

# 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  $f(\beta v) = \beta^2 v^T \overset{0}{\cancel{A_c v}} + 2\beta \underbrace{b_c^T v}_{\neq 0} \rightarrow -\infty$ ,  $\beta \rightarrow \infty$

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  $B_\varepsilon^+$  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  $c \in \mathbb{R}^m$

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  $f_i(x) = x^T A_i x + 2b_i^T x = \tilde{x}^T \tilde{A}_i \tilde{x} + 2\tilde{b}_i^T \tilde{x}$ .  $\tilde{A}_i = S^T A_i S$ .  $c\tilde{A} = \sum c_i \tilde{A}_i = \sum c_i S^T A_i S = S^T A_c S$ . 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) depends on  $\text{Rg } b$ .

#### 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  $c$ .

1. 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  $y \in Y$

$$(c_+, y) = x^T (\cancel{c_+} \cdot \overset{I}{A}) x + 2(\cancel{c_+} \cdot \overset{0}{b})^T x = \boxed{x^T x}$$

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

$$z(c) = \|A_c^g b_c\|^2 = \boxed{\left| (c \cdot A)^{-1} c \cdot b \right|^2}$$