



La matriz inversa generalizada (el espacio contravariante) A^{-1} de matrices de orden $m \times n$ con $m \geq n$ y la solución cerrada de este problema

J.L. Urrutia Galicia
Coordinación de Mecánica Aplicada
Instituto de Ingeniería, UNAM
E-mail: jurg@pumas.iingen.unam.mx

(recibido: octubre de 2001; aceptado: marzo de 2002)

Resumen

En este artículo se demuestra que dada una matriz rectangular A de orden $m \times n$ con $m \geq n$, y el conjunto $\{\phi_1, \phi_2, \dots, \phi_n\}$ formado por los vectores columna (covariantes) $\tilde{\phi}_n$ de A linealmente independientes, pero no necesariamente ortogonales, existe entonces una inversa única A^{-1} formada por la transpuesta de la matriz $B = \{\phi^1, \phi^2, \dots, \phi^n\}$ de vectores base columna (contravariantes) de cálculo tensorial. El caso de $m = n$ para matrices cuadradas es un caso particular. De esta forma, el problema clásico $Ax = b$ cuando A es rectangular, sólo se puede resolver en forma aproximada. En una etapa subsiguiente y haciendo cambios mínimos en los elementos de las columnas de la matriz A es posible obtener una solución óptima, con error cero, mediante la proyección adecuada de los vectores columna de la matriz A sobre un plano de solución óptima, como se muestra más adelante con detalle.

Descriptores: Descriptores: espacio covariante -vs- matriz rectangular, espacio contravariante -vs- inversa de una matriz rectangular, coordenadas oblicuas, producto escalar, determinante del tensor métrico, solución exacta, proyección de coordenadas.

Abstract

The objective of this paper is to show that given a rectangular $m \times n$ A matrix with $m \geq n$, and the set $\{\phi_1, \phi_2, \dots, \phi_n\}$ formed by the (covariant) column vectors $\tilde{\phi}_n$ of A , linearly independent but not necessarily orthogonal, there is a unique inverse A^{-1} formed by the transpose of the matrix $B = \{\phi^1, \phi^2, \dots, \phi^n\}$ of base contravariant columns from tensor calculus. The case when $m = n$ for square matrices is only a particular one. In this way, the classical problem $Ax = b$ when A is rectangular can only be solved in an approximate way. In a subsequent stage and making small changes in the entries of the columns of matrix A it will be seen that it is possible to obtain an optimal solution, with no error, by an adequate projection of the column vectors of matrix A on an optimum plane of solution as it will be discussed in due time.

Keywords: Keywords: covariant space -vs- rectangular matrix, contravariant space -vs- inverse of a rectangular matrix, scalar product, skew coordinates, determinant of the metric tensor, exact solution, coordinate projection.

Introducción

En la teoría actual de inversas generalizadas, una matriz \mathbf{A} de orden $m \times n$ real (rectangular) con $m \geq n$ tiene una pseudoinversa única \mathbf{A}^\dagger si, y sólo si las siguientes ecuaciones de Penrose son satisfechas según Hall (1981), Gill *et al.* (1991), Strang (1980), Ben y Greville (1974).

$$\begin{aligned} \mathbf{A} \mathbf{A}^\dagger \mathbf{A} &= \mathbf{A} \\ \mathbf{A}^\dagger \mathbf{A} \mathbf{A}^\dagger &= \mathbf{A}^\dagger \\ (\mathbf{A} \mathbf{A}^\dagger)^T &= \mathbf{A} \mathbf{A}^\dagger \\ (\mathbf{A}^\dagger \mathbf{A})^T &= \mathbf{A}^\dagger \mathbf{A} \end{aligned} \quad (1)$$

Las aplicaciones de pseudoinversas \mathbf{A}^\dagger han sido ilustradas por Forsythe *et al.* (1977) y Fraleigh y Beauregard (1989), pero probablemente la mejor exposición de ideas y aplicaciones es la de Tucker (1988), en la que abiertamente trata de “aterrizar” el Álgebra Lineal abstracta, con el objeto de demostrar que “Las aplicaciones del Álgebra Lineal son poderosas, fáciles de entender y muy diversas”. Dentro de este contexto, Tucker hace una presentación sistemática del Álgebra Lineal mediante su aplicación a programación lineal, programación entera, teoría de gráficas, circuitos eléctricos, balance de ecuaciones químicas y en ajuste de datos por el método de mínimos cuadrados. Dada una matriz \mathbf{A} de orden $m \times n$ su pseudoinversa \mathbf{A}^\dagger se define como sigue

$$\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \quad (2)$$

Desde el punto de vista geométrico, Tucker también presenta una interpretación geométrica de la solución de un sistema de ecuaciones; asimismo, en el curso de su libro demuestra que dada una matriz rectangular \mathbf{A} de $m \times n$ con $m \geq n$, la solución del sistema de ecuaciones $\mathbf{Ax} = \mathbf{b}$, con \mathbf{x} y \mathbf{b} siendo vectores columna de $n \times 1$ elementos, es sólo una aproximación $\mathbf{p} = \mathbf{A} \mathbf{A}^\dagger \mathbf{b}$. La solución \mathbf{p} es la solución con error mínimo y es la proyección perpendicular de \mathbf{b} en el espacio lineal generado por el conjunto de vectores columna de la matriz rectangular \mathbf{A} de $m \times n$ con $m \geq n$.

El desarrollo de la pseudoinversa \mathbf{A}^\dagger es el resultado de un manejo algebraico de la matriz \mathbf{A} , como se

demuestra en la ecuación (2). Con el cálculo tensorial, se puede arrojar mucha luz sobre el tema general de las inversas generalizadas. Por ejemplo, la aproximación \mathbf{p} es reconocida inmediatamente como la solución covariante del problema $\mathbf{Ax} = \mathbf{b}$.

Una segunda solución $\mathbf{p}' = \mathbf{A}^{-T} \mathbf{A}^T \mathbf{b}$, llamada solución contravariante, puede encontrarse fácilmente, de tal forma que la norma de la proyección de \mathbf{b} sobre el espacio generado por las columnas de \mathbf{A} se obtiene al efectuar el producto escalar de \mathbf{p} y \mathbf{p}' i.e. el producto escalar de las soluciones covariante y contravariante.

Los nuevos resultados de coordenadas oblicuas de Urrutia (1991-1992) en el contexto de dinámica (cuando ellos se aplican a inversas generalizadas), nos permiten reconocer a las columnas de la matriz \mathbf{A} como un espacio covariante, y a la pseudoinversa \mathbf{A}^\dagger como el correspondiente espacio contravariante. De acuerdo con Urrutia (1992), cuando estos dos espacios son conocidos, y utilizando el Análisis de Fourier Generalizado (Urrutia 1991 y 1992) en coordenadas oblicuas, nos conducirá a dos soluciones \mathbf{p} y \mathbf{p}' que son las mejores soluciones posibles (con error mínimo) que se pueden generar con las columnas de la matriz \mathbf{A} y con los renglones \mathbf{A}^\dagger , una matriz de $n \times m$ elementos. Suponemos que los vectores columna de \mathbf{A} son linealmente independientes. Con respecto a esto, se mostrará una condición que asegurará la existencia de la pseudoinversa \mathbf{A}^\dagger . La concepción geométrica de este problema, presentada ya en otros trabajos por Urrutia (1991-2001) y en problemas de dinámica estructural, permite encontrar una solución en tal forma que el problema $\mathbf{Ax} = \mathbf{b}$ para una matriz rectangular \mathbf{A} de orden $m \times n$ con $m \geq n$ puede ser resuelto en forma cerrada, con error cero.

Preliminares

Una discusión a fondo sobre la existencia de coordenadas covariantes y contravariantes, se encuentra en Urrutia (1991-1992). No obstante, se presenta un breve resumen de la teoría principal; y los aspectos físicos y matemáticos de las aproximaciones \mathbf{p} se ilustran con ejemplos.

Antes de introducirnos en la presentación de inversas generalizadas, se definirán varias operaciones como corresponde a los casos de sistemas discretos, vectores y matrices, sistemas continuos, funciones y operadores lineales, con el objeto de cubrir ambos temas en una presentación.

El producto interno de dos funciones coordenadas discretas ϕ_n y ϕ_m y la norma energética definida por Urrutia (1998) de las mismas funciones coordenadas con respecto al operador K_{nm} están definidas por las siguientes ecuaciones, respectivamente

$$\begin{aligned} \langle \phi_n, \phi_m \rangle &= \phi_n^T \phi_m \\ \langle \phi_n, K \phi_m \rangle &= \sum_{n=1}^N \phi_n^T K_{nm} \phi_m \end{aligned} \quad (3)$$

Donde ϕ_m denota vectores columna, ϕ_n^T es el vector transpuesto de ϕ_n y K_{nm} es una transformación que representa algún problema en particular.

El producto interno entre dos funciones y la norma energética definida por Mikhlin (1964) de las mismas funciones ϕ_n en un dominio Ω con respecto al operador K_{nm} está definida por las siguientes ecuaciones

$$\begin{aligned} \langle \phi_n, \phi_m \rangle &= \int_{\Omega} \phi_n(x) \phi_m(x) dx \\ \langle \phi_n, K \phi_m \rangle &= \sum_{n=1}^N \sum_{m=1}^M \int_{\Omega} \phi_n(x) K_{nm} \phi_m(x) dx \end{aligned} \quad (4)$$

A pesar de sus diferentes formas, las últimas cuatro ecuaciones denotan un proceso de integración, vía el producto interno o escalar.

Bases covariantes y contravariantes

Definamos dos espacios coordenados en un dominio Ω constituidos por un conjunto de funciones base oblicuas covariantes ϕ_n , y un segundo conjunto de funciones base contravariantes ϕ^m , de tal manera que el producto escalar entre ellos conduzca al símbolo de Kronecker δ_n^m como sigue:

$$\langle \phi_n, \phi^m \rangle = \delta_n^m = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \quad (5)$$

En donde $\langle \phi_n, \phi^m \rangle$ denota el producto escalar de vectores o funciones como en las ecuaciones (3) y (4). En lo siguiente siempre nos referiremos a vectores y matrices; sin embargo, el lector debe traducir simultáneamente la misma forma de pensar para funciones y operadores lineales.

Un vector cualquiera \tilde{v} puede descomponerse en dos espacios coordenados ϕ_n y ϕ^m como sigue

$$\begin{aligned} \tilde{v} &= \sum_{n=1}^{\infty} C_n \tilde{\phi}^n \\ \tilde{v} &= \sum_{n=1}^{\infty} C^n \tilde{\phi}_n \end{aligned} \quad (6)$$

con C_n y C^n representando las componentes covariantes y contravariantes de \tilde{v} . Cualquier vector se puede descomponer en bases covariantes o contravariantes ϕ_n y ϕ^m ; por lo tanto, un vector base covariante ϕ_n puede descomponerse en bases contravariantes ϕ^m como lo indica Flügge (1972)

$$\tilde{\phi}_n = \phi_{n1} \tilde{\phi}^1 + \phi_{n2} \tilde{\phi}^2 + \phi_{n3} \tilde{\phi}^3 + \dots etc \quad (7)$$

Que en notación tensorial se escribe

$$\tilde{\phi}_n = \phi_{nm} \tilde{\phi}^m \quad (8)$$

En general $\phi_{nm} = \phi_{mn} = \langle \phi_n, \phi_m \rangle$ es el conocido tensor métrico covariante. En forma similar, la descomposición del vector base contravariante ϕ^m es posible

$$\tilde{\phi}^n = \phi^{nm} \tilde{\phi}_m \quad (9)$$

En la última ecuación ϕ^{nm} es el tensor métrico contravariante de cálculo tensorial. Es fácil escoger un conjunto arbitrario y completo (o incompleto) de vectores base covariantes ϕ_n y con ellos podemos generar el tensor métrico $\phi_{nm} = \langle \phi_n, \phi_m \rangle$. Dada una matriz A de orden $m \times n$ con $m \geq n$, y adoptando el conjunto de n vectores columna $\{\phi_1, \phi_2, \dots, \phi_n\}$ de orden $m \times 1$, podemos llamar a este conjunto el espacio covariante S .

El conjunto de vectores $\{\phi_1, \phi_2, \dots, \phi_n\}$ es linealmente dependiente si hay una relación lineal como sigue

$$r^1 \tilde{\phi}_1 + r^2 \tilde{\phi}_2 + \dots + r^n \tilde{\phi}_n = 0 \text{ para algunos } r_i \neq 0 \quad (10)$$

Si este no es el caso, entonces el conjunto $\{\phi_1, \phi_2, \dots, \phi_n\}$ es linealmente independiente. Definiremos el conjunto de vectores columna de todas las matrices A de orden $m \times n$ con $m > n$ como el conjunto de vectores columna $\{\phi_1, \phi_2, \dots, \phi_n\}$ linealmente independiente.

Los elementos del producto matricial $(A^T A)$ son iguales a los elementos del tensor métrico $\phi_{nm} = \langle \phi_n, \phi_m \rangle$. Con este conjunto $\{\phi_1, \phi_2, \dots, \phi_n\}$ y con el tensor métrico covariante $\phi_{nm} = \langle \phi_n, \phi_m \rangle$ conocido, los elementos del tensor métrico contravariante ϕ^{nm} se pueden calcular de la siguiente ecuación

$$\begin{aligned} \phi_{11}\phi^{m1} + \phi_{12}\phi^{m2} + \phi_{13}\phi^{m3} + \dots &= \delta_1^m \\ \phi_{21}\phi^{m1} + \phi_{22}\phi^{m2} + \phi_{23}\phi^{m3} + \dots &= \delta_2^m \\ &\dots\dots\dots \\ &etc. \end{aligned} \quad (11)$$

Para ilustrar el uso de la ecuación (11) supongamos que cada uno de los espacios lineales $\tilde{\phi}_n, \tilde{\phi}^n$ tiene solamente tres componentes. Entonces la ecuación (11) nos proporcionará tres sistemas de ecuaciones. Si hacemos que $m = 1$ en la ecuación (11) obtenemos el siguiente sistema de ecuaciones

$$\begin{aligned} \phi_{11}\phi^{11} + \phi_{12}\phi^{12} + \phi_{13}\phi^{13} &= 1 \\ \phi_{21}\phi^{11} + \phi_{22}\phi^{12} + \phi_{23}\phi^{13} &= 0 \\ \phi_{31}\phi^{11} + \phi_{32}\phi^{12} + \phi_{33}\phi^{13} &= 0 \end{aligned}$$

Que en forma matricial se escribe

$$\begin{bmatrix} \phi_{11}\phi_{12}\phi_{13} \\ \phi_{21}\phi_{22}\phi_{23} \\ \phi_{31}\phi_{32}\phi_{33} \end{bmatrix} \begin{bmatrix} \phi^{11} \\ \phi^{12} \\ \phi^{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (11a)$$

En forma similar cuando hacemos $m = 2$ y 3 obtenemos

$$\begin{bmatrix} \phi_{11}\phi_{12}\phi_{13} \\ \phi_{21}\phi_{22}\phi_{23} \\ \phi_{31}\phi_{32}\phi_{33} \end{bmatrix} \begin{bmatrix} \phi^{21} \\ \phi^{22} \\ \phi^{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (11b)$$

Y finalmente a

$$\begin{bmatrix} \phi_{11}\phi_{12}\phi_{13} \\ \phi_{21}\phi_{22}\phi_{23} \\ \phi_{31}\phi_{32}\phi_{33} \end{bmatrix} \begin{bmatrix} \phi^{31} \\ \phi^{32} \\ \phi^{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (11c)$$

De donde se calculan todos los elementos de ϕ^{nm} del tensor métrico contravariante de orden 3×3 .

Con las métricas covariantes y contravariantes disponibles, los vectores base contravariantes pueden calcularse como sigue

$$\tilde{\phi}^n = \phi^{nm} \tilde{\phi}_m \quad (12)$$

Pseudoinversa de A

Definimos ahora un nuevo conjunto $B = \{\phi^1, \phi^2, \dots, \tilde{\phi}^n\}$ formado por los vectores base contravariantes dados en la ecuación (12). Sabemos que el producto escalar de $\tilde{\phi}_n$ y $\tilde{\phi}^m$ es igual a δ_n^m , i.e. la delta de Kronecker, ecuación (5). En la ecuación (12) nos damos cuenta de que los vectores columna de la matriz B , siendo combinaciones lineales de los vectores columna $\tilde{\phi}_n$, son de orden $m \times 1$. Cuando la siguiente multiplicación matricial se realiza $B^T A$, se obtiene la matriz identidad I . Es evidente que B^T es idéntica a A^\dagger , i.e. a la llamada pseudoinversa A^\dagger de la matriz A de orden $m \times n$, con $m \geq n$. A^\dagger se identifica ahora como la traspuesta de $B = \{\phi^1, \phi^2, \dots, \phi^n\}$ formado por vectores base contravariantes. Observamos que la condición dada en la ecuación (2) $A^\dagger = (A^T A)^{-1} A^T$ coincide con el algoritmo para calcular la B^T que incluye a los vectores base contravariantes como renglones.

Inversa A^{-1} , el caso general

Teorema 1

Dada una matriz A de orden $m \times n$ con $m \geq n$, y el conjunto $\{\phi_1, \phi_2, \dots, \phi_n\}$ formado por sus vectores covariantes columna $\tilde{\phi}_n$, linealmente independientes, pero no necesariamente ortogonales, existe entonces una inversa única A^{-1} formada por la traspuesta de la matriz $B = \{\phi^1, \phi^2, \dots, \phi^n\}$ de vectores base columna

contravariante, calculados mediante la ecuación (12). El caso de cuando $m=n$ para matrices cuadradas, es únicamente un caso particular.

Teorema 2

Dada una matriz A de orden $m \times n$ con $m \geq n$, el conjunto $\{\phi_1, \phi_2, \dots, \phi_n\}$, de sus vectores columna, será linealmente independiente si, y sólo si el determinante del tensor métrico ϕ_{mn} es diferente de cero. El tensor métrico siempre es cuadrado.

Ejemplo 1

Dados dos vectores base covariantes $\tilde{\phi}_1$ y $\tilde{\phi}_2$, encuentre la combinación lineal que mejor se aproxime al vector $\tilde{b}^T = (2, 2, 2)$

$$\tilde{\phi}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \tilde{\phi}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \tilde{b} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (13)$$

Solución. En forma vectorial buscamos resolver la siguiente ecuación

$$C^1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + C^2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (14)$$

en forma matricial se escribe como sigue

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C^1 \\ C^2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

Los elementos del tensor métrico ϕ_{mn} se calculan así:

$$\phi_{11} = (1, 1, 0) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 2$$

En forma similar el resto de los elementos $\phi_{mn} = \langle \tilde{\phi}_m^T, \tilde{\phi}_n \rangle$ se calculan al obtener el siguiente tensor métrico ϕ_{mn}

$$Q_{mn} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

El tensor métrico contravariante se logra por la ecuación (11) que lleva a las siguientes ecuaciones

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \phi^{11} \\ \phi^{12} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \phi^{21} \\ \phi^{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Cuando éstas se resuelven encontramos que ϕ^{mn} es igual a

$$\phi^{mn} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Los vectores base contravariantes $\tilde{\phi}^n$ se calculan con la ecuación (12) dando

$$\tilde{\phi}^1 = \phi^{11} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \phi^{12} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \therefore \rightarrow \tilde{\phi}^1 = 1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (15)$$

En forma similar, el vector base $\tilde{\phi}^2$ es calculado así

$$\tilde{\phi}^2 = \phi^{21} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \phi^{22} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \therefore \rightarrow \tilde{\phi}^2 = -1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad (16)$$

En la figura 1 se grafican las bases $\tilde{\phi}^n$ y $\tilde{\phi}_n$ junto con el vector objetivo $\tilde{b}^T = (2, 2, 2)$. Multiplicando la ecuación (14) por $\tilde{\phi}^1$

$$(1 \ 0 \ 0) \left[C^1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + C^2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right] = (1 \ 0 \ 0) \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (17)$$

Cuando realizamos la multiplicación escalar encontramos que $C^1 = 2$. Si en la ecuación (17) reemplazamos $\tilde{\phi}^1$ por $\tilde{\phi}^2$ obtenemos $C^2 = 0$. Nuestra aproximación a $\tilde{b}^T = (2, 2, 2)$ se puede escribir

$$p = 2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \therefore \rightarrow p = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (18)$$

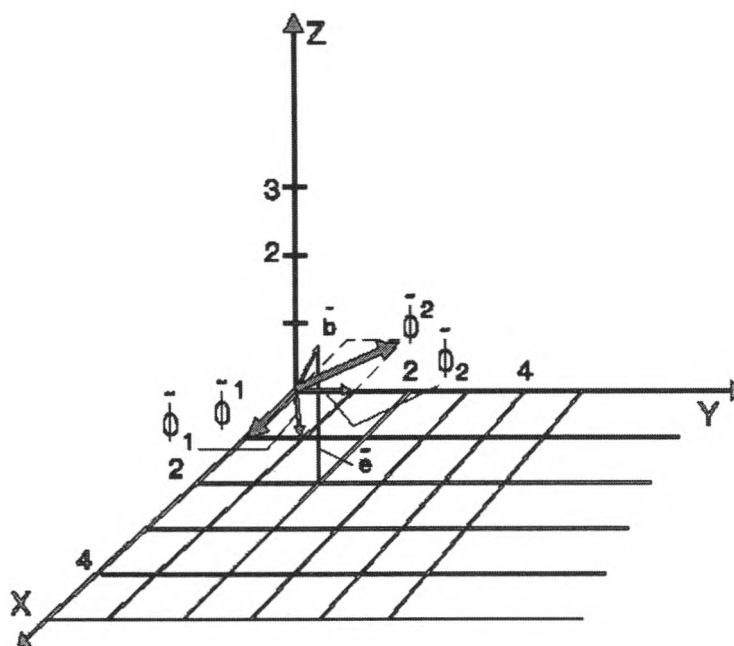


Figura 1

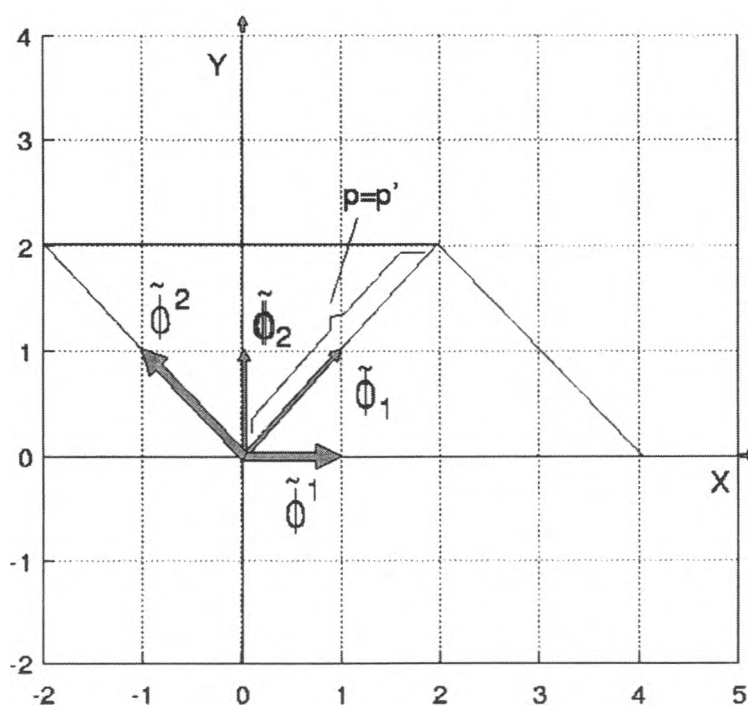


Figura 2

En las figuras 1 y 2 se observa que la proyección de \tilde{b} es dos veces el vector base covariante $\tilde{\phi}_1$ y ninguna proyección del vector $\tilde{\phi}_2$ contribuye a la solución.

La aproximación de la ecuación (18) se denomina aproximación covariante (con elementos $\tilde{\phi}_n$). Recurriendo al conjunto de bases contravariante $\tilde{\phi}^n$, existe otra solución.

Solución contravariante

Una aproximación más es posible para $\tilde{b}^T = (2, 2, 2)$ mediante una superposición lineal de bases contravariantes

$$C_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (19)$$

Si multiplicamos esta ecuación por $\phi_1^T = (1, 1, 0)$ obtenemos

$$(1 \ 1 \ 0) \left[C_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \right] \quad (20)$$

Cuando se efectúa la multiplicación escalar, se tiene que $C_1 = 4$. Si en la ecuación (20) reemplazamos $\tilde{\phi}_1^T$ por $\tilde{\phi}_2^T = (0, 1, 0)$ entonces $C_2 = 2$, esta nueva solución se ve claramente en la siguiente ecuación

$$p' = 4 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \therefore \rightarrow p' = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad (21)$$

En el ejemplo, esto significa que \tilde{b} puede ser aproximado superponiendo cuatro veces el vector base $\tilde{\phi}^1$ más dos veces el vector base $\tilde{\phi}^2$, del tal forma que se satisface **la ley del paralelogramo de composición de fuerzas de física**. Este requisito físico se observa en la figura 2 y se satisface con las descomposiciones covariante y contravariante de cualquier vector \tilde{b} .

Las soluciones covariante y contravariante conducen a idénticas aproximaciones p . Falta la norma de la solución (o de la aproximación). De acuerdo a Urrutia (1991, 1992 y 1998), cuando las descomposiciones covariante y contravariante (ecuación 6) de cualquier vector \tilde{v} se calculan, el tamaño o norma del vector \tilde{v} se obtiene de la siguiente manera

$$\|\tilde{v}\| = \left[\sum_{n=1}^N C_n C^n \right]^{1/2}$$

De las ecuaciones (18) y (21) las normas de las aproximaciones p , p' y del vector objetivo \tilde{b} son, respectivamente

$$\|\tilde{v}\| = \left[\sum_{n=1}^2 C_n C^n \right]^{1/2} = [2 \times 4 + 0 \times 2]^{1/2} = 2.828$$

$$\|\tilde{b}\| = [2^2 + 2^2 + 2^2]^{1/2} = 3.464$$

$$\text{Diferencia} = 0.636 \text{ en norma}$$

con un error en norma aproximadamente de 18.4% más pequeño que el vector original \tilde{b} .

Note de las figuras 1 y 2 que los vectores de aproximación p y p' (que son colineales) incluyen el mismo error $\tilde{\epsilon}$ de longitud mínima. Cuando obtenemos la diferencia $\tilde{\epsilon}^T = (0, 0, 2)$ entre \tilde{b} y p (o p') podemos proponer las siguientes dos soluciones para aproximar \tilde{b} .

$$2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

con vectores base covariantes $\tilde{\phi}_n$, o en la siguiente forma

$$4 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

con vectores base contravariantes $\tilde{\phi}^n$ ambas soluciones podrían ser admisibles bajo ciertas circunstancias.

Ejemplo 2

Se toma este problema de Tucker (1988), para mostrar una propuesta que mejore la solución anotada para un problema relacionado con los

niveles de producción de ciertos productos químicos en una refinería petrolera.

Problema

Encontrar la mejor solución al siguiente sistema de ecuaciones

$$\begin{aligned} 20x_1 + 4x_2 &= 500 \\ 10x_1 + 14x_2 &= 850 \\ 5x_1 + 5x_2 &= 1000 \end{aligned} \quad (22)$$

En forma vectorial, el problema se puede escribir como sigue

$$x_1 \begin{pmatrix} 20 \\ 10 \\ 5 \end{pmatrix} + x_2 \begin{pmatrix} 4 \\ 14 \\ 5 \end{pmatrix} = \begin{pmatrix} 500 \\ 850 \\ 1000 \end{pmatrix} \quad (23)$$

En forma matricial se ve así

$$\begin{pmatrix} 20 & 4 \\ 10 & 14 \\ 5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 500 \\ 850 \\ 1000 \end{pmatrix} \quad (24)$$

$$A\tilde{x} = b$$

donde las columnas de la matriz A definen un espacio covariante con sólo dos vectores. De la ecuación (12) o Tucker (1998), se tiene la inversa de A

$$A^{-1} = \begin{pmatrix} 0.0584 & -0.0164 & -0.0006 \\ -0.0435 & 0.0761 & 0.0217 \end{pmatrix} \quad (25)$$

lo cual, incluye a los vectores base contravariantes de A como renglones. Ahora, sólo se presenta la inversa de cualquier matriz A , evitando su cálculo numérico para ser concisos. Si se multiplica cada columna de (23) por el primer renglón de (25) y en seguida por el segundo renglón, se tiene que

$$\begin{aligned} x_1 &= 14.6 \\ x_2 &= 64.7 \end{aligned}$$

que permite transformar (23) en la siguiente aproximación

$$\tilde{p} = 14.6 \begin{pmatrix} 20 \\ 10 \\ 5 \end{pmatrix} + 64.7 \begin{pmatrix} 4 \\ 14 \\ 5 \end{pmatrix} = \begin{pmatrix} 551 \\ 1051 \\ 394 \end{pmatrix} \approx \begin{pmatrix} 500 \\ 850 \\ 1000 \end{pmatrix} \quad (26)$$

de acuerdo con Tucker (1998) este resultado $\tilde{p}^T = (551, 1051, 394)$ se considera como una mala aproximación. Sin embargo, no es la única solución disponible, sino la mejor que se puede generar con los vectores base covariantes de la matriz A y con los vectores renglón contravariantes de la inversa A^{-1} , que están contenidos en el mismo hiperplano, como se puede ver en la figura 3. A partir de ahora se considera el problema desde un nuevo punto de vista y se puntualiza una observación de peso.

Los elementos de los vectores columna de la matriz A , en la ecuación (24), representan los niveles de producción (de una refinería) de diferentes productos de un barril de petróleo crudo. En la primera columna, 20, 10 y 5 denotan el número de unidades de diesel, aceite lubricante y gasolina, producidos a partir de un barril de petróleo crudo. Los elementos de la columna 2 de A son los números correspondientes de diesel, aceite lubricante y gasolina, producidos por barril de petróleo de una segunda refinería.

De la ecuación (26), como se estableció por Tucker (1988), es claro que a pesar de la sobreproducción de diesel y aceite lubricante, la demanda de gasolina no puede satisfacerse en una situación real, por lo cual debe ser puesta en marcha alguna medida práctica para modificarse.

Con los niveles actuales de producción, no hay forma de satisfacer la demanda actual. Por tanto, es necesario modificar los niveles de producción en ambas refinerías. El problema es mantener las modificaciones de los niveles normales de producción en un mínimo. También podríamos construir una nueva refinería con una política diferente de producción que cubriera el error de producción de las otras dos refinerías; sin embargo, esta sería una alternativa muy costosa, por lo que se prefiere la primera opción. De la ecuación (26), el vector error es el siguiente

$$\tilde{e} = \begin{pmatrix} -51 \\ -201 \\ 606 \end{pmatrix} \quad (27) \quad \text{objeto de superar esta situación se normaliza el vector objetivo } \tilde{b} \text{ para obtener}$$

Es claro que este vector es paralelo al vector \tilde{n}

$$\tilde{b}_n = \begin{pmatrix} 0.35601 \\ 0.60522 \\ 0.71202 \end{pmatrix} \quad (29)$$

$$\tilde{n} = \begin{pmatrix} -0.07963 \\ -0.31382 \\ 0.94614 \end{pmatrix} \quad (28) \quad \text{Con los vectores } \tilde{b}_n \text{ y } \tilde{n} \text{ se obtiene la siguiente matriz auxiliar } A^*$$

que es perpendicular a los vectores columna \tilde{c}_1, \tilde{c}_2 de la matriz A . Ahora, de la figura 4 podemos ver que la aproximación \tilde{p} dada por la ecuación (26), el vector objetivo \tilde{b} y el vector normal \tilde{n} caen en un mismo plano, que es perpendicular al plano generado por \tilde{c}_1 y \tilde{c}_2 .

$$A^* = \begin{pmatrix} 0.35601 & -0.07963 \\ 0.60522 & -0.31524 \\ 0.71202 & 0.94573 \end{pmatrix} \quad (30)$$

cuya inversa esta dada por

De esta figura se puede observar que no hay ninguna combinación lineal de \tilde{c}_1 y \tilde{c}_2 que se pueda hacer para generar el vector objetivo \tilde{b} . Con el

$$A^{*-1} = \begin{pmatrix} 0.49386 & 0.94341 & 0.35562 \\ -0.30329 & -0.74405 & 0.78408 \end{pmatrix} \quad (31)$$

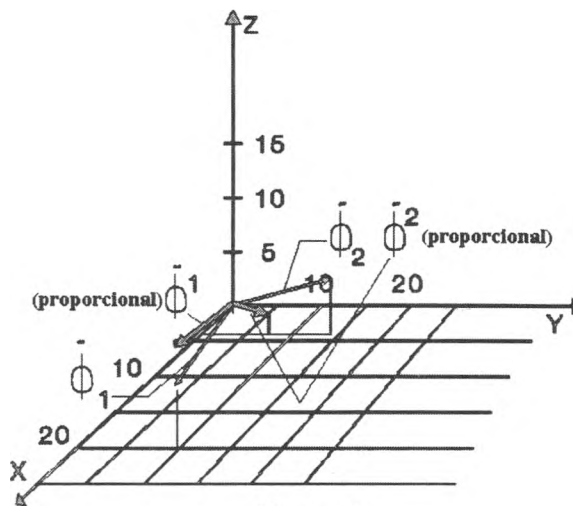


Figura 3

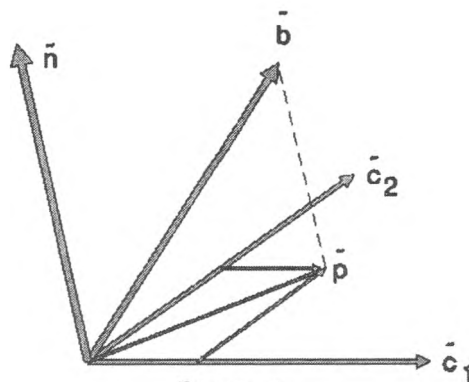


Figura 4

Se puede ver que el vector \tilde{b}_n y el segundo renglón de la matriz A^{*-1} son ortogonales. Cuando se normaliza el segundo renglón de A^{*-1} se obtiene el vector $\tilde{N}_T = (-0.27015, -0.66275, 0.69841)$ que es normal al origen del plano

$$-0.27015X - 0.66275Y + 0.69841Z = 0 \quad (32)$$

Se puede verificar que tanto el vector objetivo $\tilde{b}^T = (500, 850, 1000)$, como el vector arbitrario $\tilde{r}^T = (1, 1, 1.33582)$, satisfacen la ecuación del plano, porque ambas están contenidas en el plano dado por la ecuación (32). En este punto notamos que el plano generado por \tilde{c}_1 y \tilde{c}_2 , y el plano obtenido por la ecuación (32), son perpendiculares al plano formado por \tilde{n} y \tilde{b} , y su intersección pasa por el origen y hace un ángulo θ como se indica en la figura 5. Los vectores generadores $\tilde{c}_1^T = (20, 10, 5)$ y $\tilde{c}_2^T = (4, 14, 5)$, que contienen los niveles de producción de diesel, aceite lubricante y gasolina, tendrán que ser modificados, de tal forma que los vectores que representan tales niveles de producción estén localizados dentro del plano de la ecuación (32). De estas observaciones, vemos que las mejores aproximaciones a $\tilde{c}_1^T = (20, 10, 5)$ y $\tilde{c}_2^T = (4, 14, 5)$ deben satisfacer la ecuación (32), y la distancia de la punta de esos vectores y el plano

(32) debe de ser mínima. Esas dos aproximaciones son las proyecciones normales de $\tilde{c}_1^T = (20, 10, 5)$ y $\tilde{c}_2^T = (4, 14, 5)$ sobre el plano (32). Para obtener estas proyecciones se necesitan conocer dos vectores generadores, por ejemplo $\tilde{r}^T = (1, 1, 1.33582)$ y $\tilde{s}^T = (5, 8.5, 10)$ un vector 100 veces más pequeño que el vector objetivo $\tilde{b}^T = (500, 850, 1000)$. Esta elección se hizo por conveniencia. Notamos que la elección del vector $\tilde{r}^T = (1, 1, 1.33582)$ es arbitraria y que cualquier otro vector que satisfaga la ecuación (32) sería igualmente útil para proyectar $\tilde{c}_1^T = (20, 10, 5)$ y $\tilde{c}_2^T = (4, 14, 5)$.

Proyección de $\tilde{c}_1^T = (20, 10, 5)$

La cuestión de encontrar la mejor proyección de $\tilde{c}_1 = (20, 10, 5)$ en el plano definido por la ecuación (32), indica que hay que encontrar la combinación lineal de $\tilde{s}^T = (5, 8.5, 10)$ y $\tilde{r}^T = (1, 1, 1.33582)$ que aproxime \tilde{c}_1 , esto se logra al resolver la siguiente ecuación

$$\tilde{C}_1 = X_1 \begin{pmatrix} 5 \\ 8.5 \\ 10 \end{pmatrix} + X_2 \begin{pmatrix} 1 \\ 1 \\ 1.3358 \end{pmatrix} \approx \begin{pmatrix} 20 \\ 10 \\ 5 \end{pmatrix} \quad (33)$$

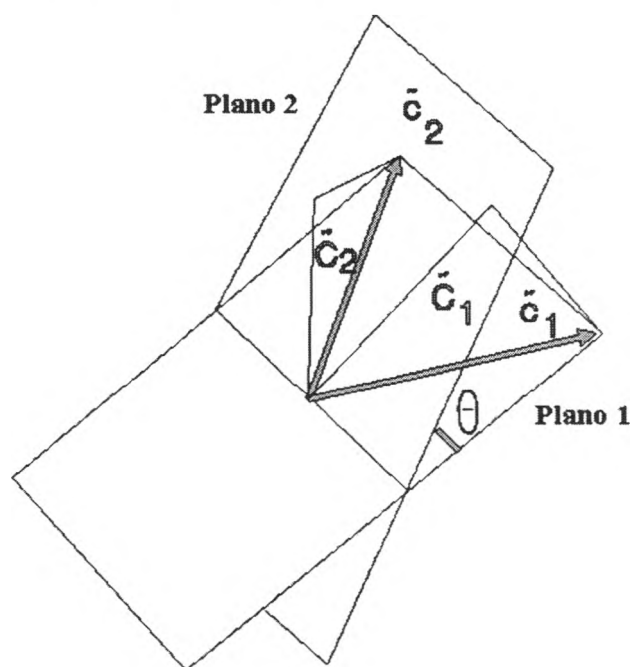


Figura 5

De esta ecuación obtenemos una nueva matriz \mathbf{B} que contiene a $\tilde{\tau}^T = (5, 8.5, 10)$ y $\tilde{\tau}^T = (1, 1, 1.33582)$ como sigue

$$\mathbf{B} = \begin{pmatrix} 5 & 1 \\ 8.5 & 1 \\ 10 & 1.33582 \end{pmatrix} \quad (34)$$

y obtenemos su inversa \mathbf{B}^{-1} como sigue

$$\mathbf{B}^{-1} = \begin{pmatrix} -0.316017 & 0.211369 & 0.0783448 \\ 2.50711 & -0.123588 & -0.203055 \end{pmatrix} \quad (35)$$

Recordemos que los renglones de \mathbf{B}^{-1} son los vectores contravariantes de \mathbf{B} . Por lo tanto, si se multiplica la ecuación (33) por el primer renglón de \mathbf{B}^{-1} , y entonces por el segundo renglón de \mathbf{B}^{-1} , desacoplamos la ecuación y obtenemos $X_1 = -3.81493$, $X_2 = 36.7681$. Cuando reemplazamos estos valores en (33), encontramos la proyección $\tilde{\mathbf{C}}_1$ de $\tilde{\tau}_1^T = (20, 10, 5)$ como sigue, (Figura 3)

$$\tilde{\mathbf{C}}_1 = \begin{pmatrix} 17.698 \\ 4.349 \\ 10.9724 \end{pmatrix} \quad (36)$$

En forma similar, la proyección $\tilde{\mathbf{C}}_2$ de $\tilde{\tau}_2^T = (4, 14, 5)$ es la siguiente

$$\tilde{\mathbf{C}}_2 = \begin{pmatrix} 2.1449 \\ 9.44877 \\ 9.796 \end{pmatrix} \quad (37)$$

Hacemos notar que los nuevos niveles de producción indicados por $\tilde{\mathbf{C}}_1$ y $\tilde{\mathbf{C}}_2$, así como el vector objetivo $\tilde{\mathbf{b}}^T = (500, 850, 1000)$ satisfacen la ecuación (32) y caen dentro del mismo plano. Por tanto, la siguiente ecuación se puede resolver en forma exacta

$$X_1 \begin{pmatrix} 17.698 \\ 4.349 \\ 10.9724 \end{pmatrix} + X_2 \begin{pmatrix} 2.1449 \\ 9.44877 \\ 9.796 \end{pmatrix} = \begin{pmatrix} 500 \\ 850 \\ 1000 \end{pmatrix} \quad (38)$$

Con la política de producción diferente, ahora calculamos los vectores contravariantes de las columnas de la ecuación (38), (o la inversa de la matriz \mathbf{C} formada por las columnas de la ecuación (38)) para obtener la siguiente solución:

$$\begin{pmatrix} 17.698 \\ 4.349 \\ 10.9724 \end{pmatrix} 18.3743 + \begin{pmatrix} 2.1449 \\ 9.44877 \\ 9.796 \end{pmatrix} 81.5015 = \begin{pmatrix} 500 \\ 850 \\ 1000 \end{pmatrix} \quad (39)$$

para el problema de la refinería presentado por Tucker (1988).

Una nota final, es que los nuevos niveles de producción prescritos por $\tilde{\mathbf{C}}_1$ y $\tilde{\mathbf{C}}_2$ son los niveles de producción que requieren los menores cambios posibles en los niveles normales de producción $\tilde{\tau}_1^T = (20, 10, 5)$ y $\tilde{\tau}_2^T = (4, 14, 5)$, esto está indicado en la figura 5.

Conclusiones

Varias conclusiones se pueden obtener de los resultados presentados, pero se iniciará con la inversa generalizada. En relación a que dada una matriz \mathbf{A} de orden $m \times n$ y con $m \geq n$, siempre existe una inversa \mathbf{A}^{-1} si el determinante del tensor métrico (generado con las columnas de la matriz \mathbf{A}) es diferente de cero. Esta afirmación es equivalente a admitir que los vectores columna de la matriz \mathbf{A} son linealmente independientes. Ahora también, se reconoce que los renglones de la matriz inversa \mathbf{A}^{-1} de orden $n \times m$ son los vectores contravariantes del espacio covariante generado por los vectores columna de la matriz \mathbf{A} .

Se puede observar otro hecho importante, en relación con el requisito de que el determinante del tensor métrico debe ser diferente de cero para garantizar la existencia de la inversa de una matriz dada \mathbf{A} de orden $m \times n$ con $m \geq n$. En vista de que su aplicación es general y no está restringida a matrices cuadradas, que el actual criterio de que el determinante de la matriz \mathbf{A} debe ser distinto de

cero se reemplaza por el criterio más general de imponer esa condición al determinante del tensor métrico (un tensor real que no depende ni de la forma ni del tamaño de la matriz \mathbf{A}) como una forma de inferir la existencia o no existencia de la matriz inversa de una matriz dada \mathbf{A} .

Para el problema lineal $\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ con \mathbf{A} con $m \times n$ elementos ($m \geq n$), si la función objetivo está contenida en el hiperplano τ generado por las columnas de la matriz \mathbf{A} , entonces siempre podemos encontrar una solución exacta. Si este hiperplano τ no contiene al vector objetivo, siempre podremos encontrar la proyección perpendicular dada por las bases covariantes o por las bases contravariantes que se obtienen de los renglones de la inversa \mathbf{A}^{-1} . Como se vio en el ejemplo 1, la aproximación obtenida dentro del hiperplano τ con las bases covariante y contravariante, satisface la ley de composición de fuerzas de física. Este hecho es muy importante porque reconoce que el álgebra generada cuando se conoce el espacio covariante (las columnas de la matriz \mathbf{A}) y el espacio contravariante (los renglones de la inversa de \mathbf{A}) satisface las leyes de mecánica y física.

Con respecto al problema de optimización relacionado con el ejemplo 2 se observó que se puede generar una solución óptima mediante la adecuada proyección de los vectores columna de la matriz \mathbf{A} sobre el plano de solución óptima. Para concluir este trabajo, se afirma que no hay ninguna razón para continuar llamando **la Pseudoinversa** \mathbf{A}^{\dagger} de la matriz \mathbf{A} de $m \times n$ elementos con $m \geq n$, ya que simplemente se le debe llamar **"la inversa \mathbf{A}^{-1} de la matriz \mathbf{A} "**.

Referencias

- Ben-Israel A. y Greville T.N.E. (1974). *Generalized Inverses: Theory and Applications*. Wiley.
- Forsythe G.E., Malcolm M.A. y Moler C.B. (1977). *Computer Methods for Mathematical Computations*. Prentice-Hall, Inc., New Jersey.
- Fraleigh J.B. y Beauregard R.A. (1989). *Álgebra lineal*, (traducción al español de *Linear Algebra*). Addison-Wesley, México.
- Flügge W. (1972). *Tensor Analysis and Continuum Mechanics*, Springer-Verlag, Berlin.

- Gill P.E., Murray W. y Wright M.H. (1991). *Numerical Linear Algebra and Optimization*. Addison-Wesley.
- Hall F.J. (1981). *Nonnegative Integral Generalized Inverses, Linear Algebra and its Applications*. Elsevier, North Holland Inc., N.Y.
- Mikhlin S.G. (1964). *Variational Methods in Mathematical Physics*. Translation from the Russian by T. Boddington, Pergamon Press.
- Strang G. (1980). *Linear Algebra and its Applications*. Academic Press, Inc., second edition.
- Tucker A. (1988). *A Unified Introduction to Linear Algebra*. Macmillan Publ. Comp.
- Urrutia-Galicia J.L. (1991). Una introducción sobre la existencia de formas modales covariantes y contravariantes en el análisis dinámico. *SISMODINAMICA*, No.2, 181-191.
- Urrutia-Galicia J.L. (1992). On the Existence of Covariant and Contravariant Modal Forms of Dynamic Analysis. *Transactions CSME (CANADA)*, No.16, 201-217.
- Urrutia Galicia J.L. (2001). *Contributions to Applied Mathematics in Engineering* (texto en inglés), Berichte aus der Konstruktiven Ingenieurbau, Universität der Bundeswehr München (Universidad de las Fuerzas Armadas Federales de Munich, Alemania), ISSN 1431-5122, Nr.01/01.
- Urrutia-Galicia J.L. (1998). *On the Absolute Form of the Theory of Dynamics for Beams, Plates and Shells, with Applications for Engineers. Algebra and Functional Analysis* (texto en inglés). Mitteilungen des Instituts für Statik, der Universität Hannover (Universidad de Hannover, Alemania).

Semblanza del autor

José Luis Urrutia-Galicia. Obtuvo el grado de ingeniero civil en la Facultad de Ingeniería, UNAM en 1975; asimismo, los grados de maestría (1979) y doctorado (1984) en la Universidad de Waterloo, en Ontario, Canadá. Es investigador del Instituto de Ingeniería, UNAM en la Coordinación de Mecánica Aplicada. Sus áreas de interés cubren: matemáticas aplicadas y mecánica teórica, análisis tensorial, estabilidad y vibraciones de sistemas discretos, vigas, placas y cascarones. Ha recibido reconocimientos como el "Premio al Mejor Artículo" de las Transacciones Canadienses de Ingeniería Mecánica (CSME) (Montreal, Canadá 1987) por el artículo "The Stability of Fluid Filled, Circular Cylindrical Pipes, part II Experimental", también le fue otorgada la "Medalla Duggan", que es la más alta distinción otorgada por la CSME (en la universidad de Toronto, Canadá, 1990) por el artículo "On the Natural Frequencies of Thin Simply Supported Cylindrical Shells".