UNIVERSIDAD NACIONAL DE INGENIERIA

FACULTAD DE CIENCIAS

SECCION DE POST GRADO Y 2DA ESPECIALIZACION PROFESIONAL

"PROGRAMACION CUADRATICA: TEORIA, ALGORITMOS Y APLICACIONES"

TESIS

PARA OPTAR EL GRADO DE MAESTRO EN CIENCIAS

MENCION EN MATEMATICAS APLICADAS

JOSE LUIS MORALES ALMORA

LIMA - PERU

1990

INTRODUCCION:

El campo de la Programación Cuadratica se inició en la década de 1950-60. En la bibliografía se indica los principales trabajos de esa época dedicados a la Programación Cuadrática, [9], [11], [13], [14], [17]. Este campo mantiene su importancia hasta la fecha, porque problemas de aproximación por el método de los mínimos cuadrados, y discretizaciones apropiadas de problemas importantes de Control Optimo, conducen a programas cuadráticos convexos. De otro lado en los últimos años se han desarrollado algoritmos eficientes de la Programación No-lineal que en sub-pasos resuelven programas cuadráticos.

En el capítulo I presentamos en forma resumida los aspectos fundamentales de las *CONDICIONES DE OPTIMIDAD* para Programas matemáticos **P**, que usan la función de *LAGRANGE* del programa **P**. Presentamos tambien las condiciones de *KUHN-TUCKER*, y las condiciones de *JOHN* en su forma global y linealizada.

En el capítulo II presentamos la teoría de dualidad de *DANTZIG, EISENBERG Y COTTLE*, que incluye la teoría de dualidas para Programas Cuadráticos Convexos como caso especial. Para desorrollar esta teoría usamos un teorema de existencia para programas cuadráticos.

En el capítulo III presentamos el algoritmo de *WOLFE* para desarrollar el problema del tipo:

$$\mathbf{P} \quad \min \{F(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} \ / \ \mathbf{A}\mathbf{x} = \mathbf{b} \quad \mathbf{x} \geq 0\}$$

donde $\mathbf{C} = \mathbf{C}^T \in M(n,n)$ es definida positiva.

El método de *WOLFE* se basa en las condicones de K-T. Este método consiste en una modificación del algoritmo *SIMPLEX*, aplicado a un programa lineal, deducido de las condiciones linealizadas de *KUHN-TUCKER*.

En el capítulo IV presentamos el algoritmo de *BEALE* para resolver el problema del tipo:

$$\mathbf{P} : \quad \min \{F(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} \ / \ \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$$

este método al igual que el método de *WOLFE* es una modificación del algoritmo de *SIMPLEX*. Este método se puede ejecutar usando tablas.

En el capítulo V presentamos el método de *ZUNGWILL* el cual considera el programa:

$$\mathbf{P} : \quad \min \{F(\mathbf{x}) \ / \ \mathbf{A}\mathbf{x} \leq \mathbf{b}\} , \quad \mathbf{A} \in M(m,n)$$

donde la función F es convexa y diferenciable. Este método resuelve una sucesión de programas:

$$\min \{F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} = b_i, \ \ \mathbf{i} \in \mathbf{I}_k\}, \ \ \mathbf{I}_k \subset \{1,2,\ldots,m\}$$

Los conjuntos de índices se cambian sucesivamente de tal modo que en un número finito de iteraciones se obtiene:

$$\mathbf{I}_k = \mathbf{I}(\hat{\mathbf{x}}) = \{j \ / \ 1 \leq j \leq n \ / \ \mathbf{a}_j^T\hat{\mathbf{x}} = b_j\}$$

donde $\hat{\mathbf{x}}$ es una solución óptima de $\mathbf{P}$.

Aplicando el algoritmo de *ZUNGWILL* a un programa cuadrático convexo, obtenemos el algoritmo de *DANTZIG Y COTTLE*.

Finalmente en el capítulo VI damos dos aplicaciones de esta teoría; específicamente un "Problema de Portafolio" (Problema de selección de inversiones) y una discretización del problema del Regulador Lineal, que es uno de los problemas mas importantes de Control Optimo. Esta discretización conduce a programas cuadráticos convexos.

Jose Luis Morales Almora

INDICE

Pgs.

CAPITULO I: CONDICIONES DE OPTIMIDAD

1.1.- Introducción 1

1.2.- Definiciones 1

CAPITULO II: TEORRIA DE DUALIDAD

2.1.- Introducción 16

2.2.- La Teoría de Dualidad de Dantzig 18

2.3.- Teoría de Dualidad para Programas cuadráticos convexos 37

CAPITULO III: EL ALGORITMO DE WOLFE

3.1.- El Algoritmo de Wolfe 42

3.1.1.- Modificación del algoritmo Simplex 45

3.1.2.- Regla adicional

3.2.- Ejemplo 51

CAPITULO IV: EL ALGORITMO DE BEALE

4.1.- El Algoritmo de Beale 57

4.2.- Resumen 66

4.3.- La convergencia del Algoritmo 68

4.4.- Ejemplo 75

CAPITULO V: LOS METODOS DE ZUNGWILL y DANTZIG-COTTLE

5.1.- El caso convexo 79

5.1.1.- El Algoritmo 83

5.2.- El caso cuadrático 87

(El Algoritmo de Dantzig y Cottle)

5.2.1. El Algoritmo (Resumen) 102

5.2.2.- Ejemplo 105

CAPITULO VI: APLICACIONES

6.1.- Un Problema de Portafolio 113

6.2.- Un Algoritmo para resolver el problema del Regulador Lineal, con dominios de control y de estado 116

6.2.1.- Una discretización del Problema 117

6.2.2.- La solución del Problema aproximado 119

6.2.3.- Ejemplo 120

BIBLIOGRAFIA

CAPITULO

CONDICIONES DE OPTIMIDAD

1.1: INTRODUCCION

En este capítulo presentaremos en forma resumida los aspectos fundamentales de las condiciones de optimidad para Programas Matemáticos **P**, que usan la *FUNCION DE LAGRANGE* del programa **P**.

1.2: DEFINICIONES:

Sean $\mathbf{C} \subset \mathbb{R}^n$ convexo, $F: \mathbf{C} \longrightarrow \mathbb{R}$

$f: \mathbf{C} \longrightarrow \mathbb{R}^m$, $g: \mathbf{C} \longrightarrow \mathbb{R}^\ell$, funciones continuas.

Consideremos el programa:

$$\mathbf{P}: \min \{F(\mathbf{x}) \;/\; \mathbf{x} \in \mathbf{C},\; f(\mathbf{x}) \le 0,\; g(\mathbf{x}) = 0\}$$

1. Si $\mathbf{u} = (u^1, u^2) \in \mathbb{R}^m \times \mathbb{R}^\ell$, entonces la función

   $\Phi: \mathbf{C} \times \mathbb{R}^m \times \mathbb{R}^\ell \longrightarrow \mathbb{R}$, definida por:

   $\Phi(\mathbf{x},\mathbf{u}) = F(\mathbf{x}) + (u^1)^T f(\mathbf{x}) + (u^2)^T g(\mathbf{x})$

   se llama la *FUNCION DE LAGRANGE* del programa **P**.

2. Si $\mathbf{v} = (v^0, v^1, v^2) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^\ell$, entonces la función

   $L: \mathbf{C} \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^\ell \longrightarrow \mathbb{R}$ definida por:

$$L(\mathbf{x},\mathbf{v}) = v^0F(\mathbf{x}) + (\mathbf{v}^1)^T f(\mathbf{x}) + (v^2)^T g(\mathbf{x})$$

se llama la *FUNCION GENERALIZADA DE LAGRANGE* para **P**.

3. El sistema:

α) $\Phi(\hat{\mathbf{x}},\hat{\mathbf{u}}) \geq \Phi(\mathbf{x},\hat{\mathbf{u}}) \qquad \forall\ \mathbf{x} \in C$

β) $\Phi(\hat{\mathbf{x}},\hat{\mathbf{u}}) = F(x)$

γ) $\hat{\mathbf{u}}^1 \geq 0$

δ) $f(\hat{\mathbf{x}}) \leq 0\ ,\quad g(\hat{\mathbf{x}}) = 0 \qquad \hat{\mathbf{x}} \in C$

Se llama *CONDICIONES DE KUHN-TUCKER* para el programa **P** en el punto $(\hat{\mathbf{x}},\hat{\mathbf{u}})$.

Si γ) y δ) son cumplidas, entonces β) se cumple exactamente si $\hat{u}^1 = 0$, siempre y cuando $f_j(\hat{x}) \quad 0$.

Si todas las funciones que intervienen en el programa **P** son diferenciables en $\hat{\mathbf{x}}$ y se sustituye la condición α) por:

α^*) $\nabla_x\Phi(\hat{\mathbf{x}},\hat{\mathbf{u}})^T(\mathbf{x}-\hat{\mathbf{x}}) \geq 0 \qquad \forall\ \mathbf{x} \in C$, entonces el sistema α^*), β), γ) y δ) se llama *CONDICIONES LINEALIZADAS DE KUHN-TUCKER* para **P**.

4. El sistema

a) $L(\hat{\mathbf{x}},\hat{\mathbf{v}}) \leq L(\mathbf{x},\hat{\mathbf{v}}) \qquad \forall\ \mathbf{x} \in C$

b) $L(\hat{\mathbf{x}},\hat{\mathbf{v}}) = \hat{v}^0F(\hat{\mathbf{x}})$

c) $(\hat{v}^0,\hat{v}^1) \geq 0\ ,\quad \mathbf{v} \neq 0$

d) $\mathbf{x}$ es admisible para **P**.

Se llama *CONDICIONES DE JOHN* para **P**, en el punto $(\hat{\mathbf{x}},\hat{\mathbf{v}})$

Si d) cumple y $\hat{v}^1 \geq 0$, entonces b) exactamente satisfecha si $\hat{v}^1_j = 0$, siempre y cuando $f_j(\mathbf{x}) \quad 0$.

Si todas las funciones que intervienen en **P** son diferenciables en $\hat{x} \in C$, definimos la condición:

$$a^{*}) \quad \nabla_x L(\hat{x}, \hat{v})(x - \hat{x}) \geq 0 \qquad \forall\, x \in C.$$

El sistema a^{*}), b), c) y d) se llama entonces *CONDICIONES LINEALIZADAS DE JOHN* para **P**.

<u>OBSERVACION</u> <u>1.1:</u> Si en $(\hat{x}, \hat{v})$ se cumplen las condiciones (linealizadas) de JOHN para **P** y si $\hat{v}^{o} > 0$, entonces el punto $(\hat{x}, \hat{u}) = \left[\hat{x}, \frac{\hat{v}^1}{\hat{v}^o}, \frac{\hat{v}^2}{\hat{v}^o}\right]$ se cumplen las condiciones (linealizadas) de KUHN-TUCKER para **P**.

5. Sean $C \subset \mathbb{R}^n$ convexo, f convexa, g afín lineal, $\hat{x}$ una solución óptima de **P** y $I(\hat{x}) = \{j \,/\, f_j(\hat{x}) = 0\}$.

Decimos que se cumplen las *CONDICIONES DE REGULARIDAD DE SLATER* para P, si existe un punto $x^o \in$ intrel(C) tal que:

$$f_j(x^o) < 0 \qquad \forall\, j \in I(x)$$

$$g(x^o) = 0$$

Si adicionalmente las funciones f_j, $j \in I(\hat{x})$ son diferenciables en **x**, entonces decimos que las CONDICIONES LINEALIZADAS DE SLATER para **P** son cumplidas, si existe un $x^o \in$ intrel(C) tal que :

$$\nabla f_j(\hat{x})^T (x^o - \hat{x}) < 0 \qquad \forall\, j \in I(\hat{x})$$

$$\nabla g_k(\hat{x})^T (x^o - x) = 0 \qquad k = 1, 2, \ldots, \ell$$

<u>OBSERVACIONES 1.2:</u>

1. Si C es abierto, por ejemplo $C = \mathbb{R}^n$, entonces las condiciones linealizadas α^{*}) respectivamente a^{*}) son equivalentes a las condiciones:

$\tilde{\alpha}$) $\nabla_x \Phi(\hat{x}, \hat{u}) =$

$\tilde{\beta}$) $\nabla_x L(\hat{x}, v) =$

DEMOSTRACION

Veamos que α^*) es equivalente a $\tilde{\alpha}$).

$\alpha^*) \Longrightarrow \tilde{\alpha})$:

Como $\hat{x} \in C = \text{int } (C)$, entonces existe $\varepsilon > 0$ tal que

$x \pm \varepsilon e_i \in C \quad j = \overline{1,n}$

- Si $x = \hat{x} + \varepsilon e_j$, entonces obtenemos de α^*)

$$0 \le \nabla_x \Phi(\hat{x},\hat{u})^T (x - \hat{x}) = \varepsilon \nabla_x \Phi(\hat{x},\hat{u})^T e_j$$
$$= \varepsilon \frac{\partial}{\partial x_j} \Phi(\hat{x},\hat{u}); \qquad j = \overline{1,n}$$

Por consiguiente : $0 \le \nabla_x \Phi(\hat{x},\hat{u})$ (i)

- Si $x = \hat{x} - \varepsilon e_j$, entonces obtenemos de α^*)

$$0 \le \nabla_x \Phi(\hat{x},\hat{u})^T (x-\hat{x}) = -\varepsilon \nabla_x \Phi(\hat{x},\hat{u})^T e_j = -\varepsilon \frac{\partial}{\partial x_j} \Phi(\hat{x},\hat{u}); \quad j = \overline{1,n}$$

Por consiguiente : $0 \ge \nabla_x \Phi(\hat{x},\hat{u})$ (ii)

De (i) y (ii) obtenemos $\tilde{\alpha}$)

Que $\tilde{\alpha}) \Longrightarrow \alpha^*)$ es evidente. ∎

2. Si $C = \mathbb{R}^n_+$, entonces las condiciones linealizadas de KUHN-TUCKER para **P** son equivalentes con el sistema

$$\left.\begin{array}{ll} \nabla_x \Phi(\hat{x},\hat{u}) \ge 0 & \hat{x}^T \nabla_x \Phi(\hat{x},\hat{u}) = 0 \\ \Phi(\hat{x},\hat{u}) = F(x) & \\ \hat{u}^1 \ge 0 & \\ \hat{x} \text{ es admisible para P} & \end{array}\right\} \quad (*)$$

DEMOSTRACION

a) Sean cumplidas las condiciones linealizadas de K-T para P en el punto $(\hat{x},\hat{u})$. Para $j = 1, n$ elegimos $x^j \in \mathbb{R}^n_+$ sigue:

$$x_i^j \quad x_i \qquad \forall\ i \neq$$

$$x_j^j = x_j +$$

Entonces se obtiene de α^*)

$$\nabla_x \Phi(\hat{x},\hat{u})^T (x^j \quad \hat{x}) - \nabla_x \Phi(\hat{x},\hat{u})_j \geq 0 \qquad \forall\ j$$

$$\nabla_x \Phi(\hat{x},\hat{u})_j \geq 0 \tag{+}$$

Si $x_j > 0$, elgimos $Y^j \in C = \mathbb{R}^n_+$ como sigue:

$$Y_i^j \quad x_i \qquad \forall\ i \neq j$$

$$Y^j = 0.$$

Entonces: $\nabla_x \Phi(\hat{x},\hat{u})^T (Y^j \quad x) = \quad \hat{x}_j\ \nabla_x \Phi(\hat{x},\hat{u})_j \geq 0, \quad (-\ x_j < 0)$

Por (+) esto solamente es posible si $\nabla_x \Phi(\hat{x},\hat{u})_j = 0.$

Por consiguiente se cumple $\hat{x}^T \nabla_x \Phi(\hat{x},\hat{u}) = 0$

b) Sean cumplidas las condiciones $\quad$ y $x \in C = \mathbb{R}^n_+$ cualquiera De (*) obtenemos

$$\nabla_x \Phi(\hat{x},\hat{u})^T (x \quad \hat{x}) = \nabla_x \Phi(\hat{x},\hat{u})^T x - \nabla_x \Phi(\hat{x},\hat{u})^T \hat{x}$$

$$= \nabla_x \Phi(\hat{x},\hat{u})^T x \geq 0 \qquad \forall\ x \in \mathbb{R}^n_+ \qquad \blacksquare$$

3. Si en el punto $(\hat{x},\hat{u})$ [respectivamente $(\hat{x},\hat{v})$] se cumplen las condiciones de K-T [condiciones de JOHN respectivamente] entonces en $(\hat{x},\hat{u})$ [$(\hat{x},\hat{v})$ respect.] se cumplen tambien las condiciones linealizadas correspondientes, si todas las funciones que intervienen en **P** son diferenciables en $\hat{\mathbf{x}}$.

DEMOSTRACION

(para el caso de las condiciones de K-T)

Supongamos que existe un $\mathbf{x} \in C$ tal que:

$$\nabla \Phi(\hat{x}, \hat{u})^T(\bar{x} - \hat{x}) = \alpha < 0$$

Como $\psi(x)$ $\Phi(x, \hat{u})$ es diferenciable en $\hat{x}$, se cumple para

$X(\lambda) = \hat{x} + \lambda(\bar{x} - \hat{x})$, $X(\lambda) \in C$; $\forall \lambda \in$ $0, 1]$

$$\Phi(X(\lambda), \hat{u}) = \Phi(\hat{x}, \hat{u}) + \nabla_x \Phi(\hat{x}, \hat{u})^T (X(\lambda) \qquad + o(\lambda)$$

$$= \Phi(\hat{x}, \hat{u}) + \lambda [\nabla_x \Phi(\hat{x}, \hat{u})^T (\bar{x} \qquad + \frac{o(\lambda)}{\lambda}], \quad \forall \lambda \in \langle 0, 1]$$

$\exists \lambda > 0$ tal que

$$\left|\frac{o(\lambda)}{\lambda}\right| < -\frac{\alpha}{2}, \quad \forall \lambda \in \langle 0, \lambda_o]$$

Entonces $\Phi(X(\lambda_o), \hat{u})$ $\Phi(\hat{x}, \hat{u}) + \lambda_o \left[\nabla_x \Phi(\hat{x}, \hat{u})^T (\bar{x} - \hat{x}) + \frac{o(\lambda_o)}{\lambda_o}\right]$

$$< \Phi(\hat{x}, \hat{u}) + \lambda_o \frac{\alpha}{2} \quad \Phi(\hat{x}, \hat{u}).$$

lo cual es una contradicción. ■

4. Si en el punto $(\hat{x}, \hat{u})$ $[(\hat{x}, \hat{v})]$ se cumplen las condiciones linealizadas de K-T [de JOHN] entonces en $(\hat{x}, \hat{u})$ $[(\hat{x}, \hat{v})]$ se cumplen también las condiciones no-linealizadas correspondientes si F y f. $j = \overline{1, m}$ son convexas y g_k; $k = \overline{1, \ell}$ son afín lineales.

DEMOSTRACION

Como $\psi(x) = \Phi(x, \hat{u})$ es convexa y $\nabla \Phi_x(\hat{x}, \hat{u})^T (x - \hat{x}) \geq 0$,

$\forall \mathbf{x} \in \mathbf{C}$ se cumple:

$$\Phi(x, \hat{u}) \geq \Phi(\hat{x}, \hat{u}) + \nabla_x \Phi(\hat{x}, \hat{u})^T (x - \hat{x}) \geq \Phi(\hat{x}, \hat{u}), \quad \forall x \in C \quad ■$$

5. Si $\forall\ j \in I(\mathbf{x})$ la función f_j es convexa y diferenciable en $\mathbf{x}$ y si g_k: $k = \overline{1, \ell}$ es afín lineal, entonces las condiciones de SLATER se cumplen exactamente si las condiciones linealizadas de SLATER se cumplen.

DEMOSTRACION

a) Sean cumplidas las condiciones de SLATER para $\mathbf{P}$. Entonces existe un $\mathbf{x}^{\mathbf{o}} \in$ intrel $(\mathbf{C})$ tal que

$$f_j(\mathbf{x}^{\mathbf{o}}) < 0 \qquad \forall\ j \in I(\mathbf{x})$$

$$g(\mathbf{x}^{\mathbf{o}}) = 0$$

Pero para $j \in I(\mathbf{x})$ se cumple:

$$0 > f_j(\mathbf{x}^{o}) \geq f_j(\mathbf{x}) + \nabla f_j(\hat{\mathbf{x}})^{\mathbf{T}}(\mathbf{x}^{o}-\hat{\mathbf{x}}) = \nabla f_j(\hat{\mathbf{x}})^{\mathbf{T}}(\mathbf{x}^{o}-\hat{\mathbf{x}}).$$

Como $g_k(\mathbf{x}) = a_k^{\mathbf{T}}\mathbf{x} + b_k$ y $\hat{\mathbf{x}}$ es admisible para $\mathbf{P}$, se cumple

$$\nabla g_k(\hat{\mathbf{x}})^{\mathbf{T}}(\mathbf{x}^{o}\quad \hat{\mathbf{x}}) = a_k^{\mathbf{T}}(\mathbf{x}^{o} - \hat{\mathbf{x}}) = a_k^{\mathbf{T}}\mathbf{x}^{o} + b_k =$$

$$= g_k(\mathbf{x}^{o}) = 0, \quad \forall\ k = \overline{1, \ell}$$

b) Sea al revés, $\mathbf{x}^{o}$ un punto que satisface las condiciones linealizadas de SLATER para $\mathbf{P}$. Como $\mathbf{x}^{o} \in$ intrel$(\mathbf{C})$, $\mathbf{x} \in \mathbf{C}$, se cumple para $0 < \lambda \leq 1$, $X(\lambda) = \hat{\mathbf{x}} + \lambda(\mathbf{x}^{o}-\hat{\mathbf{x}}) \in$ intrel$(\mathbf{C})$

Para λ suficientemente pequeño se cumple:

$$f_j(\hat{\mathbf{x}}+\lambda(\mathbf{x}^{o}-\hat{\mathbf{x}})) = \lambda\nabla f_j(\hat{\mathbf{x}})^{\mathbf{T}}(\mathbf{x}^{o}-\hat{\mathbf{x}}) + o(\lambda(\mathbf{x}^{o}-\hat{\mathbf{x}})) \quad 0; \quad \forall j \in I(\mathbf{x})$$

$$g_k(\hat{\mathbf{x}}+\lambda(\mathbf{x}^{o}-\hat{\mathbf{x}})) = g_k(\hat{\mathbf{x}}) + \lambda(g_k(\mathbf{x}^{o}) - g_k(\hat{\mathbf{x}})) = \lambda a_k^{\mathbf{T}}(\mathbf{x}^{o}-\hat{\mathbf{x}})$$

$$= \lambda\nabla g_k(\hat{\mathbf{x}})^{\mathbf{T}}(\mathbf{x}^{o}-\hat{\mathbf{x}}) = 0, \qquad k = \overline{1, \ell}$$

Por consiguiente el punto $\hat{\mathbf{x}} + \lambda(\mathbf{x}^{o}-\hat{\mathbf{x}})$ satisface las condiciones no-linealizadas de SLATER para $\mathbf{P}$, si $0 \quad \lambda \leq 1$ es suficientemente pequeño.

NOTA: Se usó $x^o \in \text{intrel}(C)$, $x \in CL(C) \Longrightarrow [x^o, \hat{x}) \subset \text{intrel}(C)$

6. De las observaciones **3.**, **4.** y **5.** se deduce lo siguiente: Si **P** es un programa convexo y todas las funciones que intervienen en **P** son diferenciables en **C**, entonces las condiciones de K-T, JOHN Y SLATER, son equivalentes con las condiciones linealizadas correspondientes. ■

Vamos a tratar ahora las condiciones de optimidad para el programa **P** que involucra la función (generalizada) de LAGRANGE.

TEOREMA 1.1.- Si en el punto $(\hat{x}, \hat{u})$ se cumplen las condiciones de K-T para **P**, donde **C** no es necesariamente convexo, entonces $\hat{x}$ es una solución óptima para **P**.

DEMOSTRACION

$f(x) \leq 0$, $g(\hat{x}) = 0$ $x \in C$, significa que x es admisible para **P**. Como $\hat{u}^1 \geq 0$ y $F(\hat{x}) = \Phi(\hat{x}, \hat{u}) \leq \Phi(x, \hat{u})$ $\forall x \in C$ se cumple para todo punto admisible z de **P**

$$F(\hat{x}) \leq F(z) + (\hat{u}^1)^T f(z) + (\hat{u}^2)^T g(z) \leq F(z)$$

Por consiguiente x es una solución óptima de P. ■

En contraste con las condiciones de K-T, las condiciones de JOHN no representan un criterio suficiente de optimidad para el programa **P**, sin embargo se cumple el

TEOREMA 1.2.- Si en punto $(\hat{x}, \hat{v})$ se cumplen las condiciones de JOHN para el programa

P: $\min\{F(\mathbf{x})/f(\mathbf{x}) \leq 0,\ \mathbf{x} \in \mathbf{C}\}$, entonces se cumple 1) o 2)

1). $\hat{\mathbf{x}}$ es una solución óptima de $\bar{\mathbf{P}}$

2). Las condiciones de SLATER para **P** no se cumplen.

DEMOSTRACION

Si 0, entonces en $(\hat{\mathbf{x}}, \dfrac{\hat{v}^1}{\hat{v}^o})$ se cumplen las condiciones de K-T para **P** y por teorema 1.1, $\mathbf{x}$ es una solución óptima de **P**.

Si $\hat{v}^o = 0$, entonces se cumple:

$$0 = (\hat{v}^o)F(\hat{\mathbf{x}}) = L(\hat{\mathbf{x}}, \hat{\mathbf{v}}) \leq L(\mathbf{x}, \hat{\mathbf{v}}) = (\hat{v}^o)F(\mathbf{x}) + (\hat{v}^1)f(\mathbf{x})$$

$$= (\hat{v}^1)^T f(\mathbf{x})$$

$$0 \leq (\hat{v}^1)^T f(\mathbf{x}), \quad \forall\ \mathbf{x} \in C$$

$$\hat{v}^1 \neq 0 \qquad (\hat{v}^1)^T f(\hat{\mathbf{x}}) = 0$$

De esto se deduce que:

$$_j \geq 0 \qquad j = 1, m$$

$$= \qquad \forall\ j \notin I(\mathbf{x})$$

$$\hat{v}_j \quad 0 \qquad \text{para al menos un } j \in I(\mathbf{x})$$

Supongamos que existe un $\mathbf{x}^o \in \text{intrel}(\mathbf{C})$ tal que

$$f_j(\mathbf{x}^o) < 0 \qquad \forall\ j \in I(\mathbf{x})$$

Entonces de (**) se deduce que

$$(\hat{v}^1)^T f(\mathbf{x}^o) = \sum_{j \in I(\hat{\mathbf{x}})} f_j(\mathbf{x}^o)\hat{v}_j < 0, \text{ lo que contradice a}$$

Por consiguiente las condiciones de SLATER para el programa **P** no son satisfechas. ■

Si **P** es un programa convexo, entonces las condiciones de JOHN son condiciones necesarias de optimidad para **P**.

TEOREMA 1.3.- Si **P** es un programa convexo y $\hat{\mathbf{x}}$ es una solución óptima de **P**, entonces existe un $\hat{\mathbf{v}} = (\hat{v}^0, \hat{v}^1, \hat{v}^2)$ tal que en $(\hat{\mathbf{x}}, \hat{\mathbf{v}})$ se cumplen las condiciones de JOHN para **P**.

DEMOSTRACION

Definimos los conjuntos:

$A = \{\mathbf{y} = (y^0, y^1, y^2) \ / \ y^0 < F(\hat{x}),\ y^1 \leq 0,\ y^2 = 0\}$

$B = \{\mathbf{y} = (y^0, y^1, y^2) \ / \exists\ x \in C \text{ con } y^0 \geq F(x),\ y^1 \geq f(x),\ y^2 = g(x)\}$

$D = \{ z \ / \ z = (F(x), f(x), g(x)),\ \ x \in C\}.$

Se cumple $D \subset B$. Es evidente que A es convexo. Como F, f y C son convexos y g es afín lineal, B es convexo:

Sean $\mathbf{y}$, $\mathbf{y} \in B$, entonces existen $\mathbf{x}$ y $\mathbf{x}$ en **C** tales que

$$F(\mathbf{x}) \leq y^0,\ \ f(\mathbf{x}) \leq y^1,\ \ g(\mathbf{x}) = y$$

$$F(\mathbf{x}) \leq \bar{y}^0,\ \ f(\mathbf{x}) \leq \bar{y}^1,\ \ g(\mathbf{x}) = \bar{y}^2$$

De esto se deducee que para $\lambda \in [0, 1]$ se cumple:

$$\lambda \mathbf{x} + (1-\lambda)\mathbf{x} \in \mathbf{C}$$

$$F(\lambda \mathbf{x} + (1-\lambda)\mathbf{x}) \leq \lambda F(\mathbf{x}) + (1-\lambda)F(\mathbf{x}) \leq \lambda y^0 + (1-\lambda)\bar{y}^0$$

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{x}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{x}) \leq \lambda y^1 + (1-\lambda)\bar{y}^1$$

$$g(\lambda \mathbf{x} + (1-\lambda)\mathbf{x}) = \lambda g(\mathbf{x}) + (1-\lambda)g(\mathbf{x}) = \lambda y^2 + (1-\lambda)\bar{y}^2$$

Por consiguiente $\lambda \mathbf{y} + (1-\lambda)\mathbf{y} \in B$ y B es convexo.

Se cumple $A \cap B = \phi$

Supongamos que existe un $\mathbf{y} \in B \cap A$. Como $\mathbf{y} \in B$

$$\left.\begin{array}{l} \text{Entonces } \exists \text{ un } \mathbf{x} \in \mathbf{C} \text{ con } \mathbf{y}^0 \geq F(\mathbf{x}),\ y^1 \geq f(\mathbf{x}),\ y^2 = g(\mathbf{x}); \\ \mathbf{y} \in A:\ y^0 < F(\hat{\mathbf{x}}),\ y^1 \leq 0,\ y^2 = 0 \end{array}\right\}$$

Entonces de (*) se deduce que

$f(x) \leq y^1 \leq 0$, $g(x) = y^2 = 0$, $y^0 \geq F(\mathbf{x})$

Por consiguiente $\mathbf{x}$ es admisible para **P**. Como $\mathbf{x}$ es una solución óptima de **P**, se cumple :

$y^o \geq F(\mathbf{x}) \geq F(\hat{\mathbf{x}})$ (contradiciendo que $y^o < F(\hat{\mathbf{x}})$)

Por consiguiente $A \cap B = \phi$.

Como A y B son convexos y $A \cap B = \phi$, por el teorema de separación existe un $\hat{\mathbf{v}} = (\hat{v}^o, \hat{v}^1, \hat{v}^2) \neq 0$ y un $\alpha \in \mathbb{R}$ tal que

$\hat{\mathbf{v}}^T z \geq \alpha \geq \hat{\mathbf{v}}^T \mathbf{y}$, $\forall\, z \in B \supset D$, $\forall\, \mathbf{y} \in A$

Como $\hat{\mathbf{y}} = (F(\hat{\mathbf{x}}), f(\hat{\mathbf{x}}), g(\hat{\mathbf{x}})) \in A \cap D$, se cumple :

$$\alpha = \hat{\mathbf{v}}^T\hat{\mathbf{y}}, \quad \hat{\mathbf{v}}^T\mathbf{z} \geq \hat{\mathbf{v}}^T\hat{\mathbf{y}} \geq \hat{\mathbf{v}}^T\mathbf{y}, \quad \forall\, \mathbf{y} \in A, \quad \forall\, z \in D \qquad (+)$$

Escribimos la primera inecuación de (+) en detalle

$$\hat{\mathbf{v}}^T\hat{\mathbf{y}} = L(\hat{\mathbf{x}}, \hat{\mathbf{v}}) = \hat{v}^o F(\hat{\mathbf{x}}) + (\hat{v}^1)^T f(\hat{\mathbf{x}}) + (\hat{v}^2)^T g(\hat{\mathbf{x}}) \leq$$

$$\leq \hat{v}^o F(\mathbf{x}) + (\hat{v}^1)^T f(\mathbf{x}) + (\hat{v}^2)^T g(\mathbf{x}) = L(\mathbf{x}, \hat{\mathbf{v}}), \quad \forall \mathbf{x} \in \mathbf{C}$$

es decir $L(\hat{\mathbf{x}}, \hat{\mathbf{v}}) \leq L(\mathbf{x}, \mathbf{v})$, $\forall\, \mathbf{x} \in \mathbf{C}$.

Elegimos ahora los puntos $\mathbf{y}$, $\mathbf{y}^j$, $\bar{\mathbf{y}}^j$; $j = 1, m$ como sigue:

$$- \ (F(\hat{\mathbf{x}}) - 1, \ f(\hat{\mathbf{x}}), \ g(\mathbf{x}))$$

$$y^j \qquad \bar{y}^j_o = F(\mathbf{x})$$

$$y^j_i = \hat{y}_i \ , \quad \forall\, i \neq j \quad y \quad y^j_j = y_j -$$

$$\bar{y}^j_i = \hat{y}_i \qquad \forall\, i \neq j \quad y \quad \bar{y}^j_i = 0$$

Estos puntos están todos en A.

De (+) se deduce:

$$\hat{\mathbf{v}}^T(\bar{\mathbf{y}} - \hat{\mathbf{y}}) = - \hat{v}^o \leq 0$$

$$\hat{\mathbf{v}}^T(\mathbf{y}^j \quad \hat{\mathbf{y}}) = \quad \hat{v}^1_i \leq 0$$

$$\hat{\mathbf{v}}^T(\bar{\mathbf{y}}^j \quad \hat{\mathbf{y}}) = - \hat{v}^1_j \hat{y}_j = - \hat{v}^1_j f_j(\hat{\mathbf{x}}) \leq$$

Como $\hat{x}$ es admisible para **P**, de esto se obtiene:

$$\hat{v} \ \geq 0, \quad \hat{v}^1 \geq 0 \ , \quad (\hat{v}^1)^T f(\hat{\mathbf{x}}) =$$

Por consiguiente en el punto $(\hat{\mathbf{x}}, \mathbf{v})$ se cumplen las condiciones de F. JOHN para **P**. ■

Las condiciones de K-T generalmente no son condiciones necesarias de optimidad para un programa convexo de la forma **P**. Sin embargo se cumple el

TEOREMA 1.4.- Sea $\hat{\mathbf{x}}$ una solución óptima del programa convexo

$$\mathbf{P}:\ \text{mín}\ \{F(\mathbf{x})\ /\ \mathbf{x} \in \mathbf{C},\ f(\mathbf{x}) \leq 0\ ,\ g(\mathbf{x})$$

Sean cumplidas las condiciones de SLATER para **P**. Entonces existe un $\mathbf{u} = (\hat{u}^1, \hat{u}^2)$ tal que en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumplen las condiciones de K-T para **P**.

DEMOSTRACION

Demostraremos el teorema solo para el programa:

$$\mathbf{P} \quad \min\ \{F(\mathbf{x})\ /\ f(\mathbf{x}) \leq 0\ ,\ \mathbf{x} \in \mathbf{C}\}$$

Por el teorema 1.3 se cumplen las condiciones de JOHN para **P** en un punto $(\hat{\mathbf{x}}, \hat{\mathbf{v}})$ $\quad \hat{\mathbf{v}} = (\hat{v}^o, \hat{v}^1)$.

Si fuera $\hat{v}^o = 0$, entonces por el teorema 1.2, no existiría ningun $\mathbf{x}^o \in \text{intrel}(\mathbf{C})$ tal que $f_j(x^o) < 0$, $\forall\ j \in I(\mathbf{x})$ y las condiciones de SLATER para **P** no se cumplirían.

Por consiguiente $\hat{v}^o > 0$ y en el punto $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ $\quad \left(\hat{\mathbf{x}}, \frac{\hat{v}^1}{\hat{v}^o}\right)$ se cumplen las condiciones de K-T para **P**. ■

<u>EJEMPLO:</u>

Indicamos un programa convexo simple **P** una única solución óptima $\hat{\mathbf{x}}$, tal que no existe $\hat{\mathbf{u}}$, de modo que en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumplen las condiciones de K-T.

Consideremos

$$\min\{x_2\ /\ x_1^2 + x_2^2 - 1 \leq 0 \quad (x_1 - 2)^2 + x_2^2 - 1 \leq 0\}$$

$$= \min\{F(\mathbf{x})\ /\ f_1(\mathbf{x}) \leq 0 \quad f_2(\mathbf{x}) \leq 0\}$$

El único punto admisible y con esto la única solución óptima es $\mathbf{x} = (1,0)$.

Veamos que las condiciones linealizadas de K-T no pueden ser cumplidas.

En efecto

De $\nabla F(\hat{\mathbf{x}}) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\quad \nabla f_1(\hat{\mathbf{x}}) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\quad \nabla f_2(\mathbf{x}) = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$

deducimos que no existe $\mathbf{u} \in \mathbb{R}^2$, tal que

$$\nabla F(\hat{\mathbf{x}}) = u_1 \nabla f_1(\hat{\mathbf{x}}) + u_2 \nabla f_2(\hat{\mathbf{x}}) \qquad \blacksquare$$

Para el resto de este capítulo asumiremos que las funciones F, f_i, g_k sean diferenciables en $\mathbf{C}$ o al menos en un punto $\mathbf{x} \in \mathbf{C}$.

Si **P** es un programa convexo, entonces las condiciones linealizadas de K-T representan un criterio suficiente de optimidad para **P**.

TEOREMA 1.5.- Sea **P** un programa convexo y sean cumplidas en el punto $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ las condiciones linealizadas de K-T para **P**. Entonces $\hat{\mathbf{x}}$ es una solución óptima de **P**.

DEMOSTRACION

Sea $\mathbf{x}$ admisible para **P**. Entonces como $\hat{u}^1 \geq 0$, se cumple

$$\Phi(\mathbf{x}, \hat{\mathbf{u}}) = F(\mathbf{x}) + (\hat{u}^1)^T f(\mathbf{x}) + (\hat{u}^2)^T g(\mathbf{x})$$

$$= F(\mathbf{x}) + (\hat{u}^1)^T f(\mathbf{x}) \leq F(\mathbf{x}) \qquad (*)$$

$= \Phi(\mathbf{y}, \mathbf{u})$ es convexa en $\mathbf{C}$ y $\nabla_x \Phi(\hat{\mathbf{x}}, \hat{\mathbf{u}})^T (\mathbf{y} \quad \mathbf{x}) \geqslant 0, \forall \mathbf{y} \in \mathbf{C}$,

se obtiene de $(*)$ para todo punto $\mathbf{x}$ admisible para **P**.

$$F(\mathbf{x}) \geq \Phi(\mathbf{x}, \hat{\mathbf{u}}) \geq \Phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + \nabla_x \Phi(\hat{\mathbf{x}}, \hat{\mathbf{u}})^T (\mathbf{x} - \hat{\mathbf{x}}) \geq \Phi(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = F(\hat{\mathbf{x}})$$

Por consiguiente $\hat{\mathbf{x}}$ es una solución óptima de **P**. $\blacksquare$

COROLARIO 1.6. - Sea

$$\mathbf{P} : \min \{F(\mathbf{x}) \ / \ f(\mathbf{x}) \leq 0, \ g(\mathbf{x}) = 0, \ \mathbf{x} \in C\}$$

un programa convexo que satisface las condiciones de SLATER. Sean todas las funciones que intervienen en **P**, diferenciables en el punto $\hat{\mathbf{x}}$.

Entonces se cumple

$\hat{\mathbf{x}}$ es solución óptima de **P** si y sólo si existe $\hat{\mathbf{u}} \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$ tal que en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se satisfacen las condiciones y las condiciones linealizadas de K-T para **P**.

DEMOSTRACION

Sabemos que en nuestro caso:

En $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumplen las condiciones linealizadas de K-T para **P** si y sólo si, en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumplen las condiciones no linealizadas de K-T para **P**.

Si en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumplen las condiciones (linealizadas) de K-T, entonces de acuerdo al teorema 1.5, $\hat{\mathbf{x}}$ es una solución óptima de **P**.

Si $\hat{\mathbf{x}}$ es una solución óptima de **P**, entonces de acuerdo al teorema 1.4, existe un $\hat{\mathbf{u}} \in \mathbb{R}^m \times \mathbb{R}^\ell$ tal que en $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ se cumple las condiciones (linealizadas) de K-T de **P**. ∎

Si en el programa **P** las aplicaciones f, g son afín lineales y F es convexa y diferenciable, entonces las condicions de K-T representan un criterio necesario de optimidad, aun en el caso donde las condiciones de SLATER para **P** no se cumplen

TEOREMA 1.7.- Sea $\mathbf{x}$ una solución óptima de **P**. Sean todas las funciones que intervienen en **P** diferenciables en $\mathbf{x}$. Sean cumplidas las condicones de SLATER para **P**.

a) Si g es afín lineal, entonces existe un u tal que en $(\hat{x}, \hat{u})$ se cumplan las condiciones linealizadas de K-T para **P**.

b) Si $\mathbf{x} \in \text{int}(\mathbf{C})$, g es continua en x y los gradientes $\nabla g_k(\hat{x})$ son linealmente independientes, entonces existe un $\hat{\mathbf{u}} \in \mathbb{R}^n \times \mathbb{R}^\ell$ tal que $(\hat{x}, \hat{\mathbf{u}})$ se cumplen las condiciones linealizadas de K-T para **P**.

<u>OBSERVACION 1.3</u>:

Si aplicamos el teorema 1.7 al programa

$$\mathbf{P} \qquad \min \{F(\mathbf{x}) \ / \ g(\mathbf{x}) = 0\}$$

entonces obtenemos la regla clásica de LAGRANGE.

CAPITULO II

TEORIA DE DUALIDAD

2.1: INTRODUCCION

La teoria de dualidad de la programación matemática se ocupa de pares de programas.

Por una prescripción explícita se asocia a un problema de minimización un problema de maximización o al revés.

Asociamos siempre a un problema de minimización **P**, un problema de maximización **P** . **P** se llamará el programa primal y **P** el programa dual.

En la mayoría de los casos se puede asociar al programa dual $\mathbf{P}^*$ otra vez un programa dual $\mathbf{P}^{**}$.

Esto se hace de la siguiente manera:

Se escribe $\mathbf{P}^*$ como problema de minimización al cual, de acuerdo con la prescripción original se asocia un problema de maximización que finalmente se formulará otra vez problema de minimización.

Al par de programas duales (P, P^*) se llama simétrico, si es posible formar P^{**} y $P^{**} = P$.

Consideramos exclusivamente el caso de pares simétricos (P, P^*).

Recordemos los resultados principales de la teoría de dualidad de la programación lineal.

Si (P, P^*) es un par de programas lineales duales, entonces (P, P^*) es simétrico y se cumplen las condiciones A) y B)

A) Si uno de los programas P, P^* tiene una solución óptima, entonces el otro también tiene solución óptima y

$$\inf(P) = \sup(P^*)$$

B) Una condición necesaria y suficiente para que los dos programas P y P^* tengan una solución consiste en que ambos programas tengan puntos admisibles.

El objetivo principal de cada teoría de dualidad es : Demostrar los enunciados A) y B) para pares de programas (P, P^*) mas generales. Pero esto no siempre se logra y debemos contentarnos con propiedades mucho mas débiles que A) y B).

En la teoría de dualidad el concepto de punto silla de una función desempeña un papel esencial. Por esto lo definimos aquí.

DEFINICION 2.1.- Sean $\mathbf{x} \in C \subset \mathbb{R}^n$, $\mathbf{y} \in D \subset \mathbb{R}^m$ y ψ: $C\times D \longrightarrow \mathbb{R}$. Entonces $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \in C\times D$ se llama un punto silla de ψ sobre $C\times D$ si:

$$\psi(\hat{x}, y) \leq \psi(\hat{x}, \hat{y}) \leq \psi(x, \hat{y}), \quad \forall\, x \in C, \quad \forall\, y \in D \tag{2.1}$$

En lo siguiente presentaremos la teoría de dualidad de *DANTZIG, EISENBERG Y COTTLE*, que incluye la teoría de dualidad de la *PROGRAMACION CUADRATICA* como caso especial.

2.2: LA TEORIA DE DUALIDAD DE DANTZIG EISENBERG Y COTTLE

Sea $\psi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$, una función diferenciable.

Para todo $\mathbf{y} \in \mathbb{R}^m$ fijo. Sea $\psi(.,\mathbf{y})$ convexa y para todo $\mathbf{x} \in \mathbb{R}^n$ fijo, sea $\psi(\mathbf{x},.)$ cóncava en $\mathbb{R}^m$. Decimos tambien que ψ es convexa-cóncava en $\mathbb{R}^n \times \mathbb{R}^m$.

Si $\mathbf{x} = (x^1, x^2)$ $\mathbf{y} = (y^1, y^2)$ son particiones arbitrarias de los vectores $\mathbf{x}$, $\mathbf{y}$; consideremos el siguiente par $(\mathbf{Q}, \mathbf{Q}^*)$ de programas

$$\mathbf{Q} : \min\{\, F(\mathbf{x},\mathbf{y}) = \psi - \mathbf{y}^T \psi_{\mathbf{y}} \;/\; \psi_{1} \leq 0 \;,\; \psi_{y^2} \quad 0 \qquad ^1 \geq 0 \,\}$$

$$\mathbf{Q}^* \quad \max\{\, g(\mathbf{x},\mathbf{y}) = \psi - \mathbf{x}^T \psi_{x} \;/\; \psi_{x^1} \geq 0 \;,\; \psi_{x^2} = 0 \qquad y^1 \geq 0 \,\}$$

Para formar $\mathbf{Q}^*$ de $\mathbf{Q}$ usaremos las siguientes reglas:

- A las restricciones en forma de inecuación $\psi_{y^1} \leq 0$ en $\mathbf{Q}$ corresponden en $\mathbf{Q}^*$ las restricciones de signo $y^1 \geq 0$.

  A las restricciones en forma de ecuación $\psi_{2} = 0$ en $\mathbf{Q}$, en $\mathbf{Q}^*$ son asociadas las variables libres y^2.

  A las restricciones de signo $x^1 \geq 0$ en $\mathbf{Q}$, en $\mathbf{Q}$ corresponden las restricciones $\psi_{x^1} \geq 0$ en forma de inecuación.

- A las variables libres x^2 de Q, en Q^* estan asociadas las ecuaciones $\psi_{x^2} = 0$.

Observamos que las funciones objetivos y g se forman usando las restricciones de Q y Q^* respectivamente y observando la misma regla. Es evidente que F y g se dejan escribir tambien como:

$$F(\mathbf{x},\mathbf{y}) = \psi - (y^1)^T \psi_{y^1}$$

$$g(\mathbf{x},\mathbf{y}) = \psi - (x^1)^T \psi_{x^1}$$

Usando las mismas reglas se puede formar Es fácil ver que $Q^{**} = Q$.

El par (Q, Q^*) es simétrico.

En lugar de (Q, Q^*) consideramos primero el siguiente par (P, P^*).

$$P \;:\; \min\{F(\mathbf{x},\mathbf{y}) - \psi - \mathbf{y}^T\psi \;/\; \psi_{y^1} \le 0,\; \psi_{y^2} = 0 \quad x^1 \ge 0,\; y^1 \ge 0\}$$

$$P^*:\; \max\{g(\mathbf{x},\mathbf{y}) = \psi - \mathbf{x}^T\psi_x \;/\; \psi_{x^1} \ge 0,\; \psi_{x^2} - 0 \quad x^1 \ge 0,\; y^1 \ge 0\}$$

Consideramos los conjuntos:

$$C = \{\mathbf{x} \;/\; x^1 \ge 0\} \quad \text{y} \quad D = \{\mathbf{y} \;/\; y^1 \ge 0\}$$

PROPOSICION 2.1.- Si $(\mathbf{x},\mathbf{y})$ es admisible para P, $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ es admible para P^*, entonces:

$$F(\mathbf{x},\mathbf{y}) \ge \psi(\mathbf{x},\mathbf{y}) \qquad \forall\, \mathbf{y} \in D$$

$$g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \le \psi(\mathbf{x},\tilde{\mathbf{y}}) \qquad \forall\, \mathbf{x} \in C$$

En particular se cumple : $F(\bar{\mathbf{x}},\bar{\mathbf{y}}) \ge \psi(\bar{\mathbf{x}},\tilde{\mathbf{y}}) \ge g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ (2.2)

DEMOSTRACION

i) Como $\psi(.,\mathbf{y})$ es convexa $\forall\, \mathbf{y} \in \mathbb{R}^m$ fijo y $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ es admisible para P^* se cumple:

$$\psi(x,\tilde{y}) \geq \psi(\tilde{x},\tilde{y}) + \psi_x(\tilde{x},\tilde{y})^T(x \quad \tilde{x})$$

$$\psi(\tilde{x},\tilde{y}) \quad \tilde{x}^T\psi_x(\tilde{x},\tilde{y}) + x^T\psi_x(\tilde{x},\tilde{y}) \qquad (2.3)$$

$$= g(\tilde{x},\tilde{y} \quad (x^1)^T\psi_{x^1}(\tilde{x},\tilde{y}) \geq g(\tilde{x},\tilde{y}) \qquad \forall\ x \in C$$

ii) Como $\psi(x,.$ es cóncava y (x,y) es admisible para P, se tiene:

$$\psi(x,y) \leq \psi(x,y) + \psi_y(\bar{x},\bar{y})^T(y - \bar{y}) =$$

$$= \psi(x,y) \quad \bar{y}^T\psi_y(\bar{x},y) + y^T\psi_y(\bar{x},\bar{y}) \qquad (2.4)$$

$$= F(x,y) + (y^1)^T\psi_{y^1}(\bar{x},\bar{y}) \leq F(\bar{x},\bar{y}) \qquad \forall\ y \in D$$

De (2.3) y (2.4) se obtiene (2.2) eligiendo $x = x$, $y = y$.

PROPOSICION 2.2.- Si (x,y) es admisible para P, $(\tilde{x},\tilde{y})$ es admisible para P y $F(x,y) = g(\tilde{x},\tilde{y})$, entonces $(\bar{x},\tilde{y})$ es un punto silla de ψ sobre $C{\times}D$ y se tiene:

$$F(\bar{x},\bar{y}) = \psi(\bar{x},\tilde{y}) = g(\tilde{x},\tilde{y})$$

$$(\tilde{y}^1)^T\psi_{y^1}(x,\bar{y}) = (x^1)^T\psi_{x^1}(\tilde{x},\tilde{y}) = 0$$

DEMOSTRACION

De (2.2) se obtiene: $F(x,y) = \psi(x,\tilde{y}) = g(\tilde{x},\tilde{y})$

De esto y de (2.3) y (2.4) se deduce que:

$$(\bar{x}^1)^T\psi_{x^1}(\tilde{x},\tilde{y}) = 0 \qquad (\tilde{y}^1)^T\psi_{y^1}(\bar{x},\bar{y}) = 0$$

De acuerdo con la proposición 2.1 se cumple

$$\psi(\bar{x},y) \leq \psi(\bar{x},\tilde{y}) \leq \psi(x,\tilde{y}),\ \forall\ y \in D,\ \ \forall\ x \in C$$

Por consiguiente $(\bar{x},\tilde{y})$ es un punto silla de ψ sobre $C{\times}D$ ∎

OBSERVACIONES 2.1:

1.- Se debe observar que $(\tilde{x},y)$ en general no es punto silla de ψ, si (x,y) es admisible para P, $(\tilde{x},\tilde{y})$ es admisible para P^* y $g(\tilde{x},\tilde{y}) - F(x,y)$.

En consecuencia el comportamiento de las variables x e y no es simétrico.

Se puede observar que las variables x tienen un papel destacado para P, mientras que las variables y tienen la misma importancia destacada para P^*.

Por esta razón llamaremos a x [respectivamente a y] las variables propias de P [respectivamente de P^*].

2.- $(\hat{x},\hat{y})$ es punto silla de ψ sobre $C \times D$ si y sólo si se cumple (2.5) y (2.6).

(2.5) $\psi_{x^1}(\hat{x},\hat{y}) \geq 0 \qquad (\hat{x}^1)^T\psi_{x^1}(\hat{x},\hat{y}) = 0$

$\psi_{x^2}(\hat{x},\hat{y}) = 0 \qquad \hat{x}^1 \geq$

(2.6) $\psi_{y^1}(\hat{x},\hat{y}) \leq 0 \qquad (\hat{y}^1)^T\psi_{y^1}(\hat{x},\hat{y}) =$

$\psi_{y^2}(\hat{x},\hat{y}) = 0 \qquad \hat{y}^1 \geq$

DEMOSTRACION

Usaremos el siguiente resultado.

Sea $E \subset \mathbb{R}^n$, convexo y cerrado

$F \quad E \longrightarrow \mathbb{R}$, convexa y diferenciable

$P : \min\{F(x) \ / \ x \in E\}$

$x \in C$. Se cumple:

$\hat{x}$ es óptimo para P si y sólo si $\nabla F(\hat{x})^T(x-\hat{x}) \geq 0, \ \forall \ x \in E$ ∎

Para la demostración de 2.

Sea $\mathbf{x}^1 = (x_1, x_2, \ldots, x_\ell)$ $\quad \mathbf{x}^2 = (x_{\ell+1}, \ldots, x_n)$

$\mathbf{y}^1 = (y_1, y_2, \ldots, y_k)$ $\quad \mathbf{x}^2 \quad (y_{k+1}, \ldots, y_m)$

Observamos que $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ es punto silla de ψ sobre $\mathbf{C}\times\mathbf{D}$, significa que $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ es una solución óptima para los programas $\mathbf{R}$ y $\mathbf{R}^*$.

$$\mathbf{R} : \min \{\psi(\mathbf{x}, \hat{\mathbf{y}}) \ / \ \mathbf{x} \in \mathbf{C}\}$$

$$\mathbf{R}^* : \max \{\psi(\hat{\mathbf{x}}, \mathbf{y}) \ / \ \mathbf{y} \in \mathbf{D}\}$$

Como $\psi(.,\hat{\mathbf{y}})$ es convexa en $\mathbf{C}$

$\psi(\hat{\mathbf{x}}, .)$ es cóncava en $\mathbf{D}$, se cumple:

$(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ es óptimo para $\mathbf{R}$ y $\mathbf{R}^*$ si y sólo si, se cumple (2.5') y (2.6').

(2.5'): $\quad \hat{x}^1 \geq 0 \ , \ \psi \ (\hat{\mathbf{x}}, \hat{\mathbf{y}})^{\mathbf{T}}(\mathbf{x} \quad \mathbf{x}) \geq 0 \quad \forall \ \mathbf{x} \in$

(2.6'): $\quad \hat{y}^1 \geq 0 \ , \ \psi_y (\hat{\mathbf{x}}, \hat{\mathbf{y}})^{\mathbf{T}}(\mathbf{y} \quad \mathbf{y}) \leq 0 \quad \forall \ \mathbf{y} \in$

[$\mathbf{R}^*$: $\min\{ -\psi(\mathbf{x},\mathbf{y}) \ / \ \mathbf{y} \in \mathbf{D}\}$, $\quad \psi(\mathbf{x}, .)$ es convexa en $\mathbf{D}$].

Basta demostrar que (2.5') [(2.6')] es equivalente con (2.5) [(2.6)].

(2.5) $\Longrightarrow$ (2.5')

Sea $\mathbf{x} \in \mathbf{C}$ cualquiera ó $x^1 \geq 0$

De (2.5) obtenemos:

$$\psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}})^{\mathbf{T}}(\mathbf{x}-\hat{\mathbf{x}}) = (x^1)^{\mathbf{T}}\psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \geq 0,$$

$$\hat{x}^1 \geq 0$$

(2.5' $\quad$ (2.5)

Para $i \leq \ell$, elegimos $\hat{\mathbf{x}} + e_i^{\ell} \in \mathbf{C}$.

De (2.5') obtenemos:

$$0 \leq \psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T(\mathbf{x}^i - \hat{\mathbf{x}}) - \psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T e_j = \psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$$

Entonces

$$\psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \geq 0 \tag{a}$$

[NOTA: $(\hat{x}^1)^T \psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = 0$ significa
$x_i > 0$, $i \leq \ell$ implica $\psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = 0$]

Sea $i \leq \ell$, $\hat{x}_i$ 0, entonces $y^i = \hat{\mathbf{x}}$ $\hat{x}_i e_1 \in C$

De (2.5') y (a) obtenemos:

$$0 \leq \psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T(y^i - \hat{\mathbf{x}}) - \hat{x}_i \psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T e_i - \hat{x}_i \psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \leq 0,$$

pues $\hat{x}_i$ 0 $\wedge$ $\psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ luego $\psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = 0$

Entonces $(\hat{x}^1)^T \psi_{x^1}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = 0$

Si i ℓ, entonces $x^i = \hat{\mathbf{x}} + e.$ C

$$i = \hat{\mathbf{x}} \quad e_i \in C$$

De (2.5') obtenemos:

$$0 \leq \psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T(x^i - \hat{\mathbf{x}}) = \psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$$

$$0 \leq \psi_x(\hat{\mathbf{x}}, \hat{\mathbf{y}})^T(y^i - \hat{\mathbf{x}}) - \psi_{x_i}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$$

Entonces $\psi_{x_i}(\hat{\mathbf{x}}, \mathbf{y}) - 0$

Es decir $\psi_{x^2}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = 0.$

Similarmente se prueba que (2.6) es equivalente a (2.6').

PROPOSICION 2.3.- Si $(\hat{x},\hat{y})$ es un punto silla de ψ sobre $C \times D$. Entonces $(\hat{x},\hat{y})$ es óptimo para P y P^*, y $F(\hat{x},\hat{y}) = g(\hat{x},\hat{y})$.

DEMOSTRACION

Como $(\hat{x},\hat{y})$ es un punto silla de ψ sobre $C \times D$ se cumple (2.5) y (2.6). Por consiguiente $(\hat{x},\hat{y})$ es admisible para P y P^* y se cumple :

$$g(\hat{x},\hat{y}) = \psi(\hat{x},\hat{y}) = F(\hat{x},\hat{y})$$

Sean (x,y) [respectivamente (u,v)] puntos admisibles arbitrarios de P [respectivamente de P^*]

i) Como (x,y) es admisible para P, $(\hat{x},\hat{y})$ es admisble para P^* de la proposición 2.1 se obtiene:

$$F(x,y) \geq \psi(x,\hat{y}) \geq g(\hat{x},\hat{y}) = F(\hat{x},\hat{y})$$

Por consiguiente $(\hat{x},\hat{y})$ es óptimo para P.

ii) Como (u,v) es admisble para P^*, $(\hat{x},\hat{y})$ es admisible para P, de la proposición 2.1 se deduce:

$$g(u,v) \leq \psi(\hat{x},v) \leq F(\hat{x},\hat{y}) = g(\hat{x},\hat{y})$$

En consecuencia $(\hat{x},\hat{y})$ es óptimo para P^*. ∎

Las proposiciones 2.1, 2.2 y 2.3 resumimos como el

TEOREMA 2.1.- Si (x,y) es admisible para P, $(\tilde{x},\tilde{y})$ es admisible para P^*, entonces $F(\bar{x},\bar{y}) \geq g(\tilde{x},\tilde{y})$.

Si ademas $F(x,y) = g(\tilde{x},\tilde{y})$, entonces $(\tilde{x},\tilde{y})$ es un punto silla de ψ sobre $C \times D$ y $(\tilde{x},\tilde{y})$ es óptimo para P y P^*

OBSERVACIONES 2.2:

1.- Si $\mathbf{C} \subset \tilde{\mathbf{C}} \subset \mathbb{R}^n$ y $\mathbf{D} \subset \tilde{\mathbf{D}} \subset \mathbb{R}^m$, entonces las proposiciones 2.1, 2.2 y 2.3 y el teorema 2.1 siguen válidas en exactamente la misma forma, si en $\mathbf{P}$ se sustituye las restricciones de signo $(\mathbf{x},\mathbf{y}) \in \mathbf{C}{\times}\mathbf{D}$ por $(\mathbf{x},\mathbf{y}) \in \mathbf{C}{\times}\mathbf{D}$ y en $\mathbf{P}^*$ por $(\mathbf{x},\mathbf{y}) \in \mathbf{C} \times \mathbf{D}$.

En consecuencia el teorema 2.1 tambien es válido para el par $(\mathbf{Q},\mathbf{Q}^*)$.

2. Si $(\hat{\mathbf{x}},\hat{\mathbf{y}})$ es un punto silla de ψ sobre $\mathbf{C} \times \mathbf{D}$, entonces en $(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{u}}) = (\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ se cumplen las condiciones linealizadas de K-T para $\mathbf{P}$ y $\mathbf{Q}$ si ψ es dos veces diferenciable.

DEMOSTRACION

Consideremos

$$\mathbf{Q} \quad \min\{\psi \quad \mathbf{y}^T\psi_y \ / \ \psi_{y^1} \leq 0, \ \psi_{y^2} = 0, \ x^1 \geq 0\}$$

La función de Lagrange de $\mathbf{Q}$ es :

$$\Phi(\mathbf{x},\mathbf{y},\mathbf{u}) = \psi(\mathbf{x},\mathbf{y}) + (\mathbf{u} - \mathbf{y}) \ \psi_y(\mathbf{x},\mathbf{y})$$

se cumple:

$$\Phi_x(\mathbf{x},\mathbf{y},\mathbf{u}) = \psi_x(\mathbf{x},\mathbf{y}) + \psi_{xy}(\mathbf{x},\mathbf{y})(\mathbf{u} - \mathbf{y})$$

$$\Phi_y(\mathbf{x},\mathbf{y},\mathbf{u}) = \psi_{yy}(\mathbf{x},\mathbf{y})(\mathbf{u} - \mathbf{y})$$

Las condiciones linealizadas de K-T para $\mathbf{Q}$ en $(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{u}})$ se dejan representar como sigue:

$$\Phi_{x^1}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{u}}) \geq \qquad (\hat{x}^1)^T\Phi_{x^1}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{u}}) = 0$$

$$\Phi \ {}_2(\hat{\mathbf{x}},\hat{\mathbf{y}},\mathbf{u}) =$$

$\Phi_y(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{u}}) = 0$

$(\hat{\mathbf{x}},\hat{\mathbf{y}})$ es admisible para Q

$\hat{u}^1 \geq 0$

De las condiciones de punto silla (2.5) y (2.6) se deduce:

$\Phi_{x^1}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}}) = \psi_x(\hat{\mathbf{x}},\hat{\mathbf{y}}) \geq 0$

$(\hat{x}^1)^T\Phi_{x^1}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}}) = (\hat{x}^1)^T\psi_{x^1}(\hat{\mathbf{x}},\hat{\mathbf{y}}) = 0$

$\Phi_{x^2}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}}) = \psi_{x^2}(\hat{\mathbf{x}},\hat{\mathbf{y}}) = 0$

$\Phi_x(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}}) = 0$

$(\hat{\mathbf{x}},\hat{\mathbf{y}})$ es admisible para Q

$\hat{y}^1 \geq 0$

Por consiguiente en $(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{y}})$ se cumplen las condiciones linealizadas de K-T para **Q.**

EJEMPLOS:

1.- Si $\mathbf{x} = x^2$ y $\psi(\mathbf{x},\mathbf{y}) = F(\mathbf{x}) + (y^1)^T f(\mathbf{x}) + (y^2)^T g(\mathbf{x})$

Se obtiene el par de programas

$$Q: \quad \min\{F(\mathbf{x}) \;/\; f(\mathbf{x}) \leq 0 \;, \quad g(\mathbf{x}) = 0\}$$

$$Q^*: \quad \max\{\psi(\mathbf{x},\mathbf{y}) \;/\; \psi_x(\mathbf{x},\mathbf{y}) = 0 \quad y^1 \geq 0\}$$

En este caso $\psi(\mathbf{x},\mathbf{y})$ es la función de Lagrange de **Q.**

Q^* se llama el programa dual de *WOLFE* de **Q** [16]

Sea $\psi(\mathbf{x},\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} - \mathbf{y}^T\mathbf{A}\mathbf{x} + \frac{1}{2}(\mathbf{x}^T\mathbf{B}\mathbf{x} - \mathbf{y}^T\mathbf{C}\mathbf{y})$

donde $\mathbf{B} = \mathbf{B}^T$ $\quad \mathbf{C} = \mathbf{C}^T$ son semi-definidas positivas.

Es evidente que ψ es convexa - cóncava en $\mathbb{R}^n \times \mathbb{R}^m$.

Si elegimos $\mathbf{x} = \mathbf{x}^1$ $\quad \mathbf{y} = \mathbf{y}^1$ obtenemos para $(\mathbf{P},\mathbf{P}^*)$ un par de programas cuadráticos, como fueron analizados por COTTLE [13]

$$P : \min\{ \frac{1}{2}(x^T B x + y^T C y) + c^T x \; / \; Ax + Cy \geq -b, \; x \geq 0, \; y \geq 0\}$$

$$P^*: \max\{ -\frac{1}{2}(x^T B x + y^T C y) - b^T y / Bx - A^T y \geq -c, \; x \geq 0, \; y \geq 0\}$$

En este caso ψ se llama "*FUNCION DE COTTLE*"

Si $C = 0$, se obtiene el par de programas cuadráticos duales que fueron estudiados por DORN [15]. En este caso ψ es la función de Lagrange de **P**.

Si además en **P** y P^* se omite las restricciones de signo de las variables no propias, obtenemos el par:

$$Q \quad \min\{ \frac{1}{2} x^T B x + c^T x \; / \; Ax + b \geq 0 \; , \; x \geq 0\}$$

$$\max\{ -\frac{1}{2} x^T B x - b^T y \; / \; Bx - A^T y + c \geq 0, \; y \geq 0\}$$

Como $B = B^T$ es definida semipositiva, existe una matriz Γ, tal que $B = \Gamma^T \Gamma$. Para $\xi = \Gamma x$, Q^* tiene la siguiente forma

$$Q^* \quad \max\{ -\frac{1}{2}\xi^T \xi - b^T y \; / \; \Gamma^T \xi - A^T y + c \geq 0, \; y \geq 0\}$$

En esta forma el programa Q^* fue analizada por PEARSON y TUCKER [12]

3.-Si se elige $x = x^1$, $y = y^1$, $\psi(x,y) = c^T x - y^T A x + b^T y$, se obtiene un par simétrico de programas lineales duales:

$$Q \quad \min\{ c^T x \; / \; Ax \geq b \; , \; x \geq 0\}$$

$$Q^*: \max\{ b^T y \; / \; A^T y \leq c \; , \; y \geq 0\}$$

TEOREMA 2.2.- Sea ψ una función cóncava-convexa en $\mathbb{R}^n \times \mathbb{R}^m$ que es dos veces diferenciable y tal que ψ_x y afin lineales en **y**. Sea (x, y) admisible para **P**. entonces a) y b) son equivalentes.

a) En el punto $(\mathbf{x}, \mathbf{y}, \mathbf{v})$ se cumplen las condiciones linealizadas de K-T para $\mathbf{P}$

b) $(\mathbf{x}, \mathbf{y})$ es óptimo para $\mathbf{P}$ y existe un $\mathbf{v}$ tal que $(\mathbf{x}, \mathbf{v})$ óptimo para $\mathbf{P}$ y $\mathbf{P}^*$ y $F(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}, \mathbf{v}) = g(\mathbf{x}, \mathbf{v})$

<u>OBSERVACION 2.3:</u> La exigencia que ψ_x y ψ_y son afin lineales en y es por ejemplo cumplida, si ψ es la función de Lagrange de un programa convexo o una función de COTTLE.

De manera general esta exigencia significa que ψ tiene la siguiente forma:

$$\psi(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^{\mathbf{T}} g(\mathbf{x}) - \mathbf{y}^{\mathbf{T}} \mathbf{C} \mathbf{y}$$

donde $\mathbf{C} = \mathbf{C}^{\mathbf{T}}$ es semi-definida positiva.

El programa $\mathbf{P}$, es en este caso un programa convexo.

DEMOSTRACION (del teorema 2.2)

a) $\Longrightarrow$ b)

Por hipótesis existen matrices $\mathbf{A}(\mathbf{x}) \in M(n,m)$, $\mathbf{B}(\mathbf{x}) \in M(m,m)$, $\alpha(\mathbf{x}) \in \mathbb{R}^n$, $\beta(\mathbf{x}) \in \mathbb{R}^m$, tales que:

$\psi_x(\mathbf{x}, \mathbf{y}) = \mathbf{A}(\mathbf{x})\mathbf{y} + \alpha(\mathbf{x})$, $\psi_y(\mathbf{x}, \mathbf{y}) = \mathbf{B}(\mathbf{x})\mathbf{y} + \beta(\mathbf{x})$

Entonces $\psi_{xy}(\mathbf{x}, \mathbf{y}) = \mathbf{A}(\mathbf{x})$, $\psi_{yy}(\mathbf{x}, \mathbf{y}) = \mathbf{B}(\mathbf{x})$

Por consiguiente $\forall\ \mathbf{x} \in \mathbb{R}^n$; $\forall\ \mathbf{y},\ \mathbf{v} \in \mathbb{R}^m$ se cumple

$$\left.\begin{aligned} \psi_x(\mathbf{x}, \mathbf{y}) + \psi_{xy}(\mathbf{x}, \mathbf{y})(\mathbf{v} \quad \mathbf{y}) &= \psi_x(\mathbf{x}, \mathbf{v}) \\ \psi_y(\mathbf{x}, \mathbf{y}) + \psi_{yy}(\mathbf{x}, \mathbf{y})(\mathbf{v} \quad \mathbf{y}) &= \psi_y(\mathbf{x}, \mathbf{v}) \end{aligned}\right\} \tag{2.7}$$

La función de Lagrange de $\mathbf{P}$ es:

$$\Phi(\mathbf{x}, \mathbf{y}, \mathbf{v}) \quad \psi(\mathbf{x}, \mathbf{y}) + (\mathbf{v} - \mathbf{y})\ \psi_y(\mathbf{x}, \mathbf{y})$$

Indicamos las condiciones linealizadas de K-T para $\mathbf{P}$

$$\Phi_{v^1}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{y^1}(\mathbf{x},\mathbf{y}) \leq 0,\ v^1 \geq 0,\ (v^1)^T\psi_{y^1}(\bar{\mathbf{x}},\bar{\mathbf{y}}) - 0$$

$$(2.8)\qquad \Phi_{2}(\mathbf{x},\mathbf{y},\mathbf{v}) = \Phi_{y^1}(\mathbf{x},\mathbf{y}) = 0$$

$$\Phi_{x^1}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{x^1}(\mathbf{x},\mathbf{y}) + \psi_{x^1 y}(\mathbf{x},\mathbf{y})(\mathbf{v} - y) \geq 0$$

$$(2.9)\qquad \left|\begin{array}{l} (\bar{x}^1)^T\Phi_{x^1}(\mathbf{x},\mathbf{y},\mathbf{v}) = 0 \qquad {}^{-1} \geq 0 \\ \Phi_{x^2}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{x^2}(\mathbf{x},\mathbf{y}) + \psi_{x^2 y}(\mathbf{x},\mathbf{y})(\mathbf{v} \quad \mathbf{y}) = 0 \\ \Phi_{y^1}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{y^1 y}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) \geq 0 \qquad \bar{y}^1 \geq 0 \end{array}\right.$$

$$(2.10)\qquad \begin{cases} (\bar{y}^1)^T\psi_{y^1 y}(\bar{\mathbf{x}},\bar{\mathbf{y}})(\mathbf{v} - \bar{\mathbf{y}}) = 0 \\ \Phi_{y^2}(\bar{\mathbf{x}},\bar{\mathbf{y}},\mathbf{v}) = \psi_{y^2 y}(\bar{\mathbf{x}},\bar{\mathbf{y}})(\mathbf{v} - \mathbf{y}) = 0 \end{cases}$$

De $v^1 \geq 0$ y (2.10) se obtiene

$$(2.10a)\qquad (\mathbf{v} \quad \bar{\mathbf{y}})^T\psi_{yy}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) \geq 0$$

Como $\psi(\mathbf{x},.)$ es cóncava en $\mathbb{R}^m$, la matriz $\psi_{yy}(\mathbf{x},\mathbf{y})$ es definida semi-negativa y de (2.10a) se obtiene

$$(2.11)\qquad (\mathbf{v} - \bar{\mathbf{y}})^T\psi_{yy}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) = 0$$

Como la forma cuadrática $\mathbf{Q}(\mathbf{z}) = \mathbf{z}^T\psi_{yy}(\mathbf{x},\mathbf{y})\mathbf{z}$, toma su mínimo en $\mathbf{z} = \mathbf{v} \quad \mathbf{y}$, se cumple:

$$(2.11a)\qquad \nabla\mathbf{Q}(\mathbf{v} \quad \mathbf{y}) = 2\,\psi_{yy}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) = 0$$

De (2.7), (2.8) y (2.9) obtenemos:

$$(2.12)\qquad \begin{cases} \psi_{x^1}(\mathbf{x},\mathbf{v}) \geq 0\ ,\quad \psi_{x^2}(\mathbf{x},\mathbf{v}) = 0 \qquad \bar{x}^1 \geq 0,\ v^1 \geq 0 \\ (\bar{x}^1)^T\psi_{x^1}(\bar{\mathbf{x}},\mathbf{v}) = 0 \end{cases}$$

Por consiguiente $(\mathbf{x},\mathbf{v})$ es admisible para $\mathbf{P}^*$

De (2.7), (2.8), (2.9) y (2.11a) obtenemos

$$(2.13) \quad \begin{cases} \psi_1(\mathbf{x},\mathbf{v}) = \psi_{y^1}(\mathbf{x},\mathbf{y}) \leq 0 & (v^1)^T\psi_1(\mathbf{x},\mathbf{v}) - \\ \psi_2(\mathbf{x},\mathbf{v}) \quad \psi_{y^2}(\mathbf{x},\mathbf{y}) = 0 & \bar{x}^1 \geq 0 \quad {}^1 \geq 0 \end{cases}$$

Por consiguiente $(\mathbf{x},\mathbf{v})$ es tambien admisible para **P**.

De (2.12) y (2.13) se deduce finalmente que

$$F(\mathbf{x},\mathbf{v}) = \psi(\mathbf{x},\mathbf{v}) \quad \mathbf{v}^T\psi\ (\mathbf{x},\mathbf{v}) \quad \psi(\mathbf{x},\mathbf{v})$$

$$- \ \psi(\mathbf{x},\mathbf{v}) \quad \bar{\mathbf{x}}^T\psi_x(\mathbf{x},\mathbf{v}) \quad g(\mathbf{x},\mathbf{v})$$

En consecuencia de acuerdo al teorema 2.1, $(\mathbf{x},\mathbf{v})$ es ópti mo para **P** y $\mathbf{P}^*$.

Como $\psi(\mathbf{x},.)$ es cóncava, se obtiene finalmente de (2.13)

$$F(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x},\mathbf{y}) \quad \bar{\mathbf{y}}^T\psi_y(\mathbf{x},\mathbf{y})$$

$$= \psi(\mathbf{x},\mathbf{y}) \quad (\mathbf{y} \quad \mathbf{v})^T\psi_y(\mathbf{x},\mathbf{y}) \leq \psi(\mathbf{x},\mathbf{v}) = F(\mathbf{x},\mathbf{v})$$

Por consiguiente $(\mathbf{x},\mathbf{y})$ es óptimo para **P** y se cumple:

$$F(\mathbf{x},\mathbf{y}) = F(\mathbf{x},\mathbf{v}) = g(\mathbf{x},\mathbf{v})$$

b) $\Longrightarrow$ a)

Por la proposición 2.2 $(\mathbf{x},\mathbf{v})$ es un punto silla de ψ sobre **C** x **D**. Como ψ es dos veces diferenciable, se cumplen en el punto $(\mathbf{x},\mathbf{v},\mathbf{v})$ las condiciones linealizadas de K-T para **P**.

Como **P** es un programa convexo, las condiciones linealiza das de K-T son equivalentes con las condiciones globales:

$$\Phi(\mathbf{x},\mathbf{v},\mathbf{v}) = F(\mathbf{x},\mathbf{v}) = \min_{\mathbf{x},\mathbf{y}} \{\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) \ / \ x^1 \geq 0, \ y^1 \geq 0\} \quad v^1 \geq 0$$

De otro lado de $v^1 \geq 0$ se obtiene $\psi_{y^1}(\mathbf{x},\mathbf{y}) \leq 0$, $\psi_{y^2}(\mathbf{x},\mathbf{y}) = 0$

$$\Phi(\bar{\mathbf{x}},\bar{\mathbf{y}},\mathbf{v}) = F(\mathbf{x},\mathbf{y}) + \mathbf{v}^T\psi_y(\mathbf{x},\mathbf{y}) = F(\mathbf{x},\mathbf{y}) + (v^1)^T\psi_{y^1}(\bar{\mathbf{x}},\bar{\mathbf{y}}) \leq$$

$$\leq F(\mathbf{x},\mathbf{v}).$$

Por consiguiente se cumple:

$$\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) = F(\mathbf{x},\mathbf{y}) = \min_{\mathbf{x},\mathbf{y}} \{\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) \;/\; x^1 \geq 0,\; y^1 \geq 0\}$$

lo que significa que las condiciones (linealizadas y glovales) de K-T se cumplen en el punto $(\mathbf{x},\mathbf{y},\mathbf{v})$. ■

Como el par de programas $(\mathbf{P},\mathbf{P}^*)$ es simétrico se cumple el siguiente teorema, que es la versión dual del teorema 2.2

TEOREMA 2.3.- Sea $\psi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$, una función convexa- cóncava, que es dos veces diferenciable y tal que ψ_x y son afin lineales en $\mathbf{y}$.

Entonces a) y b) son equivalentes

a) En $(\mathbf{x},\mathbf{y},\mathbf{v})$ se cumplen las condiicones linealizadas de K-T para $\mathbf{P}^*$.

b) $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ es óptimo para $\mathbf{P}^*$ y existe un $\mathbf{v}$ tal que $(\mathbf{v},\mathbf{y})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^*$ y $g(\mathbf{x},\mathbf{y}) = F(\mathbf{v},\mathbf{y}) = g(\mathbf{v},\mathbf{y})$

Si la condición que ψ_x y ψ_y son afin lineales, es sustituida por la exigencia que $\psi_{yy}(\mathbf{x},\mathbf{y})$ es definida negativa, entonces la tesis del teorema 2.2 se deja reforzar.

TEOREMA 2.4.- Sea $\psi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$, una función convexa-cóncava, que es dos veces diferenciable. Sea $(\mathbf{x},\mathbf{y})$ admisible para

$\mathbf{P}$ y $\psi_{yy}(\mathbf{x},\mathbf{y})$ definida negativa. Entonces se cumple:

a) Si en $(\mathbf{x},\mathbf{y},\mathbf{v})$ se cumplen las condiciones linealizadas de K-T de $\mathbf{P}$, entonces $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^{*}$ y $F(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$.

b) Si $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$, entonces en el punto $(\mathbf{x},\mathbf{y},\mathbf{v})$ se cumplen las condiciones linealizadas de K-T de $\mathbf{P}$. Entonces $(\mathbf{x},\mathbf{y})$ es también optimo para $\mathbf{P}^{*}$ y se cumple :

$$F(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$$

DEMOSTRACION

a) Como en la demostración del teorema 2.2, se deduce de las condiciones linealizadas de K-T que $\psi_{yy}(\mathbf{x},\mathbf{y})(\mathbf{v}-\mathbf{y}) = 0$ Como $\psi_{yy}(\mathbf{x},\mathbf{y})$ es definida negativa, de esto se deduce que $\mathbf{v} = \mathbf{y}$.

Entonces las restantes condiciones (2.8) y (2.9) de K-T se simplifican como sigue:

$$\psi_{y^1}(\mathbf{x},\mathbf{y}) \leq 0 \qquad \bar{y}^1 \geq 0 \qquad (\bar{y}^1)^T \psi_{y^1}(\mathbf{x},\mathbf{y}) = 0$$

$$\psi_{2}(\mathbf{x},\mathbf{y}) =$$

$$\psi_{1}(\mathbf{x},\mathbf{y}) \geq 0\ , \quad ^{-1} \geq 0 \qquad (\bar{x}^1)^T \psi_{x^1}(\mathbf{x},\mathbf{y}) = 0$$

$$\psi_{x^2}(\mathbf{x},\mathbf{y}) = 0$$

Por consiguiente $(\bar{\mathbf{x}},\mathbf{y})$ es admisible para $\mathbf{P}$ y $\mathbf{P}^{*}$ y

$$F(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y}).$$

De acuerdo con el teorema 2.1, $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^{*}$.

b) Consideramos la función generalizada de Lagrange para **P**

$$L(\mathbf{x},\mathbf{y},\mathbf{v}) = L(\mathbf{x},\mathbf{y},v^0,v^1,v^2) = v^0\psi(\mathbf{x},\mathbf{y}) + (v^1-v^0y^1)^T\psi_{y^1}(\mathbf{x},\mathbf{y}) + (v^2)^T\psi_{y^2}(\mathbf{x},\mathbf{y})$$

Como $(\mathbf{x},\mathbf{y})$ es óptimo para el programa convexo **P**, se cumplen en un punto $(\mathbf{x},\mathbf{y},\mathbf{v})$; $\mathbf{v} = (v^0, v^1, v^2)$, las condiciones linealizadas de JOHN.

$$L_{y^1}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{y^1y^1}(\bar{\mathbf{x}},\bar{\mathbf{y}})(v^1-v^0\bar{y}^1) + \psi_{y^1y^2}(\bar{\mathbf{x}},\bar{\mathbf{y}})v^2 \geq 0 \ , \quad \bar{y}^1 \geq 0$$

$$(\bar{y}^1)^T\left[\psi_{y^1y^1}(\bar{\mathbf{x}},\bar{\mathbf{y}})(v^1-v^0\bar{y}^1) + \psi_{y^1y^2}(\bar{\mathbf{x}},\bar{\mathbf{y}})v^2\right] = 0$$

$$L_{y^2}(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_{y^2y^1}(\bar{\mathbf{x}},\bar{\mathbf{y}})(v^1-v^0\bar{y}^1) + \psi_{y^2y^2}(\bar{\mathbf{x}},\bar{\mathbf{y}})v^2 = 0$$

$$(v^0, v^1) \geq 0 \qquad \mathbf{v} \neq 0.$$

Si $\mathbf{u} = (v^1-v^0\bar{y}^1, v^2)$, de esto obtenemos $\mathbf{u}^T\psi_{yy}(\mathbf{x},\mathbf{y})\mathbf{u} \geq 0$

Como $\psi_{yy}(\mathbf{x},\mathbf{y})$ es definida negativa, se obtiene de esto que

$$\mathbf{u} = 0 \quad \text{o} \quad v^1 = v^0\bar{y}^1 \qquad v^2 = 0$$

Si fuera $v^0 = 0$ se obtendría de esto que $\mathbf{v} = 0$, lo que es absurdo.

Por consiguiente $v^0 > 0$ y en el punto $\left(\mathbf{x}, \mathbf{y}, \frac{v^1}{v^0}, \frac{v}{v^0}\right)$, se cumplen las condiciones de K-T para **P**.

De acuerdo con la parte a), $(\mathbf{x},\mathbf{y})$ es óptimo para **P** y **P*** y se cumple $F(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$.

TEOREMA 2.5.- Sea $(\mathbf{x},\mathbf{y})$ admisible para **P**. Sea $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ admisible para **P***. Sean cumplidas en $(\mathbf{x},\mathbf{y},\mathbf{v})$ las condiciones de K-T para **P** y en $(\tilde{\mathbf{x}},\tilde{\mathbf{y}},\mathbf{w})$ las condiciones de K-T para **P***. Entonces $(\bar{\mathbf{x}},\tilde{\mathbf{y}})$ es óptimo para **P** y **P***, $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ es óptimo para **P** y $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$

es óptimo para $\mathbf{P}^*$. Finalmente se cumple

$$F(\bar{\mathbf{x}},\bar{\mathbf{y}}) = F(\bar{\mathbf{x}},\tilde{\mathbf{y}}) = g(\bar{\mathbf{x}},\tilde{\mathbf{y}}) = g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$$

DEMOSTRACION

La función de Lagrange de $\mathbf{P}$ es :

$\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi(\mathbf{x},\mathbf{y}) + (\mathbf{v} - \mathbf{y})^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y})$ y las condiciones de K-T

para $\mathbf{P}$ en el punto $(\mathbf{x},\mathbf{y},\mathbf{v})$ son

$$(2.14) \quad \Phi(\mathbf{x},\mathbf{y},\mathbf{v}) - \psi(\mathbf{x},\mathbf{y}) \quad \bar{\mathbf{y}}^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) =$$

$$= \min_{\mathbf{x},\mathbf{y}} \left\{\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) \ / \ x^{i} \geq 0 \quad y^{i} \geq 0\right\}, \quad {}^{i} \geq 0$$

La función de Lagrange de $\mathbf{P}^*$ es

$$\Psi(\mathbf{x},\mathbf{y},\mathbf{w}) \quad \psi(\mathbf{x},\mathbf{y}) + (\mathbf{w} - \mathbf{x})^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y})$$

Las condiciones de K-T para $\mathbf{P}^*$ en $(\tilde{\mathbf{x}},\tilde{\mathbf{y}},\mathbf{w})$ son:

$$(2.15) \quad \Psi(\tilde{\mathbf{x}},\tilde{\mathbf{y}},\mathbf{w}) = \psi(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) - \tilde{\mathbf{x}}^{\mathbf{T}}\psi_x(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$$

$$= \max_{\mathbf{x},\mathbf{y}} \left\{\Psi(\mathbf{x},\mathbf{y},\mathbf{w}) \ / \ x^{i} \geq 0 \quad y^{i} \geq 0\right\}, \quad {}^{i} \geq 0$$

De la proposición 2.1 y de (2.14) se deduce:

$$\psi(\mathbf{x},\tilde{\mathbf{y}}) \leq F(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x},\mathbf{y}) \quad \bar{\mathbf{y}}^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) =$$

$$- \psi(\mathbf{x},\mathbf{y}) + (\mathbf{v} - \bar{\mathbf{y}})^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) \leq$$

$$\leq \psi(\mathbf{x},\mathbf{y}) + (\mathbf{v} - \mathbf{y})^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}), \quad \forall \ (\mathbf{x},\mathbf{y}) \in \mathbf{C}\times\mathbf{D}$$

Como $(\mathbf{w},\mathbf{v}) \in \mathbf{C}\times\mathbf{D}$, se obtiene de esto

$$\psi(\bar{\mathbf{x}},\tilde{\mathbf{y}}) \leq F(\bar{\mathbf{x}},\bar{\mathbf{y}}) \leq \psi(\mathbf{w},\mathbf{v}) \qquad (2.15a)$$

De la proposicion 2.1 y de 2.15 se deduce que

$$\psi(\bar{\mathbf{x}},\tilde{\mathbf{y}}) \geq g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) = \psi(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) + (\mathbf{w} \quad \tilde{\mathbf{x}})^{\mathbf{T}}\psi_x(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \geq$$

$$\geq \psi(\mathbf{x},\mathbf{y}) + (\mathbf{w} - \mathbf{x})^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y}), \quad \forall \ (\mathbf{x},\mathbf{y}) \in \mathbf{C} \times \mathbf{D}$$

De esto y de (2.15a) se obtiene:

$$\psi(\bar{\mathbf{x}},\tilde{\mathbf{y}}) \ge g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \ge \psi(\mathbf{w},\mathbf{v}) \ge F(\mathbf{x},\mathbf{y}) \ge \psi(\bar{\mathbf{x}},\tilde{\mathbf{y}})$$

$$\psi(\mathbf{x},\tilde{\mathbf{y}}) = \psi(\mathbf{w},\mathbf{v}) = F(\mathbf{x},\mathbf{y}) = g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}).$$

En consecuencia de acuerdo al teorema 2.1, $(\bar{\mathbf{x}},\tilde{\mathbf{y}})$ es óptimo para **P** y $\mathbf{P}^*$ y se cumple :

$$F(\bar{\mathbf{x}},\tilde{\mathbf{y}}) = \psi(\mathbf{x},\tilde{\mathbf{y}}) = g(\mathbf{x},\tilde{\mathbf{y}}) = F(\mathbf{x},\mathbf{y}) = g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$$

Por consiguiente $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ es óptimo para **P** y $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ es óptimo para $\mathbf{P}^*$. ■

Usando las condiciones linealizadas de K-T, se obtiene el siguiente teorema menos fuerte que el teorema 2.5.

TEOREMA 2.6.- Si en los puntos $(\mathbf{x},\mathbf{y},\mathbf{v})$ y $(\mathbf{x},\mathbf{y},\mathbf{w})$ son cumplidas las condiciones linealizadas de K-T de **P** y de $\mathbf{P}^*$, entonces $(\mathbf{x},\mathbf{y})$ es óptimo para **P** y $\mathbf{P}^*$.

DEMOSTRACION

Para simplificar la demostaración asumimos que

$$\mathbf{x} = x \qquad \mathbf{y} = y^1$$

Las funciones de Lagrange de **P** y $\mathbf{P}^*$ son:

$$\Phi(\mathbf{x},\mathbf{y},\mathbf{v}) = \psi(\mathbf{x},\mathbf{y}) + (\mathbf{v} \quad \mathbf{y})^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y})$$

$$\Psi(\mathbf{x},\mathbf{y},\mathbf{w}) = \psi(\mathbf{x},\mathbf{y}) + (\mathbf{w} - \mathbf{y})^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y})$$

De las condiciones linealizadas de K-T para **P** y $\mathbf{P}^*$ deducimos

$\Phi\ (\mathbf{x},\mathbf{y},\mathbf{v}) = \psi_x(\mathbf{x},\mathbf{y}) + \psi_{xy}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) \ge 0$ a)

$\bar{\mathbf{x}}^{\mathbf{T}}\Phi_x(\bar{\mathbf{x}},\bar{\mathbf{y}},\mathbf{v}) = 0 \qquad \mathbf{x} \ge 0$ b)

$\psi\ (\mathbf{x},\mathbf{y}) \le 0 \qquad \mathbf{v}^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) = 0 \qquad \ge 0$ c)

$\Psi\ (\mathbf{x},\mathbf{y},\mathbf{w}) = \psi_y(\mathbf{x},\mathbf{y}) + \psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{w} - \mathbf{x}) \le 0$ d)

$$\bar{\mathbf{y}}^{\mathbf{T}}\Psi_x(\mathbf{x},\mathbf{y},\mathbf{w}) = 0 \qquad \mathbf{y} \geq 0 \tag{e}$$

$$\psi_x(\mathbf{x},\mathbf{y}) \geq 0 \qquad \mathbf{w}^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y}) = 0 \qquad \mathbf{w} \geq 0 \tag{f}$$

De a) y f) obtenemos :

$$\mathbf{w}^{\mathbf{T}}\psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) \geq - \mathbf{w}^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y}) = 0 \tag{g}$$

De a) y c) obtenemos :

$$\mathbf{v}^{\mathbf{T}}\psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{w} - \mathbf{x}) \leq - \mathbf{v}^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) = 0 \tag{h}$$

De la proposición 2.1, b), e), g) y h) obtenemos:

$$F(\mathbf{x},\mathbf{y}) - g(\mathbf{x},\mathbf{y}) = \bar{\mathbf{x}}^{\mathbf{T}}\psi_x(\mathbf{x},\mathbf{y}) - \bar{\mathbf{y}}^{\mathbf{T}}\psi_y(\mathbf{x},\mathbf{y}) =$$

$$= \quad \bar{\mathbf{x}}^{\mathbf{T}}\psi_{xy}(\mathbf{x},\mathbf{y})(\mathbf{v} \quad \mathbf{y}) + \bar{\mathbf{y}}^{\mathbf{T}}\psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{w} \quad \mathbf{x}) \leq$$

$$\leq (\mathbf{w} \quad \bar{\mathbf{x}})^{\mathbf{T}}\psi_{xy}(\mathbf{x},\mathbf{y})(\mathbf{v} \quad \mathbf{y}) + (\mathbf{y} - \mathbf{v})^{\mathbf{T}}\psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{w} - \mathbf{x}) =$$

$$= (\mathbf{w} - \bar{\mathbf{x}})^{\mathbf{T}}\psi_{xy}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) - (\mathbf{w} - \bar{\mathbf{x}})^{\mathbf{T}}\psi_{yx}(\mathbf{x},\mathbf{y})(\mathbf{v} - \mathbf{y}) = 0$$

Por consiguiente se cumple

$$F(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$$

y de acuerdo con el teorema 2.1, $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^*$

2.3: TEORIA DE DUALIDAD PARA PROGRAMAS CUADRATICOS CONVEXOS

Aplicaremos ahora los resultados de la teoría de dualidad a pares $(\mathbf{P},\mathbf{P}^*)$ de programas cuadráticos convexos, esto es : asumimos que ψ es una función de COTTLE.

En este caso $\mathbf{P}$ y $\mathbf{P}^*$ son programas convexos y $(\mathbf{x},\mathbf{y})$ es óptimo para uno de los programas $\mathbf{P}$ o $\mathbf{P}^*$ si y solamente si en un punto $(\mathbf{x},\mathbf{y},\mathbf{v})$ se cumplen las condiciones (linealizadas) de K-T para $\mathbf{P}$ o $\mathbf{P}^*$.

Por esta razón los teoremas 2.2, 2.4 y se dejan resumir en el siguiente teorema de dualidad de la programación cuadrática.

TEOREMA 2.7.- Sea $\psi(\mathbf{x},\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} + \frac{1}{2}(\mathbf{x}^T\mathbf{B}\mathbf{x} - \mathbf{y}^T\mathbf{C}\mathbf{y}) - \mathbf{y}^T\mathbf{A}\mathbf{x}$ una función de COTTLE. Entonces se cumple:

a) $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$ si y solamente si $(\mathbf{x},\mathbf{y})$ es admisible para $\mathbf{P}$ y existe un $\mathbf{v}$ tal que $(\mathbf{x},\mathbf{v})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^*$ y se cumple : $F(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{v}) = F(\mathbf{x},\mathbf{v}) = \psi(\mathbf{x},\mathbf{v})$.

b) Si $\mathbf{C}$ es definida positiva y $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$, entonces $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ es tambien óptimo para $\mathbf{P}^*$ y $F(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$.

c) Si $(\mathbf{x},\mathbf{y})$ es óptimo para $\mathbf{P}$, $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ es óptimo para $\mathbf{P}^*$ y entonces $(\mathbf{x},\tilde{\mathbf{y}})$ es óptimo para $\mathbf{P}$ y $\mathbf{P}^*$ y se cumple

$$F(\mathbf{x},\mathbf{y}) = F(\bar{\mathbf{x}},\tilde{\mathbf{y}}) = g(\bar{\mathbf{x}},\tilde{\mathbf{y}}) = g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$$

Para desarrollar mejor la teoría de dualidad de la programación cuadrática convexa, necesitamos el siguiente teorema de existencia, cuya demostración se encuentra en la referencia [ ].

TEOREMA 2.8.- Sea $F(\mathbf{x}) = \mathbf{p}^T\mathbf{x} + \mathbf{x}^T\mathbf{C}\mathbf{x}$ $[\mathbf{C}^T = \mathbf{C}]$ una función cuadrática no necesariamente convexa.

Sea $S = \{ \mathbf{x} \in \mathbb{R}^n \ / \ \mathbf{a}_j^T\mathbf{x} \quad \mathbf{b}_j \leq 0 \qquad j = \overline{1,m} \} \neq \phi$ y

Sea ademas, F inferiormente acotada en S. Entonces el programa : min $\{F(\mathbf{x}) \ / \ \mathbf{x} \in S\}$ tiene al menos una solución óptima.

Del teorema de dualidad 2.7 y del teorema de existencia 2.8 se deduce fácilmente el

TEOREMA 2.9.- Sea $(\mathbf{P},\mathbf{P}^*)$ un par de programas convexos y cuadráticos. Entonces se cumple :

a) Si $\inf(\mathbf{P}) = -\infty$ entonces el dominio R^* de $\mathbf{P}^*$ es vacío

b) Si el dominio admisible R de **P** es vacío y $R^* \neq \phi$, entonces $\sup(\mathbf{P}^*) = +\infty$.

DEMOSTRACION

a) Asumimos que existe un $(\mathbf{x}^o,\mathbf{y}^o) \in R^*$. Entonces se cumpliría

$F(\mathbf{x},\mathbf{y}) \geq g(\mathbf{x}^o,\mathbf{y}^o)$, $\forall$ $(\mathbf{x},\mathbf{y}) \in R$ o

$\inf(\mathbf{P}) \geq g(\mathbf{x}^o,\mathbf{y}^o) \qquad \infty$

b) Si fuera $\sup(\mathbf{P}^*) < \infty$, entonces de acuerdo al teorema 2.8, $\mathbf{P}^*$ tendría una solución óptima y por el teorema 2.7, **P** tambien tendría una solución óptima y R fuera no vacío. Por consiguiente $\sup(\mathbf{P}^*) = +\infty$. ■

Del mismo modo se demuestra el teorema dual al teorema 2.9. De los teoremas 2.8 y 2.9 se obtiene el siguiente teorema de clasificación para los programas cuadráticos convexos, que incluye el teorema de existencia y el teorema de dualidad de la programación lineal como caso especial.

TEOREMA 2.10.- Sea dado un par $(\mathbf{P},\mathbf{P}^*)$ de programas cuadráticos convexos, con los dominios admisibles R y R^*. Entonces se presenta exactamente uno de los casos siguientes:

a) $R = R^* = \phi$

b) $R \neq \phi \quad R^* = \phi$ y $\inf(\mathbf{P}) = \quad \infty$

o $R = \phi \quad R^* \neq \phi$ y $\sup(\mathbf{P}^*) = +\infty$

c) $\mathbf{P}$ y $\mathbf{P}^*$ tienen soluciones óptimas y $\inf(\mathbf{P}) = \sup(\mathbf{P}^*)$

Si ψ es la función de Lagrange del programa convexo cuadrático:

$$\mathbf{I} : \min \{ \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} \;/\; \mathbf{A}\mathbf{x} \leq b \;,\; \mathbf{x} \geq 0 \}$$

entonces $\mathbf{P}$ coincide con $\mathbf{I}$ y $\mathbf{P}^*$ tiene la forma

$$\mathbf{P}^*: \max \{- \mathbf{x}^T\mathbf{C}\mathbf{x} - \mathbf{b}^T\mathbf{u} \;/\; 2\mathbf{C}\mathbf{x} + \mathbf{A}^T\mathbf{u} + \mathbf{p} \geq 0 \;,\; \mathbf{x} \geq 0 \quad \mathbf{u} \geq 0 \}$$

Del teorema 2.7 obtenemos el siguiente criterio de optimidad para $\mathbf{I}$.

TEOREMA 2.11.- $\mathbf{x}$ es óptimo para $\mathbf{I}$ si y solamente si, existe un $\mathbf{u}$, tal que $(\mathbf{x},\mathbf{u})$ es admisible para $\mathbf{P}^*$ y $F(\mathbf{x}) = g(\mathbf{x},\mathbf{u}) = \psi(\mathbf{x},\mathbf{u})$.

Vamos a demostrar que el criterio de optimidad del teorema 2.11 es equivalente con las condiciones que anteriormente deducimos de las condiciones de K-T para I.

De acuerdo con el teorema 2.11, $\mathbf{x}$ es óptimo para I si y solamente si existe un $\mathbf{u}$ tal que :

(2.16) $\quad \bar{x}^T C\bar{x} + p^T\bar{x} = \bar{x}^T C\bar{x} + p^T\bar{x} + \bar{u}^T(A\bar{x} - b) = \bar{x}^T C\bar{x} - b^T\bar{u}$

(2.17) $\quad 2Cx + A^T\bar{u} + p \geq 0 \quad , \quad u \geq 0 \qquad Ax - b \leq 0 \qquad x \geq 0$

De (2.16) obtenemos:

(2.18) $\quad \bar{u}^T A\bar{x} = b^T\bar{u}$

(2.19) $\quad - b^T u = 2\bar{x}^T C\bar{x} + p^T\bar{x}$

De (2.18) y (2.19) obtenemos

(2.20) $\quad \bar{x}^T(2C\bar{x} + A^T\bar{u} + p) = 0$

Si ponemos $y = b - Ax$, $v = 2Cx + A^T\bar{u} + p$. Obtenemos de (2.17), (2.18) y (2.20)

(2.21)
$$\begin{aligned} &Ax + y = b \\ &2Cx \quad v + A^T\bar{u} = p \\ &x \geq 0 \ , \ v \geq 0 \ , \ u \geq 0 \ , \ y \geq 0 \\ &\bar{x}^T\bar{v} + \bar{y}^T\bar{u} = 0 \end{aligned}$$

(2.21) representa exactamente las condiciones de optimidad del capítulo precedente.

Al revés, de (2.21) se deduce fácilmente las condiciones (2.16) y (2.17).

El teorema se deja también formular de la siguiente manera: $\bar{x}$ es óptimo para I si y solamente si existe un $\mathbf{u}$ tal que $(\mathbf{x},\mathbf{u})$ resuelve el siguiente sistema lineal.

$$\mathbf{Ax} \quad \mathbf{b} \leq 0$$

$$2\mathbf{Cx} \quad \mathbf{A}^T\mathbf{u} \quad \mathbf{p} \leq 0$$

$$-\mathbf{x} \leq 0 \qquad -\mathbf{u} \leq 0$$

$$F(\mathbf{x}) - g(\mathbf{x},\mathbf{u}) = 2\mathbf{x}^T\mathbf{cx} + \mathbf{p}^T\mathbf{x} + \mathbf{b}^T\mathbf{u} \leq 0.$$

El siguiente teorema colateral es a veces útil.

TEOREMA 2.12.- a) Si $(\mathbf{x},\mathbf{u})$ es óptimo para $\mathbf{P}^*$ entonces existe un $\hat{\mathbf{x}}$, tal que $\hat{\mathbf{x}}$ es óptimo para **I** y $\mathbf{C}(\hat{\mathbf{x}} \quad \mathbf{x}) = 0$

b) Si $\hat{\mathbf{x}}$ es óptimo para **I** y $\mathbf{C}(\hat{\mathbf{x}} - \mathbf{x}) = 0$, $\mathbf{x} \geq 0$, entonces existe un $\mathbf{u}$, tal que $(\mathbf{x},\mathbf{u})$ es óptimo para $\mathbf{P}^*$.

DEMOSTRACION

a) En un punto $(\bar{\mathbf{x}},\bar{\mathbf{u}},\hat{\mathbf{x}})$ son cumplidas las condiciones linealizadas de K-T. De manera análoga con la demostración del teorema se deduce que $\hat{\mathbf{x}}$ es óptimo para **I** y que :

$$\psi_{xx}(\mathbf{x},\mathbf{u})(\hat{\mathbf{x}} \quad \mathbf{x}) = \mathbf{C}(\hat{\mathbf{x}} - \mathbf{x}) = 0$$

b) Existe un $\mathbf{u}$ tal que $(\hat{\mathbf{x}},\bar{\mathbf{u}})$ es óptimo para $\mathbf{P}^*$. Como $\mathbf{C}(\hat{\mathbf{x}} - \mathbf{x}) = 0$ y $(\hat{\mathbf{x}},\bar{\mathbf{u}})$ es admisible para $\mathbf{P}^*$, $(\mathbf{x},\mathbf{u})$ es tambien admisible para $\mathbf{P}^*$ y se cumple :

$$\begin{aligned} g(\mathbf{x},\mathbf{u}) \quad g(\hat{\mathbf{x}},\bar{\mathbf{u}}) &= \hat{\mathbf{x}}^T\mathbf{C}\hat{\mathbf{x}} \quad \bar{\mathbf{x}}^T\mathbf{C}\bar{\mathbf{x}} \\ &= \bar{\mathbf{x}}^T\mathbf{C}\hat{\mathbf{x}} \quad \bar{\mathbf{x}}^T\mathbf{C}\bar{\mathbf{x}} \\ &= \bar{\mathbf{x}}^T\mathbf{C}(\hat{\mathbf{x}} - \mathbf{x}) \\ &= 0 \end{aligned}$$

■

CAPITULO III

3.1: EL ALGORITMO DE WOLFE

Consideremos el programa:

$$\mathbf{P} : \min \{ F(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} \;/\; \mathbf{A}\mathbf{x} = \mathbf{b} \;,\; \mathbf{x} \geq 0 \}$$

donde $\mathbf{C} = \mathbf{C}^T \in M(n,n)$ es definida positiva $\mathbf{A} \in M(m,n)$, $m < n$ $ran(\mathbf{A}) = m$.

<u>NOTA:</u> El programa **P** tiene una única solución óptima siempre y cuando su dominio admisible: $S = \{ \mathbf{x} \;/\; \mathbf{A}\mathbf{x} = \mathbf{b} \;,\; \mathbf{x} \geq 0 \}$, es no vacío

<u>EXPLICACION:</u>

Sea $\phi \neq \mathbf{D} \subset \mathbb{R}^n$, convexo y cerrado,

$g : \mathbf{D} \longrightarrow \mathbb{R}$, convexa y continua.

Consideremos el siguiente programa

$$\mathbf{Q} : \min \{ g(\mathbf{x}) \;/\; \mathbf{x} \in \mathbf{D} \}$$

Es fácil demostrar el siguiente

LEMA 3.1.- **Q** tiene una solución óptima si se cumple a) o b)

a) **D** es compacto

b) **D** no es acotado y g es coercitiva respecto a **D**.

Que g sea coercitiva respecto a $\mathbf{D}$ significa:

$$\lim_{\substack{\|\mathbf{x}\| \longrightarrow \infty \\ \mathbf{x} \in \mathbf{D}}} g(\mathbf{x}) = \infty$$

En consecuencia basta demostrar que $F(\mathbf{x})$ es coercitiva respecto a S, si S no es acotado.

Sea λ el menor valor propio de $\mathbf{C}$. $\lambda > 0$ pues $\mathbf{C}$ es definida positiva.

Por consiguiente se cumple:

$$F(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\ \mathbf{x} + \mathbf{p}^T\mathbf{x} \geq \lambda\ \|\mathbf{x}\|^2 + \mathbf{p}^T\mathbf{x} \geq \lambda\ \|\mathbf{x}\|^2 - \|\mathbf{p}\|\|\mathbf{x}\| \geq$$

$$\geq \|\mathbf{x}\|(\lambda\|\mathbf{x}\| - \|\mathbf{p}\|)\ ,\ \forall\ \mathbf{x} \in \mathbb{R}^n$$

De esto se obtiene :

$$\lim_{\|\mathbf{x}\| \longrightarrow \infty} F(\mathbf{x}) \geq \lim_{\|\mathbf{x}\| \longrightarrow \infty} \|\mathbf{x}\|(\lambda\|\mathbf{x}\| \quad \|\mathbf{p}\|) = \infty$$

En consecuencia, F es coercitiva (respecto a S) ■

El algoritmo de WOLFE se basa en la condiciones linealizadas de K_T para $\mathbf{P}$.

$\Phi(\mathbf{x},\mathbf{u}) = \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} + \mathbf{u}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$, es la función de LAGRANGE para $\mathbf{P}$.

$\hat{\mathbf{x}}$ es una solución óptima de $\mathbf{P}$ si y sólo si existe $\hat{\mathbf{u}} \in \mathbb{R}^n$ tal que en $(\hat{\mathbf{x}},\hat{\mathbf{u}})$ se cumplen las condiciones linealizadas de K T.

Haciendo $\hat{\mathbf{v}} = 2\mathbf{C}\hat{\mathbf{x}} + \mathbf{p} + \mathbf{A}^T\hat{\mathbf{u}} = \nabla_x\Phi(\hat{\mathbf{x}},\hat{\mathbf{u}}) \geq 0$, estas condiciones se dejan escribir como

$$\mathbf{A}\mathbf{x} = \mathbf{b} \qquad (3.1)$$

$$2\mathbf{C}\hat{\mathbf{x}} + \hat{\mathbf{v}} \quad \mathbf{A}^T\hat{\mathbf{u}} = \mathbf{p} \qquad (3.2)$$

$$\hat{\ } \geq 0 \qquad \mathbf{v} \geq 0 \qquad (3.3)$$

$$\hat{\mathbf{x}}^T\hat{\mathbf{v}} = 0 \qquad (3.4)$$

Si encontramos una solución $\hat{\mathbf{w}} = (\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{u}})$ del sistema (3.1)-(3.4) entonces $\hat{\mathbf{x}}$ es una solución óptima del programa **P**

El algotirmo de WOLFE resuelve el sistema (3.1) ---(3.4) aplicando una versión del ALGORITMO SIMPLEX programa lineal derivado de (3.1) ---(3.4)

Determinamos primero una solución básica admisible **x** de:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \ , \quad \mathbf{x} \geq 0$$

$$\bar{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_B \\ 0 \end{bmatrix} \qquad = \mathbf{A}_B^{-1}\mathbf{b} \qquad \mathbf{A} = [\mathbf{A}_B, \mathbf{A}_N]$$

Calculamos:

$$\mathbf{h} = \nabla F(\mathbf{x}) - 2C\mathbf{x} + \mathbf{p}$$

Si $\mathbf{h} = 0$, entonces **x** es una solución óptima de **P**

Si $\mathbf{h} \neq 0$, formamos el programa **I** que es no lineal

$$\mathbf{I} : \max\ (-z)$$

$$\left.\begin{aligned} \mathbf{A}\mathbf{x} \qquad\qquad\qquad &= \mathbf{b} \\ -2C\mathbf{x} + \mathbf{v} - \mathbf{A}^T\mathbf{u} + z\mathbf{h} &= \mathbf{p} \\ \mathbf{x} \geq 0, \ \mathbf{v} \geq 0, \quad z &\geq 0 \end{aligned}\right\} \qquad (3.5)$$

$$\mathbf{x}^T\mathbf{v} = 0 \qquad (3.6)$$

Por (3.6) este programa no es lineal.

Sin embargo podemos resolver el programa **I**, aplicando una modificación del algoritmo SIMPLEX al algoritmo lineal con las restricciones (3.5)

Si el programa **P** tiene una solución óptima, entonces el programa I tiene una solución óptima $\hat{\mathbf{w}} = (\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{u}}, \hat{z})$ con $\hat{z} = 0$

Entonces $(\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{u}})$ satisface las condiciones de K-T y $\hat{\mathbf{x}}$ es una solución óptima de **P**.

3.1.1: MODIFICACION DEL ALGORITMO SIMPLEX

Se procura que todas las soluciones básicas $\mathbf{w} = (\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \tilde{\mathbf{u}}, \tilde{z})$ de (3.5) que se construyan, satisfagan (3.6).

Es necesario construir una primera solución básica $\mathbf{w}^* = (\mathbf{x}^*, \mathbf{v}^*, \mathbf{u}^*, z^*)$ admisible para (3.5) que satisface $(\mathbf{x}^*)^T\mathbf{v}^* = 0$.

Un punto admisible que satisface (3.6) es $\mathbf{x} = \mathbf{x}$, $\mathbf{v} = 0$, $\mathbf{u} = 0$, $z = 1$.

Vamos a construir una base de la matriz de restricciones $\mathbf{D}$ de (3.5) a la cual está asociada $\mathbf{x} = \mathbf{x}$, $\mathbf{v} = 0$, $\mathbf{u} = 0$, $z = 1$ como solución básica.

Tenemos:

$$\mathbf{D} = \begin{bmatrix} \mathbf{A} & 0 & 0 & 0 \\ -2\mathbf{C} & \mathbf{I} & & \mathbf{h} \end{bmatrix} \in M(m+n, 2n+m+1)$$

Si $\mathbf{A} = [\mathbf{A}_B, \mathbf{A}_N]$, entonces $\mathbf{D}$ se deja escribir como:

$$\mathbf{D} \quad \begin{bmatrix} \mathbf{A}_B & \mathbf{A}_N & 0 & 0 & 0 & 0 \\ -2\mathbf{C}_B & -2\mathbf{C}_N & \mathbf{I}_B & \mathbf{I}_N & -\mathbf{A}^T & \mathbf{h} \end{bmatrix}$$

<u>AFIRMACION:</u>

$$\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{A}_B & 0 & 0 \\ -2\mathbf{C}_B & & -\mathbf{A}^T \end{bmatrix} \in M(m+n, m+n) \text{ es una base de } \mathbf{D}$$

Como las columnas de $\mathbf{A}_B$ son linealmente independientes las columnas de $\begin{bmatrix} \mathbf{A}_B \\ -2\mathbf{C}_B \end{bmatrix}$ tambien lo son.

Podemos escribir

$$\mathbf{D'} \quad \begin{bmatrix} 0 & 0 \\ {}_{N} & -A^{T} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -A^{T} \\ I^{*}_{N} & -A^{T}_{N} \end{bmatrix}$$ donde I_N es una matriz unitaria.

Tenemos

$$\det \begin{bmatrix} 0 & -A^{T}_{B} \\ I^{*}_{N} & -A^{T}_{N} \end{bmatrix} = \pm \det \begin{bmatrix} A^{T}_{B} & 0 \\ A^{T}_{N} & I^{*}_{N} \end{bmatrix} = \pm \det A^{T}_{B} . \det I^{*}_{N} \neq 0$$

Luego: $\begin{bmatrix} 0 & -A^{T}_{B} \\ I^{*}_{N} & -A^{T}_{N} \end{bmatrix}$ es invertible, entonces las columnas de

D' son linealmente independientes.

Luego **D** es no singular y con esto una base de **D** ∎

Se resuelve entonces el programa **I** por el algoritmo SIMPLEX completado por una regla adicional que garantiza que en toda iteración se cumple $\mathbf{x}^{T}\mathbf{v} = 0$.

3.1.2: <u>REGLA ADICIONAL:</u>

Si en un cambio de base x_i [v_i] entra en la base, mientras que v_i [x_i] se mantiene (a nivel positivo) en la base, entonces este cambio de base no se ejecuta.

Si observando esta regla adicional ya no es posible ejecutar cambio de base, el algoritmo termina con una solución óptima de (3.5) con $z = 0$.

Vamos a demostrar esta última afirmación. Es decir el algoritmo de WOLFE no puede terminar con una solucion básica

$\hat{\mathbf{w}} = (\hat{\mathbf{x}}, \hat{\mathbf{v}}. \hat{\mathbf{u}}, \hat{z})$ con $\hat{z}$ $\quad 0$.

DEMOSTRACION

El algoritmo de WOLFE resuelve el programa **I**

$$\mathbf{I} : \min z$$

$$\begin{aligned} \mathbf{Ax} \qquad\qquad\qquad\qquad &= \mathbf{b} \\ -2\mathbf{Cx} + \mathbf{v} - \mathbf{A}^T\mathbf{u} + z\mathbf{h} &= \mathbf{p} \\ \mathbf{x} \geq 0, \ \mathbf{v} \geq 0, \quad z &\geq 0 \\ \mathbf{x}^T\mathbf{v} = 0 \end{aligned}$$

<u>HIPOTESIS:</u> El algoritmo de WOLFE termina con $\hat{\mathbf{w}} = (\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{u}}, \hat{z})$ con $\hat{z} > 0$.

Demostraremos al final que entonces $\hat{\mathbf{w}}$ es una solución óptima del programa lineal ($\tilde{\mathbf{I}}$)

$$(\tilde{\mathbf{I}}) \qquad \min z$$

$$\begin{aligned} \mathbf{A}x \qquad\qquad\qquad\qquad &= \mathbf{b} && \quad y \\ -2\mathbf{Cx} + \mathbf{v} - \mathbf{A}^T\mathbf{u} + z\mathbf{h} &= \mathbf{p} && \quad t \\ \hat{\mathbf{v}}^T\mathbf{x} + \hat{\mathbf{x}}^T\mathbf{v} \qquad\qquad &= 0 \\ \mathbf{x} \geq 0, \ \mathbf{v} \geq 0 \quad z &\geq 0 \end{aligned}$$

Consideramos el programa dual ($\tilde{\mathbf{I}}^*$) de ($\tilde{\mathbf{I}}$)

$$(\tilde{\mathbf{I}}^*) \qquad \max \mathbf{b}^T\mathbf{y} + \mathbf{p}^T\mathbf{t}$$

$$\begin{aligned} \mathbf{A}^T\mathbf{y} \quad 2\mathbf{C}\mathbf{t} + \xi\mathbf{v} &\leq 0 && \quad \mathbf{x} \\ + \xi\mathbf{x} &\leq 0 && \quad \mathbf{v} \\ \mathbf{At} \qquad &= 0 && \quad \mathbf{u} \\ \mathbf{h}^T\mathbf{t} \qquad &\leq 1 && \quad z \end{aligned}$$

$(\tilde{I}^*)$ tiene también una solución óptima $(\hat{y},\hat{t},\hat{\xi})$ con

$$b^T\hat{y} + p^T\hat{t} = \hat{z} > 0 \qquad (3.7)$$

Las restricciones en $(\tilde{I}^*)$ que correponden a componentes que son positivas, se cumplen en $(\hat{y},\hat{t},\hat{\xi})$ con =

$\hat{z} > 0$ implica $h^T\hat{t} = 1$ (3.8)

Para $i \leq i \leq n$, se presenta exactamente uno de los siguientes casos:

$\alpha)$ $\hat{x}_i > 0$ $\quad \hat{v}_i = 0 \Longrightarrow (A\hat{y} \quad 2C\hat{t})_i =$

$\beta)$ $\hat{x}_. = 0$ $\quad \hat{v}_i \quad 0 \Longrightarrow t_i =$

$\gamma)$ $\hat{x}_. = \hat{v}_i = 0 \quad \Longrightarrow (A^T\hat{y} \quad 2C\hat{t})_i \leq 0 \qquad t_i \leq$

de esto obtenemos:

$$\hat{t}^T(A^T\hat{y} \quad 2Ct) \geq 0$$

Como $At = 0$ esto significa que $(\hat{t})^T C\hat{t} \leq 0$

Como C es definida positiva, obtenemos de esto $\hat{t} = 0$.

Esto contradice a (3.8).

Por consiguiente la hipótesis $\hat{z} > 0$, conduce a una contradicción. Por consiguiente el algoritmo de WOLFE termina con una solución óptima de **I**.

<u>OBSERVACION</u> 3.1:

Si la matriz C solamente es definida semipositiva, entonces se cumple

$$\hat{z} > 0 \Longrightarrow (\hat{t})^T C\hat{t} \leq 0 \Longrightarrow (\hat{t})^T C\hat{t} = 0$$

$$Ct = 0 \quad [\ \psi(t) = \frac{1}{2}t^T Ct, \text{ toma su mínimo en } \hat{t}\ ]$$

Entonces de $z > 0$ se obtiene:

$$1 = h^T\hat{t} = (2C\bar{x} + p)^T\hat{t} = p^T\hat{t} + 2\bar{x}^T C\hat{t} = p^T\hat{t} \qquad (3.9)$$

En (α) se tiene ahora:

$(\alpha) \cdot x_i \quad 0 \qquad v_i = 0 \qquad (\mathbf{A}^T\hat{\mathbf{y}})_i = 0$

De esto y de $\mathbf{Ax} = \mathbf{b}$, se deduce que $\mathbf{b}^T\hat{\mathbf{y}} = \hat{\mathbf{x}}^T\mathbf{A}\hat{\mathbf{y}} \quad 0$

De esto, (3.9) y 3.7) se obtiene $1 = \mathbf{p}^T\hat{\mathbf{t}} = \hat{z}$

<u>CONSECUENCIA:</u>

Si $\mathbf{C}$ es definida semipositiva y el algoritmo de WOLFE termina en un punto $\hat{\mathbf{w}}$ con $\hat{z} > 0$, entonces $\hat{z} = 1$. Por consiguiente no ha sido posible rebajar el valor de z en absoluto [$z = 1$ es el valor inicial].

Este caso se presenta necesariamente si $F(\mathbf{x})$ no es acotada inferiormente en el dominio admisible S, del programa cuadrático original.

En este caso el rayo

$R = \{\hat{\mathbf{x}} + \lambda\mathbf{t} \ / \ \lambda > 0\} \subset S$, tal que $\mathbf{Ct} = 0$, $\mathbf{p}^T\tilde{\mathbf{t}} < 0$

Como en este caso $\inf\{F(\mathbf{x}) \ / \mathbf{x} \in S\} = -\infty$, no existe solución del sistema que representa las condiciones de K-T y el algoritmo de WOLFE debe terminar con un punto $\hat{\mathbf{w}}$ con $\hat{z} > 0$.

$(\Longrightarrow z = 1)$ ■

Por último veamos que $\hat{\mathbf{w}} = (\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{u}}, \hat{z})$ con $\hat{z} > 0$ es una solución óptima de (**I**).

$$
(\mathbf{I}) \qquad \min z
$$

$$
\begin{aligned}
\mathbf{Ax} \qquad\qquad\qquad\qquad &= \mathbf{b} \\
-2\mathbf{Cx} + \mathbf{v} - \mathbf{A}^T\mathbf{u} + z\mathbf{h} &= \mathbf{p} \\
\hat{\mathbf{v}}^T\mathbf{x} + \hat{\mathbf{x}}^T\mathbf{v} \qquad\qquad &= 0 \\
\mathbf{x} \geq 0, \quad \mathbf{v} \geq 0, \quad z &\geq 0
\end{aligned}
$$

Sea $\hat{\mathbf{w}} = (\hat{\mathbf{w}}_B, \hat{\mathbf{w}}_N)$ $\qquad \hat{\mathbf{w}}_B = (\hat{\mathbf{x}}_{B_1}, \hat{\mathbf{v}}_{B_2}, \hat{\mathbf{u}}_{B_3}, \hat{z})$

$$\hat{\mathbf{w}}_N = (\hat{\mathbf{x}}_{N_1}, \hat{\mathbf{v}}_{N_2}, \hat{\mathbf{u}}_{N_3})$$

Consideraremos sólo el caso no degenerado: $\mathbf{x}_B > 0$, $\mathbf{v}_{B2} > 0$

<u>AFIRMACION:</u>

$\hat{\mathbf{w}}$ es una solución básica de (I).

Sea $\mathbf{w}$ asociada a la base

$$\begin{bmatrix} \mathbf{A}_{B_1} & 0 & 0 & 0 \\ -2\mathbf{C}_{B_1} & & -\mathbf{A}_{B_3}^{\mathbf{T}} & \mathbf{h} \end{bmatrix} \quad \text{de } \mathbf{D}$$

juntando a $\mathbf{D}$ como última fila el vector $\mathbf{d}^{\mathbf{T}} = (\hat{\mathbf{v}}^{\mathbf{T}}, \hat{\mathbf{x}}^{\mathbf{T}}, 0^{\mathbf{T}}, 0)$, obtenemos la matriz de restriciones $\mathbf{D}$ de (I).

Observamos que : $\quad B_1 \subset N_2 \qquad B_2 \subset N_1$ $\qquad$ (3.10)

juntamos como última fila el vector $\qquad (\hat{\mathbf{v}}_{B_1}^{\mathbf{T}}, \hat{\mathbf{x}}_{B_2}^{\mathbf{T}}, 0^{\mathbf{T}}, 0)$.

Por (3.10) se cumple $\mathbf{d}_B = 0$, $\quad \mathbf{D}_B^* = \begin{bmatrix} \mathbf{D}_B \\ 0^{\mathbf{T}} \end{bmatrix}$.

Elegimos $j \in B_1$. Entonces $j \in N_2$ y a $\hat{v}_j = 0$ corresponde

la columna $\quad \begin{bmatrix} 0 \\ \mathbf{e}_j \\ \hat{\mathbf{x}}_j \end{bmatrix}$

Como x_j 0, obtenemos una base $\tilde{\mathbf{D}}_B$ de $\tilde{\mathbf{D}}$, juntando a esta columna. Se ve inmediatamente que $\mathbf{w}$ es la solución básica de (I) que es asociada a $\mathbf{D}_B$ ■

Supongamos ahora que $\hat{\mathbf{w}}$ no es óptimo para (**I**).

Entonces aun es posible realizar un cambio de base de (**I**), disminuyendo z. Pero no es posible que una variable u_i entre en la base, sino el algoritmo de WOLFE no hubiera terminado con $\hat{\mathbf{w}}$.

Asumimos entonces que : x_j , $j \in N_1$ [v_ℓ , $\ell \in N_2$] entra en la base. Si $\mathbf{w} = (\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \tilde{\mathbf{u}}, \tilde{z})$, $\tilde{z} < z$, es el nuevo vértice, entonces $\hat{\mathbf{v}}^T\tilde{\mathbf{x}} + \hat{\mathbf{x}}^T\tilde{\mathbf{v}} = 0$, esto implica que $v_j = 0$ [$\tilde{x}_\ell \quad 0$]

Si fuera $v_j > 0$ [$x_\ell > 0$], se cumpliría

$$\hat{v}_j \quad 0 \quad y \quad \hat{\mathbf{v}}^T\tilde{\mathbf{x}} + \hat{\mathbf{x}}^T\tilde{\mathbf{v}} \geq \hat{v}_j x_j > 0$$

$$[\ \hat{x}_\ell \quad 0 \quad y \quad \hat{\mathbf{v}}^T\tilde{\mathbf{x}} + \hat{\mathbf{x}}^T\tilde{\mathbf{v}} \geq \hat{x}_\ell \tilde{v}_\ell > 0\ ]$$

Pero $v_j = 0$ [$\tilde{x}_\ell = 0$] significa que el cambio de base que conduce de $\hat{\mathbf{w}}$ a $\tilde{\mathbf{w}}$ es un cambio de base legítimo para el algoritmo de WOLFE aplicado a **I**, lo que es absurdo. Por consiguiente $\hat{\mathbf{w}}$ es una solución óptima de **I**. ■

3.2: EJEMPLO PARA EL ALGORITMO DE WOLFE:

$$
\begin{aligned}
\min \quad & 2x_1^2 + x_2^2 \quad 48x_1 \quad 40x_2 \\
& x_1 + x_2 \leq \\
& x_1 \qquad\ \ \leq 6 \\
& x_1 + 3x_2 \leq 18 \\
& x_1 \ , \ x_2 \geq
\end{aligned}
$$

Introduciendo variables de holgura tenemos:

$$
\begin{aligned}
\min \quad & 2x_1^2 + \qquad - 48x_1 \quad 40x_2 \\
& x_1 + \qquad\qquad\qquad\qquad = 8 \\
& \qquad\qquad\qquad + x_4 \qquad = 6 \\
& x_1 + 3x_2 \qquad\qquad + x_5 = 18 \\
& \qquad\qquad\qquad\qquad x \geq 0, \quad i = 1, \ldots, 5
\end{aligned}
$$

El programa cuadrático convexo es de la forma

$$\min \{ F(\mathbf{x}) = \mathbf{x}^T \mathbf{C}\mathbf{x} + \mathbf{p}^T \mathbf{x} \ / \ \mathbf{A}\mathbf{x} = \mathbf{b} \ , \quad \mathbf{x} \geq 0 \}$$

$$\mathbf{C} = \begin{matrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix} \ , \quad \mathbf{p} = \begin{matrix} -48 \\ -40 \\ 0 \\ 0 \\ 0 \end{matrix} \Bigg| \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 8 \\ 6 \\ 18 \end{bmatrix}$$

Elegimos primero una solución básica admisible de

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq 0$$

$$\bar{\mathbf{x}} = (0, 0, 8, 6, 18)^T = (x_1, x_2, x_3, \bar{x}_4, \bar{x}_5)^T$$

Calculamos:

$$\mathbf{h}^T = (2\mathbf{C}\bar{\mathbf{x}} + \mathbf{p})^T = \mathbf{p}^T = (-48, -40, 0, 0, 0)^T \neq (0, 0, 0, 0, 0)^T$$

Consideramos el programa :

$$
\begin{aligned}
\mathrm{I} \qquad & \min z \\
& \mathbf{A}\mathbf{x} = \mathbf{b} \\
& -2\mathbf{C}\mathbf{x} + \mathbf{v} - \mathbf{A}^T\mathbf{u} + z\mathbf{h} = \mathbf{p} \\
& \mathbf{x}\ \mathbf{v} = 0 \\
& \mathbf{x} \geq 0, \quad \mathbf{v} \geq 0, \quad z \geq 0
\end{aligned}
$$

ó

I min z

$$
\begin{aligned}
& x_2 + x_3 && = 8 \\
& && = 6 \\
& && = 18 \\
x_1 & \qquad u_1 - u_2 - u_3 - 48z && = -48 \\
-4x_1 & \qquad u_1 \quad -3u_3 - 40z && = -40 \\
& \qquad u_1 && = 0 \\
& \qquad u_2 && = 0 \\
& \qquad - u_3 && = 0
\end{aligned}
\quad \Bigg\} \; *
$$

$$+ x_2v_2 + \qquad + x_4v_4 + x_5v_5 = 0$$

$$x_i \geq 0 \qquad v_i \geq 0 \qquad i = 1,\ldots,5 \ , \qquad z \geq 0$$

De (*) podemos eliminar las variables u_i $\quad i = 1,2,3$

Obtenemos el programa lineal

$$\mathrm{I}^* \qquad \max \ (-z$$

$$\mathbf{Dw} = \mathbf{b}$$

$$\mathbf{w} \geq 0$$

que tiene la forma:

max $\quad -z)$

$$
\begin{aligned}
& x_1 + x_2 + x_3 && = \\
& x_1 \qquad + x_4 && = 6 \\
& x_1 + 3x_2 \qquad + x_5 && = 18 \\
-4x_1 & \qquad + v_1 \qquad v_3 - v_4 - v_5 \quad 48z && = \quad 48 \\
& -2x_2 \qquad + v_2 \quad v_3 \qquad -3v_5 \quad 40z && = \quad 40
\end{aligned}
$$

$$x_i \geq 0 \qquad v_i \geq 0 \qquad i = 1,\ldots,5 \qquad z \geq 0$$

Tabla de datos

	x_1	x_2	x_3	x_4	x_5	v_1	v_2	v_3	v_4	v_5	z	
0	0	0	0	0	0	0	0	0	0	0	1	
8	1	1	1	0	0	0	0	0	0	0	0	x_3
6	1	0	0	1	0	0	0	0	0	0	0	x_4
18	1	3	0	0	1	0	0	0	0	0	0	x_5
-48	-4	0	0	0	0	1	0	-1	-1	-1	-48	v_1
-40	0	-2	0	0	0	0	1	-1	0	-3	-40	v_2

Se cumple $\mathbf{x}^T\mathbf{v} = 0$

<u>CAMBIO DE BASE:</u>

Para efectuar el algoritmo de WOLFE es necesario que z sea una variable básica

(Hagamos que z entre en la base y que v_2 salga de la base)

	x_1	x_2 ↓	x_3	x_4	x_5	v_1	v_2	v_3	v_4	v_5	z	
-1	0	$\frac{-1}{20}$	0	0	0	0	$\frac{1}{40}$	$\frac{-1}{40}$	0	$\frac{-3}{40}$	0	
8	1	1	1	0	0	0	0	0	0	0	0	x_3
6	1	0	0	1	0	0	0	0	0	0	0	x_4
18	1	3	0	0	1	0	0	0	0	0	0	x_5
→ 0	-4	$\boxed{\frac{12}{5}}$	0	0	0	1	$\frac{-6}{5}$	$\frac{1}{5}$	-1	$\frac{13}{5}$	0	v_1
1	0	$\frac{1}{20}$	0	0	0	0	$\frac{-1}{40}$	$\frac{1}{40}$	0	$\frac{3}{40}$	1	z

Se cumple $\mathbf{x}^T\mathbf{v} = 0$ pues $\mathbf{v} = 0$

<u>Cambio de base:</u> v_3 , v_4 no pueden entrar pues $x_3 > 0$, $x_4 > 0$ x_2 si puede entrar en la base, pues $v_2 = 0$ (no está en la base) sale v_1 de la base.

Nueva tabla:

	x_1	x_2	x_3	x_4	x_5	v_1	v_2	v_3	v_4	v_5	z	
-1	$\frac{-1}{12}$	0	0	0	0	$\frac{1}{48}$	0	$\frac{-1}{48}$	$\frac{-1}{48}$	$\frac{-1}{48}$	0	
8	$\boxed{\frac{8}{3}}$	0	1	0	0	$\frac{-5}{12}$	$\frac{1}{2}$	$\frac{-1}{12}$	$\frac{5}{12}$	$\frac{-13}{12}$	0	x_3
6	1	0	0	1	0	0	0	0	0	0	0	x_4
18	6	0	0	0	1	$\frac{-5}{4}$	$\frac{3}{2}$	$\frac{-1}{4}$	$\frac{15}{12}$	$\frac{-39}{12}$	0	x_5
0	$\frac{-5}{3}$	1	0	0	0	$\frac{5}{12}$	$\frac{-1}{2}$	$\frac{1}{12}$	$\frac{-5}{12}$	$\frac{13}{12}$	0	x_2
1	$\frac{1}{12}$	0	0	0	0	$\frac{-1}{48}$	0	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{48}$	1	z

Se cumple : $\mathbf{x}^T\mathbf{v} = 0$

<u>Cambio de base:</u> v_3, v_4, v_5, no pueden entrar en la base pues $x_3 > 0$, $x_4 > 0$, $x_5 > 0$. x_1 si puede netrar en la base pues $v_1 = 0$ no está en la base. Sale x_3 de la base (x_5 tambien puede) salir de la base.

Nueva tabla:

	x_1	x_2	x_3	x_4	x_5	v_1	v_2	v_3	v_4	v_5	z	
$\frac{-3}{4}$	0	0	$\frac{1}{32}$	0	0	$\frac{1}{128}$	$\frac{1}{64}$	$\frac{-3}{128}$	$\frac{-1}{128}$	$\frac{-7}{128}$	0	
3	1	0	$\frac{3}{8}$	0	0	$\frac{-5}{32}$	$\frac{3}{16}$	$\frac{-1}{32}$	$\frac{5}{32}$	$\frac{-13}{32}$	0	x_1
3	0	0	$\frac{-3}{8}$	1	0	$\frac{5}{32}$	$\frac{-3}{16}$	$\frac{1}{32}$	$\frac{-5}{32}$	$\frac{13}{32}$	0	x_4
0	0	0	$\frac{-9}{4}$	0	1	$\frac{-5}{16}$	$\frac{3}{8}$	$\frac{-1}{16}$	$\frac{5}{16}$	$\frac{-13}{16}$	0	x_5
5	0	1	$\frac{5}{8}$	0	0	$\frac{5}{32}$	$\frac{-3}{16}$	$\frac{1}{32}$	$\frac{-5}{32}$	$\frac{55}{96}$	0	x_2
$\frac{3}{4}$	0	0	$\frac{-1}{32}$	0	0	$\frac{-1}{128}$	$\frac{-1}{64}$	$\boxed{\frac{3}{128}}$	$\frac{1}{128}$	$\frac{7}{128}$	1	z

Se cumple $\mathbf{x}^T\mathbf{v} = 0$

Cambio de base: v_4 no puede entrar en la base pues x_4 0

v_3 si puede entrar en la base pues $x_3 = 0$ (no está en la base) Sale z de la base.

Nueba tabla:

	x_1	x_2	x_3	x_4	x_5	v_1	v_2	v_3	v_4	v_5	z	
0	0	0	0	0	0	0	0	0	0	0	1	
4	1	0	$\frac{1}{3}$	0	0	$\frac{-1}{6}$	$\frac{1}{6}$	0	$\frac{1}{12}$	$\frac{-1}{3}$	$\frac{4}{3}$	x_1
2	0	0	$\frac{-1}{3}$	1	0	$\frac{1}{6}$	$\frac{-1}{6}$	0	$\frac{-1}{6}$	$\frac{1}{3}$	$\frac{-4}{3}$	x_4
2	0	0	$\frac{-7}{3}$	0	1	$\frac{-1}{3}$	$\frac{1}{3}$	0	$\frac{-1}{2}$	$\frac{-2}{3}$	$\frac{8}{3}$	x_5
4	0	1	$\frac{2}{3}$	0	0	$\frac{1}{6}$	$\frac{-1}{6}$	0	$\frac{-1}{6}$	$\frac{1}{2}$	$\frac{-4}{3}$	x_2
32	0	0	$\frac{-4}{3}$	0	0	$\frac{-1}{3}$	$\frac{-2}{3}$	1	$\frac{1}{3}$	$\frac{7}{3}$	$\frac{128}{3}$	v_3

Esta tabla es óptima $T_{oj} \geq 0$.

$\hat{\mathbf{x}}^T = (4,4,0,2,2)$ $\hat{\mathbf{v}}^T = (0,0,32,0,0)$

$\hat{\mathbf{x}}^T\hat{\mathbf{v}} = 0$

$\inf(P) = F(\mathbf{x}) = -304.$

CAPITULO IV

4.1: EL ALGORITMO DE BEALE

Consideremos el programa cuadrático

$$\mathbf{P} \qquad \min\{\ F(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x}\ /\ \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$$

con $\mathbf{C} = \mathbf{C}^T \in M(n,n)$ es definida semipositiva; y

$\mathbf{A} \in M(m,n) \quad ran(\mathbf{A}) = n$

Este algoritmo trabaja con soluciones básicas de sistemas lineales que se deducen de $\mathbf{A}\mathbf{x} \leq \mathbf{b}$.

Hacemos la hipótesis que todas las soluciones básicas que el algoritmo construye, sean *NO DEGENERADAS*; Y S será el dominio admisible del programa **P**

LA ITERACION INICIAL:

El algoritmo parte de una base admisible $\mathbf{A}_B$ de **A** y de la solución básica correspondiente $\mathbf{x}^o = \mathbf{x}_B$.

Como $\mathbf{x}_B$ es no degenerada se cumple :

$$\mathbf{A}_B\mathbf{x}_B = \mathbf{b}_B \qquad \mathbf{A}_N\mathbf{x}_B < \mathbf{b}_N \tag{4.1}$$

Sea $\mathbf{A}^T = [a_1, a_2, \ldots, a_m]$ y $-\mathbf{A}_B^{-1} = [\beta_1, \beta_2, \ldots, \beta_n]$ y $\mathbf{B} = \{i_1, i_2, \ldots, i_n\}$ el conjunto de índices básicos.

Como en el algoritmo simplex se determina una dirección S^o tal que :

$$\mathbf{a}_{i_r}^T S^o = -1 \qquad \text{para un } i_r \in \mathbf{B}$$

$$\mathbf{a}_j^T S^o = 0 \qquad \forall\, j \in \mathbf{B} \setminus \{ i_r \}$$

Ahora si $\mathbf{x}^o$ no es óptimo, entonces debe existir tal dirección S^o.

Se cumple

$$\mathbf{a}_r^T \beta_r = -1 \qquad \mathbf{a}_j^T \beta_r = 0 \qquad \forall\, j \quad \mathbf{B} \setminus \{ i_r \} \tag{4.2}$$

$$\min_{1 \le r \le n} \nabla F(\mathbf{x}^o)^T \beta_r \quad 0 \qquad 1 \le r \le n$$

DEMOSTRACION (de 4.2)

Si $\mathbf{x}$ es admisible para $\mathbf{P}$ entonces

$$\mathbf{A}_B(\mathbf{x} - \mathbf{x}^o) = -\mathbf{A}_B\mathbf{x} + \mathbf{A}_B\mathbf{x}^o = -\mathbf{A}_B\mathbf{x} + \mathbf{b}_B = \mathbf{u} \ge 0$$

ó

$$\mathbf{x} - \mathbf{x} = \mathbf{A}_B^{-1}\mathbf{u} = \sum_{j=1}^{n} u_j \beta_j \qquad u_j \ge 0 \,, \quad 1 \le j \le n$$

Si fuera : $\nabla F(\mathbf{x}^o)^T \beta_r \ge 0 \quad 1 \le r \le n$, entonces se cumpliría por la convexidad de F (la ecuación de soporte)

$$F(\mathbf{x}) - F(\mathbf{x}^o) \ge \nabla F(\mathbf{x}^o)^T (\mathbf{x} - \mathbf{x}^o) = \sum_{r=1}^{n} \mathbf{u}_r \nabla F(\mathbf{x}^o)^T \beta_r \ge 0$$

$$F(\mathbf{x}) \ge F(\mathbf{x}^o).$$

y así $\mathbf{x}^o$ fuera una solución óptima de $\mathbf{P}$ ■

Si $\nabla F(\mathbf{x}^o)^T\beta_r = \min\{\nabla F(\mathbf{x}^o)^T\beta_j \ / \ 1 \le j \le n\}$, entonces elegimos la dirección $S^o = \beta_r$.

Sea $g(\lambda) = F(\mathbf{x}^o + \lambda S^o)$

si $\lambda^* = \max\{\lambda \ge 0 \ / \ \mathbf{x}^o + \lambda S^o \in S\}$ y λ^{**} es la solución de $\min\{g(\lambda) \ / \ \lambda \ge 0\}$, entonces ponemos $\lambda = \min\{\lambda^*, \lambda^{**}\}$.

Por (4.1) y (4.2) se cumple $\lambda^* > 0$. tambien se debe cumplir que : $g'(\lambda^{**}) = 0$.

Como $g'(0) = \nabla F(\mathbf{x}^o)^T S^o \quad 0$; se tiene $\lambda^{**} > 0$.

Si $\lambda = +\infty$, entonces $\inf(\mathbf{P}) = -\infty$ y $\mathbf{P}$ no tiene solución óptima.

En este caso el rayo $\mathbf{x} = \mathbf{x}^o + \lambda S^o$, $\lambda \ge 0$, se encuentra en S y se cumple

$$\begin{aligned} g'(\lambda) &= \nabla F(\mathbf{x}^o + \lambda S^o)^T S^o \\ &= [2C(\mathbf{x}^o + \lambda S^o) + \mathbf{p}]^T S^o \\ &= \nabla F(\mathbf{x}^o)^T S^o + 2\lambda (S^o)^T C S^o < 0, \quad \forall \lambda \ge 0 \end{aligned}$$

De esto y de $(S^o)^T C S^o \ge 0$, se deduce que

$(S^o)^T C \mathbf{x}^o = 0$ y $g'(\lambda) = \nabla F(\mathbf{x}^o)^T S^o = g'(0) \quad 0 \quad \forall \lambda \ge 0$

Esto implica finalmente que :

$$\lim_{\lambda \to \infty} F(\mathbf{x}^o + \lambda S^o) = \lim_{\lambda \to \infty} g(\lambda) = -\infty$$

<u>Si $0 < \lambda < \infty$</u>, elegimos $\mathbf{x}^1 = \mathbf{x}^o + \bar{\lambda} S^o$

De $g'(\lambda) = \nabla F(\mathbf{x}^o + \lambda S^o)^T S^o < 0 \quad \forall \lambda \in [0, \lambda)$

se deduce que $F(\mathbf{x}^1) = g(\lambda) \quad g(0) = F(\mathbf{x}^o)$ ∎

$$\left[\text{pues } \underset{F(\mathbf{x}^1)}{\underbrace{g(\bar{\lambda})}} \quad \underset{F(\mathbf{x}^o)}{\underbrace{g(0)}} = \int_0^\lambda g'(\lambda)\, d\lambda < 0 \right]$$

Tenemos que distinguir dos casos:

a) Si $\lambda^{*} = \bar{\lambda}$. En este caso $\mathbf{x}^1$ es el punto que genera el algoritmo simplex, si parte de $\mathbf{x}^{\bullet}$ y avanza a lo largo de la arista con dirección S^0. Por consiguiente $\mathbf{x}^1$ es solución básica admisible (y no degenerada) en este caso se puede ejecutar una $2^{\underline{a}}$ iteración.

b) Si λ -- λ^{**} $\quad \lambda^{*}$ entonces se cumple

$$
\begin{array}{lll}
a_j^T \mathbf{x}^1 = \mathbf{a}_j^T(\mathbf{x}^0 + \bar{\lambda} S^0) = b_j & \forall\ j \quad B \setminus \{ i_r \} & \\
\mathbf{a}_j \mathbf{x}^1 \quad b_j & \forall\ j \in N \quad \{ i_r \} & (4.3)
\end{array}
$$

Por consiguiente $\mathbf{x}^1$ no es solución básica. Sin embargo $\mathbf{x}^1$ satisface adicionalmente la ecuación

$$(S^0)^T \nabla F(\mathbf{x}) = 0 \qquad\qquad 2(S^0)^T \mathbf{C}\mathbf{x} = -(S^0)^T \mathbf{p}$$

$$\left[\begin{array}{l} \text{pues como } g'(\bar{\lambda}) = 0 = g'(\lambda^{**}) = \nabla F(\mathbf{x}^0 + \bar{\lambda} S^0)^T S^0 \text{ --} \\ \qquad\qquad\qquad\qquad\qquad = (S^0)^T \nabla F(\mathbf{x}^1) \end{array} \right]$$

Añadimos esta ecuación al sistema $\mathbf{Ax} \leq \mathbf{b}$ y la escribimos como $\mathbf{C}_1^T \mathbf{x} = \mathbf{d}_1$ con $\mathbf{C}_1^T = 2(S^0)^T \mathbf{C}$; $\mathbf{d}_1 = \quad (S^0)^T \mathbf{p}$.

Los vectores $\mathbf{C}_1$ y $\mathbf{a}_j \quad j \in \mathbf{B} \setminus \{ \quad$ son independientes.

Pues de:

$$\sum_{j \in \mathbf{B} \setminus \{ i_r \}} \mathbf{u}_j \mathbf{a}_j + \mathbf{u}\, \mathbf{C}_1 = 0 \qquad \text{se deduce:}$$

$$0 = \sum_{j \in \mathbf{B} \setminus \{ i_r \}} \mathbf{u}_j \mathbf{a}_j^T S^0 + 2\mathbf{u}(S^0)^T \mathbf{C} S^0 = 2\mathbf{u}(S^0)^T \mathbf{C} S^0$$

ó $2\mathbf{u}(S^0)^T \mathbf{C} S^0 = 0$ De esto obtenemos que $\mathbf{u} = 0$.

Pues si $(S^0)^T \mathbf{C} S^0 = 0$ entonces $\psi(S)$ -- $S^T \mathbf{C} S$, tomaría mínimo en S^0 y se cumpliría $2\mathbf{C}S^0 = 0$, esto implicaría que

$CS^o = 0$ y

$$(S^o)^T \nabla F(x^o) = (S^o)^T(p + 2Cx^o) = (S^o)^T p = 0$$

$$(S^o)^T \nabla F(x') = (S^o)^T \left[p + 2C(x^o + \bar{\lambda} S^o) \right] = (S^o)^T p = 0$$

lo que es absurdo. Luego $(S^o)^T CS^o > 0$.

Entonces de $u = 0$, se obtiene inmediatamente

$$u_j = 0 \quad \forall\, j \in B \setminus \{ i_r \}$$

por esta razón y por (4.3), x^1 es una solución básica admisible para el sistema ampliado de restricciones

$$\mathbf{Ax} \leq \mathbf{b} \qquad \mathbf{c}_1^T\mathbf{x} = \mathbf{d}_1$$

LA ITERACION GENERAL

Asumimos que en el k-ésimo punto de iteración $\mathbf{x}^k$ el sistema de restricciones consiste en $\mathbf{Ax} \leq \mathbf{b}$ y en g restricciones $\mathbf{c}_i^T\mathbf{x} = \mathbf{d}_i$, que se generan en el transcurso del algoritmo y que $\mathbf{x}^k$ es una solución básica admisible para este sistema.

Entonces despues de una re-enumeración eventual se cumple:

$$\mathbf{c}_i^T\mathbf{x}^k = \mathbf{d}_i \qquad i = 1, 2, \ldots, g$$

$$\mathbf{a}_i^T\mathbf{x}^k = \mathbf{b}_j \qquad j = 1, 2, \ldots, n-g$$

$$\mathbf{a}_j^T\mathbf{x}^k < \mathbf{b}_j \qquad j = n-g+1, \ldots, m$$

Esto, en forma global escribimos como:

$$\tilde{\mathbf{A}}_B\mathbf{x}^k = \tilde{\mathbf{b}}_B$$

Escribimos otra vez

$$\tilde{\mathbf{A}}^T = [\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_{m+g}] \qquad [\tilde{a}_{i_1}, \tilde{a}_{i_2}, \ldots, \tilde{a}_{i_n}]$$

$$\tilde{\mathbf{A}}_B^{-1} = [\beta_1, \beta_2, \ldots, \beta_n]$$

Si $\mathbf{x}^k$ no es óptimo para **P**, entonces existe otra vez dirección S^k que no satisface exactamente una de las ecuacio

$\tilde{\mathbf{A}}_B S^k = 0$ y para el cual $(S^k)^T \nabla F(\mathbf{x}^k) < 0$.

Evidentemente se cumple :

$$\tilde{\mathbf{a}}_i^T \beta_r = -1 \qquad \tilde{\mathbf{a}}_j^T \beta_r = 0 \qquad \forall\ j \in B \setminus \{ i_r \} \qquad \cdots n$$

Supongamos que se cumpla

$$\left.\begin{array}{ll} \beta_r^T \nabla F(\mathbf{x}^k) = 0 & r = 1,2,\ldots,g \\ \beta_r^T \nabla F(\mathbf{x}^k) \geq 0 & r = g+1,\ldots,n \end{array}\right\} \qquad (4.4)$$

Sea $\mathbf{x} \in S$ cualquiera. Entonces se cumple

$$\tilde{\mathbf{A}}_B(\mathbf{x} - \mathbf{x}^k) = \mathbf{u} \qquad u_j \geq 0 \qquad j = g+1,\ldots,n$$

ó $$\mathbf{x} - \mathbf{x}^k = -\tilde{\mathbf{A}}_B^{-1}\mathbf{u} = \sum_{j=1} u_j \beta_j \qquad u_j \geq 0\ ,\ j = g+1,\ldots,n$$

De esto y de (4.4) obtenemos (de la inecuación de soporte)

$$F(\mathbf{x}) - F(\mathbf{x}^k) \geq \nabla F(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) = \sum_{j=1}^{n} u_r \beta_r^T \nabla F(\mathbf{x}^k) \geq 0$$

$$F(\mathbf{x}) \geq F(\mathbf{x}^k)$$

Esto significa que $\mathbf{x}^k$ es una solución óptima de **P**, lo que es un absurdo. ■

<u>Elegimos S^k como sigue:</u>

Si $\max \{ |\beta_j^T \nabla F(\mathbf{x}^k)| \ /\ 1 \leq j \leq g \} = |\beta_r^T \nabla F(\mathbf{x}^k)| > 0$,

entonces elegimos $S^k = \pm\, \beta_r$ (si $\beta_r^T \nabla F(\mathbf{x}^k) > 0 \Longrightarrow S^k = -\beta_r$)

Si $\beta_j^T \nabla F(\mathbf{x}^k) = 0 \qquad j = 1,2,\ldots,g$, entonces elegimos

$S^k = \beta_r$. donde:

$$\beta_r^T \nabla F(\mathbf{x}^k) = \min \{ \beta_j^T \nabla F(\mathbf{x}^k) \ /\ g+1 \leq \quad \leq n \}$$

Sea $g_k(\lambda) = F(\mathbf{x} + \lambda S^k)$ $\quad \lambda^* = \max \{ \lambda \geq 0 \ /\ \mathbf{x}^k + \lambda S^k \in S \}$ y

λ^{**} es a solución óptima de min $\{g_k(\lambda) \quad \lambda \geq 0\}$ y

$\bar{\lambda} = \min \{\lambda^*, \lambda^{**}\}$. Entonces $0 < \bar{\lambda} \leq \infty$, puesto que

$\nabla F(\mathbf{x}^k)^T S^k = g'_k(0) < 0$ y $\mathbf{x}^k$ es no degenerado por hipótesis.

Si $\bar{\lambda} = +\infty$ entonces $\inf(P) = -\infty$ y $\mathbf{P}$ no tiene solución óptima. En este caso el rayo $\mathbf{x}^k + \lambda S^k \quad \lambda \geq 0$ se encuentra en S y se cumple :

$$g'_k(\lambda) = \nabla F(\mathbf{x}^k + \lambda S^k)^T S^k = [\, 2\mathbf{C}(\mathbf{x}^k + \lambda S^k) + \mathbf{p}\,]^T S^k$$
$$= \nabla F(\mathbf{x}^k)^T S^k + 2\lambda (S^k)^T \mathbf{C} S^k < 0 \quad , \quad \forall\, \lambda \geq 0$$

De esto y de $(S^k)^T \mathbf{C} S^k \geq 0$, se deduce que $(S^k)^T \mathbf{C}\mathbf{x}^k = 0$ y

$$g'_k(\lambda) = \nabla F(S^k)^T S^k < 0 \qquad \lambda \geq 0$$

Esto implica finalmente que :

$$\lim_{\lambda \to \infty} F(\mathbf{x}^k + \lambda S^k) = \lim_{\lambda \to \infty} g_k(\lambda) = -\infty$$

Si $0 < \bar{\lambda} \underline{< \infty}$, definimos $\mathbf{x}^{k+1} = \mathbf{x}^k + \bar{\lambda} S^k$.

Como $g'_k(\lambda) < 0 \quad , \forall\, \lambda \in [0, \bar{\lambda})$, se cumple :

$$F(\mathbf{x}^{k+1}) = g_k(\bar{\lambda}) < g_k(0) = F(\mathbf{x}^k)$$

<u>Tenemos otra vez que distinguir dos casos:</u>

a) $\bar{\lambda} = \lambda^*$ Entonces para un $s \in N$ se cumple

$$\tilde{\mathbf{a}}_j^T \mathbf{x}^{k+1} = \mathbf{b}_j \qquad \forall\, j \in B \cup \{s\} \setminus \{i_r\}$$

$$\tilde{\mathbf{a}}_j^T \mathbf{x}^{k+1} < \mathbf{b}_j \qquad \forall\, j \in N \setminus \{s\}$$

(4.5) $$\tilde{\mathbf{a}}_i^T \mathbf{x}^{k+1} \quad \mathbf{b}_{i_r} \qquad \text{si } \mathbf{a}_{i_r} \quad \mathbf{a}_i, \text{ para algun}$$

$$\tilde{\mathbf{a}}_i^T \mathbf{x}^{k+1} \neq \mathbf{b}_{i_r} \qquad \text{si } \mathbf{a}_{i_r} = \mathbf{C}_i, \text{ para algun } i$$

Si $\tilde{\mathbf{a}}_{i_r} = \mathbf{C}_i$, entonces se omite la restricción $\mathbf{C}_i^T \mathbf{x} = \mathbf{d}_i$

Por consiguiente al final de la k-ésima iteración el sistema de restricciones consiste en $\mathbf{Ax} \leq \mathbf{b}$ y en g ó g-1 restricciones adicionales $\mathbf{C}_j^T\mathbf{x} = \mathbf{d}_j$ y por (4.5) $\mathbf{x}^{k+1}$ es admisible para este sistema.

Los vectores $\tilde{\mathbf{a}}_j$ $j \in B \cup \{s\} \setminus \{i_r\}$ forman una base de este sistema de restricciones. Ya que de

$$\sum_{j\in B\setminus\{i_r\}} \mathbf{u}_j \mathbf{a}_j + \mathbf{u}\,\mathbf{a}_s = 0 \quad \text{se deduce}$$

$$\sum_{j\in B\setminus\{i_r\}} \mathbf{u}_j \tilde{\mathbf{a}}_j^T S^k \quad \mathbf{u}\,\tilde{\mathbf{a}}_s^T S^k = \mathbf{u}\mathbf{a}_s^T S^k = 0$$

Como $\lambda = \lambda^*$, se cumple $\tilde{\mathbf{a}}_s^T S^k > 0$ pues

$$\mathbf{a}_s^T x^k < \mathbf{b}_s \quad y \quad \mathbf{a}_s^T(\mathbf{x}^k + \bar{\lambda} S^k) = \mathbf{b}_s$$

Por consiguiente $\mathbf{u} = 0$, lo que a su vez implica que

$$u_j = 0 \quad \forall\, j \in B \setminus \{$$

En consecuencia en $\mathbf{x}^{k+1}$ encontramos la misma situación que en $\mathbf{x}^k$ y el algoritmo puede continuar.

b) $\lambda = \lambda^{**}$ λ^*. Entonces se cumple :

$$
\begin{array}{lll}
 & \tilde{\mathbf{a}}_j^T \mathbf{x}^{k+1} = \tilde{b}_j & \forall\, j \in B \quad \{i_r\} \\
 & \tilde{\mathbf{a}}_j^T \mathbf{x}^{k+1} < \tilde{b}_j & \forall\, j \quad N \\
(4.6) & \tilde{\mathbf{a}}_i^T\, \mathbf{x}^{k+1} < \tilde{b}_{i_r} & \text{si } \mathbf{a}_{i_r} = \mathbf{a}_i\text{, para algun } i \\
 & \tilde{\mathbf{a}}_{i_r}^T\, \mathbf{x}^{k+1} \neq \tilde{b}_{i_r} & \text{si } \tilde{\mathbf{a}}_{i_r} = C_i\text{, para algun } i
\end{array}
$$

$$g_k'(\bar{\lambda}) = (S^k)^T \nabla F(\mathbf{x}^{k+1}) = 0 \quad ó \quad C_{g+1}^T \mathbf{x}^{k+1} \quad d_{g+1}$$

Si $\mathbf{a}_i = C_i$, entonces la restricción $C_i^T\mathbf{x} = d_i$ es omitida

Se añade la restricción $C_{g+1}^{T}\mathbf{x} = d_{g+1}$ consiguiente al final de la k-ésima iteración, el sistema de restricciones consiste en $\mathbf{Ax} \leq \mathbf{b}$ y en g ó g+1 restricciones adicionales $C_i^{T}\mathbf{x} = d_i$.

Por (4.6) $\mathbf{x}^{k+1}$ es admisible para este sistema.

Los vectores $\mathbf{a}_j$ $j \in B \setminus \{ i_r \}$ y $C_{g+1} = 2CS^k$ forman una base para este nuevo sistema de restricciones. Pues

$$\sum_{j \in B\setminus\{i_r\}} u_j \tilde{\mathbf{a}}_j + \mathbf{u}(2CS^k) = 0$$

de donde se deduce:

$$0 = \sum_{j \in B\setminus\{i_r\}} u_j (S^k)^T \tilde{\mathbf{a}}_j + 2\mathbf{u}(S^k)^T CS^k = 2\mathbf{u}(S^k)^T CS^k$$

Se debe cumplir que $\mathbf{u} = 0$, pues si $(S^k)^T CS^k = 0$, se obtendría que S^k minimiza $\psi(S) = S^T CS$ y que $Cs^k = 0$

De esto se obtendría:

$$(S^k)^T \nabla F(\mathbf{x}^k) = (S^k)^T (2C\mathbf{x}^k + \mathbf{p}) = (S^k)^T \mathbf{p} < 0 \quad \text{y}$$

$$0 = (S^k)^T \nabla F(\mathbf{x}^{k+1}) = (S^k)^T [2C(\mathbf{x}^k + \bar{\lambda} S^k) + \mathbf{p}]$$

$$= 2(S^k)^T \mathbf{x}^k + 2\bar{\lambda}(S^k)^T CS^k + (S^k)^T \mathbf{p} = (S^k)^T \mathbf{p}$$

lo que es absurdo. Luego $(S^k)^T CS^k > 0$.

De $\mathbf{u} = 0$, se obtiene inmediatamente $u_j = 0$, $\forall\, j \in B\setminus\{i_r\}$ Por consiguiente $\mathbf{x}^{k+1}$ es una solución básica admisible (y no degenerada); por ésta razón es posible continuar con el algoritmo.

4.2: EL ALGORITMO DE BEALE (RESUMEN)

$$P \qquad \min \{ \mathbf{x}^T C\mathbf{x} + \mathbf{p}^T \mathbf{x} \ / \ \mathbf{A}\mathbf{x} \leq$$

Notación: $\quad - \ \mathbf{A}_B^{-1} = [\beta_1, \beta_2, \dots, \beta_n]$

$$S = \{ \mathbf{x} \ / \ \mathbf{A}\mathbf{x} \leq \mathbf{b} \} \qquad F(\mathbf{x}) = \mathbf{x}^T C\mathbf{x} + \mathbf{p}^T \mathbf{x}$$

PASO INICIAL

a) Determinar una solución básica $\mathbf{x}_B = \mathbf{A}_B^{-1} \mathbf{b}_B$ de $\mathbf{A}\mathbf{x} \leq \mathbf{b}$

$\mathbf{x}^o = \mathbf{x}_B$.

b) Calcular

$\min \{ \nabla F(\mathbf{x}^{\bullet})^T \beta_j \ / \ 1 \leq j \leq n \} = \nabla F(\mathbf{x}^o)^T \beta_r$

Si $\nabla F(\mathbf{x}^o)^T \beta_r \geq 0$, entonces $\mathbf{x}^o$ es óptimo . STOP

En otro caso se eiige $S^o = \beta_r$

c) Determinar :

$\lambda^* = \max \{ \lambda \geq 0 \ / \ \mathbf{x}^o + \lambda S^o \in S \}$

y la solución óptima λ^{**} de $\min \{ g(\lambda) = F(\mathbf{x}^o + \lambda S^o) / \ \lambda \geq 0 \}$

$\lambda = \min \{ \lambda^*, \lambda^{**} \}$

d) Si $\lambda = \infty$, entonces $\inf(\mathbf{P}) = -\infty$. STOP

En otro caso:

Si $\lambda = \lambda^*$, entonces $\mathbf{x}^1 = \mathbf{x}^o + \bar{\lambda} S^o$ es una solución básica admisible de $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. Ir al paso **b** con $\mathbf{x}^1$

Si $\lambda = \lambda^* < \lambda^{**}$ se introduce la nueva restricción

$2(S^o)^T C\mathbf{x} = -(S^o)^T \mathbf{p}$ o $C_1^T \mathbf{x} = \mathbf{d}_1$.

Entonces $\mathbf{x}^1$ es una solución básica admisible de

$$\mathbf{A}\mathbf{x} \leq \mathbf{b} \qquad C_1^T \mathbf{x} = \mathbf{d}_1.$$

Ir al paso **b**.

ITERACION GENERAL

Restricciones en $\mathbf{x}^k$

$\mathbf{Ax} \leq \mathbf{b} \qquad C_j^T\mathbf{x} = \mathbf{d}_j \qquad\qquad j = 1, 2, \ldots, g$

Base correspondiente a $\mathbf{x}^k$

$\tilde{\mathbf{A}}_B^T = [c_1, c_2, \ldots, c_g, a_1, a_2, \ldots, a_{n-g}]$

a) Determinar

$\alpha_k = \min \{ |\nabla F(\mathbf{x}^k)\beta_j| \ / \ 1 \leq j \leq g\} = |\nabla F(\mathbf{x}^k)^T\beta_r|$

Si $\alpha_k > 0$, se elige

$S_k = \beta_r \qquad\qquad$ si $\nabla F(\mathbf{x}^k)^T\beta_r$

$S_k = -\beta_r \qquad\qquad$ en otro caso

Si $\alpha_k = 0 \qquad\qquad$ se determina

$\gamma_k = \min \{ \nabla F(\mathbf{x}^k)^T\beta_j \ / \ g+1 \leq j \leq n\} = \nabla F(\mathbf{x}^k)^T\beta_t$

Si $\gamma_k \geq 0$ entonces $\mathbf{x}^k$ es óptimo

Si $\gamma_k < 0$ entonces $S_k = \beta_t$

b) Determinar

$\lambda^* = \max \{\lambda \geq 0 \ / \ \mathbf{x}^k + \lambda S^k \in S\}$

y la solución óptima λ^{**} de

$\min \{g_k(\lambda) = F(\mathbf{x}^k + \lambda S^k) \ / \ \lambda \geq 0\}$

$\lambda = \min \{\lambda^*, \lambda^{**}\}$

Si $\lambda = \infty$ entonces $\inf(\mathbf{P}) = -\infty$. STOP

∞ entonces $\mathbf{x}^{k+1} = \mathbf{x}^k + \bar{\lambda} S^k$

c) Si $\lambda = \lambda^*$, entonces $\mathbf{x}^{k+1}$ es una solución básica admisible del sistema actual de restricciones. Ir al paso k+1

Si $\lambda = \lambda^{**}$ se introduce la nueva restricción

$2(S^k)^T Cx = (S^k)^T p$ ó $C^T_{q+1} x = d_{q+1}$

x^{k+1} es entonces una solución básica admisible para el nuevo sistema de restricciones. Ir al paso k+1.

4.3: LA CONVERGENCIA DEL ALGORITMO

Vamos a demostrar que el método BEALE conduce en un número finito de iteraciones a una solución óptima de **P** o que termina con la respuesta que inf(**P**) $- \infty$

Suponer naturalmente que $S \neq \phi$

Para $x \in S$ definimos

$$I_o(x) = \{ j \ / \ 1 \leq j \leq m \quad a_j^T x = b_j \}$$

Decimos que el punto **x** es normal, si es óptimo para el programa:

$$P(x) \qquad \min \{ F(x) \ / \ a_j^T x = b_j, \ \forall \ j \in I_o(\bar{x}) \} \qquad (4.7)$$

Si $\{x^k\}$ es la sucesión de puntos de iteración del método de BEALE, se cumple $F(x^{k+1})$ $F(x^k)$, $\forall$ k . Como existe sólo número finito de programas **P(x)** de la forma (4.7), en la sucesión $\{x^k\}$ pueden presentar sólo un número finito de puntos normales.

Vamos a demostrar que en el método de BEALE, se llega siempre despues de un número finito de iteraciones a un punto normal. Con esto quedará evidentemente demostrado que el método es finito.

Sean $k_1 < \ldots < k_\ell$ tales que x^{k_j} $1 \leq j \leq \ell$, sean normales . Entonces x^{k_i} es solución óptima de

$$P(\mathbf{x}^{k_i}) \qquad \min \left\{ F(\mathbf{x}) \ / \ \mathbf{a}_j^T\mathbf{x} = \mathbf{b}_j \quad j \in I_o(\mathbf{x}^{k_i}) \right\}$$

Como $F(\mathbf{x}^{k_{i+1}}) < F(\mathbf{x}^{k_i})$, todos los programas $P(\mathbf{x}^{k_i})$ son distintos y podemos asumir que $P(\mathbf{x}^{k_i})$ $1 \leq i \leq \ell$, es la familia de todos los programas de la forma (4.7). Por consiguiente los puntos $\mathbf{x}^t$, $t > k_\ell$, son todos no normales y debe existir un $t_o \in \mathbb{N}$, t_o k_ℓ tal que en $\mathbf{x}^{t_o}$ el algoritmo termina.

Pero el algoritmo sólo puede terminar en $\mathbf{x}^{t_o}$ si $\mathbf{x}^{t_o}$ es óptimo para $\mathbf{P}$ ó $\lim_{\lambda \to \infty} F(\mathbf{x}^{t_o} + \lambda S^{t_o}) = -\infty$ y $\inf(\mathbf{P}) = -\infty$

1.- Si $\mathbf{x}^k$ no es normal, entonces en la (k+1) - ésima iteración se cancela una restricción.

Suponemos que $\mathbf{x}^k$ está asociada a la base

$$\tilde{\mathbf{A}}_B^T = [c_1, c_2, \ldots, c_g, a_{g+1}, \ldots, a_n]$$

Entonces $I_o(\mathbf{x}^k) = \{g+1, g+2, \ldots, n\}$

Si fuera $g = 0$, se obtuviera que $I_o(\mathbf{x}^k) = \{1, 2, \ldots, n\}$ y $\mathbf{x}^k$ fuera el único punto admisible de $\mathbf{P}(\mathbf{x}^k)$ y por consiguiente normal. En consecuencia $g > 0$.

Como no es normal, existe un punto $\mathbf{x}$ admisible para $I_o(\mathbf{x}^k)$ con $F(\bar{\mathbf{x}}) < F(\mathbf{x}^k)$.

Entonces

$$-\tilde{\mathbf{A}}_B(\bar{\mathbf{x}} - \mathbf{x}^k) \qquad \mathbf{u}\ , \ \mathbf{u}_j = 0 \qquad j = g+1, \ldots, n$$

De esto se obtiene:

$$\mathbf{x} \qquad \mathbf{x}^k = \qquad = \sum_{j=1}^{g} \mathbf{u}_j \beta_j$$

$$0 \qquad F(\mathbf{x}) - F(\mathbf{x}^k) \geq \nabla F(\mathbf{x}^k)^T(\bar{\mathbf{x}} - \mathbf{x}^k) = \sum_{j=1}^{g} \mathbf{u}_j \beta_j^T \nabla F(\mathbf{x}^k)$$

lo que implica que :

$$\max \{ |\beta_j^T \nabla F(\mathbf{x}^k)| \ / \ 1 \leq j \leq g \} = |\beta_r^T \nabla F(\mathbf{x}^k)|$$

Por consiguiente la restricción $C_r^T\mathbf{x} - \mathbf{d}_r$ es omitida.

EXPLICACION:

[Se cumple $C_r^T\mathbf{x}^k = \mathbf{d}_r$ $\quad C_r^T\beta_r = C_r^T S^k = -1$. Por consiguiente $C_r^T(\mathbf{x}^k + \lambda S^k) \neq \mathbf{d}_r$ $\quad \forall \lambda \neq 0$ y la restricción $C^T\mathbf{x} = \mathbf{d}_r$, se omite en la (k+1)-ésima iteración].

Por consiguiente si $\mathbf{x}^k$ no es normal, entonces la (k+1)-ésima iteración, el número de restricciones en forma de ecuación disminuye en "1" ó se mantiene constante. Este número se mantiene constante si y sólo si el caso b) se presenta, esto es : si y sólo si

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda^{**} S^k \qquad \lambda = \lambda^{**}$$

2.- Si los puntos $\mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^{k+\ell}$ no son todos normales, y si desde la (k+1)-ésima hasta la ($k+\ell$)-ésima iteración se presenta el caso b), entonces en la transición de $\mathbf{x}^{k+\ell}$ a $\mathbf{x}^{k+\ell+1}$ se cancela siempre una ecuación que ya estuvo presente en el sistema de restricciones de la k-ésima iteración.

Sea 1º $\ell = 1$. Por 1), la base en $\mathbf{x}^{k+1}$ tiene la forma

$$= [a_1, a_2, \ldots, a_{n-g}, c_1, c_2, \ldots, c_{r-1}, c_{r+1}, \ldots, c_g, c_{g+1}]$$

donde : $c_{g+1} = 2CS^k$

Si β_n es la n-ésima columna de $-\tilde{\mathbf{A}}_B^{-1}$, entonces se cumple

$$\left.\begin{array}{llll} \mathbf{a}_j^T\beta_n = 0 & \mathbf{a}_j^T S^k = 0 & j = 1,2,\ldots,n-g & \\ \mathbf{C}_j^T\beta_n = 0 & \mathbf{C}_j^T S^k = 0 & j = 1,2,\ldots,g & j \neq r \\ 2(S^k)^T \mathbf{C}\beta_n = -1 \ , & 2(S^k)^T \mathbf{C} S^k & & \end{array}\right| \quad (4.8)$$

[Hemos visto que el caso b) se debe cumplir $(S^k)^T \mathbf{C} S^k > 0$]

Como $\mathbf{A}_B$ es no singular, de (4.8) se obtiene que

$$\beta_n = \mathbf{u}S^k \qquad \mathbf{u} \neq 0$$

$$\beta_n^T \nabla F(\mathbf{x}^{k+1}) = \mathbf{u}(S^k)^T \nabla F(\mathbf{x}^{k+1}) = 0$$

$$g_k'(\lambda^{**}) = 0$$

Por consiguiente $\beta_n \neq S^{k+1}$

[pues $\max \{ |\beta_j^T \nabla F(\mathbf{x}^{k+1})| \ / \ n-g+1 \leq j \leq n \} = |\beta_r^T \nabla F(\mathbf{x}^{k+1})| > 0$

$S^k = \beta_r$] y en la (k+2)-ésima iteración debe ser omitida una de las restricciones

$$\mathbf{C}_j^T \mathbf{x} = \mathbf{d}_j \ , \ j = 1,2,\ldots,g \ . \quad j \neq r$$

Sea ahora $\ell > 1$, y suponemos que desde la (k+1)-ésima iteración hasta la (k+ℓ)-ésima haya sido siempre eliminada una de las restricciones : $\mathbf{C}_j^T \mathbf{x} - \mathbf{d}_j$; $j = 1,2,\ldots,g$

Podemos entonces asumir que en $\mathbf{x}^{k+\ell}$ la base tiene la siguiente forma :

$$[a_1, a_2, \ldots, a_{n-g}, c_{\ell+1}, \ldots, c_g, c_{g+1}, \ldots, c_{g+\ell}]$$

donde $\mathbf{C}_{g+1} = 2\mathbf{C}S^{k+i-1}$.

Si i ℓ, entonces los puntos $\mathbf{x}^{k+i}, \mathbf{x}^{k+i+1}, \ldots, \mathbf{x}^{k+\ell}$ satisfacen la restricción

$$C_{g+i}^{T}\mathbf{x} - 2(S^{k+i-1})C\mathbf{x} = \mathbf{d}_{g+i}$$

De esto se obtiene :

$$2(S^{k+i-1})^{T}C(\mathbf{x}^{k+r+1} - \mathbf{x}^{k+r}) - 2(S^{k+i-1})^{T}C(\lambda_{k+r}S^{k+r}) = 0$$

Como λ_{k+r} $\quad r = i, i+1, \ldots, \ell-1$, se obtiene de esto

$$2(CS^{k+r})^{T}S^{k+i-1} = (C_{g+r+1})^{T}S^{k+i-1} = 0, \quad r = i, i+1, \ldots, \ell-1$$

Sean $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ las ℓ-últimas, columnas de $-\tilde{A}_B^{-1}$. Entonces se cumple evidentemente para $1 \le i \le \ell$

$$\begin{array}{lll} \mathbf{a}_j^{T}\alpha_i = 0 & j = 1, 2, \ldots, n-g & \\ C_j^{T}\alpha_i = 0 & j = \ell+1, \ldots, g & \qquad +) \\ C_{g+n}^{T}\alpha_i = 0 & j = 1, 2, \ldots, \ell & \\ C_{g+i}^{T}\alpha_i = 1 & & \end{array}$$

Consideramos de otro lado S^{k+i-1}, $1 \le i \le \ell$

Como en $(k+i)$-ésima iteración las restricciones

$$\begin{array}{lll} \mathbf{a}_j^{T}\mathbf{x} = \mathbf{b} & j \; 1, 2, \ldots, n-g & \\ C_j^{T}\mathbf{x} = \mathbf{d}_j & j = \ell+1, \ldots, g+i-1 & \qquad ++) \end{array}$$

Se mantienen en la base , se cumple

$$\begin{array}{ll} \mathbf{a}_j^{T}S^{k+i-1} = 0 & j = 1, 2, \ldots, n-g \\ C_j^{T}S^{k+i-1} = 0 & j = 1, 2, \ldots, g+i-1 \end{array}$$

De (*) se deduce que :

$$(C_{g+h})^{T}S^{k+j-1} = 0 \qquad h = g+i+1, \ldots, \ell \qquad (+++)$$

Finalmente se cumple

$$(\mathbf{C}_{g+i})^{\mathsf{T}} S^{k+i-1} = 2(S^{k+i-1})^{\mathsf{T}} \mathbf{C} S^{k+i-1} > 0$$

Por consiguiente, si $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ son las últimas columnas de $-\tilde{\mathbf{A}}_B^{-1}$, se obtiene de +), ++), ++) para $1 \le i \le \ell$

$$= \mathbf{C}^{\mathsf{T}} S^{k+i-1} = 0 \qquad j = \ell+1, \ldots, g$$

$$\mathbf{a}_j^{\mathsf{T}} \alpha_i = \mathbf{a}_j^{\mathsf{T}} S^{k+i-1} = 0 \qquad j = 1, 2, \ldots, n-g$$

$$(+) \qquad \mathbf{C}_{g+h}^{\mathsf{T}} \alpha_i = \mathbf{C}_{g+h}^{\mathsf{T}} S^{k+i-1} = 0 \qquad h = 1, 2, \ldots, \ell, \qquad h \ne i$$

$$\mathbf{C}_{g+i}^{\mathsf{T}} \alpha_i = 2(S^{k+i-1})^{\mathsf{T}} \mathbf{C} \alpha_i = -1$$

$$2(S^{k+i-1}) \mathbf{C} S^{k+i-1} > 0$$

Como $\mathbf{A}_B$ es no singular, se obtiene de (+)

$$\alpha_i = \mathbf{u}_i S^{k+i-1}, \quad \mathbf{u}_i \ne 0 \qquad i \le \ell \qquad (4.9)$$

De esto y de

$$2(S^{k+i-1})^{\mathsf{T}} \mathbf{C} \mathbf{x}^{k+\ell} + (S^{k+i-1})^{\mathsf{T}} \mathbf{p} = (S^{k+i-1})^{\mathsf{T}} \nabla F(\mathbf{x}^{k+\ell}) = 0 \qquad (4.10)$$

$$\left[\mathbf{x}^{k+\ell} \text{ satisface } 2(S^{k+i-1})^{\mathsf{T}} \mathbf{C} \mathbf{x}^{k+\ell} = \mathbf{d}_{g+i} = -(S^{k+i-1})^{\mathsf{T}} \mathbf{p} \right]$$

Se obtiene que $\alpha_i^{\mathsf{T}} \nabla F(\mathbf{x}^{k+\ell}) = 0 \quad 1 \le i \le \ell$

En consecuencia $S^{k+\ell} \ne \alpha_i \qquad 1 \le i \le \ell$

Por consiguiente en la $(k+\ell+1)$-ésima iteración, ninguna de las iteraciones $\mathbf{C}_{g+i}^{\mathsf{T}} x - \mathbf{d}_{g+1}$, será omitida.

3. – Si $\tilde{\mathbf{A}}_B^{\mathsf{T}} = [a_1, a_2, \ldots a_{n-g}, c_1, c_2, \ldots c_g]$ es la base en $\mathbf{x}^k$, si todos los puntos $\mathbf{x}^{k+1}, \ldots, \mathbf{x}^{k+g-1}$ son no-normales y si desde la $(k+1)$-ésima, hasta la $(k+g)$-ésima iteración se presenta siempre el caso b), entonces $\mathbf{x}^{k+g}$ debe ser normal.

De acuerdo con 2.- la base en $\mathbf{x}^{k+g}$ tiene la forma

$\tilde{\mathbf{A}}_B^T = [a_1, a_2, \ldots a_{n-g}, c_{g+1}, c_{g+2}, \ldots c_{g+g}$

En $\mathbf{x}^{k+g}$ las restricciones en forma de ecuación tienen la forma :

$$(S^{k+i})^T \nabla F(\mathbf{x}) = 0 \qquad 0 \leq i \leq g-1$$

Si $-\tilde{\mathbf{A}}_B^{-1} = [\beta_1, \beta_2, \ldots, \beta_n]$, entonces se cumple por (4.9) y (4.10)

$$\beta_{n-g+1+i} = \mathbf{u}_i S^{k+i} \qquad \mathbf{u}_i \neq 0 \qquad 0 \leq i \leq g-1$$

$$(\beta_{n-g+1+i})^T \nabla F(\mathbf{x}^{k+g}) = \mathbf{u}_i (S^{k+i})^T \nabla F(\mathbf{x}^{k+g}) = 0 \qquad (4.11)$$

$$0 \leq i \leq g-1.$$

Si $\mathbf{x}$ es admisible para el programa :

$$\mathbf{P}(\mathbf{x}^{k+g}) \qquad \min \{ F(\mathbf{x}) \ / \ \mathbf{a}_j^T \mathbf{x} = \mathbf{b}_j \qquad j = 1, 2, \ldots, n-g \}$$

entonces $-\tilde{\mathbf{A}}_B(\mathbf{x} - \mathbf{x}^{k+g}) = \lambda \qquad \lambda_j = 0 \qquad j = 1, 2, \ldots, n-g$

ó $\mathbf{x} - \mathbf{x}^{k+g} = -\tilde{\mathbf{A}}_B^{-1} \lambda$

De esto y de (4.11) se obtiene :

$$F(\mathbf{x}) - F(\mathbf{x}^{k+g}) \geq \nabla F(\mathbf{x}^{k+g})^T (\mathbf{x} - \mathbf{x}^{k+g}) = \sum_{j=n-g+1}^{n} \lambda_j \beta_j^T \nabla F(\mathbf{x}^{k+g}) = 0$$

Por consiguiente $\mathbf{x}^{k+g}$ es óptimo para $\mathbf{P}(\mathbf{x}^{k+g})$, lo que significa que $\mathbf{x}^{k+g}$ es normal.

<u>CONCLUSION:</u>

Si en el punto no-normal $\mathbf{x}^k$ el sistema de restricciones contiene exactamenete g ecuaciones, entonces de acuerdo con 1., 2. y 3., sólo pueden presentarse los siguientes casos:

α) Existe un ℓ , $1 \leq \ell \leq g$ tal que $\mathbf{x}^{k+\ell}$ es normal.

β) Los puntos de iteración $\mathbf{x}$, $\mathbf{x}^{k+1},\ldots,\mathbf{x}^{k+g-1}$ son todos no-normales y siempre se presenta el caso b). Entonces $\mathbf{x}^{k+g}$ debe ser normal.

γ) Los puntos $\mathbf{x}^{k}$, $\mathbf{x}^{k+1},\ldots,\mathbf{x}^{k+g-1}$, $\mathbf{x}^{k+g}$ son todos no-normales

Entonces desde la (k+1)-ésima hasta la (k+g)-ésima iteración se debe presentar al menos una vez el caso a). Entonces el número de restricciones en forma de ecuación en el punto $\mathbf{x}^{k+g}$ es $\bar{g} < g$.

De α), β) y γ) se deduce que el algoritmo de BEALE siempre va a encontrar un punto normal despues de un número finito de iteraciones.

4.4: <u>Ejemplo para el algoritmo de BEALE</u>

$$\min\ 2x_1^2 + x_2^2 - 48x_1 \quad 40x_2$$

$$\begin{array}{rcll}
x_1 + x_2 & \leq & 8 & \text{1)} \\
x_1 & \leq & 6 & \text{--- 2)} \\
x_1 + 3x_2 & \leq & 18 & \text{--- 3)} \\
x_1 & \leq & 0 & \text{--- 4)} \\
x_2 & \leq & 0 & \text{--- 5)}
\end{array}
\qquad
A = \begin{matrix} 1 & 1 \\ 1 & 0 \\ 1 & 3 \\ -1 & 0 \\ 0 & -1 \end{matrix} \quad 5\times 2$$

<u>PASO INICIAL:</u>

$\mathbf{x}^{o} = \mathbf{x}_{B_o} = (0,0)^{\mathbf{T}}$ es admisible

$\mathbf{A}_{B_o} = \mathbf{I}$ $\qquad B_o = \{5,6\}$

$\beta_i = e_i \qquad i = 1,2$

$\nabla F(\mathbf{x})^{\mathbf{T}} = (4x_1 - 48,\ 2x_2 - 40)$

$\nabla F(\mathbf{x}^{o})^{\mathbf{T}} = (-48,\ -40)$

$$\left.\begin{array}{l}\nabla F(\mathbf{x}^{o})^{T}\beta_1 = 48 \\ \nabla F(\mathbf{x}^{o})^{T}\beta_2 = -40)\end{array}\right\} \Longrightarrow S^{o} = e_1$$

Determinación de λ^{*}

$$\mathbf{a}_j^{T}(\mathbf{x}^{o} + \lambda S^{o}) = \lambda \mathbf{a}_j^{T} e_1$$

$$\left.\begin{array}{l}\lambda \mathbf{a}_1 e_1 = \lambda \leq 8 \\ \lambda \mathbf{a}_2^{T} e_1 = \lambda \leq 6 \\ \lambda \mathbf{a}_3^{T} e_1 = \lambda \leq 18\end{array}\right\} \Longrightarrow \lambda^{*} = 6$$

Determinacion de λ^{**}

$$g_o(\lambda) = F(\mathbf{x}^{o} + \lambda S^{o}) = F(\lambda e_1) = 2\lambda^2 - 48\lambda$$

$$g'(\lambda) = 4\lambda - 48$$

$$\Longrightarrow \lambda^{**} = 12$$

$$\Longrightarrow \lambda = \lambda^{*} = 6$$

$\mathbf{x}^{1} = \mathbf{x}^{o} + \lambda e_1 - (6,0)^{T}$. No se introduce nueva restricción

2ª ITERACION

$\mathbf{x}^{1} = (6,0)^{T}$ $\qquad B_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$

$\mathbf{A}_B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\qquad \mathbf{A}_B^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

$$\nabla F(\mathbf{x}^{1})^{T} = (-24, -40)$$

$\beta_1 = -e_1 \qquad \beta_2 = e_2$

$$\left.\begin{array}{l}\nabla F(\mathbf{x}^{1})^{T}\beta_1 = 24 \\ \nabla F(\mathbf{x}^{1})^{T}\beta_2 = 40\end{array}\right\} \Longrightarrow S^{1} = e_2$$

$$\mathbf{x}^{1}(\lambda) = \mathbf{x}^{1} + \lambda e_2 = (6.\lambda)^{T}$$

Chequeamos las restricciones no básicas

$$\left.\begin{array}{l} a_1^T x^1(\lambda) = 6 + \lambda \leq 8 \\ a_3^T x^1(\lambda) = 6 + 3\lambda \leq 18 \end{array}\right\} \qquad \begin{array}{l} \lambda \leq 2 \\ \lambda \leq 4 \end{array}$$

$$6 \leq 0 \qquad\qquad \Longrightarrow \lambda^* = 2$$

$$g_1(\lambda) = F(x^1(\lambda)) = F(6,\lambda) = \lambda^2 - 40\lambda - 216$$

$$g_1'(\lambda) = 2\lambda - 40 = 0 \quad \Longrightarrow \lambda^{**} = 20$$

$$\Longrightarrow \lambda = \lambda^* = 2$$

$$x^2 = x^1 + 2S^1 = (6,0) + 2(0,1) = (6,2)$$

No se introduce nueva restricción.

ITERACION

$$x^2 = (6,2)^T \qquad\qquad B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$A_{B_2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad -A_B^{-1} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$\nabla F(x^2)^T = (-24, -36)$$

$$\nabla F(x^2)^T \beta_1 = (-24, -36) \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 36$$

$$\nabla F(x^2)^T \beta_2 = (-24, -36) \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -12$$

$$\Longrightarrow S^2 = \beta_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Determinación de λ^*:

Chequeamos las restricciones no básicas

$$x^2(\lambda) = x^2 + \lambda\beta_2 = (6,2)^T + \lambda(-1,1)^T = (6-\lambda,\ 2+\lambda)^T$$

$$6 - \lambda + 3(2 + \lambda) = 12 + 2\lambda \leq 18 \quad \Longleftrightarrow \quad \lambda \leq 3$$

$$-6 + \lambda \leq 0 \quad \Longleftrightarrow \quad \lambda \leq 6$$

$$-2 - \lambda \leq 0 \quad \Longleftrightarrow \quad , \forall\ \lambda \geq 0$$

$$\Longrightarrow \quad \lambda^* = 3$$

$g_3(\lambda) = F(\mathbf{x}^2(\lambda)) = F(6-\lambda,\ 2+\lambda) =$

$= 2(6-\lambda)^2 + (2+\lambda)^2 - 48(6-\lambda) - 40(2+\lambda)$

$g_3^1(\lambda) = -4(6-\lambda) + 2(2+\lambda) + 48 - 40 = 0 \iff \lambda = \lambda^{**} = 2$

$\bar{\lambda} = \lambda^{**} = 2$

$\Longrightarrow \mathbf{x}^3 = \mathbf{x}^2 + 2S^2 = (6,2)^T + 2(-1,\ 1) = (4,4)^T$

Se introduce la nueva restricción $(S^2)^T\nabla F(\mathbf{x}) = 0$

$$ó\ (-1,\ 1)^T\begin{bmatrix} 4x_1 - 48 \\ 2x_2 - 40 \end{bmatrix} = -4x_1 + 48 + 2x_2 - 40 = 0$$

Nueva restricción : $\| -4x_1 + 2x_2 = -8 \|$ $\longrightarrow$ restricción --- 6)

4ª ITERACION

$\mathbf{x}^3 = (4,4)^T$ $B_3 = \{1, 6\}$

Base asociada a $\mathbf{x}^3$:

$$\tilde{\mathbf{A}}_{B_3} = \begin{bmatrix} 1 & 1 \\ -4 & 2 \end{bmatrix}, \quad -\tilde{\mathbf{A}}_B^{-1} = \begin{bmatrix} -1/3 & 1/6 \\ -2/3 & -1/6 \end{bmatrix}$$

$\nabla F(\mathbf{x}^3) = (-32,\ -32)$

$$\nabla F(\mathbf{x}^3)^T\beta_1 = (-32,\ -32)\begin{bmatrix} -1/3 \\ -2/3 \end{bmatrix} = 32 > 0$$

$$\nabla F(\mathbf{x}^3)^T\beta_2 = (-32,\ -32)\begin{bmatrix} -1/6 \\ 1/6 \end{bmatrix} = 0$$

$\therefore\ \mathbf{x}^3$ es óptimo.

CAPITULO V

LOS METODOS DE ZUNGWILL Y DANTZIG-COTTLE

5.1: EL CASO CONVEXO

<u>HIPOTESIS:</u> $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ es convexa y diferenciable

El conjuntp R de soluciones óptimas del programa

$$\min \{ F(\mathbf{x}) \ / \ \mathbf{x} \in \mathbb{R}^n \}$$

es compacto.

$\mathbf{A} \in M(m,n)$ $\quad$ $\operatorname{ran}(\mathbf{A}) = n$

Escribimos $\quad \mathbf{A} = [a_1, a_2, \ldots, a_m]^T$

Consideramos el programa :

$$\mathbf{P} : \quad \min \{ F(\mathbf{x}) \ / \ \mathbf{Ax} \le \mathbf{b} \} \tag{5.1}$$

con el dominio admisible S.

<u>Recordemos los siguientes resultados:</u>

<u>1</u>.- $\hat{\mathbf{x}}$ es óptimo para P sí y sólo si existe $\hat{\mathbf{u}}$, tal que $m(\hat{\mathbf{x}},\hat{\mathbf{u}})$ se satisfacen las condiciones linealizadas de K-T.

Estas condiciones tienen la forma :

$\hat{\mathbf{x}} \in S$

$$\nabla F(\hat{\mathbf{x}}) = - \sum_{j \in I(\mathbf{x})} \hat{\mathbf{u}}_j \mathbf{a}_j \qquad \hat{\mathbf{u}}_j \ge 0 \qquad \forall \ j \in I(\hat{\mathbf{x}}) \tag{5.2}$$

donde $I(\hat{\mathbf{x}}) = \{ \ j \ / \ 1 \le j \le n \qquad \mathbf{a}_j^T \hat{\mathbf{x}} = \mathbf{b}_j \}$

2.- Si $\hat{x}$ es óptimo para $\hat{P}$, entonces $\hat{x}$ es también solución óptima de

$$\min \left\{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} - \mathbf{b}_i \quad \forall\ i \in I(\hat{\mathbf{x}}) \right\}$$

DEMOSTRACION

Supongamos que $\hat{\mathbf{x}}$ no sea óptimo para $\mathbf{P}$, entonces existe $\mathbf{x}$ admisible para $\hat{\mathbf{P}}$ con $F(\mathbf{x}) < F(\hat{\mathbf{x}})$

Además :

$$\mathbf{a}_i^T\hat{\mathbf{x}} - \mathbf{a}_i^T\bar{\mathbf{x}} = \mathbf{b}_i \qquad \forall\ i \in I(\hat{\mathbf{x}})$$

$$\mathbf{a}_i^T\hat{\mathbf{x}} \quad \mathbf{b}_i \qquad \forall\ i \notin I(\mathbf{x})$$

Para $\lambda \in \langle 0, 1]$, consideramos

$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda(\mathbf{x} \quad \hat{\mathbf{x}}) = \lambda\mathbf{x} + (1-\lambda)\hat{\mathbf{x}}$$

Es claro que : $\mathbf{a}_i^T\mathbf{x}(\lambda) = \mathbf{b}_i \qquad \forall\ i \in I(\mathbf{x})$

Si $i \notin I(\mathbf{x})$, entonces

$$\mathbf{a}_i^T\mathbf{x}(\lambda) = (1-\lambda)\mathbf{a}_i^T\hat{\mathbf{x}} + \lambda\mathbf{a}_i^T\bar{\mathbf{x}} = g_i(\lambda)$$

como $g_i(0) = \mathbf{a}_i^T\hat{\mathbf{x}} < \mathbf{b}_i$, entonces existe $\lambda_o \in \langle 0, 1]$ tal que :

$$\mathbf{a}_i^T\mathbf{x}(\lambda) = g_i(\lambda) < \mathbf{b}_i \qquad \forall\ \lambda \in \langle 0, \lambda_o] \qquad \forall\ i \notin I(\hat{\mathbf{x}})$$

(*) y (**) implica que $\mathbf{x}(\lambda_o)$ es admisible para $\mathbf{P}$.

F convexa implica

$$F(\mathbf{x}(\lambda_o)) \le \lambda_o F(\mathbf{x}) + (1-\lambda_o)F(\hat{\mathbf{x}}) \quad \lambda\ F(\mathbf{x}) + (1-\lambda_o)F(\hat{\mathbf{x}}) = F(\hat{\mathbf{x}})$$

Luego $\mathbf{x}$ no es óptimo para $\mathbf{P}$ (contradicción con la hipótesis)

Luego $\hat{\mathbf{x}}$ sí es óptimo para $\hat{\mathbf{P}}$ ∎

En consecuencia, si se conociera $I(\hat{\mathbf{x}})$ se pudiera resolver $\mathbf{P}$, resolviendo el programa $\hat{\mathbf{P}}$, al menos si $\mathbf{P}$ tiene una única solución.

Los resultados 1 y 2 representan la base teórica para el algoritmo de ZUNGWILL con el nombre de "*SUB-OPTIMIZACION EN VARIEDADES LINEALES*" (*"MANIFOLD SUB-OPTIMIZATION"*)

Este método resuelve una sucesión de programas

$$\mathbf{P}_k \qquad \min \{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i \qquad i \in I_k \}$$

donde $I_k \subset \{1,2,\ldots,m\}$ y S_k denotará el dominio admisible de $\mathbf{P}_k$.

Los conjuntos de índices I_k se cambian sucesivamente de tal modo que en un número finito de iteraciones se obtiene: $I_k = I(\hat{\mathbf{x}})$, donde $\hat{\mathbf{x}}$ es una solución óptima de $\mathbf{P}$.

OBSERVACION 5.1:

Como el conjunto de soluciones óptimas de $\{F(\mathbf{x}) \ / \ \mathbf{x} \in \mathbb{R}^n\}$ es compacto, los programas $\mathbf{P}_k$ tienen al menos una solución óptima, siempre y cuando su dominio admisible S_k es no vacío.

EXPLICACION:

Todos los conjuntos de nivel $\{\mathbf{x} \in \mathbb{R}^n \ / \ F(\mathbf{x}) \leq \alpha\}$ son convexos y compactos. Por consiguiente si $\mathbf{y} \in S_k$, entonces

$= S_k \cap \{\mathbf{x} \in \mathbb{R}^n \ / \ F(\mathbf{x}) \leq F(\mathbf{y})\}$ es compacto y $\mathbf{P}_k$ es equivalente a $\mathbf{P}_k$: $\quad \min \{ F(\mathbf{x}) \ / \ \mathbf{x} \in S_k\}$

Como F es continua y S_k es compacto, $\mathbf{P}_k$ y con esto $\mathbf{P}_k$ tiene al menos un solución óptima. ■

Sea $I \subset J$ e $\mathbf{y}$ una solución óptima de

$$\min \{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i \qquad \forall\ i \in I\} \tag{5.3}$$

Entonces en $\mathbf{y}$ se cumplen las condiciones linealizadas de K-T para (5.3).

$$\nabla F(\mathbf{y}) = -\sum_{j\in I} \mathbf{u}_j \mathbf{a}_j \qquad (5.4)$$

DEFINICION 5.1.- Sea $\mathbf{y}$ una solución óptima de (5.3), $I^* = I(\mathbf{y})$. Entonces $\mathbf{y}$ se llama no-degenerada, si los vectores $\mathbf{a}_j$ $j \in I^*$ son linealmente independientes.

<u>OBSERVACION 5.2:</u>

Si $\mathbf{y}$ es no degenerada, entonces como $I \subset I^*$, (5.4) es la única representación de $\nabla F(\mathbf{y})$ en la forma

$$\nabla F(\mathbf{y}) = \sum_{j\in I^*} \lambda_j \mathbf{a}_j$$

Usamos la siguiente

PROPOSICION 5.1.- Sea $\mathbf{y}$ una solución óptima no degenerada de

$$\min \{ F(\mathbf{x}) \ / \ \mathbf{a}^T \mathbf{x} - \mathbf{b}_i \ ; \ \forall\, i \quad I \} \qquad (5.3)$$

Sea en (5.4) $\mathbf{u}_j \quad 0$ para algún $j \in I$, $\bar{I} = I^* \setminus \{j\}$

Sea además $\mathbf{y}^*$ una solución óptima del programa

$$\min \{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i \quad \forall\, i \in \bar{I} \} \qquad (5.5)$$

Entonces se cumple :

$$F(\mathbf{y}^*) \quad F(\mathbf{y}) \qquad \mathbf{a}_j^T \mathbf{y}^* \quad \mathbf{b}_j$$

DEMOSTRACION

Como $\mathbf{y}$ es admisble para el programa :

$$\min \{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i, \ \forall\, i \in I^* \setminus \{j\} \qquad \mathbf{a}_j^T \mathbf{x} \geq \mathbf{b}_j \} \qquad (5.6)$$

y $\mathbf{u}_j$ 0, (5.4) representa las condiciones de K-T para (5.6)
Por consiguiente $\mathbf{y}$ es una solución óptima de (5.6).

Como $\mathbf{y}$ es no degenerada, no existen multiplicadores $\mathbf{u}_i^*$, $i \in I^*$, tales que

$$\nabla F(\mathbf{y}) = -\sum_{i \in I^*} \mathbf{u}_i^* \mathbf{a}_i, \quad \mathbf{u}_i \geq 0$$

Por consiguiente no es posible que en $\mathbf{y}$ sean cumplidas las condiciones linealizadas de K-T para el programa :

$$\min \left\{ F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i, \ \forall \ i \in I^* \setminus \{j\} \quad \mathbf{a}_j^T\mathbf{x} \leq \mathbf{b}_j \right\} \qquad (5.7)$$

En consecuencia $\mathbf{y}$ no es óptima para (5.7). Por esta razón y porque $\mathbf{y}$ es una solución óptima de (5.6), se cumple para la solución óptima de $\mathbf{y}^*$ de (5.5)

$$F(\mathbf{y}^*) < F(\mathbf{y}) \ , \quad \mathbf{a}_j^T\mathbf{y}^* < \mathbf{b}_j$$

<u>EXPLICACION:</u>

Sean α), β), γ), δ) los valores óptimos de los programas (5.3), (5.5), (5.6) y (5.7) respectivamente.
Como $\mathbf{y}$ es admisible para todos los programas y óptimo para (5.3) y (5.6), pero no óptimo para (5.7) se cumple:

$$F(\mathbf{y}) = \alpha = \gamma > \delta \geq \beta = F(\mathbf{y}^*)$$

Si fuera $\mathbf{a}_j^T\mathbf{y}^* \geq \mathbf{b}_j$, $\mathbf{y}^*$ fuera admisible para (5.6) y $F(\mathbf{y}^*) \geq F(\mathbf{y})$, lo que es absurdo

Por consiguiente : $\mathbf{a}_j^T\mathbf{y}^* \leq \mathbf{b}_j$ ■

5.1.1: EL <u>ALGORITMO:</u>

<u>Inicio:</u> Se determina un punto admisible $\mathbf{x}^0$ de $\mathbf{P}$, y se pone

$$I_0 = I(\mathbf{x}^0)$$

<u>Iteración general</u> : Si los puntos de iteración $\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^k$ y los conjuntos de índices $I_0, I_1, \ldots, I_k$ ya son construidos obtiene $\mathbf{x}^{k+1}$ y I_{k+1} de acuerdo a las siguientes reglas :

a) Se determina una solución óptima $\mathbf{y}^k$ de

$$\mathbf{P}_k : \min \{F(\mathbf{x}) \ / \ \mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i, \ \forall \ i \in I_k\}$$

b) <u>Caso I:</u> Si $\mathbf{y}^k$ es admisible para $\mathbf{P}$, entonces se resuelve el sistema

$$\nabla F(\mathbf{y}^k) = -\sum_{i \in I_k} \mathbf{u}_i \mathbf{a}_i$$

y se determina $\mathbf{u}_j = \min \{\mathbf{u}_i \ / \ i \in I_k\}$

Si $\mathbf{u}_j \geq 0$, entonces $\mathbf{y}^k$ es óptimo para $\mathbf{P}$, y el algoritmo termina.

Si $\mathbf{u}_j < 0$, se elige $\mathbf{x}^{k+1} = \mathbf{y}^k$ y $I_{k+1} = I(\mathbf{y}^k) \setminus \{j\}$

<u>Caso II:</u> Si $\mathbf{y}^k$ no es admisible para $\mathbf{P}$, entonces se elige $\mathbf{x}^{k+1}$ como el punto en $[\mathbf{x}_k, \mathbf{y}_k] \cap S$, mas cercano a $\mathbf{x}^k$ y $I_{k+1} = I(\mathbf{x}^{k+1})$

<u>Como se elige esto:</u>

Se cumple : $\mathbf{a}_j^T\mathbf{x}^k \leq \mathbf{b}_j \qquad \forall \ j \in J$

Como $\mathbf{y}^k$ no es admisible para $\mathbf{P}$, existe un $\ell \in J$ con $\mathbf{a}_\ell^T\mathbf{x}^k > \mathbf{b}_\ell$. Para todos los índices ℓ resolvemos la ecuación

$$\mathbf{a}_\ell^T(\mathbf{x}^k + \lambda(\mathbf{y}^k - \mathbf{x}^k)) = \mathbf{b}_\ell$$

Se cumple :

$$\mathbf{a}_\ell^T(\mathbf{x}^k + \lambda(\mathbf{y}^k - \mathbf{x}^k)) = \mathbf{b}_\ell \Longleftrightarrow \lambda = \frac{\mathbf{b}_\ell - \mathbf{a}_\ell^T\mathbf{x}^k}{\mathbf{a}_\ell^T \quad \mathbf{y}^k - \mathbf{x}^k}$$

Si $\lambda = \min \{\lambda_\ell \ / \ a_\ell^T y^k \ \}$

Entonces $x^{k+1} = x^k + \bar{\lambda}(y^k - x^k)$ ■

TEOREMA 5.1.- Sea F la función convexa y diferenciable en $\mathbb{R}^n$ Sean las soluciones y^k de P_k que son admisibles para P, todas no degeneradas. Entonces el algoritmo conduce en un número finito de iteraciones a una solución óptima de P.

DEMOSTRACION

a) Si y^k es admisible para P y $u_j = \min \{u_i \ / \ i \in I_k\} \geq 0$, entonces en y^k se cumplen las condiciones linealizadas de K-T para P y y^k es óptima para P. $(I_k \subset I(y^k))$

b) Observamos que x^k es admisible para P y P_k, $\forall$ k

Por consiguiente $F(y^k) \leq F(x^k)$, $\forall$ k

Como F es convexa y $x^{k+1} \in [x^k, y^k] \subset S_k$, se cumple

$$F(y^k) \leq F(x^{k+1}) \leq F(x^k) \ , \quad \forall \ k$$

Si y^{k-1} es admisible para P y si para la solución u de

$$\nabla F(y^{k-1}) = - \sum_{i \in I_{k-1}} u_i a_i$$

se cumple $u_j = \min \{u_i \ / \ i \in I_{k-1}\} < 0$, entonces se cumple

$$F(x^{k+1}) < F(x^k)$$

En este caso se cumple :

$x^k = y^{k-1} \qquad I_k = I(y^{k-1}) \setminus \{j\}$ y y^{k-1} es no degenerada.

De la proposición 5.1, se deduce entonces que

$$F(y^k) < F(x^k) = F(y^{k-1}) \qquad a_j^T y^k < b_j$$

De esto y de la convexidad de F se deduce que

$$x^{k+1} \in \langle x^k, y^k] \qquad F(x^{k+1}) < F(x^k)$$

Explicación de (*): (Es evidente que se cumple $y^k \neq x^k$)

i) Si $x^{k+1} = y^k$, entonces $x^{k+1} \in \langle x^k,$

ii) Si y^k no es admisible para P, entonces existe un índice ℓ tal que $a_\ell^T y^k > b_\ell$. Entonces $\ell \notin I(x^k)$ ó $a_\ell^T x^k$

$$\lambda_\ell = \frac{b_\ell - a_\ell^T x^k}{a_\ell^T (y^k - x^k)} > 0$$

Por consiguiente se cumple

$\bar{\lambda} = \min \{\lambda_\ell \ / \ a_\ell^T y^k > b_\ell\} > 0 \qquad x^{k+1} = x^k + \bar{\lambda}(y^k - x^k)$

c) El caso I sólo se puede presentar un número finito de veces.

Si para P_k se presenta el caso I, entonces de acuerdo a b) se cumple :

$$F(y^k) \quad F(x^{k+1}) \quad F(x^{k+2}) \quad F(y^{k+1}$$

$$F(x^{k+2}) \geq F(x^{k+\ell}) \geq F(y^{k+\ell-1}) \qquad \forall\ \ell \geq 2$$

Por consiguiente $\quad \inf(P_{k+\ell}) < \inf(P_k) \ ; \ \forall\ \ell \geq 1$

Por esta razón todos los programas $P_{k+\ell}$, $\ell \geq 1$ son distintos de P_k .

Como el número de programas posibles $P_.$ es finito, el caso I se puede presentar sólo un número finito de veces.

d) El caso II se puede presentar en a lo mas m iteraciones sucesivas.

Si y^k no es admisible para P y S_k es el dominio admisible de P_k, entonces $x^{k+1} \in [x^k, y^k\rangle \subset S_k$.

Para al menos un $i \notin I_k$. Se cumple ademas

$$b_r \qquad \mathbf{a}_r^T \mathbf{x}^{k+1} = \mathbf{b}_r$$

De esto se obtiene que $\quad I_{k+1} = I(\mathbf{x}^{k+1}) \supset I_k \cup \{r\} \supsetneqq I_k$

Por consigiente existe un $\ell \leq m-1$, tal que en el caso, donde desde la k-ésima hasta la $(k+\ell)$-ésima iteración se presenta el caso II, los vectores $\mathbf{a}_i$, $i \in I_{k+\ell}$ contienen una base de $\mathbf{A}$. Por consiguiente se cumple

$$S_{k+\ell} = \{\mathbf{x}^{k+\ell}\}, \quad \mathbf{y}^{k+\ell} = \mathbf{x}^{k+\ell}$$

En la $(k+\ell+1)$-ésima iteración debe otra vez presentarse el caso I.

De c) y d) se deduce que despues de un número finito ℓ de iteraciones se debe cumplir

$$\mathbf{y}^{\ell} \in S, \quad \mathbf{u}_j = \min \{ \mathbf{u}_i \ / \ i \in I_{\ell}\} \geq 0$$

Entonces $\mathbf{y}^{\ell}$ es una solución óptima de $\mathbf{P}$ y el algoritmo termina.

<u>OBSERVACION 5.3:</u>

El algoritmo no es finito en el sentido propio, por que para resolver los programas $\mathbf{P}_k$, se requiere generalmente de algoritmos iterativos que no son finitos.

5.2: EL CASO CUADRATICO

Sea $\mathbf{C} = \mathbf{C}^T \in M(n,n)$ definida positiva

Consideremos el programa:

$$\mathbf{Q} \qquad \min \{F(\mathbf{x}) = \tfrac{1}{2}\mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x} \ / \ \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$$

$$\mathbf{A} \in M(m,n)$$

Aplicando el algoritmo de ZUNGWILL a $\mathbf{Q}$, obtenemos el algoritmo de DANTZIG y COTTLE, si elegimos adecuadamente el punto inicial $\mathbf{x}^o$ y $\mathbf{P}_o$.

Las condiciones linealizadas de K-T para $\mathbf{Q}$ tienen la forma :

$$\mathbf{Cx} + \mathbf{p} = -\sum_{j\in I(\mathbf{x})} \mathbf{u}_j \mathbf{a}_j, \quad \mathbf{u}_j \geq 0 \qquad j \in I(\mathbf{x})$$

$$\mathbf{Ax} \leq \mathbf{b}$$

$$[\mathbf{u}_j = 0 \qquad \forall\ j \in J \setminus I(\mathbf{x})]$$

Si $\mathbf{y} = \mathbf{b} - \mathbf{Ax}$, podemos escribir estas condiciones como sigue

$$\mathbf{Ax} + \mathbf{y} = \mathbf{b} \tag{5.8}$$

$$\mathbf{Cx} + \mathbf{A}^T\mathbf{u} - \mathbf{p} \tag{5.9}$$

$$\mathbf{y} \geq 0, \ \mathbf{u} \geq 0 \qquad \mathbf{y}^T\mathbf{u} = 0 \tag{5.10}$$

Al sistema (5.8), (5.9) corresponde la siguiente matriz $\mathbf{T} \in M(m+n,\ 2m+n)$

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{C} & 0 \end{bmatrix} = [\alpha_1, \ldots, \alpha_n,\ \beta_1, \ldots, \beta_m,\ \gamma_1, \ldots, \gamma_m]$$

$\mathbf{T}_{B_1 B_2}$ denota una base de $\mathbf{T}$, que es formada por las columnas $\alpha_1, \alpha_2, \ldots, \alpha_n$, β_i; $i \in B_1$, γ_i; $i \in B_2$, donde $\mathrm{card}(B_1) + \mathrm{card}(B_2) = m$.

Fijamos primero tres conceptos que usaremos posteriormente

DEFINICION 5.2. - a) La base $\mathbf{T}_{B_1 B_2}$ es no degenerada, si para la solución básica asociada $(\mathbf{x}, \mathbf{y}, \mathbf{u})$ se cumple $y_i \neq 0$, $\forall\ i \in B_1$.

b) $T_{B_1 B_2}$ se llama *complementaria* si $B_1 \cap B_2 = \phi$.

c) $T_{B_1 B_2}$ se llama *semicomplementaria* si $B_1 \cap B_2 = \{j\}$

En el último caso, si $T_{B_1 B_2}$ es no degenerada en el sentido de la programación lineal, se cumple la solución básica asociada $(\mathbf{x}, \mathbf{y}, \mathbf{u})$

$$y_j \neq 0 \qquad u_j \neq 0$$

$$u_r = y_r = 0 \quad \text{para exactamente un } r$$

$$y_i u_i = 0 \ , \quad 1 \leq i \leq m \ , \quad i$$

<u>Hipotesis:</u> Todas las bases complementarias y semicomplementarias son no degeneradas.

lo que sigue ξ^k será la única solución óptima de $\mathbf{P}_k$.

A. Vamos a describir en detalle la primera iteración del algoritmo de DANTZIG Y COTTLE.

Elegimos el punto inicial como sigue

Determinamos una base complementaria $T_{B_1 B_2}$, tal que, para la solución básica asociada $(\mathbf{x}^o, \mathbf{y}^o, \mathbf{u}^o)$ se cumple :

$\mathbf{y}^o \geq 0$ (o que $\mathbf{x}^o$ sea admisible para $\mathbf{Q}$).

Como $T_{B_1 B_2}$ es complementaria y no degenerada, se cumple:

$$B_2 = I(\mathbf{x}^o)$$

Elegimos $\mathbf{x}^o$. Como el punto inicial y el programa inicial $\mathbf{P}_o$ tiene la forma :

$$\mathbf{P} \qquad \min \{F(\mathbf{x}) \ / \ \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i \qquad i \in B_{\cdot} \quad I(\mathbf{x}^o)\}$$

como $\mathbf{u}_i = 0 \qquad \forall\ i \quad B_1$ (5.11) se obtiene de (5.9)

$$\nabla F(x^o) = Cx^o + p - - A^T u^o - \sum_{j\in B_2} u^o_j a_j - - \sum_{j\in I(x^o)} u^o_j a_j$$

Por consiguiente en (x^o, u^o) se cumplen las condiciones linealizadas de K-T para P_o y por esto x^o es óptimo para P_o:

$$\xi^o = x^o$$

Si $u^o \geq 0$, entonces (x^o, y^o, u^o) es solución del sistema (5.8) (5.9) y (5.10).

Por consiguiente x^o es óptimo para Q y el algoritmo termina.

Si $u^o_j = \min \{u^o_i \;/\; 1 \leq i \leq m\} \lneq 0$, entonces $x^1 = \xi^o$, puesto que ξ^o es admisible para Q, y el nuevo subprograma P_1 tiene la forma

$$\min \{F(x) \;/\; a_i^T x \quad b_i \quad i \in B_2 \setminus \{j\}\}$$

Ademas los vectores γ_i, $i \in B_2$ y con esto los vectores $i \in B_2$ son linealmente independientes y se cumple la hipótesis de la proposición 5.1 y del teorema 5.1).

Para la solución óptima ξ^1 de P_1 se cumple de acuerdo con la proposición 5.1

$$F(\xi^1) < F(x^1) = F(x^o) \qquad b_j$$

Si $y^1 - b - A\xi^1$, entonces $y^1_j = b_j \qquad 0$ (5.12)

Como ξ^1 es óptimo para P_1, existe un u^1 tal que en (ξ^1, u^1) se cumplen las condicones linealizadas de K-T para P_1. Entonces (ξ^1, y^1, u^1) resuelve el sistema (5.8), (5.9) y se cumple:

$$u^1_i = 0 \;, \;\forall\; i \in B_1, \; u^1_j - 0 \qquad (5.13)$$

Si $\bar{z} = (\mathbf{x}^0, \mathbf{y}^0, \mathbf{u}^0) = (\mathbf{x}^1, \mathbf{y}^0, \mathbf{u}^0)$ y $\bar{\bar{z}} = (\xi^1, \mathbf{y}^1, \mathbf{u}^1)$ entonces se cumple por (5.11), (5.12) y (5.13)

$$T(\bar{\bar{z}}-\bar{z}) = \sum_{i=1} (\xi_i^1 \quad x_i^1)\alpha_i + \sum_{i\in B_1} (y_i^1 - y_i^0)\beta_i +$$

$$+ \sum_{i\in B_2} (u_i^1 \quad u_i^0)\gamma_i + y_j^1\beta_j = T_{B_1 B_2}(\bar{\bar{z}} - \bar{z}) + y_j^1\beta_j = 0$$

Luego

$$\bar{\bar{z}} - \bar{z} = \quad y_j^1 (T_{B_1 B_2})^{-1}\beta_j$$

Por consiguiente se obtiene $\bar{\bar{z}}$ de $\bar{z}$ de la siguiente manera:

Se hace entrar en la base $T_{B_1 B_2}$ la columna β_j avanzando la dirección de intercambio $v = -(T_{B_1 B_2})^{-1}\beta_j$, hasta que γ_j salga de la base.

Si $\bar{B}_1 = B_1 \cup \{j\}$, $\bar{B}_2 = B_2 \setminus \{j\}$, entonces $\bar{\bar{z}}$ es la solución básica asociada a $T_{\bar{B}_1 \bar{B}_2}$ que es evidentemente una base complementaria.

Sea primero ξ^1 admisible para Q ó $\mathbf{y}^1 \geq 0$

Si $\mathbf{u}^1 \geq 0$, entonces $(\xi^1, \mathbf{y}^1, \mathbf{u}^1)$ satisface el sistema (5.8), (5.9) y (5.10).

Por consiguiente ξ^1 es óptimo para Q y el algoritmo termina.

Si $\mathbf{u}_\ell^1 = \min \quad / \ 1 \leq i \leq m\} < 0$, entonces $\mathbf{x}^2 = \xi^1$ y

$$\mathbf{P}_2 \quad \min \{F(\mathbf{x}) \quad a_i^T \mathbf{x} - \mathbf{b}_i \quad \in B_2 \setminus \{\ell\}\}$$

Como $\mathbf{u}_j^1 = 0 > \mathbf{u}_\ell^1$, se tiene $j \neq \ell$, $B_2 \setminus \{j\} = B_2 \setminus \{\ell, j\}$.

Si ξ^1 no es admisible para Q; entonces $\mathbf{x}^2$ es el punto en $[\mathbf{x}^1,\xi^1] \cap S$ más cercano a $\mathbf{x}^1$.

Obtenemos entonces una solución $z = (\mathbf{x}^2,\bar{\mathbf{y}}^2,\bar{\mathbf{u}}^2)$ de (5.8), (5.9) como sigue :

$$z^2 = z + \lambda_2 \mathbf{v}$$

donde

$$\lambda_2 = \min \left\{ \frac{y_i^0}{-v_i} \; / \; i \in B_1, \; v_i \quad 0 \right\} = \frac{y_r}{-v_r}$$

Entonces β_r sale de la base $T_{B_1 B_2}$. Como $T_{B_1 B_2}$ es no degenerada, se cumple $\lambda_2 > 0$ y $\mathbf{x}^2 \in \langle \mathbf{x}^1,\xi^1 \rangle$.

Si $B_1 \quad B_1 \cup \{j\} \setminus \{r\}$, $B_2 = B_2$, entonces $\bar{z}^2 = (\mathbf{x}^2,\bar{\mathbf{y}}^2,\bar{\mathbf{u}}^2)$ es la solución básica asociada $T_{\bar{B}_1 \bar{B}_2}$ Como es más cercano a z que a $z = (\xi^1,y^1,u^1)$, $x^2 \in S$ y $\xi^1 \notin S$ se ple $y_r^1 < 0$, < 0 $\bar{\mathbf{y}}_r^2 = 0$.

Como $B_1 \cap B_2 = \{j\}$, $r \notin B_1 \cup B_2$, $T_{\bar{B}_1 \bar{B}_2}$ es semicomplementaria. Se cumple entonces : $I(\mathbf{x}^2) = J \setminus \bar{B}_1 = B_2 \cup \{r\} \setminus \{j\}$, puesto que por hipótesis $T_{\bar{B}_1 \bar{B}_2}$ es degenerada. Por consiguiente $\mathbf{P}_2$ tiene la siguiente forma

$\mathbf{P}_2 \quad \min \{ F(\mathbf{x}) \; / \; \mathbf{a}_i^T x - \mathbf{b}_i \quad i \in B_2 \cup \{r\} \setminus \{j\} \}$ ■

B. Vamos ha demostrar por inducción el siguiente

TEOREMA 5.2.

1). Para todo punto de iteración ξ^k, el punto (ξ^k, y^k, u^k) es una solución básica asociada a una base complementaria $T_{B_1 B_2}$.

2). Para todo punto de iteración x^k se presenta exactamente uno de los casos α) o β).

α). ξ^{k-1} es admisible para Q. Entonces = y $(\xi^{k-1}, y^{k-1}, u^{k-1})$ es una solución básica asociada a una base complementaria $T_{B_1 B_2}$.

β). ξ^{k-1} no es admisible para Q. Entonces a x^k corresponde una solución básica $(x^k, \bar{y}^k, \bar{u}^k)$ asociada a una base semicomplementaria $T_{B_1^* B_2^*}$.

OBSERVACIONES 5.4:

1). De acuerdo con el teorema 5.2, a todo punto de iteración x^k corresponde una solución básica $(x^k, \bar{y}^k, \bar{u}^k)$ associada a una base $T_{B_1 B_2}$ que es complementaria o semicomplementaria $T_{B_1 B_2}$ es complementaria sí y solamente si ξ^{k-1} es admisible para Q [y $x^k = \xi^{k-1}$]

2). Si ξ^k es admisible para Q y (ξ^k, y^k, u^k) es asociada a la base complementaria $T_{B_1 B_2}$, entonces $x^{k+1} = \xi^k$ y $I(x^{k+1}) = B_2$. puesto que por hipótesis $T_{B_1 B_2}$ es no degenerada. Los vectores γ_i : $i \in B_2 = I(x^{k+1})$ y con esto también los vectores $i \in B_2$ son linealmente independientes.

Por consiguiente la hipótesis del teorema 5.1 son cumplidas.

DEMOSTRACION (Del teorema 5.2, por inducción)

Al inicio $\xi^0 = \mathbf{x}^0$ es admisible para $\mathbf{Q}$ y $(\xi^0, \mathbf{y}^0, \mathbf{u}^0)$ es la solución básica asociada a una base complementaria $\mathbf{T}_{B_1 B_2}$. Por consiguiente $\mathbf{x}^1 = \xi^0$ y en la primera iteración se presenta el caso α).

Si $k \geq 1$, es arbitrario, hacemos la siguiente

Hipótesis de inducción:

Para $\ell \leq k$, la base en $(\xi^{\ell-1}, \mathbf{y}^{\ell-1}, \mathbf{u}^{\ell-1})$ es complementaria y en $\mathbf{x}^{\ell}$ se presenta uno de los casos α) o β).

Tenemos que demostrar que la base en $(\xi^k, \mathbf{y}^k, \mathbf{u}^k)$ es complementaria y que para $\mathbf{x}^{k+1}$ se presenta otra vez uno de los casos α) o β).

i) Asumimos primero que $\mathbf{x}^k$ satisface α) y que $\mathbf{T}_{B_1 B_2}$ es la base complementaria en $(\xi^{k-1}, \mathbf{y}^{k-1}, \mathbf{u}^{k-1}) = (\mathbf{x}^k, \mathbf{y}^{k-1}, \mathbf{u}^{k-1})$.

Podemos asumir que ξ^{k-1} no es óptimo para $\mathbf{Q}$

[o que $\mathbf{u}^k \ngeq 0$].

Usando exactamente los mismos argumentos como en $\underline{\mathbf{A}}$., obtenemos los siguientes resultados

Si $0 > \mathbf{u}_j^{k-1} = \min \{\mathbf{u}_i^{k-1} \ / \ 1 \leq i \leq m\}$,

$\bar{B}_1 = B_1 \cup \{j\}$, $\bar{B}_2 = B_2 \setminus \{j\}$, entonces $\mathbf{T}_{\bar{B}_1 \bar{B}_2}$ es la base en $(\xi^k, \mathbf{y}^k, \mathbf{u}^k)$, que es evidentemente complementaria.

Si ξ^k es admisible para $\mathbf{Q}$, entonces para $\mathbf{x}^k$ se presenta otra vez el caso α).

ξ^k no es admisible para $\mathbf{Q}$, entonces a ${}^{k+1}$ corresponde la solución básica $(\mathbf{x}^{k+1},\bar{\mathbf{y}}^{k+1},\bar{\mathbf{u}}^{k+1})$, asociada a la base $T_{\tilde{B}_1\tilde{B}_2}$, donde $B_1 = B_1 \cup \{j\} \setminus \{r\}$, $B_2 = B_2$, para algun $r \in B_1$, con $\mathbf{y}_r^k < 0$.

Observamos que $\bar{\mathbf{y}}_j^{k+1}$ 0, $\bar{\mathbf{u}}_j^{k+1}$ 0. Como $B_1 \cap B_2 = \{\,\}$, $r \notin B_1 \cup B_2$, $T_{\tilde{B}_1\tilde{B}_2}$ es semicomplementaria.

Por consiguiente para $\mathbf{x}^{k+1}$ se presenta el caso β)

ii). Asumimos ahora que en $\mathbf{x}^k$ se presenta el caso β)

Sean $T_{B_1B_2}$, $T_{\bar{B}_1\bar{B}_2}$, $T_{B_1^*B_2^*}$, las bases en $(\mathbf{x}^{k-1},\bar{\mathbf{y}}^{k-1},\bar{\mathbf{u}}^{k-1})$, $(\xi^{k-1},\mathbf{y}^{k-1},\mathbf{u}^{k-1})$ y $(\mathbf{x}^k,\bar{\mathbf{y}}^k,\bar{\mathbf{u}}^k)$.

iia). Consideramos primero el caso donde $B_1 \cap B_2 = \phi$. Entonces ξ^{k-1} es admisible para Q y $\mathbf{x}^{k-1} = \xi^{k-2}$. Usando mismos argumentos como en <u>A.</u>, concluimos que $\mathbf{P}_{k-1}$ es como sigue

$$\mathbf{P}_{k-1} \quad \min\{F(\mathbf{x}) \; / \; \mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i ; \; i \in B_2 \setminus \{j\}\} \tag{5.14}$$

donde $\mathbf{u}_j^{k-2} = \bar{\mathbf{u}}_j^{k-1} < 0$ y que $\bar{B}_1 = B_1 \cup \{j\}$, $\bar{B}_2 = B_2 \setminus \{j\}$

$$B_1^* = B_1 \cup \{j\} \setminus \qquad B_2^* = B_2, \text{ donde } \mathbf{y}_r^{k-1} < 0 - \bar{y}_r^k \tag{5.15}$$

Entonces $B_1^* \cap B_2^* =$ $\quad r \notin B_1^* \cup B_2^*$. Observamos en particular que $\bar{\mathbf{u}}_j^k < 0$. Como la base semicomplementaria $T_{B_1^*B_2^*}$ es no degenerada, se cumple :

$$I(x^k) = J \setminus B_1^* - \quad \cup \{r\} \setminus \{j\}$$

y P_k tiene la forma :

$$\min \{ F(x) \;/\; \mathbf{a}_i^T x = \mathbf{b}_i ;\; i \in B_2 \cup \{r\} \setminus \{j\} \} \qquad (5.16)$$

Para la solución óptima ξ^k de $\mathbf{P}_k$ de las condiciones de K-T se deduce que : $\mathbf{u}_i^k = 0$, $\forall\, i \in B_1^*$, en particular $\mathbf{u}_j^k = 0$.

<u>Vamos a demostrar que $\mathbf{u}_r^k > 0$</u>

Asumamos que $\mathbf{u}_r^k \leq 0$. Entonces en $(\xi^k, \mathbf{u}^k)$ se cumplen las condiciones linealizadas de K-T para el programa :

$$\min\{F(\mathbf{x})/\mathbf{a}_i^T\mathbf{x} = \mathbf{b}_i ;\; i \in B_2\setminus\{j\},\; \mathbf{y}_r = \mathbf{b}_r - \mathbf{a}_r^T\mathbf{x} \leq 0\} \qquad (5.17)$$

Por consiguiente ξ^k es óptimo para (5.17) y por consiguiente $F(\xi^{k-1}) \geq F(\xi^k)$. De otro lado de acuerdo con (5.17) y (5.14) se cumple

$$F(\xi^{k-1}) = \inf(\mathbf{P}_{k-1}) \leq \inf(\mathbf{p}_k^*) = F(\xi^k)$$

En consecuencia $F(\xi^k) = F(\xi^{k-1})$ y ξ^{k-1}, ξ^k son soluciones óptimas de (5.17). Como F es estrictamente convexa, (5.17) tiene exactamente una solución óptima. Entonces $\xi^k = \xi^{k-1}$ y de (5.8) obtenemos que $\bar{\mathbf{y}}^k = \mathbf{y}^{k-1}$ Pero esto es imposible, puesto que $\mathbf{y}_r^{k-1} < 0 = \bar{\mathbf{y}}_r^k$. Por consiguiente $\mathbf{u}_r^k > 0$.

Si $z^k = (\xi^k, \mathbf{y}^k, \mathbf{u}^k)$, $\bar{z}^k = (\mathbf{x}^k, \bar{\mathbf{y}}^k, \bar{\mathbf{u}}^k)$, entonces se cumple:

$$\mathbf{T}(z^k - \bar{z}^k) = \sum_{i=1}^{n} (\xi_i^k - x_i^k)\alpha_i - \sum_{i \in B_1^*} (y_i^k - \bar{x}_i^k)\beta_i$$

$$+ \quad (u_i^k - \bar{u}_i^k)\gamma_j + u_r^k\gamma_r$$

$$\mathbf{T}(z^k - \bar{z}^k) = T_{B_1^* B_2^*}(z^k - \bar{z}^k) + u_r^k\gamma_r = 0 \qquad 6$$

$$z^k = \bar{z}^k \quad u_r^k (T_{B_1^* B_2^*})^{-1} \gamma_r$$

En consecuencia se obtiene z^k de $\bar{z}^k$ como sigue

Se introduce la columna γ_r en la base $T_{B_1^* B_2^*}$ y se avanza la dirección de intercambio hasta que γ_j salga de la base, puesto que $u_j^k = 0$ y $\bar{u}_j^k < 0$.

Si $\tilde{B}_1 = B_1^* = B_1 \cup \{j\} \setminus \{r\}$ y

$\tilde{B}_2 = B_2^* \cup \{j\} \setminus \{r\} = B_2 \; \{j\} \setminus \{r\}$

entonces $T_{\tilde{B}_1 \tilde{B}_2}$ es la base en $z^k = (\xi^k, y^k, u^k)$. $T_{\tilde{B}_1 \tilde{B}_2}$ es plementaria porque $T_{B_1 B_2}$ lo es.

Si ξ^k es admisible para Q, entonces $x^{k+1} = \xi^k$ y en x^{k+1} se presenta el caso α)

- Si ξ^k no es admisible para Q, entonces se cumple

$$\bar{z}^{k+1} = (x^{k+1}, \bar{y}^{k+1}, \bar{u}^{k+1}) = \bar{z}^k + \lambda_k v$$

$$\text{donde } \lambda_k = \min \left\{ \frac{\bar{y}_i^k}{-v_i} \;/\; v_i < 0,\; i \in B_1^* \right\} = \frac{-k}{\ell \atop -v_\ell} > 0,$$

donde $y_\ell^k < 0 = {}^k_r$ y por esto $r \neq \ell$.

Entonces β_ℓ sale de la base.

Si $\tilde{B}_1 = B_1^* \setminus \{\ell\} = B_1 \cup \{j\} \setminus \{r, \ell\}$ y

$B_2 = B_2^* \cup \{r\} = B_2 \cup \{r\}$

entonces $T_{\tilde{B}_1 \tilde{B}_2}$ es la base en z^{k+1} Como $\tilde{B}_1 \cap \tilde{B}_2 = \{j\}$ y $\ell \notin \tilde{B}_1 \cup \tilde{B}_2$; la base $T_{\tilde{B}_1 \tilde{B}_2}$ es semicomplementaria. En consecuencia en x^{k+1} se presenta nuevamente el caso β).

iib) Sea ahora $T_{B_1 B_2}$ semicomplementaria. $B_1 \cap B_2 = \{j\}$ y

$r \not\in B_1 \cup B_2$. Entonces $\mathbf{x}^{k-1} \neq \xi^{k-2}$. $T_{B_1 B_2}$ es no degenerada

se cumple $I(\mathbf{x}^{k-1}) = J \setminus B_1 = B_2 \cup \{r\} \setminus \{j\}$ y $\mathbf{P}_{k-1}$ tiene la forma:

$$\mathbf{P}_{k-1} \qquad \min \{F(\mathbf{x}) \ / \ \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i \qquad i \in B_2 \cup \{r\} \setminus \{j\}\} \qquad (5.18)$$

Asumimos inductivamente que

$$B_1 \qquad B_2 = B_2 \cup \{r\} \setminus \{j\}$$

$$B_1^* = B_1 \setminus \{\ell\} \qquad B_2^* = B_2 \cup \{r\} ,$$

donde $y_\ell^{k-1} < 0$, $u_j^k < 0$.

Vamos a demostrar que en $\mathbf{x}^{k+1}$ se presenta el caso α) o que en la transición de $\mathbf{x}^k$ a $\mathbf{x}^{k+1}$ se producen los mismos cambios de base, lo que significa que en $\mathbf{x}^{k+1}$ se presenta el caso β).

Se cumple $B_1^* \cap B_2^* = \{j\}$, $\ell \not\in B_1^* \cup B_2^*$ y $\mathbf{T}_{B_1^* B_2^*}$ es en realidad semicomplementaria. Como $\mathbf{T}_{B_1^* B_2^*}$ es degenerada se cumple: $I(\mathbf{x}^k) = J \setminus B_1^* = B_2 \cup \{r, \ell\} \setminus \{j\}$

$$\mathbf{P}_k \ : \ \min \{F(\mathbf{x}) \ / \ \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i \qquad i \in B_2 \cup \{r, \ell\} \setminus \{j\}\} \qquad (5.19)$$

Para la solución ξ^k de $\mathbf{P}_k$ se deduce de las mismas condiciones linealizadas de K-T para $\mathbf{P}_k$:

$\mathbf{u}_i^k = 0$, $\forall$ i $\in B_1^* = B_1 \setminus \{j\}$, en particular $\mathbf{u}_j^k = 0$

<u>Vamos a demostrar que $\mathbf{u}_\ell^k > 0$</u>

Asumamos que $\mathbf{u}_\ell^k \leq 0$. Entonces en $(\xi^k, \mathbf{u}^k)$ se cumplen las condiciones linealizadas de K-T para el programa

$$P_k: \quad \min \left\{ F(x) \ / a_i^T x - b_i \quad i \in B_2 \cup \{r\} \setminus \{j\}, \quad y_\ell = b_\ell \quad a_\ell x \leq 0 \right. \qquad (5.20)$$

Por consiguiente ξ^k es tambien óptimo para (5.20). De $y_\ell^{k-1} < 0$ y (5.18) se deduce que ξ^{k-1} es tambien admisible para (5.20) y por esto $F(\xi^{k-1}) \geq F(\xi^k)$.

De otro lado si S_{k-1} y $\tilde{S}_k$ son los mismos dominios admisi bles de P_{k-1} y $\tilde{P}_k$. Entonces $S_{k-1} \supset \tilde{S}_k$, lo que implica que $F(\xi^{k-1}) \leq F(\xi^k)$. Por consiguiente $F(\xi^{k-1}) = F(\xi^k)$ y ξ^{k-1} y ξ^k son óptimos para (5.20), Como F es estrictamente convexa debe cumplir $\xi^{k-1} = \xi^k$ y de (5.8) se obtiene $y^k = y^{k-1}$

Pero esto no ès posible por que $y_\ell^{k-1} < 0 = y_\ell^k$.

Por consiguiente $y_\ell^k > 0$.

Para $z^k = (\xi^k, y^k, u^k)$ y $\bar{z}^k = (x^k, \bar{y}^k, \bar{u}^k)$ se cumple :

$$T(z^k - \bar{z}^k) = \sum_{i=1}^{n} (\xi_i^k \quad x_i^k)\alpha_i - \sum_{i \in B_1^*} (y_i^k - \bar{y}_i^k)\beta_i$$

$$+ \sum_{i \in B_2^*} (u_i^k \quad \bar{u}_i^k)\gamma_i + u_\ell^k \gamma_\ell = 0$$

$$z^k = \bar{z}^k \quad u_\ell^k (T_{B_1^* B_2^*})^{-1} \gamma_\ell.$$

Por consiguiente de $\bar{z}^k$ se obtiene z^k, introduciendo la columna γ_ℓ en la base $T_{B_1^* B_2^*}$, y avanzando en la dirección $v = -(T_{B_1^* B_2^*})^{-1}\gamma_\ell$ hasta que γ_j salga de la base.

Si $\tilde{B}_1 = B_1^* \setminus \{\ell\}$, $\tilde{B}_2 - B_2 \cup \{\ell\} \setminus \{j\} = B_2 \cup \{r, \ell\} \setminus \{j\}$

entonces $T_{\tilde{B}_1 \tilde{B}_2}$ es la base en $z^k = (\xi^k, \mathbf{y}^k, \mathbf{u}^k)$,

$\left[\text{como } y_{\ell}^{k-1} < 0 = y_r^{k-1} \quad \text{se tiene } r \neq \ell \right]$

Como $B_1 \cap B_2 = \{j\}$, $r \not\in B_1 \cup B_2$, se cumple $\quad B_2 - \phi$

Si ξ^k es admisible para Q, entonces $\quad - \xi^k$ y en $\mathbf{x}^{k+1}$ se presenta el caso α).

- Si ξ^k no es admisible para Q, entonces se obtiene

$z = (\mathbf{x}^{k+1}, \bar{\mathbf{y}}^{k+1}, \bar{\mathbf{u}}^{k+1})$ como sigue :

$\bar{z}^{k+1} = \bar{z}^k + \lambda_k \mathbf{v}$, donde

$$\lambda_k = \min \left\{ \frac{\bar{y}_i^k}{-\mathbf{v}_i} \;/\; i \in B_1^*,\ \mathbf{v}_i < 0 \right\} = \frac{\bar{y}_t^k}{-\mathbf{v}_t} > 0 \quad y \quad \mathbf{y}_t^k \quad 0.$$

Observamos que $\bar{\mathbf{u}}_j^{k+1} < 0$, puesto que $\bar{\mathbf{u}}_j^k < 0 = \mathbf{u}_j^k$ y $\quad {}^{-k+1}$

está más cercano a $\bar{z}^k$ que z^k. Entonces β_t sale de la base.

Si $\hat{B}_1 = B_1^* \setminus \{t\} = B_1 \setminus \{\ell, t\}$

$\hat{B}_2 = B_2^* \cup \{\ell\} = B_2 \cup \{r, \ell\}$, entonces

$T_{\hat{B}_1 \hat{B}_2}$ es la base en $\bar{z}^{k+1} = (\mathbf{x}^{k+1}, \bar{\mathbf{y}}^{k+1}, \bar{\mathbf{u}}^{k+1})$,

como $\qquad 0 = \mathbf{y}_{\ell}^k$, se tiene $\ell \neq t$.

Como $B_1^* \cap B_2^* = \{j\}$, $\ell \not\in B_1^* \cup B_2^*$ se cumple

$\hat{B}_1 \cap \hat{B}_2 = \{j\}$, $t \not\in \hat{B}_1 \cup \hat{B}_2$. Por consiguiente en $\mathbf{x}^{k+1}$ se presenta el caso β).

Para las bases $T_{B_1^* B_2^*}$, $T_{\tilde{B}_1 \tilde{B}_2}$, $T_{\hat{B}_1 \hat{B}_2}$ en los puntos

$\bar{z}^k = (\mathbf{x}^k, \bar{\mathbf{y}}^k, \bar{\mathbf{u}}^k)$, $z^k = (\xi^k, \mathbf{y}^k, \mathbf{u}^k)$ y $\bar{z}^{k+1} = (\mathbf{x}^{k+1}, \bar{\mathbf{y}}^{k+1}, \bar{\mathbf{u}}^{k+1})$

se cumple :

$$B_1^* \cap B_2^* = \{j\}, \quad \ell \notin B_1^* \quad B_2^*$$

$$B_1 = B_1^* \quad , \tilde{B}_2 = B_2^* \cup \{\ell\} \setminus \{j\}$$

$$\hat{B}_1 = B_1 \setminus \{t\} \quad \hat{B}_2 = B_2^* \cup \{\ell\}, \text{ donde } \mathbf{y}_t^k < 0, \ \bar{\mathbf{u}}_j^{k+1} \quad 0$$

$T_{B_1^* B_2^*}$ y $T_{\hat{B}_1 \hat{B}_2}$ son semicomplementarias, mientras que $T_{\tilde{B}_1 \tilde{B}_2}$ es complementaria.

En consecuencia las bases que se construyen en transición de $\mathbf{x}^k$ a $\mathbf{x}^{k+1}$ tienen exactamente las mismas características como las bases que corresponden a la transición de $\mathbf{x}^{k-1}$ a $\mathbf{x}^k$. ∎

De la demostración del teorema 5.2. podemos deducir que el algoritmo de ZUNGWILL aplicado al programa cuadrático $\mathbf{Q}$ toma una forma muy especial. Esta forma no es otra cosa que el algoritmo de DANTZIG y COTTLE.

Describimos brevemente esta forma especial

sean $T_{B_1 B_2}$ y $T_{\bar{B}_1 \bar{B}_2}$ las bases de $\mathbf{T}$ en los puntos $z^k = (\mathbf{x}^k, \bar{y}^k, \bar{\mathbf{u}}^k)$ y $\bar{z}^{k+1} = (\mathbf{x}^{k+1}, \bar{\mathbf{y}}^{k+1}, \bar{\mathbf{u}}^{k+1})$

a) Sea $\mathbf{T}_{B_1 B_2}$ complementaria

Si $\bar{\mathbf{u}}^k \geq 0$, entonces $\mathbf{x}^k$ es óptimo para $\mathbf{P}_k$ y el algoritmo termina.

Si $0 > \bar{\mathbf{u}}_j^k = \min \{\bar{\mathbf{u}}_i^k \ / \ 1 \leq i \leq m\}$, entonces se avanza desde $\bar{z}^k$ en la dirección $v = -(\mathbf{T}_{B_1 B_2})^{-1} \beta_j$, hasta que i) o ii) se cumple :

i) γ_j salga de la base

ii) La primera de las variables básicas $\bar{y}_i^k$, digamos $\bar{y}_\ell^k$ se anula.

Si primero se presenta (i) entonces

$B_1 = B_1 \cup \{j\}$, $\bar{B}_2 = B_2 \setminus \{j\}$

En otro caso : $\bar{B}_1 = B_1 \cup \{j\} \setminus \{\ell\}$

b) Si $T_{B_1 B_2}$ es semicomplementaria, $B_1 \cap B_2 = \{j\}$, $\ell \notin B_1 \cup B_2$ se avanza desde $\bar{z}^k$ en la dirección $v = (T_{B_1 B_2})^{-1} \gamma_\ell$, hasta que (i) o (ii) se cumple

(i) γ_j sale de la base

(ii) La primera digamos $\bar{y}_t^k$ de las variables positivas $\bar{y}_i^k$ $i \in B_1$ se anula.

Si (i) se presenta antes que (ii), entonces $B_1 = B_1$,

$B_2 = B_2 \cup \{\ell\} \setminus \{j\}$.

En caso contrario $B_1 = B_1 \setminus \{t\}$, $B_2 = B_2 \cup \{\ell\}$.

a) y b) nos permiten fomalizar el

5.2.1: ALGORITMO DE DANTZIG Y COTTLE (RESUMEN)

Iteración inicial

Se determina una base complementaria $T_{B_1 B_2}$ de T tal que para la solución básica asociada $z = (x^o, y^o, u^o)$ se cumple $y^o \geq 0$. z^o es el punto inicial.

Iteración General:

a) Si $T_{B_1 B_2}$ es complementaria y $\bar{u}^k \geq 0$. entonces x^k es óptimo para Q y el algoritmo termina.

Si $0 > \mathbf{u}_j^k = \min \{\mathbf{u}_i^k \;/\; 1 \le i \le m\}$ entonces se calcula

$$\mathbf{v} = -\,(\mathbf{T}_{B_1 B_2})^{-1}\beta_j$$

$$\min\left\{ \frac{\bar{y}_i^k}{-v_i} \;/\; i \in B_1 \;,\; \mathbf{v}_i < 0 \right\} = \frac{\bar{y}_\ell^k}{-v_\ell}$$

$$\min\left\{ \frac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|} \;,\; \lambda_k^* \right\} \;,\; \bar{z}^{k+1} = \bar{z}^k + \lambda_k \mathbf{v}$$

Si $\lambda_k = \dfrac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|}$, entonces $\bar{B}_1 = B_1 \cup \{j\}$, $B_2 = B_2 \setminus \{j\}$

Si $\lambda_k = \lambda_k^* \quad \dfrac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|}$, entonces $\bar{B}_1 = B_1 \cup \{j\} \setminus \{\ell\}$,

$B_2 = B_2$

b) Si $\mathbf{T}_{B_1 B_2}$ es semicomplementaria, $B_1 \cap B_2 = \{j\}$, $\ell \notin B_1 \cup B_2$ entonces se cumple

$$= -\,(\mathbf{T}_{B_1 B_2})^{-1}\gamma_\ell$$

$$\lambda_k^* = \min\left\{ \frac{\bar{y}_i^k}{-v_i} \;/\; i \in B_1 \;,\; \mathbf{v}_i < 0 \right\} = \frac{\bar{y}_t^k}{-v_t}$$

$$\lambda_k = \min\left\{ \lambda_k^* \;,\; \frac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|} \right\} \;,\; \bar{z}^{k+1} = \bar{z}^k + \lambda_k \mathbf{v}$$

Si $\lambda_k = \dfrac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|}$, entonces $B_1 = B_1$, $B_2 = B_2 \cup \{\ell\} \setminus \{j\}$

Si $\lambda_k = \lambda_k^* < \dfrac{|\bar{\mathbf{u}}_j^k|}{|\mathbf{v}_j|}$, entonces $B_1 = B_i \cup \{t\}$

$B_2 = B_2 \cup \{\ell\}$

OBSERVACIONES

1) Bajo las hipótesis del presente párrafo el algoritmo de DANTZIG y COTTLE es finito en sentido propio. Todos puntos de iteración $\bar{z}^k$ son soluciones básicas asociadas a bases de **T** de la forma $\mathbf{T}_{B_1B_2}$ que son complementarias o semicomplementarias. $\mathbf{T}_{B_1B_2}$ es exactamente complementaria si la solución óptima ξ^{k-1} de $\mathbf{P}_{k-1}$ es admisible para **Q**. En este caso se cumple : $\mathbf{x}^k = \xi^{k-1}$, $I(\mathbf{x}^k) = B_2$, de acuerdo con la demostración precedente y los vectores $\mathbf{a}_i$, $i \in B_2$ son linealmente independientes. Por consiguiente ξ^{k-1} es no degenerada siempre y cuando ξ^{k-1} es admisible para **Q**, y las hipótesis del teorema 5.1 se cumplen.

En consecuencia el algoritmo conduce en un número finito de iteraciones a una solución óptima de **Q** y cada iteración requiere de un número finito de iteraciones.

2) Como el algoritmo SIMPLEX, el algoritmo de DANTZIG y COTTLE se deja ejecutar, usando una tabla que representa el sistema : $\mathbf{T}z = \mathbf{b}^*$, $(\mathbf{b}^*)^{\mathsf{T}} =$ $- \mathbf{p}^{\mathsf{T}})$, y el algoritmo se distingue del algoritmo SIMPLEX sólo en las reglas de selección para el cambio de base.

3) La base inicial $\mathbf{T}_{B_1B_2}$ se puede obtener como sigue
Se determina una base admisible $\mathbf{A}_{B_2}$ de $\mathbf{Ax} \leq \mathbf{b}$. Si $B_1 = \{1,2,\ldots,m\} \setminus B_2$, entonces $\mathbf{T}_{B_1B_2}$ es la base inicial apropiada.

5.2.2: Ejemplo para el algoritmo de DANTZIG Y COTTLE:

El problema.

$$\min 2x_1^2 + x_2^2 - 48x_1 - 40x_2 = \frac{1}{2}\mathbf{x}^T\mathbf{C}\mathbf{x} + \mathbf{p}^T\mathbf{x}$$

$$\begin{aligned} x_1 + x_2 &\leq \\ x_1 &\leq 6 \\ x_1 + 3x_2 &\leq 18 \\ -x_1 &\leq 0, \\ -x_2 &\leq 0 \end{aligned}$$

Los datos:

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \mathbf{A}^T \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b \\ -p \end{bmatrix}$$

$$\mathbf{A}^T = \begin{bmatrix} 1 & 1 & 1 & -1 & 0 \\ 1 & 0 & 3 & 0 & -1 \end{bmatrix}, \quad \mathbf{C} \quad \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\bar{\mathbf{b}}^T = (8,\ 6,\ 18,\ 0,\ 0,\ 48,\ 40)$$

Tabla de datos

			β_1	β_2	β_3	β_4	β_5	γ_1	γ_2	γ_3	γ_4	γ_5	
	1	1	1	0	0	0	0	0	0	0	0	0	8
	1	0	0	1	0	0	0	0	0	0	0	0	6
	1	3	0	0	1	0	0	0	0	0	0	0	18
$(\mathbf{T},\mathbf{b}) =$	-1	0	0	0	0	1	0	0	0	0	0	0	0
	0	-1	0	0	0	0	1	0	0	0	0	0	0
	4	0	0	0	0	0	0	1	1	1	-1	0	48
	0	2	0	0	0	0	0	1	0	3	0	-1	40

Paso inicial:

Determinar una base complementaria $\mathbf{T}_{B_1 B_2}$ con la solución básica asociada $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{u}^0)$, tal que $\mathbf{y}^0 \geq 0$.

Elegimos la solución básica $\mathbf{x}^0 = (0,0)^T$

Entonces $\mathbf{y}^0 = \mathbf{b} \quad \mathbf{A}\mathbf{x}^0 = \mathbf{b} \geq 0$

Como $y_i^0 > 0 \quad i = 1,2,3 \qquad y_i^0 = 0$, $i = 4,\ 5.$

Elegimos: $B_1 = \{1,2,3\}$ $B_2 = \{4,5\}$. Obtenemos $T_{B_1B_2}$

$$T_{B_1B_2} = \begin{array}{c} \begin{array}{ccccccc} \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \beta_3 & \gamma_4 & \gamma_5 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \end{array}$$

$T_{B_1B_2}$ es en realidad una base, puesto que sus filas son L. I

Se cumple : $u_4^o = -48$, $u_5^o = -40$, $\underline{j = 4}$

Primera <u>iteración</u>

La base $T_{B_1B_2}$ es complementaria .

$\min\{u_4^o, u_5^o\} = \min\{-48, -40\} = -48 = u_4^o$; $\Longrightarrow$ $\underline{j = 4}$

Tenemos que resolver : $T_{B_1B_2}v = -\beta_4$ ó

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solución : $v^T = (1,0,-1,-1,-1,4,0)$

Calculamos :

$$\lambda_1^* = \min\left\{ \frac{y_i^o}{-v_i} \;/\; i \in B_1,\; v_i < 0 \right\} = \min\left\{ \frac{y_1}{-v_3}, \frac{y_2}{-v_4}, \frac{y_3}{-v_5} \right\}$$

$$= \min\left\{ \frac{6}{1}, \frac{6}{1}, \frac{18}{1} \right\} = 6 = \frac{y_2^o}{-v_4} , \quad \ell = 2$$

$$\frac{|u_4^0|}{|v_6|} = \frac{48}{4} = 12,$$

$$\lambda_1 = \min \left\{ \lambda_1^*, \frac{|u_4^0|}{|v_6|} \right\} = 6 = \lambda_1^*$$

$$\Longrightarrow B_1 := B_1 \cup \{j\} \setminus \{\ell\} = \{1,2,3\} \cup \{4\} \setminus \{2\} = \{1,3,4\}$$

$$B_2 := B_2 = \{4,5\}$$

Segunda iteración:

La base $T_{B_1B_2}$ es semicomplementaria

$$B_1 \cap B_2 = \{j\} = \{4\}, \qquad \ell = 2 \notin B_1 \cup$$

Tenemos que resolver los sistemas: $T_{B_1B_2} z = b$ $\quad T_{B_1B_2} v = -\gamma_2$

$$T_{B_1B_2} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Los sistemas son:

	sol.		sol.
$x_1 + x_2 + y_1 = 8$	$x_1 = 6$	$v_1 + v_2 + v_3 = 0$	$v_1 = 0$
$x_1 = 6$	$x_2 = 0$	$v_1 = 0$	$v_2 = 0$
$x_1 + 3x_2 + y_3 = 18$	$y_1 = 2$	$v_1 + 3v_2 + v_4 = 0$	$v_3 = 0$
$-x_1 + y_4 = 0$	$y_3 = 12$	$-v_1 + v_5 = 0$	$v_4 = 0$
$x_2 = 0$	$y_4 = 6$	$-v_2 = 0$	$v_5 = 0$
$4x_1 - u_4 = 48$	$u_4 = -24$	$4v_1 - v_6 = -1$	$v_6 = 1$
$2x_2 - u_5 = 40$	$u_5 = -40$	$2v_2 - v_7 = 0$	$v_7 = 0$

$\Longrightarrow \lambda_2^* = +\infty$ (pues no existe v_i 0

$$\frac{|u_4|}{|v_6|} = \frac{24}{1} = 24 \qquad \lambda_2 = 24$$

$\longrightarrow$ $B_1 := B_1 - \{1,3,4\}$

$B_2 := B_2 \cup \{\ell\} \setminus \{j\} = \{4,5\} \cup \{2\} \setminus \{4\} = \{2,5\}$

Tercera iteración:

$\mathbf{T}_{B_1 B_2}$ es complementaria

Tenmos que resolver $\qquad \mathbf{T}_{B_1 B_2} z = \mathbf{b}$

se cumple

$$\mathbf{T}_{B_1 B_2} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

El sistema es

solución

$$\begin{array}{rcl} x_1 + x_2 + y_1 & = & 8 \\ x_1 & = & 6 \\ x_1 + 3x_2 + y_3 & = & 18 \\ -x_1 + y_4 & = & 0 \\ -x_2 & = & 0 \\ 4x_1 + u_2 & = & 48 \\ 2x_2 - u_5 & = & 40 \end{array} \qquad \begin{array}{l} x_1 = \\ x_2 = \\ y_1 = 2 \\ y_3 = 12 \\ y_4 = \\ u_2 = 24 \\ u_5 = -40 \\ u_5 = -40 \quad 0 \Longrightarrow j = \end{array}$$

Teremos que resolver el sistema

$\mathbf{T}_{B_1 B_2} \mathbf{v} = -\beta = -$

Solución $\quad \mathbf{T} = \begin{array}{ccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 \end{array}$
$\quad \left(\begin{array}{ccccccc} 0, & 1, & -1, & -3, & 0, & 0, & \end{array} \right)$

$$\lambda_- = \min\left\{ \frac{y_1}{-v_3}, \frac{y_3}{-v_4} \right\} = \min\left\{ \frac{2}{1}, \frac{12}{3} \right\} = 2 = \frac{y_1}{-v_1} \qquad \ell =$$

$$\frac{|u_5|}{|v_7|} = \frac{40}{2} \qquad 20$$

$$\Longrightarrow \lambda_- = 2 = \lambda_3^*$$

$$\Longrightarrow B_1 := B_1 \cup \{j\} \setminus \{\ell\} = \{1,3,4\} \cup \{5\} \setminus \{1\} = \{3,4,5\}$$

$$\Longrightarrow B_2 := B_2 = \{2,5\}$$

Cuarta iteración:

La base $T_{B_1 B_2}$ es semicomplementaria

$B_1 \cap B_2 = \{j\} = \{5\}$ $\qquad \ell = 1 \notin B_1 \cup B_2$

se cumple

$$T_{B_1 B_2} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Tenemos que resover :

$$T_{B_1 B_2} z = \bar{b} \quad y \quad T_{B_1 B_2} v = -\gamma_1$$

El sistema es

solución

$$\begin{array}{rcl} x_1 + x_2 + & = & 8 \\ x_1 & = & 6 \\ x_1 + 3x_2 + y_3 & = & 18 \\ -x_1 + y_4 & = & 0 \\ -x_2 + y_5 & = & 0 \\ 4x_1 + u_2 & = & 48 \\ 2x_2 - u_5 & = & 40 \end{array} \qquad \begin{array}{rcl} x_1 & = & \\ x_2 & = & \\ y_3 & = & \\ y_4 & = & 6 \\ y_5 & = & [illegible] \\ u_2 & = & 24 \\ u_5 & = & -46 \end{array}$$

$$
\begin{aligned}
v_1 + v_2 \qquad\qquad &= 0 \\
v_1 \qquad\qquad\qquad &= 0 \\
v_1 + 3v_2 + v_3 &= \\
-v_1 + \qquad\quad v_4 &= 0 \\
-v_2 + v_5 &= 0 \\
4v_1 \qquad + v_6 &= -1 \\
2v_2 - v_7 &= -1
\end{aligned}
\qquad\qquad
\begin{aligned}
&\mathbf{v} = (0,0,0,0,0,-1,1) \\
&\lambda_4 = +\infty \\
&\frac{|u_5|}{|v_7|} = \frac{36}{1} = 36
\end{aligned}
$$

$\Longrightarrow \lambda_4 = 36$

$\Longrightarrow B_1 := B_1 = \{3,4,5\}$

$\Longrightarrow B_2 := B_2 \cup \{\ell\} \setminus \{j\} = \{2,5\} \cup \{1\} \setminus \{5\} = \{1,2\}$

Quinta iteración:

$\mathbf{T}_{B_1B_2}$ es complementaria

$$
\mathbf{T}_{B_1B_2} \quad \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}
$$

Tenemos que resover

$$\mathbf{T}_{B_1B_2} z = \mathbf{b}$$

solución

$$
\begin{aligned}
x_1 + x_2 \qquad\qquad &= 8 \\
x_1 \qquad\qquad\qquad &= 6 \\
x_1 + 3x_2 + y_3 &= 18 \\
-x_1 + \qquad\quad y_4 &= 0 \\
-x_2 + y_5 &= 0 \\
4x_1 + u_1 + u_2 &= 48 \\
2x_2 + u_1 &= 40
\end{aligned}
\qquad\qquad
\begin{aligned}
x_1 &= \\
x_2 &= 2 \\
y_3 &= \\
y_4 &= 6 \\
y_5 &= 2 \\
u_1 &= 36 \\
u_2 &= -12
\end{aligned}
$$

$\Longrightarrow u_2 = \quad 12 < 0 \quad \Longrightarrow j = 2$

$$
\begin{array}{rcrcrcl}
v_1 & + & v_2 & & & = & 0 \\
v_1 & & & & & = & 0 \\
v_1 & + & 3v_2 & + & v_3 & = & \\
-v_1 & + & & & v_4 & = & 0 \\
& & -v_2 & + & v_5 & = & 0 \\
4v_1 & & & + & v_6 & = & -1 \\
& & 2v_2 & - & v_7 & = & -1
\end{array}
\qquad
\begin{array}{l}
\mathbf{v} = (0,0,0,0,0,-1,1) \\
\lambda_4 = +\infty \\
\dfrac{|u_5|}{|v_7|} = \dfrac{36}{1} = 36
\end{array}
$$

$\Longrightarrow \lambda_4 = 36$

$\Longrightarrow B_1 := B_1 = \{3,4,5\}$

$\Longrightarrow B_2 := B_2 \cup \{\ell\} \setminus \{j\} = \{2,5\} \cup \{1\} \setminus \{5\} = \{1,2\}$

Quinta iteración:

$\mathbf{T}_{B_1B_2}$ es complementaria

$$
\mathbf{T}_{B_1B_2} \quad
\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 3 & 1 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 2 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
$$

Tenemos que resover

$$\mathbf{T}_{B_1B_2} z = \mathbf{b}$$

solución

$$
\begin{array}{rcrcrcr}
x_1 & + & x_2 & & & = & 8 \\
x_1 & & & & & = & 6 \\
x_1 & + & 3x_2 & + & y_3 & = & 18 \\
-x_1 & + & & & y_4 & = & 0 \\
& & -x_2 & + & y_5 & = & 0 \\
4x_1 & + & u_1 & + & u_2 & = & 48 \\
& & 2x_2 & + & u_1 & = & 40
\end{array}
\qquad
\begin{array}{l}
x_1 = \\
x_2 = \\
y_3 = \\
y_4 = 6 \\
y_5 = \\
u_1 = 36 \\
u_2 = -12
\end{array}
$$

$\Longrightarrow u_2 = \;\; 12 < 0 \Longrightarrow j = 2$

Lo representamos como sigue

x_2	y_2	y_3	y_4	y_5	u_1	x_1	
1	0	0	0	0	0	1	8
0	1	0	0	0	0	1	6
3	0	1	0	0	0	1	18
0	0	0	1	0	0	-1	0
-1	0	0	0	1	0	0	0
0	0	0	0	0	1	4	48
2	0	0	0	0	1	0	40

1	0	0	0	0	0	1	8
0	1	0	0	0	0	1	6
0	0	1	0	0	0	-2	-6
0	0	0	1	0	0	-1	0
0	0	0	0	1	0	1	8
0	0	0	0	0	1	4	48
0	0	0	0	0	1	-2	24
						-6	-24

$x_1 = 4$

$x_2 = 4$

$x_3 = 2x_1 - 6 =$

$y_4 = x_1 = 4$

$y_5 = - x_1 + 8 = 4$

$y_2 = - x_1 + 6 = 2$

$u_1 = - 4x_1 + 48 - 32 \geq 0$

La solución actual es óptima

$\mathbf{x}^* = (4, 4)^T$

CAPITULO VI

APLICACIONES

6.1.- UN PROBLEMA DE PORTAFOLIO

Un inversionista desea formar un portafolio de valores, haciendo una selección en adecuadas proporciones entre n diferentes valores (acciones, obligaciones, bonos etc.

Asumimos que la tasa de retorno del i-ésimo tipo de valor tiene una distribución probabilística con esperanza μ_i y varianza σ^2_{ii}.

La varianza es una medida para el riesgo que la tasa de retorno realizada se desvíe de la tasa esperada de retorno, esto es: Un valor i con una alta varianza σ^2_{ii}, de la tasa de retorno, es riesgoso. Para llegar a una razonable estimación de la tasa de retorno, consideramos lo siguiente:

Asumimos que se dispone de datos históricos sobre el precio y el dividendo de los diferentes valores para m años.

Definamos la tasa de retorno $r_j(k)$ del j-ésimo valor en el año k por :

$$r_j(k) = \frac{p_j(k) - p_j(k-1) + d_j(k)}{p_j(k-1)}$$

donde $p_j(k)$ es el precio promedio en el año k y $d_j(k)$ denota los dividendos recibidos en el año k.

Indicamos dos estimaciones razonables para la esperanza μ_j de la tasa de retorno del j-ésimo valor.

a) Promedio aritmético:

$$\mu_j = \frac{1}{m} \sum_{k=1} r_j(k)$$

b) Promedio geométrico:

$$\mu_j = \left[\prod_{k=1} (1 + r_j(k)) \right]^{m} - 1$$

Eligiendo una de estas dos estimaciones μ_j podemos medir la varianza σ_j^2 como sigue:

$$\sigma_j^2 = \frac{1}{m} \sum_{k=1}^{m} \left[r_j(k) - \mu_j \right]^2$$

Para cada (i,j) la covarianza σ_{ij} es estimada por

$$\sigma_{ij} = \frac{1}{m} \sum_{k=1}^{m} \left[r_i(k) - \mu_i \right] \left[r_j(k) - \mu_j \right]$$

σ_{ij} mide la correlación entre las tasa de retorno de los valores i y j. Una covarianza positiva σ_{ij} indica que las tasas de retorno de los valores i y j tienden a subir o bajar conjuntamente.

Una buena regla para un inversionista racional es :
No concentrar las inversiones en un conjunto de valores que tienen la tendencia de moverse todos en la misma dirección.

Es aconsejable diversificar el portafolio con objetivo de reducir el riesgo, aún sacrificando parcialmente el retorno.

La medición de las varianzas σ_{ij} nos permite acercarnos a la diversificación requerida.

Es razonabe asumir que un inversionista observe las siguientes reglas, que son en cierco modo equivalentes.

i) Entre diferentes valores con la misma tasa de retorno el inversionista prefiere el valor con la mínima varianza.

ii) Entre diferentes valores con la misma varianza el inversionista preferirá el valor con la máxima tasa de retorno.

Sea K el capital a invertir en el portafolio, $\mathbf{x}_j$ la proporción de K a invertir en el valor j.

$S = (\sigma_{ij})$ es la matriz de covarianza, entonces $\mathbf{x}^T S \mathbf{x}$ es una adecuada medida para la decisión $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

Es razonable, buscar un vector de decisión $\mathbf{x}$, que minimize "el riesgo" $\mathbf{x}^T S \mathbf{x}$, bajo la condición que el promedio de las esperanzas de las tasas de retorno no sea menor que una cota μ. Esto nos conduce al siguiente programa cuadrático.

$$\min \mathbf{x}^T S \mathbf{x}$$

$$\sum_{j=1}^{n} \mu_j \mathbf{x}_j \geq \underline{\mu}$$

$$_{j=1} \mathbf{x}_j$$

$$\mathbf{x}_j \geq \qquad j = 1, 2, \ldots, n.$$

6.2. UN ALGORITMO PARA RESOLVER EL PROBLEMA DEL REGULDOR LINEAL CON DOMINIOS DE CONTROL Y DE ESTADO

El problema :

$$\min\ y(T)^T S y(T) + \int_0 \left[y(t)^T Q(t) y(t) + \mu(t)^T R(t) \mu(t) \right] dt$$

$$\dot{x}(t) = A(t)x(t) + B(t)\mu(t)$$

$$y(t) = H(t)x(t)$$

$$x(0) = x^0$$

$$\mu(t) \in U \quad x(t) \in X \ , \ \forall\ t \in [0,T]$$

$$x(t) \in M \ , \ y(t) \in N$$

Hipótesis

$x(t) \in \mathbb{R}^n \quad y(t) \in \mathbb{R}^m \ , \ \mu(t) \in \mathbb{R}^\ell$

[$y(t)$ se llama *vector de medida*, $x(t)$ *vector de estado*, $\mu(t)$ *vector de control*]

$S = S^T \geq 0 \quad Q(t) = Q(t)^T \geq 0 \quad R(t) = R(t)^T > 0$

$Q(t)$, $R(t)$ son seccionalmente continuas en $[0,T]$

- $A(t)$, $B(t)$, $H(t)$ son continuas en $[0,T]$

<u>OBSERVACIONES:</u>

Consideramos aquí sólo los casos frecuentes, donde

$U = \prod_{i=1} [\alpha_i, \beta_i]$ ó donde U es un poliedro.

Consideramos para el conjunto de estados solamente los casos:

$$X = \prod_{i=1}^{n} [a_i, b_i] \quad ó \quad X = \mathbb{R}^n$$

El vector de control $\mu(t)$ debe ser seccionalmente continua.

6.2.1: UNA DISCRETIZACION DEL PROBLEMA

Elegimos una partición

$$0 = t_o < t_1 < \ldots < t_N = T \text{ del intervalo } [0,T]$$

Admitimos sólo controles $\mu(t)$ de la forma:

$$\mu(t) = \mu_i \in U \ , \quad \forall\ t \in [t_{i-1}, t_i] \ , \quad i = 1,2,\ldots,N \qquad (x)$$

Consideramos las siguientes funciones básicas

$$v_j(t) = \begin{cases} 1, & t_{j-1} \leq t \leq t_j \\ 0, & \text{en otro caso} \end{cases}$$

Se cumple

$$v_i(t)v_j(t) = \begin{cases} 0 & \text{si } i \neq j \\ v_j(t) & \text{si } i \quad j \end{cases} \qquad \forall\ t \in [0,T]$$

$$\int v_i(t)v_j(t)dt \quad \begin{cases} 0 & \text{si } i \neq \\ t_j \quad t_{j-1} & \text{si } i = j \end{cases}$$

Por consiguiente las funciones $v_j(t)$ son ortogonales sobre $[0,T]$.

La función de control (x) se deja representar entonces como:

$$\mu(t) = \sum_{j=1} v_j(t)\mu_j \qquad (xx)$$

Si $V(t) = (v_1(t)I_{\ell}, \ v_2(t)I_{\ell}, \ldots, v_N(t)I_{\ell})$

$$U^T = (\mu_1^T, \ \mu_2^T, \ldots, \mu_N^T),$$

entonces (xx) se deja escribir como:

$$\mu(t) = V(t)\mu$$

Sustituyendo (*) en la ecuación diferencial se obtiene:

$$\dot{x}(t) = A(t)x(t) + B(t)V(t)\mu$$

Si $\Phi(t,0)$ es la solución de la ecuación matricial

$$\dot{x} = Ax \qquad x(0) = I$$

entonces la solución de (**) se deja representar como:

$$x(t) = \Phi(t,0)x(0) + \left[\int_0^t \Phi(t,s)B(s)V(s)ds\right]\mu \qquad (***)$$

Escribimos (***) como:

$$x(t) = X(t)x(0) + X(t)\mu \qquad (+)$$

Las matrices $X(t)$, $X(t)$ son continuas en $[0,T]$

De (+) se obtiene:

$$y(t) = H(t)X(t)x(0) + H(t)X(t)\mu$$

$$= Y(t)x(0) + Y(t)\mu \qquad (++)$$

Sustituimos (+), (++) y (*) en la función objetivo :

$$Z(\mu(t)) = Z(\mu, x(0)) = (\tilde{Y}(T)x(0) + Y(T)\mu)^T S(\tilde{Y}(T)x(0) + Y(T)\mu)$$

$$+ \int^T (\tilde{Y}(t)x(0) + Y(t)\mu)^T Q(t)(\tilde{Y}(t)x(0) + Y(t)\mu)dt$$

$$+ \int_0^T \mu^T V(t)^T R(t)V(t)\mu\, dt$$

Obtenemos:

$$Z(\mu, x(0)) = x(0)^T Z(0)x(0) + 2\mu^T\beta x(0)$$

$$+ \mu^T[Z(s) + Z(Q) + Z(R)]\mu,$$

donde:

$$Z(0) = \tilde{Y}(T)^T S\tilde{Y}(T) + \int_0^T \tilde{Y}(t)^T Q(t)\tilde{Y}(t)\, dt$$

$$\tilde{Y}(T)^T S\tilde{Y}(T) + \int_0^T \tilde{Y}(t)^T Q(t)\tilde{Y}(t)\ dt$$

$$Z(s) = \tilde{Y}(T)^T S\tilde{Y}(T)$$

$$Z(Q) = \int_0^T Y(t)^T Q(t)Y(t)\ dt$$

$$Z(R) = \int_0^T V(t)^T R(t)V(t)\ dt$$

OBSERVACIONES:

Las matrices $Z(0)$, β, $Z(s)$, $Z(Q)$, $Z(R)$ se dejan calcular amalíticamente o aproximar numéricamente.

- La matrices $Z(0)$, $Z(s)$, $Z(Q)$ y $Z(R)$ son simétricas.

6.2.2: LA SOLUCION DEL PROBLEMA APROXIMADO

Consideremos algunos casos especiales del problema

(1) $X = \mathbb{R}^n$, $U = \mathbb{R}^{\ell}$, $M = \mathbb{R}^n$, $N = \mathbb{R}^m$

En este caso el problema aproximado se reduce al programa

$$\min \{ \phi(u) \ / \ u \in \mathbb{R}^{Nm} \}$$

Recordemos $u^T = (u_1^T, u_2^T, \ldots, u_N^T) \in \mathbb{R}^{Nm}$

como $\phi(u)$ es convexa (estrictamente convexa), se obtiene las soluciones de (*) resolviendo el sistema :

$\nabla\phi(u) = 0$ ó

$2\beta x(0) + 2(Z(0) + Z(Q) + Z(R))u = 0$

Si $D = Z(s) + Z(Q) + Z(R) > 0$, entonces este sistema tiene una única solución

$$u^* = - D^{-1}\beta x(0)$$

(2) $X = \mathbb{R}^n$ $\quad$ U es un poliedro

$M = \mathbb{R}^n$ $\quad$ $N = \mathbb{R}^m$

Si $P = \prod_{k=1} u_k$, $u_k = u$, $\forall$ k

entonces el problema aproximado es el siguiente programa programa cuadrático convexo:

$$\min \{ \phi(u) \ / \ u \in P \} \tag{+}$$

Los diferentes métodos que hemos analizado en el presente trabajo permiten resolver este problema.

(3) $X = \mathbb{R}^n$, $\quad$ U es un poliedro

Las condiciones finales tienen la forma

$$x_k(T) \leq a_k \qquad y_j(T) \leq b_j$$

En este caso se adicionan al programa cuadrático (+) las restricciones lineales

$$x_k(T) = (\tilde{X}(T)x(0) + X(T)u)_k \leq 0$$

$$y_j(T) = (Y(T)x(0) + Y(T)u)_j \leq 0$$

para determinados índices k, j.

6.2.3 : EJEMPLO

Considere el siguiente problema del Regulador Lineal

$$\min \int_0^4 [x_1^2(t) + x_2^2(t) + x^2(t)]dt$$

$$\dot{x}_1 = x_2 \qquad x_1(0) = 1$$

$$\dot{x}_2 = u(t) \qquad x_2(0) = -1$$

Aplicar el método directo que conduce a una función cuadrática para N = 5

SOLUCION:

Tenemos : $[0,T] = [0,4]$

Partición $t_0 = 0, \quad t_1 = 0.8, \quad t_2 = 1.6$

$t_3 = 2.4, \quad t_4 = 3.2, \quad t_5 = 4 = t_N$

$V(t) \in M(1,5)$

$$A(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad Q(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$R(t) = [1], \quad S = 0, \quad H(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Sea $\Phi(t,s)$ la matriz fundamental caracterizada por

$$\Phi(t,s) = A(t)\Phi(t,s) \ , \ \Phi(s,s) = I$$

Calculamos las matrices

$\tilde{X}(t)$, $X(t)$, $\tilde{Y}(t)$, $Y(t)$, $Z(0)$, β, D (numéricamente)

Tenemos :

$$\Phi(t,s) = \begin{bmatrix} 1 & t-s \\ 0 & 1 \end{bmatrix}, \qquad \tilde{X}(t) = \Phi(t,0) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} = \tilde{Y}(t)$$

$$X(t) = \int_0^t \Phi(t,s)B(s)V(s)ds = \Phi(t,0)\int_0^t \Phi^{-1}(s,0)B(s)V(s)ds$$

$$X(t) = Y(t) = \begin{bmatrix} 1 & t \\ 0 & t \end{bmatrix} \int_0^t \begin{bmatrix} -s \\ 1 \end{bmatrix} V(s)ds$$

$$V(s) = (v_1(s),\ v_2(s),\ v_3(s),\ v_4(s),\ v_5(s))$$

Para $0 \le t < 0.8$: $V(s) = (1,0,0,0,0)$

$$X(t) = Y(t) = \left[\begin{matrix} t^2/2 & 0 & 0 & 0 \\ t & 0 & 0 & 0 \end{matrix} \right.$$

Para $0.8 \le t < 1.6$: $V(s) = (0,1,0,0,0)$

$$X(t) = Y(t) = \begin{bmatrix} -0.32+0.8t & t^2/2-0.8t+0.32 & 0 & 0 & 0 \\ 0.8 & t-0.8 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCION:

Tenemos : $[0,T] = [0,4]$

Partición $\quad t_0 = 0, \quad t_1 = 0.8, \quad t_2 = 1.6$

$\quad t_3 = 2.4, \quad t_4 = 3.2, \quad t_5 = 4 = t_N$

$V(t) \in M(1,5)$

$$A(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad Q(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$R(t) = [1], \quad S = 0, \quad H(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Sea $\Phi(t,s)$ la matriz fundamental caracterizada por

$$\Phi(t,s) = A(t)\Phi(t,s) \ , \ \Phi(s,s) = I$$

Calculamos las matrices

$\tilde{X}(t)$, $X(t)$, $\tilde{Y}(t)$, $Y(t)$, $Z(0)$, β, D (numéricamente)

Tenemos :

$$\Phi(t,s) = \begin{bmatrix} 1 & t-s \\ 0 & 1 \end{bmatrix}, \qquad \tilde{X}(t) = \Phi(t,0) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} = \tilde{Y}(t)$$

$$X(t) = \int_0^t \Phi(t,s)B(s)V(s)ds = \Phi(t,0)\int_0^t \Phi^{-1}(s,0)B(s)V(s)ds$$

$$X(t) = Y(t) = \begin{bmatrix} 1 & t \\ 0 & t \end{bmatrix} \int_0^t \begin{bmatrix} -s \\ 1 \end{bmatrix} V(s)ds$$

$$V(s) = (v_1(s),\ v_2(s),\ v_3(s),\ v_4(s),\ v_5(s))$$

Para $0 \leq t < 0.8$: $V(s) = (1,0,0,0,0)$

$$X(t) = Y(t) = \left[\begin{matrix} t^2/2 & 0 & 0 & 0 \\ t & 0 & 0 & 0 \end{matrix} \right.$$

Para $0.8 \leq t < 1.6$: $\quad V(s) = (0,1,0,0,0)$

$$X(t) = Y(t) = \begin{bmatrix} -0.32+0.8t & t^2/2-0.8t+0.32 & 0 & 0 & 0 \\ 0.8 & t-0.8 & 0 & 0 & 0 \end{bmatrix}$$

$$\beta = \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 = \begin{bmatrix} 5.2053 & 17.7235 \\ 3.1573 & 10.9225 \\ 1.6213 & 6.9745 \\ 0.5973 & 3.0933 \\ 0.0853 & 0.6442 \end{bmatrix}$$

$$\beta x^0 = \beta \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \beta_5 = \begin{bmatrix} 12.5182 \\ 7.7652 \\ 5.3532 \\ 2.496 \\ 0.5589 \end{bmatrix}$$

<u>CALCULO DE D:</u>

$$D = \int_0^4 [X(t)^T X(t) + V(t)^T V(t)]dt$$

$$D_1 = \int_0^{0.8} [X(t)^T X(t) + V(t)^T V(t)]dt = \begin{bmatrix} 0.9870 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_2 = \int_{0.8}^{1.6} [X(t)^T X(t) + V(t)^T V(t)]dt = \begin{bmatrix} 0.86698 & 0.32427 & 0 & 0 & 0 \\ 0.32427 & 0.7959 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_3 = \begin{bmatrix} 1.85002 & 1.19466 & 4.11648 & 0 & 0 \\ 1.19466 & -1.59062 & 0.32427 & 0 & 0 \\ 4.11648 & 0.32427 & 0.98704 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_4 = \begin{bmatrix} 3.48842 & 2.50538 & 1.52234 & 0.4335 & 0 \\ 2.50538 & 1.85002 & 1.19466 & 0.37888 & 0 \\ 1.52234 & 1.19466 & 0.86698 & 0.32427 & 0 \\ 0.4335 & 0.37888 & 0.32427 & 0.98704 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_5 = \begin{bmatrix} 5.78218 & 4.47146 & 3.16074 & 1.85002 & 0.48811 \\ 4.47146 & 3.48842 & 2.50538 & 1.52234 & 0.4335 \\ 3.16074 & 2.50538 & 1.85002 & 1.19466 & 0.3788 \\ 1.85002 & 1.52234 & 1.19466 & 0.86698 & 0.32427 \\ 0.48811 & 0.4335 & 0.3788 & 0.32427 & 0.98704 \end{bmatrix}$$

$$D = D_1 + D_2 + D_3 + D_4 + D_5 =$$

$$= \begin{bmatrix} 12.9746 & 8.49577 & 8.79956 & 2.28352 & 0.48811 \\ 8.49577 & 4.54372 & 4.02431 & 1.90122 & 0.4335 \\ 8.79956 & 4.02431 & 3.70404 & 1.51893 & 0.3788 \\ 2.28352 & 1.90122 & 1.51893 & 1.85402 & 0.32427 \\ 0.48811 & 0.4335 & 0.3788 & 0.32427 & 0.98704 \end{bmatrix}$$

Luego tenemos:

$$\phi(u) = Z(x(0), u) = x(0)^T Z(0) x(0) + 2u^T \beta x(0) + u^T D u$$

Resolvemos:

$$\min \{\phi(u) \ / \ u \in \mathbf{P}\}$$

A esta programa cuadrático adicionamos las siguientes restricciones lineales:

$$X_k(T) = (X(T)x(0) + X(T)u)_k \leq 0$$

$$y_j(T) = (Y(T)x(0) + Y(T)u)_j \leq$$

para determinados índices k, j.

BIBLIOGRAFIA

1. ACTAS DE CONGRESOS:

[1] ABADIE (J.), Ed — Nonlinear Programming.
NorthHolland, Amsterdam, 1,970

[2] FLETCHER (R.), Ed — Optimization.
Academic Press, London, 1,969

II LIBROS:

[3] AVRIEL — Nonlinear Programming Analysis and Methods.
Englewood Cliffs, New Jersey, 1,976

[4] BLUM (E.), OETTLI (W.)
Nichtlineare Optimierung
Springer, Berlin, 1,975

[5] FLEMING (W.H.) — Deterministic and Stochastic Optimal Control
Springer, Berlin, 1,975

[6] KUENZI (H.P.), RISHEL (R.W.)
Nonlinear Programming
Blaisdell, Waltham, Mass., 1,966

[7] SPOSITO (V.A.) — Linear and Nonlinear Programming
Iowa State Univ., Press, 1,975

[8] ZANGWILL (W.I.) — Nonlinear Programming
Prentice Hall, Englewood Cliffs, N.J., 1969

[9] LEMKE (C.E.) — A Method of Solution for cuadratic Programs
Management Sci. 11 (1,965), 681-689

III ARTICULOS:

[10] BLUM (E.), OETTLI (W.)
Dreat proof of the existence theorem for quadratic programming
Operations Res. 20 (1,972), 165-167

[11] BEALE (E.M.L.) On quadratic programming
Naval Res. Logist Quart. <u>6</u> (1,959) 227-244

[12] Nonlinear Optimization by simplex-like methods [Fletcher 1,969], pgs, 24-36

[13] DANTZIG (G.B.), COTTLE (R.W.)
Positive (semi-) definide programming [Abadie, 1,967], pgs. 55-73

[14] DANTZIG (G.B.), EISENBERG (E.), COTTLE (R.W.)
Symmetric dual nonlinear programs
Pacific J. Math. <u>15</u> (1,965), 809-812

[15] DORN (W.S.) Duality in quadratic programming
Quart. Appl. Math. <u>18</u> (1,960), pgs. 155-162

[16] FRANK (M.), WOLFE (P.)
An algorithm for quadratic programming.
Naval Res. Log. Quart. <u>3</u> (1,956) pgs 95-110

[17] WOLFE (P.) The simplex method for quadratic programming.
Econometrica <u>27</u> (1,959), 382-398

[18] ZANGWILL (W.I.) The convex simplex method.
Management Sci. <u>14</u> (1,967), 221-238.