Problem Set 1 Solutions

Sum of Squares Seminar

September 29, 2017

Problem 1 (10 points)

Recall that given equations $s_1(x_1, \cdots, x_n) = 0, s_2(x_1, \cdots, x_n) = 0$, etc., a degree d Positivstellensatz proof of infeasibility is an expression of the form $-1 = \sum_i f_i s_i + \sum_j g_j^2$ where

1. $\forall i, deg(f_i) + deg(s_i) \leq d$
2. $\forall j, deg(g_j) \leq \frac{d}{2}$

Give a degree 2 Positivstellensatz proof that a square has no triangle, i.e. that the following equations are unsatisfiable over $\mathbb{R}$:

1. $\forall i \in [1, 4], x_i^2 - x_i = 0$
2. $x_1 x_3 = x_2 x_4 = 0$
3. $\left(\sum_{i=1}^{4} x_i\right) - 3 = 0$

Solution

Observe that

$$(x_1 + x_3 - x_2 - x_4)^2 = 2\sum_{i=1}^{4} x_i^2 - (x_1 + x_2 + x_3 + x_4)^2 + 4x_1x_3 + 4x_2x_4$$

By the problem equations,

1. $2\sum_{i=1}^{4} x_i^2 = 2\sum_{i=1}^{4} x_i = 6$
2. $-(x_1 + x_2 + x_3 + x_4)^2 = -9$
3. $4x_1x_3 = 4x_2x_4 = 0$

so $(x_1 + x_3 - x_2 - x_4)^2 = -3$ which is a contradiction.

This argument can be put in the desired form as follows:

$$-1 = (x_1 + x_3 - x_2 - x_4)^2 + 2 - 2\sum_{i=1}^{4}(x_i^2 - x_i) - 2\left(\sum_{i=1}^{4} x_i - 3\right)$$
$$+ \left(\sum_{i=1}^{4} x_i + 3\right)\left(\sum_{i=1}^{4} x_i - 3\right) - 4x_1x_3 - 4x_2x_4$$

Choose one of the following two problems:

Problem 2: Linear Programming for Vertex Cover (20 points)

The vertex cover problem is as follows. Given an input graph G, what is the smallest size of a subset V of vertices of G such that every edge of G is incident with at least one vertex of V?

(a) 10 points: Give a linear programming relaxation of the vertex cover problem.

(b) 10 points: How much smaller can the value of your linear programming relaxation be than the true size of the minimal vertex cover? In other words, letting $opt(G)$ be the true size of the minimial vertex cover of G and letting $LP(G)$ be the value the linear program gives, what is the largest α such that for all input graphs G, $\alpha \cdot opt(G) \leq LP(G) \leq opt(G)$? Prove that your answer is correct by giving examples showing that $\frac{LP(G)}{opt(G)}$ can be arbitrarily close to α and proving that for all graphs G, $opt(G) \leq \frac{1}{\alpha}LP(G)$. (Hint is available)

Solution

Part a: The actual problem can be reformulated as follows. Minimize $\sum_{i=1}^{n} x_i$ subject to

1. $\forall i, x_i \in \{0, 1\}$
2. $\forall (i, j) \in E(G), x_i + x_j \geq 1$

A natural linear programming relaxation is as follows. Minimize $\sum_{i=1}^{n} x_i$ subject to

1. $\forall i, 0 \leq x_i \leq 1$
2. $\forall (i, j) \in E(G), x_i + x_j \geq 1$

Part b: $\alpha = \frac{1}{2}$. To see this, consider $\{K_n\}$, the family of complete graphs on n vertices. The linear programming relaxation can take $x_i = \frac{1}{2}$ for all i, giving a value of $\frac{n}{2}$. However, the minimal size of a vertex cover of K_n is $n - 1$ and $\lim_{n\to\infty} \frac{\frac{n}{2}}{n-1} = \frac{1}{2}$

To see that there is always a vertex cover of size at most twice the value given by the linear programming relaxation, consider the following rounding algorithm. Take V to be the set of vertices which are given value at least $\frac{1}{2}$ by the linear program. For all $(i, j) \in E(G)$, $x_i \geq \frac{1}{2}$ or $x_j \geq \frac{1}{2}$, so V is a vertex cover of G. The rounding algorithm multiplies the weight on each vertex by at most 2, so $|V|$ has size at most twice the value given by the linear programming relaxation, as needed.

Problem 3: Linear Programming for Maximum Matching (20 points)

Recall the maximum matching problem: Given an input graph G, what is the maximum k such that there are k edges which are all disjoint (i.e. no two of the are incident with the same vertex). Further recall the linear programming relaxation for this problem: We have a variable x_{ij} for each edge $(i, j) \in E(G)$ (where edges are always ordered so that $i < j$) and we maximize $\sum_{i,j:i<j,(i,j)\in E(G)} x_{ij}$ subject to

1. $\forall i < j : (i, j) \in E(G), 0 \leq x_{ij} \leq 1$
2. $\forall i, \sum_{j<i:(j,i)\in E(G)} x_{ji} + \sum_{j>i:(i,j)\in E(G)} x_{ij} \leq 1$

(a) 10 points: Prove that if the input graph G is bipartite, this linear programming relaxation gives the correct value of k (hint is available).

(b) 10 points: For general graphs, how much larger can the value for this linear programming relaxation be than the true size of the maximum matching? In other words, letting $LP(G)$ be the value of the linear progam on G and letting $opt(G)$ be the true size of the maximum matching in G, what is the smallest α such that for all input graphs G, $opt(G) \leq LP(G) \leq \alpha \cdot opt(G)$? Give an example where $LP(G) = \alpha \cdot opt(G)$.

Challenge question: Prove that for all input graphs G, $LP(G) \leq \alpha \cdot opt(G)$.

Solution

Part a: Since this linear program is a relaxation, any actual solution gives a solution to the linear program. Thus, $LP(G) \geq opt(G)$. For the other direction, let k be the maximal number of edges in a matching of G. By König's Theorem, G has a vertex cover V of size k. Now observe that since V is a vertex cover, for any solution to the linear program,

$$\sum_{i,j:i<j,(i,j)\in E(G)} x_{ij} \leq \sum_{i\in V} \left(\sum_{j<i:(j,i)\in E(G)} x_{ji} + \sum_{j>i:(i,j)\in E(G)} x_{ij} \right) \leq |V| = k$$

Thus, $LP(G) \leq opt(G)$, as needed.

Part b: $\alpha = \frac{3}{2}$. For a tight examle, take G to be a disjoint union of triangles. The largest matching in G has $\frac{n}{3}$ edges, one for each triangle. However, the linear program can set $x_{ij} = \frac{1}{2}$ for all edges $(i, j) \in E(G)$ which gives a value of $\frac{n}{2} = \frac{3}{2}opt(G)$ for $\sum_{i,j:i<j,(i,j)\in E(G)} x_{ij}$.

Problem 4: Conditions of Von Neumann's Minimax Theorem (15 points)

Recall Von Neumann's Minimax theorem

Theorem 0.1. *If X and Y are convex, compact subsets of $\mathbb{R}^n$ and $f : X \times Y \to \mathbb{R}$ is a continuous function which is convex in x and concave in y then*

$$\min_{x\in X}\max_{y\in Y} f(x,y) = \max_{y\in Y}\min_{x\in X} f(x,y)$$

(a) 5 points: Give an example showing that the theorem may fail if X and Y are closed and convex but not bounded.

(b) 5 points: Give an example showing that the theorem may fail if X and Y are compact but not convex.

(c) 5 points: Give an example showing that the theorem may fail if f is not convex in x.

(d) Challenge question: What happens if X and Y are bounded and convex but not closed?

Hints are available

Solution

(a) $X = Y = \mathbb{R}$, $f(x, y) = x + y$ is one such example. If x plays first then y can make $f(x, y)$ arbitrarily large. If y plays first then x can make $f(x, y)$ arbitrarily negative.

(b) $X = Y = \{-1, 1\}$, $f(x, y) = xy$ is one such example. If x plays first then y can match the sign of x to obtain $f(x, y) = 1$. If y plays first then x can take the opposite sign to obtain $f(x, y) = -1$

(c) $X = Y = [-1, 1]$, $f(x, y) = xy + 1 - x^2$ is one such example. If x plays first then y can take $y = \pm 1$ with the same sign as x to obtain $f(x, y) \geq 1$. If y plays first then x can take ± 1 with the opposite sign as y to obtain $f(x, y) = -|y| + 1 - 1 = -|y| \leq 0$

Choose one of the following two problems:

Problem 5: Duality for max flow and min cut (25 points)

Recall the linear program for the max flow problem. Given non-negative capacities $\{c_{ij}\}$ (where we take $c_{n1} = \infty$), maximize x_{n1} subject to

1. $\forall i, j, 0 \leq x_{ij} \leq c_{ij}$ (no capacity is exceeded, no negative flow)
2. $\forall i, \sum_{j \neq i} x_{ji} = \sum_{j \neq i} x_{ij}$ (flow in = flow out)

(a) 10 points: What is the dual linear program?

(b) 5 points: Show that any cut separating x_1 and x_n gives a feasible solution to the dual program.

(c) 10 points: Prove the max flow/min cut theorem (assuming strong duality) by showing that for any feasible solution of the dual program, there is a cut which has equal or smaller value (hint is available).

Solution

Part a: To find the dual, we can either put this linear program in canonical form or work directly using Von Neumann's minimax theorem as a guide. Here we work directly. Consider

$$f(\{x_{ij} : i \neq j\}, \{w_{ij} : i \neq j\}, \{y_i\}) = x_{n1} + \sum_{i,j:i \neq j} w_{ij}(c_{ij} - x_{ij}) + \sum_{i=1}^{n} y_i \left(\sum_{j \neq i} x_{ji} - \sum_{j \neq i} x_{ij} \right)$$

where we restrict $\{x_{ij} : i \neq j\}$ and $\{w_{ij} : i \neq j\}$ to all be nonnegative. One player controls the x variables and is trying to maximize f. The other player controls the y and w variables and is trying to minimize f.

If we have the x player play first, this gives us the primal. If we have the w, y player play first, this gives us the dual, which is as follows: Minimize $\sum_{i,j:i \neq j} w_{ij} c_{ij}$ subject to:

1. $\forall i \neq j, w_{ij} \geq 0$
2. $\forall i \neq j : y_j - y_i - w_{ij} \leq 0$. Rearranging, this is equivalent to $w_{ij} \geq y_j - y_i$
3. $y_1 - y_n - w_{n1} + 1 \leq 0$. Since we have $c_{n1} = \infty$, we always want to take $w_{n1} = 0$. Rearranging, this is equivalent to $y_n - y_1 \geq 1$

Part b: Given a cut $C = (S, \bar{S})$ where $x_1 \in S$ and $x_n \in \bar{S}$, take the following values for the dual

1. $y_i = 0$ whenever $x_i \in S$ and $y_i = 1$ whenever $x_i \in \bar{S}$
2. $w_{ij} = \max\{y_j - y_i, 0\}$. Equivalently, $w_{ij} = 1$ if $x_i \in S, x_j \in \bar{S}$ and $w_{ij} = 0$ otherwise.

This gives a dual value of $\sum_{i,j:x_i \in S, x_j \in \bar{S}} c_{ij}$, as needed.
Part c: For any solution to the dual, it is clearly optimal to take $w_{ij} = \max\{y_j - y_i, 0\}$, so we will assume this is the case and focus on the values for $\{y_i : i \in [1, n]\}$. Assume that we have a solution to the dual with minimal value with values $\{y_i\}$. Given this solution, take C to be the cut where $S = \{x_i : y_i \leq 0\}$ and $\bar{S} = \{x_i : y_i > 0\}$. Letting $c = \min\{\min\{y_i : y_i > 0\}, \frac{1}{2}\}$, we have the following two solutions to the dual:

1. $y'_i = 0$ if $x_i \in S$ and $y'_i = 1$ if $y' \in \bar{S}$
2. $y'_i = \frac{y_i}{1-c}$ if $x_i \in S$ and $y'_i = \frac{y_i - c}{1-c}$

Note that the original solution is c times this first solution plus $1 - c$ times the second solution. Since we assumed that our original solution had minimal value, both of these solutions must have this value. Thus, the first solution to the dual has minimal value and is a cut solution, as needed.

To be added: nice alternative solutions for this part.

Problem 6: Linear Programming Strong Duality (25 points)

This problem gives a proof that strong duality holds for linear programming. Recall the slack primal and dual forms for a linear program.
Primal: Maximize $c^T x$ subject to

1. $Ax = b$
2. $x \geq 0$

where these inequalities are taken coordinatewise.
Dual: Minimize $y^T b$ subject to

1. $A^T y \geq c$

(a) 10 points: Assuming that the primal is feasible (i.e. there is an $x \geq 0$ such that $Ax = b$), prove that strong duality holds using Farka's lemma, which says the following:

Lemma 0.2 (Farka's Lemma). *If $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$ then exactly one of the following holds:*

(a) $\exists x \in \mathbb{R}^n : Ax = b, x \geq 0$
(b) $\exists y \in \mathbb{R}^m : A^T y \geq 0, b^T y < 0$

(hint is available)

Challenge question: What happens if the primal is infeasible?

(b) 10 points: Prove Farka's lemma using the following version of the separating hyperplane theorem:

Theorem 0.3. *If X, Y are convex sets such that*

(a) $X \cap Y = \emptyset$

(b) X *is compact.*

(c) Y *is closed.*

then there is a hyperplane H such that X is entirely on one side of H, Y is entirely on the other side of H, and neither X nor Y intersect H.

(hint is available)

(c) 5 points: Show that if X and Y are convex sets and $(x, y) \in X \times Y$ has minimal distance (i.e. there is no $(x', y') \in X \times Y$ with $d(x', y') < d(x, y)$ then taking H to be hyperplane perpendicular to $y - x$ which passes through the midpoint between x and y, X is entirely on one side of H, Y is entirely on the other side of H, and neither X nor Y intersect H.

If you are familiar with compactness, complete the proof of the separating hyperplane theorem by proving that if X is compact and Y is closed then there must be a pair of points $(x, y) \in X \times Y$ of minimal distance.

Solution

Part a: First recall that any feasible value for the dual gives an upper bound on the value for the primal as $y^T b = y^T Ax \geq c^T x$

To show that we have equality, let z_0 be the value of $c^T x_0$ for the given feasible point x_0. For each $z > z_0$, we can do the following:

1. Let A' be the matrix formed by appending c^T to the matrix A as a final row.
2. Let b' be the vector b formed by appending z to b.
3. Apply Farka's lemma on A' and b'.

By Farka's lemma, either $A'x = b'$ in which case $Ax = b$ and $c^T x = z$ or $\exists y' : A'^T y' \geq 0$ and $b'^T y' < 0$. The first case corresponds precisely to a feasible solution of the primal with $c^T x = z$. In the second case, observe that the final entry of y' must be negative because otherwise we would still have $A'^T y' \geq 0$ and $b'^T y' < 0$ even if we replaced z by z_0, which is impossible by Farka's lemma as there is a solution of the primal with $c^T x = z_0$.

Scaling y' so that the final entry of y' is -1 and taking y to be the vector y' with the final entry removed, we have that $A^T y - c \geq 0$ and $y^T b - z < 0$. Rearranging, $A^T y \geq c$ and $y^T b < c$. Thus, there is a feasible solution of the dual with $y^T b < z$.

We have shown that for all $z \geq z_0$ either there is a feasible solution of the primal with value z or there is a feasible solution of the dual with value less than z. In fact, precisely one of these alternatives hold because any feasible value for the dual gives an upper bound on the value for the primal. Now take z_{opt} to be the infimum of the feasible values of the dual. If this infimum does not exist becuase the dual program is infeasible, then for all $z \geq z_0$ there is a feasible solution of the primal with value z so the value of the primal can be arbitrarily large. Otherwise, note that $z_{opt} \geq z_0$ as no solution of the dual can have value less than z_0. By definition, there is no feasible value of the dual with value less than z_{opt} so there must be a solution to the primal with value z_{opt} and the duality gap is 0.

Part b: Both statements in Farka's lemma cannot hold at the same time as otherwise we would have that $0 > y^T b = y^T Ax \geq 0$, which is a contradiction. Thus, we just need to show that if the first statement in Farka's lemma does not hold, then the second one does.

Let $X = \{0\}$ (the origin) and let $Y = \{Ax - b : x \geq 0\}$. X is compact and Y is closed. If the first statement of Farka's lemma does not hold then $X \cap Y = \emptyset$. By the separating hyperplane theorem, there is a hyperplane $y^T z = c > 0$ such that for all $z \in \{Ax - b : x \geq 0\}$, $y^T z > c > 0$. Thus, for all $x \geq 0$, $y^T Ax - y^T b > 0$, which impliles that $y^T A \geq 0$ and $y^T b < 0$, as needed.

Part c: Let $x \in X$ and $y \in Y$ be a pair of points of minimal distance from each other. Take H to be the hyperplane perpendicular to the line segment between x and y which passes through its midpoints. If there was an $x' \in X$ on the opposite side of H from x, then for a sufficiently small $\lambda > 0$, $d(\lambda x' + (1-\lambda)x, y) < d(x, y)$, contradicting the choice of x, y.

To show that such a pair (x, y) exists, we can use the following argument. There must be a sequence of points (x_i, y_i) such that $d(x_i, y_i)$ monotonically decreases to the infimum d_{min} of the distance between pairs of points $x \in X$ and $y \in Y$. Since X is compact, we can take a subsequence $\{x_j, y_j\}$ where the $\{x_j\}$ converge to some $x \in X$. Now $d(x, y_j)$ must converge to d_{min}. Now if we consider the ball B of radius $2d_{min} + 1$ around x, $y_j \in B \cap Y$ for sufficiently large j. $B \cap Y$ is compact so there is a subsequence of the y_j which converges to some $y \in Y$ and we must have that $d(x, y) = d_{min}$. Thus, there must be some pair of points of minimal distance, as needed.

Problem 7: Finite two person one round simultaneous zero-sum games (30 points)

Any finite two person zero-sum game which is played in one simultaneous round (i.e. both players choose a strategy and then simultaneously reveal them) can be represented in the following way. There is a payoff matrix M whose rows are labeled by strategies for player 1 and whose columns are labeled by strategies for player 2. M_{ij} is the amount that player 2 takes from player 1 if player 1 playes strategy i and player 2 plays strategy j. Thus, the row player is trying to minimize the resulting entry of M and the column player is trying to maximize the resulting entry of M.

For example, consider the game where both players simulataneously call out 1 or 2, player 1 wins if the sum of the numbers is odd, and player 2 wins if the sum is even. This game can be represented with the following payoff matrix: $M = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

(a) 5 points: Give a payoff matrix for rock-paper-scissors.

(b) 10 points: Now consider the general game with an arbitrary payoff matrix and assume that the rules are changed so that the row player must go first but can choose a probability distribution over his/her strategies. Give a linear program which finds an optimal strategy for the row player in this setting. What is the dual linear program? How does this relate to strong duality and Nash equilibria?

(c) 15 points: Implement these linear programs to find a Nash equilibrium for the following payoff matrix:

$$M = \begin{bmatrix} 2 & 5 & 11 & 9 & 12 & 7 \\ 20 & 3 & 9 & 5 & 7 & 8 \\ 8 & 10 & 6 & 13 & 7 & 12 \\ 18 & 12 & 5 & 8 & 6 & 3 \\ 16 & 17 & 1 & 13 & 4 & 6 \\ 4 & 8 & 15 & 4 & 14 & 6 \end{bmatrix}$$

Solution

Part a: $\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ where row and column one are rock, row and column two are paper, and row and column three are scissors.

Part b: Let $M \in \mathbb{R}^{m \times n}$ be the payoff matrix. To guarantee that the column player cannot win more than z, the row player must find wieghts $x_1, \cdots, x_m$ such that $\sum_{i=1}^{m} x_i = 1$ and each coordinate of $M^T x$ (which gives the expected payoff for a column) is at most z. Thus, the row player is trying to minimize z subject to

1. $\forall i, x_i \geq 0$
2. $\sum_{i=1}^{m} x_i = 1$
3. $M^T x \leq z$ (where z is z times the all ones vector here)

There are several ways to find the dual. One way is to express the primal as follows. Consider function $f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} x_i y_j$ with the constraints that

1. $\forall i, x_i \geq 0$
2. $\sum_{i=1}^{m} x_i = 1$
3. $\forall j, y_j \geq 0$
4. $\sum_{j=1}^{n} y_j = 1$

The x player wants to minimize $f(x,y)$ and the y player wants to maximize $f(x,y)$. For the primal, the x player plays first. For the dual, the y player plays first, giving us the following linear program: Maximize z subject to

1. $\forall j, y_j \geq 0$
2. $\sum_{j=1}^{n} y_j = 1$
3. $My \geq z$ (where z is z times the all ones vector here)

A second way to find the dual is to express the primal program in canonical form and take the canonical dual. This must be done carefully (which is partly why we are presenting this method as well).

Primal program: Take $c^T = (0, \cdots, 0, -1, 1)$ and $x^T = (x_1, \cdots, x_m, z_+, z_-)$. Maximize $c^T x$ subject to:

1. $\begin{bmatrix} M^T & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} x \leq \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ where the $-1, 1$ in the first row and column are vectors rather than single entries
2. $x \geq 0$

Dual program: Minimize $(0, \cdots, 1, -1) \cdot y'$ subject to:

1. $\begin{bmatrix} M & 1 & -1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} y' \geq \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ where the $-1, 1$ in the first row and column are vectors rather than single entries
2. $y' \geq 0$

Writing $y' = (y, z_-, z_+)$, we can see that this gives an equivalent dual. The y player is trying to minimize $-z = z_- - z_+$ subject to $My \geq z$, $y \geq 0$, and $\sum_{j=1}^{n} y_j = 1$.

Strong duality tells us that the optimal valules for the primal and dual are the same. This implies that together, an optimal solution for the primal and an optimal solution for the dual form a Nash equilibrium. To see this, note that with these strategies, we have both a guarantee that the expected value for each column is at most z and the expected value for each row is at least z, so neither player can benefit by deviating from these strategies.

Part c: The optimal strategy for the row player is to take the second row with probability .1, the third row with probability .5, the fourth for with probability .1, and the last row with probability .3. This gives expected payoffs of $9, 8.9, 8.9, 9, 9, 8.9$ for the columns. The optimal strategy for the column player is to take the first column with probability .2, the fourth column with probability .3, and the fifth column with probability .5. This gives expected payoffs $-9.1, -9, -9, -9, -9.1, -9$ for the row player. Thus, in the Nash equilibrium, the column player expects to take 9 from the row player.

Hints

2b. For the examples, note that the linear programming relaxation always has a value of at most $\frac{n}{2}$ as it can always assign value $\frac{1}{2}$ to every vertex. For the proof, what happens if we take every vertex which has value at least $\frac{1}{2}$ in the output of the linear program?

3a. Use König's Theorem: For a bipartite graph G, the size of the maximum matching is equal to the size of the minimal vertex cover.

4a. There is an example where $X = Y = \mathbb{R}$ and f is linear in both x and y.

4b. Rock-paper-scissors can give such an example.

4c. Consider the function $f(x, y) = xy$. If the y player plays first then $f(x) \leq 0$ so long x or y is 0 or x is chosen to have the opposite sign as y. On the other hand, if the x player plays first, he/she had better play $x = 0$. However, if the x player is penalized for playing $x = 0$...

5c. Show that any feasible solution for the dual can be written as a positive linear combination of a cut and another solution. For the cut, your dual linear program should involve weight variables w_{ij} for the edges of the graph where the sum of the weight variables along any path from $s = x_1$ to $t = x_n$ is 1. Take the cut $(S, \bar{S})$ where S is the set of vertices reachable from $s = x_1$ using edges of 0 or negative weight.

6a. Let z be a potential value for $c^T x$ which is larger than value for the given feasible point. Add the constraint $c^T x = z$ to the program by adding c as an extra row of A and z as an extra entry in b. Now use Farka's lemma to show that if the primal is not feasible with value z then the dual is feasible with some value less than z.

6b: Apply the separating hyperplane theorem to $X = \{0\}$ (the origin) and $Y = \{(b - Ax) : x \geq 0\}$