CS 491I Approximation algorithms
Lecture notes
Feb. 6th & 8th ,2001
Dejan Desovski

1. Polyhedral Cones

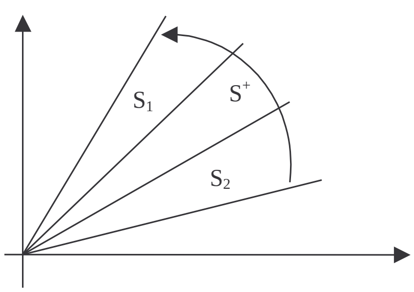
Definition 1: A **cone** is a set $C \subseteq \Re^n$ for which $\vec{x}, \vec{y} \in C$ and $\lambda, \mu \geq 0$ implies $\lambda\vec{x} + \mu\vec{y} \in C$. A cone C is said to be **generated** by $\vec{x_1}, \ldots, \vec{x_k}$ if $\vec{x_1}, \ldots, \vec{x_k} \in C$ and for any $\vec{x} \in C$ there are numbers $\lambda_1, \ldots, \lambda_k \geq 0$ with $\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{x_i}$. A cone is called **finitely generated** if some finite set of vectors generates it. A polyhedral cone is a polyhedron of type $\{\vec{x} : A\vec{x} \leq 0\}$.

Theorem 1: A polyhedral cone is indeed a cone.
Proof: Assume that C is a polyhedral cone $C = \{\vec{x} : A\vec{x} \leq 0\}$. Let $\vec{x}, \vec{y} \in C$ and λ, $\mu \geq 0$. We construct the vector $\vec{z} = \lambda\vec{x} + \mu\vec{y}$ and we should prove that $\vec{z} \in C$.

$$A\vec{z} = A(\lambda\vec{x} + \mu\vec{y}) = \lambda A\vec{x} + \mu A\vec{y} \leq 0 \quad \Rightarrow \ A\vec{z} \leq 0 \Rightarrow \vec{z} \in C.$$

Theorem 2: Polyhedral cones are convex.
Proof: Assume that C is a polyhedral cone $C = \{\vec{x} : A\vec{x} \leq 0\}$. Let $\vec{x}, \vec{y} \in C$ and $\lambda \in [0,1]$. We construct a vector that represents a convex combination $\vec{z} = \lambda\vec{x} + (1-\lambda)\vec{y}$ and we should prove that $\vec{z} \in C$.

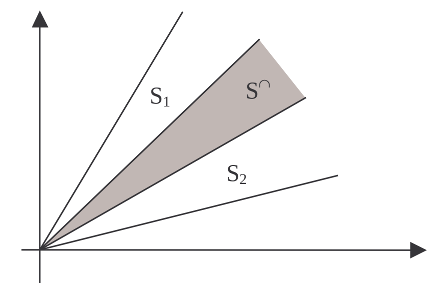
$$A\vec{z} = A(\lambda\vec{x} + (1-\lambda)\vec{y}) = \lambda A\vec{x} + (1-\lambda)A\vec{y} \leq 0 \quad \Rightarrow \ A\vec{z} \leq 0 \Rightarrow \vec{z} \in C.$$



Definition 2: A **sum cone** C^+ of two cones C_1 and C_2 is defined as $C^+ = \{\vec{x} : \vec{x} = \vec{x_1} + \vec{x_2}, \vec{x_1} \in C_1, \vec{x_2} \in C_2\}$.

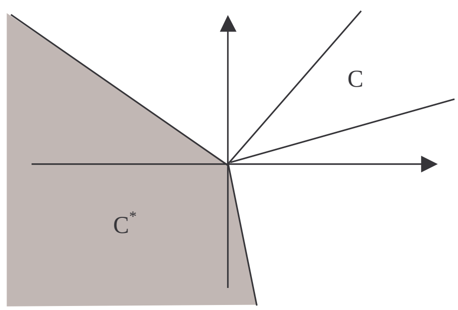
Theorem 3: A sum cone is a cone.

Proof: Let $\vec{x}, \vec{y} \in C^+$, $\lambda, \mu \geq 0$. $\vec{x} = \vec{x_1} + \vec{x_2}, \vec{y} = \vec{y_1} + \vec{y_2}$, where $\vec{x_1}, \vec{y_1} \in C_1, \vec{x_2}, \vec{y_2} \in C_2$. Then $\vec{z} = \lambda\vec{x} + \mu\vec{y} = (\lambda\vec{x_1} + \mu\vec{y_1}) + (\lambda\vec{x_2} + \mu\vec{y_2}) = \vec{z_1} + \vec{z_2}$, where $\vec{z_1} \in C_1$ and $\vec{z_2} \in C_2$. This implies that $\vec{z} \in C^+$ which proves the theorem.



Definition 3: An **intersection cone** $C^\cap$ of two cones C_1 and C_2 is defined as $C^\cap = \{\vec{x} : \vec{x} \in C_1 \wedge \vec{x} \in C_2\}$.

Exercise: Prove that an intersection cone is a cone.



Definition 4: A **polar cone** (dual cone) of a given cone C is $C^* = \{\vec{x} : \vec{x} \cdot \vec{y} \le 0, \forall \vec{y} \in C\}$.

Exercise: Prove that a polar cone is a cone.
The geometric interpretation of a polar cone: the set of all vectors that occupy obtuse angle with the vectors of a given cone.

The polar cones have the following property: $(C^*)^* = C$

2. Duality

Some motivation and introduction to the duality concept.

In order to describe an arbitrary line $ax_1 + bx_2 + c = 0$ in the $x_1 x_2$ plane we need two parameters (m, c) because each line can be represented in the form $x_2 = mx_1 + c$. On the other hand given two numbers (p, q) can represent a point in the $x_1 x_2$ plane. This is the concept of duality – we can view the same information in different way depending on the context.

Give some upper bound of the objective function for the following linear program?

$$\max z = 2x_1 + x_2$$
$$3x_1 - 4x_2 \le 7$$
$$-x_1 + 5x_2 \le -4$$
$$x_1, x_2 \ge 0$$

We notice that if we add the two constraints we will get the objective function and one upper bound is 7 + (-3) = 3. If we solve the system we will see that this is the optimum.
In this case we were fortunate and just by adding the constraints we have obtained the objective function. In general we would need to multiply each of the constraints by some appropriate number so that after adding them we would obtain the objective function. So we are interested to find this vector $\vec{y} = [y_1, \ldots, y_m]$ comprised of these numbers. This vector must be ≥ 0, if some of the components it less than zero then it will change the inequality sign of the corresponding constraint and change the original linear program.

$$\begin{array}{l} \max \vec{c}\vec{x} \\ s.t.\ A\vec{x} \le \vec{b} \\ \vec{x} \ge 0 \end{array} \quad \Rightarrow \quad \begin{array}{l} \max z = c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ y_1 * (a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n \le b_1) \\ y_2 * (a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n \le b_2) \\ \vdots \\ y_m * (a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n \le b_m) \\ \vec{x} \ge 0 \end{array}$$

If we add the constraints we will get the following.

$$\begin{array}{l} (a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_m)x_1 + \\ (a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m)x_2 + \ldots + \\ (a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m)x_n \le b_1y_1 + b_2y_2 + \ldots + b_my_m \end{array}$$

In order to get better upper bound on the original linear program, the coefficients before each of the x_i variables should be ≥ than the corresponding c_i constant, and we should find the minimum of the right hand side of the inequality. In this way we get another linear program that is called the **dual** of the original (**primal**) linear program.

$$(P)\quad \begin{array}{l} \max \vec{c}\vec{x} \\ s.t.\ A\vec{x} \le \vec{b} \\ \vec{x} \ge 0 \end{array} \qquad (D)\quad \begin{array}{l} \min \vec{y}\vec{b} \\ s.t.\ \vec{y}A \ge \vec{c} \\ \vec{y} \ge 0 \end{array}$$

Theorem 1: The dual of the dual is the primal.

Proof: We start from the dual.

$$\begin{array}{l} \min \vec{y}\vec{b} \\ s.t.\ \vec{y}A \ge \vec{c} \\ \vec{y} \ge 0 \end{array} \text{, writing in the right form } \Rightarrow \begin{array}{l} \max - \vec{y}\vec{b} \\ s.t. - A^T\vec{y} \le -\overrightarrow{c^T} \\ \vec{y} \ge 0 \end{array},$$

$$\text{writing the dual} \Rightarrow \begin{array}{l} \min - \vec{w}\overrightarrow{c^T} \\ s.t. - \vec{w}A^T \ge -\vec{b} \\ \vec{w} \ge 0 \end{array} \text{, changing the variable } w \Rightarrow \begin{array}{l} \max \vec{c}\vec{x} \\ s.t.\ A\vec{x} \le \vec{b} \\ \vec{x} \ge 0 \end{array}.$$

Rules for writing the dual.

Max Problem	**Min Problem**	**Max Problem**	**Min Problem**
Constraints	Variables	Variables	Constraints
≤	≥0	≥0	≥
≥	≤0	≤0	≤
=	unrestricted	unrestricted	=

Theorem 2 (Week Duality): Let $\vec{x}$ and $\vec{y}$ be a feasible solutions of the primal (P) linear program and its dual (D) respectively.

$$\text{(P)}\quad \begin{array}{l} \max \vec{c}\vec{x} \\ s.t.\ A\vec{x} \le \vec{b} \\ \vec{x} \ge 0 \end{array} \qquad\qquad \text{(D)}\quad \begin{array}{l} \min \vec{y}\vec{b} \\ s.t.\ \vec{y}A \ge \vec{c} \\ \vec{y} \ge 0 \end{array}$$

Then $\vec{c}\vec{x} \le \vec{y}\vec{b}$.

Proof: $\vec{c}\vec{x} \le \left(\vec{y}A\right)\vec{x} = \vec{y}\left(A\vec{x}\right) \le \vec{y}\vec{b}$.

Corollary 1: If (P) is unbounded then (D) is infeasible.

Proof: Given any $\lambda \in \Re^+$ which the primal can take because of his unboundednes, the value of the dual should be greater than λ because of the weak duality theorem. This implies that the dual is infeasible.

Corollary 2: If (D) is unbounded then (P) is infeasible.

Proof: Given any $\lambda \in \Re^-$ which the primal can take because of his unboundednes, the value of the dual should be less than λ because of the weak duality theorem. This implies that the dual is infeasible.

Observe the following linear program and its dual.

$$(1)\quad \begin{array}{l} \max x_1 + 2x_2 \\ x_1 - 2x_2 \le -2 \\ x_1 - 2x_2 \ge 2 \\ x_1, x_2 \ge 0 \end{array} \qquad\qquad (2)\quad \begin{array}{l} \min -2y_1 - 2y_2 \\ y_1 - y_2 \ge 1 \\ -2y_1 + 2y_2 \ge 2 \\ y_1, y_2 \ge 0 \end{array}$$

We note that both are infeasible. So we have the following corollary.

Corollary 3: If (P) is infeasible then (D) is infeasible or unbounded.

Corollary 4: If $\vec{x}$ is feasible to (P) and $\vec{y}$ is feasible to (D) and $\vec{c}\vec{x} = \vec{y}\vec{b}$ then $\vec{x}$ and $\vec{y}$ are optimal.

Proof: From the weak duality theorem we have that $\vec{y}\vec{b}$ is a upper bound for the (P)'s maximum and $\vec{c}\vec{x}$ is the lower bound for the (D)'s minimum. In the case when they are equal they must be optimal.

Theorem 3 (Strong Duality): Given (P) and (D), if the respective optimal points are finite, then $\vec{c}\vec{x} = \vec{y}\vec{b}$.

Proof: We apply the simplex algorithm.

$$\begin{array}{l} \max \vec{c}\vec{x} \\ A\vec{x} \le \vec{b} \\ \vec{x} \ge 0 \end{array} \qquad \Rightarrow \qquad \begin{array}{l} \max \vec{c}\vec{x} + \vec{0}\vec{x_s} \\ \begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} \vec{x} \\ \vec{x_s} \end{bmatrix} = \vec{b} \\ \vec{x}, \vec{x_s} \ge 0 \end{array}$$

Let B be an optimal basis. Then the optimal solution is given by the following.

$z^* = \overrightarrow{c_B} B^{-1} \vec{b}, \overrightarrow{x^*} = \begin{pmatrix} B^{-1}\vec{b} \\ 0 \end{pmatrix}$. Suppose we construct $\vec{y}$ s.t. $\vec{y}\vec{b} = \overrightarrow{c_B} B^{-1} \vec{b}$. That is, if we set $\vec{y} = \overrightarrow{c_B} B^{-1}$ we would have $\vec{y}\vec{b} = \overrightarrow{c_B} B^{-1} \vec{b} = \vec{c}\vec{x}$. We just need to prove that this $\vec{y}$ is feasible for (D). We know that at optimality $\frac{\partial z}{\partial x_j} = -\left(\overrightarrow{c_B} B^{-1} a_j - c_j\right) \leq 0$ for all variables. We can write all these relations together:

$$\begin{aligned}
&\overrightarrow{c_B} B^{-1} [A \quad I] - \vec{c}\,' \geq 0 \\
&\overrightarrow{c_B} B^{-1} [A \quad I] \geq [\vec{c} \quad 0] \\
&\overrightarrow{c_B} B^{-1} A \geq \vec{c}, \overrightarrow{c_B} B^{-1} \geq 0 \\
&\vec{y} A \geq \vec{c}, \vec{y} \geq 0
\end{aligned}$$

We see that the constructed $\vec{y}$ is feasible for (D).

Complementary Slackness.

Let's turn our attention to the slack variables in the primal (P) and the dual (D). By adding the slack variables in order to turn the inequalities to equalities we get:

$$A\vec{x} + I\vec{s} = \vec{b}$$

$$\Rightarrow \vec{s} = \vec{b} - A\vec{x}$$

$\vec{s}$ is the vector of slack variables and represents the degree of slackness. It is always $\vec{s} \geq 0$.

$$\vec{y}A + \vec{u}I = \vec{c}$$

$$\Rightarrow \vec{u} = \vec{y}A - \vec{c}$$

$\vec{u}$ is the vector of surplus variables.. It is always $\vec{u} \geq 0$.

Theorem 4 (Complementary Slackness): Given (P) and (D), if the respective optimal points $\left(\overrightarrow{x^*}, \overrightarrow{y^*}\right)$ are finite, then $y_i^* s_i^* = 0, i = 1, \ldots, m$ and $x_i^* u_i^* = 0, i = 1, \ldots, n$.

Proof: At optimality:

$$\vec{c}\overrightarrow{x^*} = \left(\overrightarrow{y^*} A - \overrightarrow{u^*}\right)\overrightarrow{x^*} = \overrightarrow{y^*} A \overrightarrow{x^*} - \overrightarrow{u^*}\,\overrightarrow{x^*} = \overrightarrow{y^*}\left(\vec{b} - \overrightarrow{s^*}\right) - \overrightarrow{u^*}\,\overrightarrow{x^*} = \overrightarrow{y^*}\vec{b} - \overrightarrow{y^*}\,\overrightarrow{s^*} - \overrightarrow{u^*}\,\overrightarrow{x^*}$$

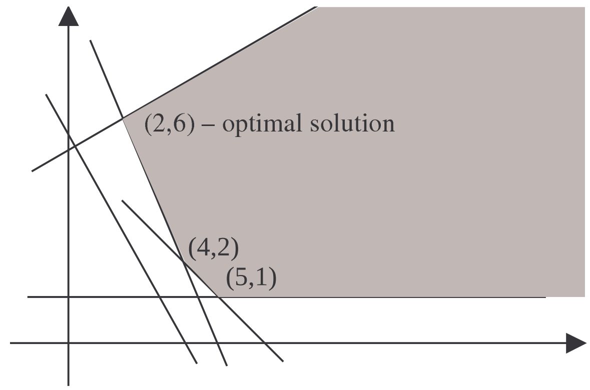
Because of strong duality theorem we have $\vec{c}\overrightarrow{x^*} = \overrightarrow{y^*}\vec{b}$. This implies $\overrightarrow{y^*}\,\overrightarrow{s^*} + \overrightarrow{u^*}\,\overrightarrow{x^*} = 0$, which is only possible if $\overrightarrow{y^*}\,\overrightarrow{s^*} = 0 \wedge \overrightarrow{u^*}\,\overrightarrow{x^*} = 0$.

Example: Solve the following linear program.

$$\begin{aligned}
&\max z = 10x_1 + 6x_2 - 4x_3 + x_4 + 12x_5 \\
&2x_1 + x_2 + x_3 + 3x_5 \leq 18 \\
&x_1 + x_2 - x_3 + x_4 + 2x_5 \leq 6 \\
&x_1, \ldots, x_5 \geq 0
\end{aligned}$$

We will try to solve the dual because it is going to have only two variables and can be solved graphically. The dual is going to be:

$$\begin{aligned} \min w &= 18y_1 + 6y_2 \\ 2y_1 + y_2 &\geq 10 \quad (1) \\ y_1 + y_2 &\geq 6 \quad (2) \\ y_1 - y_2 &\geq -4 \quad (3) \\ y_2 &\geq 1 \quad (4) \\ 3y_1 + 2y_2 &\geq 12 \quad (5) \\ y_1, y_2 &\geq 0 \end{aligned}$$



We find that (2, 6) is the optimal solution and w = 72. Duality theorem implies that the optimal value for z is going to be the same z = 72. In order to find the corresponding vector x we can use the complementary slackness theorem. The optimal solution of the dual is obtained in the intersection of the constraints (1) and (3). This implies that $u_1 = 0, u_3 = 0; u_2, u_4, u_5 > 0$. Because of the complementary slackness theorem we know that $x_2 = x_4 = x_5 = 0$. If we apply these values to the original linear program we get the following system that can be solved by Gausian elimination.

$$\begin{aligned} 2x_1 + x_3 &= 18 \\ x_1 - x_3 &= 6 \end{aligned} \Rightarrow (x_1, x_3) = (8,2)$$

And the maximum value is z = 10(8)+6(0)-4(2)+1(0)+12(0) = 72.

Corollary 5: Given (P) and (D) only four conditions are possible:

a) (P) and (D) have finite equal optimum
b) (P) is unbounded and (D) is infeasible
c) (D) is unbounded and (P) is infeasible
d) (P) and (D) are infeasible

Farkas Lemma
Given:

1) $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$
2) $\vec{b}\vec{y} < 0, \vec{y}A \geq 0, \vec{y} \geq 0$

Either 1) is feasible, or 2) is feasible, they can not be feasible at the same time.

Proof: We apply the corollary 5 to the following linear program and its dual.

(P) $\begin{aligned} &\max 0\vec{x} \\ &A\vec{x} \leq \vec{b}, \vec{x} \geq 0 \end{aligned}$ (D) $\begin{aligned} &\min \vec{b}\vec{y} \\ &\vec{y}A \geq 0, \vec{y} \geq 0 \end{aligned}$

The conditions b) and d) are not possible because (D) is feasible, $\vec{y} = 0$ is one feasible point.

If we are in case a) then (P) and (D) must have equal optimum = 0 because of (P)'s objective function. This implies that 1) is feasible and $\vec{b}\vec{y} = 0$ which makes 2) infeasible.
If we are in case c) then (D) is unbounded and (P) is infeasible which makes 1) infeasible and 2) feasible.

Geometrical interpretation.

Either $\vec{x} \geq 0, A\vec{x} = \vec{b}$ **or** $\exists y : A^T \vec{y} \geq 0, \vec{y}\vec{b} < 0$. We can prove this version of the Farkas Lemma by applying the same arguments as the previous proof to the following primal and dual.

(P) $\begin{array}{l} \max 0\vec{x} \\ A\vec{x} = \vec{b}, \vec{x} \geq 0 \end{array}$ (D) $\begin{array}{l} \min \vec{b}\vec{y} \\ \vec{y}A \geq 0 \end{array}$

The geometrical interpretation of this version of the Farkas Lemma is the following: Either $\vec{b}$ lies in the cone formed by the columns of *A*. Or there exists a witness vector $\vec{y}$ that makes acute angle with the columns of *A* and obtuse angle with $\vec{b}$ - which means that $\vec{b}$ is outside of the cone formed by the columns of *A*.