Capítulo 6

Multiplicación de matrices

Multiplicar matrices es mucho más complicado que multiplicar números normales. Las reglas de toda la vida que conoces para la multiplicación (por ejemplo, $a \times b = b \times a$) muchas veces no sirven para las matrices. Además, tendrás que aprender otras reglas nuevas y varias formas diferentes de multiplicar matrices. Y, para complicarlo aún más, no todos los pares de matrices pueden multiplicarse. Así que, respira hondo, porque en este capítulo nos sumergiremos de lleno en el álgebra lineal.

6.1 Multiplicación de matrices "estándar"

Como no tenemos un nombre mejor, a este método lo llamaremos el "estándar". A menos que se indique explícitamente, puedes asumir (en este libro y en cualquier otro sitio) que cuando hay escritas dos matrices una junto a otra (así: $\mathbf{AB}$), se trata de una multiplicación (o producto) "estándar" de matrices.

Terminología Lo primero que tienes que saber sobre la multiplicación de matrices es que no tiene la propiedad conmutativa; es decir, $\mathbf{AB} \neq \mathbf{BA}$. Hay algunos excepciones en que esta igualdad sí es cierta (por ejemplo, $\mathbf{AI} = \mathbf{IA}$), pero es raro. Esto hace que incluso la terminología de la multiplicación de matrices sea complicada .

Las siguientes cinco afirmaciones son formas de referirse a la operación $\mathbf{AB}$ con palabras (¡puedes decirlo si quieres presumir de tus conocimientos matemáticos con tus amigos y familia!).

"**A** por **B**"
"**A** multiplica a **B** por la izquierda"
"**A** premultiplica a **B**"
"**B** multiplica a **A** por la derecha"
"**B** posmultiplica a **A**"

Cuándo es válida una multiplicación Antes de aprender cómo funciona la multiplicación estándar, necesitas saber cuándo *puede hacerse*. La regla de cuándo dos matrices pueden multiplicarse es sencilla y visual, y tendrás que memorizarla antes de aprender la "mecánica" de cómo se multiplica.

Si escribes el tamaño (u orden) de las matrices debajo de ellas, la multiplicación solo podrá hacerse cuando las dos dimensiones que quedan en la parte de dentro sean iguales. El tamaño de la matriz resultante viene dada por las dimensiones de fuera. Con "de dentro" y "de fuera" me refiero a la posición donde quedan escritas, como ves en la figura 6.1.

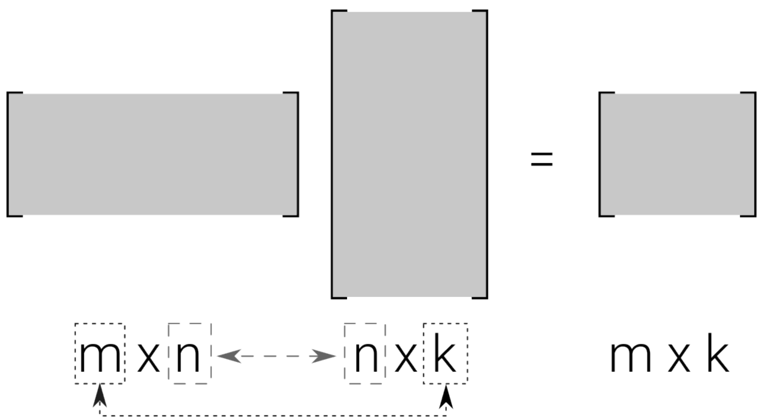
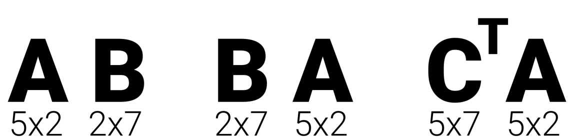


Figura 6.1. Visualización de la regla de cuándo pueden multiplicarse dos matrices. Las dimensiones "de dentro" son las enes y las "de fuera" la m y la k.

Tomemos las siguientes matrices, ¿se podrán hacer estas multiplicaciones?



La primera multiplicación (**AB**) es válida porque las dimensiones de dentro son iguales (2 en ambos casos). El orden (o tamaño) de la matriz resultante será 5×7. El segundo ejemplo permite ver que la multiplicación de matrices no tiene la propiedad conmutativa. Las dimensiones de dentro (7 y 5) no son iguales, así que la multiplicación no se puede hacer. El tercer ejemplo es un caso interesante. Podrías estar tentado de decir que está operación no puede hacerse; sin embargo, cuando calculas la traspuesta de **C**, las filas y las columnas cambian de sitio, y también las dimensiones que quedan dentro y fuera. El resultado en este caso es que las dimensiones "de dentro" después de trasponer sí son iguales (5 en ambos casos). Así que esta multiplicación es válida.

Hemos conseguido algo emocionante: Ahora tienes los conocimientos para entender la notación del producto escalar (o "producto punto") y del producto exterior. En concreto, ahora puedes ver por qué la posición del vector traspuesto ($\mathbf{v}^{\mathrm{T}}\mathbf{w}$ o $\mathbf{v}\mathbf{w}^{\mathrm{T}}$) determina si la multiplicación es un producto escalar o un producto exterior (figura 6.2).

$$\mathbf{A} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{1} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 4 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \quad \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \mathbf{1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Figura 6.2. Si tenemos dos vectores columna contiguos, trasponer el primero o el segundo es la diferencia entre un producto escalar y un producto exterior.

Problemas de práctica Indica si pueden multiplicarse las siguientes matrices y, en caso afirmativo, cuál será el tamaño de su resultado.

$$\mathbf{A} \in \mathbb{R}^{2\times4}, \quad \mathbf{B} \in \mathbb{R}^{3\times4}, \quad \mathbf{C} \in \mathbb{R}^{4\times4}$$

a) $\mathbf{AB}$ **b)** $\mathbf{AC}$ **c)** $\mathbf{AB}^{\mathrm{T}}$ **d)** $\mathbf{BCA}^{\mathrm{T}}$ **e)** $\mathbf{BB}^{\mathrm{T}}\mathbf{A}$

Respuestas

a) no **b)** 2×4 **c)** 2×3 **d)** 3×2 **e)** no

Programación Es una pena, pero la multiplicación de matrices es totalmente diferente en MATLAB y Python. Presta mucha atención a las sutiles, pero importantes, diferencias (`@` vs. `*`).

Bloque de código 6.1. Python

```
M1 = np.random.randn(4,3)
M2 = np.random.randn(3,5)
C = M1 @ M2
```

Bloque de código 6.2. MATLAB

```
M1 = randn(4,3);
M2 = randn(3,5);
C = M1 * M2
```

Ha llegado la hora de aprender a multiplicar matrices. Hay cuatro formas de ver y resolver las multiplicaciones de matrices. Los cuatro métodos dan el mismo resultado, pero son visiones diferentes de lo que significa multiplicar matrices. Resulta útil entender todos estos enfoques, ya que permiten ver los cálculos matriciales en diferentes contextos y para resolver problemas distintos. Es una pena que muchos libros de texto de álgebra lineal solo enseñen el método del producto escalar (lo que yo llamo el "enfoque por elementos").

1) Enfoque por elementos Cada uno de los elementos $c_{i,j}$ de $\mathbf{AB} = \mathbf{C}$ será el resultado del producto escalar entre la fila i de $\mathbf{A}$ y la columna j de $\mathbf{B}$. La siguiente ecuación muestra cómo se crearía el elemento superior izquierdo de la matriz producto.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1a + 2c & \\ & \end{bmatrix} \tag{6.1}$$

Por comodidad, a todas las matrices de los siguientes ejemplos las llamaré $\mathbf{AB} = \mathbf{C}$.

Se ve claramente que el elemento $c_{i,j}$ es el producto escalar de la fila $\mathbf{a}_i$ por la columna $\mathbf{b}_j$.

Te doy otro ejemplo (fíjate en que cada uno de los elementos de la matriz resultado es el producto escalar entre la fila de la matriz izquierda y la columna de la matriz derecha correspondientes):

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3{\cdot}5 + 4{\cdot}3 & 3{\cdot}1 + 4{\cdot}1 \\ -1{\cdot}5 + 2{\cdot}3 & -1{\cdot}1 + 2{\cdot}1 \\ 0{\cdot}5 + 4{\cdot}3 & 0{\cdot}1 + 4{\cdot}1 \end{bmatrix} = \begin{bmatrix} 27 & 7 \\ 1 & 1 \\ 12 & 4 \end{bmatrix}$$

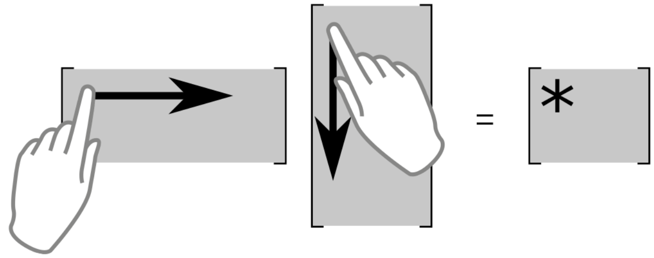


Figura 6.3. Representación visual del cálculo de los elementos de la matriz producto. Cada elemento es el resultado de multiplicar escalarmente dos vectores: la fila de la matriz de la izquierda (de izquierda a derecha) por la columna de la matriz derecha (de arriba a abajo).

Para acordarte de este método, puedes utilizar un gesto con la mano: extiende los dedos índice de las dos manos y mueve al mismo tiempo la mano izquierda hacia la derecha (siguiendo la fila de la matriz izquierda) al mismo tiempo que mueves la mano derecha hacia ti (bajando por la columna de la matriz) (figura 6.3).

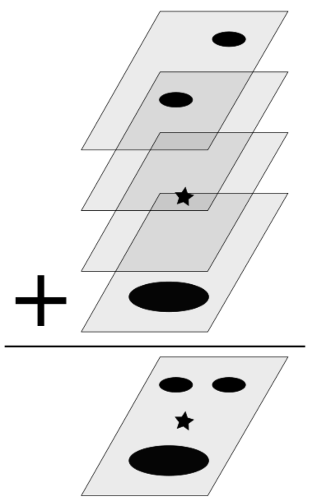


Figura 6.4. Representación de cómo crear una matriz capa a capa. Cada una de las capas tiene el mismo tamaño que el producto, pero solo proporciona información parcial.

A continuación, te doy otra visualización de la multiplicación de matrices siguiendo este enfoque por elementos en la que podrás observar tres características importantes de la multiplicación de matrices.

$$\begin{bmatrix} — & a_1 & — \\ — & a_2 & — \\ & \vdots & \\ — & a_n & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_1{\cdot}b_1 & a_1{\cdot}b_2 & \cdots & a_1{\cdot}b_n \\ a_2{\cdot}b_1 & a_2{\cdot}b_2 & & a_2{\cdot}b_n \\ \vdots & & \ddots & \vdots \\ a_n{\cdot}b_1 & a_n{\cdot}b_2 & \cdots & a_n{\cdot}b_n \end{bmatrix}$$

- La diagonal de la matriz producto **C** contiene los productos escalares de las filas y las columnas situadas en la misma posición (la fila i de **A** y la columna i de **B**).
- El triángulo inferior de **C** contiene los productos escalares de las filas *posteriores* de **A** y las columnas *anteriores* de **B** (filas i de **A** y columnas j de **B**, donde $i > j$).
- El triángulo superior de **C** contiene los productos escalares de las filas *anteriores* de **A** y las columnas *posteriores* de **B** (filas i de **A** y columnas j de **B**, donde $i < j$).

El primero de estos puntos es importante para entender las matrices de covarianzas. El segundo y tercer puntos son importantes para entender varias descomposiciones matriciales, especialmente la factorización QR y el problema de valores propios generalizado.

2) Enfoque por capas A diferencia de lo que ocurría en el enfoque por elementos, en el que cada elemento se calcula de forma independiente, en el enfoque por capas la multiplicación de matrices se conceptualiza como una serie de capas, o "sábanas" agrupadas. De forma práctica, consiste en calcular los productos exteriores de las columnas de $\mathbf{A}$ por las filas de $\mathbf{B}$, para después sumar esos productos exteriores.

Recuerda que el producto exterior es una matriz. Cada producto exterior es del mismo tamaño que $\mathbf{C}$ y puede pensarse en él como si fuera una capa. Para que puedas ver una analogía, imagínate que creas una imagen colocando hojas de papel transparentes una encima de la otra, de forma que cada hoja contenga una parte diferente de la imagen (Figura 6.4).

A continuación te doy un ejemplo utilizando la misma matriz que en la sección anterior. Asegúrate de que entiendes cómo se forman las dos columnas de los productos exteriores a partir de la columna $\mathbf{a}_i$ y la fila $\mathbf{b}_j$. Puedes utilizar también casi el mismo gesto de la mano que para el enfoque por elementos (figura 6.3), pero con los movimientos de las manos izquierda y derecha invertidos.

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 15 & 3 \\ -5 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 12 & 4 \\ 6 & 2 \\ 12 & 4 \end{bmatrix} = \begin{bmatrix} 27 & 7 \\ 1 & 1 \\ 12 & 4 \end{bmatrix}$$

Cada una de estas capas es una matriz de rango 1. Hablaremos del rango con más detalle en otro capítulo, pero por ahora puedes imaginarte una matriz de rango 1 como aquella que solo contiene información correspondiente a una columna, mientras que todas las demás son versiones a escala de ella (ella misma multiplicada por un escalar).

Fíjate en que en cada una de las matrices que forman las capas, las columnas forman un conjunto dependiente (y lo mismo ocurre con las filas). Sin embargo, cuando se suman estas matrices singulares, la matriz resultante (es decir, el producto que buscamos) tiene columnas linealmente *in*dependientes.

El enfoque por capas de la multiplicación de matrices está estrechamente relacionado con el teorema espectral de las matrices, que afirma que cualquier matriz puede representarse como una suma de matrices de rango 1. Es como si cada matriz de rango 1 fuera un solo color y la matriz el arco iris. Esta elegante idea es importante y la base de la descomposición en valores singulares que aprenderás en el capítulo 16.

3) Enfoque por columnas Desde el punto de vista de las columnas, todas las matrices (tanto las que se multiplican como su producto) se consideran como conjuntos de vectores columna. De esta manera, la matriz producto se crea columna a columna.

La primera columna de la matriz producto es una combinación lineal de todas las columnas de la matriz izquierda, utilizando para la combinación lineal los coeficientes definidos por los elementos de la primera columna de la matriz de la derecha. La segunda columna de la matriz producto es, de nuevo, una combinación lineal de todas las columnas de la matriz izquierda, excepto que ahora los coeficientes son los de la segunda columna de la matriz de la derecha. Y así sucesivamente para las n columnas de la matriz derecha. Empecemos con un ejemplo sencillo:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a\begin{bmatrix} 1 \\ 3 \end{bmatrix} + c\begin{bmatrix} 2 \\ 4 \end{bmatrix} & b\begin{bmatrix} 1 \\ 3 \end{bmatrix} + d\begin{bmatrix} 2 \\ 4 \end{bmatrix} \end{bmatrix} \tag{6.2}$$

Vayamos paso a paso por la ecuación 6.2. La primera columna de la matriz producto es la suma de todas las columnas de la matriz **A**. Pero no las sumamos simplemente: multiplicamos cada una de las columnas de **A** por un coeficiente de ponderación que obtenemos de la primera columna de la matriz **B** (porque estamos calculando la primera columna de la matriz producto). La segunda columna de la matriz **C** se crea sumando de nuevo todas las columnas de la matriz **A**, pero ahora los coeficientes son los elementos de la columna 2 de la matriz **B**. En la ecuación 6.2 solo vemos dos columnas, pero si tuviéramos más repetiríamos este procedimiento para todas las columnas de la matriz **B**.

Vayamos ahora al mismo ejemplo numérico que ya has visto con los dos enfoques anteriores:

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} & 1 \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 27 & 7 \\ 1 & 1 \\ 12 & 4 \end{bmatrix}$$

El enfoque por columnas de la multiplicación de matrices es útil en estadística, cuando las columnas de la matriz izquierda contienen un conjunto de *regresores* (un modelo simplificado de los datos), mientras que la matriz derecha contiene *coeficientes*. Los coeficientes cuantifican la importancia de cada regresor. Cuando se ajusta estadísticamente un modelo, el objetivo es encontrar los valores de los coeficientes que al multiplicarlos por los regresores se acercan más a los datos reales. ¡Te cuento más en el capítulo 14!

4) Enfoque por filas ¡Lo has adivinado! Es el mismo concepto que el enfoque por columnas, pero en este caso la matriz producto se construye fila a fila, y en el proceso se van haciendo combinaciones lineales de las filas. Por tanto: cada una de las *filas* de la matriz producto es la combinación lineal de todas las *filas* de la matriz de la derecha, mientras que los coeficientes corresponden a los elementos de cada una de las *filas* de la matriz izquierda. Empecemos con un ejemplo sencillo:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 \begin{bmatrix} a & b \end{bmatrix} + 2 \begin{bmatrix} c & d \end{bmatrix} \\ \\ 3 \begin{bmatrix} a & b \end{bmatrix} + 4 \begin{bmatrix} c & d \end{bmatrix} \end{bmatrix} \tag{6.3}$$

La fila superior de la matriz producto se crea sumando las dos filas de la matriz derecha. Pero, de nuevo, antes tenemos que "ponderar" cada una de las filas multiplicándola por el elemento correspondiente de la fila superior de la matriz izquierda. La segunda fila se construye de forma similar. Y, por supuesto, el proceso seguiría tantas veces como filas tenga la matriz de la izquierda.

No repetiré la otra multiplicación de ejemplo que he ido dando en las páginas anteriores, te dejo que lo hagas tú con lápiz y papel. (Pista: tiene que darte el mismo resultado que antes).

El enfoque por filas es útil, por ejemplo, en el análisis de componentes principales, donde las filas de la matriz de la derecha contienen datos (observaciones en las filas y características en las columnas), mientras que las filas de la matriz izquierda contienen los coeficientes que permiten combinar las características. La suma ponderada de los datos que se hace después permite obtener las puntuaciones de los componentes principales. En la figura 6.5 se resumen visualmente los diferentes enfoques.

Problemas de práctica Multiplica los siguientes pares de matrices cuatro veces, utilizando cada uno de los cuatro enfoques. Ten en cuenta que en todos los casos has de obtener el mismo resultado.

a) $\begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ 0 & 1 \\ 4 & 1 \end{bmatrix}$ **b)** $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

Respuestas

a) $\begin{bmatrix} 24 & 15 \\ 4 & 5 \end{bmatrix}$ **b)** $\begin{bmatrix} 1a+2d+3g & 1b+2e+3h & 1c+2f+3i \\ 4a+5d+6g & 4b+5e+6h & 4c+5f+6i \\ 7a+8d+9g & 7b+8e+9h & 7c+8f+9i \end{bmatrix}$

Problemas de práctica Resuelve las siguientes multiplicaciones de matrices. Usa el enfoque que más te cueste (¡porque es el que necesitas practicar!).

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

Respuestas

a) $\begin{bmatrix} 22 & -3 & 6 \\ 16 & 1 & 1 \end{bmatrix}$ **b)** $\begin{bmatrix} 3 & 3 \\ 3 & 4 \end{bmatrix}$

6.2 Multiplicación y ecuaciones

En la secundaria, ya aprendiste que puedes sumar o multiplicar términos a una ecuación, siempre que lo hagas en ambos miembros. Por ejemplo, puedo multiplicar ambos lados de esta ecuación por

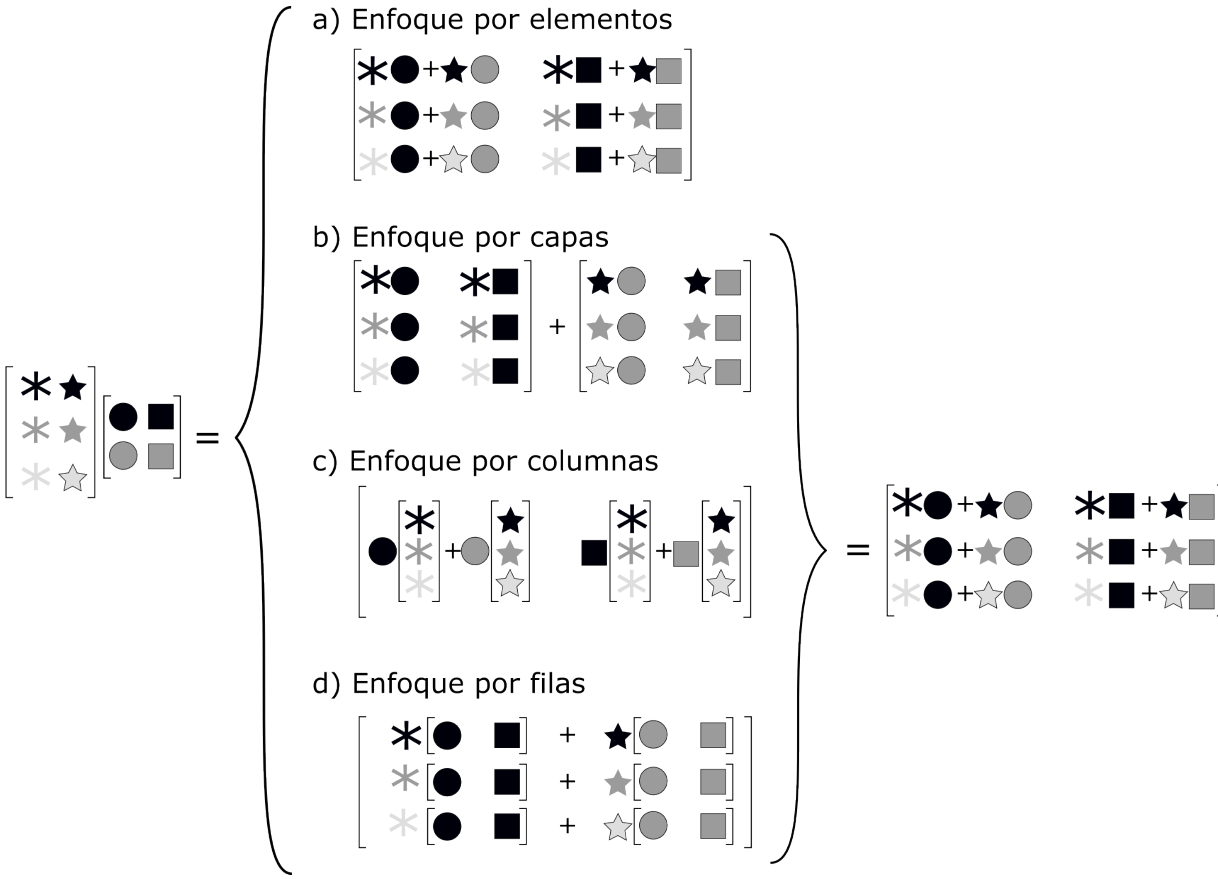


Figura 6.5. Representación visual de los cuatro enfoques de la multiplicación de matrices

7:

$$4 + x = 5(y + 3)$$

$$7(4 + x) = 7(5(y + 3))$$

$$7(4 + x) = (5(y + 3))7$$

Fíjate en que las dos últimas ecuaciones son iguales: Como los números escalares tienen la propiedad conmutativa, el 7 puede ir a la izquierda o la derecha de los paréntesis.

Sin embargo, la multiplicación de matrices no tiene la propiedad conmutativa, así que si quieres multiplicar ambos miembros de la ecuación por una matriz, asegúrate de colocarla en el mismo lado

Los tamaños de las matrices no están indicados, así que puedes asumir que son los necesarias para poder hacer las operaciones.

en los dos miembros. Por ejemplo, esto es correcto:

$$\mathbf{B} = \lambda(\mathbf{C} + \mathbf{D})$$

$$\mathbf{AB} = \lambda\mathbf{A}(\mathbf{C} + \mathbf{D})$$

$$\mathbf{AB} = \mathbf{A}(\mathbf{C} + \mathbf{D})\lambda$$

La λ se puede mover porque es un escalar, pero $\mathbf{A}$ ha de premultiplicar en ambos miembros (o bien posmultiplicar en ambos, lo importante es que introduzcas la matriz en la misma posición). A diferencia de lo anterior, esta multiplicación por una matriz en ambos miembros es **INCORRECTA**.

$$\mathbf{B} = \lambda(\mathbf{C} + \mathbf{D})$$

$$\mathbf{AB} = \lambda(\mathbf{C} + \mathbf{D})\mathbf{A}$$

En otras palabras, si premultiplicas en un miembro de una ecuación, también tienes que premultiplicar en el otro. Lo mismo pasaría con la posmultiplicación. Si en la ecuación hay matrices rectangulares, es posible que ni siquiera puedas colocar la matriz en cualquier posición porque la operación no podría hacerse.

Programación Puedes confirmar fácilmente programando que la multiplicación de matrices no es transitiva. Compara las matrices C1 y C2.

Bloque de código 6.3. Python

```
A = np.random.randn(2,2)
B = np.random.randn(2,2)
C1 = A@B
C2 = B@A
```

Bloque de código 6.4. MATLAB

```
A = randn(2,2);
B = randn(2,2);
C1 = A*B;
C2 = B*A;
```

6.3 Multiplicación por una matriz diagonal

Si una de las matrices es diagonal y la otra es una matriz densa, se dan las siguientes dos propiedades:

- Premultiplicar por una matriz diagonal multiplica las *filas* de la matriz de la derecha por los elementos de la diagonal.
- Posmultiplicar por una matriz diagonal multiplica las *columnas* de la matriz izquierda por los elementos de la diagonal.

Veamos dos ejemplos de matrices 3×3. Fíjate en que los elementos de la diagonal aparecen uno en cada fila si la colocamos antes (o sea, premultiplicamos) o en cada columna si la colocamos después (posmultiplicamos). Fíjate también en que la matriz producto es igual a la matriz densa, pero con las columnas o las filas multiplicadas por cada uno de los elementos de la diagonal de la matriz diagonal.

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1a & 2a & 3a \\ 4b & 5b & 6b \\ 7c & 8c & 9c \end{bmatrix} \tag{6.4}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} = \begin{bmatrix} 1a & 2b & 3c \\ 4a & 5b & 6c \\ 7a & 8b & 9c \end{bmatrix} \tag{6.5}$$

Por supuesto, la manera de multiplicar las matrices es exactamente la misma que la que aprendiste en la sección anterior. Lo único que ocurre ahora es que los ceros de las matrices diagonales nos permiten simplificar un poco.

Vale la pena recordar esta propiedad de la multiplicación por una matriz diagonal, porque en varios capítulos posteriores te mostraré algunas aplicaciones, por ejemplo en sistemas de ecuaciones (capítulo 10), diagonalización (capítulo 15) y descomposición en valores singulares (capítulo 16).

Orden de las matrices y modulación de filas y columnas

Si colocas la diagonal **ANTES**, afecta a las **FILAS**

Si la colocas **DESPUÉS**, afecta a las **COLUMNAS**

Problemas de práctica Resuelve las siguientes multiplicaciones de matrices. Fíjate en las diferencias entre **a** y **c**, y entre **b** y **d**.

a) $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **b)** $\begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$ **c)** $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ **d)** $\begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix}$

Respuestas

a) $\begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}$ **b)** $\begin{bmatrix} 20 & 15 \\ 18 & 24 \end{bmatrix}$ **c)** $\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}$ **d)** $\begin{bmatrix} 20 & 18 \\ 15 & 24 \end{bmatrix}$

Multiplicación de dos matrices diagonales El producto de dos matrices diagonales es otra matriz diagonal cuyos elementos son los productos de los elementos correspondientes de las diagonales de las dos matrices. ¡Menuda frase más larga me ha salido!, pero el concepto es sencillo. Te doy un ejemplo:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} d & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} ad & 0 & 0 \\ 0 & be & 0 \\ 0 & 0 & cf \end{bmatrix} \tag{6.6}$$

Tómate un momento para multiplicar estas dos matrices de la manera habitual para que veas que la respuesta es correcta. Una vez que lo hayas visto claro, en el futuro ya solo tienes que recordar que multiplicar dos matrices diagonales es fácil; de hecho, en el caso de dos matrices diagonales, la multiplicación de matrices estándar da el mismo resultado que su multiplicación elemento a elemento. Esto será relevante cuando aprendas sobre la descomposición en valores y vectores propios.

6.4 ATAR a la RATA (orden de las operaciones)

Primero te daré un ejemplo del problema que queremos resolver: Haz la siguiente multiplicación de matrices y calcula la traspuesta del resultado:

$$\left(\begin{bmatrix}2 & 4\\1 & 2\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}\right)^{\mathrm{T}} \quad = \quad \textbf{?}$$

Supongo que has obtenido la matriz $\begin{bmatrix}4 & 2\\6 & 3\end{bmatrix}$. Hagámoslo de nuevo, pero esta vez trasponiendo primero cada una de las matrices antes de multiplicarlas:

$$\begin{bmatrix}2 & 4\\1 & 2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}^{\mathrm{T}} \quad = \quad \textbf{?}$$

¿Te ha salido la misma matriz que antes? Si has hecho los cálculos correctamente, te tendría que haber dado $\begin{bmatrix}1 & 3\\2 & 6\end{bmatrix}$, que no es lo mismo que el resultado anterior. Ni siquiera es el mismo resultado, pero traspuesto. De hecho, es una matriz totalmente diferente.

Probemos otra vez, pero ahora, antes de calcular la traspuesta de cada una de las matrices, *las cambiaremos de orden*. Por tanto:

$$\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}2 & 4\\1 & 2\end{bmatrix}^{\mathrm{T}} \quad = \quad \textbf{?}$$

Como ves, si lo haces así obtienes el mismo resultado que para la primera multiplicación: $\begin{bmatrix}4 & 2\\6 & 3\end{bmatrix}$.

N.b.: ATAR a la RATA *no* es una recomendación de cómo tratar a los animales. ¡Que las matemáticas no te aparten del camino de la bondad y el respeto!

Y esto nos lleva al tema principal de esta sección: Cuando aplicas una operación a matrices que se multiplican, lo puedes hacer individualmente *siempre que inviertas el orden de las matrices.*

Es una regla un poco extraña, pero es así, que se le va a hacer. **Te doy un truco mnemotécnico para acordarte: "ATAR a la RATA"**es un truco mnemotécnico que te ayudará a recordar esta importante regla. Fíjate en que si lees ATAR al revés sale

RATA. La frase completa es un palíndromo, que de hecho utilizó Julio Cortázar en uno de sus cuentos.

La regla de ATAR a la RATA: Invertir el orden de las matrices

$$(\mathbf{A}\ldots\mathbf{B})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\ldots\mathbf{A}^{\mathrm{T}} \tag{6.7}$$

En definitiva, cuando se aplica una operación a matrices que se multiplican entre ellas, primero se invierte la posición de las matrices y después se aplica la operación a cada una de ellas. Si tuviéramos cuatro matrices, quedarían así:

$$(\mathbf{ABCD})^{\mathrm{T}} = \mathbf{D}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$$

La regla de ATAR a la RATA también puede aplicarse a otras operaciones, como el cálculo de la matriz inversa. Por ejemplo:

$$(\mathbf{ABC})^{\text{-1}} = \mathbf{C}^{\text{-1}}\mathbf{B}^{\text{-1}}\mathbf{A}^{\text{-1}}$$

No hay duda de que la regla de ATAR a la RATA es extraña y lo contrario de lo que parecería intuitivamente. Pero es importante, así que veamos otro ejemplo, que también sirve para ver por qué esta regla ha de ser así. En las siguientes ecuaciones, arriba multiplico primero las matrices y después calculo la traspuesta de su resultado, mientras que abajo hago primero la traspuesta de las matrices en el orden inverso y después la multiplicación.

$$\begin{aligned}\left(\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}e & f\\ g & h\end{bmatrix}\right)^{\mathrm{T}} &= \begin{bmatrix}ae+bg & af+bh\\ ce+dg & cf+dh\end{bmatrix}^{\mathrm{T}} &&= \begin{bmatrix}ae+bg & ce+dg\\ af+bh & cf+dh\end{bmatrix}\\ &= \begin{bmatrix}e & g\\ f & h\end{bmatrix}\begin{bmatrix}a & c\\ b & d\end{bmatrix} &&= \begin{bmatrix}ae+bg & ce+dg\\ af+bh & cf+dh\end{bmatrix}\end{aligned}$$

Con matrices cuadradas, si ignoras la regla de ATAR a la RATA y no las inviertes, obtendrás un resultado (aunque sea incorrecto). Sin embargo, si las matrices son rectangulares, sería imposible hacer la multiplicación ignorando esta regla. Esto se ilustra en la figura 6.6.

Ten cuidado con la ecuación 6.7: una operación puede ser válida para un producto de matrices, pero no serlo para cada una de las

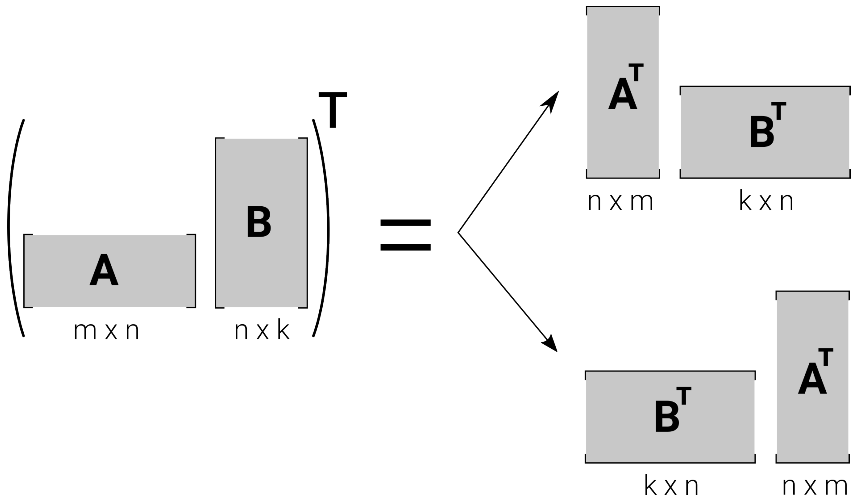


Figura 6.6. Ejemplo de la regla de ATAR a la RATA (inversión) para calcular la traspuesta de una multiplicación de matrices. Fíjate en los tamaños de las matrices: Ignorar la regla de ATAR a la RATA y trasponer sin invertir antes el orden de las matrices te lleva a una expresión inválida (arriba), mientras que la multiplicación sigue siendo válida si cambias el orden de las matrices (abajo).

matrices individualmente. Esto ocurre con frecuencia en estadística y descomposición en valores singulares cuando, por ejemplo, la expresión $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$ podría ser válida, mientras que $\mathbf{X}^{-1}\mathbf{X}^{\text{-T}}$ podría no estar definida. Verás ejemplos de esta situación en capítulos posteriores. Por suerte, todas las matrices pueden trasponerse, así que de momento puedes aplicar siempre la regla de ATAR a la RATA sin preocuparte.

Problemas de práctica Resuelve las siguientes multiplicaciones de matrices. Compáralas con los problemas y respuestas de la página 156:

a) $\begin{bmatrix} 2 & -1 & 2 \\ 4 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 3 & 4 & 0 \\ 0 & 4 & 1 \end{bmatrix}^{\mathrm{T}}$

b) $\begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{\mathrm{T}}$

Respuestas

a) $\begin{bmatrix} 22 & 16 \\ -3 & 1 \\ 6 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 3 & 3 \\ 3 & 4 \end{bmatrix}$

6.5 Multiplicación de una matriz por un vector

Multiplicar una matriz por un vector es lo mismo que multiplicar dos matrices de la forma normal; simplemente tienes que pensar en el vector como una matriz $m \times 1$ o $1 \times n$. La característica importante de la multiplicación entre una matriz y un vector (que quizá sea obvio si te paras a pensarlo, pero que vale la pena mencionarlo explícitamente), es que el resultado siempre es un vector. Esto es importante porque permite relacionar las transformaciones lineales y las matrices: Para aplicar una transformación a un vector, conviertes esa transformación en una matriz y la multiplicas por el vector. Aquí tienes dos ejemplos de multiplicaciones de una matriz por un vector:

$$\mathbf{Ab} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 4\cdot5 + 2\cdot2 \\ 1\cdot5 + 3\cdot2 \end{bmatrix} = \begin{bmatrix} 24 \\ 11 \end{bmatrix}$$

$$\mathbf{b}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} 5 & 2 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4\cdot5 + 1\cdot2 & 5\cdot2 + 2\cdot3 \end{bmatrix} = \begin{bmatrix} 24 & 16 \end{bmatrix}$$

Tres observaciones:

1. $\mathbf{bA}$ no está definido (asumiendo que $\mathbf{b}$ es un vector columna).
2. Si $\mathbf{A}$ es rectangular, entonces uno de los dos, $\mathbf{b}^{\mathrm{T}}\mathbf{A}$ o $\mathbf{Ab}$, no está definido (cuál concretamente dependerá de sus tamaños, pero las dos operaciones no pueden ser válidas a la vez).
3. $\mathbf{Ab} \neq \mathbf{b}^{\mathrm{T}}\mathbf{A}$, incluso si las dos operaciones fueran válidas.

Hay una excepción interesante a esta tercera observación, que es: si la matriz es simétrica, multiplicar el vector por la izquierda da el mismo resultado que multiplicar la traspuesta del vector por la derecha. Bueno, técnicamente los resultados no son *exactamente iguales* porque un resultado es un vector columna, mientras que el otro es un vector fila, por eso he añadido a la igualdad una traspuesta más fuera del paréntesis, pero los componentes de los vectores son idénticos.

Para una matriz simétrica por un vector:

$$\textit{Si } \mathbf{A} = \mathbf{A}^{\mathrm{T}} \textit{ entonces } \mathbf{A}\mathbf{b} = (\mathbf{b}^{\mathrm{T}}\mathbf{A})^{\mathrm{T}} \tag{6.8}$$

Veamos paso a paso una demostración de esta afirmación. En esta demostración trasponemos **Ab** y hacemos un poco de álgebra (¡que incluye aplicar la regla de ATAR a la RATA!) para simplificar y reordenar.

Fíjate en que en esta demostración hemos tenido que trasponer, expandir paréntesis y simplificar. Esta estrategia se usa en muchas demostraciones de álgebra.

$$(\mathbf{A}\mathbf{b})^{\mathrm{T}} = \mathbf{b}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{b}^{\mathrm{T}}\mathbf{A}$$

La demostración funciona porque $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$. Si la matriz no fuera simétrica, $\mathbf{A} \neq \mathbf{A}^{\mathrm{T}}$, en otras palabras, $\mathbf{A}$ y $\mathbf{A}^{\mathrm{T}}$ serían matrices diferentes. Y, por supuesto, $\mathbf{b}$ y $\mathbf{b}^{\mathrm{T}}$ son iguales, excepto por su orientación.

Veamos un ejemplo:

$$\mathbf{A}\mathbf{b} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} d \\ e \end{bmatrix} = \begin{bmatrix} ad + be \\ bd + ce \end{bmatrix}$$

$$\mathbf{b}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} ad + be & bd + ce \end{bmatrix}$$

Fíjate en que, como ya hemos mencionado, los dos resultados no son *vectores idénticos* porque uno es una columna y el otro es una fila. Sin embargo, tienen *componentes idénticos*, solo que con distinta orientación.

Ahora fíjate en qué ocurre cuando la matriz **A** no es simétrica. En las siguientes matrices, asume que $b \neq f$.

$$\mathbf{A}\mathbf{b} = \begin{bmatrix} a & b \\ f & c \end{bmatrix} \begin{bmatrix} d \\ e \end{bmatrix} = \begin{bmatrix} ad + be \\ fd + ce \end{bmatrix}$$

$$\mathbf{b}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} a & b \\ f & c \end{bmatrix} = \begin{bmatrix} ad + fe & bd + ce \end{bmatrix}$$

Problemas de práctica Resuelve las siguientes multiplicaciones de matrices.

a) $\begin{bmatrix} 4 & 3 & 0 \\ 1 & 1 & 9 \\ 0 & 3 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix}$

b) $\begin{bmatrix} 4 & 0 & 2 & 0 & 8 & 0 \\ 0 & 3 & 0 & 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$

Respuestas

a) $\begin{bmatrix} 21 \\ 16 \\ 21 \\ 8 \end{bmatrix}$

b) $\begin{bmatrix} 26 \\ 34 \end{bmatrix}$

6.6 Creación de matrices simétricas

En el capítulo anterior afirmé que aún quedaba esperanza para las matrices no simétricas que aspiran a alcanzar el estado supremo de sus colegas simétricas, las que reciben todas las atenciones y lujos del álgebra lineal. ¿Cómo puede convertirse una matriz no simétrica en una simétrica?

Hay dos métodos: el aditivo y el multiplicativo. El método aditivo no se utiliza mucho en las aplicaciones prácticas (que yo sepa), pero vale la pena aprenderlo. El método aditivo para crear una matriz simétrica a partir de una no simétrica consiste en sumar a una matriz su traspuesta. Este método solo funciona con matrices cuadradas.

$$\mathbf{C} = \frac{1}{2}(\mathbf{A}^{\mathrm{T}} + \mathbf{A}) \tag{6.9}$$

Te ilustraré con un ejemplo por qué se cumple la ecuación 6.9. Fíjate en que los elementos de la diagonal quedan multiplicados por dos, y esa es la razón por la que dividir por 2 es el factor de normalización adecuado. (En las siguientes matrices, asume que

$b \neq d$, $c \neq h$ y $f \neq i$.)

$$\begin{bmatrix} a & b & c \\ d & e & f \\ h & i & j \end{bmatrix} + \begin{bmatrix} a & d & h \\ b & e & i \\ c & f & j \end{bmatrix} = \begin{bmatrix} a+a & b+d & c+h \\ b+d & e+e & f+i \\ c+h & f+i & j+j \end{bmatrix} \tag{6.10}$$

Pero este es solo un ejemplo concreto con una matriz 3×3. ¿Y si solo se cumpliera para esta matriz? ¿Cómo sabemos que este método funcionará en todos los casos? Es importante hacerse esta pregunta, porque hay varias propiedades especiales de las matrices 2×2 o 3×3 que no pueden generalizarse a matrices más grandes.

La demostración de que la ecuación 6.9 siempre generará una matriz simétrica a partir de una no simétrica parte de la definición de las matrices simétricas ($\mathbf{C} = \mathbf{C}^{\mathrm{T}}$). Por tanto, la demostración consiste en trasponer los dos miembros de la ecuación 6.9, hacer un poco de álgebra para simplificar y ver qué ocurre (omito la división por el escalar 2 porque no afecta a la simetría).

$$\mathbf{C} = \mathbf{A}^{\mathrm{T}} + \mathbf{A} \tag{6.11}$$

$$\mathbf{C}^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}} + \mathbf{A})^{\mathrm{T}} \tag{6.12}$$

$$\mathbf{C}^{\mathrm{T}} = \mathbf{A}^{\mathrm{TT}} + \mathbf{A}^{\mathrm{T}} \tag{6.13}$$

$$\mathbf{C}^{\mathrm{T}} = \mathbf{A} + \mathbf{A}^{\mathrm{T}} \tag{6.14}$$

Las matrices se suman, no se multiplican, así que no hay que aplicar la regla de ATAR a la RATA.

Puesto que la suma de matrices tiene la propiedad conmutativa, $\mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{A}$. El miembro derecho de la ecuación 6.14 es igual al miembro derecho de la ecuación 6.11. Y si los dos miembros de la derecha son iguales, entonces los de la izquierda también han de ser necesariamente iguales. De esta forma queda demostrado que $\mathbf{C} = \mathbf{C}^{\mathrm{T}}$, que es la definición de una matriz simétrica. Esta demostración no depende del tamaño de la matriz, lo que demuestra que nuestro ejemplo anterior no era una casualidad.

Problemas de práctica Crea matrices simétricas a partir de las siguientes matrices utilizando el método aditivo:

a) $\begin{bmatrix} 1 & 11 & 1 \\ -5 & -2 & 1 \\ -1 & -5 & 2 \end{bmatrix}$

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

Respuestas

a) $\begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -2 \\ 0 & -2 & 2 \end{bmatrix}$

b) $\begin{bmatrix} -3 & -\frac{1}{2} & -3 & 2 \\ -\frac{1}{2} & -3 & -5 & -4 \\ -3 & -5 & -7 & \frac{9}{2} \\ 2 & -4 & \frac{9}{2} & 4 \end{bmatrix}$

Método multiplicativo En este método se multiplica una matriz por su traspuesta. De hecho, se trata de la matriz $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ que aprendiste en el capítulo anterior.

Afirmé en el capítulo anterior que $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ es siempre una matriz simétrica (y, por tanto, también cuadrada), incluso si $\mathbf{A}$ no es simétrica, ¡e incluso si $\mathbf{A}$ no es cuadrada! Ahora que sabes multiplicar matrices y conoces la regla de ATAR a la RATA, puedes demostrar estas dos afirmaciones importantes.

Primero vamos a demostrar que $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ es una matriz cuadrada. Pongamos que la matriz $\mathbf{A}$ es $m \times n$ (supongamos que $m \neq n$). Lo primero, fíjate en que $\mathbf{AA}$ no es una multiplicación válida. Ahora piensa en que $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ tiene dimensiones $(n \times m) \times (m \times n)$. Las dimensiones "de dentro" (m) coinciden y la matriz resultante será $n \times n$. ¡Pues ya lo tenemos! Acabamos de demostrar que $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ es una matriz cuadrada.

Demostremos ahora que $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ es simétrica. Para demostrarlo seguiremos la misma estrategia que aplicamos para el método aditivo: trasponer $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, hacer un poco de álgebra y ver qué pasa.

$$(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}}\mathbf{A}^{\mathrm{TT}} = \mathbf{A}^{\mathrm{T}}\mathbf{A} \tag{6.15}$$

Hemos traspuesto la matriz, después hemos aplicado la regla de ATAR a la RATA (y la propiedad de que la traspuesta de la traspuesta es la matriz original) para llegar así a la misma matriz que teníamos al principio. Es decir, hemos visto que la traspuesta es igual a la matriz original, que es exactamente la definición de una matriz simétrica.

¿Son estas dos características válidas solo para $\mathbf{A}^{\mathrm{T}}\mathbf{A}$? ¿Qué pasaría con $\mathbf{A}\mathbf{A}^{\mathrm{T}}$, es también cuadrada y simétrica? La respuesta es "Sí", pero te pido que tomes un pedazo de papel y te lo demuestres a ti mismo.

Reflexión

En las aplicaciones prácticas, especialmente en el análisis de datos, $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ aparece con la misma frecuencia que $\mathbf{A}\mathbf{A}^{\mathrm{T}}$. La forma correcta depende de cómo se almacenan los datos (p.~ej., observaciones $\times$ características o características $\times$ observaciones), que normalmente depende de cómo se prefiera programar o del formato del software. En textos escritos, yo prefiero $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ porque tener la $^{\mathrm{T}}$ en el medio queda más simétrico y por tanto más estético.

Problemas de práctica Crea dos matrices simétricas a partir de cada una de las siguientes utilizando el método multiplicativo ($\mathbf{A}^{\mathrm{T}}\mathbf{A}$ y $\mathbf{A}\mathbf{A}^{\mathrm{T}}$):

a) $\begin{bmatrix} 3 & 0 & 3 \\ 0 & 7 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 6 \\ 6 & 1 \end{bmatrix}$

Respuestas

a) $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} 9 & 0 & 9 \\ 0 & 49 & 0 \\ 9 & 0 & 9 \end{bmatrix}$, $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 18 & 0 \\ 0 & 49 \end{bmatrix}$

b) $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} 37 & 12 \\ 12 & 37 \end{bmatrix}$, $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 37 & 12 \\ 12 & 37 \end{bmatrix}$

6.7 Multiplicación de dos matrices simétricas

Si multiplicamos dos matrices simétricas, ¿será la matriz producto también simétrica? Probemos con un ejemplo para averiguarlo:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} d & e \\ e & f \end{bmatrix} = \begin{bmatrix} ad+be & ae+bf \\ bd+ce & be+cf \end{bmatrix} \tag{6.16}$$

¿Es el resultado simétrico? De entrada, parece que no. Sin embargo, esta ecuación revela una característica interesante sobre la simetría del resultado de multiplicar dos matrices simétricas 2×2: Si $a=c$ y $d=f$ (en otras palabras, si los elementos de cada una de las diagonales son iguales), el producto de dos matrices simétricas es también una matriz simétrica, ¡Observa y sorpréndete!

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} d & e \\ e & d \end{bmatrix} = \begin{bmatrix} ad+be & ae+bd \\ bd+ae & be+ad \end{bmatrix} \tag{6.17}$$

¿Es esta una regla general que funciona con matrices simétricas de cualquier tamaño? Hagamos lo mismo con matrices 3×3 utilizando la idea de que los elementos de la diagonal sean iguales.

$$\begin{bmatrix} a & b & c \\ b & a & d \\ c & d & a \end{bmatrix} \begin{bmatrix} e & f & g \\ f & e & h \\ g & h & e \end{bmatrix} = \begin{bmatrix} ae+bf+cg & af+be+ch & ag+bh+ce \\ be+af+dg & bf+ae+dh & bg+ah+de \\ ce+df+ag & cf+de+ah & cg+dh+ae \end{bmatrix}$$

Un vistazo rápido revela la falta de simetría. Por ejemplo, compara el elemento de la posición *2,1* con el de la posición *1,2*. No voy a escribir el producto de dos matrices simétricas 4×4 (puedes probar a hacerlo si quieres), pero puedes creerme: La matriz resultante no será simétrica.

Lo que queremos aprender con esto es que, en general, *el producto de dos matrices simétricas no es una matriz simétrica.* Hay excepciones a esta regla, como el caso 2×2 con todos los elementos de la diagonal iguales, o si una de las matrices es la matriz identidad o la matriz nula.

¿Te sorprende este resultado? Vuelve a la parte final del "enfoque por elementos" para multiplicar matrices (página 152) donde te cuento cómo se forman el triángulo superior y el triángulo inferior de la matriz producto a partir de las filas anteriores y posteriores de la matriz de la izquierda. En los triángulos superior e inferior de la matriz producto participan partes diferentes de las matrices que se multiplican.

También puedes ver que la matriz producto no es simétrica intentando demostrar que *sí es* simétrica (no se podrá demostrar, así que estarás demostrando lo contrario en lo que se llama una "demostración por contradicción"). Supongamos que **A** y **B** son matrices simétricas:

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{BA} \neq \mathbf{AB} \tag{6.18}$$

Para dejarlo claro, la ecuación 6.18 es totalmente válida, así como cada una de las multiplicaciones. Sin embargo, la multiplicación de matrices no tiene la propiedad conmutativa, por eso hemos escrito el signo $\neq$ al final de la ecuación. Así que no podemos asumir que $\mathbf{AB} = (\mathbf{AB})^{\mathrm{T}}$ y podemos decir que la multiplicación de dos matrices simétricas no es, en general, una matriz simétrica.

Reflexión

Esta información podría parecerte un dato académico sin ningún interés real, pero lleva a una de las mayores limitaciones del análisis de componentes principales. Y es también una de las ventajas más importantes del problema de valores propios generalizado (que se utilizan en multitud de métodos de aprendizaje automático, sobre todo de los clasificadores lineales y del análisis discriminante).

6.8 Multiplicación elemento a elemento (producto de Hadamard)

Antes de haber visto la multiplicación de matrices (en este capítulo o donde sea), si alguien te hubiera pedido adivinar cómo se multiplican dos matrices, seguramente hubieras pensado en el producto de Hadamard.

El producto de Hadamard consiste en ir multiplicando cada elemento de una de las matrices por el elemento correspondiente de la otra. Ya viste el producto de Hadamard para vectores en el capítulo 3. El concepto y la notación son los mismos para las matrices. Por tanto:

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} \tag{6.19}$$

La definición formal del producto de Hadamard es:

Multiplicación elemento a elemento (de Hadamard)

$$c_{i,j} = a_{i,j} \times b_{i,j} \tag{6.20}$$

Puede que ya hayas adivinado que el producto de Hadamard solo es válido si las dos matrices tienen las mismas dimensiones $m \times n$; además, la matriz resultante también tendrá tamaño $m \times n$.

Lo verás muy fácilmente con un ejemplo:

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 6 & 3 \end{bmatrix} \odot \begin{bmatrix} 3 & 8 & 5 \\ 4 & 1 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 8 & 10 \\ -4 & 6 & -15 \end{bmatrix} \tag{6.21}$$

Puesto que el producto de Hadamard se hace elemento a elemento, tiene la propiedad conmutativa igual que ocurre con los números sueltos (los números escalares). O sea:

$$\mathbf{A} \odot \mathbf{B} = \mathbf{B} \odot \mathbf{A}$$

También existe la división elemento a elemento, que sigue el mismo principio, pero dividiendo. Esta operación solo puede hacerse si la matriz por la que se divide no tiene ningún elemento igual a cero.

$\varnothing$ es el signo de la división elemento a elemento.

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 6 & 3 \end{bmatrix} \varnothing \begin{bmatrix} 3 & 8 & 5 \\ 4 & 1 & -5 \end{bmatrix} = \begin{bmatrix} 0/3 & 1/8 & 2/5 \\ -1/4 & 6/1 & -3/5 \end{bmatrix}$$

Se puede debatir si el producto y la división de Hadamard son realmente operaciones de *matrices*, puesto que también se podría considerar que no son más que multiplicaciones o divisiones entre escalares hechas *en bloque* utilizando una notación compacta. De hecho, en computación la multiplicación de matrices elemento a elemento normalmente se utiliza para programar más eficientemente (por ejemplo, para evitar el uso de bucles for), no tanto porque se quiera utilizar alguna propiedad matemática especial del producto de Hadamard. Dicho esto, el producto de Hadamard sí tiene aplicaciones en el álgebra lineal. Por ejemplo, es fundamental en uno de los algoritmos que se emplean para calcular la matriz inversa.

Programación También en este caso, la multiplicación de matrices se programa de forma totalmente diferente en MATLAB y Python. En MATLAB, `A*B` indica la multiplicación estándar de matrices, mientras que `A.*B` corresponde al producto de Hadamard (fíjate en el *punto seguido del asterisco*). En Python, `A@B` es la multiplicación de matrices estándar y `A*B` el producto de Hadamard.

Bloque de código 6.5. Python

```
M1 = np.random.randn(4,3)
M2 = np.random.randn(4,3)
C = M1 * M2
```

Bloque de código 6.6. MATLAB

```
M1 = randn(4,3);
M2 = randn(4,3);
C = M1 .* M2
```

Problemas de práctica Calcula el producto de Hadamard de los siguientes pares de matrices.

a) $\begin{bmatrix} -4 & -3 & -9 \\ -5 & -2 & 3 \\ 0 & -7 & 7 \end{bmatrix}, \begin{bmatrix} -1 & 4 & -3 \\ -3 & -4 & 2 \\ 3 & 2 & 5 \end{bmatrix}$

b) $\begin{bmatrix} -4 & -5 & -16 \\ -4 & -1 & 2 \\ 7 & 1 & 5 \end{bmatrix}, \begin{bmatrix} -3 & 6 & -10 & 0 \\ -4 & 9 & -6 & 4 \\ 1 & 2 & -6 & 0 \end{bmatrix}$

Respuestas

a) $\begin{bmatrix} 4 & -12 & 27 \\ 15 & 8 & 6 \\ 0 & -14 & 35 \end{bmatrix}$

b) ¡No puede hacerse!

6.9 Producto escalar de Frobenius

El producto escalar de Frobenius, también llamado producto interno, o producto interior, de Frobenius, es una operación que, como su nombre indica, produce un escalar (un único número) a partir de dos matrices del mismo tamaño ($m \times n$).

Para calcular el producto escalar de Frobenius, primero hay que *vectorizar* las dos matrices y después calcular su producto escalar igual que lo harías con vectores normales.

Vectorizar una matriz es concatenar todas sus columnas para crear un único vector columna. Es una función que transforma una matriz perteneciente a $\mathbb{R}^{m \times n}$ en un vector de $\mathbb{R}^{mn}$.

Vectorizar una matriz

$$\mathbf{v}_n = a_{i,j} \text{ , tal que}$$

$$\mathbf{v} = [a_{1,1}, \ldots, a_{m,1}, a_{1,2}, \ldots, a_{m,2}, \ldots, a_{m,n}] \tag{6.22}$$

Aquí tienes un ejemplo de una matriz y el resultado de vectorizarla:

$$vec\left(\begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}\right) = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

Igual que ocurre con muchas otras operaciones que has aprendido hasta ahora, en la vectorización hay algunas cosas arbitrarias: ¿Por qué no concatenamos las filas en lugar de las columnas? También podría hacerse así, pero respetar una convención común facilita la comprensión.

Programación Fíjate en que, por defecto, Python vectoriza por filas. Puedes cambiarlo indicando que se use la convención de Fortran.

Bloque de código 6.7. Python

```
A = np.array([ [1,2,3],[4,5,6] ])
A.flatten(order='F')
```

Bloque de código 6.8. MATLAB

```
A = [ 1,2,3; 4,5,6 ];
A(:)
```

Una vez aclarado esto, volvamos a lo nuestro, que es calcular el producto escalar de Frobenius. Te doy un ejemplo:

$$\langle \mathbf{A}, \mathbf{B} \rangle_F = \left\langle \begin{bmatrix} 1 & 5 & 0 \\ -4 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 4 & -1 & 3 \\ 2 & 6 & 7 \end{bmatrix} \right\rangle_F = 5$$

Fíjate en la notación del producto escalar de Frobenius: $\langle \mathbf{A}, \mathbf{B} \rangle_F$

Una forma curiosa, pero útil, de calcular el producto escalar de Frobenius entre dos matrices $\mathbf{A}$ y $\mathbf{B}$ es utilizar la traza de $\mathbf{A}^{\mathrm{T}}\mathbf{B}$. Si lo hacemos así, el producto escalar de Frobenius puede también escribirse de la siguiente manera:

Producto escalar de Frobenius como la traza de $\mathbf{A}^{\mathrm{T}}\mathbf{B}$

$$\langle \mathbf{A}, \mathbf{B} \rangle_F = tr(\mathbf{A}^{\mathrm{T}}\mathbf{B}) \tag{6.23}$$

En esta ecuación he omitido los ordenes de las matrices, pero puedes ver que para que la operación sea válida ambas matrices han de tener el mismo tamaño $m \times n$ porque la traza solo está definida para matrices cuadradas.

El motivo de que la ecuación 6.23 sea válida puede verse con algunos ejemplos que tendrás la oportunidad de hacer en los ejercicios.

El producto escalar de Frobenius tiene varios usos en el procesamiento de señales y el aprendizaje automático, por ejemplo para medir la "distancia" o la similitud entre dos matrices.

El producto interno de Frobenius de una matriz por sí misma es la suma de todos los elementos al cuadrado y se llama la *norma de Frobenius al cuadrado* o la *norma euclidiana al cuadrado* de la matriz. Te explico más sobre esto en la siguiente sección.

Programación El siguiente código muestra el truco de trasponer y calcular la traza para hacer el producto escalar de Frobenius.

Bloque de código 6.9. Python

```python
A = np.random.randn(4,3)
B = np.random.randn(4,3)
f = np.trace(A.T@B)
```

Bloque de código 6.10. MATLAB

```matlab
A = randn(4,3);
B = randn(4,3);
f = trace(A'*B);
```

Problemas de práctica Calcula el producto escalar de Frobenius de los siguientes pares de matrices:

a) $\begin{bmatrix} 4 & 2 \\ 3 & 2 \\ -5 & -1 \end{bmatrix}, \begin{bmatrix} 7 & -2 \\ -7 & -8 \\ -1 & 8 \end{bmatrix}$

b) $\begin{bmatrix} 6 & 1 & 2 \\ 3 & 3 & -2 \end{bmatrix}, \begin{bmatrix} 4 & -11 & 1 \\ 6 & 1 & -2 \end{bmatrix}$

Respuestas

a) -16

b) 40

6.10 Normas de las matrices

En la sección 3.9, aprendiste que la raíz cuadrada del producto escalar de un vector por sí mismo es igual al módulo/la magnitud del vector (su longitud), que también se llama la "norma" del vector.

Aunque también en este caso, para las matrices se complica el concepto de norma, como todo cuando pasas de vectores a matrices. ¡Y parte del lío de las normas de las matrices es que hay muchas! Todas tienen cosas en común, por ejemplo, todas son un único número que, de alguna manera, está asociado con la "magnitud" de la matriz, pero cada tipo de norma es una interpretación diferente del significado de "magnitud". En esta sección aprenderás algunas normas habituales de matrices y continuaremos hablando de ellas en el capítulo 16, cuando tratemos la descomposición en valores singulares.

Empecemos por la norma de Frobenius, ya que la tienes fresca de la sección anterior. La siguiente ecuación es otra forma diferente de expresar la norma de Frobenius.

Norma de Frobenius de una matriz

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{i,j})^2} \qquad (6.24)$$

Ahora ya sabes calcular la norma de Frobenius de tres maneras: 1) aplicando directamente la ecuación 6.24, 2) vectorizando la matriz y calculando el producto escalar por ella misma y 3) haciendo $tr(\mathbf{A}^{\mathrm{T}}\mathbf{A})$.

Si pensamos en un espacio matricial como un espacio euclídeo, la norma de Frobenius de la resta de dos matrices proporciona una medida de distancia euclídea entre ellas. El miembro derecho de la siguiente fórmula te debería resultar conocido del cálculo de la distancia euclídea entre dos puntos.

Distancia euclídea entre dos matrices

$$\|\mathbf{A}-\mathbf{B}\|_F = \sqrt{\sum_{i,j}(a_{i,j}-b_{i,j})^2} \qquad (6.25)$$

Por supuesto, esta fórmula solo es válida para dos matrices que tengan el mismo tamaño.

A la norma de Frobenius también se la llama norma $\boldsymbol{\ell}2$ (el símbolo $\boldsymbol{\ell}$ no es más que una ele elegante, o simplemente una ele minúscula en cursiva). También hay una norma $\boldsymbol{\ell}1$ de matrices. Para calcular la norma $\boldsymbol{\ell}1$, sumas los valores absolutos de los elementos de cada una de las columnas y después te quedas con la suma de la columna que dé un resultado más alto.

Hay muchas otras normas de matrices que tienen diferentes fórmulas. En cada aplicación se usa la norma más adecuada para satisfacer los criterios que tengamos o para minimizar las características de la matriz. Como no te quiero agobiar con una lista exhaustiva, te daré una fórmula general de la *norma p de una matriz*; puedes ver que en el caso particular de $p = 2$, la fórmula es igual a la de la norma de Frobenius.

$$\|\mathbf{A}\|_p = \left(\sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|^p \right)^{1/p} \tag{6.26}$$

Desigualdad de Cauchy-Schwarz En el capítulo 3 aprendiste qué era la desigualdad de Cauchy-Schwarz (que afirma que "el módulo del producto escalar de dos vectores no puede ser mayor que el producto de sus normas"). Existe una desigualdad similar aplicable a la norma de Frobenius del producto de una matriz por un vector:

$$\|\mathbf{Av}\| \leq \|\mathbf{A}\|_F \|\mathbf{v}\| \tag{6.27}$$

Esta desigualdad puede demostrarse combinando 1) el enfoque por filas de la multiplicación, 2) la desigualdad de Cauchy-Schwarz del producto escalar de vectores que vimos en el capítulo 3 y 3) la linealidad de la norma de Frobenius de las filas al cuadrado (es decir, que la norma de Frobenius al cuadrado de una matriz es igual a la suma de las normas de Frobenius al cuadrado de sus filas; esto sale de la suma de la ecuación 6.24).

Empecemos por reescribir la norma al cuadrado de la multiplicación de una matriz por un vector como la suma de las normas al cuadrado de los vectores que resultan de hacer el producto escalar entre cada una de las filas de $\mathbf{A}$ por el vector $\mathbf{v}$ (m es el número de filas y $\mathbf{a}_i$ es la fila i de $\mathbf{A}$).

$$\|\mathbf{Av}\|^2 = \|\mathbf{a}_1\mathbf{v}\|^2 + ... + \|\mathbf{a}_m\mathbf{v}\|^2 = \sum_{i=1}^{m} \|\mathbf{a}_i\mathbf{v}\|^2 \tag{6.28}$$

Por la desigualdad de Cauchy-Schwarz para el producto escalar podemos escribir lo siguiente:

$$\sum_{i=1}^{m} \|\mathbf{a}_i\mathbf{v}\|^2 \leq \sum_{i=1}^{m} \|\mathbf{a}_i\|^2 \|\mathbf{v}\|^2 \tag{6.29}$$

Por último, hacemos el sumatorio de las normas de las filas que hay en el miembro izquierdo para volver a obtener la norma al cuadrado de la matriz. Y así llegamos a nuestra conclusión inicial de la ecuación 6.27: La norma del producto de una matriz por un vector es menor o igual que el producto de la norma de Frobenius de la matriz por la norma del vector. (También puedes elevar al cuadrado los términos de 6.27, o hacer la raíz cuadrada en los dos miembros de la ecuación 6.29, puesto que hacerlo no modifica la desigualdad).

Programación Para obtener las diferentes normas de matrices, es necesario introducir entradas diferentes en las funciones `norm`. El siguiente código corresponde a la norma de Frobenius.

Bloque de código 6.11. Python

```
A = np.random.randn(4,3)
np.linalg.norm(A, 'fro')
```

Bloque de código 6.12. MATLAB

```
A = randn(4,3);
norm(A, 'fro')
```

Problemas de práctica Calcula la distancia euclídea entre los siguientes pares de matrices (nota: compara con los ejercicios de la página 176).

a) $\begin{bmatrix} 4 & 2 \\ 3 & 2 \\ -5 & -1 \end{bmatrix}, \begin{bmatrix} 7 & -2 \\ -7 & -8 \\ -1 & 8 \end{bmatrix}$

b) $\begin{bmatrix} 6 & 1 & 2 \\ 3 & 3 & -2 \end{bmatrix}, \begin{bmatrix} 4 & -11 & 1 \\ 6 & 1 & -2 \end{bmatrix}$

Respuestas

a) $\sqrt{322} \approx 17.94$

b) $\sqrt{162} \approx 12.73$

6.11 Índice de asimetría de una matriz

Dentro de toda matriz cuadrada no simétrica se esconde una matriz simétrica perfecta esperando a ser descubierta. De hecho, **toda matriz cuadrada puede expresarse como la suma de una matriz simétrica y una antisimétrica**. Otro sinónimo de matriz antisimétrica, que quizá hayas oído, es matriz hemisimétrica.

Esta es una afirmación en toda regla, y de entrada no parece muy obvia. Déjame que empiece con un ejemplo sencillo para ilustrar el concepto:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 4 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 7 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \tag{6.30}$$

A simple vista verás que la matriz de la izquierda no es simétrica, pero que puede expresarse como la suma de una matriz simétrica y otra antisimétrica. Puedes ver la matriz antisimétrica como si fuera un "valor residual" que añadimos a la matriz simétrica para conseguir llegar a la original. En este ejemplo, la parte no antisimétrica era "pequeña", en el sentido de que era una matriz dispersa y que había pocos elementos diferentes de cero en comparación con la cantidad de números que tiene la matriz simétrica. Un poco más adelante, en esta misma sección, cuantificaremos esta "pequeñez" calculando el *índice de asimetría de una matriz*, pero primero quiero enseñarte cómo descomponer una matriz no simétrica en la suma de una simétrica y otra no simétrica.

Recordarás que el método aditivo para crear una matriz simétrica consistía en sumar una matriz cuadrada a su traspuesta y dividir entre 2. Ahora haremos lo opuesto, crearemos una matriz antisimétrica *restando* en vez de sumar. En la siguiente ecuación, $\mathbf{K}$ es una matriz antisimétrica. Vamos a asumir que $\mathbf{A}$ es cuadrada, pero no tiene por qué ser ni simétrica ni antisimétrica.

$$\mathbf{K} = (\mathbf{A} - \mathbf{A}^{\mathrm{T}})/2 \tag{6.31}$$

Te doy un ejemplo:

$$\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \right) /2 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

Algunos comentarios: 1) Por supuesto, la diagonal debe ser cero; 2) la división entre 2 no es necesaria para la antisimetría, pero mantiene los valores numéricos de la matriz original; 3) si la matriz $\mathbf{A}$ ya es simétrica, entonces la matriz $\mathbf{K}$ será la matriz nula; 4) si la matriz $\mathbf{A}$ ya es antisimétrica, entonces $\mathbf{K} = \mathbf{A}$.

Con esto en mente, ya podemos descomponer cualquier matriz cuadrada en la suma de una matriz simétrica y una antisimétrica.

$$\mathbf{A}_k = (\mathbf{A} - \mathbf{A}^{\mathrm{T}})/2 \tag{6.32}$$

$$\mathbf{A}_s = \mathbf{A} - \mathbf{A}_k \tag{6.33}$$

$\mathbf{A}_k$ es la "capa" antisimétrica de $\mathbf{A}$ y $\mathbf{A}_s$ su "capa" simétrica.

Es evidente que $\mathbf{A}_k$ es antisimétrica, así como también es evidente que $\mathbf{A}_k + \mathbf{A}_s = \mathbf{A}$. Pero no es obvio que $\mathbf{A}_s$ sea simétrica, así que demostrémoslo. En palabras: lo que tenemos que demostrar es que esa matriz menos su traspuesta es igual a la matriz nula. Utilizando notación matemática, sería:

$$(\mathbf{A} - \mathbf{A}_k) - (\mathbf{A} - \mathbf{A}_k)^{\mathrm{T}} \tag{6.34}$$

$$= \mathbf{A} - \mathbf{A}_k - \mathbf{A}^{\mathrm{T}} + \mathbf{A}_k^{\mathrm{T}} \tag{6.35}$$

$$= (\mathbf{A} - \mathbf{A}^{\mathrm{T}}) - (\mathbf{A}_k - \mathbf{A}_k^{\mathrm{T}}) \tag{6.36}$$

$$= \mathbf{0} \tag{6.37}$$

La ecuación 6.36 es una resta de dos términos: el primero, $(\mathbf{A} - \mathbf{A}^{\mathrm{T}})$, no es más que $2\mathbf{A}_k$ por definición, mientras que el segundo término de hecho también es $2\mathbf{A}_k$, así que cancela al primero. Por tanto, a $\mathbf{A}_s$ le hemos restado su traspuesta y hemos llegado a la matriz nula, lo que demuestra que $\mathbf{A}_s$ es una matriz simétrica.

Volvamos al objetivo principal de esta sección: deducir un índice escalar que cuantifique el grado de asimetría de una matriz. Conceptualmente, se trata de cuantificar cómo de "pequeña" es la matriz antisimétrica residual en relación con la matriz original. Una matriz simétrica pura ha de tener un índice 0, mientras que una matriz antisimétrica pura debe tener un índice de 1.

A este índice se le llama el *índice de asimetría de una matriz*. Para indicarlo, usaré la notación $\mathbf{A}_\sigma$ para el índice de asimetría de $\mathbf{A}$.

Este índice se calcula cómo la razón entre la norma de la "capa" asimétrica y la de la matriz original:

$$\mathbf{A}_\sigma = \|\mathbf{A}_k\|_F^2 \;/\; \|\mathbf{A}\|_F^2 \tag{6.38}$$

¿Cómo interpretamos este índice? Empecemos por plantearnos qué ocurre si tenemos una matriz $\mathbf{A}$ totalmente simétrica. En ese caso, $\mathbf{A}_k$ será la matriz nula, que tiene norma cero, lo que significa que $\mathbf{A}_\sigma = 0$. Ahora pensemos en una matriz antisimétrica pura. $\mathbf{A}_k = \mathbf{A}$, y si las matrices son iguales, sus normas también lo serán, lo que significa que $\mathbf{A}_\sigma = 1$.

Estos serían los extremos, pero si el índice de asimetría se encontrara entre 0 y 1, significaría que la matriz no es ni totalmente simétrica ni totalmente antisimétrica, sino algo intermedio. (Por cierto, si lo que quieres medir es la *simetría* en vez de la *asimetría*, puedes usar simplemente $1 - \mathbf{A}_\sigma$).

¿Y por qué usamos la norma de Frobenius al cuadrado y no otra cualquiera? Todas las normas de las matrices tienen las propiedades de ser únicas (es decir, dos matrices iguales siempre tendrán la misma norma) y de ser cero para la matriz nula. Pero, además, la norma de Frobenius tiene la siguiente interesante propiedad relacionada con la suma de matrices:

$$\|\mathbf{A}_s + \mathbf{A}_k\|_F^2 = \|\mathbf{A}_s\|_F^2 + \|\mathbf{A}_k\|_F^2 + 2\,\langle \mathbf{A}_s, \mathbf{A}_k \rangle_F \tag{6.39}$$

Reflexión

La matrices simétricas y antisimétricas tienen interesantes propiedades, tanto por sí mismas como combinadas. Por ejemplo, las matrices simétricas y antisimétricas son ortogonales, lo que significa que el producto escalar de Frobenius entre ellas es cero ($trace(\mathbf{S}^{\mathrm{T}}\mathbf{K}) = 0$. ¡Demuéstralo tú mismo con programación y después piensa por qué es así!). Además, los valores propios del producto $\mathbf{SK}$ son pares de signo opuesto ($\lambda = \pm a$) (¡prueba con esto también!). Esta complementariedad entre los dos tipos de matrices me recuerda a esos dibujos animados donde el supervillano es el hermano gemelo del superhéroe. Dejaré que seas tú quien adivine qué matriz es el héroe y cuál el villano.

6.12 ¿Y la división de matrices?

Tanto hablar de las multiplicaciones de matrices ¡y no hemos dicho nada de la división! Aprendiste la división de matrices elemento a elemento en la sección 6.8, pero en realidad no es una división de *matrices*, sino solo una notación compacta para describir muchas divisiones por escalares. Cuando oyes el término "división de matrices", probablemente estés pensando en algo de este tipo:

$$\frac{\mathbf{A}}{\mathbf{B}}$$

que sería el equivalente de una división de escalares, como podría ser $\frac{2}{3}$. Pero la verdad es que, en el caso de las matrices, no se puede *dividir* una matriz entre otra de esta manera. Sin embargo, sí hay una operación que conceptualmente sería equivalente, basada en la idea de reescribir la división de escalares de esta forma:

$$\frac{2}{3} \;=\; 2\times 3^{-1}$$

La versión utilizando matrices sería $\mathbf{A}\mathbf{B}^{-1}$. La matriz $\mathbf{B}^{-1}$ se llama *matriz inversa* y es tan importante que se merece un capítulo propio (capítulo 12). Por ahora, te doy tres hechos importantes sobre la matriz inversa:

1. La matriz inversa es la que cumple que $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$
2. No todas las matrices tienen una inversa. Una matriz tiene una matriz completa solo si es cuadrada y de rango completo.
3. Las matrices que no tienen una "inversa real" podrían tener una "pseudoinversa". Las inversas por un lado y las pseudoinversas son parecidas a una inversa, pero no exactamente lo mismo.

6.13 Ejercicios

1. Di si las siguientes operaciones son válidas y, en caso de poderse hacer, cuál sería el tamaño (u orden) de la matriz resultante.

$$\mathbf{A} \in \mathbb{R}^{2\times 3}, \quad \mathbf{B} \in \mathbb{R}^{3\times 3}, \quad \mathbf{C} \in \mathbb{R}^{3\times 4}$$

a) $\mathbf{CB}$

b) $\mathbf{C}^{\mathrm{T}}\mathbf{B}$

c) $(\mathbf{CB})^{\mathrm{T}}$

d) $\mathbf{C}^{\mathrm{T}}\mathbf{BC}$

e) $\mathbf{ABCB}$

f) $\mathbf{ABC}$

g) $\mathbf{C}^{\mathrm{T}}\mathbf{BA}^{\mathrm{T}}\mathbf{AC}$

h) $\mathbf{B}^{\mathrm{T}}\mathbf{BCC}^{\mathrm{T}}\mathbf{A}$

i) $\mathbf{AA}^{\mathrm{T}}$

j) $\mathbf{A}^{\mathrm{T}}\mathbf{A}$

k) $\mathbf{BBA}^{\mathrm{T}}\mathbf{ABBCC}$

l) $(\mathbf{CBB}^{\mathrm{T}}\mathbf{CC}^{\mathrm{T}})^{\mathrm{T}}$

m) $(\mathbf{A} + \mathbf{ACC}^{\mathrm{T}}\mathbf{B})^{\mathrm{T}}\mathbf{A}$

n) $\mathbf{C} + \mathbf{CA}^{\mathrm{T}}\mathbf{ABC}$

o) $\mathbf{C} + \mathbf{BA}^{\mathrm{T}}\mathbf{ABC}$

p) $\mathbf{B} + 3\mathbf{B} + \mathbf{A}^{\mathrm{T}}\mathbf{A} - \mathbf{CC}^{\mathrm{T}}$

q) $\mathbf{A} \odot (\mathbf{ABC})$

r) $\mathbf{A} \odot \mathbf{ABC}(\mathbf{BC})^{\mathbf{T}}$

2. Haz las siguientes multiplicaciones de matrices dos veces, utilizando los dos enfoques indicados (#1: por elementos, #2: por capas, #3: por columnas, #4: por filas).

a) #1,2: $\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 5 \\ 2 & 2 \end{bmatrix}$

b) #2,4: $\begin{bmatrix} -3 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

c) #3,4: $\begin{bmatrix} 11 & -5 \\ 9 & -13 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -8 & .5 \end{bmatrix}$

d) #1,4: $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

e) #2,3: $\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 10 & 1 \\ -5 & 4 \end{bmatrix}$

f) #1,3: $\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

g) #2,3: $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

h) #1,2: $\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 1 \\ 3 & 3 & 0 \end{bmatrix} \begin{bmatrix} -2 & -3 & -1 \\ -1 & -9 & 3 \\ 0 & 1 & 5 \end{bmatrix}$

i) #2,3: $\begin{bmatrix} a & 0 & 1 \\ 0 & b & 0 \\ 1 & 0 & c \end{bmatrix} \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

3. Calcula los siguientes productos de una matriz por un vector (siempre que la operación sea válida).

a) $\begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix}$ **b)** $\begin{bmatrix}2\\3\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}$ **c)** $\begin{bmatrix}2 & 1\\2 & 3\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix}$

d) $\begin{bmatrix}2\\3\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}2 & 1\\2 & 3\end{bmatrix}$ **e)** $\begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\1 & 0 & 0\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}$ **f)** $\begin{bmatrix}1 & 0 & 1\\0 & -4 & 0\\1 & 0 & 1\end{bmatrix}\begin{bmatrix}5\\2\\1\end{bmatrix}$

g) $\begin{bmatrix}1\\2\\5\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\0 & -4 & 0\\1 & 0 & 1\end{bmatrix}$ **h)** $\begin{bmatrix}3\\3\\2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}1 & 3 & 2\\6 & 1 & 5\\3 & 5 & 0\end{bmatrix}$

4. Sean las matrices cuadradas $\mathbf{A}$ y $\mathbf{B}$, cuyos elementos son todos distintos de cero. ¿Qué suposiciones sobre la simetría permitirían que se cumplan las siguientes igualdades (nota: podría ser que las operaciones sean imposibles sean cuales sean las suposiciones)? Demuestra o da un ejemplo para cada una de ellas.

a) $\mathbf{AB} = \mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}$ **b)** $\mathbf{AB} = (\mathbf{AB})^{\mathrm{T}}$
c) $\mathbf{AB} = \mathbf{AB}^{\mathrm{T}}$ **d)** $\mathbf{AB} = \mathbf{A}^{\mathrm{T}}\mathbf{B}$
e) $\mathbf{AB} = \mathbf{B}^{\mathrm{T}}\mathbf{A}$ **f)** $\mathbf{AB} = (\mathbf{BA})^{\mathrm{T}}$

5. En la sección 6.7 aprendiste que el producto de dos matrices simétricas en general no es una matriz simétrica. Esto era así para la multiplicación estándar. ¿Es el *producto de Hadamard* de dos matrices simétricas simétrico? Intenta primero razonarlo, después comprueba tu hipótesis con el siguiente par de matrices.

$$\begin{bmatrix}2 & 5 & 7\\5 & 3 & 6\\7 & 6 & 4\end{bmatrix}, \begin{bmatrix}a & d & f\\d & b & e\\f & e & c\end{bmatrix}$$

6. Para los siguientes pares de matrices, primero vectorízalas y calcula el producto escalar de los dos vectores y después calcula el producto interno de Frobenius de la forma $tr(\mathbf{A}^{\mathrm{T}}\mathbf{B})$.

a) $\begin{bmatrix}a & b\\c & d\end{bmatrix}, \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$ **b)** $\begin{bmatrix}0 & 5\\7 & -2\end{bmatrix}, \begin{bmatrix}1 & 0\\13 & 14\end{bmatrix}$

c) $\begin{bmatrix} 4 & -5 & 8 \\ 1 & -1 & 2 \\ -2 & 2 & -4 \end{bmatrix}, \begin{bmatrix} 4 & -5 & 8 \\ 1 & -1 & 2 \\ -2 & 2 & -4 \end{bmatrix}$ **d)** $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ a & b \end{bmatrix}$

e) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ **f)** $\begin{bmatrix} 1 & 1 & 7 \\ 2 & 2 & 6 \\ 3 & 3 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$

7. Haz las multiplicaciones de matrices indicadas.

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 1 \\ 2 & 2 & 3 \end{bmatrix}$$

a) AB **b) AC** **c) BC** **d) CA**
e) CB **f) BCA** **g) ACB** **h) ABC**

6.14 Respuestas

1. **a)** No **b)** Sí: 4×3
c) No **d)** Sí: 4×4
e) No **f)** Sí: 2×4
g) Sí: 4×4 **h)** No
i) Sí: 2×2 **j)** Sí: 3×3
k) No **l)** No
m) Sí: 3×3 **n)** No
o) Sí: 3×4 **p)** Sí: 3×3
q) No **r)** Sí: 2×3

2. **a)** $\begin{bmatrix} 0 & 5 \\ 2 & 17 \end{bmatrix}$ **b)** $\begin{bmatrix} -3 & 4 \\ -2 & 6 \end{bmatrix}$ **c)** $\begin{bmatrix} 73 & 8.5 \\ 131 & 2.5 \end{bmatrix}$

d) $\begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix}$ **e)** $\begin{bmatrix} 10 & 10 \\ -5 & 13 \end{bmatrix}$ **f)** $\begin{bmatrix} 2a & 3a \\ 2b & 3b \end{bmatrix}$

g) $\begin{bmatrix} 2a+4b & 3a+b \\ 0 & 0 \end{bmatrix}$ **h)** $\begin{bmatrix} -2 & 1 & 19 \\ -1 & -8 & 8 \\ -9 & -36 & 6 \end{bmatrix}$ **i)** $\begin{bmatrix} a^2 & ab & ac+1 \\ b & 2b & 3b \\ a & b & 2c \end{bmatrix}$

3. **a)** $\begin{bmatrix} 4 \\ 9 \end{bmatrix}$ **b)** $\begin{bmatrix} 4 & 9 \end{bmatrix}$ **c)** $\begin{bmatrix} 7 \\ 13 \end{bmatrix}$ **d)** $\begin{bmatrix} 10 & 11 \end{bmatrix}$

e) $\begin{bmatrix} b \\ c \\ a \end{bmatrix}$ **f)** $\begin{bmatrix} 6 \\ -8 \\ 6 \end{bmatrix}$ **g)** No válida. **h)** $\begin{bmatrix} 27 & 22 & 21 \end{bmatrix}$

4. **a)** Que ambas sean simétricas. **b)** $\mathbf{A} = \mathbf{B}$, ambas simétricas. o $\mathbf{A} = \mathbf{B}^{\mathrm{T}}$
c) $\mathbf{B} = \mathbf{B}^{\mathrm{T}}$ **d)** $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$
e) No hay ninguna regla general. **f)** Que ambas sean simétricas.

5. Sí, porque la multiplicación se hace elemento a elemento.

6. **a)** $a+2b+3c+4d$ **b)** 63 **c)** 135
d) $a^2+b^2+ca+bd$ **e)** $a^2+b^2+c^2+d^2$ **f)** No definido

7.

a) $\begin{bmatrix} 6 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ **b)** $\begin{bmatrix} 6 & 3 & 9 \\ 0 & 8 & 2 \\ -2 & -2 & -3 \end{bmatrix}$

c) $\begin{bmatrix} 4 & 2 & 6 \\ 0 & 20 & 5 \\ 6 & 6 & 9 \end{bmatrix}$ **d)** $\begin{bmatrix} 6 & 2 & -3 \\ 0 & 8 & -1 \\ 6 & 4 & -3 \end{bmatrix}$

e) $\begin{bmatrix} 4 & 5 & 9 \\ 0 & 20 & 3 \\ 4 & 10 & 9 \end{bmatrix}$ **f)** $\begin{bmatrix} 12 & 4 & -6 \\ 0 & 40 & -5 \\ 18 & 12 & -9 \end{bmatrix}$

g) $\begin{bmatrix} 12 & 15 & 27 \\ 0 & 40 & 6 \\ -4 & -10 & -9 \end{bmatrix}$ **h)** $\begin{bmatrix} 12 & 6 & 18 \\ 0 & 40 & 10 \\ -6 & -6 & -9 \end{bmatrix}$

6.15 Retos de programación

1. Programa una multiplicación de una matriz 2×4 por otra 4×3 utilizando el enfoque por capas en un bucle for. Asegúrate de que obtienes el mismo resultado que cuando haces la multiplicación de matrices utilizando `@` (Python) o `*` (MATLAB).

2. Genera una matriz diagonal 4×4 y una matriz densa 4×4 de números aleatorios. Programa la multiplicación estándar de las dos matrices y su producto de Hadamard. Ya sabes que las matrices resultantes serán diferentes, ¿pero qué ocurre con sus diagonales?

3. Sean $\mathbf{C}_1 = (\mathbf{A}^{\mathrm{T}} + \mathbf{A})/2$ y $\mathbf{C}_2 = \mathbf{A}^{\mathrm{T}}\mathbf{A}$ para una matriz cuadrada no simétrica $\mathbf{A}$. Tanto $\mathbf{C}_1$ como $\mathbf{C}_2$ son simétricas y las dos formadas a partir de la misma matriz; sin embargo, en general $\mathbf{C}_1 \neq \mathbf{C}_2$. Curiosamente, si $\mathbf{A}$ es una matriz diagonal cuyos elementos son todos no negativos, entonces $\mathbf{C}_1 = \mathbf{C}_2^{1/2}$. Muestra esto con programación usando números aleatorios y explica después por qué obtienes este resultado.

4. Exploremos la desigualdad de Cauchy-Schwarz. Genera una matriz aleatoria $\mathbf{A}$ y un vector aleatorio $\mathbf{v}$, después calcula las normas de ambos miembros de la desigualdad 6.27 (página 178) y finalmente muestra que el miembro de la derecha es mayor que el de la izquierda.

6.16 Soluciones de programación

1. Nota: Para ver que los dos resultados son equivalentes, se restan las matrices para obtener la matriz cero (matemáticamente, sería $x = x \Rightarrow x - x = 0$). Sin embargo, como consecuencia de errores de precisión y de redondeo, los resultados podrían ser números muy pequeños, como `1e-16` (10^{-16}). Puedes considerar que los números menores de aproximadamente `1e-15` son, a todos los efectos, cero.

Bloque de código 6.13. Python

```python
A = np.random.randn(2,4)
B = np.random.randn(4,3)
C1 = np.zeros((2,3))
for i in range(4):
    C1 += np.outer(A[:,i],B[i,:])

C1 - A@B # muestra la igualdad
```

Bloque de código 6.14. MATLAB

```matlab
A = randn(2,4);
B = randn(4,3);
C1 = zeros(2,3);
for i=1:4
    C1 = C1 + A(:,i)*B(i,:);
end
C1 - A*B % muestra la igualdad
```

2. Las diagonales de las dos matrices resultantes son iguales.

Bloque de código 6.15. Python

```
D = np.diag(np.arange(1,5))
A = np.random.randn(4,4)
C1 = D*A
C2 = D@A
print(np.diag(C1))
print(np.diag(C2))
```

Bloque de código 6.16. MATLAB

```
D = diag(1:4);
A = randn(4);
C1 = D.*A;
C2 = D*A;
[diag(C1) diag(C2)]
```

3. $\mathbf{C}_1 = \mathbf{C}_2^{1/2}$ parece una propiedad sorprendente y extraña de la multiplicación de matrices, pero es de hecho trivial, porque $\mathbf{A}$ es diagonal. La razón de que se cumpla es la misma de que se cumpla $x = \sqrt{x^2}$ para $x \geq 0$.

Bloque de código 6.17. Python

```
A = np.diag(np.random.rand(3))
C1 = (A.T+A)/2
C2 = A.T@A
C1-np.sqrt(C2)
```

Bloque de código 6.18. MATLAB

```
A = diag(rand(3,1));
C1 = (A'+A)/2;
C2 = A'*A;
C1-sqrt(C2)
```

4. Este reto no es tan reto, la verdad, pero sí una buena excusa para ganar experiencia programando con normas. Mi estrategia para mostrar la desigualdad es ver que el miembro derecho *menos* el izquierdo tiene un valor positivo. Puedes ejecutarlo varias veces probando con diferentes números aleatorios.

Bloque de código 6.19. Python

```python
import numpy as np
m = 5
A = np.random.randn(m,m)
v = np.random.randn(m)
LHS = np.linalg.norm(A@v)
RHS = np.linalg.norm(A,ord='fro') *
      np.linalg.norm(v)
RHS-LHS # siempre es positivo
```

Bloque de código 6.20. MATLAB

```matlab
m = 5;
A = randn(m);
v = randn(m,1);
LHS = norm(A*v);
RHS = norm(A,'fro')*norm(v);
RHS-LHS % siempre es positivo
```