

Comparison of Interior-Point Methods Versus Simplex Algorithms

KUMAR MUKESH^{1*} AND RAJ SHEKHAR PRASAD²

1. Department of Mathematics, R. P. Sharma Institute of Technology, Patna, India.

2. Department of Mathematics, Government Engineering College, Buxar, India

Abstract:

During the last twenty years, there has been a revolution in the methods used to solve optimization problems. In the early 1980s, sequential quadratic programming and augmented Lagrangian methods were favoured for nonlinear problems, while the simplex method was unchallenged for linear programming. Since then, modern interior-point methods (IPMs) have infused virtually every area of continuous optimization, and have forced great improvements in the earlier methods. This paper aims to describe interior-point methods and their application to convex programming, special conic programming problems (including linear and semi-definite programming), and general possibly non-convex programming.

Keywords and phrases: Interior point Method, Self-concordance, conic programming, and non-convex programming.

AMS Subject Classification 2010: 90C05

Date of Submission: 06-11-2020

Date of Acceptance: 19-11-2020

I. Introduction

Almost twenty-five years ago, Karmarkar [11] proposed his projective method to solve linear programming problems: from a theoretical point of view, this was a polynomial-time algorithm, in contrast to Dantzig's simplex method. Moreover, with some refinements, it proved a very worthy competitor in practical computation, and substantial improvements to both interior-point and simplex methods have led to the routine solution of problems (with hundreds of thousands of constraints and variables) that were considered untouchable previously. Most commercial software, for example, CPLEX (Bixby [12]) and XpressMP (Gueret, Prins, and Sevaux [4]), includes interior-point as well as simplex options.

The majority of the early papers following Karmarkar's dealt exclusively with linear programming and its near-relatives, convex quadratic programming, and the (monotone) linear complementarity problem. Gill, Murray, Saunders, Tomlin and Wright [11] showed the strong connection to earlier barrier methods in nonlinear programming; Gonzaga et.al [4] introduced path-following methods with an improved iteration complexity; Kojima, Mizuno and Yoshise [9] realized, primal-dual versions of these algorithms, which are the most successful in practice.

At the same time, Nesterov and Nemirovski were investigating the new methods from a more fundamental viewpoint: What are the basic properties that lead to polynomial-time complexity? It turned out that the key property is that the barrier function should be self-concordant. This seemed to provide a clear, complexity-based criterion to delineate the class of optimization problems that could be solved in a provably efficient way using the new methods. The culmination of this work was the book by Nesterov and Nemirovski, whose complexity emphasis contrasted with the classic text on barrier methods by Fiacco and McCormick [2].

Fiacco and McCormick describe the history of the (exterior) penalty and barrier (sometimes called interior penalty) methods; other useful references are Nash (1998) and Forsgren, Gill, and Wright [3]. Very briefly, Courant [14] first proposed penalty methods, while Frisch [7] suggested the logarithmic barrier method and Carroll [6] the inverse barrier method (which inspired Fiacco and McCormick). While these methods were among the most successful for solving constrained nonlinear optimization problems in the 1960s, they lost favour in the late 1960s and 1970s when it became apparent that the sub problems that needed to be solved became increasingly ill-conditioned as the solution was approached.

The new research alleviated these fears to some extent, at least problems. Besides, the ill-conditioning turned out to be relatively Wright (1992) and Forsgren et al. [3].

The theory of self-concordant barriers is limited to convex optimization. However, this limitation has become less burdensome as more and more scientific and engineering problems are amenable to convex optimization formulations. Researchers in control theory have been much influenced by the ability to solve semi-definite programming problems (or linear matrix inequalities, in their terminology) arising in their field:

see Boyd, El Ghaoui, Feron, and Balakrishnan [17]. Moreover, several seemingly non-convex problems arising in engineering design can be reformulated as convex optimization problems: may be viewed in reference- Boyd, and Vandenberghe [16], Ben-Tal, and Nemirovski [1].

We have concentrated on the theory and application in structured convex programming of interior-point methods since the polynomial-time complexity of these methods and its range of applicability have been a major focus of the research of the last twenty years. For further coverage of interior-point methods for general nonlinear programming, we recommend the survey articles of Forsgren et al. [3] and Could, Orban, and Toint [10]. Also, to convey the main ideas of the methods, we have given short shrift to important topics including attaining feasibility from infeasible initial points, dealing with infeasible problems, and super-linear convergence. The literature on interior-point methods is huge, and the area is still very active.

II. The self-concordance-based approach to IPMs:

Preliminaries

The first path-following interior-point polynomial-time methods for linear programming, analyzed by Gonzaga [4], turned out to belong to the very well-known interior penalty scheme going back to Fiacco and McCormick [2]. Consider a convex program

$$\min\{c^T x : x \in X\}, \quad (1)$$

X being a closed convex domain (i.e., a closed convex set with a non-empty interior) in this is one of the universal forms of a convex program. To solve the problem with a path-following scheme, one equips X with an interior penalty or barrier function F – a smooth and strongly convex function defined on $\text{int } X$ such that on every sequence of points converging to a point and considers the barrier family of functions

$$F_t(x) = tc^T x + F(x), \quad (2)$$

where $t > 0$ is the penalty parameter. Under mild assumptions (e.g. when X is bounded), every function F attains its minimum on $\text{int } X$ at a unique point $X^*(t)$, and the central path $\{X^*(t) : t \geq 0\}$ converges, as $t \rightarrow +\infty$ to the optimal set of (1). The path-following scheme for solving (1) suggests ‘tracing’ this path as $t \rightarrow +\infty$ according to the following conceptual algorithm:

Given the current iterate $(t_k > 0, x_k \in \text{int } X)$ with x_k ‘reasonably close’ to $x_*(t_k)$, we

- (a) Replace the current value t_k of the penalty parameter with a larger value t_{k+1} ; and
- (b) Run an algorithm for minimizing $F_{t_{k+1}}(\cdot)$, starting at x_k , until a point x_{k+1} close to $x_*(t_{k+1}) = \arg \min_{\text{int } X} F_{t_{k+1}}(\cdot)$ is found.

The main advantage of the scheme described above is that $x_*(t)$ is, essentially, the unconstrained minimizer of F_t , which allows the use in (b) of basically any method for smooth convex unconstrained minimization, e.g., the Newton method. Note, however, that the classical theory of the path-following scheme did not suggest its polynomiality; rather, the standard theory of unconstrained minimization predicted slow-down of the process as the penalty parameter grows. In sharp contrast to this common wisdom, both Renegar and Gonzaga proved that, when applied to the logarithmic barrier $F(x) = -\sum_i \ln(b_i - a_i^T x)$ for a polyhedral set $X = \{x : a_i^T x \leq b_i, 1 \leq i \leq m\}$, a Newton-method-based implementation of the path-following scheme can be made polynomial. These breakthrough results were obtained via an ad hoc analysis of the behavior of the Newton method as applied to the logarithmic barrier (augmented by a linear term). In a short time Nesterov realized what intrinsic properties of the standard log-barrier are responsible for this polynomiality, and this crucial understanding led to the general self-concordance-based theory of polynomial-time interior-point methods developed in Nesterov and Nemirovski; this theory explained the nature of existing interior-point methods (IPMs) for LP and allowed the extension of these methods to the entire field of convex programming. We now provide an overview of the basic results of this theory.

III. Self-concordance

In retrospect, the notion of self-concordance can be extracted from analysis of the classical results on the local quadratic convergence of Newton’s method as applied to a smooth convex function f with nonsingular Hessian. These results state that a quantitative description of the domain of quadratic convergence depends on

(a) the condition number of $\nabla^2 f$ evaluated at the minimize x_* and (b) the Lipschitz constant of $\nabla^2 f$. In hindsight, such a description seems unnatural, since it is ‘frame-dependent’: it heavily depends on an ad hoc choice of the Euclidean structure in \mathbb{R}^n ; indeed, both the condition number of $\nabla^2 f(x_*)$ and the Lipschitz constant of $\nabla^2 f(\cdot)$ depend on this structure, which is in sharp contrast to the affine invariance of the Newton methods itself. At the same time, a smooth strongly convex function f by itself defines at every point x a Euclidean structure $\langle u, v \rangle_{f, x} = D^2 f(x)[u, v]$. With respect to this structure, $\nabla^2 f(x)$ is as well-conditioned as it could be it is just the unit matrix. The idea of Nesterov was to use this local Euclidean structure, intrinsically linked to the function f we intend to minimize, in order to quantify the Lipschitz constant of $\nabla^2 f$, with the ultimate goal of getting a ‘frame-independent’ description of the behavior of the Newton method. The resulting notion of self-concordance is defined as follows.

Definition 2.1 Let $X \subset \mathbb{R}^n$ be a closed convex domain. A function $f : \text{int } X \rightarrow \mathbb{R}$ is called self-concordant (Sc) on X if

- (i) f is a three times continuously differentiable convex function with $f(x_k) \rightarrow \infty$ if $x_k \rightarrow \bar{x} \in \partial$; and
- (ii) f satisfies the differential inequality

$$|D^3 f(x)[h, h, h]| \leq 2(D^2 f(x)[h, h])^{3/2}, \forall x \in \text{int } X, h \in \mathbb{R}^n. \quad (3)$$

Given a real $v \geq 1$, F is called a v -self-concordant barrier (v -SCB) for X if F is self-concordant on X and, in addition,

$$DF(x)[h] \leq 291/2(D^2 F(x)[h, h])^{1/2} \forall x \in \text{int } X, h \in \mathbb{R} \quad (4)$$

(As above, we will use f for a general sc function and F for an SCB in what follows.) Note that the powers $3/2$ and $1/2$ in (3) and (4) are a must, since both sides of the inequalities should be of the same homogeneity degree with respect to h . In contrast to this, the two sides of (3) are of different homogeneity degrees with respect to f , meaning that if f satisfies a relation of the type (3) with some constant factor on the right-hand side, we can always make this factor equal to 2 by scaling f appropriately. The advantage of the specific factor 2 is that with this definition, the function $x \rightarrow -\ln(x) : \mathbb{R}_{++} \rightarrow \mathbb{R}$ becomes a 1-SCB for \mathbb{R}_+ directly, without any scaling, and this function is the main building block of the theory we are presenting. Finally, we remark that (3) and (4) have a very transparent interpretation: they mean that $D^2 f$ and F are Lipschitz-continuous, with constants 2 and $v^{1/2}$, in the local Euclidean (semi) norm $\|h\|_{f, x} = \sqrt{\langle h, h \rangle_{f, x}} = \sqrt{h^T \nabla^2 f(x) h}$ defined by f or similarly by F .

It turns out that self-concordant functions possess nice local properties and are perfectly well suited to Newton minimization. We are about to present the most important of the related results. In what follows, f is an SC function on a closed convex domain X .

Bounds on third derivatives and the recession space of Sc functions For all $x \in \text{int } X$ and all $h_1, h_2, h_3 \in \mathbb{R}^n$, we have

$$|D^3 f(x)[h_1, h_2, h_3]| \leq 2 \|h_1\|_{f, x} \|h_2\|_{f, x} \|h_3\|_{f, x}.$$

The recession subspace $E_f = \{h : D^2 f(x)[h, h] = 0\}$ of f is independent of $x \in \text{int } X$, and $X = X + E_f$.

In particular, if $\nabla^2 f(x)$ is positive definite at some point in $\text{int } X$, then $\nabla^2 f(x)$ is positive definite for all $x \in \text{int } X$ (in this case, f is called a non-degenerate SC function; this is always the case when X does not contain lines).

It is convenient to write $A \succeq 0$ ($A \preceq 0$) to denote that the symmetric matrix A is positive definite (semidefinite), and $A \succeq B$ and $B \succeq A$ ($A \preceq B$ and $B \preceq A$) if $A - B \succeq 0$ ($A - B \preceq 0$).

IV. Dikin’s ellipsoid and the local behavior of f

For every $x \in \text{int } X$, the unit Dikin. ellipsoid off $\{y : \|y - x\|_{f, x} \leq 1\}$ is contained in X , and within this ellipsoid, f is nicely approximated by its second-order Taylor expansion:

$$r := \|h\|_{f,x} < 1 \Rightarrow$$

$$(1-r)^2 \nabla^2 f(x) \preceq \nabla^2 f(x+h) \preceq \frac{1}{(1-r)^2} \nabla^2 f(x) \quad (5)$$

$$f(x) + \nabla f(x)^T h + \rho(-r) \leq f(x+h) \leq f(x) + \nabla f(x)^T h + \rho(r),$$

where $\rho(s) := -\ln(1-s) - s = s^2/2 + s^3/3 + \dots$ (Indeed, the lower bound in the last line holds true for all h such that $x+h \in \text{int } X$.)

V. The Newton decrement and the damped Newton method

Let f be non-degenerate. Then $\|\cdot\|_{f,x}$ is a norm, and its conjugate norm is $\|\eta\|_{f,x}^* =$

$$\max\{h^T \eta : \|h\|_{f,x} \leq 1\} = \sqrt{\eta^T [\nabla^2 f(x)]^{-1} \eta}$$

The quantity

$$\begin{aligned} \lambda(x, f) &:= \|\nabla f(x)\|_{f,x}^* = \|\nabla^2 f(x)^{-1} \nabla f(x)\|_{f,x} \\ &= \max_h \{Df(x)[h] : D^2 f(x)[h, h] \leq 1\}, \end{aligned}$$

called the Newton decrement of f at x , is a finite continuous function of $x \in \text{int } X$ which vanishes exactly at the (unique, if any) minimizer x_f of f on $\text{int } X$; this function can be considered as the ‘observable’ measure of proximity of x to X_f . In particular, when it is at most $1/2$, the Newton decrement is, within an absolute constant factor, the same as in the damped Newton method as applied to f is the iterative process.

We get the following ‘frame- and data-independent’ description of the convergence properties of the damped Newton method as applied to a sc function f : the domain of quadratic convergence is $\{x : A(x, f) < 1/4\}$; after this domain is reached, every step of the method nearly squares the Newton decrement, the $\|x - x_f\|$ distance to the minimizer and the residual in terms of f . Before the domain is reached, every step of the method decreases the objective by at least $c_2(1) = 1/4 - \ln(5/4)$. It follows that a non-degenerate sc function admits its minimum on the interior of its domain if and only if it is bounded below, and if and only if $A(x, f) < 1$ for certain x . Whenever this is the case, for every $\epsilon \in (0, 0.1]$ the number of steps N of the damped Newton method which ensures that $f(x_k) - \min x f + \epsilon$ does not exceed $0(1)$.

Existence of the central path and its convergence to the optimal set Consider problem as above and assume that the domain X of the problem is equipped with a self-concordant barrier F , and the level sets of the objective $\{x \in X : c^T x < a\}$ are bounded. In the situation in question, F is nondegenerate, $c^T x$ attains its minimum on X , the central path $x(t) := \arg \min F(x), F(x) := t c^T x + F(x), t > 0$,

In the case of linear programming, when K (and then also K_*) is the nonnegative orthant, then whenever (P) or (D) is feasible, we have equality of their optimal values (possibly $\pm \infty$), and if both are feasible, we have strong duality: no duality gap, and both optimal values attained.

In the case of more general conic programming, these properties no longer hold (we will provide examples in the next subsection), and we need further regularity conditions. Nesterov and Nemirovski (1994, Theorem 4.2.1) derive the next result.

VI. Theorem.

If either (P) or (D) is bounded and has a strictly feasible solution (i.e., a feasible solution where x (respectively, s) lies in the interior of K (respectively, K_*)), then their optimal values are equal. If both have strictly feasible solutions, then strong duality holds.

The existence of an easily stated dual problem provides one motivation for considering problems in conic form (but its usefulness depends on having a closed form expression for the dual cone). We will also see that many important applications naturally lead to conic optimization problems. Finally, there are efficient primal-dual interior-point methods for this class of problems, or at least for important subclasses.

VII. Examples of conic programming problems

First, it is worth pointing out that any convex programming problem can be put into conic form. Without loss of generality, after introducing a new variable if necessary to represent a convex nonlinear objective function, we can assume that the original problem is

$$\min_x \{C^T X : x \in X\}$$

with X a closed convex subset of \mathbb{R}^n . This is equivalent to the conic optimization problem, but for one dimension higher:

$$\min_{x, \xi} \{C^T X : \xi = 1, (X, \xi) \in K\},$$

where $K := \text{cl}\{(x, \xi) \in \mathbb{R}^n \times \mathbb{R} : \xi > 0, x/\xi \in X\}$. However, this formal equivalence may not be very useful practically, partly because K and K may not be easy to work with. More importantly, even if we have a good self-concordant barrier for X , it may be hard to obtain an efficient self-concordant barrier for K (although general, if usually over conservative, procedures are available: see rule E in Section 2.4 and Freund, Jarre and Schaible [15]).

Let us turn to examples with very concrete and useful cones. The first example is of course linear programming,

where $K = \mathbb{R}_+^n$. Then it is easy to see that K is also \mathbb{R}_+^n , and so the dual constraints are just $A^T y \leq c$. The significance and wide applicability of linear programming are well known. Our first case with a non-polyhedral cone is what is known as second-order cone programming (SOCP). Here K is a second-order, or Lorentz, or 'icecream' cone,

$$L^q := \{(\xi, \bar{x}) \in \mathbb{R} \times \mathbb{R}^q : \|\bar{x}\| \leq \xi\}$$

or the product of such cones. It is not hard to see, using the Cauchy Schwarz inequality, that such cones are also self-dual, i.e., equal to their duals. We now provide an example showing the usefulness of SOCP problems (many more examples can be found in Ben-Tal and Nemirovski [1]), and also a particular instance demonstrating that strong duality does not always hold for such problems.

Suppose we are interested in solving a linear programming problem $\max \{b^T y : A^T y \leq c\}$, but the constraints are not known exactly: for the j th constraint $a_k^T y \leq c_j$, we just know that $(c_j; a_j) \in \{(\bar{c}_j; \bar{a}_j) + P_j u_j : \|u_j\| \leq 1\}$, an ellipsoidal uncertainty set centred at the nominal values $(\bar{c}_j; \bar{a}_j)$.

(We use the MATLAB-like notation $(u; v)$ to denote the concatenation of the vectors u and v .) Here P_j is a suitable matrix that determines the shape and size of this uncertainty set. We would like to choose our decision variable y so that it is feasible no matter what the constraint coefficients turn out to be, as long as they are in the corresponding uncertainty sets; with this limitation, we would like to maximize $b^T y$. This is (a particular case of)

the so-called robust linear programming problem. Since the minimum of $c_j - a_j^T y = (c_j; a_j)^T (1; -y)$

over the j th uncertainty set is $(\bar{c}_j; \bar{a}_j)^T (1; -y) + \min\{P_j u_j\}^T (1; -y) : \|u_j\| \leq 1\} =$

$(\bar{c}_j; \bar{a}_j)^T (1; -y) - \|P_j^T (1; -y)\|$, this robust linear programming problem can be formulated as $\max b^T y$

$$-\bar{c}_j; \bar{a}_j^T y + S_{j1} = 0, j = 1, \dots, m,$$

$$P_j^T [(1; -y) + S_{j1}] = 0, j = 1, \dots, m,$$

$$(S_{ji}; \bar{S}_j) \in K_j, j = 1, \dots, m,$$

where each K_j is a second-order cone of appropriate dimension. This is a SOCP problem in dual form.

Next, consider the SOCP problem in dual form with data

$$a = \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, c = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

and K the second-order cone in \mathbb{R}^3 . It can be checked that y is feasible in (D) if and only if y_1 and y_2 are positive, and $4y_1y_2 \geq 1$. Subject to these constraints, we wish to maximize $-y_1$, so the problem is feasible, with objective function bounded above, but there is no optimal solution! In this case, the optimal values of primal and dual are equal: $(\xi, \bar{x}) = (1/2; 0; 1/2)$ is the unique feasible solution to (P), with zero objective function value.

The second class of non-polyhedral cones we consider gives rise to semi-definite programming problems. These correspond to the case when K is the cone of positive semi definite matrices of a given order (or possibly a Cartesian product of such cones). Here we will restrict ourselves to the case of real symmetric matrices, and we use S^p to denote the space of all such matrices of order p . Of course, this can be identified with \mathbb{R}^n for $n := p(p+1)/2$, by making a vector from the entries and $\sqrt{2}m_{ij}, i < j$. We use the factor $\sqrt{2}$ so that the usual scalar product of the vectors corresponding to two symmetric matrices U and V equals the Frobenius scalar product

$$U \square V := \text{Tr}(U^T V) = \sum_{i,j} u_{ij} v_{ij}$$

of the matrices. However, we will just state these problems in terms of the matrices for clarity. We write S_+^p for the cone of (real symmetric) positive semidefinite matrices, and sometimes write $X \square 0$ to denote that X lies in this cone for appropriate P . As in the case of the non-negative orthant and the second-order cone, S_+^p is self-dual. This can be shown using the spectral decomposition of a symmetric matrix. We note that the case of complex Hermitian positive semidefinite matrices can also be considered and this is important in some applications.

In matrix form, the constraint $AX = b$ is defined using an operator A forms S^p to \mathbb{R}^m and we can find matrices $A_{ij} \in S^p, i = 1, \dots, m$, so that $AX = (A_j \square X)_{j=1}^m A^T$ is then the adjoint operator from \mathbb{R}^m to S^p defined by $A^T y = \sum_i y_i A_i$. The primal and dual semidefinite programming problems then become

$$\min C \square X, A_i \square X = b_i, i = 1, \dots, m, x \square 0 \tag{6}$$

$$\text{and} \quad \max b^T y, \sum_i y_i A_i + S = C, S \square 0. \tag{7}$$

Once again, we give examples of the importance of this class of conic optimization problems, and also an instance demonstrating the failure of strong duality.

Let us first describe a very simple example that illustrates techniques used in optimal control. Suppose we have a linear dynamical system

$$\dot{z}(t) = A(t) z(t),$$

where the $p \times p$ matrices $A(t)$ are known to lie in the convex hull of a number A_1, \dots, A_k , of given matrices. We want conditions that guarantee that the trajectories of this system stay bounded. Certainly a sufficient condition is that there is a positive definite matrix $Y \in S^p$ so that the Lyapunov function $L(z(t)) := z(t)^T Y z(t)$ remains bounded. And this will hold as long as $\dot{L}(z(t)) \leq 0$. Now using the dynamical system, we find that

$$\dot{L}(z(t)) = z(t)^T (A(t)^T Y + Y A(t)) z(t),$$

and since we do not know where the current state might be, we want $-A(t)^T Y - Y A(t)$ to be positive semidefinite whatever $A(t)$ is, and so we are led to the constraints

$$-A_i^T Y - Y A_i \square 0, i = 1, \dots, k, \quad Y - I_p \square 0,$$

where the last constraint ensures that Y is positive definite. (Here I_p denotes the identity matrix of order p . Since the first constraints are homogeneous in Y , we can assume that Y is scaled so its minimum eigenvalue is at least 1). To make an optimization problem, we could for instance minimize the condition number of Y by

adding the constraint $\eta \mathbf{I}_p - \mathbf{Y} \succeq \mathbf{0}$ and then maximizing $-\eta$. This is a semidefinite programming problem in dual form. Note that the variables y are the entries of the symmetric matrix \mathbf{Y} and the scalar η and the cone is the product of $k+2$ copies of \mathbb{S}_+^p . We can similarly find sufficient conditions for $z(t)$ to decay exponentially to zero.

Our second example is a relaxation of a quadratic optimization problem with quadratic constraints. Notice that we did not stipulate that the problem be convex, so we can include constraints like $x_j^2 = x_j$. Which implies that x_j is 0 or 1, i.e., we have included binary integer programming problems. Any quadratic function can be written as a linear function of a certain symmetric matrix.

The set of all matrices $\begin{pmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{xx}^T \end{pmatrix}$ is certainly a subset of the set of all positive semi definite matrices with top left entry equal to 1, and so we can obtain a relaxation of the original hard problem in \mathbf{x} by optimizing over a matrix \mathbf{X} that is subject to the constraints defining this superset. This technique has been very successful in a number of combinatorial problems, and has led to worthwhile approximations to the stable set problem, various satisfiability problems, and notably the max-cut problem. Further details can be found, for example, in Goemans [8] and Ben-Tal and Nemirovski [1].

Let us give an example of two dual semidefinite programming problems where strong duality fails. The primal problem is where the first constraint implies that and hence x_{12} and x_{21} , are zero, and so the second constraint implies that x_{33} is 1. Hence one optimal solution is $\mathbf{X} = \text{Diag}(0; 0; 1)$ with optimal value 1. The dual problem is so the dual slack matrix \mathbf{S} has $s_{22} = 0$, implying that s_{12} and s_{21} must be zero, so y_2 must be zero. So an optimal solution is $\mathbf{y} = (0; 0)$ with optimal value 0. Hence, while both problems have optimal solutions, their optimal values are not equal. Note the both problems have a strictly feasible solution, and arbitrary small perturbations in the data, which can make the optimal values, jump.

References:-

- [1]. A. Ben – Tal and A. S. Nemirovski (2001), Lectures on Modern Convex Optimization : Analysis, Algorithms, and Engineering Applications, SIAM, Philadelphia.
- [2]. A. V. Fiacco and G. P. McCormick (1968), Nonlinear Programming: Sequential Unconstrained Minimization Techniques, Wiley.
- [3]. A. Forsgren, P. E. Gill and M. H. Wright (2002), ‘Interior methods for nonlinear optimization’ , SIAM Review 44, 525 – 597.
- [4]. C. C. Gonzaga (1989), AN algorithm for solving linear programming problems in $O(n^3L)$ operations. In progress in Mathematical programming : Interior point and related methods (N. Megiddo, ed.), Springer, New York, pp. 1-28.
- [5]. C. Gueret, C. Prins and M. Sevaux (2002), Applications of optimization with Xpress MP, dash optimization. Translated and revised by Susanne Heipcke.
- [6]. C. W. Carroll (1961), ‘The Created response surface technique for optimizing nonlinear, restrained systems’ , Oper. Res. 9, 169-185.
- [7]. K. R. Frisch (1955), The Logarithmic potential method of convex programming. Memorandum of may 13, University Institute of Economics, Oslo, Norway.
- [8]. M. X. Goemans (1997) , ‘ Semiindefinite programming in combinatorial optimization’ , Math. Program. 79, 143 161.
- [9]. M. Kojima, S. Mizuno and A. Yoshise (1989), ‘A polynomial – time algorithm for a class of linear complementarity problems’ , Math. Program. 44, 1 – 26.
- [10]. N. I. M. Gould, D.Orban and P. L. Toint (2005), Numerical methods for large – scale nonlinear optimization. In Acta Numerica, Vol. 14, Cambridge University Press, pp. 299 – 361.
- [11]. N. Karmarkar (1984), ‘ A new polynomial – time algorithm for linear programming’ , Combinatorica 4, 373 – 395.
- [12]. P. E. Gill, Murray, M. A. Saunders, J. A. Tomlin and M.H. Wright (1986), ‘On projected Newton barrier methods for linear programming and an equivalence to Karmarkar’s projective method’ , Math. Program. 36, 183 – 209.
- [13]. R. E. Bixby (2002), solving real –world linear programmings: A decade and more of progress, Oper. Res. 50, 3-15
- [14]. R. Courant (1943), ‘Variational methods for the solution of the problems of equilibrium and vibrations’ , Bull. Amer. Math. Soc. 49, 1-23.
- [15]. R. W. Freund, F. Jarre and S. Schaible (1996), ‘On self-concordant barrier functions for conic hulls and fractional programming’ , Math Programm. 74, 237-246.
- [16]. S. Boyd and L. Vandenberghe (2004), Convex Optimization, Cambridge Press.
- [17]. S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishann (1994), Linear Matrix Inequalities in System and Control Theory, Vol. 15 of studies in Applied Mathematics, SIAM.

KUMAR MUKESH, et. al. "Comparison of Interior-Point Methods Versus Simplex Algorithms." *IOSR Journal of Mathematics (IOSR-JM)*, 16(6), (2020): pp. 53-59.