The Degree of Central Curve in Quadratic Programming

Serkan Hoşten and Dennis Schlief

Mathematics Department
San Francisco State University

15 October 2014

Quadratic Programs

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x^tQx + x^tc \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

where

- Q is $n \times n$ positive definite matrix
- $c \in \mathbb{R}^n$
- A is $d \times n$ matrix of rank d
- $b \in \mathbb{R}_+(A)$

Quadratic Programs

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^t Qx + x^t c \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

where

- Q is $n \times n$ positive definite matrix
- $c \in \mathbb{R}^n$
- A is $d \times n$ matrix of rank d
- $b \in \mathbb{R}_+(A)$

A convex optimization problem!

Log-barrier and a curve

A related convex problem to be solved for $\lambda \to 0$

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}x^t Qx + x^t c - \lambda \sum \log(x_i) \\ \text{subject to} & Ax = b \end{array}$$

Log-barrier and a curve

A related convex problem to be solved for $\lambda \to 0$

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}x^t Qx + x^t c - \lambda \sum \log(x_i) \\ \text{subject to} & Ax = b \end{array}$$

... and its Karush-Kuhn-Tucker (KKT) Equations:

$$Ax = b, \;\; x \geq 0,$$

$$Qx + c - \sum \frac{\lambda}{x_i} e_i - A^t y = 0$$

Log-barrier and a curve

A related convex problem to be solved for $\lambda \to 0$

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}x^tQx + x^tc - \lambda \sum \log(x_i) \\ \text{subject to} & Ax = b \end{array}$$

... and its Karush-Kuhn-Tucker (KKT) Equations:

$$Ax = b, \;\; x \geq 0,$$

$$Qx + c - \sum \frac{\lambda}{x_i} e_i - A^t y = 0$$

$\{(x^*(\lambda), \lambda, y^*(\lambda))\}$ for $\lambda > 0$ is a curve in $\mathbb{R}^{n+1+d}$.

Example

$$\begin{array}{lllllll} \text{minimize} & \frac{1}{2}(2x_1^2 & +3x_2^2 & +3x_3^2 & +4x_4^2) & & +4x_1 - x_2 + 3x_3 - 2x_4 \\ \text{subject to} & 2x_1 & -3x_2 & +x_3 & & = & 9 \\ & -x_1 & -2x_2 & & +x_4 & = & -6 \\ & x_1, & x_2, & x_3, & x_4 & \geq & 0 \end{array}$$

The Central Path

Definition

Assume $\{x \in \mathbb{R}^n : Ax = b, \ x \geq 0\}$ bounded. Then the projection of $\{(x^*(\lambda), \lambda, y^*(\lambda))$ for $\lambda > 0$ on $\mathbb{R}^n$ is called the central path of the quadratic program.

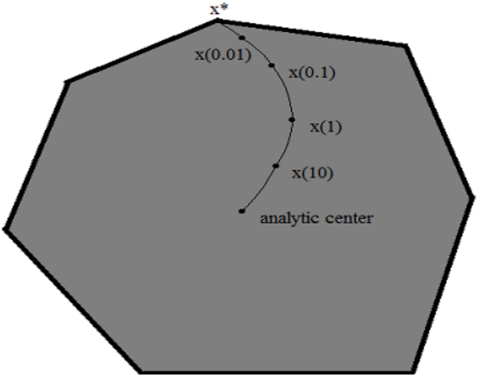
The Central Path

Definition

Assume $\{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$ bounded. Then the projection of $\{(x^*(\lambda), \lambda, y^*(\lambda))$ for $\lambda > 0$ on $\mathbb{R}^n$ is called the central path of the quadratic program.

$$\begin{array}{ccccccccccc}
2x_1 & & & & +4 & -\frac{\lambda}{x_1} & -2y_1 & +y_2 & = & 0 \\
 & 3x_2 & & & -1 & -\frac{\lambda}{x_2} & +3y_1 & +2y_2 & = & 0 \\
 & & 3x_3 & & +3 & -\frac{\lambda}{x_3} & -y_1 & & = & 0 \\
 & & & 4x_4 & -2 & -\frac{\lambda}{x_4} & & -y_2 & = & 0 \\
2x_1 & -3x_2 & +x_3 & & & & & & = & 9 \\
-x_1 & -2x_2 & & +x_4 & & & & & = & -6 \\
x_1 & x_2 & x_3 & x_4 & \geq & 0 & & & &
\end{array}$$

The Central Path



The Central Curve

Definition

The Zariski closure of the projection onto $\mathbb{C}^n$ of the solutions to

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \; i = 1, \ldots, n, \; \text{and } Ax - b = 0$$

is called the central curve of the quadradic program.

The Central Curve

Definition

The Zariski closure of the projection onto $\mathbb{C}^n$ of the solutions to

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \; i = 1, \ldots, n, \; \text{ and } Ax - b = 0$$

is called the central curve of the quadradic program.

$$\begin{array}{cccccccccc}
2x_1^2 & & & & +4x_1 & -\lambda & -2y_1x_1 & +y_2x_1 & = & 0 \\
 & 3x_2^2 & & & -x_2 & -\lambda & +3y_1x_2 & +2y_2x_2 & = & 0 \\
 & & 3x_3^2 & & +3x_3 & -\lambda & -y_1x_3 & & = & 0 \\
 & & & 4x_4^4 & -2x_4 & -\lambda & & -y_2x_4 & = & 0 \\
2x_1 & -3x_2 & +x_3 & & & & & & = & 9 \\
-x_1 & -2x_2 & & +x_4 & & & & & = & -6
\end{array}$$

The Central Curve

Definition

The Zariski closure of the projection onto $\mathbb{C}^n$ of the solutions to

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \;\; i = 1, \ldots, n, \;\; \text{and } Ax - b = 0$$

is called the central curve of the quadradic program.

The Central Curve

Definition

The Zariski closure of the projection onto $\mathbb{C}^n$ of the solutions to

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \; i = 1, \dots, n, \; \text{ and } Ax - b = 0$$

is called the central curve of the quadradic program.

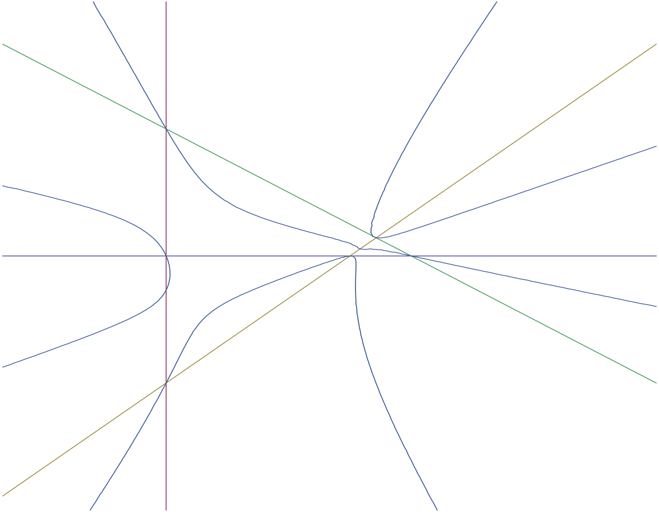
We denote the ideal generated by the above equations by J. And we denote the ideal of the central curve by $I_{\mathcal{C}}$.

The Central Curve

In our running example the central curve is defined by $Ax = b$ and

$$4x_1^2x_2x_3 - 3x_1x_2^2x_3 - 21x_1x_2x_3^2 + 6x_1^2x_2x_4 + 6x_1x_2^2x_4 + 2x_1^2x_3x_4 - 3x_2^2x_3x_4 - 6x_1x_3^2x_4 - 9x_2x_3^2x_4 + 28x_1x_2x_4^2 + 4x_1x_3x_4^2 - 8x_2x_3x_4^2 - 12x_1x_2x_3 - 4x_1x_2x_4 - 4x_1x_3x_4 - 4x_2x_3x_4 = 0$$

The Central Curve



The Degree of the Central Curve

Goal: Compute equations for and the degree (of the projective closure) of the central curve of quadratic programs (for generic Q, c, A, b).

The Degree of the Central Curve

Goal: Compute equations for and the degree (of the projective closure) of the central curve of quadratic programs (for generic Q, c, A, b).

Motivation:

- Work started by Bayer and Lagarias
- Dedieu, Malajovich, and Shub (2005) considered the total curvature of the central path for a given *linear program*
- De Loera, Sturmfels, and Vinzant (2012) determined the equations for the central curve of linear programs and computed the degree to be $\binom{n-1}{d}$. This implies a bound of $2\pi(n-d-1)\binom{n-1}{d-1}$ on the total curvature of the central curve of a generic LP
- Continuous Hirsch Conjecture is false: Allemigeon, Benchimol, Gaubert, and Joswig (2014)
- Quadratic programming is an intermediate stage from linear to semidefinite programs

Clearing Denominators

Proposition

$$\mathcal{V}(J : (x_1 x_2 \cdots x_n)^\infty) = \mathcal{V}(J)$$

Clearing Denominators

Proposition

$$\mathcal{V}(J : (x_1 x_2 \cdots x_n)^\infty) = \mathcal{V}(J)$$

- The central curve does not have components in coordinate hyperplanes.
- The LP central curve is an irreducible curve. The QP central curve *should* be irreducible as well.

Degree and the Solutions in the Torus

Theorem

When Q, A, c, and b are generic then the degree of the central curve is equal to the number of solutions in $(\mathbb{C}^*)^{n+d+1}$ to the system

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \ \ i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Degree and the Solutions in the Torus

Theorem

When Q, A, c, and b are generic then the degree of the central curve is equal to the number of solutions in $(\mathbb{C}^*)^{n+d+1}$ to the system

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0,\ \ i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Theorem

The number of solutions in $(\mathbb{C}^*)^{n+d+1}$ to the above system is equal to the number of solutions in $(\mathbb{C}^*)^{n+d+1}$ to the system when Q is generic and diagonal.

Computational Data for the Diagonal Case

d/n	3	4	5	6	7	8	9	10
1	3	7	15	31	63	127	255	511
2	1	4	11	26	57	120	247	502
3		1	5	16	42	99	219	466
4			1	6	22	64	163	382

Computing Mixed Volume: LP case

The equations:

$$c_i x_i - \lambda - (y^t A_i) x_i = 0, \;\; i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Computing Mixed Volume: LP case

The equations:

$$c_i x_i - \lambda - (y^t A_i) x_i = 0, \ \ i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Using the $(d+1)$ linear equations make a substitution $x = v_0 + t_1 v_1 + \ldots + t_{n-d-1} v_{n-d-1}$ and get a system of n equations in n unknowns: $t_1, \ldots, t_{n-d-1}, \lambda, y_1, \ldots, y_d$ where the support of each equation is

$$\lambda, 1, t_1, \ldots, t_{n-d-1}, y_1, y_1 t_1, \ldots, y_1 t_{n-d-1}, \ldots, y_d, y_d t_1, \ldots, y_d t_{n-d-1}$$

Computing Mixed Volume: LP case

So the mixed volume for the LP system is the volume of $\Delta_{n-d-1} \times \Delta_d$.

Computing Mixed Volume: LP case

So the mixed volume for the LP system is the volume of $\Delta_{n-d-1} \times \Delta_d$.

Theorem

The degree of the LP central curve for generic data is

$$\binom{n-1}{d} = \sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1}$$

Computing Mixed Volume: LP case

So the mixed volume for the LP system is the volume of $\Delta_{n-d-1} \times \Delta_d$.

Theorem

The degree of the LP central curve for generic data is

$$\binom{n-1}{d} = \sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1}$$

For the right hand side of the formula use, for instance, staircase triangulation.

Computing Mixed Volume: QP case

The equations:

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \;\; i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Computing Mixed Volume: QP case

The equations:

$$(Qx)_i x_i + c_i x_i - \lambda - (y^t A_i) x_i = 0, \; i = 1, \ldots, n,$$

$$Ax = b$$

$$ex = f$$

Using the $(d+1)$ linear equations make a substitution $x = v_0 + t_1 v_1 + \ldots + t_{n-d-1} v_{n-d-1}$ and get a system of n equations in n unknowns: $t_1, \ldots, t_{n-d-1}, \lambda, y_1, \ldots, y_d$ where the support of each equation is

$$\lambda, 1, t_1, \ldots, t_{n-d-1}, t_1^2, t_1 t_2, \ldots, t_{n-d-1}^2$$

$$y_1, y_1 t_1, \ldots, y_1 t_{n-d-1}, \ldots, y_d, y_d t_1, \ldots, y_d t_{n-d-1}$$

Computing Mixed Volume: QP case

Theorem

The degree of the QP central curve for generic data is

$$\sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1} 2^k$$

Computing Mixed Volume: QP case

Theorem

The degree of the QP central curve for generic data is

$$\sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1} 2^k$$

Again use staircase triangulation but with the right volumes of the simplices.

Equations

If a degree reverse lex $x_1 < x_2 < \cdots < x_n$ is used, the reduced Gröbner bases look like this ($d = 2, n = 5$):

$$26x_2^2x_3x_4 - 16x_2x_3^2x_4 - 44x_2x_3x_4^2 + 62x_2^2x_3x_5 + 16x_2x_3^2x_5 + \cdots$$
$$39x_1^2x_3x_4 - 42x_1x_3^2x_4 - 44x_1x_3x_4^2 + 93x_1^2x_3x_5 - 46x_1x_3^2x_5 + 96x_1^2x_4x_5 - \cdots$$
$$24x_1^2x_2x_4 - 42x_1x_2^2x_4 + 44x_1x_2x_4^2 - 24x_1^2x_2x_5 - 46x_1x_2^2x_5 + 96x_1^2x_4x_5 - \cdots$$
$$33x_1^2x_2x_3 - 22x_1x_2^2x_3 - 22x_1x_2x_3^2 + 24x_1^2x_2x_5 + 46x_1x_2^2x_5 + 93x_1^2x_3x_5 - \cdots$$

Equations

If a degree reverse lex $x_1 < x_2 < \cdots < x_n$ is used, the reduced Gröbner bases look like this ($d = 2, n = 5$):

$$26x_2^2x_3x_4 - 16x_2x_3^2x_4 - 44x_2x_3x_4^2 + 62x_2^2x_3x_5 + 16x_2x_3^2x_5 + \cdots$$
$$39x_1^2x_3x_4 - 42x_1x_3^2x_4 - 44x_1x_3x_4^2 + 93x_1^2x_3x_5 - 46x_1x_3^2x_5 + 96x_1^2x_4x_5 - \cdots$$
$$24x_1^2x_2x_4 - 42x_1x_2^2x_4 + 44x_1x_2x_4^2 - 24x_1^2x_2x_5 - 46x_1x_2^2x_5 + 96x_1^2x_4x_5 - \cdots$$
$$33x_1^2x_2x_3 - 22x_1x_2^2x_3 - 22x_1x_2x_3^2 + 24x_1^2x_2x_5 + 46x_1x_2^2x_5 + 93x_1^2x_3x_5 - \cdots$$

Initial ideal $= \langle x_2^2x_3x_4,\ x_1^2x_3x_4,\ x_1^2x_2x_4,\ x_1^2x_2x_3 \rangle =$
$\langle x_3, x_4\rangle \cap \langle x_2, x_4\rangle \cap \langle x_1^2, x_4\rangle \cap \langle x_2, x_3\rangle \cap \langle x_1^2, x_3\rangle \cap \langle x_1^2, x_2^2\rangle$.

Elimination in the Quadratic Case

In order to eliminate λ and $y_1, \ldots, y_d$ from

$$d_i x_i^2 + c_i x_i - \lambda - (y^t A_i) x_i = 0, \quad i = 1, \ldots, n$$

eliminate $y_1, \ldots, y_d$ and $z_1, \ldots, z_n$ from

$$d_i(w_i + z_i) = y^t A_i - c_i, \quad i = 1, \ldots, n \text{ and } d_1 w_1 z_1 = d_2 w_2 z_2 = \cdots = d_n w_n z_n$$

Elimination in the Quadratic Case

In order to eliminate λ and $y_1, \ldots, y_d$ from

$$d_i x_i^2 + c_i x_i - \lambda - (y^t A_i) x_i = 0, \quad i = 1, \ldots, n$$

eliminate $y_1, \ldots, y_d$ and $z_1, \ldots, z_n$ from

$$d_i(w_i + z_i) = y^t A_i - c_i, \quad i = 1, \ldots, n \text{ and } d_1 w_1 z_1 = d_2 w_2 z_2 = \cdots = d_n w_n z_n$$

- Use each *circuit* C of A to eliminate y from the first set of equations, then use $z_k = (d_1 w_1 z_1)/(d_k w_k)$ to write the resulting equation as $f_C(w_1, \ldots, w_n, z_1)$.
- each f_C is linear in z_1
- now using two carefully chosen C and C' eliminate z_1 to obtain $g_{C,C'}(w_1, \ldots, w_n)$ in the elimination ideal.

Elimination in the Quadratic Case

Theorem

The ideal $J = \langle g_{C,C'}(w_1, \ldots, w_n)\rangle$ is contained in the elimination ideal I and $M = \langle \mathrm{in}_<(g_{C,C'})\rangle$ is contained in $\mathrm{in}_<(I)$. The ideal M is equal to

$$M = \langle w_i^2 w_{j_1} w_{j_2} \ldots w_{j_d} \ : \ 1 \leq i \leq n-d-1, \ \ i < j_1 < j_2 < \cdots < j_d < n\rangle.$$

Morever, M has the irreducible decomposition

$$M = \cap_{k=0}^{n-d-1} \cap_{T \subset \{k+2,\ldots,n\}, |T|=n-d-k-1} \langle w_j^2 \ : \ 0 < j < k+1\rangle + \langle w_t \ : \ t \in T\rangle$$

Equations of the Central Curve in the Diagonal Case

Corollary

The degree of M is equal to

$$\sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1} 2^k$$

Equations of the Central Curve in the Diagonal Case

Corollary

The degree of M is equal to

$$\sum_{k=0}^{n-d-1} \binom{n-2-k}{d-1} 2^k$$

Theorem

For generic Q, A, b, c and Q a diagonal matrix the central curve for quadratic programming is defined by $J = \langle g_{C,C'}(w_1, \ldots, w_n) \rangle$