Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

Evolución y aplicación del método escalado afín, para el caso acotado y no degenerado

TESIS

Para optar el Grado Académico de Magíster en Investigación de Operaciones y Sistemas

AUTOR

Carlos Rubén GUERRERO MONCADA

ASESOR

Edinson Raúl MONTORO ALEGRE

Lima, Perú

2016

Referencia bibliográfica

Guerrero, C. (2016). *Evolución y aplicación del método escalado afín, para el caso acotado y no degenerado.* [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER

Siendo las, 16:15 horas del día jueves 30 de junio del dos mil dieciséis, en la Sala de Profesores de la Facultad de Ciencias Matemáticas, el Jurado Evaluador de Tesis, Presidido por la Dra. María del Pilar Alvarez Rivas e integrado por los siguientes miembros: Dra. María Natividad Zegarra Garay (Jurado Evaluador), Mg. Ricardo López Guevara (Jurado Informante), Mg. Sonia Esther Castro Ynfantes (Jurado Evaluador) y el Mg. Edinson Raúl Montoro Alegre como Jurado Asesor, se reunieron para la sustentación de la tesis titulada: "EVOLUCION Y APLICACIÓN DEL MÉTODO ESCALADO AFÍN, PARA EL CASO ACOTADO Y NO DEGENERADO" presentada por el Bachiller CARLOS RUBÉN GUERRERO MONCADA, para optar el Grado Académico de Magíster en Investigación de Operaciones y Sistemas.

Luego de la exposición del graduando, los Miembros del Jurado hicieron las preguntas correspondientes, así como las observaciones e inquietudes acerca del trabajo de tesis, a las cuales el Bachiller Carlos Rubén Guerrero Moncada respondió con acierto y solvencia, demostrando pleno conocimiento del tema.

A continuación se realizó la calificación correspondiente, según tabla adjunta, resultando el Bachiller Carlos Rubén Guerrero Moncada aprobado con el calificativo de ...18.....
...MUY...BUENO...................

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del **Grado Académico de Magíster en Investigación de Operaciones y Sistemas** al Bachiller Carlos Rubén Guerrero Moncada.

Siendo las 17:20 horas, se levantó la sesión, firmando para constancia la presente Acta.

Dra. María Natividad Zegarra Garay
Miembro

Dra. María del Pilar Álvarez Rivas
Presidenta

Mg. Sonia Esther Castro Ynfantes
Miembro

Mg. Ricardo López Guevara
Miembro

Mg. Edinson Raúl Montoro Alegre
Miembro Asesor

FICHA CATALOGRÁFICA

GUERRERO MONCADA CARLOS RUBÉN

Evolución y aplicación del método Escalado Afín, para el caso acotado y No degenerado

UNMSM Magister en Investigación en Investigación Operativa 2015

Tesis: Universidad Nacional Mayor de San marcos

Facultad de Ciencias Matemáticas

Investigación Operativa

I. UNMSM /F de CM

DEDICATORIA

Mi gratitud a mi hermana Juana Cabrera y mi esposa Cecilia Argote. La primera, por haberme enseñado desde niño con su ejemplo; que nada en la vida, se consigue sin esfuerzo y la segunda. Compañera inolvidable, buena amiga, y que además me dio dos hijos maravillosos; que han dado contenido a mi vida

AGRADECIMIENTOS

Mi gratitud al Magister Edson Montoro Alegre por su apoyo incondicional y permanente, durante todo el desarrollo del presente trabajo y también a todos los profesores que de diferentes formas, me apoyaron; fueron tantos que mencionar alguno de ellos, correría el riesgo de dejar fuera a otros

RESUMEN

Se presenta una variante del método de Puntos interiores para resolver un programa matemático lineal, el Método de Escalado Afín, relevando su aspecto geométrico y presentando aplicaciones

Palabras claves
Programación Lineal, esfera unitaria, elipsoide sólido, solución óptima

ABSTRACT

A variant of the method of interior points is presented to solve a linear mathematical program that is related scaling method, relieving its geometric aspect and submitting an application

Keywords
Linear programming, unit sphere , ellipsoid solid , optimal solution

CONTENIDOS

LISTA DE GRÁFICAS

CAPÍTULO 1 INTRODUCCIÓN

1.1 Situación problemática

Desde que apareciera el método símplex, que permitió resolver un Programa Matemático Lineal (PML); muchos matemáticos han contribuido a su crecimiento, ya sea al desarrollar la teoría matemática o diseñando programas computacionales, cada vez más eficientes o experimentando con nuevos algoritmos alternativos; como los métodos de "Puntos Interiores"

Con el método símplex se obtiene una solución óptima, a través de los puntos extremos del dominio admisible del programa. En la práctica este método sigue funcionando bien; incluso para programas de gran tamaño. Sin embargo la cantidad de iteraciones necesarias para llegar a la solución óptima (complejidad computacional), puede crecer teóricamente en forma exponencial.

Según (Castro, 2002)

> "El primer Método de Puntos Interiores, de gran trascendencia teórica es el debido a L. G. Khachiyan en 1979. Este autor, recopila una serie de trabajos e ideas de los autores rusos Shor, Yudin y Nemirovskii sobre un método, que resuelve un programa matemático no lineal, generando una sucesión de elipsoides, él hizo extensivo el método para un programa matemático lineal, resultando un método, que se le reconoce como "método de los elipsoides", cuya principal ventaja teórica sobre el métodos simplex radica en que resuelve programa; con menor complejidad computacional (complejidad polinómica); Sin embargo las implementaciones prácticas, han demostrado su desventaja respecto al método simplex ya que, por un lado, el número de iteraciones que requiere para llegar al óptimo es muy grande y, por otro, el trabajo que implica cada una de ellas es mucho más costoso que el requerido por el método simplex, pues no se pueden aplicar las técnicas de matrices dispersas tan características de las implementaciones en ordenador de este último".

Actualmente existen diversos textos que describen los algoritmos de puntos interiores. Unos dan mayor énfasis al rigor matemático – demostrativo; como

Tsuchiya(1995) Otros bastante introductorio, pero descuidando el aspecto intuitivo – geométrico; como Vanderbei (1996)

N. Karmarkar (1984) desarrolló un algoritmo de puntos interiores, que busca la solución óptima; a través de los puntos del dominio admisible, que no están en su frontera, construyendo simplejos en vez de elipsoides con una complejidad computacional polinomial.

Otra variante del método de puntos interiores es el escalado afín, el cual utiliza una estrategia de construir híper esferas de centro unitario y radio uno. Este algoritmo es eficaz para programas lineales extremadamente grandes.

Estas investigaciones, fueron realizadas por matemáticos de los países desarrollados; dígase: Estados Unidos de América o Europa. En el contexto Latinoamericano, sólo se investigan estos temas en países como Brasil o Argentina, se cita el trabajo de Mónica L. Ingratta, en su tesis "Implementacion del Metodo Primal-Dual para Programación Lineal" (Ingratta, 1996):

"En este trabajo se describe la implementación de un algoritmo para resolver el problema de programación lineal:

(P) min c x: s. a: $\{\mathbf{x} \in \mathbb{R}^{\mathbf{n}}\mathbf{: Ax = b, x \geq 0}\}$, donde A es matriz de m filas y n columnas y n > m

En el Perú se cita el trabajo sobre Puntos Interiores, desarrollado por (Inés Gambini 1999),

En el Perú; hay pocas investigaciones sobre este método; sí se hacen aplicaciones de la P. M. L., en el ámbito empresarial, pero no se utiliza algoritmos alternativos de puntos interiores, muchas veces por desconocimiento por parte de los profesionales que se dedican a aplicar la PML.

En este tema, las Universidades del Perú se enfocan principalmente a la enseñanza del método simplex, con literatura en español sobre puntos interiores se podría ampliar el capital humano que maneje el tema de puntos interiores, el cual está probado, que tiene una complejidad computacional menor.

En el capítulo 2 se presenta el marco teórico que incluye el problema de programación lineal y los métodos de puntos interiores.

En el capítulo 3 se presenta el método de Escalado Afín, la manera como elige el punto inicial (el que debe ser interior al dominio admisible del programa), también la forma de seleccionar la dirección del nuevo de mejora del programa. Se presentan algunos ejemplos de la forma de búsqueda.

1.2 Formulación del problema

No existe un análisis y difusión adecuada del método de Escalado Afin; variante del método de "Puntos Interiores", para la solución de un programa matemático lineal.

1.3 Justificación de la Investigación

En los textos de formación universitaria más conocidos como el de Hillier, Taha y Winston hay de dos a tres páginas dedicadas al método de puntos interiores considerando solo el de Karmakar, más no el MAE.

Según (mokhtar & bazaraa, 2010)

El texto de Bazaraa, si bien es cierto, contiene más información, tampoco presenta el MAE.

El MAE presenta una visión intuitiva – geométrica, la cual facilita el acercamiento a la temática por parte de docentes y estudiantes (Castro, 2002)

Otra ventaja es que el MAE tiene una menor complejidad computacional que el simplex y que el de Karmakar lo cual facilitará la implementación.

1.4 Objetivos de la Investigación

1.4.1 Objetivo General

- Difundir una de las alternativas, del método de Punto Interior (**MAE**), para la solución de un **P. M. L.**

1.4.2 Objetivos Específicos

- Describir el método de escalado afín
- Comparar la eficiencia del método de escalado afín con la del método simplex

CAPÍTULO 2 MARCO TEÓRICO

2.1 Antecedentes del Problema

El método del escalado afín, fue el primero de los algoritmos denominados de punto interior para resolver un programa matemático lineal, fue presentado por el matemático ruso Dikin en 1967; pero no fue conocido en occidente; sino en la década de los 80; debido a que su fortaleza mayor, era resolver un programa matemático lineal en tiempo polinomial; llamó la atención de muchos estudiosos.

Habiendo aportado en el estudio de su convergencia, entre otros: Vanderbei(1986), Tsuchiya(1995) y (Monteiro)

Según Pool David (2011, p. 612), define:

"Valores Singulares de una Matriz: Para cualquier matriz A de m x n, la matriz A^TA de n x n es simétrica y en consecuencia puede ser diagonalizable ortogonalmente, por el teorema espectral. No solo todos los eigenvalores de A^TA son reales, todos son no negativos. Para demostrar esto, sea A un eigenvalor de A^TA con su correspondiente eigenvalor unitario v. Entonces

$$0 \ll \|\mathrm{Av}\|^2 = (\mathrm{Av}).(\mathrm{Av}) = (\mathrm{Av})^{\mathrm{T}}\mathrm{Av} = \mathrm{v}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{Av}$$

$$\mathrm{v}^{\mathrm{T}}\mathrm{Av} = \lambda(\mathrm{v.v}) = \lambda\|\mathrm{v}\|^2 = \lambda$$

Por tanto tiene sentido sacar raíces cuadradas (positivas) de dichos eigenvalores".

2.2 Bases Teóricas

2.2.1 La Programación Lineal y el método de Símplex

Notación: A lo largo del presente trabajo se utilizará la siguiente notación:

Se denotará con letra minúscula el nombre del y con mayúscula el nombre de la matriz

$$x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} \Rightarrow X = \begin{bmatrix} x_{11} & x_{12} & & x_{1n} \\ x_{21} & & & \\ . & . & . & . \\ x_{nn} & & & x_{nn} \end{bmatrix}$$

Esta notación en general no es estándar en matemáticas; pero ha logrado aceptación en la comunidad, dedicada a los métodos de puntos interiores

Con x_i se denotará el i-ésimo componente del vector

Con $\mathbf{x}^*$ se denotará el vector óptimo

Con letras griegas se denotarán los escalares

Con $\mathbf{x}^k$ se denotará la solución que se tiene en la k-ésima iteración del algoritmo

Con la letra **e** se denotará al vector cuyas componentes son unos;

$\mathbf{e} = (1,1,1,..1)^t$

2.2.1.1 Forma Estándar de un Programa Matemático Lineal

Definición

Se dice que un programa matemático lineal, al que denotaremos con **p.m.l.** está en la forma Estándar, si presenta la forma:

$$\begin{aligned}
\mathbf{P:}\quad \textbf{Minimizar}\ & c_1 x_1 + c_2 x_2 + \ .\ + c_n x_n \\
\text{Sujeto a}\ & a_{11} x_1 + a_{12} x_2 + a_{13} x_3 + \ ..+\ a_{1n} x_n = b_1 \\
& a_{21} x_1 + a_{22} x_2 + a_{23} x_3 + \ ..+\ a_{2n} x_n = b_2 \\
& \qquad\qquad . \\
& \qquad\qquad . \qquad\qquad\qquad (1.1) \\
& \qquad\qquad . \\
& a_{m1} x_1 + a_{m2} x_2 + a_{m3} x_3 + \ ..+\ a_{mn} x_n = b_m \\
& x_1 \geq 0,\quad x_2 \geq 0,\quad x_3 \geq 0,\qquad x_n \geq 0
\end{aligned}$$

Aquí $c_1 x_1 + c_2 x_2 + \ ...\ .+ c_n x_n$ es la función objetivo (o función de criterio), que debe minimizarse.

Las constantes $c_1, c_2, ...\ .c_n$ que son conocidas, son los coeficientes de costo y $x_1, x_2, ..\ x_n$ son las variables de decisión (o niveles de actividad), que deben determinarse.

La ecuación $\mathbf{a}_{i1}\,\mathbf{x_1} + \mathbf{a}_{i2}\,\mathbf{x_2} + \mathbf{a}_{i3}\,\mathbf{x_3} + \,.\,.\, + \mathbf{a}_{in}\,\mathbf{x_n} = \mathbf{b}_i$, i = 1, 2, . . m, denota la i-ésima restricción.

Las constantes $\mathbf{a}_{ij}$, i = 1, 2, . . **m**, j = 1, 2, . . **n**, se les llama coeficientes tecnológicos, los que forman una matriz de restricciones **A**

$$\mathbf{A} = \begin{bmatrix} \mathbf{a_{11}} & \mathbf{a_{12}} & \mathbf{a_{13}} & . & \mathbf{a_{1n}} \\ \mathbf{a_{21}} & \mathbf{a_{22}} & \mathbf{a_{23}} & . & \\ \mathbf{a_{31}} & \mathbf{a_{32}} & \mathbf{a_{33}} & . & \\ . & . & . & . & \\ \mathbf{a_{m1}} & \mathbf{a_{m2}} & \mathbf{a_{m3}} & . & \mathbf{a_{mn}} \end{bmatrix}$$

El vector columna $\begin{bmatrix} \mathbf{b_1} \\ \mathbf{b_2} \\ \mathbf{b_3} \\ . \\ \mathbf{b_m} \end{bmatrix}$, cuya i-ésima componente es $\mathbf{b}_i$, se le denomina vector del lado derecho (o RHS), representa los requerimientos mínimos que deben satisfacerse.

Las condiciones $\mathbf{x_1 \geq 0,\ \ x_2 \geq 0,\ \ x_3 \geq 0,\ \ .\,.\,,x_n \geq 0,}$ son las restricciones de no negatividad**.**

El programa **P ,** puede reescribirse en notación matricial, como:

$$\begin{aligned} \mathbf{P: Minimizar}\ \ & \mathbf{f(x) = c^t x} \\ \mathbf{Sujeto\ a}\ \ & \mathbf{Ax = b} \\ & \mathbf{x \geq 0} \end{aligned}$$

Siendo $\mathbf{x = (x_1, x_2, x_3, .\,.\, x_n)^t}$ un vector columna y $\mathbf{c = (c_1, c_2, c_3, .\,.\, c_n)^t}$ un vector fila.

2.2.1.2 Dominio admisible de un P. M. L.

Definición Al conjunto que forman las restricciones del programa; esto es $\{\mathbf{x} \in \mathbb{R}^{\mathbf{n}}\mathbf{: Ax = b\,, x \geq 0}\}$ se le denomina: **"**Dominio admisible del programa" o "Región factible del programa"

De acuerdo a lo definido anteriormente; resolver el **P.M.L.** quiere decir encontrar un punto, entre todos los puntos del Dominio admisible del programa tal que haga mínimo el valor de **f(x)**

2.2.1.3 Rango de una matriz

Definición Sea la matriz $\mathbf{A}$ de $\mathbf{m}$ filas y $\mathbf{n}$ columnas con $\mathbf{m<n,}$ se llama rango de la matriz $\mathbf{A}$ y se le denotará con $\mathbf{r(A)}$ al número máximo de columnas linealmente independientes ($\mathbf{l.i.}$).

Ejemplo 2.2.1.1

Sean las matrices:

$$\mathbf{A}=\begin{bmatrix} 1 & 1 & 0 & 2 & 0 \\ 0 & 2 & 1 & 3 & 0 \\ 0 & 3 & 0 & 1 & 1 \end{bmatrix} \qquad \text{y } \mathbf{B}=\begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 2 & 1 \end{bmatrix}$$

Aquí se tiene que $\mathbf{r(A)=3=m}$, es decir posee tres columnas $\mathbf{l.i.}$ y $\mathbf{r(B)=2;}$ es decir posee sólo dos columnas $\mathbf{l.i.}$

2.2.1.4 Matriz de rango completo

Definición Sea la matriz $\mathbf{A}$ de $\mathbf{m}$ filas y $\mathbf{n}$ columnas con $\mathbf{m<n,}$ se dice que $\mathbf{A}$ es de **rango completo** (y se escribe $\mathbf{r(A)=m}$), si posee no más de $\mathbf{m}$ columnas $\mathbf{l.i.}$

Si se asume que la matriz $\mathbf{A}$ tiene rango completo**;** entonces se puede considerar la partición de la matriz $\mathbf{A}$ por columnas: $\mathbf{A}=[\mathbf{B},\mathbf{N}]$, donde la sub matriz $\mathbf{B}$, tiene sus columnas $\mathbf{l.i.}$

De esta manera la sub matriz $\mathbf{B}$, será una base de $\mathbb{R}^{\mathbf{m}}$ y por tanto invertible. Entonces la sub matriz **N** de **m x (n - m);** se podrá obtener como combinación de las columnas de **B**

$$A = [B, N] = \begin{array}{c} \left[\begin{array}{ccccc|ccc} a_{11} & a_{12} & a_{13} & . & a_{1m} & a_{1\,m+1} & . & a_{1n} \\ a_{21} & a_{22} & a_{23} & . & a_{2m} & a_{2\,m+1} & . & a_{2n} \\ a_{31} & a_{32} & a_{33} & . & a_{3m} & a_{3\,m+1} & . & a_{3n} \\ . & . & . & . & . & . & . & . \\ a_{m1} & a_{m2} & a_{m3} & . & a_{mm} & a_{m\,m+1} & . & a_{mn} \end{array}\right] \end{array}$$

$$\underbrace{\qquad\qquad\qquad}_{\mathbf{B}} \qquad \underbrace{\qquad\qquad}_{\mathbf{N}}$$

Ejemplo 2.2.1.2

Tómese las matrices del Ejemplo 2.2.2.1

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 & 0 \\ 0 & 2 & 1 & 3 & 0 \\ 0 & 3 & 0 & 1 & 1 \end{bmatrix} \qquad y \quad B = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 2 & 1 \end{bmatrix}$$

La matriz $\mathbf{A}$ es de rango completo, $\mathbf{r(A) = 3}$ y la matriz $\mathbf{B}$ no lo es, ya que el número de columnas **l. i.** que es 2, es menor que $\mathbf{m = 3}$.

2.2.1.5 Solución básica factible no degenerada (S. b. f. n. d.)

Considérese el **P. M. L.**

$$\begin{array}{ll} \mathbf{Min} & f(x) = c^t x \\ \mathbf{Sujeto\ a} & x \in S \end{array}$$

Siendo el conjunto poliédrico $\mathbf{S = \{x \in \mathbb{R}^n \ /. \ Ax = b, \ x \geq 0\}}$

Se dice que el vector $\mathbf{x}$ es una solución factible no degenerada de $\mathbf{P}$; si este vector posee $\mathbf{m}$ componentes estrictamente positivas y el resto de componentes nulas. Esto es

$\mathbf{x = (x_1, x_2, x_3, ..x_m, x_{m+1}, x_{m+2}, ,..x_n)^t = (x_1, x_2, x_3, ..x_m, 0, 0, ,..0)^t}$, con $\mathbf{x_j > 0}$, $\mathbf{j = 1, 2 .. \ m}$

Consideremos nuevamente el programa $\mathbf{P}$:

Asúmase además que la matriz $\mathbf{A}$ es de rango completo y sea el vector $\mathbf{x \in \mathbb{R}^n}$, una solución básica factible del programa

Luego por la **definición 1.4**

Se podrá escribir $\mathbf{x} \in \mathbb{R}^n$**,** como:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{bmatrix}, \text{ siendo } \mathbf{x_B} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_m \end{bmatrix} \geq 0 \text{ y } \mathbf{x_N} = \begin{bmatrix} x_{m+1} \\ x_2 \\ x_3 \\ . \\ x_n \end{bmatrix} \geq 0$$

Entonces $\mathbf{A}\,x = \mathbf{b}$, $x \geq \mathbf{0}$ puede escribirse como $[\mathbf{B},\mathbf{N}]\begin{bmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{bmatrix} = \mathbf{b}$ (**1**)

De (**1**), se tiene $\mathbf{B}\ \mathbf{x_B} + \mathbf{N}\ \mathbf{x_N} = \mathbf{b}$

Si $\mathbf{x_B}$ tiene todas sus componentes estrictamente positivas, esto es

$$\mathbf{x_B} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_m \end{bmatrix} > \mathbf{0}$$

Se dirá que $\mathbf{x_B}$ es una solución <u>básica factible no degenerada</u>

2.2.1.6 Dualidad

Para cada Programa Matemático Lineal **P**; existe otro programa Matemático lineal **Q** que le corresponde; al primero llamaremos P<u>rimal</u> y al otro **Q ;** simplemente el <u>Dual</u>.

La solución óptima (si es que la posee), de este nuevo programa lineal **Q** , posee algunas propiedades importantes, tales como:

i) Se puede utilizar para obtener la solución del programa primal, cuando este posee pocas variables y muchas restricciones

ii) Se puede utilizar, para hacer interpretaciones económicas, acerca de los parámetros, tales como **b** y **c** , etc.

Considérese la forma Estándar del programa lineal: "**P**", al que llamaremos "Primal":

Primal

$$\begin{array}{ll} \mathbf{P}: \textbf{Min} & \mathbf{f(x) = c^t x} \\ \textbf{Sujeto a} & \mathbf{A\,x = b} \\ & \mathbf{x \geq 0} \end{array}$$

A es una matriz de m*n, $\mathbf{c}$ es un vector de $\mathbb{R}^n$, $\mathbf{b}$ es un vector columna de $\mathbb{R}^m$ y $\mathbf{x}$ es el vector de variables perteneciente a $\mathbb{R}^n$

Entonces el programa lineal dual, al que llamaremos $\mathbf{Q}$, se define como:

$\mathbf{Q}$: **Max**	$\mathbf{g(w) = b^t\, w}$
Sujeto a	$\mathbf{w\, A^t \leq c^t}$

Siendo $\mathbf{w} \in \mathbb{R}^m$, el vector de variables duales.

2.2.1.7 Dualidad débil

Teorema

Si $\mathbf{x}$ y $\mathbf{w}$, son dos soluciones factibles de los programas $\mathbf{P}$ y $\mathbf{Q}$ respectivamente. Entonces el valor de la función objetivo del programa $\mathbf{P}$; nunca supera el valor de la función objetivo, del programa $\mathbf{Q}$. Es decir:

$$\boxed{\mathbf{wb \leq c^t x}}$$

Demostración

Sean $\bar{\mathbf{x}}$ e $\bar{\mathbf{w}}$ dos soluciones factibles de $\mathbf{P}$ y $\mathbf{Q}$ respectivamente.

Luego $\mathbf{A\,\bar{x} = b}$ y $\mathbf{\bar{w}\,A^t \leq c^t}$

También $\mathbf{\bar{w}^t\, A\,\bar{x} = \bar{w}^t\, b}$ (1)

Y por ser $\mathbf{\bar{x} \geq 0}$, en el dual se tiene $\mathbf{\bar{w}^t\, A\,\bar{x} \leq c^t\,\bar{x}}$ (2)

De (1) y (2), se concluye que = $\mathbf{\bar{w}^t\, b \leq c^t\,\bar{x}}$.

2.2.1.8 Dualidad fuerte

TEOREMA

Si el programa primal $\mathbf{P}$, tiene un punto óptimo (minimizador) $\mathbf{X}^*$; entonces el programa dual $\mathbf{Q}$, también tiene un punto óptimo (maximizador) $\mathbf{w}^*$ y además $\mathbf{c}^T\,\mathbf{X}^* = \mathbf{b}^T\,\mathbf{w}^*$.

Demostración

Supongamos que $\mathbf{B}^*$ es la matriz básica; asociada a la iteración Símplex óptima de $\mathbf{P}$.

Definamos $\quad \mathbf{w}^0 = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^* \quad \ldots \qquad (\mathbf{1})$

Luego $\quad \mathbf{w}^0\,\mathbf{A} = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^*\,\mathbf{A}$

$\mathbf{w}^0\,\mathbf{A} - \mathbf{c} = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^*\,\mathbf{A} - \mathbf{c}$:

Haciendo $\mathbf{Z} = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^*\,\mathbf{A}$; por ser $\mathbf{B}^*$ óptimo de $\mathbf{P}$, se cumple que

$\mathbf{w}^0\,\mathbf{A} - \mathbf{c} = \mathbf{Z} - \mathbf{c} \leq \mathbf{0}$

Por tanto $\quad \mathbf{w}^0$ es solución factible de $\mathbf{Q}$; ya que $\mathbf{w}^0\,\mathbf{A} \leq \mathbf{c}$

También $\quad \mathbf{w}^0\,\mathbf{b} = \mathbf{g}(\mathbf{w}^0) = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^*\,\mathbf{b} \quad$ por $\quad (\mathbf{1})$

Luego $\quad \mathbf{w}^0\mathbf{b} = \mathbf{g}(\mathbf{w}^0) = \mathbf{c}_B^T\,(\mathbf{B}^{-1})^*\,\mathbf{b} = \mathbf{f}(\mathbf{x}^*) = \mathbf{min\ f}(\mathbf{x})$

o también $\quad \mathbf{g}(\mathbf{w}^0) = \mathbf{min\ f}(\mathbf{x}) \qquad (\mathbf{2})$

Es decir: $\mathbf{w}^0$ (Llamado "Multiplicadores de Lagrange"), <u>es óptimo de **P**</u>

Faltará probar que $\mathbf{w}^0$; es también óptimo del dual

En efecto, se debe probar que: $\mathbf{g}(\mathbf{w}^0) = \mathbf{máx\ g}(\mathbf{w})$

Por el Teorema anterior $\mathbf{g}(\mathbf{w}) \leq \mathbf{f}(\mathbf{x})$.

De donde se tiene que $\mathbf{máx\ g}(\mathbf{w}) \leq \mathbf{min\ f}(\mathbf{x}) = \mathbf{g}(\mathbf{w}^0)$

Por tanto: $\mathbf{g}(\mathbf{w}^0) = \mathbf{máx\ g}(\mathbf{w})$

Concluyendo que $\mathbf{g}(\mathbf{w}^0) = \mathbf{f}(\mathbf{x}^*)$

2.2.1.9 Holgura complementaria débil

Teorema

Si el programa $\mathbf{P}$ tiene solución óptima y se halla en el punto (minimizador), $\mathbf{X}^*$; entonces el programa dual correspondiente $\mathbf{Q}$, también tiene una solución óptima, en el punto (maximizador) $\mathbf{w}^*$. Entonces si $\mathbf{x}^*$ es el vector dual asociado a $\mathbf{Q}$ y $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \geq \mathbf{0}$, es el vector de holguras duales, se verifica que:

$$\boxed{\mathbf{z}^* \mathbf{x}^* = \mathbf{0}}$$

Siendo del vector $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \geq \mathbf{0}$

Demostración

Escríbase el programa dual $\mathbf{Q}$ en la forma estándar, adicionando el vector de holguras duales $\mathbf{z}$

$$
\begin{array}{lll}
\mathbf{R}\text{: } \mathbf{Máx}\ \mathbf{Y_0} = & \mathbf{b^t\ y} & \\
\text{S. a.} & \mathbf{A^t y + z} & = \mathbf{c}^t \\
 & \mathbf{z} \geq \mathbf{0} &
\end{array}
$$

Si $\mathbf{x}$ es el vector dual asociado al programa **R;** es decir

Se tendrá $(\mathbf{A^t y + z})\ \mathbf{x} = \mathbf{C^t\ x}$

De donde se tiene $\mathbf{A^t y\ x + z\ x = c^t\ x}$, ó

$$\mathbf{A^t y\ x - c^t\ x = -z\ x} \qquad \mathbf{(1)}$$

Pero ambos programas tienen solución óptima: $\mathbf{x}^*$ para **P** y $(\mathbf{y}^*)$ para **R**

Luego en (1), se tiene $\mathbf{A^t}\ (\mathbf{y}^*)\ \mathbf{x}^* - \mathbf{c^t\ x}^* = -\mathbf{z}^*\ \mathbf{x}^*$ (2)

Pero en el programa **P**: $\mathbf{A\ x}^* = \mathbf{b}$, al reemplazar en (2):

$$(\mathbf{y}^*)^t\ \mathbf{b} - \mathbf{c^t\ x}^* = -\mathbf{z}^*\ \mathbf{x}^*$$

Luego por el Teorema 2.2.1.8: $\mathbf{c^t\ x}^* = (\mathbf{y}^*)^t\ \mathbf{b}$

Concluyendo que $\mathbf{0} = (\mathbf{y}^*)^t\ \mathbf{b} - \mathbf{C^t\ x}^* = -\mathbf{z}^*\ \mathbf{x}^*$

$$z_j^* \, x_j^* = 0 \quad \text{para } j = 1, 2, .., n$$

Observación. . Las variables w y z, se denominan multiplicadores de Lagrange.

2.2.1.10 Lectura de las variables duales en el tablero símplex

El siguiente programa, será resuelto con el método símplex:

$$\begin{array}{llllll}
\text{Maximizar } Z & 3x_1 & +2x_2 & +5x_3 & \\
\text{S. a.} & x_1 & +2x_2 & +x_3 & \leq 430 \\
 & 3x_1 & & +2x_3 & \leq 460 \\
 & x_1 & +4x_2 & & \leq 420 \\
 & & X \geq 0 & &
\end{array}$$

Adicionando las variables de holgura: x_4 para la primera restricción, x_5 para la segunda y x_6 para la tercera restricción

$$\begin{array}{lllllllll}
\text{Maximizar } Z & 3x_1 & +2x_2 & +5x_3 & +0x_4 & +0x_5 & +0x_6 & \\
\text{S. a.} & x_1 & +2x_2 & +x_3 & +1x_4 & +0x_5 & +0x_6 & =430 \\
 & 3x_1 & & +2x_3 & +0x_4 & +1x_5 & +0x_6 & =460 \\
 & x_1 & +4x_2 & & +0x_4 & +0x_5 & +1x_6 & =420 \\
 & & & X \geq 0 & & & &
\end{array}$$

Se tiene el programa escrito en la forma estándar, a continuación se escriba en formato de tablero símplex

	Z	x_1	x_2	x_3	x_4	x_5	x_6	**RHS**
Z	1	- 3	- 2	-5	0	0	0	0
x_4	0	1	2	1	1	0	0	430
x_5	0	3	0	2	0	1	0	420
x_6	0	1	4	0	0	0	1	460

Después de algunas iteraciones, se tiene el tablero óptimo

	Z	x_1	x_2	x_3	x_4	x_5	x_6	**RHS**
Z	1	4	0	0	1	2	0	**1350**
x_2^*	0	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	**100**
x_3^*	0	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	**230**
x_6^*	0	2	0	0	- 2	1	1	**20**

El valor óptimo del programa es $z^* = \$1\,350$

La solución básica factible y óptima es $\mathbf{x}_B^* = \left[\mathbf{x}_2^*,\ \mathbf{x}_3^*,\ \mathbf{x}_6^*\right]^t = \left[\mathbf{100},\ \mathbf{230},\ \mathbf{20}\right]^t$ y es además es no degenerada, ya que $\mathbf{x}_B^* > \left[\mathbf{0},\ \mathbf{0},\ \mathbf{0}\right]^t$

La solución no básica es $\mathbf{x}_N^* = \left[\mathbf{x}_1^*,\ \mathbf{x}_4^*,\ \mathbf{x}_5^*\right]^t = \left[\mathbf{0},\ \mathbf{0},\ \mathbf{0}\right]^t$

También los Multiplicadores de Lagrange son: $\left[\mathbf{y}_1^*,\ \mathbf{y}_4^*,\ \mathbf{y}_5^*\right]^t = \left[\mathbf{1},\ \mathbf{2},\ \mathbf{0}\right]^t$

El programa dual será

$$
\begin{array}{ll}
\mathbf{T}: \mathbf{Min} & \mathbf{y}^* \mathbf{b} \\
\text{Sujeto a} & \mathbf{y}^* \mathbf{A} \le \mathbf{c} \\
 & \mathbf{y}^* \ge \mathbf{0}
\end{array}
$$

Esto es

$$
\begin{array}{llllll}
\textbf{Minimizar } Z & \mathbf{430\,y_1} & \mathbf{+\ 460\,y_2} & \mathbf{+420\ y_3} & & \\
\textbf{St} & \mathbf{y_1} & \mathbf{+\ 3\,y_2} & \mathbf{+\ \ y_3} & \mathbf{\ge 3} & \\
 & \mathbf{2\ y_1} & & \mathbf{+\ 4\,y_3} & \mathbf{\ge 2} & \mathbf{y \ge 0} \\
 & \mathbf{y_1} & \mathbf{+\ 2\,y_2} & & \mathbf{\ge 5} &
\end{array}
$$

2.2.1.11 Derivada direccional

Son los costos reducidos asociados. Es la proyección del gradiente de la función objetivo, sobre el eje de la variable correspondiente.

En el presente tablero. La derivada direccional de $\mathbf{x_1}$ es el costo reducido $\mathbf{z_1^* - c_1^* = 4}$

2.2.1.12 Variables duales.

Se nombra así, a los Costo Reducidos de las variables de holgura y las variables artificiales

En el tablero simplex óptimo se lee bajo las variables de holgura

	Z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
Z	**1**	$z_1^* - c_1^* = 4$	0	0	$y_1^* = 1$	$y_2^* = 2$	$y_3^* = 0$	**1350**

2.2.1.13 Gap dual.(Intervalo de dualidad)

Sean los programas Primal **P** y su correspondiente dual **D**

$$\begin{aligned} &\mathbf{P: Min\ f(x) = c^t\ x} \\ &\quad \mathbf{S.a.} \\ &\qquad \mathbf{Ax = b} \\ &\qquad \mathbf{x \geq 0} \end{aligned}$$

$$\begin{aligned} &\mathbf{D: Máx\ g(y) = b^t y} \\ &\quad \mathbf{S.a.} \\ &\qquad \mathbf{A^t y \leq c^t} \end{aligned}$$

Al estandarizar el dual, se tiene:

$$\begin{aligned} &\mathbf{D': Máx\ g(y) = b^t y} \\ &\quad \mathbf{S.a.} \\ &\qquad \mathbf{A^t y + z = c^t} \\ &\qquad\qquad \mathbf{z \geq 0} \end{aligned}$$

El gap dual será

$$\mathbf{c^t\ x - \ b^t y \ = \ (A^t y + z)\, x - b^t y \ = \ y^t A x + \ z^t\ x - b^t y = \ b^t y + \ z^t\ x - b^t y \ = \sum_{j=1}^{n} x_j\ z_j}$$

La diferencia entre el valor de la función objetivo del programa primal $\mathbf{c^T x}$ y la función objetivo del programa dual, se irá haciendo mas pequeña, en la medida que el programa se acerque al valor óptimo; esto es la brecha (gap dual), se irá acercando al valor cero

2.2.1.14 Condiciones de Karush Khun Tucker (KKT), para un P.M.L. con restricciones de desigualdad

Considérese el P.M.L: en forma Canónica:

$$\begin{aligned} &\mathbf{Min\ f(x) = cx} \\ &\mathbf{S.a.} \\ &\qquad \mathbf{Ax \geq b} \\ &\qquad \mathbf{x \geq 0} \end{aligned}$$

Si y es el vector de variables duales, asociado al presente programa, luego el dual correspondiente será:

$$\begin{aligned} &\textbf{Máx } f(y) = b^t y \\ &\textbf{S. a.} \\ &\qquad yA \le c \\ &\qquad y \ge 0 \end{aligned}$$

Escribiendo este último programa en la forma estándar:

$$\begin{aligned} &\textbf{Máx } f(y) = b^t y \\ &\textbf{S. a.} \\ &\qquad A^t y + z = c \\ &\qquad y, z \ge 0 \end{aligned}$$

Las condiciones de KKT, plantean lo siguiente:

Si x^*, w^* son soluciones óptimas del primal y del dual respectivamente, se cumple que:

$$\begin{aligned} &Ax^* \ge b, && x^* \ge 0 \\ &A^t y^* + z^* = c, && y^* \ge 0, z^* \ge 0, \\ &y^*(Ax^* - b) = 0, && z^* x^* = 0 \end{aligned}$$

2.2.1.15 El lagrangiano de un PNL, y el teorema de Karush Khun Tucker

Teorema. Considérese el siguiente programa convexo no lineal

$$\begin{aligned} &\textbf{Min} && f(x) \\ &\text{S.a.} && g_1(x) \le 0 \\ & && g_2(x) \le 0 \\ & && .. \\ & && g_m(x) \le 0 \\ &\text{Con} && x \in C \subset \mathbb{R}^n \end{aligned}$$

El lagrangiano $L(x, \lambda)$ de un programa convexo, es la función definida por

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x)$$

Para $\mathbf{x} \in \mathbf{C}$ y $\lambda \geq 0$

Nótese que el lagrangiano $\mathbf{L(x,\ \lambda)}$, es una función de **m + n** variables, siendo **m** el número de restricciones de desigualdad y **n** es el número de variables incluidas en la función objetivo y funciones de restricción

También nótese que si **(P)** es un programa convexo súper consistente, tal que su óptimo $\mathbf{f(x^*) = MP}$ es finito, entonces existe un vector de sensibilidad $\lambda^* \in \mathbb{R}^{\mathbf{n}}$, se verifica que

$$\mathbf{MP} = \inf_{\mathbf{x} \in \mathbf{C}} \left\{ L(x, \lambda^*) \right\} \qquad (*_1)$$

2.2.1.16 Forma Punto de Silla del teorema KKT

Teorema Supóngase que el programa $\mathbf{P}$, es convexo súper consistente. Entonces el punto $\mathbf{x^*} \in \mathbf{C}$, es una solución de $\mathbf{P}$, si y sólo si existe un $\lambda^* \in \mathbb{R}$, tal que:

a. $\lambda^* \geq 0$

b. $\mathbf{L(x^*,\ \lambda) \leq L(x^*,\ \lambda^*) \leq L(x,\ \lambda^*)}$ para todo $\mathbf{x} \in \mathbf{C}$ y para todo $\lambda \geq 0$

c. $\lambda_i^* \, \mathbf{g_i(x^*) = 0}$, para $\mathbf{i = 1, 2, \cdots, m}$

Demostración

$(\Rightarrow)$

i. Se probará primeramente en **(b)** que $\mathbf{L(x^*,\ \lambda^*) \leq L(x,\ \lambda^*)}$ y **(c)** $\lambda_i^* \, \mathbf{g_i(x^*) = 0}$, para $\mathbf{i = 1, 2, \cdots, m}$

Si $\mathbf{x^*}$ es una solución del programa **(P)**, entonces $\mathbf{x^*} \in \mathbf{C}$, $\mathbf{g(x^*) \leq 0}$ y $\mathbf{f(x^*) = MP}$, debido a $(*_1)$, existe $\lambda^* \geq 0$, en $\mathbb{R}^{\mathbf{n}}$, se tendrá

$$\mathbf{f(x^*)} = \inf_{\mathbf{x} \in \mathbf{C}} (\mathbf{L}(\mathbf{x}, \lambda^*))$$

Luego

$$L(x, \lambda^*) = f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geq \inf_{x \in C}(L(x,\lambda^*)) = f(x^*)$$, para todo $x^* \in C$

En particular si $x = x^*$, se tendrá

$$L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \geq f(x^*) \ \ldots (1)$$

Por otro como $\lambda_i \geq 0$ y $g_i(x^*) \leq 0$ se concluye que $\lambda_i > 0$, para i = 1, 2, . . . , **m**

$$f(x^*) \geq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) = L(x^*, \lambda^*) \ \ldots . (2)$$

De (1) y (2), se concluye que

$f(x^*) \geq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \geq f(x^*)$; lo que quiere decir que $\sum_{i=1}^{m} \lambda_i^* g_i(x^*) = 0$; de donde se tiene $\lambda_i^* g_i(x^*) = 0$, para $i = 1, 2, \cdots, m$

Por otro lado

$f(x^*) = L(x^*, \lambda^*) = \inf_{x \in C}(L(x,\lambda^*)) \leq L(x, \lambda^*)$, para todo $x \in C$

Se tiene $L(x^*, \lambda^*) \leq L(x, \lambda^*)$

Además $\lambda^* \geq 0$ y $g_i(x^*) \leq 0$, para $i = 1, 2, \cdots, m$

$$f(x^*) \geq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) = L(x^*, \lambda^*)$$

Se sabe que

$$L(x, \lambda^*) = f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geq \inf_{x \in C}(L(x,\lambda^*)) = f(x^*)$$

Al tomar $x = x^* \in C$, se tiene que

$$f(x^*) \geq L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \geq \inf_{x \in C}(L(x,\lambda^*)) = f(x^*)$$

De donde se tiene

$$\mathbf{f(x^*) \geq L(x^*,\lambda^*) = f(x^*) + \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) \geq \inf_{x\in C}(L(x,\lambda^*)) = f(x^*)}$$

$$\mathbf{f(x^*) \geq f(x^*) + \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) \geq f(x^*)}$$

Restando de cada miembro $\mathbf{f(x^*)}$, se obtiene

$$\mathbf{0 \geq \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) \geq 0} \Rightarrow \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) = 0$$

Si y sólo si $\mathbf{\lambda_i^* g_i(x^*) = 0}$, para todo $\mathbf{i = 1, 2, \cdots, m}$

Luego

$\mathbf{f(x^*) = L(x^*,\lambda^*) = \inf_{x\in C}(L(x,\lambda^*))}$; concluyendo que

$$\mathbf{L(x^*,\lambda^*) \leq L(x\ \ ,\lambda^*)}, \text{ para todo } \mathbf{x \in C}$$

ii. Seguidamente en **(b)** que $\mathbf{L(x^*,\lambda\ ) \leq L(x\ \ ,\lambda^*)}$

$$\mathbf{L(x^*,\lambda^*) = f(x^*) + \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*)} \text{ y}$$

$$\mathbf{L(x^*,\lambda\ ) = f(x^*) + \sum_{i=1}^{m}\lambda_i\, g_i(x^*)}$$

Al restar $\mathbf{L(x^*,\lambda^*) - L(x^*,\lambda\ ) = f(x^*) + \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) - \left(f(x^*) + \sum_{i=1}^{m}\lambda_i\, g_i(x^*)\right)}$

$$\mathbf{L(x^*,\lambda^*) - L(x^*,\lambda\ ) = \sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) - \sum_{i=1}^{m}\lambda_i\, g_i(x^*)}$$

Pero $\mathbf{\sum_{i=1}^{m}\lambda_i^*\, g_i(x^*) = 0}$

La relación anterior se vuelve

$$\mathbf{L(x^*,\lambda^*) - L(x^*,\lambda\ ) = -\sum_{i=1}^{m}\lambda_i\, g_i(x^*)}$$

Al tomar en cuenta que $\mathbf{g(x^*) \leq 0}$; se tendrá que $\mathbf{g_i(x^*) \leq 0}$, para todo $\mathbf{i = 1, 2, \cdots, m}$

Esto es $g_1(x^*) \leq 0$, $g_2(x^*) \leq 0$, . . . , $g_m(x^*) \leq 0$

Y también $\lambda \geq 0$, se tendrá que $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, . . . , λ_m

Al multiplicar $\lambda_1 g_1(x^*) \leq 0$, $\lambda_2 g_2(x^*) \leq 0$ . . . , $\lambda_m g_m(x^*) \leq 0$

De donde se obtiene que $\sum_{i=1}^{m} \lambda_i \; g_i(x^*) \leq 0 \Rightarrow -\sum_{i=1}^{m} \lambda_i \; g_i(x^*) \geq 0$

Concluyendo que

$$L(x^*, \lambda^*) - L(x^*, \lambda) = -\sum_{i=1}^{m} \lambda_i \; g_i(x^*) \geq 0$$

Es decir

$$L(x^*, \lambda) \leq L(x^*, \lambda^*)$$

Lo que completa la prueba

$(\Leftarrow)$

Supóngase que $x^* \in C$ y $\lambda^* \geq 0$ en $\mathbb{R}^m$; satisfacen **(b)**, para todo $x \in C$ y todo $\lambda \geq 0$ en $\mathbb{R}^m$

Para un i dado, tal que $1 \leq i \leq m$, sea

$$\lambda_j^{(i)} = \begin{cases} \lambda_j^* & \text{si } j \neq i \\ \lambda_j^* + 1 & \text{si } j = i \end{cases}$$

O también $0 \leq \lambda^{(i)} = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_{i-1}^*, \lambda_i^* + 1, \lambda_{i+1}^*, . . . + \lambda_m^*)$

Luego $\lambda^{(i)} = (\lambda_j^{(i)}) \geq 0$ y **(b)**, implica que de la relación $L(x^*, \lambda) \leq L(x^*, \lambda^*)$; se tenga

$$0 \geq L(x^*, \lambda^{(i)}) - L(x^*, \lambda^*) = g_i(x^*)$$

Esto prueba que x^* es factible para el programa **(P)**; más aún **(b)** implica que

$$f(x^*) = L(x^*, 0) \le L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*)$$

Pero $\lambda^* \ge 0$, $g_i(x) \le 0$, para $i = 1, 2, \cdots, m$

Así que $\lambda_i^* g_i(x^*) \le 0$; para todo $i = 1, 2, \cdots, m$

Lo que implica que $\sum_{i=1}^{m} \lambda_i^* g_i(x^*) \le 0$; sumando $f(x^*)$ a cada lado se tiene:

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \le f(x^*);$$

Concluyendo que

$$f(x^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) = L(x^*, \lambda^*);$$

Y así que $\lambda_i^* g_i(x^*) = 0$, para $i = 1, 2, \cdots, m$

Si se aplica **(b)** y se utiliza el hecho respecto al ínfimo de una función sobre un conjunto:

(i). Si $A \subset B$, entonces $\inf_{x \in B} \{h(x)\} \le \inf_{x \in A} \{h(x)\}$

(ii). Si $h(x) \le k(x)$, para todo $x \in A$, entonces

$$\inf_{y \in A} \{h(x)\} \le \inf_{x \in A} \{k(x)\}$$

En esta relación se puede apreciar que

$$\begin{aligned} f(x^*) = L(x^*, \lambda^*) = \; & \inf \left\{ L(x^*, \lambda^*) : x \in C \right\} \\ \le \; & \inf \left\{ L(x, \lambda^*) : x \in C, g_1(x) \le 0, \cdots, g_m(x) \le 0 \right\} \\ = \; & \inf \left\{ f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) : x \in C, g_1(x) \le 0, \cdots, g_m(x) \le 0 \right\} \\ \le \; & \inf \left\{ f(x) : x \in C, g_1(x) \le 0, \cdots, g_m(x) \le 0 \right\} = MP \end{aligned}$$

Se concluye que el punto $\mathbf{x}^*$ es factible para el programa **(P)**, tal que $\mathbf{f(x^*) = MP}$, esto es $\mathbf{x}^*$ es una solución del programa **(P)**

Un punto $(\mathbf{x}^*, \lambda^*)$ tal que $\mathbf{x}^* \in \mathbf{C}$, $\lambda^* \geq \mathbf{0}$ y que satisface la desigualdad:

(a) $\mathbf{L}(\mathbf{x}^*, \lambda) \leq \mathbf{L}(\mathbf{x}^*, \lambda^*) \leq \mathbf{L}(\mathbf{x}, \lambda^*)$, para todo $\mathbf{x} \in \mathbf{C}$ y $\lambda \geq \mathbf{0}$

Se le nombra como **"un punto de silla"**, para el lagrangiano del programa **(P)**; condición **(c)**; esto es

(c) $\boldsymbol{\lambda_i^* g_i(x^*) = 0}$, para $\mathbf{i = 1, 2, \cdots, m}$

A esta relación, se le nombra como: "Condición de holgura complementaria"

El Teorema de Karush Khun Tucker, afirma que si el programa **(P)** es superconsistente y convexo; entonces $\mathbf{x}^* \in \mathbf{C}$ es una solución para **(P)**, si y sólo si existe un $\lambda \geq \mathbf{0}$ en $\mathbb{R}^m$, tal que $\mathbf{L}(\mathbf{x}^*, \lambda^*)$ es un punto silla del Langrangiano del Programa (P) **(P)** y es tal que se satisface la condición de holgura complementaria $\boldsymbol{\lambda_i^* g_i(x^*) = 0}$

Si $(\mathbf{x}^*, \lambda^*)$ es un punto de silla del Langrangiano de cualquier programa convexo **(P)**, luego, se verifica que:

i. **MP** es finito y $\mathbf{x}^*$ es una solución de **(P)**

ii. Se satisface la condición de holgura complementaria:

$$\boldsymbol{\lambda_i^* g_i(x^*) = 0}, \text{ para } \mathbf{i = 1, 2, \cdots, m} \;\blacksquare$$

Si se imponen las condiciones que la función objetivo y las funciones de restricción, tiene primeras derivadas parciales y son continuas en **(P)**; entonces se tendrá la siguiente versión del teorema de KKT

2.2.1.17 Forma de Gradiente del teorema KKT

Teorema

Supóngase que el programa P, es convexo súper consistente, tal que la función objetivo. $\mathbf{f(x)}$ y las funciones de restricción $\mathbf{g}_1(x) \le 0, \cdots, \mathbf{g}_m(x) \le 0$ son continuas, con derivadas parciales de orden uno, bajo el conjunto C, para del programa **(P)** y un punto interior de C

Entonces $\mathbf{x}^*$ es solución de **(P)**; si y sólo si existe $\lambda^* \in \mathbb{R}^{\mathbf{n}}$, tal que

1. $\lambda_i^* \ge 0$, para $\mathbf{i = 1, 2, \cdots, m}$

2. $\lambda_i^* \, \mathbf{g}_i(\mathbf{x}^*) = \mathbf{0}$, para $\mathbf{i = 1, 2, \cdots, m}$

3. $\nabla \mathbf{f}(\mathbf{x}^*) + \sum_{\mathbf{i=1}}^{\mathbf{m}} \lambda_{\mathbf{i}}^* \, \nabla \mathbf{g}_{\mathbf{i}}(\mathbf{x}^*) = 0$

Demostración

$(\Rightarrow)$

Si $\mathbf{x}^*$ es una solución de **(P)**, entonces existe $\lambda^* \in \mathbb{R}^{\mathbf{m}}$, que satisface **(1)** y **(2)**, para los cuales es punto de silla del lagrangiano de **(P)**; pero entonces

$\mathbf{L}(\mathbf{x}^*, \lambda) \le \mathbf{L}(\mathbf{x}^*, \lambda^*) \le \mathbf{L}(\mathbf{x}, \lambda^*)$, para todo $\mathbf{x} \in \mathbf{C}$.

Así $\mathbf{x}^*$ es un mínimo global de $\mathbf{h}(\mathbf{x}) = \mathbf{L}(\mathbf{x}\ , \lambda^*)$ sobre **C**

Puesto que $\mathbf{x}^*$ es un punto interior de **C**, que tiene primeras derivadas parciales continuas sobre **C**, se sigue que $\nabla \mathbf{h}(\mathbf{x}^*) = 0$; esto es la condición **(3)** se mantiene ■

$(\Leftarrow)$

Recíprocamente supóngase que $\mathbf{x}^* \in \mathbf{C}$ y $\lambda^* \in \mathbb{R}^{\mathbf{m}}$, satisfacen la condición **(1)** , **(2)** y **(3)**

Si $\mathbf{x}$ es cualquier punto factible para **(P)**, entonces

$$f(x) \geq f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x);$$

$$\geq \left[f(x^*) + \nabla f(x^*).(x - x^*) \right] + \sum_{i=1}^{m} \lambda_i^* \left[g_i(x) + \nabla f(x^*).(x - x^*) \right]$$

Debido a que $\lambda_i^* \geq 0$, $g_i(x) \leq 0$, para $i = 1, 2, \cdots, m$ y porque $f(x)$, $g_i(x)$, para $i = 1, 2, \cdots, m$ son funciones convexas, se sigue que

$$f(x) \geq \left[f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \right] + \left[\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) \right].(x - x^*)$$

$$f(x) \geq f(x^*)$$

Debido a **(2)** y **(3)**

Así x^* es un mínimo global para $f(x)$, sobre el dominio admisible de **(P)** ■

2.2.1.18 El lagrangiano, para restricciones de igualdad

Considérese el programa no lineal convexo

$$PL = \begin{cases} \min \ f(x) \\ S.a. \\ g_i(x) = 0 \quad i = 1, \cdots, m \\ \text{con } x \in C \subseteq R^n \end{cases}$$

El Langrangiano $L(x, \lambda) = f(x) + \sum_{L=1}^{m} \lambda_i g_i(x)$, para $x \in C$, $\lambda \geq 0$

Teorema

Dado $x^* \in C$ y el programa convexo PL, se cumple que x^* es una solución de PL; si existe $\lambda_i^* \in R$, para $i = 1, \cdots, n$ tal que

i. $\lambda_i^* > 0$, **para** $i = 1, \cdots, n$

ii. $\lambda_i g_i(x) = 0$, $\lambda_i^* > 0$, **para** $i = 1, \cdots, n$

iii. $\underset{x,\lambda}{\nabla} L(x^*,\lambda^*) = 0$, $\lambda_i^* > 0$, **para** $i = 1, \cdots, n$

Ejemplo 2.2.1.3

Considere el programa

$$\begin{aligned} P:\ \text{Min } Z = f(x_1, x_2) = &\quad x_1^2 - 2x_1 + x_2^2 + 1 \\ \text{Sujeto a } g_1(x_1, x_2) = &\quad x_1 + x_2 \leq 0 \\ g_2(x_1, x_2) = &\quad x_1^2 - 4 \leq 0 \end{aligned}$$

Resolución

Nótese que es un programa **superconsistente, convexo**. Además las funciones $f(x)$, $g_1(x_1, x_2)$ y $g_2(x_1, x_2)$; posee primeras derivas parciales continuas (condiciones del presente teorema y se puede aplicar el Teorema de KKT

Para hallar la solución $x^* = (x_1^*, x_2^*)$ para el programa **P;** debe existir $\lambda^* = (\lambda_1^*, \lambda_2^*)$, tal que

1. $\lambda_1^* \geq 0$ y $\lambda_2^* \geq 0$

2. $\begin{cases} \lambda_1^* g_1(x_1, x_2) = 0 \\ \lambda_2^* g_2(x_1, x_2) = 0 \end{cases}$

3. $\nabla f(x_1, x_2) + \lambda_1^* \nabla g_1(x_1, x_2) + \lambda_2^* \nabla g_2(x_1, x_2) = 0$

De donde se obtiene el sistema

$$\begin{aligned} &\lambda_1^* \left(x_1^* + x_2^* \right) = 0 \\ &\lambda_2^* \left[\left(x_1^2 \right)^* - 4 \right] = 0 \\ &2x_1^* - 2 + 2x_2 + \lambda_1^* + 2\lambda_2^* x_1^* = 0 \\ &\qquad 2x_2 + \lambda_1^* \qquad = 0 \end{aligned}$$

x_1	x_2	λ_1	λ_2	g_1	g_2	f
0	0		0	0	-4	
(1/2)	(1/2)					

El sistema, formado por las ecuaciones **(1)** , **(2)** y **(3)**, tiene por soluciones

$\mathbf{x_1^* = 1,\quad x_2^* = 0,\quad \lambda_1^* = 0,\quad \lambda_2^* = 0}$ es no factible

$\mathbf{x_1^* = \frac{1}{2},\quad x_2^* = -\frac{1}{2},\quad \lambda_1^* = 1,\quad \lambda_2^* = 0}$ es factible

$\mathbf{f(1/2, 1/2) = 1}$

Lo que significa que el presente tiene solución única

Ejemplo 2.2.1.4

Para el programa

$\mathbf{min\ f(x_1,x_2)} = (\mathbf{x_1} - 5)^2 + (\mathbf{x_2} - 5)^2$ St $\mathbf{x_1^2} - \mathbf{x_2} \leq 6$ $\mathbf{x_1 + 3x_2} \leq 12$	$\mathbf{x_1} \geq \mathbf{0}$ $\mathbf{x_2} \geq \mathbf{x0}$

Verificar si el punto $(\mathbf{3}, 3)$, es factible

Resolución

$\mathbf{g_1(x_1, x_2)} = \mathbf{x_1^2 - x_2} - 6$ $\qquad$ $\mathbf{g_2(x_1, x_2)} = \mathbf{x_1 + 3x_2 - 12}$

$\mathbf{g_3(x_1, x_2)} = \mathbf{-x_1}$ $\qquad$ $\mathbf{g_4(x_1, x_2)} = \mathbf{-x_2}$

$\mathbf{g_1(x_1, x_2)} = \mathbf{x_1^2 - x_2} - 6$ $\mathbf{g_2(x_1, x_2)} = \mathbf{x_1 + 3x_2 - 12}$ $\mathbf{g_3(x_1, x_2)} = \mathbf{-x_1}$ $\mathbf{g_4(x_1, x_2)} = \mathbf{-x_2}$ $\mathbf{f\begin{pmatrix}3\\3\end{pmatrix} = 8}$	$\nabla \mathbf{g_1} = \begin{pmatrix}\mathbf{2x_1}\\\mathbf{-1}\end{pmatrix} = \begin{pmatrix}\mathbf{6}\\\mathbf{-1}\end{pmatrix},\ \nabla \mathbf{g_2} = \begin{pmatrix}\mathbf{1}\\\mathbf{3}\end{pmatrix}$ $\nabla \mathbf{g_3} = \begin{pmatrix}\mathbf{-1}\\\mathbf{0}\end{pmatrix},\ \nabla \mathbf{g_4} = \begin{pmatrix}\mathbf{0}\\\mathbf{-1}\end{pmatrix}$ $\nabla \mathbf{f} = \begin{pmatrix}\mathbf{2(x_1 - 5)}\\\mathbf{2(x_2 - 5)}\end{pmatrix} \mathbf{=} \nabla \mathbf{f} = \begin{pmatrix}\mathbf{-4}\\\mathbf{-4}\end{pmatrix}$
$\mathbf{g_1(3, 3)} = \mathbf{0},\ \mathbf{g_3(3, 3)} = \mathbf{-3}$ $\mathbf{g_2(3, 3)} = \mathbf{0}\ \ \mathbf{g_4(3, 3)} = \mathbf{-3}$ $\lambda_1 \mathbf{g_1(3, 3)} = \mathbf{0}$, $\lambda_3 \mathbf{g_3(3, 3)} = \mathbf{-3}\lambda_3$ $\lambda_2 \mathbf{g_2(3, 3)} = \mathbf{0}\ \ \mathbf{g_4(3, 3)} = \mathbf{-3}\lambda_4$	$\nabla \mathbf{g_1} = \begin{pmatrix}\mathbf{6}\\\mathbf{-1}\end{pmatrix},\ \nabla \mathbf{g_2} = \begin{pmatrix}\mathbf{1}\\\mathbf{3}\end{pmatrix}$ $\nabla \mathbf{g_3} = \begin{pmatrix}\mathbf{-1}\\\mathbf{0}\end{pmatrix},\ \nabla \mathbf{g_4} = \begin{pmatrix}\mathbf{0}\\\mathbf{-1}\end{pmatrix}$

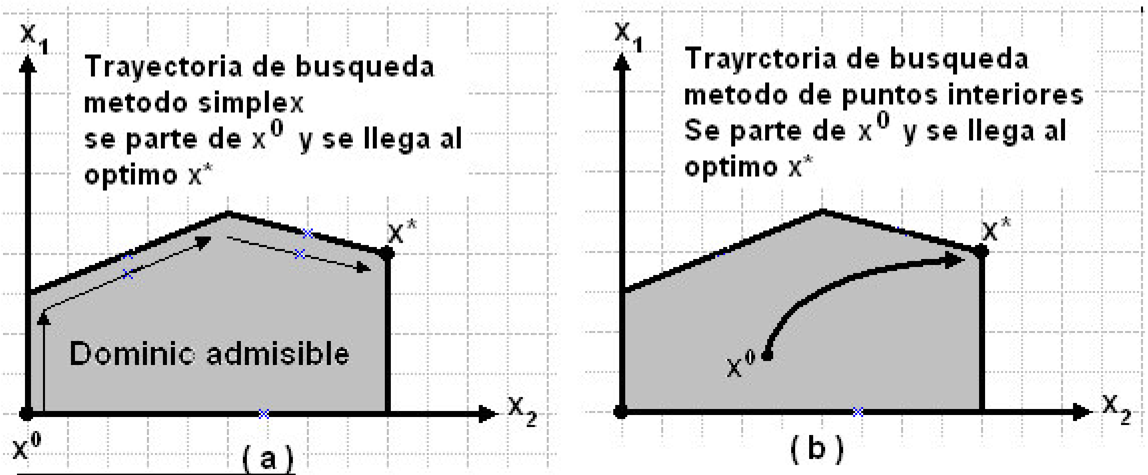
2.2.2 Método de puntos interiores

2.2.2.1 Presentación

Esta familia de métodos para resolver problemas de Programación Lineal se basa en la aplicación de métodos que originalmente se utilizaron para resolver problemas de Programación No lineal

A diferencia del método Símplex, que busca la solución por los puntos extremos del dominio admisible del programa; moviéndose de un punto extremo a otro, los Métodos de Puntos Interiores, se mueven por la parte interior del dominio admisible del programa.

Gráfica nª 1: Comparación de trayectorias de búsqueda



Fuente propia

¿Cómo buscan la solución, los métodos de Puntos Interiores?

Estos métodos basan su estrategia de búsqueda, para la solución de un P. M. L., en tres pauta principales:

i. Hallar un solución inicial en el dominio admisible del programa
ii. Definir una dirección de movimiento, tal que manteniendo la factibilidad del programa, se mueva a lo largo de ella; se avance hacia un nuevo punto, que mejore el valor actual del programa(Z)
iii. Detención de la búsqueda. Esto es cuántas veces se debe realizar la operación descrita en (ii) y ¿Cómo saber que se ha llegado al valor óptimo del programa?

No obstante estas son las tres cuestiones principales, que debe superar cualquier algoritmo de optimización. La manera de llevarlas adelante, hacen la diferencia entre esta familia de algoritmos y el Método Símplex

2.2.2.2 Clasificación de los métodos de puntos interiores

i) Método de Escalado afín

ii) Método basado en Transformaciones proyectivas

iii) Método de Path – following

iv) Método de Reducción de potencial

En el siguiente capítulo se abordará el método de Escalado Afín.

2.3 Marco Conceptual

2.3.1 Tipos de matrices

Teorema del rango

Si **A** es una matriz de m filas y n columnas; luego:

$$\mathbf{Rango(A) + nulidad(A) = n}$$

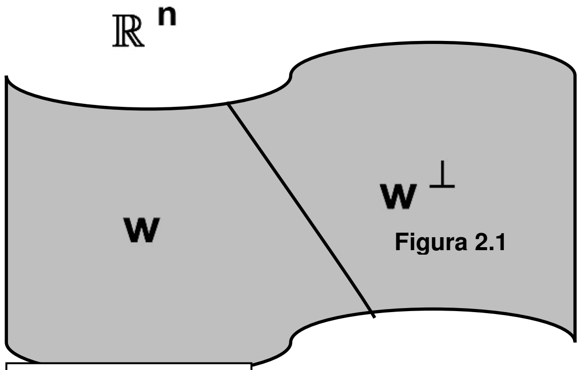
Demostración

Sea **R** , la forma escalonada reducida por renglones de la matriz **A** y supóngase que **rango(A) = r .** Luego **R** posee **r pivotes,** de modo que existen **r** variables pivotes y **n - r** **variables libres** en la solución de **Ax = 0 .** Puesto que **nulidad(A) = n - r** , se tiene

$$\mathbf{Rango(A) + nulidad(A) = r + (n - r) = n}$$

2.3.2 Complementos ortogonales

Gráfica nª 2: El sub espacio W y su complemento ortogonal



Si $v \in \mathbb{R}^n$ **y** $v \in$ **W**

$\Rightarrow$ **v** $\notin$ **W**$^{\perp}$

Por tanto

a. $\left(W^{\perp}\right)^{\perp} = W$

Fuente propia

Definición

Sea $\mathbf{W}$ un subespecie de $\mathbb{R}^{\mathbf{n}}$, se dice que un vector $\mathbf{v}$ en $\mathbb{R}^{\mathbf{n}}$ es ortogonal a $\mathbf{W}$; si $\mathbf{v}$ es ortogonal a todo vector en $\mathbf{W}$

Al conjunto de vectores que son ortogonales a $\mathbf{W}$, se les llama **complemento ortogonal de W** y se denota $\mathbf{W}^{\perp}$**;** esto es

$$\mathbf{W}^{\perp} = \left\{\mathbf{v} \text{ en } \mathbb{R}^{n} \text{: } \mathbf{v} * \mathbf{w} = \mathbf{0} \quad \textbf{para todo w en W}\right\}$$

Ejemplo 2.3.1

Sea $\mathbf{W} = \left\{ \begin{pmatrix} \mathbf{a} \\ \mathbf{a} \\ \mathbf{a} \end{pmatrix} : \mathbf{a} \in \mathbb{R} \right\}$

Hállese el complemento ortogonal de $\mathbf{W}$, esto es halle $\mathbf{W}^{\perp}$

Resolución

Una base para $\mathbf{W}$ es $\mathbf{W} = \{(\mathbf{1}, \ \mathbf{1}, \ 1)\}$, luego

$$\mathbf{W}^{\perp} = \left\{ (\mathbf{x}, \ \mathbf{y}, \ \mathbf{z}) : (\mathbf{x}, \ \mathbf{y}, \ \mathbf{z}) \perp \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} \right\}$$

$$= \left\{ (\mathbf{x}, \ \mathbf{y}, \ \mathbf{z}) : (\mathbf{x}, \ \mathbf{y}, \ \mathbf{z}) . \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} = \mathbf{0} \right\}$$

$$= \{(\mathbf{x}, \ \mathbf{y}, \ \mathbf{z}) : \mathbf{x} + \mathbf{y} + \mathbf{z} = \mathbf{0}\}$$

$$= \{(\mathbf{x}, \ \mathbf{y}, \ -\mathbf{x} - \mathbf{z})\}$$

$$= \{(\mathbf{x}, \ \mathbf{y}, \ -\mathbf{x} - \mathbf{z}) : \mathbf{x}, \mathbf{y} \in \mathbb{R}\}$$

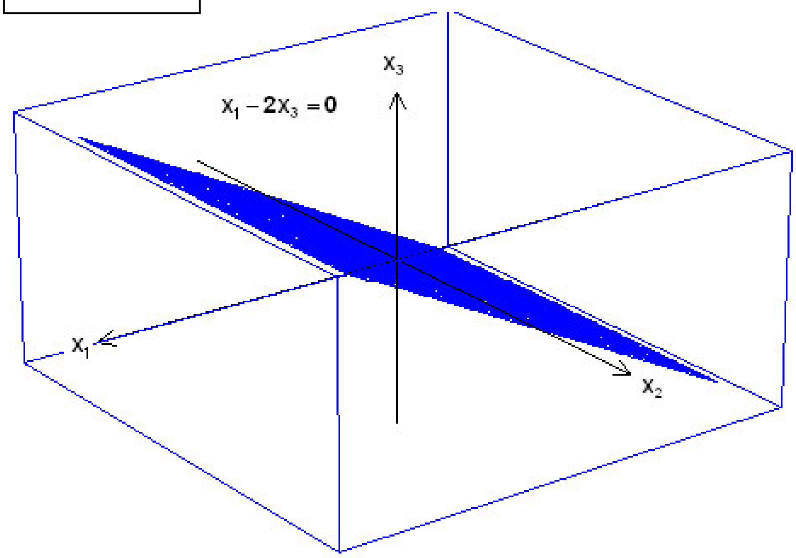
$$= \{\mathbf{x}(\mathbf{1}, \mathbf{0}, -\mathbf{1}) + \mathbf{y}(\mathbf{0}, \mathbf{1}, -\mathbf{1}) : \mathbf{x}, \mathbf{y} \in \mathbb{R}\}$$

$$\mathbf{W}^{\perp} = \{(\mathbf{1}, \mathbf{0}, -\mathbf{1}), \ (\mathbf{0}, \mathbf{1}, -\mathbf{1})\}$$

Ejemplo 2.3.2

Sea $\mathbf{W} = \left\{\mathbf{x} \in \mathbb{R}^{\mathbf{3}} : \mathbf{x}_1 - \mathbf{2x}_3 = \mathbf{0}\right\}$

Gráfica nª 3



Hiperplano $x_1 - 2x_2 = 0$

Hállese $\mathbf{W}^{\perp}$

Resolución

Primero se hallará una base para $\mathbf{W}$

En efecto

$$\mathbf{W} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x}_2 - 2\mathbf{x}_3 = 0\}$$

$$= \{\mathbf{x}_1 ,\ \ 2\mathbf{x}_2 ,\ \ \mathbf{x}_3 : \mathbf{x}_1 , \mathbf{x}_3 \in \mathbb{R}\}$$

$$= \{\mathbf{a}(1,0,0) + \mathbf{b}(0,2,1) : \mathbf{a}, \mathbf{b} \in \mathbb{R}\}$$

La base buscada es

$$\mathbf{B} = \{(1,0,\ 0),\ (0,2,1)\}$$

$$\mathbf{W}^{\perp} = \{(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) : (\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) \perp (1,0,0), (\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) \perp (0,2,1)\}$$

$$= \{(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) : (\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3).(1,0,0)^t = 0, (\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3).(0,2,1)^t = 0\}$$

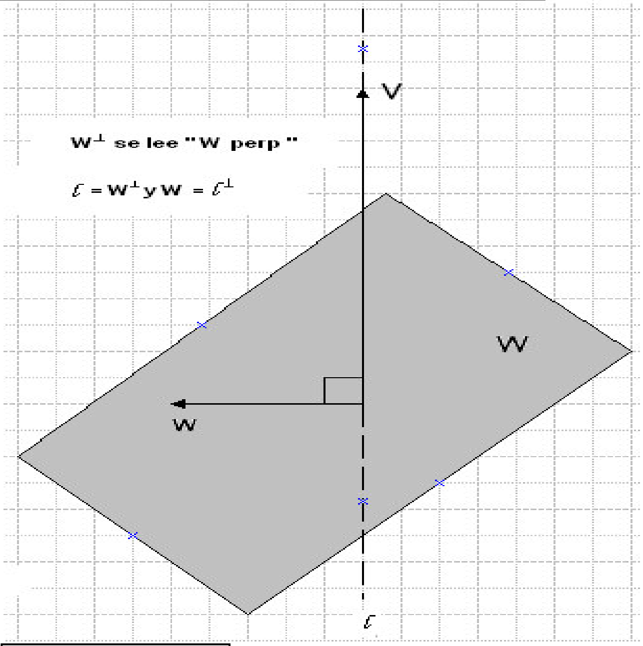
$$= \{(\mathbf{x}_1,\ \ \mathbf{x}_2,\ \ \mathbf{x}_3) : \mathbf{x}_1 = 0,\ 2\mathbf{x}_2 + \mathbf{x}_3 = 0\}$$

$$= \{\mathbf{x}_2(0,\ \ 1,\ \ -2) : \mathbf{x}_2 \in \mathbb{R}\}$$

$\mathbf{W}^{\perp} = \{(\mathbf{0},\ \mathbf{1},\ \mathbf{-2})\}$

Ejemplo 2.3.3

Gráfica n ° 4: Recta l, perpendicular al plano W



Fuente propia

Si **W** es un plano que pasa por el origen en $\mathbb{R}^3$ y l es la recta que pasa por el origen perpendicular a **W** (esto es, paralela al vector normal a **W**), entonces todo vector v en l es ortogonal a todo vector w en W. Por tanto a l = $\mathbf{W}^{\perp}$

Más aún W, consiste de aquellos vectores w que son ortogonales a todo v en l; por tanto también se tiene **W** = $l^{\perp}$

En el presente ejemplo, se evidencia que el complemento ortogonal de un subespacio es otro subespacio.

También el complemento del complemento de un subespacio; es el subespacio original

2.3.3 Proyecciones ortogonales

Teorema

(Matriz de proyección ortogonal sobre el subespacio del rango de una matriz**)**

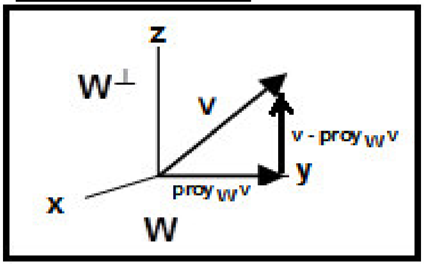
Sea $\mathbf{v}$ un vector de $\mathbb{R}^{\mathbf{n}}$ y sea $\mathbf{W}$ un subespacio de $\mathbb{R}^{\mathbf{n}}$, cuya base la integran los vectores $\{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}, \ldots, \mathbf{a_r}\}$. Si $\mathbf{A}$ es la matriz cuyas columnas son los vectores $\mathbf{a_j}$. Entonces la proyección ortogonal de $\mathbf{v}$ sobre A:

$$\mathbf{proy_W\ v} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \ldots + \lambda_k \mathbf{a}_k$$

Siendo los coeficientes λ_j, las entradas de $\left(\mathbf{A^T A}\right)^{-1} \mathbf{A^T . v}$

Demostración

Gráfica n ° 5



Proyección del vector v

Primeramente el vector $\mathbf{proy_W\ v}$ pertenece a $\mathbf{W}$; por lo que se puede escribir como combinación lineal de los vectores de la base de $\mathbf{W}$; eso es:

$$\mathbf{proy_W\ v} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \ldots + \lambda_k \mathbf{a}_k = \mathbf{A^T}\lambda$$

Siendo $\lambda \in \mathbb{R}^{\mathbf{k}}$, por lo que $\mathbf{v} - \mathbf{proy_W\ v}$ es ortogonal

A cada miembro de $\mathbf{W}$ y por lo tanto a cada miembro $\mathbf{a}_j$ de su base, esto es

$(\mathbf{v} - \mathbf{proy_W\ v}) \perp \mathbf{a_j}$ Para todo $\mathbf{j}$ = 1, 2, . . . , $\mathbf{k}$

$\langle \mathbf{v} - \mathbf{A}\lambda, \mathbf{a}_j \rangle = \mathbf{0}$, para todo $\mathbf{j}$ = 1, 2, . . . , $\mathbf{k}$

$\mathbf{a}^j \langle \mathbf{v} - \mathbf{A}\ \lambda \rangle = \mathbf{0}$, para todo i = 1, 2, . . . , $\mathbf{k}$

$\mathbf{A}^{\mathbf{T}} \langle \mathbf{v} - \mathbf{A}\ \lambda \rangle = \mathbf{0} = \mathbf{A^T v} - \mathbf{A^T A}\lambda$; de donde se concluye que

$\lambda = (\mathbf{A^T A})^{-1} \mathbf{A^T}.\ \mathbf{v}$

Siendo la matriz $(\mathbf{A^T A})$ de orden $\mathbf{k}$, simétrica, invertible con $\mathbf{rango}(\mathbf{A^T A}) = \mathbf{k}$

La matriz $(A^T A)^{-1} A^T$ se denomina: matriz de proyección ortogonal, sobre el subespacio

2.3.4 Complemento ortogonal

Teorema

Si $\mathbf{A}$ es una matriz de m filas y n columnas; luego el complemento ortogonal del espacio renglón de $\mathbf{A}$, es el espacio nulo de $\mathbf{A}$

Gráfica n ° 6

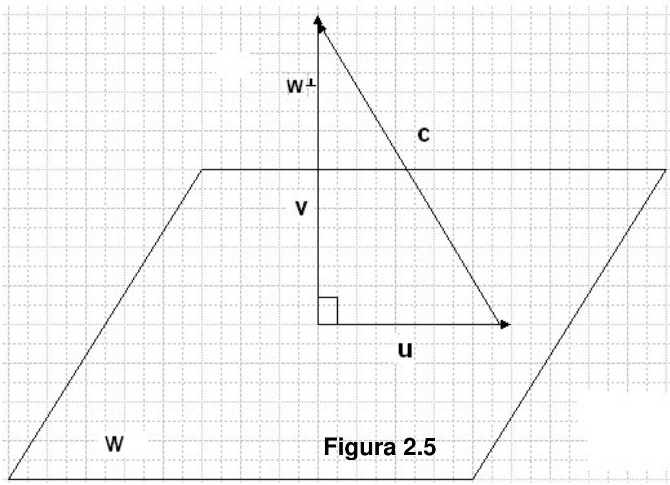


Figura 2.5

Complemento ortogonal del vector u

Demostración

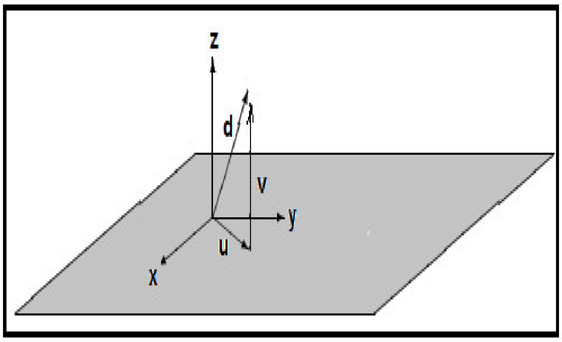
Si $\mathbf{u}$ es un vector de $\mathbb{R}^n$ luego $\mathbf{u}$ estará en $(\mathbf{renglón(A)})^{\perp}$, si y sólo si $\mathbf{u}$ es ortogonal a todo renglón de $\mathbf{A}$. Pero esto será cierto si y sólo si $\mathbf{Au = 0}$, lo cual equivale a que $\mathbf{u}$ está en $\mathbf{Null(A)}$ ■

También como $\mathbf{null(A)} = \{\mathbf{u : Au = 0,\ u} \in \mathbb{R}^n\}$ y $\mathbf{Null(A)}$ es un sub espacio de $\mathbb{R}^n$; se verifica que

$$\mathbb{R}^n = \mathbf{Null(A)} \oplus \mathbf{Renglón(A)}$$

Observación - 1

Gráfica n ° 7: Del complemento ortogonal



Fuente propia

Considérese un vector $\mathbf{d} \in \mathbb{R}^{\mathbf{n}}$, luego $\mathbf{d} = \mathbf{u} + \mathbf{v}$, con $\mathbf{u} \in \mathbf{Null(A)}$ y $\mathbf{v} \in \mathbf{renglón(A)}$.

¿Qué forma tienen los $\mathbf{v} \in \mathbf{renglón(A)}$?

Que $\mathbf{v}$ pertenezca al sub espacio generado por los vectores renglones de la matriz

$\mathbf{A}$; es equivalente a decir que $\mathbf{v} = \mathbf{gen}(\,\mathbf{a}^1, \mathbf{a}^2, \cdots, \mathbf{a}^m)$

Donde $\mathbf{a}^{\mathbf{i}}$ denota el i-ésimo renglón de la matriz $\mathbf{A}$

Luego existen reales $\mathbf{y_1}$, $\mathbf{y_2}$, ... , $\mathbf{y_m}$, tales que

$$\mathbf{v} = \mathbf{y_1 a^1 + y_2 a^2 + y_3 a^3 + \cdots + y_m a^m}$$

$$\mathbf{v} = \mathbf{y_1(a_{11}, a_{12}, a_{13}, \cdots, a_{1n}) + y_2(a_{21}, a_{22}, a_{23}, \cdots, a_{2n}) + . \quad . \quad .}$$
$$\mathbf{y_m(a_{m1}, a_{m2}, a_{m3}, \cdots, a_{mn})}$$

$$\mathbf{v} = \left(\mathbf{y_1 a_{11}}\,,\ \mathbf{y_1 a_{12}}\,,\ \mathbf{y_1 a_{13}} \cdots,\ \mathbf{y_1 a_{1n}}\right) + \mathbf{(a_{21}y_2, a_{22}y_2, a_{23}y_2, \cdots, a_{2n}y_2) + \ldots .}$$
$$.\ \mathbf{+ (a_{m1}y_m, a_{m2}y_m, a_{m3}y_m, \cdots, a_{mn}y_m)}$$

$$\mathbf{v} = \mathbf{(y_1\, a_{11} + \ y_2\, a_{21} + \ y_3\, a_{31} \cdots,\ y_m\, a_{m1}\,,}$$
$$\mathbf{y_1\, a_{12} + \ y_2\, a_{22} + \ y_3\, a_{32} \cdots,\ y_m\, a_{m2}\,,}$$
$$\mathbf{y_m\, a_{m1} + y_m\, a_{m2} + \ y_m\, a_{m3}, \cdots, y_m a_{mn})}$$

$$\mathbf{v} = \begin{bmatrix} \mathbf{a_{11}} & \mathbf{a_{21}} & . & \mathbf{a_{m1}} \\ \mathbf{a_{12}} & \mathbf{a_{22}} & . & \mathbf{a_{m2}} \\ \mathbf{a_{13}} & \mathbf{a_{23}} & . & \mathbf{a_{m3}} \\ . & . & . & . \\ \mathbf{a_{1n}} & \mathbf{a_{2n}} & . & \mathbf{a_{mn}} \end{bmatrix} \begin{bmatrix} \mathbf{y_1} \\ \mathbf{y_2} \\ \mathbf{y_3} \\ . \\ \mathbf{y_m} \end{bmatrix} = \mathbf{A^t y}$$

Equivalentemente $\mathbf{v} \in \mathbf{Im(A^t)}$

De donde todo $\mathbf{d} \in \mathbb{R}^{\mathbf{n}}$, se puede escribir como

$$\mathbf{d = u + A^t y}$$

Ejemplo 2.3.4

Para la matriz $A = \begin{bmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -9 \end{bmatrix}$ Determinar::

a. Una base para **Col(A)**, su dimensión

b. **Im(A)** y una base: $B_{Im(A)}$, para el sub espacio de las imágenes de **A**, su dimensión

c. $Im(A^T)$ y una base: $B_{Im(A^T)}$ para el sub espacio de las imágenes de A^T

d. **Null(A)** y una base: $B_{Null(A)}$, para el sub espacio nulo de **A**: **Null(A)**, su nulidad

e. ¿Qué relación guardan los elementos de **Null(A)** con los de $Im(A^T)$?

Resolución

a. Obténgase la matriz escalonada reducida

$$\begin{bmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 3 \\ 0 & 0 & 0 \\ -6 & 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow R = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow$$

$r_3 = 2r_1 + r_2$ y $r_1 = -2r_2$, luego $B_{Col(A)} = \left\{ \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix} \right\}$, su dimensión es **1**

b. $Im(A) = Rec(A) = \left\{ v \in \mathbb{R}^3 : Ax = v, \text{ para alguna } x \in \mathbb{R}^3 \right\}$

$$\begin{bmatrix} 2 & -1 & 3 & | & v_1 \\ 4 & -2 & 6 & | & v_2 \\ -6 & 3 & 9 & | & v_3 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 3 & | & v_1 \\ 0 & 0 & 0 & | & v_2 - 2v_1 \\ 0 & 0 & 0 & | & v_3 + 3v_1 \end{bmatrix} \Rightarrow$$

$$\begin{cases} v_2 - 2v_1 = 0 \\ v_3 + 3v_1 = 0 \end{cases} \text{ ó } \begin{cases} v_2 = 2v_1 \\ v_3 = -3v_1 \end{cases}$$

$$Im(A) = \left\{ v \in \mathbb{R}^3 : v_2 = 2v_1, v_3 = -3v_1 : \begin{pmatrix} v_1 \\ 2v_1 \\ -3v_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \alpha, \quad \alpha \in \mathbb{R} \right\} \text{ y}$$

$$B_{Im(A)} = \left\{ \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix} \right\}$$

Observe que $\mathbf{B_{Col(A)}} = \left\{\begin{pmatrix}\mathbf{-1}\\ \mathbf{-2}\\ \mathbf{3}\end{pmatrix}\right\} = \mathbf{B_{Im(A)}} = \left\{\begin{pmatrix}\mathbf{-1}\\ \mathbf{-2}\\ \mathbf{3}\end{pmatrix}\right\}$

c.

$$A^T = \begin{bmatrix} 2 & 4 & -6 \\ -1 & -2 & 3 \\ 3 & 6 & -9 \end{bmatrix}$$

$$\mathbf{Im(A^T) = Rec(A^T) = \left\{v \in \mathbb{R}^3 : A^T y = v,\ para\ alguna\ y \in \mathbb{R}^3\right\}}$$

$$\left[\begin{array}{ccccc} 2 & 4 & -6 & \mathrm{I} & v_1 \\ -1 & -2 & 3 & \mathrm{I} & v_2 \\ 3 & 6 & -9 & \mathrm{I} & v_3 \end{array}\right] \sim \left[\begin{array}{ccccc} 0 & 0 & 0 & \mathrm{I} & v_1 + 2v_2 \\ -1 & -2 & 3 & \mathrm{I} & v_2 \\ 0 & 0 & 0 & \mathrm{I} & v_3 + 3v_2 \end{array}\right] \Rightarrow$$

$$\begin{cases} v_1 + 2v_2 = 0 \\ v_3 + 3v_2 = 0 \end{cases} \quad \text{ó} \quad \begin{cases} v_1 = -2v_2 \\ v_3 = -\ 3v_2 \end{cases}$$

$$\mathbf{Im(A^T)} = \left\{ v \in \mathbb{R}^3 : \begin{pmatrix} -2v_2 \\ v_2 \\ -3v_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ -3 \end{pmatrix} \alpha,\ \alpha \in \mathbb{R} \right\}$$

También $\mathbf{B_{Im(A^T)}} = \left\{\begin{pmatrix}\mathbf{-2}\\ \mathbf{1}\\ \mathbf{-3}\end{pmatrix}\right\}$ y $\boldsymbol{\rho(A^T) = 1}$

d.

$$\mathbf{Null(A) = \left\{x \in \mathbb{R}^3 : Ax = 0\right\}}$$

$$\left[\begin{array}{ccccc} 2 & -1 & 3 & \mathrm{I} & 0 \\ 4 & -2 & 6 & \mathrm{I} & 0 \\ -6 & 3 & -9 & \mathrm{I} & 0 \end{array}\right] \sim \left[\begin{array}{ccccc} 2 & -1 & 3 & \mathrm{I} & 0 \\ 0 & 0 & 0 & \mathrm{I} & 0 \\ 0 & 0 & 0 & \mathrm{I} & 0 \end{array}\right] \Rightarrow \{x_2 = 2x_1 + 3x_3$$

$$\mathbf{Null(A)} = \left\{ x \in \mathbb{R}^3 : x_2 = 2x_1 + 3x_3 : \begin{pmatrix} x_1 \\ 2x_1 + 3x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}\alpha + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}\beta,\ \alpha, \beta \in \mathbb{R} \right\}$$

$\mathbf{B_{Null(A)}} = \left\{\begin{pmatrix}\mathbf{1}\\ \mathbf{2}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\ \mathbf{3}\\ \mathbf{1}\end{pmatrix}\right\}$, su nulidad **2**

e.

$$\mathbf{B}_{\mathbf{Null(A)}} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \right\} \text{ y } \mathbf{B}_{\mathbf{Im(A^T)}} = \left\{ \begin{pmatrix} -2 \\ 1 \\ -3 \end{pmatrix} \right\}$$

Se observa que son ortogonales ■

2.3.5 Descomposición ortogonal

Teorema

Sea $\mathbf{A}$ una matriz de $\mathbf{m}$ filas, $\mathbf{n}$ columnas, siendo $\mathbf{m<n}$ y $\mathbf{rango(A)=m}$. Entonces si $\mathbf{d} \in \mathbb{R}^{\mathbf{n}}$, se puede escribir como $\mathbf{d=u+v}$, con $\mathbf{u} \in \mathbf{Null(A)}$ y $\mathbf{v} \in \mathbf{renglón(A)}$.Entonces la matriz de proyección ortogonal sobre el espacio nulo de es

$$\mathbf{P = I_n - A^t \left(A A^t\right)^{-1} A}$$

Demostración

Como $\mathbf{u} \in \mathbf{Null(A)}$, $\mathbf{Au=0,}$ $\mathbf{u} \in \mathbb{R}^{\mathbf{n}}$

También como $\mathbf{v} \in \mathbf{Im(A)},$ $\mathbf{v = A^t y}$, por la **Observación - 1**

De la relación $\mathbf{d=u+v}$; se tiene $\mathbf{u=d-v=d-A^t y}$

Multiplicando ambos miembros de $\mathbf{u=d-A^t y}$, por la matriz $\mathbf{A}$, resulta

$$\mathbf{0 = Au = A\left(d - A^T y\right) = Ad - AA^T y}, \quad \mathbf{y = \left(AA^t\right)^{-1} Ad}$$

También de $\mathbf{u=d-A^t y}$ al reemplazar el valor de $\mathbf{y}$, se obtiene

$$\mathbf{u = d - A^t \left(AA^t\right)^{-1} Ad} = \mathbf{\left(I_n - A^t \left(AA^t\right)^{-1} A\right)d}$$

Por lo tanto $\mathbf{P = \left(I_n - A^t \left(AA^t\right)^{-1} A\right)}$ ■

2.3.6 Propiedades de la matriz P

Del **Teorema 2.3.5**, se tiene que $\mathbf{d} = \mathbf{u} + \mathbf{v}$, con $\mathbf{u} \in \mathbf{Null}(\mathbf{A})$ y $\mathbf{v} \in$ $\mathbf{Im}(\mathbf{A}^t)$

La matriz $\mathbf{P} = \mathbf{I}_n - \mathbf{A}^t\left(\mathbf{A}\mathbf{A}^t\right)^{-1}\mathbf{A}$, posee las siguientes propiedades

i. $\mathbf{Pd}$: Es la proyección ortogonal sobre el espacio nulo de la matriz A: **N(A)**

ii. $Pd = 0$ ¿ $\mathbf{A}\,\mathbf{P} = \mathbf{0}$

iii. $\mathbf{P}^t = \mathbf{P}$ ($\mathbf{p}$ es simétrica)

iv. $\mathbf{P}^2 = \mathbf{P}$ ($\mathbf{P}$ es idempotente)

v. ¿Qué significa que $Pd = 0$?

Demostración

i. En efecto como $\mathbf{Pd} = \left(\mathbf{I}_n - \mathbf{A}^t\left(\mathbf{A}\mathbf{A}^t\right)^{-1}\mathbf{A}\right)\mathbf{d} = \mathbf{u}$ por Teorema 2.3.5

Luego $\mathbf{Pd} \in \mathbf{Null}(\mathbf{A})$ ■

ii.
$$\begin{aligned}
\mathbf{A}\,\mathbf{Pd} &= \mathbf{A}(\mathbf{I}_n - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A})\mathbf{d} \\
&= \mathbf{A}\,\mathbf{I}_n\mathbf{d} - \mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\,\mathbf{d} \\
&= \mathbf{A}\,\mathbf{d} - \mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\,\mathbf{d} \\
&= \mathbf{A}\,\mathbf{d} - \mathbf{A}\,\mathbf{d} = \mathbf{0}
\end{aligned}$$

iii.
$$\begin{aligned}
\mathbf{P}^t &= (\mathbf{I}_n - \mathbf{A}^T(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A})^T \\
\mathbf{P}^t &= \mathbf{I}_n^T - (\mathbf{A}^T(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A})^T \\
\mathbf{P}^t &= \mathbf{I}_n - \left[(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A})\right]^T(\mathbf{A}^T)^T \\
\mathbf{P}^t &= \mathbf{I}_n - \left[\mathbf{A}^T(\mathbf{A}^T\mathbf{A})^{-1})^T\right]^T\mathbf{A} \\
\mathbf{P}^t &= \mathbf{I}_n - \left[\mathbf{A}^T(\mathbf{A}^T\mathbf{A})^T)^{-1}\right]\mathbf{A}
\end{aligned}$$

$$P^t = I_n - A^T(AA^T)^{-1}A = P$$

$$P^t = P$$

iv. $P^2 = PP = \left[I_n - A^t(AA^t)^{-1}A\right]\left[I_n - A^t(AA^t)^{-1}A\right]$

$$= I_n - A^t(AA^t)^{-1}A - A^t(AA^t)^{-1}A + A^tI_n(AA^t)^{-1}A$$

$$= I_n - A^t(AA^t)^{-1}A - A^t(AA^t)^{-1}A + A^t(AA^t)^{-1}A$$

$$= I_n - A^t(AA^t)^{-1}A - A^t(AA^t)^{-1}A + A^t(AA^t)^{-1}A$$

$$= I_n - A^t(AA^t)^{-1}A = P$$

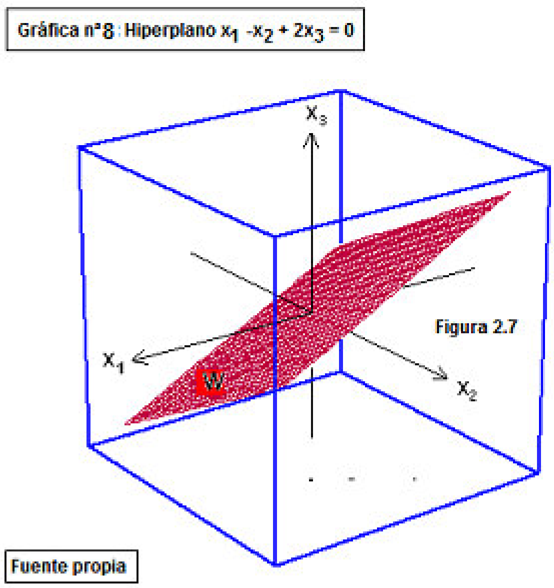
$$P^2 = P$$

v. Si Pd = 0; quiere decir que la proyección ortogonal del vector **d**, sobre el subespacio nulo de **A**; es el punto (0 , 0, .. 0); lo equivale a decir que d es ortogonal al subespacio nulo de A; es decir d está en el complemento ortogonal al subespacio nulo de A; que es el subespacio de las filas de A; esto es; existe y tal que **d** = A^ty

Ejemplo 2.3.5

Sea **W** un subespacio en $\mathbb{R}^3$, con ecuación $\mathbf{x_1 - x_2 + 2x_3 = 0}$ y sea **d** =

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$



En este caso $\mathbf{W} = \left\{ \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \mathbf{x_3} \end{bmatrix} : [\mathbf{1}, \ \ \mathbf{-1}, \ \ \mathbf{2}] \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \mathbf{x_3} \end{bmatrix} = \mathbf{0} \right\}$, siendo $\mathbf{A} = [\mathbf{1}, \ \ \mathbf{-1}, \ \ \mathbf{2}]$

a. Hallar una base para $\mathbf{W}$ y para $\mathbf{W}^{\perp}$

b. Hallar la matriz de proyección ortogonal $\mathbf{P}$ sobre $\mathbf{W}$ y $\mathbf{Q}$ sobre $\mathbf{W}^{\perp}$

c. Hallar la proyecciones ortogonales de $\mathbf{d}$ sobre $\mathbf{W}$ y sobre $\mathbf{W}^{\perp}$

Resolución

a. Hállese primero $\mathbf{W}$, que es el espacio nulo de la matriz $\mathbf{A}$ ($\mathbf{W}$ = **Null(A)**)

$$\mathbf{W} \ = \ \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{x_2} = \mathbf{x_1} + \mathbf{2x_3} : \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_1} + \mathbf{2x_3} \\ \mathbf{x_3} \end{bmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \alpha + \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \beta, \quad \alpha, \beta \in \mathbb{R} \right\}$$

Siendo una base para $\mathbf{W}$: $\mathbf{B_W} = \left\{ \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{2} \\ \mathbf{1} \end{pmatrix} \right\}$, la que se denotará con

$$\mathbf{B} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

Por tanto $W^{\perp} =$ $\mathbf{(Null(A))^{\perp} = Im(A^t) = gen}\left\{\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}\right\} = \mathbf{B_{W^T}} = \left\{\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}\right\}$

Sea **P** la matriz de proyección ortogonal sobre **W**, cuya base es

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Luego $\mathbf{P = B\left(B^tB\right)^{-1} B^t} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}\left(\frac{1}{6}\right)\begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \left(\frac{1}{6}\right)\begin{bmatrix} 5 & -2 \\ 1 & 2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} =$$

$$\mathbf{P} = \left(\frac{1}{6}\right)\begin{bmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$$

Y la matriz **Q** de proyección ortogonal sobre $\mathbf{W^{\perp}}$, cuya base es

$$\mathbf{d} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Luego por ser **C** un vector unitario; se tendrá que

$$\mathbf{Q} = \frac{\mathbf{d\,d^t}}{\|\mathbf{d}\|^2} = \frac{\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}\begin{pmatrix} 1, & -1, & 2 \end{pmatrix}}{6} = \left(\frac{1}{6}\right)\begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & -4 \end{bmatrix}$$

2.3.7 Relación entre el Rango de la matriz A y Rango de A^tA

Teorema

Sea **A** una matriz de **m** filas, **n** columnas, luego $\mathbf{rango(A^t A) = rango(A)}$

Demostración

Puesto la matriz $\mathbf{A^t A}$ es de orden $\mathbf{n}$, tiene el mismo número de columnas que $\mathbf{A}$; por el **TEOREMA 2.1**, se tiene que

$$\mathbf{rango(A) + nulidad(A) = n = rango(A^t A) + Nulidad(A^t A)}$$

Por tanto para demostrar que $\mathbf{rango(A)} = \mathbf{rango(A^t A)}$; bastará demostrar que $\mathbf{Nulidad(A) = Nulidad(A^t A)}$, será así al establecer que la dimensión de los espacios de $\mathbf{A}$ y $\mathbf{A^t}$ son iguales

Así: sea $\mathbf{x} \in \mathbf{Null(A)}$; luego $\mathbf{Ax = 0}$, de donde se tiene $\left(\mathbf{A^t Ax}\right) = \mathbf{A^t 0 = 0}$, lo que quiere decir que $\mathbf{x} \in \mathbf{Null(A^t A)}$

Recíprocamente: Sea $\mathbf{x} \in \mathbf{Null(A^t A)}$, luego $\left(\mathbf{A^t A}\right)\mathbf{x = 0}$, de donde se tiene que $\mathbf{x^t}\left(\mathbf{A^t A}\right)\mathbf{x = x^t 0 = 0}$; pero $\mathbf{x^t}\left(\mathbf{A^t A}\right)\mathbf{x = x^t A^t Ax} = (\mathbf{Ax})^{\mathbf{t}}(\mathbf{Ax})$ $= \mathbf{0}$ y en consecuencia $\mathbf{Ax = 0}$; es decir $\mathbf{x} \in \mathbf{Null(A)}$

Corolario 2.3.1

La matriz $\mathbf{A^t A}$ de orden $\mathbf{n}$ es invertible

Demostración

Por tener la matriz $\mathbf{A^t A}$ rango $\mathbf{n}$; está matriz es básica y por tanto existe $\left(\mathbf{A^t A}\right)^{\mathbf{-1}}$

Corolario 2.3.2

La matriz $\mathbf{A^t A}$ de orden $\mathbf{n}$ es simétrica

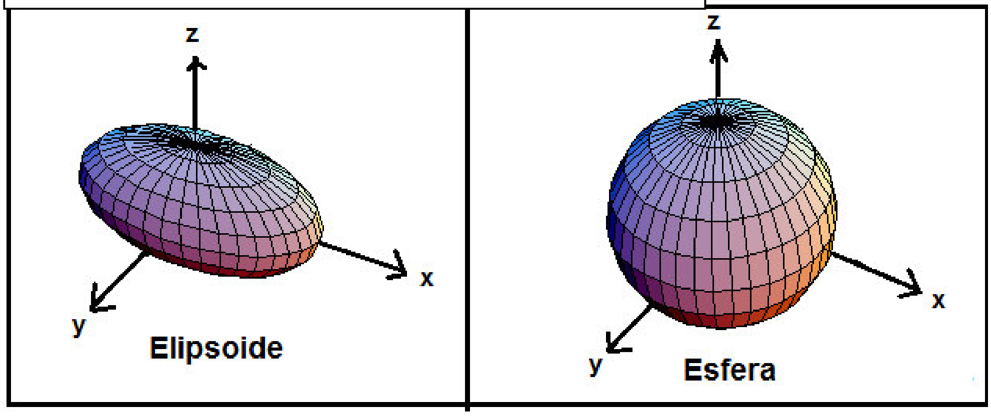
Demostración

La matriz $\mathbf{A^t A}$ será simétrica si es igual a su transpuesta

En efecto la transpuesta de $\mathbf{A^t A}$ es $\left(\mathbf{A^t A}\right)^{\mathbf{t}} \left(\mathbf{A^t}\right)^{\mathbf{t}}\left(\mathbf{A^t}\right) = \mathbf{A^t A}$

2.3.8 Elipsoide

Gráfica n ° 9: Transformación de esfera en elipse



Fuente propia

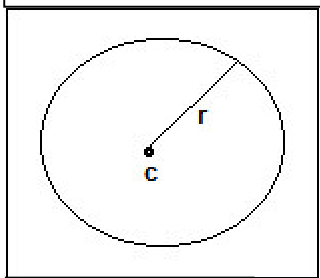
2.3.9 Bola cerrada

Definición Una bola cerrada $\mathbf{B(c,r)}$, de centro $\mathbf{c} \in \mathbb{R}^n$ **y** radio $\mathbf{r}$, es el conjunto:

$$\mathbf{B(c,r)} = \left\{ \mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \mathbf{c})^t (\mathbf{x} - \mathbf{c}) \le \mathbf{r}^2 \right\}$$

Ejemplo 2.3.6 De una bola en $\mathbb{R}^2$

Gráfica n ° 9: Circunferencia de centro c y radio r = 1



Fuente propia

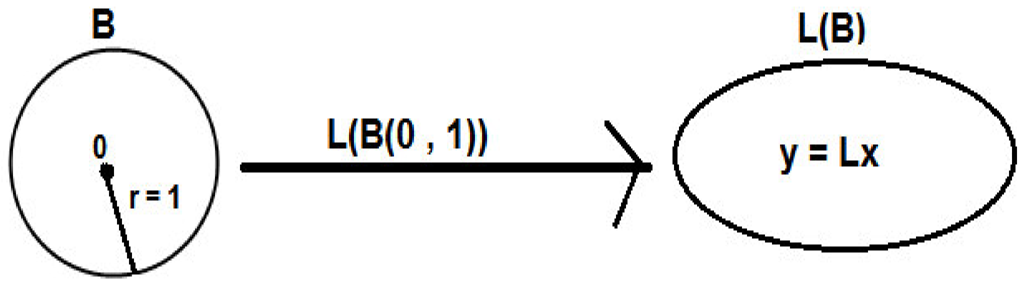
Una Bola de centro $\mathbf{c = 0}$ y radio $\mathbf{r = 1}$ es $\mathbf{B(0,r)} = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x}^t\mathbf{x} \le \mathbf{r}^2 \right\}$

2.3.10 Transformación esfera en elipsoide

Definición-1. Un elipsoide es la imagen de una bola unitaria, bajo una transformación afín lineal invertible, es decir un elipsoide $\mathbf{E}$, centrada en el origen, es la imagen $\mathbf{L(B(0,1))}$, de la bola unitaria bajo una transformación lineal invertible $\mathbf{L}:\mathbb{R}^{\mathbf{n}} \to \mathbb{R}^{\mathbf{n}}$.

Se puede escribir la definición anterior de forma más explícita como

Gráfica nº 10: Transformación de bola unitaria en elipse



Fuente propia

$$
\begin{aligned}
\mathbf{L(B(0,1))} = \; & \left\{ \mathbf{y = Lx : x \in B(0,1)} \right\} \\
= \; & \left\{ \mathbf{y : L^{-1}y \in B(0,1)} \right\} \\
= \; & \left\{ \mathbf{y : \left\| 0 - L^{-1}y \right\| \leq 1} \right\} \\
= \; & \left\{ \mathbf{y : \left\| L^{-1}y \right\| \leq 1} \right\} \\
= \; & \left\{ \mathbf{y : \left\| L^{-1}y \right\|^{2} \leq 1^{2}} \right\} \\
= \; & \left\{ \mathbf{y : \left(L^{-1}y\right)^{t} \left(L^{-1}y\right) \leq 1} \right\} \\
= \; & \left\{ \mathbf{y : y^{t}L^{-t}L^{-1}y \leq 1} \right\} \\
= \; & \left\{ \mathbf{y : y^{t} \left(L L^{t}\right)^{-1} y \leq 1} \right\}; \; \mathbf{Q = L L^{T}} \\
\mathbf{L(B(0,1))} = \; & \left\{ \mathbf{y : y^{t}Q^{-1}y \leq 1} \right\}
\end{aligned}
$$

Definición – 2. En general, un elipsoide con centro en cualquier punto $\mathbf{c} \in \mathbb{R}^{\mathbf{n}}$ es sólo la traslación $\mathbf{c + E}$, de algún elipsoide $\mathbf{E}$, con centro en $\mathbf{0}$

¿Qué se puede decir acerca de la matriz $\mathbf{Q = L L^{T}}$ asociada a una elipse $\mathbf{E}$?

Se responderán estas interrogantes, mediante las propiedades siguientes

2.3.11 Propiedades de la matriz Q

Propiedad - 1

Para una matriz simétrica $\mathbf{Q} \in \mathbb{R}^{n\,x\,n}$, las siguientes condiciones son equivalentes

i. $\mathbf{Q} = \mathbf{L}\mathbf{L}^T$ para algún $\mathbf{L} \in \mathbb{R}^{n\,x\,n}$

ii. Todos los auto valores de $\mathbf{Q}$ son no negativos

Se dirá que $\mathbf{Q}$ es **semidefinida Positiva**, si y sólo si se verifican cualquiera de las dos condiciones anteriores

Demostración

(1 ⇒ 2)

Considérese que $\mathbf{Q} = \mathbf{L}\mathbf{L}^T$ para algún $\mathbf{L} \in \mathbb{R}^{n\,x\,n}$

Supóngase que λ es un auto valor de $\mathbf{Q}$, con auto vector $\mathbf{x} \neq \mathbf{0}$; esto es $\mathbf{Qx} = \lambda\mathbf{x}$. Entonces se tiene:

$$\lambda\|\mathbf{x}\|^2 = \lambda\mathbf{x}^T\mathbf{x} = \mathbf{x}^T(\lambda\mathbf{x}^T) = \mathbf{x}^T(Q\mathbf{x}) = x^T\mathbf{L}\mathbf{L}^T x = \left(\mathbf{L}^T x\right)^T\left(\mathbf{L}^T x\right) =$$
$$\left\|\mathbf{L}^T x\right\|^2 \geq \mathbf{0}$$

Relación que demuestra que $\lambda \in \mathbb{R}$ y $\lambda \geq \mathbf{0}$ ■

(2 ⇒ 1)

Si $\mathbf{Q}$ es simétrica, entonces todos sus auto valores son reales y por hipótesis $\lambda_i \geq \mathbf{0}$, para i = 1, 2, . . . , n

Se cumple que para auto valores distintos, sus auto vectores son ortogonales

Si λ es un auto valor de multiplicidad $\mathbf{k}$, entonces existen $\mathbf{k}$ auto vectores asociados a λ, que serán **l. i.**

Como **Q** posee **n** auto valores; entonces existen **n** auto vectores linealmente independientes, con los cuales se puede formar una base **ortonormal** para $\mathbb{R}^{\mathbf{n}}$

Luego sea **Q** la matriz cuyos auto valores $\lambda_1, \lambda_2, \ldots, \lambda_n$ y sus respectivos auto vectores asociados $v_1, v_2, \ldots, v_n$ orto normales, luego se puede obtener una matriz diagonal **D**

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

Y la matriz **P**

$$\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & . & . & \mathbf{v}_n \end{bmatrix}$$

También

$$\mathbf{D}^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \sqrt{\lambda_{n-1}} & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\lambda_n} \end{pmatrix}$$

Debido a lo cual se tiene $\mathbf{Q\,P = P\,D}$ y por ser $\mathbf{P^t = P^{-1}}$ y $\mathbf{D^{1/2}\,D^{1/2} = D}$

Entonces

$$\mathbf{Q = P\,D\,P^t}$$

$$\mathbf{Q = P\,D^{1/2}\,D^{1/2}\,P^t}$$

Tomando $\mathbf{L = P\,D^{1/2}}$ y $\mathbf{L^t = D^{1/2}\,P^t}$, se obtendrá que $\mathbf{Q = L\,L^t}$ ■

Propiedad - 2

Para cualquier matriz simétrica $\mathbf{Q} \in \mathbb{R}^{\mathbf{n\,x\,n}}$, las siguientes condiciones son equivalentes

i. $\mathbf{Q = L\,L^T}$ para alguna matriz $\mathbf{L}$ **No singular**

ii. Todo auto valor de $\mathbf{Q}$ es **Estrictamente postivos**

Demostración

(1 ⇒ 2)

Por ser $\mathbf{Q = L\,L^T}$**,** es simétrica y por tanto diagonalizable, esto es existe la matriz $\mathbf{P}$**,** tal que $\mathbf{P^t Q\,P = D}$

Y como $\mathbf{Q}$ es no singular, luego todo auto valores λ_i, los que son elementos de la matriz $\mathbf{D}$, deben ser reales y distintos de cero; es decir $\lambda_i \neq \mathbf{0}$ para i = 1, 2, . . . , n

Luego utilizando la condición de que $\lambda_i \geq \mathbf{0}$, se concluye que $\lambda_i > \mathbf{0}$, para i = 1, 2, . . . , n

(2 ⇒ 1)

Si $\mathbf{Q}$ posee todos sus auto valores positivos y por ser $\mathbf{Q}$ una matriz simétrica; se verifica que $\mathbf{Q}$ es diagonalizable; esto es $\mathbf{P^t Q\,P = D}$. Entonces $\mathbf{D}$ posee inversa, lo que implica que $\mathbf{Q}$ también es invertible; es decir $\mathbf{Q}$ es no singular.

Se dice que la matriz $\mathbf{Q}$ es definida positiva; si y sólo si las condiciones anteriores se mantiene

De lo expuesto anteriormente; es claro que un elipsoide puede ser representado equivalentemente, en términos de una matriz Q definida positiva.

Ahora se puede dar otra definición equivalente a la definición 2.4

Definición. Si $\mathbf{Q} \in \mathbb{R}^{\mathbf{n\,x\,n}}$ es una matriz definida positiva. Entonces el elipsoide, asociado con $\mathbf{Q}$ de centro $\mathbf{c} \in \mathbb{R}^{\mathbf{n}}$ es

$$\mathbf{E(c\,,Q)} = \left\{\mathbf{c+y : y^T Q^{-1} y \leq 1}\right\} = \left\{\mathbf{y : (y - c)^T Q^{-1} (y - c) \leq 1}\right\}$$

Observación

i. La bola estándar $\mathbf{B(0\,,r)}$, es el elipsoide $\mathbf{E(\,0\,,r^{\,2}\,I\,)}$.

Más generalmente un **elipsoide axial**

En efecto

$$\mathbf{E}\left(0,\begin{pmatrix} \mathbf{r^2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{r^2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ . & . & . & . & . \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{r^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{r^2} \end{pmatrix}\right)$$

Donde

$$\mathbf{Q = Ir^2 =} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{r^2}$$

$$\mathbf{Q = Ir^2 =} \begin{pmatrix} r^2 & 0 & 0 & 0 & 0 \\ 0 & r^2 & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & 0 & r^2 \end{pmatrix} = \begin{pmatrix} r & 0 & 0 & 0 & 0 \\ 0 & r & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & r & 0 \\ 0 & 0 & 0 & 0 & r \end{pmatrix}\begin{pmatrix} r & 0 & 0 & 0 & 0 \\ 0 & r & 0 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & r & 0 \\ 0 & 0 & 0 & 0 & r \end{pmatrix}$$

$$\begin{aligned}
\mathbf{E(0\,,1)) =} \quad & \left\{(0 + y) : y^tQ^{-1}y \leq 1\right\} \\
= \quad & \left\{y : y^t\left(r^2 I\right)^{-1} y \leq 1\right\} \\
= \quad & \left\{y : \frac{1}{r^2}\left(y^t\, I\, y\right) \leq 1\right\} \\
= \quad & \left\{y : y^t\, y \leq r^2\right\} \\
= \quad & \left\{y : \|y\|^2 \leq r^2\right\} \\
\mathbf{E(0\,,1) =} \quad & \left\{y : \|y\| \leq r\right\} = \mathbf{B(0\,,1)}
\end{aligned}$$

ii. Será necesario recordar que el volumen de la imagen de un conjunto $\mathbf{A} \subseteq \mathbb{R}^{\mathbf{n}}$ es

$$\mathbf{Vol}(\mathbf{L}(\mathbf{A})) = |\det L|\ \mathbf{Vol}(\mathbf{A})$$

Siendo $\mathbf{L}(\mathbf{A})$, la imagen del conjunto $\mathbf{A}$, a través de la transformación lineal $\mathbf{L}$

Con este resultado se deduce que el volumen de un elipsoide $\mathbf{E(c, Q)}$, estará dado por la relación

$$\mathbf{vol}(\mathrm{E(c, Q)}) = |\det L|\ \mathrm{vol}(\mathbf{B(0, r)}) = \sqrt{|\det(Q)|}\ \mathrm{vol}(\mathbf{B(0, r)})$$

De esta manera se ha relacionado el volumen de cualquier elipsoide; con el volumen de una bola unitaria de dimensión $\mathbf{n}$

2.3.12 Descomposición de valor singular(Burden, J. Douglas Faires, 2002)

Toda matriz simétrica $\mathbf{A}$ puede factorizarse como $\mathbf{A = PDP^t}$, donde $\mathbf{P}$ es una matriz ortogonal y $\mathbf{D}$ es una matriz diagonal, que muestra los eigenvalores de la matriz $\mathbf{A}$

Si $\mathbf{A}$ no es simétrica es posible factorizar una matriz cuadrada como $\mathbf{A = PDP^{-1}}$, donde $\mathbf{D}$ es como antes; pero $\mathbf{P}$ es simplemente una matriz no singular

En general toda matriz $\mathbf{A}$ tiene una factorización de la forma $\mathbf{A} = \mathbf{PD}Q^{\mathbf{t}}$, siendo $\mathbf{P}$ y $Q^{\mathbf{t}}$ ortogonales y $\mathbf{D}$ es una matriz diagonal

2.3.13 Valores singulares de una matriz

Para cualquier matriz $\mathbf{A}$ de $\mathbb{R}^{\mathbf{m \times n}}$, la matriz $\mathbf{A^tA}$ es:

i. De orden $\mathbf{n}$

ii. Es simétrica y por ende puede diagonalizarse ortogonalmente, siendo sus eigenvalores reales no negativos

Los valores de λ son no negativos

En efecto λ un autovalor de $\mathbf{A^tA}$ y $\mathbf{v}$ su correspondiente autovector unitario, luego

$$\mathbf{0} \leq \|\mathbf{Av}\|^{\mathbf{2}} = (\mathrm{A}\,\mathrm{v})(\mathrm{A}\,\mathrm{v}) = (\mathrm{A}\,\mathrm{v})^{\mathbf{t}}(\mathrm{A}\,\mathrm{v}) = \mathbf{v}^{\,\mathbf{t}}\mathbf{A}^{\,\mathbf{t}}\mathbf{A}\,\mathbf{v} = \mathbf{v}^{\,\mathbf{t}}\lambda\,\mathbf{v} = \lambda\left(\mathbf{v}^{\,\mathbf{t}}.\mathbf{v}\right) = \lambda\|\mathbf{v}\|^{\mathbf{2}}$$

$= \lambda$, ya que $\|\mathbf{v}\|^{\mathbf{2}} = \mathbf{1}$

Concluyendo que $\mathbf{0} \leq \|\mathbf{Av}\|^{\mathbf{2}} = \lambda$, y diciendo que es posible obtener las raíces cuadradas de los eigenvalores

Definición Si A es una matriz de $\mathbb{R}^{\mathbf{mxn}}$, sus valores singulares, son las raíces cuadradas de los eigenvalores de $\mathbf{A^tA}$ y se denotan mediante $\sigma_1, \sigma_2, \ldots, \sigma_n$ Se conviene en ordenarlos de mayor a menor, esto es $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq, \ldots, \geq \sigma_n$

Ejemplo 2.3.6

Halle los valores singulares de

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \\ 3 & 0 \end{bmatrix}$$

Resolución

$$\mathbf{A^tA} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

Tiene eigenvalores $\lambda_1 = 9$ y $\lambda_2 = 4$. Por tanto **los valores singulares de A son** $\sigma_1 = 3$, y $\sigma_2 = 2$

2.3.14 Interpretación geométrica de los valores singulares

¿Qué son geométricamente, los valores singulares de una matriz $\mathbf{A} \in \mathbb{R}^{\mathbf{mxn}}$?

Para dar respuesta a esta pregunta: Consideremos los eigenvectores de la matriz $\mathbf{A^tA}$, puesto que $\mathbf{A^tA}$ es simétrica, existe una base orto normal para $\mathbb{R}^{\mathbf{n}}$, que consiste de los eigenvectores de $\mathbf{A^tA}$. Sea

$\left[v_1, \; v_2, \; . \; . \; v_n\right]$ una base de $\mathbf{A^tA}$, tal que corresponda a los eigenvalores de $\mathbf{A^tA}$, ordenados de forma que $\lambda_1 \geq \lambda_2 \geq \; \lambda_3 \geq, \; . \; . \; ,$ $\lambda_n \geq \mathbf{0}, .$

A partir de los cálculos hechos anteriormente, se tiene: De $\lambda_i = \|\mathbf{Av_i}\|^2$, esto es:

$$\sigma_i = \sqrt{\lambda_i} = \|Av_i\|$$

Ejemplo 2.3.6

Halle los valores singulares de la matriz $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

Resolución

$$A^tA = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} \end{bmatrix}$$

Tiene eigenvalores $\lambda_1 = 3$ y $\lambda_2 = 1$. Por tanto los valores singulares de A son:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{\mathbf{3}} \text{ y } \sigma_2 = \sqrt{\lambda_2} = \sqrt{\mathbf{1}} = \mathbf{1}$$

Los eigenvectores asociados de $\mathbf{A^tA}$ son

Para $\lambda_1 = 3$: $\begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = \mathbf{3} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{-1} & \mathbf{1} \\ \mathbf{1} & \mathbf{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$

Su eigenvector es $\begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$

Para $\lambda_2 = \mathbf{1}$: $\begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = \mathbf{1} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$

Su eigenvector es $\begin{bmatrix} \mathbf{-1} \\ \mathbf{1} \end{bmatrix}$

Se puede observar que los eigenvectores son ortogonales

Normalizándolos se tiene

Para $\lambda_1 = 3$: Su eigenvector ortonormal correspondiente $\mathbf{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Para $\lambda_2 = 1$: Su eigenvector ortonormal correspondiente $\mathbf{v_2} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Por tanto una base ortonormal de $\mathbb{R}^2$, para $\mathbf{A^tA}$ es $\begin{bmatrix} \mathbf{v_1}, \mathbf{v_2} \end{bmatrix}$ =

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Siendo $\mathbf{Av_1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{2}, & 1/\sqrt{2}, & 1/\sqrt{2} \end{bmatrix} \Rightarrow$

$\|\mathbf{Av_1}\| = \sqrt{3} = \sqrt{\lambda_1} = \sigma_1$

También $\mathbf{Av_2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0, & -1/\sqrt{2}, & 1/\sqrt{2} \end{bmatrix} \|\mathbf{Av_2}\| = \sqrt{1} = \sqrt{\lambda_2} = \sigma_2 = 1$

Es decir los valores singulares de la matriz $\mathbf{A}$:

La longitud del vector $\mathbf{Av_1}$: $\|\mathbf{Av_1}\| = \sqrt{3} = \sqrt{\lambda_1} = \sigma_1$ y

La longitud del vector $\mathbf{Av_2}$: $\|\mathbf{Av_2}\| = \sqrt{1} = \sqrt{\lambda_2} = \sigma_2 = 1$

En el presente ejemplo, este resultado puede tener la siguiente interpretación:

Si $\mathbf{x}$ se encuentra sobre la circunferencia unitaria en $\mathbb{R}^2$. Esto es $\|\mathbf{x}\| = 1$, luego

$$\|\mathbf{Ax}\|^2 = (\mathbf{Ax}).(\mathbf{Ax}) = (\mathbf{Ax})^t.(\mathbf{Ax}) = \mathbf{x^tA^tAx} =$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{2}x_1^2 + \mathbf{2}x_1x_2 + \mathbf{2}x_2^2$$

Relación que se reconoce como forma cuadrática

Los valores máximo y mínimo de esta forma cuadrática, sujeta a la condición $\| \mathbf{x} \| = \mathbf{1}$, son $\lambda_1 = 3$ y $\lambda_2 = 1$ respectivamente y se presentan en los correspondientes eigenvectores de $\mathbf{A^t A}$, esto es cuando **x = v1 =** $\begin{bmatrix} \mathbf{1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} \end{bmatrix}$ y **x = v2 =** $\begin{bmatrix} \mathbf{-1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} \end{bmatrix}$ respectivamente

Ya que $\|\mathbf{Av_i}\|^2 = \mathbf{v^t A^t A v} = \lambda_i$

Para i = 1: se tiene que $\boldsymbol{\sigma}_1 = \|\mathbf{Av_1}\| = \sqrt{3}$ y para $\boldsymbol{\sigma}_2 = \|\mathbf{Av_2}\| = 1$, son los valores máximo y mínimo de las longitudes $\| \mathbf{Ax} \|$, conforme $\mathbf{x}$ recorre la circunferencia de radio uno en $\mathbb{R}^2$

La transformación lineal correspondiente a la matriz $\mathbf{A}$, mapea $\mathbb{R}^2$ sobre el plano en $\mathbb{R}^3$, esto es

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = [\mathbf{x_1 + x_2}, \quad \mathbf{x_1}, \quad \mathbf{x_2}]$$

Con ecuación $\mathbf{x_1 - x_2 - x_3 = 0}$

De modo que $\boldsymbol{\sigma}_1$ y $\boldsymbol{\sigma}_2$ son las longitudes de la mitad de los ejes mayor y menor de esta elipse

2.3.15 Descomposición de valor singular de la matriz A

Teorema

Sea $\mathbf{A}$ una matriz de $\mathbf{m}$ filas y $\mathbf{n}$ columnas; con valores singulares $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq, \ldots, \sigma_r > \mathbf{0}$, y $\sigma_{r+1} = \mathbf{0}$, $\sigma_{r+2} = \mathbf{0}$, . . . , $\sigma_n = \mathbf{0}$. Entonces existe una matriz ortogonal $\mathbf{U}$ de orden $\mathbf{m}$, una matriz ortogonal $\mathbf{V}$ de orden $\mathbf{n}$ y una matriz $\sum$ de $\mathbf{m}$ filas y $\mathbf{n}$ columnas; de la forma:

$$\begin{array}{c} \\ \Sigma = \end{array}\begin{array}{cc} \mathbf{r} & \mathbf{n-r} \\ \left[\begin{matrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{matrix}\right] & \begin{matrix} r \\ \mathbf{m-r} \end{matrix} \end{array} \text{, con } \mathbf{D} = \begin{pmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ . & . & . & . & . \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_r \end{pmatrix} \quad \textbf{(1)}$$

Tal que

$$\mathbf{A = U\Sigma V^t}$$

Las columnas de **U** se llaman vectores singulares izquierdos de **A** y las columnas de **V** se llaman vectores singulares derechos de **A.**

Las matrices **U** y **V** no están determinadas exclusivamente por la matriz **A;** sino que Σ debe contener los valores singulares de **A,** como en la ecuación (1)

Demostración

Se requiere demostrar que una matriz **A**, de **m** filas y **n** columnas, puede factorizarse como

$$A = U\Sigma\, V^t$$

Siendo **U** una matriz ortogonal de orden **m**, **V** es una matriz de orden **n**

Si los valores singulares de la matriz **A**, diferentes de cero son $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq, \ldots, \sigma_r > \mathbf{0}$,

Y $\sigma_{r+1} = \mathbf{0}$, $\sigma_{r+2} = \mathbf{0}$, . . . , $\sigma_n = \mathbf{0}$, Entonces Σ , tendrá la forma

$$\begin{array}{c} \\ \Sigma = \end{array}\begin{array}{cc} \mathbf{r} & \mathbf{n-r} \\ \left[\begin{matrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{matrix}\right] & \begin{matrix} r \\ \mathbf{m-r} \end{matrix} \end{array} \quad ..\textbf{(1)}$$

Siendo

$$\mathbf{D} = \begin{pmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ . & . & . & . & . \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_r \end{pmatrix}$$

¿En el presente caso cuál es **D**?

Para construir la matriz ortogonal $\mathbf{V}$, en primer lugar se debe hallar una base de ortonormal de $\mathbb{R}^{\mathbf{n}}$: $\left[v_1, \ v_2, \ \ . \ \ . \ \ v_n \right]$, el que consiste de eigenvectores de la matriz simétrica de orden $\mathbf{n}$: $\mathbf{A^tA}$

Luego

$$\mathbf{V} = \left[v_1, \ v_2, \ \ . \ \ . \ \ v_n \right]$$

Es una matriz ortogonal de orden $\mathbf{n}$

Para la matriz ortogonal $\mathbf{U}$, nótese que $\left[Av_1, \ Av_2, \ \ . \ \ . \ \ Av_n \right]$, es un conjunto ortogonal de vectores de $\mathbb{R}^{\mathbf{m}}$. Veamos esto:

Supóngase que v_i es el eigenvector de $\mathbf{A^tA}$, correspondiente al eigenvalor λ_i, luego para $\mathbf{i \neq j}$ se tiene que

$$\left(\mathbf{A}v_{\mathbf{i}}\right)^{\mathbf{t}}\left(\mathbf{A}v_j\right) = = \mathbf{v}_{\mathbf{i}}^{\mathbf{t}}\mathbf{A}^{\mathbf{t}}\mathbf{A}\mathbf{v}_j = \mathbf{v}_{\mathbf{i}}^{\mathbf{t}}\lambda_{\mathbf{j}}\mathbf{v}_j = \lambda_j\left(\mathbf{v}_{\mathbf{i}}^{\mathbf{t}}.\mathbf{v}_j\right) = \mathbf{0}$$

Dado que los eigenvectores $\mathbf{v}_{\mathbf{i}}$ son ortogonales

Téngase en cuenta que los valores singulares, satisfacen la relación $\boldsymbol{\sigma}_{\mathbf{i}} = \left\|\mathbf{Av}_i\right\|$ y que los primeros $\mathbf{r}$, son diferentes de cero. Por tanto se puede normalizar $\mathbf{Av_1}$, $\mathbf{Av_2}$, . . . , $\mathbf{Av_r}$, al establecer

$$\mathbf{u_i} = \frac{\mathbf{1}}{\boldsymbol{\sigma}_{\mathbf{i}}}\mathbf{Av_i} \ \text{ para i = 1, 2, . . . ,} \ \mathbf{r}$$

Esto garantiza que $\left[\mathbf{u_1}, \ \mathbf{u_2}, \ \ . \ \ . \ \ \mathbf{u}_r \right]$ es un conjunto ortonormal de vectores de $\mathbb{R}^{\mathbf{m}}$.

Pero si ocurre que $\mathbf{r} < \mathbf{m}$; no será una base para $\mathbb{R}^{\mathbf{m}}$

En este caso se extiende el conjunto $\left[\mathbf{u_1}, \ \mathbf{u_2}, \ \ . \ \ . \ \ \mathbf{u}_r \right]$ a una base de ortonormal $\left[\mathbf{u_1}, \ \mathbf{u_2}, \ \ . \ \ . \ \ \mathbf{u_m} \right]$ de $\mathbb{R}^{\mathbf{m}}$ (utilizando el proceso de Gram Schmidt)

Entonces se establece

$$\mathbf{U} = \begin{bmatrix} \mathbf{u_1}, & \mathbf{u_2}, & . & . & \mathbf{u_m} \end{bmatrix}$$

Lo que resta por demostrar es que esto funciona; es decir se necesita verificar que con $\mathbf{U}$, $\sum$ y $\mathbf{V}$, como se han descrito, se tiene que $\mathbf{A = U \sum V^t}$, dado que $\mathbf{V}^t = \mathbf{V}^{-1}$, esto equivale a demostrar que

$$\mathbf{AV = U\sum}$$

Se sabe que $\mathbf{Av_i} = \boldsymbol{\sigma_i} \mathbf{u_i}$ para i = 1, 2, . . . , **r**

Y $\|\mathbf{Av}_i\| = \boldsymbol{\sigma_i} = \mathbf{0}$ para **i = r+1, . . . , n,** por tanto

$$\mathbf{Av_i = 0 \text{ para } i = r+1 , . . , n}$$

Por tanto $\mathbf{AV = A} \begin{bmatrix} v_1, & v_2, & . & . & v_n \end{bmatrix}$

$$= \begin{bmatrix} Av_1, & Av_2, & . & . & Av_n \end{bmatrix}$$

$$= \begin{bmatrix} Av_1, & Av_2, & Av_r & \mathbf{0} \dots & \mathbf{0} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 u_1, & \sigma_2 u_2, \dots & \sigma_r u_r & \mathbf{0} \dots & \mathbf{0} \end{bmatrix}$$

$$= \begin{bmatrix} u_1, & u_2 & . & . & u_m \end{bmatrix}$$

$$\begin{bmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & | & \mathbf{0} \\ \mathbf{0} & \sigma_2 & \mathbf{0} & \mathbf{0} & | & \mathbf{0} \\ . & . & . & . & | & . \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_r & | & \mathbf{0} \\ \_ & \_ & \_ & \_ & | & \_ \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & | & \mathbf{0} \end{bmatrix}$$

$$= \mathbf{U\sum}$$

Como se quería

2.3.16 Factorización de una matriz A

Definición. Una factorización de la matriz $\mathbf{A}$, realizada como en el, Teorema 2.6.3.1, se llama descomposición de valor singular (DVS), de $\mathbf{A}$.

Las columnas de $\mathbf{U}$, se llaman vectores singulares izquierdos de $\mathbf{A}$ y las columnas de $\mathbf{V}$, se llaman vectores singulares derechos de $\mathbf{A}$. Las matrices $\mathbf{U}$ y $\mathbf{V}$ no están determinadas exclusivamente por $\mathbf{A}$, sino que $\sum$ debe contener los valores singulares de $\mathbf{A}$, como en la ecuación (1)

Ejemplo 2.3.7

Halle la descomposición de valor singular $\mathbf{A = U\sum V^t}$ para la matriz

$$A = \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

Resolución

$$A^tA = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} \end{bmatrix}$$

Tiene eigenvalores $\lambda_1 = 3$ y $\lambda_2 = 1$. Por tanto los eigenvectores de A son

Para $\lambda_1 = 3$: $\begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$ y para $\lambda_2 = \mathbf{1}$: $\begin{bmatrix} \mathbf{-1} \\ \mathbf{1} \end{bmatrix}$

Al normalizarlos se obtiene $\mathbf{v_1} = \begin{bmatrix} \mathbf{1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} \end{bmatrix}$ y $\mathbf{v_2} = \begin{bmatrix} \mathbf{-1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} \end{bmatrix}$ Observe que son ortogonales

Siendo $\mathbf{V} = \begin{bmatrix} \mathbf{1/\sqrt{2}} & \mathbf{-1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} & \mathbf{1/\sqrt{2}} \end{bmatrix}$ una base ortonormal, también

$$\mathbf{V^t} = \begin{bmatrix} \mathbf{1/\sqrt{2}} & \mathbf{1/\sqrt{2}} \\ \mathbf{-1/\sqrt{2}} & \mathbf{1/\sqrt{2}} \end{bmatrix}$$

Siendo los valores singulares de **A**

$$Av_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{2} \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \Rightarrow \sigma_1 = \|Av_1\| = \sqrt{3}$$

$$Av_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \Rightarrow \sigma_2 = \|Av_2\| = 1$$

Por tanto $\sum = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

Finalmente se hallará U

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 2/\sqrt{2} \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \quad \text{y} \quad u_2 = \frac{1}{\sigma_2} Av_2 =$$

$$\frac{1}{\sqrt{1}} \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Se necesita extender $\begin{bmatrix} u_1 & u_2 \end{bmatrix}$ a una base orto normal de $\mathbb{R}^3$, para lo cual al utilizar el método de Gram Schmidt, se hará agregando $e_3 = [0, \ 0, \ 1]^t$, para obtener

$$u_3 = \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \ \begin{bmatrix} u_1 & u_2, & e_3 \end{bmatrix} \text{ estos forman una base de } \mathbb{R}^3$$

De este modo $U = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$

Siendo la DVS

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

2.3.17 La forma producto externo de la DVS

Teorema

Sea **A** una matriz de **m** filas y **n** columnas; con valores singulares $\sigma_1 \geq \sigma_2 \geq \quad \sigma_3 \geq, \ . \ . \ . \ , \ \sigma_r > \mathbf{0}$, y $\sigma_{r+1} = \mathbf{0}$, $\sigma_{r+2} = \mathbf{0}$, . . . , $\sigma_n = \mathbf{0}$. Sean $u_1, \ u_2 \quad . \quad . \quad u_r$ vectores singulares izquierdos y sean $v_1, \ v_2 \quad . \quad . \quad v_r$ vectores singulares derechos de la matriz **A,** que corresponden a dichos valores singulares, entonces:

$$\mathbf{A} = \sigma_1 u_1 v_1^\tau + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\tau + \ .. \ . \ \sigma_r \mathbf{u}_r \mathbf{v}_r^\tau$$

Observación

Si **A** una matriz simétrica positiva definida; estos dos últimos teoremas, se reducen a resultados harto conocidos.

La DVS de una matriz **A,** contiene mucha información acerca de ella**,** como lo destaca el siguiente teorema.

2.3.18 Valores singulares no nulos

Teorema

Sea $\mathbf{A} = \mathbf{U}\sum \mathbf{V}^t$, una descomposición de valor singular **(DVS)** de una matriz **A** de **m** filas y **n** columna, sean también $\sigma_1, \ \sigma_3 \ , \ . \ . \ , \ \sigma_r$ todos los valores singulares distintos de cero de la matriz **A.** Entonces, se verifican las siguientes relaciones:

i. $\left[u_1, \quad u_2 \quad . \quad . \quad u_r \right]$ es una base ortonormal, para el **col(A)**

ii. $\left[u_1, \quad u_2 \quad . \quad . \quad u_m \right]$ es una base ortonormal, para el **null(A^t)**

iii. $\left[v_1, \quad v_2 \quad . \quad . \quad v_r \right]$ es una base ortonormal, para el **fila(A)**

iv. $\left[v_1, \quad v_2 \quad . \quad . \quad v_n\right]$ es una base ortonormal, para el **null(A)**

Demostración

i. $\mathbf{rango(A) = rango(U \sum V^t)}$

$$= \mathbf{rango}(\textstyle\sum \mathbf{V}^t)$$

$$= \mathbf{rango}(\textstyle\sum\ ) = \mathbf{r}$$

ii. Se sabe que $\left[u_1, \quad u_2 \quad . \quad . \quad u_r\right]$ es un conjunto ortonormal, por lo que es **l. i.**

Debido a que $\mathbf{u_i} = \dfrac{\mathbf{1}}{\boldsymbol{\sigma}_i}\mathbf{Av_i}$, para i = 1, 2, . . . , **r**, cada $\mathbf{u_i}$ está en el espacio columna de **A y** también

$$\mathbf{r = rango(A) = dim(col(A))}$$

Por lo que $\left[u_1, \quad u_2 \quad . \quad . \quad u_r\right]$, es una base ortonormal para **col(A)**

iii. Dado que $\left[u_1, \quad u_2 \quad . \quad . \quad u_m\right]$ es una base ortonormal para $\mathbb{R}^m$ y $\left[u_1, \quad u_2 \quad . \quad . \quad u_r\right]$ es una base para **col(A),** por (**ii.**) se tiene que $\left[u_{r+1}, \quad u_{r+2} \quad . \quad . \quad u_m\right]$, es una base ortonormal, para el complemento ortogonal de **col(A)**. Pero $\mathbf{(col(A))^{\perp} = null(A^t)}$

iv. Dado que

$$\mathbf{Av_{r+1} = Av_{r+2} = \ .\ .\ = Av_n = 0}$$

El conjunto $\left[v_{r+1}, \quad v_{r+2} \quad . \quad . \quad v_n\right]$ es un conjunto ortonormal, contenido en el espacio nulo de **A**

Por tanto $\left[v_{r+1}, \quad v_{r+2} \quad . \quad . \quad v_n\right]$ es un conjunto de vectores **l.i.**, con $\mathbf{n-r}$ vectores en **null(A)**. Pero $\mathbf{dim(null(A)) = n-r}$, (por el teorema del rango), del modo que

$\left[v_{r+1}, \quad v_{r+2} \quad . \quad . \quad v_n\right]$, es una base ortonormal, para **null(A)**

v. La propiedad, se obtiene a partir de la **(iv)**

La DVS ofrece una comprensión geométrica acerca del efecto de las transformaciones matriciales. Así por ejemplo se tiene que una matriz de m filas y n columnas, transforma la esfera unitaria de $\mathbb{R}^n$ en una elipsoide de $\mathbb{R}^m$

2.3.19 Imagen de la esfera unitaria A

Teorema

Sea $\mathbf{A}$ una matriz de $\mathbb{R}^{mxn}$, con $\mathbf{rango(A) = r}$. Entonces la imagen de la esfera unitaria en $\mathbb{R}^n$, bajo la transformación matricial que mapea $\mathbf{x}$ a $\mathbf{Ax}$ es:

i. La superficie de un elipsoide en $\mathbb{R}^m$, si $\mathbf{r = n}$

ii. Un elipsoide sólido en $\mathbb{R}^m$, si $\mathbf{r < n}$

Demostración

Sea $\mathbf{A = U\sum V^t}$ una descomposición de valor singular de la matriz $\mathbf{A}$ una matriz de $\mathbb{R}^{mxn}$, sean los vectores singulares izquierdos

$u_1, \quad u_2 \quad . \quad . \quad u_m$ y derechos

$v_1, \quad v_2 \quad . \quad . \quad v_n$ respectivamente, puesto que $\mathbf{rango(A) = r}$, los valores singulares de la matriz $\mathbf{A}$, Por el **Teorema 2.6.3**; satisfacen

$\sigma_1 \geq \sigma_2 \geq \quad \sigma_3 \geq, \ldots, \sigma_r > \mathbf{0}$, y $\sigma_{r+1} = \mathbf{0}$, $\sigma_{r+2} = \mathbf{0}, \ldots, \sigma_n = \mathbf{0}$.

Sea $\mathbf{x} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ . \\ \mathbf{x_n} \end{bmatrix}$ un vector unitario de $\mathbb{R}^{\mathbf{n}}$, puesto qu**e** **V** es una matriz ortogonal, también lo es $\mathbf{V}^{\mathbf{t}}$ y en consecuencia $\mathbf{V}^{\mathbf{t}}\mathbf{x}$ será un vector unitario

Ahora

$$\mathbf{V}^{\mathbf{t}}\mathbf{x} = \begin{bmatrix} \mathbf{v_1^t} \\ \mathbf{v_2^t} \\ . \\ \mathbf{v_n^t} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{v_1^t x} \\ \mathbf{v_2^t x} \\ . \\ \mathbf{v_n^t x} \end{bmatrix}$$

De modo que

$$\left(\mathbf{v_1^t x}\right)^2 + \ldots\ldots\ldots + \left(\mathbf{v_n^t x}\right)^2 = \mathbf{1}$$

Por la forma de producto exterior de la **DVS**, se tiene que

$\mathbf{A} = \sigma_1 u_1 v_1^\tau + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\tau + \ldots\ \sigma_r \mathbf{u}_r \mathbf{v}_r^\tau$

Por tanto

$$\mathbf{Ax} = \sigma_1 u_1 v_1^\tau \mathbf{x} + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\tau \mathbf{x} + \ldots\ \sigma_r \mathbf{u}_r \mathbf{v}_r^\tau \mathbf{x}$$

$$= \left(\sigma_1 v_1^\tau \mathbf{x}\right) u_1 + \quad \ldots \quad \left(\sigma_r \mathbf{v}_r^\tau \mathbf{x}\right) \mathbf{u}_r$$

$$= y_1 u_1 + \quad \ldots \quad + \mathbf{y}_r \mathbf{u}_r$$

Siendo el escalar $y_i = \sigma_i v_i^\tau \mathbf{x}$

i. Si $\mathbf{r} = \mathbf{n}$, entonces se debe tener $\mathbf{n} \leq \mathbf{m}$ **y**

$$\mathbf{Ax} = y_1 u_1 + \quad \ldots \quad + \mathbf{y}_n \mathbf{u}_n$$

Siendo $\mathbf{y} = \begin{bmatrix} \mathbf{y_1} \\ \mathbf{y_2} \\ . \\ \mathbf{y_n} \end{bmatrix}$**,** luego $\|\mathbf{Ax}\| = \|\mathbf{Uy}\| = \|\mathbf{y}\|$, ya que **U** es ortogonal

Pero $\left(\dfrac{\mathbf{y_i}}{\sigma_i}\right) = v_i^{\tau}\mathbf{x}$, así

$$\left(\frac{\mathbf{y_1}}{\sigma_1}\right)^2 + \ . \ . \ . \ + \left(\frac{\mathbf{y_n}}{\sigma_n}\right)^2 = \left(v_1^{\tau}\mathbf{x}\right)^2 + . \ . \ . \ + \left(v_n^{\tau}\mathbf{x}\right)^2 = 1$$

Lo que demuestra que los vectores $\mathbf{Ax}$, forman la superficie de un elipsoide en $\mathbb{R}^{\mathbf{m}}$

ii. Si $\mathbf{r < n}$,la única diferencia en los pasos anteriores es que la ecuación se convierte en

$$\left(\frac{\mathbf{y_1}}{\sigma_1}\right)^2 + \ . \ . \ . \ + \left(\frac{\mathbf{y_n}}{\sigma_n}\right)^2 = \leq 1$$

Pues faltan algunos términos, esta desigualdad corresponde a un elipsoide sólido de $\mathbb{R}^{\mathbf{m}}$

Ejemplo 2.3.8

Halle la descomposición de valor singular $\mathbf{A = U\sum V^{\ t}}$ para la matriz

$$A = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

Resolución

$$\mathbf{A^t} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \text{ y } \mathbf{A^tA} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

Hállese los valores característicos de $\mathbf{A^tA}$

Ellos son $\lambda_1 = \dfrac{\mathbf{5+\sqrt{5}}}{\mathbf{2}} \simeq \mathbf{3.618}$, $\lambda_2 = \dfrac{\mathbf{5-\sqrt{5}}}{\mathbf{2}} \simeq \mathbf{1.382}$, $\lambda_3 = \mathbf{0}$ y $\lambda_4 = \mathbf{0}$

Siendo sus eigenvectores asociados:

$$\text{Para } \lambda_1\text{: } \mathbf{v}_1 = \begin{bmatrix} .4472 \\ -.2764 \\ .4472 \\ -.7236 \end{bmatrix}, \text{ para } \lambda_2\text{: } \mathbf{v}_2 = \begin{bmatrix} -.4472 \\ -.7236 \\ -.4472 \\ -.2764 \end{bmatrix}$$

Completando la base ortogonal se tiene primeramente $\mathbf{v}_3$ al resolver el sistema homogéneo

$$[\mathbf{v}_1, \mathbf{v}_2, \mathbf{0}], \text{ resultando } \mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 2 \end{bmatrix}$$

Luego se halla $\mathbf{v}_4$ al resolver el sistema homogéneo

$$[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{0}], \text{ resultando } \mathbf{v}_4 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Al normalizarlos se obtiene **A** son

$$\mathbf{v}_1 = \begin{bmatrix} .4472 \\ -.2764 \\ .4472 \\ -.7236 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -.4472 \\ -.7236 \\ -.4472 \\ -.2764 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1/\sqrt{10} \\ -2/\sqrt{10} \\ 1/\sqrt{10} \\ 2/\sqrt{10} \end{bmatrix} \text{ y } \mathbf{v}_4 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$$

Los valores singulares de la matriz **A** son $\sigma_1 = \sqrt{\lambda_1} = 1.902$ y $\sigma_2 = \sqrt{\lambda_2} = 1.176$

$\sigma_3 = \sqrt{\lambda_3} = 0$ y $\sigma_4 = \sqrt{\lambda_4} = 0$

$$\text{Siendo } \mathbf{V} = \begin{bmatrix} .4472 & -.4472 & 1/\sqrt{10} & 1/\sqrt{2} \\ -.2764 & -.7236 & -2/\sqrt{10} & 0 \\ .4472 & -.4472 & 1/\sqrt{10} & -1/\sqrt{2} \\ -.7236 & -.2764 & 2/\sqrt{10} & 0 \end{bmatrix}$$

$$\mathbf{V}^t = \begin{bmatrix} .4472 & -.2764 & .4472 & -.7236 \\ -.4472 & -.7236 & -.4472 & -.2764 \\ 1/\sqrt{10} & -2/\sqrt{10} & 1/\sqrt{10} & 2/\sqrt{10} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \end{bmatrix}$$

$$\sum = \begin{bmatrix} 1.902 & 0 & 0 & 0 \\ 0 & 1.176 & 0 & 0 \end{bmatrix}$$

También

$$\mathbf{u}_1 = \frac{1}{\sigma_1}\mathbf{A}\mathbf{v}_1 = = \begin{bmatrix} .324907 \\ -.525739 \end{bmatrix} \text{ y } \mathbf{u}_2 = \frac{1}{\sigma_2}\mathbf{A}\mathbf{v}_2 = = \begin{bmatrix} -1.376 \\ -.851 \end{bmatrix}$$

$$\text{Así } \mathbf{U} = \begin{bmatrix} .325 & -1.376 \\ -.526 & -.851 \end{bmatrix}$$

Finalmente se tiene

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \mathbf{U} \sum \mathbf{V}^t$$

$$\begin{bmatrix} .325 & -1.376 \\ -.526 & -.851 \end{bmatrix} \begin{bmatrix} 1.902 & 0 & 0 & 0 \\ 0 & 1.176 & 0 & 0 \end{bmatrix} \begin{bmatrix} .4472 & -.2764 & .4472 & -.7236 \\ -.4472 & -.7236 & -.4472 & -.2764 \\ 1/\sqrt{10} & -2/\sqrt{10} & 1/\sqrt{10} & 2/\sqrt{10} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \end{bmatrix}$$

Luego por el Teorema 2.4.3

Se tiene que Rango (A) = r = 2 < n = 4, la ecuación de la elipse es:

$\left(\frac{y_1}{\sigma_1}\right)^2 + \left(\frac{y_2}{\sigma_2}\right)^2 \leq 1$ ó $\left(\frac{y_1}{1.902}\right)^2 + \left(\frac{y_2}{1.176}\right)^2 = \frac{y_1^2}{3.618} + \frac{y_2^2}{1.382} \leq 1$

Ejemplo 2.3.9

Describa la imagen de la esfera $\mathbb{R}^3$, bajo la acción de la matriz

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Resolución

En primer lugar el rango de la matriz **A** es $\mathbf{r(A) = 2}$

$$A^t = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ y}$$

Los eigenvalores de $\mathbf{A^tA}$ son

$\lambda_1 = 2, \lambda_2 = 1$ y $\lambda_3 = 0$, siendo sus correspondientes eigenvectores:

Para $\lambda_1 = 2: \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\lambda_2 = 1: \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ y para $\lambda_3 = 0: \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

Ellos son ortogonales, al normalizarlos se obtiene

$$\mathbf{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \ \mathbf{v_2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ y } \mathbf{v_3} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Los valores singulares de **A** son $\sigma_1 = \sqrt{2}, \sigma_2 = \sqrt{1}$ y $\sigma_3 = \sqrt{0}$

Luego $\mathbf{V} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}$ y $\sum = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Finalmente se hallara **U**

$$u_1 = \frac{1}{\sigma_1} A v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ y } u_2 = \frac{1}{\sigma_2} A v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ as } \mathbf{U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Por tanto el producto buscado es

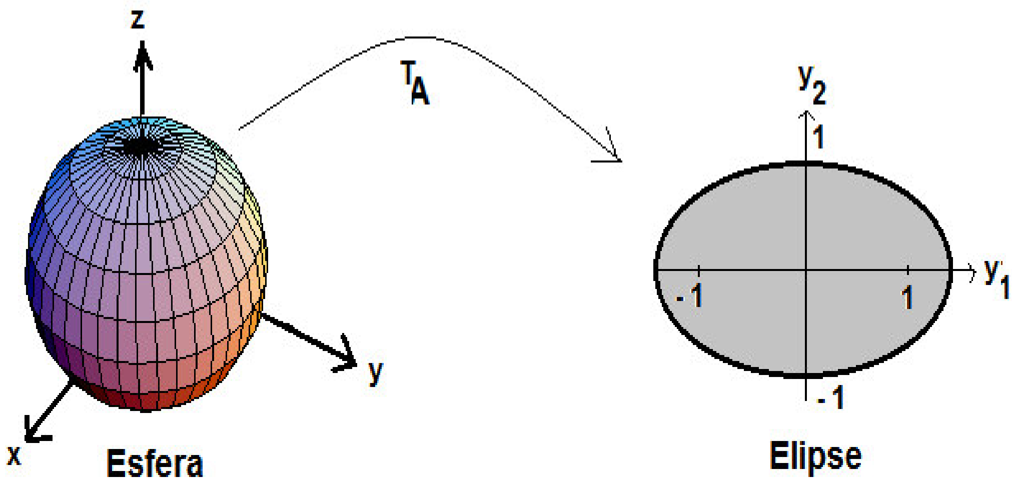
$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} = U\sum V^t$$

Ya que $r(A) = 2 < n = 3$, se aplica la segunda parte del Teorema 2.6.4; la imagen de la esfera unitaria satisface la desigualdad

$$\left(\frac{y_1}{\sqrt{2}}\right)^2 + \left(\frac{y_2}{\sqrt{1}}\right)^2 = \frac{y_1^2}{2} + y_2^2 \leq 1$$

En relación con los ejes coordenados y_1, y_2 en $\mathbb{R}^2$(correspondientes a los vectores izquierdos u_1, u_2), dado que $u_1 = e_1$ y $u_2 = e_2$, la imagen es como se muestra a continuación

Gráfica nª 11:Aplicación de los valores singulares



Fuente propia

En general se puede describir el efecto de una matriz unitaria $\mathbf{A}_{m\times n}$, sobre la esfera unitaria en $\mathbb{R}^n$, en términos del efecto de cada factor en su DVS: $\mathbf{A = U\Sigma V^t}$, de derecha a izquierda

La matriz $\mathbf{V^t}$ que ortogonal, mapea la esfera unitaria en si misma

La matriz Σ que es de **mxn**, hace dos cosas, las entradas diagonales $\sigma_{\mathbf{r+1}} = \mathbf{0}$, $\sigma_{\mathbf{r+2}} = \mathbf{0}, \ldots, \sigma_{\mathbf{n}} = \mathbf{0}$; colapsan $\mathbf{n - r}$ de las dimensiones de la esfera unitaria; lo que deja una esfera unitaria de $\mathbb{R}^r$, en la que las entradas diagonales distintas de cero $\sigma_1, \sigma_2, \ldots, \sigma_r$

devienen en una elipse, lo que hace que la matriz ortogonal **U**, alinee los ejes de este elipsoide con los vectores ortogonales $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ en $\mathbb{R}^{\mathbf{m}}$.

CAPÍTULO 3 MÉTODO DE ESCALADO AFÍN

Descripción del método: hipótesis, teoremas

Algoritmo:

Pasos

1 generar un punto interior

2 Obtener una dirección de búsqueda y determinar la longitud del paso tal que se produzca una mejora

3 Hallar el siguiente punto

4 Verificar si se cumple o no el criterio de parada sino repetir desde el paso 2

3.1 Características del método

Este método se caracteriza por su sencillez, entre todos los métodos de puntos interiores, exhibe un alto rendimiento, aunque el programa sea de gran dimensión

Originalmente fue propuesto en el año 1967 por Dikin; pero occidente no le prestó atención y fue después de 20 años redescubierto y presentado en 1986, por Barnes y Vanderbei; pero esta vez como una simplificación del algoritmo de Karmakar.

El programa que se quiere resolver es

$$\mathbf{P: Mín}\ z = \mathbf{c}^T \mathbf{x}$$

$$\mathbf{St}\ \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$$

Siendo $\mathbf{A}$ una matriz de $\mathbf{m}$ filas y $\mathbf{n}$ columnas, con $\mathbf{n} > \mathbf{m}$ y $\mathbf{A}$ de rango completo

Al conjunto $\mathbf{A\,x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$, se le llama dominio admisible del programa $\mathbf{P}$.

La aplicación del presente algoritmo, exige que se tenga un punto inicial $\mathbf{x^0} > \mathbf{0},\ \mathbf{x^0} \in \mathbb{R}^{\mathbf{n}}$: que pertenezca al **interior** del dominio admisible del programa; el que garantizará que:

$$\mathbf{A\,x^0} = \mathbf{b},\ \mathbf{x^0} > \mathbf{0}$$

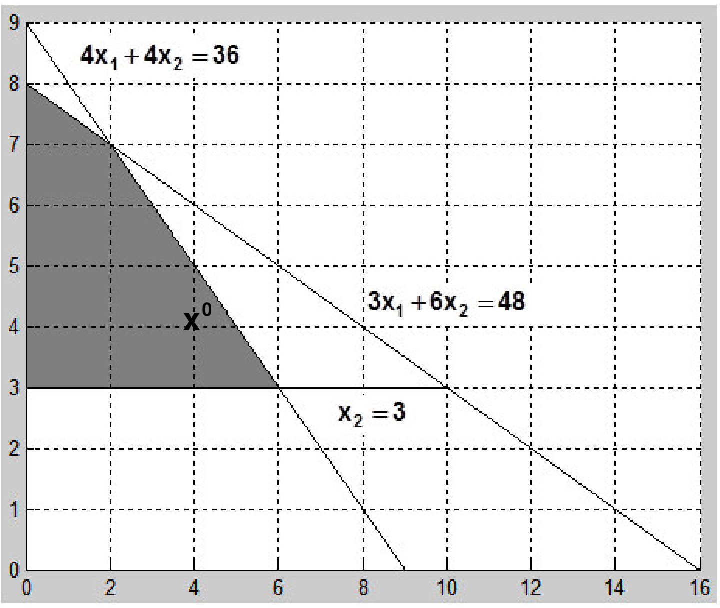
Ejemplo 3.1

• Sea el **P. M. L. P$_1$**

$$\begin{array}{llll} \textbf{Min Z} & -8x_1 + 10x_2 & & \\ \text{St} & 4x_1 + 4x_2 & \leq 36 & \\ & 3x_1 + 6x_2 & \leq 48 & \mathbf{x} \geq 0 \quad (1) \\ & x_2 & \geq 3 & \end{array}$$

Seguidamente se tiene la gráfica del dominio admisible

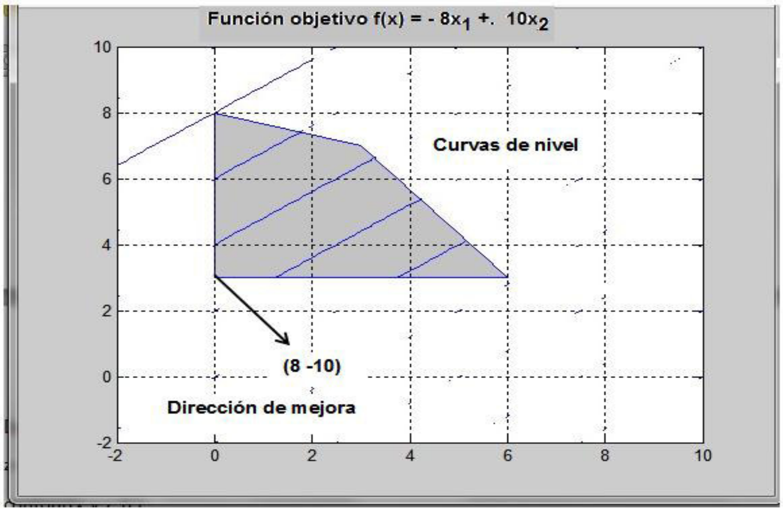
Gráfica n ª12: Dominio admisible del programa P$_1$



Fuente propia

En la presente gráfica los puntos de frontera, del conjunto solución, vienen dados por las ecuaciones asociadas a las restricciones del programa original. El punto $\mathbf{x}^0 = (4, 4)^t$ está en el interior del conjunto solución (dominio admisible del programa)

Gráfica nª 13: Curvas de nivel de la f.o. del programa P_1



Fuente propia

Al adicionar las variables de holgura $\mathbf{x}_3$, $\mathbf{x}_4$, a la primera y segunda restricción respectivamente y sustraer la variable de exceso $\mathbf{x}_5$ a la tercera restricción, se obtiene el programa estandarizado correspondiente

$$
\begin{array}{llllllll}
\mathbf{P}_{11}\text{:} & \mathbf{Min\,Z} & -8\mathbf{x}_1 & +10\mathbf{x}_2 & & & & \\
 & \text{St} & 4\mathbf{x}_1 & +4\mathbf{x}_2 & +\mathbf{x}_3 & & & = \mathbf{36} \\
 & & 3\mathbf{x}_1 & +6\mathbf{x}_2 & & +\mathbf{x}_4 & & = \mathbf{48} \\
 & & & \mathbf{x}_2 & & & -\mathbf{x}_5 & = \mathbf{3} \\
 & & & & \mathbf{x} \geq 0 & & &
\end{array} \qquad (1.1)
$$

Para el presente programa se tiene que el punto $\mathbf{x}^0 = (.5,.5,4,2.5,\ 1)$ pertenece al interior de su dominio admisible y cada una de sus componentes son estrictamente positivas.

Esto es

$$
\begin{array}{llllll}
-8(.5) & +10(.5) & +4 & & & = \mathbf{5} \\
3(.5) & +6(.5) & & -2.5 & & = \mathbf{2} \\
4(.5) & +4(.5) & & & +1 & = \mathbf{5}
\end{array}
$$

El procedimiento básico, seguido por este método, empieza la búsqueda del valor óptimo del programa, a partir del punto inicial $\mathbf{x}^0$, para continuar generando una sucesión de puntos $\{\mathbf{x}^k\}$

Esto es; partiendo de $\mathbf{x}^0$, se avanza hacia $\mathbf{x}^1$, luego a $\mathbf{x}^2$ y así hasta que se verifique alguna condición que indique que el valor actual $\mathbf{x}^k$ es solución óptima del programa

Cada nuevo punto generado, se obtendrá a partir del anterior mediante el procedimiento iterativo:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \Delta \mathbf{x}^k, \ \alpha^k \geq \mathbf{0} \qquad \textbf{(3.1)}$$

Siendo:

i. $\Delta \mathbf{x}^k$ el vector dirección de movimiento

ii. α^k la longitud de paso, indica cuanto se aleja de $\mathbf{x}^k$, a lo largo de la dirección hallada

El presente esquema de búsqueda es muy utilizado por los métodos de optimización.

En consecuencia se estará interesado en dos cuestiones centrales:

i. ¿Cómo obtener la nueva dirección de movimiento $\Delta \mathbf{x}^k$?

ii. ¿Cómo determinar la longitud de paso α^k?

La primera de ellas tiene gran importancia y cubrirá el resto del contenido del tema y la determinación de la longitud de paso, es casi inmediato; una vez que se ha obtenido la dirección de movimiento

3.2 Presentación del programa y supuestos necesarios

Se quiere resolver el Programa Matemático Lineal **(P.M.L.)**

$$\mathbf{P}: \min_{\mathbf{x}} \ \mathbf{f(x) = c^t x}$$

$$\textbf{Sujeto .a} \quad \mathbf{x} \in \mathbf{S}$$

Siendo:

i. $\mathbf{S} = \{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} : \mathbf{Ax = b, x \geq 0}\}$, un conjunto de $\mathbb{R}^{\mathbf{n}}$, no vacío, convexo, al cual se denomina: dominio admisible o región de factibilidad del **p.m.l.**

ii. El hiperplano, $\mathbf{f(x)} = \mathbf{c}^{t}\mathbf{x}$, se denomina función objetivo **(f. o.), con** $\mathbf{c} \in \mathbb{R}^{\mathbf{n}}$

La matriz $\mathbf{A}$ debe ser de valores reales, con $\mathbf{m}$ filas y $\mathbf{n}$ columnas, $\mathbf{m < n}$ y de rango completo: Véase definición 1.3.2.

Se van a considerar las siguientes hipótesis sobre el programa $\mathbf{P}$

Hipótesis 1: La solución del programa $\mathbf{P}$ debe ser factible no degenerada. Véase definición 1.4.1 y 1.4.2

Hipótesis 2: El dominio admisible(o conjunto solución) de $\mathbf{P}$, debe ser un politopo acotado.

El programa dual asociado a $\mathbf{P}$ **;** al que llamaremos $\mathbf{D}$ es

$$\boxed{\begin{array}{l} \mathbf{D}: \underset{\mathbf{y}}{\mathbf{Máx}} \ \mathbf{g(y) = b^{t}y} \\ \\ \mathbf{Sujeto\ .a} \ \mathbf{A^{t}y \leq c} \end{array}}$$

Siendo el vector de variables duales $\mathbf{y} \in \mathbb{R}^{\mathbf{m}}$

Hipótesis 3: El programa dual $\mathbf{D}$, debe tener solución factible no degenerada.

Al escribir el presente programa $\mathbf{D}$ en la forma estándar, mediante la adición del vector de holguras duales $\mathbf{z} \in \mathbb{R}^{\mathbf{n}}$, se obtiene el programa

$$\boxed{\begin{array}{l} \underset{\mathbf{y,z}}{\mathbf{Máx}} \ \mathbf{g(y) = b^{t}y} \\ \\ \mathbf{Sujeto\ .a} \ \mathbf{A^{t}y + z = c} \\ \\ \qquad\qquad\quad \mathbf{z \geq 0} \end{array}}$$

El programa dual debe tener solución Óptima no degenerada

3.3 Obtención de la dirección de mejora y su factibilidad

La dirección de movimiento $\Delta \mathbf{x}^{k}$, debe verificar en cada iteración, dos condiciones:

i. Debe preservar la factibilidad del nuevo punto (dirección admisible)

ii. Debe no empeorar el valor actual de la función objetivo

Supóngase que se está en el punto interior y admisible $\mathbf{x}^{k}$; se quiere que el siguiente punto $\mathbf{x}^{k+1}$ sea también admisible. Por el momento se centrará la atención en las restricciones de igualdad $\mathbf{A\,x} = \mathbf{b}$, las restricciones $\mathbf{x} \geq \mathbf{0}$; luego el nuevo punto $\mathbf{x}^{k+1}$ debe garantizar que

$$\mathbf{A\,x}^{k+1} = \mathbf{b}$$

Se sabe por (**3.1**), que $\mathbf{x}^{k+1} = \mathbf{x}^{k} + \alpha^{k}\,\Delta\,\mathbf{x}^{k}$ y que si $\mathbf{x}^{k}$ es admisible, satisface la relación $\mathbf{A\,x}^{k} = \mathbf{b}$

Por tanto de la condición de factibilidad $\mathbf{A\,x}^{k+1} = \mathbf{b}$, se tiene que

$$\mathbf{A\,x}^{k+1} = \mathbf{b}$$

$$\begin{aligned}
\mathbf{A}(\mathbf{x}^{k} + \alpha^{k}\,\Delta\,\mathbf{x}^{k}) &= \mathbf{A\,x}^{k} + \mathbf{A}\,\alpha^{k}\,\Delta\,\mathbf{x}^{k} = \mathbf{b} \\
&= \mathbf{A\,x}^{k} + \alpha^{k}\,\mathbf{A}\,\Delta\,\mathbf{x}^{k} = \mathbf{b} \\
&= \mathbf{b} + \alpha^{k}\,\mathbf{A}\,\Delta\,\mathbf{x}^{k} = \mathbf{b}
\end{aligned}$$

De donde se desprende que $\alpha^{k}\,\mathbf{A}\,\Delta\,\mathbf{x}^{k} = \mathbf{0}$ y por ser $\alpha^{k} \geq \mathbf{0}$, se concluye que la dirección de mejora debe verificar:

$$\boxed{\mathbf{A}\,\Delta\,\mathbf{x}^{k} = \mathbf{0} \qquad \textbf{(3.1.1)}}$$

Es decir el nuevo punto $\mathbf{x}^{k+1}$ será factible a condición que la nueva dirección de movimiento pertenezca al espacio nulo de $\mathbf{A}$,

Por el Teorema 2.4, se tiene que

$$\mathbf{P} = \mathbf{I}_n - \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A} \qquad \textbf{(3.2)}$$

Siendo $\mathbf{I}_n$: una matriz unitaria de orden $\mathbf{n}$

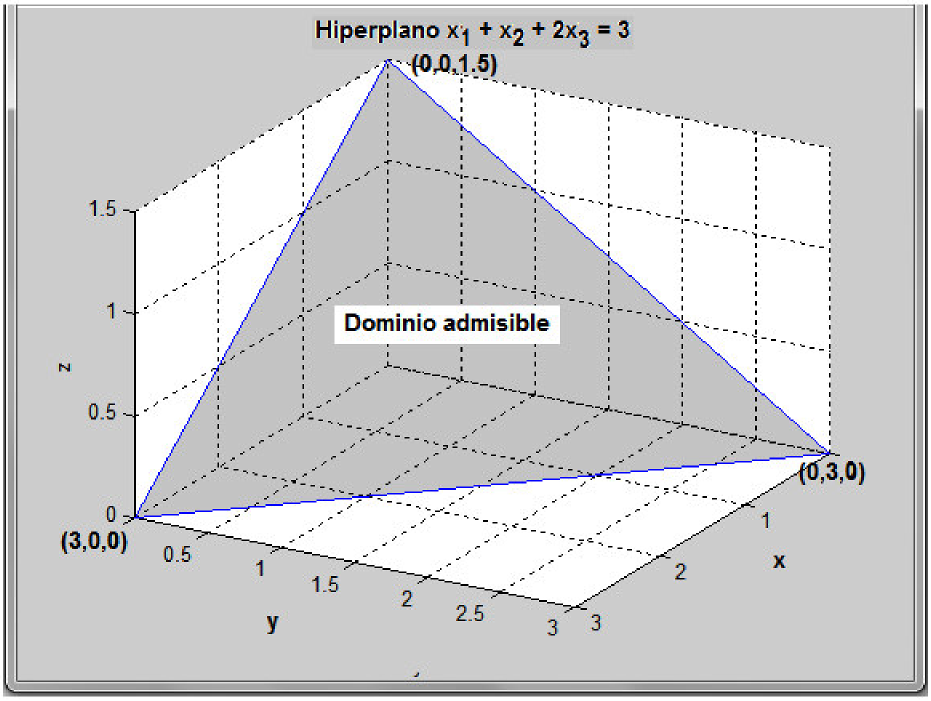
Es la matriz de proyección ortogonal, sobre el espacio nulo de la matriz $\mathbf{A}$. Es decir para cualquier vector $\mathbf{d} \in \mathbb{R}^n$; la proyección ortogonal del vector $\mathbf{d}$, sobre el espacio nulo de la matriz $\mathbf{A}$; será el vector $\mathbf{Pd}$: donde $\mathbf{Pd} \in \mathbf{Null(A)}$

Ejemplo 3.2

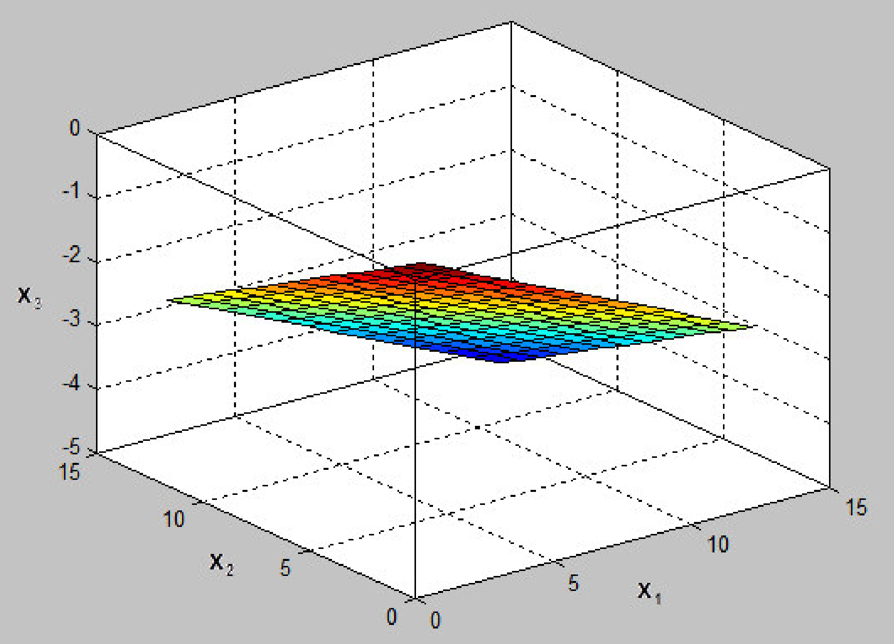
Considérese el programa lineal $\mathbf{P}_2$

$$\begin{array}{lllllll}
\textbf{Min } f(x) = & -1\,x_1 & -1.5\,x_2 & -2\,x_3 & & & \\
\text{St} & x_1 & +\,x_2 & +2\,x_3 & = & 3 \ldots. & 2 \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & & &
\end{array}$$

Gráfica n º 14: Representación del dominio admisible del programa P2



Gráfica n º 15: Representación del espacio nulo de la matriz A: $x_1 + x_2 + 2x_3 = 0$



Fuente propia

$$\mathbf{Null(A)} = \{x \in \mathbb{R}^3 : Ax = 0\} = \{x \in \mathbb{R}^3 : x_1 + x_2 + 2x_3 = 0\}$$

$$\mathbf{Null(A)} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ -\dfrac{x_1 + x_2}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1/2 \end{pmatrix} \alpha + \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix} \beta,\ \alpha, \beta \in \mathbb{R} \right\},$$

$$B_{Null(A)} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix} \right\}$$

También en el presente caso se tiene que $A = (1,1,2)$, $m = 1$, $n = 3$ y $c^t = (-1,-1.5,-2)^t$

El programa toma su mejor valor de punto extremo en $\mathbf{x}^* = \begin{pmatrix} \mathbf{0} \\ \mathbf{3} \\ \mathbf{0} \end{pmatrix}$, siendo

el valor óptimo del programa $\mathbf{f(x^*) = c\, x^* = (-1,-1.5\,,-2)} \begin{pmatrix} \mathbf{0} \\ \mathbf{3} \\ \mathbf{0} \end{pmatrix} \mathbf{= -4.5}$

Al tomar en el dominio admisible del programa, el punto $\mathbf{x^0} = \begin{pmatrix} \mathbf{0.50} \\ \mathbf{2.00} \\ \mathbf{0.25} \end{pmatrix}$,

se tiene que en este punto el valor del programa es $\mathbf{c\, x^0 = (-1,-1.5\,,-2)}$

$$\begin{pmatrix} \mathbf{0.50} \\ \mathbf{2.00} \\ \mathbf{0.25} \end{pmatrix} = \mathbf{-4.0}$$

Tómese un punto cualquiera en $\mathbf{d} \in \mathbb{R}^3$; por ejemplo $\mathbf{d} = \begin{pmatrix} \mathbf{-0.5} \\ \mathbf{-1.0} \\ \mathbf{0.0} \end{pmatrix}$; se

tiene que la proyección ortogonal sobre el espacio nulo de la matriz $\mathbf{A}$,

$\mathbf{P\,d = d - A^T (A\,A^T)^{-1} A\,d}$.

Esto es

$$\mathbf{P\,d} = \begin{pmatrix} \mathbf{-0.5} \\ \mathbf{-1.5} \\ \mathbf{0.0} \end{pmatrix} - \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{2} \end{pmatrix} \left[\underbrace{(1,1,2) \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{2} \end{pmatrix}}_{\mathbf{1/6}} \right]^{-1} \left[\underbrace{(1,1,2) \begin{pmatrix} \mathbf{-0.5} \\ \mathbf{-1.5} \\ \mathbf{0.0} \end{pmatrix}}_{\mathbf{-2}} \right] = \begin{pmatrix} \mathbf{-0.5} \\ \mathbf{-1.5} \\ \mathbf{0.0} \end{pmatrix} - \begin{pmatrix} \mathbf{-1/3} \\ \mathbf{-1/3} \\ \mathbf{-1/3} \end{pmatrix} = \begin{pmatrix} \mathbf{-1/6} \\ \mathbf{-7/6} \\ \mathbf{4/6} \end{pmatrix}$$

Así considerando $\boldsymbol{\alpha = 1}$, el nuevo punto será $\mathbf{x^1 = x^0 + P\,d}$.

a. ¿ $\mathbf{x^1}$ estará en el dominio admisible del programa?

b. ¿En el punto $\mathbf{x^1}$ el nuevo valor del programa $\mathbf{c\,x^1}$, será mejor que en el anterior $\mathbf{c\,x^0}$?

Respuesta

a. Será factible si el punto verifica que $\mathbf{x}^1 \geq \mathbf{0}$ y $\mathbf{A}\,\mathbf{x}^1 = \mathbf{b}$

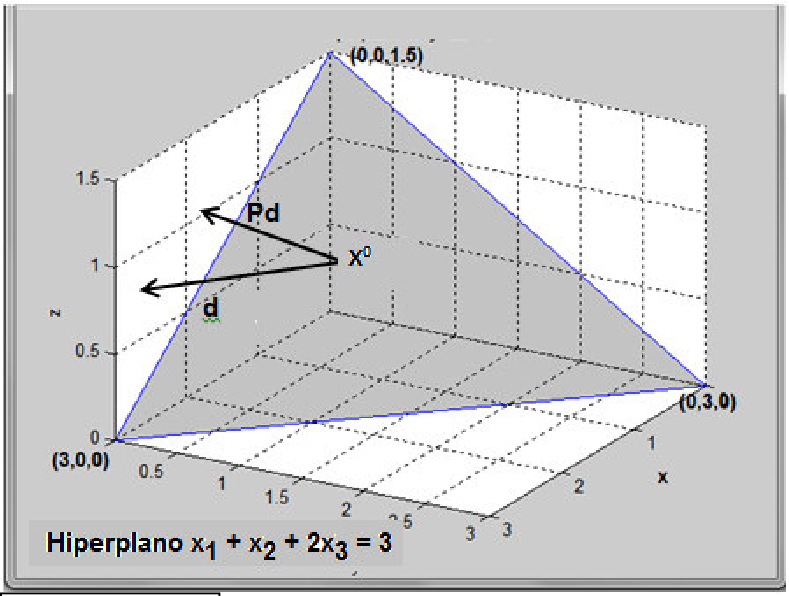
En efecto: $\mathbf{x}^1 = \mathbf{x}^0 + \mathbf{Pd} = \begin{pmatrix} 1/2 \\ 2 \\ 1/4 \end{pmatrix} + \begin{pmatrix} -1/6 \\ -7/6 \\ 4/6 \end{pmatrix} = \begin{pmatrix} 4/12 \\ 10/12 \\ 11/12 \end{pmatrix} \geq \mathbf{0}$

$\mathbf{A}\,\mathbf{x}^1 = (1,1,2)\begin{pmatrix} 4/12 \\ 10/12 \\ 11/12 \end{pmatrix} = \dfrac{4+10+22}{12} = 3 \therefore \mathbf{x}^1$ Está en el dominio admisible del programa

b. En este punto el valor del programa será:

$$\mathbf{c}\ \mathbf{x}^1 = (-1,-1.5,-2)\begin{pmatrix} 4/12 \\ 10/12 \\ 11/12 \end{pmatrix} \approx -3.417$$

Gráfica nº 16: Proyección ortogonal del vector d, sobre Null(A)



Fuente propia

Este último resultado hace notar; que al elegir cualquier valor para $\mathbf{d}$; ha llevado a que en el nuevo punto $\mathbf{x}^1$; el valor del programa empeore ($\mathbf{c}\ \mathbf{x}^1 > \mathbf{c}\,\mathbf{x}^0$); a pesar que $\mathbf{x}^1$ está en el dominio admisible del programa. Este hecho sugiere la necesidad de caracterizar la dirección de mejora **"d"**, del valor del programa

3.4 Criterio para seleccionar la dirección de mejora

Si se considera la función objetivo del programa como minimización; en el nuevo punto del programa $\mathbf{x}^{k+1}$, debería ocurrir que

$$\mathbf{c}\,\mathbf{x}^{k+1} \leq \mathbf{c}\,\mathbf{x}^{k}$$

Esto significa que; si $\mathbf{x}^{k+1} = \mathbf{x}^{k} + \alpha\,\Delta\mathbf{x}$, considerando $\alpha > \mathbf{0}$. De la relación anterior se deberá cumplir $\mathbf{c}(\mathbf{x}^{k} + \alpha\,\Delta\mathbf{x}) \leq \mathbf{c}\,\mathbf{x}^{k}$, ó

$$\boxed{\alpha\,\mathbf{c}\,\Delta\mathbf{x} \leq \mathbf{0}}$$

Esta última relación recibe el nombre de "**condición de descenso**" y el vector $\Delta\mathbf{x}$, recibirá el nombre de **"dirección de mejora"**

Ahora; considérese la siguiente dirección

$$\Delta\mathbf{x} = -\,\mathbf{P}\,\mathbf{c} \qquad \textbf{(3.3)}$$

que es la proyección ortogonal (del negativo del gradiente de la función objetivo), sobre el espacio nulo de A.

Se probará que esta elección, es una dirección de mejora del programa

En efecto:

$\mathbf{c}^t\Delta\mathbf{x} = -\mathbf{c}^t\mathbf{P}\mathbf{c}$, de donde se tiene $\mathbf{c}^t\Delta\mathbf{x} = -\mathbf{c}^t\mathbf{P}^2\mathbf{c}$, por ser $\mathbf{P}$ idempotente

Luego $\mathbf{c}^t\Delta\mathbf{x} = -\mathbf{c}^t\mathbf{P}^t\mathbf{P}\mathbf{c} = -(\mathbf{P}\mathbf{c})^t\mathbf{P}\mathbf{c} = -\|(\mathbf{P}\mathbf{c})\|^2 \leq \mathbf{0}$

Concluyendo que la condición de descenso es $\mathbf{c}^t\Delta\mathbf{x} \leq \mathbf{0}$

Por tanto $\mathbf{c}^t\,\Delta\mathbf{x} \leq \mathbf{0}$ y cumple la <u>condición de descenso</u>

Ejemplo 3.3

Considérese el programa lineal $\mathbf{P_2}$, del (Del Ejemplo 3.2); pero utilicemos como dirección de movimiento $\mathbf{d} = -\mathbf{c}^t = \begin{pmatrix} 1.0 \\ 1.5 \\ 2.0 \end{pmatrix}$,(el negativo del gradiente de la función objetivo del programa)

La proyección ortogonal de $\mathbf{d}$, sobre el espacio nulo de $\mathbf{A}$, será $\mathbf{Pd}$

$$\mathbf{Pd} = \begin{pmatrix} 1.0 \\ 1.5 \\ 2.0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \left[\underbrace{(1,1,2)\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}}_{1/6} \right]^{-1} \left[\underbrace{(1,1,2)\begin{pmatrix} 1.0 \\ 1.5 \\ 2.0 \end{pmatrix}}_{13/2} \right] = \begin{pmatrix} 1.0 \\ 1.5 \\ 2.0 \end{pmatrix} - \begin{pmatrix} 13/12 \\ 13/12 \\ 26/12 \end{pmatrix} = \begin{pmatrix} -1/12 \\ 5/12 \\ -2/12 \end{pmatrix}$$

Luego el punto $\mathbf{x}^1 = \mathbf{x}^0 + \mathbf{Pd}$, considerando nuevamente $\alpha = 1$

a. ¿Estará en el dominio admisible del programa?

b. ¿El nuevo valor del programa es?

c. ¿Qué relación guarda el valor del presente programa, con el hallado, cuando se tomó como dirección de mejora $\mathbf{Pd} = \begin{pmatrix} -1/12 \\ -5/12 \\ 0.0 \end{pmatrix}$?

Respuesta

a. Será factible si el punto verifica que $\mathbf{x}^1 \geq \mathbf{0}$ y $\mathbf{A\,x}^1 = \mathbf{b}$

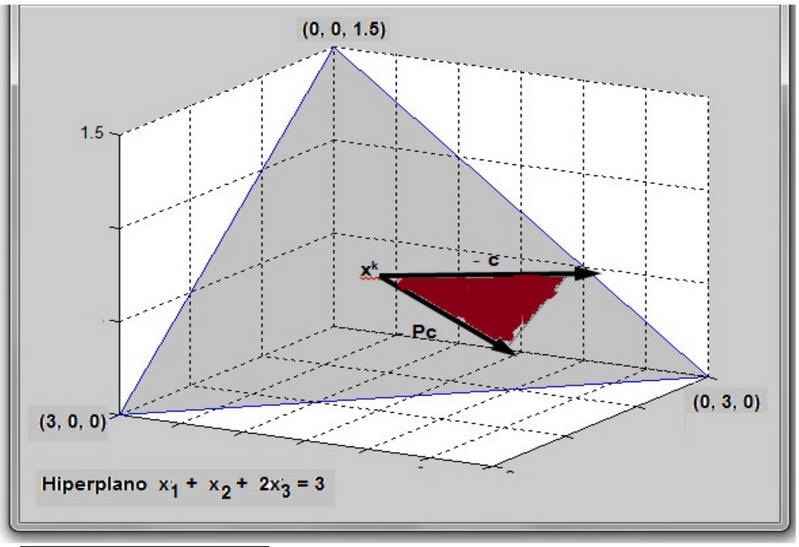
En efecto: $\mathbf{x}^1 = \mathbf{x}^0 + \mathbf{Pd} = \begin{pmatrix} 1/2 \\ 2 \\ 1/4 \end{pmatrix} + \begin{pmatrix} -1/12 \\ 5/12 \\ -2/12 \end{pmatrix} = \begin{pmatrix} 5/12 \\ 29/12 \\ 1/12 \end{pmatrix} \geq \mathbf{0}$

$\mathbf{A\,x}^1 = (1,1,2)\begin{pmatrix} 5/12 \\ 29/12 \\ 1/12 \end{pmatrix} = \dfrac{5+29+2}{12} = 3 \therefore \mathbf{x}^1$ está en el dominio admisible del programa.

b. En este punto el valor del programa será $\mathbf{c}\ \mathbf{x}^1 = (-1, -1.5, -2) \begin{pmatrix} 5/12 \\ 29/12 \\ 1/12 \end{pmatrix} \simeq \mathbf{-4.208}$

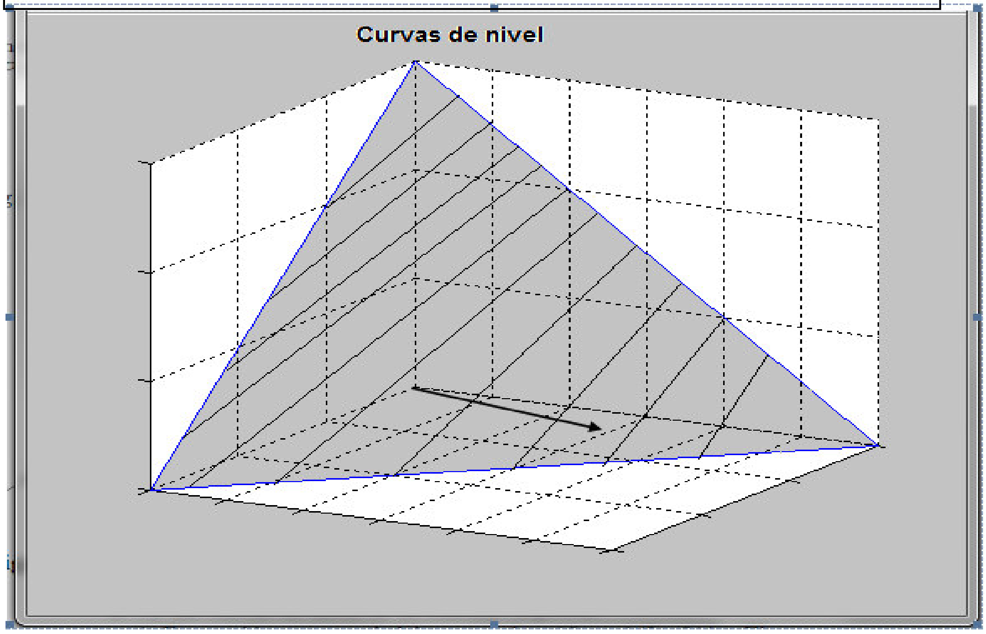
c. Al elegir el vector $\mathbf{d} = -\mathbf{c}$, como el negativo del gradiente la función objetivo; se obtiene un mejor valor del programa; que el obtenido al elegir $\mathbf{d} = \begin{pmatrix} -0.5 \\ -1.5 \\ 0.0 \end{pmatrix}$

Figura n º 17: Proyección ortogonal del vector - c



Fuente propia

Gráfica n º 18:Representación de las curvas de nivel y dirección de mejora



Fuente propia

En consecuencia si se toma como dirección de movimiento $-\mathbf{c} \in \mathbb{R}^{\mathbf{n}}$, siempre se tendrá que su proyección ortogonal $-\mathbf{Pc}$, sobre el espacio nulo de $\mathbf{A}$; será una dirección factible del programa por lo que $\mathbf{APc} = \mathbf{0}$ (

¿Qué pasaría si para la dirección de búsqueda $\Delta\mathbf{x} = \mathbf{P\,c}$; se verifica que $\|\mathbf{Pc}\|^{2} = \mathbf{0}$?

Veamos**:**

Si $\mathbf{Pc}$ la proyección ortogonal sobre **Null(A)** es nula ($\mathbf{Pc} = \mathbf{0}$**)**; por (2.3.4) significa que existe $\mathbf{y} \in \mathbb{R}^{\mathbf{m}}$, tal que

$$\mathbf{c}^{t} = \mathbf{A}^{t}\mathbf{y}$$

Esto es, los coeficientes de la función objetivo del programa original: $\mathbf{c}^{t}$, se han logrado expresar como combinación lineal de los vectores del subespacio de las filas de la matriz $\mathbf{A}$ o $\mathbf{c}^{t}$ está en el complemento ortogonal del subespacio nulo de A, o el estimador de las holguras duales z; se ha vuelto nulo

Si se considera un punto factible $\mathbf{x}$ cualquiera del programa, se tendrá que $\mathbf{c}^{t}\,\mathbf{x} = \mathbf{y}^{t}\,\mathbf{A}\,\mathbf{x}$

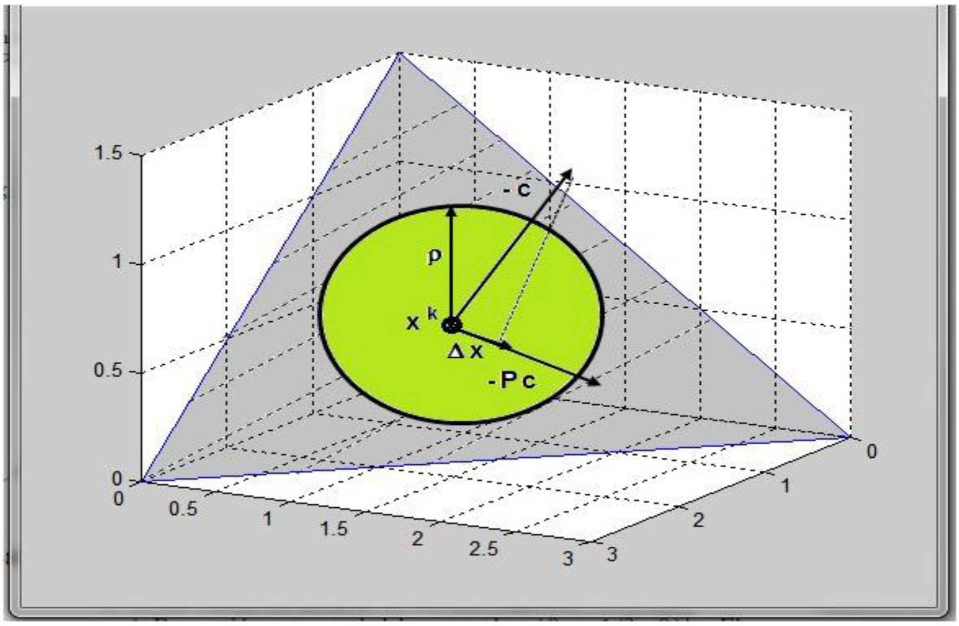
Pero por la condición de $\mathbf{x}$: $\mathbf{A\,x = b}$, se tiene que $\mathbf{y^t\,A\,x = y^t\,b}$

De donde se tiene que $\mathbf{c^t\,x} = \mathbf{y^t\,A\,x} = \mathbf{y^t\,b}$ y $\mathbf{c^t\,x} = \mathbf{y^t\,b}$

Esta relación $\mathbf{c^t\,x} = \mathbf{y^t\,b}$, significa que tanto $\mathbf{x}$ como $\mathbf{y^t}$ son óptimos de los programas **P** y **D** respectivamente

3.5 Elección de la dirección de mejoramiento

Gráfica nº 19: Representación de la dirección de mejora



Fuente propia

Si se quiere seleccionar la mejor de las direcciones $\mathbf{\Delta x}$, que sea factible y que mejore el valor actual del programa, a partir del punto $\mathbf{x^k}$, se plantea el siguiente programa:

$$\mathbf{R : \underset{\Delta x}{Min}\ f(x) = c\,(x^k + \Delta x)}$$

$$\mathbf{S.\,a.\ \ A\,(x^k + \Delta x) = b} \qquad \mathbf{(3.4)}$$

$$\mathbf{\|\Delta x\|^2 = \rho^2}$$

La segunda restricción indica que:

a. $\mathbf{\Delta x}$ debe tener una longitud ρ, valor que se ajustará convenientemente, para conseguir que el nuevo punto $\mathbf{x^k} + \mathbf{\Delta x} \geq \mathbf{0}$ y pertenezca al dominio admisible del programa

b. El nuevo valor del programa $\mathbf{f(x^k)}$, **(3.4)** sea acotado

Nótese que en este programa se quiere hallar el mínimo valor de $\mathbf{f(x)}$, en el dominio admisible, que forman el hiperplano $\mathbf{A\,(x^k + \Delta x) = b}$ y la hiperesfera de radio ρ centrada en $\mathbf{x^k}$, $\|\Delta\mathbf{x}\|^2 = \rho^2$

El presente programa puede ser resuelto, utilizando los multiplicadores de Lagrange

$$\mathbf{L(\Delta x, y, \lambda)} = \mathbf{c\,(x^k + \Delta x) - y^T(A\,(x^k + \Delta x) - b) + \lambda\,(\Delta x^T\,\Delta x - \rho^2)}$$

Donde: Los vectores $\mathbf{y^T}$ y λ son multiplicadores de Lagrange, asociado al primer y segundo conjunto de restricciones respectivamente

Derivando $\mathbf{L}$, con respecto a $\mathbf{\Delta x}$, $\mathbf{y}$ y λ e igualando a cero en cada caso

$$\frac{\partial}{\partial \Delta \mathbf{x}}\mathbf{L(\Delta x, y, \lambda)} = \mathbf{c - y^T A + 2\,\lambda\,\Delta x = 0}$$

$$\frac{\partial}{\partial \mathbf{y}}\mathbf{L(\Delta x, y, \lambda)} = \mathbf{A\,(x^k + \Delta x) - b = 0}$$

$$\frac{\partial}{\partial \lambda}\mathbf{L(\Delta x, y, \lambda)} = \mathbf{\Delta x^T\,\Delta x - \rho^2 = 0}$$

$$\frac{\partial^2}{\partial \Delta \mathbf{x}^2}\mathbf{L(\Delta x, y, \lambda) = 2\lambda}$$

En la segunda ecuación como $\mathbf{A\,x^k = b} \Rightarrow \mathbf{A\,\Delta x = 0}$

De la primera ecuación $\mathbf{c = y^T A - 2\,\lambda\,\Delta x}$

Al Multiplicar por la matriz $\mathbf{A}$ ambos miembros se obtiene

$\mathbf{A\,c = A\,A^T\,y - 2\,\lambda\,A\,\Delta x}$; pero por ser $\mathbf{A\,\Delta x = 0}$, resulta

$$\mathbf{A\,c = A\,A^T\,y}$$

Y por ser la matriz $\mathbf{A}$ de rango completo, se tendrá

$$\mathbf{y = (A\,A^t)^{-1}\,A\,c} \qquad \textbf{(3.5)}$$

De la relación $\mathbf{c} - \mathbf{y}^t \mathbf{A} + 2\lambda\, \Delta\mathbf{x} = \mathbf{0}$, se tiene $2\lambda\, \Delta\mathbf{x} = -\mathbf{c} + \mathbf{y}^t \mathbf{A} - \mathbf{c} + \mathbf{A}^t \mathbf{y}$

Reemplazando **(3.2)**, en esta última relación, se tendrá

$2\lambda\, \Delta\mathbf{x} = -\mathbf{c} + \mathbf{A}^t (\mathbf{A}\,\mathbf{A}^t)^{-1} \mathbf{A}\,\mathbf{c}$ ó

$2\lambda\, \Delta\mathbf{x} = -(\mathbf{I}_n - \mathbf{A}^t (\mathbf{A}\,\mathbf{A}^t)^{-1} \mathbf{A})\,\mathbf{c} = -\mathbf{P}\,\mathbf{c}$, de donde

$$\Delta\mathbf{x} = -\frac{1}{2\lambda}\mathbf{P}\,\mathbf{c}$$

Elevando ambos miembros al cuadrado $\rho^2 = \|\Delta\mathbf{x}\|^2 = \frac{1}{4\lambda^2}\|\mathbf{Pc}\|^2$

La relación $\rho^2 = \frac{1}{4\lambda^2}\|\mathbf{Pc}\|^2 \Rightarrow \lambda^2 = \frac{1}{4\rho^2}\|\mathbf{Pc}\|^2 \Rightarrow \lambda = \frac{1}{2\rho}\|\mathbf{Pc}\|$

De esta última relación y $\Delta\mathbf{x} = -\frac{1}{2(\frac{1}{2\rho}\|\mathbf{Pc}\|)}\mathbf{P}\,\mathbf{c}$, se obtiene

$$\boxed{\Delta\mathbf{x} = -\frac{\rho}{\|\mathbf{Pc}\|}\mathbf{P}\,\mathbf{c}} \qquad \textbf{(3.6)}$$

Por otro lado

$$\frac{\partial^2}{\partial \Delta\mathbf{x}^2}\mathbf{L}(\Delta\mathbf{x}, \mathbf{y}, \lambda) = 2\lambda > 0$$

Lo que significa que $\lambda = \frac{1}{2\rho}\|\mathbf{Pc}\| > \mathbf{0}$¸ por lo que la expresión **(3.6)**, proporciona el valor mínimo del programa

Esto justifica el signo positivo para λ en la función lagrangiana**.** Si se hubiera utilizado el signo menos; la solución hallada hubiera en dirección contraria a (**(3.6)**), luego se habría obtenido $\frac{\partial^2}{\partial \Delta\mathbf{x}^2}\mathbf{L}(\Delta\mathbf{x}, \mathbf{y}, \lambda) = -2\lambda$, lo que indicaría que el punto es de máximo

Esta es la dirección; **(-Pc)** a la que llamaremos el "negativo del gradiente proyectado", que garantiza el máximo descenso, a partir de un punto determinado; manteniendo la factibilidad del siguiente. Esta dirección coincide con la utilizada en el método del gradiente proyectado, para programación no lineal.

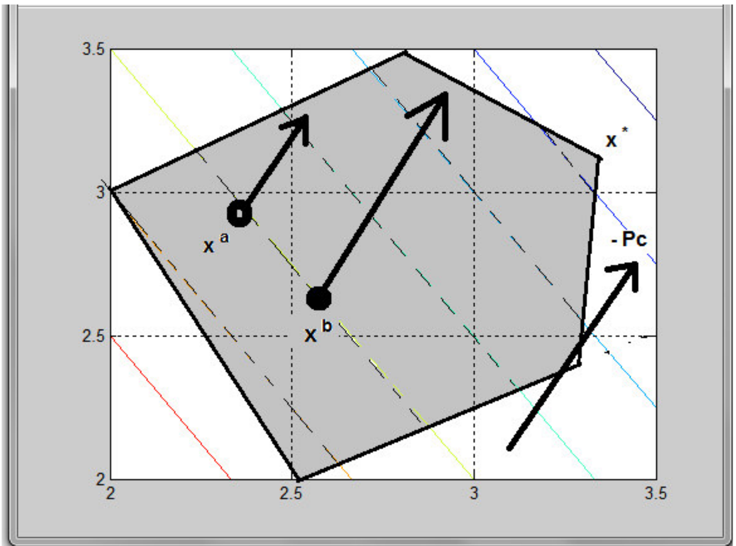
3.6 Etapas del escalado afín

El gradiente proyectado ofrece una buena dirección, en el sentido que entre todas las direcciones factibles, es la que proporciona la máxima disminución del valor del programa **P**, por unidad de longitud de paso

Sin embargo según la relación **(3.1)**, que rige el proceso iterativo de búsqueda del presente método; si se halla cerca de las caras del dominio admisible del programa; la longitud de paso α sería pequeña (de no ser así, violaría las restricciones de no negatividad de las variables). Lo que significa que no siempre es suficiente, tener una buena dirección de descenso; sino es necesario estar en el centro del dominio admisible del programa; esto es alejado de los hiperplanos $\mathbf{x} = \mathbf{0}$. Lo que obliga a tener $\alpha \geq \mathbf{0}$

Esta situación se aprecia en el siguiente gráfico

Gráfica n º 20: Movimiento en la dirección del gradiente proyectado



Fuente propia

Siendo el movimiento desde el punto $\mathbf{x}^1$ resulta pequeño comparado con el movimiento desde el punto $\mathbf{x}^2$

El método de **Escalado afín** resuelve el problema de no estar en el centro del dominio admisible del programa, por medio de un procedimiento de tres etapas: **a**, **b** y **c**; las que se repiten en cada iteración del algoritmo

a. Escalar el problema de forma tal que el punto actual se encuentre alejado de los hiperplanos generadores del dominio admisible; esto es el punto de búsqueda $\mathbf{x}^{\mathbf{k}}$ se transformará en $\mathbf{x}^{\mathbf{k}}$, en el nuevo conjunto solución formada por las variables $\mathbf{x}$

b. Obtener la dirección del gradiente proyectado en el nuevo dominio admisible; el que simbolizaremos con $\Delta\mathbf{x}$

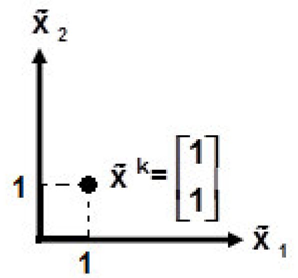
c. Trasladar la dirección obtenida $\Delta\mathbf{x}$ al problema original, deshaciendo el traslado previo. A partir de $\Delta\mathbf{x}$ se obtendrá la dirección $\Delta\mathbf{x}$ en el espacio original

¿Elección del tipo de escalado a utilizar?

Si por ejemplo se utilizara un escalado que multiplicara todas las componentes $\mathbf{x}_i^k$, i = 1, 2, . . . , n, por una misma constante; no sería de utilidad; porque se mantendría la distancia relativa de cada una de las variables, respecto de los hiperplanos generadores del dominio admisible del programa

La estrategia seguida por el presente método, consiste en transformar cada componente $\mathbf{x}_i^k$, i = 1, 2, . . . , n, del dominio admisible original en $\mathbf{x}_i^k = 1$, en el nuevo dominio admisible; garantizándose de esta forma el alejamiento de los hiperplanos generadores del dominio admisible del programa y esto se consigue mediante la relación

Gráfica n º 21 Centro del d. a. programa escalado



Fuente propia

$$\mathbf{x}_i^k = \frac{\mathbf{x}_i}{\mathbf{x}_i^k},\ \mathbf{x}_i^k,\ i = 1, 2, \ldots, \mathbf{n}$$

Siendo $\mathbf{x}_i^k$: El punto actual de la iteración

Este cambio se deshace mediante la relación:

$$\mathbf{x}_i = \mathbf{x}_i^k \mathbf{x}_i^k,\ i = 1, 2, \ldots, \mathbf{n}$$

Si se denota por X^k la matriz diagonal que tiene por componentes los términos del punto actual x_i^k; esto es

$$X^k = \begin{pmatrix} x_1^k & 0 & 0 & 0 & 0 \\ 0 & x_2^k & 0 & 0 & 0 \\ & & . & & \\ & & & . & \\ 0 & 0 & 0 & 0 & x_n^k \end{pmatrix}$$

Es posible escribir las relaciones anteriores en forma matricial:

$$\left.\begin{aligned} x &= (X^k)^{-1}x \\ x &= X^k x \end{aligned}\right\} \qquad \textbf{(3.7)}$$

Nótese que este escalado está bien definido, ya que el movimiento siempre ocurrirá por el interior del dominio admisible, con lo que se asegura que $x_i^k > 0$

Lo que implica que X^k será una matriz no singular y definida positiva

Apreciemos seguidamente. **¿c**ómo se realizan en cada iteración, las tres etapas; (**a**), (**b**) y (**c**), descritas anteriormente**?**:

3.7.1 Aplicación de las etapas.

Considérese el programa matemático lineal

$$Q_1 \ \text{Mín} \ f(x) = c^T x$$

$$\text{S.a.} \ \ A x = b,$$

$$x \geq 0$$

Etapa a: Escalamiento del programa P_1, mediante las relaciones **(3.7)** y con $x^k > 0$

$$P_1 \rightarrow P_1' : \underset{\tilde{x}}{\text{Mín}} \ f(x) = c^T X^k x$$

$$\text{S.a.} \qquad A X^k x = b$$

$$X^k x \geq 0$$

Si en el programa P_1', se toma:

$$\mathbf{c} = \mathbf{X}^{k}\mathbf{c} \;\; \text{y} \;\; \mathbf{A} = \mathbf{A}\,\mathbf{X}^{k} \qquad \textbf{(3.8)}$$

Se obtendrá el programa $\tilde{\mathbf{P}}_1^{'}$:

$$\mathbf{P}_1^{'} \rightarrow \tilde{\mathbf{P}}_1 \;\; \textbf{Mín} \;\; \mathbf{f}(\tilde{\mathbf{x}}) = \mathbf{c}\,\mathbf{x}$$

$$\textbf{S.a.} \qquad \mathbf{A}\,\mathbf{x} = \mathbf{b}$$

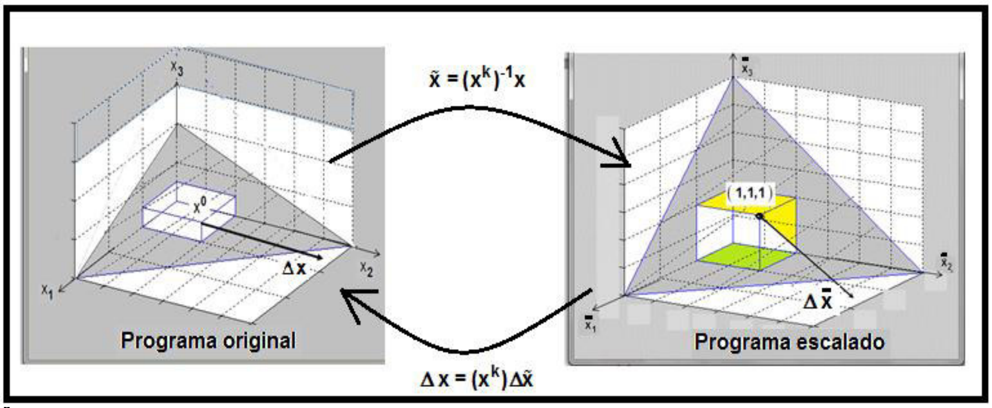
$$\mathbf{x} \geq \mathbf{0}$$

Nótese que en último paso las restricciones $\mathbf{X}^{k}\mathbf{x} \geq \mathbf{0}$, se han transformado en $\mathbf{x} \geq \mathbf{0}$. Ambos conjuntos de restricciones son equivalentes, debido a que el punto actual de iteración $\mathbf{x}^{k}$, tiene todas sus componentes positivas

Supóngase que en la búsqueda se arriba al punto $\mathbf{x}^{k}$; que está en el interior del dominio admisible del programa $\mathbf{P}_1$:

Este se transformará en el punto $\mathbf{x}^{k} = \mathbf{e} = (\mathbf{1}, \mathbf{1}, . \;\; . \;\; . \;\; , \mathbf{1})$, cuya posición es el centro del dominio admisible del programa $\tilde{\mathbf{P}}_1$

Gráfica n ª 22: Aplicación de la transformación Afín



Fuente propia

e además que este tipo de escalado, transforma también las curvas de nivel de la función objetivo, a lo largo del dominio admisible del programa

Etapa b: Obtención de $\Delta\,\mathbf{x}$

En el dominio admisible de variables $\mathbf{x}$, la dirección de descenso $\Delta\,\mathbf{x}$, será la del gradiente proyectado, la que se obtiene mediante la relación (**3.8**). Pero ahora se trabajará con el programa:

$$\tilde{\mathbf{P}}_1: \ \mathbf{Mín} \ \ \mathbf{f}(\tilde{\mathbf{x}}) = \mathbf{c}\,\mathbf{x}$$

$$\mathbf{S.a.} \qquad \mathbf{A}\,\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

Teniendo en cuenta que $\mathbf{X}^{\mathbf{k}}$ es una matriz diagonal, obténgase $\Delta\,\mathbf{x}$:

$$\Delta\,\mathbf{x} = \mathbf{-\,P\,c}$$

$$= -\left[\mathbf{I}_{\mathbf{n}} - \mathbf{A}^{t}(\mathbf{A}\mathbf{A}^{t})^{-1}\mathbf{A}\right]c$$

$$= -\left[\mathbf{I}_{\mathbf{n}} - (AX^K)^t((AX^K(AX^K)^t)^{-1}AX^K\right]X^Kc$$

$$= -\left[\mathbf{I}_{\mathbf{n}} - (AX^K)^t(AX^KX^KA^t)^{-1}AX^K\right]X^Kc$$

$$= -\left[\mathbf{I}_{\mathbf{n}} - (AX^K)^t(A(X^K)^2A^t)^{-1}AX^K\right]X^Kc$$

$$= -\left[X^K - (AX^K)^t(A(X^K)^2A^t)^{-1}AX^KX^K\right]c$$

$$= -\left[X^K - (AX^K)^t(A(X^K)^2A^t)^{-1}A(X^K)^2\right]c$$

$$= -\left[X^K - X^KA^t(A(X^K)^2A^t)^{-1}A(X^K)^2\right]c$$

$$\Delta\,\mathbf{x} = \ -\,\mathbf{X}^K\left[\mathbf{I}_n - \mathbf{A}^t(\mathbf{A}(X^K)^2\mathbf{A}^t)^{-1}\mathbf{A}(X^K)^2\right]\mathbf{c} \qquad \mathbf{(3.9)}$$

Etapa c : Regreso al programa original $\mathbf{P}: \ \mathbf{x} \rightarrow \mathbf{x}$

En el problema escalado el movimiento (sin tener en cuenta la longitud de paso) se realiza así

$$\boxed{\mathbf{x}^{\mathbf{k+1}} = \mathbf{x}^{\mathbf{k}} + \Delta\,\mathbf{x}}$$

Deshaciendo el cambio de variable: Multiplicando la relación anterior por $\mathbf{X}^{\mathbf{k}}$

$$\mathbf{X}^{\mathbf{k}}\,\mathbf{x}^{\mathbf{k+1}} = \mathbf{X}^{\mathbf{k}}\,\mathbf{x}^{\mathbf{k}} + \mathbf{X}^{\mathbf{k}}\ \Delta\,\mathbf{x}$$

$$\boxed{\mathbf{x}^{\mathbf{k+1}} = \mathbf{x}^{\mathbf{k}} + \mathbf{X}^{\mathbf{k}}\ \Delta\,\mathbf{x}}$$

En consecuencia el movimiento $\Delta\,\mathbf{x}$ realizado en el programa original, seria $\mathbf{X}^{\mathbf{k}}\ \Delta\,\mathbf{x}$

Reemplazando $\Delta\,\mathbf{x}$ de la relación **(3.9)**, se obtiene:

$$\boxed{\Delta\, \mathbf{x} = \mathbf{X}^{k}\ \Delta\, \mathbf{x}} \qquad \mathbf{(3.10)}$$

$$\Delta\, \mathbf{x} = \ -\mathbf{X}^{K}\left[\mathbf{X}^{K}\left[\mathbf{I}_{n}-\mathbf{A}^{t}(\mathbf{A}(X^{K})^{2}\mathbf{A}^{t})^{-1}\mathbf{A}(X^{K})^{2}\right]\mathbf{c}\right]$$

$$\Delta\, \mathbf{x} = \ -(\mathbf{X}^{K})^{2}\left[\mathbf{I}_{n}-\mathbf{A}^{t}(\mathbf{A}(\mathbf{X}^{K})^{2}\mathbf{A}^{t})^{-1}\mathbf{A}(\mathbf{X}^{K})^{2}\right]\mathbf{c}$$

Si se hace $\mathbf{D} = \mathbf{(X^{k})^{2}}$; se tendrá

$$\Delta\, \mathbf{x} = -\mathbf{D}\left[\mathbf{I}_{n} - \mathbf{A}^{t}(\mathbf{A\,D\,A}^{t})^{-1}\mathbf{AD}\right]\mathbf{c} = \ -\mathbf{D}\left[\mathbf{c} - \mathbf{A}^{t}(\mathbf{A\,D\,A}^{t})^{-1}\mathbf{ADc}\right]$$

$$\Delta\, \mathbf{x} = -\mathbf{D}\left[\mathbf{c} - \mathbf{A}^{t}(\mathbf{A\,D\,A}^{t})^{-1}\mathbf{ADc}\right]$$

Haciendo $\mathbf{y} = (\mathbf{A\,D\,A}^{t})^{-1}\ \mathbf{A\ D\ c}$; se obtendrá

$\Delta\, \mathbf{x} = -\ \mathbf{D}\ (\mathbf{c}\ -\ \mathbf{A}^{t}\ \mathbf{y})$ Tomando $\mathbf{z} = \left(\mathbf{c} - \mathbf{A}^{t}\mathbf{y}\right)$, resulta

$$\Delta\, \mathbf{x} = -\ \mathbf{D\ z} \qquad (3.11)$$

Adviértase que para obtener $\Delta\, \mathbf{x}$, se debe factorizar la matriz $\mathbf{A\,D\,A}^{t} = \mathbf{A\,(X^{k})^{2}\,A}^{t}$

Las variables $\mathbf{D}$, $\mathbf{y}$, $\mathbf{z}$, quedan definidas como sigue:

$$\boxed{\mathbf{D} = (\mathbf{X}^{k})^{2}}$$

El estimador de las variables duales $\mathbf{y} = \mathbf{A}^{T}\ (\mathbf{A\,D\,A}^{T})^{-1}\ \mathbf{A\ D\ c}$ y

$$\boxed{\mathbf{z} = \mathbf{c}\ -\ \mathbf{A}^{T}\mathbf{y}}$$

En el método Simplex por ser la matriz $\mathbf{A}$ de rango completo;

a. Se puede hacer una partición con sus columnas en $\mathbf{A} = [\mathbf{B,N}]$, siendo **B** una matriz básica de orden **m**, así también se descompone $\mathbf{x} = \begin{bmatrix}\mathbf{x_B}\\ \mathbf{x_N}\end{bmatrix}$, haciendo una partición equivalente $\mathbf{x}^{k} = \begin{bmatrix}\mathbf{X}^{t}_{\mathbf{B}}\\ \mathbf{X}^{t}_{\mathbf{N}}\end{bmatrix}$ y $\mathbf{c}^{t} = \begin{bmatrix}\mathbf{c}^{t}_{\mathbf{B}}\\ \mathbf{c}^{t}_{\mathbf{N}}\end{bmatrix}$ y asumiendo la ausencia de degeneración, se demostrará que el estimador de variables duales $\mathbf{y} = \mathbf{A}^{t}\ (\mathbf{A\,D\,A}^{t})^{-1}\ \mathbf{A\ D\ c}$, verifica la siguiente condición

$$\lim_{\mathbf{x_N}\to 0} \mathbf{y} = \mathbf{c_B}(\mathbf{B})^{-1}$$

El que equivale al vector de variables duales obtenido en el método símplex. En efecto

$$\mathbf{y} = (\mathbf{A}\,\mathbf{D}\,\mathbf{A}^{T})^{-1}\,\mathbf{A}\,\mathbf{D}\,\mathbf{c} = (\mathbf{A}(\mathbf{X}^{k})^{2}\,\mathbf{A}^{T})^{-1}\,\mathbf{A}\,(\mathbf{X}^{k})^{2}\,\mathbf{c}$$

$$\mathbf{y} = \left[(\mathbf{B},\mathbf{N})\begin{pmatrix} \mathbf{X}_B^k & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_N^k \end{pmatrix}^{2}\begin{pmatrix} \mathbf{B}^t \\ \mathbf{N}^t \end{pmatrix}\right]^{-1}\left[(\mathbf{B},\mathbf{N})\begin{pmatrix} \mathbf{X}_B^k & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_N^k \end{pmatrix}^{2}\begin{pmatrix} \mathbf{c}_B \\ \mathbf{c}_N \end{pmatrix}\right]$$

$$\mathbf{y} = (\mathbf{B}\mathbf{X}_B^2\,\mathbf{B}^t + \mathbf{N}\mathbf{X}_N^2\mathbf{N}^t)^{-1}(\mathbf{B}\mathbf{X}_B^2\mathbf{c}_B + \mathbf{N}^t\mathbf{X}_N^2\mathbf{c}_N)$$

Por lo expuesto en la observación previa, se tiene que $\mathbf{B}\,\mathbf{X}_B^2\,\mathbf{B}^T$ es no singular, tomando límite:

$$\lim_{x_N \to 0} \mathbf{y} = \lim_{x_N \to 0} \mathbf{y} = (\mathbf{B}\mathbf{X}_B^2\,\mathbf{B}^t + \mathbf{N}\mathbf{X}_N^2\mathbf{N}^t)^{-1}(\mathbf{B}\mathbf{X}_B^2\mathbf{c}_B + \mathbf{N}^t\mathbf{X}_N^2\mathbf{c}_N)$$

$$\lim_{x_N \to 0} \mathbf{y} = (\mathbf{B}\mathbf{X}_B^2\,\mathbf{B}^t)^{-1}\,(\mathbf{B}\mathbf{X}_B^2\,\mathbf{c}_B) = (\mathbf{B}^t)^{-1}(\mathbf{X}_B^2)^{-1}\mathbf{B}^{-1}\mathbf{B}\mathbf{X}_B^2\,\mathbf{c}_B = (\mathbf{B}^t)^{-1}(\mathbf{X}_B^2)^{-1}\,\mathbf{X}_B^2\,\mathbf{c}_B$$

$$\lim_{x_N \to 0} \mathbf{y} = (\mathbf{B}^t)^{-1}(\mathbf{X}_B^2)^{-1}\,\mathbf{X}_B^2\,\mathbf{c}_B = (\mathbf{B}^t)^{-1}\,\mathbf{c}_B$$

$$\lim_{x_N \to 0} \mathbf{y} = \mathbf{c}_B\,\mathbf{B}^{-1}$$

Observaciones previas 3.1

a. ¿Cuál es el rango de $(\mathbf{A}\,(\mathbf{X}^{k})^{2})$**?**

Respuesta. La matriz $\mathbf{X}^{k}$ es una matriz diagonal de orden $\mathbf{n}$, $r(\mathbf{X}^{k}) = \mathbf{n}$, cuyas componentes son estrictamente positivas(**es no degenerada**); también lo será $(\mathbf{X}^{k})^{2}$

Luego por propiedades del producto de matrices, se tiene que:

$r(\mathbf{A}\,(\mathbf{X}^{k})) = \text{Mínimo}\{r(\mathbf{A}), r(\mathbf{X}^{k})\} = \text{Mínimo}\{\mathbf{m}, \mathbf{n}\} = \mathbf{m}$. Por tanto $(\mathbf{A}\,(\mathbf{X}^{k}))$ es una matriz de rango completo y también lo será la matriz $(\mathbf{A}\,(\mathbf{X}^{k})^{2})$

b. ¿La matriz $(\mathbf{A}\,\mathbf{D}\,\mathbf{A}^{t}) = (\mathbf{A}\,(\mathbf{X}^{k})^{2}\,\mathbf{A}^{t})$; es no singular**?**

Respuesta. Por **hipótesis1**: La matriz $\mathbf{A}$ es de rango completo, $r(\mathbf{A}) = \mathbf{m}$ y por **definición 1.3.2.** $\mathbf{A} = [\mathbf{B},\mathbf{N}]$, $\mathbf{B}$ genera todo $\mathbb{R}^{m}$ y las columnas de la sub matriz $\mathbf{N}$, se pueden generar como combinación lineal de las columnas de $\mathbf{B}$. Así; el sub espacio generado por las columnas de la matriz $\mathbf{A}$; es el mismo que el generado por $\mathbf{B}$

Con razonamiento similar, se concluye que el sub espacio generado por la matriz $\mathbf{A}^t$ es el mismo que el generado por $\mathbf{B}^t$

Luego se concluye que $\mathbf{r(A\,(X^k)^2\,A^t) = r(B\,(X^k)^2\,B^t)}$ ■

Pero. ¿Existe $\mathbf{(B\,(X^k)^2\,B^t)^{-1}}$?

Respuesta :

$\mathbf{(B\,(X^k)^2)}$ es un matriz de $\mathbf{m \times n}$, con $\mathbf{m < n}$:

Luego $\mathbf{r(B\,(X^k)^2) = \text{Mínimo}\{r(B), r(X^k)^2\} = \text{Mínimo}\{m\ ,n\} = m}$

Es decir la matriz $\mathbf{B\,(X^k)^2}$ es de rango completo

Ahora se multiplicará la matriz $\mathbf{(B\,(X^k)^2)}$ por la sub matriz de $\mathbf{A}^t : \mathbf{B}^t$, la que es invertible y por propiedades del producto de matrices; que dice: "El resultado de multiplicar una matriz por una matriz invertible, es una matriz invertible", se concluye que existe $\mathbf{(B\,(X^k)^2\,B^t)^{-1}}$; y por lo expuesto anteriormente se concluye que también existirá $\mathbf{r(A\,(X^k)^2\,A^t)^{-1}}$ ■

En la figura 3.10.1 y 3.10.2 , se aprecia la dirección $\Delta\ \mathbf{x}$, obtenida en el programa original

La dirección del gradiente proyectado obtenida en (**3.10**). ¿Es factible y de no empeoramiento del valor actual del programa $\mathbf{P}$?. Es decir

a. ¿ $\mathbf{A\ \Delta x = 0}$? y

b. ¿ $\mathbf{c^t\ \Delta x\ A \leq 0}$?

Respuesta

a. Veamos si es factible

¿ $\Delta\ \mathbf{x} = -\ \mathbf{D}\ (\mathbf{I_n} - \mathbf{A^t\,(A\,D\,A^t)^{-1}\,A\,D})\,\mathbf{c}$ es factible?

$\mathbf{A}\ \Delta\ \mathbf{x} = -\ \mathbf{A\ D}\ (\mathbf{I_n} - \mathbf{A^t\,(A\,D\,A^t)^{-1}\,A\,D})\,\mathbf{c}$

$$-(\mathbf{A\,D\ I_n - A\,D\,A^T\,(A D A^t)^{-1}\,A\,D})\mathbf{c}$$

$$-(\mathbf{A\,D\ I_n} - (\mathbf{A\,D\,A^t})(\mathbf{A D A^t})^{-1}\,\mathbf{A\,D})\mathbf{c}$$

$$-(\mathbf{A\,D\ I_n} - \mathbf{A\,D})\mathbf{c}$$

$$-(\mathbf{A\,D} - \mathbf{A\,D})\mathbf{c} = \mathbf{0}$$

$\therefore\ \mathbf{A}\,\Delta\,\mathbf{x} = \mathbf{0}$ ■

Se sabe que $\Delta\mathbf{x} = -(\mathbf{X^k})\,\tilde{\mathbf{P}}\,\tilde{\mathbf{c}}$

Donde $\tilde{\mathbf{P}}$ es la matriz de proyección ortogonal sobre $\tilde{\mathbf{A}}$ con $\tilde{\mathbf{A}}$ y $\tilde{\mathbf{c}}$ como en 3.8.

Luego se tiene que

$$\mathbf{c^t}\Delta\mathbf{x} = -\mathbf{c^t}\,(\mathbf{X^k})\,\tilde{\mathbf{P}}\,\tilde{\mathbf{c}}$$

$$= -\mathbf{c^t}\,(\mathbf{X^k})\,\tilde{\mathbf{P}}\,(\mathbf{X^k})\,\mathbf{c} \qquad \textbf{(3.11.1)}$$

$= -\mathbf{c^t}\,(\mathbf{X^k})\,\tilde{\mathbf{P}}^{\mathbf{t}}\,\tilde{\mathbf{P}}\,(\mathbf{X^k})\,\mathbf{c}$ Por ser $\tilde{\mathbf{P}}$: simétrica e Idempotente

$$= -\mathbf{c^t}\,(\mathbf{X^k})\,\tilde{\mathbf{P}}^{\mathbf{T}}\,\tilde{\mathbf{P}}\,(\mathbf{X^k})\,\mathbf{c} = -(\tilde{\mathbf{P}}\mathbf{X^k c})^{\mathbf{t}}\,(\tilde{\mathbf{P}}\mathbf{X^k c}) = -\left\|(\tilde{\mathbf{P}}\mathbf{X^k c})\right\|^{\mathbf{2}}$$

De donde se concluye que el nuevo valor del programa en el peor de los casos, queda como está; esto es

$$\mathbf{c^t}\Delta\mathbf{x} = -\left\|(\tilde{\mathbf{P}}\mathbf{X^k c})\right\|^{\mathbf{2}} \leq 0$$

3.7 Continuidad de la función

Teorema

Si el programa $\mathbf{P}$ tiene solución no degenerada y la matriz de restricciones $\mathbf{A}$ es de rango completo; entonces $\mathbf{A}\,(\mathbf{X^k})^{\mathbf{2}}\,\mathbf{A^T}$ es una matriz no singular y los términos $\mathbf{y}$, $\mathbf{z}$ definidos en **(3.11)**, son una función continua de $\mathbf{x^k}$

Demostración

Pues todas las entradas son funciones continuas .

En cada iteración del algoritmo, las tres etapas se realizan implícitamente y sólo se calcula la dirección $\Delta\,\mathbf{x}$, según la relación 3.10**.**

Esta será la dirección a utilizar en el proceso iterativo, expuesto en 3.1

3.8 Obtención de la longitud de paso

En cada iteración el objetivo es moverse desde $\mathbf{x}^{\mathbf{k}}$, en la dirección de $\Delta\,\mathbf{x}$;(como en el método Simplex a lo largo de d) tanto como sea posible; sin violar las condiciones de no negatividad de las variables; esto es debe garantizarse en cada iteración que:

$$\boxed{\mathbf{x}^{\mathbf{k+1}} = \mathbf{x}^{\mathbf{k}} + \alpha\,\Delta\,\mathbf{x} \geq \mathbf{0}}$$

Como $\alpha > \mathbf{0}$ y las componentes $\Delta\,\mathbf{x}_{\mathbf{i}} \geq \mathbf{0}$ no pueden conducir a que $\mathbf{x}_{\mathbf{i}}$ **=** 0. Luego se debe estar preocupado sólo por las componentes que sean negativas $\Delta\,\mathbf{x}_{\mathbf{i}} < 0$

También como el tamaño máximo del paso; el que se denotará como $\bar{\alpha}$; se obtendrá mediante la relación.

$$\bar{\alpha} = \underset{\mathbf{i}}{\mathbf{mínimo}} \left\{ -\frac{\mathbf{x}_{\mathbf{i}}^{\mathbf{k}}}{\Delta\mathbf{x}_{\mathbf{i}}} : \Delta\mathbf{x}_{\mathbf{i}} < \mathbf{0} \right\} \qquad \textbf{(3.12)}$$

Nótese que con esta definición de α, la componente i-ésima de $\mathbf{x}^{\mathbf{k+1}}$; siendo i el índice de la componente, asociada al mínimo $-\frac{\mathbf{x}_{\mathbf{i}}^{\mathbf{k}}}{\Delta\mathbf{x}_{\mathbf{i}}}$, con $\Delta\mathbf{x}_{\mathbf{i}} < \mathbf{0}$ se anularía; más aún $\mathbf{x}^{\mathbf{k+1}}$; dejaría de ser interior y consecuentemente la matriz $\mathbf{x}^{\mathbf{k+1}}$ sería singular.

Para evitar la ocurrencia de esta situación, se debe reducir la longitud el paso, mediante la siguiente operación

$$\boldsymbol{\alpha = \rho\,\bar{\alpha} =} = \rho\ \underset{\mathbf{i}}{\mathbf{mínimo}} \left\{ -\frac{\mathbf{x}_{\mathbf{i}}^{\mathbf{k}}}{\Delta\mathbf{x}_{\mathbf{i}}} : \Delta\mathbf{x}_{\mathbf{i}} \leq \mathbf{0} \right\} \qquad \textbf{(3.13)}$$

Siendo **0.95 ≤ ρ ≤ 0.9995**

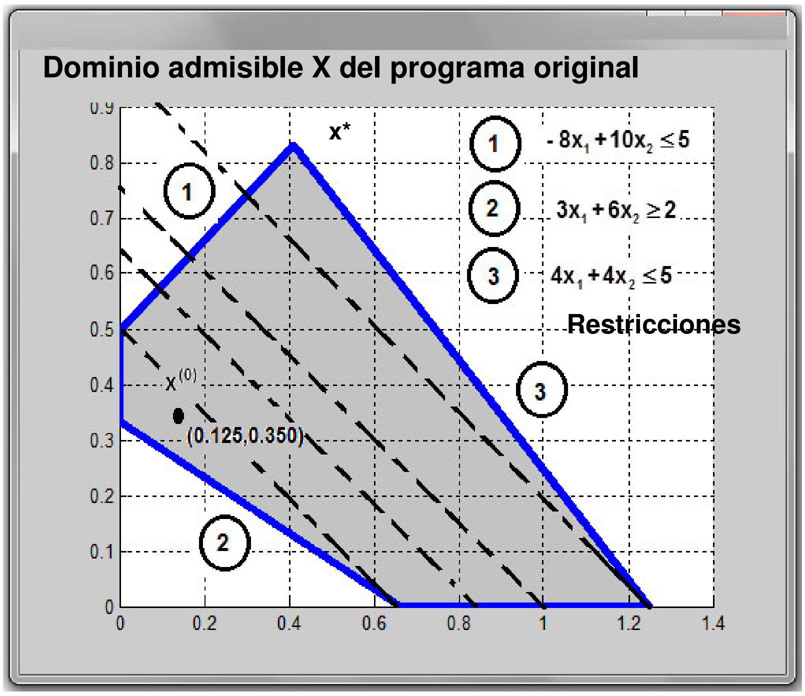
Esta forma de elegir el tamaño de la longitud de paso α; es llamado "Método de escalado de paso largo".

Considérese el programa $\mathbf{R}_1$

$$\begin{array}{llllll}
\textbf{Min f(x)}: & -\ 3x_1 & -\ 4x_2 & & & \\
\text{St} & -8x_1 & +\ 10x_2 & \leq & 5 & \\
 & 3x_1 & +\ 6x_2 & \geq & 2 & \cdots\ 3 \\
 & 4x_1 & +\ 4x_2 & \leq & 5 & \\
 & x_1 \geq 0 & x_2 \geq 0 & & &
\end{array}$$

Resolviendo gráficamente, se tiene:

Gráfica n º 23: Dominio admisible y punto inicial de búsqueda x^k



La solución óptima se encuentra en el punto

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 5/12 \\ 10/12 \end{bmatrix}$$

Curvas de nivel

Fuente propia

Escribiendo $\mathbf{R}_1$ en forma estándar:

$$\begin{array}{lllllllll}
\textbf{Min f(x)} = & -\ 2x_1 & -\ 2x_2 & +\ 0x_3 & +\ 0x_4 & +\ 0x_5 & & & \\
\text{St} & -\ 8x_1 & +\ 10x_2 & +\ 1x_3 & +\ 0x_4 & +\ 0x_5 & = & 5 & \cdots\ 3.1 \\
 & 3x_1 & +\ 6x_2 & +\ 0x_3 & -\ 1x_4 & +\ 0x_5 & = & 2 &
\end{array}$$

$$4x_1 \quad + 4x_2 \quad + 0x_3 \quad + 0x_4 \quad + 1x_5 \quad = \quad 5$$

$$x_1 \geq 0 \quad x_2 \geq 0 \quad x_3 \geq 0 \quad x_4 \geq 0 \quad x_5 \geq 0$$

Siendo

$$\mathbf{A} = \begin{bmatrix} -8 & 10 & 1 & 0 & 0 \\ 3 & 6 & 0 & -1 & 0 \\ 4 & 4 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{c}^t = \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 2 \\ 5 \end{bmatrix}$$

En este ejemplo se tomará como punto inicial

$\mathbf{x}^k = [0.5, \ 0.5, \ 4.0, \ 2.5 \ \ 1.0]^t$, el que es interior y factible al programa $\mathbf{R}_1$

En este punto el valor del programa es $\mathbf{cx}^k = -3.5$

Obténgase $\Delta\mathbf{x} = -\mathbf{D}\,\mathbf{z}$, con $\mathbf{z} = (\mathbf{c}^t - \mathbf{A}^t\mathbf{y})$ e $\mathbf{y} = \left(\mathbf{ADA}^t\right)^{-1}\mathbf{ADc}^t$

Haciendo las operaciones correspondientes, utilizando las relaciones **(3.9)**, **(3.10)** y **(3.11)**

$$\mathbf{D} = \left(\mathbf{X}^k\right)^2 = \begin{bmatrix} 0.50 & 0 & 0 & 0 & 0 \\ 0 & 0.50 & 0 & 0 & 0 \\ 0 & 0 & 4.00 & 0 & 0 \\ 0 & 0 & 0 & 2.50 & 0 \\ 0 & 0 & 0 & 0 & 1.00 \end{bmatrix}^2 =$$

$$\begin{bmatrix} 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 16.00 & 0 & 0 \\ 0 & 0 & 0 & 6.25 & 0 \\ 0 & 0 & 0 & 0 & 1.00 \end{bmatrix}^2$$

$$\mathbf{AD} = \begin{bmatrix} -8 & 10 & 1 & 0 & 0 \\ 3 & 6 & 0 & -1 & 0 \\ 4 & 4 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 16.00 & 0 & 0 \\ 0 & 0 & 0 & 6.25 & 0 \\ 0 & 0 & 0 & 0 & 1.00 \end{bmatrix}^2$$

$$\mathbf{AD} = \begin{bmatrix} -2.00000 & 2.50000 & 16.00000 & 0.00000 & 0.00000 \\ 0.75000 & 1.50000 & 0.00000 & -6.25000 & 0.00000 \\ 1.00000 & 1.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

$$\mathbf{ADA^t} = \begin{bmatrix} -2.00000 & 2.50000 & 16.00000 & 0.00000 & 0.00000 \\ 0.75000 & 1.50000 & 0.00000 & -6.25000 & 0.00000 \\ 1.00000 & 1.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

$$\begin{bmatrix} -8 & 3 & 4 \\ 10 & 6 & 4 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{ADA^t} = \begin{bmatrix} 21.00000 & 1.00000 & 0.50000 \\ 1.00000 & 12.50000 & 1.75000 \\ 0.50000 & 1.75000 & 1.50000 \end{bmatrix}$$

$\mathbf{ADA^t}$ Es una matriz de rango completo ($\mathbf{m}$), ya que $\mathbf{A}$ lo es. También la matriz $\mathbf{X^k}$ posee no menos de $\mathbf{m}$ componentes positivas; luego la matriz $\mathbf{ADA^t}$ es invertible

$$\left[\mathbf{ADA^t}\right]^{-1} = \begin{bmatrix} 0.01969 & -0.01621 & 0.01184 \\ -0.01621 & 0.13100 & -0.1274 \\ 0.01184 & -0.1274 & 0.23588 \end{bmatrix}$$

$$\mathbf{ADc^t} = \begin{bmatrix} -2.00000 & 2.50000 & 16.00000 & 0.00000 & 0.00000 \\ 0.75000 & 1.50000 & 0.00000 & -6.25000 & 0.00000 \\ 1.00000 & 1.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix} \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} -4.0000 \\ -8.2500 \\ -7.0000 \end{bmatrix}$$

$$\mathbf{y} = \left(\mathbf{ADA^t}\right)^{-1} \mathbf{ADc^t} = \begin{bmatrix} 0.01969 & -0.01621 & 0.01184 \\ -0.01621 & 0.13100 & -0.1274 \\ 0.01184 & -0.1274 & 0.23588 \end{bmatrix} \begin{bmatrix} -4.0000 \\ -8.2500 \\ -7.0000 \end{bmatrix} =$$

$$\begin{bmatrix} -0.02786 \\ -0.12412 \\ -0.64747 \end{bmatrix}$$

También $\mathbf{b^t y = -3.62489}$; luego la brecha dual será:

$$\frac{\left|\mathbf{(brecha\ dual)^k}\right|}{\mathbf{1+\left|c^T x^k\right|}} = \frac{\mathbf{|-3.5+3.62489|}}{\mathbf{1+|-3.5|}} = \mathbf{0.02775}$$, lo que significa que $\mathbf{x^k} =$ $\mathbf{[0.5,\ \ 0.5,\ \ 4.0,\ \ 2.5\ \ 1.0]^t}$, aún no es óptimo**;** luego se continuará iterando.

$$\mathbf{A^t y} = \begin{bmatrix} -8 & 3 & 4 \\ 10 & 6 & 4 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.02786 \\ -0.12412 \\ -0.64747 \end{bmatrix} = \begin{bmatrix} -2.73935 \\ -3.61318 \\ -0.02786 \\ 0.12412 \\ -0.64747 \end{bmatrix}$$

$$\mathbf{z = c^t - A^t y} = \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -2.73935 \\ -3.61318 \\ -0.02786 \\ 0.12412 \\ -0.64747 \end{bmatrix} = \begin{bmatrix} -0.26065 \\ -0.38682 \\ 0.02786 \\ -0.12412 \\ 0.64747 \end{bmatrix}$$

$$\mathbf{\Delta x = -D z} = -\begin{bmatrix} 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 16.00 & 0 & 0 \\ 0 & 0 & 0 & 6.25 & 0 \\ 0 & 0 & 0 & 0 & 1.00 \end{bmatrix}^2 \begin{bmatrix} -0.26065 \\ -0.38682 \\ 0.02786 \\ -0.12412 \\ 0.64747 \end{bmatrix}$$

$$\mathbf{\Delta x} = \begin{bmatrix} 0.06516 \\ 0.09671 \\ -0.44576 \\ 0.77572 \\ -0.64747 \end{bmatrix}$$

Al utilizar la relación **(3.12)**: $\bar{\alpha} = \min_{i} \left\{ -\frac{x_i^{(0)}}{\Delta x_i} : \Delta x_i < 0 \right\}$

$$\bar{\alpha} = \mathbf{mín}\left\{ -,\quad -,\quad -\frac{4.0}{-0.44576},\ -\quad -\frac{1.0}{-0.64747} \right\} =$$
$$\mathbf{\{8.97344\ ,\ 1.54447\} = \ 1.54447}$$

Y tomando $\rho = \mathbf{0.95}$, se tendrá $\alpha = \rho\bar{\alpha} = \mathbf{1.46725}$

De esta forma se obtiene el nuevo punto $\mathbf{x^1}$

$$x^1 = x^0 + \alpha \Delta x = \begin{bmatrix} 0.50 \\ 0.50 \\ 4.00 \\ 2.50 \\ 1.00 \end{bmatrix} + 1.46725 \begin{bmatrix} 0.06516 \\ 0.09671 \\ -0.44576 \\ 0.77572 \\ -0.64747 \end{bmatrix} = \begin{bmatrix} 0.50975 \\ 0.70150 \\ 2.06301 \\ 3.73825 \\ 0.15500 \end{bmatrix}$$

$$x^1 = \begin{bmatrix} 0.59561 \\ 0.64189 \\ 3.34596 \\ 3.63817 \\ 0.05000 \end{bmatrix}$$

El nuevo valor de la función objetivo será la función

$$cx^1 = -4.3544 < cx^k = -3.5$$

Al seguir iterando, se obtiene $b^t y = -4.52982$; luego la brecha dual será:

$\frac{|(\text{brecha dual})^k|}{1+|c^T x^k|} = \frac{|-4.3544 + 4.52982|}{1+|-4.3544|} = 0.015682$, lo que significa que

$x^k = [0.59561,\ 0.64189,\ 3.34596,\ 3.63817\ \ 0.050]^t$, aún no es óptimo; luego se continuará iterando.

El presente método trata de avanzar, tanto como se pueda, en la dirección de movimiento calculado; es por este motivo que se le llama "Método de Escalado Afín de Paso Largo"

El método de Escalado Afín de "Paso corto", como el de "Paso Largo", utiliza el cambio de variable $x = (X^k)^{-1} x$, presentada anteriormente; obteniendo el siguiente problema, en el espacio de variables x

$$\begin{aligned} P\ \underset{x}{\text{Mín}}\ Z' &= c^T x \\ \text{S.a. } A x &= b,\ x \geq 0 \end{aligned} \qquad (3.14)$$

El punto $x^k > 0$, correspondiente al programa original P_1; se transforma en $e^t = (1, 1, \ldots, 1)^t$

Considérese seguidamente la esfera $\mathbf{B}(\mathbf{x},\rho)$

$$\mathbf{B}(\mathbf{x},\rho) = \{\, \mathbf{x} : \|\mathbf{x}-\mathbf{e}\| \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\} \qquad \textbf{(3.15)}$$

Dado que $\mathbf{e}$ pertenece al interior del dominio admisible de $\mathbf{P_1}$, se tiene que la esfera $\mathbf{B}(\mathbf{x},\rho)$, estará también en el dominio admisible del programa $\mathbf{P_1}$, cuando $\rho \le \mathbf{1}$

El conjunto de puntos $\mathbf{B}(\mathbf{x},\rho)$, puede ser transformado al espacio de las variables originales $\mathbf{x}$, deshaciendo el escalado mediante las relaciones **(3.7)** **(3.8)**

$$\mathbf{B}(\mathbf{x},\rho) = \{\, \mathbf{x} : \|(\mathbf{X}^{k})^{-1}\mathbf{x}-\mathbf{e}\| \le \rho ,\ \mathbf{A}\,\mathbf{X}^{k}\,(\mathbf{X}^{k})^{-1}\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\}$$

$$= \{\, \mathbf{x} : \|(\mathbf{X}^{k})^{-1}\mathbf{x}-\mathbf{e}\| \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\}$$

$$= \{\, \mathbf{x} : \|(\mathbf{X}^{k})^{-1}\mathbf{x}-(\mathbf{X}^{k})^{-1}\mathbf{x}^{k}\| \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\}$$

$$= \{\, \mathbf{x} : \left[(\mathbf{X}^{k})^{-1}(\mathbf{x}-\mathbf{x}^{k})\right]^{t}\left[(\mathbf{X}^{k})^{-1}(\mathbf{x}-\mathbf{x}^{k})\right] \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\} \qquad \textbf{(3.16)}$$

$$= \{\, \mathbf{x} : (\mathbf{x}-\mathbf{x}^{k})^{t}(\mathbf{X}^{k})^{-1}(\mathbf{X}^{k})^{-1}(\mathbf{x}-\mathbf{x}^{k}) \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\}$$

$$= \{\, \mathbf{x} : (\mathbf{x}-\mathbf{x}^{k})^{t}\left[(\mathbf{X}^{k})\ (\mathbf{X}^{k})^{t}\right]^{-1}(\mathbf{x}-\mathbf{x}^{k}) \le \rho ,\ \mathbf{A}\,\mathbf{x} = \mathbf{b} \,\} \quad = \mathbf{E}(\mathbf{x}^{k},\rho)$$

Esta última expresión representa la intersección de la elipse centrada en $\mathbf{x}^{k}$, asociada a la matriz $\mathbf{Q} = (\mathbf{X}^{k})(\mathbf{X}^{k})^{t}$, con el dominio admisible

$\mathbf{A}\,\mathbf{x} = \mathbf{b}$

Es fácil notar que la esfera $\mathbf{B}(\tilde{\mathbf{x}},\rho)$, en el espacio $\tilde{\mathbf{x}}$, se ha convertido en un elipsoide $\mathbf{E}(\mathbf{x}^{k},\rho = \mathbf{1})$, en el en el espacio original $\mathbf{x}$, debido a la matriz de escalado $(\mathbf{X}^{k})^{-1}$

También como $\mathbf{B}(\mathbf{x}, \rho = \mathbf{1})$, estaba incluido en el dominio admisible del programa $\mathbf{P_1}$, se verifica que el elipsoide $\mathbf{E}(\mathbf{x^k}, \rho = \mathbf{1})$, queda también incluido en el dominio admisible de $\mathbf{P_1}$, condición esta que la garantiza el siguiente Teorema:

Teorema 3.8.1

El conjunto de puntos definido por el elipsoide $\mathbf{E}(\mathbf{x^k}, \rho) = \mathbf{B}(\tilde{\mathbf{x}}, \rho)$, pertenecen al dominio admisible del programa $\mathbf{P_1}$, si $\rho \leq \mathbf{1}$, es decir

$$\mathbf{B}(\mathbf{x}, \rho \leq \mathbf{1}) \subseteq \{\, \mathbf{x} : \mathbf{A}\,\mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0} \,\}$$

Y si $\rho < \mathbf{1}$, estos puntos son interiores

Demostración

Se sabe por definición, que todos los puntos $\mathbf{x}$ de: $\mathbf{E}(\mathbf{x^k}, \rho)$, satisfacen $\mathbf{A}\,\mathbf{x} = \mathbf{b}$; se debe ver que también $\mathbf{x} \geq \mathbf{0}$

Se sabe que si $\mathbf{x} \in \mathbf{B}(\mathbf{x}, \rho)$, entonces $\left\| (\mathbf{X^k})^{-1}\,\mathbf{x} - \mathbf{e} \right\| \leq \rho$

Pero para cualquier $\mathbf{u}$: $\mathbf{u_i} \leq \|\mathbf{u}\|$, se tiene

$$\left| \frac{\mathbf{x_i}}{\mathbf{x_i^k}} - \mathbf{1} \right| \leq \left\| (\mathbf{X^k})^{-1}\,\mathbf{x} - \mathbf{e} \right\| \leq \rho \leq \mathbf{1}$$

Si se multiplica la relación anterior por $\mathbf{x_i^k}$, se puede escribir como sigue:

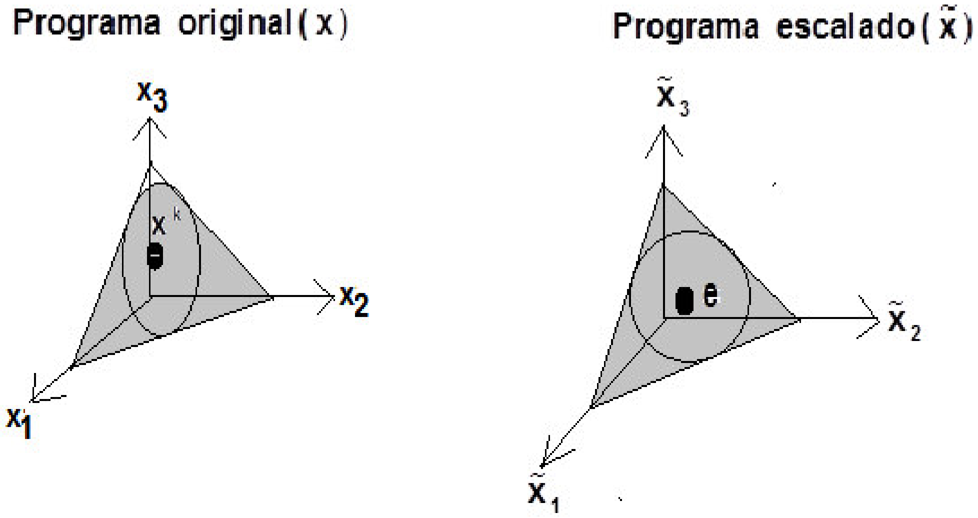
$$\left| \mathbf{x_i} - \mathbf{x_i^k} \right| \leq \rho\, \mathbf{x_i^k} \leq \mathbf{x_i^k}$$

De la que se obtienen el resultado buscado: $\mathbf{x_i} \geq \mathbf{0}$

Si $\rho < \mathbf{1}$; las relaciones anteriores, se verifican como desigualdades estrictas, teniéndose $\mathbf{x_i^k} > 0$ y por tanto interior.

La siguiente gráfica, muestra la relación entre la esfera y la elipse.

Gráfica n° 24: Relación entre dominios admisibles de ambos programas



Fuente propia

Por otro lado existe otra manera de elegir el tamaño de la longitud de paso α, que fue deducido en la sección (**3.4**). La única diferencia es que el proceso, debe ser aplicado al programa $\tilde{\mathbf{P}}_1$, de dominio admisible en $\tilde{\mathbf{x}}$. El proceso se basa en lo siguiente:

Como En sus respectivos espacios para $\rho = \mathbf{1}$, se observa como el elipsoide es factible para el programa $\mathbf{P_1}$, (tal como se demostró en el Teorema 5.1, es decir todos los puntos del elipsoide están en el dominio admisible del programa $\mathbf{P_1}$; el método iterativo de escalado afín, itera sustituyendo la restricción $\mathbf{x} \geq \mathbf{0}$, por la de pertenencia a dicho elipsoide, de la siguiente manera:

En el problema escalado, envés del elipsoide considera una esfera; planteando el siguiente programa de minimización, para obtener $\Delta\tilde{\mathbf{x}}$

$$\begin{aligned} &\underset{\Delta \mathbf{x}}{\mathbf{Mín}} \quad \mathbf{c}^{\mathbf{T}}\mathbf{x} \\ &\mathbf{St} \quad \mathbf{A}(\mathbf{x} + \Delta\tilde{\mathbf{x}}) = \mathbf{b} \\ &\quad \|\Delta\tilde{\mathbf{x}}\|^{\mathbf{2}} = \rho^2 \end{aligned} \qquad \mathbf{(3.17)}$$

Adviértase el dominio admisible de este problema, coincide con la frontera de la esfera $\mathbf{B}(\mathbf{x}, \rho)$, dado que $\tilde{\mathbf{x}}^{\mathbf{k}} = \mathbf{e}$

Si se considera ρ **≤ 1,** esta región factible, está además contenida en el dominio admisible del programa $\tilde{\mathbf{P}}_1$, como se ha expuesto anteriormente

La segunda restricción del programa, se ha podido haber escrito como $\|\Delta\tilde{\mathbf{x}}\|^2 \leq \rho^2$, logrando de esta manera como dominio admisible del programa **(3.17)** la esfera de centro $\mathbf{x}$ y radio de longitud ρ; $\mathbf{B}(\mathbf{x},\rho)$.

Sin embargo el mínimo del programa se hallará en los puntos de frontera, puesto que la función objetivo es líneal; razón por lo que la segunda restricción, se formula en forma de igualdad.

En el presente programa se está buscando la dirección $\Delta\tilde{\mathbf{x}}$, la que proporciona la máxima disminución de la función objetivo, alrededor del punto $\tilde{\mathbf{x}}^k = \mathbf{e}$

Se puede apreciar que este equivale al programa **(3.4)** y se sabe por **(3.6)**, que su solución viene dada por

$$\boxed{\Delta\tilde{\mathbf{x}} = -\frac{\rho}{\|\tilde{\mathbf{P}}\tilde{\mathbf{c}}\|}\tilde{\mathbf{P}}\,\tilde{\mathbf{c}}}$$

Entonces el programa **(3.17)**, se podría haber formulado en el espacio de variables $\mathbf{x}$; deshaciendo así el escalado y obteniendo:

$$\begin{aligned} &\underset{\Delta \mathbf{x}}{\mathbf{Mín}}\ \mathbf{c}^T(\mathbf{x}^k + \Delta\,\mathbf{x}) \\ &\mathbf{St}\ \ \mathbf{A}(\mathbf{x}^k + \Delta\,\mathbf{x}) = \mathbf{b} \\ &\|(\mathbf{X}^k)^{-1}\Delta\,\mathbf{x}\|^2 = \rho^2 \end{aligned} \qquad \mathbf{(3.18)}$$

El dominio admisible corresponde ahora a la frontera del elipsoide $\mathbf{B}(\mathbf{x},\rho) = \mathbf{E}(\mathbf{x}^k,\rho)$

Y se sabe por el teorema anterior que para ρ **≤ 1,** este elipsoide está contenido en el dominio admisible del programa original ($\mathbf{P}_1$) y que está en su interior si ρ **< 1**

Así la solución del programa **(3.18)**, se obtiene a partir de **(3.17)**; deshaciendo el escalado:

$$\Delta \mathbf{x} = \mathbf{x}^{k} \, \Delta\tilde{\mathbf{x}} = \frac{\rho}{\|\tilde{\mathbf{P}}\tilde{\mathbf{c}}\|} \mathbf{X}^{k} \tilde{\mathbf{P}} \, \tilde{\mathbf{c}}$$

Luego, el nuevo punto será:

$$\mathbf{x}^{k+1} = \mathbf{x}^{k} - \frac{\rho}{\|\tilde{\mathbf{P}}\tilde{\mathbf{c}}\|} \mathbf{X}^{k} \tilde{\mathbf{P}} \, \tilde{\mathbf{c}}$$

3.9 Condición de parada

Se trata de establecer un criterio, para determinar si el punto actual $\mathbf{x}^{k}$, se encuentra lo suficientemente cerca al óptimo $\mathbf{x}^{*}$, para la cual se utilizará el teorema de la dualidad fuerte

Si es posible obtener una estimación de la solución dual $\mathbf{y}$, asociada a $\mathbf{x}^{k}$; entonces será posible detener el proceso iterativo cuando la **brecha dual**, sea menor que un valor de tolerancia $\in$, que pertenece a un intervalo $\left[\mathbf{10}^{-6}, \mathbf{10}^{-8}\right]$

La condición para detener la búsqueda será la siguiente:

$$\frac{\left|(\text{brecha dual})^{k}\right|}{1+\left|\mathbf{c}^{T}\mathbf{x}^{k}\right|} = \frac{\left|\mathbf{c}^{T}\mathbf{x}^{k} - \mathbf{b}^{T}\mathbf{y}\right|}{1+\left|\mathbf{c}^{T}\mathbf{x}^{k}\right|} \leq \in \qquad \textbf{(3.19)}$$

En la expresión se relativiza la brecha dual, agregando **1** al denominador, para evitar problemas de precisión, cuando el valor óptimo del programa, esté cerca a **0**

Por otro lado se utilizará el teorema de la holgura complementaria débil, para obtener un estimador apropiado de la variable dual y

Se sabe que las holguras duales verifican $\mathbf{z} = \mathbf{c} - \mathbf{A}^{T}\mathbf{y}$, con $\mathbf{z} \geq \mathbf{0}$ y que en el óptimo del programa primal y el dual ocurre $\mathbf{x}_{i}\,\mathbf{z}_{i} = \mathbf{0}$, para i = 1, 2, . , **n**

Dejando a parte la condición de no negatividad $\mathbf{x} \geq \mathbf{0}$, es posible estimar $\mathbf{y}$, mediante la solución del siguiente programa

$$\underset{x,\,y}{\text{Mín}}\ \frac{1}{2}\left\|X^k z\right\|^2$$

$$\text{S.a.}\ \ z = c - A^T y$$

El cual puede ser resuelto introduciendo la función lagrangiana:

$$L(z, y, u) = \frac{1}{2}\left\|X^k z\right\|^2 - u^T(z - c - A^T y)$$

Donde $\mathbf{u} \in \mathbb{R}^n$, son los multiplicadores de lagrange. Al derivar $\mathbf{L}$ con respecto a $\mathbf{z}$, $\mathbf{y}$, $\mathbf{u}$ e igualar estas derivadas a cero se obtiene las condiciones que debe satisfacer el punto solución

$$\begin{aligned} (X^k)^2 z - u &= 0 \\ A u &= 0 \\ z - c + A^T y &= 0 \end{aligned}$$

Al multiplicar ambos miembros de la primera relación por $\mathbf{A}$, se obtiene:

$A\,(X^k)^2 z = A\,u$; pero $A\,u = 0$; debido a la segunda relación, luego $A\,(X^k)^2 z = 0$

Al multiplicar la tercera restricción por $A\,(X^k)^2$, se obtiene:

$0 = A\,(X^k)^2 z - A\,(X^k)^2 c + A\,(X^k)^2\ A^T y =$

$0 - A\,(X^k)^2 c + A\,(X^k)^2\ A^T y$

ó $0 = -A\,(X^k)^2 c + A\,(X^k)^2\ A^T y$ ó

$A\,(X^k)^2 c = A\,(X^k)^2\ A^T y$

De donde: resulta

$$y = \left(A\,(X^k)^2 A^T\right)^{-1} A\,(X^k)^2 c \qquad \mathbf{(3.20)}$$

Esta relación produce la estimación de $\mathbf{y}$

Análogamente $\mathbf{z = c - A^T y}$, es el vector que corresponderá a la estimación de las holguras duales.

Utilizando la notación $\mathbf{D} = \mathbf{(X^k)^2}$, realizada anteriormente, la estimación de $\mathbf{y}$, se puede escribir como:

$\mathbf{y} = \mathbf{(ADA^T)^{-1} A D c}$

Esta última relación coincide con el vector $\mathbf{y}$ definido en (**3.11**), que interviene para calcular la dirección de movimiento $\mathbf{\Delta x}$, lo que significa que no será necesario hacer otra operación, para obtener este estimador.

También se debe notar y valorar que la expresión (**3.20**), coincide con el vector de multiplicadores; de los programas de **Minimización** (**3.4**), (**3.17**) y (**3.18**), los que proporcionaban la dirección de movimiento hacia cada $\mathbf{x^k}$

Luego la sucesión de estimaciones duales; no es más que la sucesión de multiplicadores, de las ecuaciones $\mathbf{A\,x} = \mathbf{b}$ de estos programas.

Bajo la hipótesis que el programa es no degenerado y las, observaciones previas 3.1 – b, garantizan que $\mathbf{A\,(X^k)^2\;A^T}$ es no singular, lo que significa que el cálculo de los estimadores duales $\mathbf{y}$, está bien definido

Si el programa fuese degenerado la matriz $\mathbf{A\,(X^k)^2\;A^T}$ sería casi no singular y si la solución óptima $\mathbf{x^*}$ fuera degenerada la búsqueda se complicaría a medida que se continúe iterando

Bajo esta situación se puede utilizar un criterio, que consiste en detener la búsqueda, cuando la mejora relativa de la función sea pequeña; esto es:

$$\frac{\left|\mathbf{c^T x^k - c^T x^{k+1}}\right|}{\mathbf{1 + \left|c^T x^k\right|}} \leq \in \qquad (\mathbf{3.21})$$

3.10 Obtención de una solución factible inicial

Ya se sabe como iterar y cuándo parar; pero se ha supuesto que se parte de un punto inicial interior y factible $\mathbf{x}^0$. No es trivial hallar tal punto para el programa $(\mathbf{P_1})$, luego seguidamente se presentará una técnica, el que se denominada **técnica de la M grande** para superar esta dificultad; la que equivale a la técnica del mismo nombre en el método símplex.

Considérese un punto $\mathbf{x}^0 > \mathbf{0}$ cualquiera por ejemplo $\mathbf{e} = (1, 1, \ldots, 1)^T$. Este punto elegido aleatoriamente, es probable que sea no admisible para el programa $(\mathbf{P_1})$, esto es no verifique $\mathbf{A\,x}^0 = \mathbf{b}$; entonces se puede obtener su vector de infactibilidades $\mathbf{r} \in \mathbb{R}^m$:

$$\mathbf{r} = \mathbf{b} - \mathbf{A\,x}^0$$

Si se utiliza $\mathbf{M} \in \mathbb{R}$, $\mathbf{M} > \mathbf{0}$ y relativamente grande, defínase el siguiente programa alternativo a $(\mathbf{P_1})$

$\underset{x}{\text{Mín}} \begin{bmatrix} \mathbf{c}^T, \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_{n+1} \end{bmatrix}$ $\text{S.a.} \begin{bmatrix} \mathbf{A} & , \mathbf{r} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_{n+1} \end{bmatrix} = \mathbf{b}$ $\begin{bmatrix} \mathbf{x} \\ \mathbf{x}_{n+1} \end{bmatrix} \geq \mathbf{0}$	(3.22)

Este programa tiene una variable mas que el programa $(\mathbf{P_1})$; gracias a esta condición, es posible obtener un punto factible e interior de modo muy sencillo, esto es

$$\bar{x}^0 = \begin{bmatrix} x_1^0 \\ x_2^0 \\ \cdot \\ \cdot \\ x_n^0 \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1} \qquad (3.23)$$

Debido a que $\mathbf{x}^0$ es un punto interior, también lo es $\bar{\mathbf{x}}^0$. Y al utilizar la definición del vector $\mathbf{r}$, se comprueba que $\begin{bmatrix} \mathbf{A} \ , \mathbf{r} \end{bmatrix} \begin{bmatrix} \mathbf{x}^0 \\ \mathbf{1} \end{bmatrix} = \mathbf{A}\,\mathbf{x}^0 + \mathbf{r} =$ $\mathbf{A}\,\mathbf{x}^0 + (\mathbf{b} - \mathbf{A}\,\mathbf{x}^0) = \mathbf{b}$

Lo que prueba que $\bar{\mathbf{x}}^0$, está en el dominio admisible del programa (**3.22**)

Luego se tendrá que esperar que en el óptimo del programa (**3.22**), debe verificarse $\mathbf{x}^*_{n+1}$; por ser **M** grande

Cuando no ocurra así, se concluirá que el programa original $(\mathbf{P_1})$; es no factible, ya que no se ha podido volver cero la componente $\mathbf{x}_{n+1}$, asociada al vector de imbatibilidades **r**

Limitaciones de la técnica M

1. Se debe elegir un valor adecuado para **M**, que no sea ni muy pequeño ni muy grande.

Si fuera demasiada grande, se podría volver inestable el proceso iterativo de búsqueda y si fuera muy pequeña; se penalizaría poco la variable asociada al vector de infactibilidades **r,** lo que implicaría obtener soluciones $\mathbf{x}^*_{n+1} \neq \mathbf{0}$; sin que el programa sea infactible

En la cultura sobre esta temática, se han propuesto técnicas de adaptación de **M,** a medida que avanza la búsqueda del óptimo

2. El vector **r** introducido en la matriz de restricciones de (**3.19**); por lo general será densa; lo que provoca que la matriz $\mathbf{A\,D\,A^T}$ sea también densa.

La factorización de $\mathbf{A\,D\,A^T}$ en cada iteración, desde un punto de visto computacional; es el paso más costoso a realizar. El hecho que esta matriz sea densa incrementa ostensiblemente el tiempo de ejecución, en especial en los programas de gran dimensión.

Destaca entre sus ventajas, la que proporciona el valor óptimo del programa, al solucionar un programa único, en lugar de realizarlo en dos fases.

3.10.1 Convergencia del algoritmo de escalado afín

El análisis de la convergencia del presente algoritmo, es más complejo que el de otros métodos de punto interior.

Lo que se expondrá en este trabajo, es la convergencia, para el caso donde tanto el programa primal, como el dual son **no degenerados**.

En primer término se demostrará que si la sucesión de puntos $\mathbf{x^k}$ y ($\mathbf{y^k}$, $\mathbf{z^k}$) convergen y lo hacen hacia los puntos óptimos de los programas **(P)** y **(D)** respectivamente**. L**os puntos $\mathbf{y^k}$ y $\mathbf{z^k}$ son precisamente las variables duales definidas en (**3.11**), a las cuales por conveniencia, se les han añadido el superíndices correspondiente a la iteración **k**

Teorema 3.10.1

Dado un Programa Matemático Lineal al que llamaremos **(P),** con dominio admisible acotado y su correspondiente dual **(D).** Si la sucesión de puntos $\mathbf{x^k}$, $\mathbf{y^k}$ y $\mathbf{z^k}$ generado por el método de Escalado afín, converge hacia los puntos $\mathbf{x^*}$, $\mathbf{y^*}$ y $\mathbf{z^*}$. Entonces el punto $\mathbf{x^*}$ es óptimo para el programa **(P)** y el par ($\mathbf{y^*}$, $\mathbf{z^*}$) es óptimo para el programa **(D)**

Demostración

Consideremos dos puntos en la sucesión $\mathbf{x}^k$ y $\mathbf{x}^{k+1}$, que pertenecen al dominio admisible del programa **(P).**

Luego $\mathbf{A}\ (\mathbf{x}^k - \mathbf{x}^{k+1}) = \mathbf{A}\ \mathbf{x}^k - \mathbf{A}\ \mathbf{x}^{k+1} = \mathbf{0}$

Utilizando este resultado y las definiciones de $\mathbf{x}^k$ en (**3.11**), $\Delta\mathbf{x}$ en (**3.13**) y el proceso iterativo (**3.1**), la disminución del costo que se obtiene al pasar de $\mathbf{x}^k$ a $\mathbf{x}^{k+1}$, puede escribirse como:

$$\Delta\ \textbf{costo} = \mathbf{c}^t(\mathbf{x}^k - \mathbf{x}^{k+1}) = \mathbf{c}^t(\mathbf{x}^k - \mathbf{x}^{k+1}) - (\mathbf{y}^k)^t\ \mathbf{A}(\mathbf{x}^k - \mathbf{x}^{k+1})$$

$$= (\mathbf{c} - \mathbf{A}^t\mathbf{y}^t)^t(\mathbf{x}^k - \mathbf{x}^{k+1}) = (\mathbf{z}^k)^t(\mathbf{x}^k - \mathbf{x}^{k+1})$$

$$= (\mathbf{z}^k)^t)\left(-\rho\,\bar{\alpha}\,\Delta\,\mathbf{x}^k\right)$$

$$= -\rho\,\bar{\alpha}\,(\mathbf{z}^k)^t\,\Delta\mathbf{x}^k = -\rho\,\bar{\alpha}\,(\mathbf{z}^k)^t\left(-\mathbf{D}\mathbf{z}^k\right) \qquad (\mathbf{3.24})$$

$$= \rho\,\bar{\alpha}\left((\mathbf{z}^k)^t(\mathbf{X}^k)^2\mathbf{z}^k\right) = \rho\,\bar{\alpha}\left((\mathbf{z}^k)^t(\mathbf{x}^k)^t(\mathbf{x}^k)\mathbf{z}^k\right)$$

$$= \rho\,\bar{\alpha}\left(\mathbf{X}^k\mathbf{z}^k\right)^t\left(\mathbf{X}^k\mathbf{z}^k\right) = \rho\,\bar{\alpha}\left\|\left(\mathbf{X}^k\mathbf{z}^k\right)\right\|^2$$

$$\Delta\ \textbf{costo} = \rho\,\bar{\alpha}\left\|\left(\mathbf{X}^k\mathbf{z}^k\right)\right\|^2$$

Pero como se definió $\bar{\alpha}$ en **(3.12)** y $\Delta\mathbf{x}$ en **(3.10)**, se tiene

$$\bar{\alpha} = \min_i\left\{-\frac{\mathbf{x}_i^k}{\Delta\mathbf{x}_i} : \Delta\mathbf{x}_i \leq \mathbf{0}\right\}$$

$$= \min_i\left\{\frac{\mathbf{1}}{\mathbf{x}_i^k\mathbf{z}_i^k} : \mathbf{z}_i^k \geq \mathbf{0}\right\} \qquad (\mathbf{3.25})$$

$$= \frac{1}{\max_i \left\{ x_i^k z_i^k : z_i^k \geq 0 \right\}}$$

Ahora acotando $\bar{\alpha}$:

Si se utiliza la propiedad que para cualquier vector $\mathbf{u}$: $\mathbf{u_i} \leq \|\mathbf{u}\|$, luego se cumple:

$$\bar{\alpha} = \frac{1}{\max_i \left\{ x_i^k z_i^k \ z_i^k \geq 0 \right\}} \geq \frac{1}{\|X^k z^k\|}$$

Al sustituir esta última expresión en $(\mathbf{3.25})$, se obtiene:

$$\Delta\mathbf{costo} = \rho\,\bar{\alpha} = \|X^k z^k\|^2 \geq \rho \frac{\|X^k z^k\|^2}{\|X^k z^k\|} = \rho \|X^k z^k\|$$

Ahora como la sucesión de costos $\mathbf{c}^T \mathbf{x}^k$ es decreciente y acotada inferiormente, por ser el dominio admisible de $(\mathbf{P})$ acotado, entonces la sucesión es convergente y se verifica que:

$$\lim_{k\to\infty} \Delta\mathbf{costo} = \lim_{k\to\infty} \mathbf{c}^t (\mathbf{x}^k - \mathbf{x}^{k+1}) = \mathbf{c}^t (\mathbf{x}^k - \mathbf{x}^k) = \mathbf{c}^t(\mathbf{0})$$

Al combinar las últimas relaciones, se obtiene:

$$\lim_{k\to\infty} \rho \|X^k z^k\| = 0 \Rightarrow \lim_{k\to\infty} x_i^k z_i^k = 0, \text{ para } \mathbf{i} = 1, 2, \ldots, \mathbf{n} \qquad (\mathbf{3.26})$$

Como las sucesiones $\{\mathbf{x}^k\}$ y $\{\mathbf{z}^k\}$, convergen hacia $\mathbf{x}^*$ y $\mathbf{z}^*$ respectivamente, se tiene que:

$$x_i^* z_i^* = 0, \text{ para } \mathbf{i} = 1, 2, \ldots, \mathbf{n} \qquad (\mathbf{3.27})$$

También por ser $\{\mathbf{x}^k\}$ del dominio admisible de $(\mathbf{P})$, se tiene:

$$\mathbf{A}\,\mathbf{x}^* = \mathbf{b} \qquad (\mathbf{3.28})$$

Seguidamente se demostrará que $\mathbf{z}^* \geq 0$

Procediendo por el absurdo. Supóngase que existe aún $z_i^* < 0$; luego existe un índice k, tal que para $K > k$, se cumple que $z_i^k < 0$.

Si se escribe el proceso iterativo generador de la sucesión $\{x^k\}$ y se toma su i-ésima componente, como:

$$x_i^{k+1} = x_i^k - \alpha (x_i^k)^2 z_i^k$$

Al considerar $K > k$, se concluirá que $x_i^{k+1} > x_i^k > 0$,(pues $\alpha > 0$, $x_i^k > 0$ y la suposición que $z_i^* < 0$)

Por tanto $x_i^* z_i^* < 0$, lo que contradice la relación (**3.11.4**). Concluyendo que:

$$c^T (x^k - x^{k+1}) = c^T (x^k - x^{k+1}) - (y^k)^T A (x^k - x^{k+1})$$

$$z^* \geq 0 \qquad (3.29)$$

Como $z^* = c - A^T y^*$. Por (**3.6.5**), por las condiciones (**3.26**), (**3.27**) y (**3.28**) y el teorema de la holgura complementaria; se concluye que x^* y el par (y^*, z^*), son soluciones óptimas de los programas (**P**) y (**D**) respectivamente.

El resultado anterior supone que las sucesiones $\{x^k\}$, $\{y^k\}$ y $\{z^k\}$ son convergentes.

El siguiente teorema demuestra que si se verifica la hipótesis 2 y la hipótesis 3, de no degeneración de los programas primal y dual; las presentes sucesiones son, de hecho, convergentes.

Teorema 3.10.2

Si los programas (**P**) y (**D**) son no degenerados con y el dominio admisible del programa (**P**) es acotado, entonces para todo $\rho < 1$¸ las sucesiones $\{\mathbf{x}^{\mathbf{k}}\}$ e $\{\mathbf{y}^{\mathbf{k}}\}$, que genera el método del Escalado afín, son convergentes

Demostración

Se dividirá la demostración en tres partes:

En la parte 1: Se demostrara que la sucesión $\{\mathbf{x}^{\mathbf{k}}\}$es acotada

En la parte 2: Que cualquier subsucesión de $\{\mathbf{x}^{\mathbf{k}}\}$ converge hacia una solución básica factible

En la parte 3: Que la solución básica factible a la cual converge cualquier subsucesión de $\{\mathbf{x}^{\mathbf{k}}\}$ es única.

Queda garantizada así que la sucesión $\{\mathbf{x}^{\mathbf{k}}\}$ (y a su vez también la sucesión $\{\mathbf{y}^{\mathbf{k}}\}$ es convergente

1.- Dado que el programa $(\mathbf{P})$ tiene un dominio admisible que es acotado y al igual que su correspondiente dual $(\mathbf{D})$ son no degenerados, tiene un único punto donde el programa toma su solución óptima

Sea $\mathbf{x}^*$ tal solución óptima de $(\mathbf{P})$

Consideremos el siguiente programa

$$\begin{array}{ll} & \underset{\mathbf{x}}{\mathbf{Máx}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \mathbf{x}_{\mathbf{j}} \\ \mathbf{(Q)} & \mathbf{S.a.}\ \ \mathbf{A\,x = b} \\ & \qquad \mathbf{c^t\,x^k \leq c^t\,x^*} \\ & \qquad\quad \mathbf{x \geq 0} \end{array}$$

Es fácil notar que el dominio admisible es acotado, puesto que es un subconjunto del dominio admisible del programa **P.** Este programa tiene un valor óptimo finito y $\mathbf{x}^*$ es único.

Por otro lado si se desplaza el hiperplano $c^t x^*$ paralelo así mismo para obtener $c^t x^0$ **es** decir existe $k \in \mathbb{R}$, tal que $c^t x^* + k = c^t x^*$. Si x^0 es la solución inicial, generada por el algoritmo del escalado afín, esto implica que el siguiente programa:

$$\begin{aligned} &\underset{x}{\text{Máx}} \sum_{j=1}^{n} x_j \\ &\text{S.a.}\ \ A x = b \\ &\qquad c^t x^k \leq c^t x^0 \\ &\qquad\quad x \geq 0 \end{aligned}$$

También tiene un valor óptimo acotado.

Por tanto, el conjunto de puntos $S = \{x : A x = b,\ c^T x^k \leq c^T x^0,\ x \geq 0\}$ es acotado; pues $S \subset F$ y F es acotado (de no ser así $\underset{x}{\text{Máx}}$ $\sum_{j=1}^{n} x_j$, sería no acotado)

Por otro lado se recordará que la sucesión $\{x^k\}$, generada por el algoritmo del Escalado afín, resulta en una sucesión $\{c^k x^k\}$ decreciente. Por lo tanto pertenece al conjunto S. Luego se concluye que la sucesión $\{x^k\}$ es acotada, ya que S es acotado.

Luego por ser $\{x^k\}$una sucesión acotada y utilizando un resultado del análisis matemático (Teorema de Bolzano Weierstrass), se puede concluir que existe una subsucesión $\{x^{k_j}\}$ que converge a un punto límite $\bar{x}$.

A continuación se demostrará que este punto límite, es una solución básica factible. En efecto. Consideremos la solución del programa (P): $\beta = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_k : \bar{x}_i > 0\}$, por ser (P), no degenerado, se tiene que $k \geq m$. El teorema de existencia de soluciones factibles para un PML, afirma que si existe una solución óptima factible para el programa ($\mathbf{P}$); existirá también una solución básica factible y óptima ($\mathbf{s.b.f.o.}$), $\bar{\mathbf{x}}^*$

Sea $\mathbf{B}$ la matriz básica asociada a $\bar{\mathbf{x}}^*$. Con esta base se elegirá la variable $\bar{\mathbf{x}}_i$ asociada a $\mathbf{B}$.

Por tanto se puede seleccionar $\hat{\mathbf{x}} = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{x}}_3, . . . , \hat{\mathbf{x}}_m : \hat{\mathbf{x}}_i > \mathbf{0}\}$ del subconjunto β ; de forma que las columnas $\{\mathbf{a}_1\ \mathbf{a}_2, \mathbf{a}_3, . . . , \mathbf{a}_m\}$, de la matriz $\mathbf{A}$, asociadas a estas variables, sean linealmente independientes; lo que haría que la solución además de ser factible y no degenerada, sea básica

Sea $\mathbf{B} = \{\mathbf{a}_1\ \mathbf{a}_2, \mathbf{a}_3, . . . , \mathbf{a}_m\}$, la matriz básica asociada y $\mathbf{c}_B = \{ \mathbf{c}_{B_1}\ \mathbf{c}_{B_2}, . . . , \mathbf{c}_{B_m} \}$ el vector de coeficientes de la función objetivo asociados a las variables básicas factibles

Por otro lado, por definición se tiene: $\mathbf{z}^k = \mathbf{c} - \mathbf{A}^T \mathbf{y}^k$ y por la relación ($\mathbf{3.26}$), del Teorema 3.9.3, se cumple que

$$\mathop{\mathbf{Lím}}_{j \to \infty} \mathbf{x}_i^{k_j} (\mathbf{c}_i - \mathbf{a}_i^t \mathbf{y}^{k_j}) = \mathbf{0}, \text{ para } \mathbf{i} = 1, 2, 3, . . . , \mathbf{m}$$

Las componentes de $\mathbf{x}_{B_i}^{k_j}$ convergen hacia $\bar{\mathbf{x}}_{B_i}$, para $\mathbf{i} = 1, 2, 3, . . . , \mathbf{m}$, y son positivas porque $\mathbf{B}_i$ es un índice básico; es decir es un miembro de $\{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_m\} = \mathbf{I}_B$

Luego se concluye que $\mathbf{c}_B - \mathbf{B}\,\mathbf{y}^{k_j}$ debe converger hacia $\mathbf{0}$, para garantizar ($\mathbf{3.26}$)

De aquí debido a que $\mathbf{B}$ es no singular se deduce que $\mathbf{y}^{k_j}$ converge hacia $\mathbf{c}_B\,\mathbf{B}^{-1}$

Por lo tanto se concluye que todas las componentes $c_i - a_i^t y^{k_j}$, convergen hacia $c_i - a_i^t B^{-1} c_B$, valor que coincide con el costo reducido asociado a la componente $\bar{x}_{B_i} \in \bar{x}_B$

La no degeneración del programa dual (D), garantiza que el costo reducido $c_i - a_i^T B^{-1} c_B \neq 0$, para toda $x_i \notin \bar{x}_B$(que no es variable básica).

Entonces para las variables no básicas se tiene que el costo reducido $c_i - a_i^t B^{-1} c_B$, convergen a un valor diferente de cero. Por lo que utilizando una vez más (3.25), se tiene que $\bar{x}_i = 0$, Por lo tanto las componentes $\bar{x}_B = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_m\}$ serán estrictamente positivas; garantizando de esta forma que el punto límite $\bar{x}$ sea una **s.b.f.** Además $\bar{x} = \bar{x}^*$, por la forma como fue elegida x

Se ha demostrado que la sucesión $\{x^k\}$ tiene un punto límite $\bar{x}$ y que éste es una **s.b.f.o.** del programa (P).

Seguidamente se demostrará que este punto límite es único

Tómese el real $\delta > 0$, de tal forma que cualquier componente de una **s.b.f.** sea mayor que δ.

Este valor siempre existirá ya que la solución del programa (P) es no degenerada y cada una de sus componentes son estrictamente positivas.

Sea $\in = \delta/3$, dado que cualquier punto límite es $\{x^k\}$ es una **s.b.f.** existe un k, tal que para cualquier $k \geq K$, $\|x^k - \hat{x}\| < \in$

Procediendo por el absurdo:

Supóngase que existieran dos soluciones diferentes $\bar{x}$ y $\hat{x}$ y tales que para $k \geq K$:

$$\left\| x^k - \bar{x} \right\| \le \in \quad y \quad \left\| x^{k+1} - \hat{x} \right\| \le \in \qquad 3.30$$

Considérese ahora que ($\bar{x}_i = 0$); esto es la variable $\bar{x}_i$ no es básica pero es básica en $\hat{x}$ ($\hat{x}_i \ge \delta$). Al sustituir $\bar{x}_i = 0$ en la primera relación de (**3.30**) y utilizando el hecho que $u_i \le |u_i| \le \|u\|$, para cualquier vector u; se tiene que:

$$x_i^k < \in \qquad 3.31$$

De forma similar al sustituir $\hat{x}_i \ge \delta$ en la segunda relación de (**3.30**), se obtiene:

$$\left| x_i^{k+1} - \hat{x}_i \right| \le \in \Rightarrow - \in \le x_i^{k+1} - \hat{x}_i \le \in \Rightarrow x_i^{k+1} \ge \hat{x}_i - \in$$

Pero de $\hat{x}_i \ge \delta$, se tiene que $\hat{x}_i - \in \ge \delta - \in$; de donde se tiene que $x_i^{k+1} \ge \ \ge \delta - \in = 2\in$, por ser $\in = \delta/3$, se tiene:

$$x_i^{k+1} \ge 2\in \qquad 3.32$$

Al utilizar la ecuación del esquema iterativo (**3.1**), la definición de Δx en (**3.11**) y la expresión de longitud de paso máximo (**3.25**), se obtiene

$$x_i^{k+1} = \hat{x}_i + \rho\,\bar{\alpha}\,(- (x_i^k)^2 z_i^k) =$$

$$= x_i^k \left(1 - \rho \frac{1}{\text{máx}\left\{ x_i^k z_i^k, \forall\, x_i^k \ge 0 \right\}} x_i^k z_i^k \right)$$

$$x_i^{k+1} = x_i^k \left(1 - \rho \frac{1}{\text{máx}\left\{ x_i^k z_i^k, \forall\, x_i^k \ge 0 \right\}} x_i^k z_i^k \right) \qquad 3.33$$

Al utilizar un argumento similar al utilizado al final del Teorema 3.3, se puede afirmar que :

$x_i^k \ge 0$ (lo que posibilita asegurar que $\hat{x}_i = 0$)

De la relación (**3.33**), la suposición que $\rho < 1$ y utilizando (**3.31**) y (**3.32**), se obtiene finalmente:

$$x_i^{k+1} \leq x_i^k\ (1+\rho) < 2x_i^k \leq 2\in \leq x_i^{k+1}$$

Lo que es una contradicción. Concluyendo que el punto límite $\bar{\mathbf{x}}$ es único y que la sucesión $\{\mathbf{x}^k\}$, converge a él.

Como ha quedado demostrado anteriormente, la sucesión $\{\mathbf{y}^k\}$ converge también a $\mathbf{c}_B\ \mathbf{B}^{-1}$; la solución dual asociada a $\bar{\mathbf{x}}$.

Observación 3.10.2

Si $\rho \leq \mathbf{2/3}$ y el valor del programa es acotado, las sucesiones $\{\mathbf{x}^k\}$ y $\{\mathbf{y}^k\}$ generadas por el presente algoritmo, convergen.

CAPÍTULO 4: METODOLOGÍA

4.1 Tipo y Diseño de Investigación

La investigación es descriptiva, se presentan los métodos simplex y escalado afín para mostrar las ventajas del método de escalado afín frente al simplex.
Se compara el tiempo de ejecución del algoritmo del MAE con el del Simplex

Presentar una aplicación: un caso para dominio admisible acotado con solución óptima no degenerada y solución dual óptima única.

La que será expuesta en varias partes, la primera de ellas, empieza, comparando el método de Escalado Afín de paso corto, que fue la propuesta original de Dikin en 1967, para seguidamente presentar la manera como se elige la dirección de mejora; avanzando luego con la nueva solución del programa, para finalmente arribar a través de diversas iteraciones a la solución óptima del programa. Insistiendo en los detalles de la búsqueda, ya que el trabajo como su nombre lo indica es difusión del Mae

CAPITULO 5: RESULTADOS Y DISCUSIÓN

Primero se presenta un caso en que se aplica el escalado afín de paso corto (creado por Dikin en 1967), en el segundo caso se aplica el de paso largo (que fue una mejora al de Dikin introducida por Vanderbeit en 1986), en el tercer caso se desarrolla el procedimiento para mejorar el valor actual del programa utilizando la dirección de búsqueda (también atribuido a Vanderbeit), hasta llegar a la solución optima, utilizando el MAE.

Finalmente se comparó el tiempo de ejecución del MAE con el método simplex.

5.1. Aplicación del método escalado afín

Caso – 1

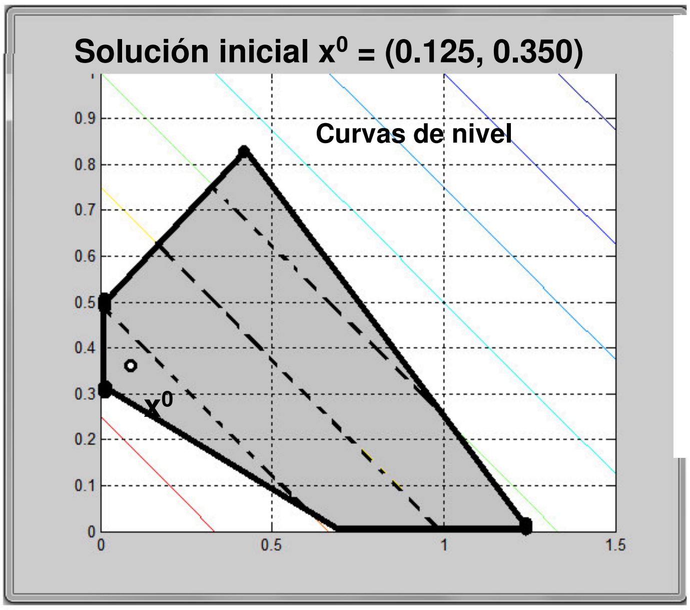
Tómese nuevamente el programa R_1

Se seguirán las etapas **a, b y c** :

Etapa a: Escalamiento del programa mediante las relaciones (3.7) y (3.8)

Tómese $\mathbf{x^k} = \begin{bmatrix} \mathbf{0.125} \\ \mathbf{0.350} \\ \mathbf{2.500} \\ \mathbf{0.475} \\ \mathbf{3.100} \end{bmatrix}$ en el interior del dominio admisible del programa $\mathbf{R_1}$

Gráfica n º 24: Curvas de nivel y punto interior, factible e inicial



Fuente propia

En este punto el valor del programa es

$$\mathbf{cx^{(0)}} = [-3, -4]\begin{bmatrix} 0.125 \\ 0.350 \end{bmatrix} = \mathbf{cx^{(0)}} = \mathbf{-1.775}$$

Al aplicar la ecuación **(3.8)**: $\mathbf{c^t = X^{(0)}c}$ y $\mathbf{A = AX^{(0)}}$, se tendrá $\mathbf{x^{(0)}}$

$$\mathbf{A = AX^{(0)}} = \begin{bmatrix} -8 & 10 & 1 & 0 & 0 \\ 3 & 6 & 0 & -1 & 0 \\ 4 & 4 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0 & 0 & 0 \\ 0 & 0 & 2.50 & 0 & 0 \\ 0 & 0 & 0 & 0.475 & 0 \\ 0 & 0 & 0 & 0 & 3.10 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} -1 & 3.50 & 2.50 & 0 & 0 \\ 0.375 & 2.10 & 0 & -0.475 & 0 \\ 0.50 & 1.40 & 0 & 0 & 3.10 \end{bmatrix}$$

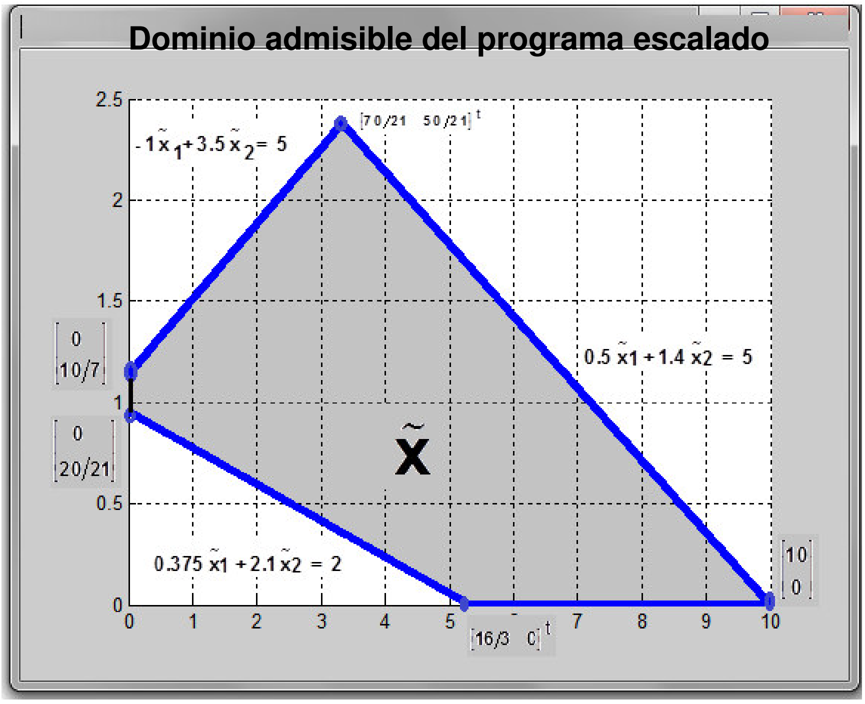
$$\mathbf{c^t = X^{(0)}c} = \begin{bmatrix} 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0 & 0 & 0 \\ 0 & 0 & 2.50 & 0 & 0 \\ 0 & 0 & 0 & 0.475 & 0 \\ 0 & 0 & 0 & 0 & 3.10 \end{bmatrix}\begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.375 \\ -1.400 \\ 0.00 \\ 0.00 \\ 0.00 \end{bmatrix}$$

De esta manera el programa escalado $\mathbf{R}_1$ será

$$\begin{array}{lrrrrrcl}
\textbf{Min f(x)} & -0.375x_1 & -1.4x_2 & +\ 0x_3 & +\ 0x_4 & +\ 0x_5 & & \\
\text{St} & -1x_1 & +3.5x_2 & +2.5x_3 & +\ 0x_4 & +\ 0x_5 & = & 5 \\
 & 0.375x_1 & +2.1x_2 & +\ 0x_3 & -0.475x_4 & +\ 0x_5 & = & 2 \qquad 5.1 \\
 & 0.5x_1 & +1.4x_2 & +\ 0x_3 & +\ 0x_4 & +3.1x_5 & = & 5 \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & &
\end{array}$$

La siguiente gráfica muestra el dominio admisible del programa escalado

Gráfica n º 25: Dominio admisible del programa escalado



Fuente propia

Etapa b: Utilizando la relaciones **(3.9)**, , **(3.10)** y **(3.11)** se obtendrá el vector gradiente proyectado; que da la dirección de movimiento $\Delta \mathbf{x}$ (para mejorar el valor actual del programa):

En efecto $\Delta \mathbf{x} = -\mathbf{X}^k(\mathbf{c}^t - \mathbf{A}^t\mathbf{y}) = -\mathbf{X}^k(\mathbf{z})$, con $\mathbf{y} = \left(\mathbf{ADA}^t\right)^{-1}\mathbf{ADc}^t$

Haciendo las operaciones correspondientes

$$\mathbf{D} = \left(\mathbf{X}^k\right)^2 = \begin{bmatrix} 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0 & 0 & 0 \\ 0 & 0 & 2.50 & 0 & 0 \\ 0 & 0 & 0 & 0.475 & 0 \\ 0 & 0 & 0 & 0 & 3.10 \end{bmatrix}^2 =$$

$$\begin{bmatrix} 0.01563 & 0 & 0 & 0 & 0 \\ 0 & 0.12250 & 0 & 0 & 0 \\ 0 & 0 & 6.2500 & 0 & 0 \\ 0 & 0 & 0 & 0.22563 & 0 \\ 0 & 0 & 0 & 0 & 9.6100 \end{bmatrix}$$

$$\mathbf{AD} = \begin{bmatrix} -8 & 10 & 1 & 0 & 0 \\ 3 & 6 & 0 & -1 & 0 \\ 4 & 4 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.01563 & 0 & 0 & 0 & 0 \\ 0 & 0.12250 & 0 & 0 & 0 \\ 0 & 0 & 6.2500 & 0 & 0 \\ 0 & 0 & 0 & 0.22563 & 0 \\ 0 & 0 & 0 & 0 & 9.6100 \end{bmatrix}$$

$$\mathbf{AD} = \begin{bmatrix} -0.125000 & 1.22500 & 6.25000 & 0.00000 & 0.00000 \\ 0.04688 & 0.73500 & 0.00000 & -0.22563 & 0.00000 \\ 0.06250 & 0.49000 & 0.00000 & 0.00000 & 9.61000 \end{bmatrix}$$

$$\mathbf{ADA^t} = \begin{bmatrix} -0.125000 & 1.22500 & 6.25000 & 0.00000 & 0.00000 \\ 0.04688 & 0.73500 & 0.00000 & -0.22563 & 0.00000 \\ 0.06250 & 0.49000 & 0.00000 & 0.00000 & 9.61000 \end{bmatrix}$$

$$\begin{bmatrix} -8 & 3 & 4 \\ 10 & 6 & 4 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{ADA^t} = \begin{bmatrix} 19.50000 & 6.975000 & 4.40000 \\ 6.97500 & 4.77625 & 3.12750 \\ 4.40000 & 3.12750 & 11.82000 \end{bmatrix}$$

$\mathbf{ADA^t}$ Es una matriz de rango completo ($\mathbf{m}$), ya que $\mathbf{A}$ lo es. También la matriz $\mathbf{X^k}$ posee no menos de $\mathbf{m}$ componentes positivas; luego la matriz $\mathbf{ADA^t}$ es invertible

$$\left[ADA^t\right]^{-1} = \begin{bmatrix} 0.107398 & -0.15804 & 0.00184 \\ -0.15804 & 0.4858 & -0.06971 \\ 0.00184 & -0.06971 & 0.10236 \end{bmatrix}$$

Por otro lado

$$ADc^t = \begin{bmatrix} -0.125000 & 1.22500 & 6.25000 & 0.00000 & 0.00000 \\ 0.04688 & 0.73500 & 0.00000 & -0.22563 & 0.00000 \\ 0.06250 & 0.49000 & 0.00000 & 0.00000 & 9.61000 \end{bmatrix} \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} -4.52500 \\ -3.08063 \\ -2.14750 \end{bmatrix}$$

$$y = \left(ADA^t\right)^{-1} ADc^t = \begin{bmatrix} 0.107398 & -0.15804 & 0.00184 \\ -0.15804 & 0.4858 & -0.06971 \\ 0.00184 & -0.06971 & 0.10236 \end{bmatrix} \begin{bmatrix} -4.52500 \\ -3.08063 \\ -2.14750 \end{bmatrix}$$

$$= \begin{bmatrix} -0.00305 \\ -0.63176 \\ -0.01339 \end{bmatrix}$$

$$A^t y = \begin{bmatrix} -8 & 3 & 4 \\ 10 & 6 & 4 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.00305 \\ -0.63176 \\ -0.01339 \end{bmatrix} = \begin{bmatrix} -1.92440 \\ -3.87466 \\ -0.00305 \\ 0.63176 \\ -0.01339 \end{bmatrix}$$

$$z = c^t - A^t y = \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -1.92440 \\ -3.87466 \\ -0.00305 \\ 0.63176 \\ -0.01339 \end{bmatrix} = \begin{bmatrix} -1.07560 \\ -0.12534 \\ 0.00305 \\ -0.63176 \\ 0.01339 \end{bmatrix}$$

$$\Delta x = -X^k(z) = \begin{bmatrix} 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0 & 0 & 0 \\ 0 & 0 & 2.50 & 0 & 0 \\ 0 & 0 & 0 & 0.475 & 0 \\ 0 & 0 & 0 & 0 & 3.10 \end{bmatrix}^2 \begin{bmatrix} -1.07560 \\ -0.12534 \\ 0.00305 \\ -0.63176 \\ 0.01339 \end{bmatrix}$$

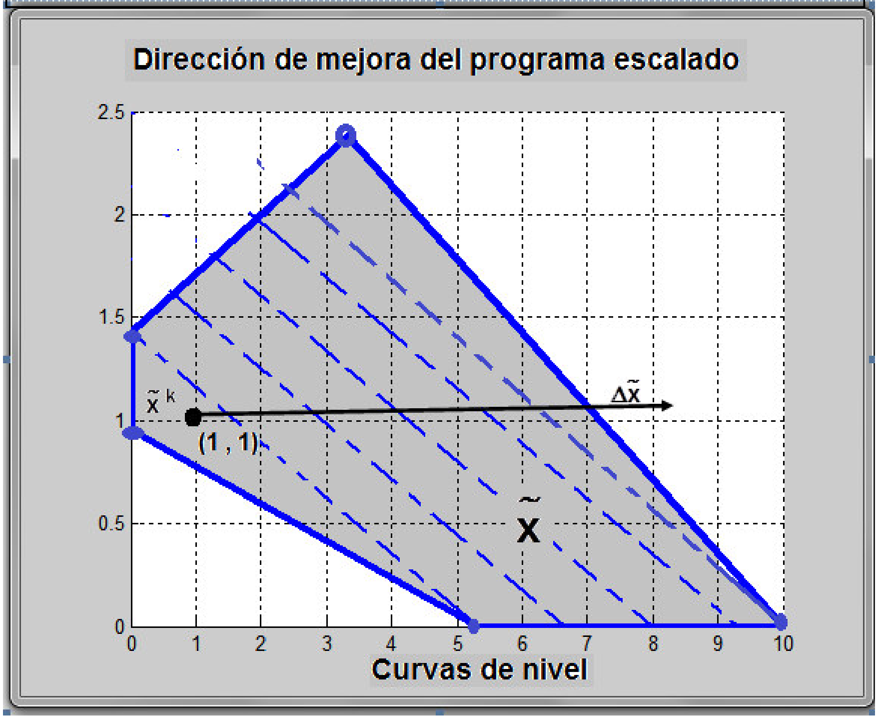
$$\Delta x = \begin{bmatrix} 0.13445 \\ 0.04387 \\ -0.00764 \\ 0.30009 \\ -0.04150 \end{bmatrix}$$

De esta manera la dirección de mejoramiento en el nuevo dominio admisible del programa $\mathbf{R}_1$; será:

$$x^1 = x^0 + \Delta x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.13445 \\ 0.04387 \\ -0.00764 \\ 0.30009 \\ -0.04150 \end{bmatrix} = \begin{bmatrix} 1.13445 \\ 1.04387 \\ 0.99236 \\ 1.30009 \\ 0.95850 \end{bmatrix}$$

La siguiente gráfica muestra la dirección $\Delta \mathbf{x}$, en que mejorará el valor actual del programa $\mathbf{cx}$.

Gráfica n º 26: Dirección de mejora en el programa escalado



Fuente propia

Pero en este dominio admisible, sólo se está interesado en obtener el gradiente proyectado $\Delta\tilde{\mathbf{x}}$; que permitirá ir mejorando el valor del programa.

En consecuencia se ejecutará la etapa C del algoritmo ; esto hallar $\mathbf{x}^{k+1}$

Etapa C: vuelta al programa original $(\tilde{\mathbf{x}} \to \mathbf{x})$

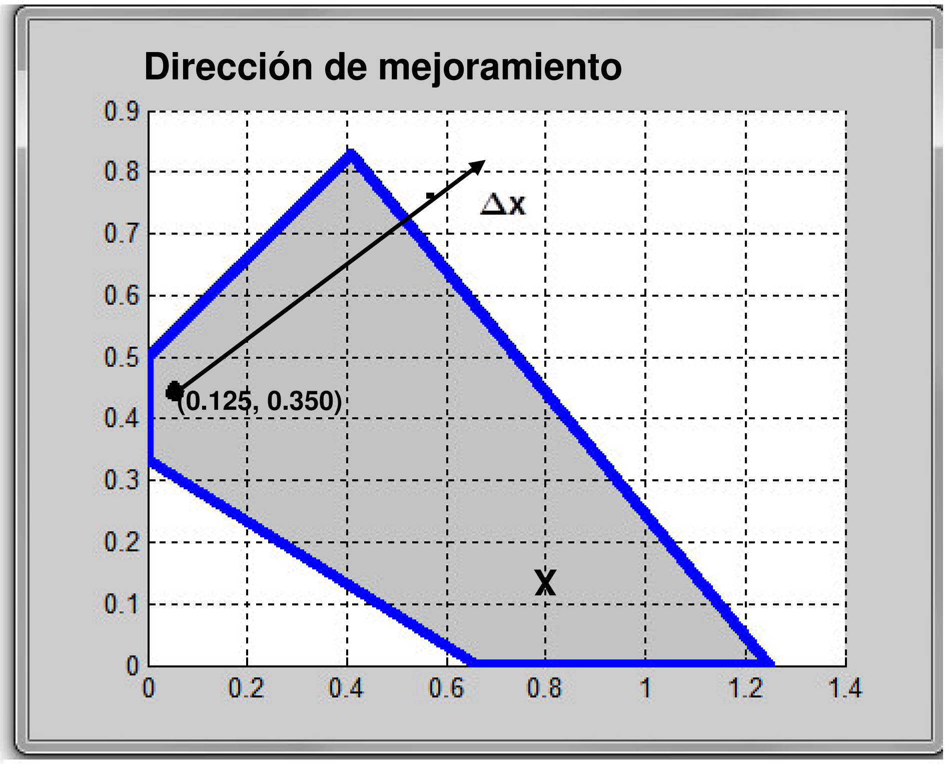
$$\Delta\mathbf{x} = \mathbf{X}^k\,\Delta\mathbf{x} = \begin{bmatrix} 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0 & 0 & 0 \\ 0 & 0 & 2.50 & 0 & 0 \\ 0 & 0 & 0 & 0.475 & 0 \\ 0 & 0 & 0 & 0 & 3.10 \end{bmatrix}^2 \begin{bmatrix} 0.13445 \\ 0.04387 \\ -0.00764 \\ 0.30009 \\ -0.04150 \end{bmatrix}$$

$$\Delta\mathbf{x} = \begin{bmatrix} 0.01681 \\ 0.01535 \\ -0.01909 \\ 0.14254 \\ -0.12864 \end{bmatrix}$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x} = \begin{bmatrix} 0.125 \\ 0.350 \\ 2.500 \\ 0.475 \\ 3.100 \end{bmatrix} + \begin{bmatrix} 0.01681 \\ 0.01535 \\ -0.01909 \\ 0.14254 \\ -0.12864 \end{bmatrix} = \begin{bmatrix} 0.50975 \\ 0.70150 \\ 2.06301 \\ 3.73825 \\ 0.15500 \end{bmatrix}$$

La siguiente gráfica, presenta la dirección de mejora $\Delta\mathbf{x}$

Gráfica n ° 27: Dirección de mejora 2da iteración



Fuente propia

Caso - 2

Considérese nuevamente el programa $\mathbf{R}_1$

$$
\begin{array}{llllll}
\textbf{Min f(x)}: & -\ 3x_1 & -\ 4x_2 & & & \\
\text{St} & -8x_1 & +\ 10x_2 & \leq & 5 & \cdots (5) \\
 & 3x_1 & +\ 6x_2 & \geq & 2 & \\
 & 4x_1 & +\ 4x_2 & \leq & 5 & \\
 & x_1 \geq 0 & x_2 \geq 0 & & &
\end{array}
$$

Estandarizando y escalándolo, se tiene el programa $\mathbf{R}_1$:

$$
\begin{array}{llllllll}
\textbf{Min f(x)} & -\ 0.375x_1 & -\ 1.4x_2 & +\ 0\ x_3 & +\ 0\ x_4 & +\ 0\ x_5 & & \\
\text{St} & -1\ x_1 & +3.5\ x_2 & +\ 2.5\ x_3 & +\ 0\ x_4 & +\ 0\ x_5 & = & 5 \\
 & 0.375\ x_1 & +\ 2.1\ x_2 & +\ 0\ x_3 & -.0.475\ x_4 & +\ 0\ x_5 & = & 2 \\
 & 0.5\ x_1 & +\ 1.4\ x_2 & +\ 0\ x_3 & +\ 0\ x_4 & +3.1\ x_5 & = & 5 \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & &
\end{array}
\qquad (5.1)
$$

Se sabe que con este escalado el punto x^k en el programa original, queda transformado $x^n = e^t = (1, . . . , 1)$, en el nuevo espacio, que es una esfera.

De acuerdo a (**3.17**), el programa escalado $\mathbf{R_1}$, toma la forma $\mathbf{R_2}$

Esto es

$$
\begin{array}{llllllll}
\textbf{Min f(x)} & -0.375\,x_1 & -1.4\,x_2 & +\ 0\,x_3 & +\ 0\,x_4 & +\ 0\,x_5 & & \\
\text{St} & -1\,x_1 & +3.5\,x_2 & +2.5\,x_3 & +\ 0\,x_4 & +\ 0\,x_5 & = & 5 \\
 & 0.375\,x_1 & +2.1\,x_2 & +\ 0\,x_3 & -.0.475\,x_4 & +\ 0\,x_5 & = & 2 \\
 & 0.5\,x_1 & +1.4\,x_2 & +\ 0\,x_3 & +\ 0\,x_4 & +3.1\,x_5 & = & 5 \\
 & (x_1-1)^2 & +(x_2-1)^2 & +(x_3-1)^2 & +(\tilde{x}_4-1)^2 & +(\tilde{x}_5-1)^2 & = & 1 \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & &
\end{array}
$$

$\tilde{5}.\tilde{2}$

Se verificará que el dominio admisible es efectivamente una esfera de centro $\mathbf{e^t}$ y radio $\mathbf{r = 1}$

De las restricciones: primeras, segunda y tercera, se tiene respectivamente

$$x_3 = \frac{5 + x_1 - 3.5x_2}{2.5}, \quad x_4 = \frac{0.375\,x_1 + 2.1\,x_2 - 2}{0.475} \quad \text{y} \quad x_5 = \frac{5 - 0.5\,x_1 - 0.5\,x_2}{3.1}$$

Al sustituir estas variables en la cuarta restricción:

$(x_1-1)^2 + (x_2-1)^2 + (x_3-1)^2 + (x_4-1)^2 + (x_5-1)^2 = 1$, (que es una hiperesfera)

Se tiene

$$(x_1-1)^2 + (x_2-1)^2 + \left(\frac{5 + x_1 - 3.5x_2}{2.5} - 1\right)^2 + \left(\frac{0.375\,x_1 + 2.1\,x_2 - 2}{0.475} - 1\right)^2 + \left(\frac{5 - 0.5\,x_1 - 0.5\,x_2}{3.1} - 1\right)^2 = 1$$

Ejecutando operaciones simplificadoras, se obtiene la ecuación de la elipse

$1.80928\,x_1^2 + 22.70966\,x_2^2 + 6.00629\,x_1\,x_2 - 9.62486x_1 - 51.42561\,x_2 + 29.52523 = 0$, que es parte constituyente del programa escalado que y que representa la intersección de la hiperesfera con los hiperplanos en el dominio admisible **x** del programa escalado, que se llamará $\mathbf{R}_3$ y que se reescribe como:

$$\text{Min } f(x) = -0.375\,x_1 - 1.4\,x_2 + 0\,x_3 + 0\,x_4 + 0\,x_5$$

S.a.

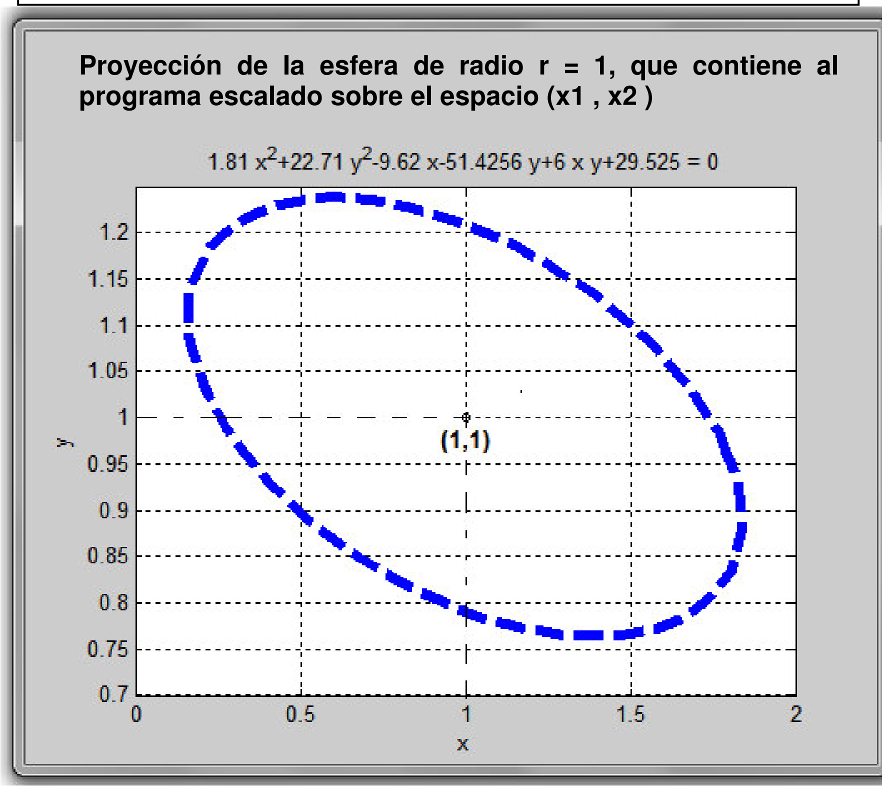
$$1.81\tilde{x}_1^2 + 22.71\tilde{x}_2^2 + 6.01\tilde{x}_1\,\tilde{x}_2 - 9.63\tilde{x}_1 - 51.43 + 29.53 = 0 - 51.42561\,x_2 + 29.52523 = 0\,,$$

$$1.81\,x_1^2 + 22.71\,x_2^2 + 6.01\,x_1\,x_2 - 9.63x_1 - 51.42561\,x_2 + 29.52523 = 0$$

$$x_1 \geq 0 \quad x_2 \geq 0 \quad x_3 \geq 0 \quad x_4 \geq 0 \quad x_5 \geq 0$$

Seguidamente se tiene la gráfica del dominio admisible del programa $\mathbf{R}_3$ sobre $\mathbb{R}^2$

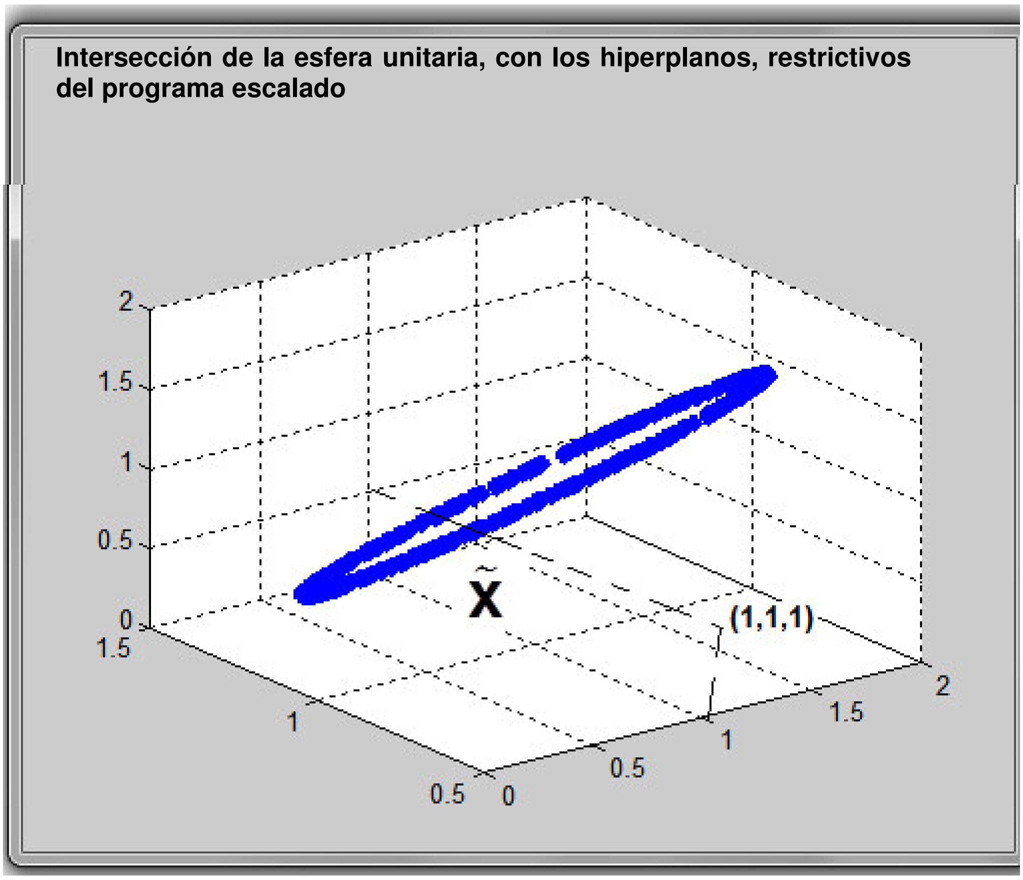
Gráfica n º 28: Proyección de la intercepción, sobre el plano x_1, x_2



Fuente propia

En $\mathbb{R}^3$, la gráfica de la intersección de la hiperesfera con los hiperplanos en el dominio admisible $\mathbf{x}$ del programa escalado, es la siguiente:

Gráfica n º 29: Proyección de la intercepción, sobre el plano x_1, x_2

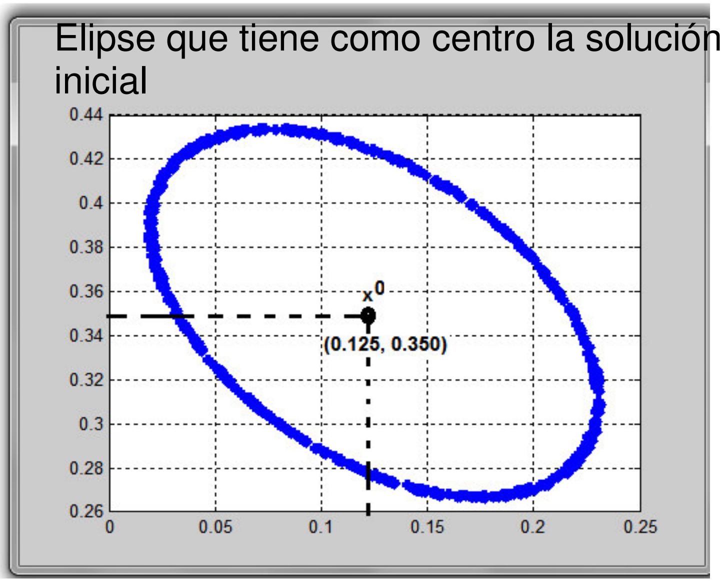


Fuente propia

También se tiene el elipsoide en el dominio admisible del programa $\mathbf{R}_1$:

$\mathbf{x} = \mathbf{X}^k\mathbf{x}$

Gráfica n º 30: Centro de la elipse del programa original



Fuente propia

También se tiene el elipsoide en el dominio admisible del programa $\mathbf{R}_1$:

$\mathbf{x} = \mathbf{X}^k\mathbf{x}$

Lo que garantiza que dicho punto sea interior, es que $\rho < 1$

En la relación **(3.11.1)** se mostró como obtener $\mathbf{X}^k\,\tilde{\mathbf{P}}\,\tilde{\mathbf{c}}$

Esta forma de elegir $\Delta\tilde{\mathbf{x}}$, para obtener

$$\tilde{\mathbf{x}}^{k+1} = \tilde{\mathbf{x}}^k + \Delta\tilde{\mathbf{x}}$$

Incluye implícitamente la elección de α en el esquema iterativo:

$$\tilde{\mathbf{x}}^{k+1} = \tilde{\mathbf{x}}^k + \alpha\,\Delta\tilde{\mathbf{x}}$$

Es llamado **"Método Iterativo de Escalado Afín de Paso Corto"**

En el **"Esquema de Paso Corto"** se puede apreciar que no se calcula α, debido a que la longitud de paso viene dada por la relación $\dfrac{\rho}{\left\|\tilde{\mathbf{P}}\tilde{\mathbf{c}}\right\|}$, la que conduce hasta la frontera del elipsoide, el término ρ desempeña un papel similar al mismo término del algoritmo de paso largo; acorta el paso realizado, alejándose de la frontera del elipsoide y en el de paso largo de alguna cara de dominio admisible, definida por $\mathbf{x} \geq \mathbf{0}$

Los elipsoides se adaptan según el escalado a la forma del dominio admisible

Obviamente si en ves de permanecer en la frontera del elipsoide, se continúa el avance, hasta alcanzar alguna cara definida por $x \geq 0$, tal y como se realiza en el algoritmo de paso largo; disminuirá el valor del programa

Caso - 3

- Tómese nuevamente el **P.M.L.** del **Ejemplo 3.4**

Se está ahora en condición de poder exponer las etapas del algoritmo

Considérese nuevamente el programa R_1

$$
\begin{array}{llllll}
\text{Min } f(x): & -3x_1 & -4x_2 & & & \\
\text{St} & -8x_1 & +10x_2 & \leq & 5 & \\
 & 3x_1 & +6x_2 & \geq & 2 & \cdots 3 \\
 & 4x_1 & +4x_2 & \leq & 5 & \\
 & x_1 \geq 0 & x_2 \geq 0 & & &
\end{array}
$$

Y su forma estándar correspondiente:

Considérese el programa R_1

$$
\begin{array}{llllll}
\text{Min } f(x): & -3x_1 & -4x_2 & & & \\
\text{St} & -8x_1 & +10x_2 & \leq & 5 & \\
 & 3x_1 & +6x_2 & \geq & 2 & \cdots 3 \\
 & 4x_1 & +4x_2 & \leq & 5 & \\
 & x_1 \geq 0 & x_2 \geq 0 & & &
\end{array}
$$

Y du forma estándar correspondiente:

$$
\begin{array}{lllllllll}
\text{Min } f(x) = & -2x_1 & -2x_2 & +0x_3 & +0x_4 & +0x_5 & & & \\
\text{St} & -8x_1 & +10x_2 & +1x_3 & +0x_4 & +0x_5 & = & 5 & \cdots 3.1 \\
 & 3x_1 & +6x_2 & +0x_3 & -1x_4 & +0x_5 & = & 2 & \\
 & 4x_1 & +4x_2 & +0x_3 & +0x_4 & +1x_5 & = & 5 & \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & & &
\end{array}
$$

Completamiento de la resolución, con el método

Se seguirán las etapas **a, b y c**:

Para el punto inicial $\mathbf{k=0}$. $\mathbf{x}^k = \mathbf{(0.125\,, 0.35\,, 2.5\,, 0.475\,, 3.1)}^t$, que pertenece al interior de su dominio admisible, se tiene $\mathbf{c\,x^0 = -1.775}$.

Tómese $\rho = \mathbf{0.95}$.

En la primera iteración del algoritmo, se tiene $\mathbf{y = (ADA^t)^{-1}ADc^t} =$

$$\begin{bmatrix} \mathbf{-0.00305} \\ \mathbf{-0.63176} \\ \mathbf{-0.01339} \end{bmatrix}$$

Así $\mathbf{b^t\,y} = [\mathbf{5} \quad ,\mathbf{2}\,, \quad \mathbf{5}] \begin{bmatrix} \mathbf{-0.00305} \\ \mathbf{-0.63176} \\ \mathbf{-0.01339} \end{bmatrix} = \mathbf{-\,1.34573}$

Obteniéndose como

Gap (brecha) dual $\dfrac{|\mathbf{c^t x^0 - b^t y}|}{\mathbf{1} + |\mathbf{c^t x^0}|} = \dfrac{|\mathbf{-1.775-(-1.34573)}|}{\mathbf{1}+|\mathbf{-1.775}|} = \dfrac{\mathbf{0.429273}}{\mathbf{2.775}} =$ $\mathbf{0.15469}$

Lo que significa que $\mathbf{x^k = x^0}$ aún no es óptimo, se hallará seguidamente $\mathbf{x^1}$

Se tiene que $\mathbf{x^1 = x^0 + \alpha\,\Delta\,x}$

Como $\alpha = \mathbf{22.89332}$

$$\mathbf{x^1 = x^0 + \alpha\,\Delta\,x} = \begin{bmatrix} \mathbf{0.125} \\ \mathbf{0.350} \\ \mathbf{2.500} \\ \mathbf{0.475} \\ \mathbf{3.100} \end{bmatrix} + \mathbf{22.89332} \begin{bmatrix} \mathbf{0.01681} \\ \mathbf{0.01535} \\ \mathbf{-0.01909} \\ \mathbf{0.14254} \\ \mathbf{-0.12864} \end{bmatrix} = \begin{bmatrix} \mathbf{0.50975} \\ \mathbf{0.70150} \\ \mathbf{2.06301} \\ \mathbf{3.73825} \\ \mathbf{0.15500} \end{bmatrix}$$

Con el nuevo punto $\mathbf{x^1} = \begin{bmatrix} \mathbf{0.50975} \\ \mathbf{0.70150} \\ \mathbf{2.06301} \\ \mathbf{3.73825} \\ \mathbf{0.15500} \end{bmatrix}$, se tendrá la siguiente iteración

Segunda Iteración

$$X^1 = \begin{bmatrix} 0.50975 & 0 & 0 & 0 & 0 \\ 0 & 0.70150 & 0 & 0 & 0 \\ 0 & 0 & 2.06301 & 0 & 0 \\ 0 & 0 & 0 & 3.73825 & 0 \\ 0 & 0 & 0 & 0 & 0.15500 \end{bmatrix}$$

Al aplicar la ecuación **(3.8)**: $\mathbf{c}^t = \mathbf{X}^{(0)}\mathbf{c}$ y $\mathbf{A} = \mathbf{A}\mathbf{X}^{(0)}$, se tendrá

$$\mathbf{A} = \mathbf{A}\mathbf{X}^{(1)} = \begin{bmatrix} -8 & 10 & 1 & 0 & 0 \\ 3 & 6 & 0 & -1 & 0 \\ 4 & 4 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0.59561 & 0 & 0 & 0 & 0 \\ 0 & 0.64189 & 0 & 0 & 0 \\ 0 & 0 & 3.34596 & 0 & 0 \\ 0 & 0 & 0 & 3.63817 & 0 \\ 0 & 0 & 0 & 0 & 0.05000 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} -4.07801 & 7.01499 & 2.06301 & 0 & 0 \\ 1.52925 & 4.20900 & 0 & -3.73825 & 0 \\ 2.03900 & 2.80600 & 0 & 0 & 0.15500 \end{bmatrix}$$

$$\mathbf{c}^t = \mathbf{X}^{(0)}\mathbf{c} = \begin{bmatrix} 0.59561 & 0 & 0 & 0 & 0 \\ 0 & 0.64189 & 0 & 0 & 0 \\ 0 & 0 & 3.34596 & 0 & 0 \\ 0 & 0 & 0 & 3.63817 & 0 \\ 0 & 0 & 0 & 0 & 0.05000 \end{bmatrix} \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1.52925 \\ -2.80600 \\ 0.00000 \\ 0.00000 \\ 0.00000 \end{bmatrix}$$

De esta manera el nuevo programa escalado $\mathbf{R}_4$ será

$$\begin{array}{llllll}
\textbf{Min } f(x) = & -1.52925\,x_1 & -2.80600\,x_2 & +\;0\,x_3 & +\;0\,x_4 & +\;0\,x_5 \\
\text{St} & -4.07801\,x_1 & +7.01499\,x_2 & +2.06301\,x_3 & +\;0\,x_4 & \\
 & 1.52925\,x_1 & +4.20900\,x_2 & +\;0\,x_3 & -3.73825\,x_4 & \\
 & 2.03900\,x_1 & +2.80600\,x_2 & +\;0\,x_3 & +\;0\,x_4 & \\
 & (x_1-1)^2 & +(x_2-1)^2 & +(x_3-1)^2 & +(x_4-1)^2 & \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 &
\end{array}$$

$$+ \quad 0\,x_5 \quad = \quad 5$$
$$+ \quad 0\,x_5 \quad = \quad 1$$
$$+ \, 0.15500\,x_5 \quad = \quad 2$$
$$+ (x_5 - 1)^2 \quad = \quad 1$$
$$x_5 \geq 0$$

Se verificará una vez más, que el dominio admisible es efectivamente una esfera de centro $\mathbf{e^t}$ y radio $\mathbf{r = 1}$

De las restricciones: primeras, segunda y tercera, se tiene respectivamente

$$x_3 = \frac{5 - 4.078x_1 + 7.015x_2}{2.063}, \qquad x_4 = \frac{1.529\,x_1 + 4.21x_2 - 2}{3.738} \qquad \text{y}$$
$$x_5 = \frac{5 - 2.039x_1 - 2.806x_2}{0.155}$$

Al sustituir estas variables en la cuarta restricción

$(x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 + (x_4 - 1)^2 + (x_5 - 1)^2 = 1$, (que es una hiperesfera)

Se tiene

$$(x_1 - 1)^2 + (x_2 - 1)^2 + \left(\frac{5 - 4.078x_1 + 7.015x_2}{2.063} - 1\right)^2 +$$
$$\left(\frac{1.529\,x_1 + 4.21x_2 - 2}{3.738} - 1\right)^2 + \left(\frac{5 - 2.039x_1 - 2.806x_2}{0.155} - 1\right)^2 = 1$$

Ejecutando operaciones simplificadoras, se nuevo programa escalado $\mathbf{R}_5$

$$\text{Min } f(x) = \; -\,1.52925\,x_1 \; - \; 2.806\,x_2 \quad + \; 0\,x_3 \quad + 0\,x_4 \; + 0\,x_5$$

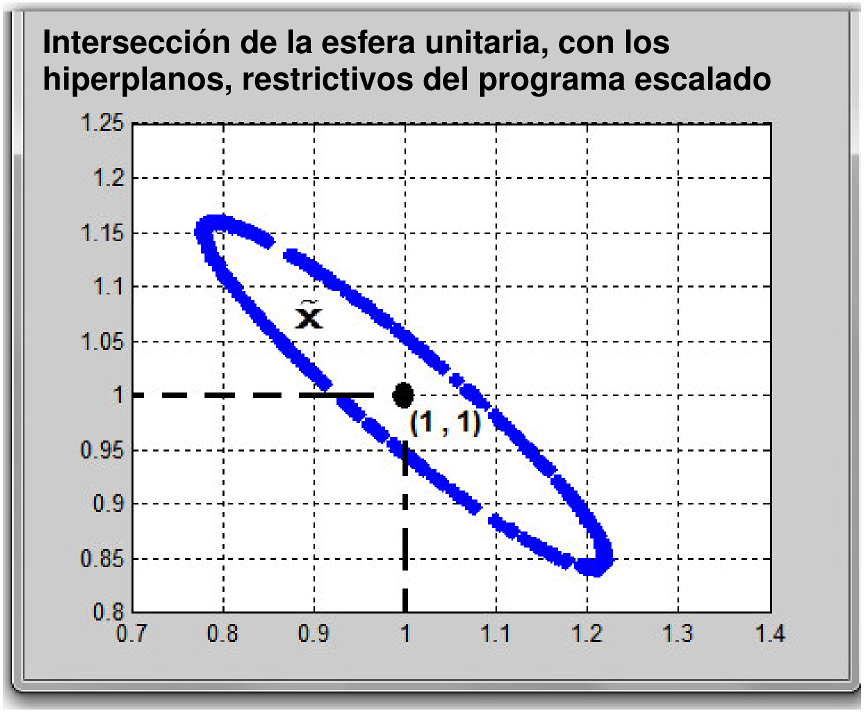
S.a.

$$178.1254\,x_1^2 + 341.55639\,x_2^2 + 463.76822\,x_1\,x_2 - 820.01822\,x_1$$
$$- 1146.88099\,x_2 + 982.44961 = 0$$

$$x_1 \geq 0 \quad x_2 \geq 0 \quad x_3 \geq 0 \quad x_4 \geq 0 \quad x_5 \geq 0$$

Seguidamente se tiene la gráfica del dominio admisible de este programa en $\mathbb{R}^2$

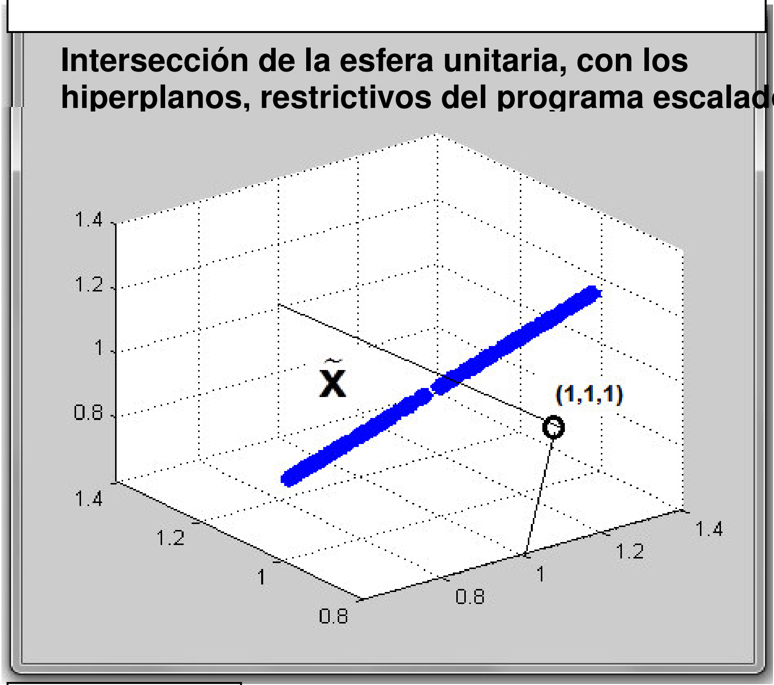
Gráfica n º 31: Proyección de la intercepción, sobre el plano x_1, x_2



Fuente propia

En $\mathbb{R}^3$, la gráfica de la intersección de la hiperesfera con los hiperplanos en el dominio admisible **x** del programa escalado, es la siguiente:

Gráfica n º32: Proyección de la intercepción sobre el plano x_1, x_2 y x_3

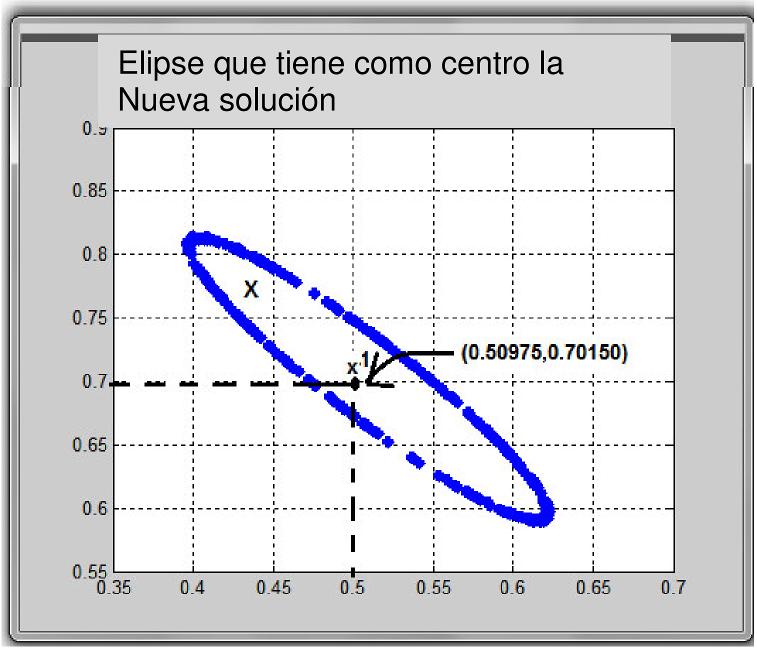


Fuente propia

También se tiene el elipsoide en el dominio admisible del programa $\mathbf{R}_1$:

$\mathbf{x} = \mathbf{X}^k\mathbf{x}$

Gráfica n º 33: Elipse cuyo centro es el avance en la 2da iteración



Fuente propia

Cuadro n º 1: Contiene todas las iteraciones, que conducen al óptimo

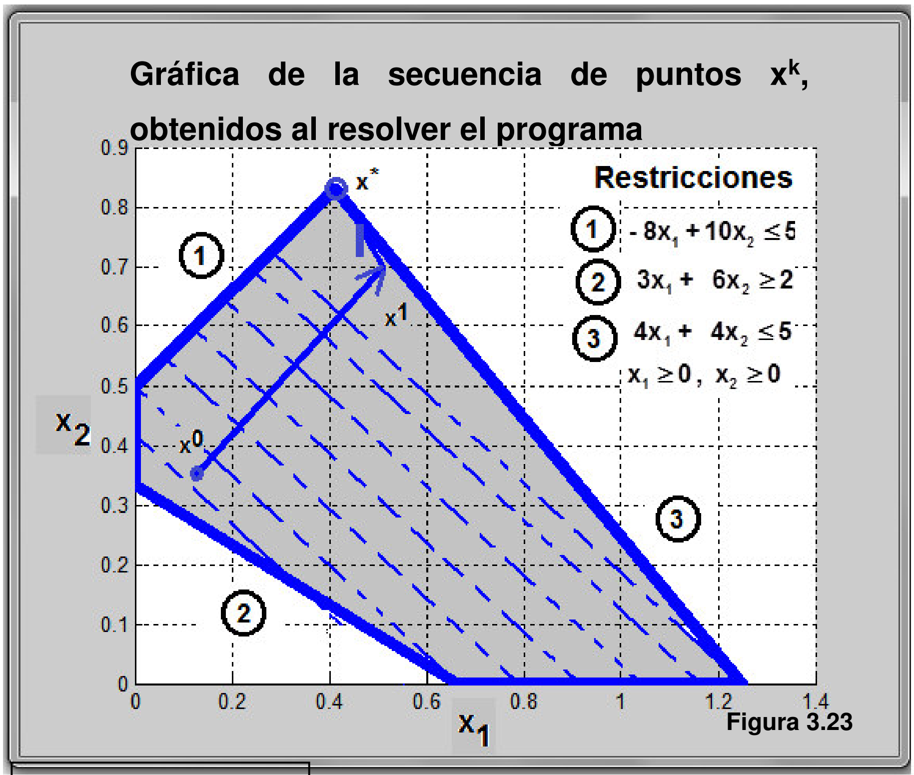
	VARIABLES				
ITERACIONES	x_1	x_2	x_3	x_4	x_5
0	0.12500	0.35000	2.50000	0.47500	3.10000
1	0.50975	0.70150	2.06301	5.73825	0.15500
2	0.41724	0.83126	0.02527	4.23930	0.00599
3	0.41767	0.83226	0.01873	4.24655	0.00030
4	0.41669	0.83326	0.00094	4.24960	0.00023

Fuente propia

F.O.	**Gap**
-1.775000	0.15469
-4.33525	0.04136
-4.54859	0.000627
-4.57.677	0.00117
-4.58203	0.00023

La siguiente gráfica muestra la secuencia de puntos seguida por el algoritmo de escalado afín, para llegar a la solución óptima

Gráfica n º34: Contiene la secuencia de puntos X^k, obtenidos al resolver el programa



Fuente propia

5.2. Análisis, interpretación y discusión de resultados

Como se ha podido apreciar, mediante las ilustraciones que se han desarrollado, el MAE ha funcionado sin perturbaciones, debido a que se han seguido escrupulosamente los pasos que este indica; sin embargo se debe remarcar los peligros de su estabilidad y convergencia, cuando una o más componentes del vector solución, se acerca a el o (los) hiperplanos, que forman la frontera del dominio admisible

También se observa que un criterio de parada, puede ser, cuando el vector de coeficientes c, de la función objetivo del programa original, se ha vuelto ortogonal al subespacio nulo de la matriz A, de coeficientes del programa; esto es; existen escalares no negativos, contenidos en el vector y, que combinados con los vectores filas de la matriz A; expresan el vector c

5.3. Comparación de resultados

Se debe decir que por ser el caso resuelto (caso 3:página 134); de pocas variables y pocas restricciones; el tiempo de ejecución es bastante pequeño; sin embargo en el libro de Vanderbei; se comenta, La buena estabilidad del MAE y su mejor eficiencia computacional que el método Símplex.

CONCLUSIONES RECOMENDACIONES

1.4.1 Objetivo General	
• Analizar y difundir el método de escalado afín, variante del método de puntos interiores	Se ha presentado la forma, como funciona el método de Escalado, los supuestos que demanda, su convergencia
1.4.2 Objetivos Específicos	
• Describir el método de escalado afín	Se há presentado la forma como funciona el método, ilustrado cada uma de sus etapas
• Comparar la eficiencia del método de escalado afín con la del método simplex	Se presenta la comparación de la eficiência computacional com rl método Simplex Se presenta uma aplicación completa del método. Caso 3: página 134

BIBLIOGRAFÍA

Aguilar, J. (s.f.). *http://wiki.foros-fiuba.com.ar/*. Obtenido de http://wiki.foros-fiuba.com.ar/: http://wiki.foros-fiuba.com.ar/_media/materias:61:08:15._descomposicion_en_valores_singulares_apunte_de_mancilla_.pdf

Bazaraa, & mokhtar, bazaraa, m. (2010). linear programming and networks flow. En mokhtar, & m. bazaraa, linear programming and networks flow.

Burden, J. Douglas Faires, B. (2002). *Análisis Numérico.* Princeton: International Thomson.

Castro, J. (15 de febrero de 2002). *https://www.google.com.pe/#q=importancia+de+estudiar+el+algoritmo+de+escalado+afin*. Recuperado el 10 de febrero de 2003, de https://www.google.com.pe/#q=importancia+de+estudiar+el+algoritmo+de+escalado+afin: http://www-eio.upc.es/~jcastro/publications/reports/dr2000-02.pdf

Christiansen, E., & Kortanek, K. (1990). Computation of the collapse state in limit. Obtenido de http://ac.els-cdn.com: http://ac.els-cdn.com/037704279190147C/1-s2.0-037704279190147C-main.pdf?_tid=397a0712-13f1-11e6-bc58-00000aacb35e&acdnat=1462583925_f044aae1e7b615921d56020e8595447a

Dikin (1967) "Iterative solution of problems of linear and quadratic" programming

Hefferon, J. (2014). LINEAR ALGEBRA. Vermont USA 05439: Mathematics, Saint Michael's Colleg

De la Fuente O'Connor, J. (s.f.). T´ecnicas de C´alculo para Sistemas de Ecuaciones, Programación Lineal y Programación Entera. Madrid.

Gambini Inés."Algoritmo de Puntos Interiores de Karmakar", Biblioteca Facultad de Ciencias Matemáticas

Monteiro, R. (s.f.). https://www.google.com.pe/?gfe_rd=cr&ei=ktAgV9WED4yw8weAs4XwCw

#q=Asimplifiedglobalconvergenceproofofthe+affinescalingalgorithm*. Recuperado el 20 de 03 de 2016, de https://www.google.com.pe/?gfe_rd=cr&ei=ktAgV9WED4yw8weAs4XwCw#q=Asimplifiedglobalconvergenceproofofthe+affinescalingalgorithm*: https://www.google.com.pe

Pool, D. (2011). Álgebra lineal Una Introducción Moderna. En P. David, Álgebra lineal Una Introducción Moderna (pág. 612). México, D.F.: Cengage Learning Editores, S.A. de C.V.

TSUCHIYA, T. (08 de 1995). *http://www3.grips.ac.jp/*. Obtenido de http://www3.grips.ac.jp/: http://www3.grips.ac.jp/~tsuchiya/papers_new/1995_Global_convergence_of_a_long-step_affine.pdf

Vanderbei, R. (2001). Linear Programming: Foundations and Extensions. PRINCETON UNIVERSITY: DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING.

Zaballa, I. (s.f.). *http://www.ehu.eus/izaballa/Cursos/valores_singulares.pdf*. Obtenido de http://www.ehu.eus/izaballa/Cursos/valores_singulares.pdf: http://www.ehu.eus/

Anexos

1.

Gráfica - 28

%PROGRAMA (MATLAB).Grafica la ecuación formada por la intersección de la esfera y los hiperplanos, en el espacio escalado

```
clear all; clc;
hL=ezplot('1.81*x^2+22.71*y^2-9.62*x-
51.4256*y+6*x*y+29.525')
set(hL,'Color','b','LineStyle','--','linewidth',5);
grid on
axis([0 2  0.7 1.25])
```

2.

Gráfica – 29

```
clear all;clc;j=0;
coefxc=1.81;coefyc=22.71;coefxy=6;coefx=-9.625;coefy=-
51.4256;indep=29.525;
x(1) =.125; x(2) =.35;x(3) =2.5;x(4) =.475;x(5) =3.1;
a1=0.125;a2=0.35;a3=2.5;a4=.475;a5=3.1
for i=1:5;xk(i,i) =x(i);end
for i=1:90000
x=2*rand;y=2*rand;
z=coefxc*x.^2+coefyc*y.^2+coefx*x+coefy*y+coefxy*x*y+inde
p;
if abs(z) <=0.02
j=j+1;C(1,j)=x;C(2,j)=y;
```

```
C(3,j)=(5+8*a1*x-10*a2*y)/a3;C(4,j)=(3*a1*x+6*a2*y-
2)/a4;C(5,j)=(5-4*a1*x-4*a2*y)/a5;
end
end
D=xk*C;
plot(D(1,:),D(2,:),' *b','linewidth',3)
grid on
```

3.

Gráfica – 30

```
clear all;clc;j=0;
coefxc=178.125;coefyc=341.5564;coefxy=463.768;coefx=-
820.018;coefy=-1146.881;indep=982.4496;
x(1) =.50975; x(2) =.7015;x(3) =2.063;x(4) =3.7383;x(5) =.155;
a1=0.50975;a2=0.7015;a3=2.063;a4=3.7383;a5=.155
for i=1:5;xk(i,i) =x(i);end
for i=1:900000
x=2*rand;y=2*rand;
z=coefxc*x.^2+coefyc*y.^2+coefx*x+coefy*y+coefxy*x*y+inde
p;
if abs(z) <=0.02
j=j+1;C(1,j)=x;C(2,j)=y;
C(3,j)=(5+8*a1*x-10*a2*y)/a3;C(4,j)=(3*a1*x+6*a2*y-
2)/a4;C(5,j)=(5-4*a1*x-4*a2*y)/a5;
end
end
D=xk*C;
plot(D(1,:),D(2,:),' *b','linewidth',3)
grid on
```

4.

Gráfica – 31

```
clear all;clc;j=0;
coefxc=178.125;coefyc=341.5564;coefxy=463.768;coefx=-
820.018;coefy=-1146.881;indep=982.4496;
x(1) =.50975; x(2) =.7015;x(3) =2.063;x(4) =3.7383;x(5) =.155;
a1=0.50975;a2=0.7015;a3=2.063;a4=3.7383;a5=.155
for i=1:5;xk(i,i) =x(i);end
for i=1:900000
x=2*rand;y=2*rand;
z=coefxc*x.^2+coefyc*y.^2+coefx*x+coefy*y+coefxy*x*y+inde
p;
if abs(z) <=0.02
j=j+1;C(1,j)=x;C(2,j)=y;
C(3,j)=(5+8*a1*x-10*a2*y)/a3;C(4,j)=(3*a1*x+6*a2*y-
2)/a4;C(5,j)=(5-4*a1*x-4*a2*y)/a5;
end
end
plot(C(1,:),C(2,:),' *b','linewidth',3)
grid on
```

5.

Gráfica – 32

```
clear all;clc;j=0;
coefxc=178.125;coefyc=341.5564;coefxy=463.768;coefx=-
820.018;coefy=-1146.881;indep=982.4496;
x(1) =.50975; x(2) =.7015;x(3) =2.063;x(4) =3.7383;x(5) =.155;
a1=0.50975;a2=0.7015;a3=2.063;a4=3.7383;a5=.155
for i=1:5;xk(i,i) =x(i);end
for i=1:900000
x=2*rand;y=2*rand;
z=coefxc*x.^2+coefyc*y.^2+coefx*x+coefy*y+coefxy*x*y+inde
p;
if abs(z) <=0.02
j=j+1;C(1,j)=x;C(2,j)=y;
C(3,j)=(5+8*a1*x-10*a2*y)/a3;C(4,j)=(3*a1*x+6*a2*y-
2)/a4;C(5,j)=(5-4*a1*x-4*a2*y)/a5;
end
end
plot(C(1,:),C(2,:),' *b','linewidth',3)
grid on
```

6.

Gráfica – 33

```
clear all;clc;j=0;
coefxc=178.125;coefyc=341.5564;coefxy=463.768;coefx=-
820.018;coefy=-1146.881;indep=982.4496;
x(1) =.50975; x(2) =.7015;x(3) =2.063;x(4) =3.7383;x(5) =.155;
a1=0.50975;a2=0.7015;a3=2.063;a4=3.7383;a5=.155
```

```
for i=1:5;xk(i,i) =x(i);end
for i=1:900000
x=2*rand;y=2*rand;
z=coefxc*x.^2+coefyc*y.^2+coefx*x+coefy*y+coefxy*x*y+inde
p;
if abs(z) <=0.02
j=j+1;C(1,j)=x;C(2,j)=y;
C(3,j)=(5+8*a1*x-10*a2*y)/a3;C(4,j)=(3*a1*x+6*a2*y-
2)/a4;C(5,j)=(5-4*a1*x-4*a2*y)/a5;
end
end
D=xk*C;
plot(D(1,:),D(2,:),' *b','linewidth',3)
grid on
```

7. Este programa resuelve, el programa 3, página 134

```
clear all;clc;
c=[ -3 -4 0 0 0];A=[ -8 10 1 0 0;3 6 0 -1 0; 4 4 0 0 1];b=[5 2
5];At=A';
sol=zeros(5,7); xk=[ .125;.35;2.5;.475;3.1];
%anillo final subindice ii
for ii=1:5;xkf=xk';
for ll=1:5;sol(ii,ll)=xkf(ll);end;sol(ii,6)=c*xk;fo=c*xk;
for i=1:5;XK(i,i)=xk(i);end
D=XK*XK;AD=A*D;ADAt=AD*At;ADc=AD*c';ADAtinv=inv(ADA
t);
y=ADAtinv*ADc;z=c'-At*y;ga=b*y;
Deltax=-
D*z;Deltasomb=inv(XK)*Deltax;coc=Deltax;coc=Deltax;
```

```
%muy bien
%Determinacion de longitud de paso
j=0;for i=1:5;
if Deltax(i)<0
j=j+1;coc(j)=-xk(i)/Deltax(i);
end
end
min=100;for i=1:j;
if coc(i)<min
min =coc(i);
end
end
alfa=.95*min;Alfadelta=alfa*Deltax;
gap=abs(ga-fo)/(1+abs(fo));sol(ii,7)=gap;
xk =xk+Alfadelta;
end
sol
```