TEMA 1: Cálculo Matricial Avanzado. Pseudoinversa de Moore-Penrose

Álgebra Lineal Numérica para Ciencia de Datos

Máster Universitario en Estadística Computacional y
Ciencia de Datos para la Toma de Decisiones
Curso 2022-2023
Prof. María Victoria Herranz

Índice

Introducción

Recordemos que una matriz cuadrada $A \in \mathbb{R}^{n\times n}$ es invertible si existe una matriz cuadrada $B \in \mathbb{R}^{n\times n}$ tal que $AB = BA = I$. En este caso, B es la matriz inversa de A y se denota por $B = A^{-1}$.
También sabemos que que dos formas equivalentes para caracterizar una matriz cuadrada invertible:

- A tiene rango completo, es decir $\operatorname{rg} A = n$.
- El determinante de A es distinto de cero, $\det(A) \neq 0$.

Introducción

Sea A una matriz de orden $m \times n$ cualquiera. Dado $b \in \mathbb{R}^n$, nos planteamos la resolución del sistema de ecuaciones lineales

$$Ax = b \qquad \text{para } x \in \mathbb{R}^m \tag{1}$$

Si A es una matriz cuadrada y A tiene inversa, entonces $Ax = b$ se cumple si y solo si $x = A^{-1}b$. Sin embargo, A puede ser de orden $m \times n$ con $m \neq n$, o A puede ser una matriz cuadrada que no es invertible. Si A no es invertible, entonces la ecuación (1) puede no tener soluciones y si hay soluciones, entonces puede haber incluso infinitas soluciones.

Introducción

Por tanto, dada $A \in \mathbb{R}^{m \times n}$, nos planteamos las siguientes cuestiones:

- ¿Existe algún tipo de inversa de A?
- Dado un vector $b \in \mathbb{R}^m$, ¿cómo podemos resolver el sistema de ecuaciones lineales $Ax = b$?

Introducción

En muchas aplicaciones, como la regresión lineal múltiple, surge el problema de los mínimos cuadrados:

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|^2 \qquad \text{donde } A \in \mathbb{R}^{m\times n},\ b\in\mathbb{R}^m \text{ son fijos}$$

Si A tiene rango de columna completo (es decir, $\mathrm{rg}(A) = n \leq m$), entonces el problema tiene una única solución

$$x^* = (A^TA)^{-1}A^Tb$$

Nuestro objetivo es analizar las propiedades de las siguientes matrices:

- $(A^TA)^{-1}A^T$ (pseudoinversa): Solución óptima $x^* = (A^TA)^{-1}A^Tb$
- $A(A^TA)^{-1}A^T$ (matriz de proyección): la aproximación más cercana a b es $Ax^* = A(A^TA)^{-1}A^Tb$.

Índice

Matrices escalonadas y escalonadas reducidas

Antes de dar la definición de inversas generalizada de una matriz, repasemos conceptos y resultados que serán útiles en el desarrollo del resto del capítulo.

Matriz escalonada

Una matriz se llama escalonada por filas o simplemente escalonada si cumple con las siguientes propiedades:

1. Todas las filas cero están en la parte inferior de la matriz (es decir, no puede seguirse una fila distina de cero después de una cero).
2. El término principal de una fila no-cero está estrictamente a la derecha del término principal de la fila de encima.

Notación matricial de los sistemas de ecuaciones lineales

Definición

Dado un sistema lineal de ecuaciones

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m, \end{array}\right.$$

Llamaremos matriz del sistema a la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Notación matricial de los sistemas de ecuaciones lineales

Definición

Llamaremos matriz ampliada del sistema a la matriz $m \times (n+1)$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix},$$

que incorpora los términos independientes. También las representaremos en la forma breve $(A|B)$

Notación matricial de los sistemas de ecuaciones lineales

Ejemplo

Consideremos el sistema

$$\begin{cases} x + y + 2z = 9, \\ 3x + 6y - 5z = 0, \\ 2x + 4y - 3z = 1. \end{cases}$$

Entonces la matriz del sistema y la matriz ampliada son:

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 6 & -5 \\ 2 & 4 & -3 \end{pmatrix}, \quad (A|B) = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right).$$

Esta notación más compacta simplificará la escritura e implementación del algoritmo de eliminación de Gauss.

Transformaciones elementales

Cada fila de la matriz del sistema corresponde a una de las ecuaciones del sistema. Las operaciones del método de Gauss se traducirán entonces en operaciones sobre las filas de la matriz. En particular, las transformaciones elementales serán, en este contexto, las siguentes:

1. Multiplicar una fila por un número $\alpha \neq 0$.
2. Intercambiar de lugar dos filas.
3. Sumar a una fila el resultado de multiplicar otra por un número cualquiera.

Matriz escalonada

Definición

Llamaremos pivote a la primera entrada no nula de cada fila de una matriz.

Definición

Llamaremos matriz escalonada a una matriz que cumpla las siguientes condiciones:

1. *todas las filas, salvo quizás la primera, comienzan con una sucesión de ceros;*
2. *cada fila tiene al principio por lo menos un cero más que la fila inmediata superior.*

Escalonar una matriz es llevarla a una forma escalonada por medio de transformaciones elementales Si A' es una matriz que se obtiene escalonando otra matriz A, entonces diremos que A' es una forma escalonada de A. Diremos que un *sistema está escalonado* si su matriz ampliada lo está. *Escalonar un sistema* es encontrar otro sistema escalonado equivalente. Naturalmente, escalonar un sistema es equivalente a escalonar la matriz ampliada del sistema.

Método de eliminación de GAuss

Ejemplo

Resolver el sistema:

$$\left\{ \begin{array}{rcr} x_1 + 2x_2 + x_3 + x_5 + x_6 & = & 1 \\ -x_1 - 2x_2 + x_3 + x_4 - x_6 & = & 0 \\ x_1 + 2x_2 + x_3 + 2x_4 + 5x_5 + 3x_6 & = & 1 \\ x_1 + 2x_2 + x_3 + x_5 + 3x_6 & = & 3 \\ x_1 + 2x_2 + x_3 + 4x_4 + 9x_5 + 3x_6 & = & -1 \end{array} \right.$$

La matriz ampliada del sistema es:

$$\left(\begin{array}{rrrrrr|r} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ -1 & -2 & 1 & 1 & 0 & -1 & 0 \\ 1 & 2 & 1 & 2 & 5 & 3 & 1 \\ 1 & 2 & 1 & 0 & 1 & 3 & 3 \\ 1 & 2 & 1 & 4 & 9 & 3 & -1 \end{array} \right).$$

Método de eliminación de GAuss

El proceso comienza fijando la entrada no nula de primera fila como pivote y utilizándola para lograr ceros en el resto de la primera columna. El resultado es la matriz

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 4 & 8 & 2 & -2 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 + F_1 \\ \longleftarrow F_3 - F_1 \\ \longleftarrow F_4 - F_1 \\ \longleftarrow F_5 - F_1. \end{array}$$

Método de eliminación de Gauss

Se debería utilizar la entrada a_{22} en el segundo paso del método para conseguir ceros en la segunda columna. Aquí esto no es posible pues la segunda columna ya tiene todas sus entradas, salvo la primera, nulas. En consecuencia, el segundo escalón deberá estar en la tercera columna, donde aparece la entrada no nula a_{23}. Por debajo de a_{23} sólo hay ceros, así que continuamos nuestro algoritmo usando el pivote 2 de la entrada a_{34}. He aquí la matriz que se obtiene, junto con la indicación de las operaciones realizadas:

$$\left(\begin{array}{cccccr|r} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 & -2 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 - 2F_3 \end{array}$$

Método de eliminación de Gauss

En el tercer paso operaremos sobre la sexta columna: Realizando transformaciones elementales obtenemos:

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \longleftarrow F_5 + F_4. \end{array}$$

Método de eliminación de Gauss

Ya tenemos la forma escalonada, con pivotes en las columnas primera, tercera, cuarta y sexta. El sistema lineal, equivalente al sistema original en el sentido de que tiene exactamente las mismas soluciones, que corresponde a esta matriz es

$$\begin{array}{rcrcrcrcrcrcl} x_1 & + & 2x_2 & + & x_3 & & & + & x_5 & + & x_6 & = & 1, \\ & & & & 2x_3 & + & x_4 & + & x_5 & & & = & 1, \\ & & & & & & 2x_4 & + & 4x_5 & + & 2x_6 & = & 0, \\ & & & & & & & & & & 2x_6 & = & 2. \end{array}$$

Método de eliminación de Gauss

Definición

Se denominan variables o incógnitas libres aquellas incógnitas que corresponden a columnas sin pivotes.

Las variables donde están los pivotes siempre se pueden despejar en función de las variables libres. Consideremos el sistema del ejemplo y asignemos los parámetros α y β a las variables libres x_2 y x_5, respectivamente. Concluimos entonces que las soluciones son

$$x_1 = -1 - 2\alpha - \frac{3}{2}\beta, \quad x_2 = \alpha, \quad x_3 = 1 + \frac{1}{2}\beta,$$
$$x_4 = -1 - 2\beta, \quad x_5 = \beta, \quad x_6 = 1.$$

donde α y β son dos parámetros reales que podemos fijar a nuestro antojo. La solución no es única, pero puede describirse completamente en términos de los parámetros α y β.

Forma normal de Hermite

Definición

Dada una matriz $A \in \mathbb{R}^{m \times n}$, si le aplicamos el método de Gauss-Jordan de modo que la matriz escalonada resultante tenga todas las entradas de las columnas correspondientes a los pivotes nulas, salvo los pivotes que son 1 , a esta matriz le llamamos Forma Normal de Hermite por filas o forma reducida por filas.

Forma normal de Hermite

Por tanto, una matriz se llama escalonada reducida por filas o es la forma normal de Hermite por filas de otra si cumple con las siguientes propiedades:

1. Todas las filas cero están en la parte inferior de la matriz (es decir, no puede seguirse una fila distina de cero después de una cero).
2. El pivote (primer elemento no nulo de cada fila) de cada fila es igual a 1.
3. El pivote de cada fila no nula está a la derecha del pivote de la fila anterior.
4. Los elementos que aparecen en la misma columna que el pivote de una fila, son todos cero. Las columnas correspondientes a los pivotes tienen todas las entradas nulas, salvo el pivote.

Forma normal de Hermite

Ejemplo

Consideremos de nuevo el sistema

$$\left\{\begin{array}{rcr} x_1 + 2x_2 + x_3 + x_5 + x_6 & = & 1 \\ -x_1 - 2x_2 + x_3 + x_4 - x_6 & = & 0 \\ x_1 + 2x_2 + x_3 + 2x_4 + 5x_5 + 3x_6 & = & 1 \\ x_1 + 2x_2 + x_3 + x_5 + 3x_6 & = & 3 \\ x_1 + 2x_2 + x_3 + 4x_4 + 9x_5 + 3x_6 & = & -1 \end{array}\right.$$

Retomamos el sistema del ejemplo anterior en el punto en que habíamos escalonado la matriz del sistema.

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

era la matriz asociada al sistema escalonado. Usaremos ahora los pivotes para hacer aparecer ceros por encima de ellos. Como paso previo dividimos las filas tercera y cuarta por sus pivotes,

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Forma normal de Hermite

Luego comenzamos por hacer aparecer ceros sobre el pivote de la última columna, y obtenemos

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - F_4, \\ \\ \longleftarrow F_3 - F_4. \\ \\ \\ \end{array}$$

A partir de esta nueva matriz operamos con el pivote de la cuarta columna:

$$\left(\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \longleftarrow F_2 - F_3, \\ \\ \\ \\ \end{array}$$

Forma normal de Hermite

Ahora lo hacemos con el de la segunda columna, y dividimos la segunda fila entre 2 para que su pivote quede igual a 1. El resultado final es

$$\left(\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & \frac{3}{2} & 0 & -1 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \longleftarrow F_1 - \frac{1}{2}F_2, \\ \longleftarrow \frac{1}{2}F_2, \\ \\ \\ \\ \end{array} \tag{2}$$

donde hemos enfatizado además la forma escalonada que tiene la matriz.

Forma normal de Hermite

El sistema de ecuaciones asociado con esta matriz es

$$\left\{\begin{array}{rcrcrcrcrcrcl} x_1 & + & 2x_2 & & & & & + & \frac{3}{2}x_5 & + & & = & -1, \\ & & & & x_3 & - & & & \frac{1}{2}x_5 & & & = & 1, \\ & & & & & & x_4 & + & 2x_5 & & & = & -1, \\ & & & & & & & & & & x_6 & = & 1, \end{array}\right.$$

y ahora es completamente directo despejar x_1, x_3 y x_4 en términos de las variables $x_2 = \alpha$ y $x_5 = \beta$ para reencontrar el conjunto de soluciones en forma paramétrica

$$\left\{\left(-1-2\alpha-\frac{3}{2}\beta, \alpha, 1+\frac{1}{2}\beta, -1-2\beta, \beta, 1\right)\ ;\ \alpha,\ \beta \in \mathbb{R}\right\}.$$

Forma normal de Hermite

Las matrices escalonadas por columnas se definen igual. De hecho, una matriz A es escalonada por columnas si su traspuesta, A^T es escalonada por filas.

Ejercicio

Reducir la matriz $\begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 8 & -3 & 1 \\ -3 & 1 & 0 & -3 \\ 0 & 4 & -1 & 1 \end{pmatrix}$ a una matriz escalonada por filas, otra por filas reducida y otra reducida por columnas.

Método de eliminación de Gauss

Hasta este momento hemos usado el algoritmo de Gauss, y resuelto unos cuantos sistemas de ecuaciones con él. El lector podría preguntarse si cualquier matriz puede ser llevada a una forma escalonada. La respuesta es que sí. Y la demostración no pasa de ser un análisis cuidadoso del método. Todo esto está contenido en el siguiente resultado teórico.

Proposición

Toda matriz puede ser transformada en una matriz escalonada mediante una cantidad finita de transformaciones elementales. En consecuencia, todo sistema lineal es equivalente a uno escalonado.

Matrices elementales

Sea e una de las transformaciones elementales, esto es,

1. multiplicar una fila (o columna) por una constante no nula.
2. intercambiar dos filas (o columnas),
3. sumarle a una fila (o a una columna) un múltiplo de otra,

Definición

La matriz elemental E asociada con la transformación e es la que se obtiene al aplicar la transformación e a la matriz identidad $n \times n$.

Matrices elementales

La matriz

$$E_1 = \begin{array}{cc} \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 & \alpha & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} & \begin{array}{l} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \longleftarrow \text{Fila i} \\ \vdots \\ \vdots \\ \vdots \end{array} \end{array} \tag{3}$$

está asociada a la transformación elemental e_1 **"multiplicar la fila i por la constante $\alpha \neq 0$"**, pues E_1 se obtiene de la matriz identidad I aplicando dicha transformación

$$I \quad \xrightarrow{\text{transf.elemental } e_1} \quad E_1$$

Matrices elementales

La matriz

$$E_2 = \begin{array}{cl} \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \cdots & \cdots & 0 & \cdots & \cdots & \cdots & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & 1 & \cdots & \cdots & \vdots \\ \vdots & \vdots & 1 & \cdots & \cdots & \cdots & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix} & \begin{array}{l} \vdots \\ \vdots \\ \longleftarrow \text{Fila j} \\ \vdots \\ \vdots \\ \vdots \\ \longleftarrow \text{Fila i} \\ \vdots \\ \vdots \end{array} \end{array} \tag{4}$$

está asociada a la transformación elemental e_2 **"intercambiar la fila i por la fila j"**, pues E_2 se obtiene de la matriz identidad I aplicando dicha transformación.

$$I \quad \xrightarrow{\text{transf.elemental } e_2} \quad E_2$$

Matrices elementales

La matriz

$$E_3 = \begin{array}{c}\left(\begin{array}{ccccccccc}
1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
\cdots & \cdots & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \cdots & 1 & \cdots & \cdots & \vdots \\
\vdots & \vdots & \alpha & \cdots & \cdots & \cdots & 1 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1
\end{array}\right)\end{array}
\begin{array}{l}
\vdots \\
\vdots \\
\longleftarrow \text{Fila j} \\
\vdots \\
\vdots \\
\vdots \\
\longleftarrow \text{Fila i} \\
\vdots \\
\vdots
\end{array} \tag{5}$$

está asociada a la transformación elemental e_3 **"sumarle a fila *i* un múltiplo de la fila *j*"**, pues E_3 se obtiene de la matriz identidad I aplicando dicha transformación

$$I \quad \xrightarrow{\text{transf.elemental } e_3} \quad E_3$$

Matrices elementales

Proposición

Sea A' la matriz (de tamaño $m \times n$) que se obtiene a partir de la matriz A (de tamaño $m \times m$) realizando la transformación elemental e con las filas de A

$$A \quad \overset{\text{transf.elemental } e}{\longrightarrow} \quad A';$$

y E la matriz elemental (de tamaño $m \times m$) asociada con la transformación e

$$I \quad \overset{\text{transf.elemental } e}{\longrightarrow} \quad E$$

Entonces se cumple que $\quad A' = EA.$

Matrices elementales

Ejemplo

Consideremos la matriz

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix}, \text{ y la matriz elemental } E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

asociada a la transformación elemental e de sumarle a la tercera fila 3 veces la primer fila. Sea A' la matriz que se obtiene de A si se le aplica la transformación elemental considerada

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \xrightarrow{\text{transf.elemental } e} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 3 \end{pmatrix} = A'$$

$$\text{y operando se verifica que } A' = EA = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 3 \end{pmatrix}$$

Matrices elementales

El último resultado es también cierto si *intercambiamos transformaciones elementales en filas por transformaciones elementales en columnas*. La única diferencia es que los productos por la izquierda de las matrices elementales pasan a productos por la derecha, y se cumpliría la relación $A' = AE$, donde $E \in \mathbb{R}^{m\times m}$.

Proposición

Toda matriz A puede reducirse a su forma escalonada A' mediante el producto a izquierda de una sucesión de matrices elementales $E_1, E_2, \ldots, E_k$, esto es:

$$E_k \ldots E_2 E_1 A = A'.$$

La matriz $P = E_k \ldots E_2 E_1$ se denomina matriz de paso. La matriz P de paso se obtiene directamente realizando a la matriz identidad I_n todas las transformaciones elementales por filas que realizamos a la matriz A para obtener la matriz A'.

Matrices elementales

El último resultado es también cierto si *intercambiamos transformaciones elementales en filas por transformaciones elementales en columnas*. La única diferencia es que los productos por la izquierda de las matrices elementales pasan a productos por la derecha, y se cumpliría la relación $A' = AE$, donde $E \in \mathbb{R}^{m \times m}$.

Matrices elementales

La matriz $P = E_k \dots E_2 E_1$ se denomina matriz de paso. La matriz P de paso se obtiene directamente realizando a la matriz identidad I_n todas las transformaciones elementales por filas que realizamos a la matriz A para obtener la matriz A'.

Ejemplo

Se considera la matriz

$$A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 1 & -2 & 2 \\ 3 & 2 & 2 & 0 \end{pmatrix}$$

Utilizando únicamente transformaciones elementales por filas a la matriz A, calcular una matriz A' triangular y hallar la matriz de paso P tal que $A' = PA$.

Matrices escalonadas y escalonadas reducidas

Ejercicio

Halla la forma normal de Hermite por filas de las matrices:

$$A = \begin{pmatrix} -2 & -4 & 2 & 2 \\ 3 & 6 & -3 & -3 \\ 1 & 2 & 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 6 & -5 & 0 \\ 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 3 & 1 & 1 \end{pmatrix}$$

Índice

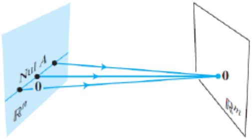
Subespacios de una matriz

En esta sección mencionaremos algunas notaciones básicas de la teoría de matrices y presentaremos sin demostración algunos resultados que involucran los espacios imagen y nulo de una matriz.

Definición

Sea $A \in \mathbb{R}^{m\times n}$.

- *Se llama espacio nulo o núcleo de A al conjunto* $ker(A) = \mathcal{N}(A) = \{x \in \mathbb{R}^n \;/\; Ax = 0\}$



- *Se llama espacio imagen de A al conjunto*

$$\mathrm{Im}(A) = \mathcal{R}(A) = \{y \in \mathbb{R}^m \;/\; y = Ax \text{ para cierto } x \in \mathbb{R}^n\}$$

Subespacios de una matriz

A partir de las definiciones anteriores, no es difícil probar que:

$$\operatorname{Im}(A^T) = \ker(A)^\perp \qquad \text{y} \qquad \ker(A^T) = \operatorname{Im}(A)^\perp$$

Más aún, algunas de las principales propiedades de estos subespacios son enumeradas en el siguiente teorema.

Teorema

Sea $A \in \mathbb{R}^{m\times n}$. Entonces se cumplen las siguiente propiedades:

1. $\ker(A^T A) = \ker(A)$.
2. $\ker(AA^T) = \ker(A^T)$.
3. $\operatorname{Im}(A^T A) = \operatorname{Im}(A^T)$.
4. $\operatorname{Im}(AA^T) = \operatorname{Im}(A)$.
5. $\operatorname{rg}(AA^T) = \operatorname{rg}(A) = \operatorname{rg}(A^T) = \operatorname{rg}(A^T A)$.

Subespacios de una matriz

Ejercicio

Calcula el núcleo y el espacio imagen de las siguientes matrices:

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ -1 & 1 & 1 & 1 \\ 0 & -2 & 0 & 1 \\ 4 & 0 & 3 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 2 & 0 & 1 \\ 2 & -4 & -1 & 1 \\ 3 & -6 & -2 & 3 \end{pmatrix}$$

Subespacios de una matriz

Otras propiedades útiles que poseen los subespacios imagen y nulo son las siguientes:

Teorema

Sean $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{p\times q}$. Entonces se cumplen las siguiente propiedades:

1. Si $n = p$, entonces $\mathrm{Im}(AB) \subset \mathrm{Im}(A)$
2. Si $m = p$, se cumple que $\mathrm{Im}(A) \subseteq \mathrm{Im}(B)$ si y sólo si existe una matriz $C \in \mathbb{R}^{q\times n}$ tal que $A = BC$.
3. Si $n = p$, entonces $\ker(B) = \ker(AB)$.
4. Si $n = p$, entonces $\mathrm{rg}(AB) \leq \min\{\mathrm{rg}(A), \mathrm{rg}(B)\}$

Subespacios de una matriz

Definición

Sea $A \in \mathbb{R}^{m\times n}$*.*

- *El espacio columna de A es el subespacio de $\mathbb{R}^m$ generado por sus columnas. Es decir, es el subespacio imagen de A*

$$colspace(A) = \operatorname{Im} A = \{Ax \ / \ x \in \mathbb{R}^n\}$$

- *El espacio fila de A es el subespacio de $\mathbb{R}^n$ generado por sus filas. Es decir, es el subespacio*

$$rowspace(A) = \{y^T A / \ y \in \mathbb{R}^m\}$$

Subespacios de una matriz

Sea $A \in \mathbb{R}^{m \times n}$.

- El conjunto de vectores columna correspondientes a aquellas columnas que contienen los pivotes en cualquier forma escalonada de la matriz A forman una base del espacio columna de A.
- El conjunto de vectores fila correspondientes a aquellas filas que contienen los pivotes en cualquier forma escalonada de la matriz A forman una base del espacio fila de A.
- El espacio columna de A es el subespacio de $\mathbb{R}^m$ que contiene a los vectores $v \in \mathbb{R}^m$ para los cuales el sistema $Ax = v$ es compatible.

Subespacios de una matriz

Si A tiene rango por columnas r, entonces:

- Cualesquiera r columnas linealmente independientes de A forman una base de $\mathrm{Im}(A)$.
- Cualquier conjunto maximal de columnas linealmente independientes de A contiene exactamente r vectores.
- Claramente, el espacio columna de A es igual al espacio fila de A^T.
- Las columnas pivote de A constituyen una base de $\mathrm{Im}(A)$.
- $\dim(\mathrm{im}(A))$ es el número de columnas pivote de A.

Subespacios de una matriz

Ejercicio

Determina una base del espacio fila y del espacio columna de las matrices

$$A = \begin{pmatrix} 1 & -1 & 1 & 3 & 2 \\ 2 & -1 & 1 & 5 & 1 \\ 3 & -1 & 1 & 7 & 0 \\ 0 & 1 & -1 & -1 & -3 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & -1 & -2 & 0 \\ 2 & 4 & -1 & 1 & 0 \\ 3 & 6 & -1 & 4 & 1 \\ 0 & 0 & 1 & 5 & 0 \end{pmatrix}$$

Subespacios de una matriz

Contrast Between Nul *A* and Col *A* for an *m* x *n* Matrix *A*

Nul A	Col A
1. Nul A is a subspace of $\mathbb{R}^n$.	**1.** Col A is a subspace of $\mathbb{R}^m$.
2. Nul A is implicitly defined; that is, you are given only a condition ($A\mathbf{x} = \mathbf{0}$) that vectors in Nul A must satisfy.	**2.** Col A is explicitly defined; that is, you are told how to build vectors in Col A.
3. It takes time to find vectors in Nul A. Row operations on $[\,A \quad \mathbf{0}\,]$ are required.	**3.** It is easy to find vectors in Col A. The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A.	**4.** There is an obvious relation between Col A and the entries in A, since each column of A is in Col A.
5. A typical vector $\mathbf{v}$ in Nul A has the property that $A\mathbf{v} = \mathbf{0}$.	**5.** A typical vector $\mathbf{v}$ in Col A has the property that the equation $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector $\mathbf{v}$, it is easy to tell if $\mathbf{v}$ is in Nul A. Just compute $A\mathbf{v}$.	**6.** Given a specific vector $\mathbf{v}$, it may take time to tell if $\mathbf{v}$ is in Col A. Row operations on $[\,A \quad \mathbf{v}\,]$ are required.
7. Nul $A = \{\mathbf{0}\}$ if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.	**7.** Col $A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.	**8.** Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$.

Subespacios de una matriz y sistemas de ecuaciones lineales

Consideremos el sistema de ecuaciones lineales

$$Ax = b$$

- Si $b = 0$, el sistema se llama homogéneo. En este caso, el conjunto solución es simplemente el espacio nulo de A, o núcleo de A.
- Cualquier sistema homogéneo tiene la solución $x = 0$, que se llama solución trivial. Geométricamente, esto significa que el conjunto solución pasa por el origen.
- Si $\text{rg}(A) = n$, entonces $Ax = 0$ tiene solo la solución trivial $x = 0$, entonces $\ker(A) = \{0\}$.
- Si $\text{rg}(A) = r < n$, entonces el sistema $Ax = 0$ tiene infinitas soluciones (obtenidas a partir de una base de $\ker(A)$)

Subespacios de una matriz y sistemas de ecuaciones lineales

Ahora considere un sistema lineal no homogéneo

$$Ax = b$$

donde A es una matriz $m \times n$ y b no es necesariamente 0.

- Si $b \notin \mathrm{Im}(A)$, entonces el sistema $Ax = b$ es incompatible.
- Si $b \in \mathrm{Im}(A)$, entonces el sistema $Ax = b$ es compatible.
  - Será compatible determinado, es decir, tendrá una única solución si y sólo si $\mathrm{rg}(A) = n$.
  - Será compatible indeterminado, es decir, tendrá infinitas soluciones si y sólo si $\mathrm{rg}(A) = r < n$.

Matrices de Rango Máximo por Filas o Columnas

Matrices de Rango Máximo por Filas

Diremos que una matriz tiene rango máximo por filas, si todas las filas son linealmente independientes, es decir, al hacer la reducción por filas no obtenemos ninguna fila nula.

Matrices de Rango Máximo por Columnas

Diremos que una matriz tiene rango máximo por columnas, si todas las columnas son linealmente independientes, es decir, al hacer la reducción por columnas no obtenemos ninguna columna nula.

La reducción por columnas la hacemos transponiendo, haciendo la reducción por filas y volviendo a transponer el resultado, por lo tanto queda claro que una matriz tiene rango máximo por columnas si y sólo si su transpuesta tiene rango máximo por filas, y viceversa.

Matrices de Rango Máximo por Filas o Columnas

- $\dim(\text{rowspace}(A)) = \text{rg}(A) = \dim(\text{Im}(A))$
- Por lo tanto, para saber si una matriz tiene rango máximo por filas y/o por columnas, lo que tenemos que hacer es calcular el rango de la matriz A, digamos que es r.
- Si r es igual al número de filas de la matriz, entonces tendrá rango máximo por filas.
- Si r es igual al número de columnas de la matriz, entonces tendrá rango máximo por columnas.
- Evidentemente, si la matriz no es cuadrada, no puede tener rango máximo por filas y columnas al mismo tiempo, y es fácil encontrar matrices cuadradas que no tienen rango máximo.

Inversas laterales

Definición

Dada una matriz $A \in \mathbb{R}^{m\times n}$,

- *Se llama inversa a la derecha de* A *a la matriz* $R \in \mathbb{R}^{n\times m}$ *que verifica* $AR = I_m$.
- *Se llama inversa a la izquierda de* A *a la matriz* $L \in \mathbb{R}^{n\times m}$ *que verifica* $LA = I_n$.

Ejemplo

Sea matriz $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$. La matriz inversa a la izquierda de A es $L = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ ya que $LA = I_2$

Caracterización de Existencia de Inversas Laterales

Teorema

Sea A una matriz. Entonces se verifica:

1. A tiene inversa por la izquierda si y sólo si tiene rango máximo por columnas.
2. A tiene inversa por la derecha si y sólo si tiene rango máximo por filas.
3. A tiene inversa si y sólo si tiene rango máximo por filas y columnas.
4. Si una matriz A rectangular tiene inversa por un lado no puede tener inversa por el otro. Las únicas matrices que pueden tener inversa por ambos lados son cuadradas y su inversa es única.

Cálculo de Inversas Laterales

- La matriz $A = \begin{pmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{pmatrix}$ tiene dos inversas por la izquierda diferentes:

$$B = \frac{1}{9}\begin{pmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{pmatrix} \quad \text{y} \quad C = \frac{1}{2}\begin{pmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{pmatrix}$$

- La matriz $A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}$ tiene rango 2, luego tiene rango máximo por filas y, en particular, tiene inversa por la derecha. Ahora bien, cualquier matriz de la forma $\begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ a & b \end{pmatrix}$ es inversa por la derecha de A, ya que

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}\begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ a & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  Observemos que la matriz A no tiene inversa por la izquierda, debido a que la última columna de A tiene todos los elementos 0.

Luego la inversa lateral por la izquierda o por la derecha, no es única, salvo el caso en el que la matriz sea invertible (en este caso, la inversa es única y coincide con las inversas laterales).

Cálculo de Inversas Laterales

Hay fórmulas simples para las inversas derecha e izquierda, si existen:

$$B = (A^T A)^{-1} A^T \qquad \text{y} \qquad C = A^T (AA^T)^{-1}$$

Efectivamente, $BA = I$ y $AB = I$.

Ahora bien, $A^T A$ y AA^T y no tienen por qué ser invertibles. Veremos que $A^T A$ tiene inversa si el rango de A es n y AA^T tiene inversa cuando el rango es m. Por lo tanto, las fórmulas tienen sentido exactamente cuando el rango es lo más grande posible.

Índice

Inversa generalizada

Definición

Sea A una matriz de orden $m \times n$. Entonces la matriz G es una inversa generalizada de A o $\{1\}$-inversa de A si es una matriz de orden $n \times m$ verificando

$$AGA = A$$

Nota

Si A es cuadrada e invertible, entonces tiene una y sólo una inversa generalizada, que coincide con la inversa ordinaria A^{-1}, ya que:

$$A^{-1}(AGA)A^{-1} = (A^{-1}A)G(AA^{-1}) = G$$

y como $AGA = A$, entonces $A^{-1}(AGA)A^{-1} = A^{-1}AA^{-1} = A^{-1}$
Luego $G = A^{-1}$, hecho que justifica el término *inversa generalizada*.

Inversa generalizada

Nota

Para una matriz general $A \in \mathbb{R}^{m \times n}$, su inversa generalizada siempre existe pero puede no ser única.

Por ejemplo, consideremos la matriz $A = \begin{pmatrix} 1 & 2 \end{pmatrix} \in \mathbb{R}^{1\times 2}$. Su inversa generalizada es una matriz $G = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2\times 1}$ tal que

$$\begin{pmatrix} 1 & 2 \end{pmatrix} = A = AGA = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = (x + 2y) \begin{pmatrix} 1 & 2 \end{pmatrix}$$

Por tanto, cualquier matriz $G = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2\times 1}$ tal que $x + 2y = 1$, es una inversa generalizada de A. Por ejemplo, $G = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $G = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ son inversas generalizadas de A.

Inversa generalizada

Ejercicio

Prueba que para la matriz $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 4 \end{pmatrix}$ dos inversas generalizadas son

$$G_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad G_2 = \begin{pmatrix} -42 & -1 \\ 5 & 3 \\ 2 & 2 \end{pmatrix}$$

Inversa generalizada

El siguiente teorema indica una forma de encontrar la inversa generalizada de cualquier matriz.

Teorema

Sea $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{R}^{m\times n}$ una matriz de rango r, siendo $A_{11} \in \mathbb{R}^{r\times r}$. Si A_{11} es invertible, entonces

$$G = \begin{pmatrix} A_{11}^{-1} & O \\ O & O \end{pmatrix} \in \mathbb{R}^{n\times m}$$

es una inversa generalizada de A.

Cualquier matriz $A \in \mathbb{R}`m \times n$ con rango r se puede reorganizar a través de permutaciones de fila y columna para tener la forma dividida anterior con una submatriz $r \times r$ invertible en la esquina superior izquierda. Este teorema esencialmente establece la existencia de una inversa generalizada para cualquier matriz.

Inversa generalizada

Ejemplo

Halla una inversa generalizada de la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

La matriz A no es invertible, puesto que $\mathrm{rg}(A) = 2$. Ahora bien, como la submatriz de orden 2 dada por $\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$ es invertible, entonces una inversa generalizada de A es

$$G = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} & 0 \\ \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Inversa generalizada

Podemos comprobar que

$$AGA = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} & 0 \\ \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Luego, efectivamente G es una pseudoinversa de A.

Inversa generalizada

Si A es una matriz $m \times n$ y de rango r , entonces sabemos que existen matrices cuadradas e invertibles P y Q tales que $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ Se puede comprobar que $G = Q \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P$ es una inversa generalizada de A. Esto demuestra que toda matriz posee alguna inversa generalizada (en general, no única).

Inversa generalizada

La inversa generalizada también puede emplearse para encontrar una solución a un sistema lineal compatible (es decir, existe al menos una solución).

Teorema

Consideremos el sistema de ecuaciones lineal $Ax = b$. Supongamos que $b \in \mathrm{Im}(A)$. Sea G la inversa generalizada de A. Entonces $x^* = Gb$ es una solución particular del sistema.

Demostración:
Multiplicando los dos miembros de la igualdad $Ax = b$ por AG, obtenemos

$$(AG)b = (AG)Ax = (AGA)x = Ax = b$$

Por tanto, $x^* = Gb$ es una solución del sistema $Ax = b$.

Inversa generalizada

Ejemplo

Halla una solución del sistema de ecuaciones lineal $Ax = b$, donde

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad b = \begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix}$$

Sabemos que una solución particular de este sistema es

$$x^* = Gb = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} & 0 \\ \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

Índice

Pseudoinversa de Moore-Penrose

Definición

Dada una matriz $A \in \mathbb{R}^{m\times n}$, llamaremos matriz inversa de Moore-Penrose o pseudoinversa de Moore-Penrose de A a una matriz de orden $n \times m$ (que notaremos por $A^{\dagger}$) que verifica

1. $AA^{\dagger}A = A$ $\quad$ ($\rightarrow A^{\dagger}$ es una inversa generalizada de A.)
2. $A^{\dagger}AA^{\dagger} = A^{\dagger}$ $\quad$ ($\rightarrow A$ es una inversa generalizada de $A^{\dagger}$)
3. $(AA^{\dagger})^T = AA^{\dagger}$ $\quad$ ($\rightarrow AA^{+}$ es simétrica)
4. $(A^{\dagger}A)^T = A^{\dagger}A$ $\quad$ ($\rightarrow A^{\dagger}A$ es simétrica)

Pseudoinversa de Moore-Penrose

- Si X es una matriz que satisface la Condición (1), se conoce como inversa generalizada de A o $\{1\}$-inversa; si X satisface las condiciones (1) y (2), se denomina inversa generalizada reflexiva o $\{1-2\}$ inversa. Solo cuando B satisface las 4 condiciones, se le llama pseudoinverso de A.
- Si A es regular, entonces A^{-1} es la pseudoinversa de A.
- Si $A = O_{m\times n}$, entonces la matriz $M = O_{n\times m}$ es la pseudoinversa de A.
- Se puede demostrar que para cualquier matriz $A \in \mathbb{R}^{m\times n}$ la pseudoinversa siempre existe y es única (como veremos a continuación).

Pseudoinversa de Moore-Penrose

Ejemplo

Comprueba que la matriz $M = \frac{1}{11}\begin{pmatrix} 3 & -7 \\ 2 & -1 \\ 3 & 4 \end{pmatrix}$ es la pseudoinversa de Moore-Penrose de la matriz $A = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}$.

Ejemplo

Comprueba que la matriz $N = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ es la pseudoinversa de Moore-Penrose de la matriz $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

Pseudoinversa de Moore-Penrose

Ejemplo

Halla la pseudoinversa de la matriz $A = \begin{pmatrix} 1 & 2 \end{pmatrix}$

Consideremos $A = \begin{pmatrix} 1 & 2 \end{pmatrix} \in \mathbb{R}^{1\times 2}$. Hemos visto que $G = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2\times 1}$ con $x + 2y = 1$ es una inversa generalizada de A, pues:

$$\begin{pmatrix} 1 & 2 \end{pmatrix} = A = AGA = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = (x + 2y) \begin{pmatrix} 1 & 2 \end{pmatrix}$$

Pseudoinversa de Moore-Penrose

Para hallar su pseudoinversa, exigimos las otras tres condiciones de la pseudoinversa:

- $\begin{pmatrix} x \\ y \end{pmatrix} = G = GAG = \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x+2y) \begin{pmatrix} x \\ y \end{pmatrix}$
- $x + 2y = (AG)^T = AG = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x + 2y$
- $\begin{pmatrix} x & y \\ 2x & 2y \end{pmatrix} = (GA)^T = GA = \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = \begin{pmatrix} x & 2x \\ y & 2y \end{pmatrix} \rightarrow 2x = y.$

Resolviendo las ecuaciones $x + 2y = 1$ y $2x = y$, obtenemos que $x = \frac{1}{5}$, $y = \frac{2}{5}$. Por tanto, la pseudoinversa de A es

$$A^{\dagger} = \begin{pmatrix} \frac{1}{5} \\ \frac{2}{5} \end{pmatrix}$$

Pseudoinversa de Moore-Penrose

Es importante observar que no toda inversa generalizada es una pseudoinversa de Moore-Penrose. Consideremos la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, que tenía como inversa generalizada la matriz $G = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} & 0 \\ \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Es decir, $AGA = A$. Además, se puede comprobar que A es también una inversa generalizada de A, pues $GAG = G$. Por tanto, G es una inversa generalizada reflexiva de A.
Ahora bien,

$$(AG)^T \neq AG$$

Por tanto, G no es la pseudoinversa de A.

Unicidad de la Pseudoinversa de Moore-Penrose

Proposición

Sea $A \in \mathbb{R}^{m\times n}$ una matriz cualquiera. Si X es una matriz real de orden $n \times m$ tal que

1. $AXA = A$
2. $XAX = X$
3. $(AX)^T = AX$
4. $(XA)^T = XA$

entonces $A^\dagger = X$.

Propiedades

Para cualquier matriz A se tiene que

1. $rg(A^\dagger) = \mathrm{rg}(A)$
2. $(A^\dagger)^\dagger = A$
3. $(\alpha A)^\dagger = \alpha^{-1} A^\dagger$ para todo escalar $\alpha \neq 0$.
4. $(A^T)^\dagger = (A^\dagger)^T$
5. $(AA^T)^\dagger = (A^T)^\dagger A^\dagger$
6. $(A^T A)^\dagger = A^\dagger (A^T)^\dagger$

Propiedades

Nota

No siempre es cierto que $(AB)^\dagger = B^\dagger A^\dagger$.

En efecto, consideremos $A = \begin{pmatrix} 1 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Entonces:

- $AB = \begin{pmatrix} 3 \end{pmatrix}$ y, por tanto, $(AB)^\dagger = \frac{1}{3}$.
- Se puede comprobar que $A^\dagger = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y $B^\dagger = \frac{1}{5}\begin{pmatrix} 1 & 2 \end{pmatrix}$.
- $B^\dagger A^\dagger = \frac{1}{5}\begin{pmatrix} 1 & 2 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 3 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} \end{pmatrix} \neq \begin{pmatrix} 3 \end{pmatrix} = (AB)^\dagger$

Propiedades

Ejercicio

Demuestra que

$$\begin{pmatrix} A_1 & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & A_m \end{pmatrix}^{\dagger} = \begin{pmatrix} A_1^{\dagger} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & A_m^{\dagger} \end{pmatrix}$$

Ejercicio

Demuestra que $(A^{\dagger})^{\dagger} = A$.

Cálculo de la pseudoinversa de Moore-Penrose

Ejercicio

Calcula la pseudoinversa de Moore-Penrose de la matriz

$$A = \begin{pmatrix} 3 & 0 & 4 \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

En el caso en que la matriz A tiene rango máximo es fácil obtener la matriz $A^\dagger$, según se ve en el siguiente resultado.

Teorema

Sea $A \in \mathbb{R}^{m\times n}$ una matriz que tiene rango máximo. Entonces,

1. *Si $m \geq n$ (es decir, si A tiene rango máximo por columnas),*
$$A^\dagger = (A^T A)^{-1} A^T.$$
2. *Si $m \leq n$ (es decir, si A tiene rango máximo por filas),*
$$A^\dagger = A^T (AA^T)^{-1}.$$

Cálculo de la pseudoinversa de Moore-Penrose

Ejercicio

Calcula la pseudoinversa de Moore-Penrose de las matrices

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1,2 \\ -2,5 & 0 \\ 3,1 & 0,4 \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

A continuación, exponemos un algoritmo para calcular la pseudoinversa de Moore-Penrose, para el caso en el que la matriz A no tenga rango máximo:
Dada $A \in \mathbb{R}^{m \times n}$:

1. Reducir A^T a su forma escalonada reducida de Hermite. Denotaremos por H_{A^T} dicha forma escalonada.
2. Seleccionar las columnas de A^T correspondientes a los pivotes de H_{A^T}. Sean $v_1, v_2, \ldots, v_r$ dichas columnas. Con ellas, formamos una matriz que denotaremos por L.
3. Calcular la matriz AL
4. Hallar la matriz $I - H_{A^T}$ y seleccionar las columnas distintas de cero de dicha matriz, que denotaremos por $w_1, w_2, \ldots, w_{n-r}$-
5. Utilizar las columnas de AL y las columnas w_i del paso anterior como columnas de una nueva matriz $M = \left(\begin{array}{c|c|c|c|c} AL & w_1 & w_2 & \cdots & w_{n-r} \end{array} \right)$ y calcular M^{-1}.
6. Usar las r primeras filas de M^{-1} en el mismo orden en el que aparecen, para formar una nueva matriz que llamaremos R.
7. Calcular $A^\dagger$ como $A^\dagger = LR$

Cálculo de la pseudoinversa de Moore-Penrose

Consideremos, por ejemplo, la matriz $A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 2 & 0 & 2 \\ 3 & 3 & 2 & 3 \\ 4 & 4 & 1 & 1 \end{pmatrix}$.

- La matriz traspuesta es $A^T = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 1 \\ 1 & 2 & 3 & 1 \end{pmatrix}$ y su forma escalonada reducida de Hermite es

$$H_{A^T} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Como podemos observar, en la primera, tercera y cuarta columna están los pivotes (observemos que $r = \mathrm{rg}(A) = \mathrm{rg}(A^T) = 3$. Seleccionamos estas columnas en la matriz A^T y formamos la matriz

$$L = \begin{pmatrix} 1 & 3 & 4 \\ 1 & 3 & 4 \\ 0 & 2 & 1 \\ 1 & 3 & 1 \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

- Calculamos el producto AL,

$$\begin{pmatrix} 3 & 9 & 9 \\ 6 & 18 & 18 \\ 9 & 31 & 29 \\ 9 & 29 & 34 \end{pmatrix}$$

- Tras realizar el cálculo de $I - H_{A^T}$, se tiene que la columna distinta de cero, denominada w_1, es la segunda,

$$\begin{pmatrix} 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \qquad w_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

- Por tanto, la matriz M es de la forma

$$M = \begin{pmatrix} 3 & 9 & 9 & -2 \\ 6 & 18 & 18 & 1 \\ 9 & 31 & 29 & 0 \\ 9 & 29 & 34 & 0 \end{pmatrix}; \qquad \text{además,} \quad M^{-1} = \begin{pmatrix} \frac{71}{120} & \frac{71}{60} & \frac{-5}{8} & \frac{-1}{4} \\ \frac{-1}{8} & \frac{-1}{4} & \frac{7}{24} & \frac{-1}{12} \\ \frac{-1}{20} & \frac{-1}{10} & \frac{-1}{12} & \frac{1}{6} \\ \frac{-2}{5} & \frac{1}{5} & 0 & 0 \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

- Las $r = 3$ primeras filas de M^{-1} forman la matriz R

$$R = \begin{pmatrix} \frac{71}{120} & \frac{71}{60} & \frac{-5}{8} & \frac{-1}{4} \\ \frac{-1}{8} & \frac{-1}{4} & \frac{7}{24} & \frac{-1}{12} \\ \frac{-1}{20} & \frac{-1}{10} & \frac{-1}{12} & \frac{1}{6} \end{pmatrix}$$

- Concluimos que la pseudoinversa de Moore-Penrose de A es

$$A^{\dagger} = LR = \begin{pmatrix} \frac{1}{60} & \frac{1}{30} & \frac{-1}{12} & \frac{1}{6} \\ \frac{1}{60} & \frac{1}{30} & \frac{-1}{12} & \frac{1}{6} \\ \frac{-3}{10} & \frac{-3}{5} & \frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{-1}{3} \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

A continuación, exponemos otro algoritmo para calcular la pseudoinversa de Moore-Penrose, para el caso en el que la matriz A no tenga rango máximo.
En este caso, descompondremos la matriz A como el producto de dos matrices, B y F, siendo B de rango máximo por columnas y F de rango máximo por filas. Es decir, $A = BF$. A partir de esta factorización, calcularemos la pseudoinversa de Moore-Penrose.

1. Reducir A a su forma escalonada reducida de Hermite. Denotaremos por H_A a dicha matriz.
2. Seleccionar las columnas de A correspondientes a los pivotes de H_A y formar una nueva matriz, denotada por B, que tendrá rango máximo por columnas (e igual al rango de A).
3. Seleccionar las filas no nulas de G_A para formar una nueva matriz, denotada por F, que tendrá rango máximo por filas (e igual al rango de A).
4. Con los pasos anteriores, hemos obtenido la factorización $A = BF$.
5. Calcular la pseudoinversa de Moore-Penrose como

$$A^{\dagger} = F^T (B^T A F^T)^{-1} B^T$$

Cálculo de la pseudoinversa de Moore-Penrose

Consideremos, por ejemplo, la matriz $A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 2 & 0 & 2 \\ 3 & 3 & 2 & 3 \\ 4 & 4 & 1 & 1 \end{pmatrix}$.

- La forma escalonada reducida de Hermite de A es

$$H_A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Como podemos observar, en la primera, tercera y cuarta columna están los pivotes (observemos que $3 = \mathrm{rg}(A)$). Seleccionamos estas columnas en la matriz A y formamos la matriz

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 3 & 2 & 3 \\ 4 & 1 & 1 \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

- Seleccionamos las filas no nulas de H_A, para forma la matriz F,

$$F = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Podemos comprobar (no es necesario), que efectivamente, $A = BF$.
- Finalmente, la pseudoinversa de A se obtendrá como

$$A^\dagger = F^T(B^TAF^T)^{-1}B^T = \begin{pmatrix} \frac{1}{60} & \frac{1}{30} & \frac{-1}{12} & \frac{1}{6} \\ \frac{1}{60} & \frac{1}{30} & \frac{-1}{12} & \frac{1}{6} \\ \frac{-3}{10} & \frac{-3}{5} & \frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{-1}{3} \end{pmatrix}$$

Cálculo de la pseudoinversa de Moore-Penrose

Ejercicio

Calcula la pseudoinversa de Moore-Penrose de las siguientes matrices:

$$\begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & 4 & 5 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 5 \\ 0 & 1 & -1 \\ 1 & 3 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 2 \\ 2 & 0 & 1 & 5 \\ 1 & 0 & 0 & 2 \end{pmatrix}$$