Ingeniería Investigación y Tecnología, ISSN 2594-0732, IV.3, 121-130, 2003
(artículo arbitrado)
DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

Sobre un enfoque vectorial para entender mejor la solución de ecuaciones lineales simultáneas

M.A. Murray-Lasso
Unidad de Enseñanza Auxiliada por Computadora
Departamento de Ingeniería de Sistemas. División de Estudios de Posgrado
Facultad de Ingeniería, UNAM

(recibido: marzo de 2002; aceptado: diciembre de 2002)

Resumen

Se presenta un enfoque alternativo para resolver conjuntos de ecuaciones lineales simultáneas, basado en la idea de considerar el problema como una ecuación entre vectores, los cuales, al premultiplicarlos con punto por un vector simultáneamente ortogonal a todas las columnas, menos una de la matriz de coeficientes, permite despejar cada una de las variables con facilidad. Se establece la relación del método con la Regla de Cramer y se presentan interpretaciones geométricas alternativas a la de la Geometría Analítica. Se comentan las conexiones con problemas de mínimos cuadrados, pseudo inversas, componentes covariantes y contravariantes de un vector visto como tensor y problemas de ecuaciones diferenciales parciales con condiciones en dos o más fronteras.

Descriptores: vector, ortogonal, ecuaciones lineales simultáneas, solución gráfica, coordenadas oblicuas.

Abstract

An alternative viewpoint for solving sets of simultaneous linear equations is presented. It is based on the idea of considering the problem as a single equation involving vectors (represented by the columns of the matrix of coefficients) that when dot multiplied by a vector simultaneously orthogonal to all the columns minus one converts the problem into one which is very easy to solve for the variable corresponding to the column which is not orthogonal to the factor vector. The connection of the method presented and Cramer's Rule is established and geometric interpretation which are alternative to that of Analytic Geometry are presented. Brief comments are made of connections with problems of minimum squares, pseudo inverses, covariant and contravariant components of a vector considered as a tensor and problems involving partial differential equations with conditions on two or more boundaries.

Keywords: vector, orthogonal, simultaneous linear equations, graphic solution, oblique coordinates.

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

Introducción

En la escuela secundaria se enseñan varios métodos para resolver ecuaciones lineales simultáneas, entre ellos están el de eliminación o reducción por igualación o por sustitución, el método gráfico y el método de determinantes (Baldor, 1980). Cualquiera que se utilice obtiene al final del proceso la misma respuesta. Desde el punto de vista didáctico, es importante que exista una variedad de posibilidades, ya que para algunos estudiantes es más fácil la comprensión de un método que otro, y para el entendimiento profundo del tema aporta algo cada uno. También desde el punto de vista computacional, existen diferencias en el volumen de cálculos y en la pérdida de precisión entre los distintos métodos. Aunque esto último hay que tomarlo en consideración en las aplicaciones reales, las cuales pueden involucrar decenas, centenas y millares de variables, en el salón de clase las ilustraciones se hacen con 2, o a lo más 3 ecuaciones e incógnitas. Para dichos problemas se pueden posponer temporalmente las consideraciones mencionadas.

El problema de encontrar la solución de un sistema de ecuaciones lineales es uno de los más importantes de las matemáticas aplicadas, el cual, no obstante la gran cantidad de investigación desarrollada sobre el tema durante siglos, sigue siendo objeto de investigación y estudio, por ejemplo, por el tamaño de las matrices que se presentan en ciertas aplicaciones, las cuales son capaces de domar a las más grandes y veloces supercomputadoras existentes. Dicho problema se presenta en estadística, en dinámica, en el análisis y síntesis de circuitos eléctricos, en el estudio por medio de ecuaciones diferenciales parciales de problemas de hidrodinámica, elasticidad, campos magnéticos, eléctricos y en problemas de mecánica cuántica y de relatividad. Se presenta también el estudio cerca de un punto estable de operación de plantas químicas y reactores nucleares, así como en estudios económicos y administrativos de programación lineal en aplicaciones industriales y gubernamentales. Cualquier función continuamente diferenciable de una o más variables es localmente para incrementos pequeños, adecuadamente representable por medio de un sistema de ecuaciones lineales, aseveración que le otorga a este tipo de sistemas una enorme aplicabilidad en toda clase de disciplinas (Aleksandrov *et al.*, 1969).

En este artículo se presenta un enfoque vectorial para la solución de problemas de ecuaciones lineales simultáneas conceptualmente distinto de los métodos mencionados, que como era de esperarse, ofrece las mismas respuestas que los otros métodos, proporciona intuición geométrica y física adicional que hace que valga la pena examinarlo por su valor conceptual.

Ecuaciones lineales simultáneas desde un punto de vista vectorial

Para mayor claridad comenzamos con un sistema de 2 ecuaciones con 2 incógnitas. Sean las dos ecuaciones

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad (1)$$

que se puede escribir en forma matricial

$$\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b} \qquad (2)$$

En todo el artículo supondremos que la matriz de coeficientes de los sistemas bajo estudio son no singulares, o lo que es lo mismo, que su determinante no es cero, o que sus columnas son vectores linealmente independientes. Si llamamos $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$ a las columnas de la matriz $\boldsymbol{A}$ numeradas de izquierda a derecha, podemos escribir

$$\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 = \boldsymbol{b}$$

la cual es una ecuación entre vectores columna, donde x_1 y x_2 son escalares desconocidos.

La idea fundamental del enfoque del artículo es buscar vectores fila $\boldsymbol{p}$ y $\boldsymbol{q}$ que sean ortogonales a los vectores columna $\boldsymbol{a}_2$ y $\boldsymbol{a}_1$ respectivamente, para que cuando toda la ecuación sea premultiplicada con punto por dichos vectores, uno de los dos

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

términos del lado izquierdo se anule y nos quedemos con una ecuación lineal en una sola variable, la cual es muy fácil de resolver. Supóngase que conocemos el vector fila $\boldsymbol{p}$ que es ortogonal al vector columna $\boldsymbol{a_2}$. Premultiplicando la última ecuación con punto por $\boldsymbol{p}$ obtenemos

$$\boldsymbol{p}.\boldsymbol{a_1} x_1 + \boldsymbol{p}.\boldsymbol{a_2} x_2 = \boldsymbol{p}.\boldsymbol{b}$$

Dado que $\boldsymbol{p}$ es ortogonal a $\boldsymbol{a_2}$ el segundo término del lado izquierdo de la ecuación se anula y quedamos con

$$\boldsymbol{p}.\boldsymbol{a_1} x_1 = \boldsymbol{p}.\boldsymbol{b}$$

Los productos punto son escalares, por lo que la última ecuación es un número por x_1 igual a otro número. La solución es simplemente

$$x_1 = \boldsymbol{p}.\boldsymbol{b} / \boldsymbol{p}.\boldsymbol{a_1} \qquad (3)$$

Si en vez de usar $\boldsymbol{p}$, usamos un vector fila $\boldsymbol{q}$ que sea ortogonal al vector $\boldsymbol{a_1}$, lo que se anula es el primer término del lado izquierdo y podemos despejar x_2 obteniendo

$$x_2 = \boldsymbol{q}.\boldsymbol{b} / \boldsymbol{q}.\boldsymbol{a_2} \qquad (4)$$

Solo resta cumplir con un método simple para obtener vectores que sean ortogonales a $\boldsymbol{a_1}$ y $\boldsymbol{a_2}$. Esto es sencillo si recurrimos a la definición de ortogonalidad. Sean a, b las componentes de un vector columna conocido; se desean obtener las componentes x, y de un vector fila ortogonal. El producto punto se anula, por lo que

$$ax + by = 0$$

de donde

$$x = -b\,y / a$$

pudiéndole dar a y un valor arbitrario. Un par de valores que simplifica las expresiones es

$$x = b, \quad y = -a$$

Podemos expresar este resultado como una regla: Para obtener en dos dimensiones un vector ortogonal a un vector dado, se puede tomar la segunda componente del vector dado y ponerla como primera componente del vector ortogonal, así, tomar la primera componente con signo opuesto y ponerla como la segunda componente del vector ortogonal.

Aplicando la regla a las ecuaciones (3) y (4) obtenemos

$$x_1 = (a_{22} b_1 - a_{12} b_2) / (a_{22} a_{11} - a_{12} a_{21})$$
$$x_2 = (a_{11} b_2 - a_{21} b_1) / (a_{11} a_{22} - a_{21} a_{12})$$

La solución coincide con la que se hubiera obtenido por cualquier otro método, por ejemplo el de determinantes.

Caso de 3 ecuaciones con 3 incógnitas

Las ecuaciones

$$a_{11} x_1 + a_{12} x_2 + a_{13} x_3 = b_1$$
$$a_{21} x_1 + a_{22} x_2 + a_{23} x_3 = b_2$$
$$a_{31} x_1 + a_{32} x_2 + a_{33} x_3 = b_3$$

se pueden poner en forma matricial $\boldsymbol{Ax} = \boldsymbol{b}$

que en forma vectorial se escribe

$$\boldsymbol{a_1} x_1 + \boldsymbol{a_2} x_2 + \boldsymbol{a_3} x_3 = \boldsymbol{b}$$

siendo $\boldsymbol{a_1}$, $\boldsymbol{a_2}$ y $\boldsymbol{a_3}$ las columnas de la matriz $\boldsymbol{A}$ numeradas de izquierda a derecha. Para resolver el sistema buscamos un vector fila $\boldsymbol{p}$ que sea simultáneamente ortogonal a $\boldsymbol{a_2}$ y $\boldsymbol{a_3}$; un vector fila $\boldsymbol{q}$ que sea simultáneamente ortogonal a $\boldsymbol{a_1}$ y $\boldsymbol{a_3}$; y un vector fila $\boldsymbol{r}$ que sea simultáneamente ortogonal a $\boldsymbol{a_1}$ y $\boldsymbol{a_2}$. Premultiplicando con punto la última ecuación por $\boldsymbol{p}$ se anulan el segundo y tercer término del lado izquierdo; premultiplicando por $\boldsymbol{q}$ se anulan el primer y tercer términos; y premultiplicando por $\boldsymbol{r}$ se anulan el primer y segundo términos. En cada caso queda un solo término expresando una ecuación lineal escalar con una sola variable que es muy sencilla de resolver. Se obtienen los resultados

$$x_1 = (\boldsymbol{p}.\boldsymbol{b}) / (\boldsymbol{p}.\boldsymbol{a_1}); \; x_2 = (\boldsymbol{q}.\boldsymbol{b}) / (\boldsymbol{q}.\boldsymbol{a_2}); \; x_3 = (\boldsymbol{r}.\boldsymbol{b}) / (\boldsymbol{r}.\boldsymbol{a_3})$$

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

Sólo resta encontrar los vectores ***p, q*** y ***r*** con las características adecuadas. En tres dimensiones podemos recurrir al producto cruz entre dos vectores cuyo resultado es un vector ortogonal a ambos vectores factor (Thomas, 1960). Por lo tanto, podemos poner

$$\boldsymbol{p} = \boldsymbol{a}_2 \times \boldsymbol{a}_3,\ \boldsymbol{q} = \boldsymbol{a}_1 \times \boldsymbol{a}_3,\ \ \boldsymbol{r} = \boldsymbol{a}_1 \times \boldsymbol{a}_2$$

Como cada uno de estos productos los vamos a multiplicar con punto con otros vectores, se van a formar productos triples escalares de la forma $\boldsymbol{u}.(\boldsymbol{v}\times\boldsymbol{w})$. De la teoría de vectores (Thomas, 1960), recordamos que se pueden intercambiar la cruz y el punto en los productos triples y que se pueden hacer permutaciones cíclicas de los vectores sin cambiar el resultado final, el cual es un escalar que se puede expresar por medio de un determinante como sigue:

$$(\boldsymbol{a}_2\times\boldsymbol{a}3\ ).\ \boldsymbol{a}_1=\boldsymbol{a}_1\ .\ (\boldsymbol{a}_2\times\boldsymbol{a}_3)=(\boldsymbol{a}_1\times\boldsymbol{a}_2)\ .\ \boldsymbol{a}_3=d=$$

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

La igualdad entre los dos últimos miembros se debe a la propiedad de los determinantes que dice que podemos intercambiar filas y columnas de un determinante sin cambiar su valor (Thomas, 1960). Veamos en detalle qué pasa al resolver x_1.

$$x_1=(\boldsymbol{p}.\boldsymbol{b})/(\boldsymbol{p}.\ \boldsymbol{a}_1)=(\boldsymbol{a}_2\times\boldsymbol{a}_3).\boldsymbol{b}/(\boldsymbol{a}_2\times\boldsymbol{a}_3).\boldsymbol{a}_1=\ d_1/d$$

donde d es el determinante de la matriz de coeficientes del sistema de ecuaciones y está dado en la penúltima ecuación, mientras que d_1 es el mismo determinante, pero en lugar de la primera columna, se ha puesto el vector ***b*** de lados derechos.

En forma similar se puede expresar x_2 y x_3 como sigue:

$$x_2=(\boldsymbol{q}.\boldsymbol{b})/(\boldsymbol{q}.\boldsymbol{a}_2)=(\boldsymbol{a}_1\times\boldsymbol{a}_3).\ \boldsymbol{b}/(\boldsymbol{a}_1\times\boldsymbol{a}_3).\boldsymbol{a}_2=d_2/d$$

$$x_3=(\boldsymbol{r}.\boldsymbol{b})/(\boldsymbol{r}.\boldsymbol{a}_3)=(\boldsymbol{a}_1\times\boldsymbol{a}_2).\ \boldsymbol{b}/(\boldsymbol{a}_1\times\boldsymbol{a}_2).\boldsymbol{a}_3=d_3/d$$

En la expresión de x_2 aparecen dos signos negativos en el numerador y el denominador (debido a la alteración del orden cíclico de los vectores) que se cancelan cuando ponemos en el mismo orden los determinantes. El determinante d_2 es igual a d pero con el vector ***b*** colocado en la segunda columna. El determinante d_3 es igual a d pero con el vector ***b*** colocado en la tercera columna. Reconocemos que las expresiones para las variables son las mismas que da la Regla de Cramer (Thomas, 1960) para la solución de un sistema de 3 ecuaciones simultáneas por determinantes.

Sistemas con *n* ecuaciones y *n* incógnitas

El último desarrollo se puede generalizar a cualquier número de ecuaciones con igual número de incógnitas. De la teoría de matrices y determinantes se puede deducir que el vector por el que hay que premultiplicar con punto a la ecuación vectorial para resolver la variable x_k es el vector de cofactores (menores con signo, Thomas,1960) de los elementos de la k-ésima columna de la matriz de coeficientes del sistema. Este vector de cofactores resulta ser simultáneamente ortogonal a todas las columnas de la matriz de coeficientes excepto la k-ésima. Esto se debe a que cuando se evalúa el producto interno entre el vector de cofactores y la j-ésima columna, $j\neq k$, el resultado es el desarrollo de un determinante por menores en función de la j-ésima columna, pero dicho determinante tiene dos columnas idénticas (la j-ésima y la k-ésima) por lo que su valor es cero, de acuerdo a un conocido teorema de determinantes (Thomas, 1960). Cuando se hace el producto punto del vector de cofactores por la k-ésima columna, el resultado es el desarrollo por menores de la k-ésima columna del determinante de la matriz de coeficientes del sistema. Del lado derecho se obtiene el producto punto del vector de cofactores con el vector del lado derecho. Dicho producto punto es igual al desarrollo por menores de la k-ésima columna de un determinante que coincide con el de

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

coeficientes del sistema, excepto que en la k-ésima columna se colocaron los elementos del vector ***b*** del lado derecho. Tenemos entonces que la k-ésima variable x_k está dada por

$$x_k = d_k / d$$

donde $d=$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13}\cdots & a_{1k}\cdots & a_{1n} \\ a_{21} & a_{22} & a_{23}\cdots & a_{2k}\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3}\cdots & a_{nk}\cdots & a_{nn} \end{vmatrix}, \quad d_k = \begin{array}{c} \begin{vmatrix} a_{11} & a_{12} & a_{13}\cdots & b_1\cdots & a_{1n} \\ a_{21} & a_{22} & a_{23}\cdots & b_2\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3}\cdots & b_n\cdots & a_{nn} \end{vmatrix} \\ \uparrow \\ \text{k-ésima columna} \end{array}$$

Esta expresión es la conocida Regla de Cramer para resolver sistemas de n ecuaciones lineales con n incógnitas (Thomas,1960).

No entraremos en detalles aquí, es conveniente que el maestro les informe a sus estudiantes que el evaluar un determinante de más de 3×3 por el desarrollo de términos cuyos signos dependen de la paridad de las permutaciones de índices es numéricamente muy ineficiente, ya que hay que evaluar $n!$ términos cada uno con n factores, cantidad que crece muy rápidamente con n, por lo que se prefieren métodos como la reducción de Gauss para calcular un determinante y la descomposición en factores triangulares de la matriz de coeficientes para resolver sistemas de ecuaciones. El método de Gauss-Jordan se puede considerar como una secuencia de transformaciones que ortogonalizan vectores. Existen otros métodos para hacerlo, entre los que están los métodos de Householder y de Gram-Schmidt (Noble, 1969). El propósito de este artículo no es ofrecer recetas prácticas para resolver problemas del tamaño de los que se presentan en las aplicaciones, sino más bien, introducir conceptos que le proporcionen orientación, intuición y entendimiento al lector interesado en sistemas de ecuaciones lineales. Para incrementar dicha intuición, se presentan ahora algunas interpretaciones geométricas de las ideas presentadas previamente. Para mayor claridad se presentan en dos dimensiones, aunque se harán algunos comentarios para tres dimensiones. Para mayor dimensionalidad se podrían introducir conceptos como hiperplanos y su intersección, así como subespacios ortogonales; sin embargo, no se mostrarán para no intimidar al lector.

Interpretación geométrica

Para ser concretos, se toma el siguiente problema como ilustrativo

$$5x_1 + x_2 = 5$$

$$x_1 + 3x_2 = 4$$

que en forma matrical es

$$\begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Desde el nivel secundaria se imparte la siguiente interpretación geométrica de la solución del sistema dado, la cual está asociada con las filas o renglones de la matriz de coeficientes, ya que las ecuaciones se pueden escribir $a^1.x = 5$, $a^2.x = 4$; donde $a^1 = (5\ \ 1)$, $a^2 = (1\ \ 3)$. El lugar geométrico de los puntos (x_1, x_2) que satisfacen la primera ecuación, es una línea recta cuya posición en un plano cartesiano se puede determinar encontrando un par de puntos de la misma. Lo más simple es encontrar los puntos de cruce con los ejes. Poniendo $x_1 = 0$, se encuentra que $x_2 = 5$, por lo que el punto $(0, 5)$ es uno de los puntos por los que pasa la recta. Poniendo $x_2 = 0$, se encuentra que $x_1 = 1$, por lo que otro punto de la recta es $(1,0)$. Haciendo lo mismo para la segunda ecuación, encontramos que los puntos $(0, 4/3)$ y $(4, 0)$ son de la segunda recta. El punto de cruce, que es el punto $(x_1, x_2) = (11/14, 15/14)$ nos da la solución del sistema como se ilustra en la figura 1. Se puede decir que este es el método de solución de la geometría analítica.

Un método vectorial como el que se presenta en este artículo, escribiría las ecuaciones

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

$$\begin{bmatrix}5\\1\end{bmatrix}x_1 + \begin{bmatrix}1\\3\end{bmatrix}x_2 = \begin{bmatrix}5\\4\end{bmatrix} \qquad (5)$$

Los vectores columna $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$ de la matriz de coeficientes, así como el vector $\boldsymbol{b}$ del lado derecho, están representados en el plano cartesiano por flechas, como se muestran en la figura 2, en la cual se han utilizado letras mayúsculas para denotar los puntos en el plano donde van las puntas de las flechas correspondientes.

La ecuación (5) indica que obedeciendo a la ley del paralelogramo, el vector $\boldsymbol{b}$ es igual a la suma de dos vectores, iguales a las columnas $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$ multiplicadas respectivamente por los escalares x_1 y x_2. Los vectores $\boldsymbol{b}$, $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$ son conocidos y se pueden representar por medio de flechas, como se muestra en la figura 2. Escogemos ejes oblicuos en las direcciones de $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$. A continuación se trazan paralelas a los ejes oblicuos que pasen por el punto B. Sus intersecciones con los ejes oblicuos marcadas con las letras E y F corresponden a los productos $x_1\,\boldsymbol{a}_1$ y $x_2\,\boldsymbol{a}_2$. [En el caso de tres dimensiones se apoyarían en la punta del vector $\boldsymbol{b}$ planos determinados por pares de flechas. Los tres planos se cortarían en el punto B y cada plano cortaría al eje oblicuo diferente del par con el que se genera el plano. Para visualizar esto con mayor facilidad se nota que los tres vectores unitarios

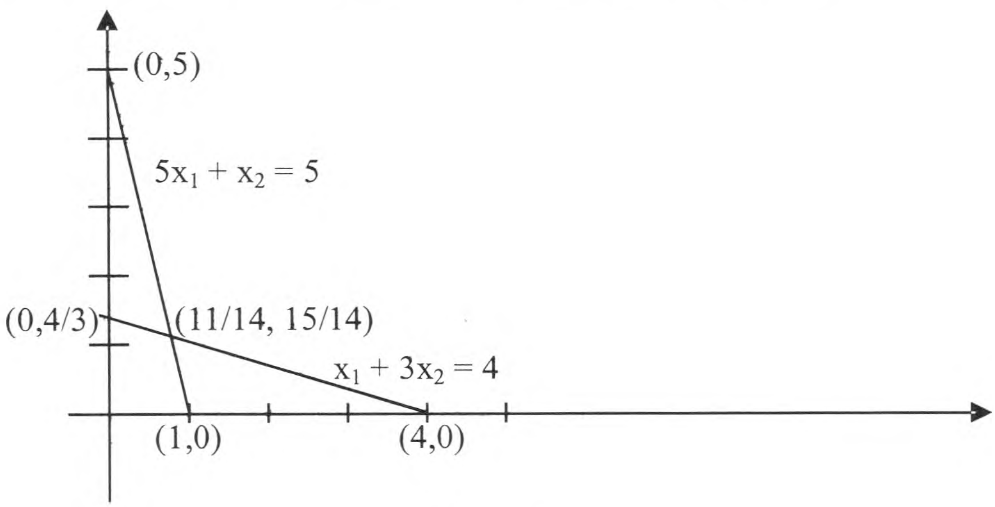

Figura 1

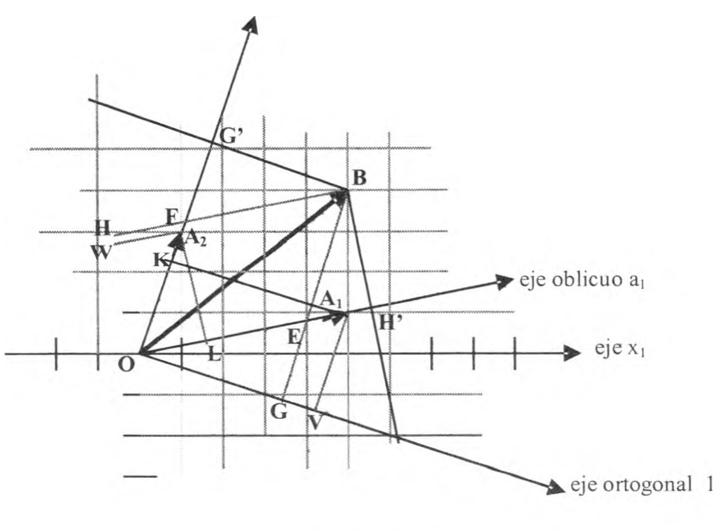

Figura 2

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

tienen como punto común el origen. Así, las puntas de dos flechas y el origen resultan en tres puntos no colineales, los cuales determinan un plano. Este es el plano generado por las dos flechas en cuestión].

Los valores de x_1 y x_2 son por ende iguales a las siguientes razones geométricas:

$$|x_1| = |OE|/|OA_1|,\ |x_2| = |OF|/|OA_2| \qquad (6)$$

La ecuación (6) expresa una relación entre magnitudes, los signos de x_1 y x_2 se determinan de la observación de los sentidos de las flechas que representan a los vectores y las flechas de los ejes. En demostraciones geométricas se acostumbra dibujar la figura de manera que los signos sean positivos (como es nuestro caso) y generalmente no se hace mayor comentario. Lo correcto sería hacer un análisis de todas las posibles posiciones de los elementos "arbitrarios" en la figura para constatar que las mismas relaciones se cumplen en todas las posiciones, o dividir el problema en casos y dar diferentes relaciones para los mismos. Por falta de espacio no se discute esto en detalle y le dejamos al lector la tarea de hacerlo (por ejemplo, poniendo algunos valores negativos en la matriz $\boldsymbol{A}$, u orientando en forma diferente los ejes oblicuos). Dejamos también como ejercicio para el lector hacer el análisis geométrico y verificar que las razones de longitudes dan como resultado $x_1 = 11/14$, $x_2 = 15/14$. Esta interpretación geométrica para hallar x_1 y x_2 es una alternativa diferente (y poco conocida) a la utilizada interpretación de la Geometría Analítica presentada anteriormente. La interpretación de la figura 2 representa una transformación de coordenadas de las originales cartesianas a unas oblicuas paralelas en un proceso al que podríamos llamar <u>Método Vectorial Oblicuo con Coordenadas Paralelas</u> (por la Ley del Paralelogramo). Con esta interpretación estamos suponiendo que el vector OB es un objeto físico que se puede expresar, ya sea en término de las coordenadas cartesianas originales o en términos de las coordenadas oblicuas paralelas, cuyos vectores unitarios son $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$. La matriz que expresa la transformación es la matriz de coeficientes. Dicha transformación queda completamente determinada viendo su efecto sobre cada uno de los vectores unitarios del sistema cartesiano original. El vector unitario $(1\ \ 0)'$ se transforma en el vector $(5\ \ 1)'$ y el vector $(0\ \ 1)'$ se transforma en el vector $(1\ \ 3)'$. Estos son precisamente los dos vectores unitarios del sistema de coordenadas (oblicuo) transformado, y el vector $\boldsymbol{b}$ en el sistema oblicuo debe ser una combinación lineal de dichos vectores. Los coeficientes de dicha combinación lineal son precisamente las desconocidas x_1 y x_2.

Ahora se presenta, utilizando la misma figura 2, la interpretación geométrica del método, que es motivo de este artículo (Thomas, 1960).

$$\boldsymbol{a}.\boldsymbol{b} = |\boldsymbol{a}|\ |\boldsymbol{b}|\ \cos\alpha$$

donde α es el ángulo entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ y las cantidades entre rayas verticales son las longitudes de las flechas correspondientes a los vectores. Por lo tanto, de las ecuaciones (3) y (4) tenemos

$$x_1 = \boldsymbol{p}.\boldsymbol{b} \,/\, \boldsymbol{p}.\boldsymbol{a}_1 = |\boldsymbol{p}|\ |\boldsymbol{b}|\ \cos <\boldsymbol{p},\boldsymbol{b}> / |\boldsymbol{p}|\ |\boldsymbol{a}_1|$$

$$\cos <\boldsymbol{p},\boldsymbol{a}_1> = |\boldsymbol{b}| \cos <\boldsymbol{p},\boldsymbol{b}> / |\boldsymbol{a}_1| \cos <\boldsymbol{p},\boldsymbol{a}_1>$$

Donde hemos adoptado la notación $<\boldsymbol{a},\boldsymbol{b}>$ para denotar el ángulo entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$. Recordamos que el vector $\boldsymbol{p}$ debe ser ortogonal a $\boldsymbol{a}_2$, por lo tanto, coincide con la dirección del <u>eje ortogonal 1</u> que se muestra en la figura 2. El numerador del último miembro de la última ecuación es la proyección del vector $\boldsymbol{b}$ sobre el eje ortogonal 1 y corresponde a la longitud del segmento OG. El denominador es la proyección del vector $\boldsymbol{a}_1$ sobre el eje ortogonal 1 y por lo tanto, es la longitud del segmento OV. Se tiene entonces que

$$x_1 = |OG| \,/\, |OV| \qquad (7)$$

En forma similar se procede para el otro eje (eje ortogonal 2) y se llega al resultado

$$x_2 = |OH| \,/\, |OW| \qquad (8)$$

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

El cálculo de la razón de los segmentos |OG| / |OV| no requiere que se tracen los ejes ortogonales, basta que se bajen líneas ortogonales a los ejes oblicuos como las líneas BH' y BG' pues notamos que son iguales |BH'| y |OH|, y también son iguales |BG'| y |OG|. En igual forma, tirando perpendiculares de la punta de la flecha que representa a los vectores unitarios OA_1 y OA_2 se obtienen segmentos A_1K y A_2L que son iguales a los segmentos OV y OW en los denominadores de las ecuaciones (7) y (8). [En el caso de tres dimensiones se tiran perpendiculares de los puntos B, A_1, A_2 y A_3 a los planos definidos por cada par de vectores oblicuos (columnas de la matriz de coeficientes del sistema de ecuaciones lineales simultáneas) y las cantidades de interés son las distancias de los puntos B, A_1, A_2 y A_3 a dichos planos]. A menudo, a estas cantidades se les conoce como las *coordenadas o componentes ortogonales* del vector B, mientras las oblicuas vistas arriba reciben el nombre de *coordenadas o componentes paralelas*. (En estas interpretaciones también aparecen los términos coordenadas *contravariantes* y *covariantes,* así como transformaciones *cogredietes* y *contragredientes,* conceptos que no se tocarán en este artículo, pero que el lector interesado puede consultar en la literatura técnica (Hoffmann, 1960; Turnbull y Aiken, 1961; Mc Connell, 1957 y Guillemin, 1949).

Es evidente de la figura 2 que se deben obtener iguales respuestas con las coordenadas oblicuas paralelas y con las ortogonales. Se observó en la ecuación (6) que

$$x_1 = |OE| \;/\; |OA_1|$$

y en la ecuación (7)

$$x_1 = \; |OG| \;/\; |OV|$$

Pero los triángulos OA_1V y OEG son semejantes por tener tres ángulos iguales, ya que los ángulos en el vértice O son comunes a los dos triángulos y los ángulos en los vértices A_1 y E, en los vértices V y G son iguales por correspondientes, toda vez que las líneas A_1V y EG son paralelas (ambas ortogonales al eje ortogonal 1). Un argumento análogo se aplica al cálculo de x_2 por la similitud entre los triángulos OWA_2 y OHF.

Conclusiones

En este artículo se ha presentado una interpretación vectorial del proceso de solución de un sistema de ecuaciones lineales simultáneas. Se amplió el panorama de la comprensión del fenómeno al establecer la relación que hay entre el método de los vectores ortogonales y la Regla de Cramer para resolver por determinantes. Para no complicar el análisis, se consideró solamente el caso de igual cantidad de ecuaciones que incógnitas cuando la matriz de coeficientes es no singular. Se compararon la interpretación geométrica de la geometría analítica para hallar la solución de ecuaciones lineales simultáneas (el método gráfico que se imparte en secundaria) con interpretaciones geométricas correspondientes a una transformación de coordenadas cartesianas a coordenadas oblicuas, tanto paralelas como ortogonales. Muchos temas se dejaron en el tintero. Por ejemplo, no se señaló la relación entre los vectores $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$ y las filas de la matriz inversa $\boldsymbol{A}^{-1}$. Consideramos que un alumno que ha comprendido lo que se presenta en este artículo queda en mucho mejores condiciones de entender posteriormente la situación que se presenta cuando se tiene un número de ecuaciones diferente al número de incógnitas. Viendo el asunto como una transformación a un sistema oblicuo, es fácil ver que solamente habrá solución si el vector del lado derecho está en el espacio generado por las columnas de la matriz de coeficientes, independientemente de si el número de ecuaciones es mayor o menor al número de incógnitas. Estos conceptos le permiten visualizar maneras convenientes de obtener la mejor solución aproximada para el caso en que no hay solución, o la más corta de las soluciones cuando hay una infinidad de soluciones, introduciendo conceptos como la *matriz pseudo inversa* de la matriz de coeficientes (Lawson y Hanson, 1974). Todas estas cuestiones tienen importancia en la solución

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

de problemas de mínimos cuadrados, en optimización lineal y cuadrática, así como en el entendimiento de la dualidad en el estudio de la investigación de operaciones y en estadística. El estudiante que asimiló lo que se presentó entenderá las sutiles diferencias entre las componentes contravariantes y covariantes de un vector (considerado como un caso particular de un tensor). Así también, para entender la diferencia entre transformaciones cogredientes y contragredientes y cómo en coordenadas oblicuas el producto interno entre dos vectores o la longitud de un vector deben permanecer invariantes es necesario introducir dos tipos de coordenadas o componentes, cosa que resulta innecesaria si los sistemas de coordenadas son ortonormales, ya que en ese caso las coordenadas paralelas y las ortogonales son iguales. En general, los estudiantes tienen dificultades para asimilar estos conceptos debido a lo abstracto de los tratamientos de los textos que dan definiciones en notaciones complicadas, saturando derivadas parciales con respecto a coordenadas con índices variables, frecuentemente sin dar interpretaciones geométricas claras. El tener un buen entendimiento de estas cosas es importante para los ingenieros que resuelven problemas de dinámica, hidrodinámica, elasticidad, mecánica de medios continuos (como los de mecánica de suelos), campos electromagnéticos y problemas que requieren física moderna. El material aquí presentado es también un útil antecedente para resolver problemas más avanzados que la solución de un sistema finito de ecuaciones lineales simultáneas. Un ejemplo se presenta en la solución de problemas de ecuaciones lineales parciales con condiciones en dos o más fronteras. En ese caso se presentan una infinidad de ecuaciones simultáneas, pero en algunos otros, en las aplicaciones, con definiciones adecuadas de ortogonalidad por medio del producto interno entre vectores. Se pueden hallar familias de vectores (funciones) biortogonales tales que la determinación de los coeficientes desconocidos es una simple extensión de los métodos presentados en este artículo. Tal es el caso de las series de Fourier y los problemas de Sturm-Liouville (Lanczos, 1961).

Dada la importancia que con los avances de la tecnología moderna han adquirido conceptos matemáticos más avanzados que los que se solían enseñar a los estudiantes de ingeniería, tales como los espacios de Hilbert y de Banach, las ecuaciones integrales, el cálculo operacional, el cálculo de variaciones, el análisis funcional, el filtrado y control óptimo, los juegos diferenciales, las funciones generalizadas o distribuciones, los procesos estocásticos y muchos otros temas, es necesario fortalecer la educación matemática de los estudiantes de ingeniería para que dichos conceptos los manejen con soltura e intuición y sean capaces de utilizar herramientas como la computadora digital para aplicárselos a la solución de problemas originales, ya sean físicos, químicos, biológicos, administrativos, logísticos o económicos. Estas matemáticas fortalecidas les permitirán practicar su profesión en el nuevo mundo globalizado en forma más competitiva.

Referencias

Aleksandrov A.D., Kolmogorov A.N. y Lavrent'ev M.A. (1969). *Mathematics: Its Content, Methods, and Meaning,* The M.I.T. Press, Cambridge, MA, Vol. 3, 37–96.

Baldor A. (1980). *Álgebra.* Ediciones y Distribuciones Códice SA., Madrid.

Guillemin E.A. (1949). *The Mathematics of Circuit Analysis*. John Wiley & Sons, Inc., New York.

Hoffmann, B. (1966). *About Vectors.* Prentice-Hall, Inc., Englewood Cliffs, NJ.

Lanczos C. (1961). *Linear Differential Operators*. D. Van Nostrand Company Limited, Londres.

Lawson C. L. y Hanson R.J. (1974). *Solving Least Square Problems*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

McConnell A.J. (1957). *Applications of Tensor Analysis*. Dover Publications, Inc., New York.

Noble B. (1969). *Applied Linear Algebra*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

DOI: http://dx.doi.org/10.22201/fi.25940732e.2003.04n3.010

Thomas G.B. (1960). *Calculus and Analytic Geometry*. 3ª. Edición, Addison-Wesley, Reading, MA.

Turnbull H.W. y Aiken A.C. (1961). *An Introduction to the Theory of Canonical Matrices*. Dover Publications, Inc., New York.

Semblanza del autor

Marco Antonio Murray-Lasso. Realizó la licenciatura en ingeniería mecánica-eléctrica en la Facultad de Ingeniería de la UNAM. El Instituto de Tecnología de Massachussetts (MIT) le otorgó los grados de maestro en ciencias en ingeniería eléctrica y doctor en ciencias cibernéticas. En México, ha laborado como investigador en el Instituto de Ingeniería y como profesor en la Facultad de Ingeniería (UNAM) durante 42 años; en el extranjero, ha sido asesor de la NASA en diseño de circuitos por computadora para aplicaciones espaciales, investigador en los Laboratorios Bell, así como profesor de la Universidad Case Western Reserve y Newark College of Engineering, en los Estados Unidos. Fue el presidente fundador de la Academia Nacional de Ingeniería de México; vicepresidente y presidente del Consejo de Academias de Ingeniería y Ciencias Tecnológicas (organización mundial con sede en Washington que agrupa las Academias Nacionales de Ingeniería) y secretario de la Academia Mexicana de Ciencias. Actualmente es jefe de la Unidad de Enseñanza Auxiliada por Computadora de la División de Estudios de Posgrado de la Facultad de Ingeniería de la UNAM, investigador nacional en ingeniería, consejero educativo del MIT y consultor de la UNESCO.