SAH

Introducción a la econometría

Un enfoque moderno

Apendice D

4a. edición

Jeffrey M. Wooldridge
Michigan State University

Traducción
Ma. del Carmen Enriqueta Hano Roa
Érika M. Jasso Hernan D'Borneville
Traductoras profesionales

Revisión técnica
Roberto Palma Pacheco
Centro de Alta Dirección en Economía y Negocios
Universidad Anáhuac

Domingo Rodríguez Benavides
Facultad de Economía
Universidad Nacional Autónoma de México

CENGAGE Learning

Australia • Brasil • Corea • España • Estados Unidos • Japón • México • Reino Unido • Singapur

Apéndice D

Resumen de álgebra matricial

Este apéndice resume los conceptos de álgebra matricial, incluida el álgebra probabilística, necesarias para el estudio de modelos de regresión lineal múltiple mediante las matrices del apéndice E. Nada de este material se utiliza en el texto principal.

D.1 Definiciones básicas

Definición D.1 (Matriz). Una **matriz** es un arreglo rectangular numérico. En otras palabras y en términos más precisos, una matriz $m \times n$ tiene m filas y n columnas. El entero positivo m se llama *dimensión de la fila o renglón*, y n es la *dimensión de la columna.*

Se usan letras mayúsculas en bold para denotar matrices. En términos generales se puede escribir una matriz $m \times n$ como

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix},$$

donde a_{ij} representa el elemento en el fila i-ésima y en la columna j-ésima. Por ejemplo, a_{25} representa el número en la fila segunda y la quinta columna de **A**. Un ejemplo específico de una matriz 2×3 es

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 7 \\ -4 & 5 & 0 \end{bmatrix}, \qquad \boxed{\textbf{D.1}}$$

donde $a_{13} = 7$. La abreviatura $\mathbf{A} = [a_{ij}]$ suele usarse para definir las operaciones de matrices.

Definición D.2 (Matriz cuadrada). Una **matriz cuadrada** tiene el mismo número de filas que de columnas. La dimensión de una matriz cuadrada es su número de filas y columnas.

Definición D.3 (Vectores)

i) Una matriz $1 \times m$ se denomina **vector fila** (de dimensión m) y se puede escribir como $\mathbf{x} \equiv (x_1, x_2, \ldots, x_m)$.

ii) Una matriz $n \times 1$ se llama **vector columna** y se puede escribir como

$$\mathbf{y} \equiv \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Definición D.4 (Matriz diagonal). Una matriz cuadrada $\mathbf{A}$ es una **matriz diagonal** si todos sus elementos fuera de la diagonal son cero, es decir, $a_{ij} = 0$ para toda $i \neq j$. La matriz diagonal se puede escribir como:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & a_{nn} \end{bmatrix}.$$

Definición D.5 (Matriz identidad y matriz cero)

i) La **matriz identidad** $n \times n$, denotada como $\mathbf{I}$, o algunas veces $\mathbf{I}_n$ para enfatizar su dimensión, es la matriz diagonal con la unidad (uno) en cada posición diagonal y cero en cualquier otro lado:

$$\mathbf{I} \equiv \mathbf{I}_n \equiv \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}.$$

ii) La **matriz cero** $m \times n$ denotada como $\mathbf{0}$, es la matriz $m \times n$ con cero para todas las entradas. No necesariamente tiene que ser una matriz cuadrada.

D.2 Operaciones matriciales

Suma matricial

Dos matrices $\mathbf{A}$ y $\mathbf{B}$, cada una con dimensión $m \times n$, se pueden sumar elemento por elemento: $\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}]$. De manera más precisa,

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \vdots & & & \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \ldots & a_{mn} + b_{mn} \end{bmatrix}.$$

Por ejemplo,

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

Las matrices de dimensiones diferentes no pueden sumarse.

Multiplicación escalar

Dado cualquier número real γ (llamado escalar), la **multiplicación escalar** se define como $\gamma\mathbf{A} \equiv [\gamma a_{ij}]$, o

$$\gamma\mathbf{A} = \begin{bmatrix} \gamma a_{11} & \gamma a_{12} & \dots & \gamma a_{1n} \\ \gamma a_{21} & \gamma a_{22} & \dots & \gamma a_{2n} \\ \vdots & & & \\ \gamma a_{m1} & \gamma a_{m2} & \dots & \gamma a_{mn} \end{bmatrix}.$$

Por ejemplo, si $\gamma = 2$ y $\mathbf{A}$ es la matriz en la ecuación (D.1), entonces

$$\gamma\mathbf{A} = \begin{bmatrix} 4 & -2 & 14 \\ -8 & 10 & 0 \end{bmatrix}.$$

Multiplicación matricial

Para multiplicar la matriz $\mathbf{A}$ por la matriz $\mathbf{B}$ y formar el producto $\mathbf{AB}$, la dimensión *columna* de $\mathbf{A}$ debe ser igual a la dimensión *fila* de $\mathbf{B}$. Por tanto, sea $\mathbf{A}$ una matriz $m \times n$ y sea $\mathbf{B}$ una matriz $n \times p$. Entonces, la **multiplicación matricial** se define como

$$\mathbf{AB} = \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right].$$

En otras palabras, el elemento $(i,j)^{\text{ésimo}}$ de una nueva matriz $\mathbf{AB}$ se obtiene al multiplicar cada elemento en la i-ésima fila de $\mathbf{A}$ por el elemento correspondiente en la j-ésima columna de $\mathbf{B}$ y después sumar esos productos n. Un esquema puede ayudar a aclarar este proceso:

$$i\text{-ésima fila} \rightarrow \overset{\mathbf{A}}{\begin{bmatrix} \\ a_{i1}\,a_{i2}\,a_{i3}\dots a_{in} \\ \\ \end{bmatrix}} \overset{\mathbf{B}}{\underset{\substack{\uparrow \\ \text{columna} \\ j\text{-ésima}}}{\begin{bmatrix} b_{1j} \\ b_{2j} \\ b_{3j} \\ \vdots \\ b_{nj} \end{bmatrix}}} = \overset{\mathbf{AB}}{\underset{\substack{\uparrow \\ \text{elemento} \\ (i,j)\text{-ésimo}}}{\begin{bmatrix} \\ \sum_{k=1}^{n} a_{ik}b_{kj} \\ \\ \end{bmatrix}}},$$

donde, por la definición del operador de suma del apéndice A,

$$\sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$

Por ejemplo,

$$\begin{bmatrix} 2 & -1 & 0 \\ -4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 6 & 0 \\ -1 & 2 & 0 & 1 \\ 3 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 12 & -1 \\ -1 & -2 & -24 & 1 \end{bmatrix}.$$

También se puede multiplicar una matriz y un vector. Si $\mathbf{A}$ es una matriz $n \times m$ matrix y $\mathbf{y}$ es un vector $m \times 1$, entonces $\mathbf{Ay}$ es un vector $n \times 1$. Si $\mathbf{x}$ es un vector $1 \times n$, entonces $\mathbf{xA}$ es un vector $1 \times m$.

La suma matricial, la multiplicación escalar y la multiplicación matricial pueden combinarse de varias formas, y estas operaciones satisfacen varias reglas que son familiares debido a las operaciones numéricas básicas. En la siguiente lista de propiedades, $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ son matrices con dimensiones apropiadas para aplicar cada operación, y α y β son números reales. La mayoría de esas propiedades son fáciles de ilustrar a partir de las definiciones.

Propiedades de la multiplicación matricial. 1) $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$; 2) $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$; 3) $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$; 4) $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B}$; 5) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$; 6) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$; 7) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$; 8) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$; 9) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$; 10) $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$; 11) $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$; 12) $\mathbf{A} - \mathbf{A} = \mathbf{0}$; 13) $\mathbf{A0} = \mathbf{0A} = \mathbf{0}$ y 14) $\mathbf{AB} \neq \mathbf{BA}$, incluso cuando ambos productos estén definidos.

La última propiedad merece un comentario adicional. Si $\mathbf{A}$ es $n \times m$ y $\mathbf{B}$ es $m \times p$, entonces $\mathbf{AB}$ está definida, pero $\mathbf{BA}$ está definida sólo si $n = p$ (la dimensión fila de $\mathbf{A}$ es igual a la dimensión columna de $\mathbf{B}$). Si $\mathbf{A}$ es $m \times n$ y $\mathbf{B}$ es $n \times m$, entonces $\mathbf{AB}$ y $\mathbf{BA}$ están definidas, pero generalmente no son iguales; de hecho, tienen diferentes dimensiones, a menos que $\mathbf{A}$ y $\mathbf{B}$ no sean matrices cuadradas. Aun cuando si $\mathbf{A}$ y $\mathbf{B}$ sean cuadradas $\mathbf{AB} \neq \mathbf{BA}$, salvo en circunstancias especiales.

Transposición

Definición D.6 (Transposición). Sea $\mathbf{A} = [a_{ij}]$ una matriz $m \times n$. La **transpuesta** de A, denotada como $\mathbf{A}'$ (llamada **A** *prima*), es la matriz $n \times m$ obtenida al intercambiar las filas y las columnas de $\mathbf{A}$. Se puede escribir esto como $\mathbf{A}' \equiv [a_{ji}]$.

Por ejemplo,

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 7 \\ -4 & 5 & 0 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} 2 & -4 \\ -1 & 5 \\ 7 & 0 \end{bmatrix}.$$

Propiedades de la transposición. 1) $(\mathbf{A}')' = \mathbf{A}$; 2) $(\alpha\mathbf{A})' = \alpha\mathbf{A}'$ para todo escalar α; 3) $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$; 4) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$, donde $\mathbf{A}$ es $m \times n$ y $\mathbf{B}$ es $n \times k$; 5) $\mathbf{x}'\mathbf{x} = \sum_{i=1}^{n} x_i^2$, donde $\mathbf{x}$ es un vector $n \times 1$; y 6) Si $\mathbf{A}$ es una matriz $n \times k$ con *filas* dadas por los vectores $1 \times k$, $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, así que se puede escribir

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix},$$

entonces $\mathbf{A}' = (\mathbf{a}_1' \, \mathbf{a}_2' \ldots \mathbf{a}_n')$.

Definición D.7 (Matriz simétrica). Una matriz cuadrada $\mathbf{A}$ es una **matriz simétrica** si, y sólo si, $\mathbf{A}' = \mathbf{A}$.

Si $\mathbf{X}$ toda matriz $n \times k$ entonces $\mathbf{X}'\mathbf{X}$ está siempre definida y es una matriz simétrica, como puede verse al aplicar la primera y cuarta propiedades transpuestas (vea el problema D.3).

Multiplicación parcial particionada

Sea $\mathbf{A}$ una matriz $n \times k$ con filas dadas por los vectores $1 \times k$ $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, y sea $\mathbf{B}$ una matriz $n \times m$ con filas dadas por los vectores $1 \times m$ $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}.$$

Entonces,

$$\mathbf{A}'\mathbf{B} = \sum_{i=1}^{n} \mathbf{a}_i'\mathbf{b}_i,$$

donde para cada i, $\mathbf{a}_i'\mathbf{b}_i$ es una matriz $k \times m$. Por tanto, $\mathbf{A}'\mathbf{B}$ puede escribirse como la suma de n matrices, cada una de las cuales es $k \times m$. Como caso especial, se tiene

$$\mathbf{A}'\mathbf{A} = \sum_{i=1}^{n} \mathbf{a}_i'\mathbf{a}_i,$$

donde $\mathbf{a}_i'\mathbf{a}_i$ es una matriz $k \times k$ para toda i.

Traza

La traza de una matriz es una operación muy simple definida sólo para matrices *cuadradas*.

Definición D.8 (Traza). Para toda matriz $\mathbf{A}$ de $n \times n$, la **traza de una matriz** $\mathbf{A}$, denotada como $\mathrm{tr}(\mathbf{A})$, es la suma de sus elementos diagonales. En términos matemáticos

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Propiedades de la traza. 1) $\mathrm{tr}(\mathbf{I}_n) = n$; 2) $\mathrm{tr}(\mathbf{A}') = \mathrm{tr}(\mathbf{A})$; 3) $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$; 4) $\mathrm{tr}(\alpha\mathbf{A}) = \alpha\mathrm{tr}(\mathbf{A})$, para toda escalar α; y 5) $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, donde $\mathbf{A}$ es $m \times n$ y $\mathbf{B}$ es $n \times m$.

Inversa

La noción de matriz inversa es muy importante para las matrices cuadradas.

Definición D.9 (Inversa). Una matriz $\mathbf{A}$ de $n \times n$ tiene una **inversa**, denotada como $\mathbf{A}^{-1}$, si $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$ y $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$. En este caso se dice que $\mathbf{A}$ es *invertible* o *no singular*. De lo contrario, se dice que es *no invertible* o *singular*.

Propiedades de la inversa. 1) Si existe una inversa, es única; 2) $(\alpha\mathbf{A})^{-1} = (1/\alpha)\mathbf{A}^{-1}$, si $\alpha \neq 0$ y $\mathbf{A}$ es invertible; 3) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, si tanto $\mathbf{A}$ como $\mathbf{B}$ son invertibles $n \times n$ y 4) $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$.

Aquí no interesa la mecánica de calcular la inversa de una matriz. Cualquier libro de álgebra matricial contiene ejemplos detallados de tales cálculos.

D.3 Independencia lineal y rango de una matriz

Para un conjunto de vectores que tienen la misma dimensión, es importante saber si un vector se puede expresar como una combinación lineal de los vectores restantes.

Definición D.10 (Independencia lineal). Sea $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ un conjunto de $n \times 1$ vectores. Éstos son **vectores linealmente independientes** si, y sólo si,

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_r\mathbf{x}_r = \mathbf{0} \tag{D.2}$$

implica que $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$. Si (D.2) aplica para un conjunto de escalares no todas iguales a cero, entonces $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ es *linealmente dependiente*.

La afirmación de que $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ es linealmente dependiente es equivalente a decir que al menos un vector en este grupo se puede escribir como una combinación lineal de los demás.

Definición D.11 (Rango)

i) Sea $\mathbf{A}$ una matriz $n \times m$. El **rango de una matriz A**, denotado como rango(**A**), es el máximo número de columnas linealmente independientes de **A**.

ii) Si **A** es $n \times m$ y rango(**A**) $= m$, entonces **A** tiene un *rango pleno por columnas.*

Si **A** es $n \times m$, su rango puede ser a lo sumo m. Una matriz tiene un rango pleno por columnas, si sus columnas forman un conjunto linealmente independiente. Por ejemplo, la matriz 3×2.

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 0 & 0 \end{bmatrix}$$

se puede tener a lo sumo un rango dos. De hecho, su rango es sólo uno porque la segunda columna es tres veces la primera.

Propiedades del rango. 1) rango($\mathbf{A}'$) = rango(**A**); 2) Si **A** es $n \times k$, entonces rango(**A**) $\leq \min(n,k)$ y 3) Si **A** es $k \times k$ y rango(**A**) $= k$, entonces **A** es invertible.

D.4 Formas cuadráticas y matrices definidas positivas

Definición D.12 (Forma cuadrática). Sea **A** una matriz $n \times n$ *simétrica*. La **forma cuadrática** asociada con la matriz **A** es la función de valor real definida para todo vector $n \times 1$ de **x**:

$$f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{j>1} a_{ij}x_ix_j.$$

Definición D.13 (Definida positiva y semidefinida positiva)

i) Se dice que una matriz **A** simétrica es **definida positiva (d.p.)** si

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \text{ para todo vector } n \times 1 \text{ de } \mathbf{x} \text{ excepto } \mathbf{x} = \mathbf{0}.$$

ii) Una matriz simétrica **A** es **semidefinida positiva (p.s.d.)** si

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \text{ para todo vector } n \times 1.$$

Si una matriz es definida positiva o semidefinida positiva, automáticamente se asume que es simétrica.

Propiedades de las matrices definidas positivas y semidefinidas positivas. 1) Una matriz definida positiva tiene elementos diagonales que son estrictamente positivos, mientas que una matriz semidefinida positiva tiene elementos diagonales no negativos; 2) si **A** es d.p., entonces $\mathbf{A}^{-1}$ existe y es su d.p.; 3) Si **X** es $n \times k$, entonces $\mathbf{X}'\mathbf{X}$ y $\mathbf{X}\mathbf{X}'$ son s.d.p.; y 4) Si **X** es $n \times k$ y rango(**X**) $= k$, entonces $\mathbf{X}'\mathbf{X}$ es d.p. (y por tanto, no singular).

D.5 Matrices idempotentes

Definición D.14 (matrices idempotentes). Sea **A** una matriz simétrica $n \times n$. Entonces, se dice que **A** es una **matriz idempotente** si y sólo si, $\mathbf{AA} = \mathbf{A}$.

Por ejemplo,

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

es una matriz idempotente, como lo comprueba la multiplicación directa.

Propiedades de las matrices idempotentes. Sea **A** una matriz $n \times n$ idempotente. 1) rango(**A**) = tr(**A**) y 2) **A** es semidefinida positiva.

Se pueden construir matrices idempotentes de manera muy general. Sea **X** una matriz $n \times k$ de rango(**X**) $= k$. Se define

$$\mathbf{P} \equiv \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$\mathbf{M} \equiv \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{I}_n - \mathbf{P}.$$

Entonces **P** y **M** son matrices simétricas, idempotentes, con rango(**P**) $= k$ y rango(**M**) $= n - k$. Los rangos se obtienen más fácilmente mediante la propiedad 1: tr(**P**) $=$ tr$[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}]$ (de la propiedad 5 de la traza) $=$ tr$(\mathbf{I}_k) = k$ (por la propiedad 1 de la traza). Se desprende que tr(**M**) $=$ tr$(\mathbf{I}_n) -$ tr(**P**) $= n - k$.

D.6 Diferenciación de formas lineales y cuadráticas

Para un vector **a** de $n \times 1$ considere la función lineal definida por

$$f(\mathbf{x}) = \mathbf{a}'\mathbf{x},$$

para todos los vectores $n \times 1$ de **x**. La derivada de f respecto a **x** es el vector de $1 \times n$ de las derivadas parciales, que simplemente es

$$\partial f(\mathbf{x})/\partial \mathbf{x} = \mathbf{a}'.$$

Para una matriz **A** simétrica de $n \times n$ se define la forma cuadrática

$$g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}.$$

Entonces,

$$\partial g(\mathbf{x})/\partial \mathbf{x} = 2\mathbf{x}'\mathbf{A},$$

que es un vector $1 \times n$.

D.7 Momentos y distribuciones de vectores aleatorios

Con el fin de derivar el valor esperado y la varianza de los estimadores MCO mediante matrices, es necesario definir el valor esperado y la varianza de un **vector aleatorio**. Como su nombre sugiere, un vector aleatorio simplemente es un vector de variables aleatorias. También es necesario definir la distribución normal multivariada. Estos conceptos son simplemente extensiones de los que se abordaron en el apéndice B.

Valor esperado

Definición D.15 (Valor esperado)

i) Si **y** es un vector $n \times 1$ aleatorio, el **valor esperado** de **y**, denotado como E(**y**), es el vector de valores esperados: $\mathrm{E}(\mathbf{y}) = [\mathrm{E}(y_1), \mathrm{E}(y_2), \ldots, \mathrm{E}(y_n)]'$.

ii) Si **Z** es una matriz aleatoria de $n \times m$, E(**Z**) es la matriz $n \times m$ de valores esperados: $\mathrm{E}(\mathbf{Z}) = [\mathrm{E}(z_{ij})]$.

Propiedades del valor esperado. 1) Si **A** es una matriz $m \times n$ y **b** es un vector $n \times 1$ ambos son no aleatorios, entonces $\mathrm{E}(\mathbf{A}\mathbf{y} + \mathbf{b}) = \mathbf{A}\mathrm{E}(\mathbf{y}) + \mathbf{b}$; y 2) Si **A** es $p \times n$ y **B** es $m \times k$, donde ambos son no aleatorios, entonces $\mathrm{E}(\mathbf{AZB}) = \mathbf{A}\mathrm{E}(\mathbf{Z})\mathbf{B}$.

Matriz varianza-covarianza

Definición D.16 (Matriz varianza-covarianza). Si **y** es un vector aleatorio $n \times 1$, su **matriz varianza-covarianza**, denotada como Var(**y**), se define como

$$\text{Var}(\mathbf{y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & & & \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{bmatrix},$$

donde $\sigma_j^2 = \text{Var}(y_j)$ y $\sigma_{ij} = \text{Cov}(y_i,y_j)$. En otras palabras, la matriz varianza-covarianza tiene las varianzas de cada elemento de **y** bajo su diagonal, con términos de covarianza fuera de las diagonales. Debido a que $\text{Cov}(y_i,y_j) = \text{Cov}(y_j,y_i)$, se desprende que la matriz varianza-covarianza es simétrica.

Propiedades de la varianza. 1) Si **a** es un vector no aleatorio de $n \times 1$ entonces $\text{Var}(\mathbf{a}'\mathbf{y}) = \mathbf{a}'[\text{Var}(\mathbf{y})]\mathbf{a} \geq 0$; 2) Si $\text{Var}(\mathbf{a}'\mathbf{y}) > 0$ para toda $\mathbf{a} \neq \mathbf{0}$, $\text{Var}(\mathbf{y})$ es definida positiva; 3) $\text{Var}(\mathbf{y}) = \text{E}[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})']$, donde $\boldsymbol{\mu} = \text{E}(\mathbf{y})$; 4) Si los elementos de **y** no se correlacionan, $\text{Var}(\mathbf{y})$ es una matriz diagonal. Si, además, $\text{Var}(y_j) = \sigma^2$ para $j = 1, 2, \ldots, n$, entonces $\text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}_n$; y 5) Si **A** es una matriz de $m \times n$ no aleatoria y **b** es un vector no aleatorio de $n \times 1$, entonces $\text{Var}(\mathbf{Ay} + \mathbf{b}) = \mathbf{A}[\text{Var}(\mathbf{y})]\mathbf{A}'$.

Distribución normal multivariada

La distribución normal de una variable aleatoria se analizó con cierto detalle en el apéndice B. Se debe ampliar la distribución normal a vectores no aleatorios. No se proporcionará una expresión para la función de distribución de probabilidad, puesto que no es necesaria. Es importante saber que un vector aleatorio normal multivariado se caracteriza completamente por su media y su matriz varianza-covarianza. Por tanto, si **y** es un vector aleatorio normal multivariado de $n \times 1$ con media $\boldsymbol{\mu}$ y matriz varianza-covarianza $\boldsymbol{\Sigma}$, se escribe $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Ahora se enunciarán las propiedades de la **distribución normal multivariada**.

Propiedades de la distribución normal multivariada. 1) Si $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, entonces cada elemento de **y** se distribuye normalmente; 2) Si $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, entonces y_i y y_j, dos elementos cualquiera de **y**, son independientes si, y sólo si, no se correlacionan, es decir, $\sigma_{ij} = 0$; 3) Si $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, entonces $\mathbf{Ay} + \mathbf{b} \sim \text{Normal}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$, donde **A** y **b** son no aleatorias; 4) Si $\mathbf{y} \sim \text{Normal}(\mathbf{0},\boldsymbol{\Sigma})$, entonces, para las matrices aleatorias **A** y **B**, **Ay** y **By** son independientes si, y sólo si, $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}' = \mathbf{0}$. En particular, si $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}_n$, entonces $\mathbf{AB}' = \mathbf{0}$ es necesaria y suficiente para la independencia de **Ay** y **By**; 5) Si $\mathbf{y} \sim \text{Normal}(\mathbf{0},\sigma^2\mathbf{I}_n)$, **A** es una matriz no aleatoria $k \times n$, y **B** es una matriz $n \times n$ simétrica e idempotente, entonces **Ay** y $\mathbf{y}'\mathbf{By}$ son independientes si, y sólo si, $\mathbf{AB} = \mathbf{0}$; y 6) Si $\mathbf{y} \sim \text{Normal}(\mathbf{0},\sigma^2\mathbf{I}_n)$ y **A** y **B** son matrices idempotentes simétricas no aleatorias, entonces $\mathbf{y}'\mathbf{Ay}$ y $\mathbf{y}'\mathbf{By}$ son independientes si, y sólo si, $\mathbf{AB} = \mathbf{0}$.

Distribución *ji*-cuadrada

En el apéndice B, se definió una **variable aleatoria *ji*-cuadrada** como la suma de las variables *cuadradas* aleatorias independientes y normales estandarizadas. En notación vectorial, si $\mathbf{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$, entonces $\mathbf{u}'\mathbf{u} \sim \chi_n^2$.

Propiedades de la distribución *ji*-cuadrada. 1) Si $\mathbf{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ y **A** es una matriz $n \times n$ simétrica e idenmpotente con $\text{rango}(\mathbf{A}) = q$, entonces $\mathbf{u}'\mathbf{Au} \sim \chi_q^2$; 2) Si $\mathbf{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ y **A** y **B**

son matrices $n \times n$ simétricas, idempotentes tal que $\mathbf{AB} = \mathbf{0}$, entonces $\mathbf{u}'\mathbf{Au}$ y $\mathbf{u}'\mathbf{Bu}$ son variables independientes aleatorias *ji*-cuadradas; y 3) Si $\mathbf{z} \sim \text{Normal}(\mathbf{0},\mathbf{C})$ donde $\mathbf{C}$ es una matriz $m \times m$ no singular, entonces, $\mathbf{z}'\mathbf{C}^{-1}\mathbf{z} \sim \chi^2_m$.

Distribución *t*

También se definió la **distribución *t*** en el apéndice **B**. Ahora se agregará una propiedad importante.

Propiedad de la distribución *t*. Si $\mathbf{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$, $\mathbf{c}$ es un vector de $n \times 1$ no aleatorio, $\mathbf{A}$ es una matriz $n \times n$ no aleatoria simétrica e idempotente con rango q, y $\mathbf{Ac} = \mathbf{0}$, entonces $\{\mathbf{c}'\mathbf{u}/(\mathbf{c}'\mathbf{c})^{1/2}\}/(\mathbf{u}'\mathbf{Au})^{1/2} \sim t_q$.

Distribución *F*

Recuerde que la **variable aleatoria *F*** se obtiene al determinar dos variables aleatorias *ji*-cuadradas *independientes* y se calcula la razón de cada una estandarizada por sus grados de libertad.

Propiedad de la distribución *F*. Si $\mathbf{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ y $\mathbf{A}$ y $\mathbf{B}$ son matrices no aleatorias simétricas e idempotentes con rango $n \times n$, rango$(\mathbf{A}) = k_1$, rango$(\mathbf{B}) = k_2$ y $\mathbf{AB} = \mathbf{0}$, entonces $(\mathbf{u}'\mathbf{Au}/k_1)/(\mathbf{u}'\mathbf{Bu}/k_2) \sim F_{k_1,k_2}$.

RESUMEN

Este apéndice contiene una forma condensada de la información general necesaria para estudiar el modelo lineal clásico mediante matrices. Aunque el material que aquí se presenta es independiente, tiene el objetivo de ser un repaso para los lectores familiarizados con el álgebra matricial y la estadística multivariada, y se usará ampliamente en el apéndice E.

TÉRMINOS CLAVE

Definida positiva
Distribución normal multivariada
Distribución *t*
Forma cuadrática
Inversa
Matriz
Matriz cero
Matriz cuadrada
Matriz diagonal
Matriz idempotente
Matriz identidad
Matriz simétrica
Matriz varianza-covarianza
Multiplicación escalar
Multiplicación matricial
Rango de una matriz
Semidefinida positiva
Transpuesta
Traza de una matriz
Valor esperado
Variable aleatoria *F*
Variable aleatoria *ji*-cuadrada
Vector aleatorio
Vector columna
Vector fila
Vectores linealmente independientes

PROBLEMAS

D.1 i) Determine el producto **AB** mediante

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 7 \\ -4 & 5 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 6 \\ 1 & 8 & 0 \\ 3 & 0 & 0 \end{bmatrix}.$$

ii) ¿Existe **BA**?

D.2 Si **A** y **B** son matrices diagonales $n \times n$, muestre que $\mathbf{AB} = \mathbf{BA}$.

D.3 Sea **X** cualquier matriz $n \times k$. Muestre que $\mathbf{X}'\mathbf{X}$ es una matriz simétrica.

D.4 i) Use las propiedades de traza para argumentar que $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}')$ para toda matriz $n \times m$ de **A**.

ii) Para $\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 3 & 0 \end{bmatrix}$, verifique que $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}')$.

D.5 i) Use la definición de la inversa para demostrar lo siguiente: si **A** y **B** son matrices no singulares $n \times n$, entonces $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

ii) Si **A**, **B** y **C** son matrices $n \times n$ no singulares, encuentre $(\mathbf{ABC})^{-1}$ en términos de $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$ y $\mathbf{C}^{-1}$.

D.6 i) Muestre que si **A** es una matriz $n \times n$ simétrica y definida positiva, entonces **A** debe tener elementos diagonales estrictamente positivos.

ii) Escriba una matriz simétrica 2×2 con elementos diagonales estrictamente positivos, que *no* sea definida positiva.

D.7 Sea **A** una matriz $n \times n$ simétrica y definida positiva. Muestre que si **P** es una matriz $n \times n$ no singular, entonces $\mathbf{P}'\mathbf{AP}$ es definida positiva.

D.8 Demuestre la propiedad 5 de las varianzas de vectores mediante la propiedad 3.