Solutions to the End-of-Chapter Excercises in "Interior-Point Algorithm: Theory and Analysis"

Yinyu Ye

November 1997

Yinyu Ye, Professor
Department of Management Sciences
The University of Iowa
Iowa City, IA 52242-1000
Phone: (319) 335-1947
E-mail: yinyu-ye@uiowa.edu

Chapter 1

Introduction and Preliminaries

1.1 Exercises

1.1 *Let* $Q \in \mathcal{R}^{n \times n}$ *be a given nonsingular matrix, and* a *and* b *be given* $\mathcal{R}^n$ *vectors. Show*

$$(Q + ab^T)^{-1} = Q^{-1} - \frac{1}{1 + b^T Q^{-1} a} Q^{-1} ab^T Q^{-1}.$$

This formula is called the Sherman-Morrison-Woodbury formula.

Solution 1.1

$$\begin{aligned}
&(Q + ab^T)(Q^{-1} - \frac{1}{1 + b^T Q^{-1} a} Q^{-1} ab^T Q^{-1}) \\
= \;& I - \frac{1}{1 + b^T Q^{-1} a} ab^T Q^{-1} + ab^T Q^{-1} - \frac{1}{1 + b^T Q^{-1} a} ab^T Q^{-1} ab^T Q^{-1} \\
= \;& I - \frac{1}{1 + b^T Q^{-1} a} ab^T Q^{-1} + ab^T Q^{-1} - -\frac{b^T Q^{-1} a}{1 + b^T Q^{-1} a} ab^T Q^{-1} \\
= \;& I.
\end{aligned}$$

1.2 *Prove that the eigenvalues of all matrices* $Q \in \mathcal{M}^{n \times n}$ *are real. Furthermore, show that* Q *is PSD if and only if all its eigenvalues are nonnegative, and* Q *is PD if and only if all its eigenvalues are positive.*

Solution 1.2

1.3 *Using the ellipsoid representation in Section ??, find the matrix Q and vector y that describes the following ellipsoids:*

1. *the 3-dimensional sphere of radius 2 centered at the origin;*
2. *the 2-dimensional ellipsoid centered at $(1;2)$ that passes the points $(0;2)$, $(1;0)$, $(2;2)$, and $(1;4)$;*
3. *the 2-dimensional ellipsoid centered at $(1;2)$ with axes parallel to the line $y=x$ and $y=-x$, and passing through $(-1;0)$, $(3;4)$, $(0;3)$, and $(2;1)$.*

Solution 1.3 *1.*

$$Q = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}, \quad y = 0;$$

2.

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix}, \quad y = (1;2);$$

3.

$$Q = \begin{pmatrix} 5/16 & -3/16 \\ -3/16 & 5/16 \end{pmatrix}, \quad y = (1;2).$$

1.4 *Show that the biggest coordinate-aligned ellipsoid that is entirely contained in $\mathcal{R}^n_+$ and has its center at $x^a \in \overset{\circ}{\mathcal{R}}{}^n_+$ can be written as:*

$$E(x^a) = \{x \in \mathcal{R}^n : \ \|(X^a)^{-1}(x - x^a)\| \le 1\}.$$

Solution 1.4 *The coordinate-aligned ellipsoid that has its center at x^a is given by*

$$\{x \in \mathcal{R}^n : \ (x - x^a)^T Q (x - x^a) \le 1\},$$

where Q is a positive diagonal matrix. The inequality can be written as

$$\sum_{j=1}^{n} q_{jj}(x_j - x^a_j)^2 \le 1.$$

In order to have it entirely contained in $\mathcal{R}^n_+$, we must have $\sqrt{q_{jj}} \ge \frac{1}{x^a_j}$ for all $j = 1, \ldots, n$. Since the volume of the ellispod is $\prod_{j=1}^{n} \frac{1}{\sqrt{q_{jj}}}$, the biggest must be $\sqrt{q_{jj}} = \frac{1}{x^a_j}$ for $j = 1, \ldots, n$.

1.5 *Show that the non-negative orthant, the positive semi-definite cone, and the second-order cone are all self-dual.*

Solution 1.5 *For a cone $C \subset \mathcal{E}$, the dual of C is the cone*

$$C^* := \{y : \ \langle x, y\rangle \ge 0 \quad \text{for all} \quad x \in C\},$$

where $\langle \cdot, \cdot \rangle$ is an inner product operation for space $\mathcal{E}$.

For $C = \mathcal{R}^n_+$,

$$C^* := \{y : \ x^T y \ge 0 \quad \text{for all} \quad x \in \mathcal{R}^n_+\}.$$

In paticular, for any $y \in C^*$ we let $x = 0$ except $x_1 = 1$. Then, we must have $y_1 \ge 0$. Similary, we have $y_j \ge 0$ for all $j = 1, \ldots, n$. Thus, $C^* = C$.

For $C = \mathcal{M}^n_+$,

$$C^* := \{Y : \ X \bullet Y \ge 0 \quad \text{for all} \quad X \in \mathcal{M}^n_+\}.$$

In paticular, for any $Y \in C^*$ we let $X = vv^T$ where v, $\|v\| = 1$, is an eigenvector of Y with eigenvalue λ. Then, we have $\lambda = v^T Y v = X \bullet Y \ge 0$. Similary, we shall have all eigenvalues of Y nonnegative. Therefore, $C^* = C$.

For $C = \{(t; x) \in \mathcal{R}^{n+1} : \ t \ge \|x\|\}$,

$$C^* := \{(s; y) : \ (t; x)^T (s; y) \ge 0 \quad \text{for all} \quad (t; x), \quad t \ge \|x\|\}.$$

In paticular, for any $(s; y) \in C^*$ we first let $(t; x) = (1; 0)$ and have $s \ge 0$. Furthermore, let $(t; x) = (\|y\|; -y)$. Then, we must have $s\|y\| \ge \|y\|^2$ which implies $s \ge \|y\|$. Thus, $C^* = C$.

1.6 *Consider the convex set $C = \{x \in \mathcal{R}^2 : \ (x_1 - 1)^2 + (x_2 - 1)^2 \le 1\}$ and let $y \in \mathcal{R}^2$. Assuming $y \notin C$,*

1. *find the point in C that is closest to y;*

2. *find a separating hyperplane vector as a function of y.*

Solution 1.6 *1.*

$$x := e + \frac{1}{\|y - e\|}(y - e).$$

2.

$$a := x - y = \frac{\|y - e\| - 1}{\|y - e\|}(e - y),$$

Note that $\|y - e\| > 1$.

1.7 *Using the idea of Exercise 1.6, prove the separating hyperplane theorem* **??**.

Solution 1.7 *Let*

$$\delta = \min_{x \in C} \; \|y - x\| > 0$$

and $\hat{x} \in C$ *such that* $\|y - \hat{x}\| = \delta$. *Such a* $\hat{x}$ *exists since* $\|y - x\|$ *achieves its minimum over any closed and bounded set that can be the intersection of* C *and the ball of radius* $\delta + 1$ *centered at* y.

For any $x \in C$ *and* $0 \le \alpha \le 1$, *the point* $\hat{x} + \alpha(x - \hat{x}) \in C$ *and thus*

$$\|\hat{x} + \alpha(x - \hat{x}) - y\|^2 \ge \|\hat{x} - y\|^2.$$

Expanding,

$$2\alpha(\hat{x} - y)^T(x - \hat{x}) + \alpha^2\|x - \hat{x}\|^2 \ge 0.$$

Considering the inequality as $\alpha \to 0$, *we have*

$$(\hat{x} - y)^T(x - \hat{x}) \ge 0$$

or

$$(\hat{x} - y)^T x \ge (\hat{x} - y)^T \hat{x} = (\hat{x} - y)^T(y + \hat{x} - y) = (\hat{x} - y)^T y + \delta^2.$$

The theorem is proved after setting $a = \hat{x} - y$.

1.8 *Given an* $m \times n$ *matrix* A *and a vector* $c \in \mathcal{R}^n$, *consider the function* $\mathcal{B}(y) = \sum_{j=1}^n \log s_j$ *where* $s = c - A^T y > 0$. *Find* $\nabla\mathcal{B}(y)$ *and* $\nabla^2\mathcal{B}(y)$ *in terms of* s.

Solution 1.8

$$\nabla\mathcal{B}(y) = e^T S^{-1} A^T$$

and

$$\nabla^2\mathcal{B}(y) = A S^{-2} A^T,$$

where

$$S = diag(s).$$

1.9 *Prove that the level set of a quasi-convex function is convex.*

Solution 1.9 *For some* z *let the level set*

$$L(z) = \{x : \quad f(x) \le z\},$$

where $f(x)$ *is a quasi-convex function, i.e., for* $0 \le \alpha \le 1$

$$f(\alpha x + (1 - \alpha)y) \le \max[f(x), f(y)].$$

Suppose $x, y \in L(z)$*, then*

$$f(\alpha x + (1-\alpha)y) \leq \max[f(x), f(y)] \leq z$$

which implies that $\alpha x + (1-\alpha)y \in L(z)$ *and thus* $L(z)$ *is a convex set.*

1.10 *Prove Propositions* **??** *and* **??** *for convex functions in Section* **??**.

Solution 1.10 *Since* f *is a convex function over a convex set* Ω*, for any* $x, y \in \Omega$ *and* $0 < \alpha < 1$

$$f(x + \alpha(y-x)) \leq (1-\alpha)f(x) + \alpha f(y).$$

Thus

$$f(y) - f(x) \geq \frac{f(x + \alpha(y-x)) - f(x)}{\alpha}.$$

Furthermore, if $f \in C^1$*, then from the Taylor expansion theorem we have for some* $0 \leq \beta \leq 1$

$$f(x + \alpha(y-x)) = f(x) + \alpha \nabla f(x + (1-\beta)\alpha(y-x))(y-x).$$

Combining the above two we have

$$f(y) - f(x) \geq \nabla f(x + (1-\beta)\alpha(y-x))(y-x).$$

Letting $\alpha \to 0^+$ *give the first result.*

If $f \in C^2$*, then for any* $x, y \in \Omega$*, from the Taylor expansion theorem there is* $0 \leq \beta \leq 1$ *such that*

$$f(y) = f(x) + \nabla f(x)(y-x) + (1/2)(y-x)^T \nabla^2 f(\beta x + (1-beta)y)(y-x).$$

From the first result, we have

$$(y-x)^T \nabla^2 f(\beta x + (1-\beta)y)(y-x) \geq 0.$$

In particular, for any $x \in \Omega$ *let* $y = x + \alpha d \in \Omega$ *for all possible direction* $d \neq 0$ *and* $\alpha > 0$*. Then we have*

$$d^T \nabla^2 f(x + (1-\beta)\alpha d)d \geq 0.$$

Let $\alpha \to 0^+$*. Then*

$$d^T \nabla^2 f(x)d \geq 0,$$

that is, the Hessian matrix $\nabla^2 f(x)$ *is positive semi-definite throughout* Ω.

1.11 *Prove the Harmonic inequality described in Section* **??**.

Solution 1.11 *We prove the inequality using mathematical inductiion. It is clearly true when $n = 1$. Suppose it is true for $n \geq 1$. Then,*

$$\begin{aligned}
&(\sum_{j=1}^{n+1} x_j)(\sum_{j=1}^{n+1} \frac{1}{x_j}) \\
&= (\sum_{j=1}^{n} x_j)(\sum_{j=1}^{n} \frac{1}{x_j}) + x_{n+1}(\sum_{j=1}^{n} \frac{1}{x_j}) \\
&\quad +(\sum_{j=1}^{n} x_j)j)\frac{1}{x_{n+1}} + 1 \\
&\geq n^2 + (\sum_{j=1}^{n} \frac{x_{n+1}}{x_j} + \frac{x_j}{x_{n+1}}) + 1 \\
&\geq n^2 + (\sum_{j=1}^{n} 2) + 1 \\
&= n^2 + 2n + 1 = (n+1)^2.
\end{aligned}$$

1.12 *Prove Farkas' lemma* **??** *for linear equations.*

Solution 1.12 *Let $A \in \mathcal{R}^{m\times n}$ and $b \in \mathcal{R}^m$. If the system $\{x : \ Ax = b\}$ has a solution, say $\bar{x}$, then for any y such that $A^T y = 0$ we have*

$$b^T y = \bar{x}^T A^T y = \bar{x}^T 0 = 0.$$

On the other hand, if the system has no solution, then b is an exterior point of the convex set $\{Ax : \ x \in \mathcal{R}^n\}$. Thus, from the separating hyperplane theorem there must be a vector y such that

$$y^T b \leq y^T(Ax) = (y^T A)x \quad \textit{for all} \quad x \in \mathcal{R}^n.$$

Since $(y^T A)x$ is bounded from below for all $x \in \mathcal{R}^n$, we must have $y^T A = 0$. Hence $y^T b < 0$.

1.13 *Prove the linear least-squares problem always has a solution.*

Solution 1.13 *We need find an $y \in \mathcal{R}^m$ such that $\|A^T y - c\|$ is minimized. It is sufficient to minimize $\|s - c\|$ where $s = A^T y$ is in the intersection of the space $\{A^T y : \ y \in \mathcal{R}^m\}$ and the ball centered at c with radius $\|c\|$. This intersection is a closed and bounded set. Thus, there must exist an $s^* = A^T y^*$ such that it is closest to c.*

1.14 *Let $P = A^T(AA^T)^{-1}A$ or $P = I - A^T(AA^T)^{-1}A$. Then prove*

1. $P = P^2$.
2. *P is positive semi-definite.*
3. *The eigenvalues of P are either 0 or 1.*

Solution 1.14 1.

$$P^2 = (I - A^T(AA^T)^{-1}A)(I - A^T(AA^T)^{-1}A)$$
$$= I - 2A^T(AA^T)^{-1}A + A^T(AA^T)^{-1}A = I - A^T(AA^T)^{-1}A.$$

2. *For any d we have*

$$d^T Pd = d^T P^2 d = ||Pd||^2 \geq 0,$$

which implies that P is a PSD matrix.

3. *Let λ be an eigenvalue of P. Then there is a vector v such that $Pv = \lambda v$. Therefore,*

$$\lambda v = Pv = P^2 v = \lambda^2 v$$

or $\lambda = \lambda^2$. Thus, λ is either 1 or 0.

1.15 *Using the separating theorem to prove Farkas' lemmas* **??** *and* **??**.

Solution 1.15 *For proving the first lemma let $A \in \mathcal{R}^{m\times n}$ and $b \in \mathcal{R}^m$. If the system $\{x: \ Ax = b, \ x \geq 0\}$ has a feasible solution $\bar{x}$, then for any y such that $A^T y \leq 0$ we have*

$$b^T y = \bar{x}^T A^T y = \bar{x}^T (A^T y) \leq 0.$$

On the other hand, if the system has no solution, then b is an exterior point of the convex set $\{Ax: \ x \in \mathcal{R}^n_+\}$. Thus, from the separating hyperplane theorem there must be a vector $-y \neq 0$ such that

$$-y^T b < -y^T(Ax) = (y^T A)x \quad \text{for all} \quad x \in \mathcal{R}^n_+.$$

Thus, $-y^T b < 0$ or $b^T y > 0$. Since $(-y^T A)x$ is bounded from below for all $x \in \mathcal{R}^n_+$, we must also have $y^T A \leq 0$.

For proving the second lemma let $A \in \mathcal{R}^{m\times n}$ and $c \in \mathcal{R}^n$. If the system $\{y: \ A^T y \leq c\}$ has a solution $\bar{y}$, then for any x such that $Ax = 0$ and $x \geq 0$ we have

$$c^T x = (c - A^T y)^T x \geq 0.$$

On the other hand, if the system has no solution, then c *is an exterior point of the convex set* $\{s: \; s \geq A^T y \text{ for some } y \in \mathcal{R}^m\}$, *which contains all nonnegative vectors. Thus, there is an* $x \neq 0$ *such that*

$$x^T c < x^T s \quad \text{for all} \quad s \geq A^T y.$$

Since $x^T s$ *is bounded from below, we must have* $x \geq 0$. *It must also be true that* $Ax = 0$, *since*

$$x^T c < x^T A^T y \quad \text{for all} \quad y \in \mathcal{R}^m.$$

Moreover, we have $c^T x < 0$.

1.16 *If a system* $A^T y \leq c$ *of linear inequalities in* m *variables has no solution, show that* $A^T y \leq c$ *has a subsystem* $(A')^T y \leq c'$ *of at most* $m+1$ *inequalities having no solution.*

Solution 1.16 *If the system* $A^T y \leq c$ *is infeasible, then we have a certificate* $x \geq 0$ *such that* $Ax = 0$ *and* $c^T x = -1$. *Thus, vector* $(0; -1) \in \mathcal{R}^{m+1}$ *is in the* $cone((a_1; c_1), ..., (a_n; c_n))$. *Since any basis of these vectors has at most* $m+1$ *vectors, from Carathéodory's theorem we have*

$$(0; -1) \in cone((a_{i_1}; c_{i_1}), ..., (a_{i_d}; c_{i_d}))$$

and $d \leq m+1$. *Thus its associated subsystem* $(A')^T y \leq c'$ *is also infeasible, where* $A' = (a_{i_1}, \ldots, a_{i_d})$ *and* $c' = (c_{i_1}, \ldots, c_{i_d})$.

1.17 *Prove the LP fundamental theorem* **??**.

Solution 1.17 *The proof of (i) can be drawn directly from Carathéodory's theorem. We include a complete proof here. Suppose* $0 \leq x \in \mathcal{R}^n$ *is a feasible point such that* $Ax = b$, *i.e.,*

$$\sum_{j=1}^{n} a_{.j} x_j = b.$$

Assume that p *of the variables* x_j *are positive, and for convenience, that they are the first* p *variables. Thus*

$$\sum_{j=1}^{p} a_{.j} x_j = b.$$

There are now two cases, corresponding as to whether the set $a_{.1}, \ldots, a_{.p}$ *is linearly dependent or not.*

Case I: If the set of vectors is linearly independent. Then, $p \le m$. *If* $p = m$, *the solution is basic. If* $p < m$, *then since* A *has rank* m, *we can add* $m - p$ *vecrtors, chosen from the remaining* $a_{.(p+1)}, \ldots, a_{.n}$, *to the set such that the* m *vectors are linear independent. Assigning the value* 0 *to the corresponding* $m - p$ *varaibles yields a (degenerate) basic feasible solution.*

Case II: If the set of vectors is linear dependent, then there is $d \neq 0$ *in the null-space of* $a_{.1}, \ldots, a_{.p}$, *i.e.,*

$$\sum_{j=1}^{p} a_{.j} d_j = 0.$$

Without of loss of generality, let $\min(d) < 0$. *Then we can set*

$$x^+ = x + \alpha d$$

where $\alpha = \min\{x_j/|d_j| : d_j < 0\}$ *such that* $x^+ \geq 0$ *and at least one component of* x^+ *is zero. Thus,* x^+ *has at most* $p - 1$ *positive variables. Repeating this process if necessary, we can eliminate positive variables until we have the corresponding columns are linearly independent.*

The proof of (ii) of the theorem is similar to the proof above. Here we let x *be an optimal feasible solution. Case 1 is identical, while in Case II we have to show that the objective does not change, i.e.,* $c^T d = 0$. *Since* $x + \alpha d$ *and* $x - \alpha d$ *are feasible solutions for a positive value of* α. *Thus we have to have* $c^T d = 0$ *for otherwise* x *is not optimal.*

1.18 *If (LP) and (LD) have a nondegenerate optimal basis* A_B, *prove that the strict complementarity partition in Theorem* **??** *is*

$$P^* = B.$$

Solution 1.18 *Since* A_B *is a nondegenerate optimal basis, we have* $A_B x_B = b$ *and* $X_B > 0$ *for the primal, and* $A_B^T y = c_B$ *and* $c_N - A_N^T y > 0$. *Thus,* (B, N) *is a strictly complementary partition. Since such partition is unique, we have* $P^* = B$.

1.19 *If* Q *is positive semi-definite, prove that* x^* *is an optimal solution for (QP) if and only if* x^* *is a KKT point for (QP).*

Solution 1.19 *A KKT point* $\bar{x}$ *of (QP) satisfies*

$$M \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} - \begin{pmatrix} \bar{s} \\ 0 \end{pmatrix} = -q, \quad \bar{x}^T \bar{s} = 0,$$

where

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

Let x be any feasible point of (QP), then

$$A(\bar{x} - x) = 0,$$

and

$$\begin{aligned} q(\bar{x}) - q(x) &= \frac{1}{2}\bar{x}^T Q\bar{x} - \frac{1}{2}x^T Qx + c^T(\bar{x} - x) \\ &= \frac{1}{2}\bar{x}^T Q\bar{x} - \frac{1}{2}x^T Qx + (\bar{s} - Q\bar{x} + A^T\bar{y})^T(\bar{x} - x) \\ &= \frac{1}{2}\bar{x}^T Q\bar{x} - \frac{1}{2}x^T Qx + (\bar{s} - Q\bar{x})^T(\bar{x} - x) \\ &= -\frac{1}{2}(\bar{x} - x)^T Q(\bar{x} - x) - \bar{s}^T x \le 0. \end{aligned}$$

Thus, $\bar{x}$ is a minimizer.

Next, let $\bar{x}$ be a minimizer and let $\bar{x}_i = 0$ for $i = 1, \ldots, t$, $t \le n$. Let d be in the null space of

$$\begin{pmatrix} A \\ I\ 0 \end{pmatrix}$$

where I is the $t \times t$ identical matrix. Withou loss of generality, we assume that this matrix has full row rank. Then $x(\alpha) = \bar{x} + \alpha d$ is a feasible point for (QP) for any $-\bar{\alpha} \le \alpha \le \bar{\alpha}$ for a sufficiently small positive $\bar{\alpha}$. Since $q(\bar{x}) \le q(x(\alpha))$ for any $\alpha \in [-\bar{\alpha}, \bar{\alpha}]$, we must have $d^T(Q\bar{x} + c) = 0$ since otherwise we can select a small α, $|\alpha| \le \bar{\alpha}$, such that $q(x(\alpha)) < q(\bar{x})$, a contradiction. This fact implies that $Q\bar{x} + c$ is orthogonal to the null space of the above matrix, or equivalently, $Q\bar{x}+c$ is in the row space of the matrix, i.e.,

$$Q\bar{x} + c - A^T\bar{y} - \bar{s} = 0$$

for some $\bar{y} \in \mathcal{R}^m$ and $\bar{s} \in \mathcal{R}^n$ where $\bar{s}_i = 0$ for $i = t+1, \ldots, n$. Thus, we have $\bar{x}^T\bar{s} = 0$.

It remains to be shown that $\bar{s}_i \ge 0$, $i = 1, \ldots, t$. Suppose not, say $\bar{s}_t < 0$. Then let d be in the null space of

$$\begin{pmatrix} A \\ I\ 0 \end{pmatrix}$$

where I is the $(t-1) \times (t-1)$ identical matrix and satisfy $d_t > 0$. (This is possible since the above two matrices have full rwo rank and the second *one is a submatrix of the first.) Thus, $\bar{x}+\alpha d$ is feasible for any sufficiently small $\alpha > 0$. Note*

$$\begin{aligned} q(\bar{x}) - q(\bar{x}+\alpha d) &= -\frac{\alpha^2}{2} d^T Q d - \alpha \bar{s}^T d \\ &= -\frac{\alpha^2}{2} d^T Q d - \alpha \bar{s}_t d_t, \end{aligned}$$

which could be positive for any small $\alpha > 0$. This contradicts to that $\bar{x}$ is the munimizer, since, for a sufficiently small $\alpha > 0$, $\bar{x}+\alpha d$ is feasible and $q(\bar{x}) - q(\bar{x}+\alpha d) > 0$.

1.20 *Show that M of (**??**) is monotone if and only if Q is positive semi-definite.*

Solution 1.20 *Since*

$$M = \begin{pmatrix} Q & -A^T \\ A & 0 \end{pmatrix},$$

for any vectors $(u;v) \in \mathcal{R}^{n+m}$, $(u;v)^T M(u;v) = u^T Q u$. Thus, $(u;v)^T M(u;v) \geq 0$ if and only if $u^T Q u \geq 0$ for any vector $u \in \mathcal{R}^n$. This implies that Q is PSD.

1.21 *Prove the monotone linear complementarity theorem* **??**.

Solution 1.21 (i)) *First, we prove that if a quadratic function $q(x) = \frac{1}{2}x^T Q x + c^T x$ is bounded from below in a convex feasible region $\mathcal{F}$, then it must have a finite (local) minimizer. Without loss of generality, we assume that $0 \in \mathcal{F}$. Let*

$$q^* = \inf\{q(x) : \ x \in \mathcal{F}\}.$$

Suppose there is no finite $x \in \mathcal{F}$ such that $q(x) = q^*$. Then there is a sequence $\{x^k \in \mathcal{F}\}$, $k = 1, \ldots,$ such that $\|x^k\| \to \infty$, $q(x^k) > q^*$, and $q(x^k) \to q^*$. Let $y^k = x^k/\|x^k\|$ and y^* be a limit point of $\{y^k\}$. Then, consider the sequence $\{\alpha^k y^*\}$ where $\alpha^k = \|x^k\|$. We must also have $\alpha^k y^* \in \mathcal{F}$, $q(\alpha^k y^*) > q^*$, and $q(\alpha^k y^*) \to q^*$. But

$$q(\alpha^k y^*) = \frac{(y^*)^T Q y^*}{2}(\alpha^k)^2 + (c^T y^*)\alpha^k$$

has a finite α^* to attain its minimal value. That is, there is a finite α^* such that $q(\alpha^* y^*) = q^*$, which is a contradiction.

Now consider the monotone LCP problem

$$\begin{array}{lll} (LCP) & \textit{minimize} & x^T s \\ & \textit{subject to} & Mx - s = -q, \ x, s \geq 0, \end{array}$$

or

$$\begin{array}{lll}(LCP) & \textit{minimize} & x^T(M+M^T)x/2+q^Tx \\ & \textit{subject to} & Mx \geq -q,\ x \geq 0,\end{array}$$

where $M+M^T$ *is a positive semi-definite matrix. This is a convex quadratic minimization problem over a convex polyhedron, and its objective value is bounded from below by* 0*. Thus, it has a finite minimizer* x *which must also be a KKT point satisfying*

$$y^T(Mx+q)+x^T((M+M^T)x+q-M^Ty)=0,$$

for some y *such that*

$$y \geq 0, \quad \textit{and} \quad (M+M^T)x+q-M^Ty \geq 0.$$

Note that from the last two inequalities

$$y^Tq \leq y^T(M+M^T)x - y^TM^Ty.$$

Thus,

$$\begin{array}{rcl}0 \leq x^T(Mx+q) & = & x^T(Mx+q)-y^T(Mx+q)-x^T((M+M^T)x+q-M^Ty) \\ & = & -x^TM^Tx-y^Tq \\ & \leq & -x^TM^Tx+y^T(M+M^T)x-y^TM^Ty \\ & = & \dfrac{-1}{2}(x-y)^T(M+M^T)(x-y) \\ & \leq & 0.\end{array}$$

That is, we have $x^T(Mx+q)=0$ *which is a solution to the original LCP.*

[ii)] Let (x^1,s^1) *and* (x^2,s^2) *be two solutions for the LCP. We now prove for any* $\alpha \in (0,1)$, $(\alpha x^1+(1-\alpha)x^2, \alpha s^1+(1-\alpha)s^2)$ *is also a solutin for the LCP. First, we have*

$$\alpha s^1+(1-\alpha)s^2 = \alpha(Mx^1+q)+(1-\alpha)M(Mx^2+q) = M(\alpha x^1+(1-\alpha)x^2)+q,$$

and

$$\alpha x^1+(1-\alpha)x^2 \geq 0 \quad \textit{and} \quad \alpha s^1+(1-\alpha)s^2 \geq 0.$$

Moreover,

$$\begin{array}{rl}0 \leq & (\alpha x^1+(1-\alpha)x^2)^T(\alpha s^1+(1-\alpha)s^2) \\ = & \alpha(1-\alpha)((x^2)^Ts^1+(x^1)^Ts^2) \\ = & \alpha(1-\alpha)((x^1)^Ts^1+(x^2)^Ts^2-(x^1-x^2)^T(s^1-s^2)) \\ \leq & \alpha(1-\alpha)((x^1)^Ts^1+(x^2)^Ts^2) \\ = & 0\end{array}$$

iii)] Since the number of positive components in $x+s$ of any solution (x,s) is at most n, the LCP has a maximal complementary solution (x^1,s^1). Let its support be

$$P^1 = \{j: \; x^1_j > 0\} \quad and \quad Z^1 = \{j: \; s^1_j > 0.\}$$

Suppose there is another maximal solution (x^2,s^2) and it has a different support from (x^1,s^1). Since the solution set is convex, $(x^1/2+x^2/2, s^1/2+s^2/2)$ is also a solution to the LCP. However, the number of positive components in $x^1/2+x^2/2+s^1/2+s^2/2$ is more than that in x^1+s^1, contradicting to that (x^1,s^1) is a maximal solutioin.

1.22 *Let $P(y) = -C + \sum_i^m y_i A_i$, where C and A_i, $i = 1,\ldots,m$, are given symmetric matrices. Formulate the minimization of the max-eigenvalue of $P(y)$ as a (PSD) problem. What does its primal problem look like?*

Solution 1.22 *The problem can be formulated as a PSD*

$$\begin{array}{ll} sup & t \\ subject\ to & tI \sum_i^m y_i A_i + S = C, \;\; S \succeq 0. \end{array}$$

The primal is

$$\begin{array}{ll} inf & C \bullet X \\ subject\ to & I \bullet X = 1, \;\; A_i \bullet X = 0, i = 1,2,...,m, \;\; X \succeq 0. \end{array}$$

1.23 *Prove $X \bullet S \geq 0$ if both X and S are positive semi-definite matrices.*

Solution 1.23 *Without losing generality, we may assume that X has an inverse.*

$$X \bullet S = trXS = trX^{-1/2}XSX^{1/2} = trX^{1/2}SX^{1/2} \geq 0.$$

1.24 *Prove that two positive semi-definite matrices are complementary to each other, $X \bullet S = 0$, if and only if $XS = 0$.*

Solution 1.24 *Let $T\Lambda T^T = X$ where diagonal matrix Λ contains the eigenvalues of X and T is its orthonormal eigenvector matrix. Then,*

$$XS = T\Lambda T^T S = T(\Lambda T^T S T)T^T.$$

Thus, $X \bullet S = \Lambda \bullet T^T S T$. Without losing generality, assume

$$\Lambda = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where D is a positive diagonal matrix. Then,

$$\Lambda T^T ST = \begin{pmatrix} D(T^T ST)_D & 0 \\ 0 & 0 \end{pmatrix}$$

and $(T^T ST)_D$ is a principle leading submatrix of $T^T ST$. If $X \bullet S = 0$, then $trD(T^T ST)_D = 0$ which implies that every diagonal component of $(T^T ST)_D$ is zero. But $(T^T ST)_D$ is a PSD matrix, so we must have $(T^T ST)_D = 0$. Hence, $\Lambda T^T ST = 0$ so that $XS = 0$.

1.25 *Prove Farkas' lemma* **??** *for positive semi-definite programming.*

Solution 1.25 *Let $A_i \in \mathcal{M}^n$, $i = 1, ..., m$, have rank m, i.e., $\sum_i^m y_i A_i = 0$ implies that $y = 0$. If there exists a symmetric matrix $X \succ 0$ with*

$$A_i \bullet X = b_i, \quad i = 1, ..., m,$$

then for any y such that $\sum_i^m y_i A_i \preceq 0$ and $\sum_i^m y_i A_i \neq 0$, we have

$$b^T y = \sum_{i=1}^m y_i A_i \bullet X = \langle \sum_i^m y_i A_i, X \rangle < 0.$$

Otherwise, b is either a boundary or exterior point to the closed cone $S := \{s : \ s_i = A_i \bullet X, \ i = 1, \ldots, m, \ x \succeq 0\}$. Thus, there is an $-y \neq 0$ such that

$$-y^T b \leq -y^T s = \langle \sum_i^m -y_i A_i, X \rangle \text{ for all } X \succeq 0.$$

Thus, we must have $-b^T y \leq 0$ or $b^T y \geq 0$. Since the right-hand is bounded from below for all $X \succeq 0$, we must have $\sum_i^m y_i A_i \preceq 0$. Moreover, $\sum_i^m y_i A_i \neq 0$ since $y \neq 0$.

1.26 *Let both (LP) and (LD) for a given data set (A, b, c) have interior feasible points. Then consider the level set*

$$\Omega(z) = \{y : \ c - A^T y \geq 0, -z + b^T y \geq 0\}$$

where $z < z^*$ and z^* designates the optimal objective value. Prove that $\Omega(z)$ is bounded and has an interior for any finite $z < z^*$, even $\mathcal{F}_d$ is unbounded.

Solution 1.26 *Let y^* be an optimal solution and y^0 be any interior point of $\mathcal{F}_d$. Then for any $\alpha \in (0, 1]$, $y(\alpha) = (1-\alpha)y^* + \alpha y^0$ is an interior point. Let α be sufficiently small but positive such that $b^T y(\alpha) > z$. Then, $y(\alpha)$ is an interior-point of $\Omega(z)$.*

Let $x^0 > 0$ be an interior-point of $\mathcal{F}_p$. For any point in $\Omega(z)$ let $s = c - A^T y \geq 0$. Then,

$$s^T x^0 = c^T x^0 - b^T y \leq c^T x^0 - z.$$

Thus, s must be a bounded vector. Since A has full row rank, y must be also bounded.

1.27 *Given an (LP) data set (A, b, c) and an interior feasible point x^0, find the feasible direction d_x ($Ad_x = 0$) that achieves the steepest decrease in the objective function.*

Solution 1.27 *It is the negative of the projection of c onto the null space of A:*

$$d_x = -(I - A^T(AA^T)^{-1}A)c.$$

1.28 *Let (LP), (QP) and (LCP) have feasible points $(x^0, y^0, s^0) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+)$, $(x^0, y^0, s^0) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+)$ and $(x^0, s^0) \in (\mathcal{R}^n_+, \mathcal{R}^n_+)$, respectively. Ignoring the nonnegativity condition, write the systems of linear equations used to calculate the Newton steps for finding points that satisfy the optimality equations (??), (??), and (??), respectively.*

Solution 1.28 *For LP*

$$\begin{array}{rcl} S^0 d_x + X^0 d_s &=& -X^0 s^0, \\ A d_x &=& 0, \\ -A^T d_y - d_s &=& 0. \end{array}$$

For QP

$$\begin{array}{rcl} S^0 d_x + X^0 d_s &=& -X^0 s^0, \\ A d_x &=& 0, \\ -A^T d_y + Q d_x - d_s &=& 0. \end{array}$$

For LCP

$$\begin{array}{rcl} S^0 d_x + X^0 d_s &=& -X^0 s^0, \\ M d_x - d_s &=& 0. \end{array}$$

1.29 *Similar to our discussion on quadratic programming and linear complementarity, demonstrate that finding a KKT point of a convex nonlinear programming problem can be reduced to solving a monotone complementarity problem with possible "free" variables.*

Solution 1.29 *Let the problem be*

$$\begin{array}{lll} (NP) & \textit{minimize} & f(x) \\ & \textit{subject to} & Ax = b,\ g(x) \geq 0, \end{array}$$

where f is convex and g is concave. Then its KKT system can be written as

$$\begin{array}{ll} \textit{minimize} & z^T s \\ \textit{subject to} & \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} \nabla^T f(x) - A^T y - \nabla^T g(x) z \\ Ax - b \\ g(x) \end{pmatrix}, \quad s, z \geq 0. \end{array} \tag{1.1}$$

The Jacobian matrix of the right-hand function is

$$\begin{pmatrix} \nabla^2 f(x) - \nabla^g(x) z & -A^T & -\nabla^T g(x) \\ A & 0 & 0 \\ \nabla g(x) & 0 & 0 \end{pmatrix},$$

which is monotone.

1.30 *Show that the ball-constrained linear problem (BP) can be written in the (BD) form and write the KKT conditions for both of them.*

Solution 1.30 *Consider*

$$\begin{array}{lll} (BP) & \textit{minimize} & c^T x \\ & \textit{subject to} & Ax = 0, \ \|x\|^2 \leq 1. \end{array}$$

We can let B be the orthornomal basis of the null space of A ($AB^T = 0$ and $BB^T = I$). Then, $x \in \{x : \ Ax = 0$ can be written as $x = B^T z$. Substituting $B^T z$ for x we have

$$\begin{array}{ll} \textit{minimize} & c^T B^T z \\ \textit{subject to} & \|z\|^2 \leq 1. \end{array}$$

The KKT condition for (BP) is

$$c - A^T y - \mu x = 0, \quad \textit{and} \quad Ax = 0$$

and

$$\mu \geq 0, \quad 1 \geq \|x\|^2, \quad \textit{and} \quad \mu(1 - \|x\|^2) = 0.$$

The KKT condition for the latter is

$$Bc - \mu z = 0, \quad 1 \geq \|z\|, \quad \textit{and} \quad \mu(1 - \|z\|^2) = 0.$$

1.31 *Given a scalar $\alpha > 0$, a positive diagonal matrix S, and an $m \times n$ matrix A, find the formula for $y \in \mathcal{R}^m$ such that*

$$\begin{array}{ll} \textit{minimize} & e^T S^{-1} A^T y \\ \textit{subject to} & \|S^{-1} A^T y\|^2 \leq \alpha^2. \end{array}$$

Solution 1.31

$$y = \frac{\alpha}{\eta_d(s)}(AS^{-2}A^T)^{-1}AS^{-1}e,$$

where

$$\eta_d(s) = \|S^{-1}A^T(AS^{-2}A^T)^{-1}AS^{-1}e\| = \sqrt{e^TA^TS^{-1}(AS^{-2}A^T)^{-1}AS^{-1}e}.$$

1.32 *Show the optimality conditions for the minimizer y^* of (BD) in Section* **??***:*

$$(Q + \mu^* I)y^* = -b, \quad \mu^* \geq 0, \quad \|y^*\| \leq 1, \quad \mu^*(1 - \|y^*\|) = 0,$$

and

$$(Q + \mu^* I) \succeq 0,$$

are necessary and sufficient.

Solution 1.32 *We basically prove these conditions are sufficient even when Q is nonconvex. Let (μ^1, x^1) and (μ^2, x^2) both satisfy these conditions. We have $\|x^1\| = \|x^2\| = 1$. If $\mu^1 > \mu^2 (> 0)$, then $Q + \mu^1 I$ is positive definite. Thus,*

$$(Q + \mu^1 I)^{-1}(Q + \mu^2 I)x^2 = -(Q + \mu^1 I)^{-1}c = x^1,$$

which implies

$$\|x^1\| = \|(Q+\mu^1 I)^{-1}(Q+\mu^2 I)x^2\| \leq \|(Q+\mu^1 I)^{-1}(Q+\mu^2 I)\|\|x^2\| < \|x^2\| = 1,$$

a contradiction. Similarly, we cannot have $\mu^1 < \mu^2$. Thus, we must have $\mu^1 = \mu^2$.

We furtherprove $q(x^1) = q(x^2)$. If $Q + \mu^1 I$ is positive definite, then $x^1 = x^2$. Hence we only need to be concerned with the case $\mu^1 = |\underline{\lambda}|$, where $\underline{\lambda} < 0$ is the least eigenvalue of Q. Then, for any solution x, $\|x\| = 1$ and satisfy $(Q + \mu^1 I)x = -c$, we have $x = v + b$ where v is a homogeneous solution of $(Q + \mu^1 I)v = 0$ and b is the minimal-norm solution with $(Q + \mu^1 I)b = -c$. Note that b is orthogonal to any homogeneous solution v. Then

$$\begin{aligned} q(x) &= -(v + b)(\frac{1}{2} + \mu^1 I)(v + b) \\ &= -\frac{1}{2}\mu^1(\|v\|^2 + \|b\|^2) - \frac{1}{2}b^T(Q + \mu^1 I)b \\ &= -\frac{1}{2}\mu^1 - \frac{1}{2}b^T(Q + \mu^1 I)b. \end{aligned}$$

This quantity is independent of v. Thus, $q(x^1) = q(x^2)$.