IJournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 11, November-2014

The relationship between theorems of the alternative, least norm problems and steepest descent directions

Author: Sunaina Hedav[1], Kirti Naik[2]

E-mail: sunainahedav@gmail.com, kirtinaik25.n@gmail.com

ABSTRACT

This paper is a partial review of reference [1]. It reviews the relationship between theorems of the alternative, least norm problems and steepest descent directions.

If a primal system characterizes a descent direction then a given point is optimal if and only if the dual system is satisfied. Each theorem of the alternative can be formulated as bounded least squares (BLS) problem. Further, each theorem of the alternative is related to a certain "steepest descent" problem whose solution can be obtained by solving the corresponding BLS problem.

If the residue vector vanishes then a given point is optimal or a minimizer of BLS and solves the dual. Converse is also true, which is obvious. Otherwise, r is a steepest descent direction and it solves the primal problem.

Keywords

Theorems of the alternative, least norm problem, steepest descent directions.

INTRODUCTION

Here we formulate BLS problem ,equivalent to Farkas' lemma and by solving this problem we will get the solution of Farkas' lemma.

Farkas' Lemma 1.1: Let A be a real $m \times n$ matrix and let g be a real non-zero n-vector. Then either the primal system

$$Ax \geq 0 \text{ and } g^t x < 0 \qquad (1)$$

has a solution $x \in R^n$, or the dual system

$$A^t y = g \text{ and } y \geq 0 \qquad (2)$$

has a solution $y \in R^m$, but never both.

Proof: [[2], corollary 7.1d].

Remark: Note that the inequality $g^t x < 0$ implies $x \neq 0$, while $A^t y = g$ implies $y \neq 0$. Hence if both (1) and (2) holds then

$$g^t x = (A^t y)^t x = y^t A x \geq 0$$

which contradicts $g^t x < 0$. Hence it is not possible that both systems are solvable.

The next theorem tells us that which of the two systems has a solution.

Theorem 1.2: Let y^* solve the bounded least squares (BLS) problem

$$\text{minimize} \|A^t y - g\|_2^2$$

$$\text{Subject to } y \geq 0 \qquad (3)$$

and let

$$r = A^t y^* - g$$

denote the corresponding residual vector. If $r = 0$ then y^* solves (2). Otherwise r solves (1) and $g^t r = -r^t r$.

Proof: Consider the function

$$\emptyset(x) = \|x - g\|_2^2$$

which is continuous and strictly convex, as so the euclidean norm.

The set

$$S = \{A^t y \mid y \geq 0 \text{ and } \|A^t y - g\|_2 \leq \|g\|_2\}$$

IJournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 11, November-2014

is non-empty ($0 \in S$), closed and bounded.

Let x be a limit point of S. Let $\{x_k\}$ be a sequence in S which converge to x. As $x_k \in S$, there exists $y_k, y_k \geq 0$ such that $x_k = A^t y_k$ and

$$\|x_k - g\|_2 \leq \|g\|_2$$

$$\Rightarrow \|x - g\|_2 \leq \|g\|_2$$ [as RHS is independent of k]

$$\Rightarrow x \in S$$

i.e., S is closed. Obviously, S is bounded. Therefore S is compact.

Therefore, by weirestrass theorem, $\emptyset(x)$ attains its minimum over the set S, and this minimizer is unique. If there were two minimizers x_1 and x_2 in S then, as S is convex, $\frac{x_1+x_2}{2} \in S$. Next, as $\emptyset$ is strictly convex

$$\left\|\frac{x_1+x_2}{2} - g\right\|_2 < \frac{1}{2}(\|x_1 - g\|_2 + \|x_2 - g\|_2) \leq \|g\|_2$$

which contradicts that x_1 (or x_2 is a minimizer. Hence minimizer is unique.

As A must be non-zero there is $y_0 \neq 0$ such that $A^t y_0 = 0$. Then

$$r = A^t(y^* + y_0) - g = A^t y^* - g.$$

Thus r is unique but not necessarily y^* (a minimizer of (3)). It is clear that, if $r = 0$ then

$$A^t y^* = g \text{ and } y^* \geq 0$$

i.e., y^* solves (2).

For the second part, consider the problem

$$\text{minimize} F_i(\theta) = \|A^t(y^* + \theta e_i) - g\|_2^2$$

$$\text{subject to } y_i^* + \theta \geq 0 \text{ for } i = 1, \dots \dots, m$$

(4)

where θ is a real variable, e_i denotes the i^{th} column of the $m \times m$ unit matrix.

Now,

$$F_i(\theta) = \|A^t y^* + \theta A^t e_i - g\|_2^2 = \|\theta a_i + r\|_2^2$$

where a_i^t denotes the i^{th} row of A. As y is minimizer of (3), we conclude that

$$F_i(\theta) = \|A^t(y^* + \theta e_i) - g\|_2^2$$

$$\geq \|A^t y \;\; - g\|_2^2 = F_i(0)$$

$\forall \theta$ reals.

Hence $\theta = 0$ is minimizer of the problem (4). This means it is not possible to obtain a better solution by changing the value of θ.

Now,

$$F_i(\theta) = \|r\|_2^2 + \theta^2\|a_i\|_2^2 + 2\theta a_i^t r$$

gives

$$F_i'(\theta) = 2\theta\|a_i\|_2^2 + 2a_i^t r .$$

Further,

$$F_i(\theta) \geq F_i(0) \Leftrightarrow \theta(\theta\|a_i\|_2^2 + 2a_i^t r) \geq 0$$

If $y_i^* > 0, \theta \in [-y_i^*, \infty)$. So there is an open interval centered at 0 and contained in $(-y_i^*, \infty)$. Hence, $F_i^{'}(0) = 2a_i^t r = 0$ [[2], corollary 7.1d].

Now, if $y_i^* = 0, \theta \in [0, \infty]$. If $a_i^t r < 0$ then for $0 < \theta < \frac{-2a_i^t r}{\|a_i\|_2^2}$, $F_i(\theta) < F_i(0)$. This is not true. Therefore, $a_i^t r \geq 0$. Thus we have: If $y_i^* > 0$ then $a_i^t r = 0$ and if $y_i^* = 0$ then $a_i^t r \geq 0$. Therefore,

$$Ar \geq 0 \text{ and } {}^*r^t A^t y^* = 0.$$

Hence $g = A^t y^* - r$ implies

$$r^t g = r^t(A^t y^* - r) = -r^t r < 0.$$

Hence r solves (1). ■

Corollary 1.3: If $r \neq 0$ then $\frac{r}{\|r\|_2}$ solves the steepest descent problem

minimize $g^t x$

subject $Ax \geq 0$ and $x^t x = 1$ (5)

Proof: We have $g = A^t y^* - r$. Now,

$$g^t x = (A^t y^* - r)^t x = y^{*t} Ax - r^t x = S(x) - r^t x$$

where $S(x) = y^{*t} Ax$.

Let x satisfy the constraints of (5). Then $S(x) \geq 0$, as $y^* \geq 0, Ax \geq 0$.

Further,

$$|r^t x| \leq \|r\|_2 . \|x\|_2 = \|r\|_2$$

$$\Rightarrow g^t x = S(x) - r^t x \geq -r^t x \geq -\|r\|_2.$$

By theorem 1.2, $y^{*t} Ar = 0$. This gives

$$g^t r = (A^t y^* - r)^t r = -\|r\|_2^2$$

$$\Rightarrow \frac{g^t r}{\|r\|_2} = -\|r\|_2.$$

i.e., $\frac{r}{\|r\|_2}$ minimizes the problem (5). ∎

Note 1.4: Theorem and its corollary constitute a simple relationship between the primal system (1), the dual system (2), the BLS problem (3) and the steepest descent problem (5).

Remark 1.5: The role of theorem 1.2 as a tool for deriving constructive optimality conditions is illustrated as below.(Remark 1.7)

Let $F(x)$ be a convex differential function. Consider the problem

minimize $F(x)$

subject to $a_i^t x \geq b_i$, for $i \in C$, (6)

where C is a finite index set, a_i are given vectors in R^n, b_i are given real numbers and

$x \in R^n$ denotes the vector of unknown. Given a feasible point $x^* \in R^n$. We define $C^* = \{i \,|\, a_i^t x^* = b_i\}$

to be the set which contains the indices of active constraints at this point.

The number of active constraints is denoted by m and active constraint matrix A, is an $m \times n$ matrix whose rows are a_i^t, $i \in C^*$. Since the order of active constraints does not matter, we assume $C^* = \{1, \dots \dots, m\}$. So,

$$A^t = [a_1, \dots \dots, a_m].$$

Definition 1.6: A vector $u \in$ is said to be a feasible descent direction at x^* if and only if $Au \geq 0$ and there exists a positive constant δ, such that

$$F(x^* + \theta u) < F(x^*) \quad \forall\, 0 < \theta < \delta.$$

Remark 1.7: If no such u exists then,

$$F(x^*) \leq F(x^* + \theta u) \quad \forall\, 0 < \theta < \delta$$

implies that x^* minimizes F i.e., x^* solves (6).

Let g be the gradient vector of $F(x)$ at the point x^*. Then the convexity assumption implies

$$F(x^* + \theta u) \geq F(x^*) + \theta g^t u \quad \forall\, u \in R^n$$

This is clear by [[3], theorem 3.2.5,theorem 3.3.3, lemma 3.3.2]

As F is differentiable then by [[3], lemma 3.3.2],the set of subgradients of F at x^* is the singleton set $\{\nabla F(x^*) = g\}$. Thus by [[3], theorem 4.1.2], $u \in$ is a feasible descent direction if and only if $Au \geq 0$ and $g^t u < 0$ (i.e., u solves (1)). Thus, from Farkas' lemma we can deduce that a feasible point x^*

solves (6) if and only if there exists $y \in$ that solves (2).

In the same way, we formulate the equivalent BLS problems for Gale's theorem, Gordan's theorem, Motzkin's theorem and Stiemke's theorem.

ACKNOWLEDGMENTS

With the blessing of Lord Krishna and the kind support of all professors of School of Mathematics, D. A. V. V., Indore, we have completed this paper. We are also thankful to our college authorities, friends and every one helped and encouraged us.

REFERENCES

[1] The relationship between theorems of the alternative, least norm problems, steepest descent directions, and degeneracy: A review, Achiya Dax, Annals of Operations Research V.46, 11-60, 1993.

[2] Theory of Linear and Intege Programming, A. Schrijver, John Wiley, © 1986.

[3] Non linear Programming : Theory and Algorithms, M. S. Bazaraa, Hanif D. Sherali, C. M. Shetty, John Wiley & Sons.