Linear Optimization

Ritesh Yadav

Assignment 2

1 Summary

A mathematical optimization model consists of an objective function and a set of constraints in the form of a system of equations or inequalities.Linear Optimization deals with a class of optimization problems, where both the objective function to be optimized and all the constraints, are linear in terms of the decision variables.
To understand how the set of all solutions look like we need a vocabulary to solve these Linear Optimization problems.This comes from Linear Algebra.

Vector Space : A vector space is defined as a set of vectors **V** and the real numbers **R** with the following operations defined :

- **Vector Addition :** Defined as $\mathbf{V}\times\mathbf{V} \rightarrow \mathbf{V}$, represented as $\mathbf{u}+\mathbf{v}$,where $\mathbf{u},\mathbf{v}$ ϵ $\mathbf{V}$.
- **Scalar Multiplication :** Defined as $\mathbf{R}\times\mathbf{V}\rightarrow\mathbf{V}$, represented as a.$\mathbf{u}$, where a ϵ $\mathbf{R}$ and $\mathbf{u}$ ϵ $\mathbf{V}$.

Subspace : $\mathbf{U} \subset \mathbf{V}$ is a subspace of $\mathbf{V}$ if $\mathbf{U}$ itself is vector space, i.e. for all $\mathbf{u}_1$,$\mathbf{u}_2$ ϵ $\mathbf{U}$ and α ϵ $\mathbf{R}$ **,** $\mathbf{u}_1$+$\mathbf{u}_2$ ϵ $\mathbf{U}$ and $\alpha.\mathbf{u}_1\epsilon$ $\mathbf{U}$.

Linear Dependent : Vectors $\mathbf{v}_1, ...\mathbf{v}_n$ are linearly dependent if their exist $\alpha_1......\alpha_n\epsilon$ $\mathbf{R}$, not all zero, such that

$$\sum_{i=1}^{n} \alpha_i.v_i = \mathbf{0}$$

Linear Independent : Vectors $\mathbf{v}_1, ...\mathbf{v}_n$ are linearly independent if they are not linearly dependent, i.e. for $\alpha_1......\alpha_n\epsilon$ $\mathbf{R}$

$$\sum_{i=1}^{n} \alpha_i.v_i = \mathbf{0} \Rightarrow \alpha_i = 0, \forall i$$

Basis : Vectors $\mathbf{v}_1, ...\mathbf{v}_n$ form the basis of a vector space $\mathbf{V}$ iff :

- they are linearly independent
- every other vector $\mathbf{w}$ which belongs to $\mathbf{V}$ can be written as

$$\mathbf{w} = \sum_{i=1}^{n} \beta_i.v_i$$

Note : Number of elements in basis is **Cardinality** of vector space.

Theorem 1 : *Suppose $S = \{u_1,....,u_n\}$ and $T = \{v_1,....,v_m\}$ be two sets of vectors such that each is basis for the vector space V. Then m=n.*

Theorem 2 : *If **U** is basis and **U** $\subset$ **V** then **V** will have some vectors that are not in U, so that vectors can be written as Linear Combination of vectors of **U**. So **V** won't be linearly independent.*

Theorem 3 : *A vector can be expressed as linear combination of basis in a unique manner only.*

Orthonormal Basis : In any vector space when two different basis are orthogonal to each other and consists of unit magnitude, they are called as orthonormal basis.

Theorem 4 : *Every vector space have a orthonormal basis.*

Theorem 5 : *In a matrix A, number of linearly independent rows is equal to number of linearly independent columns aka **rank of a matrix A**.*

Theorem 6 : *Inverse of a matrix A exist if all rows or columns are independent.*

Let A be a m $\times$ n matrix. And S be **null space** of A s.t. S = $\{$x : Ax = $\mathbf{0}\}$, where x is a n-dimensional vector.

Theorem 7 : *Suppose k is the number of linearly independent columns in the matrix A. Then, dim(S) = n-k.*

Note : Nullity is define as number of linearly independent vectors present in null space of a matrix **OR** dimension of null space.

Theorem 8 : Rank-Nullity Theorem : *Nullity of A + Rank of A = Total number of Columns of A.*

Note : An inequality constraint is tight at a certain point if the point lies on the corresponding hyperplane.

Theorem 9 : *A point is a **vertex** iff number of columns in tight constraint matrix equals to rank of a matrix.*

Theorem 10 : *In a convex problem optimal solution may exist at internal points, but we will definitely have solution at vertex.*

Theorem 11 : *Let f be a linear function over a convex set S. Then a local maximum is a global maximum.*

Theorem 12 : *Let $p_1, p_2, ..., p_n$ be extreme points of $x : Ax \leq \boldsymbol{b}$. Then every point in $x : Ax \leq \boldsymbol{b}$ can be expressed as a convex combination of the points $p_1, p_2, ..., p_n$.*

Theorem 13 : *A linear function on $S = \{x : Ax \leq \boldsymbol{b}\}$ is maximized at an extreme point i.e. vertices.*

2 The Simplex Algorithm

The Simplex Algorithm considers extreme points in a certain order by visiting extreme points one by one. It may not be able to visit extreme points of the polytope but when it terminates, it will find the optimum.

The Simplex Algorithm consists of the following two steps :

- Begin at an extreme point
- Move to a neighbouring extreme point of greater cost if one exists and repeat this step. If no such neighbour exists then exit with this point as the optimum point.

Let us proceed for implementations of the simplex algorithm on n-dimensions. Consider the usual linear optimization problem with constraints Ax $\leq$ b and cost function c^Tx. Consider an extreme points x_0 such that it satisfies

$$A^{'}x_0 = b^{'} A^{''}x_0 < b^{''} \tag{1}$$

Here $A^{'}$is the matrix formed by some n linearly independent rows of the matrix A and $b^{'}$ the corresponding entries from b. These rows are satisfied by x_0 with equality. The other rows constitute $A^{''}$. In a sense, the first set of equalities define the extreme point. These equalities are said to be the support of the extreme point.

Further we will proceed by determining the neighbours of x_0. To find one neighbour, it makes sense to remove one of the hyper-planes from $A^{'}$, add one from $A^{''}$ and check if the resultant point is feasible point. The points we get thus, which are feasible, are the neighbours.

Theorem 14 : *For each hyper-plane in $A^{'}$, there will be exactly one such hyper-plane in $A^{''}$, for which the resultant point will be feasible.*

There are n rays emerging from x_0 which connect it to its neighbours. We have to determine the vectors corresponding to these directions. As we shall see this can be determined rapidly, and once we have the directions, we can also determine the neighbours efficiently. Consider a point x_i on a line l_i passing through x_0 which connects it to some neighbour. This line is determined by the intersection of some $n-1$ rows out of the n linearly independent rows of $A^{'}$. That is, any point in this line satisfies $n-1$ of these inequalities with equality and one with strict inequality. Assume that the ith one is the strict inequality

$$A^{'}_j x_j = b^{'}_j : 1 \leq j \leq n, j \neq i \tag{2}$$

$$A^{'}_i x_i = b^{'}_i \tag{3}$$

The direction of the vector along the line l_i from the point x_0 towards x_i is $x_i - x_0$.
$A^{'}(x_i - x_0)$ is a vector with zeroes everywhere except in the ith position which contains a strictly negative value. The constant at the ith position depends on where on the line we picked x_i. The farther it is from x_0, the greater the magnitude. Let us pick it so that $A^{'}(x_i - x_0)$ = -e^i, the vector with zeroes everywhere and -1 in the ith position.

Consider a matrix Z, having as columns, vectors $(x_i - x_0)$, for i = 1, 2,....n.

$$Z = [x_1 - x_0, x_2 - x_0, ..., x_n - x_0] \tag{4}$$

$$A^{'}Z = [A^{'}x_1 - A^{'}x_0, A^{'}x_2 - A^{'}x_0, ..., A^{'}x_n - A^{'}x_0] \tag{5}$$

From the above properties of x_i we can imply that for each i, the i^{th} column of the matrix $A^{'}$Z has a -1 in the i^{th} position and zeroes in all other positions, i.e.

$$A^{'}Z = -I \tag{6}$$

Theorem 15 : *The direction vectors are the columns of the negative of the inverse of matrix $A^{'}$.*

Once we know the direction of each neighbour, we first determine which of the neighbours has greater cost. Note that if one of the neighbours x_i has greater cost then every point on the line segment joining x_i and x_0 has cost more than that of x_0. Similarly we will follow the process iteratively until we reach our optimal x_0

Theorem 16 : *If the cost does not increase along any of the columns of $-\ A^{'-1}$ then x_0 is optimal.*

We will proceed till cost vector is linear combination of non-negative tight vectors at optimal point. **Haulting Condition :** *The algorithm stops at an extreme point x_0 and returns it as optimal when the cost at x_0 is greater than or equal to the cost at the neighbouring extreme points.*

3 Duality Theory

The notion of duality within linear programming asserts that every linear program has associated with it a related linear program called its **dual**. The original problem in relation to its dual is termed the **primal**. Duality theory is the relationship between the primal and its dual, both on a mathematical and economic level, that is truly the essence of duality theory.

Standard form of Primal Problem :

$$\begin{aligned} \max_x Z = \sum_{j=1}^{n} c_j.x_j \\ a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \leq b_2 \\ ... \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \leq b_m \\ x_1, x_2, ..., x_n \geq 0 \end{aligned} \tag{7}$$

Standard form of Dual Problem :

$$\begin{aligned} \min_y w = \sum_{i=1}^{m} b_i.y_i \\ a_{11}y_1 + a_{21}y_2 + ... + a_{m1}y_m \geq c_1 \\ a_{12}y_1 + a_{22}y_2 + ... + a_{m2}y_m \geq c_2 \\ ... \\ ... \\ a_{1n}x_1 + a_{2n}y_2 + ... + a_{mn}y_m \geq c_n \\ y_1, y_2, ..., y_m \geq 0 \end{aligned} \tag{8}$$

Important points regarding primal-dual pairs :

- In the above standard forms **b** is not assumed to be non-negative.
- If the primal problem has **n variables** and **m constraints**, the dual problem will have **m variables** and **n constraints**.
- Coefficient of the objective function in the dual problem come from the right-hand side of the original problem.
- The coefficient of the first constraint function for the dual problem are the coefficients of the first variable in the constraints for the original problem, and the similarly for other constraints.
- The right-hand sides of the dual constraints come from the objective function coefficients in the original problem.
- **The dual of the dual is the Primal.**

Duality 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.*

Consider the following primal-dual pair :

$$\begin{aligned} P : \max_x c^T x \qquad\qquad & D : \min_y y^T b \\ Ax \leq \mathbf{b} \qquad\qquad & A^T y \geq \mathbf{c} \\ x \geq \mathbf{0} \qquad\qquad & y \geq \mathbf{0} \end{aligned} \tag{9}$$

The duality theorem in its complete form states the following :

- If P is feasible and has a finite maximum then D is feasible and the two optimum values coincide.
- If P is infeasible and D is feasible then D is unbounded.
- If P feasible and unbounded then D is infeasible.

Note : If the original problem is **max** model, the dual is a **min** model; if the original problem is a **min** model, the dual problem is the **max** problem.

Weak Duality : *The objective function value of the primal(dual) to be maximized evaluated at any primal(dual) feasible solution cannot exceed the dual(primal) objective function value evaluated at a dual(primal) feasible solution. In the standard equality form :*

$$c^T x \geq y^T b \tag{10}$$

Strong Duality : *If primal LP is feasible and has a finite optimum the dual LP is feasible and has a finite optimum. Furthermore, if $x*$ and $y*$ are optimal solutions for primal LP and dual LP then :*

$$c^T x^* = y^{*T} b \tag{11}$$

THE END