D

Resumo de Álgebra Matricial

Este apêndice resume os conceitos de álgebra matricial, inclusive da álgebra de probabilidade, necessária para o estudo de modelos de regressão linear múltipla usando matrizes, do Apêndice E. Nada deste material é usado no texto principal.

D.1 DEFINIÇÕES BÁSICAS

DEFINIÇÃO D.1 (MATRIZ)

Uma **matriz** é um arranjo retangular de números. Mais precisamente, uma matriz $m \times n$ tem m linhas e n colunas. O inteiro positivo m é denominado *dimensão da linha*, e n é denominado *dimensão da coluna*.

Usamos letras maiúsculas em negrito para representar as matrizes. Podemos escrever, de forma geral, uma matriz de ordem $m \times n$ como

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

onde a_{ij} representa o elemento na $i^{\text{-ésima}}$ linha e na $j^{\text{-ésima}}$ coluna. Por exemplo, a_{25} corresponde ao número na segunda linha e na quinta coluna de $\mathbf{A}$. Um exemplo específico de uma matriz 2×3 é

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

onde $a_{13} = 7$. A forma abreviada $\mathbf{A} = [a_{ij}]$ é freqüentemente usada para definir operações de matrizes.

DEFINIÇÃO D.2 (MATRIZ QUADRADA)

Uma **matriz quadrada** tem o mesmo número de linhas e colunas. A dimensão de uma matriz quadrada é dada por seus números de linhas e colunas.

DEFINIÇÃO D.3 (VETORES)

(i) Uma matriz $1 \times m$ é chamada **vetor linha** (de dimensão m) e pode ser escrita como $\boldsymbol{x} \equiv (x_1, x_2, ..., x_m)$.

(ii) Uma matriz $n \times 1$ é chamada **vetor coluna** e pode ser escrita como

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

DEFINIÇÃO D.4 (MATRIZ DIAGONAL)

Uma matriz quadrada $\boldsymbol{A}$ é uma **matriz diagonal** quando todos os seus elementos fora da diagonal forem zeros, isto é, $a_{ij} = 0$ para todos $i \neq j$. Podemos sempre escrever uma matriz diagonal como

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

DEFINIÇÃO D.5 (MATRIZES IDENTIDADE E NULA)

(i) A **matriz identidade** $n \times n$, chamada de $\mathbf{I}$, ou algumas vezes $\mathbf{I}_n$ para enfatizar sua dimensão, é a matriz diagonal com a unidade (um) em cada posição diagonal, e zero nas outras posições:

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

(ii) A **matriz nula** $m \times n$, chamada de $\mathbf{0}$, é a matriz $m \times n$ com zero em todas as entradas. Ela não precisa ser um matriz quadrada.

D.2 OPERAÇÕES COM MATRIZES

Adição de Matrizes

Duas matrizes $\mathbf{A}$ e $\mathbf{B}$, cada uma de dimensão $m \times n$, podem ser somadas elemento a elemento: $\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}]$. Mais precisamente,

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

Por exemplo,

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

Matrizes de ordens diferentes não podem ser somadas.

Multiplicação Escalar

Dado qualquer número real γ (freqüentemente chamado de escalar), uma **multiplicação escalar** é definida como $\gamma\mathbf{A} \equiv [\gamma a_{ij}]$, ou

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

Por exemplo, se $\gamma = 2$, $\mathbf{A}$ é a matriz na equação (D.1), então,

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

Multiplicação de Matrizes

Para multiplicar a matriz $\mathbf{A}$ pela matriz $\mathbf{B}$ de maneira a formar o produto $\mathbf{AB}$, a dimensão das *colunas* de $\mathbf{A}$ deve ser igual à dimensão das *linhas* de $\mathbf{B}$. Portanto, seja $\mathbf{A}$ uma matriz $m \times n$ e seja $\mathbf{B}$ uma matriz $n \times p$. Então, a **multiplicação de matrizes** será definida como

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

Em outras palavras, o $(i,j)^{\text{-ésimo}}$ elemento da nova matriz $\mathbf{AB}$ é obtido pela multiplicação de cada elemento na $i^{\text{-ésima}}$ linha de $\mathbf{A}$ pelo elemento correspondente na $j^{\text{-ésima}}$ coluna de $\mathbf{B}$ e somando-se esses n produtos. Um esquema pode ajudar a tornar esse processo mais transparente:

$$\begin{array}{cccccc} & \mathbf{A} & \mathbf{B} & & \mathbf{AB} & \\ i^{\text{-ésima linha}} \rightarrow & \begin{bmatrix} \\ a_{i1}\ a_{i2}\ a_{i3} \cdots a_{in} \\ \\ \end{bmatrix} & \begin{bmatrix} b_{1j} \\ b_{2j} \\ b_{3j} \\ \vdots \\ b_{nj} \end{bmatrix} & = & \begin{bmatrix} \\ \sum_{k=1}^{n} a_{ik} b_{kj} \\ \\ \end{bmatrix}, & \\ & & \uparrow & & \uparrow & \\ & & j^{\text{-ésima}}\text{ coluna} & & (i,j)^{\text{-ésimo}}\text{ elemento} & \end{array}$$

onde, pela definição do operador somatório no Apêndice A,

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

Por exemplo,

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

Também podemos multiplicar uma matriz e um vetor. Se $\mathbf{A}$ for uma matriz $n \times m$ e $\boldsymbol{y}$ for um vetor $m \times 1$, $\mathbf{Ay}$ será um vetor de ordem $n \times 1$. Se $\mathbf{x}$ for um vetor $1 \times n$, $\mathbf{xA}$ será um vetor $1 \times m$.

A adição de matrizes, a multiplicação escalar e a multiplicação de matrizes podem ser combinadas de várias maneiras, e essas operações satisfazem várias regras das operações básicas com números que nos são familiares. Na lista de propriedades seguinte, $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ são matrizes com dimensões apropriadas para aplicar cada operação, e α e β são números reais. A maioria dessas propriedades pode ser ilustrada facilmente a partir das definições.

PROPRIEDADES DA MULTIPLICAÇÃO DE MATRIZES: (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) $\mathbf{A0} = \mathbf{0A} = \mathbf{0}$; e (14) $\mathbf{AB} \neq \mathbf{BA}$, mesmo quando ambos os produtos forem definidos.

A última propriedade merece um comentário adicional. Se $\mathbf{A}$ for $n \times m$ e $\mathbf{B}$ for $m \times p$, então, $\mathbf{AB}$ será definida, mas $\mathbf{BA}$ somente será definida se $n = p$ (a ordem de linha da $\mathbf{A}$ é igual a ordem de coluna da $\mathbf{B}$). Se $\mathbf{A}$ for $m \times n$ e $\mathbf{B}$ for $n \times m$, então, $\mathbf{AB}$ e $\mathbf{BA}$ serão ambas definidas, mas normalmente não serão a mesma matriz; de fato, elas possuem dimensões diferentes, a menos que $\mathbf{A}$ e $\mathbf{B}$ sejam ambas matrizes quadradas. Mesmo quando $\mathbf{A}$ e $\mathbf{B}$ são quadradas, $\mathbf{AB} \neq \mathbf{BA}$, exceto sob circunstâncias especiais.

Transposta

DEFINIÇÃO D6 (TRANSPOSTA)

Seja $\mathbf{A} = [a_{ij}]$ uma matriz $m \times n$. A transposta de $\mathbf{A}$, chamada de $\mathbf{A}'$ ($\mathbf{A}$ linha), é a matriz de ordem $n \times m$, obtida intercambiando as linhas e colunas de $\mathbf{A}$. Podemos escrevê-la como $\mathbf{A}' \equiv [a_{ji}]$.

Por exemplo,

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

PROPRIEDADES DA TRANSPOSTA: (1) $(\mathbf{A}')' = \mathbf{A}$; (2) $(\alpha\mathbf{A})' = \alpha\mathbf{A}'$ para qualquer escalar α; (3) $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$; (4) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$, quando $\mathbf{A}$ for $m \times n$ e $\mathbf{B}$ for $n \times k$; (5) $\boldsymbol{x}'\boldsymbol{x} = \sum_{i=1}^{n} x_i^2$, onde $\boldsymbol{x}$ é um vetor $n \times 1$; e (6) se $\mathbf{A}$ for uma matriz $n \times k$ com *linhas* dadas pelos vetores $1 \times k$ $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$, de forma que possamos escrever

$$\mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

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

DEFINIÇÃO D.7 (MATRIZ SIMÉTRICA)
Uma matriz quadrada $\mathbf{A}$ é uma **matriz simétrica** se, e somente se, $\mathbf{A}' = \mathbf{A}$.

Se $\mathbf{X}$ for qualquer matriz $n \times k$, $\mathbf{X}'\mathbf{X}$ é sempre definida e é uma matriz simétrica, como poderá ser visto pela aplicação das primeira e quarta propriedades da transposta (veja o Problema D.3).

Multiplicação de Matrizes Particionadas

Seja $\mathbf{A}$ uma matriz $n \times k$ com linhas dadas pelos vetores $1 \times k$ $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$, e seja $\mathbf{B}$ uma matriz $n \times m$ com linhas dadas pelos vetores $1 \times m$ $\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n$;

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

Então,

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

onde para cada i, $\boldsymbol{a}_i'\boldsymbol{b}_i$ é uma matriz $k \times m$. Portanto, $\mathbf{A}'\mathbf{B}$ pode ser escrita como a soma de n matrizes, cada uma delas sendo $k \times m$. Como um caso especial, temos

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

onde $\boldsymbol{a}_i' \boldsymbol{a}_i$ é uma matriz $k \times k$ de todos os i.

Traço

O traço de uma matriz é uma operação bastante simples definida somente para matrizes *quadradas*.

DEFINIÇÃO D.8 (TRAÇO)
Para qualquer matriz $\mathbf{A}$ $n \times n$, o **traço de uma matriz** $\mathbf{A}$, representado por $\text{tr}(\mathbf{A})$, será a soma dos elementos de sua diagonal. Matematicamente,

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

PROPRIEDADES DO TRAÇO: (1) $\text{tr}(\mathbf{I}_n) = n$; (2) $\text{tr}(\mathbf{A}') = \text{tr}(\mathbf{A})$; (3) $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$; (4) $\text{tr}(\alpha\mathbf{A}) = \alpha\text{tr}(\mathbf{A})$, para qualquer escalar α; e (5) $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, onde $\mathbf{A}$ é $m \times n$ e $\mathbf{B}$ é $n \times m$.

Inversa

A noção da inversa de uma matriz é muito importante para matrizes quadradas.

DEFINIÇÃO D.9 (INVERSA)
Uma matriz $\mathbf{A}$ $n \times n$ tem uma **inversa**, representada por $\mathbf{A}^{-1}$, desde que $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$ e $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$. Nesse caso, $\mathbf{A}$ é chamada de *inversível* ou *não singular*. Caso contrário, ela será chamada de *não inversível* ou *singular*.

PROPRIEDADES DA INVERSA: (1) Se existir uma inversa, ela será única; (2) $(\alpha\mathbf{A})^{-1} = (1/\alpha)\mathbf{A}^{-1}$, se $\alpha \neq 0$ e $\mathbf{A}$ for inversível; (3) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, se $\mathbf{A}$ e $\mathbf{B}$ forem ambas $n \times n$ e inversíveis; e (4) $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$.

Não nos preocuparemos com a mecânica de cálculo da inversa de uma matriz. Qualquer texto de álgebra matricial contém exemplos detalhados de tais cálculos.

D.3 INDEPENDÊNCIA LINEAR. POSTO DE UMA MATRIZ

Para um conjunto de vetores tendo a mesma dimensão, é importante saber se um vetor pode ser expresso como uma combinação linear dos demais vetores.

DEFINIÇÃO D.10 (INDEPENDÊNCIA LINEAR)
Seja $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_r\}$ um conjunto de vetores $n \times 1$. Eles serão **vetores linearmente independentes** se, e somente se,

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

implicar que $\alpha_1 = \alpha_2 = ... = \alpha_r = 0$. Se (D.2) se sustentar para um conjunto de escalares que não sejam todos zeros, então, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_r\}$ será *linearmente dependente*.

A afirmação de que $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_r\}$ será linearmente dependente é equivalente a dizer que pelo menos um vetor no conjunto pode ser escrito como uma combinação linear dos demais.

DEFINIÇÃO D.11 (POSTO)
(i) Seja $\mathbf{A}$ uma matriz $n \times m$. O **posto de uma matriz A**, denotada posto($\mathbf{A}$), é o número máximo de colunas linearmente independentes de $\mathbf{A}$.

(ii) Se $\mathbf{A}$ for $n \times m$ e posto($\mathbf{A}$) $= m$, então, $\mathbf{A}$ terá *posto de colunas completas*.

Se $\mathbf{A}$ for $n \times m$, seu posto será no máximo m. Uma matriz terá posto de colunas completas se suas colunas formarem um conjunto linearmente independente. Por exemplo, a matriz 3×2

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

poderá ter no máximo posto 2. De fato, seu posto será apenas um porquê a segunda coluna é três vezes a primeira coluna.

PROPRIEDADES DO POSTO: (1) posto($\mathbf{A}'$) = posto($\mathbf{A}$); (2) Se $\mathbf{A}$ for $n \times k$, então, posto($\mathbf{A}$) $\leq$ min(n,k); e (3) Se $\mathbf{A}$ for $k \times k$ e posto($\mathbf{A}$) $= k$, então, $\mathbf{A}$ será não singular.

D.4 FORMAS QUADRÁTICAS E MATRIZES POSITIVAS DEFINIDAS

DEFINIÇÃO D.12 (FORMA QUADRÁTICA)

Seja $\mathbf{A}$ uma matriz simétrica $n \times n$. A **forma quadrática** associada à matriz $\mathbf{A}$ será a função de valor real definida para todos os vetores $\boldsymbol{x}$ $n \times 1$:

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

DEFINIÇÃO D.13 (POSITIVA DEFINIDA E POSITIVA SEMIDEFINIDA)

(i) Uma matriz simétrica $\mathbf{A}$ é **positiva definida (p.d.)** se

$$\boldsymbol{x}'\mathbf{A}\boldsymbol{x} > 0 \text{ para todos os vetores } \boldsymbol{x} \; n \times 1, \text{ exceto } \boldsymbol{x} = \mathbf{0}.$$

(ii) Uma matriz simétrica $\mathbf{A}$ é **positiva semidefinida (p.s.d.)** se

$$\boldsymbol{x}'\mathbf{A}\boldsymbol{x} \geq 0 \text{ para todos os vetores } n \times 1.$$

Se uma matriz for positiva definida ou positiva semidefinida, ela será automaticamente assumida como sendo simétrica.

PROPRIEDADES DAS MATRIZES POSITIVAS DEFINIDAS E POSITIVAS SEMIDEFINIDAS: (1) Uma matriz positiva definida tem elementos diagonais que são estritamente positivos, enquanto uma matriz p.s.d. tem elementos diagonais não negativos; (2) Se $\mathbf{A}$ for uma p.d., então, $\mathbf{A}^{-1}$ existe e será p.d.; (3) Se $\mathbf{X}$ for $n \times k$, então, $\mathbf{X}'\mathbf{X}$ e $\mathbf{X}\mathbf{X}'$ serão p.s.d.; e (4) Se $\mathbf{X}$ for $n \times k$ e posto($\mathbf{X}$) $= k$, então, $\mathbf{X}'\mathbf{X}$ será p.d. (e, portanto, não singular).

D.5 MATRIZES IDEMPOTENTES

DEFINIÇÃO D.14 (MATRIZ IDEMPOTENTE)

Seja $\mathbf{A}$ uma matriz simétrica $n \times n$. Então, $\mathbf{A}$ é uma **matriz idempotente** se, e somente se, $\mathbf{AA} = \mathbf{A}$.

Por exemplo,

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

é uma matriz idempotente, como é possível verificar pela multiplicação direta.

PROPRIEDADES DAS MATRIZES IDEMPOTENTES: Seja $\mathbf{A}$ uma matriz idempotente $n \times n$. (1) posto($\mathbf{A}$) $=$ tr($\mathbf{A}$), e (2) $\mathbf{A}$ é positiva semidefinida.

Podemos construir matrizes idempotentes de forma bem generalizada. Seja $\mathbf{X}$ uma matriz $n \times k$ com posto($\mathbf{X}$) $= k$. Defina

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

Então, $\mathbf{P}$ e $\mathbf{M}$ serão matrizes simétricas idempotentes, com posto($\mathbf{P}$) $= k$ e posto($\mathbf{M}$) $= n - k$. Os postos são obtidos com mais facilidade usando a Propriedade 1: tr($\mathbf{P}$) $=$ tr[$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}$] (da Propriedade 5 do traço) $=$ tr($\mathbf{I}_k$) $= k$ (pela Propriedade 1 do traço). Facilmente segue que tr($\mathbf{M}$) $=$ tr($\mathbf{I}_n$) $-$ tr($\mathbf{P}$) $= n - k$.

D.6 DIFERENCIAÇÃO DE FORMAS LINEARES E QUADRÁTICAS

Para um dado vetor $\boldsymbol{a}$ $n \times 1$, considere a função linear definida por

$$f(\boldsymbol{x}) - \boldsymbol{a}'\boldsymbol{x},$$

para todos os vetores $\boldsymbol{x}$ $n \times 1$. A derivada de f em relação a $\boldsymbol{x}$ será o vetor $1 \times n$ das derivadas parciais, que é simplesmente

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

Para uma matriz simétrica $n \times n$ $\boldsymbol{A}$, defina a forma quadrática

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

Então,

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

que é um vetor $1 \times n$.

D.7 MOMENTOS E DISTRIBUIÇÕES DE VETORES ALEATÓRIOS

Para derivar o valor esperado e a variância dos estimadores MQO usando matrizes, precisamos definir o valor esperado e a variância de um **vetor aleatório**. Como seu nome sugere, um vetor aleatório é simplesmente um vetor de variáveis aleatórias. Também precisamos definir a distribuição normal multivariada. Esses conceitos são simples extensões daqueles tratados no Apêndice B.

Valor Esperado

DEFINIÇÃO D.15 (VALOR ESPERADO)

(i) Se $\boldsymbol{y}$ for um vetor aleatório $n \times 1$, o **valor esperado** de $\boldsymbol{y}$, representado por E($\boldsymbol{y}$), é o vetor dos valores esperados: $E(\boldsymbol{y}) = [E(y_1), E(y_2),..., E(y_n)]'$.

(ii) Se $\mathbf{Z}$ for uma matriz aleatória $n \times m$, E($\mathbf{Z}$) é a matriz $n \times m$ dos valores esperados: $E(\mathbf{Z}) = [E(z_{ij})]$.

PROPRIEDADES DO VALOR ESPERADO: (1) Se $\mathbf{A}$ for uma matriz $m \times n$ e $\boldsymbol{b}$ for um vetor $n \times 1$, onde ambos são não-aleatórios, $E(\mathbf{A}\boldsymbol{y} + \boldsymbol{b}) = \mathbf{A}E(\boldsymbol{y}) + \boldsymbol{b}$ e (2) Se $\mathbf{A}$ for $p \times n$ e $\mathbf{B}$ for $m \times k$, onde ambas são não-aleatórios, $E(\mathbf{AZB}) = \mathbf{A}E(\mathbf{Z})\mathbf{B}$.

Matriz de Variância-Covariância

DEFINIÇÃO D.16 (MATRIZ DE VARIÂNCIA-COVARIÂNCIA)

Se $\boldsymbol{y}$ for um vetor aleatório $n \times 1$, sua **matriz de variância-covariância**, representada por Var($\boldsymbol{y}$) é definida como

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

onde $\sigma_j^2 = \mathrm{Var}(y_j)$ e $\sigma_{ij} = \mathrm{Cov}(y_i,y_j)$. Em outras palavras, a matriz de variância-covariância tem as variâncias de cada elemento de $\boldsymbol{y}$ em sua diagonal, com os termos de covariância fora dela. Como $\mathrm{Cov}(y_i,y_j) = \mathrm{Cov}(y_j,y_i)$, segue imediatamente que uma matriz de variância-covariância é simétrica.

PROPRIEDADES DA VARIÂNCIA: (1) Se $\boldsymbol{a}$ for um vetor não-aleatório $n \times 1$, então, $\mathrm{Var}(\boldsymbol{a}'\mathbf{y}) = \boldsymbol{a}'[\mathrm{Var}(\mathbf{y})]\boldsymbol{a} \geq 0$; (2) Se $\mathrm{Var}(\boldsymbol{a}'\mathbf{y}) > 0$ para todo $\boldsymbol{a} \neq \mathbf{0}$, Var($\boldsymbol{y}$) é positiva definida; (3) $\mathrm{Var}(\boldsymbol{y}) = E[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})']$, onde $\boldsymbol{\mu} = E(\boldsymbol{y})$; (4) Se os elementos de $\boldsymbol{y}$ forem não-correlacionados, Var($\boldsymbol{y}$) é uma matriz diagonal. Se, além disso, $\mathrm{Var}(y_j) = \sigma^2$ para $j = 1, 2, ..., n$, então $\mathrm{Var}(\boldsymbol{y}) = \sigma^2 \boldsymbol{I}_n$ e (5) Se $\boldsymbol{A}$ for uma matriz não-aleatória $m \times n$ e $\boldsymbol{b}$ for um vetor não-aleatório $n \times 1$, então, $\mathrm{Var}(\mathbf{Ay} + \boldsymbol{b}) = \boldsymbol{A}[\mathrm{Var}(\mathbf{y})]\boldsymbol{A}'$.

Distribuição Normal Multivariada

A distribuição normal de uma variável aleatória foi discutida em alguma medida no Apêndice B. Precisamos estender a distribuição normal aos vetores aleatórios. Não forneceremos uma expressão da função de distribuição de probabilidade, já que não precisaremos dela. É importante saber que um vetor aleatório normal multivariado é totalmente caracterizado por sua média e pela matriz de variância-covariância. Portanto, se $\boldsymbol{y}$ for um vetor aleatório normal multivariado $n \times 1$, com média $\boldsymbol{\mu}$ e matriz de variância-covariância $\mathbf{\Sigma}$, escrevemos $\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{\mu},\mathbf{\Sigma})$. Agora, apresentamos várias propriedades úteis da **distribuição normal multivariada**.

PROPRIEDADES DA DISTRIBUIÇÃO NORMAL MULTIVARIADA: (1) Se $\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{\mu},\mathbf{\Sigma})$, então, cada elemento de $\boldsymbol{y}$ é normalmente distribuído; (2) Se $\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{\mu},\mathbf{\Sigma})$, então, y_i e y_j, quaisquer dois elementos de $\boldsymbol{y}$, serão independentes se, e somente se, eles forem não-correlacionados, isto é, $\sigma_{ij} = 0$; (3) Se $\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{\mu},\mathbf{\Sigma})$, então, $\mathbf{A}\boldsymbol{y} + \mathbf{b} \sim \mathrm{Normal}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b},\mathbf{A\Sigma A}')$, onde $\mathbf{A}$ e $\boldsymbol{b}$ são não-aleatórios; (4) Se $\boldsymbol{y} \sim \mathrm{Normal}(\mathbf{0},\mathbf{\Sigma})$, então, para matrizes $\mathbf{A}$ e $\mathbf{B}$ não-aleatórias, $\mathbf{A}\boldsymbol{y}$ e $\mathbf{B}\boldsymbol{y}$ serão independentes se, e somente se, $\mathbf{A\Sigma B}' = \mathbf{0}$. Particularmente, se $\mathbf{\Sigma} = \sigma^2\mathbf{I}_n$, então, $\mathbf{AB}' = \mathbf{0}$ será necessária e suficiente para a independência de $\mathbf{A}\boldsymbol{y}$ e $\mathbf{B}\boldsymbol{y}$; (5) Se $\boldsymbol{y} \sim \text{Normal}(\mathbf{0},\sigma^2\mathbf{I}_n)$, $\mathbf{A}$ é uma matriz não-aleatória $k \times n$, e $\mathbf{B}$ é uma matriz simétrica idempotente $n \times n$, então, $\mathbf{A}\boldsymbol{y}$ e $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ serão independentes se, e somente se, $\mathbf{AB} = \mathbf{0}$; e (6) Se $\boldsymbol{y} \sim \text{Normal}(\mathbf{0},\sigma^2\mathbf{I}_n)$ e $\mathbf{A}$ e $\mathbf{B}$ forem matrizes simétricas idempotentes não-aleatórias, então, $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ e $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ serão independentes se, e somente se, $\mathbf{AB} = \mathbf{0}$.

DISTRIBUIÇÃO QUI-QUADRADO

No Apêndice B, definimos uma **variável aleatória qui-quadrada** como a soma dos quadrados das variáveis aleatórias normais padrão independentes. Em notação vetorial, se $\boldsymbol{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$, então, $\boldsymbol{u}'\boldsymbol{u} \sim \chi_n^2$

PROPRIEDADES DA DISTRIBUIÇÃO QUI-QUADRADO: (1) Se $\boldsymbol{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ e $\mathbf{A}$ for uma matriz simétrica idempotente $n \times n$ com posto($\mathbf{A}$) $= q$, então, $\boldsymbol{u}'\mathbf{A}\boldsymbol{u} \sim \chi_q^2$; (2) Se $\boldsymbol{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ e $\mathbf{A}$ e $\mathbf{B}$ forem matrizes simétricas idempotentes $n \times n$ de tal forma que $\mathbf{AB} = \mathbf{0}$, então, $\boldsymbol{u}'\mathbf{A}\boldsymbol{u}$ e $\boldsymbol{u}'\mathbf{B}\boldsymbol{u}$ serão variáveis aleatórias qui-quadradas independentes; e (3) Se $z \sim \text{Normal}(\mathbf{0},\mathbf{C})$ onde $\mathbf{C}$ é uma matriz não singular $m \times m$, então, $z'\mathbf{C}^{-1}z \sim \chi_m^2$.

DISTRIBUIÇÃO *t*

Também já definimos a **distribuição *t*** no Apêndice B. Agora adicionamos uma propriedade importante.

PROPRIEDADE DA DISTRIBUIÇÃO *t*: Se $\boldsymbol{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$, $\boldsymbol{c}$ é um vetor não-aleatório $n \times 1$, $\mathbf{A}$ é uma matriz não-aleatória simétrica idempotente $n \times n$ com posto q, e $\mathbf{A}\boldsymbol{c} = \mathbf{0}$, então, $\{\boldsymbol{c}'\boldsymbol{u}/(\boldsymbol{c}'\boldsymbol{c})^{1/2}\}/(\boldsymbol{u}'\mathbf{A}\boldsymbol{u})^{1/2} \sim t_q$.

DISTRIBUIÇÃO *f*

Recorde-se de que uma **variável aleatória *F*** é obtida tomando-se duas variáveis aleatórias qui-quadradas *independentes* e encontrando-se a razão entre elas, padronizadas pelos graus de liberdade.

PROPRIEDADE DA DISTRIBUIÇÃO *f*: Se $\boldsymbol{u} \sim \text{Normal}(\mathbf{0},\mathbf{I}_n)$ e $\mathbf{A}$ e $\mathbf{B}$ forem matrizes simétricas idempotentes não-aleatórias $n \times n$ com posto($\mathbf{A}$) $= k_1$, posto($\mathbf{B}$) $= k_2$ e $\mathbf{AB} = \mathbf{0}$, então, $(\boldsymbol{u}'\mathbf{A}\boldsymbol{u}/k_1)/(\boldsymbol{u}'\mathbf{B}\boldsymbol{u}/k_2) \sim F_{k_1,k_2}$.

RESUMO

Este apêndice contém uma forma condensada das informações básicas necessárias ao estudo do modelo linear clássico usando matrizes. Embora o material aqui apresentado não dependa de outros, ele é apresentado com a intenção de servir como uma revisão para os leitores familiarizados com álgebra matricial e estatística multivariada, e será amplamente usado no Apêndice E.

PROBLEMAS

D.1 (i) Encontre o produto **AB** usando

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

(ii) **BA** existe?

D.2 Se **A** e **B** forem matrizes diagonais $n \times n$, demonstre que $\mathbf{AB} = \mathbf{BA}$.

D.3 Seja **X** qualquer matriz $n \times k$. Mostre que $\mathbf{X}'\mathbf{X}$ é uma matriz simétrica.

D.4 (i) Use as propriedades do traço para demonstrar que $\text{tr}(\mathbf{A}'\mathbf{A}) = \text{tr}(\mathbf{AA}')$ para qualquer matriz **A** $n \times m$.

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

D.5 (i) Use a definição da inversa para provar o seguinte: se **A** e **B** forem matrizes não singulares $n \times n$, então, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

(ii) Se **A**, **B** e **C** forem todas matrizes não singulares $n \times n$, encontre $(\mathbf{ABC})^{-1}$ em termos de $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$ e $\mathbf{C}^{-1}$.

D.6 (i) Mostre que, se **A** for uma matriz positiva definida, simétrica $n \times n$, então, **A** deve ter elementos diagonais estritamente positivos.

(ii) Escreva uma matriz simétrica 2×2 com elementos diagonais estritamente positivos que *não* seja positiva definida.

D.7 Seja **A** uma matriz positiva definida, simétrica $n \times n$. Mostre que, se **P** for qualquer matriz não singular $n \times n$, então, $\mathbf{P}'\mathbf{AP}$ será positiva definida.

D.8 Prove a Propriedade 5 das variâncias dos vetores, usando a Propriedade 3.