Instituto de Estadística
Pontificia Universidad Católica de Valparaíso

Computación Estadística II

Operaciones matriciales

Daira Velandia[1] M. & Juan Zamora O.[2]
[1]daira.velandia , [2]juan.zamora *@pucv.cl*

4 de septiembre de 2018

Los procedimientos matemáticos que definen el álgebra lineal posibilitan la manipulación de grandes masas de datos, garantizando brevedad, simplicidad y claridad en la obtención de resultados, por tanto, constituye un soporte formal fundamental para el desarrollo de las técnicas de análisis multivariado.

Ejemplo 1

Se dispone de datos correspondientes a inversiones en los diferentes tipos de industrias de un país X, medidas en millones de dólares, desde 2007 hasta 2010. El resumen de esta información puede presentarse sintéticamente sobre la siguiente matriz:

$$\begin{pmatrix} 1000 & 800 & 500 & 600 \\ 1100 & 850 & 550 & 650 \\ 1200 & 900 & 575 & 750 \\ 1400 & 1000 & 500 & 850 \end{pmatrix}$$

En esta matriz las filas representan años y las columnas indican el tipo de industria. El elemento genérico a_{ij} representa la inversión efectuada durante el año i en la industria j.

Ejemplo 2

Considere un proceso Gaussiano bivariado estationario con media cero observado en un conjunto compacto T of $\mathbb{R}$, $Z(s)=\{(Z_1(s),Z_2(s))^\top,\ s\in T\}$ con función de covarianza indexada por el parámetro $\psi=(\sigma_1^2,\sigma_2^2,\rho,\theta)^\top\in\mathbb{R}^4$, dada por

$$\mathrm{Cov}_\psi(Z_i(s_l),Z_j(s_m))=\sigma_i\sigma_j(\rho+(1-\rho)\mathbf{1}_{i=j})e^{-\theta|s_l-s_m|},\ i,j=1,2. \qquad (1)$$

Se necesita obtener $\widehat{\psi}=(\hat{\theta},\hat{\sigma}_1^2,\hat{\sigma}_2^2,\hat{\rho})$, el estimador ML, que se obtiene maximizando $f_n(\psi)$ con respecto a ψ.
The explicit expression for the negative log-likelihood function

$$l_n(\psi)=-2\log(f_n(\psi))=2n\log(2\pi)+\log|\Sigma(\psi)|+Z_n^\top\left[\Sigma(\psi)\right]^{-1}Z_n. \qquad (2)$$

Definiciones básicas y notación I

Un **espacio vectorial** sobre un cuerpo $\mathbb{K}$ es un conjunto $\mathbb{V}$ no vacío, dotado de dos operaciones para las cuales será cerrado.

$$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$$

tal que:

1. se cumpla la propiedad conmutativa.
2. se cumpla la propiedad asociativa.
3. tenga elemento neutro.
4. tenga elemento opuesto.

Y la operación:

$$\cdot : \mathbb{K} \times \mathbb{V} \to \mathbb{V}$$

operación externa tal que:

1. se cumpla la propiedad asociativa.
2. tenga elemento neutro multiplicativo.
3. se cumpla la propiedad distributiva.

Definiciones básicas y notación II

Ejemplo

El conjunto de los polinomios P_2 de grado ≤ 2 con coeficientes reales es un espacio vectorial.

$$P_2 = ax^2 + bx + c : a, b, c \in \mathbb{R}$$

El espacio vectorial generado por las columnas de una matriz A de $n \times m$ es de orden n y de dimensión $\leq m$, y es llamado el espacio columna de A o el rango de A. Este espacio vectorial es denotado por

$$V(A)$$

Algunas matrices especiales

1. Matriz simétrica: si $a_{ij} = a_{ji} \quad \forall i, j$. (Una matriz simétrica necesariamente es cuadrada).
2. Matriz simétrica sesgada: si $a_{ij} = -a_{ji} \quad \forall i, j$. (Las entradas de la diagonal deben ser cero).
3. Matriz hermitiana: si $a_{ij} = \bar{a}_{ji} \quad \forall i, j$, donde $\bar{a}$ representa el conjugado del número complejo a. (Una matriz hermitiana necesariamente es cuadrada).
4. Matriz sparse: aquella que tiene una gran proporción de elementos que son ceros.
5. Matriz diagonal: si $a_{ij} = 0 \quad \forall i \neq j$.
6. Matriz hueca: si $a_{ij} = 0 \quad \forall i = j$.
7. Diagonal dominante por filas: Si $|a_{ii}| > \sum_{j \neq i}^{m} |a_{ij}|$ para cada $i = 1, 2, ..., n$.
8. Una matriz triangular es un tipo especial de matriz cuadrada cuyos elementos por encima o por debajo de su diagonal principal son cero.
9. Matriz banda: si todos los elementos son ceros, excepto $a_{i,i+c_k}$ para algún número pequeño de enteros c_k. Los valores no nulos son confinados en un entorno de la diagonal principal, formando una banda de valores no nulos que completan la diagonal principal de la matriz y más diagonales en cada uno de sus costados.

Operadores

1. Producto de una matriz A por un escalar c:

$$cA = (ca_{ij})$$

2. Transpuesta de una matriz: Si $A = (a_{ij})$, entonces $A^T = (a_{ji})$.
3. Conjugada transpuesta (Adjunta): si $A = (a_{ij})$ entonces $A^H = (\bar{a}_{ji})$. (Si los elementos de A son del campo de los números complejos).

Teorema

Si (y solo si) A es simétrica, $A = A^T$, si (y solo si) A es simétrica sesgada, $A^T = -A$; y si (y solo si) A es hermitiana, $A = A^H$.

1. Una matriz diagonal cuadrada se puede especificar mediante la función constructora $\mathrm{diag}(\cdot)$ que opera en un vector y forma una matriz diagonal con los elementos del vector a lo largo de la diagonal:

$$\mathrm{diag}((d_1, d_2, \ldots, d_n)) = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & d_n \end{pmatrix} \tag{3}$$

2. Si A es una matriz de $n \times m$ y $k = \min(n, m)$

$$\mathrm{vecdiag}(A) = (a_{11}, \ldots, a_{kk})$$

3. $\mathrm{Vec}(A) = (a_1^T, a_2^T, \ldots, a_m^T)$, donde $a_1, a_2, \ldots, a_m$ son los vectores columnas de la matriz A.

4. Para una matriz simétrica A con elementos a_{ij},

$$\mathrm{vech}(A) = (a_{11}, a_{21}, \ldots, a_{m1}, a_{22}, \ldots, a_{m2}, \ldots, a_{mm})$$

Sea A una matriz particionada tal que:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{4}$$

donde A_{11} y A_{12} tienen el mismo número de filas, A_{21} y A_{22} tienen el mismo número de filas, A_{11} y A_{21} tienen el mismo número de columnas y A_{12} y A_{22} tienen el mismo número de columnas.

$\operatorname{diag}(A_1, A_2, \ldots, A_k) = A_1 \oplus A_2 \oplus \ldots A_k$ denota una matriz diagonal de bloque con submatrices $A_1, A_2, \ldots, A_k$ a lo largo de la diagonal y ceros en otro lugar, donde $\oplus$ se denomina suma directa.

La transpuesta de una matriz particionada está dada por:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^T = \begin{pmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{pmatrix} \tag{5}$$

Adición de matrices

La adición de dos matrices es la matriz cuyos elementos son las sumas de los elementos correspondientes de los sumandos.

- Si todos los elementos de la matriz A son ceros, entonces A se denomina Matriz aditiva identidad.
- Si todos los elementos de A son positivos, entonces se escribe $A > 0$.
- Si todos los elementos son no negativos, entonces se escribe $A \geq 0$.
- $(A + B)^T = A^T + B^T$.
- La suma de dos matrices simétricas es también una matriz simétrica.

El **espacio vectorial** de matrices $n \times m$ es cualquier conjunto cerrado con respecto a la multiplicación por un escalar y a la adición matricial.

Operadores de valor escalar en matrices cuadradas

Hay varias asignaciones útiles de matrices a números reales, esto es, de $\mathbb{R}^{n\times n}$ a $\mathbb{R}$.

1. $\mathrm{Tr}(.) = \sum_i a_{ii}$.
2. $\mathrm{Tra}(A) = \mathrm{Tra}(A^T)$ si A es una matriz cuadrada.
3. Para un escalar c y A una matriz de $n \times n$, $\mathrm{Tra}(cA) = c\mathrm{Tra}(A)$.
4. Si
$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}, \end{pmatrix} \tag{6}$$
donde A_{11} y A_{22} son matrices cuadradas, entonces
$$\mathrm{Tr}(A) = \mathrm{Tr}(A_{11}) + \mathrm{Tr}(A_{22}).$$
5. Si A y B son matrices del mismo orden, entonces
$$\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$$

El determinante

Sea A una matriz de $n \times n$, considere el producto $a_{1j_1} a_{2j_2 nj_1}$, donde $\pi_j = (j_1, j_2, \ldots, j_n)$ es una de las $n!$ permutaciones de los enteros de 1 hasta n. Una permutación **par** es una permutación que puede ser representada por un número par de transposiciones. Una permutación **impar** es una permutación que puede ser representada por un número impar de transposiciones. Ejemplo: de las $6 = 3!$ permutaciones del conjunto $1, 2, 3$, escritas en notación de ciclos:

- $(12), (23)$ y (13) son impares.
- (123) y (132) son pares.

Sea $\sigma(\pi_j) = 1$ si $\pi_j = (j_1, j_2, \ldots, j_n)$ es una permutación par, y $\sigma(\pi_j) = -1$ de otro modo. Entonces el **determinante** de A, denotado por $|A|$, está definido por

$$|A| = \sum_{todas \quad las \quad permutaciones} \sigma(\pi_j) a_{1j_1} \ldots a_{nj_n}$$

Menores, cofactores y matriz conjugada

Sea A una matriz de 2×2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{7}$$

Sabemos que $|A| = a_{11}a_{22} + (-1)a_{21}a_{12}$. Ahora sea A una matriz de 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{8}$$

Podemos definir $|A|$ en términos de submatrices:

$$|A| = a_{11}(-1)^{1+1}\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} + a_{12}(-1)^{1+2}\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}(-1)^{1+3}\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \tag{9}$$

Sea A una matriz de $n \times n$, se denota el menor asociado con el elemento a_{ij} como $|A_{-(i)(j)}|$, donde $A_{-(i)(j)}$ es la matriz formada de A removiendo la $i-$ésima fila y la $j-$ésima columna.
Denotamos el cofactor de a_{ij} como $a_{(ij)}$:

$$a_{(ij)} = (-1)^{i+j}|A_{-(i)(j)}|$$

Entonces podemos definir el determinante de A:

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{-(i)(j)}| = \sum_{j=1}^{n} a_{ij}a_{(ij)}$$

Esta expresión se denomina **Expansión de Laplace**.
La **adjunta** de A $(n \times n)$ está definida como

$$\text{adj}(A) = (a_{(ij)})$$

La adjunta de una matriz A es la traspuesta de la matriz cofactor de A.

Propiedades del determinante

1. $|A| = |A^T|$.
2. Para un escalar c, $|cA| = c^n|A|$.
3. Si A $(n \times n)$ es una matriz triangular superior (inferior), entonces

$$|A| = \prod_{i=1}^{n} a_{ii}$$

4.

$$|A| = \begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix} = \begin{vmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{vmatrix} = |A_{11}||A_{22}| \tag{10}$$

5. $|A + B| \neq |A| + |B|$

Multiplicación de matrices

Multiplicación de matrices (Cayley): Si el número de columnas de la matriz A, con elementos a_{ij}, y el número de filas de la matriz B, con elementos b_{ij}, son iguales, entonces el producto de A por B está definido como la matriz C con elementos $c_{ij} = \sum_k a_{ik} b_{kj}$.

$$C = (A \quad B)_{n \times p} = ()_{n \times m} \, ()_{m \times p} \tag{11}$$

Propiedades de la multiplicación de matrices

1. $B^T A^T = (AB)^T$
2. $A(BC) = (AB)C$
3. $AB \neq BA$
4. $A(B + C) = AB + AC$ y $(B + C)A = BA + CA$
5. Matrices particionadas

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{21}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \tag{12}$$

Operaciones elementales en matrices

- Intercambio de dos filas (columnas).
- La multiplicación de un escalar por una fila (columna) dada.
- El reemplazo de una fila dada (columna) por la suma de esa fila (columnas) y un múltiplo escalar de otra fila (columna).

La Vec-permutación

Si A $(n \times n)$, la matriz K_{nm} ($nm \times nm$ es tal que

$$\mathrm{vec}(A^T) = K_{nm}\mathrm{vec}(A)$$

La matriz K_{nm} es llamada matriz vec-permutación.

- $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$
- $|AB| = |A||B|$

- Si representamos los vectores formados por las columnas de la matriz A $(n \times n)$ como $a_1, a_2, \ldots, a_m$, el matriz/vector producto Ax es la combinación lineal de las columnas de A:

$$Ax = \sum_{i=1}^{m} x_i a_i$$

  donde cada x_i es un escalar y cada a_i un vector.
- El producto externo de los vectores x y y es la matriz xy^T.
- $x^T Ay$ es llamada forma bilineal.
- $x^T Ax$ es una forma cuadrática.
- si A es simétrica y $x^T Ax \geq 0$, entonces A es una matriz definida no negativa.
- si A es simétrica, $x \neq 0$ y $x^T Ax > 0$, entonces A es una matriz definida positiva.

Otros tipos de producto

- Producto Hadamard: La multiplicación Hadamard se define para matrices de la misma forma como la multiplicación de cada elemento de una matriz por el elemento correspondiente de la otra matriz.
- Producto Kronecker: Sean $A\ (n \times m)$ y $B\ (p \times q)$, entonces

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \vdots & \dots & \vdots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix} \tag{13}$$

Algunas propiedades

1. $|A \otimes B| = |A|^m |B|^n$
2. $(aA) \otimes (bB) = ab(A \otimes B)$
3. $(A + B) \otimes C = A \otimes C + B \otimes C$
4. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
5. $(A \otimes B)^T = A^T \otimes B^T$
6. $(A \otimes B)(C \otimes D) = AC \otimes BC$
7. $Tr(A \otimes B) = Tr(A)Tr(B)$

En cuatro semanas, dos compañías,A y B, necesitan las siguientes cantidades de materia prima de levadura, malta y agua:

- Semana1 :
  A : 8 Barriles levadura, 4 Barriles malta, 12 Barriles agua.
  B : 6 Barriles levadura, 3 Barriles malta, 12 Barriles agua.
- Semana2 :
  A : 10 Barriles levadura, 6 Barriles malta, 5 Barriles agua.
  B : 9 Barriles levadura, 5 Barriles malta, 4 Barriles agua.
- Semana3 :
  A : 7 Barriles levadura, 8 Barriles malta, 5 Barriles agua.
  B : 7 Barriles levadura, 0 Barriles malta, 5 Barriles agua.
- Semana4 :
  A : 11 Barriles levadura, 7 Barriles malta, 9 Barriles agua.
  B : 11 Barriles levadura, 6 Barriles malta, 5 Barriles agua.

¿Qué cantidad de materia prima utilizó cada compañía en el mes?.
¿Cuál es la diferencia de consumo de ambas compañías en las dos primeras semanas?.
¿Cuál es el consumo de materia prima por semana para cinco compañías como A?