

On the feasibility for the system of quadratic equations, explanations

1. Theorem 3.2 (Sufficient condition)

Consider $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 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 $A = \inf_{y \in F} (c, y) = \inf_{y \in \text{conv } F} (c, y) = B$?

1. First, $F \subseteq \text{conv } F$, therefore, $B \leq A$.

2. Secondly, let $y_k \in \text{conv } F$ be a sequence s.t. $g_k = (c, y_k) \xrightarrow{k \rightarrow \infty} B$. $y_k = \sum_{i=1}^{n_k} \alpha_i^k y_i^k$.

$g_k(c, y_k) = \sum_{i=1}^{n_k} \alpha_i^k (c, y_i^k) = \sum_{i=1}^{n_k} \alpha_i^k g_i^k$. Define $g_0^k = \min_{i \in \{1, \dots, n_k\}} g_i^k$. Then $B \leq g_0^k \leq g^k$. Therefore, $g_0^k \rightarrow B$ also. This way, we have constructed a sequence $y_0^k \in F$ s.t. $(c, y_0^k) \rightarrow B$, therefore, $A \leq B$.

2. Minimum of $f(x)$

Consider $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. $f_i(x) = x^T A_i x + 2b_i^T x$. $A_i^T = A_i$. Let $c \in \mathbb{R}^m$

We want to find $g(c) = \inf_{x \in \mathbb{R}^n} (c, f(x))$.

Define $A_c \equiv c \cdot A = \sum_{i=1}^m c_i A_i$, $b_c = c \cdot b = \sum_{i=1}^m c_i b_i$.

$(c, f(x)) = \sum_{i=1}^m c_i f_i(x) = \sum_{i=1}^m c_i (x^T A_i x + 2b_i^T x) = x^T A_c x + 2b_c^T x$.

If $\exists v: -\alpha = v^T A_c v < 0$ then $g(c) = -\infty$: $g(\beta v) = -\beta^2 \alpha + \beta 2b_c^T v \rightarrow -\infty$, $\beta \rightarrow +\infty$.

From this point on, we assume $A_c \geq 0$. Let R_0 be a zero eigenspace (=kernel) of A_c : $R_0 = \{v: A_c v = 0\} = \ker A_c$

If $\exists v \in R_0: v^T b_c \neq 0$ then $g(c) = -\infty$: Consider $f(\beta v) = \beta^2 v^T \overbrace{(A_c v)}^0 + 2\beta \underbrace{b_c^T v}_{\neq 0} \rightarrow -\infty$, $\beta \rightarrow \infty$

Then $R_0 \subseteq \{b_c\}^\perp$

Consider $A = \sum_{i=1}^n \lambda_i s_i s_i^T = S \Lambda S^T$, $S = \|s_1 \dots s_n\|$, $S^T S = E$, $s_i^T s_j = \delta_{ij}$.

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 $\Lambda^g = \text{diag}(\lambda_1^g, \dots, \lambda_n^g)$. 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, $\underbrace{S \Lambda^g \Lambda S^T}_{\text{projector}} x^\perp = -\Lambda^g S^T b_c$.

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 $A_c A_c^g b_c = S \Lambda S^T S^{\nearrow E} \Lambda^g S^T b_c$. 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 $H: \mathbb{R}^{n+1, n+1} \rightarrow \mathbb{R}^n$ be a map s.t. $H_i(X) = \text{Tr}(H_i X)$,

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

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

$$\boxed{\begin{array}{l} \sup t \\ \left\{ \begin{array}{l} H(X) = y^0 + td \\ X \geq 0 \\ X_{n+1, n+1} = 1 \end{array} \right. \end{array}}$$

Define $f(t, X) = t$, $D_0 = \{(t, X) | X \geq 0, X_{n+1, n+1} = 1\}$, $D_1 = \{(t, X) | H(X) = y^0 + td\}$. Then supremum is equivalent to

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

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

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 Vandenbergh. Convex Optimization. Page ????. Cambridge University Press*

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

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

Now, $g(c) = \sup_{X_{n+1, n+1}=1, X \geq 0} (c, y^0 - H(X)) = (c, y^0) + \sup_{y \in \text{conv } F} -(c, y) = (c, y^0) - \inf_{y \in \text{conv } F} (c, y)$.

Then the dual problem is

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

Let us prove that $\inf_{y \in \text{conv } F} (c, y) = \inf_{H = \left\| \begin{array}{cc} A_c & b_c \\ b_c^T & \gamma \end{array} \right\| \geq 0} (-\gamma)$

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

$A_c \geq 0$ 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 $\gamma \geq b_c^T A_c b_c$ means $-\gamma \leq -b_c^T A_c b_c = \inf_{y \in \text{conv } F} (c, y)$, which means that $-\gamma$ is a lower bound

for $\inf_{y \in \text{conv } F} (c, y)$.

Then $H \geq 0 \Leftrightarrow -\gamma \leq \inf_{y \in \text{conv } F} (c, y)$.

Then $g(c) = (c, y^0) - \inf_{H \geq 0} -\gamma$.

Then the dual problem is:

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

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

4. What is z_{\max} ?

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

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

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

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

Therefore, the image of B_ε^+ is a **convex cut** $\{y \mid (c_+, y) \in (c_+, f(x_0)) + [0, z_{\max}]\}$

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

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

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

1. Condition (1). Changing variables in the x space: $x = S\tilde{x} \Leftrightarrow \tilde{x} = S^{-1}x$. Then $f_i(x) = x^T A_i x + 2b_i^T x = \tilde{x}^T S^T A_i S \tilde{x} + 2\tilde{b}_i^T S \tilde{x}$. $\tilde{A}_i = S^T A_i S$. $c \cdot \tilde{A} = \sum c_i \tilde{A}_i = \sum c_i S^T A_i S = S^T A_c S$. 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f_i(x) = x^T A_i x + 2b_i^T x$, $A_i = A_i^T$, two vectors $c, c_+ \in \mathbb{R}^m$. $c_+ A = I$, $c_+ b = 0$. Find: $z(c) = \inf_{y \in Y} (c_+, y)$ where Y is an intersection of $f(\mathbb{R}^n)$ with a tangent hyperplane defined by its normal vector c .

Define $\sigma(Q) = \{\lambda \mid \dim \text{Ker}(Q - \lambda E) > 0\}$. 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 $z(c) = +\infty$

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: $Y = \{f(x) \mid x = x^\parallel - A_c^g b_c, x^\parallel \in \text{Ker } A_c\}$. Then for

$$y \in Y \quad (c_+, y) = x^T (\cancel{c_+} \xrightarrow{A^I} x) + 2(\cancel{c_+} \xrightarrow{b^0} x) = \boxed{x^T x}$$

4. $x^T x = \|x\|^2 = \|x^\parallel\|^2 + \|A_c^g b_c\|^2$. We want to minimize (c_+, y) , therefore, we choose $x^\parallel = 0$. Then

$$z(c) = \|A_c^g b_c\|^2 = \boxed{|(c \cdot A)^g(c \cdot b)|^2}$$

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

$$\text{Now consider } z(c) = \begin{cases} |(c \cdot A)^g(c \cdot b)|^2, & \lambda_{\min}(c \cdot A) = 0 \\ +\infty, & \text{otherwise} \end{cases}$$

Consider $z(c + \alpha c_+) = |(c \cdot A + \alpha \cancel{c_+ \cdot A}^I)^g b_c|^2 = |(c \cdot A + \alpha I)^g b_c|^2$. We need α s.t. $\lambda_{\min}(c \cdot A + \alpha I) = 0$, therefore, $\alpha = -\lambda_{\min}(c \cdot A)$.

Define $\hat{z}(c) = |(c \cdot A - \lambda_{\min}(A))^g(c \cdot b)|^2$. Consider $\mathbb{R}^m \ni c = c^\parallel + c^\perp$, $c^\parallel \parallel c_+$, $(c_+, c^\perp) = 0$.

Then $\hat{z}(c) = \hat{z}(c^\perp)$, i.e. \hat{z} does not depend on c^\parallel . It depends only on c^\perp , and c^\parallel is chosen in a way that $\lambda_{\min}(c \cdot A) = 0$.

7. Theorem 3.4 (Nonconvexity certificate)

Given.

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

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

Consider $\inf_{y \in F} (c, y)$ (part 2). $x = \underbrace{x^\parallel}_{\alpha e} + \underbrace{(-A_c^g b_c)}_{e^0}$. 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 $\{f(\alpha e + e^0) | \alpha \in \mathbb{R}\} = F \cap \{y | (c, y) = \inf_{y \in F} (c, y)\}$ is nonconvex. Then F is nonconvex.

8. Equations (0.18)-(0.21)

Consider $A(t): n \times n$, $\exists \dot{A}$, $A^T = A$, $A \geq 0$, A has a simple zero eigenvalue: $\forall t \ A(t)x_0(t) = 0$, $x_0^T x_0 = 0$.

Then $A = S \Lambda S^T$, $S^T S = E$, $A^g = S \Lambda^g S^T$. Define $\lambda_i = \Lambda_{ii}$. $\Lambda_{ii}^g = \begin{cases} \frac{1}{\lambda_i}, & \lambda_i \neq 0 \\ 0, & \lambda_i = 0 \end{cases}$.

Then $AA^g = A^g A = S \Lambda S^T S \Lambda^g S^T = S \Lambda \Lambda^g S^T = \sum_{\lambda_i \neq 0} s_i s_i^T = 1 - x_0 x_0^T$ (0.20)

Consider $\frac{d}{dt} A x_0 = \dot{A} x_0 + A \dot{x}_0$. Multiplying by A^g from the left: $A^g \dot{A} x_0 + A^g A \dot{x}_0 = 0$. Consider $A^g A \dot{x}_0 = (1 - x_0 x_0^T) \dot{x}_0 = \dot{x}_0 - x_0 x_0^T \dot{x}_0$. Since $\|x_0\|^2 = x_0^T x_0 = 1$, $x_0^T \dot{x}_0 = 0$. Then $-A^g \dot{A} x_0 = \dot{x}_0$ (0.19).

Consider $\dot{A} x_0 + A \dot{x}_0 = 0$. Multiplying by x_0^T from the left: $x_0^T \dot{A} x_0 + \cancel{x_0^T A}^0 \dot{x}_0 = 0$. Then $x_0^T \dot{A} x_0 = 0$ (0.18)

Consider $AA^g = 1 - x_0 x_0^T$. Then $\dot{A} A^g + A \dot{A}^g = -\dot{x}_0 x_0^T - x_0 \dot{x}_0^T = A^g \dot{A} x_0 x_0^T + x_0 x_0^T \dot{A} A^g$ (a)

Consider $A^g x_0 = S \Lambda^g S^T x_0 = 0$.

Multiplying (a) by x_0 from the right: $A \dot{A}^g x_0 = A^g \dot{A} x_0$. Multiplying by A^g from the left: $AA^g \dot{A}^g x_0 = A^g A^g \dot{A} x_0$. Then $(1 - x_0 x_0^T) \dot{A}^g x_0 = A^g A^g \dot{A} x_0$. Then $\dot{A}^g x_0 = A^{-2} \dot{A} x_0 + x_0 x_0^T \dot{A}^g x_0$.

Let's multiply (a) by A^g from the left: $A^g \dot{A} A^g + A^g A \dot{A}^g = A^{-2} \dot{A} x_0 x_0^T + \cancel{A^g x_0}^0 x_0^T \dot{A} A^g$.

Consider $A^g A \dot{A}^g = \dot{A}^g - x_0 x_0^T \dot{A}^g = \dot{A}^g - x_0 x_0^T \dot{A} A^{-2} - x_0 x_0^T \dot{A}^g x_0 x_0^T$.

Consider $x_0^T \dot{A}^g x_0 = x_0^T (\dot{S} \Lambda^g S^T + S \Lambda^g \dot{S}^T) x_0 = x_0^T \dot{S} \Lambda^g S^T x_0 = 0$.

Then $\frac{d}{dt} A^{-1} = -A^{-1} \dot{A} A^{-1} + x_0 x_0^T \dot{A} A^{-2} + A^{-2} \dot{A} x_0 x_0^T$ (0.21).

Case $\text{Rg } A < n - 1$ is not considered since probability of such event is small.