UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS
Departamento de Estadistica

TESIS DOCTORAL

Un análisis de algunos principios Básicos de la teoría de la optimización

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Miguel Martín Dávila

Madrid, 2015

Miguel Martín Dávila

UN ANALISIS DE ALGUNOS PRINCIPIOS BASICOS EN LA TEORIA DE LA OPTIMIZACION

Departamento de Estadística
Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid
1983

Colección Tesis Doctorales. Nº 139/83

Edita e imprime la Editorial de la Universidad
Complutense de Madrid. Servicio de Reprografía
Noviciado, 3 Madrid-8
Madrid, 1983
Xerox 9200 XB 480
Depósito Legal: M-19634-1983

UN ANALISIS DE ALGUNOS PRINCIPIOS BASICOS EN LA TEORIA DE LA OPTIMI ZACION

Tesis doctoral presentada por D. Miguel Martín Dávila, en - la Facultad de CC. Matemáti-- cas de la Universidad Complu- tense y realizada bajo la di- rección del Dr. D. Ildefonso Yañez de Diego.

Madrid, Enero de 1982

I N D I C E

CAPITULO I

Teoría de las condiciones necesarias en el caso convexo. El Principio de la Proyección Maximal.

CAPITULO II

Teoría de las condiciones necesarias en el caso diferenciable. El Principio de la Proyección Maximal Tangencial

CAPITULO III

Teoría de las condiciones suficientes. El Principio de la Optimización en Desigualdades y el Principio del Minimax.

El establecimiento de dichas caracterizaciones de una manera que permita la generación de dichas condiciones necesarias es lo que nosotros hemos concretado en nuestros principios.

En relación, sin embargo, a la discusión de la generación de condiciones suficientes de optimalidad, el desarrollo ya es totalmente diferente y está conectado, como es tradicional, más con ideas intuitivas sobre las que se podrían asentar éstas, en un enfoque que podríamos denominar, tomándo un anglicismo del argot informático, "multipropósito", más que con el análisis, con más aparato técnico, de un determinado problema tipo planteado en unos términos matemáticos precisos. Una muestra de este enfoque lo sería, por ejemplo, toda la metodología elaborada a partir del Principio de Optimalidad de R. Bellman, mediante el cual lo único que se hace es enunciar una observación de sentido común. Así, aunque algo más elaborados, los principios básicos que comentamos en relación a la generación de condiciones suficientes, lo serán en términos generales, más que para alguna versión particular de un problema concreto.

Entrando más en detalles, en el primer capítulo se presenta el primero de estos principios básicos, que denominamos Principio de la Proyección Maximal, el cuál no es otro que una adaptación del conocido teorema del hiperplano soporte o de separación de conjuntos convexos.

En nuestra presentación, sin embargo, mostramos que se puede convertir en un importante instrumento operativo para generar las condiciones necesarias de optimalidad en el caso convexo desde el punto de vista geométrico, que en el caso estático tendrían como teorema más representativo el (24), y en el caso dinámico el (79) y que esta es la base de todos los demás resultados intermedios -

de interés en la generación de condiciones necesarías con convexi dad.

El segundo capítulo constituye la parte más sustancial de - este trabajo. En el se establece el que denominamos Principio de la Proyección Maximal Tengencial, que viene dado por el Teorema - (91), a partir del cuál se reconstruirian los dos resultados más importantes en este área, el Teorema de Kuhn-Tucker, y, sobre todo, el Principio de Máximo de Pontryaguin.

Este Principio (91), que aparece, dondo luego, en parte, implícito en muchos estudios del problema no se ha formulado nunca, que sepamos, en la misma manera en que lo hemos hecho nosotros -- aquí, y tampoco se ha investigado lo "próximos" que pueden estar a este Principio todos los resultados relevantes de la teoría.

La búsqueda de resultados del tipo del (91) no es, por su - puesto, reciente, pero casi todos los estudios se concretan más - en el establecimiento de estos resultados sobre el funcional que generaría el conjunto de accesibilidad más que sobre la imagen de dicho funcional en el espacio de fase, como nosotros. Esto ha dado lugar a la concreción de varias "reglas de existencia de multiplicadores en espaciosabstractos", generalización de los usuales a espacios arbitrarios, mediante las cuales podrían obtenerse todos los resultados de interés (Véase la Nota 24). Sin embargo, - para el establecimiento de dichas reglas se requiere, en general, que el funcional en cuestión verifique unas determinadas condiciones de regularidad, las cuales aparecen mucho más "diluidas" en - nuestra aproximación al imponer dichas condiciones a la imagen -- del funcional y no al propio funcional.

Todo el capítulo II se dedica, pues, tras el establecimiento de(91) a mostrar que la caracterización que representa es todo lo que se necesita para la obtención de los principales resultados de la teoría. En esa línea el resultado que correspondería al caso estático en nuestro desarrollo sería más bien el Teorema (95), más que el propio Teorema de Kuhn-Tucker, aunque respecto a este último mostramos como se puede concretar con el Teorema de la Dualidad en Programación Líneal, el cuál a su vez se seguiría del Principio de la Proyección Maximal.

La principal dificultad para la obtención de (95) y del Principio de Máximo de Pontryaguin mediante la vía de (91) estriba en la garantía de definición de conos convexos de vectores tangentes en los puntos de la frontera de los apropiados conjuntos de accesibilidad en cada caso. Por lo que respecta a (95) esta dificultad se puede sortear sin excesivo problema pero por lo que respecta al Principio de Máximo se requiere algo más de trabajo.

Es por esto por lo que la demostración del Principio de Máximo en nuestra versión pueda parecer algo larga y prolija. Sin embargo esto solo es debido a que hemos procurado discutir todos los detalles de la demostración para que ésta quedara suficientemente aclarada en todos sus términos, sobre todo en lo concerniente a la manera en como el Principio (91) podría hacerse operativo en ese caso. La demostración es pues, aunque larga, elemental en el sentido de que no incluye ninguna formulación ni resultado que no sea básico en la teoría de la integración de Lebesgue y de su aplicación en el establecimiento de teoremas de existencia y unicidad de soluciones en un sistema de ecuaciones diferenciales, y en ese nivel no es excesivamente más larga que otras demostraciones "elementales" en el sentido que hemos precisado antes, que se han elaborado del Principio de Máximo, incluída la presentada por el propio Pontryaguin (Véase la Nota 7).

Así pues, la extensión de la demostración no debería ocultar la estrecha conexión entre el Principio de Máximo y (91), del cuál, y salvo los problemas técnicos mencionados antes, el primero podría considerarse una consecuencia directa.

Por último, en el capítulo III presentamos dos principios, que denominamos el Principio de Optimización en desigualdades y el Principio del Minimax, los cuales, en la línea ya indicada antes de establecer requisitos generales de aplicación a cualquier tipo de problema de optimización, permitirían la obtención de condiciones suficientes en problemas de optimización.

Con la introducción de estos principios pretendemos dar una visión completa de la caracterización de condiciones de optimalidad en problemas de optimización al incluir los dos tipos en que podrían subdividirse estas: necesarias y suficientes.

En la concreción de estas condiciones suficientes de optimalidad para problemas tipo concretos no hemos descendido el grado de detalle en que lo hemos hecho en los capítulos I y II porque, aparte de que la menor especialización de estos principios que el (8) ó (91) de los capítulos anteriores exigiría unos desarrollos adicionales en cada caso, ello hubiera dado a esta monografía una longitud excesiva. Sin embargo, si presentamos algunos breves resultados concretos concernientes a algunos aspectos del establecimiento de condiciones suficientes en el problema estandar de Programación No Líneal y en las relaciones de nuestro enfoque con la teoría de la Dualidad.

Solo nos queda, pues, por aclarar, que, salvo en lo concerniente al enunciado en varios teoremas sobradamente conocidos en la teoría de la optimización, tales como el de Kuhn-Tucker, el --

CAPITULO I

TEORIA DE LAS CONDICIONES NECESARIAS EN EL CASO CONVEXO. EL PRINCIPIO DE LA PROYECCION MAXIMAL

1.- PLANTEAMIENTO GENERAL DEL PROBLEMA

Para normalizar el planteamiento de los problemas a investigar empezaremos conviniendo que los tipos de problemas que estudiaremos serán casos particulares del problema general de control siguiente:

(1) PROBLEMA GENERAL:

Sean $x(t)$ y $u(t)$ para cada $t \in T=[t_0,t_1]$, dos vectores n y m-dimensionales de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, conectados, para cada t, por el sistema de ecuaciones:

$$(2) \quad x(t) = \int_{t_0}^{t} f^{i1}(x(t'),\ u(t'),\ t')\,dt' + x(t_0) + f^{i2}(u(t)) \ ,$$

en donde f^{i1} y f^{i2} son funciones vectoriales de $\mathbb{R}^{n+m+1}$ y $\mathbb{R}^m$ en $\mathbb{R}^n$ cuyas propiedades se concretaran más adelante , $t_0, t_1 \in \mathbb{R}$, y la integral se interpreta en el sentido de Lebesgue.

Supondremos además que $u(t) \in \Omega$, $\forall t \in T$, y que la función $u : t \longrightarrow u(t)$, ha de pertenecer a una determinada clase de funciones admisibles Δ.

Se trata entonces de minimizar el funcional g_0 dado por:

$$(3) \quad x_0(t_1) = g_0(u(t),\ t_0, t_1) = \int_{t_0}^{t_1} f_0^{i1}(x(t),\ u(t),\ t)\,dt + f_0^{i2}(u(t_1)) \ ,$$

con las f_0^{ii} reales y las condiciones de contorno:

$$x(t_i) \in G_i \subset \mathbb{R}^n, \qquad (i=0,1)$$

Si T se redujera a un punto de $\mathbb{R}$ entonces $t_0 = t_1 = t$ y $G_1 = G_2 = G$, con lo que el problema se reduciría a uno estático.

Denotaremos de aquí en adelante:

$$(4) \qquad \hat{x} = \begin{bmatrix} x_o \\ x_1 \\ \vdots \\ x_n \end{bmatrix}; \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \qquad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix};$$

$$\hat{f}^{i} = \begin{bmatrix} f_o^{i} \\ f_1^{i} \\ \vdots \\ f_n^{i} \end{bmatrix}; \qquad f = \begin{bmatrix} f_1^{i} \\ f_2^{i} \\ \vdots \\ f_n^{i} \end{bmatrix} \quad ; i = 1,2.$$

Empezaremos nuestro estudio por el caso estático discutiendo la situación caracterizada porque el conjunto:

$$(5) \qquad C = \{\hat{x} \in \mathbb{R}^{n+1} \;/\; \hat{x} = \hat{f}^{\Omega}(u) \quad ; u \in \Omega\}$$

es un conjunto convexo y cerrado de $\mathbb{R}^{n+1}$. (Observese que en la definición de C no interviene G).

En esta primera etapa demostraremos que las condiciones necesarias de optimalidad pueden obtenerse a partir del "Principio de la Proyección Maximal" (que no es esencialmente más que el teorema de separación de conjuntos convexos) el cual comentaremos en el apartado siguiente, y que los resultados obtenidos en este caso, el estático convexo, pueden extenderse casi sin variaciones al caso dinámico continuo (en el tiempo) y convexo, sin más que exigir a f^{1} las condiciones usuales de regularidad. Luego compararemos los resultados obtenidos de esta manera en el caso dinámico convexo con los que se obtendrían utilizando el principio de Máximo de Pontryagin.

En el siguiente capítulo mostraremos como otro principio sencillo, que denominamos el "Principio de la Proyección Tangencial Maximal" pueda utilizarse también para generar condiciones necesarias de optimalidad en el caso no convexo, tanto estático como dinámico continuo, con lo que se obtendrá una demostración elemental del Principio de Máximo en el caso diferenciable.

Con esto concluiremos el estudio de la obtención de las condiciones necesarias de optimalidad en un problema de control por la via geométrica.

2.- EL CASO ESTATICO CONVEXO

De acuerdo con lo proyectado empezamos discutiendo la particularización del Problema general (1) correspondiente al caso estático, esto es:

(6) PROBLEMA:

$$\min \quad x_o(u) = f_o(u)$$

sujeto a:

$$x = f(u) \quad ; \quad u \in \Omega; \qquad x \in G$$

con la hipótesis adicional de que el conjunto C es convexo y cerrado.

Las condiciones necesarias para la existencia de un óptimo pueden obtenerse a partir de la consideración de que dicho punto óptimo $\hat{x}$, si existe, debe hallarse evidentemente en la frontera inferior del conjunto C respecto a la coordenada x_o, es decir en el conjunto:

$$(7) \qquad I = \{\hat{x} = (x_o, x_1, \ldots, x_n)' \in C \ / \text{ no existe } \hat{x}^* = (x_o^*, x_1^*, \ldots, x_n^*)' \in C, \ \text{con } x_o^* < x_o\}$$

El problema estriba pues, en caracterizar los puntos de I uno (o varios) de los cuales sería el óptimo, si éste existe.

La caracterización del conjunto I que utilizaremos es la que se sigue del que denominaremos "Principio de la Proyección Maximal" que se enuncia de la siguiente manera:

(8) TEOREMA (Principio de la Proyección Maximal)

Sea C un conjunto convexo y cerrado de $\mathbb{R}^{n+1}$ e I su frontera inferior no vacía respecto a la primera coordenada de sus puntos. Entonces:

$$(9) \qquad I = \overline{\{\hat{\bar{x}} \in C \ / \ \exists\, \hat{\eta} = (\eta_o, \eta_1, \ldots, \eta_n), \ \text{con } \eta_o < 0, \ \text{tal que } \hat{\eta}(\hat{x} - \hat{\bar{x}}) \leq 0, \ \forall \hat{x} \in C\}}$$

en donde $\overline{\{\ \}}$ denota la clausura del conjunto $\{\ \}$.
(Vease la Nota 0).

DEMOSTRACION:

(a) Supongamos que $\hat{\bar{x}} = (\bar{x}_o, \bar{x}_1, \dots, \bar{x}_n)' \in C$, verifica la condición de existencia de un $\hat{\eta}$ con $\eta_o < 0$ tal que $\eta.(\hat{x}-\hat{\bar{x}}) \leq 0$, $\forall \hat{x} \in C$. Entonces si $\hat{x} = (x_o \bar{x}_1, \dots, \bar{x}_n)' \in C$ se sigue que $x_o \geq \bar{x}_o$. Además como I es cerrado, como se puede comprobar facilmente, se deduce de lo anterior que $I \supset \overline{\{\ \}}$.

(b) Recíprocamente, supongamos que $\hat{\bar{x}} \in I$. Sea $\hat{y}^{(1}$ un punto de $\mathbb{R}^{n+1}$ no perteneciente a C. Calculemos el mínimo de las distancias de $\hat{y}^{(1}$ al conjunto C. Como C es cerrado existirá por lo menos un punto de C perteneciente a su frontera, $\hat{x}^{(1}$, tal que:

$$||\hat{x}^{(1} - \hat{y}^{(1}|| = \min_{\hat{x} \in C} ||\hat{x} - \hat{y}^{(1}|| > 0$$

en donde $||\hat{x}|| = [\Sigma\ x_i^2]^{1/2}$.

Entonces, si denotamos:

$$\hat{\eta}^{(1} = \frac{(\hat{y}^{(1} - \hat{x}^{(1})'}{||\hat{y}^{(1} - \hat{x}^{(1}||} ,$$

el hiperplano de ecuación,

$$\hat{\eta}^{(1} . \hat{x} = \hat{\eta}^{(1} . \hat{x}^{(1} = \text{cte} = c^{(1}$$

es tal que: $\hat{\eta}^{(1} . \hat{x} \leq c^{(1}$, $\forall \hat{x} \in C$, ya que, en otro caso, si $\hat{\eta}^{(1} . \hat{x}^* > c^{(1}$, para algún $\hat{x}^* \in C$, como el segmento:

$$\{\lambda \hat{x}^* + (1-\lambda)\ \hat{x}^{(1}\ /\ 0 \leq \lambda \leq 1\}$$

está contenido en C, por ser convexo, y tiene algún punto interior a la hiperesfera:

$$||\hat{x} - \hat{y}^{(1}||^2 = ||\hat{x}^{(1} - \hat{y}^{(1}||^2 ,$$

por ser $\hat{\eta}^{(1}$ perpendicular al hiperplano tangente a la hiperesfera en el punto $\hat{x}^{(1}$, habría algún punto de C a menor distancia de $\hat{y}^{(1}$ que $\hat{x}^{(1}$, en contra de lo supuesto.

Consideremos ahora una sucesión de puntos $\hat{y}^{(j}$ no pertenecientes a C y convergente a $\hat{\hat{x}}$. Sean $\hat{x}^{(j}$ y $\hat{\eta}^{(j}$ los correspondientes puntos minimizantes y vectores característicos de los hiperplanos obtenidos de la forma indicada para $j = 1$. Como $\|\eta^{(j}\| = 1$, $\forall j$, se puede extraer una subsucesión convergente de la $\{\eta^{(j}\}$, la cual volveremos a denotar por $\{\eta^{(j}\}$, de límite $\hat{\bar{\eta}}$.

Es fácil comprobar que los $x^{(j}$ correspondiente a la subsucesión anterior convergen a $\hat{\hat{x}}$, por lo que se concluiría que:

$$\hat{\bar{\eta}} \,.\, \hat{x} \leq \hat{\bar{\eta}} \,.\, \hat{\hat{x}} \quad , \quad \forall \hat{x} \in C$$

Es decir:

$$\hat{\bar{\eta}}.(\hat{x} - \hat{\hat{x}}) \leq 0, \quad \forall \hat{x} \in C$$

Por otra parte ya que $\hat{\hat{x}} \in I$, se sigue que se puede determinar el $\hat{\bar{\eta}} = (\bar{\eta}_o, \bar{\eta}_1, \ldots, \bar{\eta}_n)$ de manera que $\bar{\eta}_o \leq 0$. Esto sería evidentemente necesario si existiera algún otro $\hat{x} \in C$ tal que se diferencie de $\hat{\hat{x}}$ solo en su primera coordenada. En caso contrario bastaría tomar el límite en los $\hat{\eta}$ obtenidos para puntos de C con la propiedad anterior de existencia de otros $\hat{x} \in C$ con la primera coordenada diferente. (Si no existiera ningún punto $\hat{x} \in C$ con la propiedad anterior C se encontraría en un hiperplano de dimensión n con vector característico $\hat{\eta}$ tal que su primera coordenada η_o sería no nula, pues en otro caso la frontera inferior sería vacía, y se podría elegir el $\hat{\bar{\eta}}$ con $\bar{\eta}_o < 0$).

El párrafo anterior se puede aclarar más por medio del razonamiento siguiente: sea $\hat{x}^*$ un punto de I con la propiedad de que existen otros encima de él en la dirección de x_o. Sea $\hat{x}^{\#}$ uno de esos puntos situados encima de $\hat{x}^*$. Debido a la convexidad de C el conjunto: $\{\lambda_1 \hat{x}^* + \lambda_2 \hat{x}^{\#} + \lambda_3 \hat{\hat{x}} \;/\; \Sigma \lambda_i = 1;$ $\lambda_i \geq 0, \quad i=1,2,3\}$ está contenido en C. Si prolongamos ahora los puntos del segmento: $\{\lambda\ \hat{x}^* + (1-\lambda)\hat{\bar{x}}\ /\ 0 \leq \lambda \leq 1\}$ en la dirección x_o hasta I se sigue que existen puntos de I arbitrariamente próximos a $\hat{\bar{x}}$ con vectores característicos $\hat{\eta} = (\eta_o, \ldots, \eta_n)$ tales que $\eta_o \leq 0$. Es más, se podría asegurar que para dichos puntos debería de ser $\eta_o < 0$ ya que, en otro caso, el hipotético hiperplano vertical (es decir con $\eta_o = 0$) soporte dividiría en dos el segmento entre $\hat{x}^*$ y $\hat{\bar{x}}$ contra lo supuesto.

Finalmente si para un $\hat{\bar{x}} \in I$ dado se tuviera $\bar{\eta}_o = 0$, el razonamiento de los párrafos anteriores demuestra que existe una sucesión de puntos $\hat{x}^{(j}$ de I convergente a $\hat{\bar{x}}$ tal que los correspondientes vectores característicos se pueden elegir de manera que $\eta_o^{(j} < 0,\ \forall j$.

Se concluye entonces que $I \subset \{\overline{\ }\}$ lo que demuestra el teorema.

----

(10) Antes de mostrar algunos ejemplos en los que se aplica el resultado que acabamos de demostrar vamos a comentar algunos de sus aspectos más importantes. En primer lugar hay que señalar que la utilidad del Principio estriba en la descomposición del proceso de optimización en dos etapas. En la primera se determinaría I para el conjunto de restricciones del problema salvo la de que x pertenezca al conjunto G y en la segunda etapa se determinaría de entre los puntos del conjunto: $I \cap \{\hat{x} = \begin{bmatrix} x_o \\ x \end{bmatrix} / x \in G\}$, el óptimo.

Así, en el caso de que $G = \mathbb{R}^n$ el problema se reduciría sencillamente a determinar el mínimo de x_o con las restricciones impuestas a u, ó, lo que es equivalente, a determinar los puntos de I con vector característico del hiperplano soporte igual a $(-1,0,\ldots,0)$, mientras que en el caso opuesto, cuando G se reduce a un punto, el problema se limita a determinar el único punto de I, si existe, cuya proyección en $\mathbb{R}^n$ es G. Este segundo caso es el típico de los problemas dinámicos de control con estado final dado.

Entre los dos casos extremos anteriores hay toda una gama de casos intermedios cuya resolución, en lo que a la segunda etapa se refiere, sería tanto más difícil cuanto mayor fuera G. Por lo tanto la acotación del campo de búsqueda del óptimo en la primera etapa mediante la determinación de I sería más o menos superflua en relación al tamaño de G.

Por lo que respecta a la determinación de I en cada caso, puede observarse, de la misma presentación del Teorema (8), que de lo que se trata es de determinar previamente el conjunto:

$$(11) \quad L = \{\hat{\bar{x}} \in \mathbb{R}^{n+1} \;/\; \hat{\eta} = (\eta_o,\eta_1,\ldots,\eta_n) \quad \text{con} \quad \eta_o < 0, \quad \text{tal que} \quad \hat{\eta}.(\hat{x} - \hat{\bar{x}}) \leq 0, \quad \forall \hat{x} \in C\}$$

y luego terminar hallando I después de determinar la clausura del conjunto L. (En muchos casos el propio conjunto L será cerrado por lo que él mismo coincidiría con el conjunto I). Esta manera de proceder eliminaría la obtención de los puntos de la frontera de C con hiperplano soporte tal que $\eta_o = 0$ los cuales no necesariamente pertenecerían a I. Además, dado que $\eta_o < 0$, se pueden normalizar los rectores $\hat{\eta}$ a considerar conviniendo que $\eta_o = -1$, como es usual, lo que simplifica algo las consideraciones.

Conviene hacer notar también, a efectos de aplicación del Principio demostrado, la independencia supuesta entre el conjunto G y la variable de control u, pues en otro caso, éste podría no ser aplicable. Esta claro que en el caso de que G dependiera de u, el conjunto de condiciones: $x \in G(u)$ debería ser incorporado a las ecuaciones $\dot{x} = f(x,u)$, y quizás nuevas restricciones en los valores de u que hicieran compatibles ambas restricciones en las variables de estado, y el problema volvería a reformularse con $G = \mathbb{R}^n$.

Observese también que la condición $\hat{\eta}.(\hat{x} - \hat{\bar{x}}) \leq 0$, $\forall \hat{x} \in C$ puede enunciarse también de la siguiente manera equivalente:

$$\hat{\eta} . \hat{\bar{x}} = \max_{\hat{x} \in C} \hat{\eta} . \hat{x}$$

De ahí se sigue el nombre que se le ha dado al Principio.

Finalmente hay que indicar que, como se deduce de las características del problema, si se quisiera reducir el máximo de la expresión $\hat{\eta}.\hat{x}$ a la determinación del máximo de un funcional que dependa sólo de u, es decir si se quiere reducir el máximo en C a un máximo en u, habría que expresar $\hat{x}$ como función de u, lo que supondría sustituir todas las funciones $f_i(x,u)$, $i=0,1,\dots,n$, en la expresión $\hat{\eta}.\hat{x}$ por $f_i(x(u),u)$.

(12) EJEMPLO

El problema de la Programación Lineal.

Consideremos el siguiente problema de optimización

$$(13) \qquad \min \sum_{j=1}^{m} c_j \, x_j$$

sujeto a:

$$\sum_{j=1}^{m} a_{ij}\, x_j \geq b_i \; ; \qquad i = 1,2,\ldots.,n$$

$$x_j \geq 0 \quad ; \qquad j = 1,2,\ldots.,m$$

en donde las c_j y las a_{ij} son constantes arbitrarias.

Si definimos una nuevas variables x_i' y u_j de la siguiente manera:

$$x_i' = \sum_{j=1}^{m} a_{ij}\, x_j - b_i \; ; \qquad i = 1,2,\ldots.,n$$

$$u_j = x_j \; ; \qquad j = 1,2,\ldots.,m$$

y volvemos a poner, para normalizar la notación, $x_i = x_i'$, el problema puede replantearse dentro del esquema del Problema (6), esto es:

(14) $$\min x_o(u) = f_o(u) = \sum_{j=1}^{m} c_j\, u_j$$

sujeto a:

$$x_i = f_i(u) = \sum_{j=1}^{m} a_{ij}\, u_j - b_i; \qquad i = 1,2,\ldots,n$$

$$u \in \Omega = \{u \in \mathbb{R}^m \;/\; u_j \geq 0; \quad j = 1,2,\ldots,m\}$$

$$x \in G = \{x \in \mathbb{R}^n \;/\; x_i \geq 0; \quad i = 1,2,\ldots,n\}$$

Podemos ahora determinar la frontera inferior I del conjunto correspondiente C, en donde:

$$C = \{x \in \mathbb{R}^{n+1} \;/\; x_i = \sum_j a_{ij}\, u_j - b_i; \quad i = 1,2,\ldots,n;$$
$$x_o = \sum_j c_j\, u_j; \quad u \in \Omega\}$$

es evidentemente convexo y cerrado, ya que todas las funciones implicadas son lineales y los semiespacios $u_j \geq 0$ son convexos y cerrados. Para ello se puede aplicar el Principio de la Proyección Maximal lo que nos conduciría a resolver, para cada vector $\hat{\eta} = (\eta_o, \eta_1, \ldots, \eta_n)$ con $\eta_o < 0$, el siguiente problema de máximo:

$$\max_{\hat{x} \in C} \hat{\eta} \cdot \hat{x}$$

si, por normalizar, ponemos $\eta_o = 1$, el problema podría enunciarse así:

$$\max \left[- \sum_{j=1}^{m} c_j u_j + \sum_{i=1}^{n} \eta_i \left(\sum_{j=1}^{m} a_{ij} u_j - b_i \right) \right]$$

para $u_j \geq 0$; $j = 1, 2, \ldots, m$

es decir:

$$\max \sum_j \left[\sum_i a_{ij} \eta - c_j \right] u_j - \sum_i \eta_i b_i$$

para $u_j \geq 0$; $j = 1, 2, \ldots, m$

De acuerdo con este último planteamiento sólo habrá máximos finitos para aquellos vectores $\hat{\eta}$ tales que:

(15) $$\sum_i a_{ij} \eta_i - c_j \leq 0; \qquad j = 1, 2, \ldots, m$$

que serán, por lo tanto, los únicos admisibles, supuesto que el problema tenga solución finita.

Además, para todos aquellos $\hat{\eta}$ tales que:

$$\sum_i a_{ij} \eta_i < c_j; \qquad j = 1, 2, \ldots, m$$

el único control solución es $u_j = 0$; $j = 1, 2, \ldots, m$, sólo cuando algunas de las desigualdades $\sum_i a_{ij} \eta_i \leq c_j$ se verificara con el signo igual podrá el correspondiente u_j ser difente de cero.

Por otra parte puede comprobarse que el conjunto de los $\hat{x}$ obtenidos para los u maximizantes es cerrado y coincide, por tanto, con I.

Puesto que, en general, solo podrán verificarse a lo más n de las m desigualdades del sistema (15) con el signo igual, se concluiría que, salvo casos particulares, la solución del problema de Programación Lineal (13) solo incluiría, como máximo, tantas x_j distintas de cero como ecuaciones
lo cual podría ser significativo si n es mucho más pequeño que m. Este resultado es, desde luego, sobradamente conocido en la Teoría de la Programación Lineal.

El campo de búsqueda de posibles puntos solución $\hat{x}$ puede restringirse más si, como demostraremos en un teorema un poco más adelante,bastaraconsiderar las soluciones positivas del sistema (15). De cualquier manera, y ya que el conjunto G es demasiado grande, para solucionar este problema sería mucho más efectivo el aplicar algún algoritmo de optimización directa como los de aplicación usual.

Una de las mayores restricciones respecto a la aplicación del Principio (8) estriba en que, aparentemente, es necesario demostrar previamente la convexidad del conjunto C que corresponda. Esta comprobación puede suponer dificultades y en la mayoría de los casos, aún sencillos, será imposible al ser C no convexo.

Con el teorema que vamos a demostrar a continuación se pondrá en evidencia que el rango de aplicabilidad del Principio

(8) puede ser extendido considerablemente para cubrir incluso si tuaciones en las que C es no convexo.

El enunciado del teorema se refiere al caso estático más significativo que es el de la Programación Convexa.

(16) TEOREMA

Consideremos el siguiente problema de Programación No Lineal, denominando problema de Programación Convexa:

$$(17) \qquad \min \; x_o = f_o(u)$$

sujeto a:

$$x_i = -f_i(u); \qquad i = 1,2,\dots,n$$

$$u \in \Omega \subset \mathbb{R}^m$$

$$x \in G = \{x \in \mathbb{R}^n \; / \; x_i \geq 0; \qquad i = 1,2,\dots,n\}$$

en donde las funciones f_i; $i=0,1,2,\dots,n$ y el conjunto Ω son convexos.

Denotemos por $[C]$ la envoltura convexa del conjunto:

$$(18) \qquad C = \{\hat{x} \in \mathbb{R}^{n+1} \; / \; x_o = f_o(u),\; x_i = -f_i(u);\; i = 1,2,\dots,n; \; u \in \Omega\}$$

y sea $\hat{x} = \begin{bmatrix} \bar{x}_o \\ \bar{x} \end{bmatrix}$ un punto minimizante del Problema (17).

Entonces $\hat{x} \in I_{[C]}$, en donde $I_{[C]}$ denota la frontera inferior respecto a su primera coordenada del conjunto $[C]$.

DEMOSTRACION

Sea $\hat{x}^* = \begin{bmatrix} x_o^* \\ \bar{x} \end{bmatrix} \in [C]$. Evidentemente $[C]$ coincide con el conjunto de combinaciones convexas finitas de elementos de C, ya que este conjunto es convexo y contiene y está contenido en $[C]$. Existirán, por lo tanto, p puntos:

$$\hat{x}^{(k}, \qquad k = 1,2,\dots,p$$

pertenecientes a C tales que

$$\sum_{k=1}^{p} \lambda_k \hat{x}^{(k} = \hat{x}^*$$

con $\sum_{k=1}^{p} \lambda_k = 1; \quad \lambda_k \geq 0; \quad k = 1,2,\dots,p.$

Denotemos por $u^{(k}$ los correspondientes puntos de Ω tales que $x_o^{(k} = f_o(u^{(k})$, y $x_i^{(k} = -f_i(u^{(k})$.

Entonces, por las hipótesis de convexidad de las funciones f_i; $i = 0,1,2,\dots,n$, se obtiene:

$$x_o^* = \sum_{k=1}^{p} \lambda_k f_o(u^{(k}) \geq f_o(\sum_{k=1}^{p} \lambda_k u^{(k})$$

$$x_i = -\sum_{k=1}^{p} \lambda_k f_i(u^{(k}) \leq -f_i(\sum_{k=1}^{p} \lambda_k u^{(k})$$

por la convexidad de Ω se sigue que $\sum_k \lambda_k u^{(k} \in \Omega$. Además, de que $x_i = \bar{x}_i$; $i = 1,2,\dots,n$, se obtiene:

$$-f_i(\sum_{k=1}^{p} \lambda_k u^{(k}) \geq 0; \qquad i = 1,2,\dots,n$$

Por lo tanto, el punto $x_o = f_o(\sum_k \lambda_k u^{(k})$; $x_i = -f_i(\sum_k \lambda_k u^{(k})$, $i=1,2,\dots,n$, verifica las restricciones del Problema (17) de lo que se deduce que:

$$x_o^* \geq f_o(\sum_k \lambda_k u^{(k}) \geq \bar{x}_o$$

lo que completa la demostración.

Como consecuencia principal del Teorema (16) tendríamos la de que, de acuerdo con el Principio (8), si $\hat{\bar{x}}$ resuelve el Problema (17) y suponiendo que C es cerrado (Véase la Nota 0) lo que implica que [C] es convexo y cerrado, o bien existirá un vector $\hat{\bar{\eta}} = (-1,\bar{\eta}_1,\ldots,\bar{\eta}_n)$ tal que:

$$(19) \qquad \hat{\bar{\eta}} \cdot \hat{\bar{x}} = \max_{\hat{x} \in [C]} \hat{\bar{\eta}} \cdot \hat{x}$$

o bien existirá una sucesión de puntos $\hat{x}^{(k} \in I_{[C]}$ convergente a $\bar{x}$ y una sucesión de vectores $\hat{\eta}^{(k} = (-1,\eta_1^{(k},\ldots,\eta_n^{(k})$, de manera que:

$$(20) \qquad \hat{\eta}^{(k} \cdot x^{(k} = \max_{\hat{x} \in [C]} \hat{\eta}^{(k} \cdot \hat{x}$$

Y obsérvese que en el primer caso se podría sustituir (19) por:

$$(21) \qquad \hat{\bar{\eta}} \cdot \hat{\bar{x}} = \max_{\hat{x} \in C} \hat{\bar{\eta}} \cdot \hat{x}$$

ya que $\hat{\bar{x}} \in C$.

Podemos concluir por lo tanto que, si $\hat{\bar{x}}$ es un punto minimizante del Problema (17) y el conjunto C dado por (18) es cerrado, entonces:

$$(22) \qquad \hat{\bar{x}} \in \{\hat{x}^* \in R^{n+1} / \exists \hat{\eta}^* = (-1,\eta_1^*,\ldots,\eta_n^*) \text{ tal que: } \hat{\eta}^* \cdot \hat{x}^* = \max_{\hat{x} \in [C]} \hat{\eta}^* \cdot \hat{x}\}$$

----

(23) Vamos a completar ahora el resultado que acabamos de obtener en relación al Problema (17) con otro resultado clásico como es el de que, para este problema en particular, se puede asegurar que, si existe un punto minimizante $\hat{\bar{x}} = \begin{bmatrix} x_o \\ \bar{x} \end{bmatrix}$, existirá también un vector de multiplicadores $\hat{\bar{\eta}} = (\bar{\eta}_o,\bar{\eta})$ que verifica

la expresión (21) con la propiedad adicional de que $\hat{\eta} \leq 0$, y $\overline{\eta} \cdot \overline{x} = 0$ (en donde $\eta \leq 0$ (ó < 0) quiere indicar $\eta_i \leq 0$ (ó < 0), $i = 0,1,\ldots,n$).

Se puede demostrar también un recíproco del resultado anterior en el sentido de la existencia de dicho vector $\hat{\eta} = (\overline{\eta}_o, \overline{\eta})$ con $\overline{\eta}_o < 0$, $\overline{\eta} \leq 0$, y verificandose (21) para algún $\hat{x} = \begin{vmatrix} \overline{x}_o \\ \overline{x} \end{vmatrix}$ con $\overline{\eta} \cdot \overline{x} = 0$ garantiza la optimalidad del punto $\hat{x}$.

Vamos a establecer de una manera más precisa todo lo que acabamos de asegurar

(24) TEOREMA

(a) Sea $\hat{\overline{x}} = \begin{bmatrix} \overline{x}_o \\ \overline{x} \end{bmatrix} = \begin{bmatrix} f_o(\overline{u}) \\ -f(\overline{u}) \end{bmatrix}$ un punto minimizante solución del Problema (17).

Entonces, o bien existe un vector $\overline{\eta} = (\overline{\eta}_1, \overline{\eta}_2, \ldots, \overline{\eta}_n)$ tal que $\overline{\eta} \geq 0$ y $\overline{\eta} \cdot \overline{x} = 0$, de manera que:

$$(25) \quad \begin{cases} f_o(\overline{u}) + \overline{\eta} \cdot f(\overline{u}) \leq f_o(u) + \overline{\eta} \cdot f(u); & \forall u \in \Omega \\ f_o(\overline{u}) + \overline{\eta}\, f(\overline{u}) \geq f_o(\overline{u}) + \eta\, f(\overline{u}); & \forall \eta = (\eta_1, \ldots, \eta_n) \geq 0 \end{cases}$$

o bien existe un vector $\overline{\gamma} \in \mathbb{R}^n$, tal que $\overline{\gamma} \geq 0$; $\overline{\gamma} \neq 0$; $\overline{\gamma} \cdot f(\overline{u}) = 0$, de manera que $\overline{\gamma} \cdot f(u) \geq 0$ $\forall u \in \Omega$.

(b) Recíprocamente, supongamos que existe un punto $\hat{\overline{x}} = \begin{bmatrix} \overline{x}_o \\ \overline{x} \end{bmatrix} \in C$ tal que $\overline{x} \geq 0$ y un vector $\overline{\eta}$, con $\overline{\eta} \geq 0$ tales que se cumplen para ellos las dos desigualdades (25). Entonces $\overline{\eta} \cdot \overline{x} = 0$ y el punto $\hat{\overline{x}}$ es un punto minimizante del Problema (17).

DEMOSTRACION

(a) Supongamos previamente que C es cerrado. Sea

$\hat{x} = \begin{bmatrix} f_o(\bar{u}) \\ -f(\bar{u}) \end{bmatrix}$ un punto minimizante solución del Problema (17).

Denotemos:

$$N_1 = \{i \in \mathbb{N} \;/\; x_i(\bar{u}) = 0\}$$

$$N_2 = \{i \in \mathbb{N} \;/\; x_i(\bar{u}) > 0\}$$

en donde $\mathbb{N}$ representa el conjunto de los números naturales.

Se puede probar entonces que el conjunto de desigualdades: $x_i \geq 0$, $i \in N_2$, puede ser suprimido del problema de minimización ya que el punto $\hat{x}$ minimiza x_o para valores arbitrarios de las x_i si los subíndices de estas corresponden a N_2, y para valores positivos si los subíndices corresponden a N_1.

En efecto, supongamos que existe $\hat{x}^* = (x_o^*, x_1^*, \ldots, x_n^*)' = (f_o(u^*), -f_1(u^*), \ldots, -f_n(u^*))$, con $u^* \in \Omega$, tal que: $x_o^* < \bar{x}_o$, y $x_i^* \geq 0$ para todo $i \in N_1$. Consideremos el segmento:

$$\lambda\bar{u} + (1-\lambda)u^*, \quad \text{para} \quad 0 \leq \lambda \leq 1$$

el cual estará contenido en Ω por ser convexo.

Entonces la función

$$x(\lambda) = -f(\lambda\bar{u} + (1-\lambda)u^*)$$

es tal que:

$$x(\lambda) \geq -\lambda\ f(\bar{u}) - (1-\lambda)\ f(u^*) = \lambda\bar{x} + (1-\lambda)x^*$$

Tomando límites en la expresión anterior cuando $\lambda \uparrow 1$ (esto es cuando $\lambda \to 1$ por la izquierda) se obtiene:

$$\lim_{\lambda \uparrow 1} x(\lambda) \geq \bar{x}$$

de donde se sigue que:

$$\lim_{\lambda \uparrow 1} x_i(\lambda) > 0 \qquad \text{para todo} \quad i \in N_2$$

Por lo tanto existirán λ_i, $i \in N_2$, tales que $0 < \lambda_i < 1$ y de manera que: $x_i(\lambda) > 0$, para todo λ verificando $\lambda_i \leq \lambda \leq 1$, y para todo $i \in N_2$.

Sea $\lambda^* = \max_{i \in N_2} \lambda_i$. Tendremos entonces: $0 < \lambda^* < 1$.

Por consiguiente el punto:

$$\hat{x}(\lambda^*) = \begin{bmatrix} x_o(\lambda^*) \\ x(\lambda^*) \end{bmatrix} = \begin{bmatrix} f_o(\lambda^*\bar{u} + (1-\lambda)u^*) \\ -f(\lambda^*\bar{u} + (1-\lambda^*)u^*) \end{bmatrix}$$

sería un punto factible, como se comprueba facílmente, tal que:

$$x_o(\lambda^*) = f_o(\lambda^*\bar{u} + (1-\lambda^*)u^*) \leq \lambda^* f_o(\bar{u}) + (1-\lambda^*)\, f_o(u^*) =$$

$$= \lambda^*\bar{x}_o + (1-\lambda^*)x_o^* < \bar{x}_o$$

lo que contradiría la optimalidad de $\hat{\bar{x}}$.

Denotemos ahora por $\bar{x}_I$ y $\bar{x}_{II}$ los dos subvectores columna de $\bar{x}$ formados por las componentes de $\bar{x}$ correspondientes a subíndices pertenecientes a N_1 y N_2 respectivamente.

Según acabamos de probar el punto $\begin{bmatrix} \bar{x}_o \\ \bar{x}_I \end{bmatrix}$ resuelve el siguiente problema de minimización:

$$\min \; x_o = f_o(u)$$

sujeto a:

$$x_i = -f_i(u); \qquad i \in N_1$$

$$u \in \Omega$$

$$x_i \geq 0; \qquad i \in N_1$$

con Ω y las f_o, f_i $(i \in N_1)$, convexos.

Por el Teorema (16) existirá, por lo tanto, o bien un vector $(-1, \bar{\eta}_I)$ o bien una sucesión de vectores $(-1, \eta_I^{(k)})$ y puntos $\begin{bmatrix} x_o^{(k)} \\ x_I^{(k)} \end{bmatrix}$ de manera que se verifica (21) ó (20).

Pongamonos en el primer caso. Entonces podríamos escribir:

$$(-1, \bar{\eta}_I) \cdot \begin{bmatrix} \bar{x}_o \\ \bar{x}_I \end{bmatrix} = -f_o(\bar{u}) - \sum_{i \in N_1} \bar{\eta}_i(\bar{u}) \geq$$

$$\geq (-1, \bar{\eta}_I) \cdot \begin{bmatrix} x_o \\ x_I \end{bmatrix} = -f_o(u) - \sum_{i \in N_1} \eta_i \, f_i(u); \qquad \forall u \in \Omega$$

sigue, por tanto, que:

$$f_o(\bar{u}) + \sum_{i \in N_1} \bar{\eta}_i \, f_i(\bar{u}) \leq f_o(u) + \sum_{i \in N_1} f_i(u); \ \forall u \in \Omega \tag{26}$$

Completando el vector $\bar{\eta}_{II}$ con ceros en los lugares que corresponderían a los subíndices del conjunto N_2, de acuerdo con la partición hecha en el conjunto de subíndices del vector $\bar{x}$, se podría construir un vector $\bar{\eta}$ de manera que:

$$f_o(\bar{u}) + \bar{\eta} \cdot f(\bar{u}) \leq f_o(u) + \bar{\eta} \cdot f(u); \qquad \forall u \in \Omega \tag{27}$$

Evidentemente, por la definición de $\bar{\eta}$, se tendrá: $\bar{\eta} \cdot f(\bar{u}) = \bar{\eta} \cdot \bar{x} = 0$, de modo que (27) podría haberse escrito también así:

$$f_o(\bar{u}) \leq f_o(u) + \bar{\eta} \cdot f(u); \qquad \forall u \in \Omega \tag{28}$$

Sólo falta comprobar entonces que las componentes de $\bar{\eta}$ correspondientes a subíndices de N_1 o, lo que es equivalente, correspondientes a componentes de $f(\bar{u})$ nulas, son mayores o iguales que cero.

Supondremos que los conjuntos, para i fijo, siguientes:

$$\{u \in \Omega \ / \ f_i(u) > 0; \quad f_j(u) \leq 0 \quad \forall j \in N_1\}$$

son no vacíos, para todo $i \in N_1$, pues en otro caso la correspondiente restricción $f_i(u) \leq 0$ podría eliminarse del Problema (17) y el vector $\bar{\eta}$ de la expresión (28) podría elegirse con $\bar{\eta}_i = 0$.

Definamos, para cada $r \in \mathbb{N}$, las funciones $f_i^{(r}(u)$ para $i \in N_1$ y $u \in \Omega$, de la siguiente manera:

$$f_i^{(r}(u) = \begin{cases} f_i(u) & \text{si} \quad f_i(u) \leq 0 \\ r\ f_i(u) & \text{si} \quad f_i(u) > 0 \end{cases}$$

Estas nuevas funciones $f_i^{(r}(u)$ son convexas ya que, de $f_i(u) \leq$ $\leq f_i^{(r}(u)$, $\forall u \in \Omega$ y de $r.f_i(u) \leq f_i^{(r}(u)$ si $f_i(u) \leq 0$ se sigue:

$$f_i^{(r}(\lambda u^{(1}+(1-\lambda)u^{(2}) = \begin{cases} f_1(\lambda u^{(1}+(1-\lambda)u^{(2}) \leq \lambda f_i(u^{(1}) + (1-\lambda)f_1(u^{(2}) \leq \\ \leq \lambda f^{(r}(u^{(1}) + (1-\lambda)\ f^{(r}(u^{(2}) \\ r\ f_i(\lambda u^{(1}+(1-\lambda)u^{(2}) \leq r.\lambda.f_i(u^{(1}) + r(1-\lambda)f_i(u^{(2}) \leq \\ \leq \lambda f_i^{(r}(u^{(1}) + (1-\lambda)f_i^{(r}(u^{(2}) \end{cases}$$

El punto $\hat{x}$ continuará siendo solución del Problema (17) con las $f_i^{(r}$ en lugar de las f_i ya que el conjunto:

$$\{x \in \mathbb{R}^n \ / \ x = -f_i(u), \quad i=1,2,\ldots,n; \quad u \in \Omega; \quad x \geq 0\}$$

no cambia si se sustituyen las f_i por las $f_i^{(r}$.

Así, supuesto, como así hemos hecho, que se aplica(19), existirá una sucesión de vectores $\hat{\gamma}^{(r} = (\gamma_o^{(r}, \gamma^{(r}) \in \mathbb{R}^{n+1}$, que podemos escoger de manera que $\|\hat{\gamma}^{(r}\| = 1$, $\forall r$, siendo $\gamma_o^{(r} \geq 0, \gamma_i^{(r}=0, i \in N_2$, $\forall r$ y:

(29) $\qquad \gamma_o^{(r} \cdot f_o(\bar{u}) \leq \gamma_o^{(r} f_o(u) + \gamma^{(r} \cdot f^{(r}(u)\ ; \qquad \forall u \in \Omega$

se podría extraer de dicha sucesión $\hat{\gamma}^{(r}$ una subsucesión convergente a un vector $\hat{\gamma} = (\bar{\gamma}_o, \bar{\gamma})$ con $\bar{\gamma}_o \geq 0$ y $\|\hat{\bar{\gamma}}\| = 1$, la cual volveremos a denotar por $\hat{\gamma}^{(r}$.

Evidentemente, de la definición de las funciones $f^{(r}(u)$, y teniendo en cuenta que:

$$\lim_{r} \gamma_o^{(r} \cdot f_o(\bar{u}) = \bar{\gamma}_o \cdot f_o(\bar{u}) > -\infty$$

y que, por tanto:

$$\lim_{r \to \infty} \gamma^{(r} \cdot f^{(r}(u) > -\infty ; \qquad \forall u \in \Omega$$

se sigue de ello que: $\bar{\gamma} \geq 0$.

Además, ya que $\bar{\gamma}_i = 0$, $\forall i \in N_2$, se verificará que: $\bar{\gamma} \cdot f(\bar{u}) = 0$.

Ahora, teniendo en cuenta la positividad de $\bar{\gamma}$ y el que $f_i^{(r}(u) \geq f_i(u)$ $\forall i \in N_1$, se deduce de (29) que:

(30) $\qquad \bar{\gamma}_o \cdot f_o(\bar{u}) \leq \bar{\gamma}_o \cdot f_o(u) + \bar{\gamma} \cdot f(u); \qquad \forall u \in \Omega$

Ya que $\bar{\gamma} \cdot f(\bar{u}) = 0$, podemos volver a escribir (30) así:

(31) $\qquad \bar{\gamma}_o \cdot f_o(\bar{u}) + \bar{\gamma} \cdot f(\bar{u}) \leq \bar{\gamma}_o \cdot f_o(u) + \bar{\gamma} \cdot f(u); \qquad \forall u \in \Omega$

Si $\bar{\gamma}_o > 0$ podemos tomar como vector definitivo $\bar{\eta} \in \mathbb{R}^n$ el definido de la siguiente manera:

$$\bar{\eta}_i = \frac{\bar{\gamma}_i}{\gamma_o} \quad ; \qquad i=1,2,\dots,n$$

mientras que si $\bar{\gamma}_o = 0$, multiplicando la expresión (31) resultante por un número positivo adecuado y sumando el resultado con la expresión (27) se deduciría que, en cualquier caso, se puede asegurar la existencia de un vector $\bar{\eta} \geq 0$ tal que:

$$(32) \qquad f_o(\bar{u}) + \bar{\eta} \,.\, f(\bar{u}) \leq f_o(u) + \bar{\eta} \,.\, f(u); \qquad \forall u \in \Omega$$

con la propiedad de que $\bar{\eta} \,.\, f(\bar{u}) = 0$.

Por otra parte, en el caso de que (27) no se verificara para $\hat{\bar{x}}$, existiría una sucesión de punto $\hat{x}^{(k}$ pertenecientes a $I_{[c]}$ y una sucesión de vectores $\eta^{(k}$ tal que (27) se verificaría con $\hat{x}^{(k}$ y $\eta^{(k}$ en lugar de $\hat{\bar{x}}$ y de $\bar{\eta}$.

Por un razonamiento análogo al efectuado en el primer caso se seguiría que la tal sucesión existiría cualquiera que fuese el número $r \in \mathbb{N}$ en la modificación del problema (17) al cambiar las f_i por las $f_i^{(r}$, pues si no fuera así estaríamos en el primer caso.

De ello se deduce, por idénticos motivos a los expuestos antes, que los $\eta^{(k}$ podrían seleccionarse de manera que $\eta^{(k} \geq 0$, por lo que se concluiría que existiría una sucesión de puntos $\hat{x}^{(k} = \begin{bmatrix} x_o^{(k} \\ x^{(k} \end{bmatrix}$ correspondientes a $u^{(k} \in \Omega$, pertenecientes a la envoltura convexa del conjunto (18), y más concretamente a su frontera inferior, y una sucesión de vectores $\eta^{(k} \geq 0$, con $\eta^{(k} \,.\, \bar{x} = 0$, tales que:

$$(33) \qquad f_o(u^{(k}) + \eta^{(k} \,.\, f(u^{(k}) \leq f_o(u) + \eta^{(k} \,.\, f(u), \qquad \forall u \in \Omega, \ \forall k \in \mathbb{N}$$

y $\quad \hat{x}^{(k} \to \hat{\hat{x}}$.

Si normalizamos los vectores $(1, \eta^{(k)})$ de manera que su módulo valga la unidad se podrá extraer una subsucesión convergente al vector $(\overline{\gamma}_o, \overline{\gamma})$ con $\overline{\gamma}_o \geq 0$, $\overline{\gamma} \geq 0$, $(\overline{\gamma}_o, \overline{\gamma}) \neq (0,0)$.

Tomando límite para dicha subsucesión en (33) quedaría, por lo tanto:

$$(34) \qquad \overline{\gamma}_o \cdot f_o(\overline{u}) + \overline{\gamma} \cdot f(\overline{u}) \leq \gamma_o \cdot f(u) + \overline{\gamma} \cdot f(u); \qquad \forall u \in \Omega$$

$$\text{con} \quad \overline{\gamma} \cdot f(\overline{u}) = 0$$

Si $\overline{\gamma}_o > 0$ se puede construir un vector $(1,\overline{\eta})$ tal que se verifique (32) mientras que si $\gamma_o = 0$ resultaría que:

$$\overline{\gamma} \cdot f(u) \geq 0; \qquad \forall u \in \Omega$$

con $\qquad \overline{\gamma} \geq 0, \quad \overline{\gamma} \cdot f(\overline{u}) = 0 \quad$ y $\quad \overline{\gamma} \neq 0$

Para terminar la demostración de la parte (a) del Teorema basta comprobar que, supuesto que se verifica (32), también se cumple que:

$$f_o(\overline{u}) + \overline{\eta} \cdot f(\overline{u}) \geq f_o(\overline{u}) + \eta \cdot f(\overline{u}); \qquad \forall \eta \geq 0$$

Ahora bien, lo anterior es evidente ya que de $f(\overline{u}) \leq 0$ y $\eta \geq 0$ se sigue que:

$$\eta \cdot f(\overline{u}) \leq 0 = \overline{\eta} \cdot f(\overline{u})$$

Finalmente, supongamos ahora que C no es cerrado y planteemosnos de nuevo el Problema (17) en $\overline{C}$ (clausura de C). Si $\hat{\hat{x}}$ sigue siendo solución del problema en este conjunto valdría lo afirmado anteriormente. En otro caso existiría un

$\bar{x} = \begin{bmatrix} x_o^* \\ x^* \end{bmatrix}$ con $x^* \geq 0$ tal que $x_o^* < \bar{x}_o$. Supongamos que en este segundo caso no existiera un $\gamma \geq 0$, $\gamma \neq 0$ tal que $\gamma \, . \, f(u) \geq 0$, $\forall u \in \Omega$. Es fácil demostrar que la negación de lo anterior es equivalente a la afirmación de que existe un $u^{\#} \in \Omega$ tal que $f(u^{\#}) < 0$. Haciendo una combinación convexa adecuada con $u^{\#}$ y algún u^* tal que $f(u^*)$ fuera próximo a x^* se podría obtener entonces un $\hat{x}^{\#} \in C$ tal que $x^{\#} \geq 0$ y $x_o^{\#} < \bar{x}_o$ lo que contradiría la optimalidad de $\hat{\bar{x}}$ en C.

(b) Probaremos ahora el recíproco. Supongamos que existe $\hat{\bar{x}} \in C$ con $\hat{\bar{x}} \geq 0$ tal que se cumplen las desigualdades (25) para un vector $\bar{\eta} \geq 0$.

De la segunda desigualdad se obtiene:

$$\eta \, . \, f(\bar{u}) \leq \bar{\eta} \, . \, f(\bar{u}) \leq 0; \qquad \forall \eta \geq 0$$

ya que $\bar{\eta} \geq 0$ y $f(\bar{u}) \leq 0$.

Tomando en particular $\eta = 0$ se sigue que: $\bar{\eta} \, . \, f(\bar{u}) = 0$.

Teniendo en cuenta lo anterior de la primera desigualdad de (25) se deduciría que:

$$f_o(\bar{u}) \leq f_o(u) + \bar{\eta} \, . \, f(u) \leq f_o(u)$$

para todo $u \in \Omega$ tal que $f(u) \leq 0$, lo que completa la demostración.

----

(35) El Teorema que acabamos de demostrar juega un papel principal en la teoría de la Programación No Lineal sin diferenciabilidad (Véase la Nota 1).

Si se le hubiera exigido a las funciones f_i de (17) la condición adicional de que no existiera ningún vector $\bar{\gamma} = 0$, $\bar{\gamma} \neq 0$, tal que $\bar{\gamma} \cdot f(u) \geq 0$, $\forall u \in \Omega$ se deduciría como corolario del Teorema que si $\hat{\bar{x}}$ resuelve el problema (17) podría garantizarse la existencia de un $\bar{\eta} \geq 0$ tal que se verifica (32).

La parte (a) del Teorema (24) podría haberse establecido también de la siguiente manera:

(36) COROLARIO

Si $\hat{\bar{x}}$ es una solución del problema 17 y C es cerrado, entonces:

(37) $$\hat{\bar{x}} = \begin{bmatrix} \bar{x}_o \\ \bar{x} \end{bmatrix} \in \overline{\{\hat{x}^* \in \mathbb{R}^{n+1} / \ \exists \hat{\eta}^* = (-1, \eta_1^*, \dots, \eta_n^*) = (-1, \eta^*), \text{ tal que :}}$$

$$\overline{\hat{\eta}^* \cdot \hat{x}^* = \max_{\hat{x} \in [C]} \hat{\eta}^* \cdot \hat{x} , \qquad \text{siendo C el conjunto (18)}}$$

$$\overline{\text{y} \quad \eta^* \cdot \geq 0 \ , \quad \eta^* \cdot \bar{x} = 0\}}$$

DEMOSTRACION

Se sigue inmediatamente de la demostración de la parte (a) del Teorema (24).

----

(38) EJEMPLO

El problema de la Programación Cuadrática.

Consideremos el siguiente problema:

(39) $$\min \sum_{j,k=1}^{m} d_{jk}\, x_j\, x_k + \sum_{j=1}^{m} c_j\, x_j$$

sujeto a

$$\sum_j a_{ij}\, x_j \geq b_i \quad ; \qquad i=1,2,\ldots.,n$$

$$x_j \geq 0 \quad ; \qquad j=1,2,\ldots.,m$$

en donde la matriz $(d_{jk})_{j,k=1,\ldots,m}$ es definida positiva.

Haciendo cambios de variable análogos a los del problema de Programación Lineal podemos volver a replantear el problema así:

$$(40) \qquad \min \; x_o = \sum_{j,k=1}^{m} d_{jk}\, u_j\, u_k + \sum_{j=1}^{m} c_j\, u_j$$

sujeto a:

$$x_i = \sum_j a_{ij}\, u_j - b_i \; ; \qquad i=1,2,\ldots.,n$$

$$u \in \Omega = \{u \in \mathbb{R}^m \,/\, u_j \geq 0; \qquad j=1,2,\ldots,m\}$$

$$x \in G = \{x \in \mathbb{R}^n \,/\, x_i \geq 0; \qquad i=1,2,\ldots,n\}$$

Como el conjunto C de este problema es cerrado y convexo podemos sustituir en la expresión (37) [C] por C y entonces el máximo:

$$\hat{\eta}^* \,.\, \hat{x}^* = \max_{\hat{x} \in [C]} \hat{\eta}^* \,.\, \hat{x}$$

con $\eta_o^* \neq 0$, puede sustituirse por el mínimo:

$$(41) \qquad \hat{\eta}^* \,.\, \hat{f}\,(u^*) = \min_{u \in \Omega} \, [\, f_o(u) + \eta^* f(u)]$$

con $\eta^* \geq 0$.

Puesto que la función a minimizar es convexa y derivable la solución u^* al problema (41) vendría dada por la solución al sistema de ecuaciones:

(42) $$- \sum_{k=1}^{m} d_{jk}\, u_k^* - c_j - \sum_{i=1}^{n} a_{ij}\, \eta_i^* = 0; \qquad j=1,2,\ldots,m$$

siempre que dicha solución fuera positiva.

Si denotamos:

(43) $$(\eta_i)_{i=1,\ldots,n} = \eta; \qquad (d_{jk})_{j,k=1,\ldots,m} = D$$

$$(a_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}} = A; \qquad (c_j)_{j=1,\ldots,m} = C$$

$$(u_j)_{j=1,\ldots,m} = u \quad \text{y} \quad (b_i)_{i=1,\ldots,n} = b$$

el sistema (40) se podría escribir así:

$$-D\, u^* - c + A'\, \eta^{*\prime} = 0$$

en donde A' indica la transpuesta de A.

Se tendrá entonces:

(44) $$u^* = D^{-1}\, [c - A'\, \eta^{*\prime}]$$

La solución (43) sería válida siempre que:

(45) $$D^{-1}\, [c - A'\, \eta^{*\prime}] \geq 0$$

pues en otro caso alguno de los u_j^* debería ser nulo y habría que repetir el proceso para los restantes.

Ahora bien, aunque el conjunto de soluciones del tipo (43) con la condición (44) es cerrado cuando η^* varia en el conjunto $\{\eta \in \mathbb{R}^n \;/\; \eta \geq 0\}$ no se seguiría necesariamente de ello que el conjunto encerrado entre llaves en la expresión (37) fuera cerrado.

Por lo tanto, en principio no habría garantias de que por el procedimiento descrito se pudieran determinar todos los posibles puntos $\hat{x}$ soluciones del problema.

Sin embargo en este caso particular de la Programación Cuadrática, así como también en el de la Programación Líneal se puede establecer un resultado mediante el cual se garantizaría que si $\hat{x}$ fuera un punto solución del correspondiente problema siempre - podría determinarse el vector $\bar{\eta}$ a que se refiere el Teorema (24), como vamos a ver a continuación.

(46) TEOREMA

Consideremos el caso particular del Problema (17) en el - que las f_i, $i=1,2,\ldots,n$, con líneales, y f_o es líneal o cuadrática, y $\Omega = \{u \in \mathbb{R}^n \ / \ u_j \geq 0, \ j=1,2,\ldots,m\}$. Entonces si $\hat{\bar{x}}$ es una solución del problema existe un vector $\bar{\eta} \geq 0$ tal que se verifican las desigualdades (25).

DEMOSTRACION

Supongamos pues que $\hat{\bar{x}}$ sea una solución del problema de los Teoremas (8) y (24) se seguiría entonces que o bien se verifica el resultado enunciado por el teorema o bien existe una sucesión $\hat{x}^{(r}$ de puntos del conjunto de accesibilidad dado por $x_o = f_o; x_i = f_i$, $i=1,2,\ldots,n$ con $\hat{x}^{(r} \to \hat{\bar{x}}$ y una sucesión de vectores $\hat{\eta}^{(r*} \leq 0$, con $\eta_o^{(r*} \neq 0$, de manera que:

$$\hat{\eta}^{(*}\hat{x}^{(r} \geq \hat{\eta}^{(r*}\hat{x}$$

para todo $\hat{x}$ perteneciente a dicho conjunto de accesibilidad, con $\eta^{(r*}\bar{x} = 0$ o equivalentemente, si suponemos $\eta_o^{(r*} = -1$ por simplificar sería:

$$f_o(u^{(r}) + \eta^{(r} f\ (u^{(r}) \leq f_o(u) + \eta^{(r} f(u)$$

para todo $u \geq o$, con $u^{(r} \to \bar{u}$, en donde $\hat{\bar{x}} = \hat{f}(\bar{u})$, y $\eta^{(r} \geq 0$, para todo $r \in \mathbb{N}$.

Supongamos ahora que, para cada r, se pudiera elegir para cada $i=1,2,\ldots,n$ un $u^{(i} \geq 0$ de manera que: $f_i(u^{(i}) < 0$, con $f(u^{(i}) \leq 0$.

Tendríamos entonces que:

$$f(\frac{1}{n}\sum_{i=1}^{n} u^{(i}) \leq \frac{1}{n}\sum_{i=1}^{n} f(u^{(i}) < 0$$

con $\frac{1}{n}\sum_{i=1}^{n} u^{(i} \geq 0$, para cada r, lo que, de acuerdo con la desigualdad anterior implicaría que la sucesión $\eta^{(r}$ sería acotada ya que:

$$- \eta^{(r} f(\frac{1}{n}\sum_{i=1}^{n} u^{(i}) \leq f_o(\frac{1}{n}\sum_{i=1}^{n} u^{(i}) - f_o(u^{(r}) - \eta^{(r} f(u^{(r})$$

para todo r, con $u^{(r} \to \bar{u}$.

Se podría entonces extraer una subsucesión convergente de la $\eta^{(r}$ a un determinado vector $\bar{\eta} \geq 0$, tal que se verificaría la primera de las desigualdades (25), y $\bar{\eta}\bar{x} = 0$, mientras que la segunda se seguiría inmediatamente de que $\eta \geq 0$ y $f(\bar{u}) \leq 0$.

En el caso de que para algún i*, entre 1 y n, no se pudiera encontrar el $u^{(i}$ correspondiente con la propiedad anterior entonces tendríamos que:

$$u \in \Omega;\ f(u) \leq 0 \quad \text{implica } f_{i*}(u) = 0$$

y en las restricciones del problema se podría sustituir $f_{i*}(u) \leq 0$ por $f_{i*}(u) = 0$.

Ahora bien, en nuestro caso al ser las f_i lineales tendríamos que los anterior implicaría que:

$$\sum_{j=1}^{m} a_{i*j}\, u_j = 0$$

y puesto que $a_{i*j*} \neq 0$ para algún j pues en otro caso la restricción podría haberse eliminado totalmente del planteamiento del problema desde el principio, tendríamos entonces que:

$$u_{j*} = \sum_{j \neq j*} (a_{i*j}/a_{i*j*})\, u_j$$

La anterior igualdad permitiría replantear el problema en términos de las variables u_j, $j \neq j*$, si se modifican las restricciones $f_i(u) \leq 0$, $i \neq i*$ $i=1,2,\ldots,n$, poniendo $g_i(u) = f_i(u) + \lambda_i f_{i*}(u)$, en donde $\lambda_i \geq 0$ y las $g_i(u)$ no dependerían de u_{j*}, se añadiera la restricción:

$- \sum_{j \neq j*} (a_{i*j}/a_{i*j*})\, u_j \leq 0$, y se replanteara la fun-- ción objetivo f_o después de sustituir u_{j*} en función de las restantes u_j en terminos de una nueva función $g_o(u)$ dependiendo solo de u_j, $j \neq j*$, la cual volvería a ser cuadrática o líneal se-- gún el caso.

El razonamiento podría repetirse entonces de la misma manera hasta que, o bien se hubieran agotado todas las variables u_j menos una, o bien en alguna etapa intermedia se hubiera podido demostrar la existencia del $\bar{\eta}$ buscando, el cual, aunque en un principio obtendría para las funciones $g_i(u)$, que no incluirían a de-- terminados u_j, teniendo en cuenta que las g_i, $i=1,2,\ldots,n$ se ha-- brían obtenido como combinaciones líneales con coeficientes positivos de determinadas $f_i(u)$, $i=1,2,\ldots,n$, del planteamiento original, y que $g_o(u)$ puede reformularse equivalentemente como la $f_o(u)$ en terminos de todas la u_j, se seguiría entonces que el $\bar{\eta}$ obtenido para las g_i sería ampliable a otro $\bar{\eta}^*$ definido para las f_i iniciales y cumpliendo las desigualdades (25).

Solo quedaría pues, para completar la demostración, el - probar que si se consumieran todas las etapas y el problema acabara planteándose en términos de una sola u_j, de manera que no pudieran encontrarse los $u^{(i}$ a que hemos hecho referencia antes sería posible, de todas formas, definir el $\bar{\eta}$ en cuestión.

En efecto, en el caso de que el problema quedara planteado en términos de una sola variable es inmediato que el conjunto de restricciones podría reducirse solo a dos dadas por: $u_j \leq b$ y: $u_j \geq a$, para un determinado par de constantes a y b.

Entonces, restringidos a solo las dos restricciones anteriores la inexistencia de puntos interiores implicaría necesariamente $a=b=u_j$ con $a \geq 0$, que sería la solución del problema.

Si estuvieramos en el caso lineal y fuera $c_j^* u_j$ la función objetivo en términos de u_j con $c_j^* \geq 0$, tendríamos entonces:

$$c_j^* a \leq c_j^* u_j + c_j^* (-u_j + a)$$

y podríamos tomar $\bar{\eta}_1 = c_j^*$ para la restricción en cuestión, que supondremos que es la primera, completando con ceros para las restantes restricciones mientras que en caso contrario se podría tomar $\bar{\eta}_1 = -c_j^*$.

En el caso cuadrático, y si la función objetivo fuera: $d_j^* u_j^2 + c_j^* u_j$ y suponemos que $c_j^* \geq 0$, podríamos poner:

$$d_j^* a^2 + c_j^* a \leq d_j^* u_j^2 + c_j^* u_j + (c_j^* + 2d_j^* a)(-u_j + a)$$

ya que necesariamente, $d_j^* \geq 0$, y tomar $\bar{\eta}_1 = c_j^* + 2d_j^* a$, y $\bar{\eta}_1 = -c_j^*$; $\bar{\eta}_2 = 2d_j^* a$, en el caso en que $c_j^* < 0$, en donde consideramos asociado el subíndice 1 a la restricción $u_j - a \leq 0$, y el 2 a la $-u_j + a \leq 0$, lo que completa la demostración.

Otro resultado correspondiente a una propiedad del Problema (17) será de utilidad más adelante. Este es el siguiente:

(47) TEOREMA

Consideremos el Problema (17). Entonces, si el conjunto:

$$(48) \qquad \{\hat{x}^* \in \mathbb{R}^{n+1} \,/\, \hat{\eta}^* = (-1, \eta_1^*, \ldots, \eta_n^*) = (-1, \eta^*) \text{ tal que}$$

$$\eta^* \geq 0 \text{ y: } \hat{\eta}^* . \hat{x}^* = \max_{\hat{x} \in C} \hat{\eta}^* . \hat{x}, \text{ siendo } C \text{ el conjunto (18)}\}$$

es no vacío el Problema (17) no puede tener solución infinita.

DEMOSTRACION

Supongamos que hubiera solución infinita y sea $\hat{x}^{(r}$ una sucesión de puntos de C tal que $x_i^{(r} \geq 0$, $i=1,2,\ldots,n$, y $x^{(r} \to -\infty$.

Si existe un $\hat{x}^*$ perteneciente al conjunto (48) entonces debería ser:

$$\hat{\eta}^* \cdot \hat{x}^* \geq \hat{\eta}^* \cdot \hat{x} \tag{49}$$

para algún $\hat{\eta}^*$ y para todo $\hat{x} \in C$.

Ahora, para los $\hat{x}^{(r}$ tendríamos:

$$\hat{\eta}^* \hat{x}^* \geq - x_o^{(r} + \eta^* x^{(r} \geq - x_o^{(r}$$

lo cual es imposible ya que $-x_o^{(r} \to \infty$.

Estamos ya en condiciones de probar el Teorema de la Dualidad en Programación Lineal que tiene importancia tanto desde el punto de vista teórico como algorítmico, y el cual, a pesar de ser sobradamente conocido, presenta interés también desde el punto de vista de nuestra exposición, al ser deducido dentro de la teoría que venimos desarrollando.

(50) TEOREMA (de la Dualidad en Programación Lineal)

Consideremos los dos problemas de Programación Lineal siguientes:

$$\min f_o(u) = \sum_{j=1}^{m} c_j u_j \tag{51}$$

sujeto a:

$$-x_i = \sum_{j=1}^{m} a_{ij} u_j - b_i \leq 0; \qquad i=1,2,\ldots,n$$

$$u \in \Omega = \{u \in R^n \ / \ u_j \geq 0; \qquad j=1,2,\ldots,m\}$$

$$x \in G = \{x \in R^n \ / \ x_i \geq 0; \qquad i=1,2,\ldots,n\}$$

y

$$\min \ g_o(\eta) = \sum_{i=1}^{n} b_i \eta_i \tag{52}$$

sujeto a:

$$- y_j = - \sum_{i=1}^{n} a_{ij} \eta_i - c_j \leq 0; \qquad j=1,2,\ldots,m$$

$\eta \in \Omega^* = \{\eta \in \mathbb{R}^n \ / \ \eta_i \geq 0; \qquad i=1,2,\ldots,n\}$

$y \in G^* = \{y \in \mathbb{R}^m \ / \ y_j \geq 0; \qquad j=1,2,\ldots,m\}$

Entonces:

(a) Si uno de los dos problemas tiene solución finita también la tiene el otro y ambas funciones objetivo coinciden, salvo el signo, en el óptimo.

(b) Si uno de los dos problemas tiene solución infinita el otro no tiene solución.

DEMOSTRACION

(a) Supongamos que el problema (51) tiene la solución $\hat{x} = \begin{bmatrix} f_o(\bar{u}) \\ -f(\bar{u}) \end{bmatrix}$. Para cada $u \geq 0$ y cada $\eta \geq 0$ que verifiquen las restricciones de ambos problemas se verifica:

$$f_o(u) = \sum_{j=1}^{m} c_j u_j \geq \sum_{j=1} \left(- \sum_{i=1}^{n} a_{ij} \eta_i\right) u_j \geq - \sum_{i=1}^{n} b_i \eta_i =$$

$$= -g_o(\eta)$$

La anterior desigualdad demuestra que:

$$f_o(u) \geq -g_o(\eta)$$

(53)

$$y : g_o(\eta) \geq -f_o(u)$$

para cualquier pareja (u,η) que verifique las restricciones. En particular y en nuestra hipótesis de que $\hat{x}$ es solución de (51) se obtiene:

$$g_o(\eta) \geq -f_o(\bar{u})$$

de lo que se deduce que g_o está acotada.

Por otra parte del Teorema (46) se sigue que existe un $\bar{\eta} \geq 0$ tal que:

$$f_o(\bar{u}) \leq \sum_{j=1}^{m} c_j u_j + \sum_{i=1}^{n} \bar{\eta}_i \left(\sum_{j=1}^{m} a_{ij} u_j - b_i \right); \qquad \forall u \in \Omega$$

Obviamente el $\overline{\eta}$ anterior debe cumplir el conjunto de desigualdades del problema (52) pues si no el mínimo en Ω de la expresión del segundo miembro será $-\infty$ contra lo supuesto. Dando por verificado esto y determinando el mínimo se obtiene:

$$f_o(\overline{u}) \leq -\sum_{i=1}^{n} \overline{\eta}_i \cdot b_i = -g_o(\overline{\eta})$$

Combinando este resultado con las desigualdades (53) se obtiene que $f_o(\overline{u}) = -g_o(\overline{\eta})$ y que $\overline{\eta}$ es solución óptima de (52).

La demostración si se supone inicialmente que es el problema (52) el que tiene solución es similar.

(b) De las dos desigualdades (53) se sigue inmediatamente que si uno de los dos problemas tiene solución el otro no puede tener solución infinita.

Vamos a finalizar el estudio del caso estático convexo generalizando la aplicación del Principio de la Proyección Maximal a conjuntos C convexos cualesquiera, sean o no cerrados. El siguiente teorema se ocupa de ello.

(54) TEOREMA (Principio de Máximo en el caso estático convexo)

Si $\hat{\hat{x}}$ es solución del Problema (6) en el caso de que el conjunto C dado por (5) sea convexo, entonces $\hat{\hat{x}}$ pertenece al conjunto:

(55) $\overline{\{\hat{\overline{x}} \in C \;/\; \exists\ \hat{\eta} = (\eta_o, \eta_1, \ldots, \eta_n) \text{ con } \eta_o < 0, \text{ tal que}}$

$\overline{\hat{\eta} \cdot (\hat{x} - \hat{\overline{x}}) \leq 0, \;\; \forall \hat{x} \in C\}}$

DEMOSTRACION

Basta comprobar que $\hat{x}$ pertenece a la frontera del conjunto $\overline{C}$ porque luego se aplicaría el Teorema (8) lo que conduciría a (55). Pero el comprobar que $\hat{x}$ pertenece a la frontera de $\overline{C}$ es inmediato porque en caso contrario sería un punto interior de $\overline{C}$ y por lo tanto de C ya que ambos son convexos y un conjunto convexo y su clausura tienen los mismos puntos interiores. Por lo tanto $\hat{x}$ no podría ser punto solución de (6) en contra de lo supuesto.

Algunas aplicaciones de este Teorema en el caso dinámico convexo se verán en el apartado siguiente.

3. EL CASO DINAMICO CONTINUO CONVEXO

El tipo de problema que vamos a estudiar en este apartado es el planteado de la siguiente manera:

$$\text{(56)} \qquad \min \; x_o(t_1) = \int_{t_o}^{t_1} f_o(x,u,t)dt$$

sujeto a:

$$x(t) = \int_{t_o}^{t_1} f(x,u,s)ds + x(t_o); \qquad \forall t \in [t_o,t_1]$$

con $x(t) \in \mathbb{R}^n$; $u(t) \in \Omega \subset \mathbb{R}^m$, $\forall t \in [t_o,t_1]$, y $u \in \Delta$, siendo Δ una clase dada de funciones.

Además supondremos las siguientes condiciones de contorno:

$$x(t_i) \in G_i \subset \mathbb{R}^n; \qquad i=0,1$$

En el anterior planteamiento t_o y t_1 son, en general, dos parámetros a determinar según convenga a la minimización de x_o, pero también puede incluirse el caso en que uno de los dos o ambos sean dados.

Las condiciones iniciales que suelen solicitar a las funciones f_i, $i=0,1,\dots,n$ son las de que sean continuas y tengan primeras derivadas continuas con respecto a las x_i, $i=0,1,\dots,n$, en el espacio $\mathbb{R}^{n+1} \times \bar{\Omega} \times [t_o,t_1]$, en donde $\bar{\Omega}$ es la clausura de Ω en $\mathbb{R}^m$.

Prescindiremos, sin embargo, por el momento de imponer condiciones a las funciones f_i ó a Δ explícitamente y unicamente haremos la suposición tácita de que todas las operaciones que efectuemos con dichas funciones f_i y u son admisibles.

Así, por lo que se refiere a las integraciones podemos suponer que se entienden en el sentido de Lebesgue mientras que por lo que se refiere a la ecuación integral vectorial de (56) supondremos que f y Δ cumplen condiciones que garantizan la existencia y unicidad de la solución $x(t)$ en $[t_o, t_1]$ para $x(t_o)$ dado. La restricción de finitud de t_o y t_1 podría suprimirse en el caso de que alguno de ellos o ambos se fijaran como $-\infty$ ó $+\infty$ en cuyo caso se supondría que las funciones implicadas en el problema cumplen condiciones que garantizan el sentido y univocidad de las operaciones efectuadas. En el caso en que haga falta poner en evidencia las hipótesis que se están manejando para precisar el alcance de los teoremas o afirmaciones que se hagan éstas se enunciarán explícitamente.

Un repaso al apartado anterior muestra que se puede aplicar el Principio (8) a los casos en que el conjunto C de las variables x_i, al cual está restringido el problema, es convexo, sea o no cerrado.

Supuesta garantizada la convexidad la mayor dificultad, desde el punto de vista operativo, para obtener resultados a partir de la aplicación del Principio dependería de la posibilidad de lograr expresar de manera explícita las variables x_i en función de las u_j.

En cualquier caso el enfoque más conveniente para hacer más provechosa la aplicación del Principio en este caso dinámico es el de considerar a las x_i como variables de estado y a las u_j, t_o y t_1, como variables de control. La diferencia prin cipal con los problemas estudiados en la sección anterior es la de que las u_j son ahora un conjunto de "funciones de control" y no simplemente de variables numéricas. Sin embargo podemos encajar este nuevo problema dentro de los ya estudiados si nos limitamos a investigar las propiedades que debería cumplir el punto terminal de la trayectoria óptima para diferenciarlo de los demás puntos terminales de trayectorias no óptimas. Con esta perspectiva consideraremos el conjunto C siguiente:

$$(57) \qquad C = \{\hat{x} \in \mathbb{R}^{n+1} \,/\, \hat{x} = \int_{t_o}^{t_1} \hat{f}(x(t),u(t),t)dt + \begin{bmatrix} 0 \\ (t_o) \end{bmatrix} \quad \text{en}$$

$$\text{donde} \quad x(t) = \int_{t_o}^{t_1} f(x(s),u(s),s)ds; \quad u(t) \in \Omega,$$

$$\forall t \in [t_o,t_1], \quad u \in \Delta, \quad t_o,t_1 \in \mathbb{R}, \quad t_1 \geq t_o,$$

$$x(t_o) \in G_o\}$$

mientras que el conjunto G, de los problemas de la sección anterior, pasaría ahora a ser el G_1 del Problema (56). En el marco de este problema a la proyección del conjunto C en el espacio de sus n últimas coordenadas se la denomina el "conjunto de accesibilidad".

Con este enfoque, pues, y repasando el enunciado y demostración del Teorema (54) se puede enunciar el siguiente Teorema relativo al Problema (56):

(58) TEOREMA (Principio de Máximo para puntos terminales en el caso dinámico convexo)

Consideremos el conjunto C definido por (57) y supongamos que es convexo. Sea $\hat{\hat{x}}$ un punto solución del Problema (56). Entonces $\hat{\hat{x}}$ pertenece al conjunto (55) con C sustituido por el dado por (57) ó, equivalentemente, existe un vector $\hat{\eta} \neq 0$, con $\bar{\eta}_o \leq 0$ tal que:

$$\hat{\eta} \cdot (\hat{x} - \hat{\bar{x}}) \leq 0; \qquad \forall \hat{x} \in C$$

DEMOSTRACION

Se sigue del Principio (8) y del Teorema (54).

La desigualdad $\hat{\eta} \cdot (\hat{x} - \hat{\bar{x}}) \leq 0, \quad \forall \hat{x} \in C$, puede enunciarse también así:

$$(59) \qquad \hat{\eta} \cdot \hat{\bar{x}} = \max \sum_{i=o}^{n} \bar{\eta}_i \int_{t_o}^{t_1} f_i(x(t),u(t),t)dt + \sum_{i=1}^{n} \bar{\eta}_i \; x_i(t_o)$$

con las restricciones:

$$x(t) = \int_{t_o}^{t} f(x(s),u(s),s)ds; \qquad \forall t \in [t_o,t_1]$$

$$u(t) \in \Omega \qquad ; \qquad \forall t \in [t_o,t_1]$$

$$u \in \Delta; \quad t_o,t_1 \in \mathbb{R}; \quad t_1 \geq t_o; \quad x(t_o) \in G_o$$

si se pudiera despejar $x(t)$ en la expresión:

$$x(t) = \int_{t_o}^{t} f(x(s),u(s),s)ds + \begin{bmatrix} 0 \\ x(t_o) \end{bmatrix}; \qquad t \in [t_o,t_1],$$

es decir si se pudiera solucionar explícitamente la ecuación integral vectorial, se podría sustituir este resultado en la expresión a maximizar que ya dependería sólo, para cada $\hat{\eta}$, de $u(t)$, t_o y t_1.

La primera diferencia que se observa entre (59) y el resultado clásico del Principio de Máximo de Pontryaguin es la de que en el segundo caso el máximo se toma con respecto al producto escalar $\hat{\eta} \cdot \frac{d}{dt}(\hat{x})$ y no con respecto a $\hat{\eta} \cdot \hat{x}$. Más adelante, cuando investiguemos el caso diferenciable mostraremos como se reformularía (59) en términos de $\frac{d}{dt}(\hat{x})$ en vez de $\hat{x}$, siendo $\hat{\eta}$ en esta nueva situación una función de t a determinar para todo $t \in [t_o, t_1]$. Esta diferencia entre el distinto papel de $\hat{\eta}$ en ambos Principios, un vector de escalares, en el caso de la expresión (59), y una función vectorial, en el caso diferenciable marca la distinción fundamental entre ambos casos.

Finalmente, otra diferencia a tener en cuenta es la de que en problema de maximización (59) una de las restricciones es la de que $u \in \Delta$, la cual no se explicita en el problema de máximo puntual que habría que resolver para establecer el Principio de Pontryaguin, aunque sería una de las hipótesis a asumir para llegar a formular dicho problema. El que la clase de funciones Δ intervenga explícitamente en el problema de maximización depende, en general, de que se exijan propiedades adicionales a Δ como muestra el Teorema siguiente:

(60) TEOREMA

Consideremos el problema de maximización (59) en el que la clase no vacía de funciones Δ tiene la propiedad de que, dados t_o y t_1, si $u(t)$, $t \in [t_o, t_1]$, es un control admisible y t' y t'' son tales que: $t_o \leq t' \leq t'' \leq t_1$, el control $u^{\#}(t)$ definido de la siguiente manera:

$$u^{\#}(t) = \begin{cases} u(t) & \text{si} \quad t_o \leq t \leq t' \\ \omega & \text{si} \quad t' < t \leq t'' \quad \text{(supuesto } t' < t'' \text{, en otro caso no se áplica)} \\ u(t) & \text{si} \quad t'' < t \leq t_1 \quad \text{(supuesto } t'' < t_1 \text{, pues en otro caso no se aplicaría esta parte)} \end{cases}$$

en donde ω es un punto arbitrario de Ω, es también admisible. Supongamos también que las funciones $f_i(x,u,t)$ son continuas y con derivadas parciales continuas respecto a x y t en $\mathbb{R}^n \times \bar{\Omega} \times [t_o,t_1]$, para $i=0,1,\dots,n$ y que se verifica la siguiente propiedad de continuidad adicional:

(61) Si $x^*(t)$ y $u^*(t)$ son sendas funciones tales que:

$$x^*(t) = \int_{t_o}^{t} f(x^*(s),u^*(s),s)ds + x^*(t_o), \ \forall t \in [t_o,t_1]$$

y $u^{(k}(t)$, $k \in \mathbb{N}$, es una sucesión de funciones de Δ de manera que

$$f(x^*(t),u^{(k}(t),t) \longrightarrow f(x^*(t),u^*(t),t),$$ casi seguramente en $[t_o,t_1]$, se verifica

$$\int_{t_o}^{t_1} f(x^{(k}(t),u^{(k}(t),t)dt \longrightarrow \int_{t_o}^{t_1} f(x^*(t),u^*(t),t)dt$$

en donde

$$x^{(k}(t) = \int_{t_o}^{t_1} f(x^{(k}(t),u^{(k}(t),t)dt + x^*(t_o); \quad \forall t \in [t_o,t_1]$$

Entonces si existe una función $\bar{u} \in \Delta$ que resuelve el Problema (59) para algún $\hat{\eta}$ y unos t_o y t_1, esta misma función $\bar{u}(t)$ resuelve el Problema (59) para dichos t_o y t_1, supuesto suprimido de éste la condición de que $u \in \Delta$.

DEMOSTRACION

Es fácil comprobar que si la clase Δ tiene la propiedad del enunciado cualquier función de control constante a trozos (es decir tal que exista una partición finita de $[t_o,t_1]$ en subíntervalos: $[t^{(j},t^{(j+1}]$, $j=0,1,\dots,(d-1)$ con $t^{(o}=t_o$, $t^{(d}=t_1$, y de manera que $u(t)=\omega_j$ si $t^{(j}<t\leq t^{(j+1}$, y $u(t_o)=\omega_o$, siendo $\omega_j\in\Omega$, $\forall j$) es también admisible.

Supongamos ahora que $\bar{u}(t)$ no resuelve para dichos t_o y t_1 fijos el Problema (59) si se suprime la condición de que $u\in\Delta$. Esto implicaría que $t_o<t_1$ y que existiría una función $u^*(t)$ no perteneciente a la clase Δ tal que las funciones $f_i^*(t)=f_i(x^*(t),u^*(t),t)$, para $t_o\leq t\leq t_1$, con $i\in\{j\in\{0,1,\dots,n\}\ /\ \bar{\eta}_j\neq 0\}$ serían integrables en el sentido de Lebesgue y:

$$(62)\qquad \sum_{i=o}^{n}\bar{\eta}_i\int_{t_o}^{t_1}f_i(x^*(t),u^*(t),t)dt+\sum_{i=1}^{n}\bar{\eta}_i\,x_i^*(t_o)>$$
$$>\sum_{i=o}^{n}\bar{\eta}_i\int_{t_o}^{t_1}f_i(\bar{x}(t),\bar{u}(t),t)dt+\sum_{i=1}^{n}\bar{\eta}_i\,\bar{x}(t_o)$$

Ahora, por ser las $f_i^*(t)$ medibles, para cada $k=2^p$, $p\in\mathbb{N}$ existen conjuntos $E_i^{(r}$, $r\in\mathbb{N}$, correspondientes a cada f_i^* de manera que dichas f_i^* son contínuas en sus complementarios en $[t_o,t_1]$ y $\mu(E_i^{(r})<\frac{1}{r}$, $\forall i$, siendo μ la medida de Lebesgue (Véase la Nota 2).

Prescindamos del conjunto de puntos aislados de cada $E_i^{(r}$, que forman un conjunto de medida nula y consideremos sucesivas particiones del conjunto $E^{(r}=\bigcap_i E_i^{(r}$, correspondientes a las sucesiones de puntos:

(63) $$t_o \leq t^{(k}_{1,r} \leq \cdots \leq t^{(k}_{k,r} \leq t_1$$

en donde,

$$t^{(k}_{s,r} \in E^{(r}, \quad \text{para} \quad s=1,2,\ldots,(k-1); \quad k = 2^p, \ p \in \mathbb{N}$$

de manera que la partición (k+1)-ésima se formaría a partir de la k-ésima intercalando puntos de $E^{(k}$ en cada uno de los sub intervalos $[t^{(k}_{s-1,r}, t^{(k}_{s,r}]$, siempre que sea posible, con la propiedad de que si $t^{(k+1}_{s'} \in (t^{(k}_{s-1}, t^{(k}_{s})$ entonces $t^{(k+1}_{s'}$ pertenecería al intervalo:

$$\left(t^{(k}_{s-1} + \frac{t^{(k}_{s} - t^{(k}_{s-1}}{q^*}, \ t^{(k}_{s} - \frac{t^{(k}_{s} - t^{(k}_{s-1}}{q^*}\right)$$

en donde $q^* = \min \{q \in \mathbb{N} \ / \ q \geq 3$ y

$$\left(t^{(k}_{s-1} + \frac{t^{(k}_{s} - t^{(k}_{s-1}}{q}, \ t^{(k}_{s} - \frac{t^{(k}_{s} - t^{(k}_{s-1}}{q}\right) \cap E^{(k+1} \neq \emptyset\}$$

Si $k \to \infty$, se obtiene así un conjunto de puntos:

$$\{t \in [t_o, t_1] \ / \ t = t^{(k}_{s,r}, \quad \text{para algunos } s \text{ y } k,$$
$$t^{(k}_{s,r} \in E^{(r}, \ \forall s\}$$

que es denso en $E^{(r}$, como es fácil comprobar.

Definamos ahora para la partición correspondiente de $E^{(r}$ la función de control $u^{(k}(t)$ en $[t_o, t_1]$ siguiente:

(64)
$$u^{(k}_{r}(t) = \begin{cases} u^*(t^{(k}_{s,r}), & \text{si} \quad t \in (t^{(k}_{s-1,r}, t^{(k}_{s,r}]; \quad s=2,3,\ldots,k \\ u^*(t^{(k}_{1,r}), & \text{si} \quad t \in [t_o, t^{(k}_{1,r}] \\ u^*(t_1), & \text{si} \quad t \in (t^{(k}_{k,r}, t_1] \end{cases}$$

Cada una de las funciones de la sucesión anterior será, por tanto, medible-Lebesgue y pertenecerá a la clase Δ.

Consideremos a continuación la sucesión de funciones:

$$f^{(k}(t) = f(x^{(k}(t), u_r^{(k}(t), t), \qquad \forall t \in [t_o, t_1]$$

en donde:

$$x^{(k}(t) = \int_{t_o}^{t} f(x^{(k}(s), u^{(k)}(s), s)ds + x^*(t_o)$$

las cuales están bien definidas, al ser $u^{(k} \in \Delta, \forall k$ y por las condiciones impuestas a f.

Ahora, para cada $t \in E^{(r}$ tenemos $u^{(k}(t) = u^*(t^{(k}_{s_t,r})$ para todo k, con $t^{(k}_{s_t} \longrightarrow t$ y $t^{(k}_{s_t,r} \in E^{(r}$.

Por otra parte de la continuidad de las funciones $f(x,u,t)$ y $x^*(t)$ se deduce que para cada $\varepsilon \downarrow 0$ y k_o suficientemente grande:

$$(65) \qquad |f(x^*(t), u^*(t^{(k}_{s_t,r}), t) - f(x^*(t^{(k}_{s_t,r}), u^*(t^{(k}_{s_t,r}), t^{(k}_{s_t,r})| \leq$$

$$\leq \varepsilon \qquad \forall k \geq k_o$$

Como $f(x^*(t), u_r^{(k}(t), t) = f(x^*(t), u^*(t^{(k}_{s_t,r}), t)$, y $f^*(t)$ es continua en $E^{(r}$ tendremos que

$$f^*(t^{(k}_{s_t,r}) \longrightarrow f^*(t)$$

lo cual, junto con (65), implica que

$$f(x^*(t), u_r^{(k}(t), t) \longrightarrow f^*(t), \quad \forall t \in E^{(r}$$

Entonces podemos concluir por lo tanto que:

$$f(x^*(t), u_{2p}^{(2p}(t), t) \longrightarrow f^*(t), \quad \forall t \in \liminf_r E^{(r}$$

Ahora: $\mu([t_o,t_1] - \liminf_r E^{(r)}) = 0$, ya que todo conjunto de medida positiva tiene intersección finita con todos los $E^{(r}$ salvo un número finito.

Si tenemos en cuenta entonces (61), es fácil comprobar que lo anterior conduce a una contradicción con (62), lo que completa la demostración.

En la demostración del Teorema (60) podría haberse sustituido la clase de funciones aproximantes (continuas a trozos) por cualquier otra que también fuera aproximante y la conclusión del Teorema habría permanecido invariariable.Así, por ejemplo, podría haberse considerado Δ como la clase de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}^m$ siempre y cuando existiera la función $\bar{u}(t) \in \Delta$ del enunciado del Teorema que resolviera el Problema (59) y se verificara (61).

Otra observación al Teorema (60) es la de que éste podría extenderse al caso en que $t_1 = +\infty$ ó $t_o = -\infty$ ó ambos sin más que razonar por contradicción el resultado para un intervalo compacto lo súficientemente amplio y supuesto que se verificara (61) para todo intervalo compacto.

Por otra parte, para poder aplicar con eficacia el Teorema (58) a varios ejemplos que estudiaremos a continuación conviene proponer una versión mas operativa del Teorema como la que corresponde al siguiente corolario:

(66) COROLARIO

Denotemos por C_{t_o,t_1} la particularización del conjunto C dado por (57) a t_o y t_1 fijos y supongamos que cada uno de los conjuntos C_{t_o,t_1} con $T_o \leq t_o \leq t_1 \leq T_1$, en donde T_o y T_1 son dados, es convexo.

Entonces, si para determinados t_o y t_1, con $T_o \leq t_o \leq t_1 \leq T_1$ el punto $\hat{\bar{x}} \in C_{t_o,t_1}$ minimiza

$x_o(t_1) = \int_{t_o}^{t_1} f_o(x(t),u(t),t)dt$, con la restricción adicional de que $\hat{\bar{x}} \in G_1$, se sigue de ello que $\hat{\bar{x}}$ pertenecerá al conjunto (55) con C_{t_o,t_1} en lugar de C ó, equivalentemente, existirá un vector $\hat{\bar{\eta}} \neq 0$, con $\bar{\eta}_o \leq 0$, tal que:

$$\hat{\bar{\eta}} \cdot (\hat{x} - \hat{\bar{x}}) \leq 0, \qquad \forall \hat{x} \in C_{t_o,t_1}$$

DEMOSTRACION

Se deduce inmediatamente el resultado sin más que aplicar el Teorema (58) a la nueva situación.

Vamos a estudiar a continuación, como aplicación de los resultados anteriores el caso de mayores aplicaciones dentro de la teoría que es el de dinámica lineal.

(67) EJEMPLO (Problemas convexos con dinámica lineal)

Consideremos el problema de control caracterizado por funciones del siguiente tipo:

(68) $\qquad f(x(t),u(t),t) = A(t)\ x(t) + B(t)\ u(t) + v(t)$

en donde $A(t)$ y $B(t)$ son matrices $n \times n$ y $n \times m$ respectivamente y $v(t)$ es un vector $n \times 1$, de manera que las tres funciones multidimensionales son medibles en $\mathbb{R}$.

Supondremos también que las normas: $\|A(t)\|$, $\|B(t)\|$ y $\|v(t)\|$, definidas de la siguiente manera:

$||A(t)|| = [\sum_{ij} a_{ij}^2(t)]^{1/2}$ etc..., con integrables en cada subintervalo compacto de $\mathbb{R}$ y que Δ es la clase de funciones medibles y acotadas con valores en $\mathbb{R}^m$, con dominios en algún subintervalo $[T_o, T_1]$ de $\mathbb{R}$, incluyendo el caso en que $T_1 = +\infty$ para el que el subintervalo sería entonces $[T_o, +\infty)$, y tomando valores en un conjunto no vacío $\Omega \subset \mathbb{R}^m$.

Las hipótesis anteriores garantizan (Véase la Nota 3) la validez de la denominada "formula de variación de los parámetros" para $x(t)$, es decir, para t_o, t y t_1 tales que $T_o \leq t_o \leq t \leq t_1 \leq T_1$, tendríamos:

$$(69) \qquad x(t) = \Phi(t)x(t_o) + \Phi(t)\int_{t_o}^{t} \Phi^{-1}(s)[B(s)u(s) + v(s)]\,ds \ ; \quad \forall t \in [t_o, t_1]$$

en donde $\Phi(t)$ es la matriz fundamental, solución del sistema homogéneo:

$$(70) \qquad \frac{d}{dt}\, x(t) = A(t)\, x(t)$$

con $\Phi(t_o) = I$ (matriz unidad).

En el caso particular de que A sea constante, entonces

$$(71) \qquad \Phi(t) = e^{A(t-t_o)}$$

en donde $\quad e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \ldots\ldots..$

La aplicación del Principio de Máximo para los puntos terminales a este problema se apoya en el siguiente Teorema:

(72) TEOREMA

Supongamos dado el Problema (56) con la particularización (68) y las hipótesis sobre $A(t)$, $B(t)$, $V(t)$ y Δ enunciadas en los párrafos anteriores para el intervalo $[T_o,T_1]$. Supongamos además que el conjunto Ω es compacto y que $G_o = \{x^{(o)}\}$

Entonces, el conjunto de accesibilidad del problema para t_o y t_1 con $T_o \leq t_o \leq t_1 \leq T_1$, es decir la proyección del conjunto C_{t_o,t_1} dado por la particularización de (57) a t_o y t_1 en el espacio de sus n últimas coordenadas es compacto y convexo.

DEMOSTRACION

Véase la Nota 3 bis.

Aunque el resultado del Teorema (72) incluye la compacidad hay que observar que para la aplicación del Corolario (66) basta solo con la comprobación de la convexidad (Nótese que las hipótesis del Corolario (66) se refieren a los conjuntos C_{t_o,t_1} y no a los de accesibilidad). Estudiaremos a continuación un caso particular importante de los problemas de control con dinámica lineal como es el de los problemas de tiempo óptimo, es decir aquellos para los que $f_o(x,u,t) \equiv 1$, $\forall \in [t_o,t_1]$. En este caso la aplicación del Corolario (66) es casi inmediata a partir del Teorema (72), como muestra el siguiente Teorema:

(73) TEOREMA (Problemas lineales de tiempo óptimo)

Supongamos dadas las hipótesis del enunciado del Teorema (72) y además la de que $f_o(x(t),u(t),t) \equiv 1$, $\forall t \in [t_o,t_1]$. Em tonces C_{t_o,t_1} es compacto y convexo para todo par t_o,t_1 tal que $T_o \leq t_o \leq t_1 \leq T_1$.

DEMOSTRACION

La comprobación de que C_{t_o,t_1} es compacto y convexo se deduce del Teorema (72) sin mas que ampliar el número de variables con la dada por:

$$x_o(t) = \int_{t_o}^{t_1} ds = t_1 - t_o$$

que cumple todas las condiciones requeridas.

Como consecuencia del Teorema (73) se sigue la aplicabilidad del Corolario (66) al problema de control de tiempo óptimo con dinámica lineal. En efecto, suponiendo que $t_o = T_o$ y T_1 (incluyendo el caso $T_1 = +\infty$) son dados, que Ω es compacto y que $G_o = \{x^{(o}\}$. Entonces, de acuerdo con el Corolario (66) y en lo que se refiere al Ejemplo (67) para problemas de tiempo óptimo habrá, pues, que investigar los máximos de la función:

$$(74) \qquad -\int_{t_o}^{t_1} ds + \eta \cdot [\Phi(t_1)x^{(o} + \Phi(t_1) \int_{t_o}^{t_1} \Phi^{-1}(t)\{B(t)u(t)+v(t)\}dt]$$

para cada $t_1 \in [t_o,T_1]$ y cada $\eta = (\eta_1,\eta_2,\ldots,\eta_n)$, supuesto que $u(s) \in \Omega$ y $u \in \Delta$.

Puesto que la variable es $u(t)$, para cada t_1 fijo el máximo de (74) se reduce al máximo de la función:

$$(75) \qquad \eta\, \Phi(t_1) \int_{t_o}^{t_1} \Phi^{-1}(t)\ B(t)u(t)dt$$

Así, ya que (75) es lineal en $u(t)$ y Δ admite a las funciones constantes a trozos como admisibles es evidente que para cada η el máximo de (75) se obtiene, en general, para funciones constantes a trozos cuyos valores corresponderían a los puntos extremos de Ω (entendidos estos como los puntos extremos de $[\Omega]$) que maximizarían, para cada $t : t_o \leq t \leq t_1$, la función lineal:

$$\eta \cdot \Phi(t_1) \Phi^{-1}(t) B(t) u(t)$$

En particular si:

$$\Omega = \{u \in \mathbb{R}^m / |u_j| \leq 1, \quad j=1,2,\ldots,m\}$$

entonces:

(76) $$u(t) = \text{sgn} \{\eta \cdot \Phi(t_1) \Phi^{-1}(t) B(t)\}, \quad \forall t \in [t_o, t_1]$$

en donde "sgn" denota la función "signo" del vector correspondiente la cual se define como el vector de funciones signo de cada una de sus componentes, siendo el signo de un escalar igual a +1, 0 ó -1 según que el escalar sea positivo, nulo o negativo, respectivamente.

No discutiremos aqui la optimalidad de alguna de las funciones de control propuestas en el párrafo anterior en las hipótesis asumidas, la cual se investiga extensamente en la mayoria de los textos de teoría del control óptimo. Solo comprobaremos que (76) concuerda con el resultado que se obtiene de la aplicación del Principio de Maximo de Pontryaguin a este caso. En efecto, basta definir:

(77) $$\eta(t) = \eta \cdot \Phi(t_1) \Phi^{-1}(t), \quad \text{para} \quad t_o \leq t \leq t_1$$

Esta función vectorial verifica evidentemente el sistema de ecuaciones diferenciales siguiente:

(78) $$\frac{d\,\eta(t)}{dt} = -A(t)\,.\,\eta(t), \qquad \text{para} \quad t_o \leq t \leq t_1$$

con la condición de contorno $\eta(t_1) = \eta$.

El sistema lineal (78) es el denominado "sistema adjunto" del problema de control de dinámico lineal y tiempo óptimo.

Se podrían deducir, pues, a partir del Principio de Máximo para puntos terminales todos los teoremas que se obtienen como corolarios de la aplicación del Principio de Pontryaguin a los problemas lineales de tiempo óptimo y en particular los teoremas concernientes al número de conmutaciones del control óptimo cuando $A(t)$ y $B(t)$ son constantes.

Vamos a terminar esta sección mostrando como el Principio de máximo para puntos terminales en el caso convexo implica alguno de los resultados que se derivan del Principio de Máximo de Pontryaguin.

(79) TEOREMA

Consideremos el Problema (56) con $\hat{f}(x,u,t) = \hat{f}(x,u)$; $G_o=\{x^{(o}\}$; t_o y t_1 variables y Δ una clase de funciones tal que (Véase la Nota 4):

a) Todos los controles de Δ son medibles y acotados.

(80) b) Si $u(t)$, $t_o \leq t \leq t_1$, es un control admisible, ω es un punto arbitrario de Ω y t' y t'' son números

tales que $t_o \leq t' \leq t'' \leq t_1$, el control $u^{\#}(t)$ definido para la fórmula:

$$u^{\#}(t) = \begin{cases} \omega & \text{para } t' < t \leq t'' \\ u(t), & \text{para } t_o \leq t \leq t' \text{ y } t'' < t \leq t_1 \end{cases}$$

teniendo en cuenta que si alguno de los intervalos fuera vacio no se aplicaría la definición correspondiente, es también admisible.

c) Si el intervalo $[t_o,t_1]$ se divide por medio de puntos de subdivisión en un número finito de subintervalos en cada uno de los cuales el control $u(t)$ es admisible, entonces este control es también admisible en el intervalo completo $[t_o,t_1]$. Además, un control admisible en $[t_o,t_1]$ continuaría siendo admisible en cualquier subintervalo. También se supondrá que un control obtenido a partir de un control admisible: $u(t)$; $t_o \leq t \leq t_1$, por una traslación en el tiempo (es decir el control $u_1(t) = u(t-\alpha)$ para algún α y $t_o+\alpha \leq t \leq t_1+\alpha$) es también admisible.

Supongamos, por otra parte, que las funciones $f_i(x,u)$; $i=0,1,...,n$ son continuas y con derivadas parciales continuas respecto a x en $\mathbb{R}^n \times \bar{\Omega}$ lo que garantiza la existencia y unidad de soluciones de la ecuación integral.

Entonces, si el conjunto dado por (57) es convexo y cerrado $\hat{\bar{x}}(t)$; $t_o \leq t \leq t_1$, $t_1 > t_o$, es una solución del Problema (56) con vector de control correspondiente $\bar{u}(t)$, existe una función vectorial no nula $\hat{\bar{\eta}}(t) = (\eta_o(t),\eta_1(t),...,\eta_n(t))$ tal que:

(81) $\hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t),\bar{u}(t)) = \max\limits_{\omega \in \Omega} \hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t),\omega)$; para casi todo $t \in [t_o,t_1]$, con

(82) $\bar{\eta}_o(t_1) \leq 0$, y

(83) $\hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t), \bar{u}(t)) = 0$, para casi todo $t \in [t_o, t_1]$.

DEMOSTRACION

En primer lugar hay que observar que si $\hat{\bar{x}}(t)$ es una trayectoria optimal en $t_o \leq t \leq t_1$ cada subtrayectoria en los intervalos $[t_o, t^*]$ con $t_o \leq t^* \leq t_1$, será también optimal con $G_1 = \{\bar{x}(t^*)\}$ pues en otro caso se contradiría la optimalidad de $\hat{\bar{x}}(t)$ en el intervalo total $[t_o, t_1]$.

De hecho tendremos que $\hat{\bar{x}}(t^*)$ pertenecerá a la frontera inferior respecto a su primera coordenada del conjunto convexo cerrado C para todo $t^* \in [t_o, t_1]$.

En efecto, si no fuera asi existiría otro punto $\hat{x}^{\#}$ de C, con $x^{\#} = \bar{x}(t^*)$, tal que: $x_o^{\#} < \bar{x}_o(t^*)$. Además, como $x^{\#}$ pertenece al conjunto de accesibilidad desde t_o esto implica que existiría una función de control $u^{\#}(t)$, y un instante $t^{\#}$, de manera que la función $u^{\#}(t)$ estaría definida en $[t_o, t^{\#}]$, y se verificaría:

$$x^{\#}(t) = \int_{t_o}^{t} f(x^{\#}(s), u^{\#}(s))ds + x^{(o}; \quad \text{para} \quad t_o \leq t \leq t^{\#},$$

con $x^{\#}(t^{\#}) = x^{\#}$

Si ponemos ahora: $t^{\#} - t^* = \alpha$, y recordamos las hipótesis hechas sobre la clase Δ de los controles admisibles se sigue que la función:

$$u^{\#}(t) = \begin{cases} u^{\#}(t), & \text{si} \quad t_o \leq t \leq t^{\#} \\ \bar{u}(t-\alpha), & \text{si} \quad t^{*}+\alpha < t \leq t_1 + \alpha \end{cases}$$

sería un control admisible en el intervalo $[t_o ; t^{\#}+t_1 - t^*]$, con estado final: $x^{\#}(t^{\#}+t_1 - t^*) = \overline{x}(t_1)$, y de manera que:

$$x_o^{\#}(t^{\#} + t_1 - t^*) < \overline{x}_o(t_1),$$

lo que contradiría la optimalidad de $\hat{\overline{x}}(t)$ en $[t_o,t_1]$.

(Obsérvese que en la demostración se supone que el instante de llegada t_1 es arbitrario. Los problemas en los que el instante final es fijo pueden reducirse facilmente a éste si se añade una nueva variable de estado al problema).

De lo anterior se sigue que, por la convexidad de C y teniendo en cuenta el Teorema (58), para cada $t \in [t_o,t_1]$ existe un vector $\hat{\overline{\eta}}(t)$ tal que:

$$(84) \qquad \hat{\overline{\eta}}(t) \cdot \int_{t_o}^{t} \hat{f}(\overline{x}(s),\overline{u}(s))ds = \max_{\substack{u(s) \in \Omega \\ u \in \Delta}} \hat{\overline{\eta}}(t) \cdot \int_{t_o}^{t'} \hat{f}(x(s),u(s))ds$$

con

$$x(s) = \int_{t_o}^{s} f(x(z),u(z))dz + x^{(o}; \qquad \forall s \in [t_o,t^*];$$

$$\forall t' \in \mathbb{R}; \quad t' \geq t_o$$

(y en particular, por lo tanto, para $t' = t$).

Definamos ahora la sucesión de controles admisibles, en el intervalo $[t_o,t]$, siguiente:

$$u^{(k,\omega}(s) = \begin{cases} \overline{u}(s), & \text{si} \quad t_o \leq s \leq t - \dfrac{t-t_o}{k} \\ \omega \ , & \text{si} \quad t - \dfrac{t-t_o}{k} < s \leq t \end{cases}$$

siendo ω un elemento arbitrario de Ω y $k \in \mathbb{N}$.

De acuerdo con (84) para $t' = t$, se seguiría que:

$$(85) \qquad \hat{\eta}(t) \cdot \int_{t-\frac{t-t_o}{k}}^{t} \hat{f}(\bar{x}(s),\bar{u}(s))ds \geq \hat{\eta}(t) \cdot \int_{t-\frac{t-t_o}{k}}^{t} \hat{f}(x(s),\omega)ds; \qquad \forall k \in \mathbb{N},$$

en donde

$$x(s) = \int_{t-\frac{t-t_o}{k}}^{s} f(x(z),\omega)dz + \bar{x}(t - \frac{t-t_o}{k}); \qquad \forall s \in [t-\frac{t-t_o}{k}, t]$$

Ya que $\hat{f}(x(s),\omega)$ es una función continua de s podemos poner:

$$(86) \qquad \int_{t-\frac{t-t_o}{k}}^{t} \hat{f}(x(s),\omega)ds = \hat{f}(x(t-\frac{t-t_o}{k}),\omega)\,\frac{t-t_o}{k} + o(\frac{1}{k}, x(t - \frac{t-t_o}{k})),$$

en donde para $x(t - \frac{t-t_o}{k}) = \bar{x}(t - \frac{t-t_o}{k})$ fijo se verifica:

$$\lim_{\substack{r\to\infty \\ r\in\mathbb{N}}} r \cdot o(\frac{1}{r}, \bar{x}(t - \frac{t-t_o}{k})) = 0$$

Ahora bien, de (85) resulta también que:

$$\lim_{k\to\infty} k \;\hat{\eta}(t) \cdot o(\frac{1}{k}, \bar{x}(t - \frac{t-t_o}{k})) \geq 0,$$

pues en caso contrario, teniendo en cuenta que la función:

$$\hat{x}(t) = \int_{t_o}^{t} f(\bar{x}(s),\bar{u}(s))ds + [x^0]$$

es absolutamente continua en $[t_o,t_1]$ y que por lo tanto tendría derivada en casi todo punto que coincidiría también en casi todo punto con el integrando (Véase la Nota 5), se seguiría que:

$$\hat{\eta}(t) \int_{t-\frac{t-t_o}{k^*}}^{t} \hat{f}(\bar{x}(s),\bar{u}(s))ds < \hat{\eta}(t) \int_{t-\frac{t-t_o}{k^*}}^{t} \hat{f}(x(s),\omega)ds \;,$$

si t fuera un punto en el que existiera la derivada de $\bar{x}(t)$

y k^* fuera suficientemente grande, con lo que se contradiría (85).

Por lo tanto se puede concluir que:

$$\lim_{k} \frac{k}{t-t_o} \hat{\bar{\eta}}(t) \cdot \int_{t-\frac{t-t_o}{k}}^{t} f(\bar{x}(s),\bar{u}(s))ds = \hat{\bar{\eta}}(t).\hat{f}(\bar{x}(t),\bar{u}(t)) \geq$$

$$\geq \lim_{k} \frac{k}{t-t_o} \hat{\bar{\eta}}(t) \cdot [f(\bar{x}(t-\frac{t-t_o}{k},\omega) \frac{t-t_o}{k} +$$

$$+ o(\frac{1}{k}, \bar{x}(t-\frac{t-t_o}{k}))] \geq \lim_{k} \hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t-\frac{t-t_o}{k}, \omega) =$$

$$= \hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t),\omega); \quad \text{para casi todo } t \text{ en } [t_o,t_1].$$

Es decir:

$$\hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t),\bar{u}(t)) = \max_{\omega \in \Omega} \hat{\bar{\eta}}(t) \cdot \hat{f}(\bar{x}(t),\omega); \quad \text{para casi}$$

todo t en $[t_o,t_1]$.

En todo lo anterior el vector $\hat{\bar{\eta}}(t)$ tiene un cierto grado de indeterminación puesto que cualquier vector igual al anterior multiplicado por una constante positiva cumpliría también las condiciones enunciadas. Podemos pues suponer, sin pérdida de generalidad que: $\|\hat{\bar{\eta}}(t)\| = (\sum_{i=o}^{n} \bar{\eta}_i^{\,2})^{1/2} = 1, \ \forall t \in [t_o,t_1]$, ya que $\hat{\bar{\eta}}(t) \neq 0$. Además la propia función $\hat{\bar{\eta}}(t)$, $t_o \leq t \leq t_1$, estaría también indeterminada no sólo por la multiplicación por una función positiva sino también por pasos al límite. Esto es, si se sustituye un valor de la función $\hat{\bar{\eta}}(t)$ por el límite obtenido en alguna sucesión convergente $\{\hat{\bar{\eta}}(t_r)\}$ con $t_r \to t$, la nueva función seguiría verificando evidentemente (81).

Supongamos, pues, que se ha seleccionado un conjunto den so de medida nula en $[t_o,t_1]$, que incluya a t_1 y que se ha modificado la definición de $\hat{\eta}(t)$ en el conjunto complementario al denso sustituyendo el antiguo valor por el límite obtenido en alguna sucesión convergente a la derecha del punto en cuestión formada por valores de $\hat{\eta}(t)$ en el conjunto denso. (Esta sucesión convergente se puede obtener en cualquier caso ya que $\|\hat{\eta}(t)\| = 1, \ \forall t \in [t_o,t_1]$). Denotemos de nuevo a la función modificada por $\hat{\eta}(t)$, para $t_o \leq t \leq t_1$.

Teniendo en cuenta la redefinición de $\hat{\eta}(t)$ es fácil comprobar que se verifican (82) y (83).

En efecto, el que $\hat{\eta}_o(t_1) \leq 0$ se deduce del Teorema (8) y de la definición de la nueva función $\hat{\eta}(t)$ en t_1, coincidente con la primitiva definición de $\hat{\eta}(t)$, mientras que (83) se sigue de que:

$$\hat{\eta}(t) \,.\, \hat{x}(t) \geq \hat{\eta}(t)\, \hat{x}(t-\delta)$$

si $\delta \geq 0$ es suficientemente pequeño, y también:

$$\hat{\eta}(t-\delta) \,.\, \hat{x}(t-\delta) \geq \hat{\eta}(t-\delta) \,.\, \hat{x}(t)$$

Así, supuesto que t es un punto en el que existe $\frac{d\,\hat{x}(t)}{dt}$ y no pertenece al conjunto denso que ha servido para redefinir $\hat{\eta}(t)$ y δ se escoge de manera que los puntos $t-\delta$ correspondan a la subsucesión que ha servido para definir la nueva función $\hat{\eta}(t)$, tendremos:

$$\lim_{\delta \to 0} \hat{\eta}(t) \,.\, \frac{\hat{x}(t) - \hat{x}(t-\delta)}{\delta} = \hat{\eta}(t) \,.\, \hat{f}(\bar{x}(t),\bar{u}(t)) \geq 0$$

y, por otra parte:

$$\lim_{\delta \downarrow 0} \hat{\eta}(t-\delta) \frac{\hat{\bar{x}}(t) - \hat{\bar{x}}(t-\delta)}{\delta} = \hat{\eta}(t) \; . \; \hat{f}(\bar{x}(t),\bar{u}(t)) \leq 0$$

de donde se deduce (83).

El interés del Teorema (79) es más técnico que práctico pues si no se complementaran (81), (82) y (83) con alguna otra condición sobre la función $\bar{\eta}(t)$, tal como las que veremos en el capitulo siguiente, no parece fácil determinar $\bar{u}(t)$ a partir de sólo dichas condiciones. Está claro que la determinación efectiva de $\bar{u}(t)$ en el caso convexo que estamos estudiando dependería fundamentalmente de la aplicación del Teorema (58) o del Corolario (66).

En la demostración del Teorema (79) la hipótesis referente a la convexidad del conjunto C podría sustituirse por la de la convexidad de los conjuntos C_{t_o,t_1} sin que se modificaran los resultados (81) y (82) del teorema. Esto es así porque la hipótesis de convexidad de C solo juega un papel efectivo en la demostración de (83). Por ello se puede enunciar el siguiente corolario:

(88) COROLARIO

Supongamos dadas las hipótesis del Teorema (79) con la excepción de la hipótesis sobre la convexidad del conjunto C dado por (57) la cual se sustituye por la de la convexidad de los conjuntos C_{t_o,t_1} correspondientes a particularizaciones de C para t_o y t_1 fijos. Entonces siguen verificandose (81) y (82).

DEMOSTRACION

Es una consecuencia inmediata de la demostración del Teorema (79) ya que la utilización de (84) se hace luego sólo para el caso $t' = t$ y esta expresión (84) se deduce también, si $t' = t$, de las hipótesis de convexidad de los conjuntos C_{t_o,t_1} con $t_o \leq t \leq t_1$.

Con este resultado terminamos el estudio del caso convexo tanto estático como dinámico y pasamos a estudiar el segundo caso importante (el más investigado desde el punto de vista tradicional) que es el caso diferenciable.

Sobre las relaciones entre los desarrollos correspondientes a ambos casos, tanto en el caso estático como en el dinámico, puede verse la Nota 6.

CAPITULO II

TEORIA DE LAS CONDICIONES NECESARIAS EN EL CASO DIFERENCIABLE. EL PRINCIPIO DE LA PROYECCION MAXIMAL TANGENCIAL

1. EL CASO ESTATICO DIFERENCIABLE.

Vamos a investigar en esta sección las condiciones necesarias de optimalidad en problemas en los que el correspondiente conjunto C no es convexo aunque cumple unas condiciones determinadas de regularidad.

Para centrar la cuestión supongamos por un momento que C, ya sea el problema estático o dinámico, tiene como frontera una hipersuperficie con hiperplano tangente en cada punto. El problema de las condiciones necesarias de optimalidad vuelve a ser aquí, igual que en las secciones anteriores, el de caracterizar a los puntos de la frontera inferior de C respecto a su primera coordenada, entre los que necesariamente se haya el mínimo. En la hipótesis anterior de existencia de hiperplano tangente en todo punto de esta frontera inferior, se podrían, pues, caracterizar los puntos $\hat{\hat{x}}$ pertenecientes a esta frontera.

En efecto, consideremos el conjunto de vectores unitarios: $\frac{\hat{x} - \hat{\hat{x}}}{\|\hat{x} - \hat{\hat{x}}\|}$, con $x \in C$, $\hat{x} \neq \hat{\hat{x}}$, y determinemos el conjunto de vectores límite de los anteriores cuando $\hat{x} \to \hat{\hat{x}}$ en C, siempre que el límite exista. El haz de vectores unitarios así obtenido, con origen en el punto $\hat{\hat{x}}$ de la frontera inferior de C, cuyo elemento genérico denotaremos por $\hat{e}(\hat{\hat{x}})$, sería tal que:

$$(89) \qquad \hat{\eta}.\hat{e}(\hat{\hat{x}}) \leq 0, \qquad \forall \hat{e}(\hat{\hat{x}}) = \lim_{\substack{\hat{x} \to \hat{\hat{x}} \\ x \in C}} \frac{\hat{x} - \hat{\hat{x}}}{\|\hat{x} - \hat{\hat{x}}\|},$$

en donde $\hat{\eta}$ es el vector normal del hiperplano tangente dirigido al exterior de C (es decir tal que $\eta_0 \leq 0$ y $\hat{\hat{x}} + \hat{\eta}$ no perteneciente a C, para $\|\hat{\eta}\|$ suficientemente pequeño).

Además, si $\hat{\hat{x}}$ pertenece a la frontera inferior de C res pecto de la coordenada x_o y existe hiperplano tangente a la hipersuperficie frontera de C en $\hat{\hat{x}}$, con vector característico $\hat{\eta} = (\eta_o, \eta_1, \ldots, \eta_n)$, habría de ser evidentemente $\eta_o \leq 0$. La discusión informal anterior muestra algunas de las nuevas características del problema cuando se abandona la hipótesis de convexidad de los conjuntos C y se introducen otras que garantizan la existencia de hiperplano tangente en cada punto de la frontera inferior de C respecto a x_o.

Como se puede observar la diferencia esencial es la de que el principio de proyección maximal sobre un determinado vector $\hat{\eta}$ pasaría a ser una condición local y no ampliable a todos los vectores $(\hat{x}-\hat{\hat{x}})$ con $\hat{x} \in C$. Es más, esta condición de proyección maximal habría que entenderla no para los vectores $\hat{x}-\hat{\hat{x}}$ sino para los vectores límite $\hat{e}(\hat{\hat{x}})$, que es para los únicos que se puede garantizar. Por otra parte, toda la argumentación que hemos desarrollado para justificar (89) se ha apoyado en la existencia de un hiperplano tangente a la frontera inferior de C, lo cual suele ser difícil de probar a partir de los elementos de que se dispone inicialmente para resolver el problema, esto es los datos del Problema (1) y las hipótesis sobre las funciones y cantidades que intervienen en él.

Conviene proponer, por lo tanto, una condición a chequear del tipo de la (89) pero sin la necesidad de incluir hipótesis sobre la existencia de hiperplanos tangentes, aunque el recurso a éstos proporciona una buena imagen del significado de las condiciones que se obtendrían.

Con el fin de dar mayor manejabilidad a la condición (89) consideraremos no los límites en los cocientes $(\hat{x}-\hat{\hat{x}})/\|\hat{x}-\hat{\hat{x}}\|$, sino en los cocientes:

(90) $$\frac{\hat{x}(\delta) - \hat{\bar{x}}}{\delta},$$

cuando $\delta \downarrow 0$, es decir cuando δ decrece a 0, con la propiedad de que $\hat{x}(\delta) \in C$, para todo δ en un intervalo $(0,\delta_o)$ y de manera que el límite de (90) para $\delta \downarrow 0$ exista.

Naturalmente, los vectores límites obtenidos de esta manera incluirían a los vectores límites unitarios del tipo $\hat{e}(\hat{\bar{x}})$, pero desde este nuevo punto de vista nos ahorramos tener que restringirnos a vectores límite unitarios, lo cual tendrá su importancia en el aspecto operativo y también para caracterizar con facilidad, desde algunos puntos de vista de interés, como el de la convexidad, el conjunto de vectores límite que vamos a utilizar en cada caso.

Con estas modificaciones vamos a proponer, pues, un nuevo Principio, del mismo tipo que el (8), pero relativo a los vectores límite de los cocientes $(\hat{x}(\delta) - \hat{\bar{x}})/\delta$, en vez de a los $\hat{x}-\hat{\bar{x}}$, mediante el cual generaremos las condiciones necesarias de optimalidad en el caso diferenciable, al igual que hicimos con el (8) respecto al caso convexo. Este es el siguiente:

(91) TEOREMA (Principio de la Proyección Maximal Tangencial)

Sea C un subconjunto cerrado de puntos $\hat{x}$ de R^{n+1}, no necesariamente convexo. Supongamos que $\forall \hat{\bar{x}} \in C$, existe un conjunto de vectores $L_{\hat{\bar{x}}}$ definido de la siguiente manera:

(92) $$L_{\hat{\bar{x}}} = \{\hat{w} \in R^{n+1} \;/\; \hat{w} \in \{\hat{v} \in R^{n+1} \;/\; \hat{v} = \lim_{\substack{\delta \downarrow 0 \\ \hat{x}(\delta) \in C}} \frac{\hat{x}(\delta)-\hat{\bar{x}}}{\delta}\}\},$$

(obsérvese que $L_{\hat{\bar{x}}}$ no tendría por qué coincidir con el conjunto de vectores límite de los cocientes $(\hat{x}(\delta) - \hat{\bar{x}}) / \delta)$, el cual es convexo, e incluye el vector 0 (que se obtendría haciendo $\hat{x}(\delta) = \hat{\bar{x}}$, $\forall \delta > 0$). Entonces, si $\hat{\bar{x}}$ pertenece a la frontera inferior de C respecto a la primera coordenada de sus puntos, x_o, o bien existe un vector $\hat{\bar{\eta}} = (\bar{\eta}_o, \bar{\eta}_1, \ldots, \bar{\eta}_n) \neq 0$, de R^{n+1}, tal que:

$$(93) \qquad \hat{\bar{\eta}} \cdot \hat{w} \leq 0 \; ; \qquad \forall \hat{w} \in L_{\hat{\bar{x}}} \, ,$$

y con $\bar{\eta}_o \leq 0$, o bien existe una sucesión de puntos $\bar{x}^{(k}$, $k \in N$, pertenecientes a C, y una sucesión de vectores $\hat{\bar{\eta}}^{(k} \neq 0$, de R^{n+1}, tales que:

$$(94) \qquad \hat{\bar{\eta}}^{(k} \cdot \hat{w} \leq 0; \qquad \forall w \in L_{\hat{x}^{(k}} \, , \qquad y \qquad \forall k \in N; \qquad \eta_o^{(k} \leq 0,$$

y tal que: $\lim\limits_{k} \hat{x}^{(k} = \hat{\bar{x}}$.

DEMOSTRACION.

Supongamos que $\hat{\bar{x}} = (\bar{x}_o, \bar{x}_1, \ldots, \bar{x}_n)'$ pertenece a la frontera del conjunto C respecto a la primera coordenada de sus puntos y que no existe un vector $\hat{\eta} \neq 0$ tal que:

$$\hat{\eta} \cdot \hat{w} \leq 0; \qquad \forall \hat{w} \in L_{\hat{\bar{x}}} \, ,$$

con $\bar{\eta}_o \leq 0$.

Como $L_{\hat{\bar{x}}}$ es convexo por hipótesis, se sigue de ello que existe un vector $\hat{w}$ de $L_{\hat{\bar{x}}}$ en cada dirección del espacio euclídeo R^{n+1}. En efecto, si no fuera así existiría una dirección del espacio representada por el vector unitario $\hat{e}$, tal que $\lambda\hat{e}$ no pertenecería a $L_{\hat{\bar{x}}}$ para cualquier $\lambda > 0$. Esto es lo mismo que afirmar que $\hat{\bar{x}}$ sería un punto frontera del conjunto convexo:

$$V = \{\hat{x} \in R^{n+1} \; / \; \hat{x} = \hat{\bar{x}} + \hat{w}; \quad \hat{w} \in L_{\hat{\bar{x}}}\} \, ,$$

ya que, por hipótesis, $0 \in L_{\hat{\bar{x}}}$.

Se seguiría entonces, por el razonamiento aplicado en la demostración del Teorema (8), que existiría un vector $\hat{\eta} \neq 0$, tal que:

$$\hat{\eta} \cdot (\hat{x}-\hat{\bar{x}}) = \hat{\eta}\hat{w} \leq 0; \qquad \forall \hat{w} \in L_{\hat{\bar{x}}},$$

y además, de que $\hat{\bar{x}}$ pertenece a la frontera inferior de C respecto a x_o se deducirá que la primera coordenada de vector $\hat{\eta}$ podría seleccionarse de manera que $\eta_o \leq 0$ pues en otro caso el rayo que parte de $\hat{\bar{x}}$ en la dirección $(-1,0,\ldots,0)'$ estaría contenido, salvo $\hat{\bar{x}}$, en el interior de V, lo que contradiría la suposición anterior. Se deduce por lo tanto que la no existencia de $\hat{\eta}$, con $\eta_o \leq 0$, implica la existencia de vectores $\hat{w}$ de $L_{\hat{\bar{x}}}$ en todas las direcciones del espacio.

En particular, entonces, existiría un $\hat{w} \in L_{\hat{\bar{x}}}$, en la dirección $(-1,0,\ldots,0)'$, de lo que se deduce que, para $\delta_o > 0$ suficientemente pequeño, existirá una colección de puntos $\hat{x}(\delta)$, $0 \leq \delta \leq \delta_o$, contenida en C, tal que:

$$\lim_{\delta \downarrow 0} \frac{\hat{x}(\delta) - \hat{\bar{x}}}{\delta} = \hat{w}$$

Supongamos construida la sucesión de conos convexos cerrados y truncados $\{M_k\}$, $k \in \mathbb{N}$, de vértices en el punto $\hat{x}$, definida de la siguiente manera:

$$M_k = \{\hat{x} \in \mathbb{R}^{n+1} \;/\; \sum_{j=1}^{n} (x_j-\bar{x}_j)^2 \leq (x_o-\bar{x}_o)^2/k \quad \text{para cada}$$

$$x_o \text{ fijo}; \quad \bar{x}_o-1 \leq x_o \leq \bar{x}_o\}$$

Como $\hat{\bar{x}} \in M_k$, $\forall k \in \mathbb{N}$, ningún M_k es vacío.

Sea:

$$\inf_k = \inf \{x_o \in \mathbb{R} \;/\; \hat{x} = (x_o,x_1,\ldots,x_n)' \in C \cap M_k\}$$

Ya que C es cerrado y M_k compacto, $C \cap M_k$ será com pacto, de lo que se deduce que el ínfimo $\inf_k$ se obtendría, para cada $k \in \mathbb{N}$, en un determinado punto de C, al menos, que denominaremos $\hat{x}^{(k}$.

Entonces, existirá un k_o tal que $x_o^{(k} > x_o-1$, $\forall k \geq k_o$, pues en otro caso se podría construir una subsucesión $\{x^{(k_r}\}$, $r \in \mathbb{N}$, de la $\{x^{(k}\}$, de manera que $x_o^{(k_r} = \bar{x}_o-1$, $\forall r \in \mathbb{N}$, de la cual podría extraerse una subsucesión convergente a un punto $\hat{x}^* \in C$, tal que $x_j^* = \bar{x}_j$, $j=1,2,\ldots,n$; y $x_o^* = \bar{x}_o-1$, lo que contradiría la suposición de que $\hat{x}$ es un punto de la frontera inferior de C respecto a x_o. (Obsérvese que $x_o^{(k}$ debe de ser en cualquier caso distinto de $\bar{x}_o$ ya que hay puntos de la colección $x(\delta)$, $0 \leq \delta \leq \delta_o$, interiores a M_k, cualquiera que sea k, y con coordenada x_o inferior a $\bar{x}_o$).

Ahora bien, si $\hat{x}_o^{(k} > \bar{x}_o-1$, se sigue que el conjunto $L_{\hat{x}_o^{(k}}$ no puede contener todas las direcciones posibles del espacio. En efecto, si $\hat{x}^{(k}$ fuera interior al cono truncado M_k es claro que no habría ningún vector de $L_{\hat{x}_o^{(k}}$ en la dirección $(-1,0,\ldots,0)'$ por la minimalidad de $x_o^{(k}$.

De la misma manera, si $\hat{x}^{(k}$ es tal que: $\sum_{j=1}^{n} (x_j^{(k}-\bar{x}_j)^2 = (x_o^{(k}-\bar{x}_o)^2/k$, supuesto que $\bar{x}_o - 1 < x_o^{(k} < \bar{x}_o$, es fácil comprobar también entonces que no puede haber ningún vector de $L_{\hat{x}_o^{(k}}$ en la dirección $(-1,0,\ldots,0)'$. En efecto, en otro caso existirían puntos de C interiores a M_k con coordenada x_o inferior a $x_o^{(k}$, lo que volvería a contradecir la minimalidad de $x_o^{(k}$.

Se sigue entonces de lo anterior que existen vectores de $\mathbb{R}^{n+1}$, $\hat{\eta}^{(k}$, tales que $\hat{\eta}^{(k} \neq 0$ y $\hat{\eta}^{(k} . \hat{w} \leq 0$, $\forall \hat{w} \in L_{\hat{x}^{(k}}$ y $\forall k \in \mathbb{N}$, $k \geq k_o$.

Además, ya que $\hat{x}^{(k}$ pertenece a la frontera inferior del conjunto convexo:

$$\{\hat{x} \in \mathbb{R}^{n+1} \,/\, \hat{x} = \hat{x}^{(k} + \hat{w}; \quad \hat{w} \in L_{\hat{x}^{(k}}\},$$

respecto a la coordenada x_o, la aplicación del Teorema (8) muestra que $\eta_o^{(k} \leq 0$, $\forall k \geq k_o$.

Haciendo pues $k \to \infty$ se puede extraer entonces de la sucesión $\{\hat{x}^{(k}\}$, $k \in \mathbb{N}$, $k \geq k_o$, una subsucesión convergente a un punto del segmento:

$$\{\hat{x} \in \mathbb{R}^{n+1} \,/\, \hat{x} = \lambda\hat{\bar{x}} + (1-\lambda)(\bar{x}_o-1,\bar{x}_1,\ldots,\bar{x}_n)'; \quad 0 \leq \lambda \leq 1\},$$

que no puede ser otro que $\hat{\bar{x}}$, pues en otro caso se contradiría la minimalidad de $\bar{x}_o$.

Esto completa la demostración del Teorema.

La aplicación del Principio (91), en lo que a esta memoria concierne, se referirá a los conjuntos C relativos a los Problemas (17) y (56), dados respectivamente por (18) y (57), aunque podría pensarse en su aplicación productiva a otros tipos de problemas.

En cuanto al enunciado del Teorema (91), algunos ejemplos mostraran la conveniencia de algunas de sus afirmaciones.

En efecto, si consideramos en $\mathbb{R}^2$ el conjunto de puntos definido de la siguiente manera:

$$\{(x_1,x_2) \in \mathbb{R}^2 \,/\, \sqrt{1-(x_1-2k)^2} + x_2 \geq 0; \quad \text{para}$$
$$2k-1 \leq x_1 \leq 2k+1, \quad k \in \mathbb{N}\},$$

se observará que dicho conjunto incluye algunos puntos, tal como por ejemplo el (1,0), los cuales pertenecen a la frontera inferior del conjunto respecto a la coordenada x_2, pero con respecto a los que se pueden definir conjuntos convexos del tipo $L_{\hat{x}}$ que continen vectores en todas las direcciones del plano.

Esto demuestra que, en las hipótesis del Teorema, no se puede esperar concluir que todo punto frontera $\hat{x}$ es tal que existe algún vector $\hat{\eta} \neq 0$, con: $\hat{\eta} . \hat{w} \leq 0$, $\forall \hat{w} \in L_{\hat{x}}$. Sin embargo, como se comprueba en el ejemplo, si puede construirse una sucesión de puntos $\hat{x}^{(k}$ convergente al punto en cuestión, en este caso el (1,0), de manera que sus correspondientes $L_{\hat{x}^{(k}}$ convexos son tales que existen vectores $\hat{\eta}^{(k} \neq 0$ de manera que $\hat{\eta}^{(k} . \hat{w} \leq 0$, $\forall \hat{w} \in L_{\hat{x}^{(k}}$ y $\forall k \in \mathbb{N}$.

En cuanto a la necesidad de exigir que el conjunto C sea cerrado, ésto es del todo evidente si se considera, por ejemplo, el caso del punto (0,0) en el conjunto:

$\{(x_1,x_2) \in \mathbb{R}^2 \ / \ x_1^2 + x_2^2 \leq 1$, exceptuando el segmento semiabierto entre el punto (0,0) y el (0,-1) que incluye a (0,0)$\}$

En este ejemplo el punto (0,0) pertenece a la frontera inferior respecto de x_2, del conjunto anterior, pero se pueden definir conjuntos $L_{\hat{x}}$ tales que contengan todas las direcciones del plano cualquiera que sea el punto $\hat{x}$ del conjunto. Esto demuestra, pues, que la condición de exigir que el conjunto C sea cerrado en el enunciado del Teorema (91) no es superflua.

Este Principio (91) aparece más o menos encubierto en los autores sobre Teoría del Control Optimo, y, sobre todo, en lo que se refiere a la demostración del Principio de Máximo de Pontryagin. En los problemas de Programación No Lineal con diferenciabilidad, de los cuales nos vamos a ocupar a continuación, su exposición es menos evidente. (Sobre la conexión del Principio

(91) y los desarrollos tradicionales nos remitimos a la Nota 7).

La diferencia fundamental entre el Principio (91) y el (8) es la de que el primero es una condición local, mientras que el segundo es una condición global, como se deduce de que, mientras que en el segundo se afirmaba la existencia de un $\hat{\bar{\eta}} \neq 0$, tal que $\hat{\bar{\eta}} \cdot (\hat{x}-\hat{\bar{x}}) \leq 0$ para el $\hat{\bar{x}}$ óptimo, en el primero la desigualdad anterior queda reducida a la de que:

$$\lim_{\delta \downarrow 0} \hat{\bar{\eta}} \cdot \left(\frac{\hat{x}(\delta) - \hat{\bar{x}}}{\delta}\right) \leq 0; \quad \text{ó} \quad \hat{\bar{\eta}} . \hat{w} \leq 0, \quad \text{para} \quad \hat{v} \in L_{\hat{x}}$$

Las consecuencias de la diferencia anterior en el caso dinámico continuo se reflejaran en que la sola aplicación del Principio (91) no identificaría a la función $\hat{\bar{\eta}}(t)$, $t_o \leq t \leq t_1$, que caracteriza al problema, nada mas que localmente. Sin embargo, como más adelante veremos, las hipotesis usuales sobre las funciones f_i que caracterizan el Problema (56), y en particular la existencia de derivadas parciales continuas con respecto a $\hat{x}$, permiten deducir determinadas reglas a las que se ajusta el comportamiento de la función $\hat{\bar{\eta}}(t)$. (Vease la Nota 7 bis)

Vamos a abordar a continuación la aplicación del Principio (91) al primero de los dos problemas que nos interesa investigar aquí, el Problema (17), lo que nos permitirá obtener una condición necesaria de optimalidad.

(95) TEOREMA.

Consideremos el Problema (17) que volvemos a plantear aquí:

(96) $\min \; x_o = f_o(u)$

sujeto a:

$x = -f(u); \quad u \in \Omega; \quad x \in G,$

siendo G el conjunto:

$$\{x \in R^n \ / \ x \geq 0\},$$

y Ω un subconjunto abierto de R^m.

(No supondremos ahora la convexidad de Ω ni de las f_i).

Entonces, si $\bar{u} \in \Omega$ es una solución del Problema (96) y $\hat{f} = (f_o, f_1, \ldots, f_n)'$ es diferenciable con derivadas parciales continuas en un entorno de $\bar{u}$, existe un vector $\hat{\eta} = (\bar{\eta}_o, \bar{\eta}_1, \ldots, \bar{\eta}_n) = (\bar{\eta}_o, \bar{\eta}) \neq 0$, $\hat{\eta} \geq 0$, tal que:

(97) $\quad \bar{\eta}_o \nabla f_o(\bar{u}) + \bar{\eta} \nabla f(\bar{u}) = 0;$ y

(98) $\quad \bar{\eta} \cdot f(\bar{u}) = 0,$

en donde:

(99) $$\nabla f_i(\bar{u}) = \left(\frac{\partial f_i}{\partial u_1}, \frac{\partial f_i}{\partial u_2}, \ldots, \frac{\partial f_i}{\partial u_m}\right)_{u=\bar{u}}, \quad i=0,1,2,\ldots,n,$$

y

(100) $$\nabla f(\bar{u}) = \begin{bmatrix} \nabla f_1(\bar{u}) \\ \nabla f_2(\bar{u}) \\ \vdots \\ \nabla f_n(\bar{u}) \end{bmatrix}.$$

DEMOSTRACION.

Si $\bar{u}$ es una solución del Problema (96) también será solución del siguiente problema equivalente:

(101) $\quad \min x_o = f_o(u)$

sujeto a:

$$x_i = -(f_i(u) + u_{m+i}); \quad i=1,2,\ldots,n,$$

con $u \in \Omega$, $(u_{m+1}, u_{m+2}, \ldots, u_{m+n})$ perteneciente al subconjunto de R^n tal que $u_{m+i} \geq 0$, $i=1,2,\ldots,n$, y $x \in G^*$, en donde:

$$G^* = \{x \in R^n \ / \ x = 0\}$$

Denotemos

$$\tilde{u} = (u_1, u_2, \ldots, u_{m+n})$$

y

$$x_i = g_i\ (\tilde{u})\ ,\ i = 0,1,2,\ldots,n$$

En efecto la equivalencia de los dos problemas en cuanto a - su solución en lo que concierne a los u_j, $j = 1,2,\ldots,m$, se deduce de que para cada $u\varepsilon\Omega$ que verifique $f\ (u)\leq 0$ existe un $\tilde{u}$ tal que - $f_i\ (u) + u_{m+i} = 0$, $i = 1,2,\ldots,n$ y viceversa.

Consideremos, pues, esta segunda versión del problema. Como Ω es abierto y $\tilde{u}\varepsilon\Omega$ se puede determinar una bola cerrada de centro $\tilde{u}$ y radio estrictamente positivo en $\mathbb{R}^m$, que denotaremos por B, la cual esté totalmente contenida en Ω y en el entorno de $\tilde{u}$ en el que $\hat{f}$ es diferenciable con continuidad.

Entonces, el conjunto definido de la siguiente manera:

(102) $= \{\hat{x}\varepsilon\ \mathbb{R}^{n+1} /\ \hat{x} = \hat{g}(\tilde{u});\ u\ \varepsilon\ B\ ;\ u_{m+i}\geq 0,\ i = 1,2,\ldots,n\}.$

es un conjunto cerrado en el que el asociado al $\hat{\bar{x}}$ es un punto de - la frontera inferior respecto a la coordenada x_o.

Efectivamente, si $\hat{x}^{(k}$, $k\varepsilon\ \mathbb{N}$, es una sucesión convergente a $\hat{x}^*$, contenida en C, se verificará:

$$\hat{x}^{(k} = \hat{g}(\tilde{u}^{(k})$$

para algún $u^{(k}$ perteneciente al conjunto compacto de $\mathbb{R}^{m+n}$ dado por

(103) $B\ x\ \{(u_{m+1},\ldots,u_{m+n}) /\ u_{m+i}\geq 0,\ i = 1,2,\ldots n\}.$

Se podrá, por lo tanto extraer una subsucesión convergente de la $\tilde{u}^{(k}$, que volveremos a denotar por $\tilde{u}^{(k}$, tal que $\lim \tilde{u}^{(k} = \tilde{u}^*$.

Entonces, de la continuidad de $\hat{g}$ se deduce:

$$\hat{x}^* = \lim \hat{x}^{(k} = \hat{g}\ (\lim \tilde{u}^{(k}) = \hat{g}\ (\tilde{u}^*) \in C$$

En adelante restringiremos $\tilde{u}$ al conjunto (103). Si $\bar{u}$ es solución de (96) también lo será $\tilde{\bar{u}} = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_{m+n})$, con $\bar{u}_{m+i} = -f_i(\bar{u})$, para i= 1,2,...n, del problema modificado con la restricción adicional de pertenecer al conjunto (103)

Estaríamos pues en condiciones de aplicar el Principio (91) al problema restringido tomando, en un principio, como conjunto C el dado por (102).

Para cada $\tilde{u}$ perteneciente al conjunto (103) con $(u_1, u_2, \ldots, u_m)$ perteneciente al interior de B y $u_{m+i} \leq 0$, con $i \in N_2$, en donde - por N_2 denotamos el subconjunto de índices de $\{1,2,\ldots,n\}$ tal que $\bar{u}_{m+i} = 0$, si $i \in N_2$, se podría generar en cada correspondiente $\hat{x}$ de el siguiente conjunto convexo de vectores tangentes.

$$(104)\quad L_{\hat{x}} = \{\hat{w} \in \mathbb{R}^{n+1} / \hat{w} = \lambda_1 \frac{\partial \hat{g}(\tilde{u})}{\partial u_1} + \lambda_2 \frac{\partial \hat{g}(\tilde{u})}{\partial u_2} + \ldots + \lambda_{m+n} \frac{\partial \hat{g}(\tilde{u})}{\partial u_{m+n}}\ ;$$

con λ_j arbitrario, j = 1,2,...m; $\lambda_j \geq 0$ si $j-m \in N_1$ y λ_j arbitrario, si $j-m \in N_2$, en donde por N_1 denotamos el complementario de $N_2\}$.

Ahora bien, habría ciertas dificultades en definir conos convesos de vectores tangentes del tipo anterior en puntos $\hat{x}$ de C tales que para el $\tilde{u}$ correspondiente se verificara que algún u_j, con $1 \leq j \leq n$, perteneciera a la frontera de B, o bien que $u_{m+i} = 0$, para algún $i \in N_2$.

Supongamos en primer lugar que para una sucesión de bolas $B^{(r}$ del tipo anterior con radios $\tau_r \to 0$ se verifica que $B^{(r} \cap \{u \in \mathbb{R}^m / \hat{f}(u) = \hat{f}(\bar{u});\ u \neq \bar{u}\}$ es vacío si $r \geq r_o$, para un determinado r_o.

Por otra parte de que $f_i(\bar{u}) < 0$, para $i\in N_2$ se seguiría entonces que $\hat{x}$ tal que $x_i < 0$, para $i\in N_2$ si perteneciera a un entorno de $\hat{\bar{x}}$ de diámetro suficientemente pequeño.

Es fácil comprobar entonces que restringiendo las m primeras coordenadas $\tilde{u}$ a dicha bola se obtendría un conjunto en el que el punto de coordenadas $x_i=-f_i(\bar{u})+\bar{u}_{m+i}, i=1,2,\ldots, n$, $x_o = f_o(\bar{u})$ que vendría dado por $\hat{g}(\tilde{\bar{u}})$pertenecería a su frontera inferior respecto a su coordenada x_o, que sería cerrado, y con la propiedad de que si $\hat{x}^{(k}=\hat{g}(\tilde{u}^{(k}) \to \hat{g}(\tilde{\bar{u}})$ entonces se podría seleccionar una subsucesión de manera que los correspondientes $\tilde{u}^{(k} = (u_1^{(k},\ldots,u_m^{(k},\ldots,u_{m+n}^{(k})$, fueran tales que $(u_1^k,\ldots,u_m^{(k})$ fuera interior a la bola de radio τ_{r_o} y $u_{m+i}^{(k}> 0$ para todo k y todo $i\in N_2$.

Se seguiría de ello que en todo punto $\hat{x}$ de interés de este conjunto C se podría definir un cono convexo de vectores tangentes del tipo (104).

De la aplicación de (91) se deduce entonces que, o bien existe un $\hat{\eta}^* \neq 0$,tal que $\hat{\eta}^* \hat{w} \leq 0$, para todo $\hat{w}\in L_{\hat{g}(\tilde{\bar{u}})}$, con $\eta_o^* \leq 0$ o existe una sucesión de puntos $\hat{x}^{(k}$, $k\in\mathbb{N}$, y una sucesión de vectores $\hat{\eta}^{(k}$, $\hat{\eta}^{(k} \neq 0$, para todo $k\in\mathbb{N}$, tales que:

$$\hat{\eta}^{(k} \hat{w} \leq 0; \text{ para todo } \hat{w}\in L_{\hat{x}^{(k}} \text{ y todo } k\in\mathbb{N}$$
$$\eta_o^{(k} \leq 0; \text{ y } \lim \hat{x}^{(k} = \hat{g}(\tilde{\bar{u}})$$

Denotemos por $\tilde{u}^{(k} = (u^{(k},.)$ a los $\tilde{u}$ tales que $\hat{x}^{(k}=\hat{g}(\tilde{u}^{(k})$.

Entonces, de la arbitrariedad de los λ_j para $j = 1,2,..,m$ ó $j = m+i$ con $i\in N_2$, se sigue que:

$$\hat{\eta}^{(k} \frac{\partial \hat{g}(\tilde{u}^{(k})}{\partial u_j} = 0 \; ; \quad j = 1,2,\ldots,m$$

y: $\eta^{(k}\ \frac{\partial g(\tilde{u}^{(k)})}{\partial u_j} = -\eta_i^{(k} = 0$; para $i\varepsilon N_2$

De las últimas igualdades se deduce que

$$\hat{\eta}^{(k}\ f\ (u^{(k)}) = 0$$

Por otra parte, para $j = m+i$, con $i\varepsilon N_1$, tenemos:

$$\hat{\eta}^{(k}\left(\frac{\partial\hat{g}(\tilde{u}^{(k)})}{\partial u_1}, \frac{\partial\hat{g}(\tilde{u}^{(k)})}{\partial u_2}, \ldots, \frac{\partial\hat{g}(\tilde{u}^{(k)})}{\partial u_m}\right) = 0$$

Ahora

$$\frac{\partial g(\tilde{u}^{(k})}{\partial u_j} = \begin{bmatrix} \frac{\partial f_o(u^{(k})}{\partial u_j} \\ \\ \frac{\partial f(u^{(k})}{\partial u_j} \end{bmatrix} \quad ; j = 1,2,\ldots,m$$

Se obtendría entonces:

$$\eta_o^{(k}\ \nabla f_o\ (u^{(k})-\eta^{(k}\ \nabla f\ (u^{(k}) = 0$$

Seleccionando sendas subsucesiones convergentes de la $\eta^{(k}$, la cual podemos suponer normalizada, y de la $\tilde{u}^{(k}$, con límites: $\lim\ \hat{\eta}^{(k} = \hat{\eta}^*$, y necesariamente: $\lim\ \tilde{u}^{(k} = (\bar{u}_1,\bar{u}_2,\ldots,\bar{u}_{m+n})$, tendremos, debido a la continuidad de los gradientes que:

$$\eta_o^*\ \nabla f_o(\bar{u}) - \eta^*\ \nabla f\ (\bar{u}) = 0$$

con:

$$\eta_o^* \leq 0$$

$\eta_i^* = 0$, para los índices i tales que $f_i\ (\bar{u}) < 0$

$\eta_i^* \geq 0$, para los índices i tales que $f_i'\ (\bar{u}) = 0$

que es lo mismo que se obtiene si no se determinara el vector $\hat{\eta}^*$ mediante la sucesión $\hat{x}^{(k}$ sino directamente a partir del cono - $L_{\hat{g}(\tilde{\bar{u}})}$

En definitiva, poniendo:

$$\bar{\eta}_o = -\eta_o^*$$

$$\bar{\eta} = \eta^*$$

tenemos:

$$\eta_o \nabla f_o (\bar{u}) + \bar{\eta} \nabla f (\bar{u}) = 0$$

$$\bar{\eta} f(\bar{u}) = 0$$

$$(\bar{\eta}_o, \bar{\eta}) \geq \quad ; \hat{\bar{\eta}} \neq 0.$$

Solo queda entonces por discutir el caso en que para todo τ en un intervalo $(0,\tau_o)$ y cada bola B de radio τ exista un punto $u = (u_1, u_2, \ldots, u_m)$ de B distinto del $\bar{u}$ tal que $\hat{f}(u) = \hat{f}(\bar{u})$.

Considerando, pues, de nuevo el conjunto $\{u \in \mathbb{R}^m / \hat{f}(u) = \hat{f}(\bar{u}), u \neq \bar{u}\}$ y su intersección con una sucesión decreciente de bolas de radio τ_r.

No sería pues dificil demostrar que si $\hat{f}(u^{(r}) = \hat{f}(\bar{u})$ para todo $r \in \mathbb{N}$, en donde cada $u^{(r}$ perteneceria a la frontera de la correspondiente bola $B^{(r}$ entonces necesariamente habría de existir una combinación lineal no trivial entre las columnas de la matriz $\nabla \hat{f}(\bar{u})$.

En efecto, bastaría considerar la sucesión $(u^{(r} - \bar{u})/||u^{(r} - \bar{u}||$ y seleccionar de ella una subsucesión convergente a un determinado vector de módulo unitario que denominaremos ψ, lo cual siempre es posible.

Entonces, de que: $(u^{(r} - \bar{u}) = \psi . ||u^{(r} - \bar{u}|| + o(||u^{(r} - \bar{u}||)$ es fácil comprobar que tendría que ser: $\nabla \hat{f}(\bar{u}) \psi = 0$. Ahora bien, si por ejemplo $\psi_{i*} \neq 0$, podríamos repetir el razonamiento anterior con $u_{i*} = \bar{u}_{i*}$, ya que la existencia de un $\hat{\bar{\eta}}$ en ese caso, tal que su producto escala con cualquier $\partial \hat{f}/\partial u_i$, $i \neq i*$, se anulara podría extenderse inmediatamente al i*. Sin perdida de generalidad podríamos suponer por consiguiente que no existe ninguna combinación lineal entre las columnas de $\nabla \hat{f}(\bar{u})$ de lo que se seguiría que para algún r no existiria ningun $u \in B^{(r}$, $u \neq \bar{u}$, tal que $\hat{f}(u) = \hat{f}(\bar{u})$.

Además de que $\hat{\bar{x}}$ es tal que $x_i > 0$, si $i \epsilon N_2$, todavía podría seguir suponiéndose que λ_j sería arbitrario, si $j-m \epsilon N_2$.

Así podrían garantizarse conos convexos de vectores tangentes en cada punto $\hat{x}$, del tipo (104) y por lo tanto valdrían los resultados establecidos antes.

- - - - - - - - -

El resultado del Teorema (95) puede ser mejorado en el sentido de que se podría relajar las hipótesis de la diferenciabilidad con continuidad de $\hat{f}$ en $\bar{u}$ y dejarla solo en la diferenciabilidad de $\hat{f}$ en $\bar{u}$.

Por otra parte, con alguna restricción adicional se podría garantizar que $\bar{\eta}_o \neq 0$.

El siguiente corolario ilustra esta última observación:

(105) COROLARIO

En las mismas hipótesis del Teorema (95) si además supone mos que no existe un vector $\gamma \in R^{Car(N_1)}$, $\gamma \neq 0$, $\gamma \geq 0$, tal que: $\gamma . \nabla f_I(\bar{u}) = 0$, en donde:

$Car(N_1)$ = Cardinal de N_1 = n° de elementos de N_1,

$\nabla f_I(\bar{u}) = (\nabla f_i(\bar{u}))_{i \in N_1}$; N_1 = subconjunto de $\{1,2,\ldots,n\}$

tal que $i \in N_1 \Rightarrow f_i(\bar{u}) = 0$,

entonces existe un vector $\bar{\eta} \in R^n$, $\bar{\eta} \geq 0$, tal que:

(106) $\nabla f_o(\bar{u}) + \bar{\eta}\, \nabla f(\bar{u}) = 0$, y

(107) $\bar{\eta} \,.\, f(\bar{u}) = 0$.

DEMOSTRACION

Sea $\hat{\eta}^*$ tal que $\hat{\eta}^* \geq 0$, $\hat{\eta}^* \neq 0$, y

$$\eta_o^* \,\nabla f_o(\bar{u}) + \eta^* \,\nabla f(\bar{u}) = 0$$

se sigue entonces que, o bien $\eta^* = 0$, en cuyo caso $\eta_o^* > 0$, o bien $\eta^* \neq 0$.

Ahora bien, si en este segundo caso fuera $\eta_o^* = 0$ tendría mos:

$$\eta^* \nabla f(\bar{u}) = \eta_I^* \,\nabla f_I(\bar{u}) = 0,$$

con: $\eta_I^* = (\eta_i^*)_{i \in N_1}$, ya que $\eta_{II}^* = 0$, en donde:

$$\eta_{II}^* = (\eta_i^*)_{i \in N_2}.$$

Como $\eta^* \neq 0$ se seguiría que $\eta_I^* \neq 0$ lo que contradiría la hipótesis de no existencia del vector γ del enunciado.

Por tanto, aún en este segundo caso debería de ser

$\eta_o^* \neq 0$.

Para completar la demostración del corolario basta escoger, pues:

$$\bar{\eta}_i = \frac{\eta_i^*}{\eta_o^*} \; ; \qquad i = 1,2,\ldots,n$$

La condición de no existencia de un $\gamma \geq 0$, $\gamma \neq 0$ tal que $\gamma \nabla_I f(\bar{u}) = 0$ está en la línea de lo que en la Programación No Lineal se denomina una "restricción cualificada", las cuales se proponen para garantizar que $\bar{\eta}_o \neq 0$.

No obstante ninguna de las restricciones cualificadas que se consideran normalmente coincide con ésta (véase la Nota 8).

A continuación vamos a presentar una demostración de los resultados (97) y (98) del Teorema (95) que tiene interés porque se emplea en ella una metodología diferente a la demostración anterior, que no incluye la aplicación del Principio (91).

Esta nueva demostración se basa exclusivamente en el "Teorema de la Función Inversa" para funciones vectoriales de varias variables.

(108) PROPOSICION

En las mismas hipótesis del Teorema (95) se verifican (97) y (98) para algún vector $\hat{\bar{\eta}} \neq 0$.

DEMOSTRACION

Consideremos el siguiente problema ampliado:

$$\min \; x_o = f_o(u)$$

sujeto a:

$$x_i = -\,(f_i(u) + u_{m+i}^2), \qquad i=1,2,\ldots,n,$$

(109) con $u \in \Omega \quad R^m$; u_{m+i} arbitrario, $i=1,2,\ldots,n$,

y $x \in G^*$, en donde:

$$G^* = \{x \in R^n \;/\; x = 0\},$$

Denotemos:

$$\tilde{u} = (u_1, u_2, \ldots, u_{m+n})$$

y:

$$g_i(\tilde{u}) = -f_i(u) - u_{m+i}^2, \qquad i=1,2,\ldots,n$$

$$g_o(\tilde{u}) = f_o(u)$$

Entonces, si $\hat{\bar{x}} = \hat{f}(\bar{u})$ es una solución del Problema (96), $\tilde{\bar{u}}$, definido de la siguiente manera: $\tilde{\bar{u}} = (\bar{u}, \bar{u}_{m+1}, \bar{u}_{m+2}, \ldots, \bar{u}_{m+n})$,

con: $\bar{u}_{m+i} = (-f_i(\bar{u}))^{1/2}$, $i=1,2,\ldots,n$

será una solución del Problema (109).

Denotemos por $\nabla g_i(\bar{u})$ los vectores gradientes de las g_i, $i=0,1,2,\ldots,n$, en $\bar{u}$, y supongamos que son linealmente independientes, pues en otro caso se seguiría que hay una combinación lineal no trivial entre los vectores $\nabla g_i(\bar{u})$, $i=0,1,2,\ldots,n$, y por lo tanto que se verifican (97) y (98).

Se puede completar entonces, supuesto $m \geq 1$, la función vectorial $\hat{g}(u)$ con otras $(m-1)$ componentes, para obtener la función: $g : R^{m+n} \longrightarrow R^{m+n}$, tal que exista la matriz Jacobiana $(m+n) \times (m+n)$ $\nabla g(u)$, en un entorno de $\bar{u}$, y su determinante sea no nulo en dicho entorno. Basta para ello tener en cuenta que de la continuidad de $\nabla g_i(u)$, $i=0,1,2,\ldots,n$, se deduce que no existirá dependencia lineal entre los $\nabla g_i(u)$, en un entorno de $\bar{u}$, si no la había en $\bar{u}$, y que la matriz Jacobiana puede completarse, obviamente, con vectores fila:

$\nabla g_i(\tilde{u})$, $i = (m+1), (m+2), \ldots, (m+n-1)$, que cumplan las condi ciones requeridas.

(Si $m = 0$, no hay restricciones en el Problema (109) y la tesis de la Proposición (108) es evidentemente cierta sin más que aplicar el Principio (91) a los vectores límite $(x_o(\delta) - \bar{x}_o) / \delta)$.

La aplicación entonces del "Teorema de la Función Inver sa" (Véase la Nota 9), permite deducir que existirá un entorno de $\tilde{\bar{u}}$, $E_{\tilde{\bar{u}}}$, tal que:

$$\tilde{g}(E_{\bar{u}}) = \{\tilde{x} \in R^{m+n} \;/\; \tilde{x} = \tilde{g}(\tilde{u});\;\; \tilde{u} \in E_{\tilde{\bar{u}}}\},$$

es abierto.

Se sigue de ello que había un entorno de $\hat{g}(\tilde{\bar{u}})$ obtenido con puntos de la forma $\hat{g}(\tilde{u})$, para $u \in \Omega$. Es decir que $\hat{\bar{x}}$ no sería un punto perteneciente a la frontera inferior con respecto a x_o del conjunto:

$$C = \{\hat{x} \in R^{n+1} \;/\; x_i = g_i(\tilde{u});\;\; i=0,1,2,\ldots,n;\;\; u \in \Omega\},$$

lo que impide su optimalidad contra lo supuesto.

Se deduce entonces que debería existir una combinación lineal no trivial entre los vectores $\nabla g_i(\bar{u})$, $i=0,1,2,\ldots,n$, de lo que se siguen (97) y (98).

Como hemos indicado antes el Teorema (95) se puede mejo rar en el sentido de que basta con suponer la diferenciabilidad de $\hat{f}$ en $\bar{u}$ y no la diferenciabilidad con continuidad.

Vamos a demostrar, en ese sentido, un teorema clásico en

Programación no lineal el cual, mediante la inclusión de una "restricción cualificada" adicional, concluye además que $\overline{\eta}_o \neq 0$.

Veamos antes cual es en términos precisos dicha restricción cualificada.

(110) DEFINICION

Sea Ω un subconjunto abierto de $\mathbb{R}^m$ y $f(u)$ una función vectorial n-dimensional definida en Ω.

Consideremos el conjunto:

(111) $H = \{u \in \mathbb{R}^m \ / \ u \in \Omega; \ \ f(u) \leq 0\}$.

Entonces, se dice que f satisface la "restricción cualificada" de Kuhn-Tucker en el punto $\overline{u}$ del conjunto H si $f(u)$ es diferenciable en $\overline{u}$ y si para todo vector $\psi \in \mathbb{R}^m$ tal que:

$$\nabla f_I(\overline{u}) \cdot \psi \leq 0$$

en donde $f_I = (f_i)_{i \in N_1}$, con $i \in N_1 \iff f_i(\overline{u}) = 0$, existe una función m-vectorial $u(\delta)$ definida en: $0 \leq \delta \leq 1$, tal que: $u(0) = \overline{u}$; $u(\delta) \in H$, para $0 \leq \delta \leq 1$, y $u(\delta)$ es diferenciable en $\delta = 0$ verificandose:

$$\frac{d\ u(0)}{d\delta} = a\ \psi,$$

para algún $a > 0$.

Estamos ya en condiciones de probar el siguiente teorema:

(112) TEOREMA (de Kuhn-Tucker)

Consideremos el Problema (96) y sea $\hat{\overline{x}} = \hat{f}(\overline{u})$ una solución.

Entonces, si $f_o(u)$ es diferenciable en $\overline{u}$ y se verifica la restricción cualificada de Kuhn-Tucker para $f(u)$ en $\bar{u}$ se puede concluir que existe un vector $\hat{\eta} \geq 0$ que verifica (106) y (107).

DEMOSTRACION.

Si $\bar{u}$ es una solución del Problema (96) tendremos:

$$f_o(\bar{u}) \leq f_o(u)$$

para todo $u \in H$.

Ahora, ya que se verifica la restricción cualificada de Kuhn-Tucker, para cada vector ψ tal que $\nabla f_I(\bar{u}).\psi \leq 0$ existe una función $u(\delta)$ tal que $du(0)/d\delta = a.\psi$, con $a > 0$.

Consideremos entonces los límites, cuando $\delta \downarrow 0$, en las expresiones:

$$\frac{f_o(u(\delta)) - f_o(\bar{u})}{\delta} \quad (\geq 0, \ \forall\delta),$$

los cuales existirán por ser f_o diferenciable en $\bar{u}$, para cada ψ.

Se obtiene entonces, por la regla de la cadena, que:

$$\nabla f_I(\bar{u}).\psi \leq 0 \Longrightarrow \nabla f_o(\bar{u}).(a\psi) \geq 0 \Longrightarrow \nabla f_o(\bar{u}).\psi \geq 0$$

El anterior resultado puede expresarse también diciendo que la solución del siguiente problema de Programación Lineal:

$$\max_{\psi} \quad -\nabla f_o(\bar{u}) . \psi$$

sujeto a:

(113) $$\nabla f_I(\bar{u}).\psi \leq 0$$

ψ arbitrario,

se obtiene para $\psi = 0$.

Así, el Problema (113) tiene solución finita de lo que se deduce, teniendo en cuenta el Teorema (50) de la Dualidad en

Programación Lineal, que el problema dual del (113) dado por:

(114) $\min\ \gamma.0 \quad (=0)$

sujeto a:

$$\gamma\ \nabla f_I(\bar{u}) = -\nabla f_o(\bar{u})$$
$$\gamma \geq 0,$$

tiene solución.

Esto equivale a afirmar que hay alguna solución factible al conjunto de restricciones del Problema (114). Sea $\bar{\gamma}$ una de tales soluciones. Entonces, la conclusión del Teorema sigue pues sin mas que definir el vector $\bar{\eta}$ de la siguiente manera:

$$\bar{\eta}_I = \bar{\gamma} \geq 0,$$
$$\bar{\eta}_{II} = 0,$$

de donde $\bar{\eta}_I = (\bar{\eta}_i)_{i \in N_1}$.

La demostración del Teorema, que presenta algunas variaciones respecto a las demostraciones usuales en algunos textos (véase la Nota 10) permite hacer algunas observaciones interesantes. Una de ellas es la de que si la restricción cualificada de Kuhn-Tucker para la subfunción vectorial f_I se pudiera aplicar a cualquier otra subfunción vectorial de la f, el resultado seguiría siendo válido con las modificaciones consiguientes en la definición del vector $\bar{\eta}$.

En cuanto a la relación entre la restricción cualificada de Kuhn-Tucker y la condición de no existencia de un $\gamma \geq 0$, $\gamma \neq 0$, tal que: $\gamma\ \nabla f_I(\bar{u}) = 0$, vamos a ver que la primera no implica la segunda.

En efecto, consideremos por ejemplo el siguiente problema:

(115) $$\min \; f_o(u) = u_1 + u_2,$$

sujeto a:

$$f_1(u) = u_1 - u_2 \leq 0$$
$$f_2(u) = -u_1 + u_2 \leq 0$$
$$f_3(u) = -u_1 \leq 0, \qquad \Omega = R^2,$$

con solución: $\bar{u} = \binom{o}{o}$.

Los vectores gradientes en $\bar{u}$ son:

$$\nabla f_1(\bar{u}) = (1,-1)$$
$$\nabla f_2(\bar{u}) = (-1,1)$$
$$\nabla f_3(\bar{u}) = (-1,0)$$

y:

$$\nabla f_I(\bar{u}) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ -1 & 0 \end{pmatrix}$$

En este caso, mientras que la restricción cualificada de Kuhn-Tucker se cumple, es decir que se puede definir una función $u(\delta)$ tal que $f(u(\delta)) \leq 0$, $0 \leq \delta \leq 1$, y $\frac{du(0)}{d\delta} = a\psi$, para todo vector $\psi = \binom{\psi_1}{\psi_2}$ tal que:

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

es decir $\psi_1 = \psi_2 \geq 0$, existe un vector $\gamma = (1,1,0)$, tal que:

$$\gamma \; \nabla f_I(\bar{u}) = (1,1,0) \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ -1 & 0 \end{pmatrix} = (0,0)$$

Por supuesto, aunque exista tal vector, el cumplimiento de la restricción cualificada permite garantizar la existencia de un vector $\bar{\eta} = (\bar{\eta}_1, \bar{\eta}_2, \bar{\eta}_3)$, tal que:

$$(1,1) + (\bar{\eta}_1,\bar{\eta}_2,\bar{\eta}_3)\begin{pmatrix} 1 & -1 \\ -1 & 1 \\ -1 & 0 \end{pmatrix} = (0,0)$$

(Este sería el dado por : $\bar{\eta}_1 = 2;\ \bar{\eta}_2 = 1;\ \bar{\eta}_3 = 2$).

Del ejemplo anterior se sigue que no cabe esperar que, por el hecho de que exista el $\gamma \geq 0$, $\gamma \neq 0$ tal que $\gamma \nabla f_I(\bar{u}) = 0$, haya de ser $\bar{\eta}_o = 0$ en el Teorema (95). Queda pues bien claro que la condición del Corolario (105) es una condición suficiente y no necesaría, para garantizar la positividad del $\bar{\eta}_o$.

Lo mismo se puede decir de la restricción cualificada de Kuhn-Tucker. El siguiente ejemplo ilustra esta afirmación:

$$\min\ f_o(u) = u_1^2 ,$$

sujeto a:

$$f_1(u) = u_1^2 + u_2^2 \leq 0, \qquad \Omega = \mathbb{R}^2,$$

con solución $\bar{u} = \binom{o}{o}$.

En este problema no se verifica la restricción calificada porque mientras que el conjunto de vectores ψ viene definido por:

$$\nabla f_I(\bar{u}) \cdot \psi = (0,0) \binom{\psi_1}{\psi_2} \leq 0$$

el conjunto H se reduce al único punto $\binom{o}{o}$.

Por otra parte, de que $\nabla f_I(\bar{u}) = (0,0)$ se deduce que existirá un $\gamma > 0$ tal que $\gamma\ \nabla f_I(\bar{u}) = (0,0)$. Así, tampoco se cumple la restricción del Corolario (105).

Sin embargo se puede seleccionar un vector $\bar{\eta} \geq 0$ tal que:

$$\bar{\eta}\ \nabla f_I(\bar{u}) = -\nabla f_o(\bar{u}),$$

pues basta tomar $\bar{\eta} = 0$.

Con estos ejemplos terminamos la discusión del caso estático diferenciable. En la siguiente sección iniciaremos el estudio del caso dinámico diferenciable.

2. EL CASO DINAMICO DIFERENCIABLE

En esta sección nos vamos a ocupar de la aplicación del Principio (91) para obtener las condiciones necesarias de optimalidad en el Problema (56) que volvemos a establecer aquí:

$$(116) \qquad \min x_o(t_1) = \int_{t_o}^{t_1} f_o(x,u,t)dt$$

sujeto a:

$$x(t) = \int_{t_o}^{t} f(x,u,s)ds + x(t_o); \quad \forall t \in [t_o,t_1]$$

$$x(t) \in R^n$$

$$u(t) \in \Omega \quad R^n, \quad \forall t \in [t_o,t_1]; \quad u(.) \in \Delta$$

$$x(t_o) \in G_o \subset R^n$$

$$x(t_1) \in G_1 \subset R^n$$

en donde Δ es una determinada clase de funciones.

Las integrales se entenderan en el sentido de Lebesgue y todas las operaciones matemáticas planteadas se suponen implícitamente realizables.

En principio, el conjunto C a considerar en este problema es el siguiente:

$$(117) \quad C = \{\hat{x} \in R^{n+1} \;/\; \hat{x} = \int_{t_o}^{t_1} \hat{f}(x(t),u(t),t)dt + \begin{bmatrix} 0 \\ x(t_o) \end{bmatrix};$$

$$x(t) = \int_{t_o}^{t} f(x(s),u(s),s)ds + x(t_o), \quad u(t) \in \Omega, \quad \forall t \in [t_o,t_1];$$

$$u(.) \in \Delta, \quad t_o,t_1 \in R, \quad t_1 \geq t_o, \quad x(t_o) \in G_o\}$$

Se trata pues de caracterizar los puntos de la frontera inferior del anterior conjunto respecto a la coordenada x_o para determinar luego el óptimo seleccionando aquellos que verifican las condición $x(t_1) \in G_1$.

Una primera observación a hacer sobre el Problema (116) es la de que mediante la introducción de la variable $x_{n+1}(t)$ definida en $[t_o, t_1]$ de la siguiente manera:

$$x_{n+1}(t) = \int_{t_o}^{t} ds$$

el problema se puede volver a plantear haciendo desaparecer el argumento t de la función $\hat{f}$, cuyos únicos argumentos serían ahora el vector $(x_o, x_1, \ldots, x_n, x_{n+1})$, y el vector $u = (u_1, u_2, \ldots, u_m)'$.

Cuando el problema se plantea en términos en los que la variable t no aparece explícitamente se convierte en uno "autónomo" es decir invariante a las traslaciones en el tiempo en los argumentos de las funciones $x_i(t)$ y $u_j(t)$.

Según acabamos de mostrar, pues, basta discutir el caso autónomo del cual el general se seguiría como corolario.

Otra observación a hacer se refiere a la arbitrariedad de los dos parámetros t_o y t_1. En principio estos dos números incrementan el número de incógnitas del problema. Ahora bien, puesto que suponemos que ninguno de los conjuntos Ω, G_o y G_1 dependen de t, en el caso de problemas autónomos podemos suponer que t_o es fijo. En cuanto a t_1 puede considerarse una incógnita más del problema. Esto no excluye aquellos casos en los que t_1 es fijo. En efecto, bastará añadir una nueva variable más al problema definida como idénticamente igual al tiempo, o sea igual que antes: $x_{n+1}(t) = t$, e incrementar en una dimensión más los conjuntos G_o y G_1, en el primer caso para recoger que $x_{n+1}(t_o) = t_o$ y en el segundo que $x_{n+1}(t_1)$ debe de ser un determinado valor fijo.

Otra situación que en principio no quedaría recogida en el anterior esquema pero para la cual se generan con facilidad condiciones necesarias de optimalidad como un corolario de la situación que investigaremos aquí es la que se refiere al caso en que alguno de los t_o ó t_1 ó ambos, sean infinitos. En ese caso el Problema (116) se plantearía directamente con esas condiciones en el intervalo (t_o, t_1) que correspondiera y la misma metodología que emplearemos aquí permitiría obtener las condiciones necesarias de optimalidad. En el caso de que alguno de los límites de integración fuera infinito habría que asegurarse de la existencia de las correspondientes integrales mediante unas adecuadas restricciones en $\hat{f}$ y Δ. Por lo que respecta al conjunto Ω de valores de las funciones de control este podrá ser cualquier subconjunto arbitrario de R^m, no necesariamente convexo ni compacto. Pero lo que si supondremos es que es un conjunto independiente de t.

Los otros dos conjuntos G_o y G_1 que restringen respectivamente los valores de la variable $x(t)$ en los instantes t_o y t_1 se supone que verifican unas ciertas condiciones de regularidad a efectos de poder caracterizar los valores óptimos. Estas condiciones se concretan en que se suponen "variedades diferenciales" que pueden definirse mediante un sistema de $n-\Gamma_i$, $i=0,1$, ecuaciones en R^n, siendo Γ_i la dimensión de cada variedad, definidas por funciones diferenciables con continuidad en un abierto que contenga a la variedad y matriz Jacobiana de determinante no nulo en cada punto de la variedad (Ver Nota 11).

En las condiciones anteriores la trayectoria óptima entre los puntos $x(t_o)$ y $x(t_1)$ pertenecientes a las variedades G_o y G_1 se caracterizaría por verificar unas determinadas "condiciones de transversalidad" en cada uno de esos puntos. Esto es, con más precisión, si denominamos T_o y T_1 a los planos tangentes a las variedades G_o y G_1 en $x(t_o)$ y $x(t_1)$ de dimensiones respectivas Γ_o y Γ_1, es decir a la intersección de los hiperplanos tan gentes a las hipersuperficies dadas por cada una de las ecuaciones, en cada caso, las condiciones de transversalidad se traducen en que una determinada función vectorial temporal: $\eta(t) = (\eta_1(t),\eta_2(t),\ldots,\eta_n(t))$, definida en $t_o \le t \le t_1$, que juega un papel similar a los multiplicadores en el caso discreto y cuya caracterización y obtención se estudiaran en detalle más adelante, sea "ortogonal", para $t = t_o$ y $t = t_1$, al correspondiente plano tangente. Esto es, para todo vector $\theta = (\theta_1,\theta_2,\ldots,\theta_n)$ perteneciente a T_o habría de ser $\eta(t_o).\theta = 0$, y lo mismo para T_1.

Nosotros no nos ocuparemos de este problema y resolveremos sólo el caso relativo a extremos dados por sendos puntos $x^{(o}$ y $x^{(1}$ ya que nuestro objetivo no es el análisis exhaustivo de todos los casos sino la investigación de como unos pocos principios fundamentales permiten deducir los resultados que estan en la base de la resolución de los problemas de optimización.

El problema a estudiar será por lo tanto la versión autónoma del (116) con $G_o = \{x^{(o}\}$ y $G_1 = \{x^{(1}\}$.

Como clase Δ de funciones, la cual denominaremos "clase de los controles admisibles", escogeremos la caracterizada por las tres siguientes propiedades:

(118) (i) Todos los controles $u(t)$, $t_o \le t \le t_1$, se supondrán medibles y acotados.

(ii) Si $u(t)$, $t_o \le t \le t_1$, es un control admisible, ω es un punto arbitrario de Ω y t' y t'' son números tales que $t_o \le t' < t'' \le t_1$, el control $u^{\#}(t)$, $t_o \le t \le t_1$, definido por la fórmula:

$$u^{\#}(t) = \begin{cases} u(t) & \text{si} \quad t_o \leq t \leq t' \\ \omega & \text{si} \quad t' < t \leq t'' \\ u(t) & \text{si} \quad t'' < t \leq t_1 \end{cases}$$

es también admisible (si alguno de los subintervalos fuera vacío no se aplicaría la definición correspondiente).

(iii) Si el intervalo $t_o \leq t \leq t_1$ se particiona por medio de puntos de subdivisión en un número finito de subíntervalos en cada uno de los cuales el control $u(t)$ es admisble entonces este control es también admisible en el intervalo completo $t_o \leq t \leq t_1$. Un control obtenido a partir de un control admisible $u(t)$, $t_o \leq t \leq t_1$, por una traslación en el tiempo, es decir el control: $u^{\#}(t) = u(t-\alpha)$, $t_o+\alpha \leq t \leq t_1+\alpha$, es también admisible.

A las anteriores propiedades ya nos hemos referido antes cuando analizamos el caso dinámico continuo convexo.

(119) Estableceremos a continuación las propiedades a exigir a la función $\hat{f}$. Estas serán las de que las funciones $f_o(x,u)$, $f_1(x,u),\dots,f_n(x,u)$, sean continuas con respecto a sus dos argumentos x y u en el espacio producto $\mathbb{R}^n x\ \bar{\Omega}$ y además posean derivadas parciales continuas con respecto a las x_i, $i=1,2,\dots,n$, en $\mathbb{R}^n x\bar{\Omega}$.

El que se admita la existencia y continuidad de las derivadas parciales permitiría, ademas de garantizar la existencia y unicidad de las soluciones del sistema de ecuaciones diferenciales con valor inicial dado permitiría caracterizar el intervalo $t_o \leq t \leq t_1$ al vector de multiplicadores $\hat{\eta}(t)$, el cual, con solo la aplicación de resultados del tipo (91), quedaría solo determinado en términos de los vectores tangenciales en cada punto de la trayectoria óptima.

No entramos aquí en la discusión de otras condiciones que podrían proponerse para garantizar la existencia y unicidad de soluciones de la ecuación integral de (116). (Ver Nota 12). No hay que perder de vista que el problema de la determinación de condiciones necesarias en la teoría de la optimización arranca de la suposición de la existencia de una solución que cumple las condiciones requeridas, y no antes. Así, las condiciones expuestas cumplen el objetivo de limitar la posible gama de soluciones alternativas adicionales mas que con el fin de garantizar la existencia de alguna solución, además de permitir las manipulaciones que se precisan para establecer las condiciones buscadas.

Vamos a obtener a continuación una condición necesaria de optimalidad para el Problema (116) en las condiciones dadas. Nuestro objetivo no es tanto el propio resultado, que no es otro que el Principio de Máximo de Pontryaguin, sino el mostrar de que manera se puede obtener a partir del Principio (91).

Como primer paso vamos a establecer aquí un resultado que ya se comentó antes al estudiar el caso dinámico continuo convexo.

(120) LEMA (Principio de Evolución Optima)

Supongamos que $\hat{\bar{x}}(t) = (x_o(t), x_1(t), \ldots, x_n(t))'$, con $t_o \leq t \leq t_1$, es una trayectoria óptima para el Problema (116) en su versión autónoma, con $G_o = \{x^{(o)}\}$, y las condiciones (118) y (119). Entonces, para cada $t \in [t_o, t_1]$ el punto $\hat{\bar{x}}(t)$ pertenece a la frontera inferior, respecto a su primera coordenada, del conjunto C dado por (117).

DEMOSTRACION.

Basta reproducir la discusión de la primera parte de la demostración del Teorema (79). El que no se pueda asegurar ahora que

C sea convexo y cerrado es irrelevante en la obtención del resultado propuesto.

Recordamos ahora que el punto t en el intervalo (a,b) se denomina un "punto regular" de la función medible $u(t)$ si se verifica la relación:

$$\lim_{m(I) \downarrow 0} \frac{m(u^{-1}(E) \quad I)}{m(I)} = 1,$$

para todo entorno $E \subset \Omega$ de $u(t)$.

En la expresión anterior $u^{-1}(E)$ denota el conjunto de puntos s del intervalo (a,b) tales que $u(s) \in E$, mientras que I denota un intervalo arbitrario que contenga a t y $m(.)$ representa la medida de Lebesgue del conjunto correspondiente.

Se puede demostrar (véase la Nota 13) que el conjunto de los puntos regulares de cualquier función medible arbitraria en el intervalo (a,b) tiene medida igual a $(b-a)$, es decir, que "casi todos" los puntos del intervalo (a,b) serán puntos regulares para la función $u(t)$.

Además, si $g(s,u(s))$ es una función real y continua de las dos variables $s \in (a,b)$ y $u \in \Omega$, y $u(s)$, $a < s < b$, es una función medible y acotada con recorrido en Ω, entonces, si z es un punto regular para $u(s)$ se verifica la relación:

$$\int_{z+\theta_1\delta}^{z+\theta_2\delta} g(s,u(s))ds = \delta(\theta_2-\theta_1)\ g(z,u(z)) + o(\delta) \tag{121}$$

en donde θ_1 y θ_2 son números reales arbitrarios, δ es un número positivo suficientemente pequeño (para que quede garantizada la definición del integrando), y $o(\delta)$ es un infinitésimo de un orden mayor que δ, es decir:

$$\lim_{\delta \downarrow 0} \frac{o(\delta)}{\delta} = 0$$

La integral de (121) se entiende en el sentido de Lebesgue.

Supongamos ahora que $\hat{\bar{x}}(s)$, para $t_o \leq s \leq t_1$, es una trayectoria óptima para el Problema (116) en las condiciones dadas, con control correspondiente $\bar{u}(s)$, y sea $\hat{\bar{x}}(t)$ un punto de dicha trayectoria.

Vamos a construir progresivamente un cono convexo asociado a dicho punto el cual desempeñará un papel fundamental en la deducción de las condiciones necesarias de optimalidad.

Sean $\lambda_1, \lambda_2, \ldots, \lambda_q$, una colección de números no negativos y $\delta > 0$.

Definamos la función de control constante a trozos $u_\delta(s)$ en el intervalo $0 \leq s \leq [\sum_{i=1}^{q} \lambda_i]\delta$, de la siguiente manera:

$$u_\delta(s) = u^{(i} \in \Omega, \quad \text{para} \quad [\sum_{i=1}^{i-1} \lambda_j]\delta \leq s < [\sum_{j=1}^{i} \lambda_j]\delta$$

$$\text{si} \quad i = 2,3,\ldots,(q-1).$$

$$= u^{(1} \in \Omega, \quad \text{para} \quad 0 \leq s \leq \lambda_1\delta$$

$$= u^{(q} \in \Omega, \quad \text{para} \quad [\sum_{i=1}^{q-1} \lambda_i]\delta \leq s \leq [\sum_{i=1}^{q} \lambda_i]\delta$$

(Obsérvese que definimos la función continua por la izquierda al contrario de lo que hemos hecho en ocasiones anteriores. De cualquier manera la definición de la función de control en los puntos de discontinuidad es irrelevante).

Denotemos por $\hat{x}^*_\delta(s)$ la trayectoria cuando se aplica la función de control anterior con estado inicial $\hat{\bar{x}}(t)$, para $0 \leq s \leq [\sum_{i=1}^{q} \lambda_i]\delta$.

Tenemos:

$$\hat{x}^*_\delta([\sum_{i=1}^{q} \lambda_i]\delta) - \hat{\bar{x}}(t) = \int_0^{[\sum_{i=1}^{q} \lambda_i]\delta} \hat{f}(\hat{x}^*_\delta(s),u_\delta(s))ds =$$

$$= \int_0^{\lambda_1\delta} f(x^*(s),u_\delta(s))ds + \int_{\lambda_1\delta}^{(\lambda_1+\lambda_2)\delta} \hat{f}(\hat{x}^*_\delta(s),u_\delta(s))ds +\cdots$$

$$+\cdots+ \int_{[\sum_{i=1}^{q-1} \lambda_i]\delta}^{[\sum_{i=1}^{q} \lambda_i]\delta} \hat{f}(\hat{x}^*_\delta(s),u_\delta(s))ds$$

Ahora bien, podemos aplicar la fórmula (121) para evaluar cada una de las integrales anteriores ya que las subfunciones de la $u_\delta(s)$ definidas en los intervalos: $[\sum_{j=1}^{i-1} \lambda_j]\delta \leq s < [\sum_{j=1}^{i} \lambda_j]\delta$, se pueden prolongar por la izquierda haciendolas igual al correspondiente $u^{(i}$ para $s < [\sum_{j=1}^{i-1} \lambda_j]\delta$, de manera que los puntos: $s = [\sum_{j=1}^{i-1} \lambda_j]\delta$, para $i=2,3,\ldots,q$, y $s=0$, resultarían ser puntos regulares de las funciones ampliadas, y además la función $\hat{x}^*_\delta(s)$ es continua en s y $\hat{f}$ es continua en $\hat{x}$ y u.

Aplicando entonces la fórmula (121) con $z = \sum_{j=1}^{i-1} \lambda_j$, $\theta_1 = 0$, y $\theta_2 = \lambda_i$ se obtiene:

(122) $$\hat{x}^*_\delta([\sum_{i=1}^{q} \lambda_i]\delta) - \hat{\bar{x}}(t) = \lambda_1\delta\ \hat{f}(\hat{\bar{x}}(t),\ u^{(1}) + \sum_{i=2}^{q} \lambda_i\delta\ \hat{f}\ (\hat{x}^*_\delta([\sum_{j=1}^{i-1} \lambda_j]\delta),\ u^{(i}) + o(\delta)$$

Evidentemente $\hat{x}^*_g([\sum_{i=1}^{q} \lambda_i]\delta) \in C$, como se sigue sin más que adjuntar el control $u_\delta(s)$ al control que traslada $\begin{bmatrix} 0 \\ x^0 \end{bmatrix}$ a $\hat{\bar{x}}(t)$ y recordar las propiedades de la clase Δ.

Dividiendo por δ y haciendo $\delta \downarrow 0$ se obtiene:

(123) $$\lim_{\delta \downarrow 0} \frac{x^*_\delta([\sum_{i=1}^{q} \lambda_i]\delta) - \hat{\bar{x}}(t)}{\delta} =$$

$$= \lambda_1 \hat{f}(\hat{\bar{x}}(t), u^{(1}) + \sum_{i=2}^{q} \lambda_i \hat{f}(\lim_{\delta \downarrow 0} \hat{x}^*_\delta([\sum_{j=1}^{i-1} \lambda_j]\delta), u^{(i}) =$$

$$= \sum_{i=1}^{q} \lambda_i \hat{f}(\hat{\bar{x}}(t), u^{(i}),$$

ya que $\hat{x}^*_\delta(0) = \hat{\bar{x}}(t)$.

Es fácil comprobar que el conjunto de vectores obtenido tomando límites del tipo (123) para $\hat{\bar{x}}(t)$ fijo es un conjunto convexo (en realidad un cono convexo) ya que si $\sum_{i=1}^{q_1} \lambda_i^{(1} \hat{f}(\hat{\bar{x}}(t), u^{(i1})$ y $\sum_{i=1}^{q_2} \lambda_i^{(2} \hat{f}(\hat{\bar{x}}(t), u^{(i2})$, son dos de tales vectores, el vector:

$$\lambda \sum_{i=1}^{q_1} \lambda_i^{(1} \hat{f}(\hat{\bar{x}}(t), u^{(i1}) + (1-\lambda) \sum_{i=1}^{q_2} \lambda_i^{(2} \hat{f}(\hat{\bar{x}}(t), u^{(i2}),$$

para cualquier λ, $0 < \lambda < 1$, es evidentemente un vector del tipo anterior sin más que definir: $\lambda_i = \lambda\lambda_i^{(1}$ y $u^{(i} = u^{(i1}$, para $i=1,2,\dots,q_1$, y $\lambda_i = (1-\lambda)\lambda_i^{(2}$ y $u^{(i} = u^{(i2}$, para $i = (q_1+1),(q_1+2),\dots,(q_1+q_2)$.

Además el 0 es un vector límite sin más que tomar $\Sigma\lambda_i = 0$.

Consideramos ahora el conjunto acotado no vacio siguiente:

$$A^{(k} = \{u \in \Omega \subset \mathbb{R}^m \;/\; \sum_{i=1}^{m} u_i^2 \leq k\}$$

para $k \in \mathbb{N}$ y $k \geq \sum_{i=1}^{m} \bar{u}_i^2(t_o)$, y sea $D^{(k}$ un subconjunto denso y numerable en $A^{(k}$.

Sea, por otra parte, $D^{(kh}$, para $h \in \mathbb{N}$, una sucesión creciente de conjuntos finitos tales que:

$$\bigcup_{h \in \mathbb{N}} D^{(kh} = D^{(k}$$

y de cada $D^{(kh}$ contenga h elementos.

Nuestro objetivo es preparar la aplicación del Principio (91) problema (116) supuestas admitidas las hipótesis que ya hemos indicado antes sobre Δ, G_o y G_1. Pero ello deberemos elegir de una manera conveniente un conjunto que corresponda al C al que se refiere al Teorema (91).

Ahora, como hemos mostrado antes, se puede definir para cada t, con $t_o<t<t_1$, un haz de vectores tangentes en cada punto $\hat{x}(t)$ de la trayectoria óptima si se prolongara la función de control $\bar{u}(t)$ durante un lapso de tiempo $(\Sigma \lambda_i)\delta$, por medio de una función constante a trozos.

Tomando como referencia el resultado anterior, los conjuntos de accesibilidad que vamos a definir no van a ser otros que conjuntos de puntos de $\mathbb{R}^{n+1}$ que se pueden obtener a partir del $\hat{x}(t)$ por medio de funciones constantes a trozos, con la restricción de que los valores de las constantes $u^{(i}$ se escojan en uno de los conjuntos $D^{(kh}$, y de los sucesivos puntos que se pudieran ir obteniendo a partir de los sucesivos puntos terminales, según dos opciones. En la primera cada uno de los subintervalos de definición de la función de control correspondientes a cada $u^{(i}$ se -

iría ampliando, manteniéndose acotada la suma de amplitudes del conjunto de subintervalos, y en la segunda a cada colección de subintervalos de definición correspondientes a cada $u^{(i}$ se iría adjuntando por la derecha nuevas prolongaciones, con la restricción de que se mantuviera acotada la suma de amplitudes de todos los subintervalos.

Esta doble consideración de dos diferentes tipos de conjuntos de accesibilidad tiene como fin el de ayudar a rematar la construcción de conos convexos de vectores tangentes en los puntos de las fronteras de ambos, en las que se crean algunas dificultades adicionales para garantizar la definición de perturbaciones por medio de las cuales se pueda obtener dicho haz de vectores tangentes.

Denotemos, pues, por a_{10}, a_{11}, $a_{12}\ldots$, a_{1h}, con $a_{1j} \geq \geq a_{1(j-1)}$, para todo j, y $a_{10} = 0$, los puntos de subdivisión correspondientes a los valores $u^{(j}$, $j = 1, 2\ldots, h$, en donde, sin perdida de generalidad, podemos considerar asociado cada $u^{(j}$ a cada subintervalo $(a_{1(j-1)}, a_{1j})$, ya que esto no afectaría al vector tangente obtenido, y denotemos por $u_{a1}(s)$ a la función de control definida igual a los $u^{(j}$ en cada uno de los subintervalos indicados, la cual trasladaría el punto incial $\hat{x}(t)$ a un nuevo punto $\hat{x}^{(1*}(a_{1h})$. Estos puntos de subdivisión a_{1j} no tendrían por qué ser necesariamente distintos.

Sean, por otra parte, b_{20}, b_{21}, $b_{22}\ldots$, b_{2h}, tales que $b_{20}=0$ y $b_{1j} \geq 0$, para todo j, y definamos:

$$a_{2j} = a_{1j} + \sum_{r=1}^{j} b_{2r}$$

para todo j.

Entonces se verificaría que $a_{2j} \geq a_{2(j-1)}$, lo que permitiría definir consistentemente una nueva función de control, que denominasenos $u_{a2}(s)$, haciéndola igual a $u^{(j}$ en los intervalos $(a_{2(j-1)}, a_{2j})$, $j = 1,2\ldots, h$, la cual trasladaría el punto $\hat{x}^{(1*}(a_{1h})$ al $\hat{x}^{(2*}(a_{2h})$.

El proceso podría continuar definiendo un nuevo conjunto, $b_{30}, b_{31}, b_{32}, \ldots, b_{3h}$, con $b_{30} = 0$ y $b_{3j} \geq 0$, para todo j, y una nueva colección:

$$a_{3j} = a_{2j} + \sum_{r=1}^{j} b_{3r}$$

$j = 1,2,\ldots, h$, mediante la cual se podría definir una función de control $u_{a3}(s)$, igual a $u^{(j}$ en el intervalo $(a_{3(j-1)}, a_{3j})$, que trasladaría el punto $\hat{x}^{(2*}(a_{2h})$ al $\hat{x}^{(3*}(a_{3h})$, y así sucesivamente.

Si p fuera el número de veces que se repitiera el proceso supondremos además que se verifica $a_{ph} < \tau$, en donde $\tau > 0$ es una constante prefijada de antemano.

Además supondremos también que el punto terminal del proceso, $\hat{x}^{(p*}(a_{ph})$, y todos los intermedios de la trayectoria se mantienen en el interior de una bola de centro en el punto $\hat{\bar{x}}(t)$ y radio $\sigma^{1/2}$, menor que 1, dado, es decir, si denotamos la trayectoria por $\hat{x}^{(p*}(s)$ para $s \in [0, a_{ph}]$, sería:

$$\sum_{i=0}^{n} (x_i^{(p*}(s) - \bar{x}_i(t))^2 < \sigma < 1$$

para todo s.

Vamos a establecer, pues, de una manera precisa, el conjunto al que nos estamos refiriendo, el cual denotaremos por $C_t^{(kh}$, para $k, h \in \mathbb{N}$

(124) $C_t^{(kh} = \{\hat{x}^* = (x_o^*, x_1^*, x_2^*, \ldots, x_n^*)' \in \mathbb{R}^{n+1}$/ Existen

$a_{10}, a_{12}, \ldots, a_{1h}, a_{i0}$ y b_{ij}, para $i = 1, 2, \ldots, p$ y

$j = 0, 1, 2, \ldots, h$, de manera que: $a_{10} = 0$, $b_{1j}=0, a_{i0}= b_{i0}= 0$

$a_{1j} \geq a_{1(j-1)}$ y $b_{ij} \geq 0$, para todo i y j, y existe además una función de control $u_{ap}(s)$, definida en el intervalo:

$0 \leq s \leq a_{1h} + \sum_{i=1}^{p} \sum_{j=1}^{h} b_{ij}$ de la siguiente manera:

$u_{ap}(s) = u^{(j}$, para $a_{p(j-1)} \leq s \leq a_{pj}$, $j = 1, 2, \ldots, h$, en donde:

$a_{ij} = a_{i(j-1)} + \sum_{r=1}^{j} b_{ir}$, para $i = 1, 2, \ldots, p$; y $j = 1, 2, \ldots, h$,

siendo $u^{(j} \in D^{(kh}$ para todo j, y se verifica:

$$\hat{x}^* = \hat{x}^{(p*}(a_{ph}) = \hat{\bar{x}}(t) = \int_o^{a_{ph}} \hat{f}(\hat{x}^{(p*}(s), u_{ap}(s))\, ds$$

con:

$$\hat{x}^{(p*}(s) = \hat{\bar{x}}(t) + \int_o^{s} \hat{f}(\hat{x}^{(p*}(z), u_{ap}(z))\, dz$$

Para todo s tal que $0 \leq s \leq a_{ph}$

Además existen $\tau > 0$ y $\sigma > 0$, con $\sigma < 1$, tales que $a_{ph} < \tau$ y:

$$\sum_{i=o}^{n} (x_i^{(p*}(s) - \bar{x}_i(t))^2 < \sigma, \text{ para todo } s \in [0, a_{ph}] \}$$

(Nótese que por simplificación de la nomenclatura no mencionamos explícitamente en dependencia del conjunto $C_t^{(kh}$ de τ y de σ, y que la ambiguedad que se produciría por la doble definición de las funciones de control en los puntos de subdivisión a_{ij} sería irrelevante ya que el conjunto de todos ellos sería de medida de Lebesgue nula).

Antes de pasar a definir el segundo conjunto de accesibilidad que nos servirá de apoyo para la aplicación del Principio (91) a nuestro problema vamos a demostrar a continuación un lema previo que prepara dicha aplicación.

(125) LEMA

Si $\hat{\bar{x}}(t)$, obtenida para el control $\bar{u}(t)$, con $t_o \leq t \leq t_1$, es una trayectoria óptima para el Problema (116) en el caso autónomo, con $G_o = \{x_o\}$, las condiciones (118) y $\hat{f}$ es continua y con derivadas parciales continuas con respecto a las x_i en $\mathbb{R}^{n+1} \times \bar{\Omega}$, entonces $\hat{\bar{x}}(t)$ es, para todo t tal que $t_o < t < t_1$, un punto perteneciente a la frontera inferior respecto a su primera coordenada del conjunto $\bar{C}_t^{(kh}$, clausura del $C_t^{(kh}$ dado por (124), con h y $k \in \mathbb{N}$

DEMOSTRACION

Supongamos que para algún t, con $t_o < t < t_1$, existiera un punto $\hat{x}^{\#} \in \bar{C}_t^{(kh}$, con $k \geq \sum_{i=1}^{n} \bar{u}_1^{\,2}(t_o)$ tal que $x^{\#} = \bar{x}(t)$ y $x_o^{\#} < \bar{x}_o(t)$.

Consideremos el conjunto

(126) $\{(\hat{x}, u) \in \mathbb{R}^{n+1} \times \bar{\Omega} \subset \mathbb{R}^{n+m+1} \,/\, \sum_{i=o}^{n} (x_i - \bar{x}_i(t))^2 \leq 1;$
$\sum_{i=1}^{m} u_i^2 \leq k\}$

que es compacto en $\mathbb{R}^{n+m+1}$, ya que es cerrado y acotado.

Entonces, puesto que $\hat{f}(\hat{x}, u)$ es continua en $\mathbb{R}^{n+1} \times \bar{\Omega}$, la función: $\max\{[f_i(\hat{x}, u)], /0 \leq i \leq m\}$ alcanzará su máximo, que denominaremos M_k, en dicho conjunto (126).

Por otra parte, ya que $\hat{x}^{\#} \in \bar{C}_t^{(kh}$, existirá una sucesión - $\hat{x}^{(r}$, $r \in \mathbb{N}$, con $\hat{x}^{(r} \in C_t^{(kh}$, por todo r, tal que: $\hat{x}^{(r} \to \hat{x}^{\#}$

Consideremos ahora ampliada la definición de las funciones $\hat{x}^{(r}(s)$ dadas por:

$$\hat{x}^{(r}(s) = \overset{*}{\bar{x}}(t) + \int_o^s \hat{f}(\hat{x}^{(r}(z), u_{ap_r}^{(r}(z))\, dz$$

para $0 \leq \; \leq a_{p_r h}^{(r}$, en donde $u_{ap_r}^{(r}(Z)$ denotaria la función de con-- trol correspondiente al r-esimo punto de la sucesión con valor de p asociado que hemos denotado por p_r, y siendo $\hat{x}^{(r}(a_{p_r h}^{(r}) = \hat{x}^{(r}$, - al intervalo: $a_{p_r h}^{(r} < s \leq \tau$, haciendo $\hat{x}^{(r}(S) = \hat{x}^{(r}$ eb este último - intervalo.

Observese que la existencia y unicidad de las funciones - $\hat{x}^{(r}(s)$ en los intervalos indicados estaría garantizada por la hipótesis de derivabilidad con continuidad con respecto a $\hat{x}$ de la - función $\hat{f}$. (Véase la Nota 12).

Denominemos a la función $\hat{f}$ correspondiente a la ampliación de la definición de las $\hat{x}^{(r}(s)$, $\hat{f}^{(r}$.

Es decir, sería:

$$\hat{f}^{(r}(\hat{x}^{(r}(z), u_{ap_r}^{(r}(z)) = \hat{f}(\hat{x}^{(r}(z), u_{ap_r}^{(r}(z))$$

para $0 \leq s \leq a_{p_r h}^{(r}$

y $\hat{f}^{(r}(\hat{x}^{(r}(z), u) = 0$

para $a_{p_r h}^{(r} < s \leq \tau$

Consideremos, entonces, para cada i = 0,1,2,.....,n, la sucesión de funciones reales contínuas $x_i^{(r}(s)$, con $r \in \mathbb{N}$, definidas en $[0,\tau,]$ que acabamos de mostrar.

Esta sucesión sería equicontínua en dicho intervalo, ya que si $s < s'$ tendríamos:

$$|x_i^{(r}(s) - x_i^{(r}(s')| = |\int_s^{s'} f_i^{(r} (\hat{x}^{(r}(z), u_{ap_r}^{(r}(z)) dz| \leq$$

$$\leq \int_s^{s'} |f_i^{(r} (\hat{x}^{(r}(z), u_{ap_r}^{(r}(z)| dz \leq M_k (s'-s)$$

para todo $r \in \mathbb{N}$ y para todo $i = 0,1,2,\ldots, n$.

Además es uniformemente acotada ya que

$$|x_i^{(r}(s)| = |\bar{x}_i(t) + \int_0^s t_i^{(r}(\hat{x}^{(r}(z), u_{ap_r}^{(r}(z) dz| \leq$$

$$\leq |x_i(t)| + \int_0^s |f_i^{(r}(\hat{x}^{(r}(z), u_{ap_r}^{(r}(z)| dz \leq$$

$$\leq \max_{0 \leq i \leq n} |x_i(t)| + M_k \tau$$

para todo $s \in [0,\tau]$, todo $r \in \mathbb{N}$ y todo $i = 0,1,2,\ldots..n$

Entonces, la aplicación del teorema de Arzelá (ver Nota 14) permite concluir que existirá una subsucesión, que seguiremos denotando por $\hat{x}^{(r}(s)$, con $0 \leq s \leq \tau$, lo cual converge uniformemente en $[0,\tau]$ a una función contínua $\hat{x}^{\#}(s)$, con $0 \leq s \leq \tau$.

Supondremos en adelante la discusión restringida a dicha subsucesión. En relación a ella consideremos la correspondiente sucesión de funciones $u_{ap_r}^{(r}(z)$.

Por definición de dichas funciones tendríamos que:

$$u_{ap_r}^{(r}(z) = u^{(j}, \text{ para } a_{p_r(j-1)}^{(r} \leq z \leq a_{p_r j}^{(r}$$

(Volvemos a indicar que la ambigüedad que pudiera darse en la definición de la $u_{ap_r}^{(r}(z)$ por la duplicidad de definiciones en los $a_{p_r j}$ sería irrelevante ya que el conjunto de todos ellos sería de medida nula).

Consideremos las correspondientes sucesiones $a_{p_r j}$, $j = 1,2, \ldots.h$; $r \in \mathbb{N}$ y $p_r \in \mathbb{N}$

Ahora, como para todo, r y j se verifica: $0 \leq a^{(r}_{p_r j} \leq a^{(r}_{p_r h} \leq \tau$, podemos seleccionar una subsucesión de la $a^{(r}_{p_r h}$ convergente a un determinado valor a^*_h.

Entonces, restringidos a la subsucesión convergente anterior se podría escoger a su vez otra subsucesión de la $a^{(r}_{p_r 1}$ convergente a un determinado valor a^*_1. Restringidos a ésta última se podría escoger a su vez otra subsucesión de la $a^{(r}_{p_r 2}$ convergente a a^*_2, y así sucesivamente hasta a^*_{h-1}.

De las condiciones impuestas a las $a^{(r}_{p_r j}$ para cada r fijo se seguiría que: $a^*_j \leq a^*_{j+1}$, para j= 1,2,..., (h-1).

Escojamos ahora de las diferentes h subsucesiones inducidas en las sucesiones $u^{(r}_{ap_r}(z)$, por el procedimiento descrito antes, la última de ellas que denotaremos simplemente por $u^{(r}_a(z)$.

Vamos a construir entonces una función medible y acotada - $u^*_a(z)$, en el intervalo $[0,\tau]$,a la que convergerá casi seguramente la sucesión anterior.

Para ello definamos la función $u^*_a(z)$ en el intervalo $[0,\tau]$ de la siguiente manera:

$$u^*_a(z) = u^{(j}, \text{ para } a^*_{j-1} < z < a^*_j \quad , \text{ con } a^*_o = 0$$

$$\text{y } u^*_a(z) = u^{(h} \quad , \text{ en el resto del intervalo}[0,\tau].$$

De que $a^*_{j-1} \leq a^*_j$, para todo j, se seguiría que la función anterior estaría bien definida. (Se entiende que cuando el correspondiente conjunto fuera vacío no se aplica la definición relativa a dicho conjunto).

Consideremos ahora el conjunto unión de los conjuntos correspondientes a los puntos de discontinuidad de las funciones de la sucesión $u_a^{(r}(z)$. Como cada una de dichas funciones tendría un número finito de puntos de discontinuidad dicha unión será a lo más una colección numerable de puntos del intervalo$[0,\tau]$, y por lo tanto de medida de Lebesgue nula

Por otra parte, el conjunto de puntos límite de los anteriores, integrado por los a_j^* , es finito, por lo que la unión de ambos conjuntos será, por lo tanto, de medida nula.

Supongamos que z pertenece al complementario del conjunto anterior y $z < a_h^*$.

Entonces z será un punto interior a alguno de los subintervalos:

$$[a_{j-1}^*, a_j^*]$$

para algun j, que denominaremos j_z.

Por consiguiente z será interior a cada uno de los subintervalos:

$$[a_{p_r(j-1)}^{(r}, a_{p_r j_z}^{(r}]$$

para todo $r \geq r_o$, si r_o es suficientemente grande, pues en otro caso alguna de las correspondientes sucesiones: $a_{p_r j}^{(r}$ no convergería o bien su límite coincidiría con z , contra lo supuesto.

Se seguiría entonces que:

$$\lim_r u_a^{(r}(z) = u^{(j_z} = u_a^*(z)$$

Un razonamiento similar prueba que si $a_h^* < z \leq \tau$, supuesto que $a_h^* < \tau$, pues en otro caso no habría lugar a esta consideración, entonces:

$$\lim_r u_a^{(r}(z) = u^{(h} = u_a^*(z)$$

Así, se pondría concluir, por lo tanto, que:

$$\lim_r u_a^{(r}(z) = u_a^*(z)$$

casi seguramente en $[0,\tau]$

Además, evidentemente $u_a^*(z)$ es acotada en $[0,\tau]$, y con cada componente medible, ya que cada una de ellas sería una función "simple", es decir una para la que existiría una partición finita de $[0,\tau]$ en conjuntos medibles de manera que la función tomaría un valor constante en cada uno de dichos conjuntos.

También serán medibles, naturalmente, cada una de las funciones $f_i^{(r}(\hat{x}^{(r}(z), u_a^{(r}(z))$ para $r \in \mathbb{N}$, en $[0,\tau]$

Por otra parte, para todo $\varepsilon > 0$ suficientemente pequeño y para casi todo z tal que: $0 \leq z < a_h^* - \varepsilon$, tenemos, por la continuidad de $\hat{f}$ en $\hat{x}$ y u:

$$\lim_r \hat{f}^{(r}(\hat{x}^{(r}(z), u_a^{(r}(z)) = \hat{f}(\lim_r \hat{x}^{(r}(z), \lim_r u_a^{(r}(z)) =$$
$$= \hat{f}(\hat{x}^{\#}(z), u_a^*(z))$$

mientras que para todo z tal que: $a_h^* + \varepsilon \leq z \leq \tau$ tenemos:

$$\lim_r \hat{f}^{(r}(\hat{x}^{(r}(z), u_a^{(r}(z)) = 0$$

omo ε es arbitrario obtendríamos entonces que:

$$\lim_r \hat{f}^{(r}(\hat{x}^{(r}(z), u_a^{(r}(z)) = \hat{f}^*(\hat{x}^{\#}(z), u_a^*(z))$$

casi seguramente en $[0,\tau]$, en donde $\hat{f}^*$ vendría dada por:

$$\hat{f}^*(\hat{x}^{\#}(z), u_a^*(z)) = \hat{f}(\hat{x}^*(z), u_a(z)) \quad \text{para } z \in [0,a_h^*]$$
$$= 0, \text{ en el resto de } [0,\tau]$$

la cual será una función medible.

Entonces, ya que $|f_i(\hat{x}(z), u(z))| \leq M_k$, para todo $i = 0, 1, 2, \ldots, n$ y para todo $\hat{x}(z)$ y todo $u(z)$ pertenecientes al conjunto (126), del teorema de la convergencia dominada de Lebesgue (ver Nota 15) se obtiene:

$$(127) \qquad \hat{x}^{\#}(s) = \hat{\bar{x}}(t) + \int_0^s \hat{f}(\hat{x}^{\#}(\ ),\ u_a^*(z))\,dz$$

para $s \in [0, a_h^*]$

y: $\hat{x}^{\#}(s) = \hat{x}^{\#}(a_h^*)$

para $s \in [a_h^*, \tau]$

Pero:

$$|\hat{x}^{(r}(a^{(r}_{p_r h}) - \hat{x}^{\#}(a_h^*)| \leq |\hat{x}^{(r}(a^{(r}_{p_r h}) - \hat{x}^{\#}(a^{(r}_{p_r h})| +$$

$$+ |\hat{x}^{\#}(a^{(r}_{p_r h}) - \hat{x}^{\#}(a_h^*)|$$

de lo que se deduce:

$$\hat{x}^{\#}(a_h^*) = \lim_r \hat{x}^{(r}(a^{(r}_{p_r h}) = \lim_r \hat{x}^{(r} = x^{\#}$$

Ahora, de lo anterior se seguiría que $\hat{x}^{\#}$ sería un punto accesible para el Problema (116) en las condiciones del enunciado del Lema (125), pues muestra que existiría un instante $t + a_h^*$ y una función de control admisible, definida por $\bar{u}(s)$, para $s \in [t_0, t]$ y $u_a^*(s - t)$ para $s \in (t_1\ t + a_h^*]$, tal que para la correspondiente trayectoria $\hat{x}^{(*}(s)$ se verificaría: $\hat{x}^{(*}(t + a_h^*) = \hat{x}^{\#}(a_h^*) = \hat{x}^{\#}$, mientras que, por otra parte, se verificaría además que: $\hat{x}^{(*}(t + a_h^*) = x^{\#} = \bar{x}(t)$, y $x_0^{(*}(t + a_h^*) = x_0^{\#} < \bar{x}_0(t)$, lo que contradiría el principio de Evolución Optima (120). Esto completa la demostración.

El lema (125) ha dejado preparado ya parte del desarrollo para establecer el resultado fundamental que estaría en la base de la demostración el Principio de Máximo, el cual correspondería al de que, puesto que el punto $\hat{x}(t)$ resulta ser un punto perteneciente a la frontera inferior del conjunto $\bar{C}_t^{(kh}$, se podría garantizar la existencia de un vector no nulo de multiplicadores, que verificaría una determinada desigualdad para cierta colección de vectores tangentes definidos en $\hat{x}(t)$.

Sin embargo, el establecimiento del resultado anterior requiere alguna preparación previa debido a las dificultades que se presentarían al definir el correspondiente conjunto convexo de vectores tangentes en los puntos del conjunto $\bar{C}_t^{(kh} - C_t^{(kh}$, de manera que quedaran definidos éstos en todo punto del conjunto $\bar{C}_t^{(kh}$, como exigiría una aplicación posterior del Principio (91).

Estas dificultades se deben en primer lugar al hecho de que si el punto de $\bar{C}_t^{(kh} - C_t^{(kh}$, que denotaremos por $\hat{x}^{(q}$, fuera tal que verificara:

$$\sum_{i=o}^{n} (x_i^{(q}(s)-\bar{x}_i(t))^2 = \sigma$$

para algún s perteneciente al intervalo de definición de la trayectoria que termina en $\hat{x}^{(q}$, podría suceder entonces que alguno de los posibles puntos perturbados $\hat{x}^{(qq}$ mediante los cuales se definiria el correspondiente vector tangente, podría verificar: $\sum_{i=o}^{n} (x_i^{(qq}(s) - \bar{x}_i(t))^2 > \sigma$ para algún punto de la trayectoria $\hat{x}^{(qq}(s)$ de definición del punto $\hat{x}^{(qq}$, y no pertenecer, por lo tanto dicho $\hat{x}^{(qq}$ a $\bar{C}_t^{(kh}$, lo que invalidaría la construcción de dicho vector tangente.

Si la condición hubiera afectado solo el punto terminal no habría mayores problemas ya que en el proceso de aplicación del -

Principio (91) habría que considerar una sucesión de puntos de - $\bar{C}_t^{(kh}$, $\hat{x}^{(qr}$, tal que $\hat{x}^{(qr} \to \hat{\bar{x}}(t)$, y, dado que $\sigma > 0$, entonces, si r_o fuera suficientemente grande tendríamos que necesariamente se verificaría $\sum_{i=o}^{n} (x_i^{(qr} - \bar{x}_i(t))^2 < \sigma$, y sería irrelevante, por lo - tanto, como se definiera el cono de vectores tangentes en los puntos $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$, tales que $\sum_{i=o}^{n} (x_i^{(q} - \bar{x}_i(t))^2 = \sigma$, el cual podría considerarse simplemente coincidente con el formado exclusivamente - por el vector nulo.

Sin embargo, el hecho de que la condición $\sum_{i=o}^{n} (x_i^{(q}(S) - \bar{x}_i(t))^2 = \sigma$ se verifique en algún τ intermedio haría que la discusión - fuera algo menos simple.

Consideremos en ese caso el conjunto de puntos $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$, que denominaremos $F_{(\tau)}$ caracterizados por la propiedad de que sus trayectorias definitorias $\hat{x}^{(q}(s)$ pertenecen, en alguno de sus puntos, a la frontera de la bola de centro $\hat{\bar{x}}(t)$ y radio $\sigma^{1/2}$ en $\mathbb{R}^{n+1}$.

Asociemos a cada punto $\hat{x}^{(q}$ el valor:

$$s_q = \inf \{s \in \mathbb{R} \;/\; \sum_{i=o}^{n} (x_i^{(q}(s) - \bar{x}_i(t))^2 = \sigma\}$$

y sea $G_{(\tau)}$ el conjunto:

$$G_{(\tau)} = \{s \in \mathbb{R} \;/\; s = s_q, \text{ para algún } \hat{x}^{(q} \in F_{(\tau)}\}$$

Ahora, supuesto $G_{(\tau)}$ no vacio, pues en otro caso no habría nada que demostrar, pongamos:

$$s_\tau^* = \inf \{ \in \mathbb{N} \;/\; s \in G_{(\tau)}\}$$

Supongamos en primer lugar que $s_\tau^* > 0$. Entonces, si se sustituye τ en la definición del conjunto (124) por $s_\tau^*/2$, lo cual no modificaría en absoluto al subconjunto del $C_t^{(kh}$ original caracterizado por consistir en puntos obtenidos con controles aplicados durante el tiempo $s_\tau^*/2$, el cual coincidiría con el nuevo $C_t^{(kh}$, quedaría garantizado que el conjunto $F_{(s_\tau^*/2)}$ sería vacío, con lo que desaparecería el problema, ya que para dicho valor de τ y todo $\bar{C}_t^{(kh}$ se tendría:

$$\sum_{i=o}^{n} (x_i^{(q}(s) - \bar{x}_i(t))^2 < \sigma$$

para todo s

En caso contrario debería de ser $s_\tau^* = 0$. Pero entonces -- existiría una sucesión $\hat{x}^{(qr}$ de puntos de $\bar{C}_t^{(kh}$ y una sucesión s_r - convergente a cero de manera que:

$$\sum_{i=o}^{n} (x_i^{(qr}(s_r) - \bar{x}_i(t))^2 = \sigma,$$

para todo r.

Ahora, ya que el punto $(\hat{x}^{(qr}(s), u_a^{(qr}(s))$ pertenecerá para todo s y todo r al conjunto (126), tendremos que:

$$| x_i^{(qr}(s) - \bar{x}_i(t)| \leq M_k s_r$$

para todo i = 0, 1, 2,..., n, y todo r

Por lo tanto:

$$\lim_r \hat{x}^{(qr}(s_r) = \hat{\bar{x}}(t)$$

y debería ser:

$$\lim_r \sum_{i=o}^{n} (x_i^{(qr}(s_r) - \bar{x}_i(t))^2 = 0,$$

lo que sería imposible si $\sigma \neq 0$.

Así necesariamente tendríamos $s_\tau^* > 0$ y se podría escoger $\tau_\sigma > 0$, en donde el subindice indicaría la dependencia del τ escogido del σ, de manera que las trayectorias definitorias de todos los puntos del correspondiente conjunto $\bar{C}_t^{(kh}$ permanecieran en el interior de una bola de radio $\sigma^{1/2}$ y centro en el punto - $\hat{\bar{x}}(t)$, lo que supondremos efectuado en adelante.

En segundo lugar, una dificultad de más difícil arreglo - sería la que se originaría en el caso de que el punto $\hat{x}^{(q}$ de $\bar{C}^{(kh}$ - $C^{(kh}$, viniera definido a partir del punto $\hat{\bar{x}}(t)$ por una función de control definida en todo el intervalo $[0,\tau]$, esto es, cuando $a_h^* = \tau$.

En ese caso, ya que la definición de la colección de puntos perturbados $\hat{x}^{(qq}$ exigiría la ampliación del intervalo de definición de la función de control, no estaría garantizada entonces la pertenencia de cada uno de los puntos perturbados $\hat{x}^{(qq}$ al conjunto $\bar{C}_t^{(kh}$, ni por lo tanto la licitud de la definición del correspondiente vector tangente.

Para sortear esta dificultad es para lo que se hace necesario definir otro conjunto similar al (124) que resolvería aquellos casos en los que no se hubiera podido llegar a la conclusión buscada mediante la sola utilización de los conjuntos del tipo (124). En la generación de los conos convexos de vectores tangentes en todo punto de la clausura de estos nuevos conjuntos podría utilizarse el resultado de la inconclusión mediante los conjuntos (124).

Véamos concretamente en que resultados operativo podría traducirse la imposibilidad de la definición del correspondiente cono tangente en todo punto de interés de $\bar{C}_t^{(kh}$.

Como ya hemos observado antes, de hecho solo estaríamos interesados en la definición de los conos tangentes en aquellos puntos $\hat{x}^{(q}$ que pudieran formar una sucesión convergente a $\hat{x}(t)$, tal como se desprende del enunciado del Principio (91).

De acuerdo con lo que acabamos de exponer podrían establecerse dos posibilidades. La primera sería la de que existiera un $\tau_o > 0$ tal que para todo $\tau < \tau_o$ no existiera ninguna sucesión de puntos $\hat{x}^{(qr}$ del correspondiente $\bar{C}_t^{(kh}$ convergente a $\hat{\bar{x}}(t)$ y de manera que $a_h^{(qr} = \tau$, para todo $r \in \mathbb{N}$, en donde por $a_h^{(qr}$ denotamos el a_h^* asociado a cada $\hat{x}^{(qr}$, o bien la de que para cada uno de los términos de una sucesión de valores de τ convergente a cero existiera una sucesión de puntos $\hat{x}^{(qr}$ del correspondiente $\bar{C}_t^{(kh}$ para dicho τ tal que $a_h^{(qr} = \tau$ para todo r y $\hat{x}^{(qr} \to \hat{\bar{x}}(t)$.

En el primer caso bastaría que nos restringieramos a valores de τ menores de τ_o para que quedara garantizada la posibilidad de definición de conos de vectores tangentes en todos los puntos de inteŕes de $\bar{C}_t^{(kh}$, mientras que en el segundo caso, y para cada uno de los valores de τ en cuestión, tendríamos que existiría una sucesión de funciones: $\hat{x}^{(qr}(s)$, definidas todas ellas en el intervalo $[0,\tau]$, de manera que:

$$\hat{x}^{(qr}(s) = \hat{\bar{x}}(t) + \int_o^s \hat{f}(\hat{x}^{(qr}(z), u_a^{(qr}(z))\, dz$$

en donde las funciones de control $u^{(qr}(s)$ se definirían como:

$$u_a^{(qr}(z) = u^{(j}, \text{ para } a_{j-1}^{(qr} \leq z \leq a_j^{(qr}, \text{ y todo } r \in \mathbb{N},$$

con $a_o^{(r} = 0$, y siendo $\hat{x}^{(qr}(\tau) = \hat{x}^{(qr}$, para todo r

Entonces, por el procedimiento descrito en la demostración del Lema (125) se podrían determinar una función $\hat{x}^{(q\#}(s)$, definida en $[0,\tau]$, y una función de control asociada, $u_a^{(q*}(z)$, definida de manera análoga a las $u_a^{(qr}(z)$ pero sustiyendo las $a_j^{(r}$ por sus límites, las a_j^*, de manera que:

$$(128) \quad \hat{x}^{(q\#}(s) = \hat{\bar{x}}(t) + \int_o^s \hat{f}(\hat{x}^{(q\#}(z), u_a^{(q*}(z))\, dz$$

para todo $s \in [0,\tau]$

En particular se obtendría, por lo tanto:

$$\hat{x}^{(q\#}(\tau) = \lim_r \hat{x}^{(qr}(\tau) = \hat{\bar{x}}(t) =$$
$$= \hat{\bar{x}}(t) + \int_o^\tau \hat{f}(\hat{x}^{(q\#}(s), u_a^{(q*}(s))\, ds$$

Es decir, que:

$$(129) \quad \int_o^\tau \hat{f}(\hat{x}^{(q\#}(s), u_a^{(q*}(s))\, ds = 0$$

para una determinada función de control $u_a^{(q*}(s)$, definida de la siguiente manera:

$$u_a^{(q*}(s) = u^{(j}, \text{ para } a^*_{j-1} \leq s \leq a^*_j \text{ , con } a^*_o = 0$$

y verificando la función $\hat{x}^{(q\#}(s)$ la ecuación (128) en $[0,\tau]$.

Además (129) podría establecerse para todo τ_r en una sucesión convergente a cero.

Vamos a posponer para más adelante el análisis del caso en que se dan (128) y (129), y a examinar en primer lugar la generación de conos convexos de vectores tangentes en cada punto $\hat{x}^{(q}$ de un conjunto $\bar{C}_t^{(kh}$, clausura de uno del tipo (124), cuando para el τ en cuestión se pudiera garantizar la no existencia de una sucesión de puntos $\hat{x}^{(qr}$ pertenecientes a $\bar{C}_t^{(kh}$, con $a_h^{(qr} = \tau$, para todo r y $\hat{x}^{(qr} \to \hat{\bar{x}}(t)$. Podemos por lo tanto suponer que el punto - $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$ en el que habría que definir el haz de vectores tan-- gentes es tal que $a_h^{(q} < \tau$, en donde por $a_h^{(q}$ **denotamos** el correspondiente a_h^* asociado, y además, de acuerdo con una observación - que hemos hecho antes, que: $\sum_{i=o}^{n} (x_i^{(q}(s) - \bar{x}_i(t))^2 < \sigma$, para todo s. En esa situación, pues, se definiría el punto $\hat{x}^{(qq}$ de $\bar{C}_t^{(kh}$, mediante la siguiente ecuación:

$$\hat{x}^{(qq} = \hat{x}^{(qq}(a_h^{(q} + \sum_{j=1}^{h} b_j^{(qq}) =$$

$$= \hat{\bar{x}}(t) + \int_o^{a_h^{(q} + \sum_{j=1}^{h} b_j^{(qq}} \hat{f}(\hat{x}^{(qq}(s), u^{(qq}(s)) \, ds$$

en donde para todo $s \in [0, \sum_{j=1}^{h} b_j^{(qq}]$, sería:

$$\hat{x}^{(qq}(s) = \hat{\bar{x}}(t) + \int_o^s \hat{f}(\hat{x}^{(qq}(z), u^{(qq}(z)) \, dz$$

con:

$$u^{(qq}(z) = u^{(j},$$

$$\text{para } a_{j-1}^{(q} + \sum_{r=o}^{j-1} b_r^{(qq} \leq z \leq a_j^{(q} + \sum_{r=o}^{j} b_r^{(qq} ,$$

para $j = 1, 2,..., h$, siendo $b_r^{(qq} \geq 0$, para todo r,y $a_o^{(q}=b_o^{(qq} = 0$

Evidentemente, si $\sum_{j=1}^{h} b_j^{(qq}$ es suficientemente pequeño tendríamos que:

$$a_h^{(q} + \sum_{j=1}^{h} b_j^{(qq} < \tau, \text{ y:}$$

$$\sum_{i=o}^{n} (x_i^{(qq} - \bar{x}_i(t))^2 < \sigma,$$

debido a la dependencia continua de las soluciones de la ecuación diferencial: $\frac{d}{dt}\hat{x}(t) = \hat{f}(\hat{x}(t), u(t))$ de las condiciones iniciales, tal como se deduce de las condiciones de continuidad y derivabilidad supuestas en f, lo que garantizaría que $\hat{x}^{(qq} \in \bar{C}_t^{(kh}$,

Consideremos ahora una sucesión de puntos perturbados $\hat{x}^{(qq}$, asociados a unas sucesiones decrecientes a cero de valores de los correspondiente $b_j^{(qqr}$, $j = 1, 2, \ldots, h$.

Se trata entonces de investigar el límite:

$$\lim_{\sum_{j=1}^{h} b_j^{(qqr} \to 0} \frac{\hat{x}^{(qqr} - \hat{x}^{(q}}{\sum_{j=1}^{h} b_j^{(qqr}} =$$

$$= \lim_{r} \frac{1}{\sum_{j=1}^{h} b_j^{(qqr}} \left[\int_o^{a_h^{(q} + \sum_{j=1}^{h} b_j^{(qqr}} \hat{f}(\hat{x}^{(qqr}(s), u^{(qqr}(s))ds -\right.$$

$$\left. - \int_o^{a_h^{(q}} \hat{f}(\hat{x}^{(q}(s), u^{(q}(s))\, ds\right]$$

Ahora, para poder evaluar el anterior límite, mediante el cual definiriamos el correspondiente vector tangente en el punto $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$ necesitariamos previamente discutir algunas cuestiones relativas a la dependencia de las soluciones de la ecuación diferencial vectorial de las condiciones iniciales con algo más de detalle.

Consideremos para ello, supuesto $t_o < t < t_1$, el sistema de ecuaciones diferenciales lineal y de coeficientes variables si guiente:

(130) $\frac{d\xi(s)}{ds} = \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^*(s), u^*(s)) . \ \xi(s)$

en donde $s \in [t, t + \Delta t]$, y siendo $\partial \hat{f}/_{\partial \hat{x}}$ la matriz cuyos elementos en la fila i-esima y columna j-esima vendría dado por $\partial \hat{f}_i / \partial \hat{x}_j$, y $\xi(s) = (\xi_o(s), \xi_1(s), ..., \xi_n(s))'$, y en donde $u^*(s)$ es una determinada función de control admisible en $[t, t + \Delta t]$ siendo $\hat{x}^*(s)$ - su trayectoria asociada correspondiente a una determinada condi-- ción inicial cuya existencia estaría garantizada, como a continua ción discutiremos con algo más de detalle, por las condiciones im puestas a $\hat{f}$.

Denotemos por $\Phi_{u*}(s,t)$ a la matriz fundamental del sistema (130), cuya existencia estaría asimismo garantizada (véase la Nota 16) por la continuidad de $\hat{f}$ y $\partial \hat{f}/\partial \hat{x}$ en $\mathbb{R}^{n+1} \times \Omega$, pues de ello - se seguiría que las correspondientes normas, definidas, por ejemplo, por la raiz cuadrada positiva de la suma de las cuadrados de cada una de los componentes del vector y matriz, estarían acota-- das en $[t, t + \Delta t]$, ya que $u^*(s)$ se supone acotada en dicho inter valo.

Entonces $\Phi_{u*}(s,t)$ será, para t fijo, una función absolutamente continua de s en $[t, t + \Delta t]$. Además: $\Phi_{u*}(t,t) = I$, en - donde I denota la matriz unidad y $\Phi^{-1}_{u*}(s,t) = \Phi_{u*}(s,t)$, para todo - - - $s \in [t, t + \Delta t]$.

También se podría demostrar que si $t < s_1 < s_2 < t + \Delta t$, entonces:

$$\Phi_{u*}(s_2, t) = \Phi_{u*}(s_2, s_1) \ \Phi_{u*}(s_1, t)$$

Así, podemos poner, por lo tanto:

$$\xi(t + \Delta t) = \Phi_{u*} (t + \Delta t,\ t)\ \xi(t)$$

Supongamos ahora que $\hat{x}^{(g*}$ representa la trayectoria correspondiente a la misma función de control $u*(s)$, para $s \in [t,\ t+\Delta t]$, pero con condición inicial:

$$\hat{x}^{(\delta*}(t) = \hat{x}^{*}(t) + \delta\xi(t) + o(\delta)$$

en donde $\delta > 0$ y $o(\delta)/\delta \to 0$ cuando $\delta \downarrow 0$

Tendríamos entonces:

$$\hat{x}^{(\delta*}(s) = \hat{x}^{(\delta*}(t) + \int_t^s \hat{f}(\hat{x}^{(\delta*}(z),\ u^{*}(z))\ dz$$

para todo $s \in [t,\ t + \Delta t]$.

Ahora, teniendo en cuenta las condiciones de derivabilidad supuestas a $\hat{f}$ podríamos poner:

$$\hat{f}(\hat{x}^{(\delta*}(z),\ u^{*}(z)) = \hat{f}(\hat{x}^{*}(z) + [\hat{x}^{(\delta*}(z) - \hat{x}^{*}(z)],\ u^{*}(z)) =$$

$$= \hat{f}(\hat{x}^{*}(z),\ u^{*}(t)) + \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{*}(z),\ u^{*}(z))\ .\ [\hat{x}^{(\delta*}(z) -$$

$$- \hat{x}^{*}(z)] + o(\hat{x}^{(\delta*}(z) - \hat{x}^{*}(z))$$

Se sigue entonces que:

$$\frac{\hat{x}^{(\delta*}(s) - \hat{x}^{*}(s)}{\delta} = \frac{\hat{x}^{(\delta*}(t) - \hat{x}^{*}(t)}{\delta} +$$

$$+ \int_t^s \frac{\delta \hat{f}}{\delta \hat{x}}\ (\hat{x}^{*}(z),\ u^{*}(z))\ .\ [\frac{\hat{x}^{(\delta*}(z) - \hat{x}^{*}(z)}{\delta}]\ dz\ +$$

$$+ \int_t^s \frac{1}{\delta}\ o\ (\hat{x}^{(\delta*}(z) - \hat{x}^{*}(z))\ dz$$

para todo $\delta > 0$, ya que estaria garantizado que los dos integrandos serían funciones integrables puesto que serían medibles y acotadas en $[t,\ t + \Delta t]$.

Además, el límite:

$$\lim_{\delta \downarrow 0} \frac{\hat{x}^{(\delta *)}(s) - \hat{x}^{*}(s)}{\delta}$$

existe, para todo s, como se deduce de la dependencia de la solución de un sistema de ecuaciones diferenciales de las condiciones iniciales (véase la Nota 17). Denotemos, pues, al límite anterior por $\xi(s)$.

Por otra parte, tendríamos también:

$$\lim_{\delta \downarrow 0} \frac{1}{\delta}\, o\,(\hat{x}^{(\delta *)}(s) - \hat{x}^{*}(s)) =$$

$$= \lim_{\delta \downarrow 0} \frac{o\,(\hat{x}^{(\delta *)}(s) - \hat{x}^{*}(s))}{[\sum_{i=o}^{n} (\hat{x}_i^{(\delta *)}(s) - \hat{x}_i^{*}(s))^2]^{1/2}} \left[\sum_{i=o}^{n} \left(\frac{x_i^{(\delta *)}(s) - x_i^{*}(s)}{\delta} \right)^2 \right]^{1/2} = 0$$

Del teorema de la convergencia dominada de Lebesgue, cuya aplicación sería licita ya que los dos integrandos estarian uniformemente acotados en δ, se seguiria entonces que:

$$\lim_{\delta \downarrow 0} \frac{\hat{x}^{(\delta *)}(s) - \hat{x}^{*}(s)}{\delta} = \xi(s) = \xi(t)$$

$$+ \int_t^s \lim_{\delta \downarrow 0} \frac{\delta \hat{f}}{\delta \hat{x}}(\hat{x}^{*}(z), u^{*}(z)) \left[\frac{\hat{x}^{(\delta *)}(z) - \hat{x}^{*}(z)}{\delta} \right] dz =$$

$$= \xi(t) + \int_t^s \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{*}(z), u^{*}(z)) \,.\, \xi(t)\, dz$$

para todo $s \in [t, t + \Delta t]$

Esto es, el vector límite $\xi(s)$ verificaría precisamente el sistema lineal de ecuaciones diferenciales (130), con la condición inicial $\xi(t)$.

Podríamos poner entonces también:

$$(131) \qquad \lim_{\delta \downarrow 0} \frac{\hat{x}^{(\delta *}(s) - \hat{x}^{*}(s)}{\delta} = \Phi_{u*}(s, t)\ \xi(t)$$

para todo $s \in [t, t + \Delta t]$

Volvamos ahora de nuevo a considerar el problema de calcular el límite del cociente $(\hat{x}^{(qqr} - \hat{x}^{(q}) / \sum_{j=1}^{h} b_j^{(qqr}$, cuando $\sum_{j=1}^{h} b_j^{(qqr} \downarrow 0$.

Podemos restingirnos, en particular, al caso en que los $b_j^{(qqr}$ son de la forma: $b_j^{(qqr} = \lambda_j \delta$, en donde $\delta > 0$, y λ_j permanecerían constantes cuando $\delta \downarrow 0$.

Podemos poner entonces:

$$(132) \qquad \hat{x}^{(qqr} = \hat{x}_\delta^{(qq} = \hat{\bar{x}}(t) + \int_o^{a_1^{(q}} \hat{f}(\hat{x}_\delta^{(qq}(s), u_\delta^{(qq}(s))\, ds +$$

$$+ \int_{a_1^{(q}}^{a_1^{(q} + \lambda_1 \delta} \hat{f}(\hat{x}_\delta^{(qq}(s), u_\delta^{(qq}(s))\, ds +$$

$$+ \ldots + \int_{a_{h-1}^{(q} + \sum_{j=o}^{h-1} \lambda_j \delta}^{a_h^{(q} + \sum_{j=o}^{h-1} \lambda_j \delta} \hat{f}(\hat{x}_\delta^{(qq}(s), u_j^{(qq}(s))\, ds +$$

$$+ \int_{a_h^{(q} + \sum_{j=o}^{h-1} \lambda_j \delta}^{a_h^{(q} + \sum_{j=o}^{h} \lambda_j \delta} \hat{f}(\hat{x}_\delta^{(qq}(s), u_\delta^{(qq}(s))\, ds$$

en donde el subindice δ en $\hat{x}_\delta^{(qq}(s)$ $(= \hat{x}^{(qqr}(s))$ y $u_\delta^{(qq}(s)$ $(= u^{(qqr}(s))$denotarían que las correspondiente funciones $\hat{x}^{(qq}(s)$ y $u^{(qq}(s)$ dependerían del valor de δ escogido y además $a_o^{(q} = \lambda_o = 0$.

Vamos a comprobar a continuación que:

$$(133) \qquad \hat{x}_\delta^{(qq} = \hat{x}^{(q}(a_h^{(q}) + \xi(a_h^{(q}) \cdot \delta + o(\delta)$$

para cada $\delta > 0$, en donde:

(134) $$\xi(a_h^{(q}) = \sum_{j=1}^{h} \lambda_j \, \Phi_{u^{(q}}(a_h^{(q}, a_j^{(q}) \, . \, \hat{f}(\hat{x}^{(q}(a_j^{(q}), u^{(j)})$$

siendo $\Phi_{u^{(q}}(\cdot,\cdot)$ la matriz fundamental del sistema:

(135) $$\frac{d\xi(s)}{ds} = \frac{\partial\hat{f}}{\partial\hat{x}}(\hat{x}^{(q}(s), u^{(q}(s)) . \xi(s)$$

en el intervalo: $0 \leq s \leq a_h^{(q}$

En efecto, haremos la demostración pro recurrencia en h. - Consideremos en primer lugar que $h = 1$. Entonces tendríamos:

$$\hat{x}_\delta^{(qq} = \hat{\bar{x}}(t) + \int_o^{a_1^{(q}} \hat{f}(\hat{x}^{(qq}(s), u^{(qq}(s)) \, ds +$$

$$+ \int_{a_1^{(q}}^{a_1^{(q} + \lambda_1\delta} \hat{f}(\hat{x}_\delta^{(qq}(s), u^{(qq}(s)) \, ds =$$

$$= \hat{\bar{x}}(t) + \int_o^{a_1^{(q}} \hat{f}(\hat{x}^{(q}(s), u^{(1)}) \, ds +$$

$$+ \lambda_1\delta \, \hat{f}(\hat{x}^{(q}(a_1^{(q}), u^{(1)}) + o\,(\delta)$$

en donde sería:

$$\xi(a_1^{(q}) = \lambda_1 \, \hat{f}(\hat{x}^{(q}(a_1^{(q}), u^{(1)}) = \lambda_1 \Phi_u^{(q}(a_1^{(q}, a_1^{(q}) \, \hat{f}(\hat{x}^{(q}(a_1^{(q}, u^{(1)})$$

Supongamos entonces demostrado (133) para $h = g$ vamos a de̲mostrarlo para $h = g + 1$.

De acuerdo con (131) y si $\Phi_{u^{(q}}(\cdot,\cdot)$ denota, como antes, la matriz fundamental del sistema (135) podríamos poner entonces:

(136) $$\hat{x}_\delta^{(qq}(a_{g+1}^{(q} + \sum_{r=1}^{g} \lambda_r\delta) - \hat{x}^{(q}(a_{g+1}^{(q}) =$$

$$= \Phi_{u^{(q}}(a_{g+1}^{(q}, a_g^{(q}) \, . \, \xi(a_g^{(q}) \, . \, \delta + o\,(\delta)$$

mientras que, por otra parte:

$$\int_{a_{g+1}^{(q}+\sum_{r=1}^{g}\lambda_r\delta}^{a_{g+1}^{(q}+\sum_{r=1}^{g+1}\lambda_r\delta} f(\hat{x}_\delta^{(qq}(s),\ u_\delta^{(qq}(s))\ ds =$$

$$= \lambda_{g+1}\ .\ \delta\ .\ \hat{f}(\hat{x}_\delta^{(qq}(a_{g+1}^{(q} + \sum_{r=1}^{g} \lambda_r\delta),\ u^{(g+1}) + o(\delta),$$

y teniendo en cuenta (136) queda:

$$= \lambda_{g+1}\ .\ \delta\ \hat{f}(\hat{x}^{(q}(a_{g+1}^{(q}),\ u^{(g+1}) + o(\delta)$$

Entonces:

$$\xi(a_{g+1}^{(q}) = \Phi_{u^{(q}}(a_{g+1}^{(q},\ a_g^{(q})\ .\ \xi(a_g^{(q}) + \lambda_{g+1}\ \hat{f}(\hat{x}^{(q}(a_{g+1}^{(q}),u^{(g+1})$$

Es decir:

$$\xi(a_h^{(q}) = \Phi_{u^{(q}}(a_h^{(q},\ a_{h-1}^{(q})\ .\ \xi(a_{h-1}^{(q}) + \lambda_h\ \hat{f}(\hat{x}^{(q}(a_h^{(q}),\ u^{(h}) =$$

$$= \Phi_{u^{(q}}(a_h^{(q},\ a_{h-1}^{(q})\ \Phi_{u^{(q}}(a_{h-1}^{(q},\ a_{h-2}^{(q})\ \xi\ (a_{h-2}^{(q}) +$$

$$+ \lambda_{h-1}\ \Phi_{u^{(q}}(a_h^{(q},\ a_{h-1}^{(q})\ \hat{f}(\hat{x}^{(q}(a_{h-1}^{(q}),\ u^{(h-1}) +$$

$$+ \lambda_h\ f(\dot{x}^{(q}(a_h^{(q}),u^{(h})=\ldots= \lambda_1\ \Phi_{u^{(q}}(a_h^{(q},\ a_1^{(q})\ \hat{f}(\hat{x}^{(q}(a_1^{(q}),u^{(1})+$$

$$+ \lambda_2\ \Phi_{u^{(q}}(a_h^{(q},\ a_h^{(q})\ \hat{f}(\hat{x}^{(q}(a_2^{(q}),\ u^{(2}) + \ldots +$$

$$+ \lambda_{h-1}\ \Phi_{u^{(q}}\ (a_h^{(q},\ a_{h-2}^{(q})\ \hat{f}(\hat{x}^{(q}(a_{h-1}^{(q}),\ u^{(h-1}) +$$

$$+ \lambda_h\ \hat{f}(\hat{x}^{(q}(a_h^{(q}),\ u^{(h})$$

lo que confirma (134).

En particular, por lo tanto, el límite que nos habríamos propuesto evaluar en un principio, en el caso de que $b_j^{(qqr}=\lambda_j\delta$, valdría

$$(137) \qquad \lim_{r} \frac{\hat{x}^{(qqr} - \hat{x}^{(q}}{\sum_{j=1}^{h} b_j^{(qqr}} = \lim_{\delta \downarrow 0} \frac{\hat{x}_\delta^{(qq} - \hat{x}^{(q}}{\delta} \; \frac{1}{\sum_{j=1}^{h} \lambda_j} =$$

$$= \sum_{j=1}^{h} \frac{\lambda_j}{\sum_{j=1}^{h} \lambda_j} \; \Phi_{u^{(q}} \, (a_h^{(q}, \, a_j^{(q}) \; \hat{f}(\hat{x}^{(q}(a_j^{(q}), \, u^{(j})$$

No habría pues ninguna dificultad en comprobar que mediante el procedimiento descrito se podría definir un cono convexo de vectores tangentes del tipo (137), sin más que multiplicar el cociente en el que se pasa al límite por una constante positiva arbitraria, en cada uno de los puntos $\hat{x}^{(q}$ pertenecientes al conjunto $\bar{C}_t^{(kh} - C_t^{(kh}$, que incluiría además el vector nulo, obtenido sin más que tomar $\hat{x}_\delta^{(qq} = \hat{x}^q$, para todo δ.

Ahora bien, el anterior resultado se ha establecido sobre la base de que existe un $\tau_o > 0$ tal que si $\tau < \tau_o$ no existe ninguna sucesión de puntos $\hat{x}^{(qr}$ del correspondiente $\bar{C}_t^{(kh}$ de manera que $\hat{x}^{(qr} \to \hat{\hat{x}}(t)$ y que $a_h^{(qr} = \tau$ para todo r. En caso contrario una manera de sortear la dificultad en la definición de los conos de vectores tangentes en los puntos de $\bar{C}_t^{(kh} - C_t^{(kh}$ sería la de introducción un nuevo conjunto de accesibilidad mediante el cual quedara garantizada dicha definición en cualquier caso, tal como vamos a mostrar a continuación.

Al igual que el conjunto (124), el conjunto de accesibilidad que vamos a definir a continuación es un conjunto de puntos de $\mathbb{R}^{n+1}$ que serian alcanzables a partir del $\hat{\hat{x}}(t)$ por medio de funciones constantes a trozos con la restricción de que los valores de las constantes $u^{(j}$ mediante las que se definiría la función de control habrían de pertenecer a un determinado conjunto $D^{(kh}$, y -

de manera que estas funciones constantes a trozos se irian sola-- pando una a continuación de otra un determinado número de veces.

Con más concreción denotemos por a_{10}, a_{11}, $a_{12}, \ldots, a_{1h}$, - con $a_{1j} \geq a_{1(j-1)}$, para todo j, y $a_{10} = 0$, los puntos de subdivisión correspondientes a los valores $u^{(j}$, $j = 1, 2, \ldots, h$, en donde, sin perdida de generalidad, podemos considerar asociado cada $u^{(j}$ a cada subintervalo $(a_{1(j-1}, a_{1j})$, ya que esto no afectaría - al vector tangente obtenido. Denotemos por $u_a(s)$ a la función de control definida como igual a los $u^{(j}$ en cada uno de los subintervalos indicados, la cual trasladaría el punto inicial $\hat{\bar{x}}(t)$ a un - nuevo punto $\hat{x}^{(1*}(a_{1h})$. Estos puntos de subdivisión a_{1j} no ten- - drían por qué ser necesariamente distintos.

Sean, por otra parte a_{20}, a_{21}, $a_{22}, \ldots, a_{2h}$, con $a_{2j} \geq a_{2(j-1)}$, para todo j, y $a_{20} = 0$, los puntos de subdivisión del control - - constante a trozos que se solaparía a continuación del anterior - para llevar el punto $\hat{x}^{(1*}(a_{1h})$ al $\hat{x}^{(z*}(a_{1h} + a_{2h})$, y así sucesivamente.

Por la utilidad que más adelante se revelará de ello supondremos que se verifica la siguiente restricción:

$$a_{(i+1)h} \leq \frac{1}{2^i} \; a_{1h},$$

para todo i, con $i = 1, 2, \ldots, (p-1)$, siendo p el número de veces que se repite el proceso, y supuesto que $a_{ih} > 0$ para todo i, para evitar ambiguedades. Supondremos además que $\sum_{i=1}^{p} a_{ih} < \tau$, en - donde τ es una constante positiva arbitrariamente pequeña, y también que:

$$\sum_{i=o}^{n} (x_i^{(p*}(s) - \bar{x}_i(t))^2 < \sigma$$

para todo s, con $o \leq s \leq \sum_{i=1}^{p} a_{ih}$, y en donde $0 < \sigma < 1$.

Esto es que se supondríaque la trayectoria obtenida mediante la función de control constante a trozos a partir del $\hat{\bar{x}}(t)$ se mantendría toda ella en el interior de una bola de radio $\sigma^{1/2}$ con centro en el punto $\hat{\bar{x}}(t)$.

Establezcamos pues de una manera precisa el conjunto al que nos estamos refiriendo, el cual denotaremos por $C_t^{*(kh}$ para distinguirlo del dado mediante (124).

Este es el siguiente:

(138) $C_t^{*(kh} = \{\hat{x}^*, x_1^*, x_2^*, \ldots, x_n^*)' \in \mathbb{R}^{n+1}$ / Existen a_j, con $i = 0, 1, 2, \ldots, p$, $j = 0, 1, 2, \ldots, h$, de manera que $a_{i(j-1)} \leq a_{ij}$, $a_{0j} = 0$ y $a_{io} = 0$, para todo i y j, siendo además $a_{ih} > 0$ para todo $i = 1, 2, \ldots, p$, para evitar ambiguedades, y $\sum_{i=o}^{p} a_{ih} < \tau$, con $\tau > 0$, a los que corresponden sendas funciones: $u_{ai}(s)$, $i = 1, 2, \ldots, p$, definidas de modo que $u_{ai}(s) = u^{(j}$, para $a_{i(j-1)} \leq s \leq a_{ij}$, siendo $u^{(j} \in D^{(kh}$ para todo j, y de manera que la función definida de la siguiente forma:

$$\hat{x}^{(i*}\left(\sum_{r=o}^{i-1} a_{rh} + s\right) = \hat{x}^{((i-1)*}\left(\sum_{r=o}^{i=1} a_{rh}\right) +$$

$$+ \int_{o}^{s - \sum_{r=o}^{i-1} a_{rh}} \hat{f}\left(\hat{x}^{(i*}\left(\sum_{r=o}^{i-1} a_{rh} + z\right), u_{ai}(z)\right) dz,$$

para $i = 1, 2, \ldots, p$, y $\sum_{r=o}^{i-1} a_{rh} \leq s \leq \sum_{r=o}^{i} a_{rh}$, siendo además:

$a_{(i+1)h} \leq \frac{1}{2^i} a_{1h}$, verificaría: $\hat{x}^{(0*}(0) = \hat{x}^{(1*}(0) = \hat{\bar{x}}(t)$, y $\hat{x}^* = \hat{x}^{(p*}\left(\sum_{r=1}^{p} a_{ih}\right)$.

Además debería cumplirse que:

$$\sum_{i=o}^{n} (x_i^{(g*}(\sum_{r=o}^{g-1} a_{rh} + s) - \bar{x}_i(t))^2 < \sigma < 1,$$

para todo s, con $\sum_{r=o}^{g-1} a_{rh} \leq s \leq \sum_{r=o}^{g} a_{rh}$, y todo g = 1, 2,..., p}
Lo mismo entonces que para el conjunto $C_t^{(kh}$ dado mediante (124), se podría establecer un lema análogo al (125) relativa ahora al conjunto $C_t^{*(kh}$ dado por (138). Este sería el siguiente:

(139) LEMA

El mismo enunciado que el Lema (125) pero con $\bar{C}_t^{*(kh}$ en lugar de $\bar{C}_t^{(kh}$.

DEMOSTRACION

Puesto que la demostración de este Lema sería en bastantes puntos análoga a la del Lema (125) solo haremos mención de las diferencias e indicaremos someramente el resto de los pasos.

Se empezaría suponiendo, pues, en primer lugar, que para algún t $\in (t_0, t_1)$ existiera un punto $\hat{x}^{\#} \in \bar{C}_t^{*(kh}$, con $k \geq \sum_{i=1}^{n} \bar{u}_i^2(t_o)$, tal que $x^{\#} = \bar{x}(t)$ y $x_o^{\#} < \bar{x}_o(t)$.

De la misma manera que en la demostración del Lema (125) se demostraría, pues, que si $\hat{x}^{(r}$ es una sucesión contenida en $C_t^{*(kh}$ y convergente a $\hat{x}^{\#}$ y $\hat{x}^{(r}(s)$ y $u_a^{(r}(s)$ son la trayectoria y función de control asociadas, podría elegirse una subsucesión de la $\hat{x}^{(r}(s)$ convergente uniformemente a una función $\hat{x}^{\#}(s)$ en $[0,\tau]$. Restringidos a la subsucesión anterior se consideraría la correspondiente sucesión de funciones $u_a^{(r}(s)$, las cuales vendrían definidas de la siguiente manera:

$$u_a^{(r}(z) = u^{(j}, \text{ para } \sum_{i=o}^{g-1} a_{ih} + a_{g(j-1)} \leq z \leq \sum_{i=o}^{g-1} a_{ih} + $$

$$+ a_{gj},$$

en donde $g = 1, 2, \ldots, p_r$; $j = 1, 2, \ldots, h$, y $a_{oh} = 0$, para todo g, y siendo, por ejemplo:

$u_a^{(r}(z) = u^{(h}$, en el resto del intervalo $[0,\tau]$, supuesto que se ha considerado $\hat{x}^{(r}(s) = \hat{x}^{(r}$ para $s \in [\sum_{i=o}^{p_r} a_{ih}, \tau]$.

Consideremos ahora las correspondiente sucesiones $a_{ij}^{(r}$, para $j = 1, 2, \ldots, h$, $r \in \mathbb{N}$, y $p_r \in \mathbb{N}$.

No habría ninguna perdida de generalidad si se supone que $p_{r+1} \geq p_r \geq r$, ya que se podría considerar, a efectos de esta discusión, que en la definición de las $u_a^{(r}(z)$ se ampliará formalmente, a partir de la última $a_{ij}^{(r}$, la colección de estas añadiendo todas las $a_{ij}^{(r}$ nulas que fuesen necesarias. Nótese además que todavía se siguiría verificando la condición: $a_{(i+1)h}^{(r} \leq \frac{1}{2^i} a_{1h}^{(r}$ para las $a_{ij}^{(r}$ añadidas.

Ahora, como para toda i, j y r se verificaría que $0 \leq \sum_{i=1}^{p_r} a_{ih}^{(r} \leq \tau$, podríamos seleccionar una subsucesión de la $\sum_{i=1}^{p_r} a_{ih}^{(r}$ convergente a un determinado valor a_h^*. Entonces, restringidos a la subsucesión convergente anterior se podría escoger a su vez otra subsucesión de la $a_{11}^{(r}$ convergente o un determinado valor a_{11}^*. Restringidos a esta última se podría escoger a su vez otra subsucesión de la $a_{12}^{(r}$ convergente a $a_{12}^*, \ldots,$ restringidas a las anteriores otra subsucesión de la $a_{1h}^{(r}$ convergente a a_{1h}^*, restringidos a esta, otra subsucesión de la $a_{21}^{(r}$ convergente a a_{21}^* y así sucesivamente.

De las condiciones impuestas a las $a_{ij}^{(r}$ para cada r fijo se seguiría que:

$$a_{i(j-1)}^* \leq a_{ij}^*$$

para todo j = 1,2 ...., h, con $a^*_{io} = 0$, y:

$$a^*_{(i+1)h} \leq \frac{1}{2^i} a^*_{1h}$$

para todo i = 1, 2, ...

Además, como para r >geN tenemos

$$\sum_{i=o}^{p_r} a^{(r}_{ih} - \sum_{i=o}^{p_r} a^{(r}_{ih} = \sum_{i=g+1}^{p_r} a^{(r}_{ih} \leq \frac{1}{2^g} a^{(r}_{1h} + \frac{1}{2^{g+1}} a^{(r}_{1h} +$$

$$+\ldots+ \frac{1}{2^{p_r-1}} a^{(r}_{1h} \leq \frac{1}{2^{g-1}} a^{(r}_{1h},$$

se seguiría entonces que:

(140) $$a^*_h \leq \sum_{i=o}^{g} a^*_{it} \leq \frac{1}{2^{g-1}} a^*_{1h}, \text{ para todo } g = 1,2,...$$

Es decir que:

$$\lim_g \sum_{i=o}^{g} a^*_{ih} = a^*_h .$$

Escojamos ahora de cada una de las diferentes subsucesiones inducidas en las sucesiones $u^{(r}_a(z)$ por el procedimiento descrito antes el término que corresponde a la diagonal principal del cuadro cuyas columnas serían dichas subsucesiones inducidas. Volveremos a denotar a la sucesión diagonal resultante por $u^{(r}_a(z)$.

Vamos a construir ahora una función medibles y acotada $u^*_a(z)$, en el intervalo $[0,\tau]$ a la que convergerá casi seguramente la sucesión anterior.

Para ello definamos la función $u^*_a(z)$ en el intervalo $[0,\tau]$ de la siguiente manera:

$$u^*_a(z) = u^{(j}$$

$$\text{para } \sum_{i=o}^{g-1} a^*_{ih} + a^*_{g(j-1)} < z < \sum_{i=o}^{g-1} a^*_{ih} + a^*_{gj}$$

con $g = 1, 2, \ldots,$ y $j = 1, 2, \ldots, h$, siendo $a^*_{oh} = 0$ y - -
$a^*_{go} = 0$, para todo g,

y $u^*_a(z) = u^{(h}$, en el resto del intervalo $[0,\tau]$.

Evidentemente, ya que:

$$a^{(r}_{g(j-1)} \leq a^{(r}_{gj}$$

para todo r, todo g y todo j, se seguirá de ello que,

$$a^*_{g(j-1)} \leq a^*_{gj}$$

y la función anterior estará bien definida.

Consideremos a continuación el conjunto unión de los conjuntos correspondientes a los puntos de discontunidad de las funciones de la sucesión $u^{(r}_a(z)$. Como cada una de estas funciones tiene un número finito de esos puntos dicha unión será a lo más una colección numerable de puntos del intervalo $[0,\tau]$, y por lo tanto de medida Lebesgue nula.

Por otra parte, el conjunto de puntos límite de los anteriores, cuyos elementos son los a^*_{ij} también será numerable. La unión de ambos conjuntos será por consiguiente de medida nula.

Supongamos que z pertenece al complementario del conjunto anterior y $z < a^*_h$. Entonces de (140) se sigue que z será un punto interior a alguno de los subintervalos:

$$\left[\sum_{i=0}^{g-1} a^*_{ih}, \sum_{i=o}^{g} a^*_{ih}\right]$$

para algún $g = 1, 2, \ldots$

Sea dicho intervalo el siguiente:

$$\left[\sum_{i=o}^{g_z-1} a^*_{ih}, \sum_{i=o}^{g_z} a^*_{ih}\right]$$

Entonces, como:

$$\left[\sum_{i=o}^{g_z-1} a_{ih}^{(r}, \sum_{i=o}^{g_z} a_{ih}^{(r}\right] = \bigcup_{j=1,2,\ldots,h} \left[\sum_{i=o}^{g_z-1} a_{ir}^{(r} + a_{g_z(i-1)}^{(r},\right.$$

$$\left.\sum_{i=o}^{g_z-1} a_{ih}^{(r} + a_{g_z j}^{(r}\right]$$

dicho z será interior a alguno de los subintervalos anteriores, - tal como por ejemplo el siguiente:

$$\left[\sum_{i=o}^{g_z-1} a_{ih}^{(r} + a_{g_z(j_z-1)}^{(r}, \sum_{i=o}^{g_z-1} a_{ih}^{(r} + a_{g_z j_z}^{(r}\right]$$

para todo $r > r_o$ si r_o es suficientemente grande, pues en otro ca_ so alguna de las correspondiente sucesiones:

$$\sum_{i=o}^{g_z-1} a_{ih}^{(r} + a_{g_z j}^{(r}$$

no convergería o bien su límite coincidiría con z contra lo su-- puesto. Se sigue entonces que z será interior al intervalo:

$$\left[\sum_{i=o}^{g_z-1} a_{ih}^{*} + a_{g_z(j_z-1)}^{*}, \sum_{i=o}^{g_z-1} a_{ih}^{*} + a_{g_z j_z}^{*}\right]$$

y que:

$$\lim_r u_a^{(r}(z) = u^{(j_z} = u_a^*(z)$$

Un razonamiento similar prueba que si $a_h^* < z \leq \tau$ entonces:

$$\lim_r u_a^{(r}(z) = u^{(h} = u_a^*(z)$$

Así, se puede concluir que:

$$\lim_r u_a^{(r}(z) = u_a^*(z), \text{ casi seguramente en } [0,\tau]$$

El resto de la demostración es similar a la del Lema (125) comprobándose que, para la subsucesión adecuada se verifica:

$$\lim_r \hat{x}^{(r}(s) = \hat{x}^{\#}(s) = \hat{\hat{x}}(t) + \lim_r \int_o^s \hat{f}(\hat{x}^{(r}(z), u^{(r}(z))\, dz =$$

$$= \hat{\hat{x}}(t) + \int_o^s \hat{f}(\hat{x}^{\#}(z), u_a^*(z))\, dz$$

para $s \in [0,\tau^*]$

En la definición de los conos convexos de vectores tangentes en los puntos del conjunto $\overline{C}_t^{*(kh} - C_t^{*(kh}$ surgirían las mismas dificultades que en el caso ya discutido de la generación de estos en $\overline{C}_t^{(kh} - C_t^{(kh}$.

Sin embargo podremos utilizar ahora el resultado de que se verifica (129) para un determinado valor de $\tau>0$ lo que solucionará el problema, como vamos a ver a continuación.

En efecto, como hemos puesto en evidencia antes en el otro caso, las dificultades en la definición de los conos de vectores tangentes se darían en aquellos puntos $\hat{x}^{(q}$ de $\overline{C}_t^{*(kh}$ tales que: $\sum\limits_{i=o}^{n} (x_i^{(q}(s)-\bar{x}_i(t))^2 = \sigma$ para algún s, o bien $a_h^{(q*}=\tau$, en donde por $a_h^{(q*}$ denotamos el correspondiente valor del a_h^* asociado a $\hat{x}^{(q}$

Vamos a ver, pues, como podría sortearse la dificultad que se originaria en la definición de los nuevos puntos perturbados para construir vectores tangentes en puntos $\hat{x}^{(q}$ de $\overline{C}_t^{*(kh}$ tales que $a_h^{(q*}=\tau$.

Sea $H_{(\sigma)}$ el conjunto de los puntos de $\overline{C}_t^{*(kh}$ que verifican la condición anterior y, supuesto $H_{(\sigma)}$ no vacio, pues en otro caso no habría nada que demostrar, denotemos por $d_{(\sigma)}$ la distancia del punto $\hat{\bar{x}}(t)$ a dicho conjunto, definida de la manera usual, es decir:

$$d_{(\sigma)} = \inf_{\hat{x}^{(q}\in H_{(\sigma)}} \left[\sum_{i=o}^{n} (x_i^{(q}-\bar{x}_i(t))^2\right]^{1/2}$$

Entonces, si $d_{(\sigma)}>0$, bastaria sustituir en la definición del conjunto $C_t^{*(kh}$ σ por $d_{(\sigma)}/2$ para que quedara garantizado que ningun punto $\hat{x}^{(q}$ de la clausura del nuevo conjunto $C_t^{*(kh}$ verificara $a_h^{(q*} = \tau$.

En caso anterior consideramos una sucesión decreciente a - cero de valores de σ, que denominaremos σ_r, y supongamos que ninguno de los conjuntos $H_{(\sigma_r)}$ es no vacio, pues en otro caso bastaria sustituir σ por el σ_r para el que el $H_{(\sigma_r)}$ es vacio para - que quedara garantizada la no existencia de puntos $\hat{x}^{(q}$ con la propiedad anterior.

Sea $d_{(\sigma_r)}$ la correspondiente sucesión decreciente de dis - tancias definidas entre $\hat{\bar{x}}(t)$ y cada uno de los $H_{(\sigma_r)}$. Entonces, o bien $\lim_r d_{(\sigma_r)}>0$, con lo que nos remitiremos al caso anterior, -- pues se seguiria de ello que si se sustituye σ por $\lim_r d_{(\sigma_r)}/2$, ningún punto $\hat{x}^{(q}$ de la clausura del nuevo conjunto $C_t^{*(kh}$ verificaría que $a_h^{(q*}= \tau$, o bien $\lim_r d_{(\sigma_r)} = 0$.

Ahora bien, en este último caso sería posible encontrar - puntos $\hat{x}^{(q}$ del conjunto original $\bar{C}_t^{*(kh}$ cuyas trayectorias evolucionaran durante un tiempo τ á una distancia tan pequeña como se deseara del punto $\hat{\bar{x}}(t)$. Sea pues $\hat{x}^{(qr}$ una sucesión de puntos de $\bar{C}_t^{*(kh}$ tal que:

$$\sum_{i=o}^{n} (x_i^{(qr}(s) - \bar{x}_i(t))^2 \leq \sigma_r$$

para todo s, con $0 \leq s \leq \tau$, y todo r.

Entonces, por el procedimiento descrito en la demostración de los lemas (125) y (138) (al que el hecho de que ahora los puntos de subdivisión $a_{ij}^{(qr}$ mediante las cuales se definirian las -- funciones de control que generarían las trayectorias $\hat{x}^{(qr}(s)$ fueran, en general, infinitos no añadiria ninguna dificultad nueva, ya que, en cualquier caso, el conjunto de éstos siempre sería numerable), se podría construir una subsucesión convergente de funciones a partir de la $\hat{x}^{(qr}(s)$, para $0\leq s\leq\tau$, y una subsucesión convergente casi seguro de las correspondientes funciones de control --

$u_a^{(qr}(s)$, para $o\leq s\leq \tau$, de manera que la función límite de las $\hat{x}^{(qr}(s)$, que denotaremos por $\hat{x}^{(q\#}(s)$, y la función límite, salvo un conjunto de medida nula, de las $u_a^{(qr}(s)$, que denotaremos por $u_a^{(q\,*}(s)$, verificarían:

$$\hat{x}^{(q\#}(s) = \hat{x}(t)+ \int_o^s \hat{f}\ (\hat{x}^{(q\#}(z),u_a^{(q*}(z))dz,$$

para todo s con $o\leq s\leq \tau$

Además como:

$$\sum_{i=o}^{n} (x_i^{(qr}- \bar{x}_i(t))^2< \sigma_r,$$

para todo s, $o\leq s\leq \tau$, y todo r, se seguiría entonces que:

$$\lim_r \sum_{i=o}^{n}(x_i^{(qr}(s) - \bar{x}_i(t))^2 = \sum_{i=o}^{n}(x_i^{(q\#}(s) - \bar{x}_i(t))^2=0$$

para todo s, $o\leq s\leq \tau$.

En particular, como para cada r se verificaría que $a_{1h}^{(qr}\geq \sum_i \frac{1}{2^i}\ a_{(i+1)h}^{(qr}$, se seguiría de ello que $a_{1h}^{(qr}\geq \frac{\tau}{2}$, para todo r, y tomando límites se concluiría que $a_{ih}^{(q*}\geq \frac{\tau}{2}$, en donde las $a_{ij}^{(q*}$ denotarían, al igual que las a^*_{ij} en la demostración del ema (138) los límites de las $a_{ij}^{(qr}$, en función de las cuales se habría definido $u_a^{(q*}(s)$.

Ahora ya que:

$$\sum_{j=1}^{h} (a_{ij}^{(q*} - a_{1\,j-1}^{(q*}) = a_{1h}^{(q*} >0$$

alguna de las diferencias $a_{1j}^{(q*} - a_{1(j-1)}^{(q*}$ sería estrictamente positiva.

Supongamos pues, que $a_{1j*}^{(q*} - a_{1(j^*-1)}^{(q*} >0$. Entonces de:

$$\hat{x}^{(q\#}(a_{1(j^*-1)}^{(q*}) = \hat{\bar{x}}(t)+ \int_o^{a_{1(j^*-1)}^{(q*}} \hat{f}(\hat{x}^{(q\#}(s),u_a^{(q*}(s)\ ds=\hat{\bar{x}}(t)$$

y de que se verifica también lo anterior si se sustituye $a_{1(j^*-1)}^{(q*}$ por $a_{1j^*}^{(q*}$ se deduciría que:

$$\int_{a^{(q^*}_{1(j^*-1)}}^{a^{(q^*}_{1j^*}} \hat{f}(x^{(q\#}(s), u_a^{(q^*}(s)) ds =$$

$$= \int_{a^{(q^*}_{1(j^*-1)}}^{a^{(q^*}_{1j^*}} \hat{f}(\hat{\bar{x}}(t), u^{(j^*}) ds = (a^{(q^*}_{1j^*} - a^{(q^*}_{1(j^*-1)}) .$$

$$\hat{f}(x(t), u^{(j^*}) = 0,$$

esto es $\hat{f}(\hat{\bar{x}}(t), u^{(j^*}) = 0$

Supongamos ahora que se redefiniera de nuevo los conjuntos $C_t^{(kh}$ y $C_t^{*(kh}$ pero eliminándo el valor $u^{(j^*}$ de entre aquellos - que podría tomar la función de control.

Si lleváramos entonces a cabo de nuevo una discusión como la que acabamos de mostrar se concluiría que, o bien existiría un valor de σ, que denominaremos σ^* quedaría garantizado que no po - dría existir ningún punto $\hat{x}^{(q}$ del conjunto $\bar{C}_t^{*(kh}$ con la propiedad de que $a_h^{(q^*} = \tau$ o bien $\hat{f}(\hat{\bar{x}}(t), u^{(j}) = 0$, para otro valor $u^{(j}$ dife - rente del $u^{(j^*}$.

El proceso podría continuarse, pues, hasta que, o bien para un determinado subconjunto no vacio de las $u^{(j}$, $j=1,2,....,h$, existiera el $\sigma^*>0$ con la propiedad indicada, o bien, $\hat{f}(\hat{\bar{x}}(t), u^{(j})$ $=0$, para todo j.

Nótese, por otra parte, que la nulidad de uno o varios de los vectores $f(\hat{\bar{x}}(t), u^{(j})$ permitiría eliminarlos de la discusión del problema de la existencia de un rector no nulo de multiplicadores para la colección de vectores $\hat{f}(\hat{\bar{x}}(t), u)$, con $u \in \Omega$, pues la conclusión de su existencia para aquellos vectores $\hat{f}(\hat{\bar{x}}(t), u) \neq 0$ podría extenderse inmediatamente al conjunto total.

De acuerdo con el razonamiento anterior y mediante una prolongación sencilla de su conclusión, podríamos suponer además que se puede escoger σ lo suficientemente pequeño para que no existan puntos $\hat{x}^{(q}$ de $C_t^{*(kh}$ con $a_h^{(q*} \geq \tau_0 < \tau$, para un $\tau_0 > 0$ dado de antemano.

La única dificultad en la definición de los puntos $x^{(qq}$ estribaría entonces en que pudiera ser $\sum_{i=o}^{n}(x_i^{(q}(s) - \bar{x}_i(t))^2 = \sigma$ para algún s.

Ahora bien, si $x^{(qr} \to \hat{\bar{x}}(t)$ entonces tendrá que ser $x^{(qr}(a$ $x^{(qr}(a_h^{(qr*}) < \sigma/2$, para todo r suficientemente grande, de lo que se seguiría que, si denotamos por $a_h^{(qr}$ el límite de alguna subsucesión de la $a_h^{(qr*}$, entonces existirá un s_o tal que $x^{(qr}(s) < \sigma$ para todo s con: $0 < s_o < s < a_h^{(q*}$ y todo r suficientemente grande.

Podemos suponer que τ_0 es $\to$ arbitrariamente pequeño y en particular menor que el τ para el que se ha establecido (129).

Por otra parte, ya que el τ de (129) podría encogerse a su vez tan pequeño como se quiera, entonces, si denotamos por τ^* al correspondiente a (129) y τ al del conjunto en cuestión, podríamos poner $\tau_0 < \tau^*$ y $\tau_0 + \tau^* < \tau$.

No analizaremos en detalle la cuestión de que lo anterior nos permitiría definir los correspondientes puntos perturbados -- $\hat{x}^{(qq}$ mediante los cuales se obtendrían los conos de vectores tangentes en cada uno de los puntos $\hat{\bar{x}}^{(qr}$ de la hipotética sucesión de puntos de $\bar{C}_t^{*(kh}$ convergente a $\hat{\bar{x}}(t)$, lo cual sería factible aún en el caso de que:

$$a_{(i+1)h}^{(qr} = \frac{1}{2^i} a_{1h}^{(qr}$$

para todo $i=1,2,\ldots, p_r$ y todo r, pues se agregaría por delante - la función del control $u_a^{(q*}$ (s) dada por (129), lo que añadiría - los grados de libertad necesarios.

En cuanto el tipo de vectores tangentes obtenidos se puede comprobar facilmente aunque no lo haremos aquí para no alargar más esta discusión, que estos podría conseguirse que fueran del si- - guiente:

(141) $$\sum_{j=1}^{h} \lambda_j \, \bar{\Phi}_u(q \; (a_h^{(q*}, a_{pj}^{(q}) \; \hat{f} \; (\hat{x}^{(q} (\sum_{i=o}^{p-1} a_{ih}^{(q} + a_{pj}^{(q)}), u^{(j)})$$

para cada punto $\hat{x}^{(q}$ de la sucesión, en donde los simbolos ya han sido definidos antes, y siendo p un determinado entero tal que se pudiera garantizar que:

$$\sum_{i=o}^{n} (x_i^{(q} (\sum_{i=1}^{q} a_{ih}^{(q}) - \bar{x}_i(t))^2 < \sigma$$

para $g=p, (p+1), \ldots$

Quedaría establecido, pues, de esta manera, un cono de vectores tangentes en cada punto de $\bar{C}_t^{(kh}$, del tipo (134), o en cada punto de $\bar{C}_t^{*(kh}$ del tipo (141) lo que permite la aplicación del - Principio (91) al problema dando lugar al siguiente lema que está en la base de la verificación del Principio de Máximo.

(142) LEMA

Si $\hat{\bar{x}}^{(t)}$, obtenida para el control $\bar{u}(t)$, con $t_o \leq t \leq t_1$, es una trayectoria óptima para el Problema (116) en el caso autónomo, con $G_o = \{x^{(o)}\}$, cumpliéndose las condiciones (118), y $\hat{f}$ y $\partial\hat{f} / \partial\hat{x}$ - son continuas en $\mathbb{R}^{n+1} \times \bar{\Omega}$, entonces, para todo t tal que $t_o < t < t_1$ existe un vector: $\hat{\bar{\eta}}(t) = (\eta_o(t)), \bar{\eta}_1(t), \ldots, \bar{\eta}_n(t)) \neq 0$, de manera que:

(143) $$\hat{\bar{\eta}}(t) \; \hat{f}(\hat{\bar{x}}(t), u) \leq 0$$

para todo $u \in \Omega$, y $\eta_o(t) \leq 0$

DEMOSTRACION

Consideremos $t \in (t_o, t_1)$ y para K y $h \in \mathbb{N}$ dados, con $k \geq \sum_{i=1}^{m} \bar{u}_i^{-2}(t_o)$ sean $C_t^{(kh}$ y $C_t^{*(kh}$ los conjuntos definidos mediante (124) o (138)

Como acabamos de discutir, y supuesto que $\hat{f}(\hat{\bar{x}}(t), u^{(j)}) \neq 0$ para $j = 1,2,\ldots,h$, o bien se podría definir un cono de vectores tangentes en cada punto $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$ del tipo (134) o bien en cada punto $\hat{x}^{(q}$ de $\bar{C}_t^{*(kh}$ del tipo (141).

Por otra parte, de acuerdo con los lemas (125) y (139) el punto $\hat{x}(t)$ pertenecería a la frontera inferior con respecto a su primera coordenada de cada uno de dichos conjuntos.

El que ambos conos son convexos y que se pueden considerar generados por los vectores:

$$\Phi_{u(q}(a_h^{(q}, a_j^{(q})\,\hat{f}\,(\hat{x}^{(q}(a_j^{(q}), u^{(j)})$$

$j = 1,2,\ldots,h$, en el primer caso, y

$$\Phi_{u(q}(a_h^{(q*}, a_{pj}^{(q})\,\hat{f}\,(\hat{x}^{(q}(\sum_{i=o}^{p-1} a_{ih}^{(q} + a_{pj}), u^{(j)})$$

para $j = 1,2,\ldots,h$, es inmediato, ya que se pueden escoger los λ_j de manera que $\lambda_j = \lambda\ \lambda'_j + (1-\lambda)\dot{\lambda}''_j$, y $\lambda_j = 0$, según los casos.

Además la afirmación aludida vale no solo para los puntos $\hat{x}^{(q}$ de $\bar{C}_t^{(kh} - C_t^{(kh}$ y $\bar{C}_t^{*(kh} - C_t^{*(kh}$ sino evidentemente también para los de los propios conjuntos $C_t^{(kh}$ y $C_t^{*(kh}$.

Estamos pues en condiciones de aplicar el Principio (91) a los conjuntos cerrados $\bar{C}_t^{(kh}$ y $\bar{C}_t^{*(kh}$ y al punto $\hat{\bar{x}}(t)$ perteneciente a la frontera inferior de ambos respecto a su primera coordenada.

Por brevedad referiremos solo la discusión al conjunto $\bar{C}_t^{(kh}$ ya que su extensión al $\bar{C}_t^{*(kh}$ es inmediata.

Denotemos por $L_{\hat{x}(q}$ el correspondiente cono convexo de vectores tangentes definidos en cada punto $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$. De acuerdo pues, con (91), entonces, o bien existe un vector:

$$\hat{\bar{\eta}}^{(kh}(t) = (\bar{\eta}_o^{(kh}(t),\ \bar{\eta}_1^{(kh}(t),\ \ldots,\ \bar{\eta}_n^{(kh}(t)) \neq 0$$

con $\bar{\eta}_o^{(kh}(t) \leq 0$, tal que:

(144) $\hat{\bar{\eta}}^{(kh}(t)\ \hat{w} \leq 0$, para todo $\hat{w} \in L_{\hat{\bar{x}}(t)}$

o bien existe una sucesión de puntos $\hat{x}^{(qr}$, $r \in \mathbb{N}$, pertenecientes a $\bar{C}_t^{(kh}$, y una sucesión de vectores $\hat{\bar{\eta}}^{(khr}(t) \neq 0$,

tales que:

$$\hat{\bar{\eta}}^{(khr}(t).\ \hat{w} \leq 0.$$

para todo $\hat{w} \in L_{\hat{x}(qr}$ y todo $r \in \mathbb{N}$

con $\bar{\eta}_o^{(khr}(t) \leq 0$, para todo r, y siendo $\lim\limits_{r} \hat{x}^{(q} = \hat{\bar{x}}(t)$

En este segundo caso tendríamos:

$$\hat{\bar{\eta}}^{(khr} \sum_{j=1}^{h} \lambda_j\ \Phi_{u_{(qr}}(a_h^{(qr},\ a_j^{(qr})\ \hat{f}\ (\hat{x}^{(qr}(a_j^{(qr}),u^{(j}) \leq 0$$

para cada colección de $\lambda_j \geq 0$

Se seguiría de ello que:

(145) $$\hat{\bar{\eta}}^{(khr}\ \Phi_{u_{(qr}}(a_h^{(qr},\ a_j^{(qr})\ \hat{f}\ (\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j}) \leq 0$$

para todo $j= 1,2,\ldots,h$.

Sin pérdida de generalidad, y ya que $\hat{\bar{\eta}}^{(khr}(t) \neq 0$ para todo r y que las desigualdades (144) y (145) son invariantes a transformaciones positivas de escala en los vectores $\hat{\bar{\eta}}^{(kh}(t)$ y $\hat{\bar{\eta}}^{(khr}(t)$, podemos suponer entonces que:

$|| \hat{\bar{\eta}}^{(kh}(t)||=1$, en donde $||\hat{x}||= \{\sum_{i=o}^{n} x_i^2\}^{1/2}$, y que $||\hat{\bar{\eta}}^{(khr}(t)||=1$ para todo $r \in \mathbb{N}$, y cualesquiera que sean los valores λ y σ caracterizadores del conjunto $C_t^{(kh}$.

Supongamos ahora que no pudiera garantizarse, en principio (144). En ese caso se podría garantizar (145).

Ahora, recordando la definición de las matrices fundamentales $\Phi_{u(q}(\cdot,\cdot)$, tendremos que en particular $\Phi_{u(q_r}(\cdot,\cdot)$ sería la matriz fundamental del sistema:

(146) $$\frac{d\xi(s)}{ds} = \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{(qr}(s),\ u^{(qr}(s)).\ \xi(s)$$

en el intervalo $0 \leq s \leq\ a_h^{(qr}$, en donde $u^{(qr}(s)$ denotaria la función de control en $[0,a_h^{(qr}]$ mediante la que se habría obtenido $\hat{x}^{(qr}(s)$

Si multiplicamos por la izquierda el anterior sistema por el vector $2\xi'(s)$, en donde $\xi'(s)$ representa el vector fila: $(\xi_o(s),\ \xi_1(s),\dots,\xi_n(s))$, tendríamos:

$$2\xi'(s)\ \frac{d(\xi(s)}{ds} = \frac{d}{ds}\{\xi'(s).\xi(s)\}=$$

$$= 2\ \xi'(s)\ \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{(qr}(s),\ u^{(qr}(s)).\xi\ (s).$$

Por otra parte, como se sigue de la desigualdad de Cauchy (véase la Nota 18), podríamos poner:

$$\xi'(s)\ \frac{\partial \hat{f}}{\partial \hat{x}}\ (\hat{x}^{(qr}(s),\ u^{(qr}(s)).\xi(s) \leq |\xi'(s)\frac{\partial \hat{f}}{\partial \hat{x}}\ (\hat{x}^{(qr}(s)u^{(qr}(s)).$$

$$.\xi(s)| \leq ||\ \xi(s)||\quad ||\frac{\partial \hat{f}}{\partial \hat{x}}\ (\hat{x}^{(qr}(s),u^{(qr}(s))\xi(s)\ ||.$$

Además:

$$||\frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{(qr}(s), \quad u^{(qr}(s)).\xi (s)||$$

$$=[\sum_{i=o}^{n} [\frac{\partial f_i}{\partial \hat{x}} (\hat{x}^{(qr}(s), u^{(qr}(s)).\xi (s)]^2] ^{1/2} \leq$$

$$\leq [\sum_{i=o}^{n} ||\frac{\partial f_i}{\partial \hat{x}}(\hat{x}^{(qr}(s), u^{(qr}(s)||^2||\xi(s)||^2] ^{1/2} =$$

$$= [\sum_{i=o}^{n} \sum_{j=o}^{n} \cdot [\frac{\partial f_i}{\partial x_j} (\hat{x}^{(qr}(s), u^{(qr}(s)]^2]^{1/2}||\xi(s)||$$

$$=||\frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{(qr}(s), u^{(qr}(s)|| ||\xi (s) ||$$

en donde $\frac{\partial f_i}{\partial \hat{x}}$ (.,.) denota la fila i-esima de la matriz

$\frac{\partial \hat{f}}{\partial \hat{x}}$ (.,.) y $||\frac{\partial f}{\partial x}$ (.,.)$||$ su norma, definida como antes.

Ahora bien, de la pertenencia de cada punto de la sucesión $\hat{x}^{(qr}$ al conjunto $C_t^{-(kh}$ caracterizado por el valor σ se seguiría que:

$$\sum_{i=o}^{n} (x_i^{(qr} (s) - \bar{x}_i (t))^2 \leq \sigma <^1$$

y también que $u^{(qr}$ (s) e $D^{(kh}$, para todo $s \in [0,a_h^{(qr}]$ y todo $r \in \mathbb{N}$.

Por otro lado, de la continuidad de la matriz $\frac{\partial \hat{f}}{\partial \hat{x}}$ $(\hat{x},u)$ en $(\hat{x},u) \in \mathbb{R}^{n+1} \times \bar{\Omega}$ se seguiría que $||\frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x},u)||$ estaria acotada en el compacto de $\mathbb{R}^{n+m+1}$ dado por (126) por alguna constante positiva M_k^*.

Denotemos por $\xi^{(rj}(s)$, con j=1,2, ...,h, la solución del sistema (146) en el intervalo $[a_j^{(qr}, a_h^{(qr}]$ con la condición -- inicial:

$$\xi^{(rj}(a_j^{(qr}) = \hat{f} (\hat{x}^{(qr} (a_j^{(qr}), u^{(j}).$$

Tendremos entonces:

$$\xi^{(rj}(a_h^{(qr}) = \Phi_u(qr\,(a_h^{(qr},\ a_j^{(qr})\hat{f}(\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j}).$$

Sea, por otra parte, $\alpha^{(rj}(s)$ la solución de la ecuación diferencial escalar:

$$\frac{d\alpha(s)}{ds} = 2\ M_k^*\ \alpha(s)$$

en el intervalo $[a_j^{(qr},\ a_h^{(qr}]$ con la condición inicial:

$$\alpha^{(rj}(a_j^{(qr}) = \sum_{i=o}^{n} [f_i\ (\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j})]^2$$

De acuerdo con la desigualdad obtenida antes se tendría entonces.

(147)
$$\sum_{i=o}^{n} [\xi_i^{(rj}(s)]^2 \leq \alpha(s) = e^{2M_k^*(s-a_j^{(qr})}\ .\alpha^{(rj}(a_j^{(qr})$$

para todo $s \in [a_j^{(qr},\ a_h^{(qr}]$.

En efecto, ya que:

$$\frac{d}{ds}\ [\xi^{(rj'}(s)\ \xi^{(rj}(s)] = 2M_k^*\ [\xi^{(rj'}(s)\xi^{(rj}(s)] +$$

$$+ [2\xi^{(rj'}(s)\ \frac{\partial\hat{f}}{\partial\hat{x}}\ (\hat{x}^{(qr}(s),\ u^{(qr}(s))\ \xi^{(rj}(s) -$$

$$2M_k^*\ \xi^{(rj'}(s)\ \xi^{(rj}(s)]$$

se seguiría entonces, recordando la expresión general de la solución de una ecuación diferencial líneal no homogénea:

$$\xi^{(rj'}(s)\xi^{(rj}(s) = e^{2M_k^*\ (s-a_j^{(qr})}[\xi^{(rj'}(a_j^{(qr}).$$

$$\xi^{(rj}(a_j^{(qr})] + \int_{a_j^{(qr}}^{s} e^{2M_k^*\ (s-z)}\ y^{(rj}(z)\ dz$$

$$\leq e^{2M_k^*\ (s-a_j^{(qr})}[\xi^{(rj'}(a_j^{(qr}).\ \xi^{(rj}(a_j^{(qr})]$$

ya que:

$$y'^{(rj}(z) = [2\xi^{(rj'}(z)\ \frac{\partial \hat{f}}{\partial \hat{x}}\ (\hat{x}^{(qr}(z),\ u^{(qr}(z))\ \xi^{(rj}(z) -$$

$$2M_k^*\ \xi^{(rj'}(z)\ \xi^{(rj'}(z)\ \xi^{(rj}(z)] \leq 0$$

para todo $z \in [a_j^{(qr}, s]$.

Por lo tanto, ya que (147) es válido para todo j y r, pode mos poner entonces:

$$(148) \qquad \sum_{i=o}^{n} [\xi_i^{(rj}(s)]^2 \leq (M_k^*)^2\ e^{2M_k^* \tau}$$

para todo $s \in [a_j^{(qr}, a_h^{(qr}]$, todo j =1,2, ..., h, y todo $r \in \mathbb{N}$.

Por otra parte, podremos poner:

$$\Phi_{u^{(qr}}(a_h^{(qr}, a_j^{(qr})\ \hat{f}\ (\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j}) =$$

$$= \hat{f}(\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j}) + \int_{a_j^{(qr}}^{a_h^{(qr}} \frac{\partial \hat{f}}{\partial \hat{x}}\ (\hat{x}^{(qr}(s),\ u^{(qr}(s)\ ).$$

$$\xi^{(rj}(s)\ ds.$$

Se seguiría entonces que:

$$|\ \Phi^{(i}_{u^{(qr}}\ (a_h^{(qr}, a_j^{(qr})\ \hat{f}\ (\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j}) -$$

$$- f_i\ (\hat{x}^{(qr}(a_j^{(qr}),\ u^{(j})| \leq \int_{a_j^{(qr}}^{(qr} |\frac{\partial f_i}{\partial \hat{x}}(\hat{x}^{(qr}(s),\ u^{(qr}(s)).$$

$$.\ \xi^{(rj}(s)|ds$$

para i = 0,1,2,... n, en donde $\Phi_u{(qr}^{(i}(.,.)$ denotaria la i-esima fi la de la matriz $\Phi_{u^{(qr}}(.,.)$.

Ahora, como:

$$|\ \frac{\partial f_i}{\partial \hat{x}}\ (\hat{x}^{(qr}(s),\ u^{(qr}(s)).\xi^{(rj}(s)| \leq$$

$$\leq \left\| \frac{\partial \hat{f}}{\partial \hat{x}} (\hat{x}^{(qr}(s), u^{(qr}(s)) \right\| \, \| \xi^{(rj}(s) \| \leq$$

$$\leq M^*_k \quad (M^*_k)^2 \; e^{2M^*_k} \, \tau$$

podremos poner:

$$| \, \Phi^{(i}_{u^{(qr}} \; (a_h^{(qr}, \, a_j^{(qr}) . \; \hat{f} \; (\hat{x}^{(qr}(a_j^{(qr}), u^{(j}) \; -$$

$$- \; \hat{f}(\hat{x}^{(qr}(a_j^{(qr}), \; u^{(j}) \; | \; \leq \; \tau (M^*_k)^3 \quad e^{2M^*_k} \, \tau ,$$

para todo $i = 0,1,2,\ldots, n$ y por ello:

(149) $$\| \, \Phi_{u^{(qr}} \; (a_h^{(qr}, a_j^{(qr}) . \; \hat{f}(\hat{x}^{(qr}(a_j^{(qr}), \; u^{(j}) \quad -$$

$$\hat{f}(\hat{x}^{(qr}(a_j^{(qr}), \; u^{(j}) \| \; \leq \; \tau \; (n+1)^{1/2} \; (M^*_k)^3 \; e^{2M^*_k} \, \tau$$

para todo $j = 1,2, \ldots, h$ y todo $r \in \mathbb{N}$.

Consideremos ahora una sucesión doble simultáneamente decreciente a cero de valores de τ y de σ, en las condiciones que se han indicado en la discusión precedente al lema para garantizar la posibilidad de definición de los correspondientes conos convexos de vectores tangentes en cada punto de $\overline{C}_t^{(kh}$, la cual denotaremos por (τ_d, σ_d), $d \in \mathbb{N}$.

Entonces, para cada valor de d existiria una sucesión de puntos $x^{(qd^r}$ del correspondiente conjunto $\overline{C}_t^{(kh}$, y una sucesión de ectores $\hat{\eta}^{(khdr}_{\tau_d \sigma_d}(t)$, para $r \in \mathbb{N}$ para los que se verificará (145). Por lo tanto, para cada $d \in \mathbb{N}$ podría escogerse un punto $\hat{x}^{(qdr}$ de la correspondiente sucesión definida para dichos τ_d y σ_d, de manera que $r \leq d$ y que para dicho punto se verificara (145) para un determinado vector

$$\hat{\eta}^{(kh\,d}_{\tau_d \sigma_d}(t) \neq 0, \text{ con } \overline{\eta}^{(khd}_{\tau_d \sigma_d 0}(t) \leq 0.$$

Denotemos de nuevo por $\hat{x}^{(qr}$ a la subsucesión de las $\hat{x}^{(qdr}$, para d = 1,2, ..., así elegida y por $\hat{\bar{\eta}}^{(khr}_{\tau_d \sigma_d}(t)$ a la correspondiente subsucesión de los vectores $\hat{\bar{\eta}}^{(khd}(t)$, que supondremos normalizada, esto es $||\hat{\bar{\eta}}^{(khr}(t)||=1$, para todo r.

Tendremos entonces, ya que:

$$\sum_{i=o}^{n} (x_i^{(qr}(s) - \bar{x}_1(t))^2 \leq \sigma_r$$

para todo $s \in [0, a_h^{(qr}]$ y todo $r \in \mathbb{N}$, que:

$$\lim_r \hat{x}^{(qr}(a_j^{(qr}) = \hat{\bar{x}}(t)$$

lo que, de acuerdo con (149) y teniendo en cuenta que:

$$\lim_r \tau_r (n+1)^{1/2} (M_k^*)^3 e^{2M_k^* \tau_r} = 0$$

y la continuidad de $\hat{f}$, daría:

$$\lim_r \Phi_{u^{(qr}}(a_h^{(qr}, a_j^{(qr}). \hat{f}(\hat{x}^{(qr}(a_j.^{(qr}), u^{(j}) =$$

$$= \lim_r \hat{f}(\hat{x}^{(qr}(a_j^{(qr}), u^{(j}) = \hat{f}(\hat{\bar{x}}(t), u^{(j})$$

para todo j = 1,2, ..., h.

Entonces si se seleccionase una subsucesión convergente de la $\hat{\bar{\eta}}^{(khr}(t)$, que volveremos a denotar por $\hat{\bar{\eta}}^{(khr}(t)$, y si $\hat{\bar{\eta}}^{(kh}(t)$ fuera su límite tendríamos:

$$||\hat{\bar{\eta}}^{(kh}(t)||= 1, \text{ y } \bar{\eta}_o^{(kh^{(kh}}(t) \leq 0$$

y además, de que se verificaría (145) para todo j = 1,2, ..., h y todo $r \in N$, se obtendría finalmente que:

(150) $\hat{\bar{\eta}}(kh_{(t)} \cdot \hat{f}(\hat{\bar{x}}(t), u^{(j}) \leq 0$

para todo $j = 1,2,\ldots,h$.

Por otra parte, y mediante un razonamiento similar, podría con-- cluirse también (150) en el caso de que hubiera que remitirse al conjunto $C^{-*(kh}$ en vez del $C_t^{-(kh}$.

Además, el resultado (150) podría ampliarse evidentemente también a todos aquellos vectores nuloes $\hat{f}(\hat{\bar{x}}(t), u^{(j})$ que hubiera habi- do que suprimir para garantizar la licitud del proceso de genera ción de conos convexos de vectores tangentes en cada uno de los puntos $\hat{x}^{(q}$ de $C_t^{-*(kh} - C_t^{(kh}$, tal como ya se ha indicado.

Podemos, pues, suponer que para cada k y h existiría un vector $\hat{\bar{\eta}}^{(kh}(t)$, tal que: $||\hat{\bar{\eta}}^{(kh}(t)|| = 1$, supuesto que para todo k y - h la sucesión $\hat{\bar{\eta}}^{(khr}(t)$ se hubiera elegido de manera que $||\hat{\bar{\eta}}^{(khr}(t)|| = 1$, para todo r,y con $\bar{\eta}_o^{(kh}(t) \leq 0$, para el que se verificaría - (150) para todo $u^{(j}$ perteneciente al conjunto $D^{(kh}$.

Seleccionemos ahora para cada k una subsucesión convergente de - la $\hat{\bar{\eta}}^{(kh}(t)$, para $h \in N$, a un determinado vector, que denominare-- mos $\hat{\bar{\eta}}^{(k}(t)$, para el que se verificaría que: $||\hat{\bar{\eta}}^{(k}(t)|| = 1$, y - $\bar{\eta}_o^{(k}(t) \leq 0$.

Recordemos entonces la definición de los conjuntos $A^{(k}$ y $D^{(k}$, pa ra $k \geq \sum_{i=1}^{m} u_i^2(t_o), k \in \mathbb{N}$, obtendremos, por un paso al límite en (150), que:

(151) $\hat{\bar{\eta}}^{(k}(t) \cdot \hat{f}(\hat{\bar{x}}(t), u) \leq 0$

para todo $u \epsilon D^{(k}$, ya que para cada $u^{*} \epsilon D^{(k}$ existiría un índice h^{*} tal que $u^{*} \epsilon D^{(kh}$ si $h \geq h^{*}$.

Ahora, como $D^{(k}$ es denso en $A^{(k}$, podríamos ampliar la validez de (151) a $A^{(k}$ ya que si $u^{*} \epsilon A^{(k} - D^{(k}$, se podría construir una sucesión $u^{(r}$ contenida en $D^{(k}$ de manera que $u^{(r} \to u^{*}$.

Entonces, de que tendría que ser:

$$\hat{\overline{\eta}}^{(k}(t) \cdot \hat{f}(\hat{\overline{x}}(t), u^{(r}) \leq 0$$

para todo $u^{(r}$, y de la continuidad de $\hat{f}$, se seguiría el resultado enunciado.

Por último, podríamos seleccionar una subsucesión convergente de la $\hat{\overline{\eta}}^{(k}(t)$, para $k \in \mathbb{N}$ y $k \geq \sum_{i=1}^{m} \overline{u}_i^2(t_o)$, convergente a un vector $\hat{\overline{\eta}}^{*}(t)$ para el que se verificaría: $||\hat{\overline{\eta}}^{*}(t)|| = 1$ y $\overline{\eta}_o^{*}(t) \leq 0$.

En relación al vector anterior podría afirmarse entonces que:

$$\hat{\overline{\eta}}^{*}(t) \cdot \hat{f}(\hat{\overline{x}}(t), u) \leq 0$$

para todo $u \epsilon \Omega$, ya que para cada $u^{*} \epsilon \Omega$ existiría un índice k^{*} tal que $\mu^{*} \epsilon A^{(k}$, para todo $k \geq k^{*}$. Esto completa la demostración.

El Lema (142) nos permitirá caracterizar fácilmente el cono de vectores tangentes que podría llegar a definirse en cada punto regular t de la función de control $\overline{u}(s)$ para $s \epsilon [t_o, t_1]$, aunque hay que observar que la conclusión de dicho Lema se aplicaría ya fuera t un punto regular de $\bar{u}(s)$ o no.

En lo que sigue, y antes de proceder a enunciar en todos sus términos el Principio de Máximo, vamos a mostrar, pues, como podría ampliarse la colección de vectores tangentes definida en cada punto $\hat{\bar{x}}(t)$ de la trayectoria óptima por la envoltura convexa de los vectores $\lambda.\hat{f}(\hat{\bar{x}}(t),u)$, con $u\epsilon\Omega$ y $\lambda \geq o$, tal como hemos obtenido antes, o la envoltura convexa de todos los vectores del tipo:

(152) $\quad \lambda.\Phi(t,\ t-\Delta t)\ \hat{f}(\hat{\bar{x}}(t-\Delta t),u)$, para todo $u\epsilon\Omega$,

y:

(153) $\quad -\lambda.\Phi(t,\ t-\Delta t)\ \hat{f}(\hat{\bar{x}}(t-\Delta t),u(t-\Delta t))$,

en donde $\lambda \geq 0$, $t_o < t-\Delta t \leq t < t_1$, y siendo $t-\Delta t$, para los valores de Δt bajo consideración, un punto regular de la función de control $\bar{u}(s)$ en el intervalo $t_o \leq s \leq t_1$, mientras que $\Phi(.,.)$ denotaría la matriz fundamental del siguiente sistema lineal de ecuaciones diferenciales con coeficientes variables:

(154) $$\frac{d\xi(s)}{ds} = \frac{\partial\hat{f}}{\partial\hat{x}}(\hat{\bar{x}}(s),\ \bar{u}(s)).\ \xi(s)$$

para $s\epsilon[t_o,\ t_1]$, cuya existencia estaría garantizada dadas las condiciones de continuidad supuestas a $\partial\hat{f}/\partial\hat{x}$ y la acotación de la función $\bar{u}(s)$ en el intervalo $[t_o,\ t_1]$.

En primer lugar vamos a comprobar que si t fuera un punto regular de $\overline{u}(s)$ la colección $\hat{f}(\hat{\overline{x}}(t),u)$, a que se refiere el Lema (142), con $u\varepsilon\Omega$ puede extenderse para incluir también al vector $\hat{f}(\hat{\overline{x}}(t), \overline{u}(t))$.

Sean, pues, $t\varepsilon(t_o, t_1)$ y δ una cantidad suficientemente pequeña para que se verifique: $t_o < t - \delta < t_1$.

Podemos poner entonces:

$$\hat{\overline{x}}(t) = \hat{\overline{x}}(t-\delta) + \int_{t-\delta}^{t} \hat{f}(\hat{\overline{x}}(s),\overline{u}(s))\, ds \quad ,$$

de lo que se seguiría que:

$$\hat{\overline{x}}(t-\delta) = \hat{\overline{x}}(t) + \int_{t}^{t-\delta} \hat{f}(\hat{\overline{x}}(s), \overline{u}(s))\, ds$$

con $\hat{\overline{x}}(t-\delta)\varepsilon\hat{C}$, para todo δ suficientemente pequeño, en donde C denota el conjunto dado por (117) en el caso autónomo con $x(t_o)=x^{(o}$.

Ahora,si t es un punto regular de $u(s)$ podemos poner, de acuerdo con (121):

$$\int_{t}^{t-\delta} \hat{f}(\hat{\overline{x}}(s), \overline{u}(s))\, ds = (-\delta)\, \hat{f}(\hat{\overline{x}}(t), \overline{u}(t)) + o\,(\delta).$$

Por lo tanto tendremos:

$$(15?) \quad \lim_{\delta \downarrow 0} \frac{\hat{\overline{x}}(t-\delta) - \hat{\overline{x}}(t)}{\delta} = - \hat{f}(\hat{\overline{x}}(t), \overline{u}(t)).$$

Vamos a ver entonces como se ampliaría el cono de vectores tangentes en $\hat{x}(t)$, incluyendo (155):

Para $0 \leq \lambda \leq 1$, $\delta>0$, $t - \lambda\delta \in (t_o,t)$, $u^{(j}$, $j = 1,2,\ldots,p$, pertenecientes a Ω, y $\lambda_j \geq 0$, $j = 1,2,\ldots,p$, consideremos el punto de C definido de la siguiente manera:

$$\hat{x}_\delta(t-\lambda\delta+(1-\lambda)(\sum_{j=1}^{p}\lambda_j)\delta) = \tag{156}$$

$$= \hat{x}(t) + \int_t^{t-\lambda\delta} f(\hat{x}(s),\bar{u}(s))ds + \int_{t-\lambda\delta}^{t-\lambda\delta+(1-\lambda)(\sum_{j=1}^{p}\lambda_j)\delta} f(\hat{x}_\delta(s),\mu_\delta(s-(t-\lambda\delta)))ds$$

en donde

$$\hat{x}_\delta(s) = \hat{x}(t-\lambda\delta) + \int_{t-\lambda\delta}^{s} f(\hat{x}_\delta(z), u_\delta(z-(t-\lambda\delta)))\,dz$$

para $t-\lambda\delta \leq s \leq t-\lambda\delta+(1-\lambda)(\sum_{j=1}^{p}\lambda_j)\delta$, y

siendo $u_\delta(s)$ la función de control definida en cada uno de los subintervalos:

$$(1-\lambda)(\sum_{g=0}^{j-1}\lambda_g)\delta \leq s \leq (1-\lambda)(\sum_{g=0}^{j}\lambda_g)\delta$$

para $j = 1,2,\ldots,p$, con $\lambda_o = 0$, constantemente igual a $u^{(j}$.

Entonces, ya que t es un punto regular para $\bar{u}(s)$ tenemos:

$$\lim_{\delta\downarrow 0}\frac{1}{\delta}\ \hat{x}(t-\lambda\delta+(t-\lambda)(\sum_{j=1}^{p}\lambda_j)\delta) - \hat{x}(t) = \tag{167}$$

$$= -\lambda\ \hat{f}(\hat{\bar{x}}(t),\ \bar{u}(t)) + (1-\lambda) \sum_{i=1}^{p} \lambda_i\ \hat{f}(\hat{\bar{x}}(t),\ u^{(j})$$

Por lo tanto, podremos definir mediante el procedimiento - anterior un cono convexo de vectores tangentes en $\hat{\bar{x}}(t)$ del tipo (157) para cada $t\varepsilon(t_o,\ t_1)$ que sea un punto regular de $\bar{u}(s)$.

No obstante, debido a que la aplicación del Principio (91) exige que el correspondiente conjunto C a que se aplique deba de ser cerrado, vamos a restringir nuestra definición de conos convexos de vectores tangentes a unos ciertos subconjuntos cerrados de $\bar{C}$.

(Observese, por otra parte, que debido a la razón de que durante el intervalo $(t_o,\ t-\lambda\delta)$ para $\delta\downarrow 0$ el control aplicado es $\bar{u}(s)$ y no otro el que el cono cerrado quedará definido en el - punto $\hat{\bar{x}}(t)$ y no en otro punto $\hat{x}$ de C, como también sería posible hacerlo, y que la construcción indicada del punto $\hat{x}_\delta(.)$ dado por (156) garantizaría que $\hat{x}_\delta(.)\varepsilon C$ para todo $\delta\downarrow 0$, lo cual podría no ser así si se hubieran propuesta otras posibles definiciones de vectores perturbados a partir del $\hat{\bar{x}}(t)$, tal como podría haber sido el caso si se sustituyera la función $\bar{u}(s)$ en el primer integrando del segundo miembro de (156) por un valor constante arbitrario u^* y se pusiera:

$$\hat{x}_\delta(t-\lambda\delta) = \hat{\bar{x}}(t) + \int_t^{t-\lambda\delta} \hat{f}(\hat{\bar{x}}(s), u^*)\ ds \quad .)$$

Entonces para que el resultado del Lema (142) pudiera ampliarse a los vectores (157), bastaría introducir unas pequeñas modificaciones en la definición de los conjuntos $C_t^{(kh}$ dados por

(124) y $C_t^{*(kh}$ dados por

(138) para permitir la inclusión de vectores tangentes del tipo:

(158) $\quad - \Phi(a_h^{(q}, t-d^{(q}) \hat{f}(\hat{\bar{x}}(t-d^{(q}), \bar{u}(t-d^{(q}))$

en el caso del conjunto $C_t^{(kh}$, y del tipo:

(159) $\quad - \Phi(a_h^{(q*}, t-d^{(q}) \hat{f}(\hat{\bar{x}}(t-d^{(q}), \bar{u}(t-d^{(q})$

en los integrantes del correspondiente cono convexo, siendo - $d^{(q} \geq 0$, $t-d^{(q} \varepsilon (t_o, t_1)$ y un punto regular de $\bar{u}(s)$.

Con más precisión, la redifinición de los puntos del nuevo conjunto $C_t^{(kh}$, sería ahora la siguiente:

(160) $\quad C_t^{(kh} = \{ x^* = (x_o^*, x_1^*, x_2^*, \ldots, x_n^*)' \varepsilon R^{n+1} /$

Existen d_1, a_{10}, $a_{11}, \ldots, a_{1n}$, a_{io}, d_{ii} y b_{ij}, para $i=1,2, \ldots, p$ y $j=0,1,2,\ldots,(h+2)$ de manera que $a_{10}=0$, $b_{ij}=0$, $d_{00}=0$, - $a_{io}=b_{io}=0$, y $d_{ii} \geq 0$, $b_{ij} \geq 0$, para toda i y j, y existe además una función de control $u_a^p(s)$ definida en el intervalo:

$$t-d_p \leq s \leq t-d_p + a_{1(h+2)} + \sum_{i=1}^{p} \sum_{j=1}^{h+2} b_{ij},$$ de la siguiente manera:

$$u_a^{(p}(s) = u^{(j}, \text{ para } t-d_p+a_{p(j-1} \leq s \leq t-d_p+a_{pj}, j=1,2,\ldots,h$$

$$= \bar{u}(t), \text{ para } t-d_p+a_{ph} \leq s \leq t-d_p+a_{p(h+1)}$$

$$= \bar{u}(s-(a_{p(h+2)}-a_{p(h+1)})), \text{para: } t-d_p+a_{p(h+1)} \leq s \leq t-d_p+a_{p(h+2)}$$

en donde:

$$d_{i+1} = \sum_{r=1}^{i} d_{rr}$$

$$d_{ij} = a_{i(j-1,} + \sum_{r=1}^{j} b_{ir}, \text{ para } i=1,2\ldots,p, \text{ y } j=1,2,\ldots,(h+2)$$

siendo $u^{(j} \varepsilon D^{(kh}$, para $j=1,2,\ldots,h$, y se verifica:

$$\hat{x}^{*} = \hat{x}^{(p}(t-d_p+a_{p(h+2)}) = \hat{\bar{x}}(t-d_p) + \int_{t-d_p}^{t-d_p+a_{p(h+2)}} \hat{f}(\hat{x}^{(p}(s), u_a^{(p}(s))ds$$

con:

$$\hat{x}^{(p}(s) = \hat{\bar{x}}(t-d_p) + \int_{t-d_p}^{s} \hat{f}(\hat{x}^{(p}(z), u_a^{(p}(z))\, dz$$

para todo $s \; \varepsilon [t-d_p, t-d_p + a_{p(h+2)}]$

Además existen $\tau>0$ y $0 < \sigma < 1$, tales que $a_{p(h+1)} < 2d_p$

$$d_p < \tau \; , \; a_{p(h+1)} - a_{ph} > \frac{1}{4} d_p, \text{ y:}$$

$$\sum_{i=0}^{n} (x_i^{(p}(s) - \bar{x}_i(t))^2 < \sigma$$

para todo $s \in [t-d_p, t-d_p + a_{p(h+2)}$, y por último: $a_{p(h+2)} -$
$- a_{p(h+1)} = \frac{1}{4} d_p$, y siendo $t-d_i$, $i=1,2,\ldots,p$, y $t-\tau$ pertene-- cientes a (t_o, t_1) y puntos regulares de $\bar{u}(s)$.

Por otra parte, la definición del nuevo conjunto $C_t^{*(kh}$ se- ría ahora:

(161) $\quad C_t^{*(kh} = \{ \; x^* = (x_0^*, x_1^*, x_2^*, \ldots, x_n^*)' \in R^{n+1} /$

Existen d_p, y a_{ij}, con $i=1,2,\ldots,p$, $j=0,1,2,\ldots,(h+1)$, de manera que $a_{i(j-1)} \le a_{ij}$, $d_{oj} = 0$, $a_{io} = 0$, para todo i y j, y - $d_{ii} \ge 0$, siendo además $a_{i(h+1)} > 0$, para todo $i=1,2,\ldots,p$, para evitar ambiguedades, y siendo $d_p < \tau$, $a_{1(h+1)} < 2\, d_p$, para un determinado $\tau > 0$, y también : $a_{i(h+1)} - a_{ih}\; 1/4^i\, d_p$, a la que correspondería la función de control $u_a^{(p}(s)$, definida de la siguiente manera:

$$u_a^{(p}(s) = u^{(j}, \text{ para } t-d_p + \sum_{i=0}^{g-1} a_{i(h+1)} + a_{g(j-1)} \le s \le$$

$$\le t-d_p + \sum_{i=0}^{g-1} a_{i(h+1)} + a_{gj}$$

con $j = 1,2,\ldots,h$

$$= \overline{u}(t), \text{ para } t-d_p + \sum_{i=0}^{g-1} a_{i(h+1)} + a_{gh} \le s \le$$

$$\le t-d_p + \sum_{i=0}^{g-1} a_{i(h+1)} = a_{g(h+1)}$$

y la trayectoria $\hat{x}^{(p}(s)$, definida de la siguiente manera:

$$\hat{x}^{(p}(s) = \hat{\hat{x}}(t - \sum_{i=1}^{p} d_{ii}) + \int_{t-d_p}^{s} \hat{f}(\hat{x}^{(p}(z), u_a^{(p}(z))\, dz$$

para todo $s \in [t-d_p,\; t-d_p + \sum_{i=1}^{p} a_{i(h+1)}]$, y en donde

$a_{(i+1)(h+1)} \le \frac{1}{p\, 2^i}\, a_{i(h+1)}$, para todo i, de lo que se deducirá entonces que $\sum_{i=1} a_{i(h+1)} < 4\tau$, $t-d_p$ y $t-\tau$ pertenecientes a $(t_o,\, t_1)$

y puntos regulares de $\bar{u}(s)$, y con $\hat{x}^* = \hat{x}^{(p}(t-d_p + \sum_{i=1}^{p} a_{i(h+1)})$.

Además existiría un $\sigma>0$ de manera que:

$$\sum_{i=0}^{n} (x_i^{(p} (s) - \bar{x}_i(t))^2 < \sigma$$

para todo s. }

Comentaremos ahora brevemente el por qué de la definición del conjunto (160) en la forma que lo hemos hecho.

En primer lugar hay que observar que la restricción de que $d_p > 0$ si $\hat{x}^* \neq \hat{\bar{x}}(t)$ no impediría que se pudieran ir definiendo puntos perturbados adecuados para generar toda la gama de vectores tangentes posibles ya que este efecto quedaría compensado por la posibilidad de que $u^{(j} = \bar{u}(t)$ de manera que quedaría controlado el tamaño relativo del vector $-\hat{f}(\hat{x}^{(q}, \bar{u}(t))$, frente a los vectores $\hat{f}(\hat{x}^{(q}, u^{(j})$.

En segundo lugar observamos que es fácil demostrar que el enunciado del Lema (125) se verifica si se cambia la referencia a -(124) por la (160), pues salvo en lo que se refiere a la introducción de la función $\bar{u}(s)$ (que recordamos que estaría acotada en $[t_o, t_1]$ en algunos tramos de la definición de las funciones $u_a^{(p}(s)$, lo que no añadiría ninguna dificultad especial, el resto es igual en ambos casos.

Ahora bien, como ya hemos comentado en el momento de generar conos convexos de vectores tangentes en cada punto $\hat{x}^{(q}$ de -

$C_t^{(kh}$, aunque si τ se encogiera suficientemente pequeño se podría garantizar que:

$$\sum_{i=0}^{n} (\hat{x}^{(q}(s) - \bar{x}_i(t))^2 < \sigma$$

para todo s, siendo $\hat{x}^{(q}(s)$ la trayectoria definitoria del punto - $\hat{x}^{(q}$ de $\bar{C}_t^{(kh}$ podría darse el caso que para el τ en cuestión - - existiera una sucesión de punto $\hat{x}^{(qr}$ de $C_t^{-(kh}$ convergente a $\hat{\bar{x}}(t)$ con trayectorias definitorias en la máxima longitud admisible de su intervalo de definición.

En nuestro caso, y debido a la manera en como se ha introducido la dependencia de la longitud de dichos intervalos del retardo $t-d_p$ resultaría que ello solo podría darse si cada $\hat{x}^{(qr}$ viniera definido de la siguiente manera:

$$(162) \qquad \hat{x}^{(qr} = \hat{\bar{x}}(t-\tau) + \int_{t-\tau}^{t+\tau} \hat{f}(\hat{x}^{(qr}(s), u^{(qr}(s))\, ds + \int_{t+\tau}^{t+(5/4)\tau} \hat{f}(\hat{x}^{(qr}(s), \bar{u}(s-(5/4)\tau))\, ds$$

para todo $r \in N$, siendo $t-\tau$ un punto regular de $\bar{u}(s)$.

Tomando límites en (162), es fácil comprobar, por argumentos similares a los utilizados para establecer (129) que existiría una función de control admisible $u^{(q*}$ y una trayectoria asociada $\hat{x}^{(\#}(s)$ de manera que:

$$(163) \qquad \hat{\bar{x}}(t-\tau) + \int_{t-\tau}^{t+\tau} \hat{f}(\hat{x}^{(q\#}(s), u^{(q*}(s))ds + \int_{t+\tau}^{t+(5/4)\tau} \hat{f}(\hat{x}^{q\#}(s), \mu(s-(5/4)\tau))ds = \hat{\bar{x}}(t)$$

La demostración del Lema (142) muestra, por otra parte, que nos bastaría, para completar su demostración utilizando solo conjuntos del tipo $C_t^{(kh}$, con que quedara eliminada la posibilidad anterior para una sucesión decreciente a cero de valores de τ (y de σ). En caso contrario podríamos afirmar, por lo tanto que (163) valdría para todo τ en un entorno de cero, con $t-\tau$ un punto regular de $u(s)$.

Ahora, puesto que el conjunto de puntos no regulares de $\overline{u}(s)$ es de medida nula y los integrandos pueden modificarse en conjunto de medida nula podríamos asegurar que (163) valdría para todo τ en un intervalo cerrado a la derecha de cero en el caso de que no se hubiera podido establecer el análogo al Lema (142) en este caso.

Utilizando entonces el conjunto $C_t^{*(kh}$ dado por (161) y escogiendo σ suficientemente pequeño podría garantizarse entonces, si se hubieran eliminado aquellos vectores $\hat{f}(\hat{\overline{x}}(t), u^{(j})$ nulos, y supuesto que $\hat{f}(\hat{\overline{x}}(t), \overline{u}(t)) \neq 0$ pues si no no habría nada que demostrar, que los hipotéticos puntos $\hat{x}^{(qr}$ de una sucesión contenida en $\overline{C}_t^{*(kh}$, podrían, en el peor de los casos ser de manera que la restricción: $a^{(qr}_{(i+1)(h+1)} \leq \frac{1}{2^i} a^{(qr}_{1(h+1)}$ se cumpliera con igualdad para todo i y todo r, pero verificándose para cada uno de dichos puntos $\hat{x}^{(qr}$ que:

$$(164) \qquad \hat{x}^{(qr} = \hat{\overline{x}}(t-\tau_r^{(1}) + \int_{t-\tau_r^{(1}}^{t-\tau_r^{(1} + \tau_r^{(2}} \hat{f}(\hat{x}^{(qr}(s), u^{(qr}(s))\, ds$$

con $\tau_r^{(1} \leq \tau_o$ y $\tau_r^{(2} \leq \tau_o$, en donde $\tau_o < \tau$ siendo este τ el correspondiente al $C_t^{*(kh}$.

Por otra parte τ podría seleccionarse suficientemente pequeño de manera que quedara incluído dentro del intervalo para el que vale (163), lo que implicaría la inclusión en este intervalo también de $\tau_r^{(1}$.

En cuanto a la posible nulidad de $\tau_r^{(1}$ de la definición del conjunto $C_t^{*(kh}$ dada en (161), y ya que: $d_{i(h+1)}^{(qr} \leq \frac{1}{2^i} d_p^{(qr}$ para todo $i=1,2,\ldots$ se seguiría entonces que $\hat{x}^{(qr} = \hat{\bar{x}}(t)$, y no habría nada que demostrar. Podemos pues suponer que $\tau_r^{(1} > 0$, para todo τ lo que implicaría que $a_{i(h+1)}^{(qr} - a_{ih}^{(qr} \geq \frac{1}{4^i} \tau_r^{(1} \geq 0$, para todo $i=1,2,\ldots$, y se podría definir por lo tanto la gama completa de vectores tangentes, incluyendo los del tipo $-\hat{f}(\hat{x}^{(q}, \bar{u}(t))$, multiplicados por la izquierda por determinadas matrices $\Phi_{u^{(qr}}(.,.)$, mediante perturbaciones en valores de i lo suficientemente avanzados como para que se pueda asegurar que el punto perturbado $\hat{x}^{(qqr}$ definido mediante la trayectoria $\hat{x}^{(qqr}(s)$ verificara:

$$\sum_{i=0}^{n} (x_i^{(qqr}(s) - \bar{x}_i(t))^2 < \sigma$$

para todo s.

Solo queda pues por ver que una ampliación de la definición de cada $u^{(qr}(s)$ en esos subintervalos avanzados mediante el aumento de los correspondientes $a_{ij}^{(qr}$ es lícita.

Ahora bien, de (163), teniendo en cuenta la unicidad de las soluciones del sistema de ecuaciones diferenciales que caracteriza las trayectorias $\hat{x}(s)$ ya fuera hacia adelante, con valor inicial dado, o hacia atrás, con un valor final dado, podríamos poner:

$$\hat{x}(t-4\tau_r^{(1}) + \int_{t-4\tau_r^{(1}}^{t+4\tau_r^{(1}} \hat{f}(\hat{x}^{q\#}(s),\ u^{(q*}(s))\ ds = \hat{x}(t-\tau_r^{(1})$$

Por lo tanto podríamos poner en (164):

$$\hat{x}^{(qr} = \hat{x}(t-4\tau_r^{(1}) + \int_{t-4\tau_r^{(1}}^{t+4\tau_r^{(1}} \hat{f}(\hat{x}^{(q\#}(s),\ u^{(q*}(s))\ ds + \int_{t+4\tau_r^{(1}}^{t+4\tau_r^{(1}+\tau_r^{(2}} \hat{f}(\hat{x}^{(q\#}(s),\ u^{(q*}(s))\ ds$$

en donde la ampliación de $\hat{x}^{(q\#}(s)$ y $u^{(q*}(s)$ al intervalo $[t+4\tau_r^{(1},\ t+4\tau_r^{(1}+\tau_r^{(2}]$ se haría de manera que coincidiera con las funciones $\hat{x}^{(qr}(s)$ y $u^{(qr}(s)$ en el intervalo $[t-\tau_r^{(1},\ t-\tau_r^{(1}+\tau_r^{(2}]$.

Se comprueba entonces que esta nueva presentación de $\hat{x}^{(qr}$ está de acuerdo con la dada en (161) y además ahora se evidencia la licitud de la definición de la colección de puntos perturbados $\hat{x}^{(qq}$ mediante los cuales se definiría como convexos adecuados de vectores tangentes en los puntos $\hat{x}^{(q}$ de $C_t^{*(kh}$.

Una vez establecida esta posibilidad de definición de los conos y con una demostración similar a la del Lema (142) en la que se tendría en cuenta también la acotación de la función $\bar{u}(s)$ en $[t_o, t_1]$, se podría establecer entonces el siguiente resultado que enunciamos.

a continuación:

(165) LEMA

En las condiciones del Lema (142), si además t fuera un punto regular de $\bar{u}(s)$ en (t_o, t_1) entonces existe - -- $\hat{\bar{\eta}}^*(t) \neq 0$ tal que:

(166) $\hat{\bar{\eta}}^*(t)\ \hat{f}(\hat{\bar{x}}(t),u) \leq \hat{\bar{\eta}}^*(t)\ \hat{f}(\hat{\bar{x}}(t), \bar{u}(t) = 0$

para todo $u \in \Omega$, y $\bar{\eta}^*_o(t) \leq 0$

La ecuación (166) establece el Principio de Máximo puntualmente. Para su extensión y caracterización a lo largo de toda la trayectoria necesitamos todavía ampliar más la colección de vectores tangentes definida en cada punto $\hat{\bar{x}}(t)$ de la trayectoria óptima.

Para iniciar la discusión de esta ampliación consideremos un punto t, con $t_o < t < t_1$ y sendos decrementos no negativos $\Delta_1 t$ y $\Delta_2 t$ de manera que $t - \Delta_1 t - \Delta_2 t > t_o$ y siendo - $t-\Delta_1 t$ y $t - \Delta_1 t - \Delta_2 t$, puntos regulares de la función $\bar{u}$ (s). Empezando en el instante $t - \Delta_1 t - \Delta_2 t - \lambda^{(1}\delta$, en donde - $\delta > 0$ se supone suficientemente pequeño para que en ningún caso se den puntos a la izquierda del t_o y $0 \leq \lambda^{(1} < 1$, - consideremos la siguiente función de control:

$$
\begin{aligned}
u_{\Delta_1 \Delta_2}(s) = {}& u^{(i1} \ ; \text{ para} \\
& t - \Delta_1 t - \Delta_2 t - \lambda^{(1}\delta + (1-\lambda^{(1)}) \sum_{j=0}^{i-1} \lambda_j^{(1}\delta \leq s \leq \\
& \leq t - \Delta_1 t - \Delta_2 t - \lambda^{(1}\delta + \sum_{j=0}^{i} \lambda_j^{(1}\delta \\
& i = 1, 2, \ldots, q \ ; \ \lambda_o^{(1} = 0 \\
= {}& \bar{u}(s+\lambda^{(1}\delta - (1-\lambda^{(1)}) \sum_{i=1}^{q} \lambda_i^{(1}\delta)
\end{aligned}
$$

para $t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta \leq s \leq$

$\leq t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta$

con $0 \leq \lambda^{(2}<1$; $\lambda^{(2}\delta<\Delta_1 t$

$= u^{(i2}$; para

$t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2)})\sum_{i=0}^{i-1}\lambda_j^{(2}\delta$

$\leq s \leq t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2)})\sum_{j=0}^{i-1}\lambda_j^{(2}\delta$

con $i=1, 2, \ldots, q$; $\lambda_o^{(2}=0$

$=\bar{u}(s+\lambda^{(1}\delta-(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta+\lambda^{(2}\delta-(1-\lambda^{(2)})\sum_{i=1}^{q}\lambda_i^{(2}\delta)$;

para $t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2)})\sum_{i=1}^{q}\lambda_i^{(2}\delta \leq s$

$\leq t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2)})\sum_{i=1}^{q}\lambda_i^{(2}\delta$

Mediante la función de control anterior se trasladaría el punto $\hat{\bar{x}}(t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta)$ de la trayectoria óptima a un punto que denotaremos por:

$\hat{\bar{x}}_j(t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2)})\sum_{i=1}^{q}\lambda_i^{(2}\delta)$

y abreviadamente por $\hat{x}_{\delta\Delta_1\Delta_2}$

Se trata entonces de evaluar el siguiente límite

$$(167)\quad \lim_{\delta \downarrow 0}\frac{\hat{x}_{\delta\Delta_1\Delta_2}-\hat{\bar{x}}(t)}{\delta}$$

Ahora, el punto $\hat{x}_{\delta\Delta_1\Delta_2}$ se obtiene de la siguiente manera:

$$\hat{x}_{\delta\Delta_1\Delta_2}=\hat{\bar{x}}(t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta)+\int_{t-\Delta_1 t-\Delta_2-\lambda^{(1}\delta}^{t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1)})\sum_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta}\hat{f}(\hat{x}_\delta(s),u_\delta(s))\,ds+$$

$$+ \int_{t-\Delta_2 t - \lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda^{(1}\delta - \lambda^{(2}\delta}^{t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta - \lambda^{(2}\delta + (1-\lambda^{(2}) \sum_{i=1}^{q} \lambda_i^{(2}\delta} \hat{f}(\hat{x}_\delta(s), u_\delta(s))$$

en donde, $\hat{x}_\delta(s)$ se define de manera obvia para s en el intervalo en cuestión.

Si a la expresión anterior sumamos y restamos la cantidad:

$$\int_{t-\Delta_2 t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta- \lambda^{(2}\delta}^{t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta} \hat{f}(\hat{x}^*_\delta(S), u\,(s+\lambda^{(1}\delta- (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta)\, ds$$

en donde

$$\hat{x}^*_\delta(s) = x\,(t-\Delta_2 t- \lambda^{(1}\delta+ (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta- \lambda^{(2}\delta\,) +$$

$$+ \int_{t-\Delta_2 t-\lambda^{(1}\delta+ (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta- \lambda^{(2}\delta}^{s} \hat{f}(\hat{x}^*_\delta(z), u\,(z+\lambda^{(1}\delta - (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta)\, dz$$

para

$$t -\Delta_2 t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta-\lambda^{(2}\delta \leq s \leq t-\lambda^{(1}\delta+ (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta,$$

podemos obtener entonces el límite (167) calculando por separado los dos siguientes límites:

$$(168) \quad \lim_{\delta \downarrow 0} \Bigg[\hat{\bar{x}}(t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta) + \int_{t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta}^{t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta-\lambda^{(2}\delta} \hat{f}(\hat{x}_\delta(s), u_\delta(s))\, ds \quad +$$

$$\int_{t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta}^{t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i^{(1}\delta} \hat{f}(\hat{x}^*_\delta(s),\ \bar{u}(s+\lambda^{(1}\delta-(1-\lambda^{(1})\sum_{i=1}^{q}\lambda_i^{(1}\delta))ds-\hat{\bar{x}}(t)\Bigg]/\delta$$

y

$$(169)\quad \lim_{\delta\downarrow o}\Bigg[\int_{t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta}^{t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i^{(1}\delta-\lambda^{(2}\delta+(1-\lambda^{(2})\sum\limits_{i=1}^{q}\lambda_i^{(2}\delta} \hat{f}(\hat{x}_\delta(s),u_\delta(s))\,ds\ -$$

$$-\int_{t-\Delta_2 t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i\delta-\lambda^{(2}\delta}^{t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum\limits_{i=1}^{q}\lambda_i^{(1}\delta} \hat{f}(\hat{x}^*_\delta(s),\ u(s+\lambda^{(1}\delta-(1-\lambda^{(1})\sum_{i=1}^{q}\lambda_i^{(1}\delta))]/\delta$$

El límite (168) valdría, de acuerdo con los resultados obtenidos anteriormente y concretados en (141):

$$(170)\quad \Phi(t,\ t-\Delta_1 t-\Delta_2 t)\ \ -\lambda^{(1}\hat{f}(\hat{\hat{x}}(t-\Delta_1 t-\Delta_2 t),\ \bar{u}(t-\Delta_1 t-\Delta_2 t))+$$

$$+\ (1-\lambda^{(1})\sum_{i=1}^{q}\lambda_i^{(1}\ \hat{f}(\hat{\hat{x}}(t-\Delta_1 t-\Delta_2 t),\ u^{(i1})$$

en donde Φ (.,.) denota la matriz fundamental de (154) mientras que el límite (169), cuyo cálculo se podría llevar a cabo de una manera más cómoda sumando y restando en el numerador el siguiente término:

$$\hat{x}_\delta \; (t-\Delta_2 t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta - \lambda^{(2}\delta) =$$

$$= \hat{\bar{x}}(t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta) + \int_{t-\Delta_1 t-\Delta_2 t-\lambda^{(1}\delta}^{t-\Delta_2 t-\lambda^{(1}\delta + (1-\lambda^{(1}) \sum_{i=1}^{q} \lambda_i^{(1}\delta - \lambda^{(2}\delta} \hat{f}(\hat{x}_\delta(s), u_\delta(s))\, ds$$

vendría dado, recordando también (141), por:

$$(171) \quad \Phi(t,t-\Delta_2 t)\, [-\lambda^{(2}\, \hat{f}(\hat{\bar{x}}(t-\Delta_2 t), \bar{u}(t-\Delta_2 t)) + $$
$$+ (1-\lambda^{(2}) \sum_{i=1}^{q} \lambda_i^{(2}\, \hat{f}(\hat{\bar{x}}(t-\Delta_2 t)), u^{(i2)}]$$

El resultado obtenido combinando (170) y (171) sería generalizable, por supuesto al caso en que se consideraran g decrementos no negativos $\Delta_1 t$, $\Delta_2 t$, ..., $\Delta_g t$, siendo los g-1 primeros estrictamente positivos y el último simplemento no negativo de manera que $t - \sum_{i=1}^{g} \Delta_i t > t_o$, supuesto que δ es suficientemente pequeño de modo que todos los instantes considerados estén a la derecha de t_o, y siendo los instantes: $t - \sum_{i=1}^{e} \Delta_i t$, para $1 = 1,2$ - ..., g, puntos regulares de $\bar{u}(s)$ en (t_o, t_1).

En ese caso, sumando y restando las cantidades convenientes en el numerador de la expresión:

$$(172) \quad \frac{\hat{x}_\delta \left(t - \sum_{i=1}^{g} [\lambda^{(i}\delta - (1-\lambda^{(i}) \sum_{j=1}^{q} \lambda_j^{(i}\delta]\right) - \hat{\bar{x}}(t)}{\delta}$$

en donde $\hat{x}_\delta$ se supone obtenida con el control u_δ cuya definición correspondería a la generalización de la del caso g=2 considerado antes y siendo:

$0 \leq \lambda^{(i} < 1$; $\lambda_i^{(i} \geq 0$; $i=1,2,\ldots,g$; $j=1,2,\ldots,q$, con $\lambda^{(i}\delta < \Delta_{i-1} t$

para $i = 2,3,\ldots,g$, se podría calcular el límite (172) mediante la determinación por separado de g límites cada uno de los cuales involucraría expresiones del siguiente tipo:

$$(173)\quad \hat{x}^*_\delta\,(t-\sum_{i=1}^{d}[\lambda^{(i}\delta-(1-\lambda^{(i})\sum_{j=1}^{q}\lambda_j^{(i}\delta]) -$$
$$- \hat{x}^*_\delta\,(t-\sum_{i=1}^{d-1}[\lambda^{(i}\delta-(1-\lambda^{(i})\sum_{j=1}^{q}\lambda_j^{(i}\delta)$$

salvo el primero y el último a los que corresponderían las dos siguientes expresiones:

$$(174)\quad \hat{x}^*_\delta\,(t-\lambda^{(1}\delta+(1-\lambda^{(1})\sum_{j=1}^{q}\lambda_j^{(1}\delta) - \hat{\bar{x}}(t)$$

y,

$$(175)\quad \hat{x}_\delta(t-\sum_{i=1}^{q}[\lambda^{(i}\delta-\sum_{j=1}^{q}\lambda_j^{(i}\delta]) - \hat{x}^*_\delta(t-\sum_{i=1}^{g-1}[\lambda^{(i}\delta-\sum_{j=1}^{q}\lambda_j^{(i}\delta])$$

De acuerdo con la discusión anterior cada una de las expresiones (173) se podrían poner, para δ pequeño, de la siguiente manera:

$$\Phi(t,t-\sum_{i=d}^{g}\Delta_i t)\,[-\lambda^{(1}\,\hat{f}(\hat{x}^*_\delta(t-\sum_{i=d}^{g}\Delta_i t-\sum_{i=d}^{g-1}[\lambda^{(i}\delta-\sum_{i=1}^{q}\lambda_j^{(i}\delta]),\bar{u}(t-\sum_{i=d}^{g}\Delta_i t))+$$
$$+ (1-\lambda^{(d})\sum_{j=1}^{q}\lambda_j^{(d}\hat{f}(\hat{x}^*_\delta(t-\sum_{i=1}^{q}\Delta_i t-\sum_{i=1}^{d-1}[\lambda^{(i}\delta-\sum_{j=1}^{q}\lambda_j^{(i}\delta]),u^{(j1)}].\delta+o(\delta)$$

mientras que (174) y (175) se pueden expresar, respectivamente, así:

$$\Phi(t,t-\sum_{i=1}^{g}\Delta_i t)\,[-\lambda^{(1}\hat{f}(\hat{\bar{x}}(t-\sum_{i=1}^{g}\Delta_i t),\ \bar{u}(t-\sum_{i=1}^{g}\Delta_i t)) +$$
$$+ (1-\lambda^{(1})\sum_{j=1}^{q}\lambda_j^{(1}\hat{f}(\hat{\bar{x}}(t-\sum_{i=1}^{g}\Delta_i t),\ u^{(j1)})]\ .\delta + o(\delta)$$

y:

$$\Phi(t,t-\Delta_g t)\,[-\lambda^{(g}\hat{f}(\hat{x}^*_\delta(t-\Delta_g t-\sum_{i=1}^{g-1}[\lambda^{(i}\delta-\sum_{j=1}^{q}\lambda_j^{(i}\delta]),\ \bar{u}(t-\Delta_g t))+$$

$$+(1-\lambda^{(g})\sum_{j=1}^{q}\lambda_j^{(g}\hat{f}(\hat{x}^*_\delta(t-\Delta_g t-\sum_{i=1}^{g-1}[\lambda^{(i}\delta-\sum_{j=1}^{q}\lambda_j^{(i}\delta],\ u^{(jg})].\delta+o(\delta)$$

Por lo tanto el límite (172) vendrá dado por la siguiente expresión:

$$(176)\quad \sum_{d=1}^{g}\Phi(t,t-\sum_{i=1}^{g}\Delta_i t).[-\lambda^{(d}\hat{f}(\hat{\bar{x}}(t-\sum_{i=d}^{g}\Delta_i t),\ \bar{u}(t-\sum_{i=d}^{g}\Delta_i t))+$$

$$+\ (1-\lambda^{(d})\sum_{j=1}^{q}\lambda_j^{(d}\hat{f}(\hat{\bar{x}}(t-\sum_{i=d}^{g}\Delta_i t),\ u^{(id})]$$

Entonces, cuando los $\lambda^{(d}$ y $\lambda_j^{(d}$ varían, con $0 \leq \lambda^{(d} < 1$, $\lambda_j^{(d} \geq 0$, para unos determinados $\Delta_i t$, $i = 1,2,\ldots,g$, que se supone permanecen fijos, al igual que los $u^{(jd}$ y q, el conjunto correspondiente de vectores del tipo (176) es evidentemente un cono convexo.

Por otra parte, si $\Delta_g t = 0$ y se seleccionan los $\lambda_j^{(d}$ de manera que $\lambda_j^{(d}= 0$ para $j= 1,2,\ldots,q$, y $d= 1,2,\ldots,(g-1)$, el cono convexo anterior incluiría en particular al que hemos considerado en el establecimiento del Lema (165).

A la vista de (176) se podrían definir ahora los correspondientes conjuntos $C_t^{(kh}$ y $C_t^{*(kh}$, generalizaciones de los dados mediante (160) y (161) mediante el establecimiento de las correspondientes funciones de control $u_a^{(p}(s)$, no solo en el punto t sino en los puntos $t - \sum_{i=d}^{g}\Delta_i t$, y de manera que τ y σ fueran suficientemente pequeños para que no hubiera solapes.

El procedimiento utilizado para establecer (142) y (165) nos permitiría entonces formular el siguiente resultado:

(177) LEMA

En las mismas condiciones del Lema (142), si $\Delta_i t > 0$ para $i=1,2,\dots,(g-1)$, $\Delta_g t \geq 0$, y $t - \sum_{i=d}^{g} \Delta_i t$ para $d = 1,2,\dots,g$, son puntos regulares de la función de control $\bar{u}(s)$, entonces existe $\hat{\bar{\eta}}^*(t) \neq 0$, tal que:

$$(178)\quad \hat{\bar{\eta}}^*(t) \quad \Phi(t,t-\sum_{i=d}^{g}\Delta_i t)\ \hat{f}(\hat{\bar{x}}(t-\sum_{i=d}^{g}\Delta_i t),\ u) \leq$$

$$\leq \hat{\bar{\eta}}^*(t) \quad \Phi(t,t-\sum_{i=d}^{g}\Delta_i t)\ \hat{f}(\hat{\bar{x}}(t-\sum_{i=d}^{g}\Delta_i t),\ \bar{u}(t-\sum_{i=d}^{g}\Delta_i t)) = 0$$

para todo $d= 1,2,\dots,g$, y todo $u\epsilon\Omega$

El resultado del Lema (177) puede extenderse fácilmente a cualquier punto regular $t - \Delta_t$ de la función $\bar{u}$ (s)

En efecto, sea R_t el conjunto de puntos regulares de $\bar{u}(s)$ en $(t_o, t]$, en donde t es otro punto regular. Sea, por otra parte W un subconjunto numerable y denso de R_t y $W^{(r}, r\in N$ una sucesión creciente de subconjuntos finitos de W, esto es:

$$\lim_{r} W_t^{(r} = W_t$$

Entonces, si denotamos por $\hat{\bar{\eta}}^{(r*}(t)$ al vector a que se refiere el Lema (177) para cada $W_t^{(r}$, ya que sin pérdida de generalidad podemos suponer que $||\hat{\bar{\eta}}^{(r*}(t)||=1$, para todo r, podríamos obtener una subsucesión convergente a un determinado vector, que denotaremos por $\hat{\bar{\eta}}^*(t)$, no nulo, con $\bar{\eta}_o^*(t) \leq 0$, y de manera que:

(179) $\hat{\hat{\eta}}^*(t) \quad \Phi(t,t-\Delta t) \ \hat{f}(\hat{\bar{x}}(t-\Delta t),\ u) \leq$

$$\leq \hat{\hat{\eta}}^*(t) \ \Phi(t,t-\Delta t) \ \hat{f}(\hat{\bar{x}}(t-\Delta t),\ u(t-\Delta t) = 0$$

para todo $t - \Delta t \in W_t$ y todo $u \in \Omega$, ya que para cada $t-\Delta t \in W_t$ existirá un índice r^* de manera que $t-\Delta t \in W_t^{(r}$ si $r \geq r^*$

Sea ahora $t - \Delta t \in R_t - W_t$, y $t - \Delta^{(r} t$ una sucesión en W_t convergente a $t - \Delta t$.

De (121) se tiene:

$$\int_{t-\Delta^{(r}t}^{t-\Delta t} \hat{f}(\hat{\bar{x}}(s),\ \bar{u}(s))\ ds = (\Delta^{(r}t - \Delta t)\ \hat{f}(\hat{\bar{x}}(t-\Delta t),\ \bar{u}(t-\Delta t)) + $$
$$+ o(\Delta^{(r}t-\Delta t) =$$
$$= (\Delta^{(r}t-\Delta t)\ \hat{f}(\hat{\bar{x}}(t-\Delta^{(r}t), u(t-\Delta^{(r}t) + o\ (\Delta^{(r}t-\Delta t)$$

Dividiendo las dos últimas expresiones por $\Delta^{(r}t - \Delta t$ y haciendo tender dicha diferencia a cero se obtendría:

$$\lim_{r} \ \hat{f}(\hat{\bar{x}}(t-\Delta^{(r}t),\bar{u}(t-\Delta^{(r}t)) = \hat{f}(\hat{\bar{x}}(t-\Delta t),\ \bar{u}(t-\Delta t))$$

De ello se sigue que (179) podría establecerse para todo punto $t - \Delta t$ de R_t. Sea ahora $t^{(r}$ una sucesión de puntos regulares de $\bar{u}(s)$ en $(t_o,\ t_1)$, tal que $\lim t^{(r} = t_1$. Denotemos por $\hat{\bar{\eta}}^*(t^{(r})$ la correspondiente sucesión inducida de vectores $\hat{\bar{\eta}}^*(t)$ verificando (179) y escojamos una subsucesión convergente a un vector que denominaremos $\hat{\bar{\eta}}(t_1)$, el cual será tal que verificaría (179) para todo punto regular de $\bar{u}(s)$ en $(t_o,\ t_1)$, pues si t^* fuera un punto regular en $(t_o,\ t_1)$ de $\bar{u}(s)$ entonces existiría un r^* tal que $t^* \leq t^{(r}$ si $r \geq r^*$.

Podríamos poner entonces:

(180) $\hat{\bar{\eta}}(t_1)\ \Phi(t_1,t)\hat{f}(\hat{\bar{x}}(t),u) \leq \hat{\bar{\eta}}(t_1)\Phi\ (t_1,t)\ \hat{f}(\hat{\bar{x}}(t),\ \bar{u}(t)) = 0$

para casi todo $t \in [t_o,\ t_1]$

Hagamos ahora:

(181) $\hat{\bar{\eta}}(t) = \hat{\bar{\eta}}(t_1)\ \Phi(t_1,\ t)$

para todo $t \in [t_o,\ t_1]$.

Debido a la continuidad absoluta de $\Phi(t_1,\ t)$ quedaría definida mediante (181) una función absolutamente continua en $[t_o,\ t_1]$.

Además, debido a la caracterización de $\Phi(.,.)$ como la matriz fundamental del sistema lineal (154) se puede comprobar que la - función $\hat{\bar{\eta}}'(t)$ transpuesta de la $\hat{\bar{\eta}}(t)$ definida mediante (181) sería solución del siguiente sistema lineal:

(182) $\frac{d\xi(s)}{ds} = [-\frac{\partial t}{\partial \hat{x}}\ (\hat{\bar{x}}(s),\ \bar{u}(s))]'.\ \xi(s)$

para $s\epsilon[t_o,\ t_1]$, y en donde $[.]'$ indica la traspuesta de la matriz en el corchete, con la condición de contorno $\hat{\bar{\eta}}'(t)=\hat{\bar{\eta}}'(t_1)\neq 0$, para $t=t_1$, pues, en efecto, se demuestra facilmente que la matriz fundamental del sistema (182), denominado el "sistema adjunto" - del (154), la cual denotaremos por $\psi(.,.)$ verifica, en $[t_o,\ t_1]$, la ecuación:

$$\psi'\ (t,t_o)\ \Phi(t,t_o) = I$$

en donde ψ' denota la transpuesta de ψ e I es la matriz unidad - (Véase la Nota 19).

Por lo tanto, la solución $\xi(s)$ de (182) con condición inicial:

$$\xi(t_o) = \psi^{-1}(t_1,t_o)\ \hat{\bar{\eta}}'(t_1) = \psi(t_o,t_1)\ \hat{\bar{\eta}}'(t_1)$$

vendría dada en función de $\Phi(.,.)$ por:

$$\xi(t) = \psi(t\ ,t_o)\ \xi(t_o) = \Phi'(t_o,t)\Phi'(t_1,t_o)\hat{\bar{\eta}}'(t_1)$$

Es decir:

$$\xi'(t) = \hat{\bar{\eta}}(t) = \hat{\bar{\eta}}(t_1)\ \Phi(t_1,t_o)\ \Phi(t_o,t) =$$
$$= \hat{\bar{\eta}}(t_1)\quad \Phi(t_1,t).$$

que coincide con (181).

Observese además que al ser $\hat{\bar{\eta}}(t_1) \neq 0$ y $\Phi(.,.)$ no singular - se seguiría entonces que $\hat{\bar{\eta}}(t) \neq 0$ para todo $t \in [t_o,t_1]$.

Por otra parte, ya que la función $\hat{f}$ no depende explicitamente de x_o, tendríamos que $\partial \hat{f}/\partial x_o = 0$, y por lo tanto que $\bar{\eta}_o(t) = \bar{\eta}_o(t_1)$ para todo $t \in [t_o,t_1]$.

Por último, para cada $t \in [t_o,t_1]$ tendremos:

(183) $\quad \sup\limits_{u \in \Omega} \hat{\bar{\eta}}(t)\ .\ \hat{f}(\hat{\bar{x}}(t),\ u) = 0$

en donde $\hat{\bar{\eta}}(t)$ vendría dado por (181).

En efecto, si t fuera un punto regular de $\bar{u}(s)$ en (t_o,t_1), (183) no sería más que una consecuencia de (180), e incluso valdría (183) con "máx" en vez de "sup". Sea pues t un punto no regular de $\bar{u}(s)$ en $[t_o,t_1]$ y supongamos por ejemplo que:

$$\sup_{u \in \Omega}\ \hat{\bar{\eta}}(t).\hat{f}(\hat{\bar{x}}(t),\ u) > 0$$

Existiría entonces un $u^* \in \Omega$ tal que:

(184) $\quad \hat{\bar{\eta}}(t).\hat{f}(\hat{\bar{x}}(t),u^*) > 0$

Ahora, si $t^{(r}$ fuera una sucesión de puntos regulares convergente a t, cuya construcción siempre sería posible al ser el complementario del conjunto de puntos regulares de medida nula en $[t_o,t_1]$, entonces por (180) y la continuidad de $\hat{f}(\hat{x},u)$, $\hat{\bar{x}}(t)$ y $\hat{\bar{\eta}}(t)$, tendríamos:

$$\hat{\bar{\eta}}(t^{(r}).\hat{f}(\hat{\bar{x}}(t^{(r}),\ u^*) \leq 0, \text{ para todo } r,$$

y:

$$\lim_{r} \hat{\bar{\eta}}(t^{(r})\hat{f}(\hat{\bar{x}}(t^{(r}),\ u^*) = \hat{\bar{\eta}}(t)\hat{f}(\hat{\bar{x}}(t),\ u^*) \leq 0$$

en contradicción con (184).

Por otra parte, ya que de la continuidad de $\hat{\bar{\eta}}, \hat{f}$ y $\hat{x}$ se sigue que dado $\varepsilon > 0$,

$$|\hat{\bar{\eta}}(t^{(r})\ \hat{f}(\hat{\bar{x}}(t^{(r}),\ \bar{u}(t^{(r})) - \hat{\bar{\eta}}(t)\hat{f}(\hat{\bar{x}}(t),\bar{u}(t^{(r})| < \varepsilon$$

si $r \geq r_\varepsilon$, para un determinado r_ε, entonces:

$$\lim_{r}\ \hat{\bar{\eta}}(t)\hat{f}(\hat{\bar{x}}(t),\ \bar{u}(t^{(r})) = \lim_{r} \hat{\bar{\eta}}(t^{(r})\hat{f}(\hat{\bar{x}}(t^{(r}),\ \bar{u}(t^{(r})) = 0$$

de donde se seguiría que:

$$\sup_{u\in\Omega}\ \hat{\bar{\eta}}(t).\hat{f}(\hat{\bar{x}}(t),\ u) \geq 0$$

lo que junto con la contradicción implicada por (184) probaría (183).

Podemos ya pues resumir todos los resultados obtenidos en el Teorema siguiente:

(185) TEOREMA (Principio de Máximo de Pontryaguin)

Si $\hat{\bar{x}}(t)$ obtenida para el control $\bar{u}(t)$, con $t_o \leq t \leq t_1$, es una trayectoria óptima para el Problema (116) en el caso autónomo, con $G_o = \{x^{(o)}\}$, cumpliéndose las condiciones (118), y $\hat{f}$ y $\partial\hat{f}/\partial\hat{x}$ son continuas en $\mathbb{R}^{n+1} \times \bar{\Omega}$, entonces existe una función absolutamente continua no nula $\hat{\bar{\eta}}(t) = (\bar{\eta}_o(t), \bar{\eta}_1(t), \ldots, \bar{\eta}_n(t))$ que satisface el sistema adjunto (182) para una determinada condición inicial, de manera que:

a) La función:

$$\hat{\bar{\eta}}(t) . \hat{f}(\hat{\bar{x}}(t), u)$$

para cada t fijo, siendo t un punto regular de $\bar{u}(s)$ en (t_o, t_1), alcanzará su máximo en el punto $u = \bar{u}(t)$, lo cual implicaría que:

$$\hat{\bar{\eta}}(t) \hat{f}(\hat{\bar{x}}(t), \bar{u}(t)) = \sup_{u \in \Omega} \hat{\bar{\eta}}(t) \hat{f}(\hat{\bar{x}}(t), u) \tag{186}$$

para casi todo punto t en $[t_o, t_1]$.

b) En el punto terminal se verifican las relaciones:

$$\bar{\eta}_o(t_1) \leq 0 \quad , \text{ y } \quad \sup_{u \in \Omega} \hat{\bar{\eta}}(t_1) \hat{f}(\hat{\bar{x}}(t_1), u) = 0 \tag{187}$$

Además las funciones de t, $\bar{\eta}_o(t)$ y $\sup_{u \in \Omega} \hat{\bar{\eta}}(t) \hat{f}(\hat{\bar{x}}(t), u)$, resultan ser constantes, de manera que las condiciones (187) pueden verificarse en cualquier instante $t \in [t_o, t_1]$, y no necesariamente en t_1.

Con esto terminamos nuestra exposición de la aplicabilidad del Principio (91) en la deducción de las condiciones necesarias de optimización en el caso diferenciable. En el siguiente capítulo enunciaremos y discutiremos otros principios de interés en lo referente a la generación de condiciones suficientes.

CAPITULO III

TEORIA DE LAS CONDICIONES SUFICIENTES. EL PRINCIPIO DE LA OPTIMIZACION EN DESIGUALDADES Y EL PRINCIPIO DEL MINIMAX.

1.- INTRODUCCION.

En los capítulos I y II hemos analizado la aplicabilidad de dos principios básicos en la generación de condiciones necesarías de optimalidad desde el punto de vista geométrico, el principio de la Proyección Maximal, en el caso convexo, y el Principio de la Proyección Máximal Tangencial, en el caso diferenciable.

Se podrían proponer también principios en la generación de las condiciones necesarias de optimalidad desde el punto de vista analítico, aunque cualquiera de ellos sería probablemente una adaptación del siguiente principio básico elemental:

Si para el problema: min f(u), sujeto a uϵ $\Omega \subset W$, en donde W denota un conjunto arbitrario, y f una función real definida en su subconjunto Ω, entonces si u^* es una solución del problema y T es una transformación de W en si mismo, tal que T $(u^*) \in \Omega$, entonces necesariamente habría de ser: $f(u^*) \leq f(T(u^*))$.

Nosotros, sin embargo, vamos a pasar a proponer en primer lugar un principio que permitiría plantear condiciones suficientes de optimización, en vez de sólo condiciones necesarías como los anteriores, antes de pasar a discutir luego otra de las metodologías generadoras de condiciones suficientes de óptimo, como es la teoría del minimax para funciones de dos tipos de argumentos.

La idea que preside el primero de estos dos principios es sencilla: Ya que toda desigualdad entre dos funciones encierra la solución de un problema de óptimo se trataría de formular esa característica de una manera que fuera practicable en el análisis de los problemas de optimización. Esta idea nos fué sugerida por la lectura de algunos trabajos sobre una metodología para resolver problemas de optimización cuya puesta en práctica podia justificarse en base a la conocida desigualdad entre las medias aritméticas y geométrica por lo que ha recibido el nombre de Programación Geométrica. (Véase la Nota 20).

Ahora la Programación Geométrica ha derivado en el análisis de otras cuestiones más directamente relacionadas, con la solución efectiva de problemas que en la investigación de la formalización de un posible principio general que permita afrontar la solución de problemas de optimización o por lo menos la caracterización de dichas soluciones, mediante la utilización sistemática de desigualdades, que es el punto de vista que a nosotros nos interesa discutir aquí.

Vamos a establecer, pues, este principio de una manera precisa:

(1) TEOREMA (Principio de la Optimización en Desigualdades).

Consideremos el problema:

(2) min $f_o(u)$

sujeto a:

(3) $u \in U$

con Ω un conjunto arbitrario y f_o una función real definida en U.

Supongamos que existe para cada u o un elemento $\eta=\eta(u)$ de otro espacio H y sendas funciones reales p y q definidas en H x U, de manera que:

(4) $p(\eta, u) \leq q(\eta,u)$ para todo $u \in U$, y $\eta=\eta(u)$, y que:

$u \in U$ implica $p(\eta,u) = f(u)$, y

(5) $p(\eta,u) = q(\eta,u)$ implica $u \in U$.

Supongamos además que existe el mínimo de la función $q(\eta,u)$ para $u \in U$, el cuaÍ denotaremos por q_m, y que:

(6) $q(\eta, u) = q_m$, para todo $u \in S \subset U$.

Entonces, si $u^* \in S$ y $p(\eta^*,u^*) = q(\eta^*,u^*)$, u^* será un punto solución del Problema (2).

DEMOSTRACION.

En efecto, se verificaría:

$$f(u) = p(\eta,u) \geq q(\eta,u) \geq q_m = q(\eta^*,u^*) = p(\eta^*,u^*) = f(u^*)$$

para todo $u \in U$ y además $u^* \in U$ lo que prueba la optimalidad de u^*.

La utilidad de la aplicación del principio anterior estriba en la posibilidad de utilizar a η para determinación de q_m y S, lo que caracterizaría a la solución óptima como aquella u^* de S tal que: $p(\eta^*,u^*) = q(\eta^*,u^*)$.

Vamos a ilustrar en primer lugar la aplicación del Principio (1) con un ejemplo tomado precisamente de la Programación Geométrica.

Supongamos que se pretende resolver el siguiente problema:

(7) $\min\ 2u_1 + 4u_1^{-1}u_2^{-1} + u_2$

sujeto a: $u_1>0$, $u_2>0$.

En este ejemplo pues, $U = \{(u_1,u_2)' \in \mathbb{R}^2 / u_1>0,\ u_2>0\}$

A la función f_o de uno se le puede denominar un "posinomio" (contracción de "polinomio positivo"), queriendo indicar con ello que se trata de una generalización de las funciones polinomiales - permitiendo la posibilidad de exponentes de cualquier tipo, entero o no, pero con coeficientes restringidos a ser positivos.

Podemos entonces preparar la aplicación de (1) definiendo formalmente:

$$p\ (\eta,u) = \frac{\eta_1 2u_1}{\eta_1} + \frac{\eta_2 4u_1^{-1}u_2^{-1}}{\eta_2} + \frac{\eta_3 u_2}{\eta_3}$$

en donde $\eta=(\eta_1,\eta_2,\eta_3,)$ sería independiente de u(lo que constituiria por supuesto, una forma extrema de dependencia).

Ahora si $\eta_1+\eta_2+\eta_3= 1$ y $\eta_1 \geq 0$, de acuerdo con la desigualdad entre las medias ponderadas aritmética y geométrica podríamos poner:

$$p\ (\eta,u) \geq q(\eta,u) = \left(\frac{2u_1}{\eta_1}\right)^{\eta_1} \left(\frac{4u_1^{-1}u_2^{-1}}{\eta_2}\right)^{\eta_2} \left(\frac{u_2}{\eta_3}\right)^{\eta_3}$$

Pondríamos a continuación seleccionar las η_i de manera que se anularan los exponentes de u_1 y u_2 en $q(\eta,u)$.

Esto es, las η_i deberían verificar:

$$\begin{aligned} \eta_1+\eta_2+\eta_3 &= 1 \\ \eta_1-\eta_2 &= 0 \\ \eta_2-\eta_3 &= 0 \\ \eta_1\geq 0 \quad \eta_2\geq 0 \quad \eta_3 &\geq 0 \end{aligned}$$

lo que da $\eta_1=\eta_2= \eta_3= 1/3$

Por lo tanto $q_m = 6$, y $S = \mathbb{R}^2$

Por otra parte como la igualdad entre las medias pondera - das aritmética y geométrica solo se daría si coinciden entre sí - cada uno de los términos de ambas medias tendríamos que:

$p(\eta, u^*) = q(n,u^*)$ implicaría que:

$$\frac{2u_1^*}{1/3} = \frac{4u_1^{*-1}\ u_2^{*-1}}{1/3} = \frac{u_2^*}{1/3}$$

es decir:

$$u_1^* = 1\ ;\quad u_2^* = 2.$$

La aplicación del Principio (1) no se restringe, por su -- puesto, sólo a problemas de tipo anterior sino que podría aplicar se en general a cualquier problema para el que pudiera definirse la desigualdad (5) adecuada.

Otra observación que hay que hacer es la de que evidente - mente, se podrían modificar las condiciones de obtención del q_m - en (1) en determinadas condiciones sin que variara la conclusión del teorema. Esto es, podría haberse definido el q_m como el mínimo de q(n,u) para cada u en un conjunto más amplio que el U, - para el que estuviese todavia definida la función q(n,u), de -- acuerdo con la mayor o menor facilidad en la obtención de dicho mínimo, supuesto que el (o los) u* minimizante pertenecieran a U.

Nótese también además que no habría ninguna pérdida de generalidad si se supone que $u \in \Omega$ implicara que $p\ (\eta,u) = a\ f(u)+b$, en donde a y b serían dos constantes reales, con a>o.

Este principio podría emplearse también para garantizar - la suficiencia de soluciones que complieran las condiciones necesarias de optimización.

Como ampliación de la discusión de la aplicabilidad del Principio (1) vamos a mostrar en la sección siguiente como podría utilizarse para abordar la solución de un problema de programación no lineal.

2.- UNA APROXIMACION A LA SOLUCION DE LOS PROBLEMAS DE PROGRAMACION NO LINEAL POR MEDIO DEL PRINCIPIO DE LA OPTIMIZACION EN DESIGUALDADES.

Consideramos planteado el siguiente problema de programación no lineal:

(8) $\min f_o(u)$

sujeto a:

$g(u) \geq 0$

$u \in \Omega$

en donde $g = (g_1, g_2, \ldots, g_n)'$, $u \in \mathbb{R}^m$, Ω un subconjunto arbitrario de $\mathbb{R}^m$ con g_i, $i=0,1,2,\ldots,n$, funciones reales, e interpretándose la desigualdad $g \geq 0$, componente a componente, como hemos hecho hasta aquí. La correspondencia con el Problema (2) se establecería haciendo:

$$U = \{u \in \Omega \subset \mathbb{R}^m / \; g(u) \geq 0\}$$

Consideramos ahora para cada $\eta \in \mathbb{R}^{n+1}$, con $\eta_i \geq 0$

y $\sum_{i=o}^{n} \eta_i = 1$

la función real $p(\eta, u)$ definida de la siguiente manera:

(9) $p(\eta,u) = \eta_o f_o(u) + \sum_{i=1}^{n} \eta_i \, (f_o(u) + g_i(u)) = f_o(u) + \sum_{i=1}^{n} \eta_i g_i(u)$,

con la condición:

(10) $\eta_i \, g_i(u) = 0$

para todo $i = 1,2, \ldots, n$, $\eta \in H$ y $u \in U$.

Se puede comprobar entonces que $f_o(u)$ y $p(\eta,u)$ verifican las condiciones impuestas por el Teorema (1) ya que:

$$g(u) \geq 0 \Rightarrow p(\eta,u) = \eta_o f_o(u) + \sum_{i \in A_u} \eta_i (f_o(u)+g_i(u)) =$$

$$= \eta_o f_o(u) + \sum_{i \in A_u} \eta_i f_o(u) = f_o(u)$$

en donde:

$$A_u = i = \{1,2,\ldots,n \ / \ g_i(u) = 0\}$$

lo cuál se sigue de (10).

Intoduzcamos ahora la función $q(\eta,u)$ definida de la si - guiente manera:

$$(11) \quad q(\eta,u) = (f_o(u))^{\eta_o} . \prod_{i \in A_u} (f_o(u)+g_i(u))^{\eta_i} = f_o(u) \prod_{i \in A_u} (g_i(u))^{\eta_i}$$

Entonces, de acuerdo con la desigualdad entre las medias ponderadas aritmética y geométrica, tendríamos:

$$(12) \quad p(\eta,u) \geq q(\eta,u)$$

para todo $u \in \Omega$ tal que $g(u) \geq 0$.

Además, ya que el signo de igualdad entre los dos miembros de la desigualdad anterior se da sí y sólo si los elementos a - los que afectan las correspondientes medias ponderadas son - todos iguales tendremos entonces que:

$$p(\eta,u) = q(\eta,u) \Rightarrow f_o(u) = f_o(u) + g_i(u), \text{ para todo } i \in A_u$$

es decir $g_i(u) = 0$, para todo $i \in A_u$, lo que junto con $g_i(u) > 0$ si $i \in \{1,2,\ldots,n\} - A_u$, terminaría por implicar conjuntamente que - $f(u) \geq 0$.

(Nótese que al descartar los índices $i \in A_u - \{1,2,\ldots,n\}$ en la formulación de las funciones $p(\eta,u)$ y $q(\eta,u)$ se querría indicar con ello que habría un conjunto de g_i que obligatoriamente serían estrictamente positivas.

El resto de las g_i, esto es aquellas, con $i \in A_u$ serían simplemente no negativas, para garantizar (12), y lo que se ha analizado antes es que el signo de igualdad en (12) implicaría la anulación de todas ellas. Naturalmente, la manera de actuar para utilizar esta aproximación sería la de resolver el problema para una selección de restricciones que se propusieran "activas", es decir, verificadas con igualdad, y comprobar a posteriori que el resto de las restricciones se verifica con desigualdad estricta, para verificar luego que la elección entre restricciones "activas" y "pasivas" era la correcta).

Puesto que se verifican todas las condiciones impuestas por el Teorema (1) y supuesto que A_{u*} fuera el conjunto de índices i tales que $g_i(u^*) = o$, para una solución del Problema (8), entonces esta solución podría caracterizarse mediante la solución, si existiera, del problema:

$$(13) \quad \min f_o(u) \prod_{i \in A_{u*}} (g_i(u))^{\eta_i}$$

sujeto a: $u \in \Omega$,

el cuál podría ser más sencillo de resolver debido a la relajación del conjunto de restricciones y quizás a la mayor simplicidad de la función objetivo si el conjunto de los η_i se selecciona de manera adecuada.

Por otra parte, el establecimiento de la desigualdad (12), que en términos más explícitos vendria dada, si hacemos la conven ción de que $0^0 = 1$, por:

$$(14) \quad f_o(u) + \sum_{i=1}^{n} \eta_i \, g_i(u) \geq f_o(u) . \prod_{i=1}^{n} (g_i(u))^{\eta_i}$$

con $\sum_{i=1}^{n} \eta_i = 1$, $\eta_i \geq 0$, $i=1,2, \ldots, n$, establecería pues una acota - ción a la función objetivo para cada $u \in \Omega$, con $g(u) \geq 0$, que podría dar información sobre su posible valor óptimo independientemente de que diera lugar a su cálculo efectivo o no.

Observese que en (14) los parámetros η_i, intervendrían de una manera diferente en cada uno de los dos miembros debido a la diferencia entre las dos funciones utilizadas, p y q. Un tratamientos más simétrico de las variables se verá en el apartado - siguiente la introducción del que denominamos Principio del Minimax.

3.- EL PRINCIPIO DEL MINIMAX.

Al igual que el Principio (1) se pueden proponer otros, enfocados también a la obtención de condiciones suficientes en problemas de optimización. Uno de ellos es el que pasamos a enunciar a continuación.

(15) TEOREMA (Principio del Minimax).

Consideramos el Problema (2) con las restricciones (3), en donde los símbolos implicados tienen los mismos significados que en el Teorema (1), y sea p (η, u), con $\eta \in H$, una función - real definida en H x U, de manera que:

$$(16) \quad \sup_{\eta \in H} p(\eta, u) \leq f_o(u), \text{ para todo } u \in U$$

y que:

(17) $\inf_{u \in U} p(\eta,u) \geq d_o(\eta)$, para todo $\eta \in H$, con d_o una determinada función que supondremos "real extendida", esto es pudiendo - tomar el valor $-\infty$, y que exista $\sup_{\eta \in H} d_o(\eta) = p_m$.

Entonces, si u^* es tal que $u^* \in U$ y además $f_o(u^*) \leq p_m$, dicho u^* será una solución del problema.

DEMOSTRACION.

El resultado enunciado se sigue inmediatamente de que:

$$p(\eta,u) \geq \inf_{u \in} p(\eta,u) \geq d_o(\eta), \text{ para todo } \eta \in H \text{ y todo } u \in U$$

de donde:

(18) $$f_o(u) \geq \sup_{\eta \in H} p(\eta,u) \geq \sup_{\eta \in H} \inf_{u \in U} p(\eta,u) \geq \sup_{\eta \in H} d_o(\eta) =$$

$$= p_m \geq f(u^*).$$

para todo $u \in U$, con $u^* \in U$, lo que completa la demostración.

Observese que de (18) se seguiría que:

$$\min_{u \in U} f_o(u) \geq \inf_{u \in U} \sup_{\eta \in H} p(\eta,u) \geq$$

$$\geq \sup_{\eta \in H} \inf_{\in U} p(\eta,u) \geq \sup_{\eta \in H} d_o(\eta) = f(u^*) = \min_{u \in U} f_o(u)$$

Esto es si existieran $\min_{u \in U} p(\eta,u)$, para todo $\eta \in H$, $\max_{\eta \in H} p(\eta,u)$ para todo $u \in U$, y si $\eta^* \in H$ fuera tal que $p(\eta^*,u^*) = f_o(u^*)$, el elemento (η^*,u^*) de $H x U$ sería tal que:

(19) $$\min_{u \in U} \max_{\eta \in H} p(\eta,u) = p(\eta^*,u^*) = \max_{\eta \in H} \min_{u \in U} p(\eta,u)$$

de manera que la existencia del u^* postulado en el Teorema (15) exigiria la coincidencia entre el minimax y el maximin.

El Teorema (15) podría haberse denominado también el Principio de la Dualidad en razón a que se podría interpretar el problema de maximización en la variable dado por:

(20) $\max d_o(\eta)$

sujeto a:

$\eta \in H$,

en donde $d_o(\eta)$ es tal que:

$$d_o(\eta) \leq \inf_{u \in U} p(\eta,u), \text{ para todo } \eta \in H,$$

como un problema "dual en sentido amplio" del (2), denominándose entonces a η la variable "dual en sentido amplio" de u. (Nótese - que con la definición que hemos dado podría existir muchos problemas "duales" de uno dado no sólo por la arbitrariedad de la función p sino también por la de d_o).

Hay que observar, sin embargo que esta denominación de problema "dual en sentido amplio" que nosotros utilizamos aquí es, diferente de la usual en el sentido de que generalmente la definición de problemas "duales" suele ser mucho más rígida, pues -- supuesto planteado el problema en nuestro marco, se suele exigir respecto a ellos que exista algún par (η^*,u^*) se verifique un resultado del tipo (19). La utilización por nuestra parte de la - denominación de problema "dual en sentido amplio" para uno en el que no se pudiera garantizar, en principio, un resultado como - (19), es sólo por comodidad en la exposición.

En relación, por ejemplo, con la resolución del Problema (8) del apartado anterior, mediante la aplicación del Principio (15), una manera de proceder sería la de definir la función $p(\eta,u)$ de la misma manera que antes, esto es:

$$p(\eta,u) = f_o(u) + \sum_{i=1}^{n} \eta_i g_i(u)$$

pero estando restringida ahora la variable $\eta=(\eta_1\eta_2,\ldots,\eta_n) \in \mathbb{R}^n$ a pertenecer a un determinado conjunto H, que dependería del problema particular de que se tratara, y que, por lo tanto, podría coincidir o no con el dado por las condiciones $\Sigma\ \eta_i=1$, $\eta_i\geq 0$, $i= 1,2,\ldots,n$.

Con la definición dada de $p(\eta,u)$, un problema "dual en sentido amplio" del (8) sería entonces uno del tipo (20) en donde $d_o\ (\eta)$ fuera una determinada función tal que:

$$(21)\qquad d_o(\eta) \geq \inf_{g(u)\geq 0;\ u\in\Omega} \{f_o(u) + \sum_{i=1}^{n} \eta_i\ g_i(u)\}$$

El Problema caracterizado por (21) podría ser más sencillo de resolver que el (8), y de todas formas, aunque ese no fuera el caso aportaría información a la solución del Problema "primal", - que sería como podríamos denominar a dicho Problema (8) en rela - ción al caracterizado por (21).

Naturalmente, la definición del Problema (21) como "dual en sentido amplio" del (8) depende de la elección efectuada de la función $p(\eta,u)$. Así, para otras posibles elecciones de p se generarían otros tantos problemas duales del (8).

Tiene interés también discutir el caso particular del Pro - blema (8) en el que tanto $f_o(u)$ como $f(u)$ son funciones lineales - afines para comprobar la concordancia de la definición de proble - mas duales dada en ese caso especial en el capítulo I y la que po - díamos obtener por el método comentado aquí para unas determinadas funciones p y d_o del tipo que acabamos de considerar.

En efecto, con la misma notación utilizada en el capítulo I - podemos replantear el Problema (8) en el caso lineal de la siguiente manera:

(22) $\min f_o(u) = \sum_{j=1}^{m} c_j u_j$

sujeto a:

$$f_i(u) = \sum_{j=1}^{m} a_{ij} u_j - b_i \geq 0 \; ; \; i = 1,2,\ldots,n$$

$$u \in \Omega = \{u \in \mathbb{R}^m / u_j \geq 0, \; j = 1,2,\ldots,m\}$$

Definamos ahora H como el siguiente conjunto:

(23) $H = \{\eta \in \mathbb{R}^m / \sum_{i=1} a_{ij}\eta_i - c_j \leq 0 \;,\; j=1,2,\ldots,m; \; \eta_i \geq 0, \; i=1,2,\ldots n\}$

(y nótese que su definición no depende de u) y sea p (η,u) la función

(24) $p(\eta,u) = f_o(u) + \sum_{i=1}^{n} \eta_i g_i(u) =$

$= \sum_{j=1}^{m} (c_j - \sum_{i=1}^{n} a_{ij}\ \eta_i)\ u_j\ + \sum_{i=1}^{n} b_i \eta_i =$

$= \sum_{i=1}^{n} (\sum_{j=1}^{m} a_{ij} u_j - b_i)\eta_i + \sum_{j=1}^{m} c_j u_j$

la cual estaría definida en el producto cartesiano H x U (y no sólo en algún subconjunto de dicho producto) en donde por U denotamos al subconjunto de $\mathbb{R}^m$ para el que se verifican las restricciones del Problema (22). Entonces, dado que, como se sigue inmediatamente de las definiciones de H y U

(25) $\min_{u\epsilon U} \sum_{j=1}^{m} (c_j - \sum_{i=1}^{m} a_{ij}\eta_i)\ u_j + \sum_{i=1}^{n} b_i \eta_i \geq \sum_{i=1}^{m} b_i \eta_i$

para todo $\eta \epsilon$ H, de manera que un problema dual del (22) sería el dado por:

(26) máx $d_o(\eta) = \sum_{i=1}^{n} b_i \eta_i$

sujeto a: $\eta \epsilon$ H

Por lo tanto si existiera solución finita del Problema (26) con valor de la función objetivo: max $\sum_{i=1}^{n} b_i \eta_i = p_m$, y un u* = $= (u_1*, u_2*, \ldots, u_n*)' \epsilon$ U, tal que $\sum_{j=1}^{m} c_j u_j* = p_m$, dicho u* sería una solución óptima del Problema (22).

Naturalmente, se podrían haber planteado otros problemas duales del (22) incluso en la misma función objetivo $d_o(\eta) = \sum_{i=1} b_i \eta_i$, con tal que el nuevo conjunto H de que se tratara incluyera las - restricciones: $\sum_{i=1}^{n} a_{ij}\eta_i - c_j \leq 0$, para j = 1,2,...,m, pues es todo lo que se necesita para que se verifique (25).

Sin embargo la elección como conjunto H del dado mediante (23) permite establecer no solo la "suficiencia", para el u* que postula el Principio (15), sino también la "necesidad", supuesto que el Problema (26) tiene solución finita. Sobre la manera en co mo podría llevarse a cabo la demostración de esta segunda afirmación nos remitimos al Teorema (50) del capítulo I.

Veamos ahora como podría utilizarse el Principio (15) para obtener condiciones suficientes de óptimo en problemas del tipo (8). Respecto a ello vamos a establecer el resultado siguiente:

(27) LEMA

Consideremos el Problema (8) y sea $p(\eta,u)$ la siguiente función:

(28) $$p(\eta,u) = f_o(u) - \sum_{i=1}^{n} \eta_i\, g_i(u)$$

con $(\eta,u)\in H\times U$, siendo H un determinado subconjunto de R^n y:

$$U = \{u\in R^m / u\in\Omega;\quad g(u) \geq 0\}$$

Supongamos que para un determinado $u^*\in U$ existiera un $\eta^*\in H$ de manera que:

(29) $$f_o(u^*) - \sum_{i=1}^{n} \eta_i\, g_i\,(u^*) \leq f_o(\mu) - \sum_{i=1}^{n} \eta_i^*\, g_i\,(u)$$

para todo $\mu\in U$, y tal que

(30) $$\sup_{\eta\in H}\,(f_o(u) - \sum_{i=1}^{n} \eta_i\, g_i(u)) \leq f_o(u),\text{para todo } u\in U, \text{ y}$$

(31) $$\sum_{i=1}^{n} \eta_i^* \, g_i \, (u^*) = 0$$

Entonces u^* será una solución del Problema (8)

DEMOSTRACION

En primer lugar observamos que (30) implicaría que:

(32) $$f_o(u^*) - \sum_{i=1}^{n} \eta_i g_i(u^*) \geq f_o(u^*) - \sum_{i=1}^{n} \eta_i g_i(u^*) \ ,$$

o equivalentemente:

$$\sum_{i=1}^{n} (\eta_i - \eta_i^*) \, g_i \, (u^*) \geq 0$$

para todo $\eta \varepsilon H$.

Por otra parte, si tomamos:

$$d_o(\eta) = \inf_{u \varepsilon U} (f_o(u) - \sum_{i=1}^{n} \eta_i g_i(u))$$

para $\eta \varepsilon H$, tendríamos, de acuerdo con (29) y (30):

$$\sup_{\eta \varepsilon H} d_o(\eta) = \inf_{u \varepsilon U} (f_o(u) - \sum_{i=1}^{n} \eta_i^* g_i(u)) = f_o(u^*) - \sum_{i=1}^{n} \eta_i^* g_i(u^*)$$

$$= f_o \, (u^*)$$

mientras que, teniendo en cuenta (32) y (31):

(34) $$f_o(u^*) - \sum_{i=1}^{n} \eta_i g_i(u^*) \leq f_o(u^*) - \sum_{i=1}^{n} \eta_i^* g_i(u^*) \leq f_o(u) - \sum_{i=1}^{n} \eta_i^* g_i(u)$$

para todo $\eta \geq 0$ y todo $u \in \Omega$, entonces u^* será una solución óptima - de (8).

DEMOSTRACION

Pongamos:

$$U = \{u \in \mathbb{R}^n / u \in \Omega \ ; \ g(u) \geq 0\}$$

y:

(35) $$H = \{ \eta \in \mathbb{R}^n / \eta \geq 0 \}$$

Puesto que se verifica (29), lo que estaría implicado por la segunda de las desigualdades (34), solo hay que probar que - $g(u^*) \geq 0$, y que se verifica (30) y (31), para deducir la tesis del Teorema como una aplicación del Lema (27).

Ahora de que:

(36) $$\sum_{i=1}^{n} (\eta_i - \eta_i^*) \ g_i(u^*) \geq 0$$

para todo $\eta \in H$, se seguiría, tomando $\eta_i - \eta_j^* = 1$, y $\eta_i - \eta_i^* = 0$, si $i \neq j$, que $g_j(u^*) \geq 0$, y por consiguiente $g(u^*) \geq 0$.

Por otra parte si en (36) tomamos alternativamente, - $\eta_j = 2\eta_j^*$; $\eta_i = 0$, si $i \neq j$, y: $\eta_j = (1/2)\eta_j^*$; $\eta_i = 0$, si $i \neq j$, se

$$\inf_{u\in U} \sup_{\eta\in H} (.) \leq \sup_{\eta\in H} (f_o(u^*) - \sum_{i=1}^{n} \eta_i g_i(u^*)) =$$

$$= f_o(u^*) - \sum_{i=1}^{n} \eta_i^* g_i(u^*) = f_o(u^*)$$

El resultado del Lema se sigue entonces de que:

$$f_o(u^*) \leq \sup_{\eta\in H} d_o(\eta) = p_m \leq \inf_{u\in U} \sup_{\eta\in H} (.) \leq f_o(u^*),$$

y de la aplicación del Principio (15).

Este resultado que acabamos de obtener tiene una aplicación inmediata en la generalización de la parte b) del Teorema (24) del Capítulo I al caso de ausencia de convexidad.

En efecto, se podría establecer facilmente una condición suficiente para dicho problema como un corolario del Lema (27) como vamos a ver a continuación.

(33) TEOREMA (De condiciones suficientes para el Problema de Programación No Lineal)

Consideremos el Problema (8) con $f_o(u)$ y $g_i(u)$, $i=1,2,\ldots,n$ funciones arbitrarias definidas en el conjunto Ω.

Entonces, si existe $u^*\epsilon\Omega$, y $\eta^*\epsilon R^n$; con $\eta_i^* \geq 0$, $i=1,2,\ldots,n$, tales que:

obtendría que: $\eta_j^* g_i(x^*) = 0$, para todo j, de lo que sigue (31), mientras que (30) es evidente de acuerdo con la definición de U y H.

Observese que el Teorema (33) es el recíproco del Teorema (24) del capítulo I, pero sin la exigencia ahora de hipótesis de convexidad sobre las funciones f_o y $-g_i$ y sobre el conjunto Ω.

El resultado del Teorema (33) depende, sin embargo, de que el conjunto elegido para H sea el dado por (35), mientras que el enunciado del Lema (27) se aplicaría en Problemas de tipo (8) en las que el conjunto H de multiplicadores podría un subconjunto arbitrario de $\mathbb{R}^m$.

Por otra parte estos dos resultados no serían más que casos particulares del Principio (15) mediante el cual se podría afrontar la generación de condiciones suficientes para problemas de optimización de tipo general.

Ahora, una cuestión interesante es la de comparar los campos de aplicación de los dos Principios (1) y (15) que hemos enunciado para la generación de condiciones suficientes y sus relaciones a su vez con la teoría clásica de la dualidad, de lo cual nos ocuparemos, para terminar, en el apartado siguiente.

4. ANALOGIAS Y DIFERENCIAS ENTRE LOS DOS PRINCIPIOS BASICOS ENUNCIADOS Y SUS RELACIONES CON LA TEORIA DE LA DUALIDAD.

Si tenemos en cuenta que la mayoría de las desigualdades pueden obtenerse como subproducto de la resolución de un problema de optimización planteado de manera adecuada podría pensarse entonces, en principio, que (15) es más general que (1).

Por ejemplo, para generar la desigualdad entre las medias ponderadas aritmética y geométrica se podría plantear el siguiente problema de optimización:

(37) $$\min f_o(u) = \sum_{i=1}^{n} u_i^2 \text{;con } u \in \left[\{u \in R^n / \prod_{i=1}^{n} u_i^2 = 1 \text{ ,y } u_i \geq 0 \text{ , para todo } i\}\right]$$

en donde $[.]$ denota la clausura convexa del conjunto $\{.\}$.

Ahora tenemos que: $\sum_{i=1}^{n} u_i^2$ es estrictamente convexa, pues cada componente, u_i^2 lo es, como se comprueba de una manera elemental directa teniendo en cuenta que si $0 < \lambda < 1$ y $u_{i1} \neq u_{i2}$ - $2\lambda(1-\lambda)u_{i1}u_{i2} \leq \lambda(1-\lambda)(u_{i1}^2 + u_{i2}^2)$.

Entonces, de la simetría de la función objetivo y de la ecuación de restricción en las u_i, se seguiría que si $(u_1,u_2,\ldots,u_n)$ es una solución factible del problema también lo sería cualquier permutación circular de la anterior, verificándose además que:

$$f_o\left(\frac{1}{n}(u_1,\ldots,u_n)+\ldots+\frac{1}{n}(u_n,\ldots,u_1)\right)= \sum_{i=1}^{n}\left(\frac{1}{n}\sum_{i=1}^{n}u_i\right)^2 \leq$$

$$\leq \sum_{i=1}^{n}u_i^2 = \frac{1}{n} f_o((u_1,\ldots,u_n))+\ldots+\frac{1}{n} f_o((u_n,\ldots,u_1))=f_o(u)$$

con desigualdad estricta si $u_i \neq u_j$ para algún i y j. Además, ya que el conjunto factible es convexo el punto de coordenadas constantes igual a $(1/n)\Sigma u_i$ será factible.

Se seguiría de ello que $f_o(u)$ se minimizaría en el único punto dado por $u_i = 1$, para todo i, y por lo tanto que

$$\sum_{i=1}^{n} u_i \leq n,$$

para toda colección de u_i tales que $u_i \geq 0$ y $\prod u_i = 1$, lo que probaría que: $\Sigma \phi_i / n \leq (\prod u_i)^{1/n}$ con igualdad si y solo si u_i fuera constante para todo i, y por paso al límite en coeficientes de la ponderación racionales, que: $\Sigma \eta_i u_i \geq \prod u_i^{\eta_i}$, para toda colección de η_i, i=1,2,...,n, con $\Sigma \eta_i = 1$ y $\eta_i \geq 0$ para todo i y con igualdad si y solo si u_i es constante. (Véase la Nota 21).

Así el anterior razonamiento muestra como, en general, cabría esperar deducir cualquier disigualdad mediante el planteamiento de un problema de optimización en términos adecuados que podría involucrar explicítamente o no a una variable auxiliar η.

En el caso de (37) podríamos considerar formalmente: $p(\eta,u) = \Sigma u_i^2$, H arbitrario y $U = \{u \in \mathbb{R}^n / u \in [\{\cdot\}]\}$, y la desigualdad se habría seguido de la dada por:

$$\max_{\eta \in H} \quad p(\eta,u) \geq \max_{\eta \in H} \quad \min_{u \in U} \quad p(\eta,u)$$

para todo $u \in U$.

Ahora, hay varias diferencias importantes entre la metodología de obtención de condiciones suficientes que pudiera derivarse de cada uno de ambos principios.

La primera es la de que, mientras en el enfoque del Principio del Minimax la intervención de la variable auxiliar η es la misma en ambos miembros de la desigualdad (17), debido al planteamiento de ambos sobre la base de la misma función $p(\eta,u)$, en el caso del Principio de Optimización en Desigualdades la intervención de η en cada uno de los miembros de (4) es, en general, diferente debido a la introducción de la nueva función $q(\eta,u)$.

Otra diferencia importante es la de que, mientras en el planteamiento del Principio (15) la variable η juega un papel activo en la resolución de los problemas de optimización por medio de los cuales se procede a la generación de la condición suficiente, en el caso del Principio (1), η desempeña un papel pasivo en relación a la solución de los correspondientes problemas de minimización, que solo afectan a u, estando restringido su papel solo el de simplificar la resolución del problema de minimización en la función q.

Por último, una tercera diferencia entre ambos enfoques es la de que el resultado establecido mediante el Principio (15) depende de la hipótesis de que el minimax y el maximin se tomen en el conjunto producto $H \times U$ y no en ningún subconjunto de él. Tomando prestada la terminología del Cálculo de Probabilidades y suponiendo probabilizado el subconjunto de $H \times U$ de que se tratara con "probabilidad uniforme", podíamos decir entonces que el establecimiento de (15) dependería de que las dos "proyecciones" η y u fueran "independientes".

Sin embargo, en el planteamiento de (1) la variable η se podría definir en función de cada valor de $u \in U$, ya que su papel es meramente auxiliar y la optimización afecta solo a u.

Es por esto que las aproximaciones a la generación de condiciones suficientes en problemas de optimización derivadas de los Principios (1) y (15) podrían considerarse, en general, independientes.

En cuanto a la relación de cada uno de estos principios con la teoría clásica de la dualidad hay que decir en primer lugar que la aproximación dada por (1) no estaría, en principio, relacionada con el planteamiento de problemas duales de optimización, debido al papel meramente instrumental desempeñado por la variable η en ese caso.

Sí estaría relacionada con dicha teoría, sin embargo, la aproximación dada por (15), aunque nuestra definición de problema "dual en sentido amplio" de uno dado, es más general de las que se suelen utilizar usualmente en los textos sobre el tema.

Así en general, se suele reservar la definición de problemas "duales" solo para aquellos para los que se pueda establecer una conexión entre sus correspondientes soluciones, y se caracterizan porque la función $p(\eta,u)$ elegida suele ser lineal en η.

Como ilustración de lo que acabamos de afirmar vamos a mostrar como se puede generar el planteamiento de un problema dual típico de los que se suelen investigar en los textos sobre Programación No Lineal, en la línea de que nosotros hemos sugerido.

Antes, sin embargo, vamos a introducir unas definiciones y resultados elementales que nos serán útiles para lo que sigue.

Sea $\Omega \subset \mathbb{R}^n$ y $f_o(u)$ una función real definida en Ω. Entonces se dice que f_o es "convexa en $u^* \epsilon \Omega$", con respecto a Ω, si:

$$f_o(\lambda u^* + (1-\lambda)u) \leq \lambda f_o(u^*) + (1-\lambda)\ f_o\ (u)$$

para todo $u \epsilon \Omega$, tal que $\lambda u^* + (1-\lambda) u \epsilon \Omega$, con $0 \leq \lambda \leq 1$.

Por otra parte, se diría que f_o es convexa en $\Omega C R^n$ si fuera convexa en cada $u \epsilon \Omega$, en el sentido anterior.

Supongamos ahora que $\Omega C R^n$ es abierto y que $f_o(u)$ es diferenciable en un entorno de $u^* \epsilon \Omega$ y convexa en u^*. Se demuestra entonces fácilmente que se verificaría:

$$(38) \qquad f_o(u^*) - f_o(u) \leq \nabla f_o(u^*) \cdot (u^* - \mu)$$

para todo $u \epsilon \Omega$, en donde $\nabla f_o(u^*)$ denota el vector gradiente evaluado en el punto $u = u^*$. (Véase la Nota 22).

Consideremos ahora de nuevo el Problema (8) con $f_o(u)$ convexa y las $g_i(u)$, $i=1,2,\ldots,n$ cóncaves en Ω que supondremos abier to. Supongamos además que tanto f_o como las g_i, $i=1,2,\ldots,n$, - son diferenciables en Ω (esto es en cada punto de Ω)

Definamos, al igual que antes, la función:

$$p(\eta,u) = f_o(u) - \sum_{i=1}^{n} \eta_i g_i(u)$$

para todo $\eta\varepsilon H$ y $u\,\varepsilon U$, en donde

$$H = \{\eta\varepsilon R^n / \eta \geq 0\}$$

y:

$$U = \{u\varepsilon\Omega\ / g(u) \geq 0\}$$

Evidentemente de $\eta_i \geq 0$, para todo i se seguiría la convexidad de la función $p(\eta,u)$ en u para cada $\eta\varepsilon H$.

Supongamos ahora que para cada η fijo exista:

$$(39) \qquad d_o(\eta) = \min_{u\varepsilon\Omega} p(\eta,u)$$

Evidentemente:

$$d_o(\eta) \leq \inf_{u\varepsilon U} p(\eta,u), \text{ para todo } \eta\varepsilon H.$$

Por otra parte, es fácil comprobar que de la convexidad y diferenciabilidad de p en u , y de que Ω es un abierto convexo se tendría que si $u^*\varepsilon\Omega$ fuera tal que:

$$p(\eta, u^*) = \max_{u \epsilon \Omega} p(\eta,u) = d_o(\eta)$$

entonces u^* pertenecería al conjunto Ω_η dado por:

(40) $$\Omega_\eta = \{u \epsilon \Omega / \ \nabla f_o(\mu) - \sum_{i=1}^{n} \eta_i \nabla g_i(\mu) = 0 \}$$

En efecto $\nabla p(\eta, u^*) = 0$, será una condición necesaria de - mínimo de acuerdo con que Ω es abierto y u^* debería ser un mínimo local de $p(\eta,u)$, mientras que de la convexidad de p en u para todo $u \epsilon \Omega$ se seguiría, de acuerdo con (38) que:

$$p(\eta, u^*) - p(\eta, u) - \nabla p(\eta, u^*) \cdot (u^*-u) = 0, \text{ para todo } u \epsilon \Omega$$

esto es que u^* resolvería (39) para dicho $\eta \epsilon H$.

Así, de acuerdo con nuestro enfoque se podría definir el - siguiente problema "dual" del (8) en ese caso:

$$\max d_o(\eta)$$
$$\text{sujeto a:} \quad \eta \epsilon H ,$$

el cual, teniendo en cuenta (39) y (40) podría replantearse como:

(41) $$\max p(\eta, u) = \max f_o(u) - \sum_{i=1}^{n} \eta_i g_i(u)$$

sujeto a:

$$(\eta, u) \epsilon \{(\eta, u) \ \epsilon \ \mathbb{R}^{2n} \ / \ \eta \epsilon H, \ u \epsilon \Omega_\eta \} =$$

$$= \{(\eta, u) \epsilon \ \mathbb{R}^{2n} \ / \ u \epsilon \Omega; \ \nabla f_o(u) - \sum_{i=1}^{n} \eta_i \nabla g_i(u) = 0 \ ; \ \eta \geq 0\}$$

El desarrollo anterior justificaría la denominación de problema "dual" del (8) del dado por (41), como se hace en algunos textos sobre el tema.

Esta definición de problema dual en las condiciones enunciadas permitiría establecer, de acuerdo con el Principio (15), que si el Problema (41) tuviera solución (η^*, u^*), con $u^* \in U$ y : $\sum_{i=1}^{n} \eta_i^* \, g_i(u^*) = 0$, entonces u^* sería una solución de (8).

En general, sin embargo se suele proponer a (41) como el problema "dual" del (8) sin ninguna referencia a la convexidad de f_o y la concavidad de las g_i, con tal de que Ω fuera abierto y tanto f_o como las g_i fueron diferenciables en Ω, para que el conjunto de definición del par (η, u) en (41) tuviera sentido. (Véase la Nota 23).

La particularización de f_o y las g_i produciría entonces el planteamiento de correspondientes problemas duales para el Problema de la Programación Cuadrática u otros problemas especiales.

Nuestro interés, sin embargo, no es tanto el de indagar en qué condiciones podría asegurarse la doble solución de (8) y (41), lo que suele ser el objetivo de todos los teoremas operativos de la dualidad, elaborados a imagen del Teorema (50) del capítulo I, sino el de mostrar como el planteamiento de estos problemas duales puede considerarse un subproducto del Principio (15). Cubierto ese objetivo, pues, terminamos aquí la exposición.

NOTAS

0.- Aunque en el enunciado del Teorema se exige que el conjunto sea cerrado no habría dificultad en no considerarlo en principio asi ya que, de cualquier forma, si un punto $\hat{x}$ pertenece a la frontera inferior respecto a su primera coordenada de un conjunto convexo $C \subset R^{n+1}$, también pertenecerá, como es muy facil demostrar, a la frontera inferior respecto a su primera coordenada del conjunto convexo cerrado $\overline{C}$, y podría aplicarse el resultado del Teorema (8) a este último conjunto. Asi, lo relevante en el enunciado del Teorema es la convexidad de C y no el hecho de que sea cerrado o no.

1.- El Teorema (24) se conoce en programación no lineal como la versión del Teorema de Kuhn-Tucker sin diferenciabilidad. La condición de que no exista un vector $\overline{\gamma} \geq 0$, $\overline{\gamma} \neq 0$, tal que $\overline{\gamma}.f(u) \geq 0$, $\forall u \in \Omega$, se conoce con el nombre de "restricción cualificada" de Karlin.

Existen otras "restricciones cualificadas" equivalentes a la anterior como son las de que existe un $u^* \in \Omega$ tal que $f(u^*) < 0$, conocida como la restricción cualificada de Slater, y la de que el conjunto $\{u \in R^m \ / \ f(u) \leq 0, \ u \in \Omega\}$ contenga al menos dos puntos distintos $u^{(1}$ y $u^{(2}$ de manera que f sea estrictamente convexa en $u^{(1}$, es decir que:

$$f(\lambda\ u^{(1} + (1-\lambda)u^{(2}) < \lambda\ f(u^{(1}) + (1-\lambda)f(u^{(2})$$

para todo λ tal que: $0 < \lambda < 1$.

Puede consultarse sobre estas cuestiones el capítulo 5 del texto de O. Mangasarian (Ref. 19). La deducción que se hace en dicho texto del Teorema utiliza un teorema de separación de conjuntos convexos del mismo tipo al utilizado por nosotros. La principal diferencia estriba, sin embargo, en la metodología utilizada, pues mientras Mangasarian deduce el resultado utilizando un teorema "ex profeso" nosotros hemos procurado deducirlo como un caso particular del principio general de la proyección maximal. Esta manera de proceder intenta mantener siempre la visión de conjunto y poner en evidencia la "naturalidad" de los resultados obtenidos. (Véase el anexo de la página siguiente.)

2.- Los resultados de teoría de la medida que se aplican pueden consultarse por ejemplo en P. Halmos (Ref. 12), capítulos IV y V. El resultado concerniente a la aproximación de funciones medibles por continuas (conocido como "teorema de Lusin") puede consultarse en el texto de A.N. Kolmogorov y S.V. Fomin (Ref. 16) pág. 293.

3.- Puede consultarse por ejemplo el texto de E.B. Lee y L. Markus (Ref. 18) pág. 68.

Anexo a la Nota 1.-

Por otra parte hay que señalar que la demostración del Teorema (24) podría haberse aliviado en lo que se refiere a la demostración de la positividad del vector $\bar{\eta}$ si el Problema (17) se hubiera sustituido por el siguiente problema equivalente:

$$\min \; x_o = g_o(\tilde{\bar{u}})$$

sujeto a:

$$x_i = -g_i(\tilde{u}), \quad i=1,2,\ldots,n$$

$$\tilde{u} = (u,u_{m+1},u_{m+2},\ldots,u_{m+n}), \quad \text{con}$$

$$u \in \Omega \quad R^n, \quad \Omega \text{ convexo, y} \quad u_{m+i} \geq 0, \quad i=1,2,\ldots,n$$

$$\text{y:} \quad x \in G = \{x \in R^n \;/\; x_i = 0, \quad i=1,2,\ldots,n\},$$

siendo:

$$g_o(\tilde{u}) = f_o(u), \quad \text{convexa}$$

$$\text{y} \quad g_i(\tilde{u}) = f_i(u) + u_{m+i}, \quad i=1,2,\ldots,n, \quad \text{convexas}$$

La positividad del vector $\bar{\eta}$ se hubiera seguido entonces fácilmente de la arbitrariedad de los u_{m+i} (siempre que sean positivas) en la expresión:

$$f_o(\bar{u}) + \sum_{i\in N_1} \bar{\eta}_i f_i(\bar{u}) \leq f_o(u) + \sum_{i\in N_1} \bar{\eta}_i f_i(u) + \sum_{i\in N_1} \bar{\eta}_i u_{m+i}$$

3 bis .- La demostración de este teorema, que suele ser, en general, algo compleja, porque involucra, entre otras, la demostración previa del teorema de Liapunov sobre el recorrido convexo de un funcional vectorial sobre funciones medibles, puede consultarse en el texto de Lee y Markus. (Ref. 18) pp 69-72 y 155-168.

4.- Las condiciones a), b) y c) del enunciado del Teorema (79) se ban tomado directamente del texto de Pontryaguin (Ref. 23) pp. 75-82, aunque la versión que aqui se ofrece de su Principio de Máximo en el caso convexo es ligeramente distinta No hay que olvidar respecto a ello que el enfoque del Principio de Maximo tal como lo enunció inicialmente Pontryaguin corresponde al caso diferencible y no al caso convexo.

5.- Sobre este resultado puede consultarse el texto de Hewitt y Stromberg (Ref. 15) pp. 272-275, y el de Titchmarsh (Ref. 28) pp. 360-362.

6.- En el momento actual la teoría de las condiciones necesarias en optimización se puede subdividir, en general, en los casos convexo y diferenciable. No obstante, el enunciado inicial de los teoremas mas importantes en los casos estático y dinámico, esto es el teorema de Kuhn-Tucker y el Principio de Máximo de Pontryaguin se referían solo al caso diferenciable que se puede considerar el mas importante. Ello parece debido a la tradición sobre la exigencia de condiciones de diferenciabilidad en problemas de optimización desde la obtención de las primeras condiciones necesarias en esa línea por Fermat. Asi, la primera exposición del teorema de Kuhn-Tucker (Ref. 17) y del Principio de Máximo de Pontryaguin, en la línea, este último, del Cálculo de Variaciones clásico (Ref. 6), se reducen al caso diferenciable.

Algunas condiciones necesarias de optimalidad sin diferenciabilidad en programación no lineal se empezaron a obtener a raíz de los desarrollos de la teoría de la dualidad en programación lineal. Un buen resumen de ellas se encuentra en el capítulo 5 del texto de Mangasarian (Ref. 19). Por su parte, el estudio del caso convexo en los problemas de control arranca principalmente de la utilización del teorema de Liapunov sobre el recorrido convexo de un funcional en el caso de problemas con dinámica lineal. Un ejemplo de estos desarrollos se puede encontrar en la aproximación de Halkin (Ref. 11) a la demostración del Principio de Máximo. En este contexto la convexidad se suele considerar, en general, nada más que un paso intermedio en la demostración del Principio de Máximo por medio del teorema de Liapunov, al que seguirá una linealización del caso no lineal, como puede comprobarse también en el texto de Lee y Markus (Ref. 18), capítulos 2, 3 y 4, y no suele investigarse independientemente, debido por supuesto, a la propia naturaleza diferencial del planteamiento, lo que exigiria en cualquier caso de propiedades de diferenciabilidad a la función f(x,u) con respecto a x, sobre las que se añadirían las de convexidad. No hay que olvidar además que, en cualquier caso, el Teorema (8) o su versión conocida como "teorema del hiperplano separador", juegan un papel principal en las demostraciones clásicas del Principio de Máximo.

El caso convexo suele estudiarse también en relación con las condiciones suficientes de optimalidad, las cuales siguen facilmente de la unicidad de los mínimos relativos de las funciones convexas.

Nosotros, sin embargo, hemos presentado independientemente ambos casos, convexo y diferenciable, para mostrar hasta donde se puede llegar con las hipótesis del primer caso sin necesidad de suponer las del segundo.

7.- La demostración del Principio de Máximo de Pontryaguin con toda generalidad, tal como se expone por ejemplo en el capítulo 2 de la Ref. 23 , dedica algunos lemas previos (los lemas 3 y 4 de las páginas 94-99) a probar un resultado del mismo tipo que el Teorema (91) para unos conjuntos $L_{\hat{x}}$ definidos previamente (en el lema 2 desarrollando en las páginas 86-94). A este conjunto de lemas previos los denominan los autores "lemas fundamentales".

Quizas convenga hacer algún comentario sobre la génesis de la demostración que aparece en el texto de Pontryaguin. Siguiendo a V. Boltianskii (Ref. 5 ; nota al pie de la pág. 446) podemos decir que el Principio de Máximo se enunció en 1956 por Pontryaguin, Boltianskii y Gamkrelizde(ver pág. 429 del texto de Feldbaum, Ref. 7 en el que se cita, entre otras, la Ref. 8) en forma de conjetura para el caso particular del problema de tiempo óptimo. Es más Pontryaguin parecía bus car obtener una condición suficiente, mas bien que necesaria, de optimalidad. La demostración del Principio de Máximo y su enunciado bien precisado en la forma de "condición necesaria" de optimalidad se debe, sin embargo, según el mismo indica, a Boltianskii (la referencia de la demostración aludida es: V. Boltianskii: "Principio de Maximo en la teoría de los procesos óptimos", "Doklady" de la Academia de Ciencias de la URSS, 119, n° 6, 1958, pp. 1070-1073).

Basta una ojeada a la Referencia para comprender que el enfoque geométrico que preside la demostración del Principio de Máximo de la Ref. ha sido inspirado efectivamente por Boltionskii. En la Ref. Boltianskii desarrolla un método general para afrontar geometricamente el problema de las condiciones necesarias de optimalidad en el caso discreto, aunque el método puede aplicarse también al caso continuo. Las bases de dicho método se establecen en el capítulo IV mediante la definición de la "cúpula" de un conjunto C en un punto $\bar{x}$. Estas "cúpulas" estan directamente relacionadas con ciertos conos convexos que se pueden generar a partir de los conjuntos $L_{\hat{\bar{x}}}$ que hemos definido en (92). En este sentido un resultado crucial en toda la teoría, el cual se utiliza también en la demostración del lema 3 de la Ref. 23 , y que Boltianskii denomina "teorema de intersección", puede considerarse similar a nuestro Teorema (91), aunque no coincide exactamente con él.

Una idea de la complejidad del "teorema de intersección" la da el hecho de que su demostración, que utiliza profusamente conceptos de topología algebraica se extiende desde la pág. 290 a la 303 de la Ref. 5 e incluye sin prueba la utilización de algunos teoremas para cuya demostración remite a otros textos. Nosotros hemos preferido recorrer otro camino que evita estas complicaciones.

7 bis .- Es interesante hacer notar que la tesis del Principio (91) se mantiene aún en situaciones más generales que la expuesta En efecto basta repasar la demostración para concluir que se pueden seguir sosteniendo las afirmaciones del teorema aún en el caso de que los límites:

$$\lim_{\substack{\delta \downarrow 0 \\ \hat{x}(\delta) \in C}} \frac{\hat{x}(\delta) - \hat{x}}{\delta}$$

se sustituyan por los límites:

$$\lim_{\substack{r \to \infty,\ \delta \quad 0 \\ \hat{x}(\delta^{(r}) \in C}} \frac{\hat{x}(\delta^{(r}) - \hat{x}}{\delta^{(r}}$$

para cualquier sucesión $\delta^{(r}$ tal que el límite exista. Es más, el denominador en la expresión anterior, $\delta^{(r}$, podría sustituirse por cualquier expresión $g(\delta^{(r})$ tal que $g(\delta^{(r}) > 0$ para todo r y $g(\delta^{(r}) \to 0$, de manera que el correspondiente límite existiera.

8.- Sobre las restricciones cualificadas una buena referencia es el capítulo IV del texto antes citado de Mangasarian (Ref. 19) en el que se comentan varias de estas restricciones. Entre ellas se encuentra la de Kuhn-Tucker.

El Teorema (95), salvo en lo que se refiere a la continuidad de las derivadas parciales, se conoce con el nombre de "Teorema de Fritz John", y precede, en los textos, a la discusión de las restricciones cualificadas. Se suele prescindir de la restricción que hemos dado nosotros en el Corolario (105) porque aparenta ser menos general que la de Kuhn-Tucker. Esta restricción cualificada del Corolario (105) está en la línea de la restricción cualificada de Karlin en el caso convexo.

9.- La demostración del "teorema de la función inversa" para funciones vectoriales de varias variables puede verse en cualquier texto de cálculo avanzado. Por ejemplo son clásicos el texto de W. Fleming (Ref. 9) y el de M. Spivak (Ref. 26). Los teoremas relativos a variedades diferenciables y, en general, la Topología Diferencial, son unos instrumentos útiles en la generación de las condiciones necesarias de óptimo con diferenciabilidad.

10.- El Teorema (112) suele demostrarse utilizando el lema de Farkas, tal como lo hicieron por primera vez sus autores Kuhn y Tucker (Ref. 17). Este lema es, sin embargo, una consecuencia inmediata del teorema de la Dualidad en programación lineal.

En efecto, el lema de Farkas afirma que la desigualdad $\gamma x \leq 0$ vale para todos los n-vectores que satisfacen el sistema de m desigualdades: $Ax \geq 0$, solo si $\gamma = A'\lambda$ para algún m-vector $\lambda \geq 0$.

Veamos que se deduce del Teorema de la Dualidad.
Si el problema:

$$\begin{aligned} &\max \ \gamma x \\ &\text{sujeto a:} \\ &\quad -Ax \leq 0 \\ &\quad\ x \text{ cualquiera} \end{aligned}$$

tiene solución finita, $x = 0$, también la tendrá necesariamente el problema dual:

$$\begin{aligned} &\min \ 0.\lambda \\ &\quad A'\lambda = \gamma \\ &\quad\ \lambda \geq 0 \end{aligned}$$

de lo que sigue el lema.

De hecho, todos los teoremas conocidos como "teoremas de alternativa", no son más que corolarios del teorema de la Dualidad, siempre y cuando consideremos a éste último como el teorema primario. Naturalmente hay que señalar que el teorema de la Dualidad es relativamente reciente frente al lema de Farkas y a otros teoremas sobre las soluciones de sistemas de desigualdades lineales, tal como el de Motzkin.

11.- En el texto de Pontryaguin (Ref. 23), por ejemplo, se supone que cada uno de los conjuntos $G_i \subset R^n$ se puede definir mediante el sistema de ecuaciones:

$$f_j^{(i}(x) = 0 \ ; \quad j=1,2,\ldots,(n-\Gamma) \ ; \quad i=1,2$$

en donde las $f_j^{(i}$ se suponen continuamente diferenciables en sendos abiertos que contengan a cada G_i, y tales que las matrices jacobianas:

$$\nabla^{(i}f = \left(\frac{\partial f_k}{\partial x_j}\right) \; ; \qquad \begin{array}{l} 0 \le k \le \Gamma_i \le n \\ 1 \le j \le n \end{array}$$

tengan rango máximo (es decir Γ_i).

Los conjuntos de puntos que verifican condiciones del tipo anterior se suelen denominar en los textos anglosajones "varieties" (por ejemplo puede consultarse el texto de Austander y Mackenzie, Ref. 2) dejando la expresión "manifold" para la generalización del anterior concepto a subconjuntos de espacios generales caracterizados por la existencia de una colección de difeomorfismos locales entre cada conjunto de una colección de entornos abiertos de cada punto de la variedad y un abierto de $\mathbb{R}^{\Gamma i}$.

12.- Un análisis de diversos teoremas de existencia y unicidad para sistemas de ecuaciones diferenciales con parámetros puede verse en el apéndice del texto de Hestenes (Ref. 14)y en el de Lee-Markus (Ref. 18) pp.55.59

13.- La definición de "punto regular" y su aplicación en la evaluación de integrales en intervalos infinitesimales está tomada del texto de Pontryaguin (Ref. 23) pero se encuentra en la mayoría de los textos de análisis que incluyen la teoría de la integral de Lebesgue. Por ejemplo en el texto de W. Rudin (Ref. 25) pág. 177 se hace referencia a los "puntos de densidad" de conjuntos Lebesgue-medibles, los cuales corresponderían a "puntos regulares" de funciones medibles. Otro texto que se puede consultar es el de Hewitt y Stromberg (Ref. 15) que en la pág. 277 establece la fórmula (121). Estos autores utilizan la denominación de "punto de Lebesgue" en vez de "punto regular".

14.- Puede consultarse, por ejemplo, el texto de Kolmogorov y Fomin (Ref. 16) pág. 102.

15.- Los resultados que se utilizan de teoría de la medida e integración de Lebesgue pueden verse en cualquier texto estandar sobre el tema, y en particular en el de Halmos (Ref.12).

16.- Sobre la existencia de solución en sistemas de ecuaciones diferenciales lineales con segundos miembros medibles puede consultarse el texto de Lee y Markus (Ref. 18) pp. 59-68, y sobre relaciones entre funciones medibles e integrales y teoremas relativos a funciones integrables el texto de Halmos (Ref.12), pp 110-112.

17.- La dependencia de la solución de un sistema de ecuaciones diferenciales en el caso de segundo miembro medible se comenta en el texto de Pontryaguin (Ref. 23) pag. 83 y de una manera mas precisa, aunque sin demostración, en el texto de Lee y Markus (Ref. 18), pp. 57-58.

18.- Esta desigualdad elemental puede consultarse en cualquier texto sobre desigualdades por ejemplo el de Beckenbach y Bellman (Ref. 3), pag. 2. En este texto en las páginas 133 134 se comentan también las desigualdades entre soluciones de ecuaciones diferenciadas líneales que intervienen en nuestra discusión.

19.- Este resultado puede verse en la mayoría de los textos en ecuaciones diferenciales y teoria del control y en particular en el de Athans y Falb (Ref. 1) pags. 147-148.

20.- Una introducción actual a la Programación Geometrica puede verse en el artículo de E.L. Peterson (Ref. 22).

21.- La demostración de la desigualdad entre las medias aritmética y geométrica ha ocupado a muchos autores, porque es una desigualdad clave y porque en general cada una de dichas posibles demostraciones suelo clarificar las raices de ésta y otras desigualdades. Una colección de demostraciones de la desigualdad puede verse en el texto de Hardy, Littlewood y Polya (Ref. 13), pp. 16-21, y en el de Beckenbach y Bellman (Ref. 3) pp. 3-11. En particular allí se popone su demostra ción mediante la resolución del siguiente problema de optimi zación: min $\sum_{i=1}^{n} u_i$, con $\prod_{i=1}^{n} u_i = 1$, $u_i > 0$ para todo i, y la utilización de los multiplicadores de Lagrange. Esta demostración sería, sin embargo, menos elemental que la nuestra, que no requiere en absoluto ninguna referencia al cálculo diferencial, y más aún, que no exige ni siguiera la utilización del concepto de límite o continuidad, en el caso de las medias ponderadas en coeficientes racionales.

22.- La definición de convexidad en un punto de un conjunto Ω no necesariamente convexo y la prolongación de ésta a todo Ω puede verse en el texto de Mangasarian (Ref. 19), pag. 55. En las páginas 83 y 84 de este texto puede verse también una demostracion de (38).

23.- El planteamiento del Problema (41) como el "dual" del (8) con las solas condiciones de que Ω sea abierto y la diferenciabilidad de f_o y las g_i, i=1,2,...,n, en Ω, puede verse en el texto de Mangasarian (Ref.19), pag. 114. En este texto se enuncian también a continuación varios teoremas en los que se establecen condiciones para la existencia y coincidencia entre las soluciones de (8) y (41). Un enfoque más general del planteamiento de problemas duales en el caso de que f_o sea convexa puede verse, sin embargo, en el capítulo 5 del texto de Stoer y Witzgall (Ref. 27), pp 177-220 y en las secciones 30-31 de la Parte VI del texto de Rockafellar (Ref.24) En estos dos textos la introducción de la dualidad se hace mediante la definición previa de las denominadas"funciones conjugadas" definidas para funciones convexas y concavas, las cuales se podrian considerar en la línea de muestra $d_o(\eta)$ En ambos enfoques un resultado clave es el denominado Teorema de la Dualidad de Fenchel, mediante el que se establece en determinadas ocasiones la coincidencia entre las soluciones de dos problemas duales caracterizados por funciones objetivo consistentes en la diferencia entre una función convexa y una concava, por una parte, y las correspondientes funciones conjugadas por otra. La particularización de este Teorema permite obtener, como muestran estos autores, otros muchos teoremas especialesde la teoría de la Dualidad.

24.- Sobre los primeros inicios de la teoria actual de los multiplicadores en la teoria del control óptimo puede verse el artículo de McShane (Ref. 20) y el de Bliss (Ref. 4) del - - cual el anterior puede considerarse una prolongación . Una demostración del Principio de Maximo de Pontryaguin a partir de una regla de existencia de multiplicadores en espacios - abstractos puede verse por ejemplo en el capítulo II del texto de Fleming y Rishel (Ref. 10), mientras que una exposición integrada de este tipo de reglas, aplicadas en las obtención de condiciones necesarias en problemas de optimización planteados en espacios arbitrarios puede verse en el texto de - Neustadt (Ref. 21).

BIBLIOGRAFIA

1. M. Athans y P.L. Falb: "Optimal Control", Mc Graw-Hill, Nueva York 1966.

2. L. Auslander y R. Mackenzie: "Introductión to Differentiable Manifolds" - - Dover Publications,NuevaYork, 1963.

3. E.F. Beckenbanch y R. Bellman: "Inequalites", Springer-Verlag, Nueva York, 1971.

4. G.A. Bliss: "The Problem of Lagrange in the Calculus of Variations" , Ameri can Journal of Mathematics, vol, 52, pp 673-744, 1930.

5. V.Boltianskii:"Commande Optimale des Systemes Discrets", ediciones Mir, - - Moscu, 1976.

6. V.G. Boltyanskii, R.V. Gamkrelidze y L.S. Pontryagin: "On the Theory of Op timal Processes", en R. Bellman y R.Kalaba (editores): "Selec Papers on - Mathematical trends in control Theory",pp 172-176, Dover publications, Nue va York, 1964.

7. A. Feldbaum; "Principes Theoriques des Systemes Asservis Optimaux", Edicio nes Mir, Moscú, 1973.

8. V. Boltianskii, R. Gamkrelizde y L. Pontriaguine: "Contribución a la Teoria de los Procesos Optimos" (en ruso). Actas de la Academía de Ciencias en la U.R.S.S., 1956, t. 110, n°1, pp, 7-10.

9. W.H. Fleming : " Funciones de Varias Variables", Compañía Editorial Continen tal, México, 1969.

10. W.H. Fleming y R.W. Rishel: "Deterministic and Stochastic Opitmal Control", Springer-Verlag, Nueva York, 1975.

11. H. Halkin: "Mathematical Foundations of System Optimzation", en G. Leitmann (editor): "Topics in Optimization", pp 198-260, Academic Piess, Nueva York, 1976.

12. P.L. Halmos: "Measure Theory", Van Nostrand Reinhold, Nueva York, 1950.

13. G.H. Hardy, J.E. Littlewood y G. Polya: "Inequalities", Cambridge University Press, Londres, 1951.

14. M.R. Hestenes: "Calculus of Variations and Optimal Control Theory", John - Wiley and Sons, Nueva York, 1966.

15. E. Hewitt y K. Stromberg: "Real and Abstract Analysis", Springer-Verlag, - Nueva York, 1975.

16. A.N. Kolmogorov y S.V. Fomin: "Introductory Real Analysis", Dover Publications, Nueva York, 1970.

17. H.W. Kuhn y A.W. Tucker: "Nonlinear Programming", en J. Neyman (editor): "Proccedings of the Second Berkeley Sympsosium of Mathematical Statistics and Probability", pp 481-492, University of California Press, Bekerley, California, 1951.

18. E.B. Lee y L. Markus: "Foundations of Optimal Control Theory" John Wiley - and Sons, Nueva York, 1976.

19. O.L. Mangasarian: "Nonlinear Programming", Mcgraw-Hill, Nueva York, 1969.

20. E.J. McShane: "On multipliers for Lagrange problems", American Journal of - Mathematics, Vol 61, pp 809-819, 1939.

21. L.W. Neustadt: "Optimization", Princenton University Press, Princenton, New Jersey, 1976.

22. E.L. Peterson: "Geometric Programming" en M. Avriel (editor): "Advances in Geometric Programming", pp 31-95. Plenum Press, Nueva York, 1980.

23. L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelizde y E.F. Mishchenko: - - "The Mathematical Theory of Optimal Process", John Wiley Interscience, Nueva York, 1962.

24. R.T. Rochafellar: "Convex Analysis", Princeton University Press, Princeton New Jersey, 1970.

25. W. Rudin: "Real and Complex Analysis", McGraw-Hill, Nueva York, 1970.

26. M. Spivak: "Calculo en Variedades", Editorial Reverté, Barcelona, 1975.

27. J. Stoer, y C. Witzgall: "Convexity an Optimization in Finite Dimensions I", Springer-Verlag, Nueva York, 1970.

28. E.C. Titchmarsh: "The Theory of Functions", Oxford University Press, Londres, 1939.