UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

A PRIMAL-DUAL METHOD FOR THE DISCRETE L_1 LINEAR APROXIMATION PROBLEM

ARENALES,M. N. & TAVARES, H.M.F.

Nº 22

NOTAS

São Carlos - SP

Março 1986

A PRIMAL-DUAL METHOD FOR THE DISCRETE L_1 LINEAR APPROXIMATION PROBLEM

M. N. ARENALES+ and H. M. F. TAVARES‡

ABSTRACT. A new primal-dual method for the discrete L_1 linear approximation problem is proposed. It explores the fact that the L_1-approximation problem easily presents a basic primal feasible solution and a dual feasible solution. The method appears to be computacionally superior to the simplex method of Barrodale-Roberts.

1. INTRODUCTION. The problem of fitting the linear model

$$y = Ax + r ,$$

where y is an m-vector of observations, A is an $m \times n$ matrix of fixed known values, with rank(A) = n , x is the n-vector of parameters, and r is an m-vector of random errors with $m > n$ using the L_1 criterion, consists in determing x such that,

$$\text{(P)} \qquad \min \|r\|_1 = \sum_{i=1}^{m} |r_i|$$

$$\text{s.t.} \quad y = Ax + r.$$

We refer this problem as primal problem. The dual problem of (P) is:

$$\text{(D)} \qquad \max \; y^t\lambda$$

$$\text{s.t.} \quad A^t\lambda = 0$$

$$-\P \leq \lambda \leq \P$$

where λ is a dual variable m-vector and $\P$ an m-vector of one's.

+ Departamento de Computação e Estatística, Instituto de Ciências Matemáticas de São Carlos, Universidade de São Paulo, São Carlos-S.P., Brasil.

‡ Departamento de Engenharia Elétrica, Fac. de Eng. de Campinas, Universidade Estadual de Campinas, Campinas-S.P., Brasil.

Barrodale and Roberts (1973), Abdelmalek (1975), Robers and Ben-Israel (1969) utilized algorithms of linear programming for solving the linear L_1 approximation problem.

It's interesting to note that the feasible solutions for (P) and (D) problems are easily obtained without special procedure. However, to our knowledge, none algorithm exploits these richness of informations.

In this work we apply a primal-dual method of linear programming |2|, that iterates with a basic primal feasible solution and a dual feasible solution. The gap between the primal and dual objectives will be decreased at each iteration.

2. PRELIMINARIES

Definition: Let B be any nonsingular $n \times n$ submatrix made up of rows of A, and N the submatrix of remainder rows.

The solution $(\hat{x}, \hat{r})$ given by

$$\hat{x} = B^{-1} y_B \ , \ \hat{r}_B = 0 \ , \ \hat{r}_N = y_N - \hat{N} y_B \tag{1}$$

is a basic primal feasible solution, and the solution $\hat{\lambda}$ given by

$$\hat{\lambda}_N = \text{sgn}\ (\hat{r}_N) \ , \ \hat{\lambda}_B = - \hat{N}^t \hat{\lambda}_N \tag{2}$$

is a basic dual solution, where y_B is an m-vector formed by y-components which indexes correspond the rows of B, $\hat{N} = NB^{-1}$ and

$$\text{sgn}(\hat{r}_i) = \begin{cases} 1 & \text{if } \hat{r}_i > 0 \\ -1 & \text{if } \hat{r}_i < 0 \\ 1 \text{ or } -1 & \text{if } \hat{r}_i = 0 \end{cases}$$

Note 1 : The primal solution (1) is feasible for any basis B. However, the dual solution (2) could be infeasible: althought $A^t \hat{\lambda} = 0$, the upper and lower constraints could be violated.

Note 2 : The primal and dual objective values are equal:

$$\|\hat{r}\|_1 = y^t \hat{\lambda} .$$

If the dual solution is feasible, then $(\hat{x}, \hat{r})$ solves (P) and $\hat{\lambda}$ solves (D), since $\|r\|_1 \geq y^t \lambda$ for all (x,r) and λ feasible solutions for (P) and (D) respectively.

3. THE PRIMAL-DUAL METHOD

3.1 *Initiation*

Let

i) B,N be a partition on the rows of A, and $(\hat{x},\hat{r})$, $\hat{\lambda}$ the associated basic solutions;

ii) $\tilde{\lambda}$ be a dual feasible solution. For example: $\tilde{\lambda} = 0$.

3.2 *First Optimality Test*

If

$$\tilde{\lambda}_N = \mathrm{sgn}(\hat{r}_N) \tag{3}$$

then $(\hat{x},\hat{r})$, $\hat{\lambda}$ solve the primal and dual problems respectively.

In case of equality (3) is not verified, the solutions will be modified as bellow.

3.3 *Solutions Alterations*

We divide the modifications in two steps.

3.3.1 *Dual Step*

Let λ be a convex combination between $\hat{\lambda}$ and $\tilde{\lambda}$:

$$\lambda = \tilde{\lambda} + \varepsilon(\hat{\lambda} - \tilde{\lambda}) \ , \quad 0 \leq \varepsilon \leq 1 \tag{4}$$

This solution provides a higher value for the dual objective value, since

$$y^t \lambda = y^t \tilde{\lambda} + \varepsilon \, y^t(\hat{\lambda} - \tilde{\lambda}) \geq y^t \tilde{\lambda} \ .$$

The inequality above is resulting from

$$y^t \hat{\lambda} = \|\hat{r}\| > y^t \tilde{\lambda} \ .$$

Moreover, the λ solution satisfy:

$$A^t \lambda = 0 \ .$$

We choose ε as large as possible without it violates the upper or lower bounds on λ.

From (4) it folows:

$$\varepsilon \leq \frac{-1 - \tilde{\lambda}_i}{\hat{\lambda}_i - \tilde{\lambda}_i} \qquad \text{if} \qquad \hat{\lambda}_i - \tilde{\lambda}_i < 0 \ ,$$

$$\varepsilon \leq \frac{1 - \tilde{\lambda}_i}{\hat{\lambda}_i - \tilde{\lambda}_i} \qquad \text{if} \qquad \hat{\lambda}_i - \tilde{\lambda}_i > 0 \, .$$

Let IB and IN be the indexes sets associated with B and N respectively.

For the indexes in IN , the rations above are ever equal one (see(2)). For the indexes in IB:

$$\varepsilon \leq \varepsilon^- = \min_{i \in IB} \left\{ \frac{-1 - \tilde{\lambda}_i}{\hat{\lambda}_i - \tilde{\lambda}_i} \qquad \text{if} \qquad \hat{\lambda}_i - \tilde{\lambda}_i < 0 \right\}$$

$$\varepsilon \leq \varepsilon^+ = \min_{i \in IB} \left\{ \frac{1 - \tilde{\lambda}_i}{\hat{\lambda}_i - \tilde{\lambda}_i} \qquad \text{if} \qquad \hat{\lambda}_i - \tilde{\lambda}_i > 0 \right\}.$$

We choose ε by:

$$\hat{\varepsilon} = \min \{ 1 \, , \varepsilon^- \, , \varepsilon^+ \} \, .$$

Second Optimaly Test

If $\hat{\varepsilon} = 1$ then $\hat{\lambda}$ is a dual feasible solution. Therefore, by note 2, $(\hat{x},\hat{r})$ and $\hat{\lambda}$ are the solutions of the primal and dual problems respectively.

However if $\hat{\varepsilon} < 1$, we have found a best dual feasible solution:

$$\lambda = \tilde{\lambda} + \hat{\varepsilon}(\hat{\lambda} - \tilde{\lambda}) \, .$$

Now, the primal solution will be modified.

3.3.2 *Primal Step*

One of the two situations bellow must occur:

i) $\hat{\varepsilon} = \varepsilon^+$

or

ii) $\hat{\varepsilon} = \varepsilon^-$

We shall analyse only the first situation, the second is similar. Let k be the index that determines ε^+ :

$$\hat{\varepsilon} = \varepsilon^+ = \frac{1 - \tilde{\lambda}_k}{\hat{\lambda}_k - \tilde{\lambda}_k} \, .$$

The simplex strategy will be utilized:

$$r_k = \delta \geq 0$$

$$r_i = 0 \ , \ \forall \ i \in IB \ , \ i \neq k \ .$$

To show that the primal objective value decreases, it is sufficient to prove that the relative cost for r_k variable is negative, i.e.,

$$\Delta_o = 1 - \hat{\lambda}_k < 0 \ . \tag{5}$$

For this, we note that

$$\tilde{\lambda}_k + \hat{\varepsilon}(\hat{\lambda}_k - \tilde{\lambda}_k) = 1 \ ,$$

with $\hat{\varepsilon} < 1$ and $\hat{\lambda}_k - \tilde{\lambda}_k > 0$. Then

$$\tilde{\lambda}_k + 1.(\hat{\lambda}_k - \tilde{\lambda}_k) > \tilde{\lambda}_k + \hat{\varepsilon}.(\hat{\lambda}_k - \tilde{\lambda}_k) = 1 \ .$$

We choose δ that minimizes the primal objective value.

The basic residues are given by:

$$r_i = \hat{r}_i + \hat{N}_{ik}.\delta \qquad \forall \ i \in IN \ , \tag{6}$$

and they could change their signs as δ increases.

The values that provides a sign change to the basic residues are:

$$\delta_i = - \frac{\hat{r}_i}{\hat{N}_{ij}} \geq 0 \ , \qquad \hat{N}_{ij} \neq 0 \ . \tag{7}$$

In the denegerate case, i.e., $\hat{r}_i = 0$ for some $i \in IN$, the ratio $\delta_i = 0$ will be considerate only if $\hat{N}_{ik} < 0$ and $sgn(\hat{r}_i) = 1$, or $\hat{N}_{ik} > 0$ and $sgn(\hat{r}_i) = -1$.

We suppose there are r ratios defined by (7), and we consider an indexes permutation $i_1, i_2, \ldots, i_r$ in a way that

$$0 \leq \delta_{i_1} \leq \ldots \leq \delta_{i_r} \ .$$

In case of $0 \leq \delta \leq \delta_{i_1}$ it follows from (6) that

$$sgn(r) = sgn(\hat{r}) + \begin{pmatrix} e^k \\ o \end{pmatrix} .$$

where e^k is the k^{th} column of the n-identity matrix. So,

$$\|r\|_1 = sgn(r)^t . r = sgn(\hat{r})^t . \hat{r} + \delta(1+sgn(\hat{r}_N)^t \hat{N}^k)$$

$$= \|\hat{r}\|_1 + \delta\Delta_o .$$

where $\Delta_o < 0$ from (5), and $\hat{N}^k$ is the k^{th} row of NB^{-1}.

This result can be extended for $\delta_{i_j} \leq \delta \leq \delta_{i_{j+1}}$.

Theorem

Let

$$r^{(j)} = r^{(j-1)} + (\delta_{i_j} - \delta_{i_{j-1}})\binom{e^k}{\hat{N}^k} \qquad j = 1,2,\ldots,r$$

$$\Delta_j = \Delta_{j-1} + 2|\hat{N}_{i_j k}| \qquad (8)$$

where $\delta_{i_o} = 0$, $r^{(o)} = \hat{r}$, $\Delta_o = 1 - \hat{\lambda}_k$.

Then, for $\delta = \delta_{i_j} + \theta$, with $0 \leq \theta \leq \delta_{i_{j+1}} - \delta_{i_j}$ we have

$$\|r\|_1 = \|r^{(j)}\|_1 + \theta\Delta_j \qquad j = 0,1,\ldots,r \qquad (9)$$

where $r = r^{(j)} + \theta\binom{e^k}{\hat{N}^k}$.

We omite the proof but it can done observing that:

- $sgn(r) = sgn(r^{(j)}) + \begin{pmatrix} 0 \\ -2\, sgn(\hat{r}_{i_j}).e^{i_j} \end{pmatrix}$
- i_j^{th} component of $r_N^{(j)}$ is null due the δ_{i_j} definition.

So, if $\Delta_j < 0$ then $\|r\|_1 \leq \|r^{(j)}\|_1$ and the inequality will be strict if $\delta_{i_{j+1}} > \delta_{i_j}$.

Then, we choose s such that

$$\Delta_{s-1} < 0 \quad \text{and} \quad \Delta_s \geq 0 .$$

Update the basic and non-basic indexes: the k^{th} basic index is changed by the i_s^{th} non-basic index; and return to dual step with the new solutions.

4. COMPUTATIONAL EXPERIMENTS. We realize some experiments to compare our primal-dual method with the primal simplex method. Examples were randonly generated using the Fortran RAN function of DEC-system, and runned on a PDP 11/45 computer.

The following table (where (P) means primal and (PD) means primal-dual) furnishes the average time of CPU and the average number of iterations for 5 examples.

Table 1

Average CPU time (seconds) / average iter. number

		n = 2	n = 5	n = 10
m = 100	Primal	3.69/6.	13.97/20.	38.71/45.
	P-dual	3.13/5.	11.05/15.	34.75/37.
m = 200	Primal	21.20/8.6	54.97/26.2	121.34/49.4
	P-dual	16.11/7.6	51.54/23.6	113.18/43.2
m = 400	Primal	68.28/9.2	218.36/31.2	501.87/64.6
	P-dual	38.10/5.8	167.87/24.2	382.28/48.6

Our programs are not as efficient as the one of Barrodale and Roberts [4]. But in any case they are sufficient to compare the two methods.

It's intersting to note that the programs difer only in the determination of which variables must enter in the basis. The primal simplex method determines the most negative relative cost while the primal-dual method determines, in the dual step, the index that block the dual perturbation.

5. CONCLUSION AND FINAL REMARKS. We present in section 3 a new algorithm for the discrete L_1 linear approximation problem specializing a primal-dual method for linear programming [2] as Barrodale and Roberts [3] did with the simplex method.

In the degeneracy absence (at least $\delta_{i_1} > 0$) the method will converge in a finite number of iterations because the slack between the primal and dual objective values will decrease at each iteration and so, the optimal solution will be attained in a finite number of steps since there exists a finite number of basis.

The numerical results in section 4 encourage us to develop an efficient primal-dual code for L_1-approximation.

REFERENCES

[1] N.N. ABDELMALEK, *An efficient method for the discrete linear L_1 approximation problem*, Math. of computation, V.29, n.131, (1975), pp. 844-850.

[2] N.N. ARENALES AND H.M.F. TAVARES, *A new primal-dual method for linear programming*, to appear.

[3] I. BARRODALE AND F.D.K. ROBERTS, *An improved algorithm for discrete L_1 approximation*, SIAM J. Numer. Anal., V.10 (1973) pp. 839-848.

[4] I. BARRODALE AND F.D.K. ROBERTS, *Solution of an overdetermined system of equations in the ℓ_1 norm*, Comm. ACM, Algorithm 478, 17, pp. 319-320.

[5] P.D. ROBERS AND A. BEN-ISRAEL, *An interval programming algorithm for discrete linear L_1 approximation problems*, J. Approximation Theory, 2 (1969), pp. 323-336.