ON A CLASS OF NONLINEAR PROGRAMMING WITH EQUALITY CONSTRAINTS IN BANACH SPACES

Tagawa, Shojiro
International Institute for Advanced Study of Social Information Science

https://doi.org/10.5109/13101

出版情報：統計数理研究. 17 (1/2), pp.17-31, 1976-03. Research Association of Statistical Sciences

ON A CLASS OF NONLINEAR PROGRAMMING WITH EQUALITY CONSTRAINTS IN BANACH SPACES

By

Shôjirô TAGAWA*

(Received December 22, 1974; revised September 28, 1975)

§1. Introduction.

The author discussed nonlinear programming without equality constraints in Banach spaces in the previous paper [1]. This paper is devoted to discuss nonlinear programming with equality constraints under the same circumstances in [1]. More specifically, we shall be concerned with the problem $(\bar{P})$: Minimize $f(x)$ subject to $x\in A$, $g(x)\in B$ and $h(x)=0$, where f is a real valued function on a real Banach space X, A is a subset of X, g is a mapping of X into a Banach space Y, B is a closed convex cone with non-empty interior, and h is a function from X into R^m. We shall present necessary conditions and sufficient conditions for an optimal solution $\bar{x}\in A$ to the problem $(\bar{P})$. We also apply this nonlinear programming to the problem of linear minimum-varience unbiased estimation and to the linear optimal control problem.

We shall begin with some preliminaries and introduce the notion of a locally relatively convex set which is an extension of one given in [1] in Section 2. In Section 3, we shall present necessary conditions and sufficient conditions for an optimal solution to the problem $(\bar{P})$. In Section 4, we shall show that this argument can treat the problem of linear minimum-varience estimation and the linear optimal control problem with restricted phase coordinates.

§2. Preliminaries.

Notations and definitions are all the same as those in the author's previous paper [1] without the definition of the locally relatively convex set (see Definition 3.3 in [1]). The reader should regard this paper as an addendum to the paper [1]. We shall suppose that the reader is familiar with its notations and definitions.

First of all, we shall begin with the definition of the locally relatively convex set which is an extension of the one introduced in [1].

DEFINITION 2.1. Let A and D be arbitrary subsets of a locally convex linear topological space X, and let $\bar{x}\in A\cap D$. Then, the subset A is called to be *locally relatively convex with respect to D at $\bar{x}$* if there exists a convex neighborhood N of $\bar{x}$ such that

$$\mathrm{co}(A\cap N)\cap D=A\cap N\cap D\,.$$

* International Institute for Advanced Study of Social Information Science, Fujitsu Limited, 17–25, Shinkamata 1-Chome, Ohta-ku, Tokyo, Japan.

It should be noticed that if A is locally convex at $\bar{x}$ in the sense of Definition 3.4 in [1], then it is locally relatively convex with respect to D at $\bar{x}$ in the sense of Definition 2.1 when the space X is locally convex. Moreover, Theorem 3.3 in [1] still holds even if the condition (a) in Theorem 3.3 in [1] is replaced by the condition that A is locally relatively convex with respect to H at $\bar{x}$ in the sense of Definition 2.1.

LEMMA 2.1. *If the set A is locally relatively convex with respect to D at $\bar{x}$, then it is true that there is a convex neighborhood N of $\bar{x}$ in X such that*

$$\mathrm{LC}(A\cap D, \bar{x})=\mathrm{LC}(\mathrm{co}(A\cap N)\cap D, \bar{x}).$$

PROOF. Since A is locally relatively convex with respect to D at $\bar{x}$, there exists a convex neighborhood N of $\bar{x}$ such that $\mathrm{co}(A\cap N)\cap D=A\cap N\cap D$. It then follows that

$$\begin{aligned}\mathrm{LC}(\mathrm{co}(A\cap N)\cap D, \bar{x}) &= \bigcap_{\mathfrak{N}\in(\bar{x})} \mathrm{C}(\mathrm{co}(A\cap N)\cap D\cap U, \bar{x}) \\ &= \bigcap_{\substack{U:\ \mathrm{convex}\subset N \\ U\in\mathfrak{N}(\bar{x})}} \mathrm{C}(A\cap N\cap D\cap U, \bar{x}) \\ &= \bigcap_{\substack{U:\ \mathrm{convex}\subset N \\ U\in\mathfrak{N}(\bar{x})}} \mathrm{C}(A\cap D\cap U, \bar{x}) \\ &= \mathrm{LC}(A\cap D, \bar{x}),\end{aligned}$$

Q. E. D.

LEMMA 2.2. *Let X and Y be real linear topological spaces and let g be a mapping of X into Y. Suppose that the mapping g be differentiable at $\bar{x}$ in the sense of Neustadt and that the differential $g_{\bar{x}}(x)$ is continuous in x. If S is an arbitrary compact subset of X, then for every neighborhood W of the origin in Y there exist a positive number δ and a neighborhood N of the origin in X such that*

$$\frac{g(\bar{x}+\varepsilon y)-g(\bar{x})}{\varepsilon}\in g_{\bar{x}}(x)+W \qquad \textit{whenever} \quad x\in S,\ 0<\varepsilon<\delta,\ \textit{and}\ y\in x+N.$$

PROOF. This is an immediate extension of Lemma 3.1 described in Neustadt [2]. Hence we shall omitt the details.

§ 3. Nonlinear programming with equality constraints in Banach spaces.

In this section, we shall consider the following nonlinear programming in Banach spaces. Let X and Y be real Banach spaces, A an arbitrary subset of X, B a closed convex cone with vertex at the origin having a non-empty interior in Y, f a real valued function defined over X, g a mapping of X into Y and h a mapping of X into R^m (m-dimensional Euclidean space). Then, find a vector $\bar{x}\in A$ satisfying

$$(\bar{P}) \qquad \begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && x\in A \\ & && g(x)\in B \\ & && h(x)=0. \end{aligned}$$

We can now state the main results.

THEOREM 3.1. *Let $\bar{x}\in A$ be the optimal solution to the problem $(\bar{P})$. Assume that*

(a) *A is locally relatively convex with respect to $H\cap G$ at $\bar{x}$ where $H\equiv\{x\in X|h(x)=0\}$ and $G\equiv\{x\in X|g(x)\in B\}$,*

(b) *the mappings f, g and h are differentiable at $\bar{x}$ in the sense of Neustadt, the differential $f_{\bar{x}}(x)$ is convex continuous in x, the differential $g_{\bar{x}}(x)$ is B-convex continuous in x and the differential $h_{\bar{x}}(x)$ is linear continuous in x.*

Then, there exist a real number $\bar{\eta}$, a (row) vector $\bar{\alpha}\in R^m$ and a linear continuous functional $\bar{y}^*\in Y^*$, not all zero, such that

$$\bar{\eta}\geqq 0\,, \tag{3.1}$$

$$\bar{y}^*(y)\leqq 0 \quad \text{for all} \quad y\in B\,, \tag{3.2}$$

$$\bar{y}^*(g(\bar{x}))=0\,, \tag{3.3}$$

$$g(\bar{x})\in B\,, \qquad h(\bar{x})=0\,, \tag{3.4}$$

$$\bar{\eta}f_{\bar{x}}(x)+\bar{y}^*(g_{\bar{x}}(x))+\bar{\alpha}h_{\bar{x}}(x)\geqq 0 \quad \text{for all} \quad x\in P(A,\bar{x})\,. \tag{3.5}$$

PROOF. The condition (3.4) clearly holds for an optimal solution $\bar{x}\in A$ to the problem $(\bar{P})$. Let the sets H and G are as defined in Theorem 3.1. Since the vector $\bar{x}\in A$ is an optimal solution to the problem $(\bar{P})$, it follows from Theorem 3.2 in [1] that

$$f_{\bar{x}}(x)\geqq 0 \quad \text{for all} \quad x\in \mathrm{LC}(A\cap H\cap G,\bar{x})\,. \tag{3.6}$$

Let us define the subset M of X as follows:

$$M=\{x\in X|\, x\in P(A,\bar{x}),\ g_{\bar{x}}(x)\in \operatorname{int} B-g(\bar{x}),\ f_{\bar{x}}(x)<0\}\,.$$

It is immediate that the set M is convex. The linearity of $h_{\bar{x}}(x)$ implies that the set $K=h_{\bar{x}}(M)$ is convex in R^m. We shall now show that the origin is not an interior point of the convex set K. To show the contradiction, assume that the origin is an interior point of K. Then, there exists an m-simplex S which contains the origin in its interior and which is included in K. Let the vertices of the m-simplex S be $h_{\bar{x}}(x_1),\cdots,h_{\bar{x}}(x_{m+1})$, where $x_j\in M$ for $j=1,\cdots,m+1$. It then follows from the definition of M that

$$x_j\in \mathrm{P}(A,\bar{x})\,, \qquad g_{\bar{x}}(x_j)\in \operatorname{int} B-g(\bar{x})\,, \tag{3.7}$$

$$f_{\bar{x}}(x_j)<0 \quad \text{for} \quad j=1,\cdots,m+1\,. \tag{3.8}$$

It is valid, by Lemma 2.2 in [1], that for every j, $1\leqq j\leqq m+1$, there exist a sequence of positive numbers $\{\lambda_n^j\}$, a sequence of positive integers $\{m_n^j\}$, a sequence of vectors $\alpha_n^j=(\alpha_{n1}^j,\cdots,\alpha_{nm_n^j}^j)\in P^{m_n^j}$ and a sequence of vectors $\{y_{ni}^j\}_{i=1,\cdots,m_n^j}\subset A$ such that

$$y_{ni}^j \xrightarrow[n\to\infty]{} \bar{x} \quad \text{uniformly in } i\,,$$

$$\lambda_n^j\Big(\sum_{i=1}^{m_n^j}\alpha_{ni}^j y_{ni}^j-\bar{x}\Big) \xrightarrow[n\to\infty]{} x_j\,, \qquad j=1,\cdots,m+1\,.$$

It is easily verified that

$$\sum_{j=1}^{m+1}\beta_j\lambda_n^j\Big(\sum_{i=1}^{m_n^j}\alpha_{ni}^j y_{ni}^j-\bar{x}\Big)\xrightarrow[n\to\infty]{}\sum_{j=1}^{m+1}\beta_j x_j$$

uniformly in $\beta=(\beta_1,\cdots,\beta_{m+1})\in P^{m+1}$.

Let us define

$$\mu_n(\beta)\equiv\sum_{j=1}^{m+1}\beta_j\lambda_n^j,\qquad k_n\equiv m_n^1+m_n^2+\cdots+m_n^{m+1},$$

$$\gamma_{ni}(\beta)\equiv\begin{cases}\beta_1\lambda_n^1\alpha_{ni}^1/\mu_n(\beta) & \text{for}\quad i=1,\cdots,m_n^1,\\ \beta_2\lambda_n^2\alpha_{n(i-m_n^1)}^2/\mu_n(\beta) & \text{for}\quad i=m_n^1+1,\cdots,m_n^1+m_n^2,\\ \vdots & \\ \beta_{m+1}\lambda_n^{m+1}\alpha_{n(i-m_n^1-m_n^2-\cdots-m_n^m)}^{m+1}/\mu_n(\beta) & \text{for}\quad i=m_n^1+\cdots+m_n^m+1,\cdots,k_n,\end{cases}$$

$$x(\beta)\equiv\sum_{j=1}^{m+1}\beta_j x_j,$$

$$x_{ni}\equiv\begin{cases}y_{ni}^1 & \text{for}\quad i=1,\cdots,m_n^1,\\ y_{n(i-m_n^1)}^2 & \text{for}\quad i=m_n^1+1,\cdots,m_n^1+m_n^2,\\ \vdots & \\ y_{n(i-m_n^1-\cdots-m_n^m)}^{m+1} & \text{for}\quad i=m_n^1+\cdots+m_n^m+1,\cdots,k_n.\end{cases}$$

Then, we have

$$\mu_n(\beta)\Big(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-\bar{x}\Big)=\sum_{j=1}^{m+1}\beta_j\lambda_n^j\Big(\sum_{i=1}^{m_n^j}\alpha_{ni}^j y_{ni}^j-\bar{x}\Big)\xrightarrow[n\to\infty]{}x(\beta)=\sum_{j=1}^{m+1}\beta_j x_j \tag{3.9}$$

uniformly in $\beta\in P^{m+1}$.

It is clear that $\mu_n(\beta)$ and $\gamma_{ni}(\beta)$ are continuous in $\beta\in P^{m+1}$, that $(\gamma_{n1}(\beta),\cdots,\gamma_{nk_n}(\beta))\in P^{k_n}$ and that

$$x_{ni}\xrightarrow[n\to\infty]{}\bar{x}\qquad\text{uniformly in } i. \tag{3.10}$$

It then follows from Lemma 2.2 in [1] that $x(\beta)\in P(A,\bar{x})$ for all $\beta\in P^{m+1}$.

Since A is locally relatively convex with respect to $H\cap G$ at $\bar{x}$, there exists a convex neighborhood $\tilde{N}$ of $\bar{x}$ such that $\mathrm{co}(A\cap\tilde{N})\cap H\cap G=A\cap\tilde{N}\cap H\cap G$. It follows from (3.10) that there exists a positive integer M_1 such that

$$x_{ni}\in\tilde{N},\quad i=1,\cdots,k_n,\qquad\text{for all}\quad n\geqq M_1.$$

By virtue of (3.8), we obtain

$$f_{\bar{x}}(x(\beta))\leqq\sum_{j=1}^{m+1}\beta_j f_{\bar{x}}(x_j)<0\qquad\text{for all}\quad\beta\in P^{m+1}. \tag{3.1}$$

Since $g_{\bar{x}}(x)$ is B-convex, it is true, by virtue of (3.7), that

$$g_{\bar{x}}(x(\beta))\in\mathrm{int}\,B-g(\bar{x})\qquad\text{for all}\quad\beta\in P^{m+1}. \tag{3.12}$$

It is noticed that the set Δ in Y defined by

$$\Delta=\{g_{\bar{x}}(x(\beta))\,|\,\beta\in P^{m+1}\},$$

is compact since it is a continuous image of a compact set. By virtue of (3.12), we have $\Delta\subset\operatorname{int}B-g(\bar{x})$. Define the set Σ in Y as follows:

$$\Sigma=\text{a complement of } (\operatorname{int}B-g(\bar{x}))\,.$$

It is then clear that the set Σ is closed and that

$$\Delta\cap\Sigma=\emptyset\,. \tag{3.13}$$

If we set $\operatorname{dist}(\Delta,\Sigma)=\inf\{\|x-y\|\,|\,x\in\Delta,\ y\in\Sigma\}$, then it is valid that $\delta\equiv\operatorname{dist}(\Delta,\Sigma)>0$. Indeed, if we assume that $\operatorname{dist}(\Delta,\Sigma)=0$, then there are sequences of vectors $\{x_i\}\subset\Delta$ and $\{y_i\}\subset\Sigma$ such that $\|x_i-y_i\|\to 0$ as $i\to\infty$. The compactness of Δ implies the existence of a convergent subsequence $\{x_{i_j}\}\subset\Delta$ and a vector $x_0\in\Delta$ satisfying $x_{i_j}\to x_0$ as $j\to\infty$, and $\|x_{i_j}-y_j\|\to 0$ as $j\to\infty$. Then, we have

$$\|x_0-y_j\|\leqq\|x_0-x_{i_j}\|+\|x_{i_j}-y_j\|\to 0 \qquad \text{as} \quad j\to\infty\,.$$

This implies that $x_0\in\Delta\cap\Sigma$ since the set Σ is closed. But this contradicts to (3.13).

Let us define the neighborhood U of the origin in Y by $U=\{y\in Y\,|\,\|y\|<\delta\}$, then we have

$$g_{\bar{x}}(x(\beta))+U\subset\operatorname{int}B-g(\bar{x}) \qquad \text{for all} \quad \beta\in P^{m+1}\,.$$

We now show that

$$\mu_n(\beta)[g(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni})-g(\bar{x})] \tag{3.14}$$

$$=\frac{g\Big(\bar{x}+\dfrac{1}{\mu_n(\beta)}\mu_n(\beta)[\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-\bar{x}]\Big)-g(\bar{x})}{\dfrac{1}{\mu_n(\beta)}}\xrightarrow[n\to\infty]{} g_{\bar{x}}(x(\beta))$$

uniformly in $\beta\in P^{m+1}$.

It follows from Lemma 2.2 that to each neighborhood W of the origin in Y there correspond a positive number δ and a neighborhood N of the origin in X such that

$$\frac{g(\bar{x}+\varepsilon y)-g(\bar{x})}{\varepsilon}\in g_{\bar{x}}(x)+W \quad \text{whenever} \quad x\in\{x(\beta)\,|\,\beta\in P^{m+1}\}\,,\ 0<\varepsilon<\delta\,,\ y\in x+N\,.$$

Hence we have

$$\frac{g\Big(\bar{x}+\dfrac{1}{\mu_n(\beta)}\mu_n(\beta)[\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-\bar{x}]\Big)-g(\bar{x})}{\dfrac{1}{\mu_n(\beta)}}\in g_{\bar{x}}(x(\beta))+W$$

$$\text{whenever} \quad \beta\in P^{m+1}\,,\ 0<\frac{1}{\mu_n(\beta)}<\delta\,, \ \text{and} \ \ \mu_n(\beta)[\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-\bar{x}]\in x(\beta)+N\,.$$

Since $\dfrac{1}{\mu_n(\beta)}\to 0$ uniformly in β as $n\to\infty$, we conclude, with (3.9) that there is a positive integer $\bar{n}$ (independent of β) such that

$$0<\frac{1}{\mu_n(\beta)}<\delta \qquad \text{for all} \quad \beta\in P^{m+1}\,,$$

and

$$\mu_n(\beta)[\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-\bar{x}]\in x(\beta)+N \qquad \text{for all} \quad \beta\in P^{m+1} \text{ and all } n\geqq\bar{n}\,.$$

This implies (3.14).

Therefore, for the neighborhood U there exists a positive integer $\bar{n}_1$ (independent of β) such that

$$\mu_n(\beta)[g(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni})-g(\bar{x})]\in g_{\bar{x}}(x(\beta))+U\subset \mathrm{int}\, B-g(\bar{x})$$
$$\text{for all} \quad \beta\in P^{m+1} \text{ and all } n\geqq \bar{n}_1.$$

It is valid that there is a positive integer $\bar{n}_2$ (independent of β) such that $\mu_n(\beta)>1$ for all $\beta\in P^{m+1}$ and all $n\geqq\bar{n}_2$. If we set $\bar{n}_3=\max\{\bar{n}_1, \bar{n}_2\}$, then we have

$$\mu_n(\beta)[g(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni})-g(\bar{x})]\in \mathrm{int}\, B-g(\bar{x}) \qquad \text{for all} \quad \beta\in P^{m+1} \text{ and all } n\geqq\bar{n}_3,$$

or

$$\mu_n(\beta)g(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni})\in(\mu_n(\beta)-1)g(\bar{x})+\mathrm{int}\, B\subset \mathrm{int}\, B\subset B$$
$$\text{for all} \quad \beta\in P^{m+1} \text{ and all } n\geqq\bar{n}_3.$$

Hence, it is true that $\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}\in G$ for all $\beta\in P^{m+1}$ and all $n\geqq\bar{n}_3$. Recall that

$$\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}\in \mathrm{co}(A\cap\tilde{N}) \qquad \text{for all} \quad \beta\in P^{m+1} \text{ and all } n\geqq M_1.$$

In the similar fashion, it is verified that

$$\text{(3.15)} \qquad \mu_n(\beta)[h(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni})-h(\bar{x})] \xrightarrow[n\to\infty]{} h_{\bar{x}}(x(\beta)) \qquad \text{uniformly in } \beta\in P^{m+1}.$$

Let us define

$$t_j\equiv h_{\bar{x}}(x_j)\in R^m, \qquad j=1,\cdots,m+1,$$

$$t(\beta)\equiv\sum_{j=1}^{m+1}\beta_j t_j=\sum_{j=1}^{m+1}\beta_j h_{\bar{x}}(x_j)=h_{\bar{x}}(x(\beta)) \qquad \text{for all} \quad \beta\in P^{m+1}.$$

Let us define the mapping ψ_n of S into R^m by

$$\psi_n(t(\beta))\equiv t(\beta)-\mu_n(\beta)[h(\sum_{i=1}^{k_n}\gamma_{ni}(\beta)x_{ni}-h(\bar{x})] \qquad \text{for all} \quad \beta\in P^{m+1},$$

then it is noticed that ψ_n is continuous for each n. Since the simplex S contains the origin in its interior, there is a positive number ξ such that $t\in S$ whenever $\|t\|<\xi$, $t\in R^m$. It then follows from (3.15) that there is a positive integer $\bar{n}_4$ (independent of β) such that $\|\psi_n(t(\beta))\|<\xi$ for all $\beta\in P^{m+1}$ and all $n\geqq\bar{n}_4$, that is, each ψ_n is a continuous mapping of S into S. It is verified, on the basis of Brower fixed point theorem, that for each $n\geqq\bar{n}_4$, there is a vector $t(\beta^n)$, $\beta^n\in P^{m+1}$, such that $\psi_n(t(\beta^n))=t(\beta^n)$, i.e., $\mu_n(\beta^n)h(\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni})=0$. Since $\mu_n(\beta^n)>0$, we have

$$h(\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni})=0 \qquad \text{for every} \quad n\geqq\bar{n}_4.$$

Since the set P^{m+1} is compact, there is a convergent subsequence of $\{\beta^n\}_{n\geqq\bar{n}_4}$. We shall still denote it by $\{\beta^n\}_{n\geqq\bar{n}_4}$, i.e., there is a vector $\bar{\beta}\in P^{m+1}$ satisfying $\beta^n\to\bar{\beta}$ as $n\to\infty$.

It now follows from (3.9) that

$$\mu_n(\beta^n)(\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni}-\bar{x})-x(\bar{\beta})$$
$$=\mu_n(\beta^n)(\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni}-\bar{x})-x(\beta^n)+x(\beta^n)-x(\bar{\beta})\longrightarrow 0 \quad \text{as} \quad n\to\infty,$$

and

$$\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni}\longrightarrow\bar{x} \quad \text{as} \quad n\to\infty.$$

If we set $\bar{n}_5=\max\{M_1, \bar{n}_3, \bar{n}_4\}$, then we obtain

$$\sum_{i=1}^{k_n}\gamma_{ni}(\beta^n)x_{ni}\in\mathrm{co}(A\cap\tilde{N})\cap G\cap H \quad \text{for all} \quad n\geqq\bar{n}_5.$$

It follows from Lemma 2.1 in Varaiya [3] and Lemma 2.1 that

$$x(\bar{\beta})\in\mathrm{LC}(\mathrm{co}(A\cap\tilde{N})\cap G\cap H, \bar{x})=\mathrm{LC}(A\cap G\cap H, \bar{x}).$$

By virtue of (3.6), we have $f_{\bar{x}}(x(\bar{\beta}))\geqq 0$, which contradicts to (3.11). Consequently, we can conclude that the origin is not an interior point of the convex set K.

If the interior of K is not empty, or if the origin does not belong to K, then it follows from the separation theorem that there is a non-zero vector $\tilde{\alpha}\in R^m$ such that

$$\tilde{\alpha}t\leqq 0 \quad \text{for all} \quad t\in K. \tag{3.16}$$

If K has no interior point and if the origin belongs to K, then K is contained in some $(m-1)$-dimensional hyperplane through the origin, and hence there is again a non-zero vector $\tilde{\alpha}$ satisfying (3.16).

If we define convex sets $\tilde{B}$ and $\tilde{E}$ in $Y\times R^1\times R^1$ by

$$\tilde{B}\equiv\{\tilde{b}\in Y\,|\,\tilde{b}=b-g(\bar{x}),\ b\in\mathrm{int}\,B\}\times\{r\in R^1\,|\,r<0\}\times\{s\in R^1\,|\,s>0\},$$

and

$$\tilde{E}\equiv\{(g_{\bar{x}}(x), f_{\bar{x}}(x), \tilde{\alpha}h_{\bar{x}}(x))-(b, r, 0)\in Y\times R^1\times R^1\,|\,x\in\mathrm{P}(A, \bar{x}),\ b\in B,\ r\leqq 0\},$$

then it is clear that the set $\tilde{B}$ is open convex and that $\tilde{E}$ is convex. It is easy to verify, by virtue of (3.16), that $\tilde{B}$ and $\tilde{E}$ are disjoint, and hence it follows from the separation theorem (see Dunford and Schwartz [4]) that there exists a non-zero linear continuous functional $z^*\in(Y\times R^1\times R^1)^*$ such that

$$z^*((y', t', s'))\leqq 0\leqq z^*((y'', t'', s'')) \quad \text{for all} \quad (y', t', s')\in\tilde{B} \text{ and } (y'', t'', s'')\in\tilde{E}.$$

Furthermore, there are a linear continuous functional $\bar{y}^*\in Y^*$ and real numbers $\bar{\eta}$ and α_0, not all zero, such that

$$z^*((y, t, s))=\bar{y}^*(y)+\bar{\eta}t+\alpha_0 s \quad \text{for all} \quad y\in Y, \text{ all } t\in R^1, \text{ and all } s\in R^1.$$

Therefore we have

$$\bar{y}^*(y-g(\bar{x}))+\bar{\eta}t+\alpha_0 s\leqq 0\leqq\bar{y}^*(g_{\bar{x}}(x)-b)+\bar{\eta}(f_{\bar{x}}(x)-r)+\alpha_0\tilde{\alpha}h_{\bar{x}}(x) \tag{3.17}$$
$$\text{for all} \quad y\in\mathrm{int}\,B,\ t<0,\ s>0,\ x\in\mathrm{P}(A, \bar{x}),\ b\in B,\ r\leqq 0.$$

Since $0\in\mathrm{P}(A, \bar{x})$ and $0\in B$, it is clear that $\bar{\eta}\geqq 0$ and $\bar{y}^*(b)\leqq 0$ for all $b\in B$, which are the conditions (3.1) and (3.2).

If we set $\bar{\alpha}=\alpha_0\tilde{\alpha}\in R^m$, then it is valid that $\bar{y}^*$, $\bar{\eta}$ and $\bar{\alpha}$ are not all zero since $\tilde{\alpha}\neq 0$ and since $\bar{y}^*$, $\bar{\eta}$ and α_0 are not all zero. It then follows from (3.17) that

$$\bar{y}^*(g_{\bar{x}}(x))+\bar{\eta}f_{\bar{x}}(x)+\bar{\alpha}h_{\bar{x}}(x)\geqq 0 \qquad \text{for all} \quad x\in \mathrm{P}(A, \bar{x}),$$

which is the condition (3.5). By tending $t\to 0$ and $s\to 0$ in (3.17), we have $\bar{y}^*(g(\bar{x}))=0$ since $2g(\bar{x})\in B$. Hence the condition (3.3) holds. This completes the proof of Theorem 3.1. Q. E. D.

THEOREM 3.2. *Let us consider the problem* $(\bar{P})$. *Then, the condition that there are a real number* $\bar{\eta}$, *a vector* $\bar{\alpha}\in R$ *and a linear continuous functional* $\bar{y}^*\in Y^*$ *which satisfy the conditions* (3.1)-(3.5) *in Theorem* 3.1 *is sufficient for a vector* $\bar{x}\in A$ *to be an optimal solution to the problem* $(\bar{P})$ *if the set* $A\cap G\cap H$ *is a pseudo-cone with vertex at* $\bar{x}$, *if* f *is differentiable at* $\bar{x}$ *in the sense of Neustadt and pseudo-convex over* $A\cap G\cap H$ *at* $\bar{x}$, *if* g *is differentiable in the sense of Neustadt, and if* $\bar{\eta}\neq 0$ *where the sets* G *and* H *are as defined in Theorem* 3.1.

PROOF. It is valid that

$$x-\bar{x}\in \mathrm{P}(A, \bar{x})\cap \mathrm{LC}(G, \bar{x})\cap \mathrm{LC}(H, \bar{x}) \qquad \text{for all} \quad x\in A\cap G\cap H \tag{3.18}$$

since $A\cap G\cap H$ is a pseudo-cone with vertex at $\bar{x}$ and since

$$\mathrm{LC}(A\cap G\cap H, \bar{x})\subset \mathrm{P}(A, \bar{x})\cap \mathrm{LC}(G, \bar{x})\cap \mathrm{LC}(H, \bar{x}).$$

We now show that the following relation (3.19) holds.

$$\mathrm{LC}(H, \bar{x})\subset \{x\in X \mid h_{\bar{x}}(x)=0\}. \tag{3.19}$$

It follows from Lemma 2.1 in Varaiya [3] that for each $x\in \mathrm{LC}(H, \bar{x})$ there exist sequences $\{x_n\}$ and $\{\lambda_n \mid \lambda_n>0\}$ such that $\lim_{n\to\infty} x_n=\bar{x}$ and $\lim_{n\to\infty} \lambda_n(x_n-\bar{x})=\bar{x}$, and hence we have

$$\begin{aligned} 0&=\lim_{n\to\infty} \lambda_n(h(x_n)-h(\bar{x})) \\ &=\lim_{n\to\infty} \frac{h\left(\bar{x}+\frac{1}{\lambda_n}\lambda_n(x_n-\bar{x})\right)-h(\bar{x})}{\frac{1}{\lambda_n}} \\ &=h_{\bar{x}}(x). \end{aligned}$$

Consequently, (3.19) holds.

It is true, by (3.2) and (3.3), that

$$\bar{y}^*(g^{\bar{x}}(x))\leqq 0 \qquad \text{for all} \quad x\in \mathrm{LC}(G, \bar{x}),$$

as shown by the argument similar to Theorem 3.4 in [1]. By virtue of (3.5) and (3.19), we obtain

$$\bar{\eta}f_{\bar{x}}(x)\geqq -\bar{y}^*(g_{\bar{x}}(x))-\bar{\alpha}h_{\bar{x}}(x)=-\bar{y}^*(g_{\bar{x}}(x))\geqq 0,$$
$$\text{for all} \quad x\in \mathrm{P}(A, \bar{x})\cap \mathrm{LC}(G, \bar{x})\cap \mathrm{LC}(H, \bar{x}).$$

It is then valid, by (3.18), that $f_{\bar{x}}(x-\bar{x})\geqq 0$ for all $x\in A\cap G\cap H$. The pseudo-convexity of f implies that $f(x)-f(\bar{x})\geqq 0$ for every $x\in A\cap G\cap H$, which shows that the vector x is an optimal solution to the problem $(\bar{P})$. Q. E. D.

The condition (3.5) is different from the necessary condition given in Neustadt [2] in the sense that the inequality in (3.5) holds on the local closed convex cone $\mathrm{P}(A, \bar{x})$ of A at $\bar{x}$, whereas the corresponding inequality in [2] holds on some convex set in $\mathrm{LC}(A, \bar{x})$, i. e., the first-order convex approximation. In contrast with Varaiya [3], our hypotheses in Theorems 3.1 and 3.2 are weaker than those in [3] since the constraint qualification is not imposed. The detail discussions are stated in Section 5 in [1].

Now, we shall set up the following Regularity Assumption (RA) which corresponds to the conventional constraint qualification.

(RA): Let X and Y be Banach spaces, A a subset of X which contains a vector $\bar{x}$, B a closed convex cone with non-empty interior in Y, g a function from X into Y which is differentiable at $\bar{x}$ in the sense of Neustadt and h a function from X into R^m which is differentiable at $\bar{x}$ in the sense of Neustadt. Then, there is a vector $x_0 \in \mathrm{P}(A, \bar{x})$ such that $g(\bar{x})+g_{\bar{x}}(x_0) \in \mathrm{int}\, B$, $h_{\bar{x}}(x_0)=0$ and the origin is an interior point of $h_{\bar{x}}(\mathrm{P}(A, \bar{x}))$.

THEOREM 3.3. *Consider the nonlinear programming problem* $(\bar{P})$ *and suppose that Regularity Assumption* (RA) *holds. Let* $\bar{x}$ *be a (feasible) solution to* $(\bar{P})$ *and assume that*

(c) *A is locally relatively convex with respect to $G \cap H$ at $\bar{x}$ and $A \cap G \cap H$ is a pseudo-cone with vertex at $\bar{x}$,*

(d) *f is pseudo-convex over $A \cap G \cap H$ at $\bar{x}$ and the differential $f_{\bar{x}}(x)$ is convex continuous in x, g is differentiable at $\bar{x}$ in the sense of Neustadt and the differential $g_{\bar{x}}(x)$ is B-convex continuous in x, h is differentiable at $\bar{x}$ in the sense of Neustadt and the differential $h_{\bar{x}}(x)$ is linear continuous in x.*

Then, in order that the vector $\bar{x}$ is an optimal solution to $(\bar{P})$, it is necessary and sufficient that there exist a (row) vector $\bar{\alpha} \in R^m$ and a linear continuous functional $\bar{y}^* \in Y^*$ such that

(3.20) $$\bar{y}^*(y) \leqq 0 \quad \text{for all} \quad y \in B,$$

(3.21) $$\bar{y}^*(g(\bar{x}))=0,$$

(3.22) $$g(\bar{x}) \in B, \qquad h(\bar{x})=0,$$

(3.23) $$f_{\bar{x}}(x)+\bar{y}^*(g_{\bar{x}}(x))+\bar{\alpha} h_{\bar{x}}(x) \geqq 0 \quad \text{for all} \quad x \in \mathrm{P}(A, \bar{x}).$$

PROOF. Necessity: It follows from Theorem 3.1 that there are a real number $\bar{\eta}$, a vector $\bar{\alpha} \in R^m$ and a linear continuous functional $\bar{y}^* \in Y^*$, not all zero, such that (3.1)-(3.5) hold. Therefore, it suffices to show that $\bar{\eta} \neq 0$. To show the contradiction, suppose that $\bar{\eta}=0$. Then, by virtue of (3.5), we have

(3.24) $$\bar{y}^*(g_{\bar{x}}(x))+\bar{\alpha} h_{\bar{x}}(x) \geqq 0 \quad \text{for all} \quad x \in \mathrm{P}(A, \bar{x}).$$

For the vector x_0 in Regularity Assumption (RA), we have

$$\bar{y}^*(g_{\bar{x}}(x_0)) \geqq 0,$$

furthermore, by (3.3) and (3.2), we obtain

$$\bar{y}^*(g(\bar{x})+g_{\bar{x}}(x_0))=0.$$

Since $g(\bar{x})+g_{\bar{x}}(x_0)$ is an interior point of the convex cone B, it is true, by (3.2), that $\bar{y}^*=0$. Hence, (3.24) becomes

$$\bar{\alpha}h_{\bar{x}}(x)\geqq 0 \qquad \text{for all} \quad x\in \mathrm{P}(A, \bar{x}).$$

But, inasmuch as the origin is an interior point of the set $h_{\bar{x}}(\mathrm{P}(A, \bar{x}))$, we have $\bar{\alpha}=0$. Consequently, $(\bar{\eta}, \bar{y}^*, \bar{\alpha})=(0, 0, 0)$, which is a contradiction.

Sufficiency: This is an immediate consequence of Theorem 3.2. Q.E.D.

§4. Some applications.

In this section, we shall state some applications of our argument to minimum-variance estimation problems and to linear optimal control problems.

We first consider a problem of linear minimum-variance unbiased estimation. Let z be an m-dimensional data vector, S an $m\times n$ matrix (known), whose columns are linearly independent, λ an n-dimensional vector (unknown), and e an m-dimensional random vector. Let z be of the form $z=S\lambda+e$, where the mean $E(e)=0$ and the covariance $E(ee')=V$ which is positive definite. We shall seek a linear estimate $\hat{\lambda}$ of the form $\hat{\lambda}=Xz$, where X is an $n\times m$ matrix. Then, the problem is as follows:

$$(E_1) \qquad \begin{aligned} &\text{minimize} \quad E[(\hat{\lambda}-\lambda)'(\hat{\lambda}-\lambda)] \\ &\text{subject to} \quad E(\hat{\lambda})=\lambda, \\ &\qquad\qquad\quad \hat{\lambda}=Xz. \end{aligned}$$

If we set $f(X)=E[(X(S\lambda+e)-\lambda)'(X(S\lambda+e)-\lambda)]$, then it is easily seen that

$$f(X)=\operatorname{trace}(XVX')+\lambda'S'X'XS\lambda-2\lambda'XS\lambda+\lambda'\lambda.$$

Let $\mathscr{X}$ be a Hilbert space of all $n\times m$ matrices with the inner product $(X|Y)\equiv \operatorname{trace}(XVY')$, $\mathscr{Z}$ a Hilbert space of all $n\times n$ matrices with the inner product $(A|B)= \operatorname{trace}(AB')$, and h an affine map of $\mathscr{X}$ into $\mathscr{Z}$ defined by $h(X)=XS-I$, where I is an $n\times n$ identity matrix. Then, the problem (E_1) becomes the following problem:

$$(E_2) \qquad \begin{aligned} &\text{minimize} \quad f(X) \\ &\text{subject to} \quad X\in\mathscr{X}, \\ &\qquad\qquad\quad h(X)=0. \end{aligned}$$

It is easy to verify that $f(X)$ is a uniformly convex function in the sense of Definition 3.1 in [5] and that Assumptions 3 and 4 in [5] are satisfied. Therefore, it follows from Theorem 4.1 in [5] and Theorem 3.1 that there exists a unique optimal solution $\bar{X}\in\mathscr{X}$ to the problem (E_2) which satisfies the condition that there are a real number $\bar{\eta}$ and an $n\times n$ matrix $\bar{\alpha}\in\mathscr{Z}$, not both zero, such that

$$(4.1) \qquad \bar{\eta}\geqq 0,$$

$$(4.2) \qquad h(\bar{X})=0,$$

$$(4.3) \qquad \bar{\eta}f_{\bar{x}}(X)+\operatorname{trace}(\bar{\alpha}h_{\bar{x}}(X)')\geqq 0 \qquad \text{for all} \quad X\in\mathscr{X}.$$

By a straightforward calculation, we have

$$f_{\bar{x}}(X)=2\,\mathrm{trace}\,(\bar{X}VX')+2\lambda'(S'\bar{X}'-I)XS\lambda\,, \quad \text{and} \quad h_{\bar{x}}(X)=XS\,.$$

It is then valid, by (4.2) and (4.3), that

$$2\bar{\eta}\,\mathrm{trace}\,(\bar{X}VX')+\mathrm{trace}\,(\bar{\alpha}S'X')\geqq 0 \qquad \text{for all} \quad X\in\mathcal{X}$$

or

$$2\bar{\eta}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m}\bar{x}_{ij}v_{jk}x_{ik}+\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{\alpha}_{ij}(\sum_{k=1}^{m}x_{ik}s_{kj})\geqq 0 \qquad \text{for all} \quad X\in\mathcal{X}\,,$$

where $\bar{X}=(\bar{x}_{ij})$, $V=(v_{jk})$, $X=(x_{ik})$, $\bar{\alpha}=(\bar{\alpha}_{ij})$ and $S=(s_{kj})$. Furthermore, we have $2\bar{\eta}\bar{X}V+\bar{\alpha}S'=0$. If we set $\bar{\eta}=0$, then $\bar{\alpha}S'=0$ which contradicts to the fact that the columns of S are linearly independent. Therefore, by setting $\bar{\eta}=1$ (see (4.1)), we have

$$2\bar{X}V+\bar{\alpha}S'=0\,. \tag{4.4}$$

It is true, by (4.2) and (4.4), that

$$\bar{X}=(S'V^{-1}S)^{-1}S'V^{-1}\,. \tag{4.5}$$

Thus, we can obtain the following theorem:

THEOREM 4.1. *Consider the problem* (E_1) *which becomes the problem* (E_2). *Then, there exists a unique optimal solution* $\bar{X}\in\mathcal{X}$ *which satisfies* (4.5). *On the other hand, if the matrix* $\bar{X}$ *satisfies* (4.5), *then* $\bar{X}$ *is a unique optimal solution to the problem* (E_2). *Therefore, the linear minimum-variance unbiased estimate* $\hat{\lambda}$ *of* λ *is as follows:*

$$\hat{\lambda}=(S'V^{-1}S)^{-1}S'V^{-1}z\,.$$

The results of Theorem 4.1 were given in [7], but our proof is different from one described in [7].

Next, we shall consider a linear optimal control problem with restricted phase coordinates. Let U be a convex subset of R^m and $\mathcal{U}$ a control class which consists of all measurable m-dimensional vector valued function $u(t)$ such that $u(t)\in U$ for all $t\in I$, where $I=[t_0, t_1]$ is a compact interval. The dynamical system is defined by, for some $u\in\mathcal{U}$,

$$\dot{x}(t)=A(t)x(t)+B(t)u(t) \qquad \text{for almost all} \quad t\in I\,, \tag{4.6}$$

where $x(t)$ is an absolutely continuous n-dimensional vector valued function on I, $A(t)$ is a continuous $n\times n$ matrix valued function on I, and $B(t)$ is a continuous $n\times m$ matrix valued function on I. Let g be a real valued convex function of C^2-class on R^n, and f a real valued convex function of C^1-class on R^n. Moreover, let w_0 and w_1 be vectors in R^q and R^r, respectively, S a $q\times n$ matrix with $\mathrm{rank}\,(S)=q$ and T a $r\times n$ matrix with $\mathrm{rank}\,(T)=r$. Then, consider the following control problem (P_1):

$$(P_1)\qquad \begin{aligned} &\text{minimize} && \int_{t_0}^{t_1} f(x(t))dt \\ &\text{subject to} && x \text{ satisfies (4.6) for some } u\in\mathcal{U}\,, \\ &&& Sx(t_0)=w_0\,, \\ &&& Tx(t_1)=w_1\,, \\ &&& g(x(t))\leqq 0 \qquad \text{for all} \quad t\in I\,. \end{aligned}$$

We shall set up the following assumption (4.7):

(4.7) there is a trajectory $x_0(t)$ which satisfies (4.6) for a controller $u_0 \in \mathcal{U}$ such that $g(x_0(t))<0$ for all $t\in I$, $Sx_0(t_0)=w_0$, and $Tx_0(t_1)=w_1$.

Let X be a Banach space of all continuous n-dimensional vector valued functions on I with the uniform norm $\|x\|=\max_{t\in I}|x(t)|$, where $|\cdot|$ denotes the Euclidean norm, A a set in X of all absolutely continuous functions $x(t)$ which satisfies (4.6) for some $u\in\mathcal{U}$, Y a Banach space of all continuous real valued functions on I with the uniform norm, and B a closed convex cone with non-empty interior in Y defined by

$$B=\{y\in Y \mid y(t)\leqq 0 \text{ for all } t\in I\}\,.$$

Let $\tilde{f}$ be a real valued convex function on X defined by $\tilde{f}(x)=\int_{t_0}^{t_1} f(x(t))dt$, g a function from X into Y defined by $\tilde{g}(x)=g(x(\cdot))$ and h a function from X into R^{q+r} defined by $h(x)=(Sx(t_0)-w_0,\, Tx(t_1)-w_1)$. Then, the problem ($P_1$) is restated as follows:

$$(P_2)\qquad \begin{aligned} &\text{minimize} && \tilde{f}(x)\\ &\text{subject to} && x\in A\,,\\ &&& \tilde{g}(x)\in B\,,\\ &&& h(x)=0\,. \end{aligned}$$

It is true that the set A is convex and that the functions $\tilde{f}(x)$, $\tilde{g}(x)$ and $h(x)$ are differentiable in the sense of Neustadt. Moreover, the differentials are of the following forms:

$$\tilde{f}_{\bar{x}}(x)=\int_{t_0}^{t_1}\frac{\partial f(\bar{x}(t))}{\partial x}x(t)dt\,,$$

$$\tilde{g}_{\bar{x}}(x)=\frac{\partial g(\bar{x}(t))}{\partial x}x(t)\,,$$

$$h_{\bar{x}}(x)=(Sx(t_0),\, Tx(t_1))\,,$$

where $\dfrac{\partial f(\bar{x}(t))}{\partial x}$ and $\dfrac{\partial g(\bar{x}(t))}{\partial x}$ are n-dimensional row vectors.

LEMMA 4.1. *If the condition* (4.7) *is satisfied, then Regularity Assumption* (*RA*) *holds.*

PROOF. Let x_0 be the vector satisfying the condition (4.7). It is clear that $x_0-\bar{x}\in A-\bar{x}\subset \mathrm{P}(A,\bar{x})$ and that $\tilde{g}(x_0)\in \operatorname{int} B$. Since $\tilde{g}(x)$ is B-convex in x, we have

$$\tilde{g}(\bar{x})+\tilde{g}_{\bar{x}}(x_0-\bar{x})=\tilde{g}(x_0)+\{\tilde{g}_{\bar{x}}(x_0-\bar{x})-[\tilde{g}(x_0)-\tilde{g}(\bar{x})]\}\in \operatorname{int} B+B\subset \operatorname{int} B\,.$$

We also obtain

$$\begin{aligned} h_{\bar{x}}(x_0-\bar{x})&=(S(x_0(t_0)-\bar{x}(t_0)),\, T(x_0(t_1)-\bar{x}(t_1)))\\ &=(w_0-w_0,\, w_1-w_1)=(0,\,0)\,. \end{aligned}$$

Let $x(t\,;\xi)$ denotes the trajectory corresponding to the controller $\bar{u}(t)$ with $x(t_0\,;\xi)=\xi$. Then, we have

$$h_{\bar{x}}(x(t;\xi)-\bar{x}(t))=(S(x(t_0;\xi)-\bar{x}(t_0)),\ T(x(t_1;\xi)-\bar{x}(t_1)))$$
$$=(S(\xi-\bar{x}(t_0)),\ T\Phi(t_1)(\xi-\bar{x}(t_1))) \quad \text{for all} \quad \xi\in R^n.$$

Since rank$(S)=q$ and rank$(T)=r$, $h_{\bar{x}}(x(t;\xi)-\bar{x}(t))$ runs the whole space R^{q+r} as ξ varies the whole space R^n. Hence, the origin is an interior point of $h_{\bar{x}}(\mathrm{P}(A,\bar{x}))$.

Q. E. D.

It is easy to see that all the hypotheses in Theorem 3.3 are satisfied, and hence it is necessary and sufficient for the vector $\bar{x}$ to be an optimal solution to (P_2) that there exist a (row) vector $(\bar{\alpha},\bar{\beta})\in R^q\times R^r$ and a linear continuous functional $\bar{y}^*\in Y^*$ such that

(4.8) $$\bar{y}^*(y)\leqq 0 \quad \text{for all} \quad y\in B,$$

(4.9) $$\bar{y}^*(\tilde{g}(\bar{x}))=0,$$

(4.10) $$\tilde{g}(\bar{x})\in B, \quad h(\bar{x})=0,$$

(4.11) $$f_{\bar{x}}(x)+\bar{y}^*(\tilde{g}_{\bar{x}}(x))+(\bar{\alpha},\bar{\beta})\cdot h_{\bar{x}}(x)\geqq 0 \quad \text{for all} \quad x\in\mathrm{P}(A,\bar{x}).$$

It is valid, on the basis of the representation theorem for the space Y^*, that there is a real valued function $v(t)$ of bounded variation on I which satisfies

$$\bar{y}^*(y)=\int_{t_0}^{t_1}y(t)dv(t) \quad \text{for all} \quad y\in Y.$$

Furthermore, we can assume that $v(t)$ is continuous from right on (t_0, t_1) and that $v(t_0)=0$. By virtue of (4.8), we have

$$\int_{t_0}^{t_1}y(t)dv(t)\leqq 0 \quad \text{whenever} \quad y(t)\leqq 0 \text{ for all } t\in I,$$

which implies that $\frac{dv(t)}{dt}\geqq 0$ for almost all $t\in I$. It is immediate, by (4.9) and (4.10), that

$$\int_{t_0}^{t_1}g(\bar{x}(t))dv(t)=0 \quad \text{and} \quad g(\bar{x}(t))\leqq 0 \quad \text{for all} \quad t\in I,$$

which implies that $v(t)$ is constant on subinterval of I on which $g(\bar{x}(t))<0$. It is valid, by (4.11), that

(4.12) $$\int_{t_0}^{t_1}\frac{\partial f(\bar{x}(t))}{\partial x}(x(t)-\bar{x}(t))dt+\int_{t_0}^{t_1}\frac{\partial g(\bar{x}(t))}{\partial x}(x(t)-\bar{x}(t))dv(t)$$
$$+\bar{\alpha}S(x(t_0)-\bar{x}(t_0))+\bar{\beta}T(x(t_1)-\bar{x}(t_1))\geqq 0 \quad \text{for all} \quad x\in A.$$

Let u and $\bar{u}$ be controllers in $\mathcal{U}$ corresponding to x and $\bar{x}$, respectively. Then, we have

$$\bar{x}(t)=\Phi(t)\bar{x}_0+\Phi(t)\int_{t_0}^{t}\Phi(s)^{-1}B(s)\bar{u}(s)ds \quad \text{for all} \quad t\in I,$$

where $\Phi(t)$ is a fundamental matrix solution of the equation $\dot{x}=A(t)x$ with $\Phi(t_0)=$ identity. By the formula for integration by parts for Lebesgue-Stieltjes integral, (4.12) becomes

$$\int_{t_0}^{t_1} \frac{\partial f(\bar{x}(t))}{\partial x}(x(t)-\bar{x}(t))dt+\frac{\partial g(\bar{x}(t_1))}{\partial x}(x(t_1)-\bar{x}(t_1))v(t_1)$$
$$-\int_{t_0}^{t_1} v(t)\Big\{\frac{\partial p(\bar{x}(t), t)}{\partial x}(x(t)-\bar{x}(t))+\frac{\partial g(\bar{x}(t))}{\partial x}B(t)(u(t)-\bar{u}(t))\Big\}dt$$
$$+\bar{\alpha}S(x(t_0)-\bar{x}(t_0))+\bar{\beta}T(x(t_1)-\bar{x}(t_1))\geqq 0 \qquad \text{for all} \quad x\in A, \tag{4.13}$$

where $p(x, t)\equiv\frac{\partial g(x)}{\partial x}(A(t)x+B(t)u(t))$. It is easy to verify that for every $x\in A$ satisfying $x(t_0)=\bar{x}(t_0)$, we have, by (4.13),

$$\int_{t_0}^{t_1}\Big[\psi(t)-v(t)\frac{\partial g(\bar{x}(t))}{\partial x}\Big]B(t)[u(t)-\bar{u}(t)]dt\geqq 0 \qquad \text{for all} \quad u\in\mathcal{U}, \tag{4.14}$$

where $\psi(t)$ is an absolutely continuous, n-dimensional (row) vector valued function on I which satisfies

$$\frac{d\psi(t)}{dt}=-\psi(t)A(t)+\frac{\partial f(\bar{x}(t))}{\partial x}+v(t)\frac{\partial p(\bar{x}(t), t)}{\partial x} \qquad \text{for almost all } t\in I, \tag{4.15}$$

$$\psi(t_1)=v(t_1)\frac{\partial g(\bar{x}(t_1))}{\partial x}+\bar{\beta}T\bar{x}(t_1), \tag{4.16}$$

$$\psi(t_0)=\int_{t_0}^{t_1}\frac{\partial f(\bar{x}(t))}{\partial x}\Phi(t)dt+v(t_1)\frac{\partial g(\bar{x}(t_1))}{\partial x}\Phi(t_1)$$
$$-\int_{t_0}^{t_1}v(t)\frac{\partial p(\bar{x}(t), t)}{\partial x}\Phi(t)dt+\bar{\beta}T\bar{x}(t_1)\Phi(t_1). \tag{4.17}$$

For each $x\in A$ whose corresponding controller is $\bar{u}$, (4.13) implies that

$$\int_{t_0}^{t_1}\frac{\partial f(\bar{x}(t))}{\partial x}\Phi(t)dt+v(t_1)\frac{\partial g(\bar{x}(t_1))}{\partial x}\Phi(t_1)$$
$$-\int_{t_0}^{t_1}v(t)\frac{\partial p(\bar{x}(t), t)}{\partial x}\Phi(t)dt+\bar{\alpha}S\bar{x}(t_0)+\bar{\beta}T\bar{x}(t_1)\Phi(t_1)=0. \tag{4.18}$$

It is trivial, by (4.17) and (4.18), that

$$\psi(t_0)=-\bar{\alpha}S\bar{x}(t_0). \tag{4.19}$$

Let $\theta\in I$ be a regular point for both $v(t)$ and $\bar{u}(t)$ and ε an arbitrary positive number. Define a controller $u(t)$ by

$$u(t)=\begin{cases} w & \text{for} \quad \theta-\varepsilon<t<\theta, \\ \bar{u}(t) & \text{otherwise}, \end{cases}$$

where $w\in U$. It then follows from (4.14) that

$$\frac{1}{\varepsilon}\int_{\theta-\varepsilon}^{\theta}\Big[\psi(t)-v(t)\frac{\partial g(\bar{x}(t))}{\partial x}\Big]B(t)[w-\bar{u}(t)]dt$$
$$\xrightarrow[\varepsilon\to 0+]{} \Big[\psi(\theta)-v(\theta)\frac{\partial g(\bar{x}(\theta))}{\partial x}\Big]B(\theta)[w-\bar{u}(\theta)]\geqq 0 \qquad \text{for all} \quad w\in U.$$

Hence, we have

$$(4.20) \qquad \left[\psi(t)-v(t)\frac{\partial g(\bar{x}(t))}{\partial x}\right]B(t)\bar{u}(t) = \min_{w\in U}\left[\psi(t)-v(t)\frac{\partial g(\bar{x}(t))}{\partial x}\right]B(t)w \qquad \text{for almost all } t\in I.$$

Thus, we have arrived at the following conclusion.

THEOREM 4.2. *Consider the linear optimal control problem* (P_1) *and suppose that the condition* (4.7) *holds. Then, in order for the trajectory* $\bar{x}(t)$ *with the controller* $\bar{u}(t)$ *to be an optimal solution, it is necessary and sufficient that there exist* (*row*) *vectors* $\bar{\alpha}\in R^q$ *and* $\bar{\beta}\in R^r$, *a real valued function* $v(t)$ *of bounded variation which is continuous from right on* (t_0, t_1) *and is constant on subintervals of* I *on which* $g(\bar{x}(t))<0$ *such that* $v(t_0)=0$, $\frac{dv(t)}{dt}\geqq 0$ *for almost all* $t\in(t_0, t_1)$, *and an absolutely continuous n-dimensional* (*row*) *vector valued function* $\psi(t)$ *on* I *which satisfies* (4.15), (4.16), (4.19) *and* (4.20).

Acknowledgement. The author wishes to express his hearty thanks to Professor Nagata Furukawa of Kyushu University for his useful comments and corrections to original versions of this paper.

References

[1] S. TAGAWA, *On a class of nonlinear programming in a Banach space.* Bull. Math. Statist., 16 (1974), pp. 95-114.

[2] L.W. NEUSTADT, *An abstract variational theory with applications to a broad class of optimization problems. I. General Theory.* SIAM J. Control, 4 (1966), pp. 505-527.

[3] P. VARAIYA, *Nonlinear programming in Banach space.* SIAM J. Appl. Math., 15 (1967), pp. 284-293.

[4] N. DUNFORD and J.T. SCHWARTZ, Linear Operators, Part I, Interscience, New York, 1964.

[5] S. TAGAWA, *On a convergence of minimizing sequences for uniformly convex functions,* Bull. Math. Statist., 17 (1976), pp. 1-16.

[6] S. TAGAWA, *Generalized convexities of Continuous functions and their applications to mathematical programming.* Bull. Math. Statist., 16 (1974), pp. 115-125.

[7] D.G. LUENBERGER, Optimization by Vector Space Methods. John Wiley and Sons, 1969.