

A General Framework for Convex Relaxation of Polynomial Optimization Problems over Cones

Masakazu Kojima, Sunyoung Kim and Hayato Waki

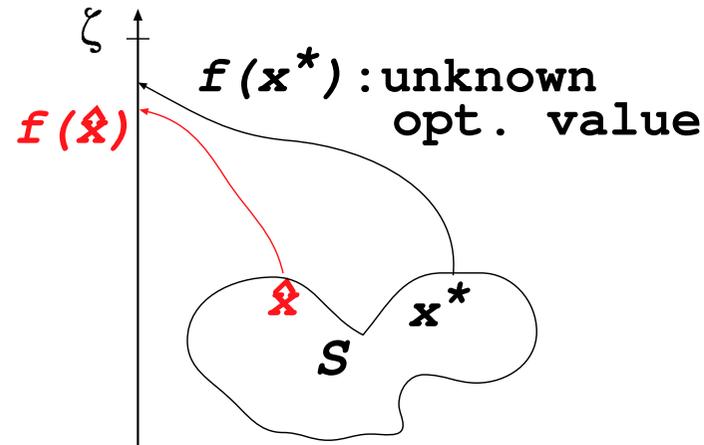
September 2002

Contents

1. Convex relaxation of global optimization problems
2. Existing convex relaxation methods
3. Polynomial optimization problems over cones and their linearization
4. General framework for convex relaxation
5. Basic theory
 - Relation to Lagrangian dual relaxation

1. Convex relaxation of global optimization problems

(1) $\max. f(x)$ sub.to $x \in S$, where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $S \subset \mathbb{R}^n$.



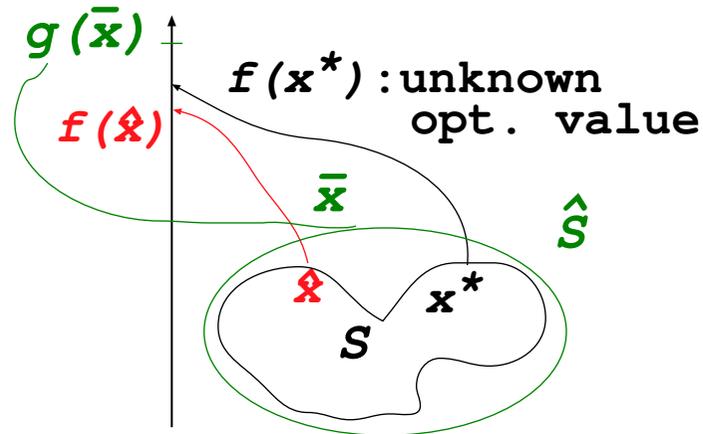
To solve (1) approximately, we need

- (a) a feasible solution $\hat{x} \in S$ with a larger objective value $f(\hat{x})$
 - (b) a smaller upper bound ζ for the unknown optimal value $f(x^*)$
- \implies a main role of convex relaxation

If $\zeta - f(\hat{x})$ is smaller, we can accept \hat{x} as a higher quality approximate optimal solution.

1. Convex relaxation of global optimization problems

(1) $\max. f(x)$ sub.to $x \in S$, where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $S \subset \mathbb{R}^n$.



The common basic idea behind convex relaxation methods is:

- (i) Replace S by a convex set \hat{S} which includes S .
- (ii) Replace f by a linear function g such that $f(x) \leq g(x)$ for $\forall x \in S$.
- (iii) Solve the resulting convex optimization problem

$$\max. g(x) \text{ sub.to } x \in \hat{S}$$

to compute an optimal solution \bar{x} whose objective value $\zeta = g(\bar{x})$ serves as an upper bound for $f(x^*)$.

2. Existing convex relaxation methods

- **One-step methods for 0-1 IPs, nonconvex QPs and polynomial programs**

- (a) SDP-based, *e.g.*, Grötschel-Lovász-Schrijver'88, Shor'90, Goemans-Williamson'95.

- (b) LP-based, *e.g.*, Reformulation-Linearization-Technique (Sherali et.al'92).

- **Successive applications of convex relaxation**

- (c) Lovász-Schrijver'91 for 0-1 IPs, the lift-and-project procedure for 0-1 IPs by Balas-Ceria-Cornuéjols'93.

- (d) SCRM (Successive Convex Relaxation Method) for QOPs by Kojima-Tunçel'00.

- (e) Hierarchical SDP relaxation by Lasserre'01 for polynomial programming.

- Theoretically very powerful: the optimal value can be approximated in arbitrary accuracy by solving a finite number of SDP relaxations under a moderate condition.

- Practically very expensive: we need to solve a sequence of large scale SDPs.

The purpose of this talk is to present

a general framework for convex relaxation methods

which includes most of the existing methods.

The main ingredients are:

(a) Polynomial Optimization Problems \supset QOPs and 0-1 IPs

\Downarrow (b) Add valid constraints and reformulate

(c) Polynomial Optimization Problems over Cones

\Downarrow (d) Linearization

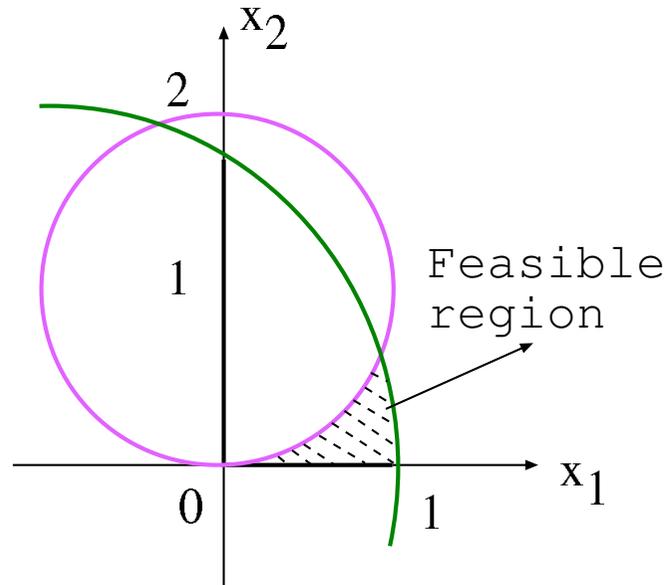
(e) Linear Optimization Problems over Cones

I will talk about

- An illustrative example
- (c) \Rightarrow (d) \Rightarrow (e)
- (b)

An illustrative example

Original problem: max. $-2x_1 + x_2$
sub.to $x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0,$
 $\left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2$ (SOCP constraint)



An illustrative example

$$\begin{aligned} \text{Original problem: max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2 \text{ (SOCP constraint)} \end{aligned}$$

⇓ Valid constraints and/or reformulation

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 \geq 0, x_1x_2 \geq 0, x_2^2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} x_1^2 + x_1 \\ x_1x_2 \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} x_1x_2 + x_2 \\ x_2^2 \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

⇓ Linearization: Keep the linear terms,
but replace **each nonlinear term** by a single independent variable

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} \geq 0, X_{12} \geq 0, X_{22} \geq 0, \\ & X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} X_{11} + x_1 \\ X_{12} \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} X_{12} + x_2 \\ X_{22} \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

An illustrative example

$$\begin{aligned} \text{Original problem: max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2 \text{ (SOCP constraint)} \end{aligned}$$

↓ Valid constraints and/or reformulation

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 \geq 0, x_1x_2 \geq 0, x_2^2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} x_1^2 + x_1 \\ x_1x_2 \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} x_1x_2 + x_2 \\ x_2^2 \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

$$\uparrow X_{11} = x_1x_1, X_{12} = x_1x_2, X_{22} = x_2x_2$$

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} \geq 0, X_{12} \geq 0, X_{22} \geq 0, \\ & X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} X_{11} + x_1 \\ X_{12} \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} X_{12} + x_2 \\ X_{22} \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

3. Polynomial optimization problems over cones and their linearization

max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m ,

$x = (x_1, \dots, x_n)$: a variable vector, $f(x) \equiv (f_1(x), \dots, f_m(x))$,

$f_j(x)$: a polynomial in x_1, \dots, x_n ($j = 0, 1, \dots, m$).

Typical examples of \mathcal{K} : \mathbb{R}_+^m : the nonnegative orthant of \mathbb{R}^m .

\mathbb{S}_+^ℓ : the cone of $\ell \times \ell$ psd symmetric matrices, where we identify each $\ell \times \ell$ matrix as an $\ell \times \ell$ dim vector.

$$\mathbb{N}_p^{1+\ell} \equiv \left\{ v = (v_0, v_1, \dots, v_\ell) \in \mathbb{R}^{1+\ell} : \sum_{i=1}^{\ell} |v_i|^p \leq v_0^p \right\}$$

: the p th order cone ($p \geq 1$).

$\mathbb{N}_2^{1+\ell}$: the second order cone.

When $f_j(x)$ ($j = 0, 1, 2, \dots, m$) are linear,

$\mathcal{K} = \mathbb{S}_+^\ell \Rightarrow$ SDP (Semidefinite Program),

$\mathcal{K} = \mathbb{N}_2^{1+\ell} \Rightarrow$ SOCP (Second-Order Cone Program)

3. Polynomial optimization problems over cones and their linearization

max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m ,

$x = (x_1, \dots, x_n)$: a variable vector, $f(x) \equiv (f_1(x), \dots, f_m(x))$,

$f_j(x)$: a polynomial in x_1, \dots, x_n ($j = 0, 1, \dots, m$).

Linearization — Keep the linear terms, but replace each nonlinear term by a single independent variable.

Example 1:

$$f(x_1, x_2) = \begin{pmatrix} 1 - 2x_1 + 3x_2 + 4x_1^2 + 5x_1x_2 + 6x_2^2 \\ 9 + 8x_1 + 7x_2 + 6x_1^2 - 5x_1x_2 - 4x_2^2 \end{pmatrix} \in \mathcal{K}$$

↓ Linearization

$$\begin{aligned} &F(x_1, x_2, X_{11}, X_{12}, X_{22}) \\ &= \begin{pmatrix} 1 - 2x_1 + 3x_2 + 4X_{11} + 5X_{12} + 6X_{22} \\ 9 + 8x_1 + 7x_2 + 6X_{11} - 5X_{12} - 4X_{22} \end{pmatrix} \in \mathcal{K} \end{aligned}$$

Here the three new variables X_{11} , X_{12} and X_{22} are introduced.

3. Polynomial optimization problems over cones and their linearization

max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m ,

$x = (x_1, \dots, x_n)$: a variable vector, $f(x) \equiv (f_1(x), \dots, f_m(x))$,

$f_j(x)$: a polynomial in x_1, \dots, x_n ($j = 0, 1, \dots, m$).

Linearization — Keep the linear terms, but replace each nonlinear term by a single independent variable.

Example 2:

$$f(x_1, x_2, x_3) = \begin{pmatrix} 1 - 2x_1 + 3x_2 + 4x_1^2x_3 + 5x_1x_2x_3 + 6x_3^4 \\ 9 + 8x_1 + 7x_2 + 6x_1^2x_3 - 5x_1x_2x_3 - 4x_3^4 \end{pmatrix} \in \mathcal{K}$$

↓ Linearization

$$\begin{aligned} & F(x_1, x_2, U, V, W) \\ &= \begin{pmatrix} 1 - 2x_1 + 3x_2 + 4U + 5V + 6W \\ 9 + 8x_1 + 7x_2 + 6U - 5V - 4W \end{pmatrix} \in \mathcal{K} \end{aligned}$$

Here the new variables U , V and W are introduced. In general, we need a systematic method of assigning a new variable to each nonlinear term.

3. Polynomial optimization problems over cones and their linearization

max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m ,

$x = (x_1, \dots, x_n)$: a variable vector, $f(x) \equiv (f_1(x), \dots, f_m(x))$,

$f_j(x)$: a polynomial in x_1, \dots, x_n ($j = 0, 1, \dots, m$).

Linearization — Keep the linear terms, but replace each nonlinear term by a single independent variable.

Systematic method of assigning a new variable to each nonlinear term:

a nonlinear term $x_1^\alpha x_2^\beta \cdots x_n^\zeta \Rightarrow y_{(\alpha, \beta, \dots, \zeta)} \in \mathbb{R}$ a new variable

For example,

$$n = 5, \quad x_1^2 x_2 x_3^3 x_5^4 = x_1^2 x_2^1 x_3^3 x_4^0 x_5^4 \Rightarrow y_{(2,1,3,0,4)}.$$

In theory, any method of assigning a new variable to each nonlinear term works. \Rightarrow This method is not essential.

4. General framework for convex relaxation

Original QOP, 0-1 IP, Polynomial programs to be solved

↓ Valid constraints and/or reformulate

POP: max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m ,

$x = (x_1, \dots, x_n)$: a variable vector, $f(x) \equiv (f_1(x), \dots, f_m(x))$,

$f_j(x)$: a polynomial in x_1, \dots, x_n ($j = 0, 1, \dots, m$).

↓ Linearization — Keep the linear terms, but replace each
↓ nonlinear term by a single independent variable.

LOP: max. $F_0(x, y)$ sub.to $F(x, y) \in \mathcal{K}$, where

y denotes a new variable vector whose elements correspond to nonlinear terms appeared in the polynomials $f_j(x)$ ($j = 0, 1, \dots, m$).

An illustrative example

$$\begin{aligned} \text{Original problem: max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2 \text{ (SOCP constraint)} \end{aligned}$$

⇓ Valid constraints and/or reformulation

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 \geq 0, x_1x_2 \geq 0, x_2^2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} x_1^2 + x_1 \\ x_1x_2 \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} x_1x_2 + x_2 \\ x_2^2 \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

⇓ Linearization: Keep the linear terms,
but replace **each nonlinear term** by a single independent variable

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} \geq 0, X_{12} \geq 0, X_{22} \geq 0, \\ & X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} X_{11} + x_1 \\ X_{12} \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} X_{12} + x_2 \\ X_{22} \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

An illustrative example

$$\begin{aligned} \text{Original problem: } \max. \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2 \text{ (SOCP constraint)} \end{aligned}$$

⇓ Valid constraints and/or reformulation

$$\begin{aligned} \max. \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1x_2 \\ x_2 & x_1x_2 & x_2^2 \end{pmatrix} \succeq O. \end{aligned}$$

⇓ Linearization

$$\begin{aligned} \max. \quad & -2x_1 + x_2 \quad \text{--- SDP} \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & X_{11} & X_{12} \\ x_2 & X_{12} & X_{22} \end{pmatrix} \succeq O. \end{aligned}$$

Given a problem, there are various ways of adding valid constraints and reformulating the problem. They usually yield different convex relaxations.

In the previous illustrative example:

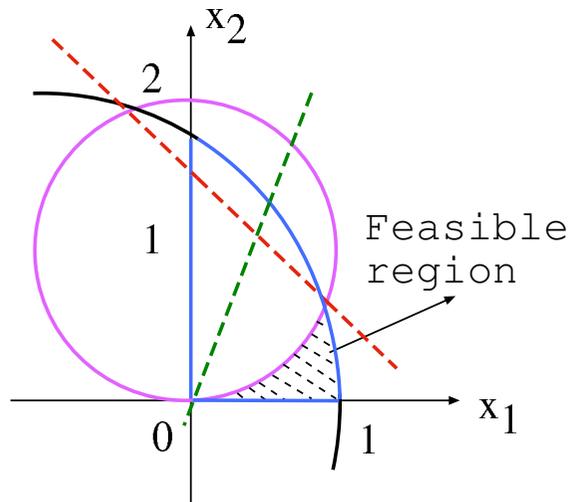
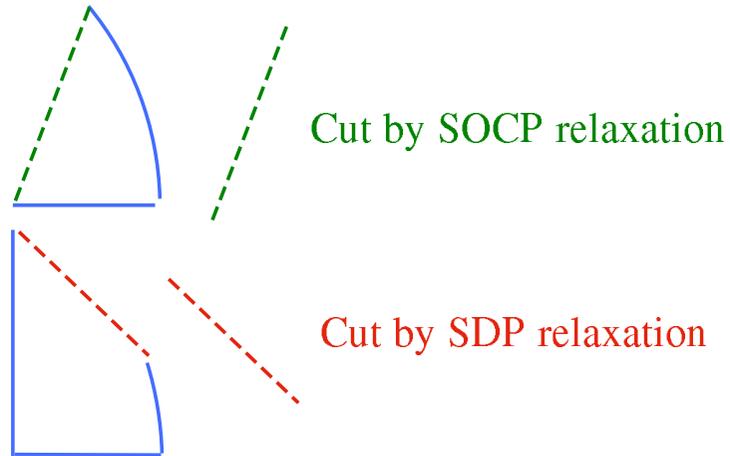
$$\begin{aligned} \text{Original problem: max.} \quad & -2x_1 + x_2 \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2 \text{ (SOCP constraint)} \end{aligned},$$

we obtained two distinct convex relaxations.

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 && \text{--- SOCP} \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} \geq 0, X_{12} \geq 0, X_{22} \geq 0, \\ & X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \left\| \begin{pmatrix} X_{11} + x_1 \\ X_{12} \end{pmatrix} \right\| \leq 2x_1, \left\| \begin{pmatrix} X_{12} + x_2 \\ X_{22} \end{pmatrix} \right\| \leq 2x_2. \end{aligned}$$

$$\begin{aligned} \text{max.} \quad & -2x_1 + x_2 && \text{--- SDP} \\ \text{sub.to} \quad & x_1 \geq 0, x_2 \geq 0, X_{11} + X_{22} - 2x_2 \geq 0, \\ & \left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2, \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & X_{11} & X_{12} \\ x_2 & X_{12} & X_{22} \end{pmatrix} \succeq O. \end{aligned}$$

Original problem: max. $-2x_1 + x_2$
sub.to $x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 - 2x_2 \geq 0,$
 $\left\| \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \right\| \leq 2$ (SOCP constraint) ,



Some examples of valid constraints — 1

- Universally valid constraints.

(a) SDP type:

$$u(x)u(x)^T = \begin{pmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{pmatrix} \succeq O,$$

where $u(x) = (1 \ x_1 \ x_2 \ x_1^2 \ x_1x_2 \ x_2^2)^T$

More generally, take a row vector consisting of a basis of the polynomials in x_1, \dots, x_n with degree ℓ for $u(x)$. [Lasserre'01].

(b) SOCP (Second-Order Cone Programming) type:

$$\forall f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \left\| \begin{pmatrix} f_1(x)^2 - f_2(x)^2 \\ 2f_1(x)f_2(x) \end{pmatrix} \right\| \leq f_1(x)^2 + f_2(x)^2$$

Some examples of valid constraints — 2

- Deriving valid constraints, “multiplication” of valid constraints:

original constraints	new constraints
$\mathbb{R} \ni f(x) \geq 0, \mathbb{R} \ni g(x) \geq 0$	$f(x)g(x) \geq 0$ [Sherali et.al'92]
$f(x) \succeq 0, G(x) \succeq O$	$f(x)G(x) \succeq 0$ [Lasserre'01]

$F(x) \succeq O, G(x) \succeq O$	$\Rightarrow F(x) \otimes G(x) \succeq 0$ (Kronecker product)
$\left. \begin{array}{l} \ f(x)\ \leq f_0(x), f(x) \in \mathbb{R}^\ell \\ \ g(x)\ \leq g_0(x), g(x) \in \mathbb{R}^\ell \end{array} \right\}$	$\Rightarrow \ f(x) \circ g(x)\ \leq f_0(x)g_0(x)$
(SOCP constraints)	(component-wise product)

5. Basic theory

POP: max. $f_0(x)$ sub.to $f(x) \in \mathcal{K}$, where

\mathcal{K} : a closed convex cone in \mathbb{R}^m , $f(x) \equiv (f_1(x), \dots, f_m(x))$.

↓ Linearization — Keep the linear terms, but replace each
↓ nonlinear term by a single independent variable.

LOP: max. $F_0(x, y)$ sub.to $F(x, y) \in \mathcal{K}$, where

y denotes a new variable vector whose elements correspond to nonlinear terms appeared in the polynomials $f_j(x)$ ($j = 0, 1, \dots, m$).

Lagrangian funct: $L(x, v) \equiv f_0(x) + \sum_{j=1}^m v_j f_j(x)$ for $\forall x \in \mathbb{R}^n, v \in \mathcal{K}^*$

Under the Slater condition ($\exists x; f(x) \in \text{int } \mathcal{K}$), if $\bar{\zeta}$ is the optimal value of LOP then there exists $\bar{v} \in \mathcal{K}^*$ satisfying

$$L(x, \bar{v}) = \bar{\zeta} \text{ for } \forall x \in \mathbb{R}^n$$

Hence $\bar{\zeta} = \max\{L(x, \bar{v}) : x \in \mathbb{R}^n\}$ (a Lagrangian relaxation)
 $\geq \min_{v \in \mathcal{K}^*} \max\{L(x, v) : x \in \mathbb{R}^n\}$ (Lagrangian dual relaxation)

6. Concluding remarks

The framework proposed in this talk for convex relaxation is **quite general**.

But we need to investigate **various issues**.

- Effectiveness — How do we generate better bounds?
- Low cost — Resulting relaxed problems need to be solved cheaply
- How to combine this framework with other methods like the branch-and-bound method
- Parallel computation?