On the feasibility for the system of quadratic equations, explanations

1. Theorem 3.2 (Sufficient condition)

Consider $f\colon \mathbb{R}^n \to \mathbb{R}^m$, s.t. $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$. Define $F = f(\mathbb{R}^n)$.
Then why $A = \inf\limits_{y \in F}(c, y) = \inf\limits_{y \in \operatorname{conv} F}(c, y) = B$?

1. First, $F \subseteq \operatorname{conv} F$, therefore, $B \leqslant A$.

2. Secondly, let $y_k \in \operatorname{conv} F$ be a sequence s.t. $g_k = (c, y_k) \underset{k\to\infty}{\longrightarrow} B$. $y_k = \sum\limits_{i=1}^{n_k} \alpha_i^k y_i^k$.

   $g_k(c, y_k) = \sum\limits_{i=1}^{n_k} \alpha_i^k (c, y_i^k) = \sum\limits_{i=1}^{n_k} \alpha_i^k g_i^k$. Define $g_0^k = \min\limits_{i \in \overline{1,n_k}} g_i^k$. Then $B \leqslant g_0^k \leqslant g^k$. Therefore, $g_0^k \to B$ also. This way, we have constructed a sequence $y_0^k \in F$ s.t. $(c, y_0^k) \to B$, therefore, $A \leqslant B$.

2. Minimum of $f(x)$

Consider $f\colon \mathbb{R}^n \to \mathbb{R}^m$. $f_i(x) = x^T A_i x + 2b_i^T x$. $A_i^T = A_i$. Let $c \in \mathbb{R}^m$
We want to find $g(c) = \inf\limits_{x \in \mathbb{R}^n}(c, f(x))$.
Define $A_c \equiv c \cdot A = \sum\limits_{i=1}^m c_i A_i$, $b_c = c \cdot b = \sum\limits_{i=1}^m c_i b_i$.
$(c, f(x)) = \sum\limits_{i=1}^m c_i f_i(x) = \sum c_i (x^T A_i x + 2b_i^T x) = x^T A_c x + 2b_c^T x$.
If $\exists v\colon -\alpha = v^T A_c v < 0$ then $g(c) = -\infty$: $g(\beta v) = -\beta^2 \alpha + \beta 2 b_c^T v \to -\infty$, $\beta \to +\infty$.
From this point on, we assume $A_c \geqslant 0$. Let R_0 be a zero eigenspace (=kernel) of A_c: $R_0 = \{v\colon A_c v = 0\} = \ker A_c$
If $\exists v \in R_0\colon v^T b_c \neq 0$ then $g(c) = -\infty$: Consider $f(\beta v) = \beta^2 v^T \cancel{(A_c v)}^{\,0} + 2\beta \underbrace{b_c^T v}_{\neq 0} \to -\infty$, $\beta \to \infty$

Then $R_0 \subseteq \{b_c\}^\perp$
Consider $A = \sum\limits_{i=1}^n \lambda_i s_i s_i^T = S \Lambda S^T$, $S = ||s_1 \ldots s_n||$, $S^T S = E$, $s_i^T s_j = \delta_{ij}$.
f is differentiable, then for finding $g(c)$ the gradiend $\nabla(c, f(x)) = 2A_c x + 2b_c = 0$.

$$S \Lambda S^T x = -b_c \Leftrightarrow \Lambda S^T x = -S^T b_c \, (*)$$

.

Let x be $x = x^\parallel + x^\perp$, $x^\parallel \in R_0$, $x^\perp \perp R_0$.
Then neither $f(x)$ nor $\Lambda S^T x$ depend on $x^\parallel$. This means that the x minimizing $g(c)$ is defined in terms of $x^\perp$ and $x^\parallel$ is arbitrary.
Define $\lambda_i^g = \begin{cases} 0, & \lambda_i = 0 \\ 1/\lambda_i, & \lambda_i \neq 0 \end{cases}$. Define $\Lambda^g = \operatorname{diag}(\lambda_1^g, \ldots, \lambda_n^g)$. Then $\Lambda \Lambda^g = \delta_{ij}[\lambda_i \neq 0]$. Then $S \Lambda^g \Lambda S^T$ is a projector on $R_0^\perp$.
Consider $(*) \Leftrightarrow \Lambda^g \Lambda S^T (x^\parallel + x^\perp) = -\Lambda^g S^T b_c$. But $\Lambda S^T x^\parallel = 0$, therefore, $\underbrace{S \Lambda^g \Lambda S^T}_{\text{projector}} x^\perp = -S \Lambda^g S^T b_c$.
But $x^\perp$ is already in $R_0^\perp$, therefore, $x^\perp = -\underbrace{S \Lambda^g S^T}_{A_c^g} b_c$. Here A_c^g is a pseudoinverse of A_c.

Therefore, $\boxed{x = -A_c^g b_c + x^\parallel}$, where $x^\parallel \in R_0$.

Consider $(c, f(x)) = (c, f(x^{\perp})) = b_c^T A_c^g A_c A_c^g b_c - 2b_c^T A_c^g b_c$. Consider $A_c A_c^g b_c = S\Lambda \overset{E}{\cancel{S^T S}} \Lambda^g S^T b_c$. Because $R_0 \subseteq \{b_c\}^{\perp}$, $\Lambda\Lambda^g S^T b_c = S^T b_c$. Therefore, $A_c A_c^g b_c = b_c$. Then $(c, f(x)) = b_c^T A_c^g b_c - 2b_c^T A_c^g b_c =$ $\boxed{-b_c^T A_c^g b_c}$

3. Finding c provided d

Let $H\colon \mathbb{R}^{n+1,n+1} \to \mathbb{R}^n$ be a map s.t. $H_i(X) = \mathrm{Tr}(H_i X)$,

$$H_i = \left\| \begin{matrix} A_i & b_i \\ b_i^T & 0 \end{matrix} \right\|^{\square}$$

.

Consider a boundary point X, which is a solution of (main article, (4)):

$$\boxed{\begin{matrix} \sup & t \\ \begin{cases} H(X) = y^0 + td \\ X \geqslant 0 \\ X_{n+1,n+1} = 1 \end{cases} & \end{matrix}}$$

Define $f(t, X) = t$, $D_0 = \{(t, X) \big| X \geqslant 0,\ X_{n+1,n+1} = 1\}$, $D_1 = \{(t, X) \big| H(X) = y^0 + td\}$.

Then supremum is equivalent to

$$\sup_{(t,X)\in D_0 \cap D_1} f(t, X)$$

Define a Lagrange function $L(c, t, X) = \underbrace{t}_{f(t,X)} + \sum\limits_{i=1}^{m} c_i(y_i^0 + td_i - H_i(X))$.

Here we divided the constraints into two parts: D_1 goes to the Lagrange function, D_0 goes to the inner supremum. *Stephen Boyd, Lieven Vandenberghe. Convex Optimization. Page ???. Cambridge University Press*

Then the dual function is $g(c) = \sup\limits_{(t,X)\in D_0} L(c, t, X)$.

Because $L = t(1 + \sum\limits_{i=1}^{m} c_i d_i) + \sum\limits_{i=1}^{n} c_i(y_i^0 - H_i(X))$, $g = +\infty$ when $(c, d) \neq -1$. From this point we assume that $\boxed{(c, d) = -1}$.

Now, $g(c) = \sup\limits_{X_{n+1,n+1}=1, X\geqslant 0} (c, y^0 - H(X)) = (c, y^0) + \sup\limits_{y\in \mathrm{conv}\, F} -(c, y) = (c, y^0) - \inf\limits_{y\in \mathrm{conv}\, F} (c, y)$.

Then the dual problem is

$$g(c) \to \inf_{(c,d)=-1}$$

Let us prove that $\inf\limits_{y\in \mathrm{conv}\, F} (c, y) = \inf\limits_{H = \left\| \begin{smallmatrix} A_c & b_c \\ b_c^T & \gamma \end{smallmatrix} \right\| \geqslant 0} (-\gamma)$

Via Schur complement $H \geqslant 0 \Leftrightarrow \begin{cases} A_c \geqslant 0 \\ \gamma - b_c^T A_c^g b_c \geqslant 0 \\ (E - A_c A_c^g) b_c = 0 \end{cases}$.

$A_c \geqslant 0$ is a necessary condition for $\exists g(c) \in \mathbb{R}$ (see part 2).

$(E - A_c A_c^g) b_c = 0$ is another necessary condition for $\exists g(c) \in \mathbb{R}$.

Statement $\gamma \geqslant b_c^T A_c b_c$ means $-\gamma \leqslant -b_c^T A_c b_c = \inf\limits_{y\in \mathrm{conv}\, F} (c, y)$, which means that $-\gamma$ is a lower bound for $\inf\limits_{y\in \mathrm{conv}\, F} (c, y)$.

Then $H \geqslant 0 \Leftrightarrow -\gamma \leqslant \inf\limits_{y\in \mathrm{conv}\, F} (c, y)$.

Then $g(c) = (c, y^0) - \inf\limits_{H\geqslant 0} -\gamma$.

Then the dual problem is:

$$\inf_{(c,d)=-1} g(c) \Leftrightarrow \inf_{(c,d)=-1} \left[(c, y^0) - \inf_{H\geqslant 0} (-\gamma) \right] \Leftrightarrow \inf_{(c,d)=-1} \inf_{H\geqslant 0} (c, y^0) + \gamma = \boxed{\begin{matrix} \inf & \gamma + (c, y^0) \\ \begin{cases} H \geqslant 0 \\ (c, d) = -1 \end{cases} & \end{matrix}}.$$

This problem is exactly (5) from main article ■.

4. What is $z_{\max}$?

Consider $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$.

Let $c_+ \in \mathbb{R}^m$ s.t. $A_+ = \sum c_i A_i > 0$. Then the minimum $\inf\limits_x (c, f(x))$ is obtained at a single point $x_0 = -A_+^{-1} b_+$, $b_+ = \sum c_i b_i$.

Consider $S_\varepsilon^+ = \{x \in \mathbb{R}^n \big| (x - x_0)^T A_+ (x - x_0) = \varepsilon^2\}$. Then $f(S_\varepsilon^+) = \{y \in \mathbb{R}^m \big| (c_+, y) = (c_+, f(x_0)) + \varepsilon^2\}$.

Indeed, if $\begin{cases} P = (c_+, f(x) - f(x_0)) \\ Q = (x - x_0)^T A_+ (x - x_0) \end{cases}$ then $P - Q = \underline{x^T A_+ x} - x_0^T A_+ x_0 + 2b_+^T x - 2b_+^T x_0 - \underline{x^T A_+ x} - x_0^T A_+ x_0 + 2x^T A_+ x_0 = \big/ x_0 = -A_+^{-1} b_+ \big/ = 2x_0^T b_+ + 2b_+^T x - 2b_+^T x_0 - 2x^T b_+ = 0.$

$\boxed{\textbf{Therefore, the image of } B_\varepsilon^+ \textbf{ is a } \textbf{\textit{convex cut}}\ \{y \big| (c_+, y) \in (c_+, f(x_0)) + [0, z_{\max}]\}}$

5. Variables s.t. $c \cdot A = I$, $c \cdot b = 0$

Given: the map $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$, vector $c \in \mathbb{R}^m\colon c \cdot A > 0$

We need to find a new pair of bases s.t. $\begin{cases} \tilde{c} \cdot \tilde{A} = I & (1) \\ \tilde{c} \cdot \tilde{b} = 0 & (2) \end{cases}$

1. Condition (1). Changing variables in the x space: $x = S\tilde{x} \Leftrightarrow \tilde{x} = S^T x$, $S^T S = E$. Then $f_i(x) = x^T A_i x + 2b_i^T x = \tilde{x}^T \tilde{A}_i x + 2\tilde{b_i}^T \tilde{x}$. $\tilde{A}_i = S^T A_i S$. $c\tilde{A} = \sum c_i \tilde{A}_i = \sum c_i S^T A_i S = S^T A_c S$. Therefore, condition (1) is equal to diagonalising A_c. Consider $||c \cdot \tilde{b}|| = ||\sum c_i S^T b_i|| = ||S^T \sum c_i b_i|| = ||\sum c_i b_i||$. Therefore, a change of variables in the x space does not affect on the value of $c \cdot b$
2. Condition (2). New variables: $\tilde{x}$, $\tilde{y}$,
$$\begin{cases} x = \tilde{x} + x^0 \\ y = \tilde{y} + y^0 \end{cases}$$
Function $y_i(x) = \tilde{y}_i(\tilde{x}) + y_i^0$. Consider $y_i(x) = x^T A_i x + 2b_i^T x = (\tilde{x} + x^0)^T A_i (\tilde{x} + x^0) + 2b_i^T (\tilde{x} + x^0) = \tilde{x}^T A_i \tilde{x} + 2x^{0T} A_i \tilde{x} + x^{0T} A_i x^0 + 2b_i^T x^0 + 2b_i^T \tilde{x} = \tilde{x}^T A_i \tilde{x} + 2\underbrace{(b_i + A_i x^0)}_{\tilde{b}_i}{}^T \tilde{x} + \underbrace{x^{0T} A_i x^0 + 2b_i^T x^0}_{y_i^0}$.

   Consider $\sum c_i \tilde{b}_i = c \cdot b + (c \cdot A)x^0$. Therefore, $x^0 = -(c \cdot A)^{-1}(c \cdot b)$

The algorithm:

1. Compute S via the eigenbasis of $c \cdot A$, $S^T (c \cdot A) S = I$
2. Compute $\tilde{A}_i = S^T A_i S$, $\tilde{b}_i = S^T b_i$
3. Compute $\tilde{x}^0 = -(c \cdot \tilde{b})$, $y_i^0 = (\tilde{x}^0)^T \tilde{A}_i \tilde{x}^0 + 2\tilde{b}_i^T \tilde{x}^0$
4. Compute $\hat{A}_i = \tilde{A}_i$, $\hat{b}_i = \tilde{b}_i + \tilde{A}_i \tilde{x}^0$

Then $\hat{y}_i = \hat{x}^T \hat{A}_i \hat{x} + 2\hat{b}_i^T \hat{x}$, $x = S(\hat{x} + \tilde{x}^0)$, $y = \hat{y} + \tilde{y}^0$

6. $z(c) =$?

Given: the map $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$, two vectors $c, c_+ \in \mathbb{R}^m$. $c_+ A = I$, $c_+ b = 0$. Find: $z(c) = \inf\limits_{y \in Y} (c_+, y)$ where Y is an intersection of $f(\mathbb{R}^n)$ with a tangent hyperplane defined by its normal vector c.

Define $\sigma(Q) = \{\lambda \big| \dim \operatorname{Ker}(Q - \lambda E) > 0\}$. Define $\lambda_{\min}(Q) = \min \sigma(Q)$ — minimal eigenvalue of Q.

1. If $\lambda_{\min}(c \cdot A) < 0$ then the tangent hyperplane does not exist, and $z(c) = +\inf$
2. Then $\lambda_{\min}(c \cdot A) > 0$, then there is no nonconvexity, and $z(c) = \inf$
3. For $\lambda_{\min} = 0$ in part 2 Y was found explicitly: $Y = \{f(x) \big| x = x^{\parallel} - A_c^g b_c, x^{\parallel} \in \operatorname{Ker} A_c\}$. Then for $y \in Y$ $(c_+, y) = x^T (\cancel{c_+ \cdot A}^{I}) x + 2(\cancel{c_+ \cdot b}^{0})^T x = \boxed{x^T x}$

4. $x^T x = ||x||^2 = ||x^{\|}||^2 + ||A_c^g b_c||^2$. We want to minimize (c_+, y), therefore, we choose $x^{\|} = 0$. Then $z(c) = ||A_c^g b_c||^2 = \boxed{\left|(c \cdot A)^g (c \cdot b)\right|^2}$

5. Consider $\inf\limits_{x \in \mathbb{R}^n} (c_+, f(x)) = -(\cancelto{0}{c_+ \cdot b})(\cancelto{I}{A_{c_+}})^{-1}(c_+ \cdot b) = 0$. Therore, $z(c) = \inf\limits_{y \in Y}(c_+, y) - \inf\limits_{y \in F}(c_+, y)$

7. Theorem 3.4 (Nonconvexity certificate)

Given.

1. The map $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$, $m, n \geqslant 3$. The vector $c \in \mathbb{R}^m$.
2. $A_c \geqslant 0$, $\dim \operatorname{Ker} A_c = 1$ (=simple zero eigenvalue), $\operatorname{Ker} A_c = \operatorname{Lin}\{e\}$
3. $b_c \perp \operatorname{Ker} A_c$
4. $b \perp \operatorname{Ker} A_c$, $e^0 = -A_c^g b_c$
5. $f(\alpha e + e^0) = f^0 + 2\alpha f^1 + \alpha^2 f^2$, $f^1 \nparallel f^2$

Then $F = \operatorname{Im} f$ is nonconvex.

Consider $\inf\limits_{y \in F}(c, y)$ (part 2). $x = \underbrace{x^{\|}}_{\alpha e} + (\underbrace{-A_c^g b_c}_{e^0})$. Then

$$f_i(\alpha e + e^0) = \alpha^2 \underbrace{e^T A_i e}_{f_i^2} + 2\alpha(\underbrace{b_i^T e + e^T A_i e^0}_{f_i^1}) + \underbrace{2b_i^T e^0 + e^{0T} A_i e^0}_{f_i^0}$$

If $f^1 \nparallel f^2$, then $\{f(\alpha e + e^0 \big| \alpha \in \mathbb{R})\} = F \cap \{y \big| (c, y) = \inf\limits_{y \in F}(c, y)\}$ is nonconvex. Then F is nonconvex.