

6. Relaxations of integer programs

Version: 20-11-2011

Contents: We consider LP-, Lagrangian-, semidefinite programming- and completely positive programming relaxations.

- For some results we refer to [FKS].

Rough idea:

- Replace (relax) an NP-hard integer program:

$$(IP) \quad \min_x f(x) \quad \text{s.t.} \quad x \in \mathcal{F}(IP)$$

- by an (easier) continuous convex program:

$$(R) \quad \min_x f(x) \quad \text{s.t.} \quad x \in \mathcal{F}(R)$$

6.1 0-1 programs:

LP- and (partial) Lagrange relaxation

We consider the *integer program*: (with $Q \succeq 0$)

$$(IP) \quad \min_x x^T Q x + 2c^T x \quad \text{s.t.} \\ Ax \leq b, \quad Bx \leq d, \\ \boxed{x \in \{0, 1\}^n}$$

with min-value $v(IP)$ and feasible set $\mathcal{F}(IP)$.

LP relaxation: Only relax $x_i \in \{0, 1\}$ to $x_i \in [0, 1]$.

$$(LP) \quad \min_x x^T Q x + 2c^T x \quad \text{s.t.} \\ Ax \leq b, \quad Bx \leq d, \\ \boxed{x \in [0, 1]^n}$$

Rem. Since $\mathcal{F}(IP) \subset \mathcal{F}(LP)$ it follows $v(LP) \leq v(IP)$.

(Partial) Lagrangian relaxation (Relax the “difficult” constraint $Ax \leq b$): For $y \geq 0$ we define the function

$$L(y) = \min_x \{x^T Qx + 2c^T x + y^T (Ax - b) \mid Bx \leq d, x \in \{0, 1\}^n\}$$

and the Lagrangian-dual: $(D_p) \quad \max_{y \geq 0} L(y)$

Ex.6.1 (“Weak duality”) For any $y \geq 0$ we have $L(y) \leq v(IP)$. Thus $v(D_p) \leq v(IP)$.

PROOF. $y \geq 0, Ax - b \leq 0$ implies $y^T (Ax - b) \leq 0$. So with $f(x) := x^T Qx + 2c^T x$

$$\begin{aligned} v(IP) &= \min_x \{f(x) \mid Ax \leq b, Bx \leq d, x \in \{0, 1\}^n\} \\ &\geq \min_x \{f(x) + y^T (Ax - b) \mid Bx \leq d, x \in \{0, 1\}^n\} \\ &= L(y) \end{aligned}$$

By taking \max_y we obtain $v(D_p) \leq v(IP)$.

By applying the strong duality result for CO (Th.2 in Chap 2) we obtain:

Th.6.1 Let $Q \succeq 0$ and let the system $Ax \leq b$, $Bx \leq d$, $x \in [0, 1]^n$ be feasible. Then

$$v(LP) \leq v(D_p) \leq v(IP)$$

PROOF. We prove $v(D_p) \geq v(LP)$: With $f(x) := x^T Qx + 2c^T x$

$$\begin{aligned} v(D_p) &= \max_{y \geq 0} \min_x \{f(x) + y^T (Ax - b) \mid Bx \leq d, \mathbf{x} \in \{0, 1\}^n\} \\ &\geq \max_{y \geq 0} \underbrace{\min_x \{f(x) + y^T (Ax - b) \mid Bx \leq d, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}}_{(*)} \end{aligned}$$

We apply “strong duality” to (the primal problem) $(*)$ with L -multipliers:

$$-x \leq 0 \Leftrightarrow \mu, \quad x - 1 \leq 0 \Leftrightarrow \lambda, \quad Bx - d \leq 0 \Leftrightarrow u;$$

Slater holds: $Bx - d \leq 0$, $0 \leq x \leq 1$ is feasible.

and find that the last right-hand-side equals

$$\begin{aligned}v(D_p) &\geq r-h-s \\ &= \max_{y \geq 0} (\max_{\mu, \lambda, u \geq 0} \min_x) \left\{ f(x) + y^T (Ax - b) + \right. \\ &\quad \left. u^T (Bx - d) - \mu^T x + \lambda^T (x - 1) \right\} \\ &= \min_x \max_{y, \mu, \lambda, u \geq 0} \left\{ f(x) + y^T (Ax - b) + \right. \\ &\quad \left. u^T (Bx - d) - \mu^T x + \lambda^T (x - 1) \right\} \\ &= \min_x \{ f(x) \mid Ax - b \leq 0, Bx - d \leq 0, 0 \leq x \leq 1 \} \\ &= v(LP)\end{aligned}$$

Here for the second “=” we applied strong duality “back” (wrt. all constraints). The Slater condition for this program holds since by assumption there is a feasible point satisfying all (linear) constraints $Ax \leq b$, $Bx \leq d$, $0 \leq x \leq 1$.

6.2 Semidefinite relaxation of quadratic problems

Notation:

- $\mathcal{S}^n = \{A \in \mathbb{R}^{n \times n} \mid A = A^T\}$ (symmetric matrices)
- $\mathcal{S}_+^n = \{A \in \mathcal{S}^n \mid A \succeq 0\}$ (pos. semidef. matrices)

We consider the *quadratic program*

$$(P_q) \quad \min q_0(x) \quad \text{s.t.} \quad q_j(x) \leq 0 \quad \forall j \in J$$

with $J = \{1, \dots, m\}$, $q_j(x) = x^T Q_j x + 2b_j^T x + c_j$. ($Q_j \in \mathcal{S}^n$)

Lifting procedure: Introduce $X = xx^T$ and use

- $x^T Q x = Q \bullet xx^T$
- $x^T Q x + 2b^T x + c = \begin{pmatrix} c & b^T \\ b & Q \end{pmatrix} \bullet \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix}$
- $X = xx^T \Leftrightarrow \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix}^T$

In the variables x, X : (P_q) is equivalent with

$$\begin{aligned}
 (P'_q) \quad \min_{x, X} \quad & \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix} \bullet \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} & \text{s.t.} \\
 & \begin{pmatrix} c_j & b_j^T \\ b_j & Q_j \end{pmatrix} \bullet \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \leq 0 \quad \forall j \in J \\
 & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix}^T \quad (*)
 \end{aligned}$$

Remark. (P'_q) has a linear objective, linear inequalities but a quadratic equality constraint $(*)$ (nonconvex!).

Semidefinite Relaxation By relaxing $(*)$ to $\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0$ we obtain the (convex) **SDP-relaxation** of (P'_q) :

$$\begin{aligned}
 (P_{SDP}) \quad \min_{x, X} \quad & \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix} \bullet \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} & \text{s.t.} \\
 & \begin{pmatrix} c_j & b_j^T \\ b_j & Q_j \end{pmatrix} \bullet \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \leq 0 \quad \forall j \in J \\
 & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0
 \end{aligned}$$

Ex.6.2 Show that the feasible set $\mathcal{F}(P_{SDP})$ of (P_{SDP}) is convex and (P_{SDP}) is a convex program. Show further:

$$X - xx^T \succeq 0 \Leftrightarrow \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0.$$

Ex.6.3 (P_{SDP}) is a relaxation of (P'_q) , i.e., $\mathcal{F}(P'_q) \subset \mathcal{F}(P_{SDP})$ and thus $v(P_{SDP}) \leq v(P'_q) = v(P_q)$

Ex.6.4 Let be given a quadratic constraint (with $Q \succeq 0$):

$$x^T Q x + 2b^T x + c \leq 0 \quad (*)$$

(a) Show: If $X - xx^T \succeq 0$ holds then

$$x^T Q x + 2b^T x + c \leq Q \bullet X + 2b^T x + c$$

with some X we have:

(b) Show: x satisfies $(*) \Leftrightarrow X - xx^T \succeq 0$ and $Q \bullet X + 2b^T x + c \leq 0$

As an implication of Ex. 6.4 it follows:

Cor. 6.1 Let (P_q) be convex (i.e., $Q_j \succeq 0$). Then the SDP relaxation (P_{SDP}) is exact: $v(SDP) = v(P_q)$

The Lagrangian dual of (P_q) : By defining for $u \geq 0$ the Lagrange-min-function wrt. (P_q) ,

$$L(u) := \min_x \{q_0(x) + \sum_j u_j q_j(x)\}$$

the Lagrangian dual of (P_q) becomes

$$(D) \quad \max_{u \geq 0} L(u)$$

Rem. 6.1 (D) can be shown to be the semidefinite dual max-problem of the primal (P_{SDP}) (see also [FKS, p195,321]). So by weak SDP-duality: $v(D) \leq v(P_{SDP})$.

Summarizing:

$$v(D) \leq v(P_{SDP}) \leq v(P_q)$$

Rem. The same holds for equality constraints $q_j(x) = 0$. Here, the L-multiplier $u \in \mathbb{R}^m$ is not constrained in $L(u)$.

Semidefinite relaxation of integer programs

Basic idea: Write the 0-1 condition $x_i \in \{0, 1\}$ equivalently as quadratic constraint $x_i^2 - x_i = 0$

We consider the 0-1 program:

$$\begin{aligned} (IP) \quad & \min_x x^T Q x + 2c^T x && \text{s.t.} \\ & Ax \leq b \\ & x_i^2 - x_i = 0 \quad \forall i \end{aligned}$$

Using again $x^T Q x = Q \bullet xx^T$, $X = xx^T$, the condition $x_i^2 = x_i$ becomes $X_{ii} = x_i^2 = x_i$ and the SDP-relaxation of (IP) reads:

$$\begin{aligned} (P_{SDP}) \quad & \min_{x, X} Q \bullet X + 2c^T x && \text{s.t.} \\ & Ax \leq b \\ & X_{ii} - x_i = 0 \quad \forall i \\ & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{aligned}$$

As in the general quadratic case we consider the (full) L-dualization:

$$(D) \quad \max_{y \geq 0, u} L(y, u)$$

with dual function

$$L(y, u) = \min_x \{ x^T Q x + 2c^T x + y^T (Ax - b) + \sum_i u_i (x_i^2 - x_i) \}$$

Again, (D) is the SDP dual of (P_{SDP}) so that:

$$(\star) \quad v(D) \leq v(P_{SDP}) \leq v(IP)$$

Case that Q is positive semidefinite

LP-relaxation (Note: $x_i \in [0, 1] \Leftrightarrow x_i^2 - x_i \leq 0$)

$$\begin{aligned} (LP) \quad & \min_x x^T Qx + 2c^T x && \text{s.t.} \\ & Ax \leq b \\ & (x_i \in [0, 1] \Leftrightarrow) \quad x_i^2 - x_i \leq 0 \quad \forall i \end{aligned}$$

By applying strong duality we can prove

Th. 6.2 Let $Q \succeq 0$ and assume that (LP) has a Slater point \bar{x} , i.e., $A\bar{x} \leq b$, $\bar{x}_i^2 - \bar{x}_i < 0$. Then $v(LP) \leq v(D)$.

PROOF: Letting $f(x) := x^T Qx + 2c^T x$ and $L(x, y, u) = f(x) + y^T (Ax - b) + \sum_i u_i (x_i^2 - x_i)$ we find:

$$v(D) = \max_{y \geq 0, \boxed{u \in \mathbb{R}^n}} \min_x L(x, y, u)$$

$$\geq \max_{y \geq 0, \boxed{u \geq 0}} \min_x L(x, y, u)$$

$$\text{apply str. dual.} \rightarrow = \min_x \max_{y \geq 0, u \geq 0} L(x, y, u)$$

$$= \min_x \{f(x) \mid Ax - b \leq 0, x_i^2 - x_i \leq 0 \forall i\}$$

$$= v(LP)$$

Summarising in this case:

$$v(LP) \leq v(D) \leq v(P_{SDP}) \leq v(IP)$$

6.3 Linear 0-1 programs

We consider as special case the 0-1 linear program:

$$(IP) \quad \min_x c^T x \quad \text{s.t.} \\ Ax \leq b \\ x_i^2 - x_i = 0 \quad \forall i \quad (\Leftrightarrow x_i \in \{0, 1\})$$

SDP-relaxation: $(P_{SDP}) \quad \min_{x, X} c^T x \quad \text{s.t.}$

$$Ax \leq b \\ X_{ii} - x_i = 0 \quad \forall i \\ \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0$$

(full) L-dualization: $(D) \quad \max_{y \geq 0, u} L(y, u)$

with dual function

$$L(y, u) = \min_x \left\{ c^T x + y^T (Ax - b) + \sum_i u_i (x_i^2 - x_i) \right\}$$

LP-relaxation:

$$(LP) \quad \min_x \quad c^T x \quad \text{s.t.} \\ Ax \leq b \\ x_i \in [0, 1] \quad \forall i$$

Remark: Obviously in (P_{SDP}) the elements X_{ij} , $i \neq j$ of X do not play any role so that $X_{ij} = x_i$ gives

$$X \succeq xx^T \Leftrightarrow X_{ii} \geq x_i^2 \Leftrightarrow x_i \geq x_i^2 \Leftrightarrow x_i \in [0, 1]$$

So (P_{SDP}) is equivalent to the LP-relaxation, *i.e.*,
 $v(LP) = v(P_{SDP})$ and thus by Rem. 6.1, p9:

Summarizing: For the linear IP we have (*without any assumptions*):

$$v(LP) = v(D) = v(P_{SDP}) \leq v(IP)$$

Dualize only $Ax \leq b$: We also consider the partial Lagrange dualization: for $y \geq 0$ we define (the partial) **L-function**

$$\begin{aligned} L_1(y) &:= \min_{x_i \in \{0,1\}} \{c^T x + y^T (Ax - b)\} \\ &= \min_{x_i \in \{0,1\}} \{-b^T y + x^T (A^T y + c)\} \\ &= -b^T y + \sum_i \min\{0, (A^T y + c)_i\} \end{aligned}$$

Here we used the obvious relation:

$$\min_{x_i \in \{0,1\}} d^T x = \min_{x_i \in \{0,1\}} \sum_i d_i x_i = \sum_i \min\{0, d_i\}$$

The corresponding L-dual is:

$$(D_1) \quad \max_{y \geq 0} L_1(y)$$

We now compare $L_1(y)$ with the full L-dual function

$$\begin{aligned} L(y, u) &= \min_x \{ \mathbf{c}^T \mathbf{x} + \mathbf{y}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \\ &\quad + \sum_i u_i (x_i^2 - x_i) \} \\ &= \min_x \{ -\mathbf{b}^T \mathbf{y} + \sum_i [x_i (\mathbf{A}^T \mathbf{y} + \mathbf{c} - \mathbf{u})_i \\ &\quad + u_i x_i^2] \} \end{aligned}$$

Ex.6.5 Let $q(\alpha, u)$, $\alpha, u \in \mathbb{R}$ be defined by

$$q(\alpha, u) := \min_{x \in \mathbb{R}} (\alpha - u)x + ux^2.$$

Show

$$q(\alpha, u) = \begin{cases} -\frac{(\alpha-u)^2}{4u} & \text{if } u > 0 \\ 0 & \text{if } u = \alpha = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\text{and: } \max_u q(\alpha, u) = \begin{cases} \alpha & \text{if } \alpha < 0 \\ 0 & \text{if } \alpha \geq 0 \end{cases} = \min\{0, \alpha\}$$

Using Ex.6.5 we find

Lem.6.1 $L_1(y) = \max_u L(y, u)$ so that $v(D_1) = v(D)$.

PROOF: $\max_u L(y, u) =$

$$= -b^T y + \sum_i \max_{u_i} \min_{x_i} \{x_i [(A^T y)_i + c_i - u_i] + u_i x_i^2\}$$

Ex. 6.5 \rightarrow $= -b^T y + \sum_i \min\{0, (A^T y)_i + c_i\}$

$$= L_1(y)$$

By taking $\max_{y \geq 0}$ we obtain $v(D_1) = v(D)$. \square

Altogether without any assumption (using (\star) on p.11) we find.

Cor. 6.2. For linear (IP) the relations hold:

$$v(LP) = v(D_1) = v(D) = v(P_{SDP}) \leq v(IP)$$

6.5 Semidefinite relaxation of max-cut

A randomized algorithm

Maxcut problem: Let $G = (V, E)$ be an undirected graph with $n = |V|$ nodes and $m = |E|$ edges. Each edge (i, j) is assigned a nonnegative weight w_{ij} with $w_{ij} = w_{ji}$ and $w_{ii} = 0$. The *maxcut problem* asks for a set $S \subseteq V$ of nodes such that the weight of the associated cut is maximized:

$$\max_{S \subseteq V} w(S), \quad \text{where} \quad w(S) = \sum_{i \in S, j \notin S} w_{ij}.$$

IP formulation: For $S \subseteq V$ we define a vector $y = (y_1, \dots, y_n)$ via

$$y_i = \begin{cases} 1 & \text{if } i \in S \\ -1 & \text{if } i \notin S. \end{cases}$$

Then

$$(1 - y_i y_j) = \begin{cases} 2 & \text{if } i \in S, j \notin S \text{ or } j \in S, i \notin S \\ 0 & \text{otherwise} \end{cases}$$

The maxcut problem is thus equivalent with

$$\max_{y \in \mathbb{R}^n} \frac{1}{4} \sum_{i,j} w_{ij}(1 - y_i y_j) \quad \text{s.t. } y_i \in \{-1, 1\} .$$

or (skipping the constant) equivalent with:

$$(IP) \quad \min_{y \in \mathbb{R}^n} \frac{1}{4} y^T W y \quad \text{s.t. } y_i \in \{-1, 1\} .$$

SDP relaxation: Writing $y^T W y = W \bullet yy^T$, putting $Y = yy^T$, using $y_i = \pm 1 \Leftrightarrow y_i^2 = 1$ we find $Y_{ii} = 1$. Replace $Y = yy^T$ by $Y \succeq 0$ to obtain the relaxation of (IP):

$$(SDP) \quad \min_Y W \bullet Y \quad \text{s.t. } Y_{ii} = 1, \forall i \\ Y \succeq 0$$

Rem. By an Ex. in Chap 3, any $n \times n$ -matrix $Y \succeq 0$ can be decomposed as $Y = U^T U$ with U a $n \times n$ -matrix.

Goemans and Williamson proposed the following algorithmic approach to the maxcut problem.

Randomized Algorithm for the Maxcut Problem:

- Compute a solution Y of the problem (SDP).
- Compute the factorization $Y = U^T U$ with u_i the columns of U .
- Choose randomly a vector r , and set $S = \{i \in V \mid u_i^T r \geq 0\}$.

Let v_{opt} be the max-value of a maximum cut in G . It is shown by Goemans and Williamson (1995) that the solution S obtained by the randomized algorithm has an expected value:

$$E(v(S)) \geq \alpha v_{opt} \quad \text{with} \quad \alpha > 0.878 .$$

6.6 Completely positive relaxation of integer programs

We reconsider the 0-1 program

$$\begin{aligned} (IP) \quad \min_x \quad & x^T Q x + 2c^T x && \text{s.t.} \\ & a_j^T x = b_j \quad j = 1, \dots, m \\ & x \geq 0 \\ & x_i^2 - x_i = 0 \quad \forall i \end{aligned}$$

Basic idea: As in the SDP relaxation we introduce $X = xx^T$ and we add the redundant conditions

$$(a_j^T x)^2 = a_j^T x x^T a_j = b_j^2 \quad \text{or} \quad \boxed{a_j^T X a_j = b_j^2} .$$

(IP) can then equivalently be written as:

$$\begin{aligned}
 (P') \quad & \min_{x, X} Q \bullet X + 2c^T x \quad \text{s.t.} \\
 & a_j^T X a_j = b_j^2, \quad a_j^T x = b_j \quad j = 1, \dots, m \\
 & x \geq 0 \\
 & X_{ii} - x_i = 0 \quad \forall i \\
 & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix}^T
 \end{aligned}$$

We again obtain a semi-definite relaxation:

(P_{SDP}) program as in (P') but

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix}^T \quad \text{replaced by} \quad \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0$$

Ex. 6.6 Show that the set \mathcal{S}_+^{n+1} of psd matrices can be written as:

$$\mathcal{S}_+^{n+1} = \left\{ \mathbf{A} = \sum_{k=1}^N \mathbf{a}_k \mathbf{a}_k^T \mid \mathbf{a}_k \in \mathbb{R}^{n+1}, N \text{ finite} \right\}$$

completely positive matrices (CP): Consider now the set \mathcal{C}^{n+1} of CP-matrices:

$$\mathcal{C}^{n+1} := \left\{ \mathbf{A} = \sum_{k=1}^N \mathbf{a}_k \mathbf{a}_k^T \mid \mathbf{a}_k \in \mathbb{R}^{n+1}, N \text{ finite}, \boxed{\mathbf{a}_k \geq \mathbf{0}} \right\}$$

Obviously: For $\mathbf{x} \geq \mathbf{0}$

$$\begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}^T \in \mathcal{C}^{n+1} \subset \mathcal{S}_+^{n+1}$$

So the original program (IP) (or P') can be replaced by the (stronger) CP -relaxation

(P_{CP}) program as in (P') (skipping $x \geq 0$) and
 $\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot \begin{pmatrix} 1 & x^T \end{pmatrix}$ replaced by $\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{C}^{n+1}$

Rem. Both, (P_{SDP}), (P_{CP}) are convex programs and obviously:

$$v(P_{SDP}) \leq v(P_{CP}) \leq v(IP)$$

In: “**Burer**, *On the copositive representation of binary and continuous nonconvex quadratic programs*, *Math. Program.* 120 (2009), no. 2, Ser. A, 479–495.” the following is shown:

Under a condition on (IP) which can be satisfied by introducing an extra slack variable we have

$$v(P_{CP}) = v(IP)$$

Some final remarks on relaxations

- LP- and Lagrange-relaxation are classical methods for (approximately) solving IP.
- SDP-relaxation are studied extensively since ≈ 20 years and often lead to sharper approximations. SDP is “polynomially approximable.”
- The recent completely positive relaxations (CP) represent continuous convex programs. CP describes some NP-hard IP’s exactly.
 - Therefore, CP must be NP-hard
 - Not every convex program is easy.
 - The convex structure of CO hopefully helps to solve IP’s “more efficiently”.