Strong Duality: Without Simplex and without theorems of alternatives

By

Somdeb Lahiri

SPM, PDPU.

email: somdeb.lahiri@yahoo.co.in

November 29, 2015.

Abstract

We provide an alternative proof of the strong duality theorem whose main step is a proposition which says that every canonical linear programming minimization problem whose image under its objective function of the set of feasible solutions is non-empty and bounded below has an optimal solution. Unlike earlier proofs, this proof neither uses the simplex method, nor does it use Farkas's lemma. We also use this proposition to obtain an independent proof of Farkas's lemma.

1. Introduction

The purpose of this note is very simple. It is to prove the strong duality theorem of linear programming (LP) without either using the simplex method or by using any theorem of alternatives.

The simplex method has its own problems related to degenerate basic feasible solutions. While such solutions are infrequent, from a theoretical standpoint a proof of the strong duality theorem that uses the simplex method is not complete until it has taken a few extra steps. Further, for economists the duality theorem is extremely important whereas the simplex method is not necessarily so. If we add to it the fact, that the simplex method has faster substitutes for computational purpose, an alternative proof of the strong duality theorem which does not use the simplex method would be very welcome.

The alternative route is to use the Farkas's lemma or a theorem of alternative than can be derived from it. Such proofs while extremely elegant pre-empt deriving Farkas's lemma itself from the strong duality theorem of LP. Thus, it would be very desirable to have a proof of the strong duality theorem of LP which does not use any theorem of alternative either. Such a proof has been provided in this paper.

The crucial step in our proof is a proposition which says the following: if a canonical LP minimization problem is such that the image under its objective function of the set of feasible solutions is non-empty and bounded below, then this LP problem has an optimal solution.

A general all purpose reference for material presented in this paper is Dorfman, Samuelson and Solow (1958). A very purposive and lucid exposition of the role of linear programming in microeconomic analysis in the past as well its future prospects is available in Thompson (1972).

2. The Primal and Dual LP problems

Let M and N be positive integers and let A be a M×N real matrix such that (i) every row of A has at least one non-zero entry, and (ii) every column of A has at least one non-zero entry. Let b be a real M-vector i.e. $b\in \mathbb{R}^M$ and c be a real N-vector i.e. $c\in \mathbb{R}^N$. All vectors are assumed to be column vectors unless otherwise mentioned. To distinguish the transpose of a column vector or a matrix from the original column vector or matrix we use a "superscript" 'T'.

The **canonical primal linear programming** is (by definition) the following:

Minimize $c^T x$

subject to $Ax = b$, $x\in \mathbb{R}^N_+$.

It is denoted by (P).

The dual of (P) is the following:

Maximize $y^T b$

subject to $y^T A \leq c^T$, $y\in \mathbb{R}^M$.

It is denoted by (D).

A vector x is said to be **feasible for (P)** if $Ax = b$, $x\in \mathbb{R}^N_+$.

A vector y is said to be **feasible for (D)** if $y^T A \leq c^T$, $y\in \mathbb{R}^M$.

A vector x which solves (P) is said to be an **optimal solution for (P)**.

A vector y which solves (D) is said to be an **optimal solution for (D)**.

The **optimal value of (P)** is the value of the objective function of (P) at any optimal solution of (P).

The **optimal value of (D)** is the value of the objective function of (D) at any optimal solution of (D).

The strong duality theorem of linear programming says that if both (P) and (D) have feasible solutions then both have optimal solutions and the optimal value of both are the same.

3. Basic Solutions

Let A^j denote the j^{th} column of A and A_i denote its i^{th} column.

A vector x is said to be **basic for P** if $x \in \mathbb{R}^N$ and the list of columns $\langle A^j | x_j > 0 \rangle$ is linearly independent. If in addition x is feasible for (P) then x is said to be a **basic feasible solution for (P)**.

If x is a basic feasible solution for (P), then $\langle A^j | x_j > 0 \rangle$ is denoted by B, x_B is the sub-vector of x corresponding to the columns in B. Further in such a situation we often write the equation Ax = b as $[B|E]\binom{x_B}{0} = b$ or $Bx_B = b$.

Note that if B has linearly independent columns then, B^TB is an invertible square matrix, (i.e. has full column and row rank).

An optimal solution for (P) is basic for (P) it is called a **basic optimal solution for (P)**.

The obvious proof of the following proposition is being omitted.

Proposition 1 If x is feasible for (P) and y is feasible for (D), then $c^Tx \geq y^Tb$.

Proposition 2 Suppose (P) has a feasible solutions and the image of the feasible set of (P) under the objective function of (P) is bounded below. Then there is a basic feasible solution for (P) such that the value of the objective function for (P) at this basic feasible solution does not exceed c^Tx.

Proof Let $F(P) = \{x \in \mathbb{R}^N_+ | Ax = b\}$. By hypothesis both F(P) is non-empty.

Let $VF(P) = \{c^Tx | x \in F(P)\}$. Clearly VF(P) is non-empty. Further by hypothesis VF(P) is bounded below. Let $\alpha \in \mathbb{R}$ such that $c^Tx \geq \alpha$ for all $x \in F(P)$.

Let x be feasible for (P). If x is basic then we are done. Hence suppose x is not basic. Thus the list of columns $\langle A^j | x_j > 0 \rangle$ are linearly dependent. Hence there exists a list of real numbers $\langle \lambda_j | x_j > 0 \rangle$ not all of which are zero such that $\sum_{x_j > 0} \lambda_j A^j = 0$.

Case 1: $\sum_{x_j > 0} c_j \lambda_j > 0$.

If $\lambda_j \leq 0$ for all j, then the N-vector x(t) whose j^{th} coordinate is 0 if $x_j = 0$, and whose j^{th} coordinate is $x_j - t\lambda_j$ if $x_j > 0$, satisfies $Ax(t) = b$ for all $t \geq 0$ and $x(t) \in \mathbb{R}^N_+$. Futher $c^Tx(t)$ diverges to $-\infty$ as $t \to \infty$, contradicting our assumption which requires $c^Tx(t) \geq \alpha$ for all $t \geq 0$.

Hence $\lambda_j > 0$ for some j.

Let $\mu = \max\{t \geq 0 | x_j - t\lambda_j \geq 0$ for all j satisfying $x_j > 0$ and $\lambda_j > 0\}$.

Then $Ax(\mu) = b$, $x(\mu) \in \mathbb{R}^N_+$ and $|\{j | x_j(\mu) > 0\}| < |\{j | x_j(\mu) > 0\}|$.

Further, $c^Tx(\mu) < c^Tx$.

Case 2: $\sum_{x_j > 0} c_j \lambda_j < 0$.

In this case $\sum_{x_j > 0} c_j(-\lambda_j) < 0$.

Repeat case 1 with λ replaced by $-\lambda$ to obtain a μ and a $x(\mu)$ as before such that j^{th} coordinate of $x(\mu)$ is 0 if $x_j = 0$, and whose j^{th} coordinate is $x_j + \mu\lambda_j$ if $x_j > 0$. Then $Ax(\mu) = b$, $x(\mu) \in \mathbb{R}_+^N$ and $|\{j|\ x_j(\mu) > 0\}| < |\{j|\ x_j(\mu) > 0\}|$.

Further, $c^T x(\mu) < c^T x$.

Case 3: $\sum_{x_j>0} c_j \lambda_j = 0$.

If $\lambda_j > 0$, then proceed as in case 1 with λ; if not consider $-\lambda$ instead of λ and proceed as before. In either case we obtain a $x(\mu) \in \mathbb{R}_+^N$ such that $Ax(\mu) = b$ and $|\{j|\ x_j(\mu) > 0\}| < |\{j|\ x_j(\mu) > 0\}|$.

Further, $c^T x(\mu) = c^T x$.

Thus, there exists an $x(\mu) \in \mathbb{R}_+^N$ such that $Ax(\mu) = b$ and $|\{j|\ x_j(\mu) > 0\}| < |\{j|\ x_j(\mu) > 0\}|$.

Further, $c^T x(\mu) \leq c^T x$.

The process terminates once we have a basic feasible solution. The value of the objective function at this basic feasible solution does not exceed $c^T x$. Q.E.D.

The following corollary of proposition 2 follows once we take notice of Proposition 1.

Corollary of Proposition 2 Suppose that both (P) and (D) have feasible solutions. Let x be a feasible solution for (P). Then there is a basic feasible solution for (P) such that the value of the objective function for (P) at this basic feasible solution does not exceed $c^T x$.

Proposition 3 Suppose (P) has a feasible solutions and the image of the feasible set of (P) under the objective function of (P) is bounded below. Then (P) has a basic optimal solution.

Proof: As in the proof of proposition 2, let $F(P) = \{x \in \mathbb{R}_+^N | Ax = b\}$. By hypothesis F(P) is non-empty.

Let $VF(P) = \{c^T x | x \in F(P)\}$. Clearly VF(P) is non-empty. Further by hypothesis VF(P) is bounded below.

Each basic feasible solution is of the form $[((B^TB)^{-1}B^Tb)^T, 0]^T$ where B is a submatrix of A whose columns are linearly independent and further all coordinates of $(B^TB)^{-1}B^Tb)$ are non-negative. Clearly, there are only a finite number of basic feasible solutions since there only finitely many collections of linearly independent columns of A.

Let $\hat{x}$ be a basic feasible solution such that $c^T\hat{x} = \min\{c^T x|\ x$ is a basic feasible solution$\}$.

Since VF(P) is bounded below, $\hat{x}$ must be an optimal solution for P, since if x is any other feasible solution with $c^T x < c^T\hat{x}$, then we would get $c^T x < c^T x'$ for all basic feasible solutions, thereby contradicting proposition 2. This proves the proposition. Q.E.D.

Corollary of Proposition 3 Suppose both (P) and (D) have feasible solutions. Then (P) has a basic optimal solution.

Proof Follows from Proposition 1 by observing that if both (P) and (D) have feasible solutions, then the image of the feasible set of (P) under the objective function of (P) is bounded below. Q.E.D.

4. Strong Duality Theorem

We begin this section with a lemma whose proof is easy.

Lemma 1 If x is a feasible solution for (P) and y is a feasible solution for (D) and if the value of the objective function for (P) at x is equal to the value of the objective function for (D) at y, then x is an optimal solution for (P) and y is an optimal solution for (D).

The next proposition is the key to the Strong Duality Theorem of LP.

Proposition 4 Suppose (P) has an optimal solution. Then (D) also has an optimal solution and the optimal value of both are the same.

Proof If (P) has an optimal solution then by Proposition 2, it has an optimal basic solution $\binom{x_B}{0}$ corresponding to the linearly independent columns B of A, i.e. A = [B|E] and $x_B = (B^TB)^{-1}Bb$.

Let x be any other feasible solution for P. Let x = $\binom{x(1)}{x(2)}$, where x(1) corresponds to the columns in E.

Then, $x(1) = (B^TB)^{-1}B^T(b-Ex(2)) = x_B - (B^TB)^{-1}B^TEx(2)$

Thus, $c_B^Tx(1) + c_E^Tx(2) = c_B^Tx_B - c_B^T(B^TB)^{-1}B^TEx(2) + c_E^Tx(2) = c_B^Tx_B + (c_E^T - c_B^T(B^TB)^{-1}B^TE)x(2)$.

Towards a contradiction suppose $(c_E^T - c_B^T(B^TB)^{-1}B^TE)_j < 0$, for some j corresponding to a non-basic column A^j of A.

Consider the vector x(2), where $x_j(2) = t \geq 0$, $x_k(2) = 0$, for all other k where k corresponds to a non-basic column A^k of A.

For t = 0, $x(1) = x_B >> 0$ and so for t > 0 sufficiently small $x(1) >> 0$. Further, $A\binom{x(1)}{x(2)} = b$.

Also, $c_B^Tx(1) + c_E^Tx(2) = c_B^Tx_B + t(c_E^T - c_B^T(B^TB)^{-1}B^TE)_j < c_B^Tx_B$ since $t(c_E^T - c_B^T(B^TB)^{-1}B^TE)_j < 0$.

This contradicts the optimality of $\binom{x_B}{0}$.

Thus it must be the case that $c_E^T - c_B^T(B^TB)^{-1}B^TE \geq 0$. Further, $c_B^T - c_B^T(B^TB)^{-1}B^TB = 0$.

Thus $c_j^T - c_B^T(B^TB)^{-1}B^TA^j \geq 0$ for all j = 1,...,N.

Thus, $c^T - c_B^T(B^TB)^{-1}B^TA \geq 0$.

Let $y^T = c_B^T(B^TB)^{-1}B^T$. Thus, $y^TA \leq c^T$. Thus y is feasible for (D).

Further $y^T b = c_B^T (B^T B)^{-1} B^T b = c_B^T x_B$.

By lemma 1, y is an optimal solution for (D) and the optimal value of (P) is equal to the optimal value of (D). Q.E.D.

Note The method we have adopted to prove proposition 4, is the one used by simplex to obtain an optimal solution for the dual given a basic optimal solution for the primal.

The following well-known result now follows immediately from the Corollary of Proposition 3 and Proposition 4.

Strong Duality Theorem of LP: If both (P) and (D) have feasible solutions then both have optimal solutions and the optimal value of both are the same.

5. Farkas's lemma

A very simple proof of the well-known Farkas's lemaa follows very easily from Proposition 3.

Farkas's lemma *Either* Ax = b has a solution in $\mathbb{R}_+^N$ *or* $y^T A \leq 0$, $y^T b > 0$ has a non-negative solution, but never both.

Proof Since the proof of "never both is standard" let us suppose Ax = b, $x \in \mathbb{R}_+^N$ does not have a solution.

If $y^T A \leq 0$, $y^T b > 0$ does not have a solution, then since $0^T A \leq 0$, $0^T b = 0$, 0 is an optimal solution for the LP problem

Maximize $y^T b$

subject to $y^T A \leq 0$.

Thus (0,0,0) is an optimal solution for the LP problem

Minimize $y_1^T(-)b + y_2^T b + w^T 0$

Subject to $(A^T | -A^T | I)\begin{pmatrix} y_1 \\ y_2 \\ w \end{pmatrix} = 0$,

$\begin{pmatrix} y_1 \\ y_2 \\ w \end{pmatrix} \in \mathbb{R}_+^N \times \mathbb{R}_+^N \times \mathbb{R}_+^M$.

By Proposition 4, its dual

Maximize $0^T x$

Subject to $x^T(A^T | -A^T | I) \leq (-b^T | b^T | 0)$

has an optimal solution.

Now $x^T(A^T| -A^T|I) \leq (-b^T|b^T|0)$ is equivalent to Ax = b and x ≤ 0.

Thus the system Ax = b and x ≤ 0 has a solution. The negative of any solution to this system satisfies will satisfy the system Ax = b, $x \in \mathbb{R}_+^N$.

Thus, Ax = b, $x \in \mathbb{R}_+^N$ has a solution leading to a contradiction. This proves the lemma. Q.E.D.

6. Conclusion

After getting the main results in the paper, I was able to locate an unpublished 2008 paper entitled "An Elementary Proof Of Optimality Conditions For Linear Programming" by Anders Forsgren, whose stated objective is similar to ours. They rely on a perturbation technique to bypass problems concerning non-degenerate basic feasible solutions. However, this non-degeneracy problem is also the shortcoming of the simplex method and our technique of proof makes no distinction between degenerate and non-degenerate basis feasible solution. The method of proof of lemma 3.1 in the Forsgren paper, is similar to the proof of our proposition 2. The real novelty of our paper is proposition 3, which to the best of my knowledge has no precedents. Everything put together our proof is simpler and shorter assuming that Forsgren has succeeded in achieving his goal.

There is a result in Frank and Wolfe (1956) which is similar in spirit to proposition 3, and says that a quadratic programming problem admits an optimal solution if the objective function is bounded from below on the feasible set. They do not show that such a solution is basic feasible and we require basic feasibility to prove the strong duality theorem. Further, to refer to the Frank and Wolfe to show that a *linear programming* problem admits an optimal solution if the objective function is bounded from below on the feasible set, would be like "using a sledge-hammer to break a pea pod".

Acknowledgment I would like to thank T Parthasarathy for very valuable comments on an earlier version of this paper which led to significant changes that have been incorporated here. I would also like to thank Roberto Serrano for his comments on this paper and to Henrik Alsing Friberg for observations which were "bang on". However, the sole responsibility for errors that do remain rests with the author.

References

1. Dorfman, R., P. Samuelson, and R. Solow (1958): "Linear Programming and Economic Analysis", The Rand Corporation.

2. Frank, M. and P. Wolfe (1956): "An algorithm for quadratic programming", Naval Research Logistics Quarterly, Vol. 3, pages 95- 110.

3. Thompson, G.E. (1972): "Linear Programming and Microeconomic Analysis", Nebraska Journal of Economics and Business, Vol. 11, No.4, Midwest Economics Association Papers, pages 25-36.