AMS 553.766: Combinatorial Optimization
Homework Problems - Week V

For the following problems, $A \in \mathbb{R}^{m \times n}$ will be $m \times n$ matrices, and $b \in \mathbb{R}^m$. An affine subspace is the set of solutions to a system of linear equations, i.e., $\{x \in \mathbb{R}^n : Ax = b\}$ is an affine subspace of $\mathbb{R}^n$. A polytope is a bounded polyhedron.

1. Show the following:

(i) The intersection of an affine subspace with a polyhedron is a polyhedron.

Solution: Express the original polyhedron as $P = \{x : Ax \leq b\}$ and the affine subspace as $L = \{x : A'x = b'\}$. The intersection $P \cap L = \{x : A'x = b', Ax \leq b\} = \{x : A'x \leq b', -A'x \leq -b', Ax \leq b\}$ which is by definition a polyhedron.

(ii) Let $P_1, P_2$ be two polytopes. Define $C = \{x + y : x \in P_1, y \in P_2\}$. Show that $C$ is a polytope.

Solution: Express $P_1 = \text{conv}\{x_1, \ldots, x_p\}$ and $P_2 = \text{conv}\{y_1, \ldots, y_q\}$. We claim that $C = \text{conv}\{x_i + y_j : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}$. Consider $z \in C$, then $z = x + y$ for some $x \in P_1$ and $y \in P_2$. We can write $x = \sum_{i=1}^{p} \lambda_i x_i$ and $y = \sum_{j=1}^{q} \mu_j y_j$, where $\lambda_i \geq 0$, $\sum_{i=1}^{p} \lambda_i = 1$, $\mu_j \geq 0$ and $\sum_{j=1}^{q} \mu_j = 1$. Observe that $\sum_{i,j} \lambda_i \mu_j = 1$. Also,

$$\sum_{i,j} \lambda_i \mu_j (x_i + y_j) = \sum_{i=1}^{p} \lambda_i x_i + \sum_{j=1}^{q} \mu_j y_j = x + y = z.$$ 

Since $\sum_{i,j} \lambda_i \mu_j (x_i + y_j) \in \text{conv}\{x_i + y_j : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}$, we have $z \in \text{conv}\{x_i + y_j : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}$. Therefore, $C \subseteq \text{conv}\{x_i + y_j : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}$. To show the reverse inclusion, consider $z \in \text{conv}\{x_i + y_j : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}$; so we can write $z = \sum_{i,j} \sigma_{ij} (x_i + y_j)$ where $\sigma_{ij} \geq 0$ and $\sum_{i,j} \sigma_{ij} = 1$. We now simply this expression:

$$\sum_{i,j} \sigma_{ij} (x_i + y_j) = \sum_{i,j} \sigma_{ij} x_i + \sum_{i,j} \sigma_{ij} y_j = \left(\sum_{i=1}^{p} \left(\sum_{j=1}^{q} \sigma_{ij}\right)x_i\right) + \left(\sum_{j=1}^{q} \left(\sum_{i=1}^{p} \sigma_{ij}\right)y_j\right) = \sum_{i=1}^{p} \lambda_i x_i + \sum_{j=1}^{q} \mu_j y_j,$$

where $\lambda_i = \sum_{j=1}^{q} \sigma_{ij}$ and $\mu_j = \sum_{i=1}^{p} \sigma_{ij}$. Observe that

$$\sum_{i=1}^{p} \lambda_i = \sum_{j=1}^{q} \sigma_{ij} = 1.$$ 

Similarly, $\sum_{j=1}^{q} \mu_j = 1$. Thus, $\sum_{i=1}^{p} \lambda_i x_i \in P_1$ and $\sum_{j=1}^{q} \mu_j y_j \in P_2$. $z = \sum_{i=1}^{p} \lambda_i x_i + \sum_{j=1}^{q} \mu_j y_j \in P_1 + P_2$.

Thus, we express $C$ as the convex hull of a finite set of points; by the Minkowski-Weyl theorem it is a polytope.

(iii) The intersection of two polyhedra is a polyhedron. Thus, show that the intersection of a polyhedron with a polytope is a polytope.

Solution: Let one polyhedron be represented by $A_1 x \leq b_1$ and the second polyhedron be $A_2 x \leq b_2$. The intersection is given by $A_1 x \leq b_1, A_2 \leq b_2$ which is another polyhedron. Since the intersection of a bounded set with another set is also bounded, the intersection of a polytope with another polyhedron is a polytope.
2. (Complementary slackness) Let \( x^* \in \mathbb{R}^n \) be an optimal solution to the problem \( \max \{c^T x : Ax \leq b\} \) and \( y^* \in \mathbb{R}^m \) be an optimal solution to the problem \( \max \{y^T b : y^T A = c^T, y \geq 0\} \). Show that for every \( i = 1 \ldots, m \) either \( a_i \cdot x^* = b_i \) or \( y_i^* = 0 \) (or both). (Here \( a_i \) denotes the \( i \)-th row of \( A \) and \( b_i \) is the \( i \)-th component of \( b \).) [Hint: consider the vector \((y^*)^T(Ax^* - b)\)]

The theorem is saying that in an optimal primal-dual pair of solutions for an LP, either a constraint is tight at the optimal primal solution or the corresponding dual multiplier is 0.

Solution: We will in fact show that two feasible solutions \( x^* \) (for the primal) and \( y^* \) (for the dual) are optimal solutions if and only if for every \( i = 1 \ldots, m \) either \( a_i \cdot x^* = b_i \) or \( y_i^* = 0 \) (or both). Observe that since \( x^* \) and \( y^* \) are feasible for the primal and dual problems respectively,

\[
c^T x^* - (y^*)^T b = ((y^*)^T A)x^* - (y^*)^T b = (y^*)^T (Ax^* - b) \leq 0
\]

Moreover, since each component of \( y^* \) is nonnegative and each component of \( Ax^* - b \) is also nonnegative, \( c^T x^* - (y^*)^T b = 0 \) if and only if for every \( i = 1 \ldots, m \) either \( a_i \cdot x^* = b_i \) or \( y_i^* = 0 \) (or both). Finally, \( c^T x^* - (y^*)^T b = 0 \) is equivalent to \( x^* \) and \( y^* \) being optimal solutions to the primal and dual problems respectively.

3. Find an example of a pair of primal and dual linear programs such that both problems are infeasible. [Hint: There is an example with 2 variables in each problem]

Solution:

Primal: \[
\begin{align*}
\text{max} & \quad x_1 \\
\text{subject to} & \quad -x_1 + x_2 \leq 1 \\
& \quad x_2 \leq 0 \\
& \quad -x_2 \leq -1
\end{align*}
\]

Dual: \[
\begin{align*}
\text{max} & \quad y_1 - y_3 \\
\text{subject to} & \quad -y_1 = 1 \\
& \quad y_1 + y_2 - y_3 = 0 \\
& \quad y_1, y_2, y_3 \geq 0
\end{align*}
\]

4. Let \( P = \{x \in \mathbb{R}^n : Ax \leq b\} \) be a nonempty polyhedron. Let \( C = \{r \in \mathbb{R}^n : x + \lambda r \in P \text{ for all } x \in P, \lambda \in \mathbb{R}_+ \} \) where \( \mathbb{R}_+ \) is the set of nonnegative real numbers. Show that

(i) \( C \) is a convex cone. [This cone is called the recession cone of the polyhedron \( P \)]

(ii) \( C = \{r \in \mathbb{R}^n : Ar \leq 0\} \).

Solution:

(i) We need to first verify that \( C \) is convex. Consider any \( r^1, r^2 \in C \) and any \( \mu \in [0, 1] \) and let \( \bar{r} = \mu r^1 + (1 - \mu) r^2 \). For any \( x \in P, \lambda \geq 0 \) we have \( x + \lambda \bar{r} = x + \lambda \mu r^1 + \lambda (1 - \mu) r^2 \). Since \( r_1 \in C \),
Let \( P \) be a polyhedron. Assume that \( r \neq 0 \) and \( r = b \).

(i) Show that \( \ker(r) \) is a linear subspace of \( \mathbb{R}^n \). (ii) Show that \( A \) is a linear subspace of \( \mathbb{R}^n \).

Solution:
(i) For any \( r_1, r_2 \in L \), \( x + \lambda r_1 + \lambda r_2 \in L \). Since \( r \neq 0 \), \( r = b \).

(ii) For any \( r \in L \), \( x + \lambda r \in L \).

If \( P \) has a vertex \( z \), then \( A_z \) has rank \( n \), thus \( A \) has rank \( n \). Therefore \( Ar = 0 \) only has the 0 solution, and so \( L = \{0\} \). So if \( P \) has a vertex, it is pointed.

6. Using the notation from Problems 5 and 6, show that \( L = C \cap (-C) \). We say that \( P \) is pointed if \( L = \{0\} \). Suppose \( P \neq \emptyset \); show that \( P \) has at least one vertex if and only if \( P \) is pointed.

Solution:
\[
C \cap (-C) = \{ r \in \mathbb{R}^n : Ar \leq 0 \} \cap \{ r \in \mathbb{R}^n : A(-r) \leq 0 \} \\
= \{ r \in \mathbb{R}^n : Ar \leq 0, Ar \geq 0 \} \\
= \{ r \in \mathbb{R}^n : Ar = 0 \} \\
= L.
\]

If \( P \) has a vertex \( z \), then \( A_z \) has rank \( n \), thus \( A \) has rank \( n \). Therefore \( Ar = 0 \) only has the 0 solution, and so \( L = \{0\} \). So if \( P \) has a vertex, it is pointed.

Assume \( P \) is pointed. So \( Ar = 0 \) has only the 0 solution and so \( A \) has rank \( n \). Let \( z \in P \) since \( P \) is nonempty. If \( A_z \) has rank \( n \), then \( z \) is a vertex and we are done. Else, consider any \( r \) such that \( A_z r = 0 \). We know that \( A \) has rank \( n \), and thus there exists a row \( a^i \) such that \( a^i \cdot r \neq 0 \). By choosing \( r \) or \(-r\), we can make sure \( a^i \cdot r > 0 \). Then let

\[
\epsilon = \min_{i \in \{1, \ldots, m\}} \left\{ \frac{b_i - a^i \cdot z}{a^i \cdot r} : i \not\in T_z, a^i \cdot r > 0 \right\}
\]

(recall that \( T_z \) is the index set of the rows in \( A_z \)). Let \( z' = z + \epsilon r \). Observe that \( A_z z' = A_z (z + \epsilon r) = A_z z + b \).

Thus, \( a^i \cdot z' = a^i \cdot (z + \epsilon r) = a^i \cdot z + (\frac{b_i - a^i \cdot z}{a^i \cdot r})(a^i \cdot r) = b_i \). Hence, we will at some point find a point \( v \in P \) where \( A_v \) has rank \( n \). At that point, \( v \) would be a vertex of \( P \).

7. (The Diet Problem) Suppose you want to design a diet for your meals. You have certain food items (e.g., spinach, chicken, rice etc.); let us label these different food types as \( f_1, \ldots, f_n \). You can choose any nonnegative amount of a food item to put in your diet. Each food item has a per unit cost \( c_1, \ldots, c_n \) associated with it. You have to meet some nutritional constraints: for example, you must have at least 5g of protein, and at most 40g of protein in the meal. Let
us say there are \( \{1, \ldots, k\} \) nutritional categories and each category has a lower bound \( \ell_i \) and an upper bound \( u_i \) that must be met. Suppose that each unit of food item \( f_j \) provides \( a_{ij} \) units of nutritional category \( i \). How will you solve the problem of designing a diet satisfying the nutritional demands that has the least cost?

Solution: Let \( x_1, \ldots, x_n \) denote the amounts of food items \( f_1, \ldots f_n \) in the best diet. Then, to satisfy the nutritional constraints for category \( i \in \{1, \ldots, k\} \), we must have

\[
\sum_{j=1}^{n} a_{ij} x_j \leq u_i, \quad \sum_{j=1}^{n} -a_{ij} x_j \leq -\ell_i
\]

Thus, we will have \( 2k \) constraints. The objective to minimize is \( \sum_{j=1}^{n} c_j x_j \). This is now a linear program (in the form \( \max \{c^T x : Ax \leq b\} \) discussed in class) with \( A = (a_{ij}) \), \( b = (u_1, \ldots, u_k, -\ell_1, \ldots, -\ell_k) \) and \( c = (-c_1, \ldots, -c_n) \).

8. (Linear Regression with different objectives) In linear regression, we have a bunch of labeled data points \( z^1, \ldots, z^k \in \mathbb{R}^n \) with real valued labels \( y_1, \ldots, y_k \). We want to fit the best linear function to this labeled data. More precisely, we want to find parameters \( \beta = (\beta_1, \ldots, \beta_n) \) so as to minimize the errors \( |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \). The typical objective is the sum of the squares of the errors, i.e., we wish to minimize \( \sum_{j=1}^{k} (y_j - \sum_{i=1}^{n} \beta_i z^j_i)^2 \). Suppose we are interested in the following variant:

Firstly, we don’t want to allow arbitrary values of the parameters; we want more controlled regression. Suppose for each parameter \( \beta_i \), we have certain preset upper and lower bounds \( u_i \) and \( \ell_i \) respectively that we want the parameter to lie within. Also, instead of minimizing the sum of squares, suppose we want to minimize the sum of the absolute values, i.e., minimize \( \sum_{j=1}^{k} |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \) subject to these bound constraints on the parameter values (\( \ell_1 \) minimization). How would you solve this problem? What if you were interested in minimizing the largest error, i.e., minimize \( \max_{j \in \{1, \ldots, k\}} |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \) (\( \ell_\infty \) minimization)?

Solution: Let \( \beta_1, \ldots, \beta_n \) be the parameter values. The bound constraints are easy:

\[
\beta_i \leq u_i, \quad -\beta_i \leq -\ell_i
\]

To model the objective \( \sum_{j=1}^{k} |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \), we introduce new auxiliary variables \( t_1, \ldots, t_k \) and impose the constraints:

\[
y_j - \sum_{i=1}^{n} \beta_i z^j_i \leq t_j, \quad -y_j + \sum_{i=1}^{n} \beta_i z^j_i \leq t_j,
\]

for every \( j = 1, \ldots, k \). This forces \( |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \leq t_j \) for all \( j = 1, \ldots, k \). If we now minimize

\[
\sum_{j=1}^{k} t_j
\]

subject to these constraints, then it is not hard argue that the \( t_j \) values will be pushed down to exactly \( |y_j - \sum_{i=1}^{n} \beta_i z^j_i| \), and out problem will be solved.

For the \( \ell_\infty \) version, we introduce a single auxiliary variable \( t \), and impose the constraints:

\[
y_j - \sum_{i=1}^{n} \beta_i z^j_i \leq t, \quad -y_j + \sum_{i=1}^{n} \beta_i z^j_i \leq t,
\]
for every $j = 1, \ldots, k$. This forces $|y_j - \sum_{i=1}^n \beta_i z_i^j| \leq t$ for all $j = 1, \ldots, k$. Now minimizing $t$ subject to these constraints will force $t$ to take the value of the maximum of $|y_j - \sum_{i=1}^n \beta_i z_i^j|$, which is exactly what we want.