THE CENTRAL PATH IN LINEAR PROGRAMMING

WORKING NOTES OF
JESÚS DE LOERA AND BERND STURMFELS

ABSTRACT. This project concerns the convex algebraic geometry of the central path of a linear programming problem. This path is an algebraic curve, described by linear and quadratic constraints arising from complementary slackness. We are interested in the defining polynomial equations and geometric invariants of this curve, such as degree, genus and singularities. These parameters are related to the curvature and thus to the performance of interior point methods. We also explore the natural extension to semidefinite programming.

1. INTRODUCTION

We consider the following linear programming problem in standard form:

$$\text{Maximize } \mathbf{c}^T\mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0. \tag{1}$$

The matrix A has n columns and rank d, and we shall assume that the feasible polyhedron P of (1) is bounded, so it is a *polytope*. The *logarithmic barrier function* is defined as

$$f_\mu(\mathbf{x}) \;:=\; \mathbf{c}^T\mathbf{x} + \mu \sum_{i=1}^{n} \log x_i,$$

where $\mu > 0$ is a real parameter. This defines a family of related optimization problems:

$$\text{Maximize } f_\mu(\mathbf{x}) \text{ subject to } A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0. \tag{2}$$

Since the logarithm is a strictly concave function on the positive real axis, the function f_μ is strictly concave, and it attains a unique maximum $\mathbf{x}^*(\mu)$ in the interior of the feasible polytope P of (1). Note that $f_\mu(\mathbf{x})$ tends to $-\infty$ when $\mathbf{x}$ approaches the boundary of P. The *central path* of the linear program (1) is the curve $\{\mathbf{x}^*(\mu) \mid \mu > 0\}$ inside the polytope P.

To gain an understanding of the geometry of the central path, we consider the dual problem

$$\text{Minimize } \mathbf{b}^T\mathbf{y} \text{ subject to } A^T\mathbf{y} \geq \mathbf{c}. \tag{3}$$

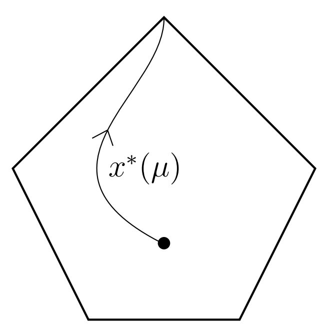


FIGURE 1. The central path: a curve through the feasible polytope.

The pair of optimal solutions, to the primal linear program (1) and the dual linear program (3), are characterized by the following *complementary slackness* condition:

$$(4) \qquad A\mathbf{x} = \mathbf{b}\,,\; A^T\mathbf{y} - \mathbf{s} = \mathbf{c}\,,\; \mathbf{x} \geq 0\,,\; \mathbf{s} \geq 0 \text{ and } x_i s_i = 0 \text{ for } i = 1, 2, \ldots, n.$$

The central path is obtained from this formulation by replacing 0 with μ:

$$(5) \qquad A\mathbf{x} = \mathbf{b}\,,\; A^T\mathbf{y} - \mathbf{s} = \mathbf{c} \text{ and } x_i s_i = \mu \text{ for } i = 1, 2, \ldots, n.$$

If we eliminate the coordinates of the vectors $\mathbf{y}$ and $\mathbf{s}$ from (5) then we get precisely the central path $\mathbf{x} = \mathbf{x}^*(\mu)$. This result follows from well-known properties of *Lagrange multipliers*. The central path $\mathbf{x}^*(\mu)$ converges for $\mu \to 0$ to the optimal solution of the initial problem (1).

Theorem 1 (The Fundamental Lemma of Interior Point Methods for LP)**.** *For all $\mu > 0$, the system of algebraic equations* (5) *has a unique real solution* $(\mathbf{x}^*(\mu), \mathbf{y}^*(\mu), \mathbf{s}^*(\mu))$ *with the properties* $\mathbf{x}^*(\mu) > 0$ *and* $\mathbf{s}^*(\mu) > 0$. *The vector* $\mathbf{x}^*(\mu)$ *is the optimal solution of* (2). *The limit* $(\mathbf{x}^*(0), \mathbf{y}^*(0), \mathbf{s}^*(0))$ *of that solution for $\mu \to 0$ is the unique solution of the system* (4).

For a derivation of this theorem and relevant background we refer to the monograph of Renegar [5], and to the work of Bayer and Lagarias [1,2] on the geometry of the central path.

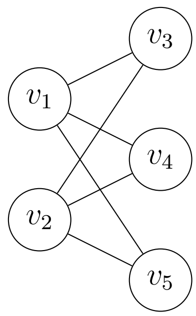
It is worth stressing that, in practical computations, the maximum in (1) is found by following a piecewise-linear approximation to the central path. This is what is called an *interior point method*. The number of steps needed to solve (1) has the order $O(\sqrt{N}L)$, where N is the number of constraints and L is the bit size of the input coefficients. However, approximating the central path with segments requires care to ensure convergence, and a judicious choice of step size and initial point. Deza, Terlaky and Zinchenko [3] constructed instances of (1) in which the number of steps for a central path approximation is greater than $O(\sqrt{N}\log N)$. This shows how little is understood of the geometry of the central path. For example, adding redundant constraints dramatically changes the central path (even though the polytope P remains unchanged) and consequently the running time of the algorithm.

In the next section we study one example which will illustrate the concepts and our questions. In Section 3 we allow A to be an arbitrary matrix A and we examine the central path under the hypothesis that $\mathbf{b}$ and $\mathbf{c}$ are generic. Under this hypothesis, the central path specifies an irreducible curve in $\mathbb{R}^n$, and we compute prime ideal of that curve. This resolves the open problem stated in the last sentence of [2, §11]. We also express the degree of that curve as a matroid invariant, and we prove the tight upper bound $\binom{n-1}{\mathrm{rank}(A)}$ for that degree.

2. Quintic Curves in Transportation Polygons

One of the best-studied family of linear programs are those arising from transportation problems. Here the decision variables are the entries in a non-negative matrix, the linear constraints specify the row sums and the column sums, and one seeks to maximize a linear function in the matrix entries. The resulting polytopes P are the *transportation polytopes*.

In this section we shall examine the central path for the transportation problem for 2×3-matrices. If we write this problem in the standard form (1), then the matrix A is the node-edge incidence matrix of the complete bipartitite graph $K_{2,3}$. This matrix has format 5×6, its entries are 0 or 1, and its rank is 4. The rows of A correspond to the vertices of $K_{2,3}$, as shown in the diagram below, and its columns correspond to the edges of $K_{2,3}$.



$$\longleftrightarrow \qquad \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

The feasible polytope P is empty unless $\mathbf{b}$ is in the column span of A, which means that

$$b_1 + b_2 \;=\; b_3 + b_4 + b_5. \tag{6}$$

We shall now assume that this holds and that $b_i > 0$ for $i = 1, 2, 3, 4, 5$. Then P is two-dimensional. It consists of all non-negative $2{\times}3$-matrices with given row and column sums:

$$\begin{array}{cc} & \begin{array}{ccc} b_3 & b_4 & b_5 \end{array} \\ \begin{array}{c} b_1 \\ b_2 \end{array} & \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{pmatrix} \end{array}$$

To be pedantically explicit, the matrix equation $A\mathbf{x} = \mathbf{b}$ translates into the linear ideal

$$I_{A,\mathbf{b}} \;=\; \langle\, x_1 + x_2 + x_3 - b_1 \,,\, x_4 + x_5 + x_6 - b_2 \,,\, x_1 + x_4 - b_3 \,,\, x_2 + x_5 - b_4 \,\rangle.$$

The remaining equations in (5) translate into the ideal

$$\tilde{J}_{A,\mathbf{c}} \;=\; \left\langle \begin{array}{c} y_1 + y_3 - s_1 - c_1 \,,\, y_1 + y_4 - s_2 - c_2 \,,\, y_1 + y_5 - s_3 - c_3 \,, \\ y_2 + y_3 - s_4 - c_4 \,,\, y_2 + y_4 - s_5 - c_5 \,,\, y_2 + y_5 - s_6 - c_6 \,, \\ x_1 s_1 - \mu, x_2 s_2 - \mu, x_3 s_3 - \mu, x_4 s_4 - \mu, x_5 s_5 - \mu, x_6 s_6 - \mu \end{array} \right\rangle.$$

The two ideals live in a polynomial ring in 18 variables,

$$K[\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu] \;=\; K[x_1, x_2, x_3, x_4, x_5, x_6, y_1, y_2, y_3, y_4, y_5, s_1, s_2, s_3, s_3, s_4, s_5, s_6, \mu],$$

over the rational function field $K = \mathbb{Q}(\mathbf{b}, \mathbf{c}) = \mathbb{Q}(b_1, b_2, b_3, b_4, b_5, c_1, c_2, c_3, c_4, c_5)$. If we eliminate $\mathbf{y}$, $\mathbf{s}$ and μ from $J_{A,\mathbf{c}}$ then we get the principal ideal in $K[\mathbf{x}] = K[x_1, x_2, x_3, x_4, x_5, x_6]$,

$$J_{A,\mathbf{c}} \;=\; \tilde{J}_{A,\mathbf{c}} \cap K[\mathbf{x}] \;\;=\;\; \langle\, f_{A,\mathbf{c}}(\mathbf{x}) \,\rangle,$$

whose generator is the following polynomial of degree five:

$$f_{A,\mathbf{c}}(\mathbf{x}) \;=\; \det \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \\ x_1^{-1} & x_2^{-1} & x_3^{-1} & x_4^{-1} & x_5^{-1} & x_6^{-1} \end{pmatrix} \cdot x_1 x_2 x_3 x_4 x_5 x_6. \tag{7}$$

The quintic hypersurface defined by this polynomial depends only on the matrix A and the cost vector $\mathbf{c}$. It intersects the two-dimensional space defined by the linear ideal $L_{A,\mathbf{b}}$ in an irreducible quintic curve. That curve is the *universal central path* over K, that is, if we specialize $\mathbf{c}$ and $\mathbf{b}$ to any particular vectors in $\mathbb{R}^6$ and in $\mathbb{R}^5$, satisfying (6), then the result will be a plane algebraic curve that contains the central path. For almost all real values of the coordinates c_i and b_j, the resulting curve is irreducible and coincides with the Zariski

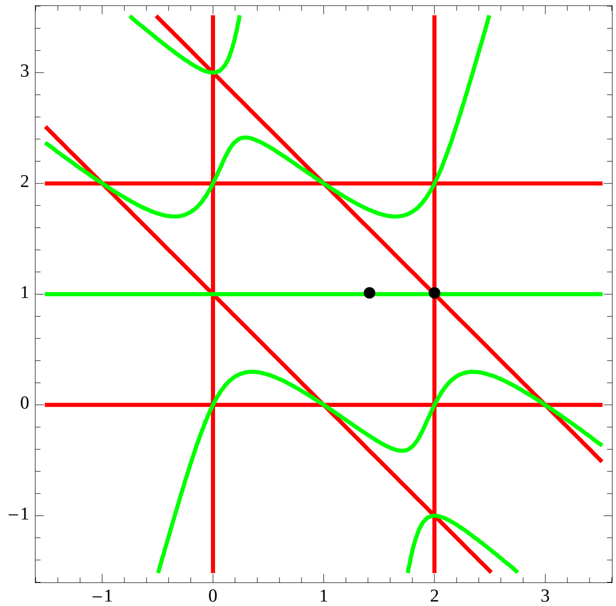


FIGURE 2. A transportation polygon and its central path.

closure of the central path. However, for special values of c_i and b_j, the ideal $I_{A,\mathbf{b}} + \langle f_{A,\mathbf{c}} \rangle$ may decompose into multiple components, only one of which represents the central path.

For a concrete numerical example we set $b_1 = b_2 = 3$ and $b_3 = b_4 = b_5 = 2$. Then the transportation polygon P is a regular hexagon, indicated in red in Figure 2. Its vertices are

$$\begin{pmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}. \tag{8}$$

Consider the two transportation problems (1) given by $\mathbf{c} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ and $\mathbf{c}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}$. In both cases, the last matrix in (8) is the unique optimal solution. Modulo the linear ideal $I_{A,\mathbf{b}}$ we can write the quintics $f_{A,\mathbf{c}}$ and $f_{A,\mathbf{c}'}$ as polynomials in only two variables x_1 and x_2. The resulting bivariate polynomials are

$$f_{A,\mathbf{c}} \quad \longrightarrow \quad \begin{array}{c} 3x_1^4x_2 + 5x_1^3x_2^2 - 2x_1x_2^4 - 3x_1^4 - 22x_1^3x_2 - 15x_1^2x_2^2 + 8x_1x_2^3 + 2x_2^4 \\ +18x_1^3 + 45x_1^2x_2 - 12x_2^3 - 33x_1^2 - 22x_1x_2 + 22x_2^2 + 18x_1 - 12x_2. \end{array}$$

$$f_{A,\mathbf{c}'} \quad \longrightarrow \quad \begin{array}{c} (x_2 - 1) \cdot (2x_1^4 + 4x_1^3x_2 + x_1^2x_2^2 - x_1x_2^3 - 12x_1^3 - 14x_1^2x_2 + x_1x_2^2 \\ +x_2^3 + 22x_1^2 + 10x_1x_2 - 5x_2^2 - 12x_1 + 6x_2) \end{array}$$

Both quintic curves pass through all intersection points of the six lines formed by the boundary edges of P. However, the first curve is irreducible, while the second curve has two components. The latter the curve shown in Figure 2. Here the Zariski closure of the central path is the straight line $x_2 = 1$ between the minimizing vertex and the maximizing vertex. For general cost functions $\mathbf{c}$, however, the central path is always an irreducible curve of degree 5.

3. THE IDEAL OF THE CENTRAL PATH

In this section we determine the vanishing ideal and the degree of the central path of the linear program (1). This answers a question raised by Bayer and Lagarias [2, §11]. The matrix A can be arbitrary subject to the standing assumption that the feasible polyhedron

P is bounded. As before, n is the number of columns of A and d is the rank of A. We work over the rational function field $K = \mathbb{Q}(\mathbf{b}, \mathbf{c})$ generated by the coordinates b_i and c_j of the right hand side $\mathbf{b}$ and the cost vector $\mathbf{c}$. This means that our result on the degree and prime ideal will remain be valid under almost all specializations $K \to \mathbb{R}$ of these coordinates.

Let $\mathcal{L}_{A,\mathbf{c}}$ denote the $(d+1)$-dimensional vector subspace of K^n spanned by the rows of A and the additional vector $\mathbf{c}$. The *central sheet* is the coordinatewise reciprocal $\mathcal{L}_{A,\mathbf{c}}^{-1}$ of that linear subspace. In precise terms, we define $\mathcal{L}_{A,\mathbf{c}}^{-1}$ to be the Zariski closure in K^n of the set

$$\left\{ \left(\frac{1}{u_1}, \frac{1}{u_2}, \dots, \frac{1}{u_n}\right) \in K^n \ : \ (u_1, u_2, \dots, u_n) \in \mathcal{L}_{A,\mathbf{c}} \quad \text{and} \quad u_i \neq 0 \ \text{ for } \ i = 1, \dots, n \right\}.$$

Lemma 2. *The Zariski closure of the central path $\{\mathbf{x}^*(\mu) \ : \ \mu \in \mathbb{R}_{\geq 0}\}$ is equal to the intersection of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$ with the affine-linear subspace defined by $A \cdot \mathbf{x} = b$.*

Proof. We eliminate $\mathbf{s}, \mathbf{y}$ and μ from the equations $A^T\mathbf{y} - \mathbf{s} = \mathbf{c}$ and $x_i s_i = \mu$ as follows. We first replace the coordinates of $\mathbf{s}$ by $s_i = \mu/x_i$. The linear system becomes $A^T\mathbf{y} - \mu\mathbf{x}^{-1} = \mathbf{c}$. This condition means that $\mathbf{x}^{-1} = (\frac{1}{x_1}, \dots, \frac{1}{x_n})$ lies in the linear space $\mathcal{L}_{A,\mathbf{c}}$ spanned by $\mathbf{c}$ and the rows of A. The result of the elimination is that $\mathbf{x}$ lies says the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$. ☐

The linear space $\{A \cdot \mathbf{x} = \mathbf{b}\}$ has dimension $n - d$, and we write $I_{A,\mathbf{b}}$ for its linear ideal. The central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$ is an irreducible variety of dimension $d+1$, and we write $J_{A,\mathbf{c}}$ for its prime ideal. This notation is consistent with that used in the example of Section 2.

The intersection of the linear space $\{A \cdot \mathbf{x} = \mathbf{b}\}$ with the central sheet is the variety of the ideal sum $I_{A,\mathbf{b}} + J_{A,\mathbf{c}}$. By slight abuse of notation, we refer to that intersection as the central path, as it is Zariski closure of the central path $\mathbf{x}^*(\mu)$. Since the right hand side vector $\mathbf{b}$ is generic over $\mathbb{Q}$, the resulting linear space is general enough, Bertini's Theorem ensures that the ideal generators and the degree of the central sheet are preserved under that intersection.

Lemma 3. *The degree of the central path coincides with the degree of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$. The prime ideal of polynomials that vanish on the central path is equal to $I_{A,\mathbf{b}} + J_{A,\mathbf{c}}$.*

At this point we are left with the problem of computing the minimal generators and the degree of the homogeneous ideal $J_{A,c}$. Luckily, this has already been done for us in the literature. The following theorem is due to Proudfoot and Speyer [4] and it refines an earlier result of Terao [7]. We refer to [6, Theorem 2] for an exposition of this result in a statistical context. The paper [6] deals primarily with positive definite matrices and it indicates how Theorem 4 should be extended from linear programming to semidefinite programming.

Consider the matroid of rank $n-d-1$ on the ground set $\{1, 2, \dots, n\}$ defined by the linear subspace $\mathcal{L}_{A,\mathbf{c}}$ of K^n. Let $\beta(A)$ denote the *β-invariant* of that matroid. The β-invariant is a non-negative integer that can be described geometrically as follows. Consider the $(d-1)$-dimensional affine-linear subspace of K^n obtained by intersecting $\{A\mathbf{x} = b\}$ with a general hyperplane $\mathbf{c} \cdot \mathbf{x} = c_0$. Consider that $(n-d-1)$-dimensional subspace consider the arrangement of n hyperplanes obtained by intersecting with $\{x_i = 0\}$ for $i = 1, 2, \dots, n$. Then $\beta(A)$ is the number of bounded regions in that hyperplane arrangement. Note that $\beta(A)$ does not depend on $\mathbf{b}$ and $\mathbf{c}$, as these vectors are generic over $\mathbb{Q}$. In the example of Section 2 we have $n - d - 1 = 1$, the hyperplane arrangement consists of six points on a general line $\mathbf{c} \cdot \mathbf{x} = c_0$ passing through Figure 2, and the number of its bounded regions is $\beta(A) = 5$.

Theorem 4. *The degree of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$ equals the beta-invariant $\beta(A)$. Its prime ideal $J_{A,\mathbf{c}}$ is generated by a universal Gröbner basis consisting of all homogeneous polynomials*

$$\sum_{i \in \mathrm{supp}(v)} v_i \cdot \prod_{j \neq i} x_j, \tag{9}$$

where $\sum v_i x_i$ runs over non-zero linear forms of minimal support that vanish on $\mathcal{L}_{A,\mathbf{c}}$.

Proof. This theorem is due to Proudfoot and Speyer [4]; see also [6, Theorem 2]. □

The polynomials in (9) correspond to the circuits of the matroid. There is at most one circuit contained in each $(d+2)$-subset of $\{x_1, \ldots, x_n\}$, so their number is bounded above by $\binom{n}{d+2}$. If the matrix A is generic then that matroid is uniform and the beta-invariant equals

$$\beta(A) \;=\; \binom{n-1}{d}.$$

For arbitrary matrices A, this binomial coefficient always furnishes an upper bound.

Corollary 5. *The degree of the central path of (1) equals the beta-invariant $\beta(A)$ and is hence at most $\binom{n-1}{d}$. The prime ideal of polynomial that vanish on the central path is generated by the circuit polynomials (9) and d linear polynomials which are coordinates of $A \cdot \mathbf{x} - \mathbf{b}$.*

References

[1] D. Bayer and J. Lagarias: The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories. *Trans. Amer. Math. Soc.* **314** (1989) 499–526.

[2] D. Bayer and J. Lagarias: The nonlinear geometry of linear programming. II. Legendre transform coordinates and central trajectories. *Trans. Amer. Math. Soc.* **314** (1989) 527–581. 90C05 (34A99 52A40)

[3] A. Deza, T. Terlaky and Y. Zinchenko: A continuous d-step conjecture for polytopes, *Discrete Comput. Geom.* **41** (2009) 318–327.

[4] N. Proudfoot and D. Speyer: A broken circuit ring, *Beiträge zur Algebra und Geometrie* **47** (2005) 161–166.

[5] J. Renegar: *A Mathematical View of Interior-point Methods in Convex Optimization*, MPS/SIAM Series on Optimization, SIAM, Philadelphia, PA, 2001

[6] B. Sturmfels and C. Uhler: Multivariate Gaussians, semidefinite matrix completion, and convex algebraic geometry, *Annals of the Institute of Statistical Mathematics* **62** (2010) 603–638.

[7] H. Terao: Algebras generated by reciprocals of linear forms, *Journal of Algebra* **250** (2002) 549–558.

Departments of Mathematics, University of California, Davis and Berkeley