Solving LPs, QPs, and LCPs

1. The Linear Complementarity Problem

The KKT conditions for quadratic programming yield an instance of a more general class of problems called *linear complementarity problems.* In order to see this connection, consider the quadratic program

$$\begin{array}{lll} \mathcal{Q} & \text{minimize} & \frac{1}{2}u^TQu - c^Tu \\ & \text{subject to} & Au \leq b, \ 0 \leq u, \end{array}$$

where $Q \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$. Define

$$M = \begin{pmatrix} Q & A^T \\ -A & 0 \end{pmatrix} \quad \text{and} \quad q = \begin{pmatrix} c \\ b \end{pmatrix}. \tag{1}$$

Then the KKT conditions for the quadratic program $\mathcal{Q}$ are equivalent to the conditions

$$y = Mx + q, \quad y^Tx = 0, \quad 0 \leq x, \quad \text{and} \quad 0 \leq y \ ,$$

where

$$x = \begin{pmatrix} u \\ v \end{pmatrix}.$$

The Linear Complementarity Problem: (LCP)
Given $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$, find x and y in $\mathbb{R}^n$ satisfying

$$y = Mx + q, \quad x^Ty = 0, \quad 0 \leq x, \quad \text{and} \quad 0 \leq y.$$

In the context of quadratic programming, we know that solving the associated LCP will solve the $\mathcal{Q}$ if the matrix Q is positive semi-definite. If this is the case, then it is easily seen that the matrix M as defined in (1) is also positive semi-definite but **not** symmetric. This gives rise to a special class of LCPs. The problem LCP is said to be a *monotone* linear complementarity problem if the matrix M is positive semi-definite. In the remainder of this section, we briefly discuss a class of methods, called *interior point methods*, for solving the monotone LCP. Note that this algorithm can be used not only to solve quadratic programs when Q is positive semi-definite, but can also be used to solve linear programs ($Q = 0$).

In these notes we will always denote the set of solutions to the problem LCP by

$$\mathcal{S} = \{(x, y) \, : \, y = Mx + q, \ 0 \leq x, \ 0 \leq y, \ x^Ty = 0\} \ .$$

Before developing an algorithm for locating points in the set $\mathcal{S}$, we study conditions under which $\mathcal{S}$ is nonempty and bounded.

2. Boundedness Properties of LCP

Monotone LCP's are naturally associated with the following quadratic program:

$$\text{(QP-LCP)} \qquad \begin{array}{l} \min x^T(Mx+q) \\ \text{subject to } 0 \le Mx+q,\ 0 \le x. \end{array}$$

Observe that this QP must always have a finite optimal value whenever it is feasible and this optimal value is non-negative. Indeed, the optimal value is zero precisely when $\mathcal{S} \neq \emptyset$ in which case $\mathcal{S}$ is also the set of solutions to QP-LCP.

In our study of (LCP) we also need to consider properties of the following sets:

$$\begin{array}{rcl} \mathcal{A} & := & \{(x,y) \in I\!R^n \times I\!R^n : Mx+q=y\} \\ \mathcal{F} & := & \{(x,y) \in I\!R^n \times I\!R^n : Mx+q=y, 0 \le x, 0 \le y\} \\ \mathcal{F}_+ & := & \{(x,y) \in I\!R^n \times I\!R^n : Mx+q=y, 0 < x, 0 < y\} \end{array}$$

and

$$\mathcal{F}(t) \ := \ \{(x,y) \in I\!R^n \times I\!R^n : Mx+q=y, 0 \le x, 0 \le y, x^Ty \le t\}.$$

The boundedness of the sets $\mathcal{F}(t)$ and $\mathcal{S}$ are related to the condition that $\mathcal{F}_+$ is nonempty.

Theorem 2.1. *If M is positive semi–definite and $\mathcal{F}_+$ is nonempty, then $\mathcal{F}(t)$ is bounded for all $t \ge 0$.*

Proof. Let $(\bar{x},\bar{y}) \in \mathcal{F}_+$ and let $(x,y) \in \mathcal{F}(t)$. Then $(x-\bar{x})^T(y-\bar{y}) \ge 0$ since M is positive semi–definite. Therefore,

$$t + \bar{x}^T\bar{y} \ge x^Ty + \bar{x}^T\bar{y} \ge \bar{x}^Ty + \bar{y}^Tx \ge \kappa\|(x,y)\|_1 \ ,$$

where $\kappa := \min_{1=1,2,\ldots,n}\{\bar{x}_i,\bar{y}_i\}$. □

3. The Central Path

Given a vector $x \in I\!R^n$ we denote by X the diagonal matrix diag(x). Hence $Y = \text{diag}(y)$, $U = \text{diag}(u)$, $W = \text{diag}(w)$, etc. Consider the function

$$F(x,y) = \left[\begin{array}{c} Mx - y + q \\ XYe \end{array} \right]$$

where $e \in I\!R^n$ is the vector of all ones. Clearly, $(x,y) \in I\!R^n \times I\!R^n$ solves (LCP) if and only if $0 \le x, y$, and $F(x,y) = 0$. The basic idea behind interior point algorithms for solving (LCP) is to apply a damped Newton's method to the function $F(x, y)$ on the interior of the cone $I\!R^n_+ \times I\!R^n_+$. In order to apply Newton's method to F we need to know when its derivative is non-singular. Here

$$F'(x, y) = \begin{bmatrix} M & -I \\ Y & X \end{bmatrix} .$$

In this regard, following result is key.

Theorem 3.1. *If M is positive semi-definite, then $F'(x, y)$ is non-singular whenever $0 < x,\ 0 < y$.*

Proof. Let $(x, y) \in \mathrm{int}(I\!R^n \times I\!R^n)$ and suppose that $F'(x, y) \begin{pmatrix} u \\ v \end{pmatrix} = 0$. Then

$$v = Mu \text{ and } v = -X^{-1}Yu.$$

Hence, $0 \geq -u^T X^{-1} Y u = u^T M u \geq 0$, so $u^T X^{-1} Y u = 0$ or $u = 0$. But then $v = 0$ as well. $\square$

Thus, the Newton step is well defined at points in $\mathrm{int}(I\!R^n_+ \times I\!R^n_+)$. Moreover, one can always choose a step length so that a damped Newton step stays in $\mathrm{int}(I\!R^n_+ \times I\!R^n_+)$. However, it may happen that the iterates approach the boundary of $I\!R^n_+ \times I\!R^n_+$ too quickly and the procedure gets bogged down. For this reason we introduce the notion of a central path.

Definition 3.2. *The set*

$$\mathcal{C} := \{(x, y) \in \mathcal{F} : XYe = te \text{ for some } t > 0\}$$

is called the central path for (LCP).

We now proceed to show that if $\mathcal{F}_+ \neq \emptyset$ and M is positive semi-definite, then $\mathcal{C}$ exists. The first step is to establish the following lemma concerning the function

$$u(x, y) = XYe.$$

Lemma 3.3. *Suppose M is positive semi–definite and $\mathcal{F}_+ \neq \emptyset$.*

(1) *The system*

$$u(x, y) = a \text{ and } (x, y) \in \mathcal{F}_+$$

has a solution for every $a > 0$.

(2) *The mapping $u : I\!R^n \times I\!R^n \to I\!R^n$ is diffeomorphism between $\mathcal{F}_+$ and $int(I\!R^n_+)$, i.e. u is a one–to–one surjective mapping between $\mathcal{F}_+$ and $int(I\!R^n_+)$ with $u \in C^\infty$ on $\mathcal{F}_+$ and $u^{-1} \in C^\infty$ on $int(I\!R^n_+)$.*

Proof. (1) Let $a > 0$ and $(\overline{x}, \overline{y}) \in \mathcal{F}_+$. Set $\overline{a} = u(\overline{x}, \overline{y})$. Consider the function

$$\widehat{F}(x, y, t) := F(x, y) - \begin{bmatrix} 0 \\ (1-t)\overline{a} + ta \end{bmatrix}.$$

Note that $\widehat{F}(\overline{x}, \overline{y}, 0) = 0$ and

$$\nabla_{(x,y)} \widehat{F}(x, y, t) = \nabla F(x, y) = \begin{bmatrix} M & -1 \\ Y & X \end{bmatrix}.$$

Hence, by the implicit function theorem, there is an open neighborhood $U \subset I\!R^n \times I\!R^n$ containing $(\overline{x}, \overline{y})$, $\delta > 0$, and a unique smooth mapping $t \mapsto (x(t), y(t))$ on $[0, \delta)$ such that

$$(x(t), y(t)) \in U \text{ and } \widehat{F}(x(t), y(t)) = 0 \text{ on } [0, \delta).$$

Let $\overline{\delta}$ be the largest such δ in $[0, 1]$. We claim that $\overline{\delta} = 1$. First observe that $(x(t), y(t)) \in \mathcal{F}(\overline{t})$ for $\overline{t} := \max\{\overline{a}^T e, a^T e\}$. Moreover, $\mathcal{F}(\overline{t})$ is a compact set by Theorem 2.1. Hence, for some sequence $t_i \uparrow \overline{\delta}$ there exists an $(\widehat{x}, \widehat{y})$ such that $(x, (t_i), y(t_i)) \to (\widehat{x}, \widehat{y})$. Clearly, $(\widehat{x}, \widehat{y}) \in \mathcal{F}_{++}$. Applying the implicit function theorem again at $(\widehat{x}, \widehat{y})$ yields a contradiction to the maximality of $\overline{\delta}$. Finally, observe that

$$F(x(1), y(1)) = a$$

which establishes the result.

(2) In Part (1) above, we have already shown that u is a surjective map from $\mathcal{F}_+$ to $I\!R^n_+$. We now show that it is one–to–one. Assume to the contrary, that $u(x^1, y^1) = u(x^2, y^2)$ for distinct points (x^1, y^1) and (x^2, y^2) in $\mathcal{F}_+$. Then

$$M(x^1 - x^2) = y^1 - y^2 \text{ and } x_i^1 y_i^1 = x_i^2 y_i^2 > 0 \; \forall \; i = 1, \dots, n.$$

Since $(x^1 - x^2)^T M (x^1 - x^2) \geq 0$, we have

$$(x^1 - x^2)^T (y^1 - y^2) \geq 0.$$

Hence for some i with $x_i^1 \neq x_i^2$ we must have $(x_i^1 - x_i^2)(y_i^1 - y_i^2) \geq 0$. If $x_i^1 > x_i^2$, then $y_i^1 \geq y_i^2 > 0$. But then $x_i^1 y_i^1 \neq x_i^2 y_i^2$. Similarly, if $x_i^1 < x_i^2$, then $0 < y_i^1 \leq y_i^2$, so again $x_i^1 y_i^1 \neq x_i^2 y_i^2$. This contradiction establishes that u is one–to–one.

Finally, it is clear that u is C^∞. To see that u^{-1} is C^∞ simply note that $(u^{-1})'(a) = [XM + Y]^{-1}$ where $u^{-1}(a) = (x, y)$. To see that $[XM + Y]^{-1}$ exists write $[XM + Y] = X[M + X^{-1}Y]$ where both X and $[M + X^{-1}Y]$ are positive definite matrices.

$\square$

An immediate consequence of this Lemma is the following existence theorem for (LCP).

Theorem 3.4. *If M is positive semi-definite and $\mathcal{F}_+ \neq \emptyset$, then there exists a solution to (LCP).*

Proof. Let $(\overline{x}, \overline{y}) \in \mathcal{F}_+$. Then $\mathcal{F}(\overline{x}^T\overline{y})$ is compact by Theorem 2.1. Moreover, the system $F(x,y) = \begin{bmatrix} 0 \\ \mu\overline{x}^T\overline{y}e \end{bmatrix}$ is solvable for all $\mu \in (0,1]$. Hence there exist $\{(x_i, y_i)\} \subset \mathcal{F}_+$, $\mu_i \downarrow 0$, and $(\widehat{x}, \widehat{y}) \in \mathcal{F}$ such that $(x_i, y_i) \to (\widehat{x}, \widehat{y})$ and $F(x_i, y_i) = \begin{bmatrix} 0 \\ \mu_i\overline{x}^T\overline{y}e \end{bmatrix}$. But then $F(\widehat{x}, \widehat{y}) = 0$ so that $(\widehat{x}, \widehat{y}) \in \mathcal{S}$. □

The existence of the central path can now also be established. The proof is similar to the proof given for Part 2 of Lemma 3.3.

Theorem 3.5. *If M is positive semi-definite and $\mathcal{F}_+ \neq \emptyset$, then the central path $\mathcal{C}$ exists as a smooth curve in $\mathcal{F}_+$.*

Proof. By Parts 1 and 2 of Lemma 3.3, for each $t > 0$ there exist a unique $(x(t), y(t)) \in \mathcal{F}_+$ with $F(x(t), y(t)) = \begin{bmatrix} 0 \\ te \end{bmatrix}$. Define

$$\widehat{F}(x,y,t) = F(x,y) - \begin{bmatrix} 0 \\ te \end{bmatrix}.$$

Given a triple $(x(\overline{t}), y(\overline{t}), \overline{t})$ for $\overline{t} > 0$, we have that $\nabla_{(x,y)}\widehat{F}(x(\overline{t}), y(\overline{t})) = \nabla F(x(\overline{t}), y(\overline{t}))$ is nonsingular by Theorem 3.1. Hence, by the implicit function theorem, there exists $t_1 \in (0, \overline{t}]$ and $t_2 > \overline{t}$ such that the mapping $t \mapsto (x(t), y(t))$ is smooth on (t_1, t_2). Let $\overline{t}_1$ be the smallest such t_1 and $\overline{t}_2$ be the largest such t_2. Due to the compactness of $\mathcal{F}(t)$ for all $t \geq 0$, we obtain as in the proof of Part (1) of Lemma 3.3 that $\overline{t}_1 = 0$ and $\overline{t}_2 = +\infty$. □

4. Asymptotic Behavior of the Central Path

In this section we study the limiting behavior of the central path as $t \downarrow 0$. In particular, we show that this limit exists and is a solution of (LCP). The key to this analysis is the potential function

$$P(x,y,t) = x^Ty - t\sum_{i=1}^{n} \ln(x_i y_i)$$

defined over the set $\mathcal{F}_+ \times \{t > 0\}$. Let us first observe that for fixed $t > 0$ the function $P(\cdot, \cdot, t)$ is strictly convex on $\mathcal{F}_+$. In order to see this observe that

$$\nabla^2_{(x,y)} P(x,y,t) = \begin{bmatrix} tX^{-2} & I \\ I & tY^{-2} \end{bmatrix}.$$

Hence if (x_1, y_1), $(x_2, y_2) \in \mathcal{F}_+$, then

$$\begin{aligned} &\begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \end{bmatrix}^T \nabla^2_{(x,y)} P(x,y,t) \begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \end{bmatrix} \\ &= t[(x_1 - x_2)^T X^{-2}(x_1 - x_2) + (y_1 - y_2)^T Y^{-2}(y_1 - y_2)] \\ &\qquad + 2(x_1 - x_2)^T (y_1 - y_2) \\ &> 0. \end{aligned}$$

Therefore, for each $t > 0$, the solution to the problem

$$(\mathrm{P}_t) \quad \begin{array}{l} \min P(x,y,t) \\ \text{subject to } (x,y) \in \mathcal{F}_+, \end{array}$$

if it exists, is unique. With this in mind we give the following theorem.

Theorem 4.1. *If M is positive semi-definite and $\mathcal{F}_+ \neq \emptyset$, then the unique solution to the problem (P_t) exists and corresponds to the unique solution of the equation $F(x,y) = \begin{bmatrix} 0 \\ te \end{bmatrix}$, i.e., it lies on the central path.*

Proof. Due to our observation concerning the strict convexity of $P(x,y,t)$, we need only show that the unique solution, $(x(t), y(t))$, to $F(x,y) = \begin{bmatrix} 0 \\ te \end{bmatrix}$ satisfies the first-order optimality conditions for (P_t). The first-order conditions for (P_t) are

$$\nabla_{(x,y)} P(x,y,t) \in \ker \begin{bmatrix} M & -I \end{bmatrix}^\perp = \operatorname{Ran} \begin{bmatrix} M \\ -I \end{bmatrix}.$$

Since $\nabla_{(x,y)} P(x,y,t) = \begin{bmatrix} y - tx^{-1} \\ x - ty^{-1} \end{bmatrix}$, where $(x^{-1})_i := (x_i)^{-1}$, these conditions imply the existence of a vector $v \in I\!\!R^n$ such that

$$\begin{aligned} y - tx^{-1} &= Mv \\ x - ty^{-1} &= -v. \end{aligned}$$

Multiplying the first of these equations by X and the second by Y, we get the system

$$\begin{aligned} Xy - te &= XMv \\ Yx - te &= -Yv. \end{aligned}$$

Therefore, $[XM + Y]v = 0$, or equivalently, $[M + X^{-1}Y]v = 0$. But $[M + X^{-1}Y]$ is a positive definite matrix so we must have $v = 0$. Consequently, the conditions $Xy = te$, $0 < x$, $0 < y$, and $Mx + q = y$, are equivalent to the first-order necessary and sufficient conditions in (P_t). The unique solution of this system is $(x(t), y(t))$ so this is the unique solution to (P_t). □

Next set

$$\begin{aligned} E &= \{i : x_i = 0 = y_i \text{ for all } (x,y) \in \mathcal{S}\}, \\ B &= \{i : x_i \neq 0 \text{ for some } (x,y) \in \mathcal{S}\}, \text{ and} \\ N &= \{i : y_i \neq 0 \text{ for some } (x,y) \in \mathcal{S}\}. \end{aligned}$$

We make the following observations about these index sets:

1. Since $\mathcal{S}$ is convex, there exists $(\widehat{x}, \widehat{y}) \in \mathcal{S}$ with

$$\begin{aligned} \widehat{x}_i > 0 \quad \forall \quad i \in B, \text{ and} \\ \widehat{y}_i > 0 \quad \forall \quad i \in N. \end{aligned}$$

   To obtain $(\widehat{x}, \widehat{y})$ just take a convex combination of points (x, y) for which $x_i > 0$ $i \in B$ and $y_i > 0$ for $i \in N$.
2. Due to the above observation we have $B \cap N = \emptyset$. This implies that the sets B, E, and N form a partition of the integers from 1 to n, i.e. $\{1, 2, \ldots, n\} = B \cup N \cup E$ with $B \cap N = \emptyset$, $B \cap E = \emptyset$, and $N \cap E = \emptyset$.
3. For all $(x, y) \in \mathcal{S}$ we have $x_i = 0$ for all $i \in \{1, 2, \ldots, n\} \setminus \bar{B}$ and $y_i = 0$ for all $i \in \{1, 2, \ldots, n\} \setminus \bar{N}$, where $\bar{B} = B \cup E$ and $\bar{N} = N \cup E$.
4. The solution set $\mathcal{S}$ has the representation

$$\mathcal{S} = \left\{ (x, y) \;\middle|\; \begin{array}{ll} 0 \le x_B, & 0 \le y_N, \\ 0 = x_{\overline{N}}, & 0 = y_{\overline{B}}, \end{array} \; M_B x_B + q = y_N \right\}, \tag{2}$$

We claim that the limit as $t \searrow 0$ in the central path is the unique solution to the problem

$$(\mathrm{P}_0) \quad \begin{array}{l} \min - \left[\sum_B \ln x_i + \sum_N \ln y_i\right] \\ \text{subject to } (x, y) \in \mathcal{S}. \end{array}$$

One can view the problem (P_0) as the limit of the problems (P_t) as $t \searrow 0$. Observe that

$$-\sum_B \ln x_i - \sum_N \ln y_i = -\ln\left[\left(\prod_B x_i\right)\left(\prod_N y_i\right)\right].$$

Therefore, since $-\ln(\mu)$ is strictly decreasing for $\mu > 0$, minimizing $-\ln\left[\left(\prod_B x_i\right)\left(\prod_N y_i\right)\right]$ over $\mathcal{S}$ is the same as maximizing $\left(\prod_B x_i\right)\left(\prod_N y_i\right)$ over $\mathcal{S}$. That is, (P_0) is equivalent to the problem

$$(3) \qquad (\hat{P}_0) \quad \max \left(\prod_B x_i\right)\left(\prod_N y_i\right) \quad \text{subject to } (x, y) \in \mathcal{S}.$$

Using this fact we can show that the problem (P_0) has a solution and that it is unique.

Lemma 4.2. *If M is positive semi-definite and $\mathcal{F}_+ \neq \emptyset$, then the solution (x^*, y^*) to (P_0) exists, is unique, and satisfies $x^*_B > 0$ and $y^*_N > 0$.*

Proof. By Theorem 2.1, $\mathcal{S}$ is a compact set. Hence the solution to $(\hat{P}_0)$, or equivalently (P_0), exists since problem $(\hat{P}_0)$ is the maximization of a continuous function over a compact set. The fact that the solution is unique is the consequence of the fact that the objective function in (P_0) is strictly convex on $\mathcal{S}$ as seen by considering the representation (2). The condition that the solution (x^*, y^*) satisfies $x^*_B > 0$ and $y^*_N > 0$ follows from the finiteness of the optimal value. □

Before proving the main result, we first establish the following technical lemma.

Lemma 4.3. *Let $(x^*, y^*) \in \mathcal{S}$ be the unique solution to (P_0), let $(x, y) \in \mathcal{C}$, and set $\mu = x^T y/n$. Then $XYe = \mu e$, $\sum_B \frac{x^*_i}{x_i} + \sum_N \frac{y^*_i}{y_i} \leq n$, $x_B \geq \frac{1}{n} x^*_B > 0$, and $y_N \geq \frac{1}{n} y^*_N > 0$.*

Proof. As usual,

$$\begin{aligned} 0 &\leq (x - x^*)^T (y - y^*) \\ &= x^T y - x^{*T} y - x^T y^* + x^{*T} y^*, \end{aligned}$$

so

$$x^{*T} y + x^T y^* \leq x^T y = n\mu.$$

Since $(x, y) \in \mathcal{C}$, we have $XYe = \mu e$ so

$$x = \mu y^{-1} \text{ and } y = \mu x^{-1}.$$

But then

$$\begin{aligned} \mu({x^*}^T x^{-1} + {y^*}^T y^{-1}) &\leq x^{*T} y + y^{*T} x \\ &= \mu n, \end{aligned}$$

or equivalently,

$$\sum_B \frac{x_i^*}{x_i} + \sum_N \frac{y_i^*}{y_i} \leq n.$$

Due to the positivity of each term in the sum, we get that

$$\frac{x_i^*}{x_i} \leq n \quad \text{for } i \in B \quad \text{and} \quad \frac{y_i^*}{y_i} \leq n \quad \text{for } i \in N,$$

or equivalently,

$$\frac{1}{n} x_B^* \leq x_B \quad \text{and} \quad \frac{1}{n} y_N^* \leq y_N.$$

□

Theorem 4.4. *Let M be positive semi-definite, $\mathcal{F}_+ \neq \emptyset$, and assume that $E = \emptyset$. Then the limit of the central path $\mathcal{C}$ exists as $t \downarrow 0$ and is the solution to the problem (P_0).*

Proof. Let $(\widehat{x}, \widehat{y})$ be any cluster point of $\mathcal{C}$ as $t \downarrow 0$. Since $(\widehat{x}, \widehat{y}) \in \mathcal{S}$, we have that $\widehat{x}_{\overline{N}} = 0$ and $\widehat{y}_{\overline{B}} = 0$. Since this is true for every cluster point, we obtain that $x_{\overline{N}}(t) \to 0$ and $y_{\overline{B}}(t) \to 0$.

Letting (x^*, y^*) be the unique solution to (P_0) and taking the limit as $t \searrow 0$, we obtain from Lemma 4.3 that

$$\sum_B \frac{x_i^*}{\hat{x}_i} + \sum_N \frac{y_i^*}{\hat{y}_i} \leq n, \tag{4}$$

$\hat{x}_B \geq \frac{1-\beta}{n} x_B^* > 0$, and $\hat{y}_N \geq \frac{1-\beta}{n} y_N^* > 0$. Thus, in particular, $(\hat{x}, \hat{y})$ is feasible for (P_0).

Next recall that the arithmetic–geometric mean inequality says that for any collection $\{\gamma_1, \gamma_2, \ldots, \gamma_N\}$ of non–negative real numbers we have that

$$\left(\prod_{i=1}^N \gamma_i \right)^{1/N} \leq \frac{1}{N} \sum_{i=1}^N \gamma_i \ .$$

Therefore, by (4) and the fact that $B \cup N = \{1, 2, \ldots, n\}$, we have

$$\left(\prod_B \frac{x_i^*}{\hat{x}_i} \prod_N \frac{y_i^*}{\hat{y}_i} \right) \leq \left(\frac{1}{n} \sum_B \frac{x_i^*}{\hat{x}_i} + \sum_N \frac{y_i^*}{\hat{y}_i} \right)^n \leq 1^n = 1 \ .$$

Consequently,

$$\begin{aligned}\left(\prod_B x_i^* \prod_N y_i^*\right) &= \left(\prod_B \hat{x}_i \prod_N \hat{y}_i\right)\left(\prod_B \frac{x_i^*}{\hat{x}_i} \prod_N \frac{y_i^*}{\hat{y}_i}\right) \\ &\leq \left(\prod_B \hat{x}_i \prod_N \hat{y}_i\right).\end{aligned}$$

But then $(\widehat{x}, \widehat{y})$ must also be a solution to $(\hat{P}_0)$ in which case $(\widehat{x}, \widehat{y}) = (x^*, y^*)$ by uniqueness. Since (x^*, y^*) is the only possible cluster point, it must be the case that the limit of the central path is (x^*, y^*). □

5. An Infeasible Interior Point Algorithm

We now build a numerical procedure for solving the monotone LCP based on the theoretical observations of the previous sections. The basic idea is to try to follow the central path to the solution. In order to do this the algorithm must be constructed so that it stays close to the central path while reducing the *homotopy* parameter t at each iteration. Then as t is reduced to zero we hopefully converge to a solution. There are several obstacles that must be overcome for this strategy to succeed. The most obvious and significant of these is that it is very difficult to locate points in the set $\mathcal{F}_+$ let alone points on the central path. For this reason we consider algorithms that initialize at points satisfying $0 < x$ and $0 < y$ but for which the affine constraint $Mx + q = y$ may be violated. Algorithms of this type are called *infeasible* interior point algorithms.

Infeasible interior point algorithms must balance reduction in the homotopy parameter t with reduction in the residual of the affine constraints $Mx + q = y$. Indeed, the overall success of the procedure depends on how this balance in achieved. In general, one must reduce these two quantities at roughly the same rate while simultaneously staying sufficiently close to the central path. An algorithm that attempts to achieve this balance is given below.

Infeasible Interior Point Algorithm for LCP

Initialization:

$$\begin{array}{lll}
\epsilon & = 10^{-8} & \left(\begin{array}{c}\text{stopping}\\ \text{tolerance}\end{array}\right)\\
\sigma & = 0 & \left(\begin{array}{c}\text{homotopy}\\ \text{scaling parameter}\end{array}\right)\\
x^0 & = 2e & (\text{initial } x)\\
(y^0)_i & = \min\{\,(Mx^0+q)_i,\ 2\},\ i=1,2,\ldots,n & (\text{initial } y)\\
\tau & = (x^0)^T y^0/n & \left(\begin{array}{c}\text{homotopy}\\ \text{parameter}\end{array}\right)\\
\rho & = \|Mx^0 - y^0 + q\|_\infty & (\text{residual})
\end{array}$$

Iteration: While $n\tau > \epsilon$ or $\rho > \epsilon$,

Step 1: (Compute the Newton Step)

Solve the linear equation

$$F(x^k, y^k) + F'(x^k, y^k)\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix} = \begin{pmatrix}0\\ \sigma\tau e\end{pmatrix},$$

or equivalently, solve the equation

$$\begin{bmatrix} M & -I\\ Y & X\end{bmatrix}\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix} = \begin{pmatrix}-Mx^k + y^k - q\\ \sigma\tau e - X_k Y_k e\end{pmatrix},$$

for Δx and Δy.

Step 2: (Compute a Feasible Steplength)

$$\begin{array}{ll}
t_x & = \min\left\{\frac{-(x^k)_i}{(\Delta x)_i}\ :\ (\Delta x)_i < 0\right\}\\
t_x & = \min\{\,1,\ 0.999 t_x\}\\[2ex]
t_y & = \min\left\{\frac{-(y^k)_i}{(\Delta y)_i}\ :\ (\Delta y)_i < 0\right\}\\
t_y & = \min\{\,1,\ 0.999 t_y\}
\end{array}$$

Step 3: (Update Iterates)

$$\begin{array}{lcl}
x^{k+1} & = & x^k + t_x \Delta x\\
y^{k+1} & = & y^k + t_y \Delta y\\
k & = & k+1\\
\tau & = & (x^k)^T y^k\\
\rho & = & \|Mx^k - y^k + q\|_\infty
\end{array}$$

Step 4: (Update Scaling Parameter)

$$\sigma = \begin{cases} 1 & \text{, if } n\tau \leq \epsilon \text{ and } \rho > \epsilon, \\ \min\left\{.5,\ (1-t_x)^2,\ (1-t_y)^2,\ \frac{|\rho - n\tau|}{\rho + 10n\tau}\right\} & \text{, otherwise.} \end{cases}$$

Exercises

In all of the following exercises assume that the matrix $Q \in \mathbb{R}^{p\times p}$ is symmetric and positive semi-definite. Also assume that $c \in \mathbb{R}^p$, $A \in \mathbb{R}^{m\times p}$, $E \in \mathbb{R}^{k\times p}$, $d \in \mathbb{R}^k$, and $b \in \mathbb{R}^m$.

1. Consider the following QP:

$$\begin{array}{lll} \mathcal{Q}_1 & \text{minimize} & \frac{1}{2}u^TQu - c^tu \\ & \text{subject to} & Au = b\ . \end{array}$$

   Show that u solves $\mathcal{Q}_1$ if and only if there exists a vector $v \in \mathbb{R}^m$ such that $Mx = q$, where

$$M = \begin{bmatrix} Q & -A^T \\ A^T & 0 \end{bmatrix},\ x = \begin{pmatrix} u \\ v \end{pmatrix},\ \text{and } q = \begin{pmatrix} c \\ b \end{pmatrix}.$$

2. Assume that $\mathrm{Nul}\,(E^T) = \{0\}$ and consider the QP

$$\begin{array}{lll} \mathcal{Q}_2 & \text{minimize} & \frac{1}{2}u^TQu - c^tu \\ & \text{subject to} & Au \le b,\ Eu = d,\ 0 \le u\ . \end{array}$$

   Related to $\mathcal{Q}_2$ is the so-called *horizontal LCP*

   The Horizontal LCP (HLCP)

   Given $M \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{k\times n}$, $h \in \mathbb{R}^k$, and $q \in \mathbb{R}^n$, find $x \in \mathbb{R}^n$, $z \in \mathbb{R}^k$, and $y \in \mathbb{R}^n$ such that

$$Mx + G^Tz + q = y,\ Gx = h,\ 0 \le x,\ 0 \le y,\ \text{and } x^Ty = 0\ .$$

   (a) Write the KKT conditions for $\mathcal{Q}_2$.
   (b) Show that the KKT conditions for $\mathcal{Q}_2$ are an instance of the HLCP by specifying $M,\ G,\ h$, and q in terms of $Q,\ A,\ E,\ c,\ b$, and d.
   (c) Show that under this specification $\mathrm{Nul}\,(E^T) = \{0\}$ if and only if $\mathrm{Nul}\,(G^T) = \{0\}$.
   (d) Again, show that under this specification the positive semi-definiteness of Q implies that M is positive semi-definite.

3. Now consider the general HLCP introduced in problem 2 above. Define

$$F(x,z,y) = \begin{bmatrix} Mx + G^Tz - y + q \\ Gx - h \\ XYe \end{bmatrix}.$$

   (a) Show that (x,z,y) solves the HLCP if and only if $F(x,z,y) = 0$ and $0 \le x, y$.
   (b) Assume that M is positive semi-definite and $\mathrm{Nul}\,(G^T) = \{0\}$. Show that $F'(x,z,y)$ is nonsingular whenever $0 < x$ and $0 < y$.

(c) Define

$$\mathcal{F}_+ = \{(x, z, y) \,:\, Mx + G^T z + q = y,\ Gx = h,\ 0 < x,\ 0 < y\}.$$

Under the assumption that M is positive semi-definite, $\mathrm{Nul}\,(G^T) = \{0\}$, and $\mathcal{F}_+ \neq \emptyset$, show that the set

$$\mathcal{F}(t) = \{(x, z, y) \,:\, Mx + G^T z + q = y,\ Gx = h,\ 0 \leq x,\ 0 \leq y,\ x^T y \leq t\}$$

is compact for all $t \geq 0$.

(*Hint*: There are a number of ways to show this. The best way to start is to first show that for all $(x, z, y) \in \mathcal{F}(t)$ (x, y) is bounded in 1-norm. This is done using exactly the same kind of argument as is used in the LCP case. To show that the z component is also bounded there are again a number of possible arguments. The most direct argument uses the fact that the matrix GG^T is invertible (see the midterm exam).)

(d) What is the appropriate definition for the *central path* for HLCP?