3

Solución de ecuaciones y cálculo de inversas: métodos

Este capítulo constituye el corazón metodológico del libro; las herramientas de la eliminación de Gauss, las operaciones elementales de renglón y la sustancia hacia atrás se utilizarán virtualmente en cada capítulo subsecuente y deberán, por lo tanto, ser dominadas por completo. El teorema 3.34 es un **teorema clave** *porque es la base de muchos de los argumentos teóricos del capítulo 4. Las secciones 3.6, 3.8 y 3.9 serán de mucho interés para aquéllos que tengan una orientación de tipo computacional, mientras que la materia de la sección 3.5 es de importancia para aquéllos con una orientación de tipo teórico aplicado o computacional. La sección 3.7 se encuentra en algún lugar entre estos dos; sin embargo, se deben conocer los* **teoremas clave 3.48** *y* **3.53** *sobre la descomposición-LU como una alternativa de la eliminación de Gauss.*

3.1 *INTRODUCCION*

En el capítulo 2, se proporcionó amplia evidencia sobre los sistemas de ecuaciones lineales que surgen a partir de la aplicación de las matemáticas en diversas áreas; ese capítulo, al igual que el capítulo 1 –véase por ejemplo, el teorema 1.38 sobre ecuaciones e inversas– demuestran también que las matrices y sus inversas están íntimamente relacionadas con el problema de la solución de sistemas de ecuaciones lineales. En este capítulo se desarrollará la metodología matricial para resolver sistemas de ecuaciones y encontrar inversas; aquí se enfatiza el *cómo*: *cómo* descubrir si un sistema tiene solución o si una matriz tiene inversa; *cómo* encontrar tales soluciones o inversas y así sucesivamente. Entonces en el capítulo 4 se reexaminarán algunos de estos procedimientos y problemas desde un punto de vista teórico para poder desarrollar un fundamento lógico riguroso para los problemas prácticos.

3.2 SOLUCION DE ECUACIONES MEDIANTE LA ELIMINACION DE GAUSS

El método de la eliminación de Gauss es un procedimiento directo pero poderoso para la reducción de sistemas de ecuaciones lineales a una *forma reducida* sencilla que se resuelve fácilmente por sustitución. Observe que muchos autores utilizan "eliminación de Gauss" para referirse tanto a la reducción como a la sustitución que le sigue, mientras que aquí sólo se hace referencia a la reducción. Para introducirlo, se empieza ignorando la notación matricial y escribiendo detalladamente las ecuaciones; más adelante se traducirá lo que se está haciendo al lenguaje matricial. Observe que hay un número de variantes igualmente efectivas en la eliminación de Gauss; se presenta una en detalle y algunas otras se mencionan en problemas posteriores.

Eliminación de Gauss para ecuaciones

Considere el sistema de ecuaciones lineales

$$
\begin{aligned}
-2x_1 + 2x_2 - 4x_3 - 6x_4 &= -4 \\
-3x_1 + 6x_2 + 3x_3 - 15x_4 &= -3 \\
5x_1 - 8x_2 - x_3 + 17x_4 &= 9 \\
x_1 + x_2 + 11x_3 + 7x_4 &= 7.
\end{aligned}
\tag{3.1}
$$

Para aplicar la *eliminacion de Gauss* a (3.1), se utiliza la primera ecuación de (3.1) para eliminar x_1 de las otras tres ecuaciones. Para hacer los cálculos más sencillos, primero se divide la primera ecuación por el coeficiente -2 de x_1, obteniendo de allí una ecuación equivalente en la cual el coeficiente x_1 es 1:

$$x_1 - x_2 + 2x_3 + 3x_4 = 2. \tag{3.2}$$

Después, se utiliza esta ecuación para eliminar x_1 de las últimas tres ecuaciones de (3.1)

> ***Nota.*** Esto se puede hacer de dos formas *aparentemente* diferentes: 1) por sustitución o 2) mediante la combinación de ecuaciones. Por sustitución, se resuelve (3.2) para obtener x_1 en términos de las otras variables y luego sustituir con esta expresión a x_1 en cada una de las otras ecuaciones; por ejemplo, la segunda ecuación de (3.1) se convierte en
>
> $$-3(2 + x_2 - 2x_3 - 3x_4) + 6x_2 + 3x_3 - 15x_4 = -3,$$

que, después de combinar términos, resulta $3x_2 + 9x_3 - 6x_4 = 3$. Con el procedimiento combinado, se agrega a cada una de las ecuaciones siguientes un múltiplo apropiado de la primera, seleccionado de tal modo, que el coeficiente de x_1 se vuelve cero en la nueva ecuación. Por ejemplo, multiplique la primera ecuación nueva (3.2), mentalmente por 3 y sume el resultado a la segunda ecuación de (3.1); obtendrá

$$-3x_1 + 6x_2 + 3x_3 - 15x_4 + (3)(x_1 - x_2 + 2x_3 + 3x_4) = -3 + (3)2,$$

que, después de combinar términos, es tan sólo $3x_2 + 9x_3 - 6x_4 = 3$ exactamente como en el primer procedimiento. Los dos procedimientos *siempre* dan el mismo resultado. Ya que es más fácil de organizar computacionalmente el procedimiento de combinación de ecuaciones y no requiere del intercambio de términos sino de cruzar el signo de igualdad, siempre se usará este método. Por lo tanto,

(3.3) cuando se dice "utilice la ecuación A para eliminar la variable B de la ecuación C", se quiere decir "reemplace la ecuación C por sí misma, más aquel múltiplo de la ecuación A que daría una nueva ecuación que no contendría explícitamente a la variable B".

Volviendo a la eliminación de Gauss en (3.1): la tarea era usar la ecuación (3.2) para eliminar x_1 de cada una de las otras tres ecuaciones. El resultado es

$$\begin{aligned} x_1 - x_2 + 2x_3 + 3x_4 &= 2 \\ 3x_2 + 9x_3 - 6x_4 &= 3 \\ -3x_2 - 11x_3 + 2x_4 &= -1 \\ 2x_2 + 9x_3 + 4x_4 &= 5. \end{aligned} \tag{3.4}$$

Observe que las últimas tres ecuaciones en (3.4) involucran sólo las tres incógnitas x_2, x_3 y x_4 mientras que la primera ecuación permite el cálculo directo de x_1 una vez que se hayan encontrado x_2, x_3 y x_4 de las ecuaciones inferiores.

El siguiente paso es reducir las tres ecuaciones inferiores con tres incógnitas a dos ecuaciones con dos incógnitas. Primero, se divide la nueva segunda ecuación de (3.4) entre su coeficiente 3, con el fin de simplificar la aritmética; esto permite que a la segunda ecuación la reemplace

$$x_2 + 3x_3 - 2x_4 = 1. \tag{3.5}$$

Ahora utilice esta segunda nueva ecuación para eliminar la segunda variable x_2 de las últimas dos ecuaciones –por ejemplo, reemplace la tercera ecuación de

(3.4) por sí misma más tres veces la segunda ecuación más nueva (3.5). Esto reemplaza el conjunto (3.4) de cuatro ecuaciones por

$$\begin{aligned} x_1 - x_2 + 2x_3 + 3x_4 &= 2 \\ x_2 + 3x_3 - 2x_4 &= 1 \\ -2x_3 - 4x_4 &= 2 \\ 3x_3 + 8x_4 &= 3. \end{aligned} \tag{3.6}$$

Ahora se tienen que manejar las últimas dos ecuaciones con dos incógnitas con el mismo proceso. Divida la tercera ecuación de (3.6) entre -2, reemplazando esta ecuación por $x_3 + 2x_4 = -1$; entonces utilice esta ecuación para eliminar x_3 de la cuarta ecuación de (3.6). Después de dividir esta nueva cuarta ecuación entre el nuevo coeficiente 2 encontrado para x_4, se ha reducido finalmente el sistema

$$\begin{aligned} x_1 - x_2 + 2x_3 + 3x_4 &= 2 \\ x_2 + 3x_3 - 2x_4 &= 1 \\ x_3 + 2x_4 &= -1 \\ x_4 &= 3. \end{aligned} \tag{3.7}$$

Esto completa la eliminación de Gauss, dando la *forma reducida* (3.7) de las ecuaciones originales (3.1). La expresión "forma reducida" se define de manera precisa en el capítulo 4; el término describe una forma que resulta de la eliminación de Gauss.

Se utiliza ahora la *sustitución en reversa* para producir la solución de (3.7) como sigue. La última ecuación en (3.7) tiene, clara y exactamente, una solución: $x_4 = 3$. Ahora, para la sustitución en reversa: al sustituir $x_4 = 3$ en la tercera ecuación de (3.7) se tiene

$$x_3 = -1 - 2x_4 = -1 - 2(3) = -7;$$

sustituyendo $x_4 = 3$ y $x_3 = -7$ en la segunda ecuación de (3.7) se obtiene $x_2 = 28$; y una última sustitución en reversa en la primera ecuación de (3.7) da como resultado $x_1 = 35$. Si, como comprobación, se sustituyen estos valores en el sistema *original* (3.1), se observará que, efectivamente, se obtuvo una solución.

Ahora podrá resolver los problemas 1 y 2.

Eliminación de Gauss para matrices aumentadas

Al realizar en el ejercicio anterior la eliminación de Gauss, todos los cálculos se hicieron con los *números* –los coeficientes y el miembro derecho de las ecuaciones– en lugar de usar para ello los símbolos x_i, aunque ciertamente se continuó escribiendo las x_i. Se puede ahorrar una cantidad considerable de escritura trabajando solamente con los números y evitando reescribir las variables todo el tiempo; sólo se debe tener cuidado de no perder la pista de qué números son coeficientes de cuáles incógnitas, o de cuáles corresponden a qué miembro derecho. Las matrices son ideales para esto.

En lugar de manipular las ecuaciones, se opera con la *matriz aumentada* $[\mathbf{A}\ \mathbf{b}]$ que se forma añadiendo la columna $\mathbf{b}$ del lado derecho a la matriz $\mathbf{A}$ de coeficientes, para formar una matriz separada; para el ejemplo anterior la matriz aumentada es

(3.8)
$$\left[\begin{array}{rrrr|r} -2 & 2 & -4 & -6 & -4 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array}\right].$$

El primer paso en la aplicación de la eliminación de Gauss para las ecuaciones en (3.1) fue reemplazar la primera ecuación por sí misma dividida entre -2; en términos de la matriz de coeficientes y de los lados derechos de las ecuaciones, esto significa simplemente reemplazar el primer *renglón*

$$\left[\begin{array}{rrrr|r} -2 & 2 & -4 & -6 & -4 \end{array}\right]$$

por sí mismo dividido entre -2:

$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \end{array}\right].$$

Observe que esto es tan sólo la matriz de números que aparece en la primera ecuación nueva (3.2). El siguiente paso, utilizando el método de *ecuaciones* fue utilizar la nueva primera *ecuación* para eliminar x_1 de las otras tres *ecuaciones*; para matrices, esto significa que se usa el primer *renglón* nuevo para eliminar los elementos en la primera columna de los otros tres *renglones*. Observe que

(3.9) cuando se dice "utilice el renglón A para eliminar el elemento en la columna B del renglón C," se quiere decir "reemplace el renglón C por sí mismo más el múltiplo del renglón A que dará un nuevo renglón con un elemento cero en la columna B";

observe la similitud con el lenguaje en (3.3) utilizado para describir la eliminación de ecuaciones. Al usar el primer renglón nuevo $\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \end{array}\right]$ para eliminar los elementos en la primera columna de los otros tres renglones de (3.8) se reemplaza a la matriz aumentada (3.8) por

(3.10)
$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & 9 & -6 & 3 \\ 0 & -3 & -11 & 2 & -1 \\ 0 & 2 & 9 & 4 & 5 \end{array}\right].$$

Observe que (3.10) es precisamente la matriz aumentada de (3.4). Después se reemplaza el segundo renglón

$$[0 \quad 3 \quad 9 \quad -6 \mid 3]$$

en (3.10) reemplazado por sí mismo y dividido entre 3 –es decir

$$[0 \quad 1 \quad 3 \quad -2 \mid 1]$$

–después utilice ese nuevo segundo renglón para eliminar los elementos en la columna 2 de los últimos dos renglones de (3.10); esto da

(3.11)
$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & -2 & -4 & 2 \\ 0 & 0 & 3 & 8 & 3 \end{array}\right].$$

Observe que (3.11) es precisamente la matriz aumentada de (3.6). Reemplace el tercer renglón de (3.11) con él mismo dividido entre -2, y utilice este nuevo renglón para eliminar el elemento en la columna 3 del último renglón, dando como resultado

(3.12)
$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right].$$

Reemplazando el último renglón con el mismo dividendo entre 2, se obtiene finalmente la matriz reducida

(3.13)
$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right].$$

Esto completa la eliminación de Gauss, dando la *forma reducida* (3.13) de la matriz aumentada original (3.8). *Observe que (3.13) es precisamente la matriz aumentada para (3.7)* Si no se hubiese realizado la eliminación en las ecuaciones, al llegar a este punto se hubiera interpretado (3.13) como la matriz aumentada de un sistema de ecuaciones –es decir (3.7)– se hubiera escrito el sistema, y después se hubiera usado la sustitución en reversa para obtener la solución al sistema de ecuaciones original.

El hecho de resolver el sistema (3.1) de dos maneras diferentes pudo haber oscurecido la facilidad de uso de la eliminación de Gauss en la matriz aumentada. Considere otro ejemplo ilustrativo.

(3.14) ***Ejemplo.*** Considere las tres ecuaciones lineales con tres incógnitas

$$\begin{aligned} -3x_1 - 3x_2 - 3x_3 &= -3 \\ -2x_1 + 2x_2 + x_3 &= 0 \\ x_1 - 3x_2 + 3x_3 &= 0. \end{aligned}$$

En seguida se escribe la matriz aumentada

(3.15)
$$\left[\begin{array}{ccc|c} -3 & -3 & -3 & -3 \\ -2 & 2 & 1 & 0 \\ 1 & -3 & 3 & 0 \end{array}\right].$$

Reemplace el primer renglón con él mismo dividido entre -3 y utilice este nuevo primer renglón para eliminar los elementos en la primera columna de otros dos renglones, lo que da

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 4 & 3 & 2 \\ 0 & -4 & 2 & -1 \end{array}\right].$$

Reemplazando el segundo renglón de esta nueva matriz aumentada con él mismo dividido entre 4 y utilizando el nuevo renglón para eliminar el elemento en la segunda columna del último renglón, da como resultado una matriz que –después de dividir su último renglón entre 5– a su vez da como resultado de la eliminación de Gauss la matriz reducida

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0.75 & 0.5 \\ 0 & 0 & 1 & 0.2 \end{array}\right].$$

Si se interpreta esta última matriz reducida como la matriz aumentada para un sistema de ecuaciones, entonces la última ecuación da claramente $x_3 = 0.2$; realizando la sustitución en reversa y poniendo este resultado en la segunda ecuación, da que $x_2 = 0.35$; nuevamente, realizando la sustitución en reversa con estos dos resultados en la primera ecuación da $x_1 = 0.45$. Si sustituye estos resultados en el sistema original de ecuaciones, verá que la eliminación de Gauss aplicada a la matriz aumentada ha producido verdaderamente una solución.

Ahora podrá resolver los problemas del 1 al 5.

Intercambios en la eliminación de Gauss

En cada uno de los casos anteriores se ha procedido de un modo sistemático, utilizando los renglones (o ecuaciones) de arriba hacia abajo para eliminar valores en las columnas (las variables) de izquierda a derecha. Sin embargo, esto no es necesario, en algunos casos es imposible, y en los cálculos prácticos –como se verá posteriormente– con frecuencia es imprudente. Si desde el principio se usa un renglón diferente del primero para realizar la eliminación, notacionalmente es más fácil *intercambiar* este renglón con el primero y después proceder de la manera acostumbrada con este nuevo primer renglón; observe que intercambiar dos renglones es lo mismo que escribir las ecuaciones en diferente orden –lo cual ciertamente no tiene ningún efecto sobre las soluciones (vea teorema 4.12).

(3.16) ***Ejemplo.*** Considere nuevamente el sistema de tres ecuaciones con tres incógnitas del ejemplo 3.14 y su matriz aumentada (3.15). Suponga que se decide utilizar el *tercer* renglón de (3.15) en lugar del primero para eliminar en la primera columna –quizá porque ya tiene un 1 como primer elemento. Por lo tanto se intercambian el primero y tercero renglones, obteniendo

$$\left[\begin{array}{rrr|r} 1 & -3 & 3 & 0 \\ -2 & 2 & 1 & 0 \\ -3 & -3 & -3 & -3 \end{array}\right].$$

El reemplazar este nuevo primer renglón por él mismo dividido entre su primer elemento 1, por supuesto no cambia nada; se usa este primer renglón para eliminar los elementos en la primera columna de los otros dos renglones y esto da

$$\left[\begin{array}{rrr|r} 1 & -3 & 3 & 0 \\ 0 & -4 & 7 & 0 \\ 0 & -12 & 6 & -3 \end{array}\right].$$

Simplemente para demostrar que se puede hacer, suponga que decide utilizar el tercer renglón para la eliminación en la segunda columna; por lo tanto intercambie el segundo y tercero renglones actuales para obtener

$$\left[\begin{array}{rrr|r} 1 & -3 & 3 & 0 \\ 0 & -12 & 6 & -3 \\ 0 & -4 & 7 & 0 \end{array}\right].$$

Reemplazando este segundo renglón por sí mismo dividido entre -12 y utilizando el nuevo segundo renglón resultante para eliminar el elemento en la segunda columna del último renglón, se obtiene una matriz que –después de dividir su último renglón entre 5– da, a su vez, la forma reducida final

$$\left[\begin{array}{ccc|c} 1 & -3 & 3 & 0 \\ 0 & 1 & -0.5 & 0.25 \\ 0 & 0 & 1 & 0.2 \end{array}\right].$$

(*Observe que esto es diferente de la matriz reducida que se obtuvo en el ejemplo 3.14; sin embargo, se obtendrá la misma solución para las ecuaciones*). Si se interpreta esta matriz reducida como la matriz aumentada de un sistema de ecuaciones, la última ecuación da por supuesto $x_3 = 0.2$; la sustitución en reversa produce $x_2 = 0.35$ y $x_1 = 0.45$ – *exactamente la misma solución que se obtuvo en el ejemplo 3.14 sin intercambios.*

El ejemplo 3.16 ilustra el hecho de que es *posible* intercambiar ecuaciones; en algunos casos es *necesario*:

(3.17) ***Ejemplo.*** Considere el sistema de cuatro ecuaciones lineales con cuatro incógnitas

$$\begin{aligned} -2x_1 + 4x_2 - 2x_3 - 6x_4 &= 4 \\ 3x_1 - 6x_2 + 6x_3 + 10x_4 &= -1 \\ -2x_1 + 6x_2 - x_3 + x_4 &= 1 \\ 2x_1 - 5x_2 + 4x_3 + 8x_4 &= -3. \end{aligned}$$

La matriz aumentada es

$$\left[\begin{array}{cccc|c} -2 & 4 & -2 & -6 & 4 \\ 3 & -6 & 6 & 10 & -1 \\ -2 & 6 & -1 & 1 & 1 \\ 2 & -5 & 4 & 8 & -3 \end{array}\right].$$

Reemplazando el primer renglón por sí mismo dividido entre -2 y haciendo uso del primer renglón resultante para eliminar los elementos en la primera columna de los últimos tres renglones da como resultado

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 3 & -2 \\ 0 & 0 & 3 & 1 & 5 \\ 0 & 2 & 1 & 7 & -3 \\ 0 & -1 & 2 & 2 & 1 \end{array}\right].$$

Ya que el elemento (2, 2) de esta matriz es cero, se puede usar el segundo renglón para eliminar en la segunda columna: no importa por cuánto se multiplique el segundo renglón, al sumar el resultado a un renglón inferior, el cero en el segundo elemento del segundo renglón deja sin cambio el segundo elemento del renglón inferior. Existe la necesidad de intercambiar el segundo renglón con el tercero o con el cuarto.

Hasta ahora los ejemplos han demostrado que siempre es *posible* y a veces *necesario* intercambiar renglones cuando se utiliza la eliminación de Gauss; se deja a las secciones 3.6 y 3.9 demostrar el hecho de que a veces es *prudente* hacerlo en los cálculos prácticos.

Ahora podrá resolver los problemas del 1 al 9.

Eliminación de Gauss para matrices aumentadas en general

Hasta aquí, sólo se han visto *ejemplos* de sistemas de ecuaciones, de matrices aumentadas y de eliminación de Gauss; ahora se examinará la situación y procedimiento generales. Vuelva a ver (1.36) en la sección 1.4, donde se demostró que el sistema general

(3.18)
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2q}x_q &= b_2 \\ \cdots\cdots\cdots\cdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q &= b_p \end{aligned}$$

de p ecuaciones lineales con q incógnitas era equivalente a $\mathbf{Ax} = \mathbf{b}$ para $p \times q$ $\mathbf{A}$ con $\langle \mathbf{A} \rangle_{ij} = a_{ij}$, $\mathbf{x}$ $q \times 1$ con $\langle \mathbf{x} \rangle_j = x_j$, y $\mathbf{b}$ $p \times 1$ con $\langle \mathbf{b} \rangle_i = b_i$. La eliminación de Gauss para matrices opera en la matriz aumentada $[\mathbf{A} \quad \mathbf{b}]$:

(3.19) ***Definición.*** La matriz separada $[\mathbf{A} \quad \mathbf{b}]$ es la *matriz aumentada* del sistema de ecuaciones descrito $\mathbf{Ax} = \mathbf{b}$.

La eliminación de Gauss procede entonces a eliminar en las columnas comenzando con la columna 1, siguiendo con la columna 2, y demás sucesivamente; se usa el renglón r_j para eliminar en la columna j, y además se puede intercambiar un renglón inferior con el renglón r_j-ésimo antes de la eliminación. Se puede describir el proceso utilizando columnas sucesivas –(3.20)– o renglones –ver (4.1).

(3.20) La eliminación de Gauss con intercambios en la matriz aumentada $p \times (q + 1)$ $[\mathbf{A} \quad \mathbf{b}]$ procede como se indica a continuación:

1. Sea $j = 1$ y $r_1 = 1$; utilice el renglón r_j para eliminar en la columna j de la matriz aumentada actual como se indica en los pasos del 2 al 6.
2. Seleccione un renglón de entre los renglones numerados $r_j, r_j + 1, \ldots, p$ para su uso en la eliminación de la columna j; llame a este renglón i, de modo que el elemento (i, j) –llamado *pivote*– en la matriz aumentada actual sea diferente de cero. Si no hay elementos diferentes de cero en esta parte inferior de la columna j, entonces no se requiere de eliminación: ponga $r_{j+1} = r_j$ para utilizar el mismo renglón, y salte directamente al paso 6.
3. Intercambie los renglones i-ésimo y r_j-ésimo.
4. Reemplace este nuevo renglón r_j-ésimo por sí mismo dividido entre el pivote (su elemento diferente de cero en su columna j-ésima).
5. Utilice este nuevo renglón r_j-ésimo para eliminar los elementos en la columna j-ésima en los renglones $r_j + 1, r_j + 2, \ldots, p$. Ponga $r_{j+1} = r_j + 1$ para utilizar el siguiente renglón.
6. a) Si $j \leq q$ y $r_{j+1} \leq p$, entonces aún es posible seguir eliminando: incremente j en 1 y vuelva al paso número 2.
   b) En caso contrario, se ha completado la eliminación de Gauss: vaya al paso número 7.
7. Interprete la matriz reducida final como la matriz aumentada para un sistema de ecuaciones y proceda a encontrar las soluciones, si existen, mediante la sustitución en reversa.

Este bosquejo se aplica para p ecuaciones con q incógnitas, donde p no es necesariamente igual a q. Todos los ejemplos presentados han tenido p igual a q; la sección 3.3 proporcionará más ejemplos generales.

Ahora podrá resolver los problemas del 1 al 12.

Eliminación de Gauss-Jordan

Antes de ver más ejemplos de la eliminación de Gauss, se considera una modificación de este proceso que –aunque ineficiente para cálculos prácticos– con frecuencia es útil para propósitos teóricos.

En la eliminación de Gauss, se utiliza un renglón en particular (ecuación) para eliminar, *pero sólo* en los *renglones (ecuaciones)* ***por abajo*** *de ese renglón (ecuación)*. Si se escoge eliminar también en los renglones (ecuaciones) ***por arriba*** de ese renglón (ecuación), el proceso se llama *eliminación de Gauss-Jordan*. Se mostrará más adelante que el método de eliminación de Gauss-Jordan involucra más aritmética que el método de eliminación de Gauss, por lo tanto, se introduce más como una herramienta teórica que como una herramienta computacional. Vea, sin embargo, el problema 14.

(3.21) ***Ejemplo.*** Considere el empleo de la eliminación de Gauss-Jordan para resolver (3.1); ya que el primer paso es idéntico al del método de la eliminación de Gauss, la eliminación de Gauss-Jordan en la matriz aumentada original (3.8) lleva primeramente a (3.10). Después de que el segundo renglón en (3.10) se divide entre 3, el segundo renglón resultante se utiliza para eliminar *por arriba* de este renglón así como por abajo; esto da

$$\left[\begin{array}{rrrr|r} 1 & 0 & 5 & 1 & 3 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & -2 & -4 & 2 \\ 0 & 0 & 3 & 8 & 3 \end{array}\right].$$

Ahora se reemplaza el tercer renglón por sí mismo dividido entre -2 y el resultado se utiliza para eliminar *por arriba* de este renglón al igual que por abajo, y esto da

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & -9 & 8 \\ 0 & 1 & 0 & -8 & 4 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right].$$

Dividir el último renglón entre 2 y usar el resultado para eliminar *por arriba* de él, produce

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 35 \\ 0 & 1 & 0 & 0 & 28 \\ 0 & 0 & 1 & 0 & -7 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right].$$

Interpretar esto como la matriz aumentada para un sistema de ecuaciones, produce un sistema resuelto de modo trivial con la misma solución encontrada anteriormente mediante la eliminación de Gauss.

La eliminación de Gauss-Jordan siempre lleva a un sistema resuelto trivialmente, como se encontró en el ejemplo 3.21; en esencia, la eliminación de Gauss-Jordan realiza la sustitución en reversa de la eliminación de Gauss conforme avanza, en lugar de esperar hasta el final. Repitiendo la eliminación de Gauss-Jordan resulta en el uso de más aritmética para obtener la solución de un sistema de ecuaciones que la eliminación de Gauss; vea sin embargo el problema 14.

PROBLEMAS 3.2

▷ **1.** Resuelva cada uno de los siguientes sistemas de ecuaciones mediante la eliminación de Gauss para ecuaciones.

(a)
$$\begin{aligned} 3x_1 - x_2 + x_3 &= -1\\ 9x_1 - 2x_2 + x_3 &= -9\\ 3x_1 + x_2 - 2x_3 &= -9 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + 2x_2 - 5x_3 + x_4 &= 3\\ 2x_1 - 3x_2 + 4x_3 - 2x_4 &= 2\\ 4x_1 + x_2 - 6x_3 + 3x_4 &= 11\\ 5x_1 + 9x_2 - 20x_3 + x_4 &= 10 \end{aligned}$$

2. Utilice la eliminación de Gauss para ecuaciones para resolver

$$\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 5\\ 4x_1 + 3x_2 + 7x_3 + 3x_4 &= 8\\ -8x_1 - x_2 - x_3 + 3x_4 &= 4\\ 6x_1 + x_2 + 2x_3 - x_4 &= 1. \end{aligned}$$

3. Cada una de las siguientes matrices es una matriz aumentada para un sistema de ecuaciones; escriba el sistema.

(a) $\left[\begin{array}{ccc|c} 5 & 0 & 2 & -1\\ -2 & 3 & 8 & 3\\ 0 & 6 & 5 & -8\\ 0 & 0 & 4 & 5 \end{array}\right]$ (b) $\left[\begin{array}{cc|c} 0 & 5 & 4\\ 2 & 4 & 3\\ 1 & 3 & 6 \end{array}\right]$

(c) $\left[\begin{array}{cccc|c} 2 & 5 & 3 & 8 & 9\\ 1 & 4 & 8 & 2 & 2 \end{array}\right]$

▷ **4.** Utilice la eliminación de Gauss para matrices aumentadas para resolver

$$\begin{aligned} -x_1 - x_2 + 2x_3 &= -5\\ -3x_1 - x_2 + 7x_3 &= -22\\ x_1 - 3x_2 - x_3 &= 10. \end{aligned}$$

5. Utilice la eliminación de Gauss para matrices aumentadas para resolver

$$\begin{aligned} -3x - 6y + 9z &= 0\\ x + 4y + z &= 6\\ 2x + 8y + 3z &= 13. \end{aligned}$$

6. Complete la solución del sistema de ecuaciones del ejemplo 3.17.

7. Intercambie la segunda y tercera ecuaciones del problema 4 antes de comenzar la eliminación de Gauss para matrices aumentadas; verifique que la solución obtenida resuelve tanto el conjunto original de ecuaciones como el conjunto con las dos ecuaciones intercambiadas.

▷ **8.** a) Utilice la eliminación de Gauss para matrices aumentadas en

$$\begin{aligned} 3x - y - 2z &= 0 \\ -6x + 2y + 6z &= 4 \\ 2x + y + 6z &= 13 \end{aligned}$$

para obtener tanto la forma reducida del sistema como la solución; tendrá que intercambiar la segunda y tercera ecuaciones *después* de eliminar x.

b) Intercambie la segunda y tercera ecuaciones *antes* de usar la eliminación de Gauss para matrices aumentadas para obtener la forma reducida del sistema y la solución; no se requiere de ningún intercambio para la eliminación.

c) Compare las formas reducidas de las ecuaciones y soluciones de los dos métodos a) y b).

9. Utilice la eliminación de Gauss para matrices aumentadas para resolver

$$\begin{aligned} 2x - y + 2z + 3w &= 3 \\ -4x + 6y - 3z - 6w &= 2 \\ 6x - y + 8z + 5w &= 9 \\ 4x - 2y + 6z + 12w &= 12. \end{aligned}$$

▷ **10.** Utilice la eliminación de Gauss para matrices aumentadas para resolver

$$\begin{aligned} -2x - 4y + 2z - 6w &= 0 \\ 3x + 6y - 2z + 13w &= 6 \\ 2x + 4y \phantom{{}-2z} + 14w &= 12 \\ 4x + 8y - 7z \phantom{{}+13w} &= -10. \end{aligned}$$

11. Utilice la eliminación de Gauss para matrices aumentadas para resolver

$$\begin{aligned} 2x + 6y - 4z &= 4 \\ -x - 3y + 5z &= 4 \\ -3x - 9y + 4z &= -11. \end{aligned}$$

▷ **12.** Encuentre todas las matrices columna $\mathbf{b}$, 3×1, para las que existe al menos una solución a $\mathbf{Ax} = \mathbf{b}$ y encuentre todas las soluciones $\mathbf{x}$ asociadas con esa $\mathbf{b}$, donde

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 2 & 6 \\ -1 & 5 & -1 & -3 \\ 3 & 4 & 1 & 3 \end{bmatrix}.$$

13. Resuelva los dos sistemas de ecuaciones del problema 1 mediante la eliminación de Gauss-Jordan para matrices aumentadas.
14. Considere el siguiente proceso para resolver las ecuaciones $\mathbf{Ax} = \mathbf{b}$. Primero realice la eliminación de Gauss para matrices aumentadas. Después, comenzando en la *derecha de abajo* y trabajando hacia la izquierda, elimine *por arriba* de la diagonal principal como se haría en la eliminación de Gauss-Jordan.
    a) Demuestre que esto produce la misma matriz aumentada resultante que si se hubiera hecho la eliminación de Gauss-Jordan desde el principio.
    b) Muestre que la aritmética que se requiere es la misma que cuando se realiza la eliminación de Gauss seguida por la sustitución en reversa.
    c) Explique por qué este método involucra menos trabajo que cuando se realiza la eliminación de Gauss-Jordan desde el principio.
    d) Utilice esto en (3.1).

3.3 *EXISTENCIA DE SOLUCIONES A SISTEMAS DE ECUACIONES: ALGUNOS EJEMPLOS Y PROCEDIMIENTOS*

Cada sistema de ecuaciones considerado en la sección 3.2 tenía tantas ecuaciones como incógnitas, y los métodos de eliminación vistos siempre arrojaban una solución. Sin embargo, los métodos trabajan de un modo más general.

El número de soluciones

Primero considere el caso, aparentemente simple, de una sola ecuación $ax = b$ con una incógnita x; inmediatamente se tiende a decir que la solución de esta ecuación es $x = b/a$, pero de hecho existen tres posibilidades:

1. Si $a \neq 0$, entonces $x = b/a$ tiene sentido, y ésta es la *única* solución de esta ecuación.
2. Si $a = 0$, entonces existen dos posibilidades:
    a) Si $b \neq 0$, entonces, en la ecuación, se pide encontrar x tal que $0x = b \neq 0$, y *no* existe una solución x. Se dice que "no existe solución" o que "la ecuación es inconsistente" ya que implica que $0 = b \neq 0$, lo cual es una contradicción.
    b) Si $b = 0$, entonces hay *infinitamente muchas* soluciones: Cada número x es una solución ya que $0x = 0 = b$ sin importar el valor de x.

Es sorprendente que las tres posibilidades –*exactamente una (única)* solución, *ninguna* solución, o *infinitamente muchas* soluciones– se cumplen para dos ecuaciones con dos incógnitas. Por ejemplo las ecuaciones

$$x_1 + x_2 = 2$$
$$x_1 - x_2 = 0$$

tienen una solución única,

$$x_1 + x_2 = 2$$
$$x_1 + x_2 = 1$$

son inconsistentes, y

$$x_1 + x_2 = 2$$
$$2x_1 + 2x_2 = 4$$

tienen infinitamente muchas soluciones, es decir $x_1 = k$, $x_2 = 2 - k$, para toda k. Aún más sorprendente es el hecho de que, precisamente, las mismas tres posibilidades se cumplen para p ecuaciones con q incógnitas: *exactamente una (única)* solución, *ninguna* solución, o *infinitamente muchas* soluciones. Esto se demostrará más adelante; por el momento, se puede confiar en que mediante la eliminación de Gauss, se pueden descubrir todas (si las hay) las soluciones a cualquier sistema concreto de ecuaciones.

Eliminación de Gauss y la existencia de soluciones: ejemplos

(3.22) ***Ejemplo.*** Considere las tres ecuaciones con tres incógnitas

$$\begin{aligned} x_1 + 2x_2 - 5x_3 &= 2 \\ 2x_1 - 3x_2 + 4x_3 &= 4 \\ 4x_1 + x_2 - 6x_3 &= 8. \end{aligned}$$

La matriz aumentada es

$$\left[\begin{array}{rrr|r} 1 & 2 & -5 & 2 \\ 2 & -3 & 4 & 4 \\ 4 & 1 & -6 & 8 \end{array}\right].$$

la eliminación de Gauss reduce esto a

$$\left[\begin{array}{rrr|r} 1 & 2 & -5 & 2 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

que se interpreta como la matriz aumentada para el sistema de ecuaciones

$$\begin{aligned} x_1 + 2x_2 - 5x_3 &= 2 \\ x_2 - 2x_3 &= 0 \\ 0x_3 &= 0. \end{aligned}$$

Esta última ecuación –en la que, por lo general, se confía para determinar x_3– se resuelve mediante $x_3 = k$ para cualquier número k arbitrario. La sustitución en reversa da $x_2 = 2k$, $x_1 = 2 + k$ y se puede escribir la solución como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + k \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \text{para } k \text{ arbitraria.}$$

Por lo tanto, existen infinitamente muchas soluciones, las cuales se encontraron mediante la eliminación de Gauss.

(3.23) ***Ejemplo.*** Considere las tres ecuaciones lineales con cuatro incógnitas

$$\begin{aligned} x_1 + 2x_2 - x_3 + 2x_4 &= 4 \\ 2x_1 + 7x_2 + x_3 + x_4 &= 14 \\ 3x_1 + 8x_2 - x_3 + 4x_4 &= 17 \end{aligned}$$

cuya matriz aumentada es

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 2 & 7 & 1 & 1 & 14 \\ 3 & 8 & -1 & 4 & 17 \end{array}\right].$$

aplicando la eliminación de Gauss, esta matriz se reduce a

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Esto se puede interpretar como la matriz aumentada para

$$\begin{aligned} x_1 + 2x_2 - x_3 + 2x_4 &= 4 \\ x_2 + x_3 - x_4 &= 2 \\ 0x_4 &= 1. \end{aligned}$$

Esta última ecuación –en la que por lo general se confía para determinar x_4– pide encontrar una x_4 tal que $0x_4 = 1$. No puede existir tal x_4, así que el sistema de ecuaciones no tiene solución.

Estos ejemplos y problemas muestran que, en general, no es posible determinar si las ecuaciones tienen una, ninguna, o infinitamente muchas soluciones sólo a partir de los *números* de las ecuaciones o incógnitas. Puede suceder que dos ecuaciones con 10 incógnitas sean inconsistentes, mientras que 10 ecuaciones con dos incógnitas tengan una solución única. La eliminación de Gauss es el método para determinar cuál de las situaciones prevalece.

Ahora podrá resolver los problemas del 1 al 7.

Primera variable y primera columna

En el ejemplo 3.22 fue posible asignar un valor arbitrario a x_3, pero las variables restantes se determinaban por completo por ese valor. Esto genera la pregunta de (una vez que se ha aplicado la eliminación de Gauss) cómo se distingue a qué variables se les pueden asignar valores arbitrarios, y cuáles, entonces, están completamente definidas. Uno de los métodos se describe fácilmente en términos de *primeras variables* (*o primeras columnas*).

(3.24) ***Definición.*** La *primera variable* en una ecuación es la primera (leyendo de izquierda a derecha) variable en esa ecuación con un coeficiente diferente a cero. La *primera columna* para un renglón de una matriz, es la columna que contiene el primer elemento (leyendo de izquierda a derecha) diferente de cero en ese renglón.

En el capítulo 4 se desarrolla la teoría necesaria para apoyar la siguiente regla:

(3.25) ***Vista previa a la solución de ecuaciones.*** Después de completar la eliminación de Gauss en la matriz aumentada $[\mathbf{A} \quad \mathbf{b}]$ del sistema de ecuaciones $\mathbf{Ax} = \mathbf{b}$, encuentre las primeras variables y (primeras columnas) de las ecuaciones reducidas (matriz aumentada reducida). Entonces.

1. No existen soluciones si y sólo si la última columna es una primera columna para cierto renglón.
2. Si la última columna no es una primera columna para algún renglón:
   a) Existe una solución única si y sólo si cada variable es una primera variable para alguna ecuación.
   b) Existen infinitamente muchas soluciones si y sólo si hay algunas variables que *no* sean primeras variables; a cada una de estas variables *no primeras* se le puede asignar un valor completamente arbitrario, entonces cada primera variable está completamente determinada en términos de los valores asignados a las variables no primeras.

(3.26) ***Ejemplo.*** Considere el sistema de ecuaciones del ejemplo 3.22, donde se aplica la eliminación de Gauss a su matriz aumentada. Las primeras columnas en la matriz reducida resultante son los números 1 y 2; la última columna, número 4, *no* es una primera columna como tampoco lo es la columna 3. Ya que x_3 no es una primera variable, de acuerdo a (3.25) 2) b) hay infinitamente muchas soluciones: se puede asignar un valor arbitrario a

x_3, en términos del cual los otros están perfectamente determinados, Esto concuerda con lo que se encontró anteriormente en el ejemplo 3.22.

PROBLEMAS 3.3

▷ **1.** Utilice la eliminación de Gauss para resolver

$$\begin{aligned} 2x - 3y &= -1 \\ 2x + y &= 3 \\ x - 3y &= -2. \end{aligned}$$

2. Utilice la eliminación de Gauss para resolver

$$\begin{aligned} 2x - 3y &= -1 \\ 2x + y &= 3 \\ x - y &= 2. \end{aligned}$$

3. Utilice la eliminación de Gauss para resolver

$$\begin{aligned} 2x - 3y &= -1 \\ -4x + 6y &= -2 \\ 12x - 18y &= -6. \end{aligned}$$

▷ **4.** Utilice la eliminación de Gauss para resolver

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 0 \\ 2x_1 + 5x_2 \qquad\qquad &= 8 \\ x_1 \qquad - 14x_3 + 8x_4 &= -15 \\ -2x_1 - 3x_2 + 14x_3 + 2x_4 &= 10. \end{aligned}$$

5. Utilice la eliminación de Gauss para resolver

$$\begin{aligned} -x_1 + x_2 &= 1 \\ 3x_1 - 2x_2 &= -1 \\ 2x_1 - x_2 &= 1. \end{aligned}$$

6. Dé un ejemplo de a) un sistema con menos ecuaciones que incógnitas pero sin solución; b) un sistema con menos ecuaciones que incógnitas pero con infinitamente muchas soluciones (vea también el problema 7).

7. Explique por qué se puede encontrar un sistema de dos ecuaciones con tres incógnitas con exactamente una solución.

8. Cada una de las matrices que se muestran abajo, es la forma reducida que resulta de la eliminación de Gauss para matrices aumentadas de cierto sistema de ecuaciones. Para cada matriz: 1) identifique sus primeras columnas; 2) identifique sus primeras variables; 3) utilice la regla de (3.25) para concluir a partir de 1) y 2) si el sistema general tiene exactamente una, infinitamente muchas o ninguna solución.

(a) $\left[\begin{array}{cc|c} 1 & -2 & 4 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{array}\right]$ (b) $\left[\begin{array}{cc|c} 1 & -2 & 4 \\ 0 & 1 & 6 \\ 0 & 0 & 2 \end{array}\right]$ (c) $\left[\begin{array}{ccc|c} 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

9. Para cada una de las formas reducidas de las matrices aumentadas del problema 8, escriba el sistema de ecuaciones que representa dichas matrices y utilice la sustitución en reversa para resolver, si es posible, el sistema.

▷**10.** Encuentre todas las matrices **X**, 3×4, para las cuales $\mathbf{AX} = \mathbf{0}$, donde **A** es

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 5 & -6 \\ 2 & -3 & 6 \end{bmatrix}.$$

11. Suponga que **A** es la matriz del problema 10 y que **B** y **C** son matrices, 3×4 tales que $\mathbf{AB} = \mathbf{AC}$. ¿Qué puede concluir a partir de la relación entre **B** y **C**? (Vea el problema 10.)

3.4 *COMO ENCONTRAR UNA INVERSA MEDIANTE LA ELIMINACION DE GAUSS*

En el capítulo 1 se demostró que las inversas de matrices y los sistemas de ecuaciones estaban estrechamente relacionados de varias maneras; en el teorema 1.38 de la sección 1.4 se explicó el modo en que las inversas proporcionan información sobre la solución de ecuaciones, mientras que en el teorema 1.44 se explicó cómo encontrar inversas mediante la solución de varios sistemas de ecuaciones. Ya que se acaba de desarrollar la eliminación de Gauss como un método sistemático para resolver sistemas de ecuaciones, se puede aplicar al problema de encontrar inversas -gracias al teorema 1.44.

De acuerdo al teorema 1.44, se puede construir una inversa derecha **R** de **A** como

$$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \cdots \quad \mathbf{r}_p]$$

si cada $\mathbf{r}_i$ resuelve $\mathbf{Ar}_i = \mathbf{e}_i$. De modo similar se puede construir una inversa izquierda **L** con

$$\mathbf{L}^T = [\mathbf{l}_1 \quad \mathbf{l}_2 \quad \cdots \quad \mathbf{l}_q]$$

si cada $\mathbf{I}_j$ resuelve $\mathbf{A}^T\mathbf{I}_j = \mathbf{e}_j$. Y, por supuesto, si $\mathbf{A}$ es cuadrada ($p = q$) y no sin gular, entonces se puede encontrar la inversa bilateral $\mathbf{A}^{-1}$ ya sea como **R** o **L**.

Para cualquier inversa de éstas, el problema computacional es similar:

(3.27) Dada una matriz general $\mathbf{C}$ y n matrices columna $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, resuelva los n sistemas de ecuaciones $\mathbf{C}\mathbf{x}_i = \mathbf{b}_i$ para $1 \leq i \leq n$ para las n soluciones $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

Para encontrar una inversa derecha $\mathbf{R}$ de $\mathbf{A}$; considere $\mathbf{C} = \mathbf{A}$; para encontrar una inversa izquierda $\mathbf{L}$ de $\mathbf{A}$, considere $\mathbf{C} = \mathbf{A}^T$, en cualquier caso, las $\mathbf{b}_i$ son matrices columna unitarias $\mathbf{e}_i$ del orden apropiado. Sin embargo, la situación en (3.27), surge también en otras circunstancias. Por ejemplo, la ecuación $\mathbf{C}\mathbf{x} = \mathbf{b}$ puede modelar algún sistema físico, con $\mathbf{b}$ representando entradas conocidas y $\mathbf{x}$ salidas desconocidas {vea (2.24), donde $\mathbf{b}$ describe las fuerzas externas en una estructura y $\mathbf{x}$ describe su desplazamiento resultante }; entonces es posible que (3.27) represente un proyecto en el cual se desea examinar la respuesta del sistema a una variedad de entradas.

Se podría atacar (3.27) un sistema de ecuaciones a la vez: formar la matriz aumentada $[\mathbf{C} \quad \mathbf{b}_1]$ y usar la eliminación de Gauss para obtener $\mathbf{x}_1$; formar $[\mathbf{C} \quad \mathbf{b}_2]$ y usar la eliminación de Gauss para obtener x_2; . . . ; y por último formar $[\mathbf{C} \quad \mathbf{b}_n]$ y usar la eliminación de Gauss para obtener $\mathbf{x}_n$. Esto es poco eficaz por la siguiente razón: *las operaciones realizadas en la fase de la eliminación, anteriores a la sustitución en reversa* (dividir un renglón entre una constante, intercambiar dos renglones, sumar el múltiplo de un renglón al múltiplo del otro) *están completamente determinadas por los coeficientes en las ecuaciones y son independientes del miembro derecho*. Por supuesto los miembros derechos de la ecuación entran en los cálculos, pero no afectan a las operaciones realizadas; esto significa que, dados todos los miembros derechos al principio como en (3.27), se puede realizar la eliminación completa *con todos los miembros derechos al mismo tiempo*. Esto es:

(3.28) Para resolver (3.27), realice la eliminación de Gauss en la matriz *múltiplo-aumentada* $[\mathbf{C} \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n]$, interprete la forma reducida como la matriz múltiplo-aumentada para n sistemas reducidos, y resuelva cada sistema reducido mediante sustitución en reversa.

(3.29) ***Ejemplo***. Suponga que se necesita resolver $\mathbf{C}\mathbf{x}_i = \mathbf{b}_i$ para $i = 1, 2, 3$, donde

$$\mathbf{C} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Se forma la matriz múltiplo-aumentada $[\mathbf{C} \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$ y se realiza la eliminación como sigue:

$$\left[\begin{array}{cc|c|c|c} 2 & 1 & 3 & 5 & 3 \\ 1 & 1 & 2 & 3 & 1 \end{array}\right]$$

entonces,

$$\left[\begin{array}{cc|c|c|c} 1 & 0.5 & 1.5 & 2.5 & 1.5 \\ 1 & 1 & 2 & 3 & 1 \end{array}\right]$$

entonces,

$$\left[\begin{array}{cc|c|c|c} 1 & 0.5 & 1.5 & 2.5 & 1.5 \\ 0 & 0.5 & 0.5 & 0.5 & -0.5 \end{array}\right]$$

y finalmente,

$$\left[\begin{array}{cc|c|c|c} 1 & 0.5 & 1.5 & 2.5 & 1.5 \\ 0 & 1 & 1 & 1 & -1 \end{array}\right].$$

Esto se interpreta como la matriz múltiplo-aumentada para tres sistemas de ecuaciones, y para cada solución se sustituye en reversa. Por ejemplo, el primero es

$$\begin{aligned} x_1 + 0.5x_2 &= 1.5 \\ x_2 &= 1, \end{aligned}$$

que da fácilmente la primera solución como $\mathbf{x}_1 = [1 \;\; 1]^T$. Al manejar el segundo sistema de forma similar se obtiene $\mathbf{x}_2 = [2 \;\; 1]^T$. La tercera solución es $\mathbf{x}_3 = [2 \;\; -1]^T$.

El método (3.28) puede ser muy efectivo al resolver (3.27); por ejemplo, si $p = q = 50$ y $n = 10$, (3.28) involucra, únicamente, alrededor del 15% del trabajo que representa resolver cada uno de los 10 conjuntos de 50 ecuaciones con 50 incógnitas repitiendo el proceso de eliminación completo en un conjunto de ecuaciones a la vez.

PROBLEMAS 3.4

▷ **1.** Encuentre la inversa derecha **R** de la matriz **A** que se muestra abajo, aplicando la eliminación de Gauss a la matriz múltiplo-aumentada,

$$[\mathbf{A} \;\; \mathbf{e}_1 \;\; \mathbf{e}_2 \;\; \mathbf{e}_3] = [\mathbf{A} \;\; \mathbf{I}_3]$$

y verifique que **R** es una inversa bilateral.

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & -2 \\ 1 & 4 & -1 \end{bmatrix}.$$

2. Utilice la eliminación de Gauss en la matriz múltiplo aumentada para encontrar las inversas derechas de las matrices que se muestran abajo.

(a) $\begin{bmatrix} 2 & 4 & 2 \\ 4 & 2 & -14 \\ 2 & 6 & 11 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 6 & -3 \\ 2 & 5 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -3 \\ -8 & 12 \end{bmatrix}$.

3. Utilice la eliminación de Gauss en la matriz múltiple-aumentada para encontrar las inversas izquierdas de las matrices que se muestran abajo.

(a) $\begin{bmatrix} 2 & -4 & -6 \\ 4 & -6 & -16 \\ -6 & 8 & 29 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 4 \\ -3 & -4 \\ 4 & 5 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -6 \\ -1 & 3 \\ 3 & -9 \end{bmatrix}$.

▷ 4. Utilice la eliminación de Gauss en la matriz múltiplo-aumentada para encontrar las tres soluciones $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, de $\mathbf{A}\mathbf{x}_i = \mathbf{b}_i$ para $i = 1, 2, 3$, donde

$$\mathbf{A} = \begin{bmatrix} 2 & -8 & 4 \\ -4 & 19 & 10 \\ 6 & -22 & 15 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} -2 \\ 25 \\ -1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -10 \\ 44 \\ -23 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 8 \\ 23 \\ 32 \end{bmatrix}.$$

5. Un distribuidor de alimentos naturistas desea crear dos suplementos alimenticios en polvo, uno (para adultos) que contenga un 100% del **RMD** (requerimiento mínimo diario) de calcio, vitamina D y fósforo, y el segundo (para niños) que contenga 75, 100 y 75% respectivamente, del **RMD.** Se dispone de tres compuestos comerciales para crear estas mezclas especiales; 1 gramo de estos compuestos contiene la fracción del **RMD** mostrada abajo.

	Compuesto 1	Compuesto 2	Compuesto 3
Calcio	0.15 RMD	0.23 RMD	0.26 RMD
Vitamina D	0.26 RMD	0.27 RMD	0 RMD
Fósforo	0.15 RMD	0.28 RMD	0.08 RMD

Determine cuántos gramos de cada compuesto deben mezclarse para cada uno de los suplementos alimenticios, de modo que cada uno cumpla exactamente con los contenidos del **RMD** ya descritos, mediante el planteamiento de dos sistemas de tres ecuaciones con tres incógnitas y el uso de la matriz múltiplo-aumentada.

3.5 *OPERACIONES DE RENGLON Y MATRICES ELEMENTALES*

La eliminación de Gauss se introdujo en la sección 3.2 como una forma de manipular ecuaciones; más adelante se vio en términos de matrices, como un proceso para aplicar a matrices aumentadas. De hecho, el *proceso en sí puede* expresarse en términos matriciales como la multiplicación por una sucesión de matrices; esto resulta ser una herramienta esencial en el estudio de algunos problemas de matrices.

Operaciones de renglón

La eliminación de Gauss consiste en una secuencia de operaciones simples en los renglones de una matriz aumentada. Estas mismas operaciones simples pueden utilizarse para aprender mucho sobre matrices diferentes de las matrices aumentadas como se verá más adelante, y por sí mismas, son de gran interés.

(3.30) ***Definición.*** Una *operación elemental de renglón* (a veces llamada sólo *operación de renglón*) en una matriz $\mathbf{A}$ es cualquiera de los tres tipos de operaciones siguientes:
1. Intercambio de dos renglones de $\mathbf{A}$.
2. Reemplazo de un renglón $\mathbf{r}$ de $\mathbf{A}$ por $c\mathbf{r}$ para algún número $c \neq 0$.
3. Reemplazo de un renglón $\mathbf{r}_1$ de $\mathbf{A}$ por la suma $\mathbf{r}_1 + c\mathbf{r}_2$ de ese renglón y el múltiplo de *otro* renglón $\mathbf{r}_2$ de $\mathbf{A}$.

Se ha asegurado que uno de los usos de la notación matricial es el de expresar relaciones u operaciones complicadas en forma concisa; aquí se usará para operaciones elementales de renglón.

Matrices elementales

(3.31) ***Definición.*** Una *matriz elemental*, $p \times p$ es una matriz que se produce aplicando exactamente una operación elemental de renglón a $\mathbf{I}_p$. $\mathbf{E}_{ij}$ es la matriz elemental que se obtiene mediante el intercambio de los renglones i-ésimo y j-ésimo de $\mathbf{I}_p$. $\mathbf{E}_i(c)$ es la matriz elemental obtenida multiplicando el i-ésimo renglón de $\mathbf{I}_p$ por $c \neq 0$. $\mathbf{E}_{ij}(c)$ es la matriz elemental que se obtiene sumando c veces el j-ésimo renglón al i-ésimo renglón de $\mathbf{I}_p$, donde $i \neq j$.

Los siguientes son ejemplos de matrices elementales, 2×2:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \text{ con } c \neq 0, \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}, \quad \text{y} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

De éstos, los tres primeros son ilustraciones directas 2×2, de las tres matrices elementales básicas $\mathbf{E}_{12}$, $\mathbf{E}_1(c)$, y $\mathbf{E}_{12}(c)$. Desde luego, la cuarta es sólo $\mathbf{I}_2$, en la que se puede pensar como $\mathbf{E}_{12}(0)$; debido a esta posibilidad, $\mathbf{I}_p$ por sí misma siempre es una matriz elemental.

Considere ahora lo que sucede cuando se premultiplica cualquier matriz **A**, 2 + q por una de estas matrices elementales de muestra. Por ejemplo, considere $\mathbf{E}_{12}(c)\mathbf{A}$:

$$\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x & y & z \\ u & v & w \end{bmatrix} = \begin{bmatrix} x + cu & y + cv & z + cw \\ u & v & w \end{bmatrix}.$$

Esto es lo mismo que aplicar a **A** precisamente la operación de renglón que se aplicó a $\mathbf{I}_2$ para obtener la matriz elemental $\mathbf{E}_{12}(c)$ que premultiplicó a **A**. Estas dos operaciones *siempre* producen el mismo resultado.

(3.32) ***Teorema*** (operaciones de renglón y matrices elementales). Suponga que **E** es una matriz elemental, $p \times p$ que se produce aplicando una operación elemental de renglón a $\mathbf{I}_p$, y que **A** es una matriz $p \times q$. Entonces **EA** es la matriz que resulta de la aplicación de esa misma operación elemental de renglón en **A**.

DEMOSTRACION. Se probó el resultado para matrices elementales $\mathbf{E}_{ij}(c)$ y las demostraciones para los otros dos tipos de matrices elementales se dejaron para los problemas. Observe que $\mathbf{E}_{ij}(c) = \mathbf{I}_p + c\mathbf{e}_i\mathbf{e}_j^T$, donde $\mathbf{e}_i$ y $\mathbf{e}_j$ son las acostumbradas matrices unitarias columna de orden p. Por lo tanto,

$$\mathbf{E}_{ij}(c)\mathbf{A} = (\mathbf{I}_p + c\mathbf{e}_i\mathbf{e}_j^T)\mathbf{A} = \mathbf{A} + c\mathbf{e}_i\mathbf{e}_j^T\mathbf{A}.$$

Ahora $\mathbf{e}_j^T\mathbf{A}$, por la definición de multiplicación de matrices, es tan sólo el j-ésimo renglón de **A**; por la misma definición, $\mathbf{e}_i$ multiplicado por este renglón es tan sólo una matriz $p \times q$ cuyo i-ésimo renglón es igual a este renglón y cuyos otros elementos son todos cero. Por lo tanto, $c\mathbf{e}_i\mathbf{e}_j^T\mathbf{A}$ es tan sólo la matriz $p \times q$ cuyo i-ésimo renglón es c multiplicado por el j-ésimo renglón de **A** y cuyos otros elementos son iguales a cero; sumando esto a **A** se realiza la operación de renglón que se discute a la matriz **A** tal y como se afirmó. ■

Este teorema dice, por ejemplo, que la eliminación de Gauss consiste en una secuencia de premultiplicaciones por matrices elementales; se hará un uso decisivo de esto más adelante, así como del hecho de que las matrices elementales son no singulares:

(3.33) ***Teorema*** (matrices elementales y la no singularidad). Cada matriz elemental es no singular, y su inversa, por sí misma es una matriz elemental. De modo más preciso:

a) $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ji}\ (=\mathbf{E}_{ij})$
b) $\mathbf{E}_i(c)^{-1} = \mathbf{E}_i(1/c)$ con $c \neq 0$
c) $\mathbf{E}_{ij}(c)^{-1} = \mathbf{E}_{ij}(-c)$ con $i \neq j$

DEMOSTRACION. Recuerde que con el fin de demostrar que una matriz cuadrada $\mathbf{G}$ es no singular y que tiene una matriz particular $\mathbf{H}$ como su inversa, es necesario demostrar que $\mathbf{GH} = \mathbf{HG} = \mathbf{I}$. En este caso, las matrices $\mathbf{G}$ dadas son matrices elementales, al igual que sus matrices inversas $\mathbf{H}$; por lo tanto se puede aplicar el teorema 3.32.

a) Para demostrar que $\mathbf{E}_{ji}$ es la inversa de $\mathbf{E}_{ij}$, primero considere $\mathbf{E}_{ij}\mathbf{E}_{ji}$. Ya que $\mathbf{E}_{ij}$ es una matriz elemental, el teorema 3.32 dice que este producto es el mismo que intercambiar los renglones i-ésimo y j-ésimo de $\mathbf{E}_{ji}$ –que lógicamente da $\mathbf{I}$. Un argumento similar muestra que $\mathbf{E}_{ji}\mathbf{E}_{ij}$ también es igual a $\mathbf{I}$, lo que demuestra a).
b) La demostración, muy parecida a a) se deja para el problema 3.
c) La demostración puede modelarse como la de a), pero se selecciona otro método. Observe que $\mathbf{E}_{ij}(c) = \mathbf{I} + c\mathbf{e}_i\mathbf{e}_j^T$, donde $\mathbf{e}_i$ y $\mathbf{e}_j$ son las matrices unitarias columna acostumbradas. Entonces,

$$\begin{aligned}\mathbf{E}_{ij}(c)\mathbf{E}_{ij}(-c) &= (\mathbf{I} + c\mathbf{e}_i\mathbf{e}_j^T)(\mathbf{I} - c\mathbf{e}_i\mathbf{e}_j^T) = \mathbf{I} + c\mathbf{e}_i\mathbf{e}_j^T - c\mathbf{e}_i\mathbf{e}_j^T - c^2\mathbf{e}_i\mathbf{e}_j^T\mathbf{e}_i\mathbf{e}_j^T \\ &= \mathbf{I} - c^2\mathbf{e}_i(\mathbf{e}_j^T\mathbf{e}_i)\mathbf{e}_j^T \\ &= \mathbf{I} - c^2\mathbf{e}_i[0]\mathbf{e}_j^T \\ &= \mathbf{I} - \mathbf{0} = \mathbf{I},\end{aligned}$$

ya que $i \neq j$. El producto en el orden inverso también produce $\mathbf{I}$, mediante el mismo argumento. ■

Ahora se reunirán varios hechos a su disposición:

1. Cada operación elemental de renglón es equivalente a la premultiplicación por una matriz elemental. (Ver teorema 3.32.)
2. Cualquier secuencia de operaciones elementales de renglón es equivalente a premultiplicarla por el producto de una sucesión de matrices elementales. {Vea 1) arriba.}
3. Cada matriz elemental es no singular. {Ver teorema 3.33.}
4. Un producto de matrices no singulares es no singular. {Ver teorema 1.35 a) y problema 21, ambos en la sección 1.4.}

Estos hechos juntos demuestran el siguiente **teorema clave:**

(3.34) **Teorema clave** (operaciones de renglón). **Suponga que $\mathbf{B}$ resulta de aplicar una secuencia de operaciones elementales de renglón en $\mathbf{A}$. Entonces existe una $\mathbf{F}$ no singular para la cual $\mathbf{B} = \mathbf{FA}$ y por lo tanto $\mathbf{F}^{-1}\mathbf{B} = \mathbf{A}$.**

Estos resultados serán básicos para nuestro análisis del uso de las operaciones elementales de renglón para resolver una amplia variedad de problemas.

PROBLEMAS 3.5

1. Demuestre el teorema 3.32 para matrices elementales $\mathbf{E}_{ij}$.

▷ 2. Demuestre el teorema 3.32 para matrices elementales $\mathbf{E}_i(c)$, $c \neq 0$.

3. Demuestre el teorema 3.33 b).

4. Si **B** se deriva de **A** aplicando una secuencia de operaciones elementales de renglón a **A**, entonces se dice que **B** es renglón-equivalente a **A** y esto se escribe como $\mathbf{B} \sim \mathbf{A}$. Demuestre que $\sim$ es una *relacion de equivalencia* verdadera en el conjunto de todas las matrices $p \times q$ en el sentido de que :
   a) $\mathbf{A} \sim \mathbf{A}$ para toda **A**
   b) Si $\mathbf{B} \sim \mathbf{A}$, entonces $\mathbf{A} \sim \mathbf{B}$.
   c) Si $\mathbf{A} \sim \mathbf{B}$ y $\mathbf{B} \sim \mathbf{C}$, entonces también $\mathbf{A} \sim \mathbf{C}$.

5. Suponga que **A** es una matriz, $p \times p$ y que, cuando la eliminación de Gauss se aplica para reducir alguna matriz aumentada $[\mathbf{A}\ \ \mathbf{b}]$ a su forma reducida, el proceso puede llevarse a cabo exitosamente hasta su término para producir una matriz reducida $[\mathbf{U}\ \ \mathbf{b}']$ (que significa que **U** es unitaria triangular superior). Demuestre que tal matriz **U** es no singular, y utilice el **teorema clave 3.34** para demostrar que **A** es, por lo tanto, no singular.

▷ 6. Utilice la eliminación de Gauss para reducir la matriz aumentada **H** mostrada abajo, a la forma reducida **G**. Para cada operación elemental de renglón realizada, encuentre la matriz elemental correspondiente. Encuentre una matriz no singular **F** de modo que $\mathbf{FH} = \mathbf{G}$.

$$\mathbf{H} = \left[\begin{array}{cc|c} 0 & 1 & 2 \\ 2 & 3 & 4 \\ 4 & 6 & 8 \end{array}\right].$$

7. Sea $\mathbf{A}$, $p \times q$ y considere la *post*multiplicación por matrices elementales, $q \times q$. Demuestre que
   a) La postmultiplicación por $\mathbf{E}_{ij}$ intercambia las *columnas* i-ésima y j-ésima de **A**.
   b) La postmultiplicación por $\mathbf{E}_i\ c)$ multiplica la i-ésima *columna* de **A** por $c \neq 0$.
   c) La postmultiplicación por $\mathbf{E}_{ij}\ c)$ suma c veces la i-ésima *columna de* **A** a la j-ésima *columna* de **A**.

▷ 8. Suponga que **A** es $p \times p$ y no singular. Utilice el problema 7 para demostrar que:
   a) Si $\mathbf{A}'$ se obtiene de **A** intercambiando dos renglones de **A**, entonces $\mathbf{A}'$ es no singular y su inversa puede obtenerse intercambiando las columnas correspondientes de $\mathbf{A}^{-1}$.
   b) si $\mathbf{A}'$ se obtiene a partir de **A**, multiplicando el renglón i-ésimo de **A** por $c \neq 0$, entonces $\mathbf{A}'$ es no singular y su inversa puede obtenerse dividiendo la columna i-ésima de $\mathbf{A}^{-1}$ entre c; y

c) Si $\mathbf{A}'$ se obtiene a partir de $\mathbf{A}$ sumando c veces el renglón j-ésimo de $\mathbf{A}$ a su renglón i-ésimo, entonces $\mathbf{A}'$ es no singular y su inversa puede obtenerse restando c veces la columna i-ésima de $\mathbf{A}^{-1}$ de su columna j-ésima.

3.6 SELECCION DE PIVOTES Y ELIMINACION DE GAUSS EN LA PRACTICA

Ahora se considera un aspecto práctico importante de la solución de p ecuaciones lineales con p incógnitas mediante la eliminación de Gauss. En ejemplos anteriores –ver ejemplos 3.16 y 3.17 de la sección 3.2– se ha demostrado que, por lo general, es posible y a veces necesario, intercambiar ecuaciones al realizar la eliminación de Gauss; por supuesto, la necesidad surge cuando un elemento que se pretende usar como el siguiente pivote, es igual a cero.

En los cálculos prácticos, sin embargo, suele ser *inteligente* intercambiar ecuaciones, aun cuando no se enfrenten con un pivote cero. Este hecho puede resultar extraño, ya que la eliminación de Gauss siempre produce el (mismo) resultado correcto sin importar qué sucuencia de pivotes diferentes de cero se utilice; pero *esto supone que la aritmética se lleve a cabo con exactitud, y las computadoras en las que se resuelven los sistemas de ecuaciones en la práctica rara vez realizan una aritmética exacta.* Este hecho sobre la aritmética computacional inexacta afecta de manera drástica la eliminación de Gauss –y cualquier otro algoritmo computacional– y requiere que se reexamine la selección de pivotes. Esta sección se concentra en el aspecto práctico de la solución de sistemas de ecuaciones lineales cuando la aritmética realizada no es exacta; observe que, por lo común, la aritmética es inexacta aun para cómputos llevados a cabo manualmente con notación decimal.

Aritmética en punto flotante

Por lo general las computadoras y calculadoras representan el cero exactamente como 0; los números diferentes de cero, por lo general, se presentan como el producto de una *parte fraccionaria* de t dígitos f y un *factor de escala* 10^e, donde $0.1 \leq |f| \leq 1$ y e está restringida a cierto rango de números enteros. En muchas microcomputadoras con coprocesadores matemáticos, t es 16 y de modo general e está restringida a $-308 \leq e \leq 308$. Por simplicidad, se supondrá que e puede tomar *cualquier* valor entero. La condición de f significa que el primer dígito diferente de cero de f ocurre inmediatamente a la derecha del punto decimal; los números que se representan en esta forma se llaman *números en punto flotante de dígito-t*. Por lo tanto, la representación en punto flotante de 3-dígitos de 0.05 es 0.500×10^{-1}.

Para manejar números tales como $\frac{1}{3} = 0.333\ldots$ con una expansión decimal infinita o aun con números que requieren de un número de dígitos muy grande en su representación decimal, en la práctica deben reemplazarse con números muy cercanos pero con un menor número de dígitos. Aunque algunas computadoras y calculadoras sólo truncan los dígitos no deseados, la mayoría *redondean* el número.

(3.35) ***Definición.*** Para *redondear* un número diferente de cero x a t *dígitos en punto flotante*: 1) represente x en la notación de punto flotante como $x = f \times 10^e$ con $0.1 \leq |f| < 1$, de modo que el primer dígito de f (a la derecha del punto decimal) sea diferente de cero; 2) retenga como f_0 los primeros t dígitos de f; 3) si el $(t + 1)$-ésimo dígito de f es 0, 1, 2, 3 o 4, entonces defina la versión redondeada $-\langle x \rangle-$ de x como $f_0 \times 10^e$; 4) en caso contrario, incremente el t-ésimo dígito de f_0 en 1 para producir f'_0, y defina la versión redondeada $-\langle x \rangle-$ de x como $f'_0 \times 10^e$. Cero se redondea a cero: $-\langle 0 \rangle- = 0$.

Habiendo introducido esta terminología, se puede describir el modelo de *aritmética computacional en punto flotante de t-dígitos*. En este modelo, por *número computacional* se quiere decir ya sea 0 o un número (en punto flotante) representado *exactamente* por una parte fraccional de, a lo mucho, t dígitos con un factor a escala. En este modelo, el resultado de cualquier operación aritmética básica $(+, -, /, \times)$ entre dos *números computacionales* u y v 41*se define como el resultado de redondear a t* dígitos (en punto flotante), o el resultado que se obtendría con aritmética perfecta:

(3.36) ***Definición.*** Sean u y v números computacionales, y sea @ cualquiera de las operaciones aritméticas básicas $+, -, /$, o $\times$. Entonces el resultado para t dígitos en punto flotante u @ v se define como $-\langle u @ v \rangle-$.

(3.37) ***Ejemplo.*** Suponga que $t = 2$, de modo que sólo se tienen dos dígitos, y se evalúan $(\frac{2}{3} + \frac{2}{3}) - \frac{1}{3}$. Primero, $\frac{2}{3}$ se evaluará como

$$-\langle 0.666\ldots \rangle- = 0.67;$$

después $0.67 + 0.67$ se evaluará como

$$-\langle 0.67 + 0.67 \rangle- = -\langle 1.34 \rangle- = 1.3;$$

después $\frac{1}{3}$ se evaluará como

$$-\langle 0.333\ldots \rangle- = 0.33;$$

y por último a la expresión general $(\frac{2}{3} + \frac{2}{3}) - \frac{1}{3}$ se le dará el valor de

$$-\langle 1.3 - 0.33 \rangle- = -\langle 0.97 \rangle- = 0.97.$$

Observe que la expresión "equivalente" $\frac{2}{3} + (\frac{2}{3} - \frac{1}{3})$ se evaluaría de modo diferente, como 1.0 –que es también el valor en aritmética perfecta. La aritmética de 16 dígitos del MATLAB evalúa la expresión original exactamente como 1; como un ejemplo más realista de su aritmética, evalúa

$$\tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6}$$

como $1 + 10^{-16}$ en lugar de 1.

Ahora podrá resolver los problemas 1 y 2.

Método de eliminación de Gauss en la aritmética computacional

Se utilizará la aritmética computacional modelo para ilustrar la importancia de la selección de pivotes cuando la eliminación de Gauss se realiza en la práctica en computadoras. Por simplicidad, suponga que $t = 2$, esto es, que se tiene aritmética de dos dígitos; los mismos argumentos surgen con una $t = 16$ más realista, pero la suposición de $t = 2$ hace que aquéllos sean más transparentes.

Considere el sistema de ecuaciones

(3.38)
$$\begin{aligned} x_1 - x_2 &= 0 \\ 0.01x_1 + x_2 &= 1. \end{aligned}$$

Observe que cada coeficiente y lado derecho en (3.38) es un número computacional de dos dígitos: 0.01, por ejemplo, se representa como 0.10×10^{-1}. La solución exacta de (3.38) es $x_1^* = x_2^* = 1/1.01 = 0.990099\ldots$; lo mejor que se le podría pedir a la eliminación de Gauss en una computadora de dos dígitos sería obtener los números computacionales de dos dígitos más cercanos a x_1^* y x_2^*, es decir $x_1' = x_2' = 0.99$. Veamos qué tan bien lo hace.

(3.39) ***Ejemplo*** (sin intercambios). Suponga que se realiza la eliminación de Gauss sin intercambios, usando la primera ecuación para eliminar 0.01 de la segunda. Sumar -0.01 veces la primera ecuación a la segunda produce $(1 + 0.01)$ como el nuevo coeficiente de x_2, *pero en aritmética de los dígitos esto se debe evaluar* como $\prec 1.01 \succ = 1.0$. De modo que el sistema reducido se calcula como

$$\begin{aligned} x_1 - x_2 &= 0 \\ x_2 &= 1. \end{aligned}$$

La sustitución en reversa (en aritmética computacional) da la solución aritmética computacional $x_1 = x_2 = 1.0$ –una muy buena aproximación a la mejor posible de $x_1' = x_2' = 0.99$.

(3.40) ***Ejemplo*** (intercambiando). Ahora suponga que se realiza la eliminación de Gauss después de intercambiar las ecuaciones a

$$\begin{aligned} 0.01x_1 + x_2 &= 1 \\ x_1 - x_2 &= 0. \end{aligned}$$

Añadiendo –100 veces la primera ecuación a la segunda, da en la segunda nueva ecuación un coeficiente para x_2 igual a $-100 + (-1)$ y un lado derecho nuevo de -100, para evaluarse nuevamente en aritmética computacional de dos dígitos; se obtiene $\prec -101 \succ = -100$, de modo que el sistema reducido calculado es

$$\begin{aligned} 0.01x_1 + \quad x_2 &= \quad 1 \\ -100x_2 &= -100. \end{aligned}$$

La sustitución en reversa da primero $x_2 = 1$ (aún una muy buena aproximación a la mejor de $x_2' = 0.99$) y entonces $x_1 = 0$ (una aproximación incomparable con la mejor de $x_1' = 0.99$).

Estos dos ejemplos ilustran que, en la aritmética en punto flotante, pivotes diferentes pueden producir respuestas dramáticamente diferentes en la eliminación de Gauss, que algunas de estas respuestas pueden ser muy aceptables, pero que otras pueden ser completamente inaceptables. Nuestro problema es entender este fenómeno con el fin de evitar pivotes que llevan a respuestas pobres.

Análisis de error en reversa

Considere nuevamente la aplicación de la eliminación de Gauss a (3.38) sin hacer intercambios, como en el ejemplo 3.39. Si la segunda ecuación en (3.38) hubiese sido $0.01x_1 + 0.99x_2$, entonces con aritmética *perfecta* en este sistema alterado se habría producido la misma forma reducida que se obtuvo con aritmética computacional en el ejemplo 3.39. Por lo tanto:

> el efecto de utilizar aritmética computacional de dos dígitos para llevar a cabo la eliminación de Gauss sin intercambiar en (3.38) es exactamente el mismo que se produciría al usar aritmética exacta en el sistema *levemente* cambiado
>
> $$\begin{aligned} x_1 - \quad x_2 &= 0 \\ 0.01x_1 + 0.99x_2 &= 1. \end{aligned}$$

Por otro lado, considere nuevamente la aplicación de la eliminación de Gauss a (3.38) con intercambio, como en el ejemplo 3.40. Al igual que arriba, se desea encontrar un sistema que, con aritmética *perfecta*, se reducirá a lo que se obtuvo en el ejemplo 3.40 en aritmética de dos dígitos. Imagine que se reemplaza el término $x_1 - x_2$ por $x_1 - \alpha x_2$; ¿qué valor debería tener α para que, con aritmética perfecta en la eliminación como en el ejemplo 3.40 resultara igual a -100 como coeficiente de x_2 en las ecuaciones reducidas? En aritmética perfecta se tiene como segunda ecuación en el conjunto reducido tan sólo $-(100 + \alpha)x_2 = -100$; por lo tanto es necesario $-(100 + \alpha)$ para igualar a -100, de manera que $\alpha = 0$. Por lo tanto:

> el efecto de utilizar aritmética computacional de dos dígitos para llevar a cabo la eliminación de Gauss con intercambios en (3.38) es precisamente el

mismo que se produciría al usar aritmética exacta en un sistema muy cambiado

$$\begin{aligned} x_1 \qquad &= 0 \\ 0.01x_1 + x_2 &= 1. \end{aligned}$$

El análisis ha mostrado que para las dos maneras de aplicar la eliminación de Gauss en (3.38), la solución *calculada* puede visualizarse como la solución *exacta* a un problema *perturbado* (es decir, cambiado); este método se llama *análisis de error en reversa*: la "culpa" de los errores en el resultado de un proceso en una computadora se *regresa* a los datos en lugar de al proceso o a la aritmética computacional.

El análisis de error en reversa es fundamental en la matemática aplicada en lo general, y particular en el análisis de procedimientos computacionales. Los datos utilizados en los cómputos ya son por lo general inexactos debido a errores experimentales y de modelaje; si el análisis de error en reversa puede mostrar que la solución calculada es la solución exacta a un problema en el cual los datos han sido cambiados por cantidades aproximadamente iguales en tamaño a los errores que existen ya en los datos, entonces el resultado de la computadora no ha creado errores mayores, que aquéllos ya inherentes al problema. Esto permite la conclusión de que el resultado de la computadora es tan bueno como razonablemente pueda pedírsele.

Sin embargo, observe, que *no* se ha demostrado que la eliminación de Gauss siempre da una solución exacta a un problema ligeramente perturbado –al contrario: La solución del ejemplo 3.40 es la solución exacta a un problema muy perturbado. Nuestro problema consiste en descubrir una estrategia para realizar los intercambios –esto es, para seleccionar pivotes– para la cual es válido el hecho anterior. Con una estrategia apropiada, el intercambiar puede mejorar los hechos en lugar de producir un desastre como en el ejemplo 3.40.

Es posible que se tenga la impresión de que la razón por la cual la solución sin intercambios fue satisfactoria, mientras que la solución intercambiada no lo fue, se relaciona con el hecho de que el elemento (1, 1) –es decir 1– utilizado como el pivote para obtener la primera solución es mucho más grande que el elemento (2, 1) –es decir 0.01– utilizado como pivote para obtener la segunda solución. Por lo tanto, bien se puede sugerir la regla: "Evite pivotes pequeños". Esta no es una regla mala, pero no puede aplicarse indiscriminadamente. Para visualizar esto, suponga que se *reescala* (3.38) multiplicando la primera ecuación por 10^{-2}, multiplicando la segunda por 10^2, reemplazando x_1 por $x_1 = 10^2z_1$, y reemplazando x_2 por $x_2 = 10^{-2}z_2$. Esto da como resultado

(3.41)
$$\begin{aligned} z_1 - 0.0001z_2 &= 0 \\ 100z_1 + \qquad z_2 &= 100. \end{aligned}$$

Si cree que los pivotes pequeños son malos y los utiliza en el elemento mayor 100, en aritmética de dos dígitos se obtiene $z_2 = 100$ y entonces $z_1 = 0$, que a su vez da la misma solución errónea de $x_1 = 0$ y $x_2 = 1$ que se encontró en el ejemplo 3.40. La estrategia de evitar pivotes pequeños no es necesariamente exitosa.

Ahora podrá resolver los problemas del 1 al 4.

Estrategia de selección de pivotes

Considerando las ecuaciones equivalentes (3.38) y (3.41) que difieren sólo en escala, se demuestra que el hecho de evitar pivotes pequeños no es necesariamente una buena estrategia: dio buenos resultados en (3.38) pero resultados pobres en (3.41). Otra manera de visualizar esto sería que evitar pivotes pequeños podría ser una buena estrategia si se *escalaran las ecuaciones de modo* ''*apropiado*''. Con ''escalar'' nuevamente se quiere decir multiplicar cada ecuación por una constante diferente de cero y reemplazar cada variable por un múltiplo diferente de cero de una nueva variable. Muchos –no todos– buenos programas computacionales para la solución de sistemas de ecuaciones lineales (ver sección 3.9), escalan las ecuaciones de algún modo antes de usar la eliminación de Gauss; los pivotes pequeños se evitan entonces mediante una de dos estrategias:

1. *Pivoteo parcial*, en el cual las incógnitas se eliminan del modo común x_1, $x_2, \ldots, x_p$ y en el cual el pivote para la eliminación con x_i es el coeficiente de x_i en las ecuaciones enumeradas $i, i + 1, \ldots, p$ que tiene el mayor valor absoluto.
2. *Pivoteo completo*, en el que las incógnitas no se eliminan necesariamente en el orden acostumbrado. La primera variable eliminada es aquella variable x_i con el coeficiente mayor (en valor absoluto), y ese coeficiente se utiliza como pivote. La variable r-ésima eliminada es aquella variable con el coeficiente mayor (en valor absoluto) que el de todas aquellas $p - r + 1$ variables aún sin eliminar, en las ecuaciones $p - r + 1$ restantes y ese coeficiente se utiliza como el pivote.

La experiencia indica que en la práctica, por lo general, es suficiente utilizar un pivoteo parcial; las ventajas posibles del pivoteo completo tienden a ser abrumadas por las desventajas de una ''contabilidad'' incrementada en su implementación y el hecho de tener que examinar $(p - r + 1)^2$ coeficientes para determinar en cada ocasión el pivoteo en vez de sólo revisar $p - r + 1$ como en el pivoteo parcial. Los sistemas de ecuaciones que surgen en la práctica –a diferencia de aquéllos creados para los libros de texto con el fin de ilustrar las anomalías– con frecuencia parecen tener un escalamiento "natural" interconstruido que impide que el pivoteo parcial lleve a una selección desastrosa de pivotes; en topografía, por ejemplo, raramente se miden algunas distancias en millas y otras en pulgadas. Otro factor es que es innecesario encontrar el *mejor* pivote; sólo es necesario evitar pivotes *pésimos*.

También existen razones teóricas para usar pivoteo parcial o completo. El análisis de error en reversa puede ser utilizado para demostrar que la solución calculada $\mathbf{x}'$ resuelve exactamente un sistema perturbado en el que el tamaño de las perturbaciones ("tamaño" se mide en relación a los coeficientes originales) se relaciona con el tamaño de los pivotes utilizados durante la eliminación. En la práctica, los pivotes rara vez se vuelven grandes. De hecho:

(3.42) La eliminación de Gauss con pivoteo parcial realizado en aritmética en punto flotante de t dígitos produce una solución aproximada $\mathbf{x}'$, $p \times 1$, para $\mathbf{Ax} = \mathbf{b}$ que es la solución exacta a un problema perturbado $\mathbf{A}'\mathbf{x}' = \mathbf{b}$, donde en la práctica es por lo general el caso que el mayor elemento de la perturbación $\mathbf{A} - \mathbf{A}'$ no es mayor, alrededor de $p10^{1-t}$ veces, que el mayor elemento de $\mathbf{A}$.

Por lo tanto, si $\mathbf{A}$ se escala de modo que todos sus elementos sean de aproximadamente el mismo tamaño, entonces la eliminación de Gauss con pivoteo parcial produce la solución exacta a un sistema de ecuaciones, cada uno de cuyos coeficientes, por lo general, está sólo ligeramente perturbado de su valor correcto.

La necesidad de exhibir la $\mathbf{A}'$ en (3.42) hace difícil ilustrar (3.42) con un ejemplo realista. Sin embargo, existe una consecuencia derivada de (3.42), que por sí misma es interesante y más fácil de ilustrar. Si se calcula el residuo $\mathbf{r} = \mathbf{b} - \mathbf{Ax}'$ para ver qué tanto se acerca $\mathbf{x}'$ a la solución de $\mathbf{Ax} = \mathbf{b}$, entonces de (3.42) se puede escribir

$$\mathbf{r} = \mathbf{b} - \mathbf{Ax}' = \mathbf{b} - \mathbf{A}'\mathbf{x}' + \mathbf{A}'\mathbf{x}' - \mathbf{Ax}' = (\mathbf{A}' - \mathbf{A})\mathbf{x}'.$$

De esto se puede concluir a partir de (3.42) que

(3.43) La eliminación de Gauss con pivoteo parcial realizado en aritmética en punto flotante de t dígitos produce una solución aproximada $\mathbf{x}'$, $p \times 1$ para $\mathbf{Ax} = \mathbf{b}$ que es la solución exacta a un problema perturbado $\mathbf{Ax}' = \mathbf{b} - \mathbf{r}$ donde en la práctica se da por lo general el caso de que el mayor elemento en $\mathbf{r}$ no es mayor que $p10^{1-t}$ veces el mayor elemento en $\mathbf{A}$ por el mayor elemento en $\mathbf{x}'$.

$\mathfrak{M}$ (3.44) ***Ejemplo.*** Considere solucionar $\mathbf{Ax} = \mathbf{b}$, donde

$$\mathbf{A} = \begin{bmatrix} -4 & 7 & 1 & -3 & 71 & 6 & 5 & -2 & 9 & 8 \\ 2 & 1 & -6 & 3 & 5 & -60 & 9 & -8 & -4 & 7 \\ 3 & 53 & 2 & -7 & -6 & 4 & -9 & 8 & 5 & 1 \\ -9 & 3 & -5 & 8 & 2 & -1 & 7 & 4 & 6 & -61 \\ 7 & -5 & 4 & -6 & -1 & 2 & -8 & 51 & 9 & -3 \\ -5 & -8 & 2 & -50 & 9 & -7 & -3 & -1 & 6 & 4 \\ -48 & -1 & 3 & -7 & 9 & -2 & 4 & -6 & 8 & 5 \\ 1 & -9 & -63 & -2 & 8 & -3 & -6 & -7 & -4 & 5 \\ 8 & -6 & 1 & 4 & 7 & 3 & -46 & 2 & 9 & 5 \\ 2 & 1 & 9 & 5 & -3 & 8 & 6 & 4 & -73 & 7 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 10 \\ 4 \\ -6 \\ -3 \\ -2 \\ 9 \\ -8 \\ 1 \\ -5 \\ -7 \end{bmatrix}.$$

se resolvió el sistema utilizando MATLAB y tuvo lugar una $\mathbf{x}'$ cuyo elemento mayor es de alrededor 0.3; ya que $p = 10$ y el elemento mayor en $\mathbf{A}$ es de 73, se espera de (3.43) que el elemento mayor del residuo $\mathbf{r} = \mathbf{b} - \mathbf{Ax}'$ deba ser menor que alrededor de $(10)10^{1-16}(73)(0.3)$, que es aproximadamente 2.2×10^{-13}. Con el MATLAB se calculó el residuo $\mathbf{r}$ –usando aritmética computacional, y, por lo tanto, se obtuvo un resultado imperfecto– teniendo como elemento mayor a aproximadamente 2×10^{-15}, bastante menor que el límite superior de (3.43).

PROBLEMAS 3.6

▷ **1.** Evalúe en aritmética de punto flotante de dos dígitos:
   a) $(\frac{1}{3} + \frac{1}{3}) + \frac{1}{3}$.
   b) $(0.58 + 0.53) - 0.53$.
   c) $0.58 + (0.53 - 0.53)$—compare este resultado con el de b)

2. Evalúe en aritmética de punto flotante de dos dígitos:
   a) $6(\frac{1}{6})$
   b) $3(\frac{1}{3})$

3. Considere el sistema (3.41) –una versión escalada de (3.38).
   a) Encuentre su solución exacta.
   b) Encuentre la mejor aproximación de dos dígitos para esta solución.
   c) Encuentre la solución que se obtiene de la eliminación de Gauss en aritmética de punto flotante de dos dígitos, suponiendo que el mayor coeficiente de cualquier variable se usa como el primer pivote.

▷ **4.** Utilizando aritmética de punto flotante de dos dígitos, aplique la eliminación de Gauss sin intercambios para resolver el sistema

$$0.98x_1 + 0.43x_2 = 0.91$$
$$-0.61x_1 + 0.23x_2 = 0.48.$$

b) Compare su resultado con la solución real $x_1^* = 0.005946\ldots$, $x_2^* = 2.102727\ldots$ y con la aproximación de dos dígitos posible $x_1' = 0.0059$, $x_2' = 2.1$;

c) Utilice el análisis de error en reversa y encuentre un sistema perturbado del cual la eliminación de Gauss en aritmética perfecta produciría x_1' y x_2' como la solución real.

(Observe que los errores absolutos pequeños en las respuestas numéricas representan de hecho errores relativos grandes y podrían no ser aceptables.)

5. Identifique el elemento que se utilizaría como el primer pivote en: 1) pivoteo parcial y 2) pivoteo completo para cada una de las matrices aumentadas mostradas abajo.

a) $\left[\begin{array}{ccc|c} -8 & 1 & 3 & 10 \\ 7 & 2 & 2 & 2 \\ 2 & 4 & 6 & 3 \end{array}\right]$

b) $\left[\begin{array}{ccc|c} 2 & 3 & 4 & 1 \\ 8 & 1 & 0 & 20 \\ -9 & 2 & 6 & 3 \end{array}\right]$

c) $\left[\begin{array}{ccc|c} 2 & 3 & 2 & 1 \\ 8 & 6 & 9 & 2 \\ 4 & 1 & 3 & 30 \end{array}\right]$

▷ **6.** Demuestre que el residuo $\mathbf{b} - \mathbf{Ax}$ es pequeño aun para la inaceptable solución aproximada de (3.41) que se encuentra en el libro; demuestre que la regla general (3.43) se cumple para este caso.

7. Muestre que el residuo $\mathbf{b} - \mathbf{Ax}$ es pequeño aun para la solución aproximada a (3.38) encontrada en el ejemplo 3.39; verifique que la regla general (3.43) se cumple para este caso.

8. Demuestre que cada uno de los tres pivotes diferentes del mayor en (3.41) –incluyendo el más pequeño, 0.0001– da una aproximación satisfactoria a x_1^* y x_2^* en aritmética de punto flotante de dos dígitos.

▷ **9.** La regla general (3.43) indica que la solución calculada produce un pequeño residuo; intuitivamente esto significa que "la solución casi resuelve $\mathbf{Ax} =$

b'', lo que puede ser diferente del enunciado ''la solución calculada está cerca de la solución verdadera''. Para visualizar esto, considere el sistema

$$0.89x_1 + 0.53x_2 = 0.36$$
$$0.47x_1 + 0.28x_2 = 0.19,$$

cuya solución exacta es $x_1^* = 1, x_2^* = -1$. Demuestre que la ''solución aproximada'' $x_1 = 0.47, x_2 = -0.11$ tiene un pequeño residuo y por lo tanto ''casi resuelve'' la ecuación, aunque la solución aproximada está lejos de la solución real. Este es un ejemplo de un problema mal acondicionado –ver la sección 6.4.

𝔐 10. a) Utilice el MATLAB o software similar para obtener una solución aproximada $\mathbf{x}'$ para el sistema del problema 9.
b) Encuentre el error $\mathbf{x}' - \mathbf{x}^*$ entre la solución calculada y la solución real.
c) Compare el tamaño de este error con el tamaño del residuo $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}'$, pero con el fin de calcular $\mathbf{r}$ de modo exacto, observe que

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}' = \mathbf{A}\mathbf{x}^* - \mathbf{A}\mathbf{x}' = \mathbf{A}(\mathbf{x}^* - \mathbf{x}')$$

y calcule $\mathbf{r}$ como $\mathbf{A}(\mathbf{x}^* - \mathbf{x}')$. Nuevamente esto demuestra que el sistema está *mal acondicionado*.

11. Ni el pivoteo parcial ni el pivoteo completo están *garantizados* para producir buenos resultados. Para visualizar esto, muestre que ambos producen resultados poco satisfactorios cuando se aplica la eliminación de Gauss en aritmética de punto flotante de t dígitos a

$$2x_1 + x_2 + x_3 = 1$$
$$x_1 + \epsilon x_2 + \epsilon x_3 = 2\epsilon$$
$$x_1 + \epsilon x_2 - \epsilon x_3 = \epsilon$$

cuando ϵ es muy pequeño y se le conoce con exactitud. Demuestre que al reescalar mediante $z_1 = x_1/\epsilon$, $z_2 = x_2$, $z_3 = x_3$ y dividir la segunda y tercera ecuación entre ϵ para obtener

$$2\epsilon z_1 + z_2 + z_3 = 1$$
$$z_1 + z_2 + z_3 = 2$$
$$z_1 + z_2 - z_3 = 1$$

se logra que ambos métodos funcionen satisfactoriamente.

3.7 LA DESCOMPOSICION-LU

Hasta ahora se ha enfatizado en el *proceso* de la eliminación de Gauss; el énfasis en esta sección se hará sobre el *resultado* –que es lo que se consigue con la eliminación de Gauss. Los hechos básicos se resumen en (3.56) y se ilustran en los ejemplos 3.45 y 3.55.

Recuerde del ejemplo 3.17 que la eliminación de Gauss *sin intercambios* no siempre puede completarse satisfactoriamente porque pueden aparecer ceros sobre la diagonal principal. Ya que es más fácil de comprender, primero se trata el caso en que tales ceros no aparecen y se puede llevar a cabo la eliminación de Gauss satisfactoriamente sin intercambios.

LU y eliminación de Gauss sin intercambios

Es fácil enunciar el resultado principal:

> Si la eliminación de Gauss en la matriz **A**, $p \times p$ puede completarse satisfactoriamente como en (3.20) pero sin intercambios, entonces el proceso es equivalente a escribir **A** como un producto **LU** de una matriz triangular inferior **L** y una matriz triangular superior **U** (vea la definición 1.3)

Recuerde el proceso (3.20) de la eliminación de Gauss, pero en el caso especial cuando **A** *es* $p \times p$, *todos los pivotes son* diferentes de *cero, y no se usan intercambios*. Al eliminar en la i-ésima columna, primero se dividen los elementos del i-ésimo renglón entre el pivote, el elemento (i, i) de la matriz en su forma parcialmente reducida; denote este i-ésimo pivote como α_{ii}. Entonces se reemplaza el renglón j-ésimo (para $j > i$) por el renglón j-ésimo más un múltiplo m_{ji} del i-ésimo renglón a modo de producir un cero en el elemento (j, i); los números m_{ji} se conocen como *multiplicadores*. Las matrices triangulares involucradas en el producto antes mencionado se define a partir de los números producidos durante la eliminación.

(3.45) ***Ejemplo.*** Considere la matriz **A**, 4×4 utilizada para introducir la eliminación de Gauss en la sección 3.2; vea (3.8). Una matriz triangular inferior **L** y una matriz unitaria triangular superior **U**. **U** son simplemente la forma reducida (3.13) de **A** producida durante la eliminación. Los elementos de la diagonal principal de **L** son los pivotes: $\langle \mathbf{L} \rangle_{ii} = \alpha_{ii}$; los elementos subdiagonales de **L** son los negativos de los multiplicadores: $\langle \mathbf{L} \rangle_{ji} = -m_{ji}$. Se pueden leer los multiplicadores m_{ji} y los pivotes α_{ii} de los pasos (3.8), (3.10), (3.11), (3.12) en la eliminación: $\alpha_{11} = -2$, $m_{21} = 3$, $m_{31} = -5$, y $m_{41} = -1$ de (3.8); $\alpha_{22} = 3$, $m_{32} = 3$, y $m_{42} = -2$ de (3.10); $\alpha_{33} = -2$, y $m_{43} = -3$ de (3.11); y $\alpha_{44} = 2$ de (3.12). Al formar **L** y **U** como se describió y después de comprobar la multiplicación, queda claro que la descomposición de **A** como $\mathbf{A} = \mathbf{LU}$ es correcta:

(3.46) **A** es igual a **L** por **U** ya que

$$\begin{bmatrix} -2 & 2 & -4 & -6 \\ -3 & 6 & 3 & -15 \\ 5 & -8 & -1 & 17 \\ 1 & 1 & 11 & 7 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 5 & -3 & -2 & 0 \\ 1 & 2 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Esta es la *descomposición LU* de **A**. De ella se puede derivar una descomposición triangular relacionada. Sea $\mathbf{D}_0$ la matriz diagonal formada por los pivotes; en este caso

$$\mathbf{D}_0 = \mathbf{diag}(-2, 3, -2, 2).$$

Ya que $\mathbf{A} = \mathbf{LU} = \mathbf{LD}_0^{-1}\mathbf{D}_0\mathbf{U}$, también se puede escribir $\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$, donde $\mathbf{L}_0 = \mathbf{LD}_0^{-1}$ y $\mathbf{U}_0 = \mathbf{D}_0\mathbf{U}$. Se encuentra $\mathbf{L}_0$ a partir de **L** dividiendo cada columna de **L** entre el pivote en esa columna, mientras que $\mathbf{U}_0$ se encuentra a partir de **U** multiplicando cada renglón de **U** por el pivote de ese renglón. $\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$ es la *descomposición L_0U_0* de **A**:

(3.47) **A** es igual a $\mathbf{L}_0$ por **U** ya que

$$\begin{bmatrix} -2 & 2 & -4 & -6 \\ -3 & 6 & 3 & -15 \\ 5 & -8 & -1 & 17 \\ 1 & 1 & 11 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{2} & 1 & 0 & 0 \\ -\frac{5}{2} & -1 & 1 & 0 \\ -\frac{1}{2} & \frac{2}{3} & -\frac{3}{2} & 1 \end{bmatrix} \begin{bmatrix} -2 & 2 & -4 & -6 \\ 0 & 3 & 9 & -6 \\ 0 & 0 & -2 & -4 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Primero se describió la eliminación de Gauss como la división del renglón entre el pivote, con el fin de obtener un 1 en ese elemento; esto hace los cálculos más fáciles a mano. Por otro lado, en los programas computacionales, por lo general *no* se divide primero entre este pivote, sino que se comienza a eliminar. El efecto es que la forma reducida que se obtiene es la $\mathbf{U}_0$ de arriba en este caso, en lugar de **U**; los elementos en $\mathbf{L}_0$ por debajo de la diagonal principal son tan sólo los negativos de los multiplicadores utilizados durante esta forma de eliminación. Antes se mostró cómo encontrar $\mathbf{L}_0$ y $\mathbf{U}_0$ a partir de $\mathbf{D}_0$, la matriz diagonal formada de los elementos de la diagonal principal de **L**. A la inversa, al tener $\mathbf{L}_0$ y $\mathbf{U}_0$, se podría encontrar $\mathbf{L} = \mathbf{L}_0\mathbf{D}_0$ y $\mathbf{U} = \mathbf{D}_0^{-1}\mathbf{U}_0$, donde $\mathbf{D}_0$ es la matriz diagonal que se forma de los elementos de la diagonal principal de **U**.

En el ejemplo 3.45 se pudo descomponer **A** como $\mathbf{A} = \mathbf{LU}$ y como $\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$, donde **L** es la matriz triangular inferior con elemento no cero en la diagonal principal y $\mathbf{U}$ es la matriz unitaria triangular superior, donde $\mathbf{L}_0$ es la matriz unitaria triangular inferior y $\mathbf{U}_0$ es la matriz triangular superior con elementos en la diagonal principal no cero. El siguiente teorema dice que esto puede hacerse siempre y cuando la eliminación de Gauss pueda completarse sin intercambio, y lo caracteriza cuando esto ocurre; la demostración, un tanto teórica, puede quedar omitida por la mayoría de los lectores.

(3.48) **Teorema clave** (*LU* **sin intercambios). Suponga que A es una matriz** $p \times p$. **Entonces**

a) la eliminación de Gauss sin intercambios puede llevarse a término con pivotes no cero a modo de producir una forma reducida unitaria triangular superior U si y sólo si las ***submatrices principales*** $\mathbf{A}_k$ **son no singulares para** $1 \leq k \leq p$, **donde** $\mathbf{A}_k$ **es la esquina superior izquierda,** $k \times k$ **de A:** $\langle \mathbf{A}_k \rangle_{ij} = \langle \mathbf{A} \rangle_{ij}$ **para** $1 \leq i, j \leq k$ **(en particular A es no singular)**

b) la eliminación de Gauss sin intercambios puede completarse como en a) si y sólo si $\mathbf{A} = \mathbf{LU}$, **donde U es la matriz reducida unitaria triangular superior de a) y L es una matriz triangular inferior con los pivotes** α_{ii} **de la eliminación de Gauss sobre la diagonal principal y los negativos de los multiplicadores** m_{ji} **de la eliminación de Gauss en las subdiagonales;** $\langle \mathbf{L} \rangle_{ii} = \alpha_{ii}$, $\langle \mathbf{L} \rangle_{ji} = -m_{ji}$. **De modo equivalente,**

(3.49)
$$\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$$

donde $\mathbf{L}_0 = \mathbf{L}\mathbf{D}_0^{-1}$ **unitaria triangular inferior y** $\mathbf{U}_0 = \mathbf{D}_0\mathbf{U}$ **triangular superior con los pivotes** α_{ii} **sobre la diagonal principal;**

$$\mathbf{D}_0 = \mathbf{diag}(\alpha_{11}, \ldots, \alpha_{pp}) \text{ y } \alpha_{ii} = \langle \mathbf{U}_0 \rangle_{ii} = \langle \mathbf{L} \rangle_{ii}$$

DEMOSTRACION. La demostración se hace por inducción en p. Si $p = 1$ y $\mathbf{A} = [a]$, entonces $\mathbf{L} = [a]$, $\mathbf{U} = [1]$, $\mathbf{L}_0 = [1]$, $\mathbf{U}_0 = [a]$, y $a \neq 0$ si y sólo si los pivote(s) –esto es, a– es no cero. Suponga que el teorema es verdadero para $p - 1$ y que $\mathbf{A}$ es $p \times p$. Efectúe la eliminación de Gauss para $\mathbf{A}$, mediante la eliminación únicamente de los primeros $p - 1$ renglones usando las primeras $p - 1$ columnas y luego regresando para eliminar los elementos en el renglón inferior en cada columna de izquierda a derecha; esto es equivalente a la eliminación de Gauss común. La primera parte de este proceso involucra eliminación de Gauss regular para $\mathbf{A}_{p-1}$, una matriz $(p - 1) \times (p - 1)$ para la cual se supone que se cumple el teorema. De tal modo, completar la eliminación de Gauss para $\mathbf{A}$ es equivalente a completarla en los primeros $p - 1$ renglones y después terminarla; completarla en $\mathbf{A}_{p-1}$ es equivalente a $\mathbf{A}_{p-1} = \mathbf{L}'\mathbf{U}'$ de la forma apropiada. Considere tratar de escribir $\mathbf{A} = \mathbf{LU}$ con $\mathbf{L}$ y $\mathbf{U}$ como sigue:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{p-1} & \mathbf{w} \\ \mathbf{v}^T & a_{pp} \end{bmatrix} = \begin{bmatrix} \mathbf{L}' & \mathbf{0} \\ \mathbf{l}^T & \alpha \end{bmatrix} \begin{bmatrix} \mathbf{U}' & \mathbf{u} \\ \mathbf{0}^T & 1 \end{bmatrix},$$

donde $a_{pp} = \langle \mathbf{A} \rangle_{pp}$. Esto requiere que $\mathbf{A}_{p-1} = \mathbf{L}'\mathbf{U}'$, que por su construcción resulta cierto; $\mathbf{L}'\mathbf{u} = \mathbf{w}$, lo cual se cumple si y sólo si $\mathbf{U} = \mathbf{L}'^{-1}\mathbf{w}$, ya que $\mathbf{L}'$ es no singular (los elementos en su diagonal principal son los pivotes no cero); $\mathbf{l}^T\mathbf{U}' = \mathbf{v}^T$, que se cumple si y sólo si $\mathbf{l}^T = \mathbf{v}^T\mathbf{U}'^{-1}$ ya que $\mathbf{U}'$ es no singular; y $a_{pp} = \mathbf{l}^T\mathbf{u} + \alpha$, que se cumple si y sólo si $\alpha = a_{pp} - \mathbf{l}^T\mathbf{u}$ con $\mathbf{l}^T$ y $\mathbf{u}$ como ya han sido determinados. Esto da la descomposición LU deseada.

En seguida se relaciona la descomposición con la eliminación de Gauss. La eliminación (como se modificó para tratar primero con los primeros $p - 1$ renglones) redujo los primeros $p - 1$ renglones de $\mathbf{A}$ de $[\mathbf{A}_{p-1} \quad \mathbf{w}]$ a $[\mathbf{U}' \ \mathbf{u}]$ sin alterar el último renglón $[\mathbf{v}^T \quad a_{pp}]$ de $\mathbf{A}$. Al eliminar ahora en este último renglón, simplemente se agrega al último renglón cierta combinación lineal de los primeros $p - 1$ nuevos renglones $[\mathbf{U}' \ \mathbf{u}]$ a modo de tener $[\mathbf{0}^T \quad \alpha_{pp}]$:

$$\mathbf{z}^T[\mathbf{U}' \quad \mathbf{u}] + [\mathbf{v}^T \quad a_{pp}] = [\mathbf{0}^T \quad \alpha_{pp}],$$

que da como resultado $\mathbf{z}^T = -\mathbf{v}^T\mathbf{U}'^{-1} = -\mathbf{l}^T$ (es decir, el nuevo renglón en la descomposición LU que ya se ha encontrado es el negativo del renglón de multiplicadores, como se indicaba) y

$$\alpha_{pp} = a_{pp} + \mathbf{z}^T\mathbf{u} = a_{pp} - \mathbf{l}^T\mathbf{u},$$

que es igual a α, el nuevo elemento en la diagonal principal encontrado para $\mathbf{L}$. Este pivote $\alpha_{pp} = \alpha$ será diferente de cero como se requiere si y sólo si $\mathbf{L}$ es no singular; ya que $\mathbf{U}$ es no singular y $\mathbf{L} = \mathbf{AU}^{-1}$, $\mathbf{L}$ es no singular si y sólo si $\mathbf{A}$ lo es –lo cual es la condición para $\mathbf{A}_p = \mathbf{A}$ que se necesitaba. La construcción de $\mathbf{L}_0$ y $\mathbf{U}_0$ a partir de $\mathbf{L}$ y de $\mathbf{U}$ demuestra la equivalencia de las dos descomposiciones triangulares. ■

Ahora podrá resolver los problemas del 1 al 10.

LU y eliminación de Gauss con intercambios

En el ejemplo 3.17 se revela que no toda matriz puede reducirse mediante eliminación de Gauss sin intercambios. Sin embargo, si se permiten los intercambios de renglones, toda matriz *no singular* puede reducirse mediante eliminación de Gauss a una matriz unitaria triangular superior; esto puede relacionarse como en el **teorema clave 3.48** con una descomposición triangular. Sin embargo, si **A** es singular, el proceso estándar (3.20) puede realizarse a cabo *pero no producirá una matriz* **unitaria** *triangular superior.*

(3.50) ***Ejemplo.*** Considere la matriz que es evidentemente singular

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

Con o sin intercambios, la eliminación de Gauss reduce esto a una matriz triangular superior **R** con un *cero* sobre la diagonal principal en lugar del uno deseado. Un argumento simple (problema 19) muestra que tal vez no pueda escribirse **A** como el producto de una matriz triangular inferior por una matriz unitaria triangular superior. De modo que **A** no puede descomponerse en $\mathbf{A} = \mathbf{LU}$ como en el teorema 3.48; pero **A** *sí puede descomponerse en* $\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$ *del segundo tipo, en el teorema con* $\mathbf{L}_0$ *unitaria triangular inferior, si se permiten elementos cero sobre la diagonal principal de* $\mathbf{U}_0$:

$$\mathbf{A} = \mathbf{L}_0\mathbf{U}_0 \qquad \text{con } \mathbf{L}_0 = \mathbf{I} \text{ y } \mathbf{U}_0 = \mathbf{A}$$

en este caso.

En el ejemplo 3.50 se deja entrever que la segunda descomposición $-\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$ con $\mathbf{U}_0$ posiblemente singular (como **A**), y $\mathbf{L}_0$ unitaria triangular inferior– permite que exista la descomposición de ciertas matrices singulares (la razón para la "$_0$" en $\mathbf{L}_0$ y $\mathbf{U}_0$) sin intercambios. Y, por supuesto, si **A** fuese no singular de modo que los elementos diagonales de $\mathbf{U}_0$ fueran diferentes de cero, se podrían construir **L** y **U** de $\mathbf{L}_0$ y $\mathbf{U}_0$ mediante el uso de la matriz diagonal $\mathbf{D}_0$ (que se forma de los elementos en la diagonal principal de $\mathbf{U}_0$ en este caso) usando (3.49). Pero, ¿qué pasa si ocurren intercambios? Nuevamente, es fácil enunciar los hechos básicos:

> Si la eliminación de Gauss en la matriz **A**, $p \times p$ puede completarse satisfactoriamente como en (3.20) con intercambios, entonces el proceso es equivalente a escribir una modificación de **A** llamada **A**$'$ (= **A** con sus renglones reordenados) como un producto **LU** de una matriz triangular inferior **L** y una triangular superior **U**.

Para verificar este enunciado para la eliminación de Gauss con intercambios, imagine que, de algún modo, se sabe *antes* del proceso de eliminación qué intercambios tendrán que hacerse, y que esos intercambios se realizan en los renglones de la *matriz original* **A**, *obteniendo una matriz* **A**$'$, *antes de comenzar la eliminación*. El hecho es que la eliminación de Gauss puede entonces realizarse en **A**$'$ *sin intercambios*, y que se usan *exactamente los mismos pivotes y multiplicadores para* **A**$'$ *que para* **A**, y que resulta *exactamente la misma forma reducida de* **A**$'$ *que de* **A**. Por lo tanto se puede utilizar la información del **teorema clave 3.48** para entender los efectos de la eliminación de Gauss con intercambios, visualizándola con eliminación de Gauss sin intercambios sobre una matriz obtenida de la original *permutando* –es decir, reordenando– sus renglones. Una definición ayudará a hacer más precisa esta reordenación.

(3.51) ***Definición.*** Una *matriz de permutación* **P**, $p \times p$ es cualquier matriz $p \times p$ que resulta de *permutar*–es decir, reordenar el orden de los renglones de $\mathbf{I}_p$. De modo más preciso: Cada renglón de **P** contiene exactamente un elemento diferente de cero, es decir 1; y cada columna de **P** contiene exactamente un elemento diferente de cero, es decir 1.

La sola definición de la multiplicación de matrices aclaró el hecho de que premultiplicar una matriz **A** por una matriz elemental daba el mismo resultado que aplicar la operación elemental de renglón correspondiente a **A**. De manera similar, resulta claro que premultiplicar **A** por una matriz de permutación **P** produce el mismo resultado que permutar los renglones de **A** exactamente del mismo modo en el cual los renglones de $\mathbf{I}_p$ para producir **P**.

(3.52) ***Teorema*** (matrices de permutación). Sea **P** una matriz de permutación. Entonces:

a) Para cualquier **A**, **PA** puede obtenerse a partir de **A** permutando los renglones de **A** exactamente como se permutaron los renglones de **I** para obtener **P**.
b) **P** es no singular, y $\mathbf{P}^{-1} = \mathbf{P}^T$: $\mathbf{P}\mathbf{P}^T = \mathbf{P}^T\mathbf{P} = \mathbf{I}$.

DEMOSTRACION

a) Esto se sigue fácilmente de las definiciones de la multiplicación de matrices y de matrices de permutación.
b) Separe **P** en sus renglones $\mathbf{r}_1, \ldots, \mathbf{r}_p$, que son tan sólo los renglones $\mathbf{e}_i^T$ de **I** en cierto orden. Entonces $\mathbf{P}^T$ tiene como sus columnas $\mathbf{r}_i^T$. La definición de la multiplicación de matrices implica que el elemento (i, j) de $\mathbf{P}\mathbf{P}^T$ es tan sólo (el elemento en la matriz 1×1) $\mathbf{r}_i\mathbf{r}_j^T$, y esto es 1 si $i = j$ y 0 si $i \neq j$; esto es, $\mathbf{P}\mathbf{P}^T = \mathbf{I}$. Un argumento similar en términos de las columnas de **P** demuestra que también $\mathbf{P}^T\mathbf{P} = \mathbf{I}$. ■

Considere una vez más la eliminación de Gauss con intercambios.

(3.53) **Teorema clave** (*LU* con intercambios). Suponga que **A** es $p \times p$. Entonces:

a) **A** se puede reducir mediante la eliminación de Gauss con intercambios de la manera usual, a una matriz *unitaria* triangular superior **U** si y sólo si **A** es no singular. Cuando se obtiene dicha **U**, existe una matriz de permutación **P** y una matriz triangular inferior **L** tal que

$$\mathbf{PA} = \mathbf{LU} \quad \text{y} \quad \mathbf{A} = \mathbf{P}^T\mathbf{LU},$$

con **L** y **U** como en el **teorema clave 3.48 b)**.

b) Existen una matriz de permutación **P**, una matriz *unitaria* triangular inferior $\mathbf{L}_0$, y una matriz triangular superior $\mathbf{U}_0$ tales que

$$\mathbf{PA} = \mathbf{L}_0\mathbf{U}_0 \quad \text{y} \quad \mathbf{A} = \mathbf{P}^T\mathbf{L}_0\mathbf{U}_0$$

A es no singular si y sólo si $\mathbf{U}_0$ es no singular; para **A** no singular, $\mathbf{L}_0$ y $\mathbf{U}_0$ se relacionan con las **L** y **U** de a) por

(3.54) $$\mathbf{U} = \mathbf{D}_0^{-1}\mathbf{U}_0 \quad \text{y} \quad \mathbf{L} = \mathbf{L}_0\mathbf{D}_0$$

donde $\mathbf{D}_0 = \mathbf{diag}(\alpha_{11}, \ldots, \alpha_{pp})$ y $\alpha_{ii} = \langle \mathbf{U}_0 \rangle_{ii} = \langle \mathbf{L} \rangle_{ii}$.

DEMOSTRACION. Si **A** puede escribirse como $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ como en a), entonces ciertamente **A** es no singular ya que $\mathbf{P}^T$, **L** y **U** son no singulares. La demostración del resto del teorema se hace por inducción en p. Para $p = 1$, el resultado es justamente el mismo que en el **teorema clave 3.48.**

a) Suponga que a) es verdadero para $p - 1$. Si **A** es no singular, entonces su primera columna es diferente de cero y los renglones pueden permutarse de modo que el nuevo elemento (1, 1) sea diferente de cero, entonces se puede llevar a cabo la eliminación de Gauss del modo usual en esa columna. La matriz $(p - 1) \times (p - 1)$ que permanece para su reducción también es no singular ya que **A** es no singular y se ha transformado a esta forma parcialmente reducida mediante matrices elementales (no singulares). Mediante la hipótesis inductiva, puede llevarse a cabo el resto de la eliminación de Gauss con intercambios, siendo el resultado una forma reducida unitaria triangular superior para **A.** Pero este proceso de eliminación es equivalente (se omite la demostración) a la eliminación sin intercambios en **PA,** y entonces el **teorema clave 3.48** se aplica a **PA,** obteniendo los resultados que se requieren.

b) Problema 18. ■

(3.55) ***Ejemplo.*** Como se indica, la eliminación de Gauss comienza con un intercambio en la matriz **A** que se muestra abajo, pero un poco más adelante se topa con un pivote cero:

$$\begin{bmatrix} 0 & 0 & 4 \\ 2 & 4 & 6 \\ -4 & -8 & -10 \end{bmatrix} \xrightarrow{\text{intercambio}} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 0 & 4 \\ -4 & -8 & -10 \end{bmatrix} \xrightarrow[m_{31}=2]{m_{21}=0} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{bmatrix}.$$

Después, en la eliminación de Gauss estándar (3.20), se usaría el elemento (2, 3), 4 como un pivote y se eliminaría el 2; sin embargo, esto no llevará a una descomposición LU. La modificación que se necesita para demostrar el **teorema clave 3.53 b)** requiere que, en caso de encontrar una columna inferior cero como en la segunda columna de este caso, se brinque simplemente esa columna (es decir, que tome los multiplicadores iguales a cero) y continúe a la siguiente columna y el siguiente renglón. En este caso, es el elemento (3, 3),de modo que se ha completado la eliminación. Se obtiene la descomposición L_0U_0 $\mathbf{A} = \mathbf{P}^T\mathbf{L}_0\mathbf{U}_0$ mediante la formación de **P** como la matriz de permutación que intercambia los dos renglones superiores, $\mathbf{U}_0$ como la forma triangular superior a la cual se ha reducido **A** y $\mathbf{L}_0$ como una matriz unitaria triangular inferior cuyos elementos son los negativos de los multiplicadores. En este caso, esto da $\mathbf{A} = \mathbf{P}^T\mathbf{L}_0\mathbf{U}_0$ como sigue:

$$\begin{array}{ccccccc} \mathbf{P}^T & \text{por} & \mathbf{L}_0 & \text{por} & \mathbf{U}_0 & & \text{igual a } \mathbf{A}\text{:} \\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} & & \begin{bmatrix} 2 & 4 & 6 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{bmatrix} & = & \begin{bmatrix} 0 & 0 & 4 \\ 2 & 4 & 6 \\ -4 & -8 & -10 \end{bmatrix}, \end{array}$$

como puede comprobarse mediante multiplicación.

Los detalles de las demostraciones oscurecen un tanto las ideas básicas, así que se resume

(3.56)
1. **A** es no singular si y sólo si **A** puede escribirse en una descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ siendo **P** una matriz de permutación, **U** una matriz unitaria triangular superior, y **L** una matriz triangular inferior con elementos en la diagonal principal diferentes de cero.
2. Cada matriz **A**, $p \times p$, puede escribirse como una descomposición L_0U_0, $\mathbf{A} = \mathbf{P}^T\mathbf{L}_0\mathbf{U}_0$ con **P** una matriz de permutación, $\mathbf{L}_0$ una matriz unitaria triangular inferior (y por lo tanto no singular), y $\mathbf{U}_0$ una matriz triangular superior. **A** es no singular si y sólo si $\mathbf{U}_0$ es no singular.
3. Para **A** no singular, las descomposiciones LU y L_0U_0 de 1) y 2) pueden relacionarse mediante (3.54).

4. Estas descomposiciones pueden encontrarse a partir de la eliminación de Gauss.

PROBLEMAS 3.7

▷ **1.** Para la matriz **A** que se muestra abajo, realice la eliminación de Gauss sin intercambios y dé los multiplicadores m_{21}, m_{31} y m_{32}, y los pivotes α_{11}, α_{22} y α_{33}, al igual que las matrices elementales que logran que se realice cada operación de renglón.

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -2 \\ 4 & -1 & 2 \\ 2 & -1 & 1 \end{bmatrix}$$

2. Verifique (3.46).

3. Verifique (3.47).

▷ **4.** Demuestre que una matriz triangular inferior **L** es no singular si y sólo si los elementos de su diagonal principal son todos diferentes de cero.

5. Demuestre que cada matriz unitaria triangular superior es no singular y que su inversa es unitaria triangular superior.

▷ **6.** Demuestre que el producto de muchas matrices arbitrarias unitarias triangulares superiores (o inferiores) es unitario triangular superior (o inferior).

7. Demuestre que la inversa de una matriz triangular inferior (o superior) no singular –ver problema 4– es triangular inferior (o superior).

8. Demuestre que el producto de muchas matrices arbitrarias triangulares superiores (o inferiores) es triangular superior (o inferior).

9. Para cada matriz **A** mostrada abajo, encuentre su descomposición LU, $\mathbf{A} = \mathbf{LU}$ y verifique que $\mathbf{LU} = \mathbf{A}$.

a) $\begin{bmatrix} 2 & 4 \\ 2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} -2 & -4 & 2 & -2 \\ 1 & 5 & 5 & -8 \\ -1 & 0 & 7 & -11 \\ 2 & 7 & 3 & -3 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & -6 \\ -3 & 4 & 7 \\ 2 & 4 & 3 \end{bmatrix}$

▷ **10.** Suponga que $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$ son dos descomposiciones LU de **A**, teniendo ambas matrices $\mathbf{L}_i$ elementos no cero en la diagonal principal. Mediante el examen de $\mathbf{L}_2^{-1}\mathbf{L}_1$ y $\mathbf{U}_2\mathbf{U}_1^{-1}$, muestre que $\mathbf{L}_1 = \mathbf{L}_2$ y $\mathbf{U}_1 = \mathbf{U}_2$. En otras palabras, la descomposición LU, $\mathbf{A} = \mathbf{LU}$ de una matriz **A** no singular es única.

11. Para cada matriz **A** mostrada abajo, encuentre su descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ realizando la eliminación de Gauss con intercambios.

a) $\begin{bmatrix} 0 & 2 \\ 2 & 4 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 6 & -4 \\ -4 & -12 & 11 \\ 3 & 14 & -16 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 4 & 0 & -2 \\ -4 & -8 & 0 & 3 \\ 3 & 7 & 2 & 4 \\ 5 & 6 & 1 & -8 \end{bmatrix}$

12. La matriz **A** se define mediante la descomposición LU mostrada abajo. Realice la eliminación de Gauss sin intercambios y verifique que los multiplicadores, pivotes, y forma reducida final son como los contenidos en la descomposición LU.

$$\mathbf{A} = \mathbf{LU} = \begin{bmatrix} -6 & 0 & 0 \\ 2 & 4 & 0 \\ -2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

13. Para cada una de las matrices del problema 11, para el cual se encontró $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$, calcule la descomposición LU sin intercambios; para la matriz $\mathbf{A}'$ que se define como $\mathbf{A}' = \mathbf{PA}$ y verifique que se obtienen las mismas matrices **L** y **U**.

▷**14.** Demuestre que la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ no es única ni aun para las matrices no singulares, encontrando dos de tales descomposiciones diferentes para

$$\mathbf{A} = \begin{bmatrix} 2 & -4 \\ 2 & -7 \end{bmatrix}.$$

15. En las aplicaciones, muchas matrices son *esparcidas*: la mayoría de sus elementos son cero. En muchos casos son *matrices de banda*, es decir, que sus elementos diferentes de cero están en una banda vecina a la diagonal principal. Una matriz tiene un *ancho de banda inferior* l si $\langle \mathbf{A} \rangle_{ij} = 0$ para $j < i - 1$ y *ancho de banda superior* u si $\langle \mathbf{A} \rangle_{ij} = 0$ para $j > i + u$. Una matriz con ancho de banda superior e inferior 1 se llama *tridiagonal*.

a) Demuestre que si se realiza la eliminación de Gauss para producir la descomposición *LU sin intercambios* para una matriz tridiagonal **A**, $\mathbf{A} = \mathbf{LU}$, entonces **L** tiene un ancho de banda inferior 1 y **U** un ancho de banda superior 1.

b) Demuestre que si se requiere de intercambios, entonces **U** puede tener ancho de banda superior 2 mientras que **L** puede tener ancho de banda inferior $p - 1$ para $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$, $p \times p$.

c) Generalice a) y b) para matrices generales de banda.

16. Encuentre la descomposición L_0U_0 de cada matriz del problema 11 a), b) y c).

▷**17.** Encuentre la descomposición L_0U_0 de **A** si **A** es:

a) $\begin{bmatrix} 2 & 0 \\ 6 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 2 & 3 & 4 \\ -4 & -4 & 1 & 4 \\ -2 & -2 & 4 & 8 \\ 6 & 6 & 2 & 2 \end{bmatrix}$

18. Demuestre el **teorema clave 3.53 b)** como sigue, mediante inducción en p, donde el resultado está claro para $p = 1$. Suponga que es verdadero para $p - 1$ y sea $\mathbf{A}$, $p \times p$.

a) Demuestre qué eliminación en la primera columna de $\mathbf{A}$ es equivalente a premultiplicar primero $\mathbf{A}$ por una matriz de permutación $\mathbf{P}_1$ y después por una matriz unitaria triangular inferior $\mathbf{L}_1$ cuyos únicos posibles elementos en la subdiagonal, diferentes de cero, están en la primera columna y son los negativos de los multiplicadores $\mathbf{m}$; ya que la matriz restante $(p - 1) \times (p - 1)$ puede escribirse como $\mathbf{P}'^T\mathbf{L}'\mathbf{U}'$ con $\mathbf{L}'$ unitaria triangular inferior mediante la hipótesis inductiva, demuestre que este producto puede ser separado como $\mathbf{P}_2\mathbf{L}_2\mathbf{U}_2$, dando

$$\mathbf{L}_1\mathbf{P}_1\mathbf{A} = \mathbf{P}_2\mathbf{L}_2\mathbf{U}_2,$$

donde

$$\mathbf{L}_1 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{m} & \mathbf{I} \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{P}'^T \end{bmatrix},$$

$$\mathbf{L}_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{L}' \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} \alpha & \mathbf{w}^T \\ \mathbf{0} & \mathbf{U}' \end{bmatrix}.$$

Premultiplicando por $\mathbf{P}_2^T\mathbf{L}_1^{-1}$ esta ecuación cambia a

$$\mathbf{P}_2^T\mathbf{P}_1\mathbf{A} = \mathbf{P}_2^T\mathbf{L}_1^{-1}\mathbf{P}_2\mathbf{L}_2\mathbf{U}_2.$$

b) Muestre que $\mathbf{L}_1^{-1}$ se parece a $\mathbf{L}_1$ a excepción de que los multiplicadores $\mathbf{m}$ se reemplazan por $-\mathbf{m}$.

c) Tomando ventaja de la estructura especial de los términos $\mathbf{P}_2^T$, $\mathbf{L}_1^{-1}$, $\mathbf{P}_2$, y $\mathbf{L}_2$ a la derecha y usando multiplicación separada, demuestre que esta ecuación puede escribirse como

$$(\mathbf{P}_2^T\mathbf{P}_1)\mathbf{A} = \begin{bmatrix} 1 & \mathbf{0}^T \\ -\mathbf{m}' & \mathbf{L}' \end{bmatrix}\begin{bmatrix} \alpha & \mathbf{w}^T \\ \mathbf{0} & \mathbf{U}' \end{bmatrix},$$

donde $\mathbf{m}' = \mathbf{P}'\mathbf{m}$; siendo ésta la $\mathbf{PA} = \mathbf{L}_0\mathbf{U}_0$ deseada

19. Mediante la multiplicación e igualación de elementos, demuestre que la $\mathbf{A}$ abajo no puede escribirse como $\mathbf{A} = \mathbf{LU}$:

$$\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \neq \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ u & 1 \end{bmatrix} \quad \text{para toda } a,\ b,\ c, \text{ y } u.$$

3.8 *MEDIDAS DE TRABAJO Y SOLUCION DE SISTEMAS LIGERAMENTE MODIFICADOS*

Un modelo matemático dado de cierto sistema de interés, comúnmente se utiliza de manera extensa para explorar las propiedades de ese sistema: se puede examinar la salida del sistema para varias entradas diferentes; se pueden estudiar los efectos de ligeras modificaciones en el propio modelo y así sucesivamente. Si el modelo involucra un sistema de ecuaciones lineales, entonces tales sondeos requieren la solución de varios sistemas de ecuaciones que difieren muy ligeramente uno de otro: se puede requerir de soluciones para varios miembros derechos diferentes pero la misma matriz de coeficientes; se puede requerir de soluciones después de unos pocos cambios en los coeficientes; y así sucesivamente.

En tales situaciones que involucran ecuaciones lineales, por lo general es posible ahorrarse una cantidad de trabajo haciendo uso de cálculos anteriores para resolver, más adelante, sistemas de ecuaciones ligeramente modificados. Con el fin de visualizar el menor esfuerzo posible, primero es importante saber qué trabajo involucra la solución de sistemas de ecuaciones lineales.

Medidas de trabajo para eliminación y sustitución en reversa

Suponga que el problema es resolver para $\mathbf{x}$ en $\mathbf{Ax} = \mathbf{b}$, donde $\mathbf{A}$ es $p \times p$, $\mathbf{b}$ es $p \times 1$, y $\mathbf{x}$ es $p \times 1$. ¿Cuánto esfuerzo involucra la eliminación de Gauss para producir un sistema reducido, y cuánto involucra la sustitución en reversa para producir $\mathbf{x}$? Por supuesto, el esfuerzo mayor está en la aritmética –las sumas, restas, multiplicaciones y divisiones esenciales al proceso. Para la mayoría de las calculadoras y computadoras, al igual que para las personas las sumas y restas consumen aproximadamente el mismo tiempo, pero son apreciablemente más rápidas que las multiplicaciones y divisiones (cuestan aproximadamente lo mismo). Por lo tanto, al medir el trabajo se contará la suma/resta separadamente de las multiplicaciones/divisiones pero se agrupará aquél en estas dos categorías.

Considere la eliminación de Gauss en la columna k de la matriz $\mathbf{A}$, $p \times p$; se contabilizará el trabajo de procesar $\mathbf{b}$ más adelante. El renglón que contiene el pivote debe dividirse entre ese pivote; se puede evitar dividir el pivote entre sí mismo ya que se conoce que la respuesta es 1. Por lo tanto, se realizarán $p - k$ divisiones. A cada renglón de abajo se le debe sumar un múltiplo m del renglón pivotal; el múltiplo m es tan sólo el negativo del coeficiente que se busca eliminar (por lo tanto, la obtención de m es gratuita) y no se tiene que calcular ese elemento eliminado ya que se sabe que el resultado es 0. Por lo tanto, se deben realizar $p-k$ multiplicaciones y $p-k$ sumas para *cada uno* de los $p-k$ renglones debajo del renglón pivotal.

El costo total de la eliminación de Gauss es la sumatoria de todos estos costos:

$$\sum_{k=1}^{p}\left\{\{(p-k)^2 \text{ en } + \text{ y } -\} + \{(p-k)+(p-k)^2 \text{ en } \times \text{ y } /\}\right\}.$$

Usando los hechos de que

$$1+2+\cdots+n=\frac{n(n+1)}{2} \quad \text{y} \quad 1+4+\cdots+n^2=\frac{n(n+1)(2n+1)}{6},$$

se puede calcular el trabajo total como $p^3/3 - p/3$ multiplicaciones/divisiones y $p^3/3 - p^2/2 + p/6$ sumas/restas.

Por supuesto, resolver $\mathbf{Ax} = \mathbf{b}$ requiere que se procese $\mathbf{b}$ y también que se realice la sustitución en reversa; esto se expresa fácilmente en términos de la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ del **teorema clave 3.53.** $\mathbf{Ax} = \mathbf{b}$ significa que $\mathbf{P}^T\mathbf{LUx} = \mathbf{b}$; entonces $\mathbf{LUx} = \mathbf{Pb}$, lo que se escribe como $\mathbf{Ly} = \mathbf{Pb}$ y $\mathbf{Ux} = \mathbf{y}$, cada una de las cuales es simplemente un sistema *triangular* que puede resolverse inmediatamente con sustitución hacia adelante o en reversa. La matriz columna $\mathbf{y}$ es precisamente lo que se hubiera obtenido si se hubieran realizado todos los pasos de eliminación en $\mathbf{b}$ durante la eliminación de Gauss.

(3.57) Para resolver $\mathbf{Ax} = \mathbf{b}$ dada la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$:
1. Permute los elementos de $\mathbf{b}$ para obtener $\mathbf{b}' = \mathbf{Pb}$.
2. Resuelva el sistema triangular inferior $\mathbf{Ly} = \mathbf{b}'$ para $\mathbf{y}$ mediante la sustitución hacia adelante.
3. Resuelva el sistema unitario triangular superior $\mathbf{Ux} = \mathbf{y}$ para $\mathbf{x}$ mediante la sustitución en reversa.

(3.58) ***Ejemplo.*** Considere solucionar el sistema de ecuaciones del ejemplo 3.17 usando (3.57); la descomposición LU, $\mathbf{P}^T\mathbf{LU}$ de $\mathbf{A}$ es

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} -2 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ -2 & 2 & 5 & 0 \\ 3 & 0 & 3 & -5.6 \end{bmatrix}\begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 1 & 2.2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ya que $\mathbf{b} = [4 \ \ -1 \ \ 1 \ \ -3]^T$ y $\mathbf{b}' = \mathbf{Pb}$ es tan sólo $\mathbf{b}' = [4 \ \ -3 \ \ 1 \ \ -1]^T$. Se calcula fácilmente la solución $\mathbf{y}$ de $\mathbf{Ly} = \mathbf{b}'$ como $\mathbf{y} = [-2 \ \ -1 \ \ -\frac{1}{5} \ \ -1]^T$, y la solución $\mathbf{x}$ a $\mathbf{Ux} = \mathbf{y}$ se obtiene fácilmente como $\mathbf{x} = [1 \ \ 1 \ \ 2 \ \ -1]$. Esto resuelve verdaderamente el sistema de ecuaciones en el ejemplo 3.17.

Es fácil contabilizar el trabajo involucrado en (3.57) para una matriz $\mathbf{A}$, $p \times p$.

El paso 1 involucra intercambios pero nada de aritmética.

En el paso 2 se requiere encontrar $\langle \mathbf{y} \rangle_k$ en el orden de $k = 1, 2, \ldots, p$. Al encontrar $\langle \mathbf{y} \rangle_k$ los $k - 1$ valores anteriores de $\mathbf{y}$ se sustituyen, se multiplican por un coeficiente, y se restan al miembro derecho; entonces una división produce $\langle \mathbf{y} \rangle_k$. Por lo tanto, para cada k el costo es de k multiplicaciones/divisiones y $k - 1$ sumas/restas. Esto da un costo total de $p(p + 1)/2$ multiplicaciones/divisiones y $p(p - 1)/2$ sumas/restas para el paso 2.

En el paso 3 se requiere encontrar $\langle \mathbf{x} \rangle_k$ en el orden de $k = p, p - 1, \ldots, 1$. El trabajo es similar al del paso 2 a excepción de que la igualdad $\langle \mathbf{U} \rangle_{kk} = 1$ elimina la necesidad de una división para cada variable. De allí que el costo total sea $p(p - 1)/2$ multiplicaciones/divisiones y $p(p - 1)/2$ sumas/restas.

Ahora se pueden calcular fácilmente los costos combinados de los tres pasos de (3.57).

(3.59) **Teorema** (trabajo en eliminación y solución). Suponga que $\mathbf{A}$ es una matriz no singular, $p \times p$. Entonces:

a) El costo de calcular la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ de $\mathbf{A}$ es

$p^3/3 - p/3$ multiplicaciones/divisiones y
$p^3/3 - p^2/2 + p/6$ sumas/restas

b) Dada la descomposición LU, el costo de resolver $\mathbf{Ax} = \mathbf{b}$ mediante el proceso de (3.57) es de

p^2 multiplicaciones/divisiones y
$p^2 - p$ sumas/restas

Los puntos claves a observar son que el trabajo en la eliminación de Gauss para producir la descomposición LU varía con el *cubo* de p, mientras que el costo de resolver un sistema –una vez que se dispone de la descomposición LU– varía con el *cuadrado* de p. El costo total para un solo sistema, por lo tanto, varía con p^3: Duplicar el número de ecuaciones y variables hace que el trabajo sea *ocho* veces mayor. Sin embargo, observe que el costo de multiplicar simplemente dos matrices $p \times p$ es de p^3 multiplicaciones y $p^3 - p^2$ sumas –ambas mayores que en (3.59). Observe también, que el resolver un solo sistema de 100 ecuaciones con 100 incógnitas requiere de 343,300 multiplicaciones/divisiones y 338,250 sumas/restas– una tarea enorme para la mayoría de los seres humanos, por lo que es importante resaltar que muchas microcomputadoras realizan 10,000 operaciones aritméticas por segundo, y podrían hacer esto en aproximadamente 70 segundos, mientras que las supercomputadoras modernas requerirían de una pequeñísima fracción de segundo.

Puede usarse el mismo tipo de análisis para contabilizar el trabajo en la eliminación de Gauss-Jordan; el costo total para resolver un solo sistema de p ecuaciones lineales con p incógnitas es $p^3/2 + p^2/2$ multiplicaciones/divisiones y

$p^3/2 - p/2$ sumas/restas. Aquí los términos dominantes son $p^3/2$ contra $p^3/3$ para la eliminación de Gauss; en la eliminación de Gauss-Jordan se consume alrededor de 50% más del tiempo que en la eliminación de Gauss para resolver un sistema de ecuaciones. De hecho, se sabe que ningún método que utilice sólo operaciones de renglón (o de columna) puede tomar menos trabajo para matrices generales que la eliminación de Gauss.

Ahora podrá resolver los problemas del 1 al 6.

Modificación de los miembros derechos en ecuaciones

Suponga ahora que es necesario resolver $\mathbf{Ax}_i = \mathbf{b}_i$ con varios miembros derechos diferentes $\mathbf{b}_i$, $i = 1, 2, \ldots, n$; esto podría fácilmente ocurrir en un estudio de diseño, como se indicó antes. Si desde el comienzo se conocen todas las $\mathbf{b}_i$, entonces, de acuerdo a (3.28), simplemente se pueden usar la eliminación de Gauss y la sustitución en reversa en la matriz múltiplo-aumentada

$$[\mathbf{A} \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n].$$

En estudios de diseño y otras aplicaciones, sin embargo, se desconoce con frecuencia el siguiente miembro derecho $\mathbf{b}_{i+1}$ hasta haber obtenido los resultados $\mathbf{x}_i$ del miembro derecho presente; por lo tanto, no se puede proponer la matriz múltiplo-aumentada. Pero aún así se desea evitar realizar la eliminación de Gauss otra vez para cada uno de los n diferentes sistemas $p \times p$, ya que esto costaría alrededor de $n(p^3/3)$ operaciones aritméticas de acuerdo al teorema (3.59).

La descomposición LU de $\mathbf{A}$ proporciona una resolución eficiente a este problema. Primero se calcula la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ de $\mathbf{A}$, y después se utiliza (3.57) en cada uno de los lados derechos $\mathbf{b}_i$ conforme se va disponiendo de ellos. De acuerdo a las medidas de trabajo en el teorema (3.59):

(3.60) Se puede obtener la descomposición LU: $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ de $\mathbf{A}$ y resolver n conjuntos de ecuaciones diferentes $\mathbf{Ax}_i = \mathbf{b}_i$, para $1 \le i \le n$ a un costo total de

$$p^3/3 - p/3 + np^2 \text{ multiplicaciones/divisiones, y}$$
$$p^3/3 - p^2/2 + p/6 + n(p^2 - p) \text{ sumas/restas}$$

Una de las razones para querer resolver $\mathbf{Ax}_i = \mathbf{b}_i$ para varias $\mathbf{b}_i$ podría ser la de obtener la inversa $\mathbf{A}^{-1}$, ya que

$$\mathbf{A}^{-1} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_p]$$

si $\mathbf{b}_i = \mathbf{e}_i$, la matriz columna unitaria de orden p –ver **teorema clave 1.44.** En este caso $n = p$ y las medidas de trabajo en (3.60) son de aproximadamente $4p^3/3$ operaciones. Sín embargo, cada $\mathbf{e}_i$ es un tanto especial, ya que, básicamente, contiene ceros. Si se puede evitar sumar los ceros y evitar multiplicar y dividir por ceros, se puede reducir bastante el trabajo al aplicar (3.57) a estos miembros derechos $\mathbf{e}_i$ especiales. El resultado es:

(3.61) La inversa de una matriz no singular, $p \times p$ puede calcularse mediante la eliminación de Gauss y la sustitución en reversa con

$$p^3 \text{ multiplicaciones/divisiones, y}$$
$$p^3 - 2p^2 + p \text{ sumas/restas}$$

Aquí hay un hecho sorprendente que se sigue de estas medidas de trabajo: No importa el número n de sistemas $\mathbf{Ax}_i = \mathbf{b}_i$ que se necesiten resolver con la misma matriz **A**, ya que siempre involucra menor esfuerzo en el caso de una **A** general encontrar la descomposición LU de **A** y utilizar (3.57) para cada lado derecho, que el calcular $\mathbf{A}^{-1}$ mediante eliminación de Gauss y después calcular cada $\mathbf{x}_i$ como $\mathbf{A}^{-1}\mathbf{b}_i$. Vea el problema 7. Esta es una de las razones por las que en la práctica, es raro que la inversa $\mathbf{A}^{-1}$ sea realmente útil –la mayor parte de lo que se puede lograr con $\mathbf{A}^{-1}$ puede lograrse de manera más eficiente a través de otros medios, tales como la descomposición LU. Por supuesto, existen excepciones, como cuando **A** tiene una estructura muy especial o cuando realmente necesitan verse los elementos en $\mathbf{A}^{-1}$; aun así es una buena regla general.

Ahora podrá resolver los problemas del 1 al 7.

Modificación de los coeficientes en las ecuaciones

Ahora se pondrá atención al caso en el cual no **b** sino **A** cambia de algún modo un tanto simple. Si se cambia el papel que juega una variable en particular o una ecuación en un modelo, por ejemplo, entonces quizá sea necesario investigar cómo resolver sistemas después de haber cambiado un renglón o una columna de una matriz. Es fácil describir esta situación en lenguaje matricial. Para reemplazar la i-ésima columna $\mathbf{c}_i$ de **A** por la matriz columna $\mathbf{c}_i'$, sólo se necesita sumar a **A** la matriz $(\mathbf{c}_i' - \mathbf{c}_i)\mathbf{e}_i^T$; sumar por otro lado $\mathbf{e}_i(\mathbf{r}_i - \mathbf{r}_i)^T$ en su lugar, reemplaza el i-ésimo renglón r_i por r_i'.

(3.62) ***Ejemplo.*** Sean

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 0 & -2 & 4 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Por lo tanto

$$\mathbf{A} + \mathbf{d}\mathbf{e}_1^T = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 0 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} [1 \quad 0 \quad 0]$$

$$= \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 0 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ 4 & 1 & 2 \\ 1 & -2 & 4 \end{bmatrix}$$

Mientras que

$$\mathbf{A} + \mathbf{e}_2\mathbf{d}^T = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 0 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} [2 \quad 1 \quad 1] = \begin{bmatrix} 1 & 2 & -1 \\ 5 & 2 & 3 \\ 0 & -2 & 4 \end{bmatrix}.$$

De manera más general, considere sumar a **A** una matriz (construida de unas cuantas matrices columna y unas cuantas matrices renglón) de la forma **CR**, donde **C** es $p \times n$ y **R** es $n \times p$, con n mucho menor que p. El siguiente resultado describe la inversa de la nueva matriz $\mathbf{A}' = \mathbf{A} + \mathbf{CR}$ y, quizá lo más importante, cómo obtener la solución $\mathbf{x}'$ de $\mathbf{A}'\mathbf{x}' = \mathbf{b}$ de la de $\mathbf{Ax} = \mathbf{b}$.

(3.63) ***Teorema*** (inversas de matrices modificadas). Suponga que **A** es $p \times p$ y no singular, que **C** es $p \times n$ y **R** es $n \times p$ con $n \leq p$ (por lo general n mucho menor que p), y que la matriz $\mathbf{K} = \mathbf{I}_n + \mathbf{RA}^{-1}\mathbf{C}$ $n \times n$ es no singular. Entonces:

a) $\mathbf{A}' = \mathbf{A} + \mathbf{CR}$ es no singular, y

$$\mathbf{A}'^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{CK}^{-1}\mathbf{RA}^{-1}.$$

b) La solución $\mathbf{x}'$ de $\mathbf{A}'\mathbf{x}' = \mathbf{b}$ puede expresarse en términos de la solución $\mathbf{x}$ de $\mathbf{Ax} = \mathbf{b}$ como

$$\mathbf{x}' = \mathbf{x} - \mathbf{A}^{-1}\mathbf{CK}^{-1}\mathbf{Rx}.$$

c) Dada $\mathbf{A}^{-1}$, se puede calcular $\mathbf{A}'^{-1}$ mediante a) con aproximadamente $3np^2 + 2n^2p + n^3/3$ multiplicaciones/divisiones y un número comparable de sumas/restas.

d) Dada la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$, y la solución $\mathbf{x}$ de $\mathbf{Ax} = \mathbf{b}$, se puede calcular $\mathbf{x}'$ mediante b) con aproximadamente $np^2 + n^2p + n^3/3$ multiplicaciones/divisiones y un número comparable de sumas/restas.

DEMOSTRACION

a) Sea **H** la inversa de $\mathbf{A}'$ mencionada; debe demostrarse que $\mathbf{A}'\mathbf{H} = \mathbf{HA}' = \mathbf{I}_p$ —esto es, el problema 8.

b) $\mathbf{x}' = \mathbf{A}'^{-1}\mathbf{b} = (\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}\mathbf{K}^{-1}\mathbf{R}\mathbf{A}^{-1})\mathbf{b} = \mathbf{A}^{-1}\mathbf{b} - \mathbf{A}^{-1}\mathbf{C}\mathbf{K}^{-1}\mathbf{R}\mathbf{A}^{-1}\mathbf{b}$
$= \mathbf{x} - \mathbf{A}^{-1}\mathbf{C}\mathbf{K}^{-1}\mathbf{R}\mathbf{x}$, como se enunció.

c) y d) se basan en las medidas de trabajo antes mencionadas y en el hecho de que requiere alrededor de rst multiplicaciones (y un número similar de sumas) calcular el producto de una matriz $r \times s$ por una matriz $s \times t$ –éste es el problema 9. ■

(3.64) ***Ejemplo.*** Se ilustra el punto sobre las medidas de trabajo del teorema 3.36. Suponga que $p = 100$ y $n = 5$; suponga que ya se ha obtenido la descomposición LU de la matriz $\mathbf{A}$, 100×100. Ahora, suponga que es necesario resolver $\mathbf{A}'\mathbf{x}'' = \mathbf{b}$ para alguna matriz columna $\mathbf{b}$. Si se ignora el teorema 3.63 y se aplica la eliminación de Gauss a $\mathbf{A}'\mathbf{x}' = \mathbf{b}$, se requiere de alrededor de $p^3/3$ multiplicaciones/divisiones, ($\approx 333{,}000$). Suponga que en lugar de ello se utiliza el teorema 3.63 b). Primero, se necesita $\mathbf{x}$, la solución a $\mathbf{A}\mathbf{x} = \mathbf{b}$; ya que se tiene la descomposición LU de $\mathbf{A}$; esto se puede obtener a un costo de alrededor de p^2 multiplicaciones/divisiones. Entonces se puede usar el método del teorema 3.63 d) con el fin de obtener $\mathbf{x}'$ con un costo de alrededor de $np^2 + n^2p + n^3/3$ multiplicaciones/divisiones. El esfuerzo total con este método es, por lo tanto, de aproximadamente

$$p^2 + np^2 + n^2p + \frac{n^3}{3} \approx 62{,}500$$

multiplicaciones/divisiones –ahorrando alrededor del 80% con respecto al primer método. Para $p = 100$ y $n = 2$, como otro ejemplo, el ahorro es de alrededor del 90%: el método directo toma aproximadamente 10 veces más tiempo.

A primera instancia puede parecer que las circunstancias del teorema 3.63 aparecen raras veces en la práctica, pero éste no es el caso. Con mucha frecuencia, modelos extremadamente grandes de sistemas complicados –grandes edificios, redes de distribución eléctrica, grandes sistemas económicos, etc., consisten en un número pequeño de subsistemas muy grandes y casi independientes, con un pequeño número de interconexiones entre ellos. El teorema 3.63 puede usarse con frecuencia en este caso, siendo $\mathbf{A}$ la representación de la situación en la que los subsistemas son realmente independientes con $\mathbf{C}$ y $\mathbf{R}$ proporcionando los ajustes que reflejen las interconexiones. Por ejemplo, con sólo dos subsistemas, se podría tener que resolver un sistema de 2000 ecuaciones con 2000 incógnitas, con una matriz de coeficientes $\mathbf{A}'$ y la forma separada no estándar siguiente:

$$\mathbf{A}' = \begin{array}{c} \\ 1000\langle \\ 1000\langle \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1000 & 1000 \\ \wedge & \wedge \end{array} \\ \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \end{array} + \begin{array}{c} \begin{array}{cccc} 990 & 10 & 10 & 990 \\ \wedge & \wedge & \wedge & \wedge \end{array} \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{array},$$

con las dimensiones de los bloques como se indican. Sea **A** la primera de las dos matrices separadas, y suponga que **B** y **D** –y por lo tanto **A**– son no singulares. Se concibe **B** como la representación de las relaciones internas dentro de un subsistema grande, **D** como la relación interna dentro del otro subsistema grande, y **E** y **F** como las interconexiones entre los dos subsistemas. Como en el teorema 3.63, se puede escribir $\mathbf{A}'$ con $\mathbf{A}' = \mathbf{A} + \mathbf{CR}$, donde

$$\mathbf{C} = \begin{array}{c} \\ 1000\langle \\ 1000\langle \end{array}\begin{array}{c} \begin{array}{cc} 10 & 10 \\ \wedge & \wedge \end{array} \\ \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ \mathbf{F} & \mathbf{0} \end{bmatrix} \end{array} \quad \text{y} \quad \mathbf{R} = \begin{array}{c} \\ 10\langle \\ 10\langle \end{array}\begin{array}{c} \begin{array}{cccc} 990 & 10 & 10 & 990 \\ \wedge & \wedge & \wedge & \wedge \end{array} \\ \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix} \end{array}.$$

En tales circunstancias los ahorros al usar el teorema 3.63 pueden ser enormes: pueden resolverse algunos problemas que quizá no pudieran resolverse de modos más directos.

Se han desarrollado varias técnicas especiales basadas esencialmente en este método, las cuales tienen una amplia variedad de áreas de aplicación. Esta idea básica lleva, por ejemplo, al *método de rompimiento* en análisis de reses, a *métodos de matrices de capacitancia* para la solución numérica directa de ciertas ecuaciones diferenciales, a varios *métodos modernos* para la optimización de programación no lineal restringida y no restringida, y a ciertas implementaciones del *método simplex* en la programación lineal

PROBLEMAS 3.8

▷ **1.** Demuestre que con el fin de multiplicar una matriz $r \times s$ por una matriz $s \times t$, se requieren de, a lo mucho, rst multiplicaciones y $rt(s - 1)$ sumas.

2. Utilice las fórmulas

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

y

$$1 + 4 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

para demostrar el teorema 3.59 a).

3. Utilice las fórmulas dadas en el problema 2 para demostrar el teorema 3.59 b).

4. Deduzca las medidas de trabajo enunciadas en el texto para la eliminación de Gauss-Jordan.

▷ **5.** Utilice el teorema 3.59 a) para calcular el número de multiplicaciones/divisiones que se necesitan para encontrar la descomposición LU y resolver un sistema de ecuaciones para una matriz $p \times p$, si:

a) $p = 10$ b) $p = 30$ c) $p = 50$
d) $p = 70$ e) $p = 90$

6. Considere la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{L}\mathbf{U}$ que se encontró en el ejemplo 3.45 de la sección 3.7. Para dicha matriz $\mathbf{A}$, utilice (3.57) para resolver

$$\mathbf{A}\mathbf{x} = \mathbf{b} = [-4 \quad 5 \quad 13 \quad 20]^T$$

y verifique que la $\mathbf{x}$ encontrada resuelve las ecuaciones originales.

▷ **7.** Demuestre que el trabajo que se requiere para obtener la descomposición LU, $\mathbf{A} = \mathbf{P}^T\mathbf{L}\mathbf{U}$ para una $\mathbf{A}$ general es siempre menor que el que se requiere para calcular $\mathbf{A}^{-1}$ mediante la eliminación de Gauss, y utilícelo para resolver k sistemas $\mathbf{A}\mathbf{x}_i = \mathbf{b}_i$ para $1 \le i \le k$, y calcular $\mathbf{x}_i$ como $\mathbf{x}_i = \mathbf{A}^{-1}\mathbf{b}_i$ para $1 \le i \le k$.

8. Complete la demostración del teorema 3.63 a).

9. Complete la demostración del teorema 3.63 c) y d).

10. Suponga que $p = 1000$ y $n = 10$. Al igual que como en el ejemplo 3.64, calcule el ahorro al usar el método del teorema 3.63 d) para resolver $\mathbf{A}'\mathbf{x}' = \mathbf{b}$.

▷ **11.** Sea $\mathbf{A}$ una matriz no singular $p \times p$ y $\mathbf{r}$, $1 \times p$. Describa el uso del teorema 3.63 a) para encontrar la inversa, si existe, de:
a) $\mathbf{A} + \mathbf{e}_i\mathbf{r}$ b) $\mathbf{A} + \mathbf{r}^T\mathbf{e}_j^T$

12. Sea $\mathbf{A}$ la matriz no singular $p \times p$. Describa el uso del teorema 3.63 a) para encontrar la inversa, si existe, de la matriz que se forma sumando el número α al elemento (i, j) de $\mathbf{A}$.

13. Utilice el teorema 3.63 a) del modo descrito en el párrafo siguiente a (3.64) para encontrar la inversa de

$$\mathbf{A}' = \left[\begin{array}{cc|cc} 3 & 5 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right] + \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

𝔐 **14.** Considere la matriz $\mathbf{A}'$, 50×50 con elementos en la diagonal principal $\langle \mathbf{A}' \rangle_{ii} = 2$ para toda i, elementos arriba de la diagonal $\langle \mathbf{A}' \rangle_{i,i+1} = -1$ para toda i, y elementos abajo de la diagonal $\langle \mathbf{A}' \rangle_{i-1,i} = -1$ para toda i, y cualquier otro elemento igual a cero. Igual que en el párrafo siguiente a (3.64), visualice $\mathbf{A}'$ como la suma $\mathbf{A} + \mathbf{C}\mathbf{R}$ de la forma

$$\begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

donde **B** es la matriz 25 × 25 que parece una versión más pequeña de **A**′, donde **F** y **E** son 25 × 1 con $\mathbf{F} = [-1 \ 0 \ \cdots \ 0]^T$ y $\mathbf{E} = [0 \ 0 \ \cdots \ 0 \ -1]^T$.

a) Utilice el MATLAB para encontrar la inversa de **A**′ directamente –utilizando el comando *flops* del MATLAB inmediatamente antes y después del cálculo– y encuentre el número de operaciones en punto flotante requeridas.

b) Utilice el MATLAB para encontrar la inversa de **A**′ por medio del teorema 3.63 a) –calcule la inversa de **A** formando la matriz separada evidente usando la inversa $\mathbf{B}^{-1}$– y compare el número de operaciones en punto flotante requeridas por los dos métodos.

▷ **15.** Considere el problema 15 de la sección 3.7, donde se consideró la eliminación de Gauss con y sin intercambios para matrices tridiagonales y para matrices de banda.

a) Muestre que el trabajo para encontrar la descomposición *LU* de una matriz tridiagonal $p \times p$ *sin intercambios* es igual a $(2p - 2)$ multiplicaciones/divisiones y $(p - 1)$ sumas/restas.

b) Demuestre que si se permiten los intercambios, el trabajo que se requiere es a lo mucho $(4p - 6)$ multiplicaciones/divisiones y $(p - 1)$ sumas/restas.

c) Generalice a) y b) para matrices generales de banda.

3.9 PROGRAMAS COMPUTACIONALES PARA LA ELIMINACION DE GAUSS

> **ADVERTENCIA:** ¡Siempre utilice un programa estándar en lugar de escribir su propio programa para la solución de ecuaciones lineales o para encontrar inversas!
>
> Si no tiene un programa estándar copie uno de Forsythe, Malcolm y Moler (42).

En las aplicaciones prácticas, la resolución de sistemas de ecuaciones lineales y el cálculo de inversas se llevan a cabo casi siempre mediante el uso de computadoras –supercomputadoras, macrocomputadoras estándar o microcomputadoras. A través de los años los expertos han desarrollado programas computacionales de propósito general excelentes para la resolución eficiente y precisa de sistemas de ecuaciones y para el cálculo de inversas. En la sección 3.6 se dio una ligera indicación sobre las sutilezas involucradas en esta tarea. Ya que los resultados de dichos esfuerzos pueden encontrarse ya sea gratis o a bajo costo, no hay razón para que los inexpertos tengan que crear sus propios programas para estos problemas. Y existen buenas razones para *no* crear su propio programa –es muy probable que proporcione respuestas inexactas (y costosas) sin ningún aviso de que las respuestas son poco precisas. Aunque pueda encontrar ilustrativo el escribir un programa sencillo para implementar la eliminación de Gauss con o sin intercambios como una herramienta pedagógica para ayudarlo en la comprensión de lo que el proceso hace exactamente, *no* utilice tal programa en problemas reales cuyas respuestas le sean importantes.

Qué exigir del software

Lo mínimo que se puede exigir es que el software sea eficiente, exacto, sencillo y fácil de usar, además de informativo con respecto a la confianza con que los resultados puedan aceptarse en cada caso específico. Características adicionales podrían incluir la capacidad de utilizar métodos especiales para problemas especiales (tales como matrices simétricas, matrices cuyos elementos diferentes de cero están cerca de la diagonal principal, y así sucesivamente), la capacidad de evaluar determinantes (ver sección 4.5), y así sucesivamente. No todo el software para las ecuaciones e inversas cumple con este criterio, pero la mayoría lo hace; vea la siguiente subsección para sugerencias específicas.

Lo que se quiere decir con eficiencia y precisión puede parecer claro, aunque de hecho aquí existen algunos puntos oscuros; sin embargo, estos temas están más allá del alcance de este libro. Los detalles pueden encontrarse en Golub y Van Loan (46).

Un software que es fácil y sencillo de usar debe permitir que las matrices de datos puedan introducirse en algún estilo natural y directo, y no debe pedirle al usuario que proporcione mucha información adicional.

Un aspecto especialmente importante es el de proporcionar información sobre la confiabilidad de los resultados. Muchos buenos programas en esta área dan un indicio de los llamados *números condicionales* $c(\mathbf{A})$ de la matriz $\mathbf{A}$ que es la matriz de coeficientes de las ecuaciones e inversas que se están buscando; vea el teorema 6.29. Este número $c(\mathbf{A})$ mide la cantidad por la cual las inexactitudes en los datos de los problemas se magnifican en la respuesta; si los datos contienen errores de tamaño relativo, por ejemplo, 10^{-3}, entonces puede ser que las respuestas tengan errores de tamaño relativo tan grande como $c(\mathbf{A}) \times 10^{-3}$. Estos errores no son "culpa" del programa, más bien son inherentes a las respuestas debido a las inexactitudes de los datos. Ya que una realización cuidadosa de la eliminación de Gauss permite estimar este número condicional, el proporcionar tal estimador es una característica razonable a exigir.

Qué software obtener

Un software de alta calidad que cumpla con el criterio general anterior, ciertamente puede encontrarse comercialmente; el sistema MATLAB de The Math Works es un ejemplo evidente pero se dispone de sistemas excelentes donde quiera como en las International Mathematical and Statistical Libraries {**IMSL**

–vea (57)} (Bibliotecas Internacionales de Matemáticas y Estadística **BIME**) y en la Numerical Algorithms Group {**NAG** –ver (59)}, (Grupo de Algoritmos Numéricos **GAN**). Para macrocomputadoras, se dispone de subrutinas actualizadas sin costo en el sistema LINPACK.

El LINPACK se desarrolló como un esfuerzo cooperativo entre muchos analistas numéricos en muchas instituciones; la mayoría del trabajo se concentró en el Laboratorio Nacional de Argonne (Argonne National Laboratory), y el apoyo financiero lo proporcionó la Fundación Nacional de Ciencias (National Science Foundation). El resultado fue una excelente colección de programas portátiles FORTRAN de los que se puede disponer gratuitamente en el Laboratorio Nacional de Argonne; los detalles de los programas y de su adquisición están contenidos en el *Manual del Usuario LINPACK* (40) (*LINPACK User's Guide*). El LINPACK contiene programas para la mayoría de los cómputos matriciales relacionados con las ecuaciones, incluyendo, por supuesto, la solución de sistemas de ecuaciones y la inversión de matrices.

Actualmente existen unas cuantas docenas de programas en LINPACK para la solución de sistemas o para encontrar inversas; varían por las estructuras especiales supuestas para la matriz (general, simétrica, y así sucesivamente), el tipo de los números usados como elementos de las matrices (real, complejo, etc), el modo en que los elementos de las matrices se almacenan en la computadora, y la precisión deseada. Las rutinas son de dos tipos básicos: aquéllas que calculan las descomposiciones LU, y aquéllas que utilizan las descomposiciones LU (para resolver sistemas, encontrar inversas, evaluar determinantes, y así sucesivamente). La metodología básica es la eliminación de Gauss y la sustitución en reversa, semejante en mucho a la que se ha descrito aquí, pero una variante es que $\mathbf{L}$ es *unitaria* triangular inferior, mientras que $\mathbf{U}$ tiene elementos generales en la diagonal principal –es decir, la descomposición L_0U_0 (3.49) descrita aquí. Las rutinas para encontrar descomposiciones LU estiman el número condicional e informan de esta estimación como un parámetro de calidad accesible al usuario; esto permite que el usuario determine la precisión de cualquier resultado a generarse. Cualquiera que necesite solucionar sistemas de ecuaciones –o realizar otros cómputos matriciales– en una macrocomputadora debería considerar adquirir el sistema LINPACK.

PROBLEMAS 3.9

1. Determine si la colección de subrutinas LINPACK está disponible para su uso.
2. Determine de qué programas se dispone en su localidad, para la solución de p ecuaciones lineales con p incógnitas y de qué posibilidades especiales se dispone (tales como elementos complejos, estructuras especiales para la matriz de coeficientes, etc.).

3. Determine de qué programas se dispone en su localidad, para invertir matrices, $p \times p$ y de qué posibilidades especiales se dispone (tales como elementos complejos, estructuras especiales para la matriz, etc.).
4. Determine si el software disponible en su localidad para sistemas lineales e inversas, provee una estimación del número condicional para el problema.

3.10 PROBLEMAS VARIOS

PROBLEMAS 3.10

1. Suponga que $\mathbf{A}$ es $p \times p$, $\mathbf{B}$ es $p \times q$, y $\mathbf{AB} = \mathbf{0}$. Demuestre que o bien $\mathbf{A}$ es singular o bien $\mathbf{B} = \mathbf{0}$.

▷ 2. La matriz mostrada abajo es la matriz aumentada para un sistema de tres ecuaciones con tres incógnitas. Demuestre que:
   a) Si $k = 0$, entonces el sistema tiene un número infinito de soluciones.
   b) Para otro valor específico de k, que debe determinar, el sistema no tiene soluciones.
   c) Para cualquier otro valor de k, el sistema tiene una solución única

$$\mathbf{A} = \left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 2 & k & 6 & 6 \\ -1 & 3 & k-3 & 0 \end{array}\right]$$

3. a) Suponga que $\mathbf{A}$ es $p \times q$, $\mathbf{B}$ es $q \times q$, y $\mathbf{AB} = \mathbf{0}$. Demuestre que, ya sea $\mathbf{B}$ es singular o $\mathbf{A} = \mathbf{0}$.
   b) Suponga que $\mathbf{A}$ y $\mathbf{B}$ son $p \times p$, y $\mathbf{AB} = \mathbf{0}$. Utilice a) y el problema 1 para demostrar que ya sea $\mathbf{A} = \mathbf{0}$ o $\mathbf{B} = \mathbf{0}$ o bien, que $\mathbf{A}$ y $\mathbf{B}$ son ambas singulares.

4. Encuentre la descomposición LU de la matriz $\mathbf{A}$ que se muestra abajo, no mediante eliminación de Gauss, sino escribiendo $\mathbf{A} = \mathbf{LU}$ con $\mathbf{L}$ y $\mathbf{U}$ como se muestra abajo y después resolviendo para los elementos en $\mathbf{L}$ y $\mathbf{U}$.

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 4 & -1 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix}.$$

▷ 5. Generalice el problema 4 como sigue. Suponga que $\mathbf{A}$ es $p \times p$ y que tiene una descomposición LU, $\mathbf{A} = \mathbf{LU}$ para $\mathbf{L}$ triangular inferior con elementos en la diagonal principal diferentes de cero y $\mathbf{U}$ unitaria triangular superior. Escriba la forma general para $\mathbf{L}$ y $\mathbf{U}$ con elementos l_{ij} y u_{ij} como en el problema 4. Se conoce el elemento (1, 1) de $\mathbf{U}$. Demuestre entonces que esto le permite encontrar la primera columna de $\mathbf{L}$, la cual permite encontrar el primer renglón de $\mathbf{U}$, el cual permite encontrar la segunda columna de $\mathbf{L}$, y así sucesivamente, hasta que se determinan $\mathbf{L}$ y $\mathbf{U}$. Escriba las ecuaciones para calcular $\mathbf{L}$ y $\mathbf{U}$ de este modo.