Álgebra Matricial

Maestría en Cómputo Estadístico

CIMAT - MCE

Espacios vectoriales

Definición

Sea A un conjunto. Una operación en A es una función $f : A \times A \to A$.

Definición

Sea V un conjunto. Se dice que V es un espacio vectorial sobre K ($K = \mathbb{R}$ ó $\mathbb{C}$) si existe una operación $+$ en V tal que

i) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ *para cualesquiera* $v_1, v_2, v_3 \in V$

ii) *Existe un elemento* 0 *en* V *tal que* $0 + v = v + 0 = v$ *para todo* $v \in V$

iii) *Dado* $v \in V$ *existe un* $u \in V$ *tal que* $v + u = u + v = 0$

iv) $v_1 + v_2 = v_2 + v_1$, $\forall v_1, v_2 \in V$.

Definición

(cont.) Existe además una función $\cdot : K \times V \to V$ *tal que*

i) $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$ *para cualesquiera* $\alpha \in K$, $v_1, v_2 \in V$

ii) $(\alpha_1 + \alpha_2) \cdot v = \alpha_1 \cdot v + \alpha_2 \cdot v$ *para cualesquiera* $\alpha_1, \alpha_2 \in K$, $v \in V$

iii) $\alpha_1 \cdot (\alpha_2 \cdot v) = (\alpha_1 \alpha_2) \cdot v$, *para cualesquiera* $\alpha_1, \alpha_2 \in K$, $v \in V$

iv) $1 \cdot v = v$, $\forall v \in V$.

Ejemplo

$V = \mathbb{R}^n$.

Ejemplo

El espacio trivial $V = \{0\}$.

Ejemplo

$V = M_{m\times n}(\mathbb{R})$.

Definición

Sea V un espacio vectorial sobre K. $W \subset V$ es un subespacio de V si W es un espacio vectorial con las operaciones inducidas por V.

Proposición

Sea V un espacio vectorial sobre K, $W \subset V$, $W \neq \emptyset$. W es un subespacio de V si

i) *Si $w_1, w_2 \in W$ entonces $w_1 + w_2 \in W$*

ii) *Si $w \in W$, $\alpha \in K$ entonces $\alpha w \in W$*

Ejemplo

El subespacio trivial $W = \{0\}$ *de un espacio* V.

Ejemplo

Lineas en $\mathbb{R}^2$

Ejemplo

Hiperplanos en $\mathbb{R}^n$

Proposición

Sea V un espacio vectorial. Si W_1, W_2 son subespacios de V entonces $W_1 \cap W_2$ es subespacio de V.

La unión de subespacios no es en general un subespacio.

Definición

Sea V un espacio vectorial, W_1, W_2 subespacios de V. La suma de W_1 y W_2 es

$$W_1 + W_2 = \{v \in V \mid v = w_1 + w_2, w_1 \in W_1, w_2 \in W_2\}$$

Definición

Sea V un espacio vectorial, W, U subespacios de V. Se dice que V es suma directa de W y U si $V = W + U$ y $W \cap U = \{0\}$. En cuyo caso se escribe $V = W \oplus U$.

Proposición

$V = W \oplus U$ si y solo si todo $v \in V$ se escribe de manera única como $v = w + u$ con $w \in W$ y $u \in U$.

Definición

Sea V un espacio vectorial, W_i, $i = 1, \ldots, r$ subespacios de V. Se dice que V es suma directa de los subespacios W_i si todo $v \in V$ se escribe de manera única como $v = w_1 + \cdots + w_r$ con $w_i \in W_i$, lo cual se escribe como $V = W_1 \oplus \cdots \oplus W_r$.

Proposición

Sea V un espacio vectorial, W_1, W_2 subespacios de V. Entonces $W_1 + W_2$ es un subespacio de V. De hecho, es el espacio más pequeño de V que contiene a $W_1 \cup W_2$.

Definición

Sea V un espacio vectorial y v_1, ..., v_n vectores en V. Si α_1, ..., α_n son escalares, el vector

$$\alpha_1 v_1 + \cdots + \alpha_n v_n$$

es una combinación lineal de v_1, ..., v_n.

Definición

Sea V un espacio vectorial y $S \subset V$. El espacio generado por S es el conjunto

$$\operatorname{gen}(S) = \{v \in V \mid v = \alpha_1 v_1 + \cdots + \alpha_n v_n, a_i \in S, \alpha_i \in K\}$$

Proposición

Sea V un espacio vectorial y $S \subset V$, $S \neq \emptyset$. $\operatorname{gen}(S)$ *es un subespacio de V. Más aún, es el subespacio más pequeño que contiene a S.*

.

Definición

Sea V un espacio vectorial. Sea $S = \{v_1, \ldots, v_n\} \subset V$. S es linealmente independiente si

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

implica que $\alpha_1 = \ldots = \alpha_n = 0$. Si el conjunto no es linealmente independiente se dice que es linealmente dependiente.

Proposición

Sea V un espacio vectorial. Si $S = \{v_1, \ldots, v_n\} \subset V$ y v_j, $1 \leq j \leq n$ se puede escribir como una combinación lineal de los otros elementos de S entonces $\mathrm{gen}(S \setminus \{v_j\}) = \mathrm{gen}(S)$.

Proposición

Sea V *un espacio vectorial y sea* $S = \{v_1, \ldots, v_n\} \subset V$*. Si* S *contiene un subconjunto linealmente dependiente, entonces* S *es linealmente dependiente.*

Proposición

Sea V *un espacio vectorial y sea* $S = \{v_1, \ldots, v_n\} \subset V$*. Si* S *es linealmente independiente, entonces cualquier subconjunto de* S *también es linealmente independiente.*

$\{0\}$ *es siempre linealmente dependiente, por lo que un conjunto linealmente independiente no puede contener a* 0*.*

Proposición

Sea V un espacio vectorial y sea $S = \{v_1, \ldots, v_n\} \subset V$.

1. *Si S es linealmente independiente y $v \in V$, entonces $S \cup \{v\}$ es linealmente independiente si y solo si $v \notin \mathrm{gen}(S)$.*
2. *Si $v_1 \neq 0$, S es linealmente dependiente si y solo si $v_j \in \mathrm{gen}\{v_1, \ldots, v_{j-1}\}$ para algún $2 \leq j \leq n$.*

Proposición

Sea $S = \{v_1, \ldots, v_n\} \subset \mathbb{R}^m$. Si $n > m$ entonces S es linealmente dependiente.

Definición

Sea V un espacio vectorial, $W \subset V$ un subespacio. Una base de W es un subconjunto de W linealmente independiente que genera V.

Proposición

Sea V un espacio vectorial y sea $S = \{v_1, \ldots, v_n\} \subset V$ un subconjunto linealmente independiente. Si $v \in \mathrm{gen}(S)$ y $v = \alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n$, entonces $\alpha_i = \beta_i$, $i = 1, \ldots, n$, es decir, la expresión lineal de v como combinación lineal de los vectores en S es única.

Definición

Sea V un espacio vectorial y sea $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$ una base. Si $v \in V$, las coordenadas de v respecto a $\mathcal{B}$ son los escalares α_i, $i = 1, \ldots, n$, que aparecen en

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

Proposición

Sean V un espacio vectorial, W un subespacio de V, $S = \{v_1, \ldots, v_r\}$ un subconjunto de W linealmente independiente y $W = \text{gen}\{w_1, \ldots, w_s\}$. Entonces $s \geq r$.

Proposición

Si $\mathcal{B}_1$, $\mathcal{B}_2$ *son dos bases del mismo espacio* V, *entonces* $\#(\mathcal{B}_1) = \#(\mathcal{B}_2)$.

Definición

Sea V *un espacio vectorial. La dimensión de* V *es el número de vectores en cualquier base de* V.

Observemos que a una base no se le pueden quitar vectores que generen el espacio ni agregar vectores para mantener los vectores linealmente independientes.

Definición

Sea V un espacio vectorial. La dimensión de V, $\dim V$, es el número de vectores en una base, y por lo tanto cualquiera, de V.

Definición

Sea V un espacio vectorial. La dimensión de V, $\dim V$, es el número de vectores en una base, y por lo tanto cualquiera, de V.

Proposición

Sea V un espacio vectorial, $\dim V = r$, $S \subset V$. Si $V = \mathrm{gen}(S)$, entonces existe un $\mathcal{B} \subseteq S$ tal que $\mathcal{B}$ es base de V.

Proposición

Sea V un espacio vectorial, $\dim V = r$, $S \subset V$. Si S es linealmente independiente, entonces existe $\mathcal{B} \supseteq S$ tal que $\mathcal{B}$ es base de V.

Proposición

Sea V un espacio vectorial y $W \subset V$ un subespacio. Entonces $dim W \leq dim V$.

Proposición

Sea V un espacio vectorial y $W \subset V$ un subespacio. Si $dim W = dim V$, entonces $V = W$.

Si un espacio vectorial V es de dimensión finita n y $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$, $\mathcal{B}_2 = \{u_1, \ldots, u_n\}$ son dos bases distintas, entonces un vector $v \in V$ tendrá distintas coordenadas dependiendo de la base usada. Es decir

$$v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i u_i$$

Es posible encontrar una relación entre ambas coordenadas.

Proposición

Si $(v_i)_{\mathcal{B}_2}$ *es el vector (en* $\mathbb{R}^n$*) de coordenadas de* v_i *con respecto a la base* $\mathcal{B}_2$, $v_{\mathcal{B}_1}$ *es el vector de coordenadas de* v *con respecto a la base* $\mathcal{B}_1$ *y* $v_{\mathcal{B}_2}$ *es el vector de coordenadas de* v *con respecto a la base* $\mathcal{B}_2$ *, entonces existe una matrix invertible* A *dada por* $A = ((v_1)_{\mathcal{B}_2} \cdots (v_2)_{\mathcal{B}_2})$ *(es decir, tiene los vectores de coordenadas como columnas) tal que*

$$v_{\mathcal{B}_2} = A v_{\mathcal{B}_1}$$

A es la matriz de cambio de coordenadas de la base $\mathcal{B}_1$ a la base $\mathcal{B}_2$.

$$A_{\mathcal{B}_1 \to \mathcal{B}_2}$$

Proposición

Si A es la matriz de cambio de coordenadas de la base $\mathcal{B}_1$ a la base $\mathcal{B}_2$ entonces A^{-1} es la matriz de cambio de coordenadas de la base $\mathcal{B}_2$ a la base $\mathcal{B}_1$.

$$A_{\mathcal{B}_2 \to \mathcal{B}_1} = (A_{\mathcal{B}_1 \to \mathcal{B}_2})^{-1}$$

Matrices

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}_{m\times n}$$

$a_{ij} \in F,\ i = 1, \cdots, m,\ j = 1, \cdots, n,\ F = \mathbb{R},\ \mathbb{C}$

$$(a_{ij})_{1\le i\le m, 1\le j\le n}$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}_{m\times n} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}_{p\times q}$$

si y solo si $m = p$, $n = q$ y $a_{ij} = b_{ij}$, $\forall i, j$.

Suma de matrices

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+a_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

$$(c_{ij}), \quad c_{ij} = a_{ij} + b_{ij}$$

Definición

La matriz cero es

$$0_{m\times n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Si $A = (a_{ij})$,

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

Proposición

Sean A, B, C matrices del mismo tamaño. Entonces

i) $A + B = B + A$

ii) $A + (B + C) = (A + B) + C$

iii) $A + 0 = 0 + A = A$

iv) $A + (-A) = (-A) + A = 0$

$A - B := A + (-B)$

$A = (a_{ij})$

$r \in \mathbb{R}$

$$rA = \begin{pmatrix} ra_{11} & ra_{12} & \dots & ra_{1n} \\ ra_{21} & ra_{22} & \dots & ra_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \dots & ra_{mn} \end{pmatrix}$$

Proposición

Sean $A = (a_{ij})$, $r, s \in \mathbb{R}$

i) $r(A + B) = rA + rB$

ii) $(r + s)A = rA + sA$

iii) $r(sA) = (rs)A$

Multiplicación de matrices

Definición

Sea $A = (a_{ij})$ *una matriz de tamaño* $m \times n$ *y* $B = (b_{ij})$ *una matriz de tamaño* $n \times p$*. El producto de* A *y* B *es la matriz de de tamaño* $m \times p$

$$AB := (c_{ij})$$

donde

$$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$$

Definición

La matriz identidad $n \times n$ es

$$I_{n\times n} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Proposición

Sean A, B y C matrices para las cuales las operaciones abajo están bien definidas.

i) $A(BC) = (AB)C$

ii) $IA = AI$

iii) $A(B + C) = AB + AC$

iv) $(A + B)C = AC + BC$

Para una matriz cuadrada A, la potencia k-ésima de A es $A^k = AA \cdots A$ (k veces)

Transpuesta de una matriz

Definición

Sea $A = (a_{ij})$ *una matriz de* $m \times n$. *La transpuesta de* A *es la matriz de* $n \times m$ *dada por*

$$A^t := (b_{ij})$$

donde $b_{ij} = a_{ji}$.

Proposición

Sean A y B matrices para las cuales las operaciones abajo están bien definidas.

i) $(AB)^t = B^t A^t$

ii) $(A^t)^t = A$

iii) $(A + B)^t = A^t + B^t$

iv) $(rA)^t = rA^t, \forall r \in \mathbb{R}$

Definición

El determinante de

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

está dado por el $\det(A) = ad - bc$.
También se usa la notación $|A|$.

Definición

Sea $A = (a_{ij})$ *una matriz* $n \times n$, $n > 2$. *El determinante de* A *es*

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}$$

para un valor de i *fijo. Donde,*

$$C_{ij} = (-1)^{i+j} \det(M_{ij})$$

y M_{ij} *la matriz resultante de descartar la fila* i *y la columna* j.

Propiedades de los determinantes

1. $|I_n| = 1$
2. $|A^T| = |A|$
3. $|A^{-1}| = |A|^{-1}$
4. $|AB| = |A||B|$
5. $|cA| = c^n|A|$
6. Si A es triangular, $A = \prod_i a_{ii}$
7. Si A tiene un renglón o columna de 0's, $|A| = 0$
8. Si intercambiamos dos renglones o columnas, el determinante cambia de signo
9. Si A tiene dos columnas iguales, $|A| = 0$
10. Si A tiene columnas que son combinaciones lineales de otras columnas, $|A| = 0$
11. Si a un renglón dde A le agregamos otro renglón de A, multiplicado por una constante, el determinante no se altera

Definición

Sea A una matriz de tamaño $m \times n$. Se dice que A es una matriz particionada o por bloques si se puede escribir como una matriz

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1q} \\ A_{21} & A_{22} & \dots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \dots & A_{pq} \end{pmatrix}_{m \times n}$$

donde cada una de las entradas A_{ij} es a su vez una matriz de tamaño $m_i \times n_j$ y $\sum_{i=1}^{p} m_i = m$, $\sum_{j=1}^{q} n_j = n$.

Proposición

$$A^t = \begin{pmatrix} A_{11}^t & A_{21}^t & \cdots & A_{p1}^t \\ A_{12}^t & A_{22}^t & \cdots & A_{p2}^t \\ \vdots & \vdots & \ddots & \vdots \\ A_{1q}^t & A_{2q}^t & \cdots & A_{pq}^t \end{pmatrix}_{n\times m}$$

Sumar y multiplicar matrices por bloques para que los bloques sean tratados como si fueran elementos requiere condiciones adicionales a las que ya se tienen.

Ejemplo

$$A = \begin{pmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 4 & 0 \end{pmatrix}_{n\times m}$$

$$A = \begin{pmatrix} 4 & -1 & 6 \\ 3 & 8 & 3 \\ 0 & 0 & 0 \\ 4 & 3 & -5 \end{pmatrix}_{n\times m}$$

Sea A una matriz cuadrada de tamaño $m \times n$ y x un vector $n \times 1$. Si A está dada por bloques columna por

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix},$$

es decir a_i es un vector columna $m \times 1$ y

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

entonces

$$\begin{aligned} Ax &= \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1 a_1 + x_2 a_2 + \ldots + x_n a_n \end{aligned}$$

Una combinación lineal de matrices es una suma del tipo

$$\alpha_1 A_1 + \alpha_2 A_2 + \ldots + \alpha_k A_k$$

donde las A_i son matrices y los α_1 números reales.

Ax es una combinación lineal de las columnas de A

Análogamente sea A una matriz cuadrada de tamaño $m \times n$ y y un vector horizontal $1 \times m$. Si A está dada por bloques renglón por

$$A = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix},$$

es decir b_i es un vector renglón $1 \times n$ y

$$y = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix},$$

entonces

$$\begin{aligned} yA &= \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \\ &= y_1 u_1 + y_2 u_2 + \ldots + y_m u_m \end{aligned}$$

Luego, yA es una combinación lineal de los renglones de A.

En general, un producto matricial puede escribirse como

$$AB = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_p \end{pmatrix} = \begin{pmatrix} u_1 b_1 & u_1 b_2 & \cdots & u_1 b_p \\ u_2 b_1 & u_2 b_2 & \cdots & u_2 b_p \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

donde A tiene bloques de vectores renglón y B tiene bloques de vectores columna.

Si ahora A tiene bloques de vectores columna y B tiene bloques de vectores renglón, entonces el producto está dado por la suma de matrices dadas por el producto exterior

$$AB = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum a_i v_i$$

Definición

Sea $A = (a_{ij})$ *una matriz cuadrada de tamaño* n*. La traza de* A *es*

$$\operatorname{tr}(A) := \sum_{i=1}^{n} a_{ii}$$

Proposición

Sean A*,* B *matrices cuadradas de tamaño* n*,* $\alpha \in \mathbb{R}$*.*

i) $\operatorname{tr}(A) = \operatorname{tr}(A^t)$

ii) $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr}(A)$

iii) $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$

Proposición

Sean $A_{m\times n}$, $B_{n\times m}$ *matrices. Entonces* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$,

Proposición

Sean u, v *vectores* $m \times 1$. *Entonces* $\operatorname{tr}(uv^t) = u^t v = v^t u$.

$$\operatorname{tr}(ABC) = \operatorname{tr}(CAB) = \operatorname{tr}(BCA)$$

Ejemplo

Sea A *una matriz cuadrada. Encuentre una matriz cuadrada* X *tal que* $AX - XA = I$

Definición

P es una matriz de permutación de tamaño n si se obtiene de permutar las columnas o renglones de la matriz identidad I_n.

Equivalentemente:

Definición

Una matriz de permutación de tamaño n es una matriz cuadrada tal que cada columna cada renglón contiene exactamente un 1 y 0 en los demás lados.

Si e_i es el vector canónico y

$$A_{m\times n} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

por bloques renglón y columna respectivamente y P es una matriz de permutación, entonces

$$PA = \begin{pmatrix} e_{i1}^t \\ e_{i2}^t \\ \vdots \\ e_{im}^t \end{pmatrix} A = \begin{pmatrix} Ae_{i1}^t \\ Ae_{i2}^t \\ \vdots \\ Ae_{im}^t \end{pmatrix} = \begin{pmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{im} \end{pmatrix}$$

$$\begin{aligned} AQ = A\begin{pmatrix} e_{j1} & e_{j2} & \cdots & e_{jn} \end{pmatrix} &= \begin{pmatrix} Ae_{j1} & Ae_{j2} & \cdots & Ae_{jn} \end{pmatrix} \\ &= \begin{pmatrix} a_{j1} & a_{j2} & \cdots & a_{jn} \end{pmatrix} \end{aligned}$$

Definición

Una matriz cuadrada $A = (a_{ij})$ *es triangular superior si* $a_{ij} = 0$, $i > j$. *A es triangular inferior si* $a_{ij} = 0$, $j > i$. *A es triangular si es triangular inferior o superior*

Proposición

Sean A, B *matrices cuadradas de tamaño* n.

i) *Si* A *es triangular superior (inferior) entonces* A^t *es triangular inferior(superior).*

ii) *Si* A *y* B *son triangulares superiores (inferiores) entonces* AB *es triangular superior(inferior).*

Definición

Sea A una matriz cuadrada. Si existe un entero positivo k tal que $A^k = 0$, se dice que A es nilpotente.

Proposición

Sean $A_{n\times n}$ una matriz triangular con entradas diagonales cero. Entonces A es nilpotente. De hecho, $A^n = 0$.

Sistema de ecuaciones lineales

Un sistema de ecuaciones lineales de m ecuaciones y n variables es un sistema del tipo

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Los a_{ij} son escalares fijos llamados coeficientes y las x_i son las variables. Si $m = n$ el sistema se llama cuadrado y rectangular de otra manera.

Una solución es una n-eada $(s_1, \cdots, s_n)$ que hace que las ecuaciones sean ciertas al mismo tiempo.

Representación matricial de un sistema lineal

Todo sistema se puede representar por una ecuación matricial del tipo $Ax = b$.

La matriz de coeficientes del sistema A está dada por

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}_{m\times n}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n\times 1} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m\times 1}$$

En representación matricial, una solución es un vector

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix}$$

que satisface

$$As = b$$

Un sistema $Ax = b$ puede tener

i) Solución única
ii) Un número infinito de soluciones
iii) No solución

Proposición

Si un sistema $Ax = b$ tiene más de una solución entonces tiene un número infinito de soluciones.

Definición

Un sistema $Ax = b$ es consistente si tiene al menos una solución e inconsistente si no.

Definición

Dada una matriz, una operación elemental por renglones es una de las siguientes:

i) *Tipo I: intercambiar dos renglones de la matriz*
ii) *Tipo II: multiplicar un renglón por un escalar distinto de cero.*
iii) *Tipo III: reemplazar un renglón por la suma de ese renglón con el múltiplo escalar otro renglón*

Proposición

Si se aplica las mismas operaciones por renglón a A y b en el sistema $Ax = b$, la solución del sistema sigue siendo la misma.

La matriz de coeficientes aumentada del sistema está dada por

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

De esta forma aplicar operaciones por renglones a la matriz aumentada $(A|b)$ es equivalente a aplicarlas en ambos lados de la ecuación $Ax = b$.

Sea U la matriz dada por bloques renglón o columna

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} = (v_1, v_2, \ldots, v_n)$$

Una matriz U tiene forma escalonada por renglones si se cumple:

- Si el primer elemento no cero en un renglón u_i está en la posición j, entonces todas las entradas abajo de la posición i en las columnas $v_1, \ldots, v_j$ son cero.
- Si el renglón u_i es cero, entonces todos los renglones abajo de él son vectores de ceros.

Los pivotes son las primeras entradas no cero en cada renglón.

El método de Gauss o de eliminación Gaussiana consiste en llevar una matriz aumentada correspondiente a un sistema $Ax = b$ a una forma escalonada por renglones.

En una matriz cuadrada $A_{n \times n}$, correspondiente al sistema $Ax = b$, obtenemos un nuevo sistema $Ux = b'$ con matriz aumentada

$$\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} & b'_1 \\ 0 & 0 & \dots & u_{2n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & u_{nn} & b'_n \end{pmatrix}_{m \times n}$$

Las soluciones del sistema cuadrado están dadas entonces por: 1. $x_n = b'_n / u_{nn}$.

2. Recursivamente:

$$x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{j=i+1}^{n} u_{ij} x_j \right)$$

para $i = n, n-1, \ldots, 2, 1$.

Si el sistema es rectangular $n > m$, el sistema reducido tendrá mas incógnitas que ecuaciones por lo que se deben seleccionar algunas variables que se llamaran básicas y las restantes serán libres. Usualmente se escogen los pivotes como básicas y las restantes como libres. Una vez seleccionadas se hace sustitución hacia atrás y se encuentra una solución general.

Por medio de eliminación Gaussiana podemos saber si un sistema es inconsistente o no: Un sistema es inconsistente si y solo si al reducirlo se encuentra un renglón (en la matriz aumentada) del tipo $(0 \ldots 0, \alpha)$

Teorema

Sea A una matriz y sean U_1, U_2 dos formas escalonadas por renglones de A distintas (i.e. se emplearon dos sucesiones distintas de operaciones elementales). Entonces el número de pivotes de U_1 y U_2 es el mismo.

De esta manera, no importa que sucesión de operaciones elementales usemos, el número de pivotes permanece constante, y en consecuencia en las mismas posiciones. Por lo tanto el número de pivotes es un invariante de A.

El número de pivotes cuenta el número de variables básicas por lo que estás son siempre el mismo número y similarmente para las libres.

Sea U cualquier matriz en forma reducida por renglones obtenida de A, entonces:

$$\begin{aligned} \text{Número de pivotes de } A &= \text{Número de pivotes de } U \\ &= \text{Número de renglones no cero de } U \\ &= \text{Número de columnas básicas de } U \end{aligned}$$

Definición

El rango de A es el número de pivotes de A.

Una matriz E tiene forma escalonada reducida por renglones si se cumple:

- E Está en forma escalonada por renglones
- El primer elemento no cero en cada renglón es 1.
- Todas las entradas arriba de cada pivote son 0.

El método de Gauss-Jordan o de eliminación Gauss-Jordan consiste en llevar una matriz aumentada correspondiente a un sistema $Ax = b$ a una forma escalonada por renglones.

Las columnas básicas de $E_{m\times n}$ son r vectores canónicos en R^m.

Siempre podemos encontrar una matriz de permutación P tal que

$$EP = \begin{pmatrix} I_r & J \\ 0 & 0 \end{pmatrix}$$

Administración de Recursos

Un departamento de pesca y caza del estado proporciona tres tipos de comida a un lago que alberga a tres especies de peces. Cada pez de la especie 1 consume cada semana un promedio de 1 unidad del alimento A, 1 unidad del alimento B y 2 unidades del alimento C. Cada pez de la especie 2 consume cada semana un promedio de 3 unidades del alimento A, 4 del B y 5 del C. Para un pez de la especie 3, el promedio semanal de consumo es de 2 unidades del alimento A, 1 unidad del alimento B y 5 unidades del C. Cada semana se proporcionan al lago 25 000 unidades del alimento A, 20 000 unidades del alimento B y 55 000 del C. Si suponemos que los peces se comen todo el alimento, ¿cuántos peces de cada especie pueden coexistir en el lago?

Proposición

Sea A una matriz $m \times n$. Hacer una operación elemental por renglones en A es lo mismo que multiplicar A por la izquierda por la correspondiente matriz elemental.

Teorema

Toda matriz elemental es invertible. El inverso de una matriz elemental es una matriz del mismo tipo.

Además, si:

- Si B multiplica el renglón i de A por $c \neq 0$, entonces B^{-1} multiplica el renglón i de A por ${}^{1}/_{c}$.
- Si B multiplica el renglón i de A por c y lo suma al renglón j, entonces B^{-1} multiplica el renglón i de A por $-c$ y lo suma al renglón j.
- Si B permuta los renglones i y j de A, entonces B^{-1} también permuta los renglones i y j de A.

Sistemas lineales homogéneos

Definición

Un sistema lineal $Ax = 0$ se llama un sistema lineal homogéneo

Siempre tiene la solución trivial $x = 0$ por lo que es consistente.

Dados v_1, v_2, ..., v_n vectores, una combinación lineal es una expresión del tipo

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

donde los α_i son escalares.

La solución general de un sistema homogéneo es una combinación lineal de las soluciones particulares que se obtienen al hacer una variable libre igual a 1 y 0 en las demás, tomando las variables libres como coeficientes.

Proposición

Si A es $m \times n$ y $n > m$, entonces $Ax = 0$ tiene soluciones no triviales.

En general, si A es $m \times n$ la solución general del sistema homogéno es

$$x = x_{f_1} h_1 + x_{f_2} h_2 + \ldots + x_{f_{n-r}} h_{n-r}$$

Para el sistema $Ax = b$, la solución general es del tipo

$$x = x_p + x_h$$

donde x_h es la solución general del sistema homogéneo asociado y x_p es una solución particular que se obtiene haciendo todas las variables libres igual a cero.

Definición

Sea A una matriz cuadrada $n \times n$. Se dice que A es invertible o no singular si existe una matriz B tal que $AB = I_n$, $BA = I_n$. B es una inversa de A.

Si A no es invertible se le llama singular.

Proposición

Si A es invertible, la matriz inversa es única.

La inversa de A se denota por A^{-1}.

Si una matriz es invertible, entonces la solución del sistema $Ax = b$ esta dada por $x = A^{-1}b$.

El sistema se llama sistema no singular cuando A es no singular.

Proposición

Sen A, B matrices $n \times n$. Entonces $AB = I$ si y solo si $BA = I$.

Si A es una matriz $m \times n$, una inversa izquierda de A es una matriz C, $n \times m$ tal que $CA = I_n$. Una inversa derecha es una matriz B, $m \times n$ tal que $AB = I_m$.

Una matriz rectangular puede tener inversa por un lado, pero no del otro y no tienen por que coincidir.

Proposición

Sean A una matriz $n \times n$. Entonces

i) $(AB)^{-1} = B^{-1}A^{-1}$

ii) $(A^{-1})^{-1} = A$

iii) $(A^{-1})^t = (A^t)^{-1}$

Para encontrar A^{-1}, tenemos la ecuación $AX = I$ que puede verse como un conjunto de sistemas de ecuaciones $Ax_i = e_i$ donde x_i son las columnas de X.

Luego con el método de Gauss, si es posible resolver cada uno de estos sistemas con solución x_i, encontramos la inversa de A dada por $X = (x_1 \ldots x_n)$. Usando Gauss Jordan

$$L(A|I) = (LA|LI) = (I|L)$$

implica que $L = A^{-1}$

Proposición

A es invertible si y solo si el sistema homogéneo asociado solo tiene la solución $x = 0$.

Proposición

Sea $A \neq 0$ una matriz $m \times n$. Entonces:

i) *Existe una matriz invertible G tal que $GA = U$, donde U está en forma escalonada por renglones.*

ii) *Existe una matriz invertible H tal que $HA = E$, donde E está en forma escalonada reducida por renglones.*

Proposición

A es no singular si y solo si A se puede representar mediante el producto de matrices elementales de tipo I, II, ó III.

Proposition

Si $A_{n\times n}$ es triangular inferior (superior) con todas sus entradas diagonales distintas de cero, entonces es invertible y su inversa es triangular inferior (superior). Más aún, cada elemento de la diagonal de la matriz inversa es el recíproco del elemento correspondiente de A.

Proposition

El producto de dos matrices de permutación es una matriz de permutación.

Proposition

Una matriz de permutación es invertible y su inversa es su transpuesta.

Definición

Sea A una matriz $n \times n$. Se dice que A tiene una descomposición LU si se puede factorizar como $A = LU$ donde

i) $L = (l_{ij})$ *es una matriz triangular inferior tal que* $l_{ii} = 1$, $i = 1, \ldots, n$

ii) $U = (u_{ij})$ *es una matriz triangular superior tal que* $u_{ii} \neq 0$, $i = 1, \ldots, n$

L es el factor inferior y U es el factor superior.

Proposición

Si A es una matriz y U es la matriz que se obtiene usando eliminación Gaussiana y no se necesitaron intercambios de renglones (i.e. no aparecieron pivotes cero) entonces A tiene una descomposición LU, dada por $A = LU$.

Los elementos de la diagonal de U son los pivotes de A.

Para resolver el sistema lineal $Ax = b$ cuando $A = LU$ se resuelven dos sistemas triangulares $Ly = b$ y $Ux = y$.

Primero se usa sustitución hacia adelante en $Ly = b$:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

$$y_1 = b_1$$

$$y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_i$$

$i = 2, 3, \ldots, n.$

Luego se resuelve el sistema $Ux = y$ usando sustitución hacia atras:

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$$

$$x_n = y_n / u_{nn}$$

$$x_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_i$$

$$x_i = \frac{1}{u_{ii}} \left(y_i - \sum_{j=i+1}^{n} u_{ij} x_j \right)$$

$i = n-1, n-2, \ldots, 1.$

Una submatriz de una matriz A es una matriz que se obtiene al eliminar renglones y columnas de A.

Si A es de tamaño $m \times n$, $I_r \subset \{1, \ldots, m\}$, $I_c \subset \{1, \ldots, n\}$, se denota por $A_{I_r, \cdot}$ a la matriz que se forma al dejar solo los renglones indexados por I_c y $A_{\cdot, I_c}$ a la matriz que se forma al dejar solo las columnas indexadas por I_c.

Si A es $n \times n$, una una submatriz principal si se obtiene de A eliminando los mismos renglones y columnas, i.e. $I_c = I_r$. Una submatriz principal lider se obtiene cuando es principal y $I_c = I_r = \{1, 2, \cdots, k\}$, $k < n$.

Proposición

Sea A una matriz no singular. A tiene una factorización LU si y solo si todas sus submatrices principales líderes son no singulares.

Proposición

Si A una matriz no singular y $A = LU$ es una factorización LU de A, entonces L y U son únicas.

No todas la matrices no singulares tienen una descomposición LU, por ejemplo

$$\begin{pmatrix} 0 & 1 \\ 1 & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix}$$

Sin embargo

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

Es decir, podemos multiplicar una matriz A por una matriz de permutación P de tal manera que $PA = LU$.

En general esto es cierto, siempre podemos encontrar una matriz de permutación P tal que PA tiene una descomposición LU.

Proposición

Si A es no singular, entonces existe una matriz de permutación P tal que PA tiene una descomposición LU

$$PA = LU$$

donde L es triangular inferior con unos en la diagonal y U es triangular superior.

Proposición

Sea A no singular de tamaño $n \times n$. Entonces $PA = LDU$ donde P es una matriz de permutación $n \times n$, L es $n \times n$ es una matriz triangular inferior con unos en la diagonal, D es $n \times n$ diagonal con elementos diagonales no cero y U es $n \times n$ es una matriz triangular superior con unos en la diagonal.

Los elementos de D son distintos de cero pues son los pivotes de A.

L y U son únicas.

Lema

Sea A una matriz simétrica de tamaño $n \times n$ tal que $A = LDU$ donde L es triangular inferior con unos en la diagonal, U es triangular superior con con unos en la diagonal, D es diagonal con elementos diagonales no cero. Entonces $U = L^t$ y $A = LDL^t$.

Proposición

Sea A una matriz simétrica de tamaño $n \times n$ que tiene una descomposición LU con pivotes estrictamente positivos. Entonces existe una matriz triangular inferior T tal que $A = TT^t$ y los elementos diagonales de T son positivos.

Definición

Sea A una matriz simétrica de tamaño $n \times n$. Se dice que A es positiva definida si todos sus pivotes son estrictamente positivos.

La descomposición $A = TT^t$ de una matriz positiva definida se llama la descomposición de Cholesky.

Definición

Sea A una matriz de tamaño $m \times n$. El espacio columna de A es

$$\mathcal{C}(A) = \{y \in \mathbb{R}^m \mid y = Ax \textit{ para algún } x \in \mathbb{R}^n\}$$

Dicho de otra manera $\mathcal{C}(A)$ es el espacio generado por las columnas de A de donde automáticamente obtenemos que $\mathcal{C}(A)$ es subespacio de $\mathbb{R}^m$.

Definición

Sea A una matriz de tamaño $m \times n$. El espacio renglón de A es

$$\mathcal{R}(A) = \{y \in \mathbb{R}^n \mid y = A^t x \textit{ para algún } x \in \mathbb{R}^m\}$$

$\mathcal{R}(A) = \mathcal{C}(A^t)$ por lo que es subespacio de $\mathbb{R}^n$.

Definición

Sea A una matriz de tamaño $m \times n$. El espacio nulo de A es

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

$\ker(A) = \mathcal{N}(A)$

Proposición

Sean A una matriz $m \times n$. $\mathcal{N}(A)$ es subespacio de $\mathbb{R}^n$.

En general, dados vectores v_1,..., v_n, en $\mathbb{R}^m$, estos pueden acomodarse para formar las columnas de una matriz A de tamaño $m \times n$. Luego, un vector b estará en gen$\{v_1$,..., $v_n\}$ si y solo si $Ax = b$ es consistente.

Proposición

Sea $S = \{a_1, \ldots, a_n\} \subset \mathbb{R}^m$. S es linealmente independiente si y solo si la matriz formada con los a_i como columnas tiene que $\ker(A) = \{0\}$.

Corolario

Sea $S = \{a_1, \ldots, a_n\} \subset \mathbb{R}^n$. S es linealmente independiente si y solo si la matriz formada con los a_i como columnas es invertible.

Corolario

Sea A una matriz $m \times n$. Un subconjunto de columnas de A es linealmente independiente si y solo si las columnas correspondientes en las mismas posiciones de la matriz escalonada por renglones de A son linealmente independientes.

Las columnas donde están los pivotes son linealmente independientes y son las columnas básicas.

Proposición

Sea A una matriz $m \times n$. El número de renglones linealmente independientes de A es igual al número de sus columnas linealmente independientes.

Definición

Sea A una matriz $m \times n$. El rango de A es el número (máximo) de renglones linealmente independientes de A.

Proposición

Sea A una matriz $m \times n$ con r columnas linealmente independientes. Entonces existe un conjunto de $n - r$ vectores linealmente independientes que son solución del sistema homogéneo dado por A y cualquier otra solución se puede expresar como una combinación lineal de esas $n - r$ soluciones linealmente independientes.

Definición

El rango de una matriz A es la dimensión del espacio columna de A.

$\rho(A)$

Definición

La nulidad de una matriz A es la dimensión del espacio nulo (o kernel) de A.

$\nu(A)$

Ejemplo

$\rho(A) = 0$ *si y solo si* $A = 0$.

Ejemplo

T *una matriz triangular*

Ejemplo

I_n

Proposición

$\rho(A) =$ *número de renglones linealmente independientes de A.*

Corolario

Si A es cuadrada, entonces $\rho(A) = \rho(A^t)$

Corolario

Si A es de tamaño $m \times n$, *entonces* $\rho(A) \leq \min\{m, n\}$

Teorema

Sea A un matriz de tamaño $m \times n$. Entonces

$$\rho(A) + \nu(A) = n$$

Proposición

Sea A de tamaño $n \times n$. A es invertible si y solo si $\nu(A) = 0$.

Proposición

Sea A de tamaño $n \times n$. A es invertible si y solo si $\rho(A) = n$.

Aplicación

Sistemas dinámicos y los búhos manchados

En 1990, el búho manchado del norte se convirtió en el centro de una controversia nacional en Estados Unidos acerca del uso de los majestuosos bosques del Noroeste pacífico. Los ecologistas convencieron al gobierno federal de que el búho manchado estaría condenado a la extinción si continuaba la tala en los bosques de crecimiento viejo (con árboles de más de 200 años), donde prefi eren vivir los búhos. La industria maderera, anticipando la pérdida de entre 30,000 y 100,000 empleos como resultado de las nuevas restricciones a la tala por parte del gobierno, argumentó que el búho no debería clasifi carse como una "especie en peligro de extinción" y citó varios informes científi cos publicados para apoyar su caso.

Atrapados en el fuego cruzado de los dos grupos en confl icto, los ecologistas matemáticos intensifi caron sus esfuerzos por entender la dinámica poblacional del búho manchado. El ciclo de vida de un búho manchado se divide naturalmente en tres etapas: juvenil (hasta 1 año de edad), subadulto (1 a 2 años), y adulto (más de 2 años).

El búho se aparea de por vida durante las etapas de subadulto y adulto, empieza a reproducirse en la edad adulta, y vive hasta 20 años. Cada pareja de búhos requiere aproximadamente de 1000 hectáreas (4 millas cuadradas) como territorio base. Un momento crítico en el ciclo de vida es cuando los búhos jóvenes abandonan el nido. Para sobrevivir y convertirse en subadulto, un búho joven debe encontrar un nuevo territorio base (y generalmente una pareja).

Un primer paso para estudiar la dinámica poblacional es confi gurar un modelo de la población a intervalos anuales, en tiempos denotados mediante $k = 0, 1, 2, ...$ Por lo general, se supone que existe una relación 1:1 de machos a hembras en cada etapa de vida, y se cuentan exclusivamente las hembras. La población en el año k puede describirse por medio de un vector $x_k = (j_k, s_k, a_k)$, donde j_k, s_k y a_k son las cantidades de hembras existentes en las etapas juvenil, subadulto y adulto, respectivamente.

Utilizando datos de campo de estudios demográfi cos reales, R. Lamberson y colaboradores consideraron el siguiente modelo de matrices por etapas:

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

Aquí, la cantidad de nuevas hembras juveniles en el año $k+1$ es .33 veces la cantidad de hembras adultas en el año k (con base en el índice de natalidad promedio por pareja de búhos). También, el 18% de los jóvenes sobrevive para convertirse en subadultos, y un 71% de los subadultos y el 94% de los adultos sobreviven para ser contados como adultos.

El modelo de matrices por etapas es una ecuación en diferencias de la forma $x_{k+1} = Ax_k$. Una ecuación de este tipo se llama sistema dinámico (o sistema dinámico lineal discreto) porque describe los cambios experimentados en un sistema al paso del tiempo.

La tasa de supervivencia juvenil del 18% en la matriz por etapas de Lamberson es la entrada más afectada por la cantidad de bosque viejo disponible. De hecho, el 60% de los búhos juveniles normalmente sobrevive para dejar el nido, pero en la región de Willow Creek, California, estudiada por Lamberson y sus colegas, sólo el 30% de los jóvenes que dejaron el nido pudo encontrar nuevos territorios base; el resto pereció durante el proceso de búsqueda.

Una razón importante del fracaso de los búhos al tratar de encontrar nuevos territorios es el aumento en la fragmentación de las áreas con árboles viejos debido a la tala total de áreas diseminadas en los terrenos de crecimiento viejo. Cuando un búho deja la protección del bosque y cruza un área devastada, el riesgo de que sea atacado por un depredador aumenta de modo impresionante.

El modelo descrito anteriormente predice la eventual extinción del búho manchado, pero también que si el 50% de los búhos juveniles que sobreviven para dejar el nido encuentra nuevos territorios, entonces la población de búho manchado prosperará.

Definición

Sea A una matriz $n \times n$, un vector $x \in \mathbb{R}^n$, $x \neq 0$ es un vector propio de A si $Ax = \lambda x$ para algún escalar λ, que en tal caso es llamado un valor propio.

λ es un valor propio de A si y solo si la ecuación

$$(A - \lambda I)x = 0$$

tiene una solución no trivial.

Definición

El espacio propio de A correspondiente a λ es

$$E_\lambda := \mathcal{N}(A - \lambda I)$$

Proposición

Si A es triangular, sus valores propios son las entradas de su diagonal principal.

Proposición

0 *es un valor propio de A si y solo si A es no invertible.*

Proposición

Sea A una matriz $n \times n$. Si el valor propio λ_i corresponde al vector propio v_i, $i = 1, \ldots, r$ y $\lambda_i \neq \lambda_j$, $i \neq j$, entonces el conjunto $\{v_1, \ldots, v_r\}$ es linealmente independiente.

Definición

Sea A una matriz $n \times n$. La ecuación característica de A es

$$\det(A - \lambda I) = 0$$

Proposición

λ es un valor propio de A si y solo si λ satisface la ecuación característica.

Proposición

Sea A una matriz $n \times n$. $p(\lambda) := \det(A - \lambda I)$ es un polinomio de grado n, llamado el polinomio característico de A.

Definición

Sean A y B matrices $n \times n$. Se dice que A es similar a B si existe una matriz invertible P tal que $A = PBP^{-1}$ o, de manera equivalente, $B = P^{-1}AP$.

Si A es similar a B entonces B es similar a A, por lo que simplemente decimos que A y B son similares. De hecho, la similaridad es una relación de equivalencia.

Definición

**Alternativa*

A y B son matrices semejantes si y sólo si existe una matriz invertible P tal que $PA = BP$ o, de manera equivalente, $PB = AP$.

Proposición

Si A y B son similares entonces tienen el mismo polinomio característico y por tanto tienen los mismos valores propios.

Definición

Una matriz es diagonalizable si es similar a una matriz diagonal D.

Proposición

Sea A una matriz $n \times n$. Entonces A es diagonalizable si y solo si A tiene n vectores propios linealmente independientes. De hecho, $A = PDP^{-1}$ si y solo si las columnas de P son los n vectores propios de A linealmente independientes y las entradas de la matriz diagonal D son los valores propios correspondientes a los vectores propios.

Proposición

Sea A una matriz $n \times n$. Si A tiene n valores propios distintos, entonces A es diagonalizable.

Definición

Sea A una matriz $n \times n$ y λ_0 un valor propio de A. La multiplicidad algebraica de λ_0 es el número m de veces que aparece como raíz del polinomio característico, $p(\lambda) = (\lambda - \lambda_0)^m q(\lambda)$. La multiplicidad geométrica de λ_0 es $\dim E_{\lambda_0} = \dim \mathcal{N}(A - \lambda_0 I)$.

Proposición

La multiplicidad geométrica de un valor propio es menor o igual que su multiplicidad algebraica.

Proposición

Sea A una matriz $n \times n$. A es diagonalizable si y solo si la multiplicidad geométrica de cada valor propio es igual a su multiplicidad algebraica; es decir si la suma de las dimesiones de los espacios propios es igual a n.

Proposición

Sea A una matriz diagonalizable $n \times n$ de rango r. Entonces existe una matriz no singular V tal que

$$A = V \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} V^{-1}$$

donde Λ_1 es diagonal de tamaño $r \times r$ con elementos distintos de cero en la diagonal.

Proposición

Sea A una matriz diagonalizable $n \times n$ de rango r. Entonces existen escalares λ_1, ..., λ_r tales que

$$A = \lambda_1 v_1 w_1^t + \lambda_2 v_2 w_2^t + \cdots + \lambda_r v_r w_r^t$$

donde los v_i, w_j son vectores en $\mathbb{C}^n$ tales que $w_i^t v_i = 1$, $w_i^t v_j = 0$, $i \neq j$, $i, j = 1, \ldots, r$.

Proposición

Sea A una matriz diagonalizable $n \times n$ de rango r. Entonces existen una matriz S de tamaño $n \times r$, una matriz W de tamaño $r \times n$ y una matriz diagonal Λ_1, $r \times r$ tales que

$$A = S \Lambda_1 W$$

donde $WS = I_r$.

Aplicación

Reajuste del Nivel de Referencia Norteamericano

Imagine comenzar un proyecto imponente que se estima tomará diez años concluir y requiere el esfuerzo de decenas de personas para estructurar y resolver un sistema de 1,800,000 por 900,000 ecuaciones lineales. Esto es exactamente lo que se hizo en 1974 durante el llamado Sondeo Geodésico Nacional, cuando se propuso actualizar el Nivel de Referencia Norteamericano (NAD, por sus siglas en inglés) —una red con 268,000 puntos de referencia cuidadosamente medidos y marcados que cubren todo el territorio de América del Norte situado al norte del Istmo de Panamá, junto con Groenlandia, Hawai, las Islas Vírgenes, Puerto Rico y otras islas del Caribe.

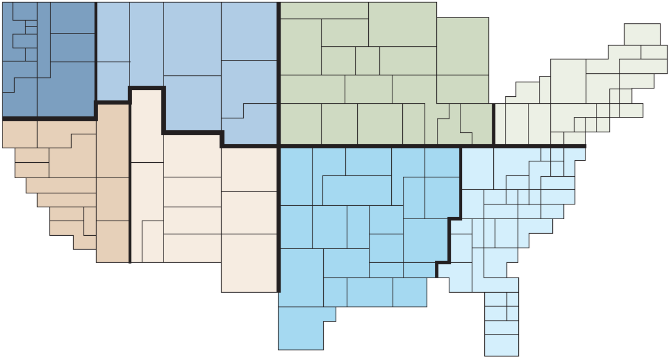
Las latitudes y longitudes registradas en el NAD deben determinarse con exactitud de unos pocos centímetros puesto que forman la base para trazar todos los planos, mapas, límites legales de la propiedad, fronteras estatales y regionales, y organizar proyectos de ingeniería civil como carreteras y líneas públicas de transmisión de electricidad. Desde el último ajuste —de los puntos de referencia geodésicos realizado en 1927, más de 200,000 nuevos puntos habían sido añadidos a un viejo conjunto de mediciones. Los errores se fueron acumulando gradualmente a través de los años, y en algunos lugares el terreno mismo se ha desplazado (hasta 5 centímetros por año). Hacia 1970 ya era urgente reacondicionar el sistema por completo y se hicieron planes para determinar un nuevo conjunto de coordenadas para los puntos de referencia.

Los datos de mediciones recopilados a lo largo de un periodo de 140 años debían convertirse a una forma legible por computadora, y los propios datos tenían que estandarizarse. (Por ejemplo, se usaron modelos matemáticos de los movimientos de la corteza terrestre para actualizar las mediciones efectuadas años atrás a lo largo de la falla de San Andrés en California.) Después de eso, había que comparar las mediciones para identifi car errores surgidos de los datos originales o de los introducidos en la computadora. Los cálculos fi nales comprendían aproximadamente 1.8 millones de observaciones, ponderadas según su precisión relativa y cada una dando lugar a una ecuación.

El sistema de ecuaciones del NAD no tenía solución en el sentido común, pero sí una solución por mínimos cuadrados, la cual asignaba latitudes y longitudes a los puntos de referencia de tal modo que correspondieran de la mejor manera posible a los 1.8 millones de observaciones. Se encontró la solución de mínimos cuadrados al resolver un sistema lineal relacionado de ecuaciones normales, el cual incluía 928,735 ecuaciones con 928,735 variables.

Como las ecuaciones normales eran demasiado grandes para las computadoras existentes, se descompusieron en sistemas más pequeños mediante una técnica llamada bloqueo de Helmert, la cual partía de manera recursiva la matriz de coefi cientes en bloques cada vez más pequeños. Los bloques menores proporcionaban ecuaciones para bloques geográfi camente contiguos de 500 a 2000 puntos de referencia del NAD. En la figura 1 se muestra cómo se subdividió Estados Unidos para conformar estos bloques de Helmert. Luego de algunos pasos intermedios se utilizaron las soluciones de los sistemas más pequeños para producir los valores fi nales de todas las 928,735 variables.

En 1983 se completó la base de datos para el reajuste del NAD. Tres años después, tras un análisis extenso y más de 940 horas de procesamiento en computadora, se resolvió el mayor problema de mínimos cuadrados jamás intentado.



Fronteras de los bloques de Helmert contiguos para Estados Unidos.

Definición

Sea X un espacio vectorial sobre K. Un producto interno sobre X es una función $\langle , \rangle : X \times X \to K$ que cumple que:

i) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

ii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

iii) $\langle x, y \rangle = \langle y, x \rangle$

iv) $\langle x, x \rangle \geq 0$ *y* $\langle x, x \rangle = 0 \Leftrightarrow x = 0$

para todo x, y, $z \in X$, $\alpha \in K$.

Definición

Sea X un espacio vectorial. X es un espacio con producto interno si tiene un producto interno definido sobre X.

Todo producto interno define una función

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

llamada la norma de v.

Un vector cuya longitud es 1, es decir, $\|v\| = 1$, se llama *vector unitario*. Si se divide un vector v diferente de cero entre su longitud - esto es, se puntiblica por $1/\|v\|$ - se obtiene un vector unitario u porque la longitud de u es $(1/\|v\|)\|v\|$. El proceso de crear a u a partir de v se denomina *normalización de* v, y se dice que u está en la misma dirección que v.

Proposition (Desigualdad de Schwarz)

Sea X un espacio con producto interno. Entonces

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

para cualesquiera x, $y \in X$. La igualdad se cumple si y solo si x y y son paralelos.

Corolario (Desigualdad del triángulo)

Sea X un espacio con producto interno. Entonces

$$\|x+y\| \leq \|x\| + \|y\|$$

para cualesquiera x, $y \in X$. La igualdad se cumple si y solo si x y y son paralelos.

La distancia entre dos vectores está definida por

$$d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle}.$$

En $\mathbb{R}^n$, un producto interno está definido por

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j$$

donde $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. interpretando x, y como vectores

$$\langle x, y \rangle = x \cdot y = x^t y$$

Definición

Sea X un espacio con producto interno. u, $v \in X$ son ortogonales si $\langle u, v \rangle = 0$.

$u \perp v$

Proposición (Ley del paralelogramo)

u y v son ortogonales si y solo si

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Definición

Sea V un espacio con producto interno y $W \subset V$ un subespacio. El complemento ortogonal de W es

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \quad \forall w \in W\}$$

Proposición

$W^{\perp}$ es un subespacio de V.

Definición

Dos conjuntos son ortogonales si todos los elementos de uno son ortogonales con todos los elementos del otro.

Proposición

Sea V un espacio con producto interno y sea $S = \{u_1, \ldots, u_p\}$ un conjunto ortogonal de vectores en V distintos de cero. Entonces S es linealmente independiente.

Proposición

Sea V un espacio con producto interno y $W \subset V$ un subespacio. Una base ortogonal para W es una base para W que también es un conjunto ortogonal.

Proposición

Sea V un espacio con producto interno y $W \subset V$ un subespacio. Sea $\{u_1, \ldots, u_p\}$ una base ortogonal para W. Para cada y en W, los pesos en la combinación lineal son:

$$y = c_1 u_1 + \cdots + c_p u_p$$

donde

$$c_j = \frac{\langle y, u_j \rangle}{\langle u_j, u_j \rangle}$$

Teorema

Sea u un vector diferente de cero en $\mathbb{R}^n$. Entonces cualquier y de $\mathbb{R}^n$ se puede escribir de manera única como

$$y = \hat{y} + z$$

donde $\hat{y} = \alpha u$ para algún escalar α y z es un vector ortogonal a u.
Dado cualquier escalar α, sea $z = y - \alpha u$ de manera que $y = \hat{y} + z$ se cumple. Entonces $y - \hat{y}$ es ortogonal a u si y sólo si:

$$0 = (y - \alpha u) \cdot u = y \cdot u - (\alpha u) \cdot u = y \cdot u - \alpha (u \cdot u)$$

Esto es, $y = \hat{y} + z$ se cumple con z ortogonal a u si y sólo si

$$\alpha = \frac{y \cdot u}{u \cdot u}$$

$$\hat{y} = \frac{y \cdot u}{u \cdot u} u$$

$\hat{y}$ se denota por $\text{proy}_L y$ y se le llama **proyección ortogonal de** y **sobre** L, es decir:

$$\hat{y} = \text{proy}_L y = \frac{y \cdot u}{u \cdot u} u$$

Sea V un espacio con producto interno y $W \subset V$. Además, $y \in V$. En la descomposición $y = w + z$, al vector w se le llama la proyección ortogonal de y sobre W, $w = \text{proy}_W y$.

Proposición

i) $proy_W(proy_W y) = proy_W y$

ii) *Si* $y \in W$, $proy_W y = y$.

Proposición

Un conjunto $\{u_1, \ldots, u_p\}$ *es un conjunto ortonormal si es un conjunto ortogonal de vectores unitarios. Si* W *es el subespacio generado por un conjunto de este tipo, entonces* $\{u_1, \ldots, u_p\}$ *es una base ortonormal para* W*, puesto que el conjunto es, de manera automática, linealmente independiente.*

Proposición

Una matriz U de $m \times n$ tiene columnas ortonormales si y solo si $U^t U = I$.

Teorema

Sea U una matriz de $m \times n$ con columnas ortonormales, y sean x y y vectores en $\mathbb{R}^n$. Entonces:

- $\|Ux\| = \|x\|$
- $(Ux) \cdot (Uy) = x \cdot y$
- $(Ux) \cdot (Uy) = 0$ *si y sólo si* $x \cdot y = 0$

Proposición

Una matriz cuadrada es ortogonal si y solo si tiene columnas ortonormales.

El teorema es cierto también para los renglones de U.

Las columnas (renglones) de una matriz ortogonal forman una base ortonormal de $\mathbb{R}^n$ si U es de tamaño $n \times n$

Definición

Una matriz ortogonal es una matriz U cuadrada invertible tal que si $U^{-1} = U^t$.

Proposition (Descomposición ortogonal)

Sea W un subespacio de $\mathbb{R}^n$. Entonces toda y en $\mathbb{R}^n$ puede escribirse únicamente en la forma

$$y = \hat{y} + z$$

donde $\hat{y}$ está en W y z en $W^{\perp}$. De hecho, si $\{u_1, \ldots, u_p\}$ es cualquier base ortogonal de W, entonces

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \ldots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$$

y

$$z = y - \hat{y}$$

Proposición (La mejor aproximación)

Sea V un espacio con producto interno, $W \subset V$ un subespacio, $y \in V$ y $\hat{y} = proy_W y$. Entonces $\hat{y}$ es el vector en W más cercano a y, en el sentido que

$$\|y - \hat{y}\| < \|y - v\|$$

para todo $v \in W$, $v \neq \hat{y}$.

Es decir, $\hat{y}$ es la mejor aproximación a y de los elementos de W.

Proposición

Sea V un espacio con producto interno y $W \subset V$ un subespacio. Si $\{u_1, \ldots, u_p\}$ es una base ortonormal de W, entonces dado $y \in V$

$$proy_W y = (y \cdot u_1)u_1 + \cdots (y \cdot u_p)u_p$$

y

$$proy_W y = UU^t y$$

donde $U = (u_1 \cdots u_p)$.

El proceso de Gram-Schmid es un algoritmo sensillo para producir una base ortogonal u ortonormal para cualquier subespacio diferente de cero de $\mathbb{R}^n$.

Recordar que para normalizar vectores, basta dividirlos entre su norma.

Proceso de ortogonalización de Gram-Schmidt

A partir de una base $\{x_1, x_2, \ldots, x_p\}$ para un subespacio W de $\mathbb{R}^n$, se define:

$$
\begin{aligned}
v_1 &= x_1 \\
v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
v_4 &= x_4 - \frac{x_4 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_4 \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{x_4 \cdot v_3}{v_3 \cdot v_3} v_3 \\
&\vdots \\
v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}
\end{aligned}
$$

Entonces, $\{v_1, v_2, \ldots, v_p\}$ es una base ortogonal para W. Además,

$$\text{Gen}\{x_1, x_2, \ldots, x_k\} = \text{Gen}\{v_1, v_2, \ldots, b_k\}$$

para $1 \leq k \leq p$.

Factorización QR

Si A es una matriz de $m \times n$ con columnas linealmente independientes, entonces A se puede factorizar como $A = QR$, donde Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para $\mathcal{C}(A)$, y R es una matriz triangular superior invertible de $n \times n$ con entradas positivas en su diagonal.

Mínimos cuadrados

Si el sistema $Ax = b$ no tiene solución, i.e., es inconsistente, podríamos tratar de encontrar x tal que Ax este tan cercano a b como sea posible, es decir, la mejor aproximación a la solución. El problema general de mínimos cuadrados es encontrar x de tal manera que la distancia $\|b - Ax\|$ sea tan pequeña como sea posible.

Definición

Sea A una matriz $m \times n$ y $b \in \mathbb{R}^m$. Una solución de mínimos cuadrados del sistema $Ax = b$ es un $\hat{x} \in \mathbb{R}^n$ tal que

$$\|b - A\hat{x}\| \leq \|b - Ax\|$$

para todo $x \in \mathbb{R}^n$.

El vector Ax está en $\mathcal{C}(A)$. El vector en $\mathcal{C}(A)$ más cercano a b está dado por la proyección ortogonal de b sobre $\mathcal{C}(A)$, $\hat{b}$.

Si $A\hat{x} = \hat{b}$, entonces $\hat{x}$ satisface

$$A^t A x = A^t b$$

Definición

El sistema $A^t A x = A^t b$ se llama el sistema de ecuaciones normales de $Ax = b$.

Proposición

El conjunto de soluciones de mínimos cuadrados del sistema $Ax = b$ es el conjunto de soluciones del sistema de ecuaciones normales $A^t A x = A^t b$.

Proposición

Son equivalentes

1. *La solución de mínimos cuadrados de $Ax = b$ es única*
2. *Las columnas de A son linealmente independientes*
3. *$A^t A$ es invertible.*

En cualquier caso de los anteriores

$$\hat{x} = (A^t A)^{-1} A^t b$$

En general $\|b - A\hat{x}\|$ es llamado el error de mínimos cuadrados.

Proposición

Sea A una matriz $m \times n$ con columnas linealmente independientes y $A = QR$, donde QR es su factorización. Entonces dado $b \in \mathbb{R}^m$, la solución de mínimos cuadrados de $Ax = b$ está dada por

$$\hat{x} = R^{-1} Q^t b$$

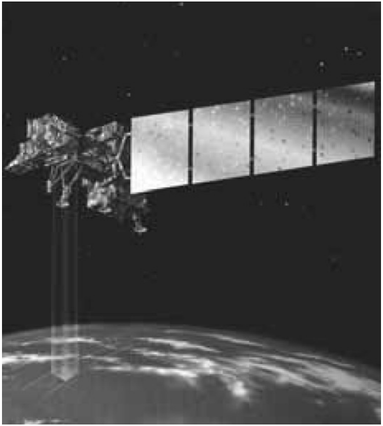
Dado que R es triangular superior, es mucho más rápido resolver:

$$Rx = Q^t b$$

Aplicación

Procesamiento de imágenes multicanal

Dando la vuelta al mundo en poco más de 80 minutos, los dos satélites Landsat cruzan el cielo como un rayo silencioso con órbitas casi polares, graban imágenes del terreno y de las líneas costeras en franjas de 185 kilómetros de ancho. En periodos de 16 días, estos satélites pasan sobre casi todos los kilómetros cuadrados de la superficie terrestre, de modo que cualquier lugar se puede monitorear cada 8 días.



Las imágenes Landsat son útiles para muchos propósitos. Los desarrolladores y planifi cadores urbanos las usan para estudiar el ritmo y la dirección del crecimiento urbano, el desarrollo industrial, y otros cambios en el uso del suelo. Las comunidades rurales pueden analizar la humedad del suelo, clasifi car la vegetación de áreas remotas, y localizar depósitos y corrientes de agua tierra adentro. Los gobiernos pueden detectar y estimar los daños debidos a desastres naturales, como incendios forestales, fl ujos de lava, inundaciones y huracanes. Las agencias de protección del medio ambiente pueden identifi car la contaminación por emisiones de chimeneas y medir la temperatura del agua de lagos y ríos cercanos a plantas de energía.

Los sensores colocados a bordo de los satélites obtienen siete imágenes simultáneas de cualquier región de la Tierra que se vaya a estudiar. Estos sensores registran la energía en diferentes bandas de longitud de onda —tres en el espectro de luz visible y cuatro en las bandas de infrarrojo y térmico—. Cada imagen se digitaliza y archiva como una formación rectangular de números, donde cada número indica la intensidad de la señal en un pequeño punto (o píxel) correspondiente de la imagen. Cada una de las siete imágenes es un canal de una imagen multicanal o multiespectral.

Las siete imágenes Landsat de una región fija suelen contener mucha información redundante, puesto que algunas características aparecen en varias imágenes. Sin embargo, otras características, por su color o temperatura, pueden refl ejar luz que registran únicamente uno o dos sensores. Una meta del procesamiento de imágenes multicanal es la de visualizar los datos de manera que la información se extraiga de mejor modo que estudiando cada imagen por separado.

El análisis de componentes principales es una manera efectiva de eliminar información redundante y de proporcionar en una sola o en dos imágenes compuestas la mayor parte de la información proveniente de los datos iniciales. A grandes rasgos, el objetivo principal es encontrar una combinación lineal especial de las imágenes, es decir, una lista de pesos que combinen en cada píxel los siete valores correspondientes de las imágenes en un nuevo valor. Los pesos se eligen de tal manera que hagan al intervalo de intensidades de luz —la varianza de la escena— de la imagen compuesta (llamada primera componente principal) mayor que en cualquiera de las imágenes originales. También se pueden estructurar imágenes adicionales de componentes

El análisis de componentes principales se ilustra en las siguientes fotografías, tomadas sobre el valle Railroad en Nevada, EUA. En (a), (b) y (c) se muestran las imágenes provenientes de tres bandas espectrales Landsat. La información total de las tres bandas se reacomoda en tres imágenes de componentes (d), (e) y (f). El primer componente (d) despliega (o "explica") 93.5 % de la varianza de la escena presente en los datos iniciales. De esta manera, los datos iniciales de tres canales se han reducido a datos de un canal, con una pérdida en algún sentido de sólo el 6.5 % de la varianza de la escena.

La empresa Earth Satellite Corporation de Rockville, Maryland, que amablemente proporcionó las fotografías mostradas, está experimentando con imágenes de 224 bandas espectrales individuales. El análisis de componentes principales, que resulta indispensable al tratar con tales conjuntos masivos de datos, a menudo reduce los datos a aproximadamente 15 componentes principales utilizables.

a) Banda espectral 1: Azul visible.

b) Banda espectral 4: Casi infrarrojo.

c) Banda espectral 7: Infrarrojo medio.

$$S = \begin{pmatrix} 2382.78 & 2611.84 & 2136.20 \\ 2611.84 & 3106.47 & 2553.90 \\ 2136.20 & 2553.90 & 2650.71 \end{pmatrix}$$

Aplicación

$$\lambda_1 = 7614.23 \qquad \lambda_2 = 427.63 \qquad \lambda_3 = 98.10$$

$$u_1 = \begin{pmatrix} 0.5417 \\ 0.6295 \\ 0.5570 \end{pmatrix} \quad u_2 = \begin{pmatrix} -0.4894 \\ -0.3026 \\ 0.8179 \end{pmatrix} \quad u_3 = \begin{pmatrix} 0.6834 \\ -0.7157 \\ 0.1441 \end{pmatrix}$$

El primer componente captura el 93.5% de la varianza total.

$$y_1 = 0.5417x_1 + 0.6295x_2 + 0.5570x_3$$

Aplicación

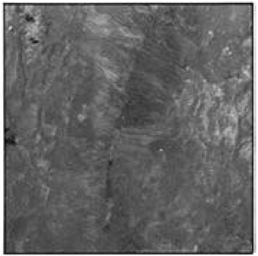
a) Banda espectral 1: Azul visible.

b) Banda espectral 4: Casi infrarrojo.

c) Banda espectral 7: Infrarrojo medio.

d) Componente principal 1: 93.5%.

e) Componente principal 2: 5.3%.



f) Componente principal 3: 1.2%.

Matrices simétricas y formas cuadráticas

Proposición

Si A es simétrica real, todos sus valores propios son reales.

Proposición

Si A es simétrica real, λ_1, λ_2 son valores propios distintos y $x_1 \in E_{\lambda_1}$, $x_2 \in E_{\lambda_2}$, entonces x_1 y x_2 son ortogonales.

Si A es diagonalizable, $A = PDP^{-1}$ con P ortogonal, entonces $A = PDP^t$. En tal caso diremos que A es diagonalizable ortogonalmente.

Proposición

Si A es diagonalizable ortogonalmente, entonces A es simétrica

Proposición

Si A es simétrica entonces existe una matriz ortogonal P y una matriz diagonal D tal que $A = PDP^t$.

En el caso de una matriz simétrica, la descomposición espectral toma la forma

$$A = \lambda_1 u_1 u_1^t + \lambda_2 u_2 u_2^t + \cdots \lambda_n u_n u_n^t$$

donde los u_i son los vectores de P en la descomposición $A = PDP^t$ de A.

En una matriz simétrica la multiplicad algebraica es la misma que la multiplicidad geométrica.

Definición

Una forma cuadrática es una función $Q : \mathbb{R}^n \to \mathbb{R}$ *de la forma*

$$Q(x) = x^t Bx$$

donde B *es una matriz* $n \times n$ *y* x *es el vector correspondiente al valor de* $\mathbb{R}^n$.

Proposición

Dada una forma cuadrática $Q(x) = x^t Bx$, *siempre se puede encontrar una matriz simétrica* A *tal que* $Q(x) = x^t Ax$ *para todo* $x \in \mathbb{R}^n$.

Definición

Una forma cuadrática es una función $Q : \mathbb{R}^n \to \mathbb{R}$ *de la forma*

$$Q(x) = x^t A x$$

donde A *es una matriz simétrica* $n \times n$ *y* x *es el vector correspondiente al valor de* $\mathbb{R}^n$.

Proposición

Si A *es simétrica, entonces* $x^t A x = 0$, $\forall x \in \mathbb{R}^n$ *si y solo si* $A = 0$.

Proposición

Si $Q(x) = x^t A x$ *es una forma cuadrática, con* A *simétrica, entonces existe una matriz ortogonal* P *tal que* $A = PDP^t$, *haciendo el cambio de variable* $y = P^{-1}x$, $y^t D y$ *toma los mismos valores que* Q *sin tener productos cruzados.*

Definición

Una forma cuadrática Q es:

i) *positiva definida si* $Q(x) > 0, \forall x \neq 0$

ii) *no negativa definida si* $Q(x) \geq 0, \forall x \neq 0$

iii) *negativa definida si* $Q(x) < 0, \forall x \neq 0$

iv) *no positiva definida si* $Q(x) \leq 0, \forall x \neq 0$

v) *indefinida si asume ambos valores, i.e. no es n.p.d ni n.n.d*

Definición

Una matriz simétrica A es positiva definida, no negativa definida, negativa definida, no positiva definida o indefinida si su correspondiente forma cuadrática lo es.

Proposición

Sea A una matriz simétrica, Q su correspondiente forma cuadrática. Entonces A (o la forma Q) es:

i) *positiva definida si y solo si todos sus valores propios son positivos*
ii) *no negativa definida si y solo si todos sus valores propios son no negativos*
iii) *negativa definida si y solo si todos sus valores propios son negativos*
iv) *no positiva definida si y solo si todos sus valores propios son no positivos*
v) *indefinida ssi y solo si sus valores propios son negativos y positivos*

Descomposición en valores singulares

Sea A una matriz $m \times n$. Observemos que:

1. $A^t A$ es simétrica
2. Los valores propios de $A^t A$ son no negativos

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0$$

Definición

Sea A una matriz $m \times n$. Los valores singulares de A son $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, n$.

Observación: Los valores singulares son las longitudes de los vectores Av_i para $\{v_i\}$ los vectores propios correspondientes.

Proposición

Sean $\{v_i\}$ *una base ortonormal de* $\mathbb{R}^n$ *de los vectores propios correspondientes a los valores propios* $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n$ *de* $A^t A$. *Si* A *tiene* r *valores singulares diferentes de cero, entonces* $\{Av_1, \ldots, Av_r\}$ *es una base ortogonal de* $\mathcal{C}(A)$ *y* $\rho(A) = r$.

Teorema (La descomposición en valores singulares)

Sea A una matriz $m \times n$ de rango r. Entonces existe una matriz Σ de tamaño $m \times n$ de la forma

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

donde D es una matriz diagonal $r \times r$ que tiene como entradas los primeros r valores singulares de A, $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$ (en ese orden), y existen una matriz ortogonal U de $m \times m$ y una matriz ortogonal V de $n \times n$ tal que

$$A = U\Sigma V^t$$

Definición

La descomposición en valores singulares de una matriz no es única. Si $r = \rho(A)$ y $A = U\Sigma V^t$,

$$U = \begin{pmatrix} U_r & U_{m-r} \end{pmatrix}$$

donde $U_r = (u_1 \cdots u_r)$, $m \times r$, está formada por las primeras r columnas,

$$V = \begin{pmatrix} V_r & V_{n-r} \end{pmatrix}$$

donde $V_r = (v_1 \cdots v_r)$, $n \times r$, entonces

$$A = \begin{pmatrix} U_r & U_{m-r} \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_r^t \\ V_{n-r}^t \end{pmatrix} = U_r D V_r^t$$

Esta es la descomposición en valores singulares reducida de A.

Definición

Sea A una matriz $m \times n$. La seudoinversa o la inversa de Moore-Penrose de A está dada por

$$A^{\dagger} = V_r D^{-1} U_r^t$$

donde $A = U_r D V_r^t$ es una descomposición en valores singulares reducida de A.

Proposición

Propiedades de la inversa de Moore-Penrose.
Sea $A \in \mathbb{C}^{m\times n}$*, entonces:*

- $A^* = A^*AA^\dagger = A^\dagger AA^*$. *En consecuencia, si* $\rho(A) = n$ *entonces* $A^\dagger = (A^*A)^{-1}A^*$*, y si* $\rho(A) = m$ *entonces* $A^\dagger = A^*(AA^*)^{-1}$.
- $Im(I - AA^\dagger) = [Im(A)]^\perp$ *y* $Nul(AA^\dagger) = Nul(A^\dagger)$.
- $Im(A^\dagger) = Im(A^*)$. *En consecuencia,* $\rho(A^\dagger) = \rho(A)$.
- $Nul(A^\dagger) = Nul(A^*)$.

Considerando de nuevo el sistema $Ax = b$, sea $\hat{x} = A^\dagger b$.

Entonces

$$A\hat{x} = U_r U_r^t b$$

de donde $U_r U_r^t b$ es la proyección ortogonal de b sobre $\mathcal{C}(A)$. Por lo tanto $\hat{x}$ es una solución de mínimos cuadrados de $Ax = b$.

Definición

$\hat{x} = A^\dagger b$ es la solución de mínimos cuadrados de $Ax = b$ de norma mínima entre todas las soluciones de mínimos cuadrados.