Journal of the Operations Research
Society of Japan
2003, Vol. 46, No. 1, 22-34

© 2003 The Operations Research Society of Japan

ON THE STABILITY OF STATIONARY SOLUTIONS OF NONLINEAR POSITIVE SEMIDEFINITE PROGRAMS

Toshihiro Matsumoto
Teikyo University of Science and Technology

(Received Received June 12, 2000; Revised August 2, 2002)

Abstract In this paper we deal with strong stability of stationary solutions of nonlinear positive semidefinite programs. We prove two convergent properties of matrices sequences, and we give a sufficient condition for strong stability under the Linear Independence Constraint Qualification (LICQ) and the transversality condition.

Keywords: Nonlinear programming, optimization

1. Introduction

In this section we introduce nonlinear positive semidefinite programs. For its definition we prepare some notations:

$\boldsymbol{R}$: the field of all real numbers,

$\boldsymbol{R}^\ell$: the ℓ dimensional Euclidean space,

$\mathcal{M}(m,n)$: the set of all $m \times n$ real matrices,

$\mathcal{M}(n)$: the set of all $n \times n$ real matrices,

$S(n)$: the set of all $n \times n$ symmetric real matrices,

$S_+(n)$: the set of all $n \times n$ positive semidefinite symmetric real matrices,

$S_-(n)$: the set of all $n \times n$ negative semidefinite symmetric real matrices,

$S_{r,s}(n)$: the set of all $n \times n$ symmetric real matrices with r positive eigenvalues and s negative eigenvalues,

$O(n)$: the set of all $n \times n$ orthogonal real matrices,

$D(n)$: the set of all $n \times n$ diagonal real matrices,

$Diag(\gamma_1, \cdots, \gamma_n)$: an $n \times n$ diagonal matrix whose (i,i) component is γ_i $(1 \leq i \leq n)$,

$\boldsymbol{X}^T$: the transposition of the matrix $\boldsymbol{X}$,

$\boldsymbol{A} \bullet \boldsymbol{B}$: the trace form of $m \times n$ matrices $\boldsymbol{A} = (a_{ij})$ and $\boldsymbol{B} = (b_{ij})$, i.e.,

$$\boldsymbol{A} \bullet \boldsymbol{B} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij},$$

$t^+ = max\{t, 0\}$ for a real number t,

$t^- = min\{t, 0\}$ for a real number t,

$$\begin{aligned} int(A) &: \text{ the interior of a subset } A \text{ of a topological space } X, \\ c\ell(A) &: \text{ the closure of a subset } A \text{ of a topological space } X, \\ A+B &= \{a+b : a \in A, b \in B\} \text{ for subsets } A \text{ and } B \text{ of a vector space } V, \\ \mathcal{F} &= \{(f,h) = (f,h_1,\cdots,h_\ell) : f,h_1,\cdots,h_\ell \in C^2(S(n))\}, \\ & \quad \text{where } C^2(S(n)) \text{ is the set of all functions on } S(n) \text{ of } C^2 \text{ class.} \end{aligned}$$

Linear positive semidefinite programs (LSDP) are defined as follows:

$$\mathbf{Pro}_{(1)}(\boldsymbol{A},\boldsymbol{b},\boldsymbol{C}) \quad \left\| \begin{array}{ll} \text{minimize} & \boldsymbol{C} \bullet \boldsymbol{X} \\ \text{subject to} & \boldsymbol{X} \in S_+(n) \\ & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \quad (i = 1,\cdots,\ell) \end{array} \right\}, \tag{1}$$

where $\boldsymbol{C}, \boldsymbol{A}_i \in S(n)$ $(1 \leq i \leq \ell)$ and $\boldsymbol{b} = (b_1,\cdots,b_\ell) \in \boldsymbol{R}^\ell$.

LSDP has intensively been studied for this decade. For details, we recommend the bibliography of the paper [8].

We identify functions on $S(n)$ as those on $\mathcal{M}(n)$ satisfying $f(\boldsymbol{X}) = f(\boldsymbol{X}^T)$ $(\forall \boldsymbol{X} \in \mathcal{M}(n))$. In this situation, it is easily seen that $D_{\boldsymbol{X}} f(\boldsymbol{X}) \in S(n)$. We refer to the following programs as nonlinear positive semidefinite programs (NSDP):

$$\mathbf{Pro}_{(2)}(f,h) \quad \left\| \begin{array}{ll} \text{minimize} & f(\boldsymbol{X}) \\ \text{subject to} & \boldsymbol{X} \in S_+(n), \\ & h_i(\boldsymbol{X}) = 0 \quad (i = 1,\cdots,\ell). \end{array} \right\}, \tag{2}$$

where $(f,h) \in \mathcal{F}$.

In the paper [7] Kojima introduced for the first time the concept of strong stability of nonlinear programs which have finite equality constraints $h_i(\boldsymbol{x}) = 0$ $(i = 1,\cdots,\ell)$ and finite inequality constraints $g_j(\boldsymbol{x}) \geq 0$ $(j = 1,\cdots,m)$ with $h_i(\boldsymbol{x})$ and $g_j(\boldsymbol{x})$ twice continuous differentiable functions on $\boldsymbol{R}^n$ and satisfying the so called Mangasarian-Fromovitz condition, and gave an algebraic condition which is necessary and sufficient for strong stability by means of Jacobian and Hessian matrices. However, since LSDP and NSDP do not have such finite inequality constraints of C^2 class, we cannot apply Kojima's theory directly to LSDP and NSDP.

Definition 1.1. Let $\boldsymbol{X} \in S_+(n)$. The normal cone $\sigma(\boldsymbol{X})$ of $S_+(n)$ at $\boldsymbol{X}$ is defined by $\sigma(\boldsymbol{X}) = \{\boldsymbol{G} \in S(n) : (\boldsymbol{Y} - \boldsymbol{X}) \bullet \boldsymbol{G} \leq 0 \ (\forall \boldsymbol{Y} \in S_+(n))\}$ and $\boldsymbol{R}\sigma(\boldsymbol{X})$ denotes the affine space spanned by $\sigma(\boldsymbol{X})$.

$\boldsymbol{O}_r$ and $\boldsymbol{E}_r$ denote the $r \times r$ zero matrix and the $r \times r$ identity matrix respectively, and $\boldsymbol{O}$ without index denotes the zero matrix of an appropriate size. For $\mathcal{S} \subset S(r)$ and $\mathcal{T} \subset S(n-r)$, we define a set $\mathcal{S} \times \mathcal{T} = \left\{ \begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix} : \boldsymbol{A} \in \mathcal{S}, \boldsymbol{B} \in \mathcal{T} \right\}$. We abbreviate $\{\boldsymbol{O}_r\} \times S_{n-r}(n)$ to $\boldsymbol{O}_r \times S_{n-r}(n)$. The next lemma is easily proved and we omit its proof.

Lemma 1.2. *Let $\boldsymbol{X} \in S_+(n)$. Then the following (i), (ii), and (iii) hold.*

(i) $\sigma(\boldsymbol{X}) = \{\boldsymbol{G} \in S_-(n) : \boldsymbol{G} \bullet \boldsymbol{X} = 0\}$.

(ii) $\sigma(\boldsymbol{P}\boldsymbol{X}\boldsymbol{P}^T) = \boldsymbol{P}\sigma(\boldsymbol{X})\boldsymbol{P}^T = \{\boldsymbol{P}\boldsymbol{G}\boldsymbol{P}^T : \boldsymbol{G} \in \sigma(\boldsymbol{X})\}$ *holds for any* $\boldsymbol{P} \in O(n)$.

(iii) Let $\boldsymbol{X} \in S_+(n)$ *and* $rank\boldsymbol{X} = r$. *Suppose* $\boldsymbol{X} = \boldsymbol{P}\Gamma\boldsymbol{P}^T$, *where* $\boldsymbol{P} \in O(n)$ *and* $\Gamma = Diag(\alpha_1,\cdots,\alpha_r,0,\cdots,0)$ *with* $\alpha_1,\cdots,\alpha_r > 0$. *Then* $\sigma(\boldsymbol{X}) = \boldsymbol{P}(\boldsymbol{O}_r \times S_-(n-r))\boldsymbol{P}^T$ ∎

According to Sylvester's inertia law ([9]), for $\boldsymbol{X}, \boldsymbol{Y} \in S_{r,s}(n)$, there exists a nonsingular matrix $\boldsymbol{G}$ satisfying $\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X}\boldsymbol{G}^T$. By taking a local coordinate system as in the proof of Lemma 1.3 below, it is easily seen that $S_{r,s}(n)$ is an $\frac{(r+s)(2n-r-s+1)}{2}$ dimensional C^∞ submanifold of $S(n)$. Denote by $T_{\boldsymbol{X}}S_{r,s}(n)$ the tangent space of the manifold $S_{r,s}(n)$ at $\boldsymbol{X} \in S_{r,s}(n)$ and by $(T_{\boldsymbol{X}}S_{r,s}(n))^\perp = \{\boldsymbol{Z} \in S(n) : \boldsymbol{Z} \bullet \boldsymbol{Y} = 0 \ (\forall \boldsymbol{Y} \in T_{\boldsymbol{X}}S_{r,s}(n))\}$ the orthogonal complementary space of $T_{\boldsymbol{X}}S_{r,s}(n)$ in $S(n)$ with respect to the inner product defined by the trace form. The next lemma plays an important role in proving Lemma 2.24.

Lemma 1.3. *Let* $\boldsymbol{X} \in S_{r,0}(n)$ *and suppose that* $\boldsymbol{X} = \boldsymbol{P}\begin{pmatrix} \Gamma_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}\boldsymbol{P}^T$, *where* $\boldsymbol{P} \in O(n)$ *and* $\Gamma_{11} \in S_{r,0}(r)$. *Then the following (i) and (ii) hold.*

(i) $T_{\boldsymbol{X}}S_{r,0} = \left\{\boldsymbol{P}\begin{pmatrix} \dot{\boldsymbol{Y}}_{11} & \dot{\boldsymbol{Y}}_{21}^T \\ \dot{\boldsymbol{Y}}_{21} & \boldsymbol{O} \end{pmatrix}\boldsymbol{P}^T : \dot{\boldsymbol{Y}}_{11} \in S(r) \text{ and } \dot{\boldsymbol{Y}}_{21} \in \mathcal{M}(n-r,r)\right\}$.

(ii) $\boldsymbol{R}\sigma(\boldsymbol{X}) = (T_{\boldsymbol{X}}S_{r,0})^\perp$.

Proof: From the representation $\boldsymbol{X}$ as $\boldsymbol{X} = \boldsymbol{P}\begin{pmatrix} \Gamma_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}\boldsymbol{P}^T$,

$\mathcal{W}(\boldsymbol{X}) = \left\{\boldsymbol{P}\begin{pmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{21}^T \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{21}\boldsymbol{Y}_{11}^{-1}\boldsymbol{Y}_{21}^T \end{pmatrix}\boldsymbol{P}^T : \boldsymbol{Y}_{11} \in S_{r,0}(r) \text{ and } \boldsymbol{Y}_{21} \in \mathcal{M}(n-r,r)\right\}$ is an open neighborhood of $\boldsymbol{X}$ in $S_{r,0}(n)$, which implies that $S_{r,0}(n)$ is a smooth submanifold of $S(n)$. Since we can take $(\boldsymbol{Y}_{11}, \boldsymbol{Y}_{21}) \in S_{r,0}(r) \times \mathcal{M}(n-r,r)$ as a local coordinate system of $S_{r,0}(n)$ around $\boldsymbol{X}$, it is easily proved that

$$T_{\boldsymbol{X}}S_{r,0} = \left\{\boldsymbol{P}\begin{pmatrix} \dot{\boldsymbol{Y}}_{11} & \dot{\boldsymbol{Y}}_{21}^T \\ \dot{\boldsymbol{Y}}_{21} & \boldsymbol{O} \end{pmatrix}\boldsymbol{P}^T : \dot{\boldsymbol{Y}}_{11} \in S(r) \text{ and } \dot{\boldsymbol{Y}}_{21} \in \mathcal{M}(n-r,r)\right\}.$$

Hence $(T_{\boldsymbol{X}}S_{r,0})^\perp = \boldsymbol{P}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{P}^T$ holds. On the other hand, $\boldsymbol{R}\sigma(\boldsymbol{X}) = \boldsymbol{P}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{P}^T$ holds immediately from Lemma 1.2, which implies the assertion of this lemma. ∎

2. Strong Stability of Stationary Solutions of the Program $\mathbf{Pro}_{(2)}(f,h)$ under the LICQ Condition

In this section we investigate strong stability of stationary solutions of the program $\mathbf{Pro}_{(2)}(f,h)$. We will prove some convergent properties of matrix sequences by the inequality estimate which follows from Lemma 1.3, and by their means we will give a sufficient condition for strong stability in the sense of Kojima for stationary solutions of $\mathbf{Pro}_{(2)}(f,h)$ under the LICQ condition.

Definition 2.1.

(1) Let $\boldsymbol{Z} \in S(n)$. We define $\boldsymbol{Z}^+ \in S_+(n)$ and $\boldsymbol{Z}^- \in S_-(n)$ as follows. First we represent $\boldsymbol{Z}$ as $\boldsymbol{Z} = \boldsymbol{P}\Gamma\boldsymbol{P}^T$ with $\boldsymbol{P} \in O(n)$ and $\Gamma = Diag(\gamma_1, \cdots, \gamma_n)$. Then define $\Gamma^+ = Diag(\gamma_1^+, \cdots, \gamma_n^+)$, $\Gamma^- = Diag(\gamma_1^-, \cdots, \gamma_n^-)$, $\boldsymbol{Z}^+ = \boldsymbol{P}\Gamma^+\boldsymbol{P}^T$ and $\boldsymbol{Z}^- = \boldsymbol{P}\Gamma^-\boldsymbol{P}^T$. Those definitions of $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$ are independent of the representation $\boldsymbol{Z} = \boldsymbol{P}\Gamma\boldsymbol{P}^T$. In fact, if $\boldsymbol{Q} \in O(n)$ and $\boldsymbol{M} \in D(n)$ with $\boldsymbol{Q}\Gamma\boldsymbol{Q}^T = \boldsymbol{M}$, then it is easily proved that $\boldsymbol{Q}\Gamma^+\boldsymbol{Q}^T = \boldsymbol{M}^+$ and $\boldsymbol{Q}\Gamma^-\boldsymbol{Q}^T = \boldsymbol{M}^-$ hold. This leads to the well-definedness of $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$.

(2) We define $\rho^+, \rho^- : S(n) \to S(n)$ by $\rho^+(\boldsymbol{Z}) = \boldsymbol{Z}^+$ and $\rho^-(\boldsymbol{Z}) = \boldsymbol{Z}^-$.

Define $\|\boldsymbol{X}\| = \sqrt{\boldsymbol{X} \bullet \boldsymbol{X}}$ for $\boldsymbol{X} \in S(n)$. $\|\cdot\|$ is clearly a norm on $S(n)$. The following lemma is well-known and it characterizes $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$ ([9]).

Lemma 2.2. *Let* $\boldsymbol{Z} \in S(n)$ *and* $\boldsymbol{P} \in O(n)$*. Then following (i),(ii), and (iii) hold.*

(i) $(\boldsymbol{PZP}^T)^+ = \boldsymbol{PZ}^+\boldsymbol{P}^T$ *and* $(\boldsymbol{PZP}^T)^- = \boldsymbol{PZ}^-\boldsymbol{P}^T$.
$\boldsymbol{Z}^+ \in S_+(n), \boldsymbol{Z}^- \in S_-(n)$ *and* $\boldsymbol{Z}^+ \bullet \boldsymbol{Z}^- = 0$.

(ii) $\boldsymbol{Z}^+$ *is a unique matrix* $\boldsymbol{Y} \in S_+(n)$ *that minimizes* $\|\boldsymbol{Y} - \boldsymbol{Z}\|$.

(iii) $\boldsymbol{Z}^-$ *is a unique matrix* $\boldsymbol{Y} \in S_-(n)$ *that minimizes* $\|\boldsymbol{Y} - \boldsymbol{Z}\|$.

From the lemma we can see that $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$ are continuous with respect to $\boldsymbol{Z}$.

Definition 2.3. Let $\mathcal{H} = \{(\boldsymbol{X}, \boldsymbol{G}) \in S(n)\times S(n) : \boldsymbol{X} \in S_+(n) \text{ and } \boldsymbol{G} \in \sigma(\boldsymbol{X})\}$. Then from Lemma 1.2, $\mathcal{H} = \{(\boldsymbol{X}, \boldsymbol{G}) \in S(n) \times S(n) : \boldsymbol{X} \in S_+(n) \text{ and } \boldsymbol{G} \in S_-(n) \text{ and } \boldsymbol{X} \bullet \boldsymbol{G} = 0\}$. We define $\eta : \mathcal{H} \to S(n)$, $\rho : S(n) \to \mathcal{H}$ by $\eta(\boldsymbol{X}, \boldsymbol{G}) = \boldsymbol{X} + \boldsymbol{G}$ and $\eta(\boldsymbol{Z}) = (\boldsymbol{Z}^+, \boldsymbol{Z}^-)$.

Both ρ and η are continuous, and it is easily proved that $\rho \circ \eta = Id$ and $\eta \circ \rho = Id$, where Id's denote the identity maps on appropriate spaces. The next lemma follows from these relations.

Lemma 2.4. $\mathcal{H}$ *and* $S(n)$ *are homeomorphic to each other by* ρ *and* η*.*

Remark 2.5. Let $(\boldsymbol{X}, \boldsymbol{G}) \in \mathcal{H}$ and $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{G}$. Since both $(\boldsymbol{PZP}^T)^+ = \boldsymbol{PXP}^T$ and $(\boldsymbol{PZP}^T)^- = \boldsymbol{PGP}^T$ hold for $\boldsymbol{P} \in O(n)$, it is easily seen that $\boldsymbol{X}$ and $\boldsymbol{G}$ can be simultaneously diagonalized, i.e., there exists $\boldsymbol{P} \in O(n)$ satisfying $\boldsymbol{PXP}^T, \boldsymbol{PGP}^T \in D(n)$.

Definition 2.6. Let $(f, h) \in \mathcal{F}$. $\boldsymbol{R}D_{\boldsymbol{x}}h(\boldsymbol{X})$ denotes the affine space spanned by $\{D_{\boldsymbol{x}}h_i(\boldsymbol{X}) : i = 1, \cdots, \ell\}$. Then $\bar{\boldsymbol{X}} \in S_+(n)$ is called a stationary solution of the program $\mathbf{Pro}_{(2)}(f, h)$ if $-D_{\boldsymbol{x}}f(\bar{\boldsymbol{X}}) \in \boldsymbol{R}D_{\boldsymbol{x}}h(\bar{\boldsymbol{X}}) + \sigma(\bar{\boldsymbol{X}})$ holds. Also $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{G}}, \bar{\lambda}) \in \mathcal{H} \times \boldsymbol{R}^\ell$ is called a stationary point of the program $\mathbf{Pro}_{(2)}(f, h)$ if $D_{\boldsymbol{x}}f(\bar{\boldsymbol{X}}) + \sum_{i=1}^{\ell} \bar{\lambda}_i D_{\boldsymbol{x}}h_i(\bar{\boldsymbol{X}}) + \bar{\boldsymbol{G}} = \boldsymbol{O}$ holds. Identifying $\mathcal{H} \times \mathcal{F}$ with $S(n) \times \mathcal{F}$ by Lemma 2.4, $(\bar{\boldsymbol{Z}}, \bar{\lambda}) \in S(n) \times \boldsymbol{R}^\ell$ is also called a stationary point of the program $\mathbf{Pro}_{(2)}(f, h)$ if $(\rho(\bar{\boldsymbol{Z}}), \bar{\lambda})$ is a stationary point of the program $\mathbf{Pro}_{(2)}(f, h)$.

We prepare some notations for the remainder of this paper. For $(f, h) \in \mathcal{F}$, we define $\phi(\cdot, \cdot, \cdot; f, h) : S(n) \times S(n) \times \boldsymbol{R}^\ell \to S(n) \times \boldsymbol{R}^\ell$, $\psi(\cdot, \cdot; f, h) = \phi(\cdot, \cdot, \cdot; f, h) \circ (\rho \times Id) : S(n) \times \boldsymbol{R}^\ell \to S(n) \times \boldsymbol{R}^\ell$, $\Omega \subset S(n) \times \boldsymbol{R}^\ell \times \mathcal{F}$, $\Xi \subset S(n) \times \mathcal{F}$ and $\chi : \Omega \to \Xi$ as follows.

$$
\begin{aligned}
\phi(\boldsymbol{X}, \boldsymbol{G}, \lambda; f, h) &= (D_{\boldsymbol{x}}f(\boldsymbol{X}) + \sum_{i=1}^{\ell} \lambda_i D_{\boldsymbol{x}}h_i(\boldsymbol{X}) + \boldsymbol{G}, h(\boldsymbol{X})), \\
\psi(\boldsymbol{Z}, \lambda; f, h) &= \phi(\boldsymbol{Z}^+, \boldsymbol{Z}^-, \lambda; f, h) \\
&= (D_{\boldsymbol{x}}f(\boldsymbol{Z}^+) + \sum_{i=1}^{\ell} \lambda_i D_{\boldsymbol{x}}h_i(\boldsymbol{Z}^+) + \boldsymbol{Z}^-, h(\boldsymbol{Z}^+)), \\
\Omega &= \left\{ (\boldsymbol{Z}, \lambda, f, h) \in S(n) \times \boldsymbol{R}^\ell \times \mathcal{F} : \begin{array}{l} (\boldsymbol{Z}, \lambda) \text{ is a stationary point of} \\ \mathbf{Pro}_{(2)}(f, h) \end{array} \right\} \\
&= \{(\boldsymbol{Z}, \lambda, f, h) \in S(n) \times \boldsymbol{R}^\ell \times \mathcal{F} : \psi(\boldsymbol{Z}, \lambda, f, h) = (\boldsymbol{O}, \boldsymbol{0})\}, \\
&\quad \text{where } \boldsymbol{0} \text{ denotes the zero vector of } \boldsymbol{R}^\ell, \\
\Xi &= \{(\boldsymbol{X}, f, h) \in S(n) \times \mathcal{F} : \boldsymbol{X} \text{ is a stationary solution of } \mathbf{Pro}_{(2)}(f, h)\}, \\
\chi(\boldsymbol{Z}, \lambda, f, h) &= (\boldsymbol{Z}^+, f, h), \text{ i.e., } \chi : \Omega \to \Xi \text{ is a natural projection.}
\end{aligned}
$$

Remark 2.7. It is easily seen that $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{G}}, \bar{\lambda}) \in \mathcal{H} \times \boldsymbol{R}^\ell$ is a stationary point of the program $\mathbf{Pro}_{(2)}(f, h)$ if and only if $\phi(\bar{\boldsymbol{X}}, \bar{\boldsymbol{G}}, \bar{\lambda}; f, h) = (\boldsymbol{O}, \boldsymbol{0})$, and that $(\bar{\boldsymbol{Z}}, \bar{\lambda}) \in S(n) \times \boldsymbol{R}^\ell$ is a stationary point of the program $\mathbf{Pro}_{(2)}(f, h)$ if and only if $\psi(\bar{\boldsymbol{Z}}, \bar{\lambda}; f, h) = (\boldsymbol{O}, \boldsymbol{0})$.

Let $\boldsymbol{e}_{ij} \in \mathcal{M}(n)$ be the elementary matrix whose (i,j)-component is 1 and other components are all 0's. Then the Jacobian matrix of $f \in C^2(S(n))$ can be represented as

$$D_{\boldsymbol{X}} f(\boldsymbol{X}) = \sum_{i=1}^{n}\sum_{j=1}^{n} D_{x_{ij}} f(\boldsymbol{X})\boldsymbol{e}_{ij} \in S(n),$$

and the Hessian matrix of f can be represented as

$$D^2_{\boldsymbol{X}} f(\boldsymbol{X}) = \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{x_{pq}} D_{x_{ij}} f(\boldsymbol{X})\boldsymbol{e}_{pq} \otimes \boldsymbol{e}_{ij} \in S(n) \otimes S(n),$$

where $\otimes$ denotes the Kronecker product ([3]). In this situation, for symmetric matrices $\boldsymbol{A} = (a_{pq})$ and $\boldsymbol{Y} = (y_{ij})$, we obtain

$$\boldsymbol{A} \bullet D^2_{\boldsymbol{X}} f(\boldsymbol{X}) = \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{pq} D_{x_{pq}} D_{x_{ij}} f(\boldsymbol{X})\boldsymbol{e}_{ij} \in S(n), \text{ and}$$

$$\boldsymbol{A} \bullet D^2_{\boldsymbol{X}} f(\boldsymbol{X}) \bullet \boldsymbol{Y} = \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{pq} y_{ij} D_{x_{pq}} D_{x_{ij}} f(\boldsymbol{X}) \in \boldsymbol{R}.$$

The norms $\|D_{\boldsymbol{X}} f(\boldsymbol{X})\|$ and $\|D^2_{\boldsymbol{X}} f(\boldsymbol{X})\|$ are induced by the trace form, i.e.,

$$\|D_{\boldsymbol{X}} f(\boldsymbol{X})\| = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |D_{x_{ij}} f(\boldsymbol{X})|^2}, \text{ and}$$

$$\|D^2_{\boldsymbol{X}} f(\boldsymbol{X})\| = \sqrt{\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} |D_{x_{pq}} D_{x_{ij}} f(\boldsymbol{X})|^2}.$$

For $f \in C^2(S(n))$ and a subset $B \subset S(n)$, a norm $\|f\|_B$ is defined by

$$\|f\|_B = sup\{|f(\boldsymbol{X})|, \|D_{\boldsymbol{X}} f(\boldsymbol{X})\|, \|D^2_{\boldsymbol{X}} f(\boldsymbol{X})\| : \boldsymbol{X} \in B\}.$$

For $(f,h) \in \mathcal{F}$ and a subset $B \subset S(n)$, a norm $\|\cdot\|_B$ is defined by

$$\|(f,h)\|_B = max\{\|f(\boldsymbol{X})\|_B\ ,\ \|h_i(\boldsymbol{X})\|_B : 1 \leq i \leq \ell\}.$$

We denote by $\mathcal{F}_B$ the space $\mathcal{F}$ with $\|\cdot\|_B$-topology.

In general, given a normed vector space V with its norm $\|\cdot\|$, we define a closed ball and an open ball by $B_\delta(x) = \{y \in V : \|y - x\| \leq \delta\}$ and $int(B_\delta(x)) = \{y \in V : \|y - x\| < \delta\}$ for $x \in V$ and a positive real number $\delta > 0$.

Definition 2.8. ([7]) Let $\bar{\boldsymbol{X}} \in S_+(n)$ be a stationary solution of $\mathbf{Pro}_{(2)}(\bar{f},\bar{h})$. $\bar{\boldsymbol{X}}$ is said to be strongly stable if there exists $\delta^* > 0$ satisfying the following statement (*).

(*) For any real number δ with $0 < \delta \leq \delta^*$, there exists a real number $\alpha(\delta) > 0$ such that, for all $(f,h) \in \mathcal{F}$ satisfying $\|(f,h) - (\bar{f},\bar{h})\|_{B_{\delta^*}(\bar{\boldsymbol{X}})} < \alpha(\delta)$, $\mathbf{Pro}_{(2)}(f,h)$ has a unique stationary solution in $B_\delta(\bar{\boldsymbol{X}})$.

Remark 2.9. In the above statement (*), if we take $\alpha(\delta)$ satisfying $0 < \alpha(\delta) \leq \alpha(\delta^*)$, we see that $\mathbf{Pro}_{(2)}(f,h)$ has a unique stationary solution $\boldsymbol{X}$ in $B_{\delta^*}(\bar{\boldsymbol{X}})$ and that $\boldsymbol{X} \in B_\delta(\bar{\boldsymbol{X}})$ for all $(f,h) \in \mathcal{F}$ satisfying $\|(f,h) - (\bar{f},\bar{h})\|_{B_{\delta^*}(\bar{\boldsymbol{X}})} < \alpha(\delta)$.

From this remark, the above definition of strong stability is readily rephrased as follows.

Definition 2.10. Let $\bar{\boldsymbol{X}} \in S_+(n)$ be a stationary solution of $\mathbf{Pro}_{(2)}(\bar{f}, \bar{h})$. $\bar{\boldsymbol{X}}$ is said to be strongly stable if there exist a neighborhood $U = B_{\delta^*}(\bar{\boldsymbol{X}})$ of $\bar{\boldsymbol{X}}$ in $S(n)$ and a neighborhood V of $(\bar{f}, \bar{h})$ in $\mathcal{F}_U$ such that the natural projection $pr : \Xi \bigcap (U \times V) \to V$ is bijective and $pr^{-1} : V \to \Xi \bigcap (U \times V)$ is continuous at $(\bar{f}, \bar{h})$.

Next we introduce another stability which is a little stronger in its definition than strong stability.

Definition 2.11. Let $\bar{\boldsymbol{X}} \in S_+(n)$ be a stationary solution of $\mathbf{Pro}_{(2)}(\bar{f}, \bar{h})$. $\bar{\boldsymbol{X}}$ is said to be strictly strongly stable if there exist a neighborhood $U = B_{\delta^*}(\bar{\boldsymbol{X}})$ of $\bar{\boldsymbol{X}}$ in $S(n)$ and a neighborhood V of $(\bar{f}, \bar{h})$ in $\mathcal{F}_U$ such that the natural projection $pr : \Xi \bigcap (U \times V) \to V$ is a homeomorphism.

We will investigate relations between these stabilities for a while. The next condition is called the Mangasarian-Fromovitz condition.

Condition 2.12.

(i) $D_{\boldsymbol{X}} h_i(\boldsymbol{X}) \ \ (1 \leq i \leq \ell)$ are linearly independent.

(ii) There exists $\boldsymbol{W} \in S(n)$ satisfying

(a) $D_{\boldsymbol{X}} h_i(\boldsymbol{X}) \bullet \boldsymbol{W} = 0 \ \ (1 \leq i \leq \ell)$

(b) $\boldsymbol{G} \bullet \boldsymbol{W} > 0 \ \ (\forall \boldsymbol{G} \in \sigma(\boldsymbol{X}) : \boldsymbol{G} \neq \boldsymbol{O})$.

Proposition 2.13. *Under Condition 2.12, strong stability is equivalent to strictly strong stability.*

Proof: Strictly strong stability obviously implies strong stability. We prove only the converse below. Let $\bar{\boldsymbol{X}}, \bar{f}, \bar{h}, pr, U, V$ be as in Definition 2.10. Since $pr : \Xi \bigcap (U \times V) \to V$ is bijective, for any $(f, h) \in V$ there exists a unique element $\boldsymbol{X} \in U$ such that $(\boldsymbol{X}, f, h) \in \Xi$. Then define $\tau(f, h) = \boldsymbol{X}$, i.e., we define $\tau : V \to U$ by a relation $pr^{-1}(f, h) = (\tau(f, h), f, h) \ (\forall (f, h) \in V)$. Take a relatively compact neighborhood U_0 of $\bar{\boldsymbol{X}} = \tau(\bar{f}, \bar{h})$ with $c\ell(U_0) \subset U$. Because of the continuity of τ at $(\bar{f}, \bar{h})$, there exists a neighborhood V_0 of $(\bar{f}, \bar{h})$ in $\mathcal{F}_U$ satisfying $V_0 \subset V$ and $\tau(V_0) \subset U_0$. In the below we will prove that pr^{-1} is continuous on V_0.

On the contrary suppose there should exist $(f, h) \in V_0$ where pr^{-1} is not continuous. Then we may assume that there exists a sequence $(\boldsymbol{X}^{(k)}, f^{(k)}, h^{(k)}) \in \Xi \bigcap (U_0 \times V_0)$ satisfying $\lim_{k \to \infty} (f^{(k)}, h^{(k)}) = (f, h)$ in $\mathcal{F}_U$ and that the sequence $\boldsymbol{X}^{(k)}$ does not converge to $\boldsymbol{X} = \tau(f, h)$. Choose a sequence $(\boldsymbol{Z}^{(k)}, \lambda^{(k)}, f^{(k)}, h^{(k)}) \in \Omega$ such that $\boldsymbol{Z}^{(k)+} = \boldsymbol{X}^{(k)}$. Since $c\ell(U_0)$ is compact, taking a subsequence we may assume that $\lim_{k \to \infty} \boldsymbol{Z}^{(k)+} = \lim_{k \to \infty} \boldsymbol{X}^{(k)}$ exists $= \hat{\boldsymbol{X}} \in c\ell(U_0) \subset U$. We have to consider the following two cases (a) and (b).

(a): In case that $\{(\boldsymbol{Z}^{(k)}, \lambda^{(k)}) : k = 1, 2, \cdots\}$ is a bounded set of $S(n) \times \boldsymbol{R}^\ell$, taking a subsequence we may assume that $\lim_{k \to \infty} (\boldsymbol{Z}^{(k)}, \lambda^{(k)})$ exists $= (\boldsymbol{Z}, \lambda)$. Then $(\boldsymbol{Z}, \lambda, f, g) \in \Omega \bigcap ((\rho^+)^{-1}(U) \times \boldsymbol{R}^\ell \times \mathcal{F}_U)$ is easily proved, which means $\hat{\boldsymbol{X}} = \boldsymbol{Z}^+ = \tau(f, g) = \boldsymbol{X}$. This contradicts $\hat{\boldsymbol{X}} \neq \boldsymbol{X}$.

(b): In case that $\{(\boldsymbol{Z}^{(k)}, \lambda^{(k)}) : k = 1, 2, \cdots\}$ is not a bounded set of $S(n) \times R^\ell$, taking a subsequence we may assume $\lim_{k \to \infty} \|(\boldsymbol{Z}^{(k)-}, \lambda^{(k)})\| = \infty$. Taking a subsequence again, we may assume $\lim_{k \to \infty} \dfrac{(\boldsymbol{Z}^{(k)-}, \lambda^{(k)})}{\|(\boldsymbol{Z}^{(k)-}, \lambda^{(k)})\|}$ exists $= (\boldsymbol{Z}^-, \lambda) \neq (\boldsymbol{O}, \boldsymbol{0})$. It is easily seen that $\boldsymbol{Z}^- \in \sigma(\boldsymbol{Z}^+)$. Since $(\boldsymbol{Z}^{(k)}, \lambda^{(k)}, f^{(k)}, h^{(k)}) \in \Omega$, it follows that

$$\frac{D_{\boldsymbol{X}} f^{(k)}(\boldsymbol{Z}^{(k)+}) + \sum_{i=1}^{\ell} \lambda_i^{(k)} D_{\boldsymbol{X}} h_i^{(k)}(\boldsymbol{Z}^{(k)+}) + \boldsymbol{Z}^{(k)-}}{\|(\boldsymbol{Z}^{(k)-}, \lambda^{(k)})\|} = \boldsymbol{O}.$$

Taking a limit we have $\sum_{i=1}^{\ell} \lambda_i D_{\boldsymbol{X}} h_i(\boldsymbol{Z}^+) + \boldsymbol{Z}^- = \boldsymbol{O}$, which contradicts Condition 2.12.

We have proved the continuity of τ and pr^{-1} at (f, h). ∎

Hirabayashi et al.[5] pointed out that Condition 2.12 holds for every strongly stable point when linearly constrained nonlinear programs are concerned. We refer to the part (i) of Condition 2.12 as Condition 2.14.

Condition 2.14. $D_{\boldsymbol{X}} h_i(\boldsymbol{X}) \ \ (1 \le i \le \ell)$ are linearly independent.

Under Condition 2.14, $\mathcal{N}(h) = \{\boldsymbol{X} \in S(n) : h_i(\boldsymbol{X}) = 0 \ \ (1 \le i \le \ell)\}$ is an $(\frac{n(n+1)}{2} - \ell)$ dimensional C^2 submanifold of $S(n)$.

Condition 2.15. Suppose that Condition 2.14 holds. Let $\boldsymbol{X} \in S_{r,0}(n) \bigcap \mathcal{N}(h)$. Denote tangent spaces of manifolds $\mathcal{N}(h), S_{r,0}(n)$, and $S(n)$ at $\boldsymbol{X}$ by $T_{\boldsymbol{X}}\mathcal{N}(h)$, $T_{\boldsymbol{X}} S_{r,0}(n)$, and $T_{\boldsymbol{X}} S(n)$ respectively. If $T_{\boldsymbol{X}}\mathcal{N}(h) + T_{\boldsymbol{X}} S_{r,0}(n) = T_{\boldsymbol{X}} S(n)$ is satisfied, then we state that $\mathcal{N}(h)$ and $S_{r,0}(n)$ intersect transversally at $\boldsymbol{X}$ ([2][11]).

Since the orthogonal complementary space of $T_{\boldsymbol{X}} S_{r,0}(n)$ is $\boldsymbol{R}\sigma(\boldsymbol{X})$ from Lemma 1.3, Condition 2.15 is equivalent to that $\boldsymbol{R} D_{\boldsymbol{X}} h(\boldsymbol{X}) \bigcap \boldsymbol{R}\sigma(\boldsymbol{X}) = \{\boldsymbol{O}\}$. It is easily seen that a pair of Conditions 2.14 and 2.15 is stronger than Condition 2.12 and takes a role in the program $\mathbf{Pro}_{(2)}(f, h)$ as the LICQ condition does in the setting of the paper [7]. We assume Conditions 2.14 and 2.15 throughout in the remainder of this paper. Under these two conditions, any stationary solution corresponds to a unique stationary point. In fact, we can prove the next proposition. In order to prove Proposition 2.17, we need the next lemma whose proof is easy.

Lemma 2.16. *Let $\boldsymbol{A} = (a_{ij}) \in \mathcal{M}(n)$ and $\Gamma = Diag(\gamma_1, \cdots, \gamma_n)$. Suppose that there exist indices p, q such that $\gamma_p \ne \gamma_q$. Then $\boldsymbol{A}\Gamma = \Gamma\boldsymbol{A}$ implies $a_{pq} = a_{qp} = 0$.* ∎

Proposition 2.17. *Under Conditions 2.14 and 2.15, for any subset $U \subset S(n)$, $\chi : \Omega \bigcap ((\rho^+)^{-1}(U) \times \boldsymbol{R}^\ell \times \mathcal{F}_U) \to \Xi \bigcap (U \times \mathcal{F}_U)$ is a homeomorphism.*

Proof: Since χ is bijective, it suffices to prove the continuity of χ^{-1} with respect to $(\boldsymbol{Z}^+, f, h)$. Therefore we will prove that $\boldsymbol{Z}^-$ and λ are continuous with respect to $(\boldsymbol{Z}^+, f, h)$ in the below. Let $k = \frac{(n-r)(n-r+1)}{2}$ and $Grass(k, S(n))$ be a Grassmannian manifold which consists of all linear subspaces of dimension k in $S(n)$. Suppose that $(\bar{\boldsymbol{Z}}^+, \bar{f}, \bar{h}) \in \Xi \bigcap (U \times \mathcal{F}_U)$ and that $\bar{\boldsymbol{X}} = \bar{\boldsymbol{Z}}^+ \in S_{r,0}(n), i.e., rank\bar{\boldsymbol{X}} = r$. Let us represent $\bar{\boldsymbol{X}} = \bar{\boldsymbol{Z}}^+ = \bar{\boldsymbol{P}}\bar{\Gamma}\bar{\boldsymbol{P}}^T$ with $\bar{\Gamma} = Diag(\bar{\gamma}_1, \cdots, \bar{\gamma}_r, 0, \cdots, 0)$ and $\bar{\boldsymbol{P}} \in O(n)$. Set $\bar{\gamma}_{r+1} = \cdots = \bar{\gamma}_n = 0$. Define an open neighborhood $U_1(\epsilon)$ of $\bar{\boldsymbol{X}}$ by $U_1(\epsilon) = \{\boldsymbol{P} Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{P}^T : \boldsymbol{P} \in O(n)$ and $|\gamma_i - \bar{\gamma}_i| < \epsilon \ \ (\forall i = 1, \cdots, n)\}$. Choose ϵ with $0 < \epsilon < \frac{1}{2} min\{\bar{\gamma}_i : 1 \le i \le r\}$. Now suppose that $\boldsymbol{P} Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{P}^T \in U_1(\epsilon)$ for $\boldsymbol{P} \in O(n)$ with $\gamma_1 \ge \cdots \ge \gamma_n$. Then $\gamma_i \ne \gamma_j \ (1 \le \forall i \le r < \forall j \le n)$ follows from the definition of $U_1(\epsilon)$. Therefore if $\boldsymbol{P} Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{P}^T = \boldsymbol{Q} Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{Q}^T$, or equivalently $Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{P}^T\boldsymbol{Q} = \boldsymbol{P}^T\boldsymbol{Q} Diag(\gamma_1, \cdots, \gamma_n)$ holds for $\boldsymbol{P}, \boldsymbol{Q} \in O(n)$, then we can derive from Lemma 2.16 that $\boldsymbol{P}^T\boldsymbol{Q} \in O(r) \times O(n-r)$. Hence $\boldsymbol{P}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{P}^T = \boldsymbol{Q}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{Q}^T$. Define two continuous distributions $\boldsymbol{V}$ and $\boldsymbol{W}$ on $U_1(\epsilon)$, i.e., two continuous maps $\boldsymbol{V} : U_1(\epsilon) \to Grass(\ell, S(n))$ and $\boldsymbol{W} : U_1(\epsilon) \to Grass(k, S(n))$ by

$$\begin{aligned} \boldsymbol{V}_{\boldsymbol{X}} &= \boldsymbol{R} D_{\boldsymbol{X}} h(\boldsymbol{X}), \text{ and} \\ \boldsymbol{W}_{\boldsymbol{X}} &= \boldsymbol{P}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{P}^T \\ &= \left\{ \boldsymbol{P} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix} \boldsymbol{P}^T : \boldsymbol{C} \in S(n-r) \right\} \end{aligned}$$

for $\boldsymbol{X} = \boldsymbol{P} Diag(\gamma_1, \cdots, \gamma_n)\boldsymbol{P}^T \in U_1(\epsilon)$. Hence $\boldsymbol{W}_{\boldsymbol{X}}$ is well-defined and continuous with respect to $\boldsymbol{X} \in U_1(\epsilon)$. It is easily seen that $\sigma(\boldsymbol{X}) \subset \boldsymbol{W}_{\boldsymbol{X}}$. Since $\boldsymbol{V}_{\boldsymbol{X}} = (T_{\boldsymbol{X}}\mathcal{N}(h))^\perp$

and $\boldsymbol{W}_{\boldsymbol{X}} = (T_{\boldsymbol{X}} S_{r,0}(n))^{\perp}$, it is clear that $\boldsymbol{V}_{\boldsymbol{X}} \cap \boldsymbol{W}_{\boldsymbol{X}} = (T_{\boldsymbol{X}}\mathcal{N}(h) + T_{\boldsymbol{X}} S_{r,0}(n))^{\perp}$. From this relation we can deduce that $\boldsymbol{V}_{\bar{\boldsymbol{X}}} + \boldsymbol{W}_{\bar{\boldsymbol{X}}}$ is a direct sum. Therefore there exists a neighborhood $U_2 \subset U_1(\epsilon)$ of $\bar{\boldsymbol{X}}$ where $\boldsymbol{V}_{\boldsymbol{X}} + \boldsymbol{W}_{\boldsymbol{X}}$ is a direct sum, by continuity of $\boldsymbol{V}_{\boldsymbol{X}}$ and $\boldsymbol{W}_{\boldsymbol{X}}$ with respect $\boldsymbol{X}$. Hence we have a continuous map $F : U_2 \times S(n) \to S(n)$ such that $F(\boldsymbol{X}, \cdot) : S(n) \to S(n)$ is a linear map, $F(\boldsymbol{X}, \boldsymbol{Y}) \in \boldsymbol{W}_{\boldsymbol{X}} (\forall \boldsymbol{Y} \in S(n))$, $F(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{O} (\forall \boldsymbol{Y} \in \boldsymbol{V}_{\boldsymbol{X}})$ and $F(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y} (\forall \boldsymbol{Y} \in \boldsymbol{W}_{\boldsymbol{X}})$. Since $\psi(\boldsymbol{Z}, \lambda; f, h) = (\boldsymbol{O}, \boldsymbol{0})$ for $(\boldsymbol{Z}, \lambda; f, h) \in \Omega$, $-D_{\boldsymbol{X}} f(\boldsymbol{Z}^+) = \sum_{i=1}^{\ell} \lambda_i D_{\boldsymbol{X}} h_i(\boldsymbol{Z}^+) + \boldsymbol{Z}^- \in \boldsymbol{V}_{\boldsymbol{Z}^+} + \sigma(\boldsymbol{Z}^+) \subset \boldsymbol{V}_{\boldsymbol{Z}^+} + \boldsymbol{W}_{\boldsymbol{Z}^+}$. Therefore $\boldsymbol{Z}^- = F(\boldsymbol{Z}^+, -D_{\boldsymbol{X}} f(\boldsymbol{Z}^+))$ is continuous with respect to $(\boldsymbol{Z}^+, f, h)$ and so is $\sum_{i=1}^{\ell} \lambda_i D_{\boldsymbol{X}} h_i(\boldsymbol{Z}^+) = -D_{\boldsymbol{X}} f(\boldsymbol{Z}^+) - \boldsymbol{Z}^-$. It is also easily shown that λ is continuous with respect to $(\boldsymbol{Z}^+, f, h)$. ∎

Definition 2.18. We refer to $(\boldsymbol{Z}, \lambda)$ as a strongly stable stationary point of $\mathbf{Pro}_{(2)}(f, h)$ if and only if $\boldsymbol{Z}^+$ is a strongly stable stationary solution of $\mathbf{Pro}_{(2)}(f, h)$. From Propositions 2.13 and 2.17 we can restate the strong stability as follows.

> Let $(\bar{\boldsymbol{Z}}, \bar{\lambda}) \in S(n)$ be a stationary point of $\mathbf{Pro}_{(2)}(\bar{f}, \bar{h})$. Under Conditions 2.14 and 2.15, $(\bar{\boldsymbol{Z}}, \bar{\lambda})$ is strongly stable if and only if there exist a neighborhood $U = B_{\delta^*}(\bar{\boldsymbol{Z}}^+)$ of $\bar{\boldsymbol{Z}}^+$ in $S(n)$ and V of $(\bar{f}, \bar{h})$ in $\mathcal{F}_U$ such that the natural projection $\pi : \Omega \bigcap ((\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell} \times V) \to V$ is a homeomorphism.

The next proposition gives a sufficient condition for strong stability from a little different point of view, and it takes an important role in proving Theorem 2.21.

Proposition 2.19. *Let $(\bar{f}, \bar{h}) \in \mathcal{F}$ and $(\bar{\boldsymbol{X}}, \bar{\lambda}) \in S(n) \times \boldsymbol{R}^{\ell}$ be a stationary point of $\boldsymbol{Pro}_{(2)}(\bar{f}, \bar{h})$. Suppose there exists a neighborhood $U = B_{\delta^*}(\bar{\boldsymbol{X}}^+)$ of $\bar{\boldsymbol{X}}^+$ such that $V = \{(f, h) \in \mathcal{F} : \psi(\cdot, \cdot; f, h)$ is one-to-one on $(\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell}\}$ is a neighborhood of $(\bar{f}, \bar{h})$ in $\mathcal{F}_U$. Then $(\bar{\boldsymbol{X}}, \bar{\lambda})$ is strongly stable.*

Proof: Let ϵ be any positive real number with $0 < \epsilon \leq \delta^*$, and let $U_\epsilon = int(B_\epsilon(\bar{\boldsymbol{X}}^+))$. Since $\psi(\cdot, \cdot; f, h)$ is one-to-one on $(\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell}$, it follows from the Brouwer's invariance theorem of domain ([6]) that $\psi((\rho^+)^{-1}(U_\epsilon) \times \boldsymbol{R}^{\ell}; f, h)$ is an open set in $S(n) \times \boldsymbol{R}^{\ell}$ which is homeomorphic to $(\rho^+)^{-1}(U_\epsilon) \times \boldsymbol{R}^{\ell}$. Therefore $\psi((\rho^+)^{-1}(U_\epsilon); \bar{f}, \bar{h})$ includes $B_{\delta_0}((\boldsymbol{O}, \boldsymbol{0}))$ for some $\delta_0 > 0$ because $\psi(\bar{\boldsymbol{X}}, \bar{\lambda}; \bar{f}, \bar{h}) = (\boldsymbol{O}, \boldsymbol{0})$. Let $K = ((\rho^+)^{-1}(U_\epsilon) \times \boldsymbol{R}^{\ell}) \bigcap \psi(\cdot, \cdot; \bar{f}, \bar{h})^{-1}(B_{\delta_0}((\boldsymbol{O}, \boldsymbol{0})))$. Since K is homeomorphic to $B_{\delta_0}((\boldsymbol{O}, \boldsymbol{0}))$, K is compact. Define $d(f, h) = \inf\{\|\psi(\boldsymbol{X}, \lambda; f, h)\| : (\boldsymbol{X}, \lambda) \in K\}$, where $\|\cdot\|$ is the norm on $S(n) \times \boldsymbol{R}^{\ell}$. Since $\psi(\boldsymbol{X}, \lambda; f, h)$ is continuous with respect to $(\boldsymbol{X}, \lambda; f, h) \in (\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell} \times \mathcal{F}_U$ and K is compact, $d(f, h)$ is continuous with respect to $(f, h) \in \mathcal{F}_U$. $d(\bar{f}, \bar{h}) = 0$ follows from $\psi(\bar{\boldsymbol{X}}, \bar{\lambda}; \bar{f}, \bar{h}) = (\boldsymbol{O}, \boldsymbol{0})$, which implies that there exists a positive real number $\delta > 0$ such that $V_\delta = \{(f, h) \in \mathcal{F} : \|(f, h) - (\bar{f}, \bar{h})\|_U < \delta\} \subset V$ and $d(f, h) = 0$ $(\forall (f, h) \in V_\delta)$. Suppose that $(f, h) \in V_\delta$. Then, since $d(f, h) = 0$ and $\psi(\cdot, \cdot; f, h)$ is one-to-one on $(\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell}$, there exists a unique $(\boldsymbol{X}(f, h), \lambda(f, h)) \in (\rho^+)^{-1}(U) \times \boldsymbol{R}^{\ell}$ such that $\psi(\boldsymbol{X}(f, h), \lambda(f, h); f, h) = (\boldsymbol{O}, \boldsymbol{0})$. Clearly $(\boldsymbol{X}(f, h), \lambda(f, h)) \in K$. Therefore, $\boldsymbol{X}(f, h)^+ \in U_\epsilon$. We conclude that, for any positive real number ϵ with $0 < \epsilon \leq \delta^*$, there exists $\delta > 0$ such that $\boldsymbol{X}(f, h)^+ \in U_\epsilon$ and $\boldsymbol{X}(f, h)^+$ is a unique stationary point in U for $(f, h) \in V_\delta$. Therefore, $(\bar{X}, \bar{\lambda})$ is strongly stable by Definition 2.10 of the strong stability. ∎

Definition 2.20. Let $\boldsymbol{Z} = \boldsymbol{P}\Gamma\boldsymbol{P}^T \in S_-(n)$, where $\boldsymbol{P} \in O(n)$ and $\Gamma = Diag(\gamma_1, \cdots, \gamma_n)$ with $\gamma_1 = \cdots = \gamma_s = 0 > \gamma_{s+1}, \cdots, \gamma_n$. Define a linear subspace $\mathcal{V}(\boldsymbol{Z})$ of $S(n)$ by

$$\mathcal{V}(\boldsymbol{Z}) = \left\{ \boldsymbol{P} \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{21}^T \\ \boldsymbol{X}_{21} & \boldsymbol{O} \end{pmatrix} \boldsymbol{P}^T : \boldsymbol{X}_{11} \in S(s) \text{ and } \boldsymbol{X}_{21} \in \mathcal{M}(n-s, s) \right\}.$$

From Lemma 2.16, we see that $\mathcal{V}(\boldsymbol{Z})$ is independent of the representation $\boldsymbol{Z} = \boldsymbol{P}\Gamma\boldsymbol{P}^T$, and hence is well-defined. We give a sufficient condition for strong stability in the following theorem.

Theorem 2.21. *Let* $L(\boldsymbol{X},\lambda; f,h) = f(\boldsymbol{X}) + \sum_{i=1}^{\ell}\lambda_i h_i(\boldsymbol{X})$ *and* $(\bar{\boldsymbol{Z}},\bar{\lambda})$ *be a stationary point for* $\boldsymbol{Pro}_{(2)}(\bar{f},\bar{h})$*. Suppose that* $D^2_{\boldsymbol{X}}L(\bar{\boldsymbol{Z}}^+,\bar{\lambda};\bar{f},\bar{h})$ *be positive definite on the space* $Ker(D_{\boldsymbol{X}}h(\bar{\boldsymbol{Z}}^+))\bigcap\mathcal{V}(\bar{\boldsymbol{Z}}^-)$*. Then* $(\bar{\boldsymbol{Z}},\bar{\lambda})$ *is strongly stable.*

To prove this theorem we will prepare a series of lemmas.

Lemma 2.22. *Suppose there exist two sequences* $\boldsymbol{Z}^{(k)},\boldsymbol{W}^{(k)} \in S(n)$ $(k=1,2,\cdots)$ *such that* $\lim_{k\to\infty}\boldsymbol{Z}^{(k)} = \lim_{k\to\infty}\boldsymbol{W}^{(k)} = \bar{\boldsymbol{Z}}$*. If* $\lim_{k\to\infty}\dfrac{\boldsymbol{Z}^{(k)+}-\boldsymbol{W}^{(k)+}}{\|\boldsymbol{Z}^{(k)+}-\boldsymbol{W}^{(k)+}\|} = \boldsymbol{A}$*, then* $\boldsymbol{A}\in\mathcal{V}(\bar{\boldsymbol{Z}}^-)$

Proof: Represent $\bar{\boldsymbol{Z}} = \bar{\boldsymbol{P}}\bar{\Gamma}\bar{\boldsymbol{P}}^T$ and $\boldsymbol{Z}^{(k)} = \boldsymbol{T}^{(k)}\Gamma^{(k)}\boldsymbol{T}^{(k)T}$ and $\boldsymbol{W}^{(k)} = \boldsymbol{Q}^{(k)}\boldsymbol{M}^{(k)}\boldsymbol{Q}^{(k)T}$, where $\bar{\boldsymbol{P}},\boldsymbol{T}^{(k)},\boldsymbol{Q}^{(k)}\in O(n)$ and $\bar{\Gamma} = Diag(\bar{\gamma}_1,\cdots,\bar{\gamma}_n), \Gamma^{(k)} = Diag(\gamma_1^{(k)},\cdots,\gamma_n^{(k)}), \boldsymbol{M}^{(k)} = Diag(\mu_1^{(k)},\cdots,\mu_n^{(k)})$ with $\bar{\gamma}_1\ge\cdots\ge\bar{\gamma}_s\ge 0>\bar{\gamma}_{s+1}\ge\cdots\ge\bar{\gamma}_n$ and $\gamma_1^{(k)}\ge\cdots\ge\gamma_n^{(k)}$ and $\mu_1^{(k)}\ge\cdots\ge\mu_n^{(k)}$. Clearly $\lim_{k\to\infty}\Gamma^{(k)} = \lim_{k\to\infty}\boldsymbol{M}^{(k)} = \bar{\Gamma}$ holds from the continuity of eigenvalues. Hence we may assume $\gamma_i^{(k)}\ne\gamma_j^{(k)},\ \mu_i^{(k)}\ne\mu_j^{(k)}, (i\le s<j)$ and $0>\gamma_{s+1}^{(k)}\ge\cdots\ge\gamma_n^{(k)}$ and $0>\mu_{s+1}^{(k)}\ge\cdots\ge\mu_n^{(k)}$.

By taking a subsequence, we may assume that $\lim_{k\to\infty}\boldsymbol{T}^{(k)} = \bar{\boldsymbol{T}}$ and $\lim_{k\to\infty}\boldsymbol{Q}^{(k)} = \bar{\boldsymbol{Q}}$ exist, so that we have $\bar{\boldsymbol{T}},\bar{\boldsymbol{Q}}\in O(n)$ and $\bar{\boldsymbol{T}}\bar{\Gamma}\bar{\boldsymbol{T}}^T = \bar{\boldsymbol{Q}}\bar{\Gamma}\bar{\boldsymbol{Q}}^T = \bar{\boldsymbol{P}}\bar{\Gamma}\bar{\boldsymbol{P}}^T$. Since $\bar{\gamma}_i\ne\bar{\gamma}_j$ $(1\le\forall i\le s<\forall j\le n)$, it follows from Lemma 2.16 that $\bar{\boldsymbol{T}}^T\bar{\boldsymbol{P}},\bar{\boldsymbol{Q}}^T\bar{\boldsymbol{P}},\bar{\boldsymbol{Q}}^T\bar{\boldsymbol{T}}\in O(s)\times O(n-s)$. Let $\boldsymbol{C}^{(k)} = \boldsymbol{T}^{(k)T}\boldsymbol{Q}^{(k)} = \begin{pmatrix}\boldsymbol{C}_{11}^{(k)} & \boldsymbol{C}_{12}^{(k)}\\ \boldsymbol{C}_{21}^{(k)} & \boldsymbol{C}_{22}^{(k)}\end{pmatrix}$ and $\bar{\boldsymbol{C}} = \bar{\boldsymbol{T}}^T\bar{\boldsymbol{Q}} = \lim_{k\to\infty}\boldsymbol{C}^{(k)} = \begin{pmatrix}\bar{\boldsymbol{C}}_{11} & \bar{\boldsymbol{C}}_{12}\\ \bar{\boldsymbol{C}}_{21} & \bar{\boldsymbol{C}}_{22}\end{pmatrix} = \begin{pmatrix}\bar{\boldsymbol{C}}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \bar{\boldsymbol{C}}_{22}\end{pmatrix}$, where $\boldsymbol{C}_{11}^{(k)}\in\mathcal{M}(s)$, $\boldsymbol{C}_{12}^{(k)}\in\mathcal{M}(s,n-s)$, $\boldsymbol{C}_{21}^{(k)}\in\mathcal{M}(n-s,s)$, and $\boldsymbol{C}_{22}^{(k)}\in\mathcal{M}(n-s)$.

It is clear that $\boldsymbol{Z}^{(k)+} = \boldsymbol{T}^{(k)}\Gamma^{(k)+}\boldsymbol{T}^{(k)T}$ and $\boldsymbol{W}^{(k)+} = \boldsymbol{Q}^{(k)}\boldsymbol{M}^{(k)+}\boldsymbol{Q}^{(k)T}$, where $\Gamma^{(k)+} = Diag(\gamma_1^{(k)+},\cdots,\gamma_s^{(k)+},0,\cdots,0)$ and $\boldsymbol{M}^{(k)+} = Diag(\mu_1^{(k)+},\cdots,\mu_s^{(k)+},0,\cdots,0)$. Now consider the following limit

$$
\begin{aligned}
\lim_{k\to\infty}\frac{\boldsymbol{C}^{(k)T}\Gamma^{(k)+}\boldsymbol{C}^{(k)}-\boldsymbol{M}^{(k)+}}{\|\boldsymbol{Z}^{(k)+}-\boldsymbol{W}^{(k)+}\|} &= \lim_{k\to\infty}\frac{\boldsymbol{Q}^{(k)}(\boldsymbol{Z}^{(k)+}-\boldsymbol{W}^{(k)+})\boldsymbol{Q}^{(k)T}}{\|\boldsymbol{Z}^{(k)+}-\boldsymbol{W}^{(k)+}\|}\\
&= \bar{\boldsymbol{Q}}\boldsymbol{A}\bar{\boldsymbol{Q}}^T\\
&= \begin{pmatrix}\boldsymbol{D}_{11} & \boldsymbol{D}_{12}\\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22}\end{pmatrix}.
\end{aligned}
$$

By block calculation of matrices, we have $\boldsymbol{C}^{(k)T}\Gamma^{(k)+}\boldsymbol{C}^{(k)}-\boldsymbol{M}^{(k)+} = \begin{pmatrix}\boldsymbol{X}_{11}^{(k)} & \boldsymbol{X}_{21}^{(k)T}\\ \boldsymbol{X}_{21}^{(k)} & \boldsymbol{X}_{22}^{(k)}\end{pmatrix}$, where

$$
\left.\begin{aligned}
\boldsymbol{X}_{21}^{(k)} &= \boldsymbol{C}_{12}^{(k)T}Diag(\gamma_1^{(k)+},\cdots,\gamma_s^{(k)+})\boldsymbol{C}_{11}^{(k)},\\
\boldsymbol{X}_{22}^{(k)} &= \boldsymbol{C}_{12}^{(k)T}Diag(\gamma_1^{(k)+},\cdots,\gamma_s^{(k)+})\boldsymbol{C}_{12}^{(k)}.
\end{aligned}\right\}
$$

Comparing these two equations, it is easily seen that $\boldsymbol{X}_{22}^{(k)} = \boldsymbol{X}_{21}^{(k)}\boldsymbol{C}_{11}^{(k)-1}\boldsymbol{C}_{12}^{(k)}$ holds for every sufficiently large k, from which we can derive $\boldsymbol{D}_{22} = \boldsymbol{D}_{21}\bar{\boldsymbol{C}}_{11}^{-1}\bar{\boldsymbol{C}}_{12} = \boldsymbol{D}_{21}\bar{\boldsymbol{C}}_{11}^{-1}\boldsymbol{O} = \boldsymbol{O}$. Since $\bar{\boldsymbol{Q}}^T\bar{\boldsymbol{P}}\in O(s)\times O(n-s)$ and $\boldsymbol{A} = \bar{\boldsymbol{P}}(\bar{\boldsymbol{P}}^T\bar{\boldsymbol{Q}})(\bar{\boldsymbol{Q}}^T\boldsymbol{A}\bar{\boldsymbol{Q}})(\bar{\boldsymbol{Q}}^T\bar{\boldsymbol{P}})\bar{\boldsymbol{P}}^T$, it follows that $\boldsymbol{A}\in\mathcal{V}(\bar{\boldsymbol{Z}}^-)$. ∎

Lemma 2.23. *Let* $\boldsymbol{X},\boldsymbol{Y}\in S_+(r)$*. Then* $\sigma(\boldsymbol{X}+\boldsymbol{Y}) = \sigma(\boldsymbol{X})\bigcap\sigma(\boldsymbol{Y})$ *holds.*

Proof: Suppose $\boldsymbol{G}\in\sigma(\boldsymbol{X}+\boldsymbol{Y})$, i.e., $\boldsymbol{G}\in S_-(n)$ and $(\boldsymbol{X}+\boldsymbol{Y})\bullet\boldsymbol{G} = 0$. Since $\boldsymbol{X}\in S_+(n)$ and $\boldsymbol{Y}\in S_+(n)$ and $\boldsymbol{G}\in S_-(n)$, $\boldsymbol{X}\bullet\boldsymbol{G}\le 0$ and $\boldsymbol{Y}\bullet\boldsymbol{G}\le 0$ hold. Therefore $(\boldsymbol{X}+\boldsymbol{Y})\bullet\boldsymbol{G} = 0$ implies $\boldsymbol{X}\bullet\boldsymbol{G} = \boldsymbol{Y}\bullet\boldsymbol{G} = 0$, i.e., $\boldsymbol{G}\in\sigma(\boldsymbol{X})\bigcap\sigma(\boldsymbol{Y})$. The converse can be similarly proved.

∎

Lemma 2.24. *Suppose there exist two sequences* $\boldsymbol{Z}^{(k)}, \boldsymbol{W}^{(k)} \in S(n)$ $(k = 1, 2, \cdots)$ *such that* $\lim_{k\to\infty} \boldsymbol{Z}^{(k)} = \lim_{k\to\infty} \boldsymbol{W}^{(k)} = \bar{\boldsymbol{Z}}$*. If* $\lim_{k\to\infty} \frac{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = 0$ *and* $\lim_{k\to\infty} \frac{\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = \boldsymbol{B}$*, then* $\boldsymbol{B} \in \boldsymbol{R}\sigma(\bar{\boldsymbol{Z}}^{+})$*.*

Proof: Represent $\bar{\boldsymbol{Z}}, \boldsymbol{W}^{(k)}$ such as $\bar{\boldsymbol{Z}} = \bar{\boldsymbol{P}}\bar{\Gamma}\bar{\boldsymbol{P}}^T, \boldsymbol{W}^{(k)} = \boldsymbol{Q}^{(k)}\boldsymbol{M}^{(k)}\boldsymbol{Q}^{(k)T}$, where $\bar{\boldsymbol{P}}, \boldsymbol{Q}^{(k)} \in O(n)$ and $\bar{\Gamma} = Diag(\bar{\gamma}_1, \cdots, \bar{\gamma}_n), \boldsymbol{M}^{(k)} = Diag(\mu_1^{(k)}, \cdots, \mu_n^{(k)})$ with $\bar{\gamma}_1 \ge \cdots \ge \bar{\gamma}_r > 0 \ge \bar{\gamma}_{r+1} \ge \cdots \ge \bar{\gamma}_n$ and $\mu_1^{(k)} \ge \cdots \ge \mu_n^{(k)}$. Clearly $\lim_{k\to\infty} \boldsymbol{M}^{(k)} = \bar{\Gamma}$ and we may assume $\mu_i^{(k)} > 0$ $(\forall i \le r, \forall k)$, which implies $\boldsymbol{R}\sigma(\boldsymbol{W}^{(k)+}) \subset \boldsymbol{Q}^{(k)}(\boldsymbol{O}_r \times S(n-r))\boldsymbol{Q}^{(k)T}$. By taking a subsequence, we may assume that $\lim_{k\to\infty} \boldsymbol{Q}^{(k)} = \bar{\boldsymbol{Q}}$. Then $\bar{\boldsymbol{Q}} \in O(n)$ and $\bar{\boldsymbol{Q}}\bar{\Gamma}\bar{\boldsymbol{Q}}^T = \bar{\boldsymbol{P}}\bar{\Gamma}\bar{\boldsymbol{P}}^T$ hold. Since $\bar{\gamma}_i \ne \bar{\gamma}_j$ $(1 \le \forall i \le r < \forall j \le n)$, it follows from Lemma 2.16 that $\bar{\boldsymbol{Q}}^T\bar{\boldsymbol{P}} \in O(r) \times O(n-r)$.

Let $\epsilon^{(k)} = \frac{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|}$. From the assumption, $\lim_{k\to\infty} \epsilon^{(k)} = 0$ holds. Using $\boldsymbol{Q}^k$ above represent $\boldsymbol{Z}^{(k)+}$ as $\boldsymbol{Z}^{(k)+} = \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} & \boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{K}_{21}^{(k)} & \boldsymbol{K}_{22}^{(k)} \end{pmatrix} \boldsymbol{Q}^{(k)T}$, where $\boldsymbol{K}_{11}^{(k)} \in S(r)$, $\boldsymbol{K}_{21}^{(k)} \in \mathcal{M}(n-r, r)$, and $\boldsymbol{K}_{22}^{(k)} \in S(n-r)$. Since $\lim_{k\to\infty} \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} & \boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{K}_{21}^{(k)} & \boldsymbol{K}_{22}^{(k)} \end{pmatrix} = \bar{\boldsymbol{Q}}^T\bar{\boldsymbol{P}}\bar{\Gamma}^+\bar{\boldsymbol{P}}^T\bar{\boldsymbol{Q}}$, we may suppose $\boldsymbol{K}_{11}^{(k)}$ be positive definite. So we can calculate

$$\begin{aligned} \boldsymbol{Z}^{(k)+} &= \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} & \boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{K}_{21}^{(k)} & \boldsymbol{K}_{22}^{(k)} \end{pmatrix} \boldsymbol{Q}^{(k)T} \\ &= \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{O} & \boldsymbol{E}_{n-r} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{K}_{22}^{(k)} - \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \end{pmatrix} \\ &\quad \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{O} & \boldsymbol{E}_{n-r} \end{pmatrix} \boldsymbol{Q}^{(k)T}, \end{aligned}$$

from which it follows $\boldsymbol{H}_{22}^{(k)} = \boldsymbol{K}_{22}^{(k)} - \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \in S_+(n-r)$. Define

$$\begin{aligned} \hat{\boldsymbol{Z}}^{(k)+} &= \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} & \boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{K}_{21}^{(k)} & \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \end{pmatrix} \boldsymbol{Q}^{(k)T} \in S_{r,0}(n), \text{ and} \\ \boldsymbol{H}^{(k)} &= \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{O} & \boldsymbol{E}_{n-r} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{H}_{22}^{(k)} \end{pmatrix} \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{O} & \boldsymbol{E}_{n-r} \end{pmatrix} \boldsymbol{Q}^{(k)T} \\ &= \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{H}_{22}^{(k)} \end{pmatrix} \boldsymbol{Q}^{(k)T} \in S_+(n). \end{aligned}$$

The equation $\boldsymbol{Z}^{(k)+} = \hat{\boldsymbol{Z}}^{(k)+} + \boldsymbol{H}^{(k)}$ implies $\sigma(\boldsymbol{Z}^{(k)+}) \subset \sigma(\hat{\boldsymbol{Z}}^{(k)+})$ by virtue of Lemma 2.23.

Represent $\boldsymbol{Z}^{(k)-}$ as $\boldsymbol{Z}^{(k)-} = \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{Y}_{11}^{(k)} & \boldsymbol{Y}_{21}^{(k)T} \\ \boldsymbol{Y}_{21}^{(k)} & \boldsymbol{Y}_{22}^{(k)} \end{pmatrix} \boldsymbol{Q}^{(k)T}$. Then, from the above, we readily get $\boldsymbol{Z}^{(k)-} \in \sigma(\boldsymbol{Z}^{(k)+}) \subset \sigma(\hat{\boldsymbol{Z}}^{(k)+})$. Since $\boldsymbol{R}\sigma(\hat{\boldsymbol{Z}}^{(k)+}) = (T_{\hat{\boldsymbol{Z}}^{(k)+}}S_{r,0}(n))^{\perp}$ from Lemma 1.3 and $S_{r,0}(n)$ can be represented as a manifold $\mathcal{W}(\hat{\boldsymbol{Z}}^{(k)+})$ around $\hat{\boldsymbol{Z}}^{(k)+}$ as in the proof of Lemma 1.3, we can prove

$$\boldsymbol{R}\sigma(\hat{\boldsymbol{Z}}^{(k)+}) = \left\{ \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{21}^T \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22} \end{pmatrix} \boldsymbol{Q}^{(k)T} : \begin{array}{c} \begin{pmatrix} \dot{\boldsymbol{K}}_{11} & \dot{\boldsymbol{K}}_{21}^T \\ \dot{\boldsymbol{K}}_{21} & \dot{\boldsymbol{K}}_{22} \end{pmatrix} \bullet \begin{pmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{21}^T \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22} \end{pmatrix} = 0 \\ (\forall \dot{\boldsymbol{K}}_{11} \in S(r), \forall \dot{\boldsymbol{K}}_{21} \in \mathcal{M}(n-r, r)) \end{array} \right\}, \quad (3)$$

where $\dot{\boldsymbol{K}}_{22} = \dot{\boldsymbol{K}}_{21}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} + \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\dot{\boldsymbol{K}}_{21}^{T} - \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\dot{\boldsymbol{K}}_{11}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T}$. From (3) and $\boldsymbol{Z}^{(k)+} \in \sigma(\hat{\boldsymbol{Z}}^{(k)+})$, we can derive

$$\left.\begin{array}{rcl} \dot{\boldsymbol{K}}_{11} \bullet \boldsymbol{Y}_{11}^{(k)} - \boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\dot{\boldsymbol{K}}_{11}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \bullet \boldsymbol{Y}_{22}^{(k)} &=& 0, \\ \dot{\boldsymbol{K}}_{21} \bullet \boldsymbol{Y}_{21}^{(k)} + \dot{\boldsymbol{K}}_{21}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \bullet \boldsymbol{Y}_{22}^{(k)} &=& 0 \end{array}\right\}. \tag{4}$$

Write explicitly $\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}$ and $\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}$ as

$$\left.\begin{array}{rcl} \boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+} &=& \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{K}_{11}^{(k)} - Diag(\mu_1, \cdots, \mu_r) & \boldsymbol{K}_{21}^{(k)T} \\ \boldsymbol{K}_{21}^{(k)} & \boldsymbol{K}_{22}^{(k)} - Diag(\mu_{r+1}^{+}, \cdots, \mu_n^{+}) \end{pmatrix} \boldsymbol{Q}^{(k)T}, \\ \boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-} &=& \boldsymbol{Q}^{(k)} \begin{pmatrix} \boldsymbol{Y}_{11}^{(k)} & \boldsymbol{Y}_{21}^{(k)T} \\ \boldsymbol{Y}_{21}^{(k)} & \boldsymbol{Y}_{22}^{(k)} - Diag(\mu_{r+1}^{-}, \cdots, \mu_n^{-}) \end{pmatrix} \boldsymbol{Q}^{(k)T}. \end{array}\right\}$$

By the first equality above, we have the inequality $\|\boldsymbol{K}_{21}^{(k)}\| \leq \|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|$. From the equation (4), we can prove

$$\begin{array}{rcl} \lim\limits_{k\to\infty} \dfrac{|\dot{\boldsymbol{K}}_{11} \bullet \boldsymbol{Y}_{11}^{(k)}|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} &\leq& \lim\limits_{k\to\infty} \epsilon^{(k)} \dfrac{|\dot{\boldsymbol{K}}_{11} \bullet \boldsymbol{Y}_{11}^{(k)}|}{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|} \\ &=& \lim\limits_{k\to\infty} \epsilon^{(k)} \dfrac{|\boldsymbol{K}_{21}^{(k)}\boldsymbol{K}_{11}^{(k)-1}\dot{\boldsymbol{K}}_{11}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \bullet \boldsymbol{Y}_{22}^{(k)}|}{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|} \\ &=& M \lim\limits_{k\to\infty} \epsilon^{(k)} \|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\| = 0, \end{array}$$

where M is an appropriate constant. It follows $\lim\limits_{k\to\infty} \dfrac{\|\boldsymbol{Y}_{11}^{(k)}\|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = 0$.
Similarly we can prove

$$\begin{array}{rcl} \lim\limits_{k\to\infty} \dfrac{|\dot{\boldsymbol{K}}_{21} \bullet \boldsymbol{Y}_{21}^{(k)}|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} &\leq& \lim\limits_{k\to\infty} \epsilon^{(k)} \dfrac{|\dot{\boldsymbol{K}}_{21} \bullet \boldsymbol{Y}_{21}^{(k)}|}{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|} \\ &\leq& \lim\limits_{k\to\infty} \epsilon^{(k)} \dfrac{|\dot{\boldsymbol{K}}_{21}\boldsymbol{K}_{11}^{(k)-1}\boldsymbol{K}_{21}^{(k)T} \bullet \boldsymbol{Y}_{22}^{(k)}|}{\|\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}\|} \\ &\leq& N \lim\limits_{k\to\infty} \epsilon^{(k)} = 0, \end{array}$$

where N is an appropriate constant. It also follows $\lim\limits_{k\to\infty} \dfrac{\|\boldsymbol{Y}_{21}^{(k)}\|}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = 0$.
Therefore we can write $\lim\limits_{k\to\infty} \dfrac{\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = \bar{\boldsymbol{Q}} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & * \end{pmatrix} \bar{\boldsymbol{Q}}^{T}$.

Since $\bar{\boldsymbol{Q}}^{T}\bar{\boldsymbol{P}} \in O(r) \times O(n-r)$, we have proved
$$\lim_{k\to\infty} \frac{\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}}{\|\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}\|} = \bar{\boldsymbol{P}}(\bar{\boldsymbol{P}}^{T}\bar{\boldsymbol{Q}} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & * \end{pmatrix} \bar{\boldsymbol{Q}}^{T}\bar{\boldsymbol{P}})\bar{\boldsymbol{P}}^{T} = \bar{\boldsymbol{P}} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & ** \end{pmatrix} \bar{\boldsymbol{P}}^{T} \in \boldsymbol{R}\sigma(\bar{\boldsymbol{Z}}^{+}). \quad \blacksquare$$

Proof of Theorem 2.21:
For any positive integer k, define

$$V_k = \left\{(f, h) \in \mathcal{F} : \psi(\cdot, \cdot; f, h) \text{ is one-to-one on } (\rho^{+})^{-1}(B_{\frac{1}{k}}(\bar{\boldsymbol{X}}^{+})) \times \boldsymbol{R}^{\ell}\right\}.$$

Suppose that $(\bar{\boldsymbol{Z}}, \bar{\lambda})$ is not a strongly stable stationary point for $\mathbf{Pro}_{(2)}(\bar{f}, \bar{h})$. Then, from Proposition 2.19, V_k is not a neighborhood of $(\bar{f}, \bar{h})$ in $\mathcal{F}_{B_{\frac{1}{k}}(\bar{\boldsymbol{X}}^{+})}$ for any positive integer k.

Hence there exists a sequence $(f^{(k)}, h^{(k)}) \notin V_k (k = 1, 2, \cdots)$ satisfying $\lim_{k\to\infty} \|(f^{(k)} - \bar{f}, h^{(k)} - \bar{h})\|_{B_{\frac{1}{k}}(\bar{\boldsymbol{X}}^+)} = 0$. Therefore there exist a distinct pair of elements $(\boldsymbol{Z}^{(k)}, \lambda^{(k)}), (\boldsymbol{W}^{(k)}, \mu^{(k)}) \in (\rho^+)^{-1}(B_{\frac{1}{k}}(\bar{\boldsymbol{X}}^+)) \times \boldsymbol{R}^\ell$ such that $\psi(\boldsymbol{Z}^{(k)}, \lambda^{(k)}; f^{(k)}, h^{(k)}) = \psi(\boldsymbol{W}^{(k)}, \mu^{(k)}; f^{(k)}, h^{(k)})$, i.e.,

$$\left.\begin{array}{rcl} D_{\boldsymbol{x}} L(\boldsymbol{Z}^{(k)+}, \lambda^{(k)}; f^{(k)}, h^{(k)}) + \boldsymbol{Z}^{(k)-} & = & D_{\boldsymbol{x}} L(\boldsymbol{W}^{(k)+}, \mu^{(k)}; f^{(k)}, h^{(k)}) + \boldsymbol{W}^{(k)-}, \\ h^{(k)}(\boldsymbol{Z}^{(k)+}) & = & h^{(k)}(\boldsymbol{W}^{(k)+}). \end{array}\right\}$$

By the mean value theorem, we can symbolically write

$$\left.\begin{array}{rcl} D_{\boldsymbol{x}} h^{(k)}(\gamma^{(k)})(\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}) & = & \boldsymbol{0}, \\ (\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}) \bullet D_{\boldsymbol{x}}{}^2 L(\alpha^{(k)}, \xi^{(k)}; f^{(k)}, h^{(k)}) & & \\ +(\lambda^{(k)} - \mu^{(k)}) D_{\boldsymbol{x}} h(\beta^{(k)}) + (\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}) & = & \boldsymbol{O}. \end{array}\right\} \tag{5}$$

Taking a subsequence we may assume that

$$\lim_{k\to\infty} \frac{(\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}, \boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}, \lambda^{(k)} - \mu^{(k)})}{\|(\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}, \boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}, \lambda^{(k)} - \mu^{(k)})\|} \text{ exists } = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{u}) \neq (\boldsymbol{O}, \boldsymbol{O}, \boldsymbol{0}).$$

Then from Lemma 2.22, it follows $\boldsymbol{A} \in \mathcal{V}(\bar{\boldsymbol{Z}}^-)$.
Since $(\boldsymbol{Z}^{(k)-} - \boldsymbol{W}^{(k)-}) \bullet (\boldsymbol{Z}^{(k)+} - \boldsymbol{W}^{(k)+}) = -\boldsymbol{Z}^{(k)-} \bullet \boldsymbol{X}^{(k)+} - \boldsymbol{Z}^{(k)-} \bullet \boldsymbol{W}^{(k)+} \geq 0$, we have $\boldsymbol{A} \bullet \boldsymbol{B} \geq 0$. Factoring the equation (5) by $\|(\boldsymbol{Z}^{(k)} - \boldsymbol{W}^{(k)}, \lambda^{(k)} - \mu^{(k)})\|$ and taking a limit, we get

$$\boldsymbol{A} \in Ker(D_{\boldsymbol{x}} h(\bar{\boldsymbol{Z}}^+)) \bigcap \mathcal{V}(\bar{\boldsymbol{Z}}^-) \text{ and } \boldsymbol{A} \bullet D^2_{\boldsymbol{x}} L(\bar{\boldsymbol{Z}}^+, \bar{\lambda}; \bar{f}, \bar{h}) + \boldsymbol{u} D_{\boldsymbol{x}} h(\bar{\boldsymbol{Z}}^+) + \boldsymbol{B} = \boldsymbol{O}. \tag{6}$$

Suppose $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{O}$. Then the equation (6) contradicts Condition 2.14. Suppose $\boldsymbol{A} = \boldsymbol{O}$. Since $\boldsymbol{B} \in \boldsymbol{R}\sigma(\bar{\boldsymbol{Z}}^+)$ follows from Lemma 2.24, the equation (6) contradicts Condition 2.15. Suppose $\boldsymbol{A} \neq \boldsymbol{O}$. In this case $\boldsymbol{A} \bullet D^2_{\boldsymbol{x},\lambda} L(\bar{\boldsymbol{Z}}^+, \bar{\lambda}; \bar{f}, \bar{h}) \bullet \boldsymbol{A} = -\boldsymbol{u} D_{\boldsymbol{x}} h(\bar{\boldsymbol{Z}}^+) \bullet \boldsymbol{A} - \boldsymbol{A} \bullet \boldsymbol{B} = -\boldsymbol{A} \bullet \boldsymbol{B} \leq 0$. This contradicts that $D^2_{\boldsymbol{x}} L(\bar{\boldsymbol{Z}}^+, \bar{\lambda}; \bar{f}, \bar{h})$ is positive definite on the space $Ker(D_{\boldsymbol{x}} h(\bar{\boldsymbol{Z}}^+)) \bigcap \mathcal{V}(\bar{\boldsymbol{Z}}^-)$. We have proved this theorem. ∎

3. Conclusions

We have proved a sufficient condition for strong stability in the sense of Kojima for stationary solutions of nonlinear positive semidefinite programs under the LICQ condition. For the forthcoming study of strong stability, we would like to propose the following problem:

Problem 3.1. Give an algebraic condition that is necessary and sufficient for strong stability by means of Hessians, Jacobians, etc.

Acknowledgement The author gratefully acknowledges the referees who patiently pointed out many mistakes in the original version of this paper and gave me a lot of precious mathematical advice.

References

[1] G. P. Barker: Theory of cones. *Linear Algebra and its Applications*, **39** (1981) 263-291.

[2] J. F. Bonnans and A. Shapiro: *Perturbation Analysis of Optimization Problems* (Springer, New York, 2000).

[3] A. Graham: *Kronecker Products and Matrix Calculus with Applications* (J. Wiley and Sons, New York, NY, 1981).

[4] R. D. Hill and S. R. Waters: On the cone of positive semidefinite matrices. *Linear Algebra and its Applications*, **90** (1987) 81-88.

[5] R. Hirabayashi, H. Th. Jongen and M. Shida: Stability for linearly constrained optimization problems. *Mathematical Programming*, **66** (1994) 351-360.

[6] B. Iversen: *Cohomology of Sheaves* (Springer-Verlag, New York, 1986).

[7] M. Kojima: Strongly stable stationary solutions in nonlinear programs. in S. M. Robinson (ed.): *Analysis and Computation of Fixed Points* (Academic Press, New York, 1980), 93-138.

[8] M. Kojima: Semidefinite programming and interior-point methods (in Japanese). *Bulletin of the Japan Society for Industrial and Applied Mathematics*, **6** (1996) 16-25.

[9] S. Lang: *Algebra* (Addison-Wesley, New York, 1993).

[10] O.L. Mangasarian: *Nonlinear Programming* (McGraw-Hill, New York, 1969).

[11] A. Shapiro: First and second order analysis of nonlinear semidefinite programs. *Mathematical Programming*, **77** (1997) 301-320.

Toshihiro Matsumoto
Department of Media Science
Teikyo University of Science and Technology
Yatsuzawa 2525, Uenohara-machi,
Kitatsuru-gun, Yamanashi 409-0193, Japan
E-mail: `matsu@ntu.ac.jp`