Linear Optimization

CS5040 - JAN 2020

Assignment 2:

Write-Up

INSTRUCTOR: Sobhan Babu

Report By:
Vijay Tadikamalla
CS17BTECH11040

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Set of Theorems

Basic Linear Algebra:

- Given a set of equations from $A_{1..n}$, suppose A_j is replaced by $\alpha A_i + A_j$ then the solution set does not change.
- Suppose $S = \{u_1, u_2, \ldots, u_m\}$ and $T = \{v_1, v_2, \ldots, v_n\}$ be two sets of vectors such that each is a basis for the vector space V . Then m = n.
- Let S = {x: Ax = 0} and suppose k is the number of linearly independent columns in the matrix A. Then dim(S) = n - k.
- Let x_0 be a vector in R^n such that $Ax_0 = b$. Then every solution to Ax = b can be written in the form $x_o + x'$, where x_0 is any vector satisfying $Ax' = 0$.
- **Linear Independence and Rank:** In a matrix, no. of linearly independent rows is equal to no. of linearly independent columns. (Rank of the matrix)

Theorems on Vector spaces, subspaces:

- **Dimension theorem for vector spaces -** It states that all bases of a vector space have equally many elements. This number of elements may be finite or infinite (in the latter case, it is a cardinal number), and defines the dimension of the vector space.
- A vector can be expressed as a linear combination of the vectors that form a basis for its vector space in exactly one way.
- Every vector space has an orthonormal basis. An orthonormal basis for an inner product space V with finite dimension is a basis for V whose vectors are orthonormal, that is, they are all unit vectors and orthogonal to each other. For example, the standard basis for a Euclidean space R^n is an orthonormal basis, where the relevant inner product is the dot product of vectors.
- **Rank Nullity theorem:** The theorem states that the rank and the nullity (the dimension of the kernel) sum to the number of columns in a given matrix, i.e. for a matrix A_{m*n}, $rank(A) + nullity(A) = n$.

Some theorems related to Linear Optimization:

- A point is a vertex iff number of columns in a tight constraint is equal to the rank of the matrix. (**Tight constraint:** An inequality constraint is tight at a certain point if the point lies on the corresponding hyperplane.)
- In a Linear Optimisation problem, the optimal point is present at one of the vertices.

Simplex algorithm

Goal

Simplex algorithm is a method to solve the optimization problem of linear programming.

Approach

- Identify the first feasible point. A feasible point is a point satisfying all the constraints of a linear optimization problem.
- Start moving from this point until you reach an extreme point. An extreme point is a solution to some *n* equations of $AX \leq b$.
- Move to a neighboring extreme point of greater cost if one exists and repeat this step. If no such neighbor exists then exit with this point as the optimum point

Algorithm (Non - Degenerate Case)

We first reformulate the problem into the standard form, i.e

$$max \;\; c^T x \;\; given \; Ax \leq b$$

To find the initial feasible point, we apply the following method:

1. If $b_i > 0 \;\forall\; i$, then the origin is the initial feasible point.
2. Else we add an extra term, $\min(b_i)$ in each equation and add a new equation $z_i \leq -\min(b_i)$. So, the new initial point is equal to z_0.

Let the initial feasible point obtained from the above process be z_0. Now to find the initial vertex/extreme point, we apply the following method:

1. If z_0 does not satisfy any constraint, we take a random vector and move in its direction until some constraint is satisfied, say at y_0.

2. Now there exists a constraint satisfied at y_0. Now we find a vector from the null space and move in its direction until the number of constraints satisfied at this point becomes equal to n, i.e rank of the matrix.

Let the initial vertex/extreme point obtained from the above process be x_0.

There are n planes intersecting at that extreme point, with n neighboring points and there are n neighboring points. Among these n points, the points with lower costs are the only desired.
So, we have to find the directions that lead to all neighboring points.

The point x_0 satisfies n out of m inequalities. We call the rows satisfying equality as **tight rows** and others as **non-tight rows**. Let the constraints formed using tight rows be

$A'\ x_0 \le b'$ and due to non-tight rows be $A''\ x_0 \le b''$

Consider a point x_i in a line l_i passing through x_0 which connects it to some neighbor. This line is determined by the intersection of some (n − 1) rows out of the n linearly independent rows of *A'*. Any point in this line satisfied
(n − 1) if these inequalities with equality and one with strict inequality. Assume that the i^{th} one is strict inequality.

$$A'_j\, x_j = b'_j,\ 1 \le j \le n,\ j \ne i$$
$$A'_i\, x_i < b'_i$$

The direction of the vector along the line l_i front the point x_0 towards x_i is $x_i - x_0$.
Using the above constraints, we can show that $A'(x_i - x_0)$ is a vector of zeros with only i^{th} position as negative. The magnitude of this negative value depends on the distance of x_i from x_0. Consider this value as −1.
Thus, $A'(x_i - x_0) = -ei$ which is a vector with −1 at i^{th} index and rest all zeros.
Now consider the matrix $Z = [x_1 - x_0, x_2 - x_0, \ldots, x_n - x_0]$.
$A'Z = [A'x_1 - A'x_0, \ldots A'x_n - A'x_0]$. With the above discussed properties of x_i's

$$(A'x_i)_j = (A'x_0)_j,\ j \ne i$$
$$(A'x_i)_I - (A'x_0)_i = -1$$

For each i, the i^{th} column of the matrix *A'Z* has *a − 1* in the i^{th} position and zeros in all other positions. Thus. $A'Z = -I$.

Theorem: The direction vectors are the columns of the negative of the inverse of matrix A'.

Now we know the direction vertex to the neighbors. In this way, when we know a vertex x_0, we can move to another neighboring vertex at a higher cost. This can be continued till all the neighboring points satisfying the constraints have a lower cost.

The general problem can also be degenerate. This means, some constraints can have the same RHS. In such cases, we perturb the vector b slightly and then solve the problem. The solution obtained is an approximate solution. However, considering smaller perturbations can make the solution closer to the theoretical solution.

Vertex Cover Problem

The vertex cover problem is **NP-hard.** However, it can be approximately solved using the simplex algorithm. Consider the graph *G = (V, E)*. Each vertex is given a variable x_i. Considering the problem statement, we have to select at least one vertex on each edge. Thus, for an edge with vertices x_i. and x_j.,and add constraints, $x_i + x_j \geq 1$, $x_i \geq 0$ and $x_j \geq 0$. Here, each variable denotes the number of selections that happen for a vertex. We add a similar constraint for each edge in the graph.

The vertex cover problem tries to minimize the number of vertices covered. Since every vertex has been allocated a variable that denotes the number P of selections, we minimize the sum of variables. Thus, our objective is

$$z = \Sigma x_i$$

Consider the case where the vertices of the graph are weighted, with weights w_i. In this case, the constraints remain the same. However, the objective is $z = \Sigma w_i x_i$.

The solution obtained from this problem will contain floating-point numbers. However, the selection of vertices cannot be a fraction. In order to get integral solutions, we round off the x_i to the nearest integer. It can also be proved that the rounding does not affect constraints. But it can only affect the cost.

Primal is equal to dual

Separating Hyperplane Theorem: Let A and B be two convex sets in R^n that do not intersect (i.e., A ∩ B = φ). Then, there exists, $a \in R^n$, $a \neq 0$, $b \in R$, such that $a^T x \leq b \ \forall \ x \in A$ and $a^T x \geq b \ \forall \ x \in B$.

Primal: maximize

$$z = \sum_{j=1}^{n} c_j x_j$$

Constraint:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, 2, \ldots, m$$

where

$$x_j \geq 0, j = 1, 2, \ldots, n$$

The associated dual problem is given by
Dual: minimize:

$$v = \sum_{i=1}^{m} b_i y_i$$

Constraint:

$$\sum_{i=1}^{m} a_{ij} y_i \geq c_j, j = 1, 2, \ldots, n$$

where

$$y_i \geq 0, i = 1, 2, \ldots, m$$

Theorem: If the primal is feasible and has a finite optimum then the dual is also feasible, has a finite optimum and their optimums must be equal.

Proof: We first show that the dual is feasible. This means we can find a y that satisfies $A^T y = c$, $y \geq 0$. Rows of A are also the outward normals to the hyperplanes in $Ax \leq b$. The inequality $y \leq 0$ means the coefficients should be non-negative.

It can be shown that an optimum point of the primal can be written as a non-negative linear combination of the normals to the defining hyperplanes. The coefficients of this non-negative linear combination yield a feasible point in the dual.

In other words, we can show that for each point infeasible region F, we can find a feasible point in the dual. The set F is non-empty since it contains at least the optimum point x_0. Hence we infer that the dual is feasible.
Let y be any feasible point in the dual. Let x_0 be an optimal point in the primal, then the cost of y in the dual is $y^Tb \geq y^TAx_0 = c^Tx_0$. Hence the dual cost is bounded from below by c^Tx_0 . To prove the last part, we only need to observe that since x_0 is in F there is a feasible y_0 corresponding to it at the same cost. This y_0 has to be the optimum point. Hence the two optimums are equal.