On the feasibility for the system of quadratic equations, explanations

TODO: fix basis issue in 2 (with $(x, y) = x^T\Gamma y$), fix S issue $S^T A_c S = I$, check gradient equation.

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

Then $R_0 \subseteq \{b_c\}^\perp$
Consider $A = \sum\limits_{i=1}^{n} \lambda_i s_i s_i^T = S\Lambda S^T$, $S = ||s_1 ... s_n||$, $S^T S = E$, $s_i^T s_j = \delta_{ij}$.
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 $\Lambda^g = \operatorname{diag}(\lambda_1^g, ..., \lambda_n^g)$. 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:

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

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 $\inf_x (c, f(x))$ 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 $P - Q = \underline{x^T A_+ x} - x_0^T A_+ x_0 + 2b_+^T x - 2b_+^T x_0 - \underline{x^T A_+ x} - x_0^T A_+ x_0 + 2x^T A_+ x_0 = \big/ x_0 = -A_+^{-1} b_+ \big/ = 2x_0^T b_+ + 2b_+^T x - 2b_+^T x_0 - 2x^T b_+ = 0$.

Therefore, the image of B_ε^+ is a *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. $\begin{cases} \tilde{c} \cdot \tilde{A} = I & (1) \\ \tilde{c} \cdot \tilde{b} = 0 & (2) \end{cases}$ *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 $f_i(x) = x^T A_i x + 2b_i^T x = \tilde{x}^T S^T A_i S \tilde{x} + 2\tilde{b_i}^T S\tilde{x}$. $\tilde{A}_i = S^T A_i S$. $c\tilde{A} = \sum c_i \tilde{A}_i = \sum c_i S^T A_i S = S^T A_c S$. 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: $z(c) = \inf_{y \in Y}(c_+, y)$ where Y is an intersection of $f(\mathbb{R}^n)$ with a tangent hyperplane defined by its normal vector c.

Define $\sigma(Q) = \{\lambda \big| \dim \mathrm{Ker}(Q - \lambda E) > 0\}$. 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 $y\in Y$ $(c_+,y)=x^T(\cancelto{I}{c_+\cdot A})x+2(\cancelto{0}{c_+\cdot b})^T x=\boxed{x^Tx}$

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

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, $z(c)=\inf\limits_{y\in Y}(c_+,y)-\inf\limits_{y\in F}(c_+,y)$

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 $z(c+\alpha c_+)=|(c\cdot A+\alpha\cancelto{I}{c_+\cdot A})^g b_c|^2=|(c\cdot A+\alpha I)^g b_c|^2$. 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 $f\colon\mathbb{R}^n\to\mathbb{R}^m$, $f_i(x)=x^TA_ix+2b_i^Tx$, $A_i=A_i^T$, $m,n\geqslant 3$. 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. $b\perp\operatorname{Ker}A_c$, $e^0=-A_c^gb_c$
5. $f(\alpha e+e^0)=f^0+2\alpha f^1+\alpha^2f^2$, $f^1\nparallel f^2$

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

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

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

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.

8. Equations (0.18)-(0.21)

Consider $A(t)\colon n\times n$, $\exists\dot{A}$, $A^T=A$, $A\geqslant 0$, A has a simple zero eigenvalue: $\forall t\, A(t)x_0(t)=0$, $x_0^Tx_0=0$.

Then $A=S\Lambda S^T$, $S^TS=E$, $A^g=S\Lambda^gS^T$. 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 $AA^g=A^gA=S\Lambda S^TS\Lambda^gS^T=S\Lambda\Lambda^gS^T=\sum\limits_{\lambda_i\neq 0}s_is_i^T=1-x_0x_0^T$ (0.20)

Consider $\frac{d}{dt}Ax_0=\dot{A}x_0+A\dot{x}_0$. Multiplying by A^g from the left: $A^g\dot{A}x_0+A^gA\dot{x}_0=0$. Consider $A^gA\dot{x}_0=(1-x_0x_0^T)\dot{x}_0=\dot{x}_0-x_0x_0^T\dot{x}_0$. Since $||x_0||^2=x_0^Tx_0=1$, $x_0^T\dot{x}_0=0$. 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: $x_0^T\dot{A}x_0+\cancelto{0}{x_0^TA}\dot{x}_0=0$. Then $x_0^T\dot{A}x_0=0$ (0.18)

Consider $AA^g=1-x_0x_0^T$. Then $\dot{A}A^g+A\dot{A}^g=-\dot{x}_0x_0^T-x_0\dot{x}_0^T=A^g\dot{A}x_0x_0^T+x_0x_0^T\dot{A}A^g$ (a)

Consider $A^gx_0=S\Lambda^gS^Tx_0=0$.

Multiplying (a) by x_0 from the right: $A\dot{A}^gx_0=A^g\dot{A}x_0$. Multiplying by A^g from the left: $AA^g\dot{A}^gx_0=A^gA^g\dot{A}x_0$. Then $(1-x_0x_0^T)\dot{A}^gx_0=A^gA^g\dot{A}x_0$. Then $\dot{A}^gx_0=A^{-2}\dot{A}x_0+x_0x_0^T\dot{A}^gx_0$.

Let's multiply (a) by A^g from the left: $A^g\dot{A}A^g+A^gA\dot{A}^g=A^{-2}\dot{A}x_0x_0^T+\cancelto{0}{A^gx_0}x_0^T\dot{A}A^g$.

Consider $A^gA\dot{A}^g=\dot{A}^g-x_0x_0^T\dot{A}^g=\dot{A}^g-x_0x_0^T\dot{A}A^{-2}-x_0x_0^T\dot{A}^gx_0x_0^T$.

Consider $x_0^T \dot{A}^g x_0 = x_0^T(\dot{S}\Lambda^g S^T + S\dot{\Lambda}^g S^T + S\Lambda^g \dot{S}^T)x_0 = x_0^T \dot{S}\Lambda^g S^T x_0 = 0$.
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 Rg$A < n-1$ is not considered since probability of such event is small.

9. Gradient descent

1. (0.5): $\underbrace{(c \cdot A - \lambda_{\min}(c \cdot A))}_{Q} x_0 = 0$, then $\dot{x}_0 = -Q^{-1}\dot{Q}x_0$. $\frac{d}{dt}\lambda_{\min} = \frac{d}{dt}x_0^T(c \cdot A)x_0 = 2\dot{x}_0 \underbrace{(c \cdot A)x_0}_{\lambda_{\min}x_0} + x_0^T(\dot{c} \cdot A)x_0 = 2\lambda_{\min}\cancelto{0}{\dot{x}_0^T x_0} + x_0^T(\dot{c} \cdot A)x_0$.
   Then $\dot{x}_0 = -(A_c - \lambda_{\min}(Ac))^{-1}(\dot{c} \cdot A - x_0^T(\dot{c} \cdot A)x_0)x_0$ (correct).
2. (0.6). $x_0^T(c \cdot b) = 0$, use (0.5) (correct).
3. (0.7). $\frac{\partial}{\partial t}||v(c)||^2 = \frac{\partial}{\partial t}\sum_j v_j^2(c) = 2\sum_j v_j \frac{\partial}{\partial t}v_j = 2v^T(c)\frac{d}{dt}v(c(t))$ (correct). $v(c) = (\underbrace{c \cdot A - \lambda_{\min}(c \cdot A)}_{Q})^{-1}(c \cdot b) = Q^{-1}(c \cdot b)$.
4. (0.8) Define $Q = c \cdot A - \lambda_{\min}(c \cdot A)$. Define $v = Q^{-1}(c \cdot b)$. Then $z(c) = ||v||^2$ and $\dot{z} = 2v^T\dot{v}$.
   $\dot{v} = \dot{Q}^{-1}(c \cdot b) + Q^{-1}(\dot{c} \cdot b)$.
   Consider (0.21) $\dot{Q}^{-1} = -Q^{-1}\dot{Q}Q^{-1} + x_0 x_0^T \dot{Q}Q^{-2} + Q^{-2}\dot{Q}x_0 x_0^T$.
   Then $\dot{z} = 2v^T\left(Q^{-1}(\dot{c} \cdot b) + (-Q^{-1}\dot{Q}Q^{-1} + x_0 x_0^T\dot{Q}Q^{-2} + Q^{-2}\dot{Q}x_0x_0^T)(c \cdot b)\right) \boxed{=}$
   $\boxed{=} 2v^TQ^{-1}(\dot{c} \cdot b) - 2v^TQ^{-1}\dot{Q}Q^{-1}(c \cdot b) + 2\cancelto{0}{v^T x_0}\, x_0^T\dot{Q}Q^{-2}(c \cdot b) + 2v^TQ^{-2}\dot{Q}x_0\cancelto{0}{x_0^T(c \cdot b)} \boxed{=}$
   Since $x_0 \in \operatorname{Ker} Q \perp (c \cdot b)$, we have $x_0^T(c \cdot b) = 0$.
   Since $Qx_0 = 0$, $Q^{-1}x_0 = 0$: $Q^{-1}x_0 = S\Lambda^{-1}S^Tx_0 = S * 0 = 0$. Since $v^T = (c \cdot b)^TQ^{-1}$. Then $v^Tx_0 = 0$
   $\boxed{=} 2v^TQ^{-1}(\dot{c} \cdot b) - 2v^TQ^{-1}\dot{Q}\underbrace{Q^{-1}(c \cdot b)}_{v} = 2v^TQ^{-1}(\dot{c} \cdot b) - 2v^TQ^{-1}\dot{Q}v = \boxed{\dot{z} = 2v^TQ^{-1}(\dot{c} \cdot b - \dot{Q}v)}$,
   $\dot{Q} = \dot{c} \cdot A - x_0^T(\dot{c} \cdot A)x_0$
   Since $\dot{z} = \sum \frac{\partial z}{\partial c_i}\frac{\partial c_i}{\partial t}$, $\frac{\partial z}{\partial c_i}$ can be found as a coefficient at $\dot{c}_i$ in $\dot{z}$
   $\dot{z} = 2v^TQ^{-1}\sum_i\left(\dot{c}_ib_i - \dot{c}_iA_iv + x_0^T\dot{c}_iA_ix_0v\right) = \sum_i \dot{c}_i\left[2v^TQ^{-1}(b_i - (A_i - x_0^TA_ix_0)v)\right]$
   Thus, $\boxed{\dfrac{\partial z}{\partial c_i} = 2v^TQ^{-1}(b_i - (A_i - x_0^TA_ix_0)v)}$, $Q = c \cdot A - \lambda_{\min}(c \cdot A)$, $v = Q^{-1}(c \cdot b)$, $x_0 \in \operatorname{Ker} Q$, $||x_0|| = 1$
   not the same as (0.8) in draft.pdf:
   $\dot{z}_{(0.8)} = 2\underbrace{(c \cdot b)^TQ^{-2}}_{v^TQ^{-1}}(\dot{c} \cdot b) - v^T(Q^{-1}\dot{Q} + \dot{Q}Q^{-1})v = 2v^TQ^{-1}(\dot{c} \cdot b) - v^TQ^{-1}\dot{Q}v - v^T\dot{Q}Q^{-1}v$
5. (0.10). If $\dot{c} = \beta c_+$, then $\dot{z} = 2v^TQ^{-1}(\dot{c} \cdot b - \dot{Q}v) = \boxed{=}$. Since $c_+ \cdot b = 0$, $c_+ \cdot A = I$, $\dot{Q} = c_+ \cdot A - x_0^T(c_+ \cdot A)x_0 = I - x_0^Tx_0 = I - 1 = 0$. And $\boxed{=}0$ (correct with new $\dot{z}$).
6. (0.14) $n_i = \left(b_i^T - v^T(A_i - x_0^TA_ix_0)\right)x_0$
7. (0.16) $P(\lambda) = Q^{-1}Q = S\Lambda^gS^TS\Lambda S^T = S\Lambda\Lambda^gS^T = 1 - x_0x_0^T$ — projector on $(\operatorname{Ker} Q)^\perp$ (correct)
8. (0.15) $P(\lambda)(c(\lambda) \cdot b) = c(\lambda) \cdot b \Leftrightarrow \underbrace{(c(\lambda) \cdot b) \perp \operatorname{Ker} Q}_{\Leftrightarrow c(\lambda) \in c_{\text{bad}}} \Leftrightarrow c(\lambda) \cdot b \in \operatorname{Im}Q \Leftrightarrow \exists\hat{x}\colon Q\hat{x} = c(\lambda) \cdot b$ (0.17) (correct)

10. Gradient descent. Projection

We have $c \in \mathbb{R}^m$, $c \in c_{\text{bad}} = \{c \big| ||c|| = 1, \operatorname{Ker} Q(c) \perp (c \cdot b)\}$. $Q = c \cdot A - \lambda_{\min}(c \cdot A)$, $v = Q^{-1}(c \cdot b)$, $x_0 \in \operatorname{Ker} Q$, $||x_0|| = 1$

1. Calculate $\frac{\partial z}{\partial c_i} = 2v^T Q^{-1}(b_i - (A_i - x_0^T A_i x_0)v)$. Define $\Delta c = -\nabla z(c)$
2. Calculate $\hat{n}_i = \left(b_i^T - v^T(A_i - x_0^T A_i x_0)\right) x_0$, define $n_i = \frac{\hat{n}_i}{|\hat{n}|}$
3. Define $c' = c + \Delta c - n(\Delta c, n)$
4. Define $c(\lambda) = c' + \lambda n$. Define $x_0(\lambda)$ s.t. $x_0(\lambda) \in \operatorname{Ker} Q(\lambda)$, $x_0(\lambda)^T x_0 > 0$, $||x_0(\lambda)|| = 1$. Define $m(\lambda) = (c(\lambda) \cdot b)^T x_0(\lambda)$. Beware of $\mathrm{Rg} Q < n - 1$.
5. Find root of $m(\lambda)$ using binary search on $[-\lambda^0, \lambda^0]$, $\lambda^0 = ||c - c'||$.
6. Next c: $c(\lambda)$

$m(\lambda) = 0 \Leftrightarrow (c(\lambda) \cdot b) \perp \{x_0\} = \operatorname{Ker} Q$ if $\mathrm{Rg} Q = n - 1$.