A Steepest Feasible Direction Extension of the Simplex Method

Biressaw C. Wolde and Torbjörn Larsson

Abstract We present a feasible direction approach to general linear programming, which can be embedded in the simplex method although it works with non-edge feasible directions. The feasible direction used is the steepest in the space of all variables, or an approximation thereof. Given a basic feasible solution, the problem of finding a (near-)steepest feasible direction is stated as a strictly convex quadratic program in the space of the non-basic variables and with only non-negativity restrictions. The direction found is converted into an auxiliary non-basic column, known as an external column. Our feasible direction approach allows several computational strategies. First, one may choose how frequently external columns are created. Secondly, one may choose how accurately the direction-finding quadratic problem is solved. Thirdly, near-steepest directions can be obtained from low-dimensional restrictions of the direction-finding quadratic program or by the use of approximate algorithms for this program.

Keywords Linear program · Steepest-edge · Feasible direction · External pivoting

1 Derivation

Let $A \in \mathbb{R}^{m \times n}$, with $n > m$, have full rank, and let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Consider the Linear Program (LP)

$$z^\star = \min z = c^\mathrm{T} x$$

$$\text{s.t. } Ax = b$$

$$x \geq 0,$$

B. C. Wolde · T. Larsson (✉)
Department of Mathematics, Linköping University, Linköping, Sweden
e-mail: biressaw.wolde@liu.se; torbjorn.larsson@liu.se,tolar@mai.liu.se

whose feasible set is assumed to be non-empty. Let a non-optimal non-degenerate basic feasible solution be at hand. With a proper variable ordering, the solution corresponds to the partitioning $A = (B, N)$ and $c = (c_B^T, c_N^T)^T$, where $B \in \mathbb{R}^{m \times m}$ is the non-singular basis matrix and $N \in \mathbb{R}^{m \times (n-m)}$ is the matrix of non-basic columns. Let $x = (x_B^T, x_N^T)^T$. Introducing the reduced cost vector $\bar{c}_N^T = c_N^T - u^T N$, where $u^T = c_B^T B^{-1}$ is the complementary dual solution, and letting I_m be the identity matrix of size m, problem LP is equivalent to

$$z^\star = \min z = c_B^T B^{-1} b + \bar{c}_N^T x_N$$

$$\text{s.t. } I_m x_B + B^{-1} N x_N = B^{-1} b$$

$$x_B, \ x_N \geq 0.$$

The basic solution is $x_B = B^{-1}b$ and $x_N = 0$. By assumption, $B^{-1}b > 0$ and $\bar{c}_N \ngeq 0$ hold.

Let $\mathcal{N} = \{m+1, \ldots, n\}$, that is, the index set of the non-basic variables, and let a_j be column $j \in \mathcal{N}$ of the matrix A. Geometrically, the given basic feasible solution corresponds to an extreme point of the feasible polyhedron of problem LP, and a variable x_j, $j \in \mathcal{N}$, that enters the basis corresponds to the movement along a feasible direction that follows an edge from the extreme point. The edge direction is given by, see e.g. [4],

$$\eta_j = \begin{pmatrix} -B^{-1} a_j \\ e_{j-m} \end{pmatrix} \in \mathbb{R}^n,$$

where $e_{j-m} \in \mathbb{R}^{n-m}$ is a unit vector with a one entry in position $j - m$. The directional derivative of the objective function along an edge direction of unit length is $c^T \eta_j / \|\eta_j\| = (c_B^T, c_N^T)\eta_j / \|\eta_j\| = (-c_B^T B^{-1} a_j + c_j)/\|\eta_j\| = \bar{c}_j / \|\eta_j\|$ (where $\|\cdot\|$ is the Euclidean norm). This is the rationale for the *steepest-edge criterion* [3], which in the simplex method finds a variable x_r, $r \in \mathcal{N}$, to enter the basis such that $\bar{c}_r / \|\eta_r\| = \min_{j \in \mathcal{N}} \{\bar{c}_j / \|\eta_j\|\}$.

We consider feasible directions that are constructed from non-negative linear combinations of the edge directions. To this extent, let $w \in \mathbb{R}^{n-m}_+$ and consider the direction

$$\eta(w) = \sum_{j \in \mathcal{N}} w_j \eta_j = \begin{pmatrix} -B^{-1} N \\ I_{n-m} \end{pmatrix} w,$$

where I_{n-m} is the identity matrix of size $n - m$. Note that any feasible solution to LP is reachable from the given basic feasible solution along some direction $\eta(w)$, and that $c^T \eta(w) = \bar{c}_N^T w$. Our development is founded on the problem and theorem below.

Define the Steepest-Edge Problem (SEP)

$$\min \bar{c}_N^{\mathrm{T}} w$$
$$\text{s.t. } \|\eta(w)\|^2 \leq 1$$
$$\text{supp}(w) = 1$$
$$w \geq 0,$$

where supp$(\cdot)$ is the support of a vector, that is, its number of nonzero components.

Theorem 1 *An index $r \in \mathcal{N}$ fulfils the steepest-edge criterion if and only if the solution*

$$w_j = \begin{cases} 1/\|\eta_j\| \ if \ j = r \\ 0 \ otherwise \end{cases} , \quad j \in \mathcal{N},$$

solves SEP.

The theorem follows directly from an enumeration of all feasible solutions to problem SEP. Note that the optimal value of SEP is the steepest-edge slope $\bar{c}_r/\|\eta_r\|$.

To find a feasible direction that is steeper than the steepest edge, the support constraint in problem SEP is relaxed. The relaxed problem, called Direction-Finding Problem (DFP), can be stated as

$$\min \bar{c}_N^{\mathrm{T}} w$$
$$\text{s.t. } w^{\mathrm{T}} Q w \leq 1 \tag{1}$$
$$w \geq 0,$$

where the matrix $Q = N^{\mathrm{T}}(B^{-1})^{\mathrm{T}} B^{-1} N + I_{n-m} \in \mathbb{R}^{(n-m)\times(n-m)}$, and it gives the *steepest* feasible direction. Because Q is symmetric and positive definite, and $\bar{c}_N \not\geq 0$ holds, the problem DFP has a unique, nonzero optimal solution, which fulfils the normalization constraint (1) with equality and has a negative objective value. Further, the optimal solution will in general yield a feasible direction that is a non-trivial non-negative linear combination of the edge directions, and has a directional derivative that is strictly better than that of the steepest edge.

As an example, we study the linear program

$$\min \left\{ -x_1 - 2x_2 \mid 5x_1 - 2x_2 \leq 10; \ -2x_1 + 4x_2 \leq 8; \ 2x_1 + x_2 \leq 6; \ x_1, x_2 \geq 0 \right\},$$

which is illustrated to the left in Fig. 1. For the extreme point at the origin, which has the slack basis with $B = I_3$, we have $\eta_1 = (-5, 2, -2, 1, 0)^{\mathrm{T}}$ and $\eta_2 = (2, -4, -1, 0, 1)^{\mathrm{T}}$, with $\bar{c}_1/\|\eta_1\| = -1/\sqrt{34}$ and $\bar{c}_2/\|\eta_2\| = -2/\sqrt{22}$. If using the steepest-edge criterion, the variable x_2 would therefore enter the basis. The feasible set of DFP is shown to the right in Fig. 1. The optimal
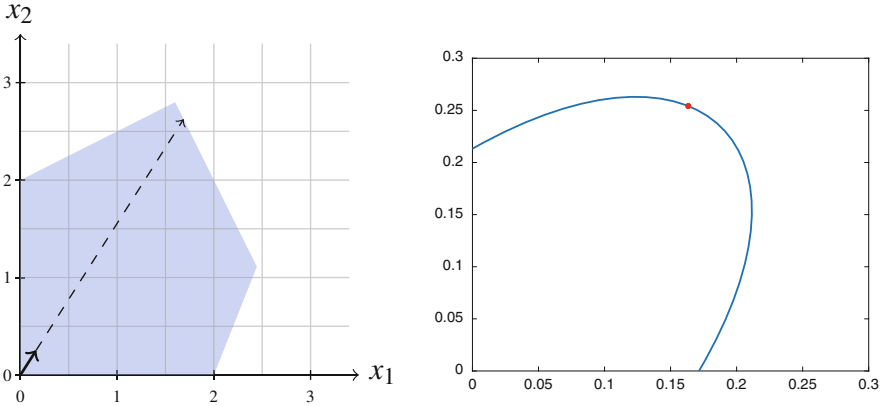
Fig. 1 Illustration of steepest feasible direction and problem DFP

solution of DFP is $w^* \approx (0.163, 0.254)^\mathrm{T}$ with $\bar{c}_N^\mathrm{T} w^* = c^\mathrm{T}\eta(w^*) \approx -0.672$, which should be compared with $-2/\sqrt{22} \approx -0.426$. The feasible direction is $\eta(w^*) \approx (-0.309, -0.690, -0.581, 0.163, 0.254)^\mathrm{T}$. The maximal feasible step in this direction yields the boundary point $(x_1, x_2) \approx (1.687, 2.625)$, whose objective value is better than those of the extreme points that are adjacent to the origin. (Further, this boundary point is close to the extreme point $(1.6, 2.8)^\mathrm{T}$, which is optimal.)

We next establish that problem DFP can be solved by means of a relaxed problem. Let $\mu/2 > 0$ be an arbitrary value of a Lagrangian multiplier for the constraint (1), and consider the Lagrangian Relaxation (LR)

$$\min_{w \geq 0} r(w) = \bar{c}_N^\mathrm{T} w + \frac{\mu}{2} w^\mathrm{T} Q w$$

of problem DFP (ignoring the constant $-\mu/2$). Since r is a strictly convex function, problem LR has a unique optimum, denoted $w^*(\mu)$. The following result can be shown.

Theorem 2 *If $w^*(\mu)$ is optimal in problem LR, then $w^* = w^*(\mu)/\|\eta(w^*(\mu))\|$ is the optimal solution to problem DFP.*

The proof is straightforward; both problems are convex and have interior points, and it can be verified that if $w^*(\mu)$ satisfies the Karush–Kuhn–Tucker conditions for problem LR then $w^* = w^*(\mu)/\|\eta(w^*(\mu))\|$ satisfies these conditions for problem DFP.

Hence, the steepest feasible direction can be found by solving the simply constrained quadratic program LR, for any choice of $\mu > 0$. The following result, which is easily verified, gives an interesting characterization of the gradient of the

objective function r. It should be useful if LR is approached by an iterative descent method, such as for example a projected Newton method [1].

Proposition 1 *Let $\eta_B(w) = -B^{-1}Nw$ and $\Delta u^{\mathrm{T}} = \mu\eta(w)^{\mathrm{T}}B^{-1}$. Then*

$$\nabla r(w) = c_N - N^{\mathrm{T}}(u + \Delta u) + \mu w.$$

Note that the expression for Δu is similar to that of the dual solution, and that the pricing mechanism of the simplex method is used to compute $\nabla r(w)$, but with a modified dual solution. Further, $-\eta_B(w) = B^{-1}Nw = B^{-1}\sum_{j\in\mathcal{N}} w_j a_j$, that is, a non-negative linear combination of the original columns $(a_j)_{j\in\mathcal{N}}$ expressed in the current basis.

In order to use a feasible direction within the simplex method, it is converted into an *external column* [2], which is a non-negative linear combination of the original columns in LP. Letting w^* be an (approximate) optimal solution to problem DFP, we define the external column as $c_{n+1} = c_N^{\mathrm{T}}w^*$ and $a_{n+1} = Nw^*$. Letting $\bar{c}_{n+1} = c_{n+1} - u^{\mathrm{T}}a_{n+1}$, the problem LP is augmented with the variable x_{n+1}, giving

$$z^{\star} = \min z = c_B^{\mathrm{T}}B^{-1}b + \bar{c}_N^{\mathrm{T}}x_N + \bar{c}_{n+1}x_{n+1}$$

$$\text{s.t. } I_m x_B + B^{-1}Nx_N + B^{-1}a_{n+1}x_{n+1} = B^{-1}b$$

$$x_B, \ x_N, \ x_{n+1} \geq 0,$$

where $\bar{c}_{n+1} < 0$. By letting the external column enter the basis, the feasible direction will be followed. Note that the augmented problem has the same optimal value as the original one (If $B^{-1}a_{n+1} \leq 0$ holds, then $z^{\star} = -\infty$.) Further, if the external column is part of an optimal solution to the augmented problem, then it is easy to recover an optimal solution to the original problem [2].

The approach presented above is related to those in [5] and [2], which both use auxiliary primal variables for following a feasible direction. (The term external column is adopted from the latter reference.) These two works do however use *ad hoc* rules for constructing the feasible direction, for example based on only reduced costs, instead of solving a direction-finding problem with the purpose of finding a steep direction.

2 Numerical Illustration and Conclusions

It is in practice reasonable to use a version of LR that contains only a restricted number of edge directions. Letting $J \subseteq \mathcal{N}$, $w_J = (w_j)_{j\in J}$, $\bar{c}_J = (\bar{c}_j)_{j\in J}$, $N_J = (a_j)_{j\in J}$ and $Q_J = N_J^{\mathrm{T}}B^{-\mathrm{T}}B^{-1}N_J + I_{|J|}$, the restricted LR is given by $\min_{w_J \geq 0} r_J(w_J) = \bar{c}_J^{\mathrm{T}}w_J + \frac{\mu}{2}w_J^{\mathrm{T}}Q_J w_J$.

An external column should be advantageous to use compared to an ordinary non-basic column, but it is computationally more expensive. Further, pivots on external columns will lead to points that are *not* extreme points in the original polyhedron. It is therefore reasonable to combine pivots on external columns with ordinary pivots. Considering the choice of $|J|$, a high value makes the restricted LR computationally more expensive, but it then also has the potential to yield a better feasible direction. Hence, there is a trade-off between the computational burden of creating the external columns, both with respect to frequency of external columns and the choice of $|J|$, and the reduction in simplex iterations that they may lead to.

To make an initial assessment of the potential usefulness of generating external columns as described above, we made a simple implementation in MATLAB of the revised simplex method, using the Dantzig entering variable criterion but with the option to replace ordinary pivots with pivots on external column, which are found by solving restricted versions of LR. Letting k be the number of edge directions to be included in the restriction, we consider two ways of making the selection: by finding the k most negative values among $\{\bar{c}_j\}_{j \in \mathcal{N}}$ or among $\{\bar{c}_j/\|\eta_j\|\}_{j \in \mathcal{N}}$. These ways are called *Dantzig selection* and *steepest-edge selection*, respectively. The latter is computationally more expensive. The restricted problem LR is solved using the built-in solver `quadprog`.

We used test problem instances that are randomly generated according to the principle in [2], which allows the simplex method to start with the slack basis. The external columns are constructed from original columns only (although it is in principle possible to include also earlier external columns). We study the number of simplex iterations and running times required for reaching optimality when using the standard simplex method, and when using the simplex method with external columns with different numbers of edge directions used to generate the external columns and when generating external column only once or repeatedly. Table 1 shows the results. Figure 2 shows the convergence histories for the smallest problem instance when using $k = 200$.

Our results indicate that the use of approximate steepest feasible directions can considerably reduce both the number of simplex iterations and the total running times, if the directions are based on many edge directions and created seldomly; if the directions are based on few edge directions and created frequently, then the overall performance can instead worsen. These findings clearly demand for more investigations. Further, our feasible direction approach must be extended to properly handle degeneracy, and tailored algorithms for fast solution of the restricted problem LR should be developed.

Table 1 Simplex iterations and running times

Size		Std simplex		External columns: Dantzig selection				External columns: steepest-edge selection			
m	n	Iterations	Time	k	nmax	Iterations	Time	k	nmax	Iteration	Time
1000	2000	35,181	418.0	20	∞	28,198	310.6	20	∞	28,091	302.4
				20	50	43,999	720.4	20	50	27,261	552.8
				100	∞	20,200	211.3	100	∞	23,455	245.0
				100	500	21,237	231.5	100	500	20,133	225.9
				200	∞	20,561	217.7	200	∞	20,891	219.7
				200	500	18,926	210.2	200	500	19,168	220.9
1000	3000	47,647	715.2	20	∞	38,545	571.0	20	∞	36,324	511.2
				20	50	41,217	754.4	20	50	34,367	998.1
				100	∞	30,676	440.5	100	∞	22,886	305.8
				100	500	29,212	424.6	100	500	22,799	315.6
				200	∞	24,504	365.0	200	∞	26,150	342.0
				200	500	24,405	344.0	200	500	21,372	303.1
1000	4000	57,572	1096.1	20	∞	50,150	842.1	20	∞	41,857	713.9
				20	50	65,081	1593.8	20	50	61,379	2725.6
				100	∞	35,819	599.7	100	∞	33,346	574.5
				100	500	40,863	674.0	100	500	32,866	588.1
				200	∞	34,441	586.6	200	∞	31,678	536.8
				200	500	33,430	544.3	200	500	25,350	462.6

Here, $(m, n) = $ problem size; $k = $ number of edge directions used to generate the external columns; $nmax = $ number of simplex iterations between external columns, where "∞" means that an external column is generated only at the initial basis
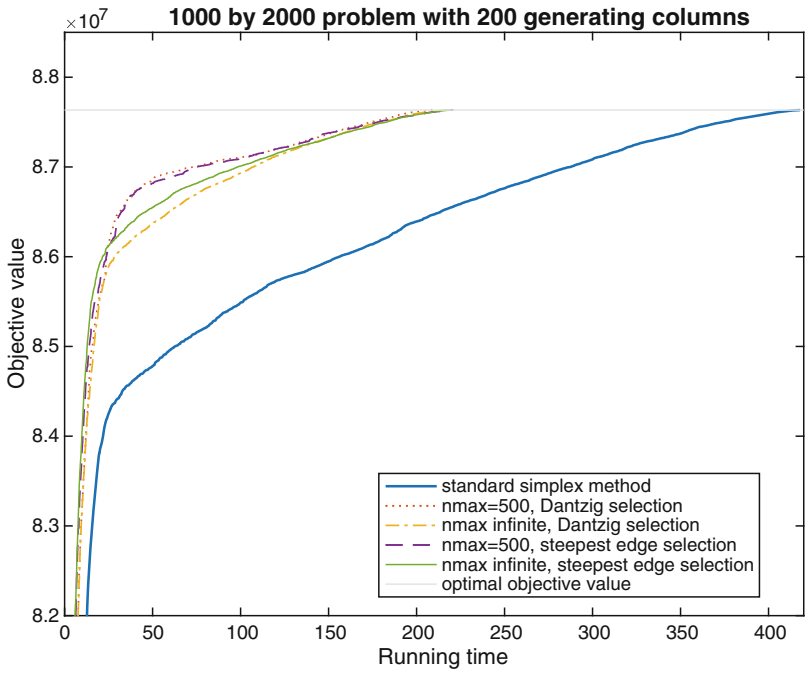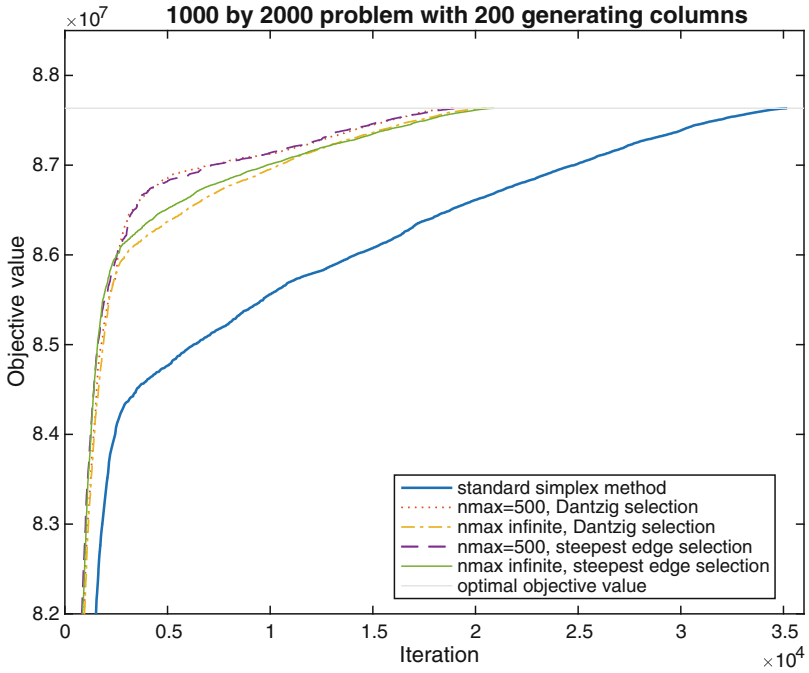
Fig. 2 Objective value versus iteration and running time, respectively

References

1. Bertsekas, D.P.: Projected Newton methods for optimization problems with simple constraints. SIAM J. Control Optim. **20**, 221–246 (1982)
2. Eiselt, H.A., Sandblom, C.-L.: Experiments with external pivoting. Comput. Oper. Res. **17**, 325–332 (1990)
3. Goldfarb, D., Reid, J.K.: A practicable steepest-edge simplex algorithm. Math. Program. **12**, 361–371 (1977)
4. Murty, K.G.: Linear Programming. Wiley, New York (1983)
5. Murty, K.G., Fathi, Y.: A feasible direction method for linear programming. Oper. Res. Lett. **3**, 121–127 (1984)