, X \in \mathcal{S}_+^3
 \end{aligned}$$

A standard linear cone problem with

$$\begin{aligned}
 x &= (z_1 \quad z_2 \quad z_3 \quad X_{11} \quad X_{21} \quad X_{31} \quad X_{12} \quad X_{22} \quad X_{23} \quad X_{13} \quad X_{23} \quad X_{33})^T \\
 c &= (1 \quad 0 \quad 0 \quad 2 \quad 1 \quad 0 \quad 1 \quad 2 \quad 1 \quad 0 \quad 1 \quad 2)^T \\
 A &= \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\
 b &= (1 \quad 1/2)^T
 \end{aligned}$$

```

>> [x(1:3), s(1:3)]
ans =
    0.2544    0.4552
    0.1799   -0.3219
    0.1799   -0.3219

>> X
X =
    0.2173   -0.2600    0.2173
   -0.2600    0.3111   -0.2600
    0.2173   -0.2600    0.2173

>> S
S =
    1.1333    0.6781   -0.3219
    0.6781    1.1333    0.6781
   -0.3219    0.6781    1.1333

>> c'*x - b'*y
ans =
    4.3340e-08

>> norm(A*x-b)
ans =
    2.4379e-08

>> norm(c-A'*y-s)
ans =
    1.6585e-08

>> x'*s
ans =
    2.9971e-08

>> [eig(X), eig(S)]
ans =
    0.0000   -0.0000
    0.0000    1.4552
    0.7456    1.9448

>> x(1)-norm(x(2:3))
ans =
    8.0901e-09

>> s(1)-norm(s(2:3))
ans =
    2.2202e-08
  
```


Conic optimization

Convex cones

Semidefinite matrices

Linear cone problems

Conic modeling

Simple cones

Nearest correlation

Linear matrix inequalities

Eigenvalue optimization

Combinatorial relaxations

Sum-of-squares relaxations

Nonnegative polynomials

Conclusions

Conic
optimization

Convex cones

Semidefinite
matrices

Linear cone
problems

Conic
modeling

Simple cones

Nearest
correlation

Linear matrix
inequalities

Eigenvalue
optimization

Combinatorial
relaxations

Sum-of-squares
relaxations

Nonnegative
polynomials

Conclusions

Simple examples of quadratic cones

- Absolute values

$$|x| \leq t \iff (t, x) \in \mathcal{Q}^2$$

- Euclidean norms

$$\|x\| \leq t \iff (t, x) \in \mathcal{Q}^{n+1}$$

- Squared euclidean norms

$$\|x\|^2 \leq t \iff (t, 1/2, x) \in \mathcal{Q}_r^{n+2}$$

- Hyperbolic sets

$$\frac{1}{x} \leq t, x > 0 \iff (t, x, \sqrt{2}) \in \mathcal{Q}_r^3$$

Simple examples of quadratic cones

- Absolute values

$$|x| \leq t \iff (t, x) \in \mathcal{Q}^2$$

- Euclidean norms

$$\|x\| \leq t \iff (t, x) \in \mathcal{Q}^{n+1}$$

- Squared euclidean norms

$$\|x\|^2 \leq t \iff (t, 1/2, x) \in \mathcal{Q}_r^{n+2}$$

- Hyperbolic sets

$$\frac{1}{x} \leq t, x > 0 \iff (t, x, \sqrt{2}) \in \mathcal{Q}_r^3$$

Simple examples of quadratic cones

- Absolute values

$$|x| \leq t \iff (t, x) \in \mathcal{Q}^2$$

- Euclidean norms

$$\|x\| \leq t \iff (t, x) \in \mathcal{Q}^{n+1}$$

- Squared euclidean norms

$$\|x\|^2 \leq t \iff (t, 1/2, x) \in \mathcal{Q}_r^{n+2}$$

- Hyperbolic sets

$$\frac{1}{x} \leq t, x > 0 \iff (t, x, \sqrt{2}) \in \mathcal{Q}_r^3$$

Simple examples of quadratic cones

- Absolute values

$$|x| \leq t \iff (t, x) \in \mathcal{Q}^2$$

- Euclidean norms

$$\|x\| \leq t \iff (t, x) \in \mathcal{Q}^{n+1}$$

- Squared euclidean norms

$$\|x\|^2 \leq t \iff (t, 1/2, x) \in \mathcal{Q}_r^{n+2}$$

- Hyperbolic sets

$$\frac{1}{x} \leq t, x > 0 \iff (t, x, \sqrt{2}) \in \mathcal{Q}_r^3$$

Very simple examples of semidefinite cones

- Nonnegativity

$$x \succeq 0 \iff \text{diag}(x) \succeq 0.$$

- Quadratic cones

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff x_1 x_2 \geq x_3^2, \quad x_1, x_2 \geq 0,$$

in other words,

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff (x_1, x_2, x_3/\sqrt{2}) \in Q_r^3.$$

A similar result for n -dimensional quadratic cones.

Very simple examples of semidefinite cones

- Nonnegativity

$$x \succeq 0 \iff \text{diag}(x) \succeq 0.$$

- Quadratic cones

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff x_1 x_2 \geq x_3^2, \quad x_1, x_2 \geq 0,$$

in other words,

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff (x_1, x_2, x_3/\sqrt{2}) \in Q_r^3.$$

A similar result for n -dimensional quadratic cones.

Very simple examples of semidefinite cones

- Nonnegativity

$$x \succeq 0 \iff \text{diag}(x) \succeq 0.$$

- Quadratic cones

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff x_1 x_2 \geq x_3^2, \quad x_1, x_2 \geq 0,$$

in other words,

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff (x_1, x_2, x_3/\sqrt{2}) \in Q_r^3.$$

A similar result for n -dimensional quadratic cones.

Very simple examples of semidefinite cones

- Nonnegativity

$$x \succeq 0 \iff \text{diag}(x) \succeq 0.$$

- Quadratic cones

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff x_1 x_2 \geq x_3^2, \quad x_1, x_2 \geq 0,$$

in other words,

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq 0 \iff (x_1, x_2, x_3/\sqrt{2}) \in Q_r^3.$$

A similar result for n -dimensional quadratic cones.

A picture is worth a thousand words...

The pillow

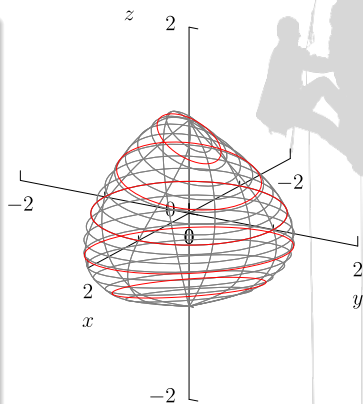
Consider the set:

$$\begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0.$$

- Exterior is a *spectrahedron*.
- Can be characterized as

$$x^2 + y^2 + z^2 - 2xyz = 1.$$

- **Ellipsoids** for fixed $z \in [-1, 1]$.



A picture is worth a thousand words...

The pillow

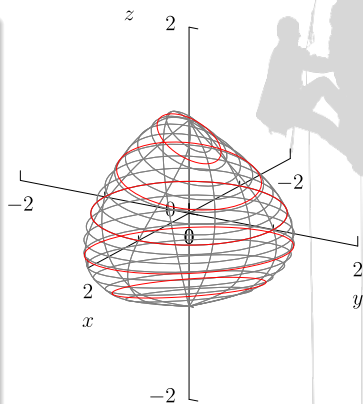
Consider the set:

$$\begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0.$$

- Exterior is a *spectrahedron*.
- Can be characterized as

$$x^2 + y^2 + z^2 - 2xyx = 1.$$

- Ellipsoids for fixed $z \in [-1, 1]$.



A picture is worth a thousand words...

The pillow

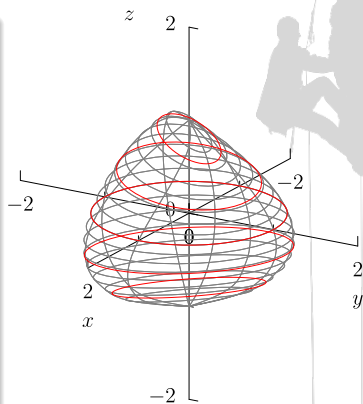
Consider the set:

$$\begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0.$$

- Exterior is a *spectrahedron*.
- Can be characterized as

$$x^2 + y^2 + z^2 - 2xyz = 1.$$

- **Ellipsoids** for fixed $z \in [-1, 1]$.



Nearest correlation problem

Consider the set

$$S = \{X \in \mathcal{S}_+^n \mid q_{ii} = 1, i = 1, \dots, n\}.$$

For $A \in \mathcal{S}^n$ the *nearest correlation matrix* is

$$X^* = \arg \min_{X \in S} \|A - X\|_F.$$

A conic formulation

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|\text{vec}(A - X)\| \leq t \\ & \text{diag}(X) = e \\ & X \succeq 0 \end{array}$$

where $\text{vec}(X) = (x_{11}, x_{21}, \dots, x_{n1}, x_{12}, \dots, x_{nn})$.

Nearest correlation problem

Consider the set

$$S = \{X \in \mathcal{S}_+^n \mid q_{ii} = 1, i = 1, \dots, n\}.$$

For $A \in \mathcal{S}^n$ the *nearest correlation matrix* is

$$X^* = \arg \min_{X \in S} \|A - X\|_F.$$

A conic formulation

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|\text{vec}(A - X)\| \leq t \\ & \text{diag}(X) = e \\ & X \succeq 0 \end{array}$$

where $\text{vec}(X) = (x_{11}, x_{21}, \dots, x_{n1}, x_{12}, \dots, x_{nn})$.

Linear matrix functions

Consider a matrix-valued function $F : \mathbb{R}^m \mapsto \mathcal{S}^n$,

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m$$

where $F_i \in \mathcal{S}^n$.

- The inequality

$$F_0 + x_1 F_1 + \cdots + x_m F_m \succeq 0$$

is called a *linear matrix inequality* (LMI).

- Corresponds to conic dual constraints,

$$C - (y_1 A_1 + \cdots + y_m A_m) = S, \quad S \succeq 0.$$

Linear matrix functions

Consider a matrix-valued function $F : \mathbb{R}^m \mapsto \mathcal{S}^n$,

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m$$

where $F_i \in \mathcal{S}^n$.

- The inequality

$$F_0 + x_1 F_1 + \cdots + x_m F_m \succeq 0$$

is called a *linear matrix inequality* (LMI).

- Corresponds to conic dual constraints,

$$C - (y_1 A_1 + \cdots + y_m A_m) = S, \quad S \succeq 0.$$

Linear matrix functions

Consider a matrix-valued function $F : \mathbb{R}^m \mapsto \mathcal{S}^n$,

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m$$

where $F_i \in \mathcal{S}^n$.

- The inequality

$$F_0 + x_1 F_1 + \cdots + x_m F_m \succeq 0$$

is called a *linear matrix inequality* (LMI).

- Corresponds to conic dual constraints,

$$C - (y_1 A_1 + \cdots + y_m A_m) = S, \quad S \succeq 0.$$

Eigenvalue optimization

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && \gamma I \succeq F(x), \end{aligned}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{aligned} & \text{maximize} && \gamma \\ & \text{subject to} && F(x) \succeq \gamma I, \end{aligned}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma - \lambda \\ & \text{subject to} && \gamma I \succeq F(x) \succeq \lambda I, \end{aligned}$$

Eigenvalue optimization

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && \gamma I \succeq F(x), \end{aligned}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{aligned} & \text{maximize} && \gamma \\ & \text{subject to} && F(x) \succeq \gamma I, \end{aligned}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma - \lambda \\ & \text{subject to} && \gamma I \succeq F(x) \succeq \lambda I, \end{aligned}$$

Eigenvalue optimization

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && \gamma I \succeq F(x), \end{aligned}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{aligned} & \text{maximize} && \gamma \\ & \text{subject to} && F(x) \succeq \gamma I, \end{aligned}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{aligned} & \text{minimize} && \gamma - \lambda \\ & \text{subject to} && \gamma I \succeq F(x) \succeq \lambda I, \end{aligned}$$

Minimizing matrix norms

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathbb{R}^{n \times p}.$$

- (Standard) matrix norm: $\|F(x)\|_2 = \max_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} t & F(x)^T \\ F(x) & t \end{bmatrix} \preceq 0, \end{array}$$

- Nuclear norm: $\|F(x)\|_* = \sum_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & \text{trace}(U + V)/2 \\ \text{subject to} & \begin{bmatrix} U & F(x)^T \\ F(x) & V \end{bmatrix} \preceq 0. \end{array}$$

Minimizing matrix norms

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathbb{R}^{n \times p}.$$

- (Standard) matrix norm: $\|F(x)\|_2 = \max_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} t & F(x)^T \\ F(x) & t \end{bmatrix} \succeq 0, \end{array}$$

- Nuclear norm: $\|F(x)\|_* = \sum_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & \text{trace}(U + V)/2 \\ \text{subject to} & \begin{bmatrix} U & F(x)^T \\ F(x) & V \end{bmatrix} \succeq 0. \end{array}$$

A binary quadratic problem

Binary quadratic problem

We consider a binary problem.

$$\begin{aligned} & \text{minimize} && x^T Q x + c^T x \\ & \text{subject to} && x_i \in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned}$$

where Q can be indefinite.

- Very difficult non-convex problem.
- In general we have to explore 2^n different objectives.
- Instead use a *semidefinite relaxation*.

A binary quadratic problem

Binary quadratic problem

We consider a binary problem.

$$\begin{aligned} & \text{minimize} && x^T Q x + c^T x \\ & \text{subject to} && x_i \in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned}$$

where Q can be indefinite.

- Very difficult non-convex problem.
- In general we have to explore 2^n different objectives.
- Instead use a *semidefinite relaxation*.

A binary quadratic problem

Binary quadratic problem

We consider a binary problem.

$$\begin{aligned} & \text{minimize} && x^T Q x + c^T x \\ & \text{subject to} && x_i \in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned}$$

where Q can be indefinite.

- Very difficult non-convex problem.
- In general we have to explore 2^n different objectives.
- Instead use a *semidefinite relaxation*.

A binary quadratic problem

Binary quadratic problem

We consider a binary problem.

$$\begin{aligned} & \text{minimize} && x^T Q x + c^T x \\ & \text{subject to} && x_i \in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned}$$

where Q can be indefinite.

- Very difficult non-convex problem.
- In general we have to explore 2^n different objectives.
- Instead use a *semidefinite relaxation*.

Lifting of binary constraints

Rewrite binary constraints $x_i \in \{0, 1\}$:

$$x_i^2 = x_i \iff X = xx^T, \quad \text{diag}(X) = x.$$

Still non-convex, since $\text{rank}(X) = 1$.

Semidefinite relaxation of binary constraints

$$X \succeq xx^T, \quad \text{diag}(X) = x.$$

Note that:

$$X \succeq xx^T \iff \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0,$$

which is a linear matrix inequality.

Lifting of binary constraints

Rewrite binary constraints $x_i \in \{0, 1\}$:

$$x_i^2 = x_i \iff X = xx^T, \quad \text{diag}(X) = x.$$

Still non-convex, since $\text{rank}(X) = 1$.

Semidefinite relaxation of binary constraints

$$X \succeq xx^T, \quad \text{diag}(X) = x.$$

Note that:

$$X \succeq xx^T \iff \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0,$$

which is a linear matrix inequality.

Semidefinite relaxation of binary QP

Semidefinite
optimization

Joachim Dahl

The lifted non-convex problem

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && X = xx^T \end{aligned}$$

Conic
optimization

Convex cones

Semidefinite
matricesLinear cone
problemsConic
modeling

Simple cones

Nearest
correlationLinear matrix
inequalitiesEigenvalue
optimization**Combinatorial
relaxations**Sum-of-squares
relaxationsNonnegative
polynomials

Conclusions

The semidefinite relaxation

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxations for combinatorial optimization.

Semidefinite relaxation of binary QP

Semidefinite
optimization

Joachim Dahl

The lifted non-convex problem

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && X = xx^T \end{aligned}$$

Conic
optimization

Convex cones

Semidefinite
matrices

Linear cone
problems

Conic
modeling

Simple cones

Nearest
correlation

Linear matrix
inequalities

Eigenvalue
optimization

Combinatorial
relaxations

Sum-of-squares
relaxations

Nonnegative
polynomials

Conclusions

The semidefinite relaxation

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxations for combinatorial optimization.

Semidefinite relaxation of binary QP

The lifted non-convex problem

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && X = xx^T \end{aligned}$$

The semidefinite relaxation

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxations for combinatorial optimization.

Semidefinite relaxation of binary QP

The lifted non-convex problem

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && X = xx^T \end{aligned}$$

The semidefinite relaxation

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxations for combinatorial optimization.

Semidefinite relaxation of binary QP

The lifted non-convex problem

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && X = xx^T \end{aligned}$$

The semidefinite relaxation

$$\begin{aligned} & \text{minimize} && \langle Q, X \rangle + c^T x \\ & \text{subject to} && \text{diag}(X) = x \\ & && \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxations for combinatorial optimization.

Relaxations for boolean optimization

Same approach used for boolean constraints $x_i \in \{-1, +1\}$.

Lifting of boolean constraints

Rewrite boolean constraints $x_i \in \{-1, 1\}$:

$$x_i^2 = 1 \iff X = xx^T, \quad \text{diag}(X) = e.$$

Semidefinite relaxation of boolean constraints

$$X \succeq xx^T, \quad \text{diag}(X) = e.$$

Sum-of-squares relaxations

- f : multivariate polynomial of degree $2d$.
- $v_d = (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_n^2, \dots, x_n^d)$.
Vector of monomials of degree d or less.

Sum-of-squares representation

f is a sum-of-squares (SOS) iff

$$f(x_1, \dots, x_n) = v_d^T Q v_d, \quad Q \succeq 0.$$

If $X = LL^T$ then

$$f(x_1, \dots, x_n) = v_d^T LL^T v_d = \sum_{i=1}^m (l_i^T v_d)^2.$$

Is obviously **sufficient** for $f(x_1, \dots, x_n) \geq 0$.

Sum-of-squares relaxations

- f : multivariate polynomial of degree $2d$.
- $v_d = (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_n^2, \dots, x_n^d)$.
Vector of monomials of degree d or less.

Sum-of-squares representation

f is a sum-of-squares (SOS) iff

$$f(x_1, \dots, x_n) = v_d^T Q v_d, \quad Q \succeq 0.$$

If $X = LL^T$ then

$$f(x_1, \dots, x_n) = v_d^T LL^T v_d = \sum_{i=1}^m (l_i^T v_d)^2.$$

Is obviously **sufficient** for $f(x_1, \dots, x_n) \geq 0$.

A simple example

Consider

$$f(x, z) = 2x^4 + 2x^3z - x^2z^2 + 5z^4,$$

homogeneous of degree 4, so we only need

$$v = (x^2 \quad xz \quad z^2).$$

Comparing coefficients of $f(x, z)$ and $v^T Q v = \langle Q, v v^T \rangle$,

$$\langle Q, v v^T \rangle = \left\langle \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix}, \begin{pmatrix} x^4 & x^3z & x^2z^2 \\ x^3z & x^2z^2 & xz^3 \\ x^2z^2 & xz^3 & z^4 \end{pmatrix} \right\rangle$$

we see that $f(x, z)$ is SOS iff $Q \succeq 0$ and

$$q_{00} = 2, \quad 2q_{10} = 2, \quad 2q_{20} + q_{11} = -1, \quad 2q_{21} = 0, \quad q_{22} = 5.$$

A simple example

Consider

$$f(x, z) = 2x^4 + 2x^3z - x^2z^2 + 5z^4,$$

homogeneous of degree 4, so we only need

$$v = \begin{pmatrix} x^2 & xz & z^2 \end{pmatrix}.$$

Comparing coefficients of $f(x, z)$ and $v^T Q v = \langle Q, v v^T \rangle$,

$$\langle Q, v v^T \rangle = \left\langle \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix}, \begin{pmatrix} x^4 & x^3z & x^2z^2 \\ x^3z & x^2z^2 & xz^3 \\ x^2z^2 & xz^3 & z^4 \end{pmatrix} \right\rangle$$

we see that $f(x, z)$ is SOS iff $Q \succeq 0$ and

$$q_{00} = 2, \quad 2q_{10} = 2, \quad 2q_{20} + q_{11} = -1, \quad 2q_{21} = 0, \quad q_{22} = 5.$$

Applications in global optimization

$$f(x, z) = 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xz - 4z^2 + 4z^4$$

Global lower bound

Replace non-tractable problem,

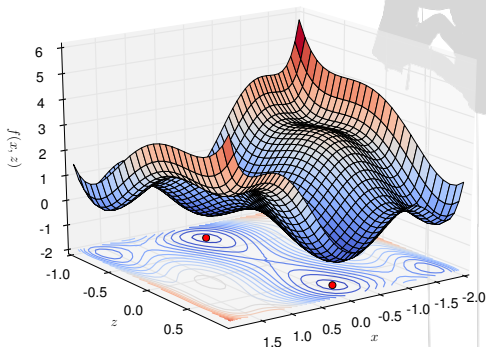
$$\text{minimize } f(x, z)$$

by a tractable lower bound

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & f(x, z) - t \text{ is SOS.} \end{array}$$

Relaxation finds the global optimum $t = -1.031$.

Essentially due to Shor, 1987.



$$f(x, z) - t = 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xz - 4z^2 + 4z^4 - t$$

$$vv^T = \begin{pmatrix} 1 & x & z & x^2 & xz & z^2 & x^3 & x^2z & xz^2 & z^3 \\ x & x^2 & xz & x^3 & x^2z & xz^2 & x^4 & x^3z & x^2z^2 & xz^3 \\ z & xz & z^2 & x^2z & xz^2 & z^3 & x^3z & x^2z^2 & xz^3 & z^4 \\ x^2 & x^3 & x^2z & x^4 & x^3z & x^2z^2 & x^5 & x^4z & x^3z^2 & x^2z^3 \\ xz & x^2z & xz^2 & x^3z & x^2z^2 & xz^3 & x^4z & x^3z^2 & x^2z^3 & xz^4 \\ z^2 & xz^2 & z^3 & x^2z^2 & xz^3 & z^4 & x^3z^2 & x^2z^3 & xz^4 & y^5 \\ x^3 & x^4 & x^3z & x^5 & x^4z & x^3z^2 & x^6 & x^5z & x^4z^2 & x^3z^3 \\ x^2z & x^3z & x^2z^2 & x^4z & x^3z^2 & x^2z^3 & x^5z & x^4z^2 & x^3z^3 & x^2z^4 \\ xz^2 & x^2z^2 & xz^3 & x^3z^2 & x^2z^3 & xz^4 & x^4z^2 & x^3z^3 & x^2z^4 & xz^5 \\ z^3 & xz^3 & z^4 & x^2z^3 & xz^4 & z^5 & x^3z^3 & x^2z^4 & xz^5 & z^6 \end{pmatrix}$$

By comparing coefficients of $v^T Qv$ and $f(x, z) - t$:

$$q_{00} = -t, \quad (2q_{30} + q_{11}) = 4, \quad (2q_{72} + q_{44}) = -\frac{21}{10}, \quad q_{77} = \frac{1}{3}$$

$$2(q_{51} + q_{32}) = 1, \quad (2q_{61} + q_{33}) = -4, \quad (2q_{10,3} + q_{66}) = 4$$

$$2q_{10} = 0, \quad 2q_{20} = 0, \quad 2(q_{71} + q_{42}) = 0, \quad \dots$$

A standard SDP with a 10×10 variable and 27 constraints.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1x + \cdots + c_{2n}x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \quad \iff \quad f(x) = v^T Q v, \quad Q \succeq 0$$

with $v = (1, x, \dots, x^n)$.

- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Nesterov, Y. Squared functional systems and optimization problems, 2000.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1x + \cdots + c_{2n}x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \quad \iff \quad f(x) = v^T Q v, \quad Q \succeq 0$$

with $v = (1, x, \dots, x^n)$.

- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Nesterov, Y. Squared functional systems and optimization problems, 2000.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1x + \cdots + c_{2n}x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \quad \iff \quad f(x) = v^T Q v, \quad Q \succeq 0$$

with $v = (1, x, \dots, x^n)$.

- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Nesterov, Y. Squared functional systems and optimization problems, 2000.

Polynomial interpolation

Fit a polynomial of degree n to a set of points (x_j, y_j) ,

$$f(x_j) = y_j, \quad j = 1, \dots, m,$$

i.e., linear equality constraints in c ,

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Semidefinite shape constraints

- Nonnegativity $f(x) \geq 0$.
- Monotonicity $f'(x) \geq 0$.
- Convexity $f''(x) \geq 0$.

Polynomial interpolation

Fit a polynomial of degree n to a set of points (x_j, y_j) ,

$$f(x_j) = y_j, \quad j = 1, \dots, m,$$

i.e., linear equality constraints in c ,

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Semidefinite shape constraints

- Nonnegativity $f(x) \geq 0$.
- Monotonicity $f'(x) \geq 0$.
- Convexity $f''(x) \geq 0$.

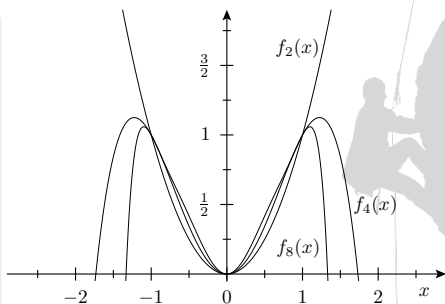
Smooth interpolation

Minimize largest derivative,

$$\begin{array}{ll} \text{minimize} & \max_{x \in [-1, 1]} |f'(x)| \\ \text{subject to} & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1 \end{array}$$

or equivalently

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & -z \leq f'(x) \leq z \\ & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1. \end{array}$$



$$\begin{array}{ll} f_2(x) = x^2 & f_4(x) = \frac{3}{2}x^2 - \frac{1}{2}x^4 \\ f_2'(1) = 2 & f_4'(\frac{1}{\sqrt{2}}) = \sqrt{2} \end{array}$$

Conic optimization

Convex cones

Semidefinite matrices

Linear cone problems

Conic modeling

Simple cones

Nearest correlation

Linear matrix inequalities

Eigenvalue optimization

Combinatorial relaxations

Sum-of-squares relaxations

Nonnegative polynomials

Conclusions

Conic
optimization

Convex cones

Semidefinite
matrices

Linear cone
problems

Conic
modeling

Simple cones

Nearest
correlation

Linear matrix
inequalities

Eigenvalue
optimization

Combinatorial
relaxations

Sum-of-squares
relaxations

Nonnegative
polynomials

Conclusions

Conic optimization

Convex cones

Semidefinite matrices

Linear cone problems

Conic modeling

Simple cones

Nearest correlation

Linear matrix inequalities

Eigenvalue optimization

Combinatorial relaxations

Sum-of-squares relaxations

Nonnegative polynomials

Conclusions

Conic
optimization

Convex cones

Semidefinite
matrices

Linear cone
problems

Conic
modeling

Simple cones

Nearest
correlation

Linear matrix
inequalities

Eigenvalue
optimization

Combinatorial
relaxations

Sum-of-squares
relaxations

Nonnegative
polynomials

Conclusions