MATRICES

Fundamentos, Transformaciones y Álgebra Espectral

Editor:

William Alexander Ávila Aguilar

Autores:

Gladis Del Rocio Mora Veintimilla
Eduardo Antonio Vázquez Aldaco
Lizania Godina Young
Lethzaira Mercedes Contreras Borjon
Hernán Alonso Mancipe Bohórquez
Marcos Chacón Castro

MATRICES TRANSFORMACIONES LINEALES EIGENVALORES

PRIMERA EDICIÓN • 2026

MATRICES

William Alexander Ávila Aguilar
(Comp)

Gladis Mora
Eduardo Vázquez
Lizania Godina
Lethzaira Contreras
Hernán Mancipe
Marcos Chacón

Didaxis
Editorial

Montecristi – Manabí - Ecuador

Del autor: William Alexander Ávila Aguilar (Compilador), Gladis Del Rocio Mora Veintimilla, Eduardo Antonio Vázquez Aldaco, Lizania Godina Young, Lethzaira Mercedes Contreras Borjon, Hernán Alonso Mancipe Bohórquez y Marcos Chacón Castro

Matrices/ Ávila Aguilar et al.- Editorial Didaxis, 2025

135 p.; 14,8 cm x 21 cm

ISBN 978-9942-7491-2-3

1. Matrices 2. Transformaciones lineales 3. Eigenvalores, 4. Ortogonalidad

370 – Educación / JNUM – Recursos y materiales didácticos para docentes

DOI: 10.64325/1fjan565

Titulo original:
Matrices
Copyright © Editorial Didaxis, 2025
Publicado en español
Reservados todos los derechos
Prohibida la reproducción total o parcial de esta obra sin permiso escrito de la Editorial

C.C. 1 190587
ISBN 978-9942-7491-2-3

Palabras al lector

Matrices presenta un tratamiento riguroso y sistemático de la teoría de matrices desde una perspectiva algebraica lineal. Cada noción matricial se ancla en el marco axiomático de los espacios vectoriales y las transformaciones lineales sobre un campo F, estableciendo una correspondencia precisa entre objetos algebraicos abstractos y sus representaciones matriciales concretas. Esta elección metodológica dota al lector de una comprensión estructural que trasciende la mera manipulación simbólica, siendo de especial relevancia para aplicaciones en ingeniería, análisis numérico, optimización y sistemas de control.

La obra se despliega en una progresión conceptual coherente que abarca desde la definición formal de matriz y sus propiedades dimensionales, pasando por la multiplicación matricial como expresión de la composición de

transformaciones lineales, el análisis del núcleo e imagen mediante el Teorema de la Dimensión, la teoría de productos escalares, ortogonalidad y el proceso de Gram-Schmidt, hasta el estudio del rango y la solución de sistemas de ecuaciones lineales. En los tramos finales, el texto desarrolla la factorización de polinomios sobre campos complejos, el Teorema Fundamental del Álgebra y la teoría espectral de operadores lineales, incluyendo eigenvalores, eigenvectores, subespacios invariantes y la noción de similaridad matricial derivada del cambio de base.

Cada definición precede a los resultados que de ella se derivan, y cada demostración se construye de manera autocontenida haciendo explícito el rol de cada hipótesis. Las demostraciones recurren a inducción matemática, prueba por contradicción y argumentos de dimensión, patrones de razonamiento que el lector internaliza

progresivamente. Esta rigurosidad se complementa con ejemplos concretos y ejercicios de complejidad graduada, incluyendo problemas de alta exigencia que requieren la combinación de múltiples herramientas teóricas.

William Alexander Ávila Aguilar
Compilador

Contenido

Matrices 12

Definición 1 12

Teorema 1 13

Definición 2 14

Multiplicación Matricial 15

Composición Matricial, generalización 16

Teorema 2 19

Ejercicios 21

Kernel e Imagen de una transformación lineal. Análisis matricial. 22

Teorema 3 24

Teorema 4 25

Imagen de una transformación lineal. Análisis en matrices. 27

Teorema 5 28

Productos Escalares y Ortogonalidad 32

Definición 3 33

Ortogonal 39

Proyección de un elemento Y de Fn a lo largo de un elemento X de Fn (cX) 46

Definición 4 49

Teorema 6 49
Corolario 1 51
Teorema 7 54
Productos Hermitianos 56
Aplicaciones a ecuaciones lineales; El Rango 60
Teorema 8 60
El Rango 64
Teorema 7 64
Factorización Única 69
Teorema de división 69
Número complejo 70
Propiedades de los números complejos 71
Polinomios 73
Ceros de un Polinomio 74
Factorización de polinomios sobre Complejos (C) 80
Mapa Identidad 85
Mapa Inverso 86
Isomorfismo 88
Cambio de Base 89
Eigenvalores, Eigenvectores y Similaridad 99
Cambio de base y similaridad 99

Máximo restringido y eigenvalores 100
Operador 106
Subespacio Invariante 107
Espacios Invariantes Especiales 108
Eigenvalor 109
Condiciones Equivalentes 109
Eigenvector 111
Eigenvectores Linealmente Independientes 111
Un operador no puede tener más eigenvalores que la dimensión del espacio vectorial 113
Polinomios Aplicados a Operadores 113
Notación: Tm 114
Notación: $p(T)$ 115
Producto de Polinomios 117
Propiedades Multiplicativas 117
Espacio Nulo y Rango de $p(T)$ son invariantes bajo T 119
Ejercicios de Aplicación 121
Ejercicio de Alta Complejidad 128
Bibliografía 131
Sobre los autores: 132

Matrices

Definición 1

Una matriz se puede definir como un arreglo de elementos. Ahora bien, para establecer cuáles son estos elementos, se debe aclarar sobre qué tipo de conjunto se está trabajando, por ende, por fines prácticos, estos conjuntos de interés son: el conjunto de números reales y el conjunto de los números complejos denotados por $\mathbb{R}$ y $\mathbb{C}$ respectivamente. Cuando no se especifique sobre qué conjunto de elementos se está estudiando, se utilizará la letra F, que denota o bien los reales o complejos. Con estas aclaraciones, una matriz es un arreglo de elementos sobre F cuya representación o estructura es de la siguiente forma:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \qquad 1.1$$

Para poder interpretar al arreglo 1.1 se debe tener en cuenta que la letra n hace referencia a la cantidad de filas de dicho arreglo, mientras que la letra m hace referencia a la cantidad de columnas de la matriz. Por ejemplo, el elemento a_{11} que pertenece al campo F se encuentra representado en la matriz A en la fila 1 columna 1. Más aún, el elemento a_{ij} se encuentra ubicado en la fila i columna j. En este capítulo de libro, se asumirá que el lector conoce o tiene noción de los conceptos de espacios de vectoriales y transformaciones lineales. Resulta

de gran interés para el estudio de problemas de ingeniería y matemática cómo se vincula los arreglos matriciales. Se definirá a continuación ciertos detalles que hacen relevancia en el armado de una estructura matricial.

Sea un arreglo matricial $\boldsymbol{A = [a_{ij}]}$ de tamaño de $\boldsymbol{nxm}$. Se define como columnas de $\boldsymbol{[a_{ij}]}$ a los elementos que conforman todas y cada una de las columnas de $\boldsymbol{[a_{ij}]}$. Se denotan por:

$$\boldsymbol{A = [A^1\ A^2 \dots A^i \dots A^m],\quad i = 1, \dots, m \quad 1.2}$$

Es importante señalar a qué nos referimos cuando hablamos sobre el tamaño $\boldsymbol{nxm}$. Es fundamental establecer y definir lo que significa realmente el tamaño de una matriz $\boldsymbol{nxm}$. Para ello, se establece el siguiente teorema.

Teorema 1

Sea una matriz $\boldsymbol{A = [a_{ij}]}$ de n filas y m columnas sobre F. Entonces su dimensión es $(n)(m)$.

Demostración. Sea una matriz $\boldsymbol{A = [a_{ij}]}$ definida como:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

Esta matriz se puede escribir de la siguiente manera:

$$[a_{ij}] = a_{11}\begin{bmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix} + \cdots a_{1m}\begin{bmatrix}0 & \cdots & 1\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix}$$
$$+ a_{21}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 1 & 0 & \ldots & 0 & 0\\ 0 & 0 & \ldots & 0 & 0\end{bmatrix}$$
$$+ a_{22}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 0 & 1 & & 0 & 0\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ 0 & 0 & \ldots & 0 & 0\end{bmatrix}$$
$$+ \cdots a_{2m}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 0 & 0 & & 0 & 1\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ 0 & 0 & \ldots & 0 & 0\end{bmatrix} + \cdots$$
$$+ a_{n1}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 0 & 0 & & 0 & 0\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ 1 & 0 & \ldots & 0 & 0\end{bmatrix}$$
$$+ a_{n2}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 0 & 0 & & 0 & 0\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ 0 & 1 & \ldots & 0 & 0\end{bmatrix}$$
$$+ \cdots a_{nm}\begin{bmatrix}0 & 0 & \ldots & 0 & 0\\ 0 & 0 & & 0 & 0\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ 0 & 0 & \ldots & 0 & 1\end{bmatrix}$$

Por lo tanto, hay de tal forma $(n)(m)$ matrices lo que iguala la dimensión $(m)(n)$.

Definición 2

Suponga que $T \epsilon L(V, W)$ donde T se define como una transformación lineal sobre los espacios vectoriales V, W sobre un campo F. Sean $v_1, v_2, \ldots, v_n$ una base de V y sean $w_1, w_2, \ldots, w_m$ una base de W. La matriz de T con respecto a

estas bases es la matriz $m\,x\,n$ definida como $M(T)$ cuyas entradas a_{ij} están definidas por:

$$Tv_k = a_{1k}w_1 + a_{2k}w_2 + \cdots a_{mk}w_m$$

Multiplicación Matricial

Considere los mapas lineales $T: U \to V$ y $S: V \to W$ en donde U, V, W son espacios vectoriales sobre un campo F. Considere sus respectivas bases vectoriales $u_1, u_2, \ldots, u_p$ como la base de U, en la que $v_1, v_2, \ldots, v_n$ es una base de V, y en la que $w_1, w_2, \ldots, w_m$ es una base de W. Resulta que para la composición ST, para $1 \leq k \leq p$ tenemos que:

$$\begin{aligned}(ST)u_k = S(Tu_k) &= S(c_{1k}v_1 + c_{2k}v_2 + \cdots c_{nk}v_n)\\ &= S(c_{1k}v_1) + S(c_{2k}v_2) + \cdots + S(c_{nk}v_n)\\ &= S\left(\sum_{r=1}^{n} c_{rk}v_r\right) = \sum_{r=1}^{n} c_{rk}S(v_r)\\ &= \sum_{r=1}^{n} c_{rk}(a_{1r}w_1 + a_{2r}w_2 + \cdots + a_{mr}w_m)\\ &= \sum_{r=1}^{n} c_{rk}\left(\sum_{j=1}^{m} a_{jr}w_j\right) = \sum_{j=1}^{m}\left(\sum_{r=1}^{n} a_{jr}c_{rk}\right)w_j\end{aligned}$$

De hecho, se forma el arreglo matricial o la matriz de la composición de transformaciones lineales $M(ST)$ cuyo tamaño es $m\ x\ p$ cuya entrada en la fila j columna k iguala a

$$\sum_{r=1}^{n} a_{jr}c_{rk}$$

Bajo esta decisión sobre la composición de transformaciones lineales usando sus respectivas definiciones, se produce la multiplicación entre matrices. Cabe notar que para que la multiplicación pueda ejecutarse la cantidad de columnas de la primera matriz a la izquierda debe ser la misma cantidad de filas que la matriz a la derecha en la se realiza el producto matricial.

Composición Matricial, generalización

Sean U, V, W conjuntos. Sea $F: U \to V$ y $G: V \to W$ mapas. Entonces podemos formar el mapeo compuesto desde $U \to W$ denotado por $G°F$. Es por definición el mapeo definido por:

$$G°F(t) = G(F(t))$$

Es bastante interesante asociar esta definición de regla de correspondencia o mapeo en la composición matricial. Es exactamente el análisis anterior realizado, la diferencia es que a cada función o mapeo le corresponde una matriz asociada o la matriz de esa transformación lineal y su composición no es otra cosa más que la multiplicación entre esas dos matrices. Es por ello que la cantidad o longitud de la base de V debe ser exactamente la misma que la longitud de la base del espacio vectorial del dominio de G que en este caso es V. Es claro que por ello la multiplicación matricial puede ejecutarse bajo esa condición del número de columnas de la matriz izquierda

debe ser igual al número de filas de la matriz derecha. En otras palabras, sea A una matriz mxn y sea S una matriz sobre un campo F. Resulta que la multiplicación entre la matriz A y S puede ejecutarse y darse bajo la condición de que la cantidad de filas de S sea exactamente n. Por ende,

$$A_{mxn}S_{nxp} = H_{mxp}$$

De cierta forma, podemos analizar el siguiente ejemplo:

$$2x - 7y = 0$$

Esta misma expresión en forma matricial podría expresarse como:

$$[2 \ -7]_{1x2}\begin{bmatrix} x \\ y \end{bmatrix}_{2x1} = [0]_{1x1}$$

El ejemplo mostrado para esa ecuación pudo generarse de forma intuitiva para asociarlo a un análisis matricial. Lo interesante es que el resultado de su multiplicación genera de hecho una matriz de tamaño $1x1$. En ciertas ocasiones, utilizando las reglas del algebra lineal, podemos generalizar y motivar las definiciones correspondientes a la multiplicación entre matrices la cual hace válido los esquemas y estructuras establecidos.

De forma más general, sea F un campo y A un vector fijo en F^n. Tenemos un mapa lineal:

$$L_A: F^n \to F$$

Tal que $L_A(X) = XA$ para todo $X \ que \ pertenece \ a \ F^n$.

Dicho de otra manera, sea A una matriz de tamaño mxn, es decir con m filas y n columnas sobre un campo F. Obtenemos un mapa lineal

$$L_A: F^n \to F^m$$

Tal que $L_A(X) = AX$ para todo vector columna X en K^n. De nuevo, como se ha mencionado anteriormente, la linealidad sigue de las propiedades de multiplicación de matrices.

$$1.\, L_A(X + Y) = L_A(X) + L_A(Y)$$

$2.\, L_A(zX)$
$= zL_A(X), para\ todo\ z\ que\ pertenece\ a\ un\ campo\ F\ siendo\ un\ escalar$

La multiplicación entre AX se muestra algo así:

$$AX = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \cdot\cdot & \dots & \cdot\cdot \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot\cdot \\ x_n \end{bmatrix}$$

Tomar en cuenta y nuevamente resulta de interés, que al multiplicar AX que es la ejecución de multiplicación de matrices, debido a que el conjunto de matrices es un espacio vectorial, significa que su multiplicación entre matrices no es un espacio vectorial ya que la conmutatividad no se cumple. Por ejemplo, no es necesariamente cierto que $AX = XA$. De hecho, no se cumple para el caso mencionado anteriormente. Sin embargo, la multiplicación entre AX da como resultado una nueva matriz. En general la multiplicación de matrices es un conjunto de matrices de un tamaño especifico.

Teorema 2

Sean V y W espacios vectoriales. Sean $v_1, \dots, v_n$ una base de V y sean $w_1, \dots, w_n$ elementos arbitrarios de W. Entonces existe una **única transformación lineal** $T: V \to W$ tal que:

$$T(v_1) = w_1, \dots, T(v_n) = w_n$$

Si $c_1, \dots, c_n$ son números, entonces

$$T(c_1 v_1 + \cdots c_n v_n) = c_1 w_1 + \cdots + c_n w_n$$

Reescribiendo este teorema al espacio vectorial de matrices, se lo podría asociar de la siguiente manera.

Sean F^n y F^m espacios vectoriales. Sean $E^1, \dots, E^n$ una base de F^n y sean $K_1, \dots, K_n$ elementos arbitrarios de F^m. Tomar en consideración que por definición y notación F^n son aquellos vectores de elementos sobre F que constan de n filas y F^m son aquellos vectores cuyos elementos sobre F resultan ser de m filas. Entonces existe una única transformación lineal $L_A: F^n \to F^m$ tal que

$$L_A(E^1) = K_1, \dots, L_A(E^n) = K_n$$

Si $x_1, \dots, x_n$ son números $x_1, \dots, x_n$ son únicos debido a que los elementos $E^1, \dots, E^n$ son linealmente independientes, entonces

$$L_A(x_1 E^1 + \cdots + x_n E^n) = x_1 K_1 + \cdots x_n K_n$$

Demostración. En primer lugar, probaremos que tal mapa L_A satisfaciendo la requerida condición existe. Sea E un elemento de F^n, y sean $x_1, \dots, x_n$ los únicos números tales que $E =$

$x_1E^1 + \cdots + x_nE^n$. De hecho, es factible pensar que esos números Dejamos como:

$$L_A(E) = x_1K_1 + \cdots + x_nK_n$$

Hemos entonces definido un mapa L_A desde F^n hacia F^m, y afirmamos que L_A es una transformación lineal. Si E' es un elemento de F^n, y si $E' = y_1E^1 + \cdots + y_nE^n$. Entonces:

$$E + E' = (x_1 + y_1)E^1 + \cdots (x_1 + y_n)E^n$$

Por definición, obtenemos:

$$\begin{aligned} L_A(E + E') &= (x_1 + y_1)K_1 + \cdots + (x_n + y_n)K_n \\ &= x_1K_1 + y_1K_1 + \cdots + x_nK_m + y_nK_m \\ &= L_A(E) + L_A(E') \end{aligned}$$

Sea c un número. Entonces $cE = cx_1E^1 + \cdots + cx_nE^n$. Aplicando la transformación lineal L_A tenemos:

$$L_A(cE) = cx_1K_1 + \cdots + cx_nK_n$$

En consecuencia, hemos probado que L_A es lineal. Ahora bien, para el concepto de matrices, la definición de L_A en donde por definición denotamos como A a la matriz de la transformación lineal asociada, resulta que:

$$L_A(E) = AE = K$$

Es más, A o siendo más claro, el tamaño de A debe ser de $m\ x\ n$ con el fin de obtener el siguiente único resultado asociado a la matriz K:

$$L_A(E) = A_{mxn}E_{nx1} = K_{mx1}$$

Tal mapa es único, porque para cualquier elemento $x_1E^1 + \cdots + x_nE^n$ de F^n, resulta que cualquier mapa lineal $L_B: F^n \rightarrow F^m$ tal que $L_B(E^i) = K_i$ para $i = 1, \ldots, n$ debe también satisfacer:

$$L_B(x_1E^1 + \cdots + x_nE^n) = x_1L_B(E^1) + \cdots + x_nL_B(E^n)$$
$$= x_1K_1 + \cdots + x_nK_n$$

En consecuencia:

$$L_A(E) = L_B(E) = AE = BE = BE = AE = K$$

Por tanto, encontramos otro resultado más interesante. Habiendo ya definido la matriz de la transformación lineal, su matriz asociada es única para esa transformación lineal L_A donde A es la matriz asociada a esa transformación lineal.

Ejercicios

1. **Sea $S: V \rightarrow W$ el mapa lineal desde un espacio vectorial a otro. Muestre que $S(O) = O$**

Desarrollo. Asociado este problema a carácter matricial, se debe esclarecer que la matriz asociada a S debe estar especificada según la base del espacio vectorial V y W. Por ende, definiendo la base de $v_1, \ldots, v_n$ y la base de W como $w_1, \ldots, w_m$ la representación del elemento 0 se da bajo esta condición:

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

Bajo la definición de que la matriz de la transformación lineal o mapa lineal S denotada por $M(S)$ está considerada como la matriz de tamaño mxn tal que:

$$Sv_k = s_{1k}w_1 + s_{2k}w_2 + \cdots + s_{mk}w_m$$

Los elementos de $M(S)$ son los que están descritos como s_{ij} en el que i corresponde a la fila i columna j. Ahora bien, lo que solicita el problema, es demostrar que la transformación de S sobre el elemento neutro definido como O es O. No está solicitando información de la matriz de la transformación lineal, lo que también podría hacerse pensar, cual es la matriz O. En consecuencia,

$$S(O + 0) = S(O) + S(O) = S(O) + S(O) - S(O)$$
$$= S(O + O) - S(O) = S(O) - S(O) = O = S(O)$$

Kernel e Imagen de una transformación lineal. Análisis matricial.

Considerando que se define como kernel o espacio nulo, tal como esas palabras sin importar lo que conlleve a pensar después, al conjunto de elementos $v\ que\ pertenecen\ a\ V$ tal que $F(v) = O$. Ahora bien, en el contexto de matrices, consideremos el siguiente ejemplo. Si A es un vector arbitrario en R^n, se puede definir el mapa lineal $L_A: R^n \to R$, en el que su

kernel puede ser interpretado como el conjunto de todos los elementos X los cuales son perpendiculares a A, estableciendo que $L_A(X) = AX$. Por ejemplo, particularizando, sea $L: R^3 \to R$ el mapa tal que:

$$L(x_1, x_2, x_3) = 3x_1 - 2x_2 + x_3$$

Por ende, si $A = (3, -2, 1)$, podemos escribir $L(X) = XA = AX$. La multiplicación está siendo obvia con respecto al producto punto:

$$(x_1, x_2, x_3)(a, b, c) = x_1a + x_2b + x_3c = (a, b, c)(x_1, x_2, x_3)$$
$$= ax_1 + bx_2 + cx_3$$

La necesidad de analizar cual es el espacio nulo de una transformación lineal, en el que se asocia su matriz, podría interpretarse de la siguiente manera. Sea S matriz de la trasformación lineal descrita por $S: F^n \to F^m$ ya definida sus bases. Entonces esta transformación lineal:

$$L_S: F^n \to F^m$$

En el que $L_S: SE$ corresponde a multiplicar la matriz de L_S cuyo tamaño es de mxn.

Su kernel o espacio nulo se define como aquellos elementos E tales que $SE = O$. La intención de este capítulo de libros es analizar que características debe cumplir esa matriz S para ciertos casos particulares.

Teorema 3

Sea $L_S: F^n \rightarrow F^m$ un mapa lineal cuyo espacio nulo es 0. Si $H^1, \dots, H^k$ son linealmente independientes en F^n, entonces $L_S(H^1), \dots, L_S(H^k)$ son linealmente independientes en F^m.

Demostración. Sean $f_1, \dots, f_k$ números tales que:

$$f_1 L_S(H^1) + \cdots + f_k L_S(H^k) = 0$$

Es decir, existen escalares tales que al multiplicar por los elementos $L_s(H^1), \dots, L_S(H^k)$ se obtiene el elemento 0 de F^m. En primer lugar, como sería posible aceptar que esa suposición es factible. F^m es un espacio vectorial, por ende 0 está en F^m. De cuantas maneras sería posible escribir dicho elemento neutro en F^m. Puede ser escrito en una infinidad de formas, por tanto la generalización descrita no está mal propuesta. Ahora bien, por linealidad tenemos lo siguiente:

$$L_S(f_1 H^1 + \cdots + f_k H^k) = 0$$

Por ende, $f_1 H^1 + \cdots + f_k H^k = 0$. Debido a que $H^1, \dots, H^k$ son linealmente independientes, $f_1, \dots, f_k$ son iguales a 0. Por tanto, la expresión:

$$f_1 L_S(H^1) + \cdots + f_k L_S(H^k) = 0 L_S(H^1) + \cdots + 0 L_S(H^k) = 0$$

Manifiesta que $L_S(H^1), \dots, L_S(H^k)$ son linealmente independientes.

Cabe mencionar, que la representación para

$$f_i L_S(H^i) = f_i S H^i = f_i S_{mxn} H^i_{nx1}$$

En donde S es una matriz de tamaño mxn. Por tanto, los elementos obtenidos $SH^1, \dots, SH^k$ son linealmente independientes.

En general, sea $L_S: F^n \to F^m$ un mapa. Se dice que L_S es inyectiva si para cualquier $A, B\ en\ V$ y $A \neq B$, entonces $L_S(A) \neq L_S(B)$. En otras palabras, que L_S sea inyectiva, significa que L_S tome distintos valores a distintos elementos de F^n. Dicho de otra manera, también se lo puede parafrasear en el que L_S es inyectiva si y solo si dado $A, B\ en\ V$

$$L_S(A) = L_S(B), \qquad implica\ que\ A = B$$

Teorema 4

Sea $L_S: F^n \to F^m$ un mapa lineal, entonces $Kernel(L_S) = \{O\} \equiv$ $L_S\ es\ inyectiva.$

Demostración. Para probarlo en primer lugar partimos de $Kernel(L_S) = O$, y ahora supongamos que H, T son elementos de F^n tales que $L_S(H) = L_S(T)$. Analicemos por qué razón podría asumirse que dichos elementos pueden cumplir tal suposición. Se tiene conocimiento que F^m es un espacio vectorial, por tanto, el elemento O está en F^m. De igual forma O es único en F^m. Así mismo el inverso aditivo es único. Más aún, lo que se está tratando de escribir como suposición a partir de que F^m es un espacio vectorial, es de que $L_S(T)$ es el inverso aditivo de $L_S(H)$. En conclusión, no hay nada de malo en esta asunción. Por lo tanto,

$$L_S(H - T) = O$$

Por condición inicial $Ker(L_S) = Kernel(L_S) = O$, es decir que $H - T = O$, por tanto $H = T$ concluyendo que L_S es inyectiva.

Para probar la otra parte, ahora la condición inicial es que L_S es inyectiva. Esto quiere decir que si H es tal que $L_s(H) = L_S(O)$, hay tomar nuevamente en consideración si es factible asumir tal posibilidad. Se conoce que F^n es un espacio vectorial y en consecuencia contiene al elemento neutro O. Además, si L_S es una transformación lineal, siempre $L_S(O) = O$. Por ende, la consideración de asumirla factible no está demás al utilizar L_S. En consecuencia, por ser inyectiva, y parafraseando la definición de $Ker(L_S) = \{H \in F^n : L_S(H) = O\}$, la única posibilidad o resultado bajo esa definición es que entonces si $L_S(H) = L_S(O)$ implica que $H = O$ por ser inyectiva. Por otra parte,

$$L_S(H) = SH, S_{mxn}$$

Conlleva a responder que si L_S es inyectiva, el espacio nulo de L_S es la matriz nula. Es de interés analizar que relevancia está asociada con la matriz de la transformación lineal S considerando que es el vector nulo $O \in F^n$ si el mapa lineal es inyectivo.

Imagen de una transformación lineal. Análisis en matrices.

Sea una transformación lineal $L_S: F^n \rightarrow F^m$. Se define como la imagen de L_S al conjunto de elementos $R\ en\ F^m$ tal que existe un elemento $E\ de\ F^n$ en el que $L_S(E) = R$. Tener en consideración que resulta de particular interés que la imagen de L_S es un subespacio de F^m. Para probarlo o darnos cuenta de que tal aseveración es cierta, en primer lugar, se debe estar seguros de que el elemento neutro de F^m pertenece a la imagen de L_S. Ya se probó en líneas anteriores $L_S(O) = O$ y según la definición de imagen de L_S, es el conjunto de elementos en $K\ en\ F^m$ tal que existe un elemento $E\ en\ F^n$ en el que $L_S(E) = R$, por otro lado, dicho elemento en F^n al ser tomado en consideración al elemento neutro, al utilizar $L_S(O) = O$. Más aún, O está en F^m por ser un espacio vectorial, pero se consigue con mayor seguridad de que O está en la imagen de L_S por su definición, de que existe un elemento, tal que al utilizar L_S sobre O que está en F^m nos garantiza que dicho elemento neutro O está la imagen de L_S, en otras palabras $O \in Im(L_S)$. Ahora debemos comprobar que la adición sobre F^m y multiplicación por un escalar sobre F^m también son válidas en la imagen de L_S.

Sean R_1, R_2 elementos de la imagen de L_S. Consideremos por un instante, si se puede suponer tal cosa. Debido a que L_S es una transformación lineal desde un espacio vectorial a otro, tal consideración si es posible que existan elementos R_1 y R_2

que estén en la imagen de L_S. Por ende, existen elementos E, H tal que $L_S(E) = R_1$ y $L_S(H) = R_2$. En consecuencia:

$$L_S(E + H) = L_S(E) + L_S(H) = R_1 + R_2$$

Ahora bien, teniendo en consideración que L_S es un mapa lineal, tal condición es verdadera, y además $L_S(E) \in F^m$, $L_S(H) \in F^m$ entonces por la definición de imagen de L_S la suma de esos dos elementos también está en la imagen de L_S.Por otra parte $E \in F^n$ y también $H \in F^n$. Finalmente revisando la multiplicación por escalar, debemos verificar que la multiplicación por escalar este en la imagen de L_S. Sea c es un número, y teniendo en cuenta que R_1 está en la imagen de L_S, por tanto:

$$L_S(cE_1) = cL_S(E_1) = cR_1$$

Por tanto, la multiplicación por escalar está en la imagen de L_S o en otras palabras cR_1 está en la imagen. Sin dejar atrás el significado de L_S en el que, asociando su matriz de la transformación lineal, se tiene que para un elemento E de F^n, $L_S(E) = S_{mxn}E_{nx1} = K_{mx1}$. Por lo tanto, la imagen de L_S consta de todos los elementos $K \in F^m$ en el que existe un elemento $E_{nx1} \in F^n$ en el que $S_{mxn}E_{nx1} = K_{mx1}$.

Teorema 5

Sea F^n un espacio vectorial. Sea $L_S: F^n \to F^m$ un mapa lineal de F^n hacia otro espacio F^m. Sea n la dimensión de F^n, q la dimensión del espacio nulo o kernel de L_S, y p la dimensión

de la imagen de L_S. Entonces resulta que $n = q + p$. En otras palabras,

$$\dim F^n = \dim Ker L_S + \dim Im L_S$$

Demostración. Si la imagen de L_S consta de solamente O, entonces nuestra aserción es trivial. Ahora bien, cabe mencionar que el teorema entrega la dimensión de la imagen de L_S por ende significa que consta de una cantidad de elementos que constituyen una base de la imagen de L_S sin olvidar la definición de lo que es la imagen de L_S, recordándola por si acaso, como aquellos elementos $R \in F^m$ en el que existen un elemento $E \in F^n$ en el que $L_S(E) = R$. Ahora, asumamos para evitar el caso trivial, que $p > 0$.Estableciendo de forma general, sea una base de la imagen de L_S definida por $K_1, \dots, K_p$. Como pertenecen a la imagen de L_S, sean los elementos de V aquellos $E^1, \dots, E^p$ tales que $L_S(E^i) = K_i,\ \ i = 1, \dots, s$. Si el kernel de L_S no es O, sean $U^1, \dots, U^q$ una base del kernel de L_S. En el caso de que el kernel de L_S sea O, está entendido que la referencia en lo que respecta a $U^1, \dots, U^q$ será omitida para lo que se seguirá probando. De hecho, aseguramos bajo estas consideraciones que están generalizadas que $E^1, \dots, E^p, U^1, \dots, U^q$ es una base de F^n. Es más, esto será suficiente para probar el teorema. Sea E cualquier elemento de F^n. Entonces existen elementos $x_1, \dots, x_p$ tal que:

$$L_S(E) = x_1 K_1 + \cdots + x_p K_p$$

Debido que $K_1, \dots, K_p$ constituye una base de la imagen de L_S. Por linealidad, dicho resultado es obtenido por:

$$L_S(E) = L_S(x_1E^1 + \cdots + x_pE^p)$$

Ahora, restando a ambos lados $L_S(x_1E^1 + \cdots + x_pE^p)$:

$$\begin{aligned} L_S(E) - L_S\left(x_1E^1 + \cdots + x_pE^p\right) \\ = L_S\left(x_1E^1 + \cdots + x_pE^p\right) \\ - L_S\left(x_1E^1 + \cdots + x_pE^p\right) = O \end{aligned}$$

En consecuencia, nunca olvidando que L_S es una transformación lineal

$$L_S\left(E - x_1E^1 - \cdots - x_pE^p\right) = O$$

Por ende, $E - x_1E^1 - \cdots - x_pE^p$ está en el kernel de L_S. Lo que falta probar para que $E^1, \dots, E^p$ sea una base del kernel es que extienda al kernel y sean linealmente independientes. Entonces existen elemento $y_1, \dots, y_q$ tal que:

$$E - x_1E^1 - \cdots - x_pE^p = y_1U^1 + \cdots + y_qE^q$$

Por ende, reorganizando el elemento $E \in F^n$

$$E = x_1E^1 + \cdots + x_pE^p + y_1U^1 + \cdots + y_qE^q$$

Esto logra probar que los elementos $E^1, \dots, E^p, U^1, \dots., E^q$ extienden a F^n. Ahora toca demostrar que sean linealmente independientes. Debido a que F^n es un espacio vectorial, el elemento O está en F^n, por tanto, la consideración siguiente existe:

$$O = x_1E^1 + \cdots + x_pE^p + y_1U^1 + \cdots + y_qE^q$$

Aplicando L_S a ambos lados,

$$\begin{aligned} L_S(O) = L_S\left(x_1E^1 + \cdots + x_pE^p + y_1U^1 + \cdots + y_qU^q\right) \\ = x_1L_S(E^1) + \cdots + x_pL_S(E^p) + y_1L_S(U^1) + \cdots \\ + y_qL_S(U^q) = O \end{aligned}$$

Analicemos un poco esta expresión que es bastante interesante y un resultado maravilloso. Nuevamente recordando la definición de $Ker(L_S) = \{E \in F^n : L_S(E) = O\}$. Debido a que la base de $Ker(L_S)$ es $U^1, \dots, U^q$ estos elementos están en ese subespacio vectorial, por lo tanto

$$L_S\left(U^i\right) = O, \qquad i = 1, \dots, q$$

Y obtenemos ahora,

$$O = x_1L_S(E^1) + \cdots + x_pL_S(E^p) = x_1K_1 + \cdots + x_pK^p$$

Lo que significa que $x_i = 0,\ i = 1, \dots, p$ debido a que los elementos $L_S\left(E^i\right) = K_i,\ i = 1, \dots, p$ definidos anteriormente por ser una base de $Im(L_S)$ son linealmente independientes por definición. Incluso concuerda con la cantidad de elementos definidos como dicha base de la imagen de L_S. En consecuencia, se obtiene el siguiente resultado volviendo a la condición inicial de haber expresado al elemento O

$$\begin{aligned} y_1L_S(U^1) + \cdots + y_qL_S(U^q) = O = L_S(O) \\ = L_S(y_1U^1 + \cdots + y_qU^q) \end{aligned}$$

Es decir, lo único que se ha probado es que los escalares para la combinación lineal que involucra a los elementos de la base

de la imagen de L_S son 0. Por ende, dejando como estaba anteriormente

$$y_1 U^1 + \cdots + y_q U^q = O$$

Sin dejar a un lado que dichos elementos $U^1, \ldots, U^q$ son una base del kernel de L_S, son linealmente independientes, por ende, $y_i = 0,\ i = 1, \ldots q$. Esto concluye la prueba.

Como este capítulo de libro está enfocado a la matriz de la transformación lineal como única, hasta el momento, en líneas posteriores se vinculará el significado de rango columna y rango fila con la dimensión de la imagen de L_s, que en consecuencia significan lo mismo.

Productos Escalares y Ortogonalidad

En esta sección, se estudiará y analizará cada enunciado de los teoremas a continuación con el objetivo nuevamente de vincular al análisis matricial. Resulta que la parte o palabra ortogonalidad da un resultado bastante curioso cuando se trata de sistemas de ecuaciones lineales y esto es sumamente útil para problemas asociados a análisis numérico, optimización, sistemas de control, física, economía entre otros. Se definirá sobre un campo F. Para no perder la desilusión sobre todos los conceptos, se da un breve recuerdo del significado un campo F.

Definición 3

Sea F un subconjunto de los números complejos $\mathbb{C}$. Los elementos de F también se llamarán **números o escalares.** Diremos que F es un **campo** si satisface las siguientes condiciones:

C*F*1. Si x, y son elementos de F entonces $x + y$ y xy son también elementos de F.

C*F*2. Si x está en F, entonces $-x$ es también un elemento de F. Si más aún, x es diferente de 0, entonces x^{-1} es también un elemento de F.

C*F*3. Los elementos 0 y 1 son elementos de F.

Con esta definición, siempre se definirá un espacio vectorial V sobre un campo F como un conjunto de objetos los cuales pueden ser añadidos y multiplicados por elementos de F, en tal forma que la suma de dos objetos de V es nuevamente un elemento de V, el producto de un elemento de K por un elemento de V es un elemento de V y las siguientes propiedades son satisfechas:

EV1. Dado elementos u, v, w de V, tenemos

$$(u + v) + w = u + (v + w)$$

EV2. Hay un elemento de V denotado por O tal que:

$$0 + u = u + O = u$$

Para todos los elementos u que están en V.

E$V3$. Dado un elemento u de V, existe un elemento $-u$ en V tal que

$$u + (-u) = O$$

E$V4$. Para todos los elementos u, v de V tenemos

$$u + v = v + u$$

E$V5$. Si c es un número, entonces $c(u + v) = cu + cv$.

E$V6$. Si a, b son dos números, entonces $(a + b)v = av + bv$.

E$V7$. SI a, b son dos números, entonces $(ab) = a(bv)$

E$V8$. Para todos los elementos u que están en V, tenemos $1u = u$.

Estas propiedades como es el caso de estudio en este capítulo de libro para matrices también pueden ser verificadas en ese conjunto A de matrices de tamaño mxn como un espacio vectorial. Es decir:

Un espacio vectorial A_{mxn} sobre un campo F es un conjunto de objetos los cuales pueden ser añadidos y multiplicados por elementos de F, en tal forma que la suma de dos objetos de A_{mxn} es nuevamente un elemento de A_{mxn}, el producto de un elemento de K por un elemento de A_{mxn} es un elemento de A_{mxn} y las siguientes propiedades son satisfechas:

E$V1$. Dado elementos $V_{mxn}, U_{mxn}, W_{mxn}$ de A_{mxn}, tenemos

$$(V_{mxn} + U_{mxn}) + W_{mxn} = V_{mxn} + (U_{mxn} + W_{mxn})$$

EV2. Hay un elemento de A_{mxn} denotado por O tal que:

$$0 + U_{mxn} = U_{mxn} + O = U_{mxn}$$

Para todos los elementos U_{mxn} que están en A_{mxn}.

EV3. Dado un elemento U_{mxn} de A_{mxn}, existe un elemento $-U_{mxn}$ en A_{mxn} tal que

$$U_{mxn} + (-U_{mxn}) = O$$

EV4. Para todos los elementos U_{mxn}, V_{mxn} de A_{mxn} tenemos

$$U_{mxn} + V_{mxn} = V_{mxn} + U_{mxn}$$

EV5. Si c es un número, entonces $c(U_{mxn} + V_{mxn}) = cU_{mxn} + cV_{mxn}$.

EV6. Si a, b son dos números, entonces $(a + b)V_{mxn} = aV_{mxn} + bV_{mxn}$.

EV7. SI a, b son dos números, entonces $(ab) = a(bV_{mxn})$

EV8. Para todos los elementos u que están en A_{mxn} , tenemos $1V_{mxn} = V_{mxn}$.

Analicemos la siguiente transformación lineal: Sea V_{mxn} un espacio vectorial y sea P_{mxn}. La intención es verificar si $T: V_{nxn} \rightarrow P_{nxm}$ tal que $T(V_{mxn}) = V_{nxn}R_{nxm}$ es un espacio vectorial. En este caso si lo es. Ahora se analizará si el conjunto de elementos que pertenecen al conjunto de n tuplas de elementos en F denotado como $V = F^n$ es un espacio

vectorial. Entonces, sea $V = F^n$ el conjunto de $n - tuplas$ de elementos en F denotados por $X = (x_1, \dots, x_n)$ o llamados vectores. Existe un especial interés por un vector de esa estructura definida con un vector columna. Para ello, se define como vector columna a $X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$. Sin embargo, al utilizar los principios de espacio vectorial, la representación de X como vector columna no es congruente con la representación de X como vector fila, pero si a X lo representamos como este vector $X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}^T = (x_1, \dots, x_n)$ si se puede sumar y multiplicar por escalar por elementos de K de tal manera que la suma de dos elementos estén nuevamente de F^n de la misma clase y la multiplicación por un escalar a $\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}^T$ este nuevamente en F^n. Ahora bien, para probar que este conjunto es un espacio vectorial sobre F, es decir, al añadir y multiplicar por elementos de F, la suma de dos objetos de F^n es nuevamente un elemento que está en F^n, la multiplicación de un elemento de F por un elemento de F^n es nuevamente un elemento que está en F^n y en el que las siguientes propiedades son satisfechas:

EV1. Dado elementos $X = (x_1, \dots, x_n), R = (r_1, \dots, r_n), P = \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T$ de F^n, tenemos

$$(X + R) + P = X + (R + P)$$

$$\left((x_1, \dots, x_n) + (r_1, \dots, r_n)\right) + \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T$$

$$= (x_1, \dots, x_n) + \left((r_1, \dots, r_n) + \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T \right)$$

EV2. Hay un elemento de F^n denotado por 0 tal que:

$$X + 0 = X + 0 = X$$

Por ejemplo, sea $\begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T + 0 = \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T + 0 = \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T$

Para todos los elementos X que están en F^n.

EV3. Dado un elemento $\begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T$ de F^n, existe un elemento $-\begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T$ en F^n tal que

$$\begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T + \left(-\begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix}^T \right) = 0$$

EV4. Para todos los elementos X, W de F^n tenemos

$$\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}^T + (w_1, \dots, w_n) = (w_1, \dots, w_n) + \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}^T$$

E$V5$. Si c es un número, entonces $c\left(\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T + (w_1, \ldots, w_n)\right) = c\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T + c(w_1, \ldots, w_n)$.

E$V6$. Si a, b son dos números, entonces $(a+b)\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T = a\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T + b\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T$.

E$V7$. SI a, b son dos números, entonces $(ab) = a\left(b\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T\right)$

E$V8$. Para todos los elementos u que están en F^n , tenemos $1\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T = 1(x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}^T$.

Regresando al caso de productos escalares y ortogonalidad, establezca un espacio vectorial F^n sobre un campo F. Un **producto escalar** sobre F^n es una asociación en cual para cualquier par de elementos E, A que están en F^n, denotado por $< E, A >$ o también EA, que satisface las siguientes propiedades:

PE1. Tenemos $< E, A >=< A, E >$ para todo E, A que están en F^n

PE2. Si E, A, R son elementos de F^n entonces,

$$< E, A + R >=< E, A > +< E, R >$$

PE3. Si x está en F, entonces $< xE, A >= x < E, A >$ y $< E, xA >= x < E, A >$

El producto escalar es dicho como **no degradado** si en adición también la satisface la condición:

$$Si\ E\ es\ un\ elemento\ de\ F^n, y\ < E, A > = 0\ para\ todo\ A\ que\ está\ en\ F^n, entonces\ E = O$$

Ortogonal

Tómese en consideración el siguiente ejemplo. Sea F^n. Entonces el mapa

$$(X, Y) \longmapsto X \cdot Y^T$$

En la cual X, Y son elementos vectores filas de F^n. La intención es comprobar si esta asociación entre X, Y tal que el segundo elemento del producto sea la transpuesta del vector escogido y se consiga un producto escalar. Comprobando las propiedades que debe satisfacer:

PE1. Tenemos $< E, A >=< A, E >$ para todo E, A que están en F^n

$$< X, Y >= X \cdot Y^T = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} =< Y, X >= Y \cdot X^T$$
$$= (y_1, \dots, y_n) \cdot \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = y_1 x_1 + \cdots + y_n x_n$$

PE2. Si X, Y, W son elementos de F^n entonces,

$$\begin{aligned} < X, Y + W > &= (x_1, \dots, x_n) \cdot [(y_1, \dots, y_n) + (w_1, \dots, w_n)]^T \\ &= (x_1, \dots, x_n) \cdot [(y_1 + w_1, \dots, y_n + w_n)]^T \\ &= (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 + w_1 \\ \dots \\ y_n + w_n \end{pmatrix} \\ &= x_1 y_1 + x_1 w_1 + \cdots + x_n y_n + x_n w_n \\ &= x_1 y_1 + \cdots + x_n y_n + x_1 w_1 + \cdots + x_n w_n = \\ &< X, Y > + < X, W > \end{aligned}$$

PE3. Si r está en F, entonces $< rE, A >= r < E, A >$ y $< E, rA > = r < E, A >$

$$< rX, Y >= X \cdot Y^T = (rx_1, \dots, rx_n) \cdot \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} = r < X, Y > \; y$$

$$< X, rY > = X \cdot rY^T = (x_1, \dots, x_n) \cdot \begin{pmatrix} ry_1 \\ \dots \\ ry_n \end{pmatrix} = r < X, Y >$$

Analizar si el mapa descrito anteriormente es una transformación o mapa lineales, se debe probar los siguiente. Sea el mapa definido anteriormente como una función $P: F^n \times F^n \to F$ tal que $P(X, Y) = X \cdot Y^T$ para todos los elementos $X, Y \in F^n$. En ese caso,

$$P\big((X, Y) + (R, S)\big) = P(X + R, Y + S) = (X + R) \cdot (Y + S)^T$$

$$P(X, Y) = X \cdot Y^T, \qquad P(R, S) = R \cdot S^T$$

$$\begin{aligned} X \cdot Y^T + R \cdot S^T &= (x_1 y_1 + \cdots + x_n y_n) + (r_1 s_1 + \cdots + r_n s_n) \\ &\neq (X + R) \cdot (Y + S)^T \\ &= (x_1 + r_1)(y_1 + s_1) + \cdots + (x_n + r_n)(y_n + s_n) \end{aligned}$$

Por tanto, el mapa definido como la asociación de elementos $X, Y \in F^n$ hacia un escalar no es un mapa lineal. Incluso podría generarse la pregunta si $F^n \times F^n \to F$ en el que $Q(X,Y) = X \cdot Y^T$ es un espacio vectorial sobre F. Ahora surge la pregunta si Q forma algún conjunto de elementos. Se podría definir como el conjunto P tal que:

$$P = \{Q \in [F^n \times F^n \longmapsto F : Q(X,Y) = X \cdot Y^T]\}$$

Cuál sería la suma de elementos tal que al escoger dos elementos en P, nuevamente se tenga un elemento en P. Por ejemplo, si Q_1, Q_2 están en P,

$$Q_1(X,Y) + Q_2(X,Y)$$

De esta manera se define la suma $Q_1 + Q_2$ como el mapa en el que su valor en un elemento (X,Y) de $F^n \times F^n$ es $Q_1(X,Y) + Q_2(X,Y)$. De igual manera definimos al producto de Q_1 por un elemento de F como el mapa cuyo valor en un elemento (X,Y) de $F^n \times F^n$ es $cQ_1(X,Y)$ para c en F

Verificando las propiedades para espacio vectorial:

EV1. Dado elementos tenemos Q_1, Q_2, Q_3

$$(Q_1 + Q_2) + Q_3 = Q_1 + (Q_2 + Q_3)$$

EV2. Hay un elemento de P denotado por O tal que:

$$O + Q = Q + O = Q$$

Por ejemplo, sea $Q(X,Y) + O(X,Y) = X \cdot Y^T + O(X,Y) = O(X,Y) + X \cdot Y^T = X \cdot Y^T$

Para todos los elementos X, Y que están en $F^n \times F^n$. Por tanto, dicho elemento $O(X,Y)$ es 0.

EV3. Dado un elemento Q de P, existe un elemento $-Q$ en P tal que:

$$Q(X,Y) + \big(-Q(X,Y)\big) = O(X,Y)$$

$$X \cdot Y^T + \big(-Q(X,Y)\big) = O(X,Y) = 0$$
$$= (x_1y_1 + \cdots + x_ny_n) + \big(-O(X,Y)\big) = 0$$

$-O(X,Y) = -x_1y_1 - \cdots - x_ny_n = -1(x_1y_1 + \cdots + x_ny_n) = -1Q(X,Y)$

EV4. Para todos los elementos Q_1, Q_2 de P tenemos

$$Q_1(X,Y) + Q_2(X,Y) = Q_2(X,Y) + Q_1(X,Y) = 2(X \cdot Y^T)$$

EV5. Si c es un número, entonces $c\big(Q_1(X,Y) + Q_2(X,Y)\big) = cQ_1(X,Y) + cQ_2(X,Y)$

EV6. Si a, b son dos números, entonces $(a+b)Q(X,Y) = aQ(X,Y) + bQ(X,Y)$.

EV7. SI a, b son dos números, entonces $(ab) = a\big(bQ(X,Y)\big)$

EV8. Para todos los elementos u que están en F^n , tenemos $1Q(X,Y) = 1(x_1y_1 + \cdots + x_ny_n) = 1x_1y_1 + \cdots + 1x_ny_n = x_1y_1 + \cdots + x_ny_n = Q(X,Y)$.

Sea F^n un espacio vectorial con un producto escalar. Como siempre, definimos elementos E, A de F^n los cuales son **ortogonales** ($E \perp A$) si $< E, A >= 0$. Si S es un subconjunto de F^n, denotamos como $S^\perp$ al conjunto de todos los elementos E

que están en F^n los cuales son perpendiculares a todos los elementos de S, es decir:

$$S^{\perp} = \{E \in F^n : < E, W >= 0, \forall W \in S\}$$

De hecho, se puede utilizar las propiedades PE2 y PE3 para probar que $S^{\perp}$ es un subespacio de F^n, el cual se denomina **espacio ortogonal** de S. Sea U el subespacio de F^n generado por los elementos de S. Si W es perpendicular a S, y si V_1, V_2 están en S, entonces

$$< W, V_1 + V_2 > =< W, V_1 > +< W, V_2 >= 0$$

Si c es un número o escalar, entonces

$$< W, cV_1 > = c < W, V_1 >$$

Decir que U es el subespacio de F^n ya se está afirmado que contiene el elemento 0 de F^n. Ahora si también es generado por los elementos de S, definamos como $B_1, \dots, B_p$ elementos de S que extienden a U, en otras palabras:

$$U = span\{B_1, \dots, B_p\}$$
$$= \{x_1 B_1 + \cdots + x_p B_p : x_i \in F\,, \qquad i = 1, \dots, p\}$$

Sea F^n un espacio vectorial con un producto escalar. Entonces, en primer lugar el elemento 0 está en F^n. Ahora bien,

$$< V, V > =< V, V + 0 > =< V, V > +< V, 0 >$$

Debido a que el producto escalar sobre un campo F asocia un escalar, estos tienen su inverso aditivo, por tanto

$$< V, V > - < V, V > = < V, V > + < V, O > - < V, V > = < V, V > - < V, V > + < V, O > = 0$$

$$< V, O > = < O, V > = 0$$

Hay algo bastante interesante respecto al análisis matricial de este capítulo. Resulta que la generalización para todo $\forall$ conlleva a expresar algebraicamente un elemento vector columna de F^n lo más general posible, por ejemplo, $R = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$. Ahora bien, escogiendo un elemento Q de F^n y si en adición $< Q, R > = 0$ para todo R que está en F^n resulta ser que siempre el elemento Q es el elemento neutro, sin importar cual sea R. Por ejemplo, escogiendo $R = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ que está en R^2 y siendo $T = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$

$< T, R > = 0$, tanto como para $T = O$ en el que sea O de R^2 se cumple de que ese espacio vectorial con un producto escalar es 0 sin embargo si se escoge $T = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ también el producto escalar es 0. Lo que pasa es que la condición de no degradado es que el elemento R tiene que ser generalizado para todo, y en adición el producto escalar es 0 por tanto, R es expresado de esa manera para que se cumpla la condición de no degradado.

Más general para matrices, $[a_{ij}]$ representa una matriz de tamaño mxn en un campo F, y sean $A_1, \dots, A_m$ sus vectores filas en el que $X = [x_1, \dots, x_n]^t$ como lo usual. Resulta que el sistema de ecuaciones lineales

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0$$

$$\dots$$

$$a_{m1}x_1 + \cdots + a_{mn}x_n = 0$$

Puede ser abreviado de la siguiente manera: $A_1X = 0, \dots, A_mX = 0$. Si analizamos un momento, la abreviación realizada por el sistema de ecuaciones lineales es por medio del producto escalar definido anteriormente, es decir, la primera componte es un vector fila y su segunda la transpuesta.

El conjunto de soluciones X de este sistema homogéneo es un espacio vectorial sobre F, lo cual es cierto, basta con empezar que el elemento O que está en F^n es valido para ese conjunto denotado por X. Sea W el espacio generado por $A_1, \dots, A_m$. Es decir;

$$\begin{aligned} W = span(A_1, \dots, A_m) \\ &= \{c_1A_1 + \cdots + c_mA_m : c_1, \dots, c_m \ estan \ en \ F\} \\ &= \{c_1(a_{11}, \dots, a_{1n}) + \cdots + c_m(a_{m1}, \dots, a_{mn})\} \end{aligned}$$

Ahora sea U el espacio que consiste en todos los vectores en F^n perpendiculares a $A_1, \dots, A_m$. Es decir, $U = \{Y \in F^n : < Y, A_i > = 0, \ A_i \in F^n, \ i = 1, \dots, m\}$. Entonces U es precisamente el espacio de soluciones del sistema de

ecuaciones lineales mostrado anteriormente. Los vectores $A_1, \dots, A_m$ podrían no ser linealmente independientes. Notamos que $\dim W \leq m$, y llamamos

$$\dim U = \dim W^{\perp}$$

Como la dimensión del espacio de soluciones del sistema de ecuaciones lineales.

Sea F^n de nuevo un espacio vectorial sobre un campo F, con un producto escalar. **Se llama a F^n un espacio vectorial con un producto escalar fijo positivo**, si $< X, X > \geq 0$ para todo X que pertenece a F^n, y en el que $< X, X > > 0$ si $X \neq O$

Proyección de un elemento Y de F^n a lo largo de un elemento X de F^n (cX)

Sea X un elemento de F^n tal que $\|X\| \neq 0$, definiéndose la norma como $\|X\| = \sqrt{< X, X >}$. Para cualquier Y existe un único número c tal que $Y - cX$ es perpendicular a X. Se genera la pregunta si cualquier espacio vectorial con un producto escalar fijo positivo debe tener elementos perpendiculares entre sí. La respuesta es correcta debido que la generación del producto escalar entre esos vectores de F^n contienen al elemento O y es posible dicha situación. Ahora debido a que el producto escalar asocia en efecto un número, teniendo a c

como el siguiente valor lo cual también puede suceder con normalidad:

$$c = \frac{< Y, X >}{< X, X >}$$

Muestra que $Y - cX$ es perpendicular a X. Dicho de otra manera:

$$c < X, X > = < Y, X >$$

$$c < X, X > - < Y, X > = c < X, X > - < X, Y > = c < X, X > + < X, -1 \cdot Y > = < X, cX > + < X, -1 \cdot Y > = 0$$

$$< X, cX > + < X, -1 \cdot Y > = 0$$

$$- < X, cX > - < X, -1 \cdot Y > = 0$$

$$< X, -cX > + < X, Y > = 0$$

$$< X, Y - cX > = 0$$

Lo que prueba que existe un único c en el que $X, Y - cX$ son perpendiculares. Se llama a c como la componente de Y a lo largo de X. Recordar que X, Y están en F^n. Ahora bien, supongamos que existe otro escalar d tal que,

$$d = \frac{< X, Y >}{< X, X >}$$

Entonces por las propiedades de producto escalar se tiene

$$d < X, X > = < X, Y > = < Y, X >$$

Por ende, se obtiene lo siguiente,

$$d = \frac{< Y, X >}{< X, X >} = c$$

En donde, la expresión $< X, Y > = < Y, X >$ se cumple debido a las propiedades de producto escalar. En efecto, la unicidad del número c es factible para la aserción sobre la proyección de Y de F^n.

Sean $X_1, \dots, X_p$ elementos diferentes de cero que pertenecen a F^n los cuales son mutuamente perpendiculares, es decir

$$< X_i, X_j > = 0, \qquad si\ i \neq j$$

Tómese los números c_i como la componente de Y a lo largo de X_i. Entonces,

$$Y - c_1 X_1 - \cdots - c_n X_n$$

Es perpendicular a $Y_1, \dots, Y_n$. Para observar esto, todo lo que tenemos que hacer es tomar el producto con X_i para cualquier j. Todos los términos que involucran $< X_i, X_j >$ darán 0 si $i \neq j$, y tendremos dos términos restantes

$$< Y, X_j > \ -c_j < X_j, X_j >$$

Por ejemplo, escójase X_j y realice el producto escalar entre $Y - c_1 X_1 - \cdots - c_n Y_n$ con X_j. Entonces

$$< Y - c_1 X_1 - \cdots - c_n X_n, X_j > = < Y, X_j > - < c_1 X_1, X_j > - \cdots - < c_n X_n, X_j >$$

$$< Y - c_1X_1 - \cdots - c_nX_n, X_j > = < Y, X_j > - c_1 < X_1, X_j > - \cdots - c_n < X_n, X_j > = < Y, X_j > - c_j < X_j, X_j >$$

Definición 4

Sea F^n un espacio vectorial con un producto escalar fijo positivo. Una base $\{X_1, \ldots, X_n\}$ de F^n se dice ser ortogonal si sus elementos son mutuamente perpendiculares, es decir, $< X_i, X_j > = 0$ siempre que $i \neq j$. Si en adición cada elemento de la base tiene norma 1, entonces la base es llamada **ortonormal**.

Teorema 6

Sea F^n un espacio vectorial de dimensión finita, con un producto escalar fijo positivo. Sea W un subespacio de F^n, y sea $\{Y_1, \ldots, Y_m\}$ una base ortogonal de W. Si $W \neq V$, entonces existen elementos $Y_{m+1}, \ldots, Y_n$ de F^n tal que $\{Y_1, \ldots, Y_n\}$ es una base ortogonal de F^n.

Demostración. El método de la prueba es importante como el teorema y es llamado proceso de Ortogonalización de Gram – Schmidt. Sabemos que se pueden encontrar elementos $Y_{m+1}, \ldots, Y_n$ tales que

$$\{Y_1, \ldots, Y_m, X_{m+1}, \ldots, X_n\}$$

Es una base de F^n. Si por un momento se analiza esta aserción, en efecto es posible debido a que en primera instancia se denota como $\{Y_1, \ldots, Y_m\}$ una base de W y siendo W un

subespacio de F^n en el que por definición su dimensión, es decir, $m \leq \dim F^n \leq n$. En este caso ya se conoce la dimensión de F^n y W. En consecuencia, m al ser distinto de la dimensión de F^n, $m < n$ ($W \neq W$), para definir la base de F^n por ende se debe encontrar más elementos $X_{m+1}, \dots, X_n$.

Por supuesto que $\{Y_1, \dots, Y_m, X_{m+1}, \dots, X_n\}$ no es una base ortogonal. Sea W_{m+1} el espacio generado por $Y_1, \dots, Y_m, X_{m+1}$. Primero obtendremos un base ortogonal de W_{m+1}, es decir, elementos que están en W_{m+1} tales que su producto escalar es igual a 0.

$$\begin{aligned} W_{m+1} = span\{Y_1, \dots, Y_m, X_{m+1}\} \\ = \{c_1 Y_1 + \cdots + c_m Y_m \\ + c_{m+1} X_{m+1} : c_1, \dots, c_m, c_{m+1} \in F\} \end{aligned}$$

La idea es tomar X_{m+1} y extraer su proyección a lo largo de $Y_1, \dots, Y_m$. Por ende, dejamos

$$c_1 = \frac{< X_{m+1}, Y_1 >}{< Y_1, Y_1 >}, \dots, c_m = \frac{< X_{m+1}, Y_m >}{< Y_m, Y_m >}$$

Sea

$$Y_{m+1} = X_{m+1} - c_1 Y_1 - \cdots - c_m Y_m$$

Entonces Y_{m+1} es perpendicular a $\{Y_1, \dots, Y_m\}$. Más aún, $Y_{m+1} \neq O$ ya que, si fuese O, entonces X_{m+1} se podría escribir como combinación lineal de $Y_1, \dots, Y_m$ lo cual estaría contradiciéndose a la definición previamente establecida de que $\{Y_1, \dots, Y_m, X_{m+1}, \dots, X_n\}$ es una base de F^n, es decir, $\{Y_1, \dots, Y_m, X_{m+1}, \dots, X_n\}$ son linealmente independientes por

teorema ya demostrado. En consecuencia, X_{m+1} está en el espacio generado por $\{Y_1, \dots, Y_m, Y_{m+1}\}$ debido a que

$$X_{m+1} = Y_{m+1} + c_1 Y_1 + \cdots + c_m Y_m$$

Por ende $\{Y_1, \dots, Y_m, Y_{m+1}\}$ es una base de W_{m+1}.

Podemos ahora proceder por inducción, mostrando que el espacio W_{m+s} generado por

$$\{Y_1, \dots, Y_m, X_{m+1}, \dots, X_{m+s}\}$$

Tiene una base ortogonal

$$\{Y_1, \dots, Y_{m+1}, \dots, Y_{m+s}\}$$

Con $s = 1, \dots, n - m$. Esto concluye la prueba.

Corolario 1

Sea F^n un espacio vectorial de dimensión finita con un producto escalar fijo positivo. Asuma que $F^n \neq \{O\}$. Entonces F^n tiene una base ortogonal.

Demostración. Por hipótesis, existe un elemento Y^1 de F^n tal que $Y^1 \neq O$. Sea $Y^1, Y^2, \dots, Y^n$ una base de F^n. Dejamos como W al subespacio generado por Y^1, es decir:

$$W = span(Y^1) = \{cY^1 : c \in F\}$$

Podemos encontrar elementos $X^1, X^2, \dots X^n$ de F^n en el que sean una base de ortogonal de F^n. Recordando que una base ortogonal, son aquellos elementos de F^n que son una base de

F^n y que en adición son mutuamente perpendiculares, es decir, $< X^i, X^j > = 0$, siempre que $i \neq j$. Un conjunto de un solo vector diferente del nulo es linealmente independiente, por ende, si se define a $X^1 = Y^1$ no habría ningún problema en el que para

$$X^1 = cY^1, \ \ con\, c = 1$$

En consecuencia, X^1 está en la extensión de Y^1. Ahora bien, sea el espacio generado W_2 por

$$W_2 = span(Y^1, Y^2) = \{a_1 Y^1 + a_2 Y^2 : a_1, a_2 \in F\}$$

Estableciendo,

$$c_2 = \frac{< Y^2, Y^1 >}{< Y^1, Y^1 >}$$

Sea el elemento

$$X^2 = Y^2 - c_2 Y^1$$

Entonces X^2 es perpendicular a Y^1 y reescribiendo:

$$X^2 = Y^2 - c_2 X^1$$

se consigue que X^2 es perpendicular a X^1. Por otra parte, X^2 no puede ser el elemento neutro caso contrario Y^2 se escribiría como combinación lineal de $X^1 = Y^1$, por ello,

$$Y^2 = X^2 + c_2 X^1$$

lo que significa que Y^2 cae en el espacio generado por X^1, X^2, y en consecuencia, $\{X^1, X^2\}$ es una base ortogonal para W_2.

Nuevamente repitiendo el proceso, sea W_3 el espacio generado por

$$W_3 = span(X^1, X^2, Y^3) = \{a_1X^1 + a_2X^2 + a_3Y^3 : a_1, a_2, a_3 \in F\}$$

Estableciendo,

$$d_3 = \frac{< Y^3, X^2 >}{< X^2, X^2 >}, d_2 = \frac{< Y^3, X^1 >}{< X^1, X^1 >}$$

Sea el elemento

$$X^3 = Y^3 - d_3X^2 - d_2X^1$$

se consigue que X^3 es perpendicular a X^1, X^2. Por otra parte X^3 no puede ser el elemento neutro caso contrario Y^3 se escribiría como combinación lineal de X^1, X^2 lo que contradice ya que son linealmente independientes por ser base ortogonal de W_2. Por ello

$$Y^3 = X^3 + d_3X^2 + d_2X^1$$

lo que significa que Y^3 cae en el espacio generado por X^1, X^2, X^3, en consecuencia, $\{X^1, X^2, X^3\}$ es una base ortogonal para W_3.

Se puede realizar el mismo procedimiento hasta obtener como base ortogonal $\{X^1, X^2, X^3, \dots, X^n\}$. La estructura de pasos realizada fue por inducción.

Teorema 7

Sea R^n un espacio vectorial sobre R con un producto escalar positivo fijo de dimensión n. Sea W un subespacio de R^n de dimensión r. Sea $W^\perp$ es subespacio de R^n que consiste en todos los elementos los cuales son perpendiculares a W. Entonces R^n es la suma directa de W y $W^\perp$, y $W^\perp$ tiene dimensión $n - r$. En otras palabras,

$$\dim W + \dim W^\perp = \dim R^n = n$$

Demostración. Si W consiste sólo del elemento neutro O, o si $W = R^n$, entonces nuestra aserción es obvia, es decir, para el primer caso,

Si $W = O$, entonces una base de W es el vacío, por ende, su dimensión es cero. En consecuencia $r = 0$ y la dimensión de $W^\perp$ es n.

Si $W = R^n$, entonces su dimensión es $r = n$ y por tanto la dimensión de $W^\perp$ es 0, lo que significa que $W^\perp = O$.

Se asumirá por ende que $W \neq R^n$ y que $W \neq O$. Sea $\{W^1, \dots, W^r\}$ una base ortonormal de W; por el teorema anterior existen elementos $U^{r+1}, \dots, U^n$ de R^n tal que

$$\{W^1, \dots, W^r, U^{r+1}, \dots, U^n\}$$

Probaremos que $\{U^{r+1}, \dots, U^n\}$ es una base ortonormal de $W^\perp$.

Sea U un elemento de $W^\perp$. Entonces existen números $x_1, \dots, x_r, x_{r+1}, \dots, x_n$ tal que

$$U = x_1W^1 + \cdots + x_rW^r + x_{r+1}U^{r+1} + \cdots + x_nU^n$$

Debido a que U es perpendicular a W, tomando el producto con cualquier W^i para $i = 1, \ldots, r$ se encuentra que

$$0 = < x_1W^1 + \cdots + x_rW^r + x_{r+1}U^{r+1} + \cdots + x_nU^n, W^i >$$

$$< x_1W^1, W^i > + \cdots + < x_rW^r, W^i > + < x_{r+1}U^{r+1}, W^i > + \cdots + < x_iW^i, W^i > + \cdots + < x_nU^n, W^i > = 0$$

En donde todos los términos se cancelan debido a que U es perpendicular a W a excepción de:

$$< x_iW^i, W^i > = x_i < W^i, W^i > = x_i = 0$$

Cabe recalcar que debido a que cada W^i conforman una base ortonormal, $< W^i, W^i > = 1$, por ende, $x_i = 0$ para $i = 1, \ldots, r$. En consecuencia U es una combinación lineal de $U^{r+1}, \ldots, U^n$.

Contrariamente, sea $U = x_{r+1}U^{r+1} + \cdots + x_nU^n$ una combinación lineal de $U^{r+1}, \ldots, U^n$. Tomando el producto con cualquier W^i se obtiene 0. Por ende, U es perpendicular a todo W^i para $i = 1, \ldots, r$, y por tanto es perpendicular a W. Esto prueba que $U^{r+1}, \ldots, U^n$ extiende a $W^{\perp}$. Debido a que ellos son mutuamente perpendiculares, y de norma 1, ellos forman una base ortonormal de $W^{\perp}$, cuya dimensión es en consecuencia $n - r$.

Más aún, un elemento de R^n tiene una única expresión como combinación lineal

$$x_1W^1 + \cdots + x_rW^r + x_{r+1}U^{r+1} + \cdots + x_nU^n$$

y por ende una única expresión como una suma $W + U$ con $W \in W$ y $U \in W^{\perp}$. Por tanto, R^n es la suma directa de $W\ y\ W^{\perp}$.

El espacio $W^{\perp}$ es llamado **el complemento ortogonal** de W.

Productos Hermitianos

Ahora se describirá la modificación necesaria para adaptar los resultados precedentes a espacios vectoriales de R^n sobre C^n, es decir, números complejos. Se desea preservar la notación de producto escalar fijo positivo cuanto sea posible. Debido a que el producto punto de vectores con coordenadas complejas puede ser cero sin que ellos sean igual al vector nulo O, se debe cambiar algo en su definición. Se vuelve a que su cambio es muy ligero y se presenta a continuación. En este capítulo de libro cabe mencionar que la intención sobre aplicaciones en ingeniería se basa exclusivamente sobre matrices, su comportamiento ante la representación de fenómenos físicos, poder interpretar de forma matricial sus aplicaciones y en consecuencia utilizando lógica algebraica rigurosa como lo es algebra lineal, se puede cuantificar toda su estructura.

Sea C^n un espacio vectorial sobre los números complejos. Un **producto hermitiano** sobre C^n es una regla la cual para cualquier para de elementos X, Y de C^n asocia un número complejo, denotado de nuevo por $< X, Y >$, que satisface las siguientes propiedades:

HP1. Tenemos $< X, Y > = \overline{< Y, X >}$, para todo X, Y que está en C^n. (Aquí la barra denota la conjugada compleja).

HP2. Si X, Y, S son elementos de C^n, entonces

$$< X, Y + S > = < X, Y > \ + < X, S >$$

HP3. Si $\alpha \in C$, entonces

$< \alpha X, Y > = \alpha < X, Y >$ y además $< X, \alpha Y > = \bar{\alpha} < X, Y >$

El producto hermitiano se dice ser **fijo positivo** si $< X, X > \geq 0$ para todo $X \in C^n$ y $< X, X > > 0$ si $X \neq O$.

Se define de igual manera como antes las palabras **ortogonal, perpendicular, base y complemento ortogonal.** No hay nada que cambiar en nuestra definición de **componente** y **proyección de** X a **lo largo de** Y.

Ejemplo. Sea $V = C^n$. Si $X = (x_1, \dots, x_n)$ y $Y = (y_1, \dots, y_n)$ son vectores en C^n, definimos su **producto hermitiano** como

$$< X, Y > = x_1\overline{y_1} + \cdots + x_n\overline{y_n}.$$

Se desea verificar si las condiciones **HP1, HP2 y HP3** se cumplen.

Solución.

Se procederá a validar la propiedad **HP1** por medio de inducción matemática. Uno de los primeros matemáticos quien verificó este método y confirmó su poder fue Peano el cual consiste en que si es válido para $i = 1$ e $i = n$ entonces es

válido para $i = n + 1$ con n que pertenece a los números enteros positivos.

HP1. Tenemos $< X, Y > = \overline{< X, Y >}$, para todo X, Y que está en C^n. (Aquí la barra denota la conjugada compleja).

$$< X, Y > = x_1\overline{y_1} + \cdots + x_n\overline{y_n}.$$

$$\overline{< Y, X >} = \overline{y_1\overline{x_1} + \cdots + y_n\overline{x_n}}.$$

Se probará que ambas expresiones son iguales. Para $i = 1$ se tiene lo siguiente:

$$< X, Y > = x_1\overline{y_1} = \overline{< Y, X >} = \overline{y_1\overline{x_1}}$$

Sea $x_1 = a_1 + b_1 i$ en donde $i = \sqrt{-1}$ y de igual manera sea $y_1 = c_1 + d_1 i$. Por conjugada se entiende que $\overline{x_1} = a_1 - b_1 i$ y $\overline{y_1} = c_1 - d_1 i$.

$$\begin{aligned} x_1\overline{y_1} = (a_1 + b_1 i)(c_1 - d_1 i) &= a_1 c_1 - a_1 d_1 i + b_1 c_1 i + b_1 d_1 \\ &= (a_1 c_1 + b_1 d_1) - (a_1 d_1 - b_1 c_1)i \end{aligned}$$

$$\begin{aligned} y_1\overline{x_1} = (c_1 + d_1 i)(a_1 - b_1 i) &= a_1 c_1 - c_1 b_1 i + a_1 d_1 i + d_1 b_1 \\ &= (a_1 c_1 + b_1 d_1) + (a_1 d_1 - c_1 b_1)i \end{aligned}$$

$$\overline{y_1\overline{x_1}} = (a_1 c_1 + b_1 d_1) - (a_1 d_1 - c_1 b_1)i$$

Es válido para $i = 1$. Si es válido para $i = n$ entonces es válido para $i = n + 1$.

$$< Y_{n+1}, X_{n+1} > = y_1\overline{x_1} + \cdots + y_n\overline{x_n} + y_{n+1}\overline{x_{n+1}}$$

$$\begin{aligned} \overline{< Y_{n+1}, X_{n+1} >} &= \overline{y_1\overline{x_1} + \cdots + y_n\overline{x_n} + y_{n+1}\overline{x_{n+1}}} \\ &= \overline{y_1\overline{x_1}} + \cdots + \overline{y_n\overline{x_n}} + \overline{y_{n+1}\overline{x_{n+1}}} \end{aligned}$$

$$\overline{< Y_{n+1}, X_{n+1} >} = \overline{y_1}\overline{x_1} + \cdots + \overline{y_n}\overline{x_n} + \overline{y_{n+1}}\overline{x_{n+1}}$$
$$= x_1\overline{y_1} + \cdots + x_n\overline{y_n} + x_{n+1}\overline{y_{n+1}}$$

En consecuencia, $< X, Y > = \overline{< Y_{n+1}, X_{n+1} >}$.

HP2. Si X, Y, S son elementos de C^n, entonces

$$< X, Y + S > = < X, Y > \ + < X, S >$$

$$< X, Y + S > = x_1\overline{(y_1 + s_1)} + \cdots + x_n\overline{(y_n + s_n)}$$
$$= (x_1\overline{y_1} + x_1\overline{s_1}) + \cdots + (x_n\overline{y_n} + x_n\overline{s_n})$$

$$< X, Y + S > = (x_1\overline{y_1}) + (x_1\overline{s_1}) + \cdots + (x_n\overline{y_n}) + (x_n\overline{s_n}) =$$
$$< X, Y > \ + < X, S >$$

HP3. Si $\alpha \in C$, entonces

$< \alpha X, Y > = \alpha < X, Y >$ y además $< X, \alpha Y > = \bar{\alpha} < X, Y >$

$< \alpha X, Y > = \alpha x_1\overline{y_1} + \cdots + \alpha x_n\overline{y_n} = \alpha(x_1\overline{y_1} + \cdots + x_n\overline{y_n}) = \alpha < X, Y >$

$$< X, \alpha Y > = x_1\overline{\alpha y_1} + \cdots + x_n\overline{\alpha y_n} = \bar{\alpha}(x_1\overline{y_1} + \cdots + x_n\overline{y_n}) = \bar{\alpha} < X, Y >$$

Aplicaciones a ecuaciones lineales; El Rango

Considérese el siguiente sistema:

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0$$
$$\cdots$$
$$a_{m1}x_1 + \cdots + a_{mn}x_n = 0$$

se puede interpretar el espacio de soluciones, de tres maneras:

a. Consiste de aquellos vectores X que dan relaciones lineales

$$x_1A^1 + \cdots + x_nA^n = O$$

entre las columnas de A.

b. Las soluciones forman el espacio ortogonal a los vectores fila de la matriz A.

c. Las soluciones forman el kernel del mapa lineal representado por A, es decir, las soluciones de la ecuación $AX = O$.

Ahora analicemos un teorema aplicado al conjunto de vectores en R^n en el que interviene su dimensión.

Teorema 8

Sea F^m un espacio vectorial sobre un campo K. Sea $\{X^1, \ldots, X^m\}$ una base de F^m sobre K. Sean $W^1, \ldots, W^n$

elementos de F^m, y asuma que $n > m$. Entonces $W^1, \dots, W^n$ son linealmente dependientes.

Demostración. Asuma que $W^1, \dots, W^n$ son linealmente independientes, es decir, al tomar en consideración esta proposición significa que se quiere llegar a una contradicción, por tanto, se demostrará que en efecto $W^1, \dots, W^n$ son linealmente dependientes.

Debido a que $\{X^1, \dots, X^m\}$ es una base, existen elementos $a_1, \dots, a_m \in F$ tal que

$$W^1 = a_1 X^1 + \cdots + a_m X^m$$

Por asunción, sabemos que $W^1 \neq O$ y por ende algún $a_i \neq 0$. Después de enumerar $X^1, \dots, X^m$ si es necesario, sin perdida de generalidad podemos asumir que $a_1 \neq 0$. Podemos resolver para X^1 y obtener

$$a_1 X^1 = W^1 - a_2 X^2 - \cdots - a_m X^m,$$

$$X^1 = \frac{1}{a_1} W^1 - \frac{a_2}{a_1} X^2 - \cdots - \frac{a_m}{a_1} X^m$$

El subespacio de F^m generado por $W^1, X^2, X^3, \dots, X^m$ contiene X^1 y por ende todo F^m debido a que $X^1, X^2, \dots, X^m$ genera F^m. La idea es continuar nuestro procedimiento paso a paso y reemplazar sucesivamente $X^2, X^3, \dots$ por $W^2, W^3, \dots$ hasta que todos los elementos $X^1, \dots, X^m$ sean agotados y $W^1, \dots, W^m$ generen a F^m. Ahora asumamos por inducción que existe un entero r con $1 \leq r < m$ tal que, después enumerar correctamente $X^1, X^2, \dots, X^m$, los elementos

$W^1, \dots, W^r, X^{r+1}, \dots, X^m$ generan F^m. Entonces existen elementos

$$b_1, \dots, b_r, c_{r+1}, \dots, c_m$$

en F tal que

$$W^{r+1} = b_1 W^1 + \cdots + b_r W^r + c_{r+1} X^{r+1} + \cdots + c_m X^m$$

Por ejemplo, para $r = 1$

$$W^2 = b_1 W^1 + c_2 X^2 + \cdots + c_m X^m$$

De igual manera para $r = 2$

$$W^3 = b_1 W^1 + b_2 W^2 + c_3 X^3 + \cdots + c_m X^m$$

No podemos tener $c_j = 0$ para $j = r + 1, \dots, m$ porque de otra manera obtendríamos una relación de dependencia lineal entre $W^1, \dots, W^{r+1}$, contradiciendo nuestra asunción. Después de renumerar $X^{r+1}, \dots, X^m$ de ser necesario, podríamos asumir sin pérdida de generalidad que $c_{r+1} \neq 0$. Obtenemos entonces,

$$c_{r+1} X^{r+1} = W^{r+1} - b_1 W^1 - \cdots - b_r W^r - c_{r+2} X^{r+2} - \cdots - c_m X^m$$

Dividiendo entre c_{r+1} se obtiene

$$X^{r+1} = \frac{1}{c_{r+1}} W^{r+1} - \frac{b_1}{c_{r+1}} W^1 - \cdots - \frac{b_r}{c_{r+1}} W^r - \frac{c_{r+2}}{c_{r+1}} X^{r+2} - \cdots - \frac{c_m}{c_{r+1}} X^m$$

lo que concluye que X^{r+1} está en el subespacio generado por $W^1, \dots, W^{r+1}, X^{r+2}, \dots, X^m$. Debido a que $1 \leq r < m$ que está en nuestra asunción, se sigue que $W^1, \dots, W^{r+1}, X^{r+2}, \dots, X^m$ genera a F^m. Por ejemplo, si $r = m - 1$, entonces $W^1, \dots, W^m$ genera de igual manera a F^m. Por tanto, por inducción hemos probado que $W^1, \dots, W^m$ generan a F^m. Si $n > m$. Si $n > m$, entonces existen elementos

$$d_1, \dots, d_m \in F$$

tal que

$$W^n = d_1 W^1 + \cdots + d_m W^m$$

por consiguiente, probando que $W^1, \dots, W^n$ son linealmente dependientes. Esto prueba el teorema.

Retomando la interpretación a ecuaciones lineales, se asume que estas expresiones tienen coeficientes a_{ij} en un campo K. Por ejemplo, sea G el subespacio de F^n y sea $G^\perp$ el subconjunto de todos los elementos $X \in F^n$ tal que

$$X \cdot Y = 0 \; para \; todo \; Y \in G$$

Entonces $G^\perp$ es un subespacio de F^n.

Teorema. Sea G el subespacio de F^n. Entonces

$$\dim G + \dim G^\perp = n$$

El Rango

Si $A = (a_{ij})$ es una matriz de tamaño $m\ x\ n$, entonces las columnas $A^1, \dots, A^n$ generan un subespacio, cuya dimensión es llamado el **rango columna** de A. Las filas $A_1, \dots, A_m$ de A generan un subespacio cuya dimensión es llamada el **rango fila** de A. Podríamos también decir que el rango columna de A es el máximo número de columnas linealmente independientes, y el que el rango fila de A es el máximo número de filas linealmente independientes.

Teorema 7

Sea $A = (a_{ij})$ una matriz de tamaño $m\ x\ n$. Entonces el rango fila y rango columna de A son iguales al mismo número r. Más aún, $n - r$ es la dimensión del espacio soluciones del sistema de ecuaciones lineales que se muestra a continuación

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0$$
$$\dots$$
$$a_{m1}x_1 + \cdots + a_{mn}x_n = 0$$

Demostración. Probaremos todos nuestros estamentos simultáneamente. Consideremos el mapa

$$L: F^n \to F^m$$

dado por

$$L(X) = x_1A^1 + \cdots + x_nA^n$$

Este mapa obviamente es lineal. Su imagen consiste en el espacio generado por los vectores columnas de A. Su kernel es por definición el espacio de soluciones del sistema de ecuaciones lineales. Por el teorema de mapas lineales obtenemos que

$$rango\ columna + \dim espacio\ soluciones = n$$

Por otra parte, interpretando el espacio soluciones como el espacio ortogonal de vectores fila, y usando el teorema de la dimensión de un subespacio ortogonal, obtenemos

$$rango\ fila + \dim espacio\ soluciones = n$$

De esto, todas nuestras aserciones se siguen de una sola manera y el teorema es probado.

Sean $b_1, \dots, b_m$ números, y considere el sistema no homogéneo de ecuaciones

$$A_1 \cdot X = b_1$$

...

$$A_m \cdot X = b_m$$

Podría pasar que este sistema no tenga una solución del todo, es decir, que sus ecuaciones sean inconsistentes. Por ejemplo, el sistema

$$2x + 3y - z = 1$$

$$2x + 3y - z = 2$$

no tiene solución. Sin embargo, si hay al menos una solución, entonces todas las soluciones son obtenidas de la que al menos hay mediante la adición de una solución arbitraria del sistema homogéneo asociado. En consecuencia, en este caso, podemos hablar de la dimensión del conjunto de solución, y esta es la dimensión del sistema homogéneo asociado.

Ejemplo 1. Encuentre el rango de la matriz

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

Solución.

Hay solamente dos filas, así que su rango es a lo mucho 2, debido a que tanto el rango fila como rango columna están asociados al mismo número, dígase r, que denota el máximo número de filas y columnas linealmente independientes y que por defecto constituye una base.

Por otra parte, las dos columnas

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

son linealmente independientes, para a y b que son números tal que

$$a\begin{pmatrix} 2 \\ 0 \end{pmatrix} + b\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

entonces

$2a + b = 0;$

$b = 0;$

Así que $a = 0$. Por ende, las dos columnas son linealmente independientes.

Ejemplo 2. Encuentre la dimensión del conjunto soluciones del siguiente sistema de ecuaciones, y determine este conjunto en R^3

$$2x + y + z = 1$$

$$y - z = 0$$

Solución.

Vemos por inspección que al menos existe una solución, dígase $x = \frac{1}{2}, y = z = 0$.

El rango de la matriz

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

es igual a 2. Por ende, la dimensión del conjunto de soluciones es 1. El espacio vectorial de soluciones del sistema homogéneo tiene dimensión 1, y una solución es fácilmente encontrada como

$$y = z = 1, \quad x = -1$$

Por ende, el conjunto de soluciones del sistema no homogéneo es el conjunto de todos los vectores

$$\left(\frac{1}{2}, 0,0\right) + t(-1,1,1)$$

donde t varía sobre todos los números reales. Observamos que nuestro conjunto soluciones es una línea recta.

Por ejemplo, para $t = 0$:

$$(x_1, y_1, z_1) = \left(\frac{1}{2}, 0,0\right) + 0(-1,1,1) = \left(\frac{1}{2}, 0,0\right)$$

$$2\left(\frac{1}{2}\right) + 0 + 0 = 1$$

$$0 - 0 = 0$$

Por ejemplo, para $t = 1$:

$$(x_2, y_2, z_2) = \left(\frac{1}{2}, 0,0\right) + 1(-1,1,1) = \left(-\frac{1}{2}, 1,1\right)$$

$$2\left(-\frac{1}{2}\right) + 1 + 1 = 1$$

$$1 - 1 = 0$$

Por ejemplo, para $t = -1$:

$$(x_3, y_3, z_3) = \left(\frac{1}{2}, 0,0\right) - 1(-1,1,1) = \left(\frac{3}{2}, -1, -1\right)$$

$$2\left(\frac{3}{2}\right) - 1 - 1 = 1$$

$$-1 + 1 = 0$$

Composición e inversa de mapas correspondientes a $\boldsymbol{L: F^n \rightarrow F^m}$

Simplemente corresponde a la composición de mapas lineales y su inversa se define como la obtención tanto por derecha como por izquierda del mapa identidad.

Factorización Única

A continuación, se mostrará características que permitirán demostrar el famoso teorema fundamental del álgebra. Estos requisitos nos favorecerán al momento de analizar matrices con valores propios. Como apartado y para análisis futuros, esto debe tenerse en consideración.

Teorema de división

Polinomio, definición.

Sea $p(x) = a_d x^d + \cdots + a_1 x + a_0$ un polinomio con coeficientes en un campo F. Recordando que si $a_d \neq 0$, entonces d es el grado de $p(x)$, y además a_d es llamado el coeficiente que lidera a $p(x)$. Si $p(x)$ tiene grado menos que cero (< 0), entonces $p(x) = a_0$, así, puede ser considerado como un elemento de un campo F, una constante o un escalar.

Si d es el grado de $p(x)$ y el coeficiente principal o líder es igual a uno, es decir, $a_d = 1$ entonces $p(x)$ es Mónico.

Número complejo

Suponga que $z = a + bi$, donde a y b son números reales.

- La parte real de z, denotada por $Re\ z$, está definida por $Re\ z = a$.
- La parte imaginaria de z, denotada por $Im\ z$, está definida por $Im\ z = b$.

Para todo número complejo z, tenemos

$$z = Re\ z + (Im\ z)i$$

Complejo conjugado, $\bar{z}$, valor absoluto, $|z|$

Suponga que $z \in C$

- El complejo conjugado de $z \in C$, denotado por $\bar{z}$, está definido por

$$\bar{z} = Re\ z - (Im\ z)i$$

- El valor absoluto del número complejo z, denotado por $|z|$, está definido por

$$|z| = \sqrt{(Rez)^2 + (Imz)^2}$$

Propiedades de los números complejos

Suponga que $w, z \in C$. Entonces las siguientes igualdades y desigualdades se cumplen

1. $z + \bar{z} = 2Rez$

Demostración. Sea $z = a + bi$ y $\bar{z} = a - bi$, entonces

$$a + bi + a - bi = a + a + 0 = 2a + 0 = 2Rez$$

2. $z - \bar{z} = 2(Imz)i$

Demostración. Sea $z = a + bi$ y $\bar{z} = a - bi$, entonces

$$a + bi - (a - bi) = a + bi - a + bi = 2bi = 2(Imz)i$$

3. $z\bar{z} = |z|^2$

Demostración. Sea $z = a + bi$ y $\bar{z} = a - bi$, entonces

$$(a + bi)(a - bi) = a^2 - abi + abi - (bi)^2 = a^2 - (bi)^2$$

$$= a^2 - \left(b^2 \left(-1^{\frac{1}{2}} \right)^{\frac{1}{2}} \right) = a^2 + b^2 = z\bar{z}$$

$$z\bar{z} = |z|^2$$

4. $\overline{w + z} = \bar{w} + \bar{z}$

Demostración. Sean $w = a + bi$ y $z = c + di$. Entonces,

$$w + z = (a + c) + (b + d)i$$

$$\overline{w + z} = (a + c) - (b + d)i = a - bi + c - di = \bar{w} + \bar{z}$$

5. $\overline{wz} = \bar{w}\bar{z}$

Demostración. Sean $w = a + bi$ y $z = c + di$. Entonces,

$$wz = (a + bi)(c + di) = ac + adi + bci + bdi^2$$
$$= ac - bd + (ad + bc)i$$

$$\overline{wz} = (ac - bd) - (ad + bc)i = ac - bd - adi - bci$$
$$= (a - bi)(c - di) = \bar{w}\bar{z}$$

6. $\bar{\bar{z}} = z$

Demostración. Sea $z = a + bi$ y $\bar{z} = a - bi$, entonces

$$\bar{\bar{z}} = a - (-bi) = a + bi = z$$

7. $|a| \leq |z|$ y $|b| \leq |z|$, tal que $a = Rez$ y $b = Imz$ con $z = a + bi$

Demostración. $|z| = \sqrt{a^2 + b^2} \geq 0$ debido a que a, b son números reales. En consecuencia,

$$|a| \leq \sqrt{a^2 + b^2}$$

Si $b = 0$, entonces $|a| = |z|$, por otra parte, si $b \neq 0$ entonces $|a| < |z|$. Lo mismo sucede para $|b| \leq |z|$.

8. $|w + z| \leq |a| + |b|$

Demostración. Sean $w = a + bi$ y $z = c + di$. Entonces,

$$|w + z|^2 = (w + z)(\bar{w} + \bar{z})$$
$$= \big((a + bi) + (c + di)\big)\big((a - bi) + (c - di)\big)$$

$$|w + z|^2 = \big((a + c) + (bi + di)\big)\big((a + c) - (bi + ci)\big)$$

$$|w+z|^2 = w\bar{w} + z\bar{z} + w\bar{z} + z\bar{w}$$

$$|w+z|^2 = |w|^2 + |z|^2 + 2Re(w\bar{z}) \leq |w|^2 + |z|^2 + 2|w\bar{z}|$$

$$|w+z|^2 \leq |w|^2 + |z|^2 + 2|w\bar{z}| = |w|^2 + |z|^2 + 2|w||z|$$

$$|w+z|^2 \leq |w|^2 + |z|^2 + 2|w||z| = (|w| + |z|)^2$$

$$|w+z|^2 \leq (|w| + |z|)^2$$

Tomando la raíz cuadrada en ambos lados,

$$|w+z| \leq |w| + |z|$$

Polinomios

Retomando que una función $p: F \to F$ es llamada un polinomio de grado m si existen $a_0, \dots, a_m \in F$ con $a_m \neq 0$ tal que

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

para todo $z \in F$. Un polinomio podría tener más de un grado si la representación de p en la forma de arriba no fuera única. La primera tarea es mostrar que esto no podría pasar. Por ejemplo, supongamos que hay un polinomio $q(z)$ que puede representarse de la siguiente manera

$$q(z) = b_0 + b_1 z + \cdots + b_m z^m$$

y asumamos que $q(z)$ también puede representarse de la siguiente manera

$$q(z) = c_0 + c_1 z + \cdots + c_n z^n$$

Por lo tanto, $q(z)$ tiene más de un grado, es decir, tiene por una parte grado m y por otra parte tiene grado n. La intención es probar que esto no puede pasar.

Ceros de un Polinomio

Un número $\alpha \in F$ es llamado un cero (o raíz) de un polinomio $p \in P(F)$ si

$$p(\alpha) = 0$$

Las soluciones de la ecuación $p(\alpha) = 0$ juegan un papel crucial en el estudio de un polinomio $p \in P(F)$.

Cada cero de un polinomio corresponde a un factor de 1 grado

Suponga que m es un entero positivo y que $p \in P(F)$ es un polinomio de grado m. Suponga que $\alpha \in F$. Entonces $p(\alpha) = 0$ si y solo si existe un polinomio $q \in P(F)$ de grado $m - 1$ tal que

$$p(z) = (z - \alpha)q(z)$$

para todo $z \in F$.

Demostración. Primero suponga que $p(\alpha) = 0$. Sean $a_0, a_1, \ldots, a_m \in F$ tales que

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

para todo $z \in F$. Entonces

exp 1. $p(z) = p(z) - p(\alpha) = (a_0 + a_1 z + \cdots + a_m z^m) - (a_0 + a_1\alpha + \cdots + a_m\alpha^m)$

$$p(z) = p(z) - p(\alpha) = a_1(z - \alpha) + \cdots + a_m(z^m - \alpha^m)$$

para todo $z \in F$. Para cada $k \in \{1, \ldots m\}$, la ecuación

$$z^k - \alpha^k = (z - \alpha)\sum_{j=1}^{k} \alpha^{j-1} z^{k-j}$$

muestra que $z^k - \alpha^k$ iguala a un polinomio de grado $k - 1$ por $(z - \alpha)$. Por ejemplo,

para $k = 1$

$$(z - \alpha)\sum_{j=1}^{1} \alpha^{j-1} z^{k-j} = (z - \alpha)\alpha^{1-1} z^{1-1} = (z - \alpha)$$

Para otra prueba, $k = 2$

$$\begin{aligned}(z - \alpha)\sum_{j=1}^{2} \alpha^{j-1} z^{k-j} &= (z - \alpha)[\alpha^{1-1} z^{2-1} + \alpha^{2-1} z^{2-2}] \\ &= (z - \alpha)(z^1 + \alpha^1) = (z^2 - \alpha^2) \\ &= (z - \alpha)(z + \alpha)\end{aligned}$$

Por tanto, la expresión (exp. 1) muestra que p iguala a un polinomio de grado $m - 1$ por $(z - \alpha)$ como era lo deseado.

Ahora para demostrar la implicación en la otra dirección, suponga que el polinomio $q \in P(F)$ es tal que $p(z) = (z - \alpha)q(z)$ para todo $z \in F$. Entonces

$$p(\alpha) = (\alpha - \alpha)q(\alpha) = 0$$

cómo era lo deseado.

El grado de un polinomio implica hasta $\boldsymbol{m}$ ceros

Suponga que m es un entero positivo y $p \in P(F)$ es un polinomio de grado m. Entonces p tiene máximo, m ceros en F.

Demostración. Usaremos inducción sobre m. Para ello, se prueba para $m = 1$:

$$p(z) = a_0 + a_1 z$$

Si $a_1 \neq 0$ entonces el polinomio $p(z)$ para $m = 1$ tiene únicamente un solo cero. Ahora si es válido para $m > 1$ y para $m - 1$, se probará que es válido para m.

Si p no tiene ceros en F, entonces el resultado se mantiene y se acaba la demostración, por ejemplo, el polinomio $x^2 + 1$ no tiene ceros con $F = R$.

Por ende, suponga que p tiene un cero $\alpha \in F$. Por el teorema anterior, hay un polinomio $q \in P(F)$ de grado $m - 1$ tal que

$$p(z) = (z - \alpha)q(z)$$

para todo $z \in F$. Nuestra hipótesis de inducción implica que q tiene máximo $m - 1$ ceros en F. La ecuación de arriba muestra que los de p en F son exactamente los de q en F en conjunto con α. Por tanto, p tiene máximo m ceros en F.

El resultado de arriba implica que los coeficientes de un polinomio son únicamente determinados (porque si un

polinomio tuviese dos diferentes conjuntos de coeficientes, entonces restando las representaciones del polinomio, entregaría como resultado un polinomio con algunos coeficientes diferentes de cero, pero con infinitos ceros). Por ejemplo, sea $p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$ y suponga que el mismo polinomio tiene otra representación $p(z) = b_0 + b_1 z + b_2 z^2 + \cdots + b_m z^m$. Restando ambas expresiones

$$0 = (a_0 - b_0) + z(a_1 - b_1) + z^2(a_2 - b_2) + \cdots + z^m(a_m - b_m)$$

se obtiene un polinomio con algunos coeficientes diferentes de cero que puede darse el caso, pero la particularidad es que otorga infinitos ceros. Para verificarlo, suponga que $z = 0$, entonces

$$a_0 = b_0$$

con $a_0, b_0 \in F$. Es decir, si $a_0 = b_0$ entonces $z = 0$ es un cero del polinomio obtenido, sin embargo, tener presente que toda la representación $(a_0 - b_0) + z(a_1 - b_1) + z^2(a_2 - b_2) + \cdots + z^m(a_m - b_m)$ es el polinomio cero y por definición **el grado del polinomio cero está definido como** $-\infty$. Incluso la expresión $(a_0 - b_0) + z(a_1 - b_1) + z^2(a_2 - b_2) + \cdots + z^m(a_m - b_m)$ podría expresarse como

$$\begin{aligned} 0 + 0z + 0z^2 + \cdots + 0z^m \\ = (a_0 - b_0) + z(a_1 - b_1) + z^2(a_2 - b_2) + \cdots \\ + z^m(a_m - b_m) = 0 \end{aligned}$$

Por ejemplo, supongamos que $z = 1$. Entonces

$$0 + 0(1) + 0(1)^2 + \cdots + 0(1)^m = 0$$

Por tanto, $z = 1$ es un cero. Supongamos ahora que $z = 3$, entonces

$$0 + 0(3) + 0(3^2) + \cdots + 0(3^m) = 0$$

Por tanto $z = 3$ es otro cero. En consecuencia, el polinomio cero tiene infinitos ceros.

Algoritmo de división para polinomios

La intención es tratar con números enteros positivos, por ejemplo, si p y s son números no negativos, con $s \neq 0$ existe un entero no negativo q y r tal que

$$p = sq + r$$

y $r < s$. Se piensa en el sentido de dividir p entre s, obteniendo como **cociente** q y con **residuo** r. Entonces, se tiene el siguiente resultado.

Suponga que $p, s \in P(F)$, con $s \neq 0$. Entonces existen los únicos polinomios $q, r \in P(F)$ tal que

$$p = sq + r$$

y $\deg r < \deg s$

Demostración. Sea $n = \deg p$ y $m = \deg s$. Si $n < m$, entonces se toma $q = 0$ y $r = p$ para obtener la ecuación deseada $p = sq + r$ con $\deg r < \deg s$. En otras palabras, tomando esta consideración en general, el grado de $\deg r = n$. Por tanto, ahora asumamos que $n \geq m$.

La lista

Exp 2. $1, z, \dots, z^{m-1}, s, zs, \dots, z^{n-m}s$

es linealmente independiente en $P_n(F)$ porque cada polinomio en esta lista tiene un grado diferente. Por ejemplo,

$$1, z, \dots, z^{m-1}, a_0 + a_1 z + \cdots + a_m z^m, z(a_0 + a_1 z + \cdots + a_m z^m), \dots, z^{n-m}(a_0 + a_1 z + \cdots + a_m z^m)$$

es la representación general de cada polinomio y en efecto tiene grados diferentes. Además, la lista tiene longitud (cantidad de elementos) $n + 1$ lo cual iguala a la dim $P_n(F)$. En consecuencia, la Exp.2 es una base de $P_n(F)$.

Debido a que $p \in P_n(F)$ y Exp 2. es una base de $P_n(F)$ existen únicas constantes $c_0, c_1, \dots, c_{m-1} \in F$ y $d_0, d_1, \dots, d_{n-m} \in F$ tal que

Exp 3. $p = c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + d_0 s + d_1 zs + \cdots + d_{n-m} z^{n-m} s$

$$p = (c_0 + c_1 z + \cdots + c_{m-1} z^{m-1}) + (d_0 s + d_1 zs + \cdots + d_{n-m} z^{n-m} s) = r + sq$$

con r y q definidos arriba, observamos que p puede ser escrito como $p = sq + r$ con $\deg r < \deg s$ como lo era deseado.

La **unicidad** de $q, r \in P(F)$ que satisfacen estas condiciones viene de la unicidad de las constantes $c_0, c_1, \dots, c_{m-1} \in F$ y $d_0, d_1, \dots, d_{n-m} \in F$ que satisfacen la expresión Exp 3.

Factorización de polinomios sobre Complejos (C)

Se tratará ahora sobre el teorema fundamental del algebra el cual implícitamente usa como resultados funciones reales evaluadas sobre un disco cerrado en R^2 donde se alcanza a un mínimo valor.

Si k es un entero positivo y $\theta \in R$ entonces

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$$

Demostración. Se lo probará por inducción.

Para $i = 1$

$$(\cos\theta + i\sin\theta)^1 = \cos 1\theta + i\sin 1\theta = \cos\theta + i\sin\theta$$

Ahora por hipótesis es válido para $i = n$. Entonces para $i = n + 1$ se tiene

$$(\cos\theta + i\sin\theta)^{n+1} = \cos((n+1)\theta) + i\sin((n+1)\theta)$$

En consecuencia,

$$\cos((n+1)\theta) + i\sin((n+1)\theta) = \cos(n\theta + \theta) + i\sin(n\theta + \theta)$$

$$\begin{aligned}\cos(n\theta + \theta) + i\sin(n\theta + \theta) &= \cos(n\theta)\cos(\theta) - \sin(n\theta)\sin(\theta) \\ &\quad + i\sin(n\theta)\cos(\theta) + i\cos(n\theta)\sin(\theta)\end{aligned}$$

$$= \cos(\theta)\,[\cos(n\theta) + i\sin(n\theta)] + \sin(\theta)\,[i\cos(n\theta) - \sin(n\theta)]$$

$$= \cos(\theta)\,[\cos(n\theta) + i\sin(n\theta)] + i\sin(\theta)\,[\cos(n\theta) + i\sin(n\theta)]$$

$$= [\cos(\theta) + i\sin(\theta)]^1[\cos(n\theta) + i\sin(n\theta)]$$

$$= [\cos(\theta) + i\sin(\theta)]^1[\cos(\theta) + i\sin(\theta)]^n$$

$$= [\cos(\theta) + i\sin(\theta)]^{n+1}$$

Teorema fundamental del algebra, primera versión

Todo polinomio que no sea una constante con coeficientes complejos tiene un cero en C.

Demostración. El teorema de Moivre, el cual puede ser probado por inducción sobre k y las fórmulas de adición para seno y coseno, dictamina que si k es un entero positivo y $\theta \in R$, entonces

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$$

Suponga que $w \in C$ y que k es un entero positivo. Usando coordenadas polares, sabemos que existe un $r \geq 0$ y $\theta \in R$ tal que

$$r(\cos\theta + i\sin\theta) = w$$

El teorema de Moivre implica que

$$\left(r^{\frac{1}{k}}\left(\cos\left(\frac{\theta}{k}\right) + i\sin\left(\frac{\theta}{k}\right)\right)\right)^k = w$$

Por ende, todo número complejo tiene una $k - ésima$ raíz, un hecho que se usará luego.

Suponga que p es un polinomio que no es una constante con coeficientes complejos y con el término más alto diferente de

cero como $c_m z^m$. Entonces $|p(z)| \to \infty$ mientras $|z| \to \infty$ (porque $\frac{|p(z)|}{|z^m|} \to |c_m|$ mientras $|z| \to \infty$). Por ejemplo,

$$p(z) = c_0 + c_1 z + \cdots + c_m z^m$$

$$z = a + bi$$

$$|z| = \sqrt{a^2 + b^2}$$

Si se reemplazara a z como un número complejo cualquiera $a + bi$ en el que $\forall a, b \in R$ el polinomio $p(z)$ podría expresarse como

$$p(z) = c_0 + c_1(a + bi) + \cdots + c_m(a + bi)^m$$

Si se sacara el valor absoluto a $p(z)$ se obtendría

$$\begin{aligned} |p(z)| &= |c_0 + c_1(a + bi) + \cdots + c_m(a + bi)^m| \\ &= |c_0 + c_1(z) + \cdots + c_m(z)^m| \end{aligned}$$

Si se divide por $|z^m|$, se obtiene

$$\begin{aligned} &\frac{|c_0 + c_1(a + bi) + \cdots + c_m(a + bi)^m|}{|z^m|} \\ &\quad = \frac{|c_0|}{|z^m|} + \frac{|c_1(a + bi)|}{|z^m|} + \cdots + \frac{|c_m(a + bi)^m|}{|z^m|} \end{aligned}$$

Aplicando el límite cuando z^m tiende al infinito se consigue

$$\lim_{|z^m|\to\infty}\left[\frac{|c_0|}{|z^m|}+\frac{|c_1(a+bi)|}{|z^m|}+\cdots+\frac{|c_m(a+bi)^m|}{|z^m|}\right]$$
$$=\lim_{|z^m|\to\infty}\frac{|c_0|}{|z^m|}+\lim_{|z^m|\to\infty}\frac{|c_1(a+bi)|}{|z^m|}+\cdots$$
$$+\lim_{|z^m|\to\infty}\frac{|c_m(a+bi)^m|}{|z^m|}$$
$$=\lim_{|z^m|\to\infty}\frac{|c_0|}{|z^m|}+\lim_{|z^m|\to\infty}\frac{|c_1(z)|}{|z^m|}+\cdots$$
$$+\lim_{|z^m|\to\infty}\frac{|c_m(z)^m|}{|z^m|}=0+0+\cdots+|c_m|=|c_m|$$

Por ende, la función continua $z\to|p(z)|$ tiene un mínimo global en algún punto $\delta\in C$. Para lograr mostrar que $p(\delta)=0$, suponga que $p(\delta)\neq 0$.

Defínase un nuevo polinomio q por

$$q(z)=\frac{p(z+\delta)}{p(\delta)}$$

para todo $z\in C$.

La función $z\to|q(z)|$ tiene un valor mínimo global de 1 en $z=0$, por ejemplo:

$$q(0)=\frac{p(\delta)}{p(\delta)}=1,\qquad p(\delta)\neq 0$$

El valor absoluto es entonces $|q(0)|=|1|=1$. Al referirse a su valor mínimo se puede concluir que $|q(0)|<|q(z)|,\forall\, z\in C$. El polinomio $q(z)$ entonces se puede escribir

$$q(z)=\frac{p(z+\delta)}{p(\delta)}=1+a_kz^k+\cdots+a_mz^m$$

donde k es el entero positivo más pequeño tal que el coeficiente que acompaña z^k es diferente de cero, en otras palabras, $q(z)$ solamente se lo reescribió con el fin de que $a_k \neq 0$.

Ahora bien, sea $\beta \in C$ tal que $\beta = -\frac{1}{a_k}$. Existe una constante $c > 1$ tal que si $t \in (0,1)$, entonces

$$|q(t\beta)| = \left|1 + a_k(t\beta)^k + \cdots + a_m(t\beta)^m\right|$$
$$= \left|1 + a_k(t \times -\frac{1}{a_k})^k + \cdots + a_m(t \times -\frac{1}{a_k})^m\right|$$

$$|q(t\beta)| = \left|1 + a_k(t\beta)^k + a_{k+1}(t\beta)^{k+1} + \cdots + a_m(t\beta)^m\right|$$

Utilizando la desigualdad triangular en el que $|w + r| \leq |w| + |r|$, para el caso de que $r = s + z$, entonces $|w + s + z| \leq |w| + |s + z| \leq |w| + |s| + |z|$. En otras palabras

$$\left|1 + a_k(t\beta)^k + a_{k+1}(t\beta)^{k+1} + \cdots + a_m(t\beta)^m\right|$$
$$\leq \left|1 + a_k(t\beta)^k\right|$$
$$+ \left|a_{k+1}(t\beta)^{k+1} + \cdots + a_m(t\beta)^m\right|$$

$$\left|1 + a_k(t\beta)^k\right| + \left|a_{k+1}(t\beta)^{k+1} + \cdots + a_m(t\beta)^m\right|$$
$$= \left|1 + a_k(t\beta)^k\right|$$
$$+ \left|t^{k+1}\left(a_{k+1}\beta^{k+1} + \cdots + \frac{a_m t^m \beta^m}{t^{k+1}}\right)\right|$$

$$\left|1 + a_k(t\beta)^k\right| + \left|t^{k+1}\left(a_{k+1}\beta^{k+1} + \cdots + \frac{a_m t^m \beta^m}{t^{k+1}}\right)\right|$$
$$= \left|1 + a_k(t\beta)^k\right|$$
$$+ \left|t^{k+1}\right|\left|a_{k+1}\beta^{k+1} + \cdots + \frac{a_m t^m \beta^m}{t^{k+1}}\right|$$
$$= \left|1 + a_k(t\beta)^k\right| + t^{k+1}c$$

En consecuencia,

$$|q(t\beta)| \leq \left|1 + a_k(t\beta)^k\right| + t^{k+1}c = 1 - t^k(1 - tc)$$

$$|q(t\beta)| \leq 1 - t^k(1 - tc)$$

Por ende, si se toma $t = \frac{1}{2c}$

$$\left|q\left(\frac{1}{2c}\beta\right)\right| \leq 1 - \left(\frac{1}{2c}\right)^k\left(1 - \frac{1}{2c}c\right) = 1 - \left(\frac{1}{2c}\right)^k\left(\frac{1}{2}\right)$$
$$= 1 - \frac{1^{k+1}}{2^{k+1}c^k} < 1$$

La desigualdad para $t = \frac{1}{2c}$

$$\left|q\left(\frac{1}{2c}\beta\right)\right| < 1$$

contradice que el mínimo global de $z \to |q(z)|$ es 1. Esta contradicción implica necesariamente que $p(\delta) = 0$, mostrado que $p(z)$ tiene un cero tal como era lo deseado.

Mapa Identidad

Si S es cualquier conjunto, se define como al mapa identidad denotado por I_S aquel que cumple con $I_S(x) = x$ para todo $x \in$

S. Notamos que el mapa identidad es tanto inyectivo como sobreyectivo. Si no se especifica la referencia para S (debido a que suele estar clara por contexto), entonces escribimos I en lugar de I_S.

Mapa Inverso

Sea $F: S \to S'$ un mapa desde un conjunto hacia otro conjunto. Decimos que F tiene una inversa si existe algún mapa $G: S' \to S$ tal que

$$G \circ F = I_S \quad \& \quad F \circ G = I_{S'}$$

Sea

$$f: S \to S'$$

Entonces existe $\exists x, y \in S$ tal que $f(x) = f(y)$.

Demostración.

Se utilizará el método por contradicción, es decir, asumir que el consecuente es falso y luego resumir que la conclusión contradice al consecuente como falso. Por ende, supongamos que $f(x) \neq f(y)$. En consecuencia,

$$f(x) - f(y) \neq 0$$

Sea $x = y$, entonces $f(x) - f(y) = f(x) - f(x) = f(y) - f(y) = 0$ lo cual conlleva a decir que $0 \neq 0$ lo cual no es posible. En consecuencia, existe $\exists x, y \in S$ tal que $f(x) = f(y)$.

Sea

$$f: S \to S'$$

un mapa el cual tiene un mapa inverso g. Entonces f es tanto inyectivo como sobreyectivo, es decir, f es inyectiva.

Demostración.

Sea $x, y \in S$. Sea $g: S' \to S$ la inversa de f. Si $f(x) = f(y)$, cual asunción no tiene pérdida de generalidad ya que f es un mapa y es posible sin lugar a duda tal asunción, entonces debemos tener

$$x = g\big(f(x)\big) = g\big(f(y)\big) = y$$

y por ende f es inyectiva. Para probar que f es sobreyectiva, sea $z \in S'$. Entonces

$$f\big(g(z)\big) = z$$

por definición de lo que es un mapa inverso, y por tanto $z = f(x)$ donde $x = g(z)$. Esto prueba que f es sobreyectiva.

Sea $f: S \to S'$ un mapa el cual es biyectivo. Entonces f tiene un mapa inverso.

Demostración.

Dado $z \in S'$, debido a que f es sobreyectiva, existe $x \in S$ tal que $f(x) = z$. Debido a que f es inyectiva, este elemento x es únicamente determinado por z, y podemos entonces definir

$$g(z) = x$$

Por la definición de g, encontramos que $f\big(g(z)\big) = z$, y $g\big(f(x)\big) = x$, y así, g es un mapa inverso para f.

Isomorfismo

Un mapa lineal $F: U \to V$ el cual tiene una inversa $G: V \to U$ (también decimos que es invertible) es llamado un **isomorfismo.** Por ejemplo, el mapa

$$Q: V \to F^n$$

en el que $\dim V = n$ y $Q(v) = x_1 w_1 + \cdots + x_n w_n$ con $v_1, \dots, v_n$ una base de V es un isomorfismo.

Demostración.

Sea $v_1, \dots, v_n$ una base de V y sea $w_1, \dots, w_n$ una base de F^n, entonces

$$Q(x_1 v_1 + \cdots + x_n v_n) = x_1 w_1 + \cdots + x_n w_n$$

es un mapa bien definido sin ambigüedad. Además, Q es sobreyectivo debido a que $w_1, \dots, w_n$ extiende a F^n. Más aún, el espacio nulo de Q es O debido a que $w_1, \dots, w_n$ son linealmente independientes y por tanto Q es inyectiva. En consecuencia, Q tiene una inversa y por ende es un isomorfismo.

Cambio de Base

Sea V un espacio vectorial de dimensión n sobre un campo F, y sea la lista $B_1 = v_1, v_2, \dots, v_n$ una base de V. Cualquier vector $x \in V$ puede ser representado como $x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ porque B_1 extiende a V. Si existiera otra representación de $x = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$ en la misma base, entonces

$$0 = x - x = (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n$$

de lo cual se sigue que todo $\alpha_i - \beta_i = 0$ debido a que la lista B_1 es linealmente independiente. Dada la base B_1, el mapa lineal

$$x \to [x]_{B_1} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad en\ la\ cual\ x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

desde V hacia F^n está bien definido, no existe ambigüedad del caso y es inyectivo y sobreyectivo. Los escalares α_i son las coordenadas de x con respecto a la base B_1, y el vector columna $[x]_{B_1}$ es el único $B_1 - representación\ coordenada$ de x.

Sea $T: V \to V$ una transformación lineal dada. La acción de T sobre cualquier $x \in V$ es determinada una vez que se conoce los vectores $Tv_1, Tv_2, \dots, Tv_n$, debido a que cualquier $x \in V$ tiene una única representación $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$ y $Tx = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = T(\alpha_1 v_1) + \cdots + T(\alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n)$ por linealidad. Por ende, el valor de Tx es determinado una vez que $[x]_{B_1}$ es conocido.

Sea $B_2 = \{w_1, w_2, \dots, w_n\}$ también una base de V y además $Tv_1, Tv_2, \dots, Tv_n \in V$ entonces cualquier $Tv_j,\ j = 1, \dots, n$ puede ser expresado como combinación lineal de los elementos de B_2. Por ejemplo

$$Tv_1 = t_1 w_1 + t_2 w_2 + \cdots + t_n w_n$$

De igual manera Tv_2 puede ser expresado como

$$Tv_2 = k_1 w_1 + k_2 w_2 + \cdots + k_n w_n$$

En general para definir solamente una letra, se puede utilizar t_{1j} en el que j se asocia a cada elemento $Tv_j,\ j = 1, \dots, n$. Entonces

$$Tv_j = t_{1j} w_1 + \cdots + t_{nj} w_n$$

Sea $B_2 = \{w_1, w_2, \dots, w_n\}$ también una base de V (ya sea diferente o de la misma base B_1) y suponga que la representación $B_2 - coordenada$ de Tv_j es

$$\left[Tv_j\right]_{B_2} = \begin{bmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{bmatrix}, \quad j = 1,2, \dots, n$$

Entonces, para cualquier $x \in V$, tenemos

$$[Tx]_{B_2} = \left[\sum_{j=1}^{n} \alpha_j\, Tv_j\right]_{B_2} = \sum_{j=1}^{n} \alpha_j \left[Tv_j\right]_{B_2}$$

$$[Tx]_{B_2} = \sum_{j=1}^{n} \alpha_j \begin{bmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{bmatrix} = \begin{bmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

La matriz de tamaño $n \times n$ definida como $[t_{ij}]$ depende de T y sobre la selección de la base B_1 y B_2, pero no depende de x. Se define como la *representación de base* $B_1 - B_2$ *de* T como

$$ {}_{B_2}[T]_{B_1} = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{bmatrix} = [[Tv_1]_{B_2} \ldots [Tv_n]_{B_2}] $$

Se ha mostrado que $[Tx]_{B_2} = {}_{B_2}[T]_{B_1}[x]_{B_1}$ para cualquier $x \in V$. En el importante especial caso de que $B_2 = B_1$, tenemos que ${}_{B_1}[T]_{B_1}$ la cual es llamada *representación de la base* B_1 *de* T.

Considere la transformación lineal identidad $I: V \to V$ definida por $Ix = x, \forall x$. Entonces

$$ \begin{aligned} [x]_{B_2} = [Ix]_{B_2} = {}_{B_2}[I]_{B_1}[x]_{B_1} = {}_{B_2}[I]_{B_1}[Ix]_{B_1} \\ = {}_{B_2}[I]_{B_1\,B_1}[I]_{B_2}[x]_{B_2} \end{aligned} $$

para todo $x \in V$. Sucesivamente escogiendo $x = w_1, w_2, \ldots, w_n$, esta identidad permite identificar cada columna de ${}_{B_2}[I]_{B_1\,B_1}[I]_{B_2}$ y muestra que

$$ {}_{B_2}[I]_{B_1\,B_1}[I]_{B_2} = I_n $$

Por ejemplo, con $x = w_1$

$$ [w_1]_{B_2} = {}_{B_2}[I]_{B_1\,B_1}[I]_{B_2}[w_1]_{B_2} = {}_{B_2}[I]_{B_1\,B_1}[I]_{B_2} \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} $$

y en consecuencia se deduce por inspección que ${}_{B_2}[I]_{B_1\,B_1}[I]_{B_2} = I_n$.

Si se realiza el mismo cálculo empezando con $[x]_{B_1} = [Ix]_{B_1} = \cdots$, se encuentra que

$$ {}_{B_1}[I]_{B_2}\, {}_{B_2}[I]_{B_1} = I_n $$

Por tanto, toda matriz de la forma ${}_{B_2}[I]_{B_1}$ es invertible y ${}_{B_1}[I]_{B_2}$ es su inversa.

Contrariamente, toda matriz invertible $S = [s_1\ s_2 \ldots s_n] \in M_n(F)$ (es decir, S tiene inversa) tiene la forma ${}_{B_1}[I]_B$ para alguna base B.

Podríamos tomar a B como los vectores $\{\widetilde{s_1}, \widetilde{s_2}, \ldots, \widetilde{s_n}\}$ definidos por $[\widetilde{s_i}]_{B_1} = s_i,\ \ i = 1,2, \ldots, n.$ Por ejemplo, supongamos que

$$ s_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} $$

Entonces al definir la lista B por $[\widetilde{s_1}]_{B_1} = s_1$, se tiene

$$ s_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = [\widetilde{s_1}]_{B_1} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} $$

La matriz invertible S entonces está definida por

$$ \begin{aligned} S = \left[[\widetilde{s_1}]_{B_1}\ [\widetilde{s_2}]_{B_1} \ldots [\widetilde{s_n}]_{B_1}\right] &= {}_{B_1}[I]_B \\ &= \left[[I\widetilde{s_1}]_{B_1}\ [I\widetilde{s_2}]_{B_1} \ldots [I\widetilde{s_n}]_{B_1}\right] \end{aligned} $$

$$ S = \left[[I\widetilde{s_1}]_{B_1}\ [I\widetilde{s_2}]_{B_1} \ldots [I\widetilde{s_n}]_{B_1}\right] = [s_1\ s_2 \ldots s_n] $$

Sin embargo, S es invertible, por ende

$$\left[[\widetilde{s_1}]_{B_1}\ [\widetilde{s_2}]_{B_1} \dots [\widetilde{s_n}]_{B_1}\right]S^{-1} = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{bmatrix} S^{-1}$$

$$= I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{bmatrix} S^{-1}$$

$$= \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} \alpha_{11}^{-1} & \dots & \alpha_{1j}^{-1} & \cdots & \alpha_{1n}^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1}^{-1} & \cdots & \alpha_{nj}^{-1} & \cdots & \alpha_{nn}^{-1} \end{bmatrix}$$

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} h_{11} & \dots & h_{1j} & \cdots & h_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{n1} & \cdots & h_{nj} & \cdots & h_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

La lista B es linealmente independiente porque S es invertible a pesar de que ya se menciona que B es una base. Note que

$$_{B_2}[I]_{B_1} = \left[[Iv_1]_{B_2} \dots [Iv_n]_{B_2}\right] = \left[[v_1]_{B_2} \dots [v_n]_{B_2}\right]$$

así $_{B_2}[I]_{B_1}$ describe como los elementos de la base B_1 son formados por los elementos de la base B_2.

Ejemplo. Sea la transformación $T: R^2 \to R^2$ en donde R denota el conjunto de los números reales tal que

$$T(x, y) = (2x, 2y), \quad \forall (x, y) \in R^2$$

a. Verifique si T es lineal

b. Defina una base para tanto R^2 en el dominio y R^2 en la imagen de T.

c. Verifique si se cumple la siguiente igualdad

$$_{B_1}[T]_{B_2}[x]_{B_2} = {}_{B_2}[T]_{B_1}[x]_{B_1}$$

Para $x = (1,1)$.

Solución.

Para verifique que T sea lineal se debe probar dos cosas: $T[(x_1, y_1) + (x_2, y_2)] = T(x_1, y_1) + T(x_2, y_2)$

$T[(x_1, y_1) + (x_2, y_2)] = T(x_1 + x_2, y_1 + y_2) = \big(2(x_1 + x_2), 2(y_1 + y_2)\big)$

$$T(x_1, y_1) = (2x_1, 2y_1)$$

$$T(x_2, y_2) = (2x_2, 2y_2)$$

$$\begin{aligned} T(x_1, y_1) + T(x_2, y_2) &= (2x_1, 2y_1) + (2x_2, 2y_2) \\ &= (2x_1 + 2x_2, 2y_1 + 2y_2) \\ &= \big(2(x_1 + x_2), 2(y_1 + y_2)\big) \\ &= T[(x_1, y_1) + (x_2, y_2)] \end{aligned}$$

Para la homogeneidad se debe cumplir que $T\big(r(x,y)\big) = rT(x,y) \quad r \in F$

$$T\big(r(x,y)\big) = T(rx, ry) = (2rx, 2ry) = r(2x, 2y) = rT(x,y)$$

Se ha probado a.

Para probar b se utilizan las siguientes bases denotadas por B_1 y B_2. Para la primera, se utilizará la base estándar o canónica

y para la segunda dos elementos que no sean múltiplos del otro.

$$B_1 = \{(1,0),(0,1)\} \qquad B_2 = \{(2,1),(3,5)\}$$

La primera base B_1 es la trivial sin embargo la segunda es fácil comprobar.

$$(0,0) = a(2,1) + b(3,5)$$

$$(0,0) = (2a + 3b, a + 5b)$$

De donde $2a + 3b = 0,\ \ a + 5b = 0$. Se consigue que $a = -5b$ y reemplazando

$$2(-5b) + 3b = 0,\ \ b = 0$$

Con lo que se concluye que $a = 0$. Por otra parte, todo vector $v \in R^2$ se puede expresar de la forma

$$v = x(2,1) + y(3,5)$$

lo que significa que B_2 extiende a R^2 y por tanto forma una base para R^2.

Para probar c, ahora se tiene que encontrar

$$_{B_1}[T]_{B_2} = \left[[T(2,1)]_{B_1}[T(3,5)]_{B_1}\right]$$

$$T(2,1) = (4,2) = a_1(1,0) + a_2(0,1), \quad a_1 = 4, a_2 = 2$$

$$T(3,5) = (6,10) = b_1(1,0) + b_2(0,1), \quad b_1 = 6, b_2 = 10$$

$$_{B_1}[T]_{B_2} = \begin{bmatrix} 4 & 6 \\ 2 & 10 \end{bmatrix}$$

Para encontrar la otra matriz, se debe expresar como

$$_{B_2}[T]_{B_1} = \left[[T(1,0)]_{B_2}[T(0,1)]_{B_2}\right]$$

$$T(1,0) = (2,0) = a_1(2,1) + a_2(3,5) = (2a_1 + 3a_2, a_1 + 5a_2) = (2,0)$$

$$2a_1 + 3a_2 = 2, \quad a_1 + 5a_2 = 0, \quad a_1 = -5a_2$$

$$2(-5a_2) + 3a_2 = -7a_2 = 2, \quad a_2 = -\frac{2}{7}, \quad a_1 = -5\left(-\frac{2}{7}\right) = \frac{10}{7}$$

$$T(0,1) = (0,2) = b_1(2,1) + b_2(3,5) = (2b_1 + 3b_2, b_1 + 5b_2) = (0,2)$$

$$2b_1 + 3b_2 = 0, \quad b_1 + 5b_2 = 2, \quad 2b_1 = -3b_2$$

$$2b_1 + 10b_2 = 4, \qquad -3b_2 + 10b_2 = 7b_2 = 4, \quad b_2 = \frac{4}{7}$$

$$b_1 = -\frac{3}{2}b_2 = -\frac{3}{2}\left(\frac{4}{7}\right) = -\frac{12}{14} = -\frac{6}{7}$$

$$_{B_2}[T]_{B_1} = \begin{bmatrix} \frac{10}{7} & -\frac{6}{7} \\ -\frac{2}{7} & \frac{4}{7} \end{bmatrix}$$

Se debe comprobar que

$$_{B_2}[T]_{B_1}[x]_{B_1} = {}_{B_1}[T]_{B_2}[x]_{B_2}$$

Sea $x = (1,1)$

$$(1,1) = c_1(1,0) + c_2(0,1), \quad c_1 = 1, c_2 = 1$$

Entonces $[x]_{B_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Sea $x = (1,1)$

$$(1,1) = d_1(2,1) + d_2(3,5) = (2d_1 + 3d_2, d_1 + 5d_2) = (1,1)$$

$$2d_1 + 3d_2 = 1, \;\; d_1 + 5d_2 = 1$$

$$2d_1 = 1 - 3d_2, \;\; 2d_1 + 10d_2 = 2$$

$$1 - 3d_2 + 10d_2 = 1 + 7d_2 = 2, \;\; d_2 = \frac{1}{7}$$

$$d_1 = 1 - 5\left(\frac{1}{7}\right) = \frac{7}{7} - \frac{5}{7} = \frac{2}{7} = d_1$$

Entonces $[x]_{B_2} = \begin{bmatrix} \frac{2}{7} \\ \frac{1}{7} \end{bmatrix}$. Verificando la comprobación se tiene que

$${}_{B_2}[T]_{B_1}[x]_{B_1} = \begin{bmatrix} \frac{10}{7} & -\frac{6}{7} \\ -\frac{2}{7} & \frac{4}{7} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5714 \\ 0.2857 \end{bmatrix}$$

$${}_{B_2}[T]_{B_1}[x]_{B_1} = \begin{bmatrix} 4 & 6 \\ 2 & 10 \end{bmatrix} \begin{bmatrix} \frac{2}{7} \\ \frac{1}{7} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Basta con haber encontrado un solo contraejemplo que demuestra que

$${}_{B_2}[T]_{B_1}[x]_{B_1} \neq {}_{B_1}[T]_{B_2}[x]_{B_2}$$

Ejemplo 3. Sea la transformación $S: R^3 \to R^3$ definida por

$$S(x, y, z) = (x + y, y, z)$$

Determine lo siguiente:

1. Linealidad de S sobre la función $R^3 \to R^3$
2. Defina una base para R^3 en el dominio y una para la imagen de S sobre R^3.
3. Sea el elemento $x = (0,0,0)^T = O$, verifique si ${}_{B_1}[T]_{B_2}[x]_{B_2} = {}_{B_2}[T]_{B_1}[x]_{B_1}$.

Se deja como ejercicio al lector.

Ahora sea $x \in V$ y compute

$$\begin{aligned} {}_{B_2}[T]_{B_2}[x]_{B_2} = [Tx]_{B_2} = [I(Tx)]_{B_2} &= {}_{B_2}[I]_{B_1}[Tx]_{B_1} \\ &= {}_{B_2}[I]_{B_1}{}_{B_1}[T]_{B_1}[x]_{B_1} \end{aligned}$$

$$\begin{aligned} {}_{B_2}[T]_{B_2}[x]_{B_2} &= {}_{B_2}[I]_{B_1}{}_{B_1}[T]_{B_1}[Ix]_{B_1} \\ &= {}_{B_2}[I]_{B_1}{}_{B_1}[T]_{B_1}{}_{B_1}[I]_{B_2}[x]_{B_2} \end{aligned}$$

Escogiendo $x = w_1, w_2, \dots, w_n$ sucesivamente, se concluye que

$${}_{B_2}[T]_{B_2} = {}_{B_2}[I]_{B_1}{}_{B_1}[T]_{B_1}{}_{B_1}[I]_{B_2}$$

Esta identidad muestra como la base B_1 que representa a T cambia si la base es cambiada por B_2. Por esta razón, la matriz ${}_{B_2}[I]_{B_1}$ es llamada la matriz de cambio de base de $B_1 - B_2$.

Cualquier matriz $A \in M_n(F)$ es una representación de la base de alguna transformación lineal $T: V \to V$, por si B es cualquier base de V, podemos determinar Tx mediante $[Tx]_B = A[x]_B$.

Para esta transformación lineal T, una computación revela que ${}_B[T]_B = A$.

Eigenvalores, Eigenvectores y Similaridad

Cambio de base y similaridad

Toda matriz invertible es una matriz de cambio de base, y toda matriz de cambio de base es invertible según lo que se demostró en líneas anteriores. Por ende, si B es una base dada de un espacio vectorial V, si T es una transformación lineal dada sobre V, y si $A = {}_B[T]_B$ es la base de la representación de T, el conjunto de todas las posibles bases que representan a T es

$$\{{}_{B_1}[I]_B \, {}_B[T]_B \, {}_B[I]_{B_1} : \ B_1 \ es \ una \ base \ de \ V\}$$

$$= \{S^{-1}AS : S \in M_n(F) es \ invertible\}$$

Este solo es el conjunto de todas las matrices que son *similares* a una matriz A dada. Similar pero no matrices idénticas son entonces diferentes bases de representación de una sola transformación lineal. La noción de similaridad es un concepto clave para este apartado.

Máximo restringido y eigenvalores

Un segundo concepto clave es la noción de $eigenvector$ y $eigenvalor$. Vectores diferentes de cero x tal que Ax es un múltiplo escalar de x juegan un rol mayor en analizar la estructura de una matriz o transformación lineal, pero aquellos vectores surgen en el contexto más elemental de maximizar o minimizar una forma cuadrática simétrica sujetos a una restricción geométrica: Para una matriz real simétrica $A \in M_n(\boldsymbol{R})$,

Maximizar $x^T Ax$,	Sujeto a $x \in \boldsymbol{R}^n$,	$x^T x = 1$

Un enfoque convencional para tal problema restringido de optimización es introducir el $Lagrangiano\ L = x^T Ax - \lambda x^T x$. Condiciones necesarias para un extremo son

$$0 = \nabla L = 2(Ax - \lambda x) = 0$$

Por ende, si un vector $x \in R^n$ con $x^T x = 1$ (y en consecuencia $x \neq 0$) es un extremo de $x^T Ax$, debe satisfacer la ecuación $Ax = \lambda x$.

Resulta que un escalar λ tal que $Ax = \lambda x$ para algún vector diferente de cero x es llamado un $eigenvalor$ de A.

Ejemplo. Use el teorema de Weirstrass para explicar por qué el problema de un extremo restringido tiene una solución. El problema dice lo siguiente

Vectores diferentes de cero x tal que Ax es un múltiplo escalar de x juegan un rol mayor en analizar la estructura de una matriz o transformación lineal, pero aquellos vectores surgen en el contexto más elemental de maximizar o minimizar una forma cuadrática simétrica sujetos a una restricción geométrica, es decir, para una matriz real simétrica $A \in M_n(R)$

Maximizar $x^T Ax$,	Sujeto a $x \in \boldsymbol{R}^n$,	$x^T x = 1$

Concluya que toda matriz real simétrica tiene al menos un eigenvalor real.

Solución.

El teorema de Weirstrass dice lo siguiente: Sea S un subconjunto compacto de un espacio vectorial V real o complejo de dimensión finita con una normal $\|\cdot\|$, y sea $f: S \to R$ una función continua. Existe un punto $x_{\min} \in S$ tal que

$$f(x_{min}) \leq f(x) \ \ para\ toda\ x \in S$$

y un punto $x_{max} \in S$ tal que

$$f(x) \leq f(x_{max}) \ \ para\ toda\ x \in S$$

Esto es, f alcanza su mínimo y máximo valor sobre S.

Para comprender un poco mejor este teorema, se escribe ciertas referencias de acuerdo con un espacio vectorial real o complejo de dimensión finita. Sea V un espacio vectorial real o complejo de dimensión finita con una norma dada $\|\cdot\|$.

La $bola\ cerrada\ de\ radio\ \epsilon$ alrededor de $x \in V$ es $B_\epsilon(x) = \{y \in V: \|y - x\| \leq \epsilon\}$; la correspondiente $bola\ abierta$ es $B_\epsilon(x) = \{y \in V: \|y - x\| < \epsilon\}$. Por ejemplo, el espacio vectorial de los números reales para R^1, tiene definido su producto escalar como

$$\langle x, y\rangle = x \cdot y, \quad x, y \in R$$

Ahora, su norma se define por $\|x\| = \sqrt{< x, x >}$. Para $y \in R$ tal que $y = \frac{1}{2}$ y definiendo la bola cerrada de radio $\epsilon = 1$ se tiene

$$B_{\epsilon=1}(x) = \left\{\frac{1}{2} \in R^1: \left\|\frac{1}{2} - x\right\| \leq 1\right\}$$

$$= \left\{\frac{1}{2} \in R^1: \sqrt{< \frac{1}{2} - x, \frac{1}{2} - x >} \leq 1\right\}$$

$$B_{\epsilon=1}(x) = \left\{\frac{1}{2} \in R^1: \sqrt{< \frac{1}{2} - x, \frac{1}{2} - x >} \leq 1\right\}$$

$$\left(\frac{1}{2} - x\right)\left(\frac{1}{2} - x\right) = \frac{1}{4} - 2x + x^2 = \left(\frac{1}{2} - x\right)^2$$

de tal forma

$$B_{\epsilon=1}(x) = \left\{\frac{1}{2} \in R^1: \sqrt{\left(\frac{1}{2} - x\right)^2} \leq 1\right\} = \left\{\frac{1}{2} \in R^1: \left|\frac{1}{2} - x\right| \leq 1\right\}$$

Por otra parte, si fuese la bola abierta con $\epsilon = 1$ se tiene

$$B_{\epsilon=1}(x) = \left\{ \frac{1}{2} \in R^1 : \sqrt{\left(\frac{1}{2} - x\right)^2} \leq 1 \right\} = \left\{ \frac{1}{2} \in R^1 : \left| \frac{1}{2} - x \right| < 1 \right\}$$

Otra acotación es sobre la definición de un conjunto y características de este. Es decir, un conjunto $S \subseteq V$ es abierto si, para cada $x \in S$, existe un $\epsilon > 0$ tal que $B_\epsilon(x) \subseteq S$. Por otra parte, un conjunto $S \subseteq V$ es cerrado si el complemento de S en V es abierto.

Un conjunto S es $acotado$ si hay un $r > 0$ tal que $S \subseteq B_r(0)$.

Ejemplo. Sea $V = \boldsymbol{R^1}$ y un conjunto $S \subseteq V$ definido como

$$S = \{y \in V : 3 > y > 1\}$$

Determine si S es abierto o cerrado y acotado.

Solución.

Definición de conjunto abierto: Un conjunto $S \subseteq V$ es abierto si para cada x que está en S existe un $\epsilon > 0$ tal que $B_\epsilon(x) \subseteq S$. El conjunto S cumple con la definición de ser abierto ya que nunca ningún x que está en S iguala a sus extremos, entonces habrá infinidades de bolas abiertas que estén dentro del intervalo. Para el caso de conjunto cerrado, no se cumple. Ahora para verificar que sea restringido, debe cumplirse lo siguiente

$$\{y \in V : 3 > y > 1\} \subseteq B_r(0) : \ r > 0$$

Por tanto, si es restringido.

De forma equivalente, un conjunto $S \subseteq V$ es cerrado si y solo si el límite de cualquier secuencia convergente de puntos (con respecto a $\|\cdot\|$) en S está a sí misma en S.

S está acotado si y solo si está contenido en alguna forma de radio finito. Un conjunto $S \subseteq V$ es compacto si es tanto cerrado como acotado. Para un conjunto dado $S \subseteq V$ y una función real evaluada f definida sobre S, $\inf_{x \in S} f(x)$ y $\sup_{x \in S} f(x)$ no necesitan ser finitos, y si es que lo son, podrían ser o no puntos $x_{\min}$ y $x_{\max}$ en S tal que $f(x_{\min}) = \inf_{x \in S} f(x)$ y $f(x_{\max}) = \sup_{x \in S} f(x)$, es decir, f no necesita alcanzar un máximo o mínimo valor sobre S. Sin embargo, bajo ciertas circunstancias, podríamos estar seguros de que f alcanza ambos máximos y mínimos sobre S.

Una función f real o compleja evaluada definida sobre un conjunto $S \subseteq V$ es continua en $x_0 \in S$ si, para cada $\epsilon > 0$, hay un $\delta > 0$ tal que $|f(x) - f(x_0)| < \epsilon$ siempre que $x \in S$ y $\|x - x_0\| < \delta$; f es continua sobre S si es continua en cada punto de S; f es uniformemente continua sobre S, si para $\epsilon > 0$, hay $\delta > 0$ tal que $|f(x) - f(y)| < \epsilon$ siempre que $x, y \in S$ y siempre que $\|x - y\| < \delta$.

Volviendo al ejemplo, para explicar que toda matriz real simétrica tiene al menos un eigenvalor, se debe cumplir que

$$Ax = \lambda x, \ \ con\ x \neq 0$$

La información que proporciona el teorema de Weierstrass es de que una función definida sobre un subconjunto compacto

en un espacio vectorial que en este caso son los números reales R^n con $f: S \to \boldsymbol{R}$, en el que $f(x) = Ax = \lambda x$ y con una norma, está acotada por

$$f(x_{\min}\ ) \leq f(x) \leq f(x_{\max}\ )$$

La idea sobre maximizar debe concretarse. Para el término $f(x_{\max}\ )$ se refiere a que la función alcanza su punto superior más alto debido a $x_{\max}$. Entonces si

$$f(x_{\max}\ ) = x_{\max}^T\ \ Ax_{\max}$$

se cumple que $x_{\max}^T\ \ Ax_{\max}\ \ \geq x^T Ax, \quad \forall x \in S$, siendo S un conjunto compacto, es decir, cerrado y acotado.

$$x_{\max}^T\ \ Ax_{\max}\ \ \geq x^T Ax$$

$$x_{\max}^T\ \ [A^1 x_{1-max} + \cdots + A^n x_{n-max}] \geq x^T [A^1 x_1 + \cdots A^n x_n]$$

$$x_{\max}^T\ \ Ax_{\max}\ \ - x^T Ax \geq 0$$

Con $x^T x = 1$, se puede simplificar a

$$Ax(x^T x) = Ax(1) = (Axx^T)x = Ax$$

$$(Axx^T)x - Ax = O, \quad x \neq O$$

$$Axx^T - A = O$$

En consecuencia, debido a que el tamaño de Ax es de $n \times 1$, la complicidad de la igualdad

$$(Axx^T)x = (A)x, \ \ x \neq O$$

Significa que debe existir algún escalar $\lambda \in F$ tal que

$$Ax = \lambda x$$

y por ende, toda matriz real simétrica tiene un eigen valor. Para mostrar que es un número real se debe tener en consideración que

$$Axx^T \in \boldsymbol{R},\ \ x \in \boldsymbol{R},\ \ x^T \in \boldsymbol{R}$$

Ejemplo. Suponga que $A \in M_n(\boldsymbol{R})$ es simétrica. Muestre que $\max\{x^T Ax : x \in \boldsymbol{R^n}, x^T x = 1\}$ tiene el eigenvalor más alto de A.

Solución.

Por el teorema de Weierstrass

$$f(x_{\max}\ \ ) \geq f(x)$$

Para la matriz A está asociado su correspondiente eigenvalor, por ende, al estar acotada corresponde a su valor más alto.

Operador

Una transformación lineal desde un espacio vectorial hacia sí mismo es llamado un Operador.

Subespacio Invariante

Suponga que $T \in L(V)$. Un subespacio U de V es llamado **invariante** bajo T si $Tu \in U$ para todo $u \in U$.

Por tanto, U es invariante bajo T si T_{U} es un operador sobre U.

Ejemplo. Subespacio invariante bajo el operador diferenciación

Suponga que $T \in L\big(P(R)\big)$ está definido por $Tp = p'$. Los polinomios definidos sobre $P_4(R)$ se pueden representar como

$$p_4(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

con $a_0, a_1, a_2, a_3, a_4 \in R$ debido a que P_4 está restringido a pertenecer a los números reales. Por ejemplo, si los coeficientes fueran $\{a_0, a_1, a_2, a_3, a_4\} = \{i, 1,1,1,1\}$ respectivamente, dicho conjunto está en el conjunto de los números complejos, sin embargo,

$$p(0) = i$$

lo que significa que para $x = 0 \in R$, $p(0)$ no pertenece a $P_4(R)$. En consecuencia, definir $P_4(R)$ restringe a que sus coeficientes estén en el conjunto de los números reales.

Por otra parte, $P_4(R) \subseteq P(R)$, por lo tanto $P_4(R)$ es invariante bajo T porque si $p \in P(R)$ tiene a lo mucho grado 4, su derivada p' tiene también a lo mucho grado 4. Por ejemplo,

$$(Tp_4)(x) = T\big(p_4(x)\big) = p_4'(x) = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3$$

Espacios Invariantes Especiales

Sea $T \in L(V)$, entonces los siguientes subespacios son invariantes bajo T.

- **Subespacio nulo**

El subespacio $\{0\}$ es invariante bajo T porque si $u \in \{0\}$, entonces $Tu = 0 \in \{0\}$.

- **El Subespacio V**

El subespacio V es invariante bajo T porque si $u \in V$ entonces $Tu \in V$.

- ***kernel T***

El subespacio nulo $null\ T$ es invariante bajo T porque si $u \in null\ T$, entonces $Tu = 0 \in null\ T$.

- ***Range T***

El subespacio $rango\ T$ también es invariante bajo T porque si $u \in rango\ T$, entonces $Tu \in rango\ T$. Hay que recordar que $Rango\ T = \{Tu : u \in V\}$.

Surge la siguiente pregunta, ¿Debe un operador $T \in L(V, V) = L(V)$ tener algún otro subespacio invariante diferente del $\{0\}$ y V?. Esta pregunta resulta ser afirmativa si V es dimensión finita y si $\dim V > 0$ (para $F = C$) o $\dim V > 2$ (para $F = R$).

El subespacio invariante más simple diferente del nulo de dimensión uno es uno de los más interesantes debido a que si

se toma cualquier $v \in V$ con $v \neq 0$ y sea U igual al conjunto de todos los múltiplos escalares de v, se puede definir U como:

$$U = \{\lambda v : \lambda \in F\} = span(v)$$

Debido que $v \neq O$, este es linealmente independiente y constituye una base para U. Ahora bien, si U es invariante bajo $T \in L(V)$, entonces $Tv \in U$, y en consecuencia hay un escalar $\lambda \in F$ tal que

$$Tv = \lambda v$$

Contrariamente, si $Tv = \lambda v$ para algún $\lambda \in F$, entonces $span(v)$ es un subespacio de V invariante bajo T de dimensión uno porque $Tv = \lambda v \in span(v)$ para todo $v \in U = span(v)$.

La ecuación $Tv = \lambda v$, la cual la acabamos de ver, inmediatamente está conectada con subespacios invariantes de dimensión uno y aquellos vectores v y escalares λ reciben nombres especiales.

Eigenvalor

Suponga que $T \in L(V)$. Un número $\lambda \in F$ es llamado **eigenvalor** de T si existe un vector $v \in V$ tal que $v \neq 0$ y $Tv = \lambda v$.

Condiciones Equivalentes

Suponga que V es de dimensión finita, $T \in L(V)$, y $\lambda \in F$. Entonces las siguientes sentencias son equivalentes.

λ es un eigenvalor de $T \equiv T - \lambda I$ no es inyectiva

Se puede traducir a que si $T - \lambda I$ no es inyectiva entonces λ es un eigenvalor de T. Para ello debe cumplirse que $kernell(T - \lambda I)v = \{v \in V: (T - \lambda I)v = O\} \neq \{O\}$, lo que ocurre que al ser $v \neq 0$, el valor de λ corresponde a un eigenvalor. Por otra parte, si λ es un eigenvalor entonces $T - \lambda I$ no es inyectiva, es decir,

$$kernel(T - \lambda I) = \{v \in V: (T - \lambda I)v = 0\}$$

Debido a que $\lambda \in F$ corresponde a un elemento de V tal que $v \neq 0$, entonces $kernel(T - \lambda I) \neq \{O\}$ lo que significa que $(T - \lambda I)$ no es inyectiva.

λ es un eigenvalor de $T \equiv T - \lambda I$ no es sobreyectiva

Se puede traducir a que si $T - \lambda I$ no es sobreyectiva, entonces λ es un eigenvalor de T. La definición de sobreyectiva bajo $T - \lambda I$ es de que $range\,(T - \lambda I) = V = \{(T - \lambda I)v: v \in V\}$. En consecuencia $\{(T - \lambda I)v: v \in V\} \neq V$ significa que al menos un vector en V no logra completar al espacio vectorial V y dicho vector es el nulo, lo que significa que si $(T - \lambda I)v = 0$ con $v \neq 0$, λ es un eigenvalor y por ende $(T - \lambda I)$ no es sobreyectiva. Por otro lado, si λ es un eigenvalor, $(T - \lambda I)$ no es sobreyectiva.

λ es un eigenvalor de $T \equiv T - \lambda I$ no es invertible

Se puede traducir a que si $T - \lambda I$ no es invertible, es decir no es sobreyectiva e inyectiva, λ es un eigenvalor lo cual se probó en líneas anteriores. Así mismo si λ es un eigenvalor la transformación $(T - \lambda I)$ no es invertible.

Eigenvector

Suponga que $T \in L(V)$, y $\lambda \in F$ es un eigenvalor de T. Un vector $v \in V$ es llamado eigenvector de T correspondiente a λ si $v \neq 0$ y $Tv = \lambda v$.

Eigenvectores Linealmente Independientes

Suponga que $T \in L(V)$. Entonces toda lista de eigenvectores de T correspondientes a distintos valores de T es linealmente independiente.

Demostración.

Suponga que el resultado deseado es falso. Entonces existe un entero positivo más pequeño m tal que existe una lista linealmente dependiente de eigenvectores $v_1, \dots, v_m$ de T correspondiente a diferentes eigenvalores $\lambda_1, \dots, \lambda_m$ de T (note que $m \geq 2$ porque un eigenvector por definición es distinto de cero). Por ende, existen $a_1, \dots, a_m \in F$, ninguno de los cuales son 0 (debido al minimización de m descrita), tal que

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0$$

Aplicando $T - \lambda_m I$ a ambos lados de la ecuación

$$\begin{aligned}(T - \lambda_m I)(a_1 v_1 + a_2 v_2 + \cdots + a_m v_m) \\ = (T - \lambda_m I)(a_1 v_1) + (T - \lambda_m I)(a_2 v_2) + \cdots \\ + (T - \lambda_m I)(a_m v_m)\end{aligned}$$

$$\begin{aligned}(T - \lambda_m I)(a_1 v_1) + (T - \lambda_m I)(a_2 v_2) + \cdots + (T - \lambda_m I)(a_m v_m) \\ = T(a_1 v_1) - \lambda_m I(a_1 v_1) + T(a_2 v_2) - \lambda_m I(a_2 v_2) \\ + \cdots + T(a_m v_m) - \lambda_m I(a_m v_m)\end{aligned}$$

$$\begin{aligned}T(a_1 v_1) - \lambda_m I(a_1 v_1) + T(a_2 v_2) - \lambda_m I(a_2 v_2) + \cdots + T(a_m v_m) \\ - \lambda_m I(a_m v_m) \\ = \lambda_1 a_1 v_1 - \lambda_m a_1 v_1 + \lambda_2 a_2 v_2 - \lambda_m a_2 v_2 + \cdots \\ + \lambda_m a_m v_m - \lambda_m a_m v_m\end{aligned}$$

$$\begin{aligned}\lambda_1 a_1 v_1 - \lambda_m a_1 v_1 + \lambda_2 a_2 v_2 - \lambda_m a_2 v_2 + \cdots + \lambda_m a_m v_m \\ - \lambda_m a_m v_m \\ = \lambda_1 a_1 v_1 - \lambda_m a_1 v_1 + \lambda_2 a_2 v_2 - \lambda_m a_2 v_2 + \cdots \\ + \lambda_{m-1} a_{m-1} v_{m-1} - \lambda_m a_{m-1} v_{m-1} = 0\end{aligned}$$

$$\begin{aligned}\lambda_1 a_1 v_1 - \lambda_m a_1 v_1 + \lambda_2 a_2 v_2 - \lambda_m a_2 v_2 + \cdots + \lambda_{m-1} a_{m-1} v_{m-1} \\ - \lambda_m a_{m-1} v_{m-1} \\ = a_1(\lambda_1 - \lambda_m) v_1 + \cdots + a_{m-1}(\lambda_{m-1} - \lambda_m) v_{m-1} \\ = (T - \lambda_m I)(0) = 0\end{aligned}$$

Debido a que los eigenvalores $\lambda_1, \ldots, \lambda_m$ son distintos, ninguno de los coeficientes de arriba son iguales a cero. Por ende, $v_1, \ldots, v_{m-1}$ es una lista linealmente dependiente de $m - 1$ eigenvectores de T correspondiente a distintos eigenvalores, contradiciendo a la minimización de m. Esta contradicción completa la prueba.

Un operador no puede tener más eigenvalores que la dimensión del espacio vectorial

Suponga que V es de dimensión finita. Entonces cada operador sobre V tiene a lo mucho $\dim V$ distintos eigenvalores.

Demostración.

Suponga que V es de dimensión finita. Suponga que $\lambda_1, \dots, \lambda_m$ son distintos eigenvalores de T. Sea $v_1, \dots, v_m$ correspondiente a eigenvectores. Entonces la sentencia anterior implica que la lista $v_1, \dots, v_m$ es linealmente independiente. Por ende, $m \leq \dim V$ según lo deseado.

Polinomios Aplicados a Operadores

La principal razón por la cual una teoría más enriquecedora para operadores existe, considerando que un operador es un mapa que transforma hacia sí mismo, en términos más generales, es que los operadores pueden ser elevados a potencias. Este concepto será clave para probar después que todo operador sobre un espacio vectorial diferente de cero complejos de dimensión finita tiene un eigenvalor.

Notación: T^m

Suponga que $T \in L(V,V)$ y además que m es un entero positivo.

- $T^m \in L(V)$ está definido por $T^m = T \cdots T$, es decir, multiplicar T m veces.
- T^0 está definido como el operador identidad I sobre V
- Si T es invertible con inversa T^{-1}, entonces $T^{-m} \in L(V)$ está definido

por $T^{-m} = (T^{-1})^m$.

Se podría verificar que si T es un operador, entonces

$T^m T^n = T^{m+n}$ y $(T^m)^n = T^{mn}$

Por ejemplo,

$$T^m = T \cdots T = T^{r+s} = T^r \cdot T^s, \quad m = r + s$$

Por otra parte,

$$T^r = T \cdots T = S$$

$$S^m = S \cdots S = T^r \cdots T^r = (T^r)^m = T^{rm}$$

Notación: $p(T)$

Suponga que $T \in L(V,V)$ y que $p \in P(F)$ es un polinomio dado por

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

para todo $z \in F$. Entonces $p(T)$ es un operador sobre V definido por

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

Ejemplo. Suponga que $D \in (P(R))$ es el operador diferenciación definido por $Dq = q'$ y que p es el polinomio definido por $p(x) = 7 - 3x + 5x^2$. Escriba la aplicación del operador D sobre $p(x)$.

Solución.

$$p(D) = 7I - 3D + 5D^2$$

Ahora aplicando sobre q se tiene

$$\big(p(D)\big)q = (7I - 3D + 5D^2)q = (7I)q + (-3D)q + (5D^2)q$$

$$= 7(D^0 q) - 3(Dq) + 5(D^2 q) = 7q - 3q' + 5\big(D(Dq)\big)$$

$$= 7q - 3q' + 5(Dq')$$

$$= 7q - 3q' + 5(Dq') = 7q - 3q' + 5q''$$

para todo $q \in P(R)$.

Es importante recalcar que si se fija un operador $T \in L(V,V)$, entonces la función de $P(F)$ a $L(V)$ dada por $p \longmapsto p(T)$ es

lineal. Por ejemplo, sea la función $S: P(F) \longrightarrow L(V)$, entonces para todo $r, q \in P(F)$:

Sean los polinomios

$r(z) = a_0 z^0 + a_1 z^1 + a_2 z^2 + \cdots + a_m z^m, \quad q(z) = b_0 z^0 + b_1 z + b_2 z^2 + \cdots + b_m z^m$

Entonces,

$$S(r+q) = S(a_0 z^0 + a_1 z + a_2 z^2 + \cdots + a_m z^m + b_0 z^0 + b_1 z + b_2 z^2 + \cdots + b_m z^m)$$

$$\begin{aligned} S[(a_0 + b_0)z^0 + (a_1 + b_1)z^1 + (a_2 + b_2)z^2 + \cdots \\ + (a_m + b_m)z^m] \\ = (a_0 + b_0)I + (a_1 + b_1)T + (a_2 + b_2)T^2 + \cdots \\ + (a_m + b_m)T^m \end{aligned}$$

Debido a que $I, T, T^2, \ldots, T^m$ es un operador sobre V,

$$\begin{aligned} (a_0 + b_0)I + (a_1 + b_1)T + (a_2 + b_2)T^2 + \cdots + (a_m + b_m)T^m \\ = a_0 I + b_0 I + a_1 T + b_1 T + a_2 T^2 + b_2 T^2 + \cdots \\ + a_m T^m + b_m T^m \end{aligned}$$

$$\begin{aligned} = (a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m) \\ + (b_0 I + b_1 T + b_2 T^2 + \cdots + b_m T^m) \\ = S(r) + S(q) \end{aligned}$$

En consecuencia, se cumple para la adición la función de $S: P(F) \longrightarrow L(V)$. Por otra parte,

$$Con\ \alpha \in F, \qquad S(\alpha q) = S\big(\alpha(b_0 z^0 + b_1 z + b_2 z^2 + \cdots + b_m z^m)\big)$$

$$S\big(\alpha(b_0z^0 + b_1z + b_2z^2 + \cdots + b_mz^m)\big) = S[(\alpha b_0z^0) + (\alpha b_1z) + (\alpha b_2z^2) + \cdots + (\alpha b_mz^m)]$$

$$S[(\alpha b_0z^0) + (\alpha b_1z) + (\alpha b_2z^2) + \cdots + (\alpha b_mz^m)] = \alpha b_0I + \alpha b_1T + \alpha b_2T^2 + \cdots + \alpha b_mT^m$$

$$\alpha b_0I + \alpha b_1T + \alpha b_2T^2 + \cdots + \alpha b_mT^m = \alpha(b_0I + b_1T + b_2T^2 + \cdots + b_mT^m) = \alpha S(q)$$

Finalmente, la función $S: p \longmapsto p(T)$ es lineal.

Producto de Polinomios

Si $p, q \in P(F)$, entonces $pq \in P(F)$ es el polinomio definido por

$$(pq)(z) = p(z)q(z)$$

para todo $z \in F$.

Propiedades Multiplicativas

Suponga que $p, q \in P(F)$ y $T \in L(V)$. Entonces

- $(pq)(T) = p(T)q(T)$
- $p(T)q(T) = q(T)p(T)$

Demostración.

Suponga que $p(z) = \sum_{j=0}^{m} a_jz^j$ y que $q(z) = \sum_{k=0}^{n} b_kz^k$ para todo $z \in F$. Entonces

$$(pq)(z) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k z^{j+k}$$

Por ende,

$$(pq)(T) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k}$$

$$= \left(\sum_{j=0}^{m} a_j T^j \right) \left(\sum_{k=0}^{n} b_k T^k \right)$$

$$(pq)(T) = p(T)q(T)$$

Usando el análisis anterior dos veces, tenemos en primer lugar $p(T)q(T) = (pq)(T)$, y en segundo, $(qp)(T) = q(T)p(T)$. En consecuencia,

$$\left(\sum_{j=0}^{m} a_j T^j \right) \left(\sum_{k=0}^{n} b_k T^k \right) = p(T)q(T) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k}$$

$$\left(\sum_{k=0}^{n} b_k T^k \right) \left(\sum_{j=0}^{m} a_j T^j \right) = q(T)p(T) = \sum_{k=0}^{n} \sum_{j=0}^{m} b_k a_j T^{k+j}$$

$$= \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k}$$

$$p(T)q(T) = q(T)p(T)$$

Espacio Nulo y Rango de $p(T)$ son invariantes bajo T

Suponga que $T \in L(V)$ y $p \in P(F)$. Entonces $null\, p(T)$ y $rango\ p(T)$ son invariantes bajo T.

Demostración.

Suponga que $u \in null\, p(T)$. Recordando la notación de $p(T)$,

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$$

entonces es posible reescribir

$$T\big(p(T)\big) = T(a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m)$$

$$T(a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m) \\ = T(a_0 I) + T(a_1 T) + T(a_2 T^2) + \cdots + T(a_m T^m)$$

$$T(a_0 I) + T(a_1 T) + T(a_2 T^2) + \cdots + T(a_m T^m) \\ = (a_0 I)T + (a_1 T)T + \cdots + (a_m T^m)T = p(T)T$$

En consecuencia,

$$p(T)T(u) = Tp(T)(u) = T(0) = 0$$

Por ende $Tu \in null\, p(T)$. En consecuencia, $null\, p(T)$ es invariante bajo T porque $p(T)(Tu) \in null\, p(T)$ para todo $Tu \in null\, p(T)$. Con esta demostración, se puede representar a $null\, p(T)$ como

$$null\, p(T) = \{Tu \in V : p(T)(Tu) = 0\}$$

Por ejemplo, sea el polinomio $p(x) = x^2 + x - 2,\ \forall x \in R$. Defínase el operador diferenciación como $Dq = q'$. Entonces $p(D)$ se puede representar como

$$p(D) = D^2 + D - 2I$$

Aplicando el operador $p(D)$ sobre $P(F)$ con $q \in P(F)$

$$p(D)q = (D^2 + D - 2I)q = D^2q + Dq - 2Iq = q'' + q' - 2q$$

Sería interesante definir $null\ p(D)$. Entonces

$$null\ p(D) = \{Du \in P(R): p(D)(Du) = 0\}$$

Si $u \in P(R)$ tal que $u(x) = x$, entonces $Du = 1$. Con ello se puede escribir

$$\begin{aligned} p(D)(Du) = p(D)(1) &= (D^2 + D - 2I)(1) \\ &= D^2(1) + D(1) - 2I(1) = 0 + 0 - 0 = 0 \end{aligned}$$

entonces $Du \in P(R)$ está en $null\ p(D)$.

Suponga ahora que $u \in rango\ p(T)$. Entonces existe un elemento $v \in V$ tal que $u = p(T)v$. Por ende,

$$Tu = T(p(T)v) = p(T)(Tv)$$

En consecuencia, $Tu \in rango\ p(T)$. Por ende, el $rango\ p(T)$ es invariante bajo T, como lo era deseado.

Ejercicios de Aplicación

1. **Suponga que $T \in L(V)$ y U es un subespacio de V.**

- **Pruebe que si $U \subseteq null\, T$, entonces U es invariante bajo T.**
- **Pruebe que si $range\, T \subseteq U$, entonces U es invariante bajo T**

Desarrollo.

En primer lugar, si U es un subespacio de V, se dice que si $Tu \in U,\ \forall u \in U$ entonces U es invariante bajo T. La condición del problema es que dice que $U \subseteq null\, T$ entonces, el conjunto de elementos $\{e_i\},\ \forall i \in N^+$ que están en U están en el espacio nulo de T. En consecuencia, $Te_i = 0 \in null\, T, \forall i \in N^+$ y debido a que $null\, T$ es invariante bajo T, por ende, se concluye que $Te_i \in U, \forall e_i \in U$ lo que significa que es invariante U bajo T.

Por otra parte, $range\, T$ es invariante bajo T si $Tv \in range\, T, \forall v \in range\, T$, en consecuencia $range\, T$ subespacio de V es invariante bajo T, sin embargo, definiendo nuevamente $U = \{e_i,\ \forall i \in N^+\}$, se obtiene que

$$range\, T = \{Tv : v \in V\} \subseteq \{Te_i, \qquad \forall i \in N^+, e_i \in U\} \in U \subseteq V$$

Por ende, U es invariante bajo T.

2. Suponga que $T \in L(V)$. Pruebe que la intersección de cada colección de subespacios de V invariantes bajo T es invariante bajo T.

Desarrollo.

Suponga que U, Q son subespacios invariantes bajo T. En consecuencia, $\{Tu \in U, \forall u \in U\}$ y además $\{Tq \in Q, \forall q \in Q\}$. Si se realiza la intersección de estos subespacios, se puede generar un nuevo subespacio

$$S = \{Tu \in U, \forall u \in U\} \cap \{Tq \in Q, \forall q \in Q\}$$

Es decir que S es el conjunto de elementos $Tu \vee Tq, \ \forall u \in U \vee \forall q \in Q$. Por ende, cualquiera de las dos opciones puede ser válida y en consecuencia produce un subespacio invariante bajo T. Lo más curioso es que puede haber más subespacios que se intersecan entre sí generando nuevos subespacios de V invariantes bajo T.

3. Pruebe o de un contra ejemplo: Si V es de dimensión finita y U es un subespacio de V que es invariante bajo todo operador sobre V, entonces $U = \{O\}$ o $U = V$.

Desarrollo.

Si V es de dimensión finita entonces existen elementos $\{v_1, v_2, \dots, v_n\}$ que conforman una base para V. Por otra parte, sea U un subespacio de V y $T \in L(V,V)$. Si se define que

$$S = \{Tu \in U, \qquad \forall u \in U\}$$

entonces el conjunto S representa que U es invariante bajo T. ¿Significa que $U = \{O\}$ o V?

Si $U = O,\ S = \{T(O) = O \in U,\ O \in U\}$ de igual manera la representación de S indica que U es invariante bajo todo operador T sobre V.

Si $U = V, S = \{T(v) \in V,\ \forall v \in V\}$ de igual manera la representación de S indica que U es invariante bajo todo operador T sobre V.

Sin embargo, defínase como $V = \boldsymbol{R}$ y por otra parte a la transformación lineal

$$T: \boldsymbol{R} \to \boldsymbol{R},\ T(x) = \alpha x, \qquad \forall x \in \boldsymbol{R}\ , \boldsymbol{\alpha} \in \boldsymbol{R}$$

Ahora sea $U = \{n,\ \forall n \in \boldsymbol{N}\}$ el cual es un subespacio de V pero no es invariante bajo todo operador sobre V es decir que $T(n) \in U,\ \forall n \in U$ no se cumple, por ejemplo, $\alpha = 0.1$, $T(x) = 0.1x$ y si se escoge cualquier $n \in U$ resulta que Tn no pertenece a U. Con este pequeño ejemplo se puede garantizar que para que un subespacio sea invariante bajo cualquier operador, este tiene que ser el subespacio trivial o todo el espacio completo.

4. Suponga que $T \in L(R^2)$ está definido por $T(x, y) = (-3y, x)$. Encuentre los eigenvalores de T.

Desarrollo.

Los eigenvalores de T corresponden a aquellos $\lambda \in F$ tales que $T(v) = \lambda v, \ \ con\ v \neq (0,0)$. En consecuencia, si existen valores tales que

$$\lambda(x, y) = (-3y, x)$$

con $\lambda \in F$ entonces se denomina a λ eigenvalor de T.

$$(\lambda x, \lambda y) = (-3y, x)$$

se debe cumplir entonces que

$$\lambda x = -3y$$

$$\lambda y = x, \ \ \lambda^2 y = (\lambda x), \ \ \lambda \in \boldsymbol{R}$$

$$\lambda^2 y = -3y$$

$$\frac{y}{y} = 1, \ \ y \neq 0$$

$$\lambda^2 = -3$$

$$\lambda = \pm\sqrt{3}i$$

A pesar de que el ejercicio ya está completo para el enunciado, causa interés saber cuales son aquellos elementos que están en R^2 de tal manera que sean eigenvectores.

Sea

$$a = \pm\sqrt{3}i$$

entonces

$$a(x, y) = (ax, ay) = (-3y, x)$$

Con ello se obtiene que $ax = -3y$ y $ay = x$. Se puede reescribir

$$ay + ax = x - 3y$$

$$ay = x - ax - 3y = x(a - 1) - 3y$$

$$a = \frac{x}{y}(a - 1) - 3, \quad x, y \neq 0$$

$$\frac{a + 3}{a - 1} = \frac{x}{y}$$

Entonces los vectores de la forma

$$(x, y) = \left(\frac{a + 3}{a - 1}y, y\right), \ con\ (x, y) \in \boldsymbol{R^2}$$

son todos los eigenvectores de T correspondientes a los eigenvalores $\lambda = \pm\sqrt{3}i$.

5. Defínase $\boldsymbol{T \in L(F^2)}$ por $\boldsymbol{T(w, z) = (z, w)}$. Encuentre todos los eigenvalores y eigenvectores de $\boldsymbol{T}$.

Desarrollo.

Encontrar todos los eigenvalores de T conlleva utilizar todos los elementos $(w, z) \neq 0$ tales que

$$T(w,z) = (z,w) = \lambda(w,z)$$

$$(z,w) = (\lambda w, \lambda z)$$

$$z = \lambda w, \qquad w = \lambda z, \;\; \lambda w = \lambda(\lambda z)$$

$$z = \lambda^2 z, \;\; \lambda^2 = 1, \;\; \lambda = \pm 1$$

Entonces aquellos eigenvalores que corresponden a T son $+1$ y -1.

Para encontrar todos los eigenvectores correspondientes a T asociados a su eigenvalor se debe encontrar todas las posibles representaciones de (z,w) que se presenten por

$$(z,w) = (\lambda w, \lambda z)$$

$$z = \lambda w, \qquad w = \lambda z$$

$$z + w = \lambda w + \lambda z$$

$$z - \lambda z + w = \lambda w$$

$$z(1-\lambda) = w(\lambda - 1)$$

$$\frac{z}{w} = \frac{\lambda - 1}{1 - \lambda} = -\frac{\lambda - 1}{\lambda - 1} = -1, \;\; z = -w$$

En consecuencia, la representación de los eigenvectores de T asociados a λ tienen la forma $(z,w) = (-w,w)$.

6. Suponga que $P \in L(V)$ es tal que $P^2 = P$. Pruebe que si λ es un eigenvalor de P entonces $\lambda = 0$ o $\lambda = 1$.

Desarrollo.

Si λ es un eigenvalor de T entonces existe un $v \in V$ tal que $P(v) = \lambda v$. Por otra parte, también se tiene que $P\big(P(v)\big) = \lambda v$.

$$P\big(P(v)\big) = P(\lambda v) = \lambda P(v) = \lambda v$$

$$\lambda(\lambda v) = \lambda v$$

Se tiene que con $\lambda \in F$,

$$\lambda^2 v = \lambda v.$$

Debido a que $v \neq O$, para que se cumpla la igualdad o bien $\lambda = 0$ o $\lambda = 1$.

Comprobando,

$$\lambda = 0, \;\; (0)^2 v = (0)v = 0 = 0$$

$$\lambda = 1, \;\; (1)^2 v = (1)v = v = v$$

Ejercicio de Alta Complejidad

7. Suponga que $v_1, v_2, \dots, v_n$ es una base de V y $T \in L(V)$. Pruebe que si λ es un eigenvalor de T, entonces

$$|\lambda| \leq n \max\{|M(T)_{j,k}| : 1 \leq j, k \leq n\}$$

donde $M(T)_{j,k}$ denota la entrada en la fila j, columna k de la matriz de T con respecto a las bases $v_1, \dots, v_n$.

Desarrollo.

Se intentará demostrarlo por inducción. Entonces para $n = 1$.

$$v_1 \in V \; con \; T \in L(V)$$

Si λ es un eigenvalor, entonces $T(v) = \lambda v$. Para determinar la matriz de la transformación lineal, se deben expresar a cada elemento de la base de V como combinación lineal de los elementos del conjunto de llegada de la base del espacio vectorial.

$$T(v_1) = a_{11} v_1$$

entonces $M(T) = (a_{11})_{1x1}$. Incluso se puede observar que $\lambda = a_{11}$ corresponde a un eigenvector de v_1.

$$|a_{11}| \leq n \max\{|M(T)_{j,k}| : 1 \leq j, k \leq n\} = n \max\{|a_{11}|\}$$
$$= 1(|a_{11}|) = |a_{11}|$$

El método de inducción conlleva a demostrar que si es válido para $i = 1$ y para $i = n$, entonces es válido para $i = n + 1$. Se ha demostrado que es válido para $i = 1$ y se presume que es válido para $i = n$. Entonces

Sea $v_1, v_2, \ldots v_n, v_{n+1}$ una base de V con $T \in L(V)$.

$$T(v_1) = a_{11}v_1 + a_{21}v_2 + \cdots + a_{n+1,1}v_{n+1}$$

$$T(v_2) = a_{12}v_1 + a_{22}v_2 + \cdots + a_{n+1,2}v_{n+1}$$

$$\ldots$$

$$T(v_{n+1}) = a_{1,n+1}v_1 + a_{2,n+1}v_2 + \cdots + a_{n+1,n+1}v_{n+1}$$

La matriz de T está definida entonces por

$$M(T) = \begin{bmatrix} a_{11} & \cdots & a_{1,n+1} \\ \vdots & \ddots & \vdots \\ a_{n+1,1} & \cdots & a_{n+1,n+1} \end{bmatrix}, \quad a_{ij} \in F, i, j = 1, \ldots, n+1$$

Sea λ un eigenvalor de T. Entonces existe un elemento $v \neq 0$ tal que

$$\begin{aligned} T(v) = \lambda v &= \lambda(c_1v_1 + \cdots + v_{n+1}v_{n+1}) \\ &= T(c_1v_1 + \cdots + c_{n+1}v_{n+1}), \quad c_1, \ldots, c_{n+1} \in F \end{aligned}$$

$$\begin{aligned} \lambda(c_1v_1 + \cdots + v_{n+1}v_{n+1}) &= T(c_1v_1 + \cdots + c_{n+1}v_{n+1}) \\ &= c_1T(v_1) + \cdots + c_{n+1}T(v_{n+1}) \end{aligned}$$

$$\begin{aligned} \lambda(c_1v_1 + \cdots + v_{n+1}v_{n+1}) & \\ &= c_1\left(a_{11}v_1 + a_{21}v_2 + \cdots + a_{n+1,1}v_{n+1}\right) + \cdots \\ &+ c_{n+1}\left(a_{1,n+1}v_1 + a_{2,n+1}v_2 + \cdots \right. \\ &\left. + a_{n+1,n+1}v_{n+1}\right) \end{aligned}$$

$$\lambda(c_1v_1 + \cdots + v_{n+1}v_{n+1})$$
$$= \left[c_1a_{11}v_1 + \cdots + c_{n+1}a_{1,n+1}v_1\right] + \cdots$$
$$+ \left[c_1a_{n+1,1}v_{n+1} + \cdots + c_{n+1}a_{n+1,n+1}v_{n+1}\right]$$

Los vectores $v_1, \ldots, v_{n+1}$ son linealmente independientes, entonces

$$\lambda c_1v_1 = c_1a_{11}v_1 + \cdots + c_{n+1}a_{1,n+1}v_1$$
$$= v_1(c_1a_{11} + \cdots + c_{n+1}a_{1,n+1})$$

En consecuencia,

$$\lambda c_1 = c_1a_{11} + \cdots + c_{n+1}a_{1,n+1}$$

Enfocando la definición de que F pueden ser los números reales o complejos, se puede generalizar de que $F = C$ y en caso de que sean los reales, simplemente su parte imaginaria es igual a cero, entonces

$$|\lambda| \leq |\lambda c_1| = |\lambda||c_1| = \left|c_1a_{11} + \cdots + c_{n+1}a_{1,n+1}\right|$$
$$\leq |c_1a_{11}| + \cdots + \left|c_{n+1}a_{1,n+1}\right|$$

Debido a que es válido para $i = n$

$$|\lambda| \leq \frac{|c_1||a_{11}|}{|c_1|} + \cdots + \frac{|c_{n+1}|\left|a_{1,n+1}\right|}{|c_1|}$$
$$\leq |a_{11}| + \cdots + \frac{|c_{n+1}|\left|a_{1,n+1}\right|}{|c_1|}$$
$$\leq (n+1)\max\{|a_{11}|, \ldots, \left|a_{1,n+1}\right|\}$$

Bibliografía

Analysis, C. (1987). Complex Analysis (W. Rudin (ed.); Third). McGraw Hill.

Angarone, R., & Soskin, D. (2025). Generalized diagonals in positive semi-definite matrices. 130, 1–8. https://doi.org/https://doi.org/10.1016/j.ejc.2025.104220

Axler, S. (2024). Sheldon Axler About the Author (S. Axler (ed.); Fourth, Issue January). https://linear.axler.net

Axler, S., & Edition, T. (2015). Linear Algebra Done Right (S. Axler (ed.); Third). https://doi.org/10.1007/978-3-319-11080-6

Cherney, D., Denton, T., Thomas, R., & Waldron, A. (2014). Linear Algebra (D. Cherney, T. Denton, R. Thomas, & A. Waldron (eds.); First). Creative Commons Attribution-NonCommercial- ShareAlike 3.0 Unported License.

Hefferon, J. (2021). Linear algebra (J. Hefferon (ed.); Fourth). http://joshua.smcvt.edu/linearalgebra

Sobre los autores:

William Alexander Ávila Aguilar, Universidad Técnica de Machala

Docente Ocasional a tiempo completo de la Universidad Técnica de Machala en la Facultad de Ciencias Químicas y de la salud especialización en ingeniería en matemática y computación aplicada en Ciencias de la Ingeniería Mecánica con mención en Sistemas Mecánicos

Gladis Del Rocio Mora Veintimilla, Universidad Técnica de Machala

Licenciada en Enfermería. Magister en Enfermería Clínico-Quirúrgica. Magister en Gerencia y Administración de Salud. Docente Titular de la Carrera de Enfermería de la UTMACH

Eduardo Antonio Vázquez Aldaco, CUT LAGUNA UNIVERSIDAD

Docente Titular "B" de Tiempo Completo con más de 20 años de servicio en el CBTis 97 (DGETI). Ingeniero Electrónico egresado del Instituto Tecnológico de Saltillo y Maestro en Educación (CUT Laguna). Cuenta con publicación académica en la revista Ciencia Latina y actualmente se desempeña como

secretario Local de la Academia de Cultura Digital en su institución.

Lizania Godina Young, CUT LAGUNA UNIVERSIDAD

Docente de Tiempo Completo con más de 15 años de servicio en el CBTis 97, perteneciente a la Dirección General de Educación Tecnológica Industrial (DGETI). Contadora Pública con Doctorado en Educación, se desempeña en el área Económico-Administrativa. Cuenta con publicación académica en la revista Ciencia Latina y actualmente funge como secretaria de Academia Local en su institución.

Lethzaira Mercedes Contreras Borjon, CUT LAGUNA UNIVERSIDAD

Química Farmacobióloga con más de 15 años de servicio en el CBTis 97, perteneciente a la Dirección General de Educación Tecnológica Industrial (DGETI). Cuenta con experiencia en el área de análisis clínicos y una sólida trayectoria en la docencia de ciencias experimentales en Educación Media Superior. Actualmente se desempeña como docente y jefa del Departamento de Servicios Docentes en el turno vespertino.

Hernán Alonso Mancipe Bohórquez, Fundación Universitaria internacional de La Rioja - Colombia

Docente de modalidad presencial y virtual con énfasis en el área de Ciencias Básicas e Industria 4.0

Marcos Chacón Castro, Fundación Universitaria internacional de La Rioja - Colombia

Licenciado en Matemática, Magíster en Educación con mención en Investigación y Doctor en Educación. Docente e investigador con más de ocho años de experiencia en instituciones de educación superior a nivel regional, nacional e internacional. Director y fundador del Congreso Internacional de STEAM y Metodologías Activas. Embajador de Genially en Latinoamérica, integrante del grupo de docentes en línea de Latinoamérica y representante de la Red de Círculos y Festivales Matemáticos (Cyfemat). Especializado en la dirección y asesoramiento de trabajos de investigación educativa, con énfasis en la elaboración y publicación de artículos científicos de alto impacto en revistas académicas indexadas. Su práctica docente integra herramientas tecnológicas para la divulgación científica en entornos virtuales y presenciales, promoviendo el trabajo colaborativo y la investigación aplicada en el campo de la educación matemática.

Esta obra se publicó en primera edición durante el año 2026.

Colección: Educación

DOI: 10.64325/1fjan565

ISBN: 978-9942-7491-2-3



Esta obra se edita bajo una Licencia Creative Commons BY.4.0