



Universidad del Bío-Bío

Facultad de Educación y Humanidades

ESCUELA DE PEDAGOGÍA EN EDUCACIÓN MATEMÁTICA

Álgebra Lineal con Aplicaciones

Autores: JESSICA CARINA GATICA CUEVAS
CRISTIAN MANUEL MARDONES SEPÚLVEDA
Orientador: Prof. Ivo Basso Basso

SEMINARIO PARA OPTAR AL TÍTULO DE PROFESOR DE EDUCACIÓN MEDIA EN
EDUCACIÓN MATEMÁTICA

CHILLÁN, 2007.

Agradecimientos

A mi familia por el apoyo brindado desde que ingresé a la carrera, por ayudarme a levantarme en los momentos que caí, y es por su apoyo incondicional que les dedico este seminario. A Paula por ser mi amiga durante estos cinco años, por reunirnos todos los domingos a comparar las guías de estudio, por su apoyo moral, muchas gracias. A Cristian mi pololo y compañero de tesis, por el logro de haber terminado nuestra Actividad de Titulación y otros aspectos de nuestra vida sentimental.

También he de agradecer a nuestros formadores, en el Área de Pedagogía en especial a la Profesora Paula Solo, por el apoyo brindado y por su formación como mujer; pero por sobre todo a los docentes de la Especialidad muchas gracias por la formación entregada y entre ellos al Profesor Wilfred por habernos entregado una parte de sus conocimientos, abrirnos de alguna forma la mente, mostrarnos nuevos horizontes y junto a la Profesora Yuri Haragushi enseñarnos la importancia de la Geometría en el diario vivir; en particular a ella, muchos compañeros le agradecemos por su calidad como persona y personalmente digo gracias por las oportunidades y por creer siempre en nosotros.

Al Profesor Roberto Cabrales que a pesar de haberlo conocido cuando estábamos terminando nuestros cursos de especialidad, nos orientó como un segundo profesor guía, nos facilitó textos para la creación del seminario y ayudó en reiteradas ocasiones a resolver algunos problemas con la edición del presente texto. Disculpe todas las molestias y muchas gracias por su buena voluntad.

Finalmente me resta agradecer a nuestro Profesor Guía por habernos aceptado como sus alumnos tesis, porque tuvimos la posibilidad de compartir fuera de las aulas, en mi caso por ser en varios cursos su Alumna Ayudante y porque fue él quien nos empujó a formar desde el segundo año de Universidad y que nos exigió un poco de rigor matemático para escribir y expresarnos y quien como nuestro profesor de Álgebra Lineal nos entregó una mirada diferente de ésta. Desde ya muchas gracias.

Más que un agradecimiento quiero mencionar un saludo a Leslie y sus secuaces de Álgebra Lineal 2007 y decirles si tienen interés en Lineal tendrán un texto en el cual verán sus aplicaciones.

Me resta agradecer al Sr. Ernesto Figueroa M. porque a pesar de no ser profesor nuestro, nos dio la idea del tema que motiva nuestra tesis y por facilitarnos el primer libro para empezar a trabajar en ella.

Jessy

A los primeros que quiero expresar mi más sincera admiración y más profundo agradecimiento es a los profesores, señor Ivo Basso Basso quien durante los 5 años de carrera de Pedagogía en Educación Matemática, nos apoyó en nuestra formación, no sólo como docente, sino también en la ampliación de nuestros horizontes, conocimientos y dándonos la posibilidad de asistir a Congresos; gracias a sus avisos y contactos, tener la enorme disponibilidad de ser nuestro director de tesis y al considerarnos en todos los años de carrera como alumnos regulares, además de su trato diferente con todos mis compañeros en general, su apoyo y calidad.

También pero no menos importante, al señor Roberto Carlos Cabrales, a quien hace dos años, desde su llegada nos ha apoyado con su nueva visión de las matemáticas, su buena disponibilidad y el apoyo fundamental en la elaboración de este texto al proporcionarnos las herramientas para su escrito. Además de la paciencia que nos tuvo, al ir a preguntarle su visión de nuestra tesis, a pesar de no ser nuestro director de Tesis, fue gracias a su calidad humana que se aprendió a manejar software necesarios para la formación de un profesor de Matemáticas y más adelante si Dios quiere de alumno de Magister. Por ello, desde ya, se agradece la iniciativa a la Universidad del Bío-Bío y al profesorado que realizará la escuela de verano del 2008, para los alumnos que quierán nivelarse, este proceso será de gran ayuda para los que queremos seguir estudios de postgrado en el futuro.

Quisiera también agradecer a aquellas personas (Profesores y alumnos en especial) que durante los años nos ayudaron y contribuyeron de alguna u otra forma en la realización de este trabajo, gracias a los conocimientos adquiridos, en especial a la señorita Yuri Haraguchi y al señor Wilfred Reyes Scantlebury quienes nos acompañaron gran parte de la carrera, nos apoyaron, destacamos su gran calidad humana.

Por último al señor Luis Frix por su apoyo en la rama de análisis y buena disposición.

Por parte de mis compañeros sólo quiero desearles en sus nuevos proyectos los más buenos deseos, en especial a Paula Verdugo, Horacio Hormaxábal y Pablo Molina con quienes más compartí durante la carrera.

Además quiero agradecer a nuestras familias en especial a nuestros Padres quienes durante los años de escolaridad nos han apoyado en todos los proyectos que hemos emprendido, dándonos la educación Universitaria, a partir de esfuerzo, sudor y lágrimas, deseándonos lo mejor en lo profesional y en nuestro futuro, para ellos los cariños más efusivos, y sólo puedo decir que los quiero en lo más profundo. Además quisiera agradecer a la familia de mi compañera por acogerme unos días para el avance de este seminario.

Por parte de Cristian Manuel Mardones Sepúlveda, quiero agradecer a su compañera tesisista, quien gracias a su apoyo y paciencia han realizado este seminario, no sólo se le agradece como compañera sino como su polela, sin ella, la vida universitaria habría sido monótona, casi vacía, gracias a ella conocí el verdadero amor sin precedentes y lo que es la lealtad de una compañera por completo. A la señorita Jessica Gatica Cuevas le deseo lo mejor en su futuro académico y laboral, espero que ese futuro lo comparta a mi lado, porque en realidad es el amor de mi vida. Pd: Cielo Te Amo.

Cristian

Resumen

El Álgebra Lineal permite y también impulsa a combinar de modo muy satisfactorio dos elementos de la matemática: la abstracción y la aplicación, es por esto que el tema del seminario es *Álgebra Lineal con Aplicaciones*.

En el presente texto se pretende mostrar algunos tópicos de Álgebra Lineal, siguiendo un nivel creciente de complejidad, los que se han organizado por capítulos de tal modo que al final de cada capítulo hay una sección de aplicaciones. Los tópicos se han diferenciado en seis capítulos: Sistemas de Ecuaciones, Matrices y sus propiedades, Espacios Vectoriales, Espacio con Producto Interno, Valores y Vectores Propios y Programación Lineal. Todos comienzan con una breve introducción al tema, luego se definen algunos conceptos, en algunos casos se enuncian propiedades o proposiciones y se citan teoremas los cuales en su mayoría fueron demostrados.

Las aplicaciones son de distinta naturaleza; en el primer capítulo hay una aplicación en Circuitos Eléctricos y en Flujo de Tráfico, en el segundo capítulo las aplicaciones son en Criptografía, Seriación en Arqueología, Sociología y Cadenas de Markov, en el tercer capítulo se aproximarán curvas por medio de polinomios, en el cuarto capítulo se aproximarán funciones por medio de polinomios y polinomios trigonométricos, en el quinto capítulo las aplicaciones son en Formas Cuadráticas, un ejemplo de Crecimiento Poblacional, una aplicación a la Meteorología y Sistemas Oscilantes; finalmente en el sexto capítulo se muestra un ejemplo de Programación Lineal a la Nutrición.

Abstract

Linear Algebra allows and also it impels to combine of very satisfactory way two elements of the mathematical one: the abstraction and the application, are by which the subject of the seminary is *Álgebra Linear with Applications*.

In the present text it is tried to show some topics of Linear Algebra, following a level increasing of complexity, those that have been organizado by chapters in such a way that at the end of each chapter there is a section of applications. The topics have been different in six chapters: Vectorial systems of Equations, Matrices and their properties, Spaces, Own Space with Internal Product, Values and Vectors and Linear programming. All begin with a brief introduction to the subject, soon define some concepts, in some cases properties are enunciated or proposals and mention theorems which in their majority were demonstrated.

The applications are of different nature; in the first chapter there is an application in Electrical Circuits and in Flow of Traffic, in the second chapter the applications are in Cryptography, Seriación in Archaeology, Sociology and Chains of Markov, in the third chapter will come near curves by means of polynomials, in the fourth chapter will come near functions by means of trigonometrical polynomials and polynomials, in the fifth chapter the applications are in Quadratic Forms, an example of Population Growth, an Oscillating application to Meteorology and Systems; finally in the sixth chapter an example of Linear programming is the Nutrition.

La filosofía está escrita en ese grandísimo libro abierto ante los ojos; quiero decir, el universo, pero no se puede entender si antes no se aprende a entender la lengua, a conocer los caracteres en los que está escrito. Está escrito en lengua matemática y sus caracteres son triángulos, círculos y otras figuras geométricas, sin las cuales es imposible entender ni una palabra; sin ellos es como girar vanamente en un oscuro laberinto;

Galileo Galilei

Introducción

Muchos cursos de la formación de especialidad de los estudiantes de Pedagogía en Educación Matemática se quedan principalmente en lo teórico, dejando de lado las aplicaciones tanto en Matemática como en otros campos. Tal motivo es el que mueve a los alumnos-tesistas a presentar este seminario: Álgebra Lineal y con Aplicaciones, que puede ser de carácter interesante tanto para docentes que ya se encuentran ejerciendo, como para estudiantes secundarios o de pregrado que tienen nociones de los conceptos básicos del Álgebra Lineal; esto porque su nivel de complejidad es en forma creciente; desde una aplicación en Circuitos Eléctricos donde se utilizan sistemas de ecuaciones; hasta una aplicación en Programación Lineal donde se utilizan algunos conceptos aprendidos en Enseñanza Media (inecuaciones lineales de dos incógnitas), algunas nociones de cálculo y finalmente algunos conceptos de Álgebra Lineal.

Es importante mencionar que la historia del Álgebra Lineal Moderna se remonta a los años de 1843 cuando William Rowan Hamilton (de quien proviene el uso del término vector) creó los cuaterniones; y de 1844 cuando Hermann Grassmann publicó su libro *Die Lineare Ausdehnungslehre*, con lo cual se da inicio al estudio de numerosas áreas en donde el Álgebra Lineal toma gran importancia.

El Álgebra Lineal tiene sus orígenes en el estudio de los vectores en el plano y en el espacio tridimensional cartesiano. Aquí, un vector es un segmento, caracterizado por su longitud (o magnitud), dirección y sentido (u orientación). Los vectores pueden entonces utilizarse para representar ciertas magnitudes físicas como las fuerzas, pueden sumarse y ser multiplicados por escalares; formando entonces el primer ejemplo de espacio vectorial real. Estas fueron las primeras conclusiones con las cuales, se comienza a considerar otros campos de estudio. Definiéndose hoy en día Álgebra Lineal como la rama de la Matemática que concierne al estudio de vectores, espacios vectoriales, transformaciones lineales y sistemas de ecuaciones lineales, entre otros.

ÍNDICE DE FIGURAS

1.1. Solución gráfica del sistema (1.1)	2
1.2. Solución gráfica del sistema (1.2)	3
1.3. Solución gráfica del sistema (1.3)	3
1.4. Ramal de un circuito eléctrico	6
1.5. Dirección de la corriente a lo largo de una resistencia	8
1.6. Ejemplo 1.4	8
1.7. Ejemplo 1.5	10
1.8. Flujo de tráfico en las principales calles de Chillán	12
2.1. Paraderos de micros P_1, P_2, P_3 y P_4	28
2.2. Representación gráfica de la Tabla N°1	30
2.3. Diagrama de transición entre estados	31
3.1. $y - A\tilde{u}$ es ortogonal Au	50
3.2. Gráfica de $2x^2 - 3x + 5 = y$	53
3.3. Representación gráfica de la Tabla N°1	54
5.1. Elipse	80
5.2. Elipse rotada en $63,4^\circ$	81
5.3. Hipérbola	82
5.4. Hipérbola rotada en 45°	83
5.5. Par de rectas	84
5.6. Elipsoide $\frac{(x')^2}{1/3} + \frac{(y')^2}{1/6} + \frac{(z')^2}{1/9} = 1$	88
5.7. Probabilidad Total	101
5.8. Muelle	111
5.9. Muelles acoplados	111
5.10. Muelle desde x_1 a x_2	112
5.11. Muelle desde x_{i-1} a x_{i+1}	112
5.12. Muelle desde x_{N-1} a x_N	113
5.13. Muelle con 2 partículas	114

5.14. Muelle con 3 partículas	115
5.15. Muelle con 4 partículas	115
5.16. Sistema formado por tres partículas	116
6.1. Solución de la inecuación lineal	127
6.2. Solución del sistema de inecuaciones lineales	128
6.3. Región Factible	130
6.4. Región factible y vector de la función objetivo	131
6.5. Solución gráfica. Paralelas al vector por los vértices.	131
6.6. Región Factible	133
6.7. Solución gráfica	139
6.8. Solución gráfica	140

ÍNDICE GENERAL

Agradecimientos	III
Resumen	V
Abstract	VII
Índice de figuras	XI
Índice general	XIII
1. Sistemas de Ecuaciones	1
1.1. Introducción	1
1.2. ¿Qué es un Sistema de Ecuaciones Lineales?	2
1.3. Tipos de Sistemas de Ecuaciones Lineales:	2
1.4. ¿Qué es una matriz?	4
1.5. Operaciones Elementales	5
1.6. Eliminación Gaussiana	5
1.7. Aplicaciones	6
1.7.1. Análisis de Redes Eléctricas	6
1.7.2. Flujo de Tráfico	12
1.8. Ejercicios Propuestos	15
2. Matrices y sus Propiedades	17
2.1. Introducción	17
2.2. Operaciones con matrices	18
2.2.1. Suma de Matrices	18
2.2.2. Multiplicación de Matrices	18
2.2.3. Multiplicación por un escalar	18
2.3. Matriz Identidad	19
2.4. Matriz Transpuesta, Simétrica y sus propiedades	19

2.5. Matrices Estocásticas	20
2.6. Determinantes y sus Propiedades	20
2.6.1. Determinantes de orden 2 y 3	21
2.6.2. Propiedades de los determinantes	21
2.7. Inversa de una Matriz	23
2.7.1. Cálculo de la inversa de una matriz	23
2.8. Aplicaciones	24
2.8.1. Criptografía	24
2.8.2. Seriación en Arqueología	25
2.8.3. Relaciones de Grupo en Sociología	27
2.8.4. Cadenas de Markov	31
2.9. Ejercicios Propuestos	35
3. Espacios Vectoriales	39
3.1. Introducción	39
3.2. ¿Qué es un Espacio Vectorial?	40
3.3. Vectores Linealmente Independientes y Vectores Linealmente Dependientes	40
3.4. ¿Qué es una Base de un Espacio Vectorial V ?	41
3.5. Transformaciones Lineales	41
3.6. Bases Ortonormales y Proyecciones en \mathbb{R}^n	42
3.7. Aproximación por mínimos cuadrados	48
3.7.1. Aproximación por una Recta	48
3.7.2. Aproximación Cuadrática	52
3.7.3. Aplicación a la Geología Minera	54
3.7.4. Aplicación al Movimiento Rectilíneo Uniforme	56
3.8. Ejercicios Propuestos	57
4. Espacios con Producto Interno y Proyecciones	59
4.1. Introducción	59
4.2. Aproximación por mínimos cuadrados a una función continua	61
4.3. Aplicaciones	62
4.3.1. Mínimos Cuadrados	62
4.3.2. Aproximaciones de Fourier	65
4.4. Ejercicios Propuestos	67
5. Valores y Vectores Propios	69
5.1. Introducción	69
5.2. ¿Qué son los valores y vectores propios?	70
5.3. Matriz Característica, Ecuación Característica y Polinomio Característico	70
5.3.1. Cálculo de los valores y vectores propios de una matriz	70
5.4. Formas Cuadráticas	72
5.4.1. Matrices Semejantes y Diagonalización	73
5.4.2. Matrices Simétricas y Diagonalización Ortogonal	75
5.4.3. Formas Cuadráticas	76
5.5. Aplicaciones	88

5.5.1. Un Modelo de Crecimiento de Población	88
5.5.2. Cadenas de Markov	95
5.5.3. Aplicación a la Demografía	96
5.5.4. Aplicación a la Metereología	100
5.5.5. Sistemas Oscilantes	108
5.6. Ejercicios Propuestos	121
6. Programación Lineal	125
6.1. Introducción	125
6.2. ¿Qué es la programación lineal?	126
6.3. Inecuaciones Lineales	126
6.3.1. Inecuaciones lineales con 2 variables	126
6.3.2. Sistema de inecuaciones lineales con 2 variables	127
6.4. Problemas de optimización de una función sujeta a restricciones	129
6.4.1. Forma geométrica	130
6.4.2. Forma algebraica	131
6.5. Aplicación a problemas concretos de 2 variables	132
6.6. ¿Qué es el Método Simplex?	134
6.6.1. Interpretación geométrica del método simplex	138
6.7. Sugerencias para formular modelos de programación lineal	141
6.8. Tipos de soluciones	141
6.9. Algoritmo Simplex	141
6.10. Aplicación	142
6.10.1. Análisis Matricial del método simplex	142
6.11. Ejercicios Propuestos	147
Conclusiones	149
Bibliografía	151

CAPÍTULO 1

Sistemas de Ecuaciones

1.1. Introducción

La motivación fundamental para el estudio de Matrices se basa en el análisis de la solución de Sistemas de Ecuaciones Lineales. Los algoritmos que simplifican la búsqueda de soluciones están basados en representar dichos sistemas por medio de una matriz y trabajar con las ecuaciones por medio de operaciones elementales en la matriz que lo representa.

En las secciones (1.2), (1.3), (1.4) y (1.5) comenzaremos definiendo conceptos básicos como Sistemas de Ecuaciones y Tipos de Soluciones, Matriz, Operaciones Elementales, entre otros.

En la sección (1.6) analizaremos un método para resolver sistemas de ecuaciones lineales por medio de su representación matricial: la Eliminación Gaussiana. La idea subyace en manipular matrices por medio de operaciones de filas hasta llegar a la forma escalonada reducida de dicha matriz y finalmente utilizar sustitución regresiva para encontrar todas las soluciones.

En la última sección (1.7) se mostrarán aplicaciones en *Circuitos Eléctricos* y *Flujo de Tráfico*.

1.2. ¿Qué es un Sistema de Ecuaciones Lineales?

Definición 1.1 *Un sistema de ecuaciones lineales es un conjunto de ecuaciones que presentan soluciones comunes.*

1.3. Tipos de Sistemas de Ecuaciones Lineales:

Definición 1.2 (Sistema Compatible Determinado o Consistente) *Es aquel en que la solución de dicho sistema es única.*

Ejemplo 1.1

$$\begin{aligned} 2x - 3y &= 9 \\ 3x + 4y &= 5 \end{aligned} \tag{1.1}$$

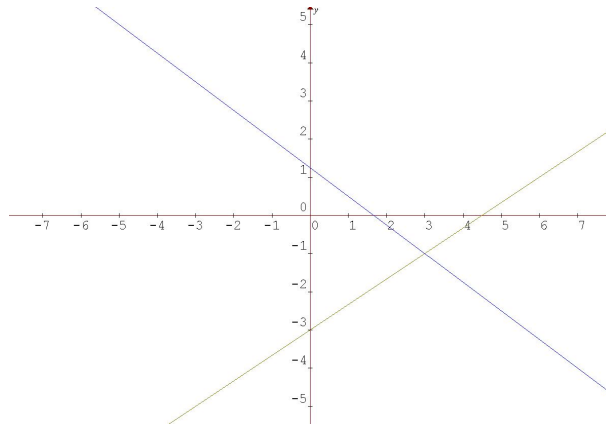


Figura 1.1: Solución gráfica del sistema (1.1)

este sistema tiene una única solución $x = 3$, $y = -1$ por lo tanto es compatible determinado

Definición 1.3 (Sistema Compatible Indeterminado) *Es aquel en que no existe una única solución, vale decir, tiene infinitas soluciones.*

Ejemplo 1.2

$$\begin{aligned} 2x - \frac{1}{3}y &= 4 \\ x - \frac{1}{6}y &= 2 \end{aligned} \tag{1.2}$$

se puede notar que la segunda ecuación es idéntica a la primera multiplicada por 2, por lo tanto el sistema tiene infinitas soluciones

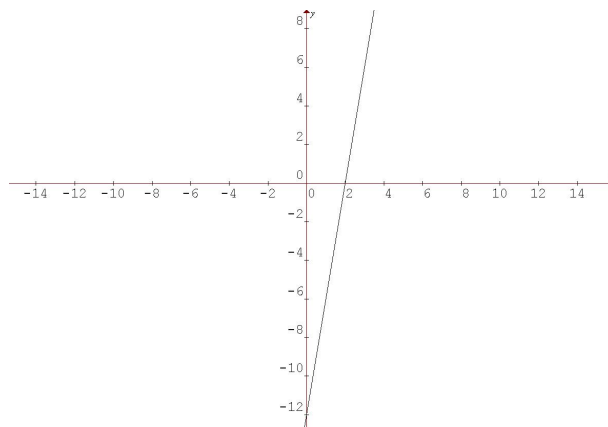


Figura 1.2: Solución gráfica del sistema (1.2)

Definición 1.4 (Sistema Incompatible o Inconsistente) *Es aquel que no tiene solución.*

Ejemplo 1.3

$$\begin{aligned} 2x - 6y &= 2 \\ x - 3y &= 3 \end{aligned} \tag{1.3}$$

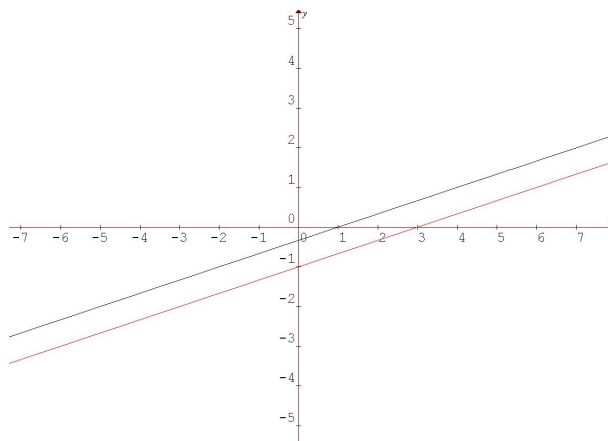


Figura 1.3: Solución gráfica del sistema (1.3)

este sistema es incompatible

1.4. ¿Qué es una matriz?

Definición 1.5 Una matriz A de $m \times n$ es un arreglo rectangular de mn números dispuestos en m filas y n columnas

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \dots & \dots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

donde los a_{ij} se llaman elementos de la matriz y que generalmente son números reales o complejos. Y donde la componente o elemento ij de A , denotado por a_{ij} es el número que aparece en la fila i y la columna j de A . Por lo general las matrices se denotan con letras mayúsculas.

Observación 1.1 Si A es una matriz de $m \times n$ con $m = n$, entonces A se llama matriz cuadrada y su orden es n .

Representación de un sistema utilizando matrices

El sistema:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1j}x_j + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2j}x_j + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{in}x_n &= b_i \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mj}x_j + \dots + a_{mn}x_n &= b_m \end{aligned}$$

puede reescribirse así:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix} \bullet \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_j \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_j \\ \dots \\ b_n \end{pmatrix}$$

o sea $A \bullet X = B$ donde $A = [a_{ij}]$ es la matriz de los coeficientes; X es una matriz columna que representa las incógnitas y B es una matriz columna que representa los términos independientes.

Observación 1.2 Dos sistemas de ecuaciones son equivalentes \Leftrightarrow tienen la misma solución.

1.5. Operaciones Elementales

Una operación elemental sobre filas en una matriz A de dimensión $m \times n$ es una de las siguientes operaciones:

Operación tipo 1 Intercambiar la fila i , con la m

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix} E_{im} \sim B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & \dots & a_{in} \end{pmatrix}$$

Operación tipo 2 Multiplicar la fila i por un escalar $\lambda \neq 0$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix} E_i(\lambda) \sim B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \lambda a_{i1} & \lambda a_{i2} & \dots & \lambda a_{ij} & \dots & \dots & \lambda a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix}$$

Operación tipo 3 Suma λ veces la fila m , a la fila i ($i \neq m$)

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} E_{im}(\lambda) \sim B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} + \lambda a_{m1} & a_{i2} + \lambda a_{m2} & \dots & a_{ij} + \lambda a_{mj} & \dots & a_{in} + \lambda a_{mn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

Notar que una matriz A de orden $m \times n$ se dice equivalente por filas a una matriz B de orden $m \times n \Leftrightarrow B$ proviene de realizar una sucesión finita de operaciones elementales sobre las filas de A .

1.6. Eliminación Gaussiana

Consiste en reducir por filas la matriz de coeficientes a la forma escalonada por filas, luego se despeja el valor de la última incógnita y después se usa la sustitución hacia atrás para las demás

incógnitas.

Procedimiento:

- Anotar la matriz ampliada $[A:b]$
- Escalonar la matriz ampliada $[A:b]$
- Resolver este nuevo sistema equivalente

Una matriz está en la forma escalonada por filas si cumple las siguientes condiciones:

- Todas las filas (si las hay) cuyos elementos son todos ceros aparecen en la parte inferior de la matriz.
- El primer número diferente de cero (comenzando por la izquierda) en cualquier fila cuyos elementos no todos son cero es 1.
- Si dos filas sucesivas tienen elementos distintos de cero, entonces el primer 1 en la fila de abajo, está más hacia la derecha que el primer 1 en la fila de arriba.

Cualquier columna que contiene al primer 1 en una fila tiene ceros en el resto de sus componentes. El primer número diferente de cero en una fila (si la hay) se llama pivote para esa fila.

1.7. Aplicaciones

1.7.1. Análisis de Redes Eléctricas

Los sistemas de ecuaciones en la física son una de las herramientas más utilizadas para determinar la corriente a través de varios ramales de una determinada red eléctrica, desde luego existen más métodos matemáticos para resolverlos, de hecho el señor Gustav Robert Kirchhoff determinó el uso de leyes para determinar la corriente de un ramal determinado, las cuales utilizaremos para dar forma a los sistemas que resolveremos en los próximos problemas, sin embargo, antes tenemos que considerar algunas definiciones previas.

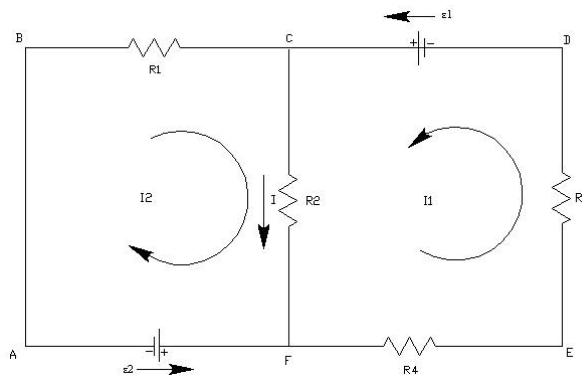


Figura 1.4: Ramal de un circuito eléctrico

1.7.1.1. Definiciones Previas

En este tipo de circuitos es necesario introducir algunos términos:

Definición 1.6 (Nodo) *Es aquel punto del circuito donde se unen tres o más conductores.*

Por ejemplo, los puntos C y F del circuito anterior son nodos. Las corrientes (cargas eléctricas) pasan por los nodos pero no se acumulan en ellos. Las corrientes que llegan al nudo se consideran positivas y las que salen de él, se consideran negativas, para los efectos de las leyes de Kirchhoff.

Definición 1.7 (Malla o Red) *Es todo recorrido cerrado por los conductores, de tal modo que cada uno de ellos se recorre una sola vez.*

En el circuito anterior se pueden identificar las mallas ABCFA y DEFCD. En cambio, el recorrido ABCFEDCF no es una malla, pues el conductor CF se ha recorrido dos veces. En la resolución de muchos problemas relacionados con circuitos eléctricos complejos son muy útiles las llamadas Leyes de Kirchhoff, que no son otra cosa que diferentes expresiones de la ley de conservación de la energía (en este caso, de la energía eléctrica) y de la carga eléctrica.

1.7.1.2. Leyes de Kirchhoff

1. Primera ley de Kirchhoff o ley de nodos

La suma algebraica de las intensidades de corriente en cualquier nodo de un circuito es igual a cero.

Consideremos el esquema de la figura siguiente en la cual N representa el nudo; I_1 , I_2 y I_3 representan las intensidades de las corrientes que entran (I_1 e I_3), e I_2 la corriente que sale por dicho nudo, indicadas por las correspondientes flechas. Entonces, de acuerdo al convenio de los signos, tendremos la siguiente ecuación:

$$I_1 - I_2 + I_3 = 0$$

o lo que es lo mismo $I_1 + I_3 = I_2$. En otras palabras, la primera ley de Kirchhoff también puede enunciarse diciendo que la suma de todas las corrientes que llegan al nodo es igual a la suma de todas las corrientes que salen de él.

2. Segunda ley de Kirchhoff o ley de voltajes (o ley de las mallas)

La suma algebraica de las subidas y caídas de potencial en cualquier malla es igual a cero.

- a) Cuando se recorre la malla en dirección de la corriente a lo largo de una resistencia se tiene una bajada de potencial, pues se va de un potencial mayor a otro menor, es decir, en el sentido en el que convencionalmente fluye la corriente. Por lo tanto, se tiene una bajada o caída de potencial igual a menos el producto de la intensidad de la corriente I que circula por la resistencia misma R : $-R$. En cambio, si se recorre la malla en el sentido contrario, se tiene una subida de potencial igual a $+R$.

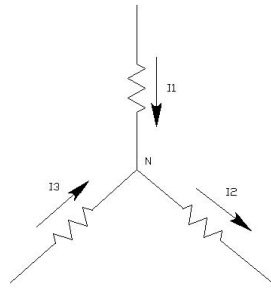


Figura 1.5: Dirección de la corriente a lo largo de una resistencia

- b) Cuando se recorre una fuente en el sentido de la fuerza electromotriz (ε), se tendrá una subida de potencial igual a $+\varepsilon$. En cambio, al recorrer en el sentido contrario, se tendrá una bajada de potencial igual a $-\varepsilon$.

Por ejemplo, en la malla ABCFA de la figura anterior, aplicando la segunda ley de Kirchhoff, se cumple que: $-\varepsilon_2 + I_2R_1 + IR_3 = 0$

Ejemplo 1.4 *Considérese la siguiente red eléctrica y determine la corriente a través de cada ramal de esta red.*

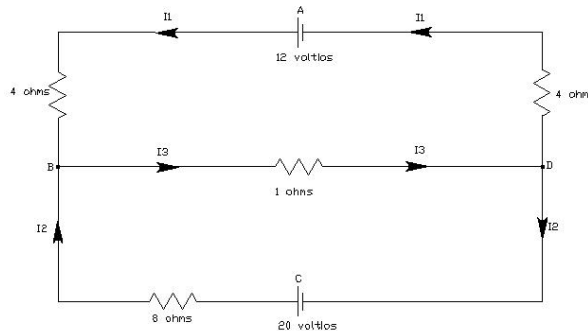


Figura 1.6: Ejemplo 1.4

Dentro de esta red eléctrica podemos identificar las baterías que son denotadas por $(|I)$, medidas en voltios y también las resistencias $(\diagdown \diagup)$, medidas en OHMs. Una observación que mencionaremos es que la corriente que entra a cada batería será la misma que la que sale de ella, respetando las leyes mencionadas al principio.

Las corrientes en los distintos ramales del circuito son: I_1, I_2 e I_3 . Las leyes de Kirchhoff se refieren a las uniones (nodos) y a los caminos cerrados (mallas). Tenemos B y D dos uniones y además tres caminos cerrados ABDA, CBDC y ABCDA.

Las ecuaciones de las uniones son:

$$\begin{aligned} \text{B: } & I_1 + I_2 = I_3 \\ \text{D: } & I_3 = I_1 + I_2 \end{aligned}$$

Despejando resulta $I_1 + I_2 - I_3 = 0$

Luego, las ecuaciones de los caminos son:

$$\begin{aligned} \text{ABDA: } & 4I_1 + 1I_3 + 4I_1 = 12 \\ \text{CBDC: } & 8I_2 + 1I_3 = 20 \end{aligned}$$

De esto, tenemos un sistema de tres ecuaciones lineales con tres incógnitas.

$$\begin{aligned} I_1 + I_2 - I_3 &= 0 \\ 4I_1 + 4I_1 + 1I_3 &= 12 \\ +8I_2 + 1I_3 &= 20 \end{aligned}$$

Que escritas en forma de matriz ampliada nos queda: $\left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 8 & 0 & 1 & 12 \\ 0 & 8 & 1 & 20 \end{array} \right)$

Si utilizamos la Eliminación Gaussiana se tiene:

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 8 & 0 & 1 & 12 \\ 0 & 8 & 1 & 20 \end{array} \right) E_{12}(-8) \sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -8 & 8 & 12 \\ 0 & 8 & 1 & 24 \end{array} \right) E_{23}(1) \sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -8 & 8 & 12 \\ 0 & 0 & 9 & 36 \end{array} \right)$$

Luego

$$\begin{aligned} I_1 + I_2 - I_3 &= 0 \\ -8I_2 + 8I_3 &= 12 \\ 9I_3 &= 36 \end{aligned}$$

Se tiene I_3 y luego usando sustitución regresiva obtenemos:

$$\begin{aligned} I_3 &= 4 \\ I_2 &= 5/2 = 2,5 \\ I_1 &= 3/2 = 1,5 \end{aligned}$$

En efecto, las unidades son amperes. Por lo tanto la solución es única.

Observación 1.3 *Se entiende por ampere la unidad que mide la intensidad de una corriente eléctrica.*

Ejemplo 1.5 *Considérese la siguiente red eléctrica y determine la corriente a través de cada ramal de esta red.*

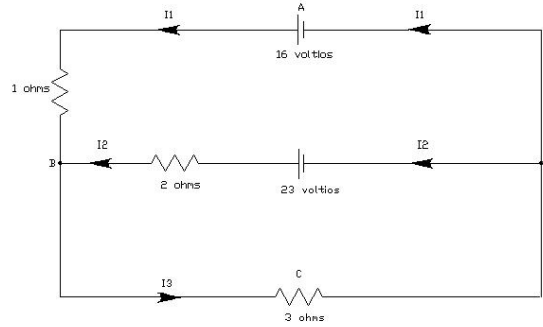


Figura 1.7: Ejemplo 1.5

Observación 1.4 *Se ha considerado este ejemplo, puesto que el grado de dificultad es mayor que el anterior, debido a que uno de sus ramales va en sentido contrario y el sistema ya no es tan evidente.*

Uniones (Nodos):

$$\begin{aligned} \text{B: } I_1 + I_2 &= I_3 \\ \text{D: } I_3 &= I_1 + I_2 \end{aligned}$$

despejando resulta: $I_1 + I_2 - I_3 = 0$

Caminos:

$$\begin{aligned} \text{ABCD: } I_1 + 3I_3 &= 16 \\ \text{ABDA: } I_1 + 2(-I_2) &= 16 + (-23) \end{aligned}$$

Cabe destacar que la corriente a lo largo del ramal BD está en dirección $-I_2$, y el voltaje es -22 . Luego, las ecuaciones son:

$$\begin{aligned} I_1 + I_2 - I_3 &= 0 \\ I_1 - 2I_2 &= -7 \\ I_1 + 3I_3 &= 16 \end{aligned}$$

Si utilizamos la Eliminación Gaussiana se tiene:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 1 & -2 & 0 & -7 \\ 1 & 0 & 3 & 16 \end{array} \right) E_{12}(1) &\sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 1 & -7 \\ 1 & 0 & 3 & 16 \end{array} \right) E_{13}(-1) &\sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 1 & -7 \\ 0 & -1 & 4 & 16 \end{array} \right) E_{32}(-3) &\sim \\ \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 0 & -11 & -55 \\ 0 & -1 & 4 & 16 \end{array} \right) E_{32} &\sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -1 & 4 & 16 \\ 0 & 0 & -11 & -55 \end{array} \right) \end{aligned}$$

$$I_3 = 5$$

$$I_2 = 4$$

$$I_1 = 1$$

Donde $I_1 = 1$ Ampere, $I_2 = 4$ Ampere y $I_3 = 5$ Ampere.

1.7.2. Flujo de Tráfico

Considérese la red de calles del centro de Chillán. Todas las calles son de un sólo sentido, las flechas indican la dirección del flujo de tráfico. El flujo de tráfico que entra y que sale de esta red se mide en vehículos por hora (vph).

Los números dados aquí se suponen en las horas pick 7A.M a 9A.M a media semana. En la noche del viernes se considera un aumento en todo el tráfico. Construyendo un modelo matemático para analizar esta red.

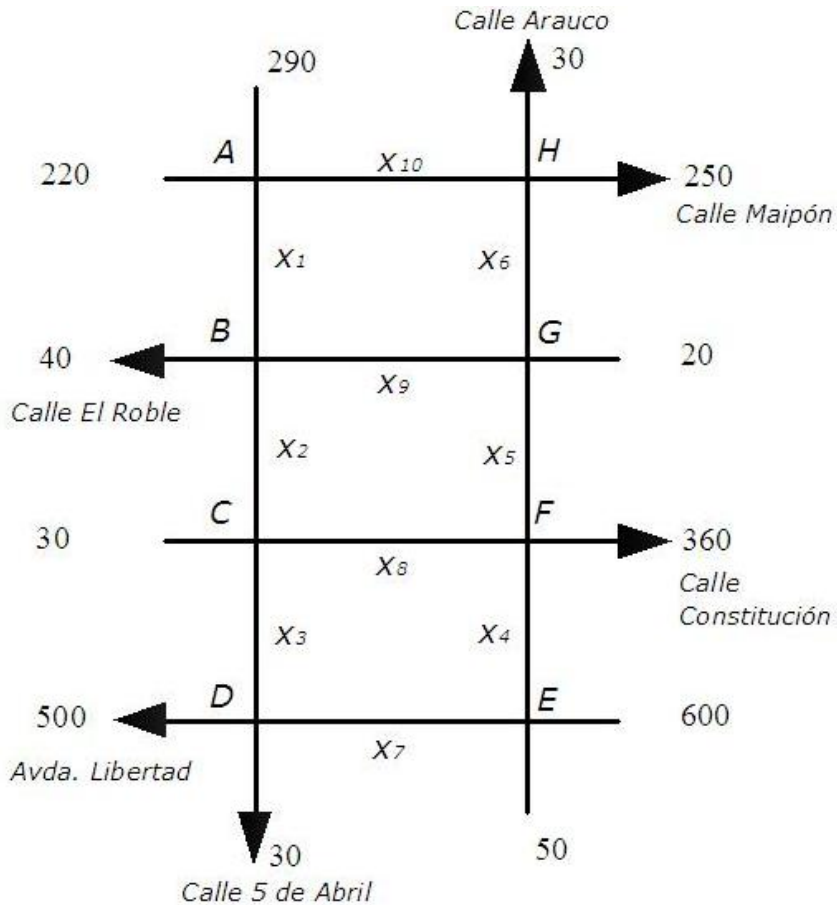


Figura 1.8: Flujo de tráfico en las principales calles de Chillán

$$\begin{aligned} \text{Unión A: Tráfico de Entrada} &= 290 + 220 \\ \text{Tráfico de Salida} &= x_1 + x_{10} \\ \Rightarrow x_1 + x_{10} &= 510 \end{aligned}$$

$$\begin{aligned} \text{Unión B: Tráfico de Entrada} &= x_1 + x_9 \\ \text{Tráfico de Salida} &= x_2 + 40 \\ \Rightarrow x_1 + x_9 &= x_2 + 40 \end{aligned}$$

$$\begin{aligned} \text{Unión C: Tráfico de Entrada} &= x_2 + 30 \\ \text{Tráfico de Salida} &= x_3 + x_8 \\ \Rightarrow x_3 + x_8 &= x_2 + 30 \end{aligned}$$

$$\begin{aligned} \text{Unión D: Tráfico de Entrada} &= x_3 + x_7 \\ \text{Tráfico de Salida} &= 500 + 30 \\ \Rightarrow x_3 + x_7 &= 530 \end{aligned}$$

$$\begin{aligned} \text{Unión E: Tráfico de Entrada} &= 50 + 600 \\ \text{Tráfico de Salida} &= x_7 + x_4 \\ \Rightarrow x_7 + x_4 &= 650 \end{aligned}$$

$$\begin{aligned} \text{Unión F: Tráfico de Entrada} &= x_4 + x_8 \\ \text{Tráfico de Salida} &= x_5 + 360 \\ \Rightarrow x_4 + x_8 &= x_5 + 360 \end{aligned}$$

$$\begin{aligned} \text{Unión G: Tráfico de Entrada} &= x_5 + 20 \\ \text{Tráfico de Salida} &= x_9 + x_6 \\ \Rightarrow x_9 + x_6 &= x_5 + 20 \end{aligned}$$

$$\begin{aligned} \text{Unión H: Tráfico de Entrada} &= x_6 + x_{10} \\ \text{Tráfico de Salida} &= 250 + 30 \\ \Rightarrow x_6 + x_{10} &= 280 \end{aligned}$$

Esto nos entrega el siguiente sistema de ecuaciones lineales:

$$\begin{array}{rcccccccccc}
 \text{Unión A:} & x_1 & & & & & & & & & +x_{10} & = & 510 \\
 \text{Unión B:} & x_1 & -x_2 & & & & & & & & +x_9 & = & 40 \\
 \text{Unión C:} & & -x_2 & +x_3 & & & & & & & +x_8 & = & 30 \\
 \text{Unión D:} & & & x_3 & & & & & & & +x_7 & = & 530 \\
 \text{Unión E:} & & & & x_4 & & & & & & +x_7 & = & 650 \\
 \text{Unión F:} & & & & x_4 & -x_5 & & & & & +x_8 & = & 360 \\
 \text{Unión G:} & & & & & -x_5 & +x_6 & & & & +x_9 & = & 20 \\
 \text{Unión H:} & & & & & & x_6 & & & & +x_{10} & = & 280
 \end{array}$$

Observación 1.5 *Note que $x_5 = 0$ debido a que el flujo de tráfico en el paseo Arauco es 0.*

En forma matricial el sistema se puede representar de la forma $A \cdot x = b$; donde $[A:b]$ está dada por la matriz

$$\left(\begin{array}{cccccccccc|c}
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 510 \\
 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 40 \\
 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 30 \\
 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 530 \\
 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 650 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 360 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 20 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 280
 \end{array} \right)$$

realizando las siguiente operaciones elementales por filas:

$$E_{21}(-1)E_{78}(-1) \sim E_{27}(-1)E_{56}(-1) \sim E_{32}(-1)E_{45}(-1) \sim E_{43}(-1) E_2(-1)E_{68} \sim E_{48} E_{56} \sim$$

se obtiene la matriz equivalente:

$$\left(\begin{array}{cccccccccc|c}
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 510 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 210 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 240 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 360 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 280 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 290 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -260 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right)$$

Luego el sistema de ecuaciones que corresponde a esta forma escalonada reducida es:

$$\begin{array}{rcl}
 x_1 + x_{10} & = & 510 \\
 x_2 & = & 210 \\
 x_3 + x_8 & = & 240 \\
 x_4 + x_8 & = & 360 \\
 x_6 + x_{10} & = & 280 \\
 x_7 - x_8 & = & 290 \\
 x_9 - x_{10} & = & -260
 \end{array}$$

Expresando cada variable principal en términos de las otras variables:

$$\begin{aligned}x_1 &= 510 - x_{10} \\x_2 &= 210 \\x_3 &= 240 - x_8 \\x_4 &= 360 - x_8 \\x_6 &= 280 - x_{10} \\x_7 &= 290 + x_8 \\x_9 &= x_{10} - 260\end{aligned}$$

De aquí podemos inferir que las variables independientes son x_{10} y x_8 y los valores que pueden tomar están entre los intervalos:

$$x_{10} \in [260, 280]$$

$$x_8 \in [0, 240]$$

Por lo tanto si Ud. da un valor a las variables x_{10} y x_8 entonces obtendrá un flujo de tráfico, pero si asigna otro valor pertinente de acuerdo a los intervalos mencionados obtendrá un flujo de tráfico distinto. Existen infinitos flujos de tráfico para esta red de calles.

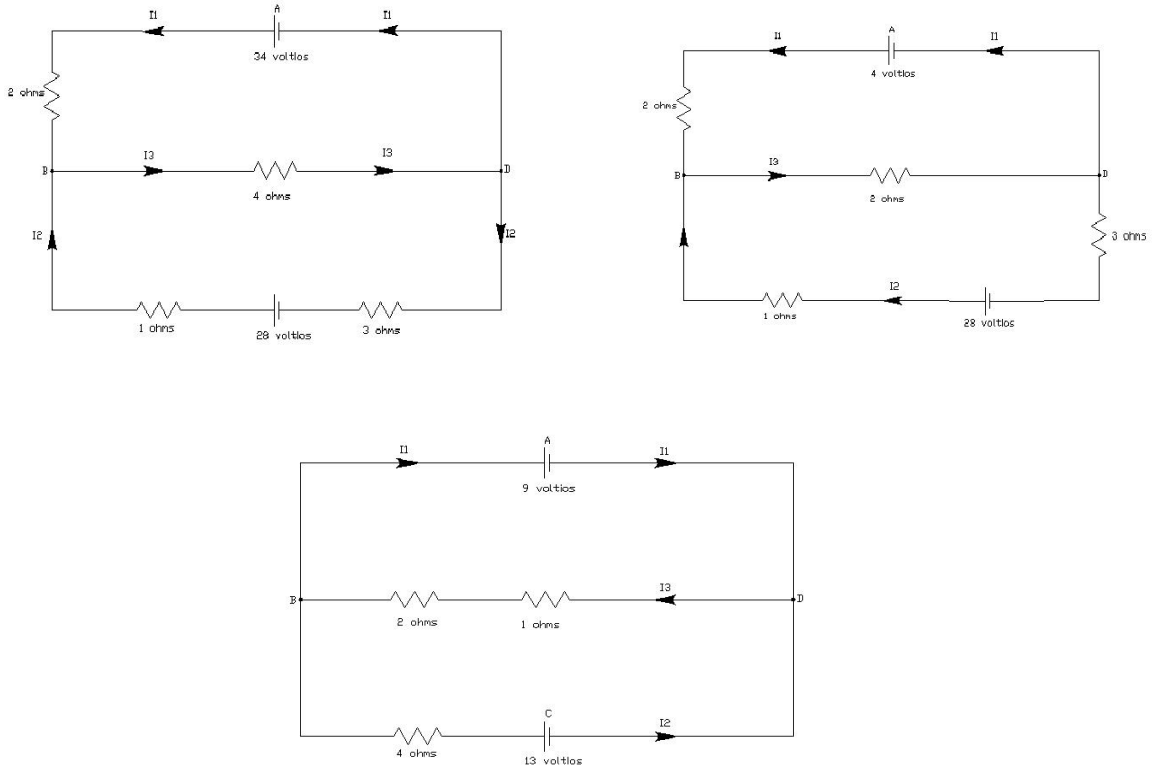
Por ejemplo un flujo de tráfico posible sería tomando $x_{10} = 275$ y $x_8 = 136$ y se tendría que:

$$\begin{aligned}x_1 &= 235 \\x_2 &= 210 \\x_3 &= 104 \\x_4 &= 224 \\x_6 &= 5 \\x_7 &= 426 \\x_9 &= 15\end{aligned}$$

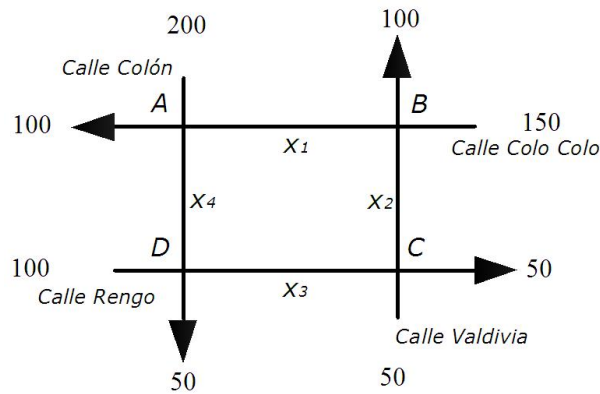
Observación 1.6 *Para recabar estos datos se necesitaron doce personas, para así determinar el flujo de tráfico de entradas y salidas en las horas pick. Piense en que si se hubiesen considerado las calles correspondientes al centro de la Ciudad, las variables aumentarían a 16 y la matriz aumentaría considerablemente de tamaño.*

1.8. Ejercicios Propuestos

1. Considere las siguientes redes eléctricas y determine la corriente (en Amperes) a través de cada ramal.



2. Construya un modelo matemático que describa el flujo de tráfico en la red de las calles de Los Angeles mencionadas. Todas las calles son de un solo sentido, en la dirección indicada. Las unidades se dan en vehículos por hora. De tres flujos de tráfico posibles (distintos). ¿Cuál es la afluencia mínima posible que puede esperarse a lo largo de ramal AB?



CAPÍTULO 2

Matrices y sus Propiedades

2.1. Introducción

En el presente capítulo se introducirán Operaciones con Matrices ($+$, \bullet) lo que nos permite abarcar más temas relacionados con Matrices o que se puedan representar por medio de ellas. Además se presentarán algunos tipos de matrices y sus propiedades y algo fundamental: se enunciará la condición necesaria para que una matriz tenga inversa y en caso de que esta exista se dará un ejemplo para ilustrar el método para encontrarla.

El capítulo termina en la sección (2.8) con el estudio de algunas aplicaciones a la *Sociología*, *Criptografía*, *Seriación en Arqueología* y *Cadenas de Markov*.

Al leer este capítulo el lector puede notar que se han omitido algunas demostraciones de propiedades y proposiciones, puesto que se realizan en cualquier curso habitual de Álgebra Lineal; no obstante, en otros capítulos se enuncian teoremas los cuales incluyen su demostración.

2.2. Operaciones con matrices

2.2.1. Suma de Matrices

Sean A y $B \in M_{m \times n}$ donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix} \text{ y } B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{i1} & b_{i2} & \dots & b_{ij} & \dots & \dots & b_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mj} & \dots & \dots & b_{mn} \end{pmatrix}$$

Se define $A + B =$

$$\begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1j} + b_{1j} & \dots & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2j} + b_{2j} & \dots & \dots & a_{2n} + b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \dots & a_{ij} + b_{ij} & \dots & \dots & a_{in} + b_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mj} + b_{mj} & \dots & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

2.2.2. Multiplicación de Matrices

Sean $A = [a_{ij}]_{m \times n}$ y $B = [b_{ij}]_{n \times p}$

Se define

$$A \bullet B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \bullet \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b_{i1} & b_{i2} & \dots & b_{ij} & \dots & b_{ip} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1j} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2j} & \dots & c_{2p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{i1} & c_{i2} & \dots & c_{ij} & \dots & c_{ip} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mj} & \dots & c_{mp} \end{pmatrix}_{m \times p}$$

en donde c_{ij} es el producto punto de la fila i de A y la columna j de B ; esto es: $c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj}$

2.2.3. Multiplicación por un escalar

Sea $A = [a_{ij}] \in M_{m \times n}$ y α un escalar, entonces la matriz $\alpha \cdot A$, está dada por

$$\alpha \cdot A = \begin{pmatrix} \alpha \cdot a_{11} & \alpha \cdot a_{12} & \dots & \alpha \cdot a_{1j} & \dots & \dots & \alpha \cdot a_{1n} \\ \alpha \cdot a_{21} & \alpha \cdot a_{22} & \dots & \alpha \cdot a_{2j} & \dots & \dots & \alpha \cdot a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha \cdot a_{i1} & \alpha \cdot a_{i2} & \dots & \alpha \cdot a_{ij} & \dots & \dots & \alpha \cdot a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha \cdot a_{m1} & \alpha \cdot a_{m2} & \dots & \alpha \cdot a_{mj} & \dots & \dots & \alpha \cdot a_{mn} \end{pmatrix}$$

2.3. Matriz Identidad

La Matriz identidad I_n es una matriz de $n \times n$ cuyos elementos de la diagonal principal son iguales a 1 y todos los demás elementos son 0.

Así:

$$I_n = [a_{ij}]$$

donde

$$a_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En las matrices $M_{4 \times 4}$ la matriz identidad es:

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Proposición 2.1 Sea A una matriz cuadrada de $n \times n$. Entonces $AI_n = I_nA = A$

2.4. Matriz Transpuesta, Simétrica y sus propiedades

Definición 2.1 (Matriz Transpuesta) Sea $A = [a_{ij}]$ una matriz de $m \times n$. Entonces la transpuesta de A , que se escribe A^t , es la matriz de $n \times m$ obtenida al intercambiar las filas de A por las columnas de A ; así $A^t = [a_{ji}]$. Luego si

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

Propiedades

Sean $A = [a_{ij}]$ una matriz de $n \times m$ y $B = [b_{ij}]$ una matriz de $m \times p$. Entonces

1. $(A^t)^t = A$
2. $(AB)^t = B^tA^t$
3. Si A y B son de $n \times m$, entonces $(A + B)^t = A^t + B^t$
4. Si A es invertible, entonces A^t es invertible y $(A^t)^{-1} = (A^{-1})^t$

Definición 2.2 (Matriz Simétrica) Una matriz A de $n \times n$ se llama simétrica si $A^t = A$. Es decir, las columnas de A son también las filas de A .

Ejemplo de una matriz simétrica:

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & 4 \\ 3 & 4 & 7 \end{pmatrix}$$

2.5. Matrices Estocásticas

Definición 2.3 (Matriz Estocástica) Una matriz estocástica es una matriz cuadrada cuyos elementos son probabilidades y la suma de sus columnas es igual a 1.

Una matriz estocástica general de 2×2 se puede escribir como:

$$A = \begin{pmatrix} x & y \\ 1 - x & 1 - y \end{pmatrix}$$

donde $0 \leq x \leq 1$ y $0 \leq y \leq 1$.

Las matrices estocásticas tienen las siguientes propiedades que se utilizarán en el transcurso de este tema.

Sea $A = [a_{ij}]_{n \times n}$ una matriz estocástica, entonces:

- $a_{ij} \geq 0$ para todo i, j
- $a_{1j} + a_{2j} + a_{3j} + \dots + a_{nj} = 1$ para cada j

En consecuencia, una matriz estocástica T es una matriz con elementos no negativos tal que la suma de los números en cada columna es 1.

Ejemplo

$$T = \begin{pmatrix} 0,25 & 0,60 & 0,25 \\ 0,60 & 0,60 & 0,30 \\ 0,15 & 0,10 & 0,45 \end{pmatrix}$$

es una matriz estocástica.

Teorema 2.1 Si A y B son matrices estocásticas del mismo tamaño, entonces $A \cdot B$ es una matriz estocástica. Así, si A es una matriz estocástica, entonces $A^2, A^3, A^4, \dots, A^n$ son estocásticas.

2.6. Determinantes y sus Propiedades

Definición 2.4 (Determinante) Sea $A \in M_{n \times n}$, $(a_{i1} \ a_{i2} \ a_{i3} \ \dots \ a_{in})$ una fila cualquiera y $(A_{i1} \ A_{i2} \ A_{i3} \ \dots \ A_{in})$ los cofactores de dicha fila. Entonces el determinante de A , denotado por $\det(A)$ o $|A|$, está dado por:

$$\det(A) = |A| = \sum_{k=1}^n a_{ik} A_{ik}$$

que corresponde al desarrollo por la i -ésima fila.

Este desarrollo del $\det(A)$ se conoce como **desarrollo por cofactores**. Se puede demostrar que es el mismo cualquiera sea la fila que se elija o también es válido si se desarrolla por una columna.

2.6.1. Determinantes de orden 2 y 3

determinante de 2×2

Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ una matriz de 2×2 . El determinante de A es:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

determinante de 3×3

Sea $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ una matriz de 3×3 . El determinante de A es:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Definición 2.5 (Cofactor) Sea A una matriz de $n \times n$. El cofactor ij de A , denotado por A_{ij} , está dado por

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

Observe que:

$$(-1)^{i+j} = \begin{cases} 1 & \text{si } i+j \text{ es par} \\ -1 & \text{si } i+j \text{ es impar} \end{cases}$$

Definición 2.6 (Matriz triangular) Una matriz cuadrada se llama triangular superior si todas sus componentes abajo de la diagonal son cero ($a_{ij}=0$ para $i > j$). Es una matriz triangular inferior si todas sus componentes arriba de la diagonal son cero ($a_{ij}=0$ para $i < j$).

Teorema 2.2 Sea $A = [a_{ij}]$ una matriz de $n \times n$ triangular superior o inferior. Entonces

$$\det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdot \dots \cdot a_{nn}$$

2.6.2. Propiedades de los determinantes

Teorema 2.3 Sean A y B dos matrices de $n \times n$. Entonces

$$\det(AB) = \det(A) \cdot \det(B)$$

Es decir, el determinante del producto es el producto de los determinantes.

Teorema 2.4 El determinante de una matriz A es igual al determinante de su transpuesta

$$\det(A^t) = \det(A)$$

Propiedad 2.1 Si cualquier fila o columna de A es un vector cero, entonces $\det(A) = 0$

Propiedad 2.2 Si la fila i o la columna j de A se multiplica por un escalar c , entonces $\det(A)$ se multiplica por c . Es decir, si se denota por B esta nueva matriz, entonces

$$|B| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ c \cdot a_{i1} & c \cdot a_{i2} & \dots & c \cdot a_{in} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = c \cdot \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = c|A|$$

Propiedad 2.3 Sean $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$, $B = \begin{pmatrix} a_{11} & a_{12} & \dots & \alpha_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \alpha_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & \alpha_{nj} & \dots & a_{nn} \end{pmatrix}$,

y $C = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} + \alpha_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} + \alpha_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} + \alpha_{nj} & \dots & a_{nn} \end{pmatrix}$

Entonces

$$\det(C) = \det(A) + \det(B)$$

En otras palabras, suponga que A , B y C son idénticas excepto por la columna j y que la columna j de C es la suma de las j -ésimas columnas de A y B . Entonces, $\det(C) = \det(A) + \det(B)$. La misma afirmación es cierta para las filas.

Propiedad 2.4 Sea B la matriz que se obtiene intercambiando dos filas o columnas de la matriz cuadrada A , entonces el determinante de B es el opuesto al determinante de A .

$$\det(B) = -\det(A)$$

Propiedad 2.5 Si A tiene dos filas o columnas iguales, entonces:

$$\det(A) = 0$$

Propiedad 2.6 Si una fila (columna) de A es un múltiplo escalar de otra fila(columna), entonces:

$$\det(A) = 0$$

Propiedad 2.7 Si se suma un múltiplo escalar de una fila(columna) de A a otra fila(columna) de A , entonces el determinante no cambia.

2.7. Inversa de una Matriz

Sean A y B dos matrices de $n \times n$. Suponga que $AB = BA = I$, entonces B es llamada inversa de A y se denota como $B = A^{-1}$. Por ende se tiene que:

$$AA^{-1} = A^{-1}A = I$$

donde I es la identidad de orden n .

Observación 2.1 Si A tiene inversa, entonces se dice que A es invertible.

Teorema 2.5 Sean A y B invertibles, entonces

1. Si una matriz es invertible, entonces su inversa es única
2. $A \cdot B$ también es invertible y además $(AB)^{-1} = B^{-1}A^{-1}$.
3. Como A es invertible, también lo es A^{-1} ; más aún, $(A^{-1})^{-1} = A$.
4. Como A es invertible y λ es un escalar no nulo; $(\lambda A)^{-1} = (1/\lambda)A^{-1}$.

2.7.1. Cálculo de la inversa de una matriz

Sea $A \in M_{n \times n}$ queremos encontrar A^{-1} de tal forma que $A \cdot A^{-1} = I_n$.

El procedimiento que utilizaremos es escribir la matriz ampliada $[A|I]$ y por medio de operaciones elementales se escalonará la matriz para finalmente convertir A en I_n , cuando esto ocurra, la ampliada que originalmente era I_n se habrá convertido en la matriz inversa.

Ejemplo 2.1 Sea

$$A = \begin{pmatrix} 3 & 1 & 4 & 3 \\ 7 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Nuestro objetivo es encontrar A^{-1} .

Entonces:

$$\begin{pmatrix} 3 & 1 & 4 & 3 & | & 1 & 0 & 0 & 0 \\ 7 & 0 & 1 & 2 & | & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \end{pmatrix} E_{13}E_{14} \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 3 & 1 & 4 & 3 & | & 1 & 0 & 0 & 0 \\ 7 & 0 & 1 & 2 & | & 0 & 1 & 0 & 0 \end{pmatrix} E_{31}(-3)E_{41}(-7) \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & -5 & 1 & 3 & | & 1 & 0 & -3 & 0 \\ 0 & -14 & -6 & 2 & | & 0 & 1 & -7 & 0 \end{pmatrix} E_{32}(5)E_{42}(14) \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 8 & | & 1 & 0 & -3 & 5 \\ 0 & 0 & -6 & 16 & | & 0 & 1 & -7 & 14 \end{pmatrix} E_{43}(6) \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 8 & | & 1 & 0 & -3 & 5 \\ 0 & 0 & 0 & 64 & | & 6 & 1 & -25 & 44 \end{pmatrix} E_4(\frac{1}{64}) \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 8 & | & 1 & 0 & -3 & 5 \\ 0 & 0 & 0 & 1 & | & 6/64 & 1/64 & -25/64 & 44/64 \end{pmatrix} E_{24}(-1)E_{34}(-8) \sim$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & | & -6/64 & -1/64 & 25/64 & 20/64 \\ 0 & 0 & 1 & 0 & | & 16/64 & -8/64 & 8/64 & -32/64 \\ 0 & 0 & 0 & 1 & | & 6/64 & 1/64 & -25/64 & 44/64 \end{pmatrix} E_{12}(-2) \sim$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 & | & 12/64 & 2/64 & 14/64 & -40/64 \\ 0 & 1 & 0 & 0 & | & -6/64 & -1/64 & 25/64 & 20/64 \\ 0 & 0 & 1 & 0 & | & 16/64 & -8/64 & 8/64 & -32/64 \\ 0 & 0 & 0 & 1 & | & 6/64 & 1/64 & -25/64 & 44/64 \end{pmatrix} E_{13}(-1) \sim$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & | & -4/64 & 10/64 & 6/64 & -8/64 \\ 0 & 1 & 0 & 0 & | & -6/64 & -1/64 & 25/64 & 20/64 \\ 0 & 0 & 1 & 0 & | & 16/64 & -8/64 & 8/64 & -32/64 \\ 0 & 0 & 0 & 1 & | & 6/64 & 1/64 & -25/64 & 44/64 \end{pmatrix}$$

por lo tanto ya hemos encontrado la inversa, ya que A por medio de operaciones elementales se convirtió en la Identidad y así la ampliada (la identidad) se ha convertido por medio de operaciones elementales en A^{-1} .

2.8. Aplicaciones

2.8.1. Criptografía

La Criptografía es el arte o ciencia de cifrar y descifrar información utilizando técnicas que hagan posible el intercambio de mensajes de manera que sólo puedan ser leídos por las personas a quienes van dirigidos. Es por eso que se han ideado distintas técnicas para cifrar los mensajes de tal forma que sólo el receptor sea capaz de descifrarlo. Una de las técnicas es la que presentamos a continuación, que consiste en cifrar el mensaje utilizando matrices.

1. Sea el mensaje: **CADA SOL TIENE SU OCASO** y la matriz codificadora sea

$$\begin{pmatrix} 3 & 1 & 4 & 3 \\ 7 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

A cada letra del alfabeto le asignaremos un número; en este caso para mayor simplicidad a cada letra se le asociará su posición en el alfabeto: la letra A es 1, la B es 2 y así sucesivamente. Al espacio entre palabras lo denotaremos con el número 0.

El mensaje entonces será:

C A D A S O L T I E N E S U O C A S O
 3 1 4 1 0 19 15 12 0 20 9 5 14 5 0 19 21 0 15 3 1 19 15

Como se utilizará una matriz de 4×4 para codificar el mensaje, divide el mensaje numerado en matrices columnas de 4×1 , de esta forma:

$$\begin{pmatrix} 3 \\ 1 \\ 4 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 19 \\ 15 \\ 12 \end{pmatrix} \begin{pmatrix} 0 \\ 20 \\ 9 \\ 5 \end{pmatrix} \begin{pmatrix} 14 \\ 5 \\ 0 \\ 19 \end{pmatrix} \begin{pmatrix} 21 \\ 0 \\ 15 \\ 3 \end{pmatrix} \begin{pmatrix} 1 \\ 19 \\ 15 \\ 0 \end{pmatrix}$$

En nuestro caso fue necesario agregar un espacio al final del mensaje para completar la matriz. Para eso se utiliza el mismo número designado para establecer el espacio entre palabras. Para poner el mensaje en código se multiplica cada una de las matrices columna, anteriores, por la matriz codificadora. Para mayor facilidad escribiremos las matrices columna como columnas de una matriz, y multiplicando, por la izquierda, por la matriz codificadora; se tiene:

$$\begin{pmatrix} 3 & 1 & 4 & 3 \\ 7 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 0 & 14 & 21 & 1 \\ 1 & 19 & 20 & 5 & 0 & 19 \\ 4 & 15 & 9 & 0 & 15 & 15 \\ 1 & 12 & 5 & 19 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 29 & 115 & 71 & 104 & 132 & 82 \\ 27 & 39 & 19 & 136 & 168 & 22 \\ 9 & 53 & 49 & 24 & 36 & 54 \\ 2 & 31 & 25 & 24 & 3 & 19 \end{pmatrix}$$

Las columnas de esta matriz producen el mensaje codificado. El mensaje se transmite en forma lineal de la manera siguiente: 29, 27, 9, 2, 115, 39, 53, 31, 71, 19, 49, 25, 104, 136, 24, 24, 132, 168, 36, 3, 82, 22, 54, 19.

Para decodificar el mensaje, el receptor escribe esta lista como una sucesión de matrices columna de 4×1 y repite la técnica anterior usando la inversa de la matriz codificadora; es decir, la matriz decodificadora. Para determinar quien es esta matriz; véase la sección referida a la inversa de una matriz; en donde se explicitan las operaciones necesarias para encontrar la inversa de una matriz, en caso de que exista. En nuestro caso, la matriz decodificadora es:

$$\begin{pmatrix} -4/64 & 10/64 & 6/64 & -8/64 \\ -6/64 & -1/64 & 25/64 & 20/64 \\ 16/64 & -8/64 & 8/64 & -32/64 \\ 6/64 & 1/64 & -25/64 & 44/64 \end{pmatrix} = 1/64 \begin{pmatrix} -4 & 10 & 6 & -8 \\ -6 & -1 & 25 & 20 \\ 16 & -8 & 8 & -32 \\ 6 & 1 & -25 & 44 \end{pmatrix}$$

Para decodificar el mensaje, se multiplican:

$$\begin{aligned} &= 1/64 \begin{pmatrix} -4 & 10 & 6 & -8 \\ -6 & -1 & 25 & 20 \\ 16 & -8 & 8 & -32 \\ 6 & 1 & -25 & 44 \end{pmatrix} \cdot \begin{pmatrix} 29 & 115 & 71 & 104 & 132 & 82 \\ 27 & 39 & 19 & 136 & 168 & 22 \\ 9 & 53 & 49 & 24 & 36 & 54 \\ 2 & 31 & 25 & 24 & 3 & 19 \end{pmatrix} \\ &= 1/64 \begin{pmatrix} 192 & 0 & 0 & 896 & 1344 & 64 \\ 64 & 216 & 1280 & 320 & 0 & 1216 \\ 256 & 960 & 576 & 0 & 960 & 960 \\ 64 & 768 & 320 & 1216 & 192 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 & 14 & 21 & 1 \\ 1 & 19 & 20 & 5 & 0 & 19 \\ 4 & 15 & 9 & 0 & 15 & 15 \\ 1 & 12 & 5 & 19 & 3 & 0 \end{pmatrix} \end{aligned}$$

El mensaje original se obtiene al escribir en forma lineal las columnas de la última matriz:

3 1 4 1 0 19 15 12 0 20 9 5 14 5 0 19 21 0 15 3 1 19 15
C A D A S O L T I E N E S U O C A S O

2.8.2. Seriación en Arqueología

El 05 de Noviembre de 1922 fue descubierta la tumba del Rey Tutankamón por el arqueólogo británico Howard Carter; sin embargo, el 05 de Noviembre del presente año fue exhibida por primera vez ante el

público, 85 años después de que la tumba del niño faraón fuese descubierta en el Valle de los Reyes de la ciudad de Luxor, según informó la cadena británica BBC. Por estos motivos es que resulta tan interesante descubrir las datas de las tumbas y sus contenidos. Esto nos motiva a presentar la siguiente aplicación sobre fechado secuencial.

1. **La siguiente matriz describe los objetos arqueológicos encontrados en varias tumbas. Determine el orden cronológico de las tumbas y de los objetos arqueológicos.**

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

Esta matriz A cuyos elementos son 1 ó 0; describe el contenido arqueológico de las tumbas. Numerando las tumbas con 1, 2, 3, 4 y los tipos de objetos arqueológicos con 1, 2, 3, 4. La matriz A se define:

$$a_{ij} = \begin{cases} 1 & \text{si la tumba } i \text{ contiene objetos arqueológicos del tipo } j; \\ 0 & \text{si la tumba } i \text{ no contiene objetos arqueológicos del tipo } j. \end{cases}$$

La matriz A contiene toda la información acerca del contenido arqueológico de las diversas tumbas. El siguiente enunciado dice cómo obtener información a partir de la matriz A.

El elemento g_{ij} de la matriz $G = A \cdot A^t$ es igual al número de tipos de objetos arqueológicos comunes a la tumba i y a la tumba j.

De manera que entre mayor sea g_{ij} más cercanas, cronológicamente, se encontrarán las tumbas i y j. Examinando los elementos de G, el arqueólogo puede determinar el orden cronológico de las tumbas. Se comprueba esto

$$\begin{aligned} g_{ij} &= \text{elemento de la fila } i, \text{ columna } j \text{ de } G \\ &= (\text{fila } i \text{ de } A) \cdot (\text{columna } j \text{ de } A^t) \\ &= \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{jn} \end{pmatrix} \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \dots + a_{in}a_{jn} \end{aligned}$$

Cada término en esta suma será 1 ó 0.

Observación 2.2 *La matriz $P = A^t \cdot A$ lleva de una manera análoga la información acerca del fechado secuencial de los objetos arqueológicos.*

Supuesto Arqueológico: Se supone que entre mayor sea el número de tumbas en las que se encuentran dos tipos de objetos arqueológicos, más cercanos serán estos cronológicamente.

El elemento p_{ij} de la matriz $P = A^t \cdot A$ da el número de tumbas en las que se encuentran objetos arqueológicos tanto del tipo i, como del tipo j. Así entre mayor sea p_{ij} , los objetos arqueológicos de tipo i y de tipo j estarán cronológicamente más cercanos. Examinado los elementos de P se ordenan cronológicamente los objetos arqueológicos.

Se puede demostrar matemáticamente que las matrices $G(= A \cdot A^t)$ y $P(= A^t \cdot A)$ son matrices simétricas. Esta simetría implica que toda la información está contenida en los elementos sobre la diagonal principal de las matrices. En los elementos debajo de la diagonal principal se duplica la información.

En nuestro caso: $a_{11} = 1$ implica que la tumba 1 contiene objetos arqueológicos de tipo 1.

$a_{24} = 0$ implica que la tumba 2 no contiene objetos arqueológicos de tipo 4.

$$G = A \cdot A^t = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}$$

Observe que G es simétrica; la información contenida en los elementos encima de la diagonal principal se duplican en los elementos debajo de ella. Analizando los elementos encima de la diagonal principal:

$g_{12} = 1$: las tumbas 1 y 2 tienen un tipo de objetos arqueológicos en común

$g_{13} = 0$: las tumbas 1 y 3 no tienen ningún tipo de objetos arqueológicos en común

$g_{14} = 1$: las tumbas 1 y 4 tienen un tipo de objetos arqueológicos en común

$g_{23} = 0$: las tumbas 2 y 3 no tienen ningún tipo de objetos arqueológicos en común

$g_{24} = 0$: las tumbas 2 y 4 no tienen ningún tipo de objetos arqueológicos en común

$g_{34} = 1$: las tumbas 3 y 4 tienen un tipo de objetos arqueológicos en común

Las tumbas 1 y 2 tienen objetos arqueológicos en común; por lo tanto, son cercanas cronológicamente. Se inicia con las tumbas 1 y 2 y se construye un diagrama.

1-2

Se agrega la tumba 4 a este diagrama, $g_{14} = 1$; mientras que $g_{24} = 0$. La tumba 4 está cercana a la tumba 1, pero no a la tumba 2.

4-1-2

Finalmente se agrega la tumba 3 al diagrama, $g_{34} = 1$; mientras que $g_{13} = 0$ y $g_{23} = 0$. La tumba 3 está cercana a la tumba 4 pero no a la tumba 1 y 2. Tenemos:

3-4-1-2

Hay dos posibilidades:

$3 \rightarrow 4 \rightarrow 1 \rightarrow 2$

o

$3 \leftarrow 4 \leftarrow 1 \leftarrow 2$

Los arqueólogos saben por otras fuentes cuál de las dos tumbas de los extremos (3 y 2) es la primera. De esta manera se conoce el orden cronológico de las tumbas.

2.8.3. Relaciones de Grupo en Sociología

En las ramas de las ciencias físicas, sociales y matemáticas se usan modelos de la teoría de grafos para analizar relaciones. Es por esto que en las siguientes aplicaciones se utilizará la teoría de grafos para resolver este tipo de problemas; pero antes definiremos dígrafo y matriz de adyacencia.

Definición 2.7 Un dígrafo es una colección finita de vértices P_1, P_2, \dots, P_n , junto con arcos dirigidos que unen algunos pares de vértices. Un camino es una sucesión de arcos que le permiten a uno pasar de forma continua de un vértice a otro. La longitud de un camino es el número de sus arcos. A un camino de longitud N se le llama un camino n .

Definición 2.8 Considere un dígrafo con vértices P_1, \dots, P_n . La matriz de adyacencia del dígrafo es tal que

$$a_{ij} = \begin{cases} 1 & \text{si hay un arco del vértice } P_i \text{ al } P_j ; \\ 0 & \text{de otra manera.} \end{cases}$$

Para entender mayormente los conceptos, vamos a considerar 4 paraderos de micros P_1, P_2, P_3 y P_4 , algunas de estos paraderos están unidos por una comunicación unidireccional y otros bidireccionalmente. En el siguiente esquema se representa lo expuesto anteriormente:

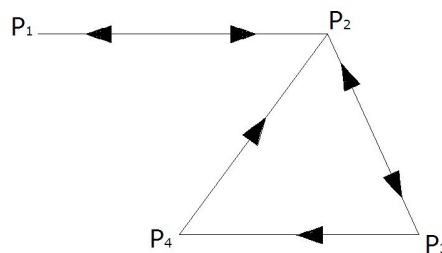


Figura 2.1: Paraderos de micros P_1, P_2, P_3 y P_4

Como podemos ver hay paraderos intermedios, y otros que son de comunicación directa; las flechas dan la dirección de esos vínculos. P_1 y P_2 tienen comunicación bidireccional, mientras que P_4 tiene dirección unidireccional con P_2 . Por el paradero P_4 sólo pueden pasar micros a P_2 . De esta forma podemos vincular los paraderos, con su debida orientación, es decir:

$$P_1 \leftrightarrow P_2 \leftrightarrow P_3 \rightarrow P_4 \rightarrow P_2 \leftrightarrow P_1$$

$$P_2 \leftrightarrow P_3 \rightarrow P_4 \rightarrow P_2 \rightarrow P_1$$

$$P_3 \rightarrow P_4 \rightarrow P_2 \rightarrow P_1$$

$$P_4 \rightarrow P_2 \rightarrow P_1$$

A través de esto podemos dar un ejemplo de un dígrafo, el interés de la aplicación radica en que, si queremos enviar un mensaje de un paradero a otro, deberíamos saber el recorrido de las micros para que el mensaje llegase por el camino más conveniente.

Considerando el ejemplo anterior construiremos la llamada matriz de adyacencia.

La matriz de comunicación entre los paraderos es la siguiente:

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

De esta forma la matriz de comunicación anterior es la matriz de adyacencia del esquema.

Ahora si queremos enviar un mensaje de un paradero a otro de la manera más rápida posible, tendríamos que ubicar en que dirección se encuentra cada paradero, observando la figura resulta sencillo elegir el camino, pero para 40 paraderos con distintas direcciones resultaría complejo. Por ello enunciaremos el siguiente teorema.

Teorema 2.6 Si A es la matriz de adyacencia de un dígrafo, sea $(a_{ij})^{(m)}$ el elemento de la fila i , columna j de A^m .

$$\text{El número de caminos } m \text{ de } P_i \text{ a } P_j = (a_{ij})^{(m)}$$

Es decir, si tenemos la matriz anterior de adyacencia, podemos determinar sus potencias sucesivas.

En nuestro ejemplo:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 2 & 1 & 2 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 4 & 1 & 2 \\ 2 & 3 & 2 & 1 \\ 2 & 1 & 2 & 0 \end{bmatrix} \quad A^5 = \begin{bmatrix} 0 & 4 & 0 & 2 \\ 4 & 4 & 1 & 1 \\ 3 & 5 & 3 & 2 \\ 1 & 4 & 1 & 2 \end{bmatrix} \dots$$

Definición 2.9 (Distancia de un dígrafo) La distancia de un vértice a otro en un dígrafo es la longitud del camino más corto desde ese vértice al otro. Si no hay camino de un vértice al otro, se dice que la distancia está indefinida. Las distancias entre los diferentes vértices de un dígrafo forman los elementos de una matriz. La matriz distancia D está definida como sigue:

$$d_{ij} = \begin{cases} \text{Número de arcos en el camino más corto del vértice } P_i \text{ al vértice } P_j ; \\ 0 & i = j ; \\ x & \text{si no hay camino } P_i \text{ a } P_j . \end{cases}$$

De este modo, podemos construir la matriz distancia de matrices más grandes, en el caso que tengamos más vértices se emplean las potencias de la matriz de adyacencia. En el caso de los paraderos la matriz de adyacencia es:

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

Con esta información podemos resolver aplicaciones en relaciones de grupo en sociología.

Ejemplo 2.2 Si consideramos a 5 personas en un salón de clases. A un sociólogo le interesa encontrar cuál de las cinco personas tiene más influencia sobre las otras. Se le pide al grupo llenar el siguiente cuestionario.

Nombre:

¿Cuál es la persona que cree usted es líder del grupo?

Los datos recabados se resumen en la siguiente tabla:

Tabla N°1

Miembro del Grupo	Persona Líder
M_1	M_5
M_2	M_1
M_3	M_2
M_4	M_2
M_5	M_4

El psicólogo supone que una persona tiene influencia sobre otros miembros del grupo, considerándolo como un posible líder. De manera que la tabla anterior puede representarse mediante un dígrafo. Los miembros del grupo representan vértices y la influencia directa se observa mediante arcos. Siendo la dirección de la influencia, la dirección del arco. A continuación la representaremos en forma de dígrafo.

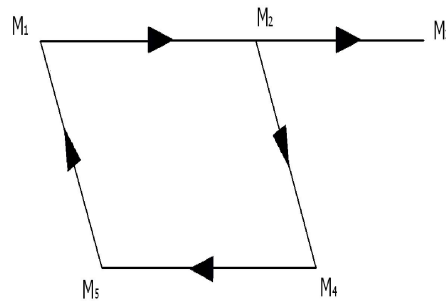


Figura 2.2: Representación gráfica de la Tabla N°1

Luego, la matriz distancia del dígrafo es la siguiente:

$$\begin{matrix}
 & M_1 & M_2 & M_3 & M_4 & M_5 & \\
 M_1 & \left(\begin{matrix} 0 & 1 & 2 & 2 & 3 \\ 3 & 0 & 1 & 1 & 2 \\ x & x & 0 & x & x \\ 2 & 3 & 4 & 0 & 1 \\ 1 & 2 & 3 & 3 & 0 \end{matrix} \right) & 8 \\
 M_2 & & & & & & 7 \\
 M_3 & & & & & & 4x \\
 M_4 & & & & & & 10 \\
 M_5 & & & & & & 9
 \end{matrix}$$

En este grafo podemos observar que los arcos corresponden a influencia directa, es decir, que entre menor sea la distancia de M_i a M_j , mayor será la influencia que tiene M_i sobre M_j , o sea, entre menor sea la suma de la fila i , mayor es la influencia que M_i tiene sobre los otros miembros del grupo. Como la segunda fila tiene una suma de 7 implica que M_2 tiene mayor influencia sobre los otros individuos del grupo (M_1, M_3, M_4 y M_5).

2.8.4. Cadenas de Markov

En el estudio de los fenómenos aleatorios en los que los resultados no se saben con exactitud, se puede elaborar una probabilidad que ayuda a determinar el suceso. Aquí el uso de las llamadas matrices estocásticas es imprescindible para obtener la probabilidad de las variables que intervienen en alguna aplicación o ejercicio.

Hemos de recordar entonces algunas ideas básicas de la probabilidad que serán de utilidad.

- Si un resultado de un evento es seguro que ocurra, se dice que la probabilidad de que ocurra es 1
- Si no ocurrirá, se dice que la probabilidad es 0.
- Otras probabilidades se representan entre 0 y 1
- Entre mayor sea la fracción, mayor será la probabilidad p de que ese resultado ocurra, es decir, $0 \leq p \leq 1$.
- Si n resultados completamente independientes tienen todos la misma probabilidad de ocurrir, y si interesan m de estos resultados, entonces la probabilidad p de que ocurra uno de estos m resultados es m/n .

Ejemplo 2.3 *Suponga que el flujo probable de los votos por los partidos 1, 2 y 3 de las elecciones del año 2000 al 2001, primera elección, está dado por el diagrama siguiente:*

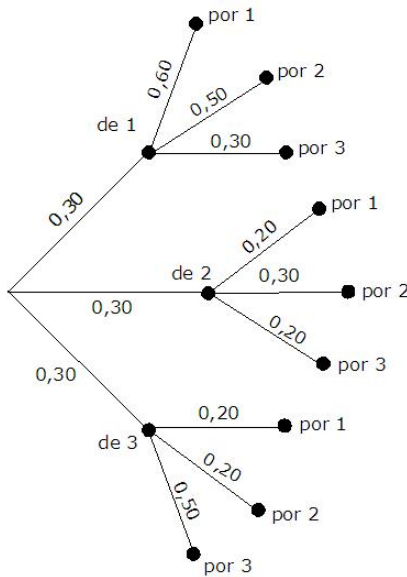


Figura 2.3: Diagrama de transición entre estados

El diagrama de transición entre estados indica, por ejemplo que el 30 % de las personas que pertenecían al partido 2 en el 2000 votarán de nuevo por el partido 2 en el 2001, mientras que el 50 % votará por el partido 1 y el 20 % por el partido 3. El resto del diagrama es fácil de interpretar. Esta información puede representarse en la matriz estocástica:

$$T = \begin{matrix} & \begin{matrix} \text{de 1} & \text{de 2} & \text{de 3} \end{matrix} \\ \begin{pmatrix} 0,60 & 0,50 & 0,30 \\ 0,20 & 0,30 & 0,20 \\ 0,20 & 0,20 & 0,50 \end{pmatrix} & \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix} \end{matrix}$$

donde el espacio de estados es {de 1, de 2, de 3} y de votación {por 1, por 2, por 3}.

Por simplicidad asumiremos que en el siguiente período no contabilizaremos a los nuevos votantes. Estudiemos a la matriz de transición T^2 .

$$T^2 = \begin{matrix} & \begin{matrix} \text{de 1} & \text{de 2} & \text{de 3} \end{matrix} \\ \begin{pmatrix} 0,52 & 0,51 & 0,43 \\ 0,22 & 0,23 & 0,22 \\ 0,26 & 0,26 & 0,35 \end{pmatrix} & \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix} \end{matrix}$$

Es de nuevo una matriz estocástica.

Asumiremos que el número total de votantes N , permanece constante de año en año. Suponga que la distribución de votos en el año 2000 y la esperada en el 2001 está dada por:

	2000	2001
Por 1	x_0 %	x_1 %
Por 2	y_0 %	y_1 %
Por 3	z_0 %	z_1 %

En este caso:

$$\begin{aligned} x_1 &= 0,60x_0 + 0,50y_0 + 0,30z_0 \\ y_1 &= 0,20x_0 + 0,30y_0 + 0,20z_0 \\ z_1 &= 0,20x_0 + 0,20y_0 + 0,50z_0 \end{aligned}$$

Es decir que: $T \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$

Si la matriz T se utiliza para calcular la distribución porcentual del año 2002

	2002
Por 1	x_0 %
Por 2	y_0 %
Por 3	z_0 %

Tendríamos que:

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = T \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = T^2 \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

Asumiendo que esta matriz de transición se pudiese aplicar en años venideros, tendríamos que en n años, la distribución de los votos estaría dada por

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = T^n \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

En particular, la relación de los votantes con los resultados de las elecciones, está dado por:

$$\begin{matrix} T & T^2 & T^3 & T^n \\ 2000 \rightarrow 2001 & 2000 \rightarrow 2002 & 2000 \rightarrow 2003 & 2000 \rightarrow 200(n) \end{matrix}$$

Las matrices

$$T \qquad T^2 \qquad T^3 \dots\dots\dots T^n$$

son todas, matrices estocásticas.

Observación 2.3 *la suma de los elementos de cada columna es 1; es decir, 100 %.*

De la matriz T sacamos la siguiente información:

$$T = \begin{pmatrix} \text{de 1} & \text{de 2} & \text{de 3} \\ 0,60 & 0,50 & 0,30 \\ 0,20 & 0,30 & 0,20 \\ 0,20 & 0,20 & 0,50 \end{pmatrix} \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix}$$

- El 60 % de los votantes de 1 votarán por 1 en la siguiente elección
- El 20 % de los votantes de 1 votarán por 2 en la siguiente elección
- El 20 % de los votantes de 1 votarán por 3 en la siguiente elección
- El 50 % de los votantes de 2 votarán por 1 en la siguiente elección
- El 30 % de los votantes de 2 votarán por 2 en la siguiente elección
- El 20 % de los votantes de 2 votarán por 3 en la siguiente elección
- El 30 % de los votantes de 3 votarán por 1 en la siguiente elección
- El 20 % de los votantes de 3 votarán por 2 en la siguiente elección
- El 50 % de los votantes de 3 votarán por 3 en la siguiente elección

De T^2 concluimos que:

$$T^2 = \begin{pmatrix} \text{de 1} & \text{de 2} & \text{de 3} \\ 0,52 & 0,51 & 0,43 \\ 0,22 & 0,23 & 0,22 \\ 0,26 & 0,26 & 0,35 \end{pmatrix} \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix}$$

- El 52 % de los votantes originales de 1 votarán por 1 en la siguiente elección
- El 22 % de los votantes originales de 1 votarán por 2 en la siguiente elección
- El 26 % de los votantes originales de 1 votarán por 3 en la siguiente elección
- El 51 % de los votantes originales de 2 votarán por 1 en la siguiente elección
- El 23 % de los votantes originales de 2 votarán por 2 en la siguiente elección
- El 26 % de los votantes originales de 2 votarán por 3 en la siguiente elección
- El 43 % de los votantes originales de 3 votarán por 1 en la siguiente elección
- El 22 % de los votantes originales de 3 votarán por 2 en la siguiente elección

El 35 % de los votantes originales de 3 votarán por 3 en la siguiente elección

Por ende, si la matriz T es una matriz estocástica regular, T^n tiende a una matriz con todas sus columnas iguales a medida que n crece.

Con un programa computacional puede comprobarse que:

$$T^5 = \begin{matrix} & \begin{matrix} \text{de 1} & \text{de 2} & \text{de 3} \end{matrix} \\ \begin{pmatrix} 0,49276 & 0,49275 & 0,49033 \\ 0,22222 & 0,22223 & 0,22222 \\ 0,28502 & 0,28502 & 0,28745 \end{pmatrix} & \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix} \end{matrix}$$

y

$$T^{17} = \begin{matrix} & \begin{matrix} \text{de 1} & \text{de 2} & \text{de 3} \end{matrix} \\ \begin{pmatrix} 0,492063492 & 0,492063492 & 0,492063492 \\ 0,222222222 & 0,222222222 & 0,222222222 \\ 0,285714286 & 0,285714286 & 0,285714286 \end{pmatrix} & \begin{matrix} \text{por 1} \\ \text{por 2} \\ \text{por 3} \end{matrix} \end{matrix}$$

Por lo que T^{18} tiene la forma

$$T^{18} = \begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & \beta & \beta \\ \gamma & \gamma & \gamma \end{pmatrix}$$

Se concluye que

$$T^{19} = T^{18} \cdot T = \begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & \beta & \beta \\ \gamma & \gamma & \gamma \end{pmatrix} \cdot \begin{pmatrix} 0,60 & 0,50 & 0,30 \\ 0,20 & 0,30 & 0,20 \\ 0,20 & 0,20 & 0,50 \end{pmatrix} = \begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & \beta & \beta \\ \gamma & \gamma & \gamma \end{pmatrix}$$

Por lo tanto, a partir de $n=18$:

$$T^n \approx T^{n+1} \approx T^{n+2} \approx \dots$$

Por consiguiente:

Si las expectativas de votación (flujo de votantes de año en año), se mantuviesen igual por años, se concluiría que a partir de $n > 18$, los partidos se repartirían los votos de una manera que puede predeterminarse, quedando el partido 1 con α votantes, el partido 2 con β votantes y el partido 3 con γ votantes.

La matriz T^n con n suficientemente grande, sirve para **predecir** como terminarán las cosas si la situación planteada por la matriz de transición T , se mantiene igual por n años.

Por supuesto, en realidad la matriz de transición cambia de año en año, y a la final, la matriz T_n que relacionaría el flujo de votantes entre los partidos a partir del censo inicial hasta la n -ésima votación, se calcularía así:

$$\begin{aligned} T^{(n)} &= T^{(n-1)} \cdot T_n \\ &= T^{(n-2)} \cdot T_{n-1} \cdot T_n \\ &= T^{(n-3)} \cdot T_{n-2} \cdot T_{n-1} \cdot T_n \\ &= T_1 \cdot T_2 \cdot T_3 \cdot T_4 \cdot \dots \cdot T_n \end{aligned}$$

Cada matriz T_k es estocástica y en particular lo es $T^{(n)}$.

El vector $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ se denomina el **vector de estado inicial**, cada uno de los vectores $X_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$ es también un vector de estado. Las siguientes ecuaciones son válidas:

$$X_1 = T \cdot X_0; \quad n \geq 1, \quad X_n = T \cdot X_{n-1}; \quad n \geq 1, \quad X_n = T^{(n)} \cdot X_0$$

2.9. Ejercicios Propuestos

1. Codifica el mensaje **soy un genio** usando la matriz

$$A = \begin{pmatrix} 2 & -1 \\ 4 & 2 \end{pmatrix}$$

matriz codificadora.

2. Decodifique el mensaje 35, 46, -8, 65, -24, 34, 73, -43, 45, 35, -30, 30, 27, 21, 3, 18, -6, 12, 52, 38, 0, 44, 47, -5, 56, -34, 36, codificado, utilice la matriz

$$B^{-1} = \begin{pmatrix} 1 & 5 & 4 \\ 0 & -1 & -1 \\ -1 & -9 & -7 \end{pmatrix}$$

donde B^{-1} es la matriz decodificadora.

3. La siguiente matriz describe los objetos arqueológicos encontrados en varias tumbas. Determine el orden cronológico de las tumbas y de los objetos arqueológicos.

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

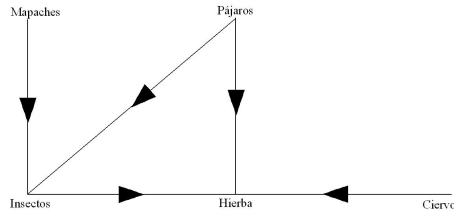
4. El modelo presentado para arqueología se usa también en sociología para analizar las relaciones entre las personas de un grupo. Consideremos la relación "conocidos" entre las personas de un grupo. Supongamos que la relación "conocidos" es recíproca. Numere a las personas con $1, 2, \dots, n$ y defina una matriz como sigue:

$$a_{ij} = \begin{cases} 1, & \text{si } i \text{ y } j \text{ son conocidos;} \\ 0, & \text{si } i \text{ y } j \text{ no se conocen.} \end{cases}$$

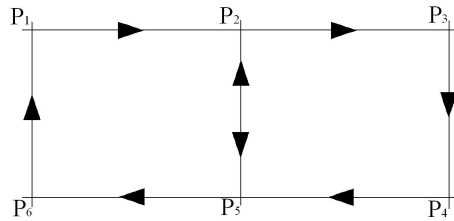
a) Demuestre que si $P = A \cdot A^t$, entonces p_{ij} es el número de conocidos en común de i y j .

b) Suponga que la relación "conocidos" no es recíproca para todos. ¿Cómo afecta esto al modelo?

5. La teoría de grafos se aplica en modelos matemáticos para entender el delicado balance de la naturaleza. La siguiente figura da el dígrafo que describe la red alimenticia de una comunidad ecológica. Determine la matriz de adyacencia.



6. La red describe un sistema de avenida en el centro de una ciudad. Muchas de las avenidas son de un sólo sentido. Interprete la red como un dígrafo, dé la matriz de adyacencia y la matriz distancia.



7. Las tablas siguientes representan la información obtenida de cuestionarios dados a grupos de personas. En cada caso construya el dígrafo que describe la relación de liderazgo dentro del grupo. Ordene a los miembros de acuerdo con su influencia en el grupo.

Miembro del grupo	Persona cuya opinión considera más valiosa
M_1	M_5
M_2	M_1
M_3	M_2
M_4	M_3
M_5	M_4

8. Las siguientes matrices estocásticas dan probabilidades de transición ocupacionales. Responda las siguientes preguntas que se hacen de cada matriz.

$$\begin{matrix}
 & \text{Profesional} & \text{No Profesional} \\
 \begin{pmatrix} 1 & 0,2 \\ 0 & 0,8 \end{pmatrix} & \text{Profesional} \\
 & \text{No Profesional}
 \end{matrix}$$

- a) Si el padre es un trabajador no profesional, ¿Cuál es la probabilidad de que el hijo sea un trabajador profesional?
- b) Si hay 10.000 en la categoría de profesional y 20.000 en la categoría no profesional, ¿Cuál será la distribución en la generación siguiente?
9. Construya un modelo de Movimineto de población entre las ciudades, los alrededores y las áreas no metropolitanas de Estados Unidos. Sus poblaciones respectivas son 58 millones, 142 millones y 60 millones. La matriz estocástica que da las probabilidades de los movimientos es

$$\begin{array}{c}
 \begin{array}{ccc}
 & \begin{array}{c} (de) \\ ciudad \quad alrededor \quad no metropolitana \end{array} & \\
 \left(\begin{array}{ccc}
 0,52 & 0,51 & 0,43 \\
 0,22 & 0,23 & 0,22 \\
 0,26 & 0,26 & 0,35
 \end{array} \right. & \begin{array}{c} (a) \\ ciudad \\ alrededor \\ no metropolitana \end{array}
 \end{array}
 \end{array}$$

Este modelo es un refinamiento del modelo anterior en el sentido de que la población metropolitana se divide en ciudad y alrededores.

Prediga las poblaciones de la ciudad, de los alrededores y de las áreas y de las no metropolitanas para el año 2001 y 2002. ¿Cuál es la probabilidad de que una persona que vive en la ciudad en el año 2000 viva en un área no metropolitana en el año 2002?

CAPÍTULO 3

Espacios Vectoriales

3.1. Introducción

En este tercer capítulo se definirá Espacio Vectorial \mathbb{V} , se enunciarán sus propiedades; además se definirá cuando un conjunto de vectores $\{v_1, v_2, \dots, v_n\} \in \mathbb{V}$ es Linealmente Dependiente o Linealmente Independiente y ¿Qué condiciones debe cumplir para ser una Base de \mathbb{V} ?. También se definirá cuando una función es una Transformación Lineal.

En la sección (3.6) se enunciarán las condiciones para ver que un conjunto de vectores $\{v_1, v_2, \dots, v_n\} \in \mathbb{V}$ es Ortonormal; en la misma sección se enunciarán algunos teoremas referidos a proyección ortogonal, lo que nos lleva a un resultado importante: *la Aproximación por Mínimos Cuadrados a una serie de datos.*

3.2. ¿Qué es un Espacio Vectorial?

Definición 3.1 Un Espacio Vectorial \mathbb{V} sobre un cuerpo \mathbb{F} se dice un $\mathbb{V}(\mathbb{F})$ espacio vectorial, si $(\mathbb{V}, +)$ es un grupo abeliano y dados $x, y \in \mathbb{V}$ y α, β escalares pertenecientes a \mathbb{F} se cumplen las siguientes propiedades:

1. $\forall x \in \mathbb{V}$ y $\alpha \in \mathbb{F}$, se tiene $\alpha \cdot x \in \mathbb{V}$ (cerradura bajo la multiplicación por un escalar)
2. $\forall x, y \in \mathbb{V}$ y $\alpha \in \mathbb{F}$, se tiene $\alpha(x + y) = \alpha \cdot x + \alpha \cdot y$ (ley distributiva)
3. $\forall x \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{F}$, se tiene $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot y$ (ley distributiva)
4. $\forall x \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{F}$, se tiene $\alpha \cdot (\beta \cdot x) = (\alpha \cdot \beta) \cdot x$ (ley asociativa de la multiplicación por escalares)
5. Para cada vector $x \in \mathbb{V}$, $1 \cdot x = x$

Observación 3.1 Se define $(\mathbb{V}, +)$ grupo abeliano si $\forall x, y, z \in \mathbb{V}$ se cumplen las siguientes propiedades:

1. Cerradura bajo la operación previamente definida $x + y \in \mathbb{V}$
2. Asociatividad $(x + y) + z = x + (y + z)$
3. Existencia de Neutro; existe un vector $0 \in \mathbb{V}$ tal que $x + 0 = 0 + x = x$
4. Inverso Aditivo; $\forall x \in \mathbb{V} \exists! -x \in \mathbb{V}$ tal que $x + (-x) = 0$
5. Conmutatividad $x + y = y + x$

3.3. Vectores Linealmente Independientes y Vectores Linealmente Dependientes

Sean v_1, v_2, \dots, v_n , n vectores en un Espacio Vectorial \mathbb{V} . Se dice que estos vectores son linealmente dependientes si existen escalares $\alpha_1, \alpha_2, \dots, \alpha_n$ no todos iguales a cero simultáneamente, tales que,

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

Los vectores que no son linealmente dependientes, se llaman linealmente independientes.

Siendo los vectores v_1, v_2, \dots, v_n linealmente dependientes $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$ y siendo por ejemplo $\alpha_n \neq 0$, se tiene

$$v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \frac{\alpha_2}{\alpha_n} v_2 - \dots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1}$$

Si tiene lugar la última igualdad se dice que el vector v_n es una combinación lineal de los vectores v_1, v_2, \dots, v_{n-1} . Luego se puede concluir que siendo los vectores v_1, v_2, \dots, v_n linealmente dependientes, al menos uno de ellos se expresa linealmente en términos de los restantes.

Ejemplo 3.1 Los vectores $(3, 4), (2, 1), (-1, 2)$ son linealmente dependientes, puesto que:

$$(-1, 2) = 1 \cdot (3, 4) + (-2) \cdot (2, 1)$$

Ejemplo 3.2 Los vectores $(1, 1)$ y $(-1, 1)$ son linealmente independientes, puesto que $\nexists \alpha \in \mathbb{R}$ tal que

$$(1, 1) = \alpha \cdot (-1, 1)$$

3.4. ¿Qué es una Base de un Espacio Vectorial \mathbb{V} ?

Definición 3.2 Sea \mathbb{V} un Espacio Vectorial y $S = \{v_1, v_2, \dots, v_n\}$ un conjunto de vectores en \mathbb{V} . Entonces:

$$\begin{aligned} \langle S \rangle &= \langle \{v_1, v_2, \dots, v_n\} \rangle \\ &= \{ \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n / \alpha_i \in \mathbb{F} \} \end{aligned}$$

es un Subespacio de \mathbb{V} , llamado el engendrado o generado por S .

Además cualquier vector $\in S$ se puede escribir como combinación lineal de los vectores de S .

Definición 3.3 Un conjunto finito de vectores $\{v_1, v_2, \dots, v_n\}$ es una base para un Espacio Vectorial \mathbb{V} si

- $\{v_1, v_2, \dots, v_n\}$ es linealmente independiente
- $\{v_1, v_2, \dots, v_n\}$ genera a \mathbb{V} .

Ejemplo 3.3 Sea $A = \{M_{2 \times 2} \mid A = -A^t\}$.

Una base de este subespacio es:

$$\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

Observación 3.2 Queda al lector demostrar que efectivamente A así definida es un Subespacio Vectorial de $M_{2 \times 2}$.

3.5. Transformaciones Lineales

Definición 3.4 Sean \mathbb{V} y \mathbb{W} espacios vectoriales reales. Una transformación lineal T de \mathbb{V} en \mathbb{W} es una función que satisface, para cada u, v en \mathbb{V} y cada escalar α :

- $T(u + v) = Tu + Tv$
- $T(\alpha v) = \alpha Tv$

Ejemplo 3.4 Sea $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $L(x, y) = (x, -y)$.

Es fácil para el lector demostrar que efectivamente L es una transformación lineal.

Teorema 3.1 Sea $T: \mathbb{V} \rightarrow \mathbb{W}$ una transformación lineal. Entonces para todos los vectores $u, v, v_1, v_2, \dots, v_n$ en \mathbb{V} y todos los escalares $\alpha_1 \alpha_2 \dots \alpha_n$

- $T(0) = 0$
- $T(u - v) = Tu - Tv$
- $T(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = \alpha_1 Tv_1 + \alpha_2 Tv_2 + \dots + \alpha_n Tv_n$

Ejemplo 3.5 Tomando en consideración el ejemplo anterior con $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$L(x, y) = (x, -y)$, verificaremos que efectivamente:

$$L(x, y) = xL(1, 0) + yL(0, 1)$$

En efecto,

$$\begin{aligned} L(x, y) &= L(x \cdot (1, 0) + y \cdot (0, 1)) \\ &= x \cdot L(1, 0) + y \cdot L(0, 1) \\ &= x \cdot (1, 0) + y \cdot (0, -1) \\ &= (x, -y) \end{aligned}$$

3.6. Bases Ortonormales y Proyecciones en \mathbb{R}^n

Definición 3.5 (Conjunto Ortonormal en \mathbb{R}^n) Se dice que un conjunto de vectores $S = \{u_1, u_2, \dots, u_k\}$ en \mathbb{R}^n es un conjunto ortonormal si:

$$\begin{aligned} u_i \cdot u_j &= 0 & \text{si } i \neq j \\ u_i \cdot u_i &= 1 \end{aligned}$$

Si solo satisface la primera condición, se dice que el conjunto es ortogonal.

Ejemplo 3.6 En \mathbb{R}^2 la base canónica es $\{(1, 0), (0, 1)\}$. Verifiquemos que esta base es un conjunto ortonormal:

En efecto

$$\begin{aligned} (1, 0) \cdot (0, 1) &= 1 \cdot 0 + 0 \cdot 1 \\ &= 0 \\ (1, 0) \cdot (1, 0) &= 1 \\ &= (0, 1) \cdot (0, 1) \end{aligned}$$

como satisface las dos condiciones mencionadas en la definición, entonces $\{(1, 0), (0, 1)\}$ es ortonormal.

Definición 3.6 Sea \mathbb{V} un espacio vectorial sobre \mathbb{R} (ó \mathbb{C}). Una norma vectorial en \mathbb{V} , denotada por $\|\cdot\|$, $\|\cdot\|_v$ es una función

$$\|\cdot\|_v : \mathbb{V} \rightarrow [0, +\infty[$$

que satisface las siguientes propiedades:

1. $\forall \vec{v} \in \mathbb{V} \quad \|\vec{v}\| \geq 0$ y $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$
2. $\forall \vec{v} \in \mathbb{V}$ y $\forall \alpha \in \mathbb{R} \quad \|\alpha \vec{v}\| = |\alpha| \cdot \|\vec{v}\|$
3. $\forall \vec{v}, \vec{w} \in \mathbb{V}$ se tiene que: $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$

Se define la norma euclidiana: $\|\vec{x}\|_2 = [x_1^2 + x_2^2 + \dots + x_n^2]^{\frac{1}{2}}$

Teorema 3.2 Si $S = \{v_1, v_2, \dots, v_k\}$ es un conjunto ortogonal de vectores diferentes de cero, entonces S es linealmente independiente.

Demostración:

Supongamos que $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$ debemos demostrar que $\alpha_i = 0$ para todo $i = 1, 2, \dots, k$

$$\begin{aligned} \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k &= 0 \\ (\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k) \cdot v_j &= 0 \cdot v_j \\ (\alpha_1 v_1) \cdot v_j + (\alpha_2 v_2) \cdot v_j + \dots + (\alpha_j v_j) \cdot v_j + \dots + (\alpha_k v_k) \cdot v_j &= 0 \\ \alpha_1 (v_1 \cdot v_j) + \alpha_2 (v_2 \cdot v_j) + \dots + \alpha_j (v_j \cdot v_j) + \dots + \alpha_k (v_k \cdot v_j) &= 0 \\ \alpha_1 \cdot 0 + \alpha_2 \cdot 0 + \dots + \alpha_j (v_j \cdot v_j) + \dots + \alpha_k \cdot 0 &= 0 \\ \alpha_1 \cdot 0 + \alpha_2 \cdot 0 + \dots + \alpha_j \cdot 1 + \dots + \alpha_k \cdot 0 &= 0 \\ 0 + 0 + \dots + \alpha_j + \dots + 0 &= 0 \\ \Rightarrow \alpha_j &= 0 \end{aligned}$$

esto es cierto para $j = 1, 2, \dots, k$ lo que completa la demostración.

Definición 3.7 (Matriz Ortogonal) Una matriz Q de $n \times n$ se llama ortogonal si Q es invertible y

$$Q^{-1} = Q^t$$

donde Q^t es la transpuesta de Q

Teorema 3.3 La matriz Q de $n \times n$ es ortogonal \Leftrightarrow las columnas de Q forman una base ortonormal para \mathbb{R}^n .

Demostración:

$$\text{Sea } Q = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \text{ entonces } Q^t = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

Así:

$$\begin{aligned} Q^t \cdot Q &= \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}a_{11} + a_{21}a_{21} + \dots + a_{n1}a_{n1} & \dots & a_{11}a_{1n} + a_{21}a_{2n} + \dots + a_{n1}a_{nn} \\ \vdots & & \vdots \\ a_{1n}a_{11} + a_{2n}a_{21} + \dots + a_{nn}a_{n1} & \dots & a_{1n}a_{1n} + a_{2n}a_{2n} + \dots + a_{nn}a_{nn} \end{pmatrix} \\ \Rightarrow Q^t \cdot Q &= [c_{ij}] \end{aligned}$$

con i, j variando de 1 hasta n donde $c_{ij} = a_{1i}a_{1j} + a_{2i}a_{2j} + \dots + a_{ni}a_{nj}$

Si las columnas de Q son ortonormales, implica que

$$c_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Por lo tanto $C = I_n$ y así queda demostrado que Q es ortogonal.

Por otro lado si $Q^{-1} = Q^t$ entonces $C = I_n$; esto nos quiere decir que:

$$c_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

y por lo tanto las columnas de Q son ortonormales.

Queda demostrado el teorema.

Definición 3.8 : Sean \vec{v} y $\vec{w} \in \mathbb{R}^n$, se define el producto interno:

$$(x, y) = xy^t = x_1y_1 + x_2y_2 + \dots + x_ny_n$$

Definición 3.9 Sea \mathbb{H} un subespacio de \mathbb{R}^n con base ortonormal $\{u_1, u_2, \dots, u_k\}$. Si $v \in \mathbb{R}^n$, entonces la proyección ortogonal de v sobre \mathbb{H} , denotada por $Proy_{\mathbb{H}}v$, está dada por:

$$Proy_{\mathbb{H}}v = (v \cdot u_1) \cdot u_1 + (v \cdot u_2) \cdot u_2 + \dots + (v \cdot u_k) \cdot u_k$$

Observación 3.3 *Notar que $\text{Proy}_{\mathbb{H}}v \in \mathbb{H}$.*

Definición 3.10 *Sea \mathbb{H} un subespacio de \mathbb{R}^n . El complemento ortogonal de \mathbb{H} , denotado por \mathbb{H}^\perp , está dado por:*

$$\mathbb{H}^\perp = \{x \in \mathbb{R}^n : x \cdot h = 0 \text{ para toda } h \in \mathbb{H}\}$$

Teorema 3.4 *Si \mathbb{H} es un subespacio de \mathbb{R}^n , entonces:*

1. \mathbb{H}^\perp es un subespacio de \mathbb{R}^n
2. $\mathbb{H} \cap \mathbb{H}^\perp = 0$
3. $\dim \mathbb{H}^\perp = n - \dim \mathbb{H}$

Demostración (i):

Sean $x, y \in \mathbb{H}^\perp$ y $h \in \mathbb{H}$. Debemos demostrar que $x + y \in \mathbb{H}^\perp$.
Por hipótesis:

$$\begin{aligned} x \cdot h &= 0 \\ y \cdot h &= 0 \\ (x + y) \cdot h &= x \cdot h + y \cdot h \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

$$\therefore x + y \in \mathbb{H}^\perp$$

Sea $x \in \mathbb{H}^\perp$ y $\alpha \in \mathbb{R}^n$. Debemos demostrar que $\alpha x \in \mathbb{H}^\perp$.
Por hipótesis

$$\begin{aligned} x \cdot h &= 0 \\ (\alpha x) \cdot h &= \alpha \cdot (x \cdot h) \\ &= \alpha \cdot 0 \\ &= 0 \end{aligned}$$

$$\therefore \alpha x \in \mathbb{H}^\perp$$

Demostración (ii):

Sea $z \in \mathbb{H} \cap \mathbb{H}^\perp$. Debemos demostrar que $z = 0$.
Por hipótesis $z \in \mathbb{H}$ y $z \in \mathbb{H}^\perp$. Eso implica que $z \cdot z = 0$ y

$$\therefore z = 0$$

lo que demuestra que

$$\mathbb{H} \cap \mathbb{H}^\perp = 0.$$

Observación 3.4 *La demostración de (iii) se ha omitido, pues resulta innecesaria ya que no se utilizará a futuro.*

Teorema 3.5 (Teorema de Proyección) Sea \mathbb{H} un subespacio de \mathbb{R}^n y sea $v \in \mathbb{R}^n$. Entonces existe un par único de vectores h y p tales que $h \in \mathbb{H}$, $p \in \mathbb{H}^\perp$ y $v = h + p$. En particular,

$$h = \text{Proy}_{\mathbb{H}} v \quad \text{y} \quad p = \text{Proy}_{\mathbb{H}^\perp} v$$

de manera que:

$$\begin{aligned} v &= h + p \\ &= \text{Proy}_{\mathbb{H}} v + \text{Proy}_{\mathbb{H}^\perp} v \end{aligned}$$

Demostración:

Sea $h = \text{Proy}_{\mathbb{H}} v \Rightarrow h \in \mathbb{H}$ y sea $p = v - h$. Debemos demostrar que $p \in \mathbb{H}^\perp$.

Sea $\{u_1, u_2, \dots, u_k\}$ una base ortonormal para \mathbb{H} . Entonces

$$h = (v \cdot u_1) \cdot u_1 + (v \cdot u_2) \cdot u_2 + \dots + (v \cdot u_k) \cdot u_k$$

Por otro lado, sea x un vector en \mathbb{H} , entonces existen α_i tales que:

$$x = \alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \dots + \alpha_k \cdot u_k$$

$$\Rightarrow p \cdot x = [v - (v \cdot u_1) \cdot u_1 - (v \cdot u_2) \cdot u_2 - \dots - (v \cdot u_k) \cdot u_k] \cdot [\alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 + \dots + \alpha_k \cdot u_k]$$

Como $\{u_1, u_2, \dots, u_k\}$ es un conjunto ortonormal

$$\Rightarrow u_i \cdot u_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

$$\therefore p \cdot x = \sum_{i=1}^k \alpha_i (v \cdot u_i) - \sum_{i=1}^k \alpha_i (v \cdot u_i) = 0$$

Así $p \cdot x = 0 \quad \forall x \in \mathbb{H}$, lo que implica que $p \in \mathbb{H}^\perp$. Para demostrar que $p = \text{Proy}_{\mathbb{H}^\perp} v$ se amplía a una base ortonormal en \mathbb{R}^n : $\{u_1, u_2, \dots, u_k, u_{k+1}, \dots, u_n\}$. Entonces $\{u_{k+1}, \dots, u_n\}$ es una base para \mathbb{H}^\perp y por teorema anterior tenemos:

$$\begin{aligned} v &= (v \cdot u_1) \cdot u_1 + (v \cdot u_2) \cdot u_2 + \dots + (v \cdot u_k) \cdot u_k + (v \cdot u_{k+1}) \cdot u_{k+1} + \dots + (v \cdot u_n) \cdot u_n \\ &= \text{Proy}_{\mathbb{H}} v + \text{Proy}_{\mathbb{H}^\perp} v \end{aligned}$$

Esto prueba la igualdad.

La unicidad se deja como tarea al lector.

Teorema 3.6 (Teorema de Aproximación de la Norma) Sea \mathbb{H} un subespacio de \mathbb{R}^n y sea $v \in \mathbb{R}^n$. Entonces $\text{Proy}_{\mathbb{H}} v$ es la mejor aproximación para $v \in \mathbb{H}$, en el sentido que, si h es cualquier otro vector en \mathbb{H} , entonces:

$$|v - \text{Proy}_{\mathbb{H}} v| < |v - h|$$

Demostración:

Del teorema de la proyección $v - \text{Proy}_{\mathbb{H}} v \in \mathbb{H}^\perp$.

Escribiremos

$$v - h = (v - \text{Proy}_{\mathbb{H}} v) + (\text{Proy}_{\mathbb{H}} v - h)$$

Notar que: $v - \text{Proy}_H v$ está en \mathbb{H}^\perp , mientras que $\text{Proy}_H v - h$ está en \mathbb{H} , por la definición (3.10) se cumple:

$$(v - \text{Proy}_H v) \cdot (\text{Proy}_H v - h) = 0$$

Ahora:

$$\begin{aligned} |v - h|^2 &= (v - h) \cdot (v - h) \\ &= [(v - \text{Proy}_H v) + (\text{Proy}_H v - h)] \cdot [(v - \text{Proy}_H v) + (\text{Proy}_H v - h)] \\ &= |v - \text{Proy}_H v|^2 + 2(v - \text{Proy}_H v) \cdot (\text{Proy}_H v - h) + |\text{Proy}_H v - h|^2 \\ &= |v - \text{Proy}_H v|^2 + |\text{Proy}_H v - h|^2 \end{aligned}$$

Pero:

$$h \neq \text{Proy}_H v \Rightarrow |\text{Proy}_H v - h|^2 > 0$$

$$\therefore |v - h|^2 > |v - \text{Proy}_H v|^2$$

o sea

$$|v - h| > |v - \text{Proy}_H v|$$

\therefore Queda demostrado el teorema.

Teorema 3.7 (Desigualdad de Cauchy - Schwarz en \mathbb{R}^n) Sean u y v vectores en \mathbb{R}^n . Entonces:

- $|u \cdot v| \leq |u| \cdot |v|$
- $|u \cdot v| = |u| \cdot |v| \Leftrightarrow v = \lambda \cdot u \quad \lambda \in \mathbb{R}$

Demostración:

Observación 3.5 Si $u = 0$ ó $v = 0$ (o ambos a la vez) entonces $|u \cdot v| = |u| \cdot |v| = 0$

Debido a la observación anterior se supone que $u \neq 0$ y $v \neq 0$. Entonces:

$$\begin{aligned} 0 &\leq \left| \frac{u}{|u|} - \frac{v}{|v|} \right|^2 \\ &= \left(\frac{u}{|u|} - \frac{v}{|v|} \right) \cdot \left(\frac{u}{|u|} - \frac{v}{|v|} \right) \\ &= \frac{u \cdot u}{|u|^2} - \frac{2 \cdot u \cdot v}{|u| \cdot |v|} + \frac{v \cdot v}{|v|^2} \\ &= \frac{|u|^2}{|u|^2} - \frac{2 \cdot u \cdot v}{|u| \cdot |v|} + \frac{|v|^2}{|v|^2} \\ &= 2 - \frac{2 \cdot u \cdot v}{|u| \cdot |v|} \end{aligned}$$

Lo anterior implica:

$$\begin{aligned}
 0 &\leq 2 - \frac{2 \cdot u \cdot v}{|u| \cdot |v|} \\
 \frac{2 \cdot u \cdot v}{|u| \cdot |v|} &\leq 2 \\
 \text{luego: } \frac{u \cdot v}{|u| \cdot |v|} &\leq 1 \\
 u \cdot v &\leq |u| \cdot |v|
 \end{aligned}$$

Análogamente; partiendo

$$0 \leq \left| \frac{u}{|u|} + \frac{v}{|v|} \right|^2$$

se termina en

$$\begin{aligned}
 \frac{u \cdot v}{|u| \cdot |v|} &\geq -1 \\
 \text{es lo mismo } u \cdot v &\geq -|u| \cdot |v|
 \end{aligned}$$

Por lo tanto

$$-|u| \cdot |v| \leq u \cdot v \leq |u| \cdot |v|$$

Así queda demostrado

$$|u \cdot v| \leq |u| \cdot |v|$$

Por otro lado si $u = \lambda v$ entonces:

$$\begin{aligned}
 |u \cdot v| &= |\lambda v \cdot v| \\
 &= |\lambda| \cdot |v|^2
 \end{aligned}$$

y

$$\begin{aligned}
 |u| \cdot |v| &= |\lambda v| \cdot |v| \\
 &= |\lambda| \cdot |v| \cdot |v| \\
 &= |\lambda| \cdot |v|^2 \\
 &= |u \cdot v|
 \end{aligned}$$

Inversamente si $|u \cdot v| = |u| \cdot |v|$ con $u \neq 0$ y $v \neq 0$, entonces:

$$\begin{aligned}
 \left| \frac{|u| \cdot |v|}{|u \cdot v|} \right| &= 1 \\
 \text{de manera que } \frac{|u| \cdot |v|}{|u \cdot v|} &= \pm 1
 \end{aligned}$$

■

$$\begin{aligned}
 \frac{|u| \cdot |v|}{|u \cdot v|} &= 1 \\
 \Rightarrow \left| \frac{u}{|u|} - \frac{v}{|v|} \right|^2 &= \left(\frac{u}{|u|} - \frac{v}{|v|} \right) \cdot \left(\frac{u}{|u|} - \frac{v}{|v|} \right) \\
 &= 2 - \frac{2|u| \cdot |v|}{|u \cdot v|} \\
 &= 0
 \end{aligned}$$

Así

$$\frac{u}{|u|} = \frac{v}{|v|}$$

$$u = \frac{|u|}{|v|} \cdot v = \lambda v$$

■

$$\begin{aligned} \frac{|u| \cdot |v|}{|u \cdot v|} &= -1 \\ \Rightarrow \left| \frac{u}{|u|} + \frac{v}{|v|} \right|^2 &= \left(\frac{u}{|u|} + \frac{v}{|v|} \right) \cdot \left(\frac{u}{|u|} + \frac{v}{|v|} \right) \\ &= 2 + \frac{2|u| \cdot |v|}{|u \cdot v|} \\ &= 0 \end{aligned}$$

Así

$$\frac{u}{|u|} = -\frac{v}{|v|}$$

$$u = -\frac{|u|}{|v|} \cdot v = \lambda v$$

∴ Queda demostrado el Teorema.

3.7. Aproximación por mínimos cuadrados

En muchos experimentos de la vida cotidiana se hace una sobreabundante recogida de datos, más de los estrictamente necesarios, y se quiere ajustar esa nube de puntos por una función más sencilla (normalmente un polinomio) que los represente lo mejor posible. Esa función de ajuste y simple no trata de satisfacer las ecuaciones, de pasar por los puntos, sino más bien de minimizar cierto error global y resolver el problema globalmente.

Por tal razón, nuestro objetivo es encontrar la curva del tipo específico que se **ajuste mejor** a los datos dados.

3.7.1. Aproximación por una Recta

Para mayor simplicidad enunciaremos el caso en \mathbb{R}^2 y luego generalizaremos para \mathbb{R}^n . Supongamos que buscamos una recta de la forma $y = mx + b$ que mejor represente a los n datos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Como se supuso anteriormente x e y están relacionadas por la ecuación $y = mx + b$, así por ejemplo, para $x = x_1$ el valor correspondiente de y es $mx_1 + b$; esto difiere del valor

real $y = y_1$.

En \mathbb{R}^2 la distancia entre los puntos (a_1, b_1) y (a_2, b_2) está dada por $d = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$. Por lo tanto, al determinar la manera de elegir la recta que mejor se aproxima a los datos dados, es razonable usar el criterio de escoger aquella que minimiza la suma de los cuadrados de las diferencias entre los valores y_j de los puntos y el valor y correspondiente a la recta. Notar que la distancia entre los puntos (x_1, y_1) y $(x_1, mx_1 + b)$ es $y_1 - (mx_1 + b)$, por lo tanto el problema (para los n datos) puede establecerse como sigue:

Encontrar números m y b tales que la suma

$$[y_1 - (mx_1 + b)]^2 + [y_2 - (mx_2 + b)]^2 + \dots + [y_n - (mx_n + b)]^2 \tag{3.1}$$

sea mínima.

Para estos valores de m y b , la recta $y = mx + b$ se llama aproximación por la recta de mínimos cuadrados a los datos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

Para simplicidad de la notación, escribiremos lo anterior en forma matricial. Si los puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ están todos sobre la recta $y = mx + b$ (es decir, si son colineales) se tiene:

$$\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ &\vdots \\ y_n &= mx_n + b \\ \therefore y &= Au \end{aligned}$$

donde

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, \quad u = \begin{pmatrix} m \\ b \end{pmatrix}$$

Si los puntos no son colineales, entonces $y - Au \neq 0$ y el problema sería encontrar u tal que la norma euclideana

$$|y - Au| \tag{3.2}$$

sea mínima; lo que constituye la Forma Vectorial del problema de Mínimos Cuadrados. Luego minimizar (3.2) es equivalente a minimizar la suma de cuadrados en (3.1).

Notar que A es una matriz de $n \times 2$ y u de 2×1 , por lo tanto, el vector Au es un vector en \mathbb{R}^n cuya dimensión es a lo más dos (observe que como máximo dos columnas de A son linealmente independientes). Así por el teorema de Aproximación de la Norma $|y - Au|$ es mínimo cuando

$$Au = Proj_{\mathbb{H}} y$$

donde \mathbb{H} es la imagen de A .

Para comprender mejor, mostraremos su interpretación geométrica, como vemos en la figura (3.1), en el caso de \mathbb{R}^3 la imagen será un plano o una recta que pasa por el origen (pues estos son los únicos subespacios de \mathbb{R}^3 de dimensión uno o dos). El vector que minimiza se denota \tilde{u} . Del Teorema de Pitágoras se deduce $|y - Au|$ es mínima cuando $y - Au$ es ortogonal a la imagen de A .

Observación 3.6 $y - Au$ es ortogonal a Au

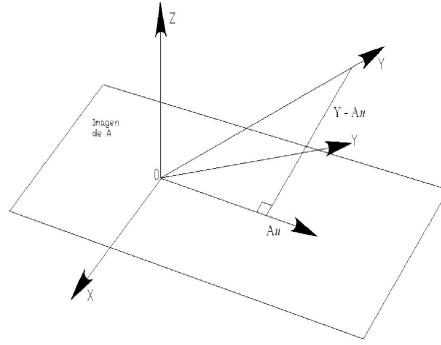


Figura 3.1: $y - A\tilde{u}$ es ortogonal Au

Luego, si \tilde{u} es el vector que minimiza, entonces para todo vector $u \in \mathbb{R}^2$ tenemos

$$Au \perp (y - A\tilde{u}).$$

Usando la definición de producto escalar en \mathbb{R}^n ; tenemos:

$$\begin{aligned} Au \cdot (y - A\tilde{u}) &= 0 \\ (Au)^t \cdot (y - A\tilde{u}) &= 0 \\ (u^t A^t) \cdot (y - A\tilde{u}) &= 0 \end{aligned}$$

o

$$u^t (A^t y - A^t A \tilde{u}) = 0$$

La última ecuación se cumple $\forall u \in \mathbb{R}^2$ sólo si

$$(A^t y - A^t A \tilde{u}) = 0$$

Al despejar \tilde{u} , se obtiene la *Solución al problema de mínimos cuadrados para un ajuste por línea recta*.

Si

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, \quad u = \begin{pmatrix} m \\ b \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

entonces la recta $y = mx + b$ da el mejor ajuste (en el sentido de mínimos cuadrados) para los puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ cuando

$$\begin{aligned} \tilde{u} &= \begin{pmatrix} m \\ b \end{pmatrix} \\ &= y \\ \tilde{u} &= (A^t A)^{-1} A^t y \end{aligned}$$

Se supuso que $A^t A$ es invertible, en el caso que los n datos no son colineales.

Teorema 3.8 Sean $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ n puntos en \mathbb{R}^2 y suponga que no todas las x_i son iguales. Entonces si

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$$

la matriz $A^t A$ es una matriz invertible de 2×2 .

Observación 3.7 Si $x_1 = x_2 = \dots = x_n$ entonces todos los datos están sobre la recta vertical $x = x_1$ y la mejor aproximación es obviamente esta recta.

Demostración por reducción al absurdo:

Observe $A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$ y note que no todas las x_i son iguales, por lo tanto las columnas de A son linealmente independientes.

Veamos quien es $A^t A$.

$$\begin{aligned} A^t A &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \end{aligned}$$

Supongamos que $A^t A$ no es invertible, eso implicaría que $\det(A^t A) = 0$

$$\begin{aligned} \Leftrightarrow n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 &= 0 \\ \Leftrightarrow n \sum_{i=1}^n x_i^2 &= \left(\sum_{i=1}^n x_i \right)^2 \end{aligned}$$

Sea: $u = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ y $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ entonces

$$|u|^2 = u \cdot u = n \quad |x|^2 = \sum_{i=1}^n x_i^2 \quad \text{y} \quad u \cdot x = \sum_{i=1}^n x_i$$

por lo tanto se puede reescribir como:

$$\begin{aligned} |u|^2 \cdot |x|^2 &= |u \cdot x|^2 \\ \Leftrightarrow |u| \cdot |x| &= |u \cdot x| \end{aligned}$$

Utilizando la desigualdad de Cauchy Schwartz que dice: $|u| \cdot |x| \leq |u \cdot x|$ en donde la desigualdad se cumple, si y solo si x es un múltiplo de u . Pero u y x son linealmente independientes; ya que son las columnas de A

$\Rightarrow \Leftarrow$

por hipótesis.

Por lo tanto queda demostrado el Teorema.

3.7.2. Aproximación Cuadrática

Recordemos que una curva cuadrática tiene la forma:

$$y = a_2x^2 + a_1x + a_0.$$

Deseamos ajustar esta curva cuadrática a n datos. La ecuación anterior, es la ecuación de una parábola en el plano. Si los n datos estuvieran sobre la parábola se tendría lo siguiente:

$$\begin{aligned} y_1 &= a_2x_1^2 + a_1x_1 + a_0 \\ y_2 &= a_2x_2^2 + a_1x_2 + a_0 \\ &\vdots \\ y_n &= a_2x_n^2 + a_1x_n + a_0 \end{aligned}$$

$$\text{Para } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad A = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \quad u = \begin{pmatrix} a_2 \\ a_1 \\ a_0 \end{pmatrix}$$

El sistema se puede reescribir como: $y = Au$ al igual que en el caso de Aproximación por una recta.

Si los datos están todos sobre la curva se tiene $y - Au = 0$ y para eso se encuentra el polinomio de grado dos que pasa por los puntos; por ejemplo para: $(1,4); (2,7); (3,14)$ se quiere encontrar un polinomio de grado dos (uno menos que el número de puntos dados como datos) que se ajuste a esos puntos. Sea el polinomio buscado:

$$a_2x^2 + a_1x + a_0$$

Se dieron tres puntos; los cuales se utilizarán para determinar las incógnitas a_2, a_1 y a_0 . Sustituyendo en nuestro polinomio:

$$\begin{aligned} x = 1, y = 4 \\ x = 2, y = 7 \\ x = 3, y = 14 \end{aligned}$$

Obtenemos el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} a_2x^2 + a_1x + a_0 &= 4 \\ 4a_2x^2 + 2a_1x + 1a_0 &= 7 \\ 9a_2x^2 + 3a_1x + 1a_0 &= 14 \end{aligned}$$

Podemos representar el sistema de ecuaciones por medio de la matriz

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 4 & 2 & 1 & 7 \\ 9 & 3 & 1 & 14 \end{array} \right)$$

efectuando las siguientes operaciones elementales:

$$E_{21}(-4) E_{21}(-9) \sim E_{32}(-3) \sim E_{23}(3) E_{13}(-1) \sim E_2(1/2) \sim E_{12}(1) \sim$$

obtenemos la matriz en su forma escalonada reducida

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & -1 & 0 & 3 \\ 0 & 0 & 1 & 5 \end{array} \right)$$

De donde: $a_2 = 2$, $a_1 = -3$ y $a_0 = 5$

Por lo tanto la parábola que pasa por los tres puntos dados es $2x^2 - 3x + 5 = y$

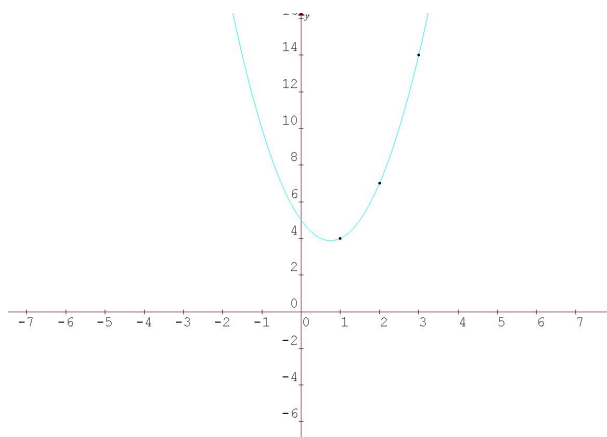


Figura 3.2: Gráfica de $2x^2 - 3x + 5 = y$

Observación 3.8 *El lector puede verificar que efectivamente $y - Au = 0$ lo que nos quiere decir que los datos están sobre la parábola.*

Si los datos no están todos sobre la parábola; entonces $y - Au \neq 0$ para cualquier vector u , y luego el problema se convierte en:

Encontrar un vector $u \in \mathbb{R}^3$ tal que $|y - Au|$ sea mínima.

Usando el mismo razonamiento que en el caso de la recta, se puede demostrar que si al menos tres de las x_i son diferentes, entonces $A^t A$ es invertible y el vector que minimiza al vector \tilde{u} está dado por:

$$\tilde{u} = (A^t A)^{-1} \cdot A^t y$$

3.7.3. Aplicación a la Geología Minera

Los geólogos estudian la composición de rocas y minerales en las formaciones para reunir información sobre ellas. Estudiando las rocas metamórficas y determinando aspectos como temperatura y presión a la que se formaron se obtendrá información útil sobre las condiciones presentes en el momento de formación. Un mineral común es el granate. Se sabe que el coeficiente de distribución de $F_e - M_g$ del granate es altamente dependiente de la temperatura a la que se formó. (Aquí, el coeficiente de distribución $F_e - M_g$ se relaciona con las proporciones de hierro (F_e) y magnesio (M_g) en el granate). Sin embargo, la cantidad de calcio (C_a) en el granate también afecta el coeficiente de distribución ($F_e - M_g$). Se pueden hacer correcciones a las estimaciones de temperatura, si la relación entre calcio presente y el coeficiente $F_e - M_g$ del granate se pueden determinar. Se reunieron los siguientes datos de las muestras del granate formadas en las montañas de Esplanade en British Columbia.

Tabla N°1

Fracción Molecular de C_a	Coficiente de distribución $F_e - M_g$
0,1164	0,12128
0,0121	0,17185
0,0562	0,13365
0,0931	0,1485
0,0664	0,12637
0,1728	0,10406
0,1793	0,10703
0,1443	0,1189
0,1824	0,09952

Encontraremos la recta de mínimos cuadrados y la graficaremos. Para eso utilizaremos la fracción molecular de C_a para las coordenadas x y el coeficiente de distribución $F_e - M_g$ para las coordenadas y . ¿Tienen los datos, en apariencia, una relación lineal?

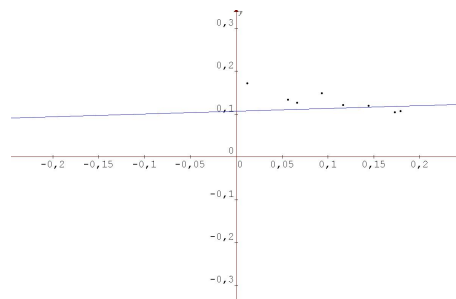


Figura 3.3: Representación gráfica de la Tabla N°1

Construyendo el sistema:

$$\begin{aligned}
 0,12128 &= m \cdot 0,1164 + b \\
 0,17185 &= m \cdot 0,0121 + b \\
 0,13365 &= m \cdot 0,0562 + b \\
 0,1485 &= m \cdot 0,0931 + b \\
 0,12637 &= m \cdot 0,0664 + b \\
 0,10406 &= m \cdot 0,1728 + b \\
 0,10703 &= m \cdot 0,1793 + b \\
 0,1189 &= m \cdot 0,1443 + b \\
 0,09952 &= m \cdot 0,1824 + b
 \end{aligned}$$

Luego su representación matricial es: $y = \begin{pmatrix} 0,12128 \\ 0,17185 \\ 0,13365 \\ 0,1485 \\ 0,12637 \\ 0,10406 \\ 0,10703 \\ 0,1189 \\ 0,09952 \end{pmatrix} \quad A = \begin{pmatrix} 0,1164 & 1 \\ 0,0121 & 1 \\ 0,0562 & 1 \\ 0,0931 & 1 \\ 0,0664 & 1 \\ 0,1728 & 1 \\ 0,1793 & 1 \\ 0,1443 & 1 \\ 0,1824 & 1 \end{pmatrix} \quad u = \begin{pmatrix} m \\ b \end{pmatrix}$

De donde:

$$\begin{aligned}
 A^t A &= \begin{pmatrix} 0,1164 & 0,0121 & 0,0562 & 0,0931 & 0,0664 & 0,1728 & 0,1793 & 0,1443 & 0,1824 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0,1164 & 1 \\ 0,0121 & 1 \\ 0,0562 & 1 \\ 0,0931 & 1 \\ 0,0664 & 1 \\ 0,1728 & 1 \\ 0,1793 & 1 \\ 0,1443 & 1 \\ 0,1824 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 0,14603096 & 1,023 \\ 1,023 & 10 \end{pmatrix}
 \end{aligned}$$

Calculando el determinante de $A^t A$ se obtiene: $\det(A^t A) = 0,4137806$

$$\Rightarrow (A^t A)^{-1} = \frac{1}{0,4137806} \begin{pmatrix} 10 & -1,023 \\ -1,023 & 0,14603096 \end{pmatrix}$$

$$\tilde{u} = (A^t A)^{-1} A^t y$$

$$\tilde{u} = \frac{1}{0,4137806} \begin{pmatrix} 10 & -1,023 \\ -1,023 & 0,14603096 \end{pmatrix} \cdot \begin{pmatrix} 0,11840559 \\ 1,13116 \end{pmatrix}$$

$$\tilde{u} = \frac{1}{0,4137806} \begin{pmatrix} 0,02687922 \\ 0,044055462 \end{pmatrix}$$

$$\tilde{u} = \begin{pmatrix} 0,064960077 \\ 0,106470584 \end{pmatrix}$$

Por lo tanto, la recta que **mejor se ajusta** a los datos, está dada por la ecuación:

$$y = 0,064960077x + 0,106470584$$

3.7.4. Aplicación al Movimiento Rectilíneo Uniforme

Una persona lanza una pelota al aire hacia abajo. La altura que alcanza está dada por

$$s(t) = s_o + v_o t + \frac{1}{2}gt^2.$$

Se toman las siguientes mediciones:

Tabla N°2

Tiempo transcurrido(segundos)	Altura(pies)	Altura(en metros)
1	57	17,3736
1,5	67	20,4216
2,5	68	20,7264
4	9,5	2,8956

Usando los datos, estimar:

- La altura a la que se dejó caer la pelota.
- La velocidad inicial.
- g en (en m/seg^2)

Los datos dados se pueden reescribir como el siguiente sistema de ecuaciones:

$$\begin{aligned} 17,3736 &= s_0 + v_0 + \frac{1}{2}g \\ 20,4216 &= s_0 + 1,5v_0 + \frac{1}{2}2,25g \\ 20,7264 &= s_0 + 2,5v_0 + \frac{1}{2}6,25g \\ 2,8956 &= s_0 + 4v_0 + \frac{1}{2}16g \end{aligned}$$

y que en forma matricial se representa:

$$y = Au$$

$$\text{donde } y = \begin{pmatrix} 17,3736 \\ 20,4216 \\ 20,7264 \\ 2,8956 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1,5 & 2,25 \\ 1 & 2,5 & 6,25 \\ 1 & 4 & 16 \end{pmatrix} \quad u = \begin{pmatrix} s_0 \\ v_0 \\ g \end{pmatrix}$$

Queremos saber quien es $A^t A$ y $A^t y$

$$A^t A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1,5 & 2,5 & 4 \\ 1 & 2,25 & 6,25 & 16 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1,5 & 2,25 \\ 1 & 2,5 & 6,25 \\ 1 & 4 & 16 \end{pmatrix} = \begin{pmatrix} 4 & 9 & 25,5 \\ 9 & 25,5 & 84 \\ 25,5 & 84 & 301,125 \end{pmatrix}$$

Haciendo una serie de cálculos se obtuvo que $(A^t A)^{-1} = \begin{pmatrix} 8,38636362 & -7,65151513 & 1,42424242 \\ -7,65151515 & 7,46349206 & -1,43434343 \\ 1,42424242 & -1,43434343 & 0,28282828 \end{pmatrix}$

Por otro lado $A^t y = \begin{pmatrix} 61,4172 \\ 111,4044 \\ 239,1918 \end{pmatrix}$

Finalmente como $\tilde{u} = (A^t A)^{-1} \cdot A^t y$ tenemos:

$$\begin{aligned} \tilde{u} &= \begin{pmatrix} 8,38636362 & -7,65151513 & 1,42424242 \\ -7,65151515 & 7,46349206 & -1,43434343 \\ 1,42424242 & -1,43434343 & 0,28282828 \end{pmatrix} \cdot \begin{pmatrix} 61,4172 \\ 111,4044 \\ 239,1918 \end{pmatrix} \\ &= \begin{pmatrix} 3,321627649 \\ 18,44803413 \\ -4,668982271 \end{pmatrix} \end{aligned}$$

En consecuencia los datos se ajustan a la ecuación cuadrática

$$s(t) = 3,32163 + 18,44803t - 4,66898t^2$$

lo que nos indica:

- que el término $3,32163$ apróx. representa la altura inicial de la pelota (medida en metros)
- que el término $18,44803t$ representa la velocidad inicial (hacia arriba) de $18,45m/seg$
- como

$$\begin{aligned} \frac{1}{2}g &\approx 4,668982271 \\ g &\approx 2 \cdot 4,668982271 \\ g &\approx 9,34m/seg^2 \end{aligned}$$

este valor es cercano a $9,81m/seg^2$.

Para obtener una aproximación más exacta de g tendrían que haberse tomado los datos con mayor precisión.

3.8. Ejercicios Propuestos

1. Encuentre el polinomio de grado dos, cuya gráfica pasa a través de los puntos dados: $(1,14)$; $(2,22)$; $(3,32)$. Grafíquelo y además verifique que $y - Au = 0$.
2. Encuentre el polinomio de grado tres que mejor se ajusta a los puntos dados: $(-1,0)$; $(1,1)$; $(2,2)$; $(3,5)$. Grafíquelo y además verifique que $y - Au \neq 0$.
3. Un fabricante compra grandes cantidades de refacciones para cierta máquina. Él encuentra que este costo depende del número de cajas compradas al mismo tiempo y que el costo por unidad disminuye conforme el número de cajas aumenta. Supone que el costo es una función cuadrática del volumen y de las facturas anteriores, obtiene la siguiente tabla:

N°de cajas compradas	Costo Total(dólares)
10	150
30	260
50	325
100	500
175	670

Encuentre su función de costo total.

CAPÍTULO 4

Espacios con Producto Interno y Proyecciones

4.1. Introducción

En este capítulo se definirá Espacio Vectorial con Producto Interno $\langle u, v \rangle$ y se enunciarán sus propiedades. Además se presentará el Espacio $\mathbb{V} = C[a, b]$ de las funciones continuas de valores reales en el intervalo $[a, b]$ con el fin de definir el producto interno en este Espacio.

En la sección (3.6) estudiamos la proyección ortogonal de v sobre \mathbb{H} , en el capítulo presente se extienden los resultados del capítulo anterior al Espacio con Producto Interno.

Al final del capítulo se presentarán las aplicaciones relacionadas con Mínimos Cuadrados: *Aproximación de una función por medio de polinomios y Aproximaciones de Fourier de una función f en $C[\pi, -\pi]$ por medio de polinomios trigonométricos.*

Definición 4.1 (Espacio con producto interno) *Un espacio vectorial complejo \mathbb{V} se llama espacio con producto interno si para cada par ordenado de vectores u y $v \in \mathbb{V}$, existe un número complejo único $\langle u, v \rangle$, llamado producto interno de u y v , tal que si u, v y w están en \mathbb{V} , y $\alpha \in \mathbb{C}$, entonces:*

- $\langle v, v \rangle \geq 0$
- $\langle v, v \rangle = 0 \Leftrightarrow v = 0$
- $\langle u, v+w \rangle = \langle u, v \rangle + \langle u, w \rangle$
- $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
- $\langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle$

Ejemplo 4.1 (Un producto interno) *Sea $\mathbb{V} = C[a, b]$ el espacio de las funciones continuas de valores reales en el intervalo $[a, b]$ (suponiendo que $a < b$). Se define*

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

Queda al lector probar que efectivamente $\langle f, g \rangle = \int_a^b f(t)g(t)dt$ es un producto interno.

Observación 4.1 *Esta es la manera más común de definir el producto interno en $C[a, b]$; pero no es la única.*

Definición 4.2 *Sea \mathbb{V} un espacio con producto interno y suponga que u y v están en \mathbb{V} . Entonces:*

1. *u y v son ortogonales si $\langle u, v \rangle = 0$*
2. *La norma de u , denotada por $\|u\|$, está dada por $\|u\| = \sqrt{\langle u, u \rangle}$*

Definición 4.3 *El conjunto de vectores $\{v_1, v_2, \dots, v_n\}$ es un conjunto ortonormal en \mathbb{V} si:*

- $\langle v_i, v_j \rangle = 0$ para $i \neq j$
- $\|v_i\| = \sqrt{\langle v_i, v_i \rangle} = 1$

Si solo se cumple la primera condición se dice que el conjunto es ortogonal.

Teorema 4.1 *Cualquier conjunto finito de vectores ortogonales diferentes de cero en un espacio con producto interno es linealmente independiente.*

Definición 4.4 *Sea \mathbb{H} un subespacio del espacio con producto interno \mathbb{V} con una base ortonormal $\{u_1, u_2, \dots, u_k\}$ si $v \in \mathbb{V}$, entonces la proyección ortogonal de \mathbb{V} sobre \mathbb{H} denotada por*

$$Proy_{\mathbb{H}}v = \langle v, u_1 \rangle \cdot u_1 + \langle v, u_2 \rangle \cdot u_2 + \dots + \langle v, u_k \rangle \cdot u_k$$

Observación 4.2 *Notar que $Proy_{\mathbb{H}}v \in \mathbb{H}$.*

Definición 4.5 Sea \mathbb{H} un subespacio del espacio con producto interno \mathbb{V} . Entonces el complemento ortogonal de \mathbb{H} , denotado por \mathbb{H}^\perp , está dado por:

$$\mathbb{H}^\perp = \{x \in \mathbb{V} : \langle x, h \rangle = 0 \text{ para todo } h \in \mathbb{H}\}$$

Teorema 4.2 Si \mathbb{H} es un subespacio del espacio con producto interno \mathbb{V} , entonces:

1. \mathbb{H}^\perp es un subespacio de \mathbb{V}
2. $\mathbb{H} \cap \mathbb{H}^\perp = 0$
3. $\dim \mathbb{H}^\perp = n - \dim \mathbb{H}$ si $\dim \mathbb{V} = n < \infty$

Teorema 4.3 (Teorema de Proyección) Sea \mathbb{H} un subespacio de dimensión finita del espacio con producto interno \mathbb{V} y suponga que $v \in \mathbb{V}$. Entonces existe un par único de vectores h y p tales que $h \in \mathbb{H}$, $p \in \mathbb{H}^\perp$. Y $v = h + p$. En particular, $h = \text{Proy}_{\mathbb{H}} v$ y si \mathbb{V} tiene dimensión finita, entonces $p = \text{Proy}_{\mathbb{H}^\perp} v$.

Teorema 4.4 (Teorema de Aproximación de la Norma) Sea \mathbb{H} un subespacio de dimensión finita del espacio con producto interno \mathbb{V} y sea v un vector en \mathbb{V} . Entonces, en \mathbb{H} , $\text{Proy}_{\mathbb{H}} v$ es la mejor aproximación a v en el siguiente sentido: si h es cualquier vector en \mathbb{H} , entonces:

$$\|v - \text{Proy}_{\mathbb{H}} v\| < \|v - h\|$$

4.2. Aproximación por mínimos cuadrados a una función continua

Sea $f \in C[a,b]$. Se quiere aproximar f por un polinomio de grado n . Luego surge la pregunta

¿Cuál es el polinomio que se aproxima **mejor** a la función?.

Se entiende por **mejor aproximación** aquella que minimiza el error al aproximar.

Definición 4.6 (Tipos de Error) Definiremos los tres tipos de error más comunes:

- *Error Máximo:* $\max |f(x) - g(x)|$ para $x \in [a,b]$
- *Error de Área:* $\int_a^b |f(x)g(x)| dx$
- *Error Cuadrático Medio:* $\int_a^b |f(x)g(x)|^2 dx$

Se puede utilizar el teorema de aproximación de la norma para encontrar el polinomio único de grado n que se aproxima a una función continua dada con el error cuadrático medio más pequeño.

Para todo $n \in \mathbb{N}$, $P_n[a,b]$ el espacio de los polinomios de grado menor o igual a n definidos sobre $[a,b]$ es un subespacio de dimensión finita de $C[a,b]$. Se puede calcular, para $f \in C[a,b]$ y $p_n \in P_n$.

$$\begin{aligned} \|f - p_n\|^2 &= \langle f - p_n, f - p_n \rangle \\ &= \int_a^b [(f(t) - p_n(t))(f(t) - p_n(t))] dt \\ &= \int_a^b |f(t) - p_n(t)|^2 dt \\ &= \text{error cuadrático medio} \end{aligned}$$

(4.2)

Por lo tanto el polinomio de grado n que se aproxima a una función continua con el error cuadrático medio más pequeño, está dado por:

$$p_n = \text{Proy}_{P_n} f$$

4.3. Aplicaciones

4.3.1. Mínimos Cuadrados

Encuentre un polinomio de segundo grado que mejor se aproxime a $\sin(\frac{\pi}{2}x)$ en el intervalo $[0,1]$ en el sentido del error cuadrático medio.

Primero debemos encontrar una base ortonormal en $P_2[0,1]$, para así gracias a un teorema mencionado anteriormente sabremos quien es la $\text{Proy}_{P_2[0,1]} \sin(\frac{\pi}{2}x)$.

Verifiquemos que el conjunto $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$ es una base ortonormal en $P_2[0,1]$.

En efecto :

▪

$$\begin{aligned} \langle 1, \sqrt{3}(2x-1) \rangle &= \int_0^1 [1 \cdot \sqrt{3}(2x-1)] dx \\ &= \sqrt{3} \int_0^1 (2x-1) dx \\ &= \sqrt{3}(x^2-x) \Big|_0^1 \\ &= \sqrt{3}[(1^2-1) - (0^2-0)] \\ &= 0 \end{aligned}$$

▪

$$\begin{aligned} \langle \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1) \rangle &= \int_0^1 [\sqrt{3}(2x-1) \cdot \sqrt{5}(6x^2-6x+1)] dx \\ &= \sqrt{15} \int_0^1 (12x^3 - 18x^2 + 8x - 1) dx \\ &= \sqrt{15} \int_0^1 (3x^4 - 6x^3 + 4x^2 - x) \Big|_0^1 \\ &= \sqrt{15}[(3-6+4-1) - 0] \\ &= 0 \end{aligned}$$

▪

$$\begin{aligned} \langle 1, \sqrt{5}(6x^2-6x+1) \rangle &= \int_0^1 [1 \cdot \sqrt{5}(6x^2-6x+1)] dx \\ &= \sqrt{5}[2x^3 - 3x^2 + x] \Big|_0^1 \\ &= \sqrt{5}[(2-3+1) - 0] \\ &= 0 \end{aligned}$$

$\therefore \{1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1)\}$ es ortogonal.

Solo falta verificar la segunda condición de conjunto ortonormal.

▪

$$\begin{aligned} \|1\|^2 &= \int_0^1 1 \cdot 1 dx \\ &= x \Big|_0^1 \\ &= 1 - 0 \\ &= 1 \end{aligned}$$

$$\therefore \|1\| = 1$$

▪

$$\begin{aligned} \|\sqrt{3}(2x - 1)\|^2 &= \int_0^1 [\sqrt{3}(2x - 1) \cdot \sqrt{3}(2x - 1)] dx \\ &= 3 \int_0^1 (4x^2 - 4x + 1) dx \\ &= 3 \left[\frac{4}{3} x^3 - 2x^2 + x \right]_0^1 \\ &= 3 \left[\left(\frac{4}{3} - 2 + 1 \right) - 0 \right] \\ &= 3 \cdot \frac{1}{3} \\ &= 1 \end{aligned}$$

$$\therefore \|\sqrt{3}(2x - 1)\| = 1$$

▪

$$\begin{aligned} \|\sqrt{5}(6x^2 - 6x + 1)\|^2 &= \int_0^1 [\sqrt{5}(6x^2 - 6x + 1) \cdot \sqrt{5}(6x^2 - 6x + 1)] dx \\ &= 5 \int_0^1 (36x^4 - 72x^3 + 48x^2 - 12x + 1) dx \\ &= 5 \left[\frac{36}{5} x^5 - 18x^4 + 16x^3 - 6x^2 + x \right]_0^1 \\ &= 5 \left[\left(\frac{36}{5} - 18 + 16 - 6 + 1 \right) - 0 \right] \\ &= 5 \left(\frac{36}{5} - 7 \right) \\ &= 1 \end{aligned}$$

$$\therefore \|\sqrt{5}(6x^2 - 6x + 1)\| = 1$$

\therefore El conjunto $\{1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1)\}$ es ortonormal.

Calculando:

$$\begin{aligned}
 \text{Proy}_{P_2}[0, 1] &= \langle \sin(\frac{\pi}{2}x, 1) \rangle \cdot 1 + \langle \sin(\frac{\pi}{2}x, \sqrt{3}(2x-1)) \rangle \cdot \sqrt{3}(2x-1) + \langle \sin(\frac{\pi}{2}x, \sqrt{5}(6x^2-6x+1)) \rangle \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \int_0^1 \sin(\frac{\pi}{2}x) \cdot 1 dx + \int_0^1 [\sin(\frac{\pi}{2}x) \cdot \sqrt{3}(2x-1)] dx \cdot \sqrt{3}(2x-1) + \int_0^1 [\sin(\frac{\pi}{2}x) \cdot \sqrt{5}(6x^2-6x+1)] dx \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} \int_0^1 \frac{\pi}{2} \sin(\frac{\pi}{2}x) dx + \sqrt{3} \int_0^1 \sin(\frac{\pi}{2}x) \cdot (2x-1) dx \cdot \sqrt{3}(2x-1) + \sqrt{5} \int_0^1 \sin(\frac{\pi}{2}x) \cdot (6x^2-6x+1) dx \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} \int_0^1 \sin(u) du + \sqrt{3} [\int_0^1 2x \sin(\frac{\pi}{2}x) dx - \int_0^1 \sin(\frac{\pi}{2}x) dx] \cdot \sqrt{3}(2x-1) + \sqrt{5} [\int_0^1 6x^2 \sin(\frac{\pi}{2}x) dx - \int_0^1 6x \sin(\frac{\pi}{2}x) dx + \int_0^1 \sin(\frac{\pi}{2}x) dx] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} \cdot [-\cos(u)]_0^1 + \sqrt{3} [2 \cdot \frac{4}{\pi^2} \int_0^1 \frac{\pi}{2} \cdot (\frac{\pi}{2}x) \sin(\frac{\pi}{2}x) dx - \frac{2}{\pi} \int_0^1 \frac{\pi}{2} \sin(\frac{\pi}{2}x) dx] \sqrt{3}(2x-1) + \sqrt{5} [6 \cdot \frac{4}{\pi^2} \cdot \frac{2}{\pi} \int_0^1 \frac{\pi}{2} (\frac{\pi}{2}x)^2 \sin(\frac{\pi}{2}x) dx - 6 \cdot \frac{4}{\pi^2} \int_0^1 \frac{\pi}{2} (\frac{\pi}{2}x) \sin(\frac{\pi}{2}x) dx + \frac{2}{\pi} \int_0^1 \frac{\pi}{2} \sin(\frac{\pi}{2}x) dx] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{-2}{\pi} \cdot \cos(\frac{\pi}{2}) + \sqrt{3} [\frac{8}{\pi^2} \int_0^1 u \sin(u) du - \frac{2}{\pi} \cdot [-\cos(\frac{\pi}{2})]] \cdot \sqrt{3}(2x-1) + \sqrt{5} [\frac{48}{\pi^3} \int_0^1 u^2 \sin(u) du - \frac{24}{\pi^2} \int_0^1 u \sin(u) du + \frac{2}{\pi} \cdot [-\cos(\frac{\pi}{2})]] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + \sqrt{3} [\frac{8}{\pi^2} \{-\frac{\pi}{2}x \cdot \cos(\frac{\pi}{2}x) + \sin(\frac{\pi}{2}x)\}]_0^1 - \frac{2}{\pi} \cdot \sqrt{3}(2x-1) + \sqrt{5} [\frac{48}{\pi^3} \{-\frac{1}{2}(\frac{\pi}{2}x)^2 \cdot \cos(\frac{\pi}{2}x) + 2(\frac{\pi}{2}x) \sin(\frac{\pi}{2}x) + 2 \cdot \cos(\frac{\pi}{2}x)\}]_0^1 - \frac{24}{\pi^2} \{-\frac{1}{2}(\frac{\pi}{2}x) \cos(\frac{\pi}{2}x) + \sin(\frac{\pi}{2}x)\}]_0^1 + \frac{2}{\pi} \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + \sqrt{3} [\frac{8}{\pi^2} \{1\} - \frac{2}{\pi}] \sqrt{3}(2x-1) + \sqrt{5} [\frac{48}{\pi^3} \{\pi-2\} - \frac{24}{\pi^2} \{1\} + \frac{2}{\pi}] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + \sqrt{3} [\frac{8}{\pi^2} - \frac{2}{\pi}] \sqrt{3}(2x-1) + \sqrt{5} [\frac{48}{\pi^3} \{\pi-2\} - \frac{24}{\pi^2} + \frac{2}{\pi}] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + \sqrt{3} [\frac{8}{\pi^2} - \frac{2}{\pi}] \sqrt{3}(2x-1) + \sqrt{5} [\frac{48}{\pi^3} - \frac{96}{\pi^3} - \frac{24}{\pi^2} + \frac{2}{\pi}] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + \sqrt{3} [\frac{8}{\pi^2} - \frac{2}{\pi}] \sqrt{3}(2x-1) + \sqrt{5} [-\frac{96}{\pi^3} - \frac{24}{\pi^2} + \frac{2}{\pi}] \cdot \sqrt{5}(6x^2-6x+1) \\
 &= \frac{2}{\pi} + [\frac{24}{\pi^2} - \frac{6}{\pi}](2x-1) + [-\frac{480}{\pi^3} - \frac{120}{\pi^2} + \frac{10}{\pi}] \cdot (6x^2-6x+1)
 \end{aligned}$$

$$= \left[\frac{-2880}{\pi^3} + \frac{720}{\pi^2} + \frac{60}{\pi} \right] \cdot x^2 + \left[\frac{2880}{\pi^3} - \frac{672}{\pi^2} - \frac{72}{\pi} \right] \cdot x + \left[\frac{-480}{\pi^3} + \frac{96}{\pi^2} + \frac{18}{\pi} \right]$$

$$\approx -0.834574 \cdot x^2 + 1,87827 \cdot x - 0,024325$$

4.3.2. Aproximaciones de Fourier

Sea f una función en $C[-\pi, \pi]$ con producto interno definido:

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$$

El objetivo es determinar la aproximación por mínimos cuadrados de f en el espacio $T[-\pi, \pi]$ de polinomios trigonométricos generados por el conjunto $\{1, \cos x, \sin x, \dots, \cos nx, \sin nx\}$ donde $n \in \mathbb{N}$.

Queda al lector demostrar que efectivamente $\{1, \cos x, \sin x, \dots, \cos nx, \sin nx\}$ son ortogonales.

Además las magnitudes de estos vectores son:

$$\begin{aligned} \|1\|^2 &= \int_{-\pi}^{\pi} (1 \cdot 1)dx = 2\pi \\ \|\cos nx\|^2 &= \int_{-\pi}^{\pi} (\cos nx \cdot \cos nx)dx = \pi \\ \|\sin nx\|^2 &= \int_{-\pi}^{\pi} (\sin nx \cdot \sin nx)dx = \pi \end{aligned}$$

Por lo tanto

$$\{g_0, \dots, g_{2n}\} = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin x, \dots, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx \right\}$$

es una base ortonormal para $T[-\pi, \pi]$.

De manera similar a mínimos cuadrados, se utilizará una base ortonormal para determinar la proyección de f sobre $T[-\pi, \pi]$ y así determinar la aproximación por mínimos cuadrados g de f

$$g(x) = \text{Proy}_{T[-\pi, \pi]} f = \langle f, g_0 \rangle g_0 + \dots + \langle f, g_{2n} \rangle g_{2n}$$

es decir:

$$g(x) = \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \left\langle f, \frac{1}{\sqrt{\pi}} \cos x \right\rangle \frac{1}{\sqrt{\pi}} \cos x + \left\langle f, \frac{1}{\sqrt{\pi}} \sin x \right\rangle \frac{1}{\sqrt{\pi}} \sin x + \dots + \left\langle f, \frac{1}{\sqrt{\pi}} \cos nx \right\rangle \frac{1}{\sqrt{\pi}} \cos nx + \left\langle f, \frac{1}{\sqrt{\pi}} \sin nx \right\rangle \frac{1}{\sqrt{\pi}} \sin nx$$

Usando la siguiente notación:

$$\begin{aligned} a_0 &= \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} = \int_{-\pi}^{\pi} \left(f(x) \cdot \frac{1}{\sqrt{2\pi}} \right) dx \cdot \frac{1}{\sqrt{2\pi}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \\ a_k &= \left\langle f, \frac{1}{\sqrt{\pi}} \cos jx \right\rangle \frac{1}{\sqrt{\pi}} = \int_{-\pi}^{\pi} \left(f(x) \cdot \frac{1}{\sqrt{\pi}} \cos jx \right) dx \cdot \frac{1}{\sqrt{\pi}} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos jx dx \\ b_j &= \left\langle f, \frac{1}{\sqrt{\pi}} \sin jx \right\rangle \frac{1}{\sqrt{\pi}} = \int_{-\pi}^{\pi} \left(f(x) \cdot \frac{1}{\sqrt{\pi}} \sin jx \right) dx \cdot \frac{1}{\sqrt{\pi}} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin jx dx \end{aligned}$$

La aproximación trigonométrica de $f(x)$ se puede expresar como:

$$g(x) = a_0 + \sum_{j=1}^n (a_j \cos jx + b_j \sin jx)$$

donde $g(x)$ recibe el nombre de **aproximación de Fourier** de orden n de $f(x)$.

Los coeficientes $a_0, a_1, b_1, \dots, a_n, b_n$ reciben el nombre de coeficientes de Fourier y obviamente a medida que n se incrementa la aproximación es cada vez mejor, en el sentido que $\|f - g\|$ se reduce.

La suma infinita

$$g(x) = a_0 + \sum_{j=1}^{\infty} (a_j \cos jx + b_j \sin jx)$$

recibe el nombre de Serie de Fourier de f en el intervalo $[-\pi, \pi]$.

Ejemplo 4.2 *Se quiere aproximar la función $f(x) = x^2$ en el intervalo $[-\pi, \pi]$.*

Para eso, se calcularán los coeficientes de Fourier, es decir, a_0, a_j y b_j

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 dx \\ &= \frac{1}{2\pi} \left[\frac{x^3}{3} \right]_{-\pi}^{\pi} \\ &= \frac{1}{2\pi} \left(\frac{\pi^3}{3} - \frac{(-\pi)^3}{3} \right) \\ &= \frac{2\pi^3}{6\pi} \\ &= \frac{\pi^2}{3} \end{aligned}$$

$$\begin{aligned} a_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos(jx) dx \\ &= \frac{1}{j^3 \pi} \int_{-\pi}^{\pi} u^2 \cos(ju) du \\ &= \frac{1}{j^3 \pi} (u^2 \sin u - 2 \sin u + 2u \cos u) \Big|_{-\pi}^{\pi} \\ &= \frac{1}{j^3 \pi} ((j\pi)^2 \sin(j\pi) - 2 \sin(j\pi) + 2(j\pi) \cos(j\pi)) \Big|_{-\pi}^{\pi} \\ &= \frac{1}{j^3 \pi} [((j\pi)^2 \sin(j\pi) - 2 \sin(j\pi) + 2(j\pi) \cos(j\pi)) - ((-j\pi)^2 \sin(-j\pi) - 2 \sin(-j\pi) + 2(j\pi) \cos(-j\pi))] \\ &= \frac{1}{j^3 \pi} [2j\pi \cos(j\pi) + 2 \cos(-j\pi)] \\ &= \frac{4j\pi}{j^3 \pi} \cos(j\pi) \\ &= \frac{4}{j^2} \cos(j\pi) \\ &= \frac{4}{j^2} (-1)^j \end{aligned}$$

$$\begin{aligned}
 b_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \sin(jx) dx \\
 &= \frac{1}{j^3 \pi} \int_{-\pi}^{\pi} u^2 \sin(u) du \\
 &= \frac{1}{j^3 \pi} (-u^2 \cos u + 2u \sin u + 2 \cos u) \Big|_{-\pi}^{\pi} \\
 &= \frac{1}{j^3 \pi} (-(j\pi)^2 \cos(j\pi) + 2(j\pi) \sin(j\pi) + 2 \cos(j\pi)) \Big|_{-\pi}^{\pi} \\
 &= \frac{1}{j^3 \pi} [((-j\pi)^2 \cos(j\pi) + 2(j\pi) \sin(j\pi) + 2 \cos(j\pi)) - ((-j \cdot -\pi)^2 \cos(-j\pi) + 2(-j\pi) \sin(-j\pi) \\
 &\quad + 2 \cos(-j\pi))] \\
 &= \frac{1}{j^3 \pi} [-(j\pi)^2 \cos(j\pi) + 2 \cos(j\pi) + (j\pi)^2 \cos(-j\pi) - 2 \cos(-j\pi)] \\
 &= 0
 \end{aligned}$$

Por lo tanto, la aproximación de Fourier de f es:

$$g(x) = \frac{\pi^2}{3} + \sum_{j=1}^n \frac{4}{j^2} (-1)^j \cdot \cos(jx)$$

Ahora, en el caso que se pidiese la aproximación de Fourier de orden cuatro de $f(x) = x^2$ en el intervalo $[-\pi, \pi]$; se debe reemplazar $j = 1, 2, 3, 4$; y así se obtiene la aproximación de orden cuatro de f .

$$g(x) = 3,284 + 4 \left[-\cos(x) + \frac{1}{4} \cos(2x) - \frac{1}{9} \cos(3x) + \frac{1}{16} \cos(4x) \right]$$

4.4. Ejercicios Propuestos

- Encuentre un polinomio de quinto grado que mejor se aproxime a $\cos(\frac{\pi}{2}x)$ en el intervalo $[0, 1]$ en el sentido del error cuadrático medio.
- Encuentre la serie de Fourier y aproxime las siguientes funciones:
 - $f(x) = |x|$ en $[\pi, -\pi]$
 - $f(x) = e^{\alpha x}$ en $[\pi, -\pi]$
 - $f(x) = |\sin x|$ en $[\pi, -\pi]$

CAPÍTULO 5

Valores y Vectores Propios

5.1. Introducción

El capítulo comenzará definiendo valores y vectores propios?, la ecuación característica de una matriz y su respectivo polinomio característico, todo esto acompañado de un ejemplo ilustrativo para determinar cada uno de ellos.

En la sección (5.4) se analizarán las Formas Cuadráticas, se demostrarán algunos teoremas relevantes de las subsecciones (5.4.1) y (5.4.2) correspondientes a matrices semejentes y diagonalización; con el fin de aplicarlos para determinar dada una ecuación de dos variables, su forma canónica por medio de diagonalización utilizando valores y vectores propios y finalmente decir ¿qué forma geométrica es? (elipse, círculo, hipérbola, un par de rectas o una cónica degenerada).

Finalmente veremos algunas aplicaciones orientadas al *Crecimiento Poblacional*, *Cadenas de Markov* y *Física Mecánica (sistemas oscilantes de muelles y partículas y su generalización)*.

5.2. ¿Qué son los valores y vectores propios?

Definición 5.1 Sea $L : \mathbb{V} \rightarrow \mathbb{V}$ un operador lineal, $\dim \mathbb{V} = n$. Un escalar $\lambda \neq 0$ se llama un valor propio de L si existe $v \in \mathbb{V}$, $v \neq 0$, tal que $L(v) = \lambda v$. Todo vector que cumpla esa condición se llama **vector propio** asociado al valor propio λ .

Nota: Si $L : \mathbb{V} \rightarrow \mathbb{V}$ es un operador lineal, $\dim \mathbb{V} = n$, entonces L se puede representar por una matriz AL . Por esto se habla indistintamente de valores y vectores propios de un operador lineal o de la respectiva matriz.

5.3. Matriz Característica, Ecuación Característica y Polinomio Característico

Definición 5.2 Sea A una matriz de orden n . La matriz polinomial $C = A - \lambda I$ se llama matriz característica de A (I es la matriz identidad). El determinante de $A - \lambda I$ es un polinomio de grado n , llamado el polinomio característico de A . Así $f(\lambda) = |A - \lambda I|$ y la ecuación $|A - \lambda I| = 0$ se llama ecuación característica de la matriz A .

Proposición 5.1 Un escalar λ es un valor propio de un operador $L \leftrightarrow \lambda$ es una solución de la ecuación característica de una matriz que representa a L .

5.3.1. Cálculo de los valores y vectores propios de una matriz

Dada la matriz

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 4 & 0 & 4 \\ 1 & -1 & 4 \end{pmatrix}$$

Si λ es un valor propio de A , entonces tenemos el sistema:

$$\begin{aligned} Av &= \lambda v \\ A \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ A \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \lambda I_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

1. La matriz característica se encuentra viendo quien es $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} 1 & 1 & -2 \\ 4 & 0 & 4 \\ 1 & -1 & 4 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1 - \lambda & 1 & -2 \\ 4 & -\lambda & 4 \\ 1 & -1 & 4 - \lambda \end{pmatrix}$$

2. El polinomio característico es:

$$\begin{aligned}
 f(\lambda) &= |A - \lambda I| \\
 &= \begin{vmatrix} 1 - \lambda & 1 & -2 \\ 4 & -\lambda & 4 \\ 1 & -1 & 4 - \lambda \end{vmatrix} \\
 &= (1 - \lambda) \begin{vmatrix} -\lambda & 4 \\ -1 & 4 - \lambda \end{vmatrix} - 1 \cdot \begin{vmatrix} 4 & 4 \\ 1 & 4 - \lambda \end{vmatrix} + (-2) \cdot \begin{vmatrix} 4 & -\lambda \\ 1 & -1 \end{vmatrix} \\
 &= (1 - \lambda)[-\lambda(4 - \lambda) + 4] - [4(4 - \lambda) - 4] - 2[-4 + \lambda] \\
 &= (1 - \lambda)[-4\lambda + \lambda^2 + 4] - [16 - 4\lambda - 4] + 8 - 2\lambda \\
 &= (1 - \lambda)[-4\lambda + \lambda^2 + 4] - 16 + 4\lambda + 4 + 8 - 2\lambda \\
 &= -4\lambda + \lambda^2 + 4 + 4\lambda^2 - \lambda^3 - 4\lambda - 16 + 4\lambda + 12 - 2\lambda \\
 &= -\lambda^3 + 5\lambda^2 - 6\lambda \\
 &= -\lambda(\lambda^2 - 5\lambda + 6) \\
 &= -\lambda(\lambda - 3)(\lambda - 2)
 \end{aligned}$$

3. La ecuación característica se obtiene igualando el polinomio característico a cero

$$f(\lambda) = 0$$

$$\begin{aligned}
 \Leftrightarrow -\lambda(\lambda - 3)(\lambda - 2) &= 0 \\
 \Rightarrow \lambda_1 &= 0 \\
 \lambda_2 &= 3 \\
 \lambda_3 &= 2
 \end{aligned}$$

y así obtuvimos los valores propios de la matriz A.

4. Corresponde buscar los vectores propios de A, para eso se recuerda el sistema inicial $(A - \lambda I)x = 0$, representado por la matriz ampliada:

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

reemplazando $\lambda_1 = 0$

$$\begin{aligned}
 \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 4 & 0 & 4 & 0 \\ 1 & -1 & 4 & 0 \end{array} \right) E_2\left(\frac{1}{4}\right) &\sim \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & 4 & 0 \end{array} \right) E_{12}(-1)E_{32}(-1) \sim \\
 \left(\begin{array}{ccc|c} 0 & 1 & -3 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 3 & 0 \end{array} \right) E_{31}(1) &\sim \left(\begin{array}{ccc|c} 0 & 1 & -3 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)
 \end{aligned}$$

se obtiene el nuevo sistema en su forma reducida:

$$\begin{aligned}
 x + z &= 0 \\
 y - 3z &= 0 \\
 \text{de donde } x &= -z \\
 y &= 3z
 \end{aligned}$$

Por lo tanto

$$\begin{aligned} S_{\lambda=0} &= \{(-z, 3z, z) \in \mathbb{R}^3 / z \neq 0\} \\ S_{\lambda=0} &= \langle (-1, 3, 1) \rangle \end{aligned}$$

reemplazando $\lambda_2 = 3$

$$\begin{aligned} \left(\begin{array}{ccc|c} -2 & 1 & -2 & 0 \\ 4 & -3 & 4 & 0 \\ 1 & -1 & 1 & 0 \end{array} \right) E_{13}(2)E_{23}(-4) &\sim \left(\begin{array}{ccc|c} 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{array} \right) E_{12}(1) \sim \\ \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{array} \right) E_{32}(1) &\sim \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array} \right) \end{aligned}$$

se obtiene el nuevo sistema en su forma reducida:

$$\begin{aligned} y &= 0 \\ x + z &= 0 \\ \text{de donde } x &= -z \end{aligned}$$

Por lo tanto

$$\begin{aligned} S_{\lambda=3} &= \{(-z, 0, z) \in \mathbb{R}^3 / z \neq 0\} \\ S_{\lambda=3} &= \langle (-1, 0, 1) \rangle \end{aligned}$$

reemplazando $\lambda_3 = 2$

$$\left(\begin{array}{ccc|c} -1 & 1 & -2 & 0 \\ 4 & -2 & 4 & 0 \\ 1 & -1 & 2 & 0 \end{array} \right) E_{13}(1)E_{23}(-4) \sim \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 2 & -4 & 0 \\ 1 & -1 & 2 & 0 \end{array} \right)$$

se obtiene el sistema en su forma reducida:

$$\begin{aligned} 2y - 4z &= 0 \\ x - y + 2z &= 0 \\ \text{de donde } y &= 2z \\ x &= 0 \end{aligned}$$

Por lo tanto

$$\begin{aligned} S_{\lambda=2} &= \{(0, 2z, z) \in \mathbb{R}^3 / z \neq 0\} \\ S_{\lambda=2} &= \langle (0, 2, 1) \rangle \end{aligned}$$

Finalmente hemos encontrado los vectores propios correspondientes a la matriz A.

5.4. Formas Cuadráticas

Teorema 5.1 *Los valores propios de una matriz triangular son las componentes de la diagonal de la matriz.*

Demostración:

Si A es triangular superior tiene la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

entonces su matriz característica es $A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} - \lambda \end{pmatrix}$

y como Ud. lector puede demostrar previamente el determinante de una matriz triangular es igual al producto de los componentes de la diagonal. Eso implica que:

$$|A - \lambda I| = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda)$$

Luego las soluciones del polinomio característico $f(\lambda) = 0$ son:

$$\begin{aligned} \lambda_1 &= a_{11} \\ \lambda_2 &= a_{22} \\ &\vdots \\ \lambda_n &= a_{nn} \end{aligned}$$

La demostración para una matriz triangular inferior es idéntica.

Por lo tanto queda demostrado el teorema.

5.4.1. Matrices Semejantes y Diagonalización

Definición 5.3 Se dice que dos matrices A y B de $n \times n$ son semejantes si existe una matriz invertible C de $n \times n$ tal que: $B = C^{-1}AC$

Observación 5.1 De la definición se deduce que: $B = C^{-1}AC$, así que si se multiplica por C a la izquierda

$$\begin{aligned} CB &= C(C^{-1}AC) \\ &= (CC^{-1})AC \\ &= I_n AC \\ &= AC \end{aligned}$$

que constituye una definición alternativa de semejanza.

Definición 5.4 (Definición Alternativa) Dos matrices A y B son semejantes \Leftrightarrow existe una matriz invertible C tal que $CB = AC$

Ejemplo 5.1 Verifiquemos que A y B son semejantes; donde

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

Entonces

$$\begin{aligned}
 C \cdot B &= \begin{pmatrix} 4 & -1 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix} \\
 &= \begin{pmatrix} 20 & 2 \\ 15 & -2 \end{pmatrix} \\
 A \cdot C &= \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -1 \\ 3 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 20 & 2 \\ 15 & -2 \end{pmatrix} \\
 &\therefore CB = AC
 \end{aligned}$$

lo que verifica que A y B son semejantes.

Teorema 5.2 *Dos matrices A y B se dicen semejantes de $n \times n$ si representan la misma transformación lineal. Entonces A y B tienen el mismo polinomio característico, y por lo tanto los mismos valores propios.*

Hipótesis: A y B son semejantes

Tesis: $|A - \lambda I_n| = |B - \lambda I_n|$

Demostración:

Por hipótesis A y B son semejantes $\Rightarrow B = C^{-1}AC$

$$\begin{aligned}
 |B - \lambda I_n| &= |C^{-1}AC - \lambda I_n| \\
 &= |C^{-1}AC - I_n \lambda I_n| \\
 &= |C^{-1}AC - C^{-1}C \lambda I_n| \\
 &= |C^{-1} \cdot | \cdot | AC - C \lambda I_n | \\
 &= |C^{-1} \cdot | \cdot | AC - \lambda C I_n | \\
 &= |C^{-1} \cdot | \cdot | AC - \lambda I_n C | \\
 &= |C^{-1} \cdot | \cdot | A - \lambda I_n | \cdot | C | \\
 &= |C^{-1} \cdot | \cdot | C | \cdot | A - \lambda I_n | \\
 &= |C^{-1} \cdot C| \cdot | A - \lambda I_n | \\
 &= |I_n| \cdot | A - \lambda I_n | \\
 &= |A - \lambda I_n|
 \end{aligned}$$

Por lo tanto queda demostrado el teorema.

Observación 5.2 *Queda al lector verificar que los valores propios de una matriz diagonal son los componentes de la diagonal.*

Definición 5.5 *Una matriz A de $n \times n$ es diagonalizable si existe una matriz diagonal D tal que A es semejante a D .*

Corolario: Si A es diagonalizable $\Rightarrow A$ es semejante a una matriz diagonal cuyas componentes en la diagonal son los valores propios de A .

Teorema 5.3 Una matriz A de $n \times n$ es diagonalizable \Leftrightarrow tiene n vectores linealmente independientes. En tal caso, la matriz diagonal D semejante a A , está dada por

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

donde $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores propios de A .

Si C es una matriz cuyas columnas son vectores propios linealmente independientes de A

$$\Rightarrow D = C^{-1}AC.$$

5.4.2. Matrices Simétricas y Diagonalización Ortogonal

Teorema 5.4 Sea A una matriz simétrica real de $n \times n$. Entonces los valores propios de A son reales.

Teorema 5.5 Sea A una matriz simétrica real de $n \times n$. Si λ_1 y λ_2 son valores propios diferentes con vectores propios reales correspondientes v_1 y v_2 , entonces v_1 y v_2 son ortogonales.

Demostración: Si λ_1 es un valor propio de A , entonces:

$$\begin{aligned} Av_1 &= \lambda_1 v_1 \\ (Av_1) \cdot v_2 &= (\lambda_1 v_1) \cdot v_2 \\ A(v_1 \cdot v_2) &= \lambda_1 (v_1 \cdot v_2) \\ & y \\ A(v_1 \cdot v_2) &= v_1 (A^t v_2) \\ &= v_1 (A v_2) \\ &= v_1 (\lambda_2 v_2) \\ &= \lambda_2 (v_1 \cdot v_2) \end{aligned}$$

Así

$$\lambda_1 (v_1 \cdot v_2) = \lambda_2 (v_1 \cdot v_2)$$

eso implica que $v_1 \cdot v_2 = 0$ pues $\lambda_1 \neq \lambda_2$ ya que son valores propios diferentes.

Así queda demostrado que v_1 y v_2 son ortogonales.

Definición 5.6 Se dice que una matriz A de $n \times n$ es diagonalizable ortogonalmente si existe una matriz Q tal que

$$Q^t A Q = D$$

donde D es una matriz diagonal, cuyos elementos de la diagonal son $\lambda_1, \lambda_2, \dots, \lambda_n$ (los valores propios de A), y que se denota $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$

Teorema 5.6 Sea A una matriz real de $n \times n$. Entonces A es diagonalizable ortogonalmente \Leftrightarrow A es simétrica

5.4.3. Formas Cuadráticas

Definición 5.7 Una Ecuación Cuadrática en dos variables sin términos lineales, es una ecuación de la forma

$$ax^2 + bxy + cy^2 = d$$

donde al menos uno de los coeficientes a, b, c es diferente de cero. Mientras que una forma cuadrática en dos variables es una expresión de la forma

$$F(x, y) = ax^2 + bxy + cy^2$$

donde al menos uno de los coeficientes a, b, c es diferente de cero.

Si $F(x, y) = ax^2 + bxy + cy^2$ es una forma cuadrática, se puede reescribir como

$$F(x, y) = v^t A v$$

donde $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ y $v = \begin{pmatrix} x \\ y \end{pmatrix}$

En efecto tenemos:

$$\begin{aligned} v^t A v &= \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} xa + y\frac{b}{2} & x\frac{b}{2} + yc \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \left(xa + y\frac{b}{2}\right) \cdot x + \left(x\frac{b}{2} + yc\right) \cdot y \\ &= ax^2 + \frac{b}{2}xy + \frac{b}{2}xy + cy^2 \\ &= ax^2 + bxy + cy^2 \\ &= F(x, y) \end{aligned}$$

donde A es una matriz simétrica asociada a esta forma cuadrática, denominada matriz de la forma cuadrática.

Ejemplo 5.2 Escribiremos la siguiente forma cuadrática en términos matriciales:

$$F(x, y) = -x^2 + 2xy - y^2$$

Note que $a = -1, b = 2, c = -1$, entonces la forma matricial de la forma cuadrática es:

$$F(x, y) = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

Como este caso es simplista, se generaliza al caso en que hay término libre, para eso abordaremos el problema como una ecuación cuadrática: $v^t A v = d$.

Como A es simétrica, existe una matriz ortogonal C tal que $C^t A C$ es una matriz diagonal D (cuyos elementos de la diagonal son los valores propios de A).

De lo anterior tenemos

$$\begin{aligned} C^t AC &= D \\ C(C^t AC) &= CD \\ (CC^t)(AC) &= CD \\ I_n(AC) &= CD \\ (AC)C^{-1} &= (CD)C^{-1} \\ A(CC^{-1}) &= CDC^{-1} \\ AI_n &= CDC^{-1} \\ A &= CDC^t \end{aligned}$$

Luego tenemos que:

$$v^t Av = d \Rightarrow v^t CDC^t v = d.$$

Estas matrices definen una rotación de coordenadas $v^t = C^t v$ del sistemas de coordenadas rectangulares xy a un sistema de coordenadas $x'y'$; en la cual falta el término $x'y'$.

En el sistema de coordenadas $x'y'$, la ecuación $v^t Av = d$ se convierte en

$$v'^t D v' = d$$

Sean $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ $v' = \begin{pmatrix} x' \\ y' \end{pmatrix}$

Entonces

$$\begin{aligned} \begin{pmatrix} x' & y' \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= d \\ \begin{pmatrix} \lambda_1 \cdot x' & \lambda_2 \cdot y' \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= d \\ \lambda_1 \cdot (x')^2 + \lambda_2 \cdot (y')^2 &= d \end{aligned}$$

Si $d \neq 0 \Rightarrow$

$$\frac{(x')^2}{\frac{d}{\lambda_1}} + \frac{(y')^2}{\frac{d}{\lambda_2}} = 1$$

esta es la ecuación de una cónica.

Observación 5.3 Si λ_1 y λ_2 tienen el mismo signo, entonces $ax^2 + bxy + cy^2 = d$ define una elipse (o un círculo) o una cónica degenerada. Si λ_1 y λ_2 tienen signo contrario, entonces $ax^2 + bxy + cy^2 = d$ es la ecuación de una hipérbola.

Teorema 5.7 (De los ejes principales en \mathbb{R}^2) Sea $ax^2 + bxy + cy^2 = d$ una ecuación cuadrática en las variables (x, y) . Entonces existe un número único θ en $[0, 2\pi[$ tal que la ecuación se puede reescribir de la forma $\lambda_1 \cdot (x')^2 + \lambda_2 \cdot (y')^2 = d$ donde (x', y') son los ejes obtenidos al rotar los ejes (x, y) en un ángulo θ en el sentido contrario a las manecillas del reloj y donde λ_1 y λ_2 son los valores propios de la matriz

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

Los ejes (x', y') reciben el nombre de ejes principales de la ecuación cuadrática.

Demostración:

Como C es una matriz real y ortogonal, entonces $C \cdot C^{-1} = I_n$

$$\begin{aligned} C \cdot C^{-1} &= I_n \\ \Rightarrow \det(C \cdot C^{-1}) &= 1 \\ &= \det(C \cdot C^t) \\ &= \det(C) \cdot \det(C^t) \\ &= \det(C) \cdot \det(C) \\ &= \det(C)^2 \\ \Rightarrow \det(C) &= \pm 1 \end{aligned}$$

Si $\det(C) = -1$ se pueden intercambiar las filas de C para hacer $\det(C) = 1$. Queda al lector demostrar que $C = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ para algún $\theta \in [0, 2\pi[$. Como puede notar C es una matriz de rotación.

Por lo tanto queda demostrado el teorema.

Teorema 5.8 Si $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ entonces la ecuación cuadrática $ax^2 + bxy + cy^2 = d$ es la ecuación de:

1. Una hipérbola si $d \neq 0$ y $\det(A) < 0$
2. Una elipse, círculo o sección cónica degenerada si $d \neq 0$ y $\det(A) > 0$
3. Un par de rectas o una sección cónica degenerada si $d \neq 0$ y $\det(A) = 0$
4. Si $d = 0 \Rightarrow ax^2 + bxy + cy^2 = d$ es la ecuación de dos rectas si $\det(A) \neq 0$ y la ecuación de una sola recta si $\det(A) = 0$

Demostración

La ecuación característica de A es

$$\begin{aligned} (A - \lambda I) &= \begin{vmatrix} a - \lambda & b/2 \\ b/2 & c - \lambda \end{vmatrix} \\ &= (a - \lambda)(c - \lambda) - \frac{b^2}{4} \\ &= \lambda^2 - (a + c)\lambda + \left(ac - \frac{b^2}{4}\right) \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2) \end{aligned}$$

De aquí:

$$\lambda_1 \cdot \lambda_2 = ac - \frac{b^2}{4}$$

- Si $d \neq 0$ y $\det(A) < 0 \Rightarrow ac - \frac{b^2}{4} < 0$

Por lo que si $ac - \frac{b^2}{4} < 0 \Rightarrow \lambda_1$ y λ_2 tienen distinto signo y la ecuación $ax^2 + bxy + cy^2 = d$ es la ecuación de una hipérbola.

- Si $d \neq 0$ y $\det(A) > 0$

Entonces λ_1 y λ_2 tienen el mismo signo y por tanto la ecuación $ax^2 + bxy + cy^2 = d$ sería la ecuación de una elipse, círculo o sección cónica degenerada.

- Si $d \neq 0$ y $\det(A) = 0$

$\Rightarrow ac - \frac{b^2}{4} = 0$ y por tanto $\lambda = 0$ es un valor propio de A; así la ecuación $\lambda_1 x^2 + \lambda_2 y^2 = d$ se convierte en $\lambda_1 x^2 = d$ ó $\lambda_2 y^2 = d$. Sin pérdida de generalidad supongamos que $\lambda_1 x^2 = d$ tenemos $x = \pm \sqrt{d/\lambda_1}$ que es la ecuación de dos rectas en el plano \mathbb{R}^2 . Note que d/λ_1 no puede ser menor que cero, pues implica que $x^2 < 0 \Rightarrow \Leftarrow$. Así tendríamos una cónica degenerada; lo mismo sucede al analizar $\lambda_2 y^2 = d$.

- Si $d = 0$ tenemos $\lambda_1(x')^2 + \lambda_2(y')^2 = 0$ si $\det(A) \neq 0 \Rightarrow \lambda_1 \neq 0$ y $\lambda_2 \neq 0$ y además λ_1 y λ_2 no pueden tener el mismo signo pues implica que $(x', y') = (0, 0)$. En caso de que λ_1 y λ_2 tienen signos opuestos se tiene:

$$\begin{aligned} \lambda_1(x')^2 &= -\lambda_2(y')^2 \\ x' &= \pm \sqrt{\frac{-\lambda_2}{\lambda_1}} y' \end{aligned}$$

que son las ecuaciones de un par de rectas en el plano xy .

Y si $d = 0$ y $\det(A) = 0 \Rightarrow$ que λ_1 ó λ_2 es igual a cero y por lo tanto $x' = 0$ ó $y' = 0$ que es la ecuación de la recta.

Por lo tanto queda demostrado el Teorema.

Ejemplo 5.3 Identificar la cónica que representa la ecuación

a) $17x^2 - 12xy + 8y^2 = 20$

Identificando los coeficientes: $a = 17, b/2 = -6, c = 8$ y $d = 20$

La matriz que representa la ecuación anterior es $\Rightarrow A = \begin{pmatrix} 17 & -6 \\ -6 & 8 \end{pmatrix}$

Gracias a los teoremas demostrados anteriormente se sabe que A es semejante a una matriz D (diagonal) cuya forma es $D = \text{diag}(\lambda_1, \lambda_2)$ donde los λ_i son los valores propios de A respectivamente.

Buscando los valores propios de A.

La matriz característica de A es

$$|A - \lambda I| = \begin{vmatrix} 17 - \lambda & -6 \\ -6 & 8 - \lambda \end{vmatrix} = (17 - \lambda)(8 - \lambda) - 36$$

Ahora el polinomio característico $f(\lambda) = 0 \Leftrightarrow$

$$\begin{aligned} 136 - 17\lambda - 8\lambda + \lambda^2 - 36 &= 0 \\ \lambda^2 - 25\lambda + 100 &= 0 \\ \lambda &= \frac{25 \pm \sqrt{25^2 - 400}}{2} \\ \lambda &= \frac{25 \pm \sqrt{225}}{2} \\ \lambda &= \frac{25 \pm 15}{2} \end{aligned}$$

Así los valores propios son los ceros del polinomio característico:

$$\begin{aligned}\lambda_1 &= 5 \\ \lambda_2 &= 20\end{aligned}$$

Finalmente se escribe la nueva representación de la ecuación:

$$\begin{aligned}(x' \ y') \cdot \begin{pmatrix} 5 & 0 \\ 0 & 20 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 20 \\ (5x' \ 20y') \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 20 \\ 5(x')^2 + 20(y')^2 &= 20 \\ \frac{(x')^2}{4} + \frac{(y')^2}{1} &= 20\end{aligned}$$

que es la ecuación de una elipse.

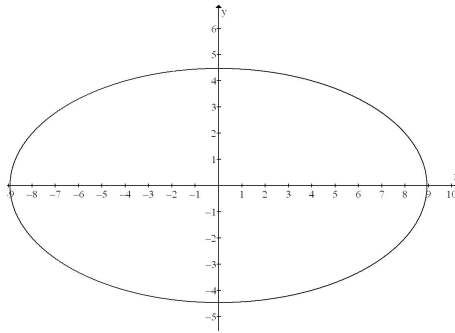


Figura 5.1: Elipse

Como esta ecuación está en términos de las nuevas variables (x', y') debemos encontrar la matriz de rotación para así determinar en cuántos grados se ha rotado la elipse. Para eso encontraremos los vectores propios correspondientes.

Nuestro sistema original era:

$$\begin{pmatrix} 17 - \lambda & -6 \\ -6 & 8 - \lambda \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego para encontrar el vector propio correspondiente a $\lambda = 5$ se debe reemplazar en la matriz $\begin{pmatrix} 17 - 5 & -6 & | & 0 \\ -6 & 8 - 5 & | & 0 \end{pmatrix} = \begin{pmatrix} 12 & -6 & | & 0 \\ -6 & 3 & | & 0 \end{pmatrix} E_1(\frac{1}{12}) \sim \begin{pmatrix} 1 & -1/2 & | & 0 \\ -6 & 3 & | & 0 \end{pmatrix} E_{21}(6) \sim \begin{pmatrix} 1 & -1/2 & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}$

$$\begin{aligned}\Rightarrow x - 1/2y &= 0 \\ 2x &= y \\ \therefore S_{\lambda=5} &= \{(x, 2x)/x \in \mathbb{R}, x \neq 0\}\end{aligned}$$

Luego si $x = 1 \Rightarrow v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ es un vector propio.

Si reemplazamos $\lambda = 20$ en el sistema tenemos:

$$\begin{aligned} \left(\begin{array}{cc|c} 17-20 & -6 & 0 \\ -6 & 8-20 & 0 \end{array} \right) &= \left(\begin{array}{cc|c} -3 & -6 & 0 \\ -6 & -12 & 0 \end{array} \right) E_1\left(\frac{-1}{3}\right) \sim \left(\begin{array}{cc|c} 1 & 2 & 0 \\ -6 & -12 & 0 \end{array} \right) E_{21}(6) \sim \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array} \right) \\ \Rightarrow x + 2y &= 0 \\ y &= x/2 \\ \therefore S_{\lambda=20} &= \{(x, -x/2)/x \in \mathbb{R}, x \neq 0\} \end{aligned}$$

Luego si $x = 1 \Rightarrow v_2 = \begin{pmatrix} 1 \\ -1/2 \end{pmatrix}$ es un vector propio.

Al normalizar estos vectores propios y escribirlos como las columnas de una matriz ortogonal C , se tiene:

$$C = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}$$

La rotación de coordenadas está definida por esta matriz, por lo que se concluye

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}$$

Esto nos dice que $\cos \theta = 1/\sqrt{5}$, $\sin \theta = 2/\sqrt{5}$ y $\therefore \theta = 63,4^\circ$

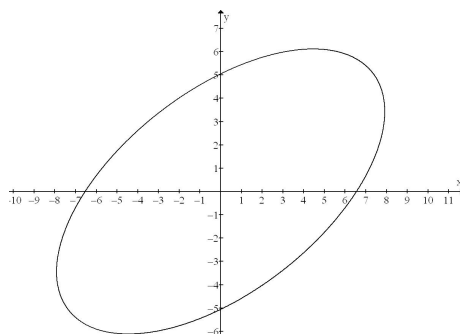


Figura 5.2: Elipse rotada en $63,4^\circ$

Así la coordenada de rotación es aproximadamente $63,4^\circ$.

b) $xy = 2$

Identificando los coeficientes: $a = 0, b/2 = 1/2, c = 0$ y $d = 2$

La matriz que representa la ecuación anterior es $\Rightarrow A = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$

Gracias a los teoremas demostrados anteriormente se sabe que A es semejante a una matriz D (diagonal) cuya forma es $D = \text{diag}(\lambda_1, \lambda_2)$ donde los λ_i son los valores propios de A respectivamente.

Buscando los valores propios de A .

La matriz característica de A es

$$|A - \lambda I| = \begin{vmatrix} 0 - \lambda & 1/2 \\ 1/2 & 0 - \lambda \end{vmatrix} = (-\lambda)^2 - 1/4$$

Ahora el polinomio característico $f(\lambda) = 0 \Leftrightarrow$

$$\begin{aligned} (-\lambda)^2 - 1/4 &= 0 \\ (\lambda)^2 &= 1/4 \\ \lambda_1 &= 1/2 \\ \lambda_2 &= -1/2 \end{aligned}$$

Así los valores propios son los ceros del polinomio característico.

Finalmente se escribe la nueva representación de la ecuación:

$$\begin{aligned} (x' \ y') \cdot \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 2 \\ (x'/2 \ -y'/2) \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 2 \\ \frac{(x')^2}{2} - \frac{(y')^2}{2} &= 2 \\ \frac{(x')^2}{4} - \frac{(y')^2}{4} &= 1 \end{aligned}$$

que es la ecuación de una hipérbola.

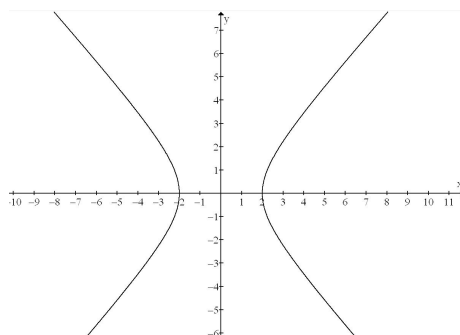


Figura 5.3: Hipérbola

Como esta ecuación está en términos de las nuevas variables (x', y') debemos encontrar la matriz de rotación para así determinar en cuántos grados se ha rotado la hipérbola. Para eso encontraremos los vectores propios correspondientes.

Nuestro sistema original era:

$$\begin{pmatrix} 0 - \lambda & 1/2 \\ 1/2 & 0 - \lambda \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego para encontrar el vector propio correspondiente a $\lambda = 1/2$ se debe reemplazar en la matriz

$$\begin{aligned} \left(\begin{array}{cc|c} 0 - 1/2 & 1/2 & 0 \\ 1/2 & 0 - 1/2 & 0 \end{array} \right) &= \left(\begin{array}{cc|c} -1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \end{array} \right) E_1(2) \sim \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1/2 & -1/2 & 0 \end{array} \right) E_{21}(1/2) \sim \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right) \\ \Rightarrow -x + y &= 0 \\ x &= y \\ \therefore S_{\lambda=1/2} &= \{(x, x) / x \in \mathbb{R}, x \neq 0\} \end{aligned}$$

Luego si $x = 1 \Rightarrow v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ es un vector propio.

Si reemplazamos $\lambda = -1/2$ en el sistema tenemos:

$$\begin{pmatrix} 0 + 1/2 & 1/2 & | & 0 \\ 1/2 & 0 + 1/2 & | & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & | & 0 \\ 1/2 & 1/2 & | & 0 \end{pmatrix} E_{1(2)} \sim \begin{pmatrix} 1 & 1 & | & 0 \\ 1/2 & 1/2 & | & 0 \end{pmatrix} E_{21(-1/2)} \sim \begin{pmatrix} 1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}$$

$$\Rightarrow x + y = 0$$

$$y = -x$$

$$\therefore S_{\lambda=-1/2} = \{(x, -x)/x \in \mathbb{R}, x \neq 0\}$$

Luego si $x = 1 \Rightarrow v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ es un vector propio.

Al normalizar estos vectores propios y escribirlos como las columnas de una matriz ortogonal C , se tiene:

$$C = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

La rotación de coordenadas está definida por esta matriz, por lo que se concluye

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

Esto nos dice que $\cos \theta = 1/\sqrt{2}$, $\sin \theta = 1/\sqrt{2}$ y $\therefore \theta = 45^\circ$

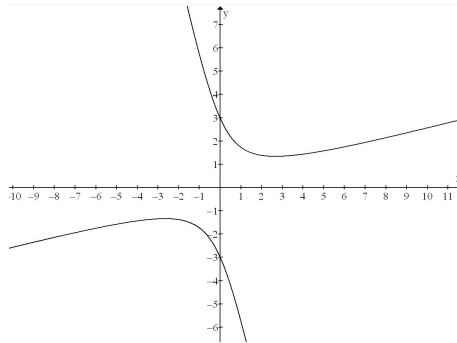


Figura 5.4: Hipérbola rotada en 45°

Así la coordenada de rotación es 45° .

b) $4x^2 - 12xy + 8y^2 = 10$

Identificando los coeficientes: $a = 4, b/2 = -6, c = 8$ y $d = 10$

La matriz que representa la ecuación anterior es $\Rightarrow A = \begin{pmatrix} 4 & -6 \\ -6 & 8 \end{pmatrix}$

El lector puede verificar que el $\det(A) = 0$ a diferencia de los dos casos anteriores. Podrá inferir después de este ejemplo que es trascendental conocer la naturaleza del determinante de A .

Por otro lado se sabe que A es semejante a una matriz D (diagonal) cuya forma es $D = \text{diag}(\lambda_1, \lambda_2)$ donde los λ_i son los valores propios de A respectivamente.

Buscando los valores propios de A .

La matriz característica de A es

$$|A - \lambda I| = \begin{vmatrix} 4 - \lambda & -6 \\ -6 & 8 - \lambda \end{vmatrix} = (4 - \lambda)(8 - \lambda) - 36$$

Ahora el polinomio característico $f(\lambda) = 0 \Leftrightarrow$

$$\begin{aligned} (4 - \lambda)(8 - \lambda) - 36 &= 0 \\ 36 - 12\lambda + \lambda^2 - 36 &= 0 \\ \lambda(\lambda - 12) &= 0 \\ \lambda_1 &= 0 \\ \lambda_2 &= 12 \end{aligned}$$

Así los valores propios son los ceros del polinomio característico.

Finalmente se escribe la nueva representación de la ecuación:

$$\begin{aligned} \begin{pmatrix} x' & y' \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 12 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 10 \\ \begin{pmatrix} 0 & 12y' \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} &= 10 \\ 12(y')^2 &= 10 \\ (y') &= \sqrt{\frac{5}{6}} \\ (y') &= -\sqrt{\frac{5}{6}} \end{aligned}$$

que son las ecuaciones de un par de rectas.

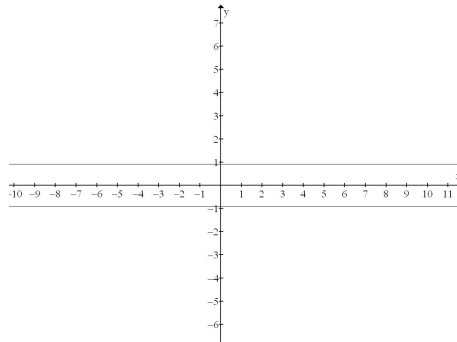


Figura 5.5: Par de rectas

Observación 5.4 De la definición de valor propio y vector propio se considera como un escalar $\lambda \neq 0$ y un vector $v \neq 0$ que cumplen cierta condición.

De hecho el lector puede verificar que

$$S_{\lambda=0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Si reemplazamos $\lambda = 12$ en el sistema tenemos:

$$\left(\begin{array}{cc|c} 4-12 & -6 & 0 \\ -6 & 8-12 & 0 \end{array} \right) = \left(\begin{array}{cc|c} -8 & -6 & 0 \\ -6 & -4 & 0 \end{array} \right) E_1\left(\frac{-1}{8}\right) \sim \left(\begin{array}{cc|c} 1 & 3/4 & 0 \\ -6 & -4 & 0 \end{array} \right) E_{21}(6) \sim \left(\begin{array}{cc|c} 1 & 3/4 & 0 \\ 0 & -13/4 & 0 \end{array} \right)$$

$$\Rightarrow y = 0$$

$$x = 0$$

$$\therefore S_{\lambda=12} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Por ende

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

no es invertible; eso implica que no es una matriz ortogonal y por lo tanto no representa la matriz de rotación.

En general si tenemos una ecuación de la forma

$$ax^2 + bxy + cy^2 + 2dx + 2ey = f$$

intentaremos eliminar los términos libres con el fin de simplificar la ecuación. Para eso haremos una traslación, sustituyendo en la ecuación

$$x = x' + h$$

$$y = y' + k$$

donde (h, k) representa el centro de los ejes de coordenadas (x', y') .

Sustituyendo en la ecuación las variables anteriores tenemos:

$$a(x' + h)^2 + b(x' + h)(y' + k) + c(y' + k)^2 + 2d(x' + h) + 2e(y' + k) = f$$

$$a(x')^2 + 2ax'h + ah^2 + bx'y' + bx'k + by'h + bhk + c(y')^2 + 2cy'k + ck^2 + 2dx' + 2dh + 2ey' + 2ek = f$$

$$a(x')^2 + bx'y' + c(y')^2 + 2x'(ah + b/2k + d) + 2y'(b/2h + ck + e) + (ah^2 + bhk + ck^2 + 2dh + 2ek) = f$$

Observemos que para eliminar los términos lineales se tiene que hacer:

$$ah + b/2k + d = 0 \tag{5.1}$$

$$b/2h + ck + e = 0 \tag{5.2}$$

Resolviendo el sistema queda que:

$$h = \frac{b/2e - dc}{ac - (b/2)^2} \quad k = \frac{ae - b/2d}{ac - (b/2)^2}$$

Después de encontrar el valor de (h, k) se reemplaza en

$$a(x')^2 + bx'y' + c(y')^2 + 2x'(ah + b/2k + d) + 2y'(b/2h + ck + e) + (ah^2 + bhk + ck^2 + 2dh + 2ek) = f$$

y obtiene una ecuación de la forma:

$$a(x')^2 + bx'y' + c(y')^2 = f'$$

a la cual ya sabemos encontrarle su representación como cónica.

Ejemplo 5.4 Identifique la cónica que representa la ecuación:

$$17(x')^2 - 12x'y' + 8(y')^2 + 12x - 16y = 12$$

Acá $a = 17, b/2 = -6, c = 8, d = 6, e = -8$ y $f = 12$

Como hay términos lineales, debemos desaparecerlos y para eso se sustituye

$$\begin{aligned} x &= x' + h \\ y &= y' + k \end{aligned}$$

Luego se encuentra el valor de h y k y se reemplaza en la ecuación original para así eliminarlos y obtener una ecuación más reducida. Sustituyendo:

$$\begin{aligned} 17(x' + h)^2 - 12(x' + h)(y' + k) + 8(y' + k)^2 + 12(x' + h) - 16(y' + k) &= 12 \\ 17(x')^2 + 34x'h + 17h^2 - 12x'y' - 12x'k - 12y'h - 12hk + 8(y')^2 + 16y'k + 8k^2 + 12x' + 12h - 16y' - 16k &= 12 \\ 17(x')^2 - 12x'y' + 8(y')^2 + 2x'(17h - 6k + 6) + 2y'(-6h + 8k - 8) + (17h^2 - 12hk + 8k^2 + 12h - 16k) &= 12 \end{aligned}$$

Para eliminar los términos lineales debemos encontrar (h, k) que satisfaga ambas ecuaciones:

$$17h - 6k + 6 = 0 \tag{5.3}$$

$$-6h + 8k - 8 = 0 \tag{5.4}$$

Al resolver este sistema homogéneo se ha encontrado que $(h, k) = (0, 1)$, por lo tanto la ecuación

$$17(x')^2 - 12x'y' + 8(y')^2 + 12x - 16y = 12$$

se transforma en

$$17(x')^2 - 12x'y' + 8(y')^2 = 20$$

que como Ud. puede reconocer es uno de los ejemplos anteriores y que se mostró representa la ecuación de una elipse.

Observación 5.5 En caso de que la ecuación

$$ax^2 + bxy + cy^2 + 2dx + 2ey = f$$

tenga los coeficientes $a = c = 1$ y $b = 0$, entonces representa la ecuación de un círculo. En efecto:

$$x^2 + y^2 + 2dx + 2ey = f$$

es lo mismo que

$$(x + d)^2 + (y + e)^2 = f + d^2 + e^2$$

y ahora haciendo la traslación

$$\begin{aligned} x' &= x + d \\ y' &= y + e \\ r &= \sqrt{f + d^2 + e^2} \end{aligned}$$

obtenemos

$$(x')^2 + (y')^2 = r^2$$

que es la ecuación de un círculo.

En el ejemplo siguiente el lector puede notar que la ecuación está en términos de tres variables, por lo que se trata de una Superficie Cuadrática.

Ejemplo 5.5

$$5x^2 + 6y^2 + 7z^2 + 4xy + 4yz - 1 = 0$$

Sea

$$a = 5 \quad \frac{b}{2} = 2 \quad c = 6 \quad d = 0 \quad e = 2 \quad f = 7$$

$$A = \begin{pmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{pmatrix} \quad y \quad v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

En efecto

$$\begin{aligned} v^t A v &= (x \ y \ z) \begin{pmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= (5x + 2y \quad 2x + 6y + 2z \quad 2y + 7z) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= (5x + 2y)x + (2x + 6y + 2z)y + (2y + 7z)z \\ &= 5x^2 + 2xy + 2xy + 6y^2 + 2yz + 2yz + 7z^2 \\ &= 5x^2 + 4xy + 6y^2 + 4yz + 7z^2 \\ &= F(x, y, z) \end{aligned}$$

Por lo tanto debemos diagonalizar A

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 5 - \lambda & 2 & 0 \\ 2 & 6 - \lambda & 2 \\ 0 & 2 & 7 - \lambda \end{vmatrix} = (5 - \lambda) \begin{vmatrix} 6 - \lambda & 2 \\ 2 & 7 - \lambda \end{vmatrix} - 2 \begin{vmatrix} 2 & 2 \\ 0 & 7 - \lambda \end{vmatrix} \\ &= (5 - \lambda)[(6 - \lambda)(7 - \lambda) - 4] - 2[2(7 - \lambda) - 0] \\ &= (5 - \lambda)[42 - 6\lambda - 7\lambda + \lambda^2 - 4] - 2[14 - 2\lambda] \\ &= (5 - \lambda)[\lambda^2 - 13\lambda + 38] - 28 + 4\lambda \\ &= 5\lambda^2 - 65\lambda + 190 - \lambda^3 + 13\lambda^2 - 38\lambda - 28 + 4\lambda \\ &= -\lambda^3 + 18\lambda^2 - 99\lambda + 162 \\ &= (\lambda - 3)(-\lambda^2 + 15\lambda - 54) \end{aligned}$$

de donde $f(\lambda) = 0$, obtenemos $\lambda_1 = 3, \lambda_2 = 9, \lambda_3 = 6$

$$\begin{aligned} \Rightarrow (x' \quad y' \quad z') \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} &= 1 \\ (3x' \quad 6y' \quad 9z') \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} &= 1 \\ 3(x')^2 + 6(y')^2 + 9(z')^2 &= 1 \\ \frac{(x')^2}{1/3} + \frac{(y')^2}{1/6} + \frac{(z')^2}{1/9} &= 1 \end{aligned}$$

que es la ecuación canónica de un elipsoide.

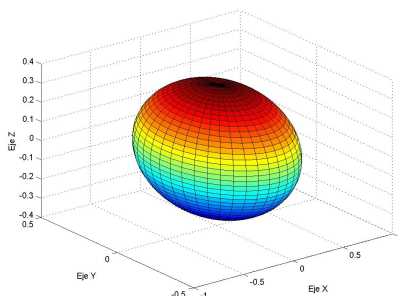


Figura 5.6: Elipsoide $\frac{(x')^2}{1/3} + \frac{(y')^2}{1/6} + \frac{(z')^2}{1/9} = 1$

5.5. Aplicaciones

5.5.1. Un Modelo de Crecimiento de Población

Se supone que cierta especie crece a una tasa constante; es decir, la población de la especie después de un período (entendiéndose este como una hora, un mes, una semana, un año, etc.) es un múltiplo constante de la población del período anterior. Una forma de que esto suceda, por ejemplo, es que cada generación es distinta y cada organismo produce r -críos y después muere. Si p_n denota la población después de n -períodos, se puede tener $p_n = r \cdot p_{n-1}$.

Este modelo puede describir el clásico ejemplo de una población de bacterias, donde, en un tiempo dado, un organismo se divide en dos organismos separados. Entonces $r = 2$. Si p_0 es la población inicial, entonces $p_1 = r \cdot p_0$, $p_2 = r \cdot p_1 = r \cdot (r \cdot p_0) = r^2 \cdot p_0$ y así sucesivamente, de manera que

$$p_n = r^n \cdot p_0 \tag{5.5}$$

De este modo se ve que la población aumenta sin cota si $r > 1$ y disminuye a cero si $r < 1$. Si $r = 1$ la población se mantiene en un valor constante p_0 .

Queda en evidencia la simplicidad de este modelo, puesto que no considera que el número de críos depende de las edades de los adultos como lo que sucede con la población humana, donde las hembras adultas de más de 50 años promedio producen considerablemente menos que las hembras de 18 años

promedio. Para manejar esta dificultad, se introduce un modelo que permita agrupar por edades y asignar tasas de fertilidad diferentes.

Mostraremos un modelo de crecimiento de población para una especie de venados. En esta población supondremos que el número de venados hembras es igual al número de venados machos. Sea $p_{j,n-1}$ la población juvenil (inmadura) de hembras en el año $(n - 1)$ y sea $p_{a,n-1}$ el número de hembras adultas en el mismo año. Algunos de los venados jóvenes morirán durante el año. Supondremos que cierta proporción α de los venados jóvenes sobrevivirán para llegar a adultos en la primavera del año n . Cada hembra que sobrevive engendra, en promedio, κ venados hembras jóvenes en la siguiente primavera. Los adultos también mueren y la proporción de adultos que sobreviven de una primavera a otra es β .

Esta tasa constante de supervivencia de los venados no es una suposición simplista, a diferencia de la que denota el crecimiento de las bacterias. Este modelo ocurre en una gran cantidad de poblaciones de pájaros en su estado natural. Esto significa que la tasa de supervivencia de los adultos en muchas especies es independiente de la edad.

En la notación introducida $p_{j,n}$ y $p_{a,n}$ representan la población de hembras jóvenes y adultas en el año, respectivamente. Ahora que contamos con toda la información construimos el siguiente sistema de 2×2 .

$$\begin{aligned} p_{j,n} &= \kappa p_{a,n-1} \\ p_{a,n} &= \alpha p_{j,n-1} + \beta p_{a,n-1} \end{aligned}$$

$$p_n = A \cdot p_{n-1} \tag{5.6}$$

donde $p_n = \begin{bmatrix} p_{j,n} \\ p_{a,n} \end{bmatrix}$ y $A = \begin{bmatrix} 0 & \kappa \\ \alpha & \beta \end{bmatrix}$

Es evidente de (5.6) que

$$\begin{aligned} p_1 &= A \cdot p_0 \\ p_2 &= A \cdot p_1 = A \cdot (A \cdot p_0) = A^2 \cdot p_0 \\ &\vdots = \vdots \\ p_n &= A^n \cdot p_0 \end{aligned} \tag{5.7}$$

donde p_0 es el vector de las poblaciones iniciales de hembras jóvenes y adultas.

Si observan la ecuación (5.7) es bastante parecida a la ecuación (5.5), con la diferencia que ahora se puede distinguir entre las tasas de supervivencia de venados jóvenes y adultos.

Suponga que la información que sigue representa una población de venados hembras:

$$A = \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \text{ y } p_0 = \begin{bmatrix} 100 \\ 200 \end{bmatrix} \text{ Demostraremos que a la larga la población crecerá por un}$$

factor aproximado de 1,27 y nos colocaremos en el caso en que los granjeros y otras personas del área no quieren que la población crezca. Para esto permitieron la cosecha (caza de los venados). Si h es la proporción de población cosechada en cada período, analizaremos por

qué la matriz de este modelo sería $A = \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 - h \end{bmatrix}$ y además demostraremos que $h = 0,6$ es una cosecha demasiado grande; es decir, a la larga la población de venados se extinguiría y eso no es lo que el que queremos. Finalmente seleccionaremos h de tal forma que la población no crezca ni desaparezca.

Primeramente explicaremos la información entregada:

$$A = \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix}$$

significa que cada venado hembra engendra un venado hembra y como se supone que el número de venados machos es igual al número de hembras, entonces tiene al menos dos crías. También podemos ver que tanto $\alpha = 0,6$ como $\beta = 0,8$ están en el intervalo $[0,1]$.

Recordemos que α es la proporción de venados jóvenes que sobrevivirán para llegar a adultos en la primavera del año n y que β es la proporción de adultos que sobreviven de una primavera a la siguiente. Como no es tan probable que sobrevivan los venados jóvenes como los adultos, se debe tener $\alpha < \beta$.

Por otro lado: $p_0 = \begin{bmatrix} 100 \\ 200 \end{bmatrix}$ indica que en un principio hay 200 venados hembras (y 200 venados machos) adultos y 100 venados hembras jóvenes (y 100 venados machos jóvenes).

Analizemos la población:

En el Primer Año: $p_1 = \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 100 \\ 200 \end{bmatrix} = \begin{bmatrix} 200 \\ 220 \end{bmatrix}$ de manera que $p_{j,1} = 200$ y $p_{a,1} = 220$ así el total de población hembras después de un año es 420 y la razón de hembras jóvenes a adultas es $0,90909090 \dots$

$$\begin{aligned} \text{En el Segundo Año: } p_2 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 200 \\ 220 \end{bmatrix} = \begin{bmatrix} 220 \\ 296 \end{bmatrix} \\ \text{En el Tercer Año: } p_3 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 200 \\ 220 \end{bmatrix} = \begin{bmatrix} 220 \\ 296 \end{bmatrix} \\ \text{En el Cuarto Año: } p_4 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 296 \\ 396 \end{bmatrix} = \begin{bmatrix} 396 \\ 473 \end{bmatrix} \\ \text{En el Quinto Año: } p_5 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 396 \\ 473 \end{bmatrix} = \begin{bmatrix} 473 \\ 600 \end{bmatrix} \\ \text{En el Sexto Año: } p_6 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 473 \\ 600 \end{bmatrix} = \begin{bmatrix} 600 \\ 764 \end{bmatrix} \\ \text{En el Séptimo Año: } p_7 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 600 \\ 764 \end{bmatrix} = \begin{bmatrix} 764 \\ 971 \end{bmatrix} \\ \text{En el Octavo Año: } p_8 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 764 \\ 971 \end{bmatrix} = \begin{bmatrix} 971 \\ 1235 \end{bmatrix} \\ \text{En el Noveno Año: } p_9 &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 971 \\ 1235 \end{bmatrix} = \begin{bmatrix} 1235 \\ 1571 \end{bmatrix} \\ \text{En el Décimo Año: } p_{10} &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 1235 \\ 1571 \end{bmatrix} = \begin{bmatrix} 1571 \\ 1998 \end{bmatrix} \end{aligned}$$

Observación 5.6 *Casi todos los $p_{a,i}$ dieron decimal, así que se aproximaron al entero superior, ya que por ejemplo no se puede tener 472,8 venados.*

Año n	N° de Jóvenes $p_{j,n}$	N° de Adultos $p_{a,n}$	Pobl. Total de hembras T_n en el año n	$p_{j,n} / p_{a,n}$	T_n / T_{n-1}
0	100	200	300	0,5
1	200	220	420	0,909090...	1,4
2	220	296	516	0,743...	1,229
3	296	369	665	0,802...	1,289
4	369	473	842	0,780...	1,266
5	473	600	1073	0,788...	1,2743
6	600	764	1364	0,785...	1,2712
7	764	971	1735	0,7868...	1,2719
8	971	1235	2206	0,7862...	1,2715
9	1235	1571	2806	0,7861...	1,2719
10	1571	1998	3569	0,7862...	1,2719

Observación 5.7 *Puede Notar que los cálculos de la última columna se han aproximado.*

Se percibe en la tabla que la razón $p_{j,n}/p_{a,n}$ se acerca a la constante 0,78 mientras que la población total parece aumentar a una tasa constante de 27 % anual. Demostraremos nuestra percepción utilizando los valores y vectores propios.

Primero regresaremos al caso general $p_n = A^n \cdot p_0$. Suponga que A tiene dos valores propios reales y distintos λ_1 y λ_2 con sus vectores propios correspondientes v_1 y v_2 . Como v_1 y v_2 son linealmente independientes, forman una base para \mathbb{R}^2 y se puede escribir entonces

$$p_0 = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \tag{5.8}$$

para algunos números reales α_1 y α_2

Luego se convierte en

$$p_n = A^n(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) \tag{5.9}$$

Pero

$$\begin{aligned} A \cdot v_1 &= \lambda_1 \cdot v_1 \\ A^2 \cdot v_1 &= A \cdot (A \cdot v_1) \\ &= A \cdot (\lambda_1 \cdot v_1) \\ &= \lambda_1 \cdot A \cdot v_1 \\ &= \lambda_1 \cdot (\lambda_1 \cdot v_1) \\ &= \lambda_1^2 \cdot v_1 \\ &\vdots = \vdots \end{aligned}$$

Así se puede ver que:

$$\begin{aligned} A^n \cdot v_1 &= \lambda_1^n \cdot v_1 \\ A^n \cdot v_2 &= \lambda_2^n \cdot v_2 \end{aligned}$$

Por lo tanto

$$p_n = \alpha_1 \cdot \lambda_1^n \cdot v_1 + \alpha_2 \cdot \lambda_2^n \cdot v_2 \tag{5.10}$$

Además sabemos que la ecuación característica de A es $\begin{vmatrix} -\lambda & \kappa \\ \delta & \beta - \lambda \end{vmatrix} = \lambda^2 - \beta \cdot \lambda - \kappa \cdot \delta = 0$

o sea, $\lambda = \frac{\beta \pm \sqrt{\beta^2 + 4 \cdot \delta \kappa}}{2}$ por suposición, $\kappa > 0$, $0 < \delta < 1$ y $0 < \beta < 1$ Entonces: $4\delta\kappa > 0$ y $\beta^2 + 4\delta\kappa > 0$; esto significa que, sin duda, los valores propios son reales y diferentes y que un valor propio, λ_1 es positivo; el otro, λ_2 es negativo y $|\lambda_1| > |\lambda_2|$.

La ecuación (6) se puede escribir como: $p_n = \lambda_1^n \cdot [\alpha_1 \cdot v_1 + (\frac{\lambda_2}{\lambda_1})^n \cdot \alpha_2 \cdot v_2]$

Como $|\frac{\lambda_2}{\lambda_1}| < 1$, es evidente que $(\frac{\lambda_2}{\lambda_1})^n$ se vuelve muy pequeña cuando n crece. Luego para n grande $p_n \approx \alpha_1 \lambda_1^n \cdot v_1$.

Esto quiere decir que, a la larga, la distribución de edades se estabiliza y es proporcional a v_1 . Cada grupo de edad cambiará por un factor de λ_1 cada año. Así, a la larga, la ecuación $p_n = A^n \cdot p_0$ actúa igual que la ecuación $p_n = r^n \cdot p_0$ en el corto plazo, es decir, antes de alcanzar la **estabilidad** los números oscilan. La magnitud de esta oscilación depende de la magnitud de λ_2/λ_1 (que como es negativa se explica la oscilación).

Como nuestro sistema en cuestión $\begin{bmatrix} 0 - \lambda & 1 \\ 0,6 & 0,8 - \lambda \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ es homogéneo, sabemos que tiene solución trivial, pero en nuestro caso queremos infinitas soluciones; para eso, hacemos

$$\begin{aligned} p(\lambda) &= 0 \\ \rightarrow \begin{vmatrix} 0 - \lambda & 1 \\ 0,6 & 0,8 - \lambda \end{vmatrix} &= 0 \\ \rightarrow -\lambda(0,8 - \lambda) - 0,6 &= 0 \\ \rightarrow \lambda^2 - 0,8\lambda - 0,6 &= 0 \\ \lambda &= \frac{0,8 \pm \sqrt{0,8^2 + 4 \cdot 0,6}}{2} \\ &= \frac{0,8 \pm \sqrt{3,04}}{2} \\ &= \frac{0,8 \pm 1,743559577}{2} \end{aligned}$$

Por lo tanto:

$$\lambda_1 = 1,27177979 \text{ y } \lambda_2 = -0,47177979$$

Esto explica el 27% de aumento en la población que se observa en la última columna de la tabla.

Nuevamente recordemos el sistema de ecuaciones:

$$\begin{bmatrix} -\lambda & 1 & | & 0 \\ 0,6 & 0,8 - \lambda & | & 0 \end{bmatrix}$$

para encontrar los vectores propios, reemplazamos $\lambda_1 = 1,27177979$,

$$\begin{bmatrix} -1,27177979 & 1 & | & 0 \\ 0,6 & 0,8 - 1,27177979 & | & 0 \end{bmatrix} \sim \begin{bmatrix} -1,27177979 & 1 & | & 0 \\ 0,6 & -0,47177979 & | & 0 \end{bmatrix} E_{12}(\frac{1,27177979}{0,6})$$

$$\sim \begin{bmatrix} 0 & 0 & | & 0 \\ 0,6 & -0,47177979 & | & 0 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow 0,6 \cdot x - 0,47177979 \cdot y &= 0 \\ \Rightarrow 0,6 \cdot x / 0,47177979 &= y \\ \Rightarrow 1,27177979x &= y \end{aligned}$$

Luego

$$S_{\lambda=1,27177979} = \{(x, 1,27177979 \cdot x) / x \in \mathbb{R}; x \neq 0\}$$

Así dándole el valor $x = 1$, obtenemos $v_1 = \begin{bmatrix} 1 \\ 1,27177979 \end{bmatrix}$ que es un vector propio. De la misma forma si reemplazamos $\lambda = -0,47177979$ tendríamos

$$\begin{aligned} & \left[\begin{array}{cc|c} 0,47177979 & 1 & 0 \\ 0,6 & 0,8 + 0,47177979 & 0 \end{array} \right] \sim \left[\begin{array}{cc|c} 0,47177979 & 1 & 0 \\ 0,6 & 1,27177979 & 0 \end{array} \right] E_{12} \left(\frac{-0,47177979}{0,6} \right) \\ & \sim \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0,6 & 1,27177979 & 0 \end{array} \right] \\ & \Rightarrow 0,6 \cdot x + 1,27177979 \cdot y = 0 \\ & \Rightarrow 0,6 \cdot x / 1,27177979 = y \\ & \Rightarrow -0,47177979x = y \end{aligned}$$

Luego

$$S_{\lambda=-0,47177979} = \{(x, -0,47177979 \cdot x) / x \in \mathbb{R}; x \neq 0\}$$

Así dándole el valor $x = 1$ obtenemos $v_2 = \begin{bmatrix} 1 \\ -0,47177979 \end{bmatrix}$ que es su segundo vector propio.

Observamos que en v_1 se tiene $\frac{1}{1,27177979} = 0,78629965$ esto explica la razón $p_{j,n}/p_{a,n}$ en la quinta columna de la tabla.

Es de gran interés saber si la población crecerá o decrecerá. Notemos que la población aumentará si $\lambda_1 > 1$ y la condición para que esto ocurra es $\frac{\beta \pm \sqrt{\beta^2 + 4 \cdot \delta \cdot \kappa}}{2} > 1$ ó que $\beta^2 + 4 \cdot \alpha \cdot \kappa > (2 - \beta)^2 = 4 - 4 \cdot \beta + \beta^2$. Esto conduce a $4 \cdot \alpha \cdot \kappa > 4 - 4 \cdot \beta$ o sea

$$\kappa > \frac{1 - \beta}{\alpha} \tag{5.11}$$

En nuestro ejemplo, se tenía $\beta = 0,8$ y $\alpha = 0,6$, eso quiere decir que (5.11) se cumple si $\kappa > (1 - 0,8)/0,6$ o sea $\kappa > 0,33333\dots$ y en efecto se puede comprobar que se cumple, pues el κ dado en nuestra matriz A, es $\kappa = 1$.

Observación 5.8 *En resumen si $\kappa > \frac{1-\beta}{\alpha}$ la población aumenta, mientras que si $\kappa < \frac{1-\beta}{\alpha}$ la población disminuye.*

La segunda parte de nuestra aplicación se refiere a demostrar que $h = 0,6$ es una cosecha demasiado grande, es decir, la población se extinguiría.

Veamos lo que sucede si

$$\begin{aligned} h &= 0,6 \\ \rightarrow A &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,8 - 0,6 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 0,6 & 0,2 \end{bmatrix} \\ |A - \lambda I| &= 0 \\ \rightarrow -\lambda(0,2 - \lambda) - 0,6 &= 0 \\ \lambda^2 - 0,2\lambda - 0,6 &= 0 \\ \lambda &= \frac{0,2 \pm \sqrt{0,2^2 + 4 \cdot 0,6}}{2} = \frac{0,2 \pm \sqrt{2,44}}{2} = \frac{0,2 \pm 1,56204994}{2} \end{aligned}$$

$\lambda_1 \approx 0,88$ y $\lambda_2 \approx 0,78$ como ambos λ_1 y λ_2 son menores que 1, implica que la población se extinguiría si se permite $h = 0,6$ como cosecha.

Finalmente para que la población se mantenga $\kappa > (1 - \beta)/\alpha$, esto ocurre si $\kappa = 1$; o sea tendríamos que $1 = (1 - 0,8 + h)/0,6$ es decir, si $h = 0,4$. El lector interesado puede verificar que el valor propio más grande es exactamente 1.

5.5.2. Cadenas de Markov

Si recordamos en el capítulo 2 estudiamos cadenas de markov, además la aplicamos a una situación concreta, ahora se dará una explicación formal del tema en cuanto a la Probabilidad propiamente tal y por ende también se aplicará a otros temas como son la *Demografía* y *Metereología*.

5.5.2.1. Conocimientos Previos

Teorema 5.9 *Sea T una matriz estocástica regular $n \times n$. Entonces cuando $n \rightarrow \infty$ T tiende a una matriz $L_{n \times n}$, de la forma*

$$L = (v \ v \ \dots \ v)$$

donde v es un vector- n de probabilidad, con todos sus elementos mayores que cero.

Así para cualquier matriz estocástica regular existe el Límite L de potencias.

Teorema 5.10 *Si A es una matriz estocástica regular, y L y v tiene la forma mencionada anteriormente, entonces:*

1. *Para cualquier vector x_0 de probabilidad inicial, $T^n x_0$ tiende a v cuando $n \rightarrow \infty$, esto es,*

$$\lim_{n \rightarrow \infty} T^n x_0 = v$$

2. *v es el único vector de probabilidad que satisface*

$$Tv = v$$

Así, v es un vector propio de T , con el valor propio $\lambda = 1$

Definición 5.8 *La matriz A cuadrada es diagonalizable si y sólo si existe una matriz D diagonal tal que A y D resulten semejantes.*

Teorema 5.11 (Potencias de matrices diagonalizables) *Si A es diagonalizable, entonces*

$$A^n = Q \cdot D^n \cdot Q^{-1}$$

Definición 5.9 *Una matriz de transición de orden $n \times n$ es aquella cuyos elementos no son negativos y tal que la suma de los elementos de cada una de sus columnas es uno. También se la suele llamar matriz estocástica o de probabilidad.*

Las matrices de transición se emplean para estudiar una determinada situación de una población clasificada en distintas categorías o estados. Interesa en particular en nuestro texto, estudiar como la distribución de una población clasificada en distintos estados puede cambiar durante un período de tiempo.

5.5.3. Aplicación a la Demografía

Ejemplo 5.6 *Supongamos que en cierta ciudad un residente tiene el 65 % de probabilidad de permanecer en ella después de un año y un 85 % de mudarse a un área suburbana. Con estos datos se responderá a los siguientes enunciados:*

a) Formular una matriz estocástica de transición de probabilidad para este caso.

La matriz de transición es: $T = \begin{pmatrix} 0,65 & 0,15 \\ 0,35 & 0,85 \end{pmatrix}$

b) ¿Cuál es la distribución a largo plazo de una población que vive en esta ciudad y en el área suburbana que la rodea?

Ahora para responder a la pregunta, debemos saber que sucederá en n años.

La matriz queda

$$\begin{pmatrix} 0,65 & 0,15 \\ 0,35 & 0,85 \end{pmatrix}^n \begin{pmatrix} p \\ q \end{pmatrix}$$

Primero notemos que $p + q = 1$, para cumplir la condición de probabilidad del vector inicial.

Por ende, necesitamos el vector inicial cuando $n \rightarrow \infty$ y además la diagonalización de la matriz estocástica T , queda:

$$|T - \lambda I| = \begin{vmatrix} 0,65 - \lambda & 0,15 \\ 0,35 & 0,85 - \lambda \end{vmatrix} = \lambda^2 - 1,5\lambda + 0,5$$

luego el polinomio característico es

$$\lambda^2 - 1,5\lambda + 0,5 = 0$$

las raíces (valores propios) son:

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 1/2 \end{aligned}$$

Para $\lambda_1 = 1$

$$[T - \lambda_1 I | 0] = \left(\begin{array}{cc|c} -0,35 & 0,15 & 0 \\ 0,35 & -0,15 & 0 \end{array} \right) = \left(\begin{array}{cc|c} -7/20 & 3/20 & 0 \\ 7/20 & -3/20 & 0 \end{array} \right) \sim \left(\begin{array}{cc|c} -7 & 3 & 0 \\ 7 & -3 & 0 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & -3/7 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

$$\Rightarrow x - \frac{3}{7}y = 0$$

$$\Rightarrow x = \frac{3}{7}y$$

$$S_{\lambda=1} = \left\{ \left(\frac{3}{7}y, y \right) / y \in \mathbb{R}, y \neq 0 \right\}$$

Luego si $y = 7 \Rightarrow v_1 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$ es un vector propio.

Para $\lambda_2 = 1/2$

$$[T - \lambda_2 I | 0] = \left(\begin{array}{cc|c} 0,15 & 0,15 & 0 \\ 0,35 & 0,35 & 0 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array} \right)$$

$$\begin{aligned} \Rightarrow x + y &= 0 \\ \Rightarrow x &= -y \\ S_{\lambda=1/2} &= \{(-y, y)/y \in \mathbb{R}, y \neq 0\} \end{aligned}$$

Luego si $y = 1 \Rightarrow v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ es un vector propio.

Así hemos encontrado los vectores propios de la matriz Q

Luego, por teorema (5.11) se tiene:

$$\begin{aligned} T^n &= QD^nQ^{-1} \\ T^n &= \begin{pmatrix} 3 & -1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}^n \begin{pmatrix} 3 & -1 \\ 7 & 1 \end{pmatrix}^{-1} \\ \lim_{n \rightarrow \infty} T^n &= \begin{pmatrix} 3 & -1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^n \begin{pmatrix} 3 & -1 \\ 7 & 1 \end{pmatrix}^{-1} \\ \lim_{n \rightarrow \infty} T^n &= \begin{pmatrix} 0,3 & 0,7 \\ 0,3 & 0,7 \end{pmatrix} \end{aligned}$$

Notamos que la observación que $p + q = 1$ es cierta. Entonces concluimos que a largo plazo, el 30 % de los residentes permanecieron en la ciudad y en comparación el 70 % se mudará a una área suburbana que los rodea.

Ejemplo 5.7 *Se considera que transcurre una generación humana cada 20 años aproximadamente; se desea obtener e interpretar la matriz de transición que describe la siguiente situación:*

$$\begin{aligned} E_1 &= \text{Estado 1: pobre} \\ E_2 &= \text{Estado 2: clase media} \\ E_3 &= \text{Estado 3: rico} \end{aligned}$$

De la población del E_1 , el 19 % pasó al E_2 y el 1 % al E_3 .

De la población del E_2 , el 15 % pasó al E_1 y el 10 % al E_3 .

De la población del E_3 , el 5 % pasó al E_1 y el 30 % al E_2 .

La matriz de transición que refleja la situación planteada durante una generación es:

$$T = \begin{pmatrix} 0,80 & 0,15 & 0,05 \\ 0,19 & 0,75 & 0,30 \\ 0,01 & 0,10 & 0,65 \end{pmatrix}$$

Interpretación: Un elemento cualquiera t_{ij} indica con j el estado actual al que pertenece un individuo y con i el estado al que pasa. Es decir, un elemento t_{ij} se conoce como la probabilidad de pasar del estado j al estado i en un período de tiempo. De acuerdo a lo expresado:

$T_{21} = 0,19$ indica que el 19 % pasó del estado $j = 1$ al estado $i = 2$.

$T_{31} = 0,01$ indica que el 1 % pasó del estado $j = 1$ al estado $i = 3$.

$T_{11} = 0,80$ se obtuvo como resultado de la operación $1 \wedge (0,19 + 0,01)$ pues de acuerdo a la definición de matriz de transición cada columna debe sumar uno.

Si T es la matriz de transición que describe una situación reiterada, generación tras generación, x_0 es el vector de orden $n \times 1$ que describe el tamaño de la población inicial al comenzar el estudio y x_n es el vector de orden $n \times 1$ que representa el tamaño de la población luego de transcurrir n períodos iguales de tiempo, se tiene que:

$$\begin{aligned} x_1 &= Tx_0 \\ x_2 &= Tx_1 = T(Tx_0) = T^2x_0 \\ x_3 &= Tx_2 = T(T^2x_0) = T^3x_0 \\ &\vdots = \vdots \\ x_n &= T^n x_0 \end{aligned}$$

Ahora, si T es diagonalizable, entonces

$$T^n = Q \cdot D^n \cdot Q^{-1}$$

Con los datos del Ejemplo, estudiaremos el comportamiento a largo plazo de la distribución de dicha población.

Sea

$$T^n x_0 = \begin{pmatrix} 0,80 & 0,15 & 0,05 \\ 0,19 & 0,75 & 0,30 \\ 0,01 & 0,10 & 0,65 \end{pmatrix}^n \cdot x_0$$

Como vemos, al tratar de diagonalizar la matriz T y encontrar la matriz Q los resultados son números decimales con varias cifras significativas. Para que el cálculo no sea tan engorroso se enunciará solo el polinomio característico, la matriz Q (raíces), D y Q^{-1} .

Observación 5.9 Debido a que la cantidad de cifras significativas eran 16, se han realizado los cálculos utilizando programas computacionales; sin embargo, se presentan acá solo con 4 cifras significativas para mayor claridad.

Resolveremos el sistema con la matriz aumentada $[T - \lambda I|0]$

$$[T - (\lambda)I|0] = \left(\begin{array}{ccc|c} 0,80 - \lambda & 0,15 & 0,05 & 0 \\ 0,19 & 0,75 - \lambda & 0,30 & 0 \\ 0,01 & 0,10 & 0,65 - \lambda & 0 \end{array} \right)$$

Luego el polinomio característico es: $\lambda^3 + 0,6\lambda^2 - 0,6625\lambda - 0,3765 = 0$
 las raíces (Valores propios) son:

$$\begin{aligned} \lambda_1 &= 0,804705 \\ \lambda_2 &= -0,542898 \\ \lambda_3 &= -0,861807 \end{aligned}$$

por lo tanto los vectores propios son: $v_1 = \begin{pmatrix} -0,6201 \\ -0,7495 \\ -0,2319 \end{pmatrix}$, $v_2 = \begin{pmatrix} -0,8137 \\ 0,3479 \\ 0,4657 \end{pmatrix}$ y $v_3 = \begin{pmatrix} 0,3148 \\ -0,8098 \\ 0,4950 \end{pmatrix}$

y tenemos que:

$$Q = \begin{pmatrix} -0,6201 & -0,8137 & 0,3148 \\ -0,7495 & 0,3479 & -0,8098 \\ -0,2319 & 0,4657 & 0,4950 \end{pmatrix}, Q^{-1} = \begin{pmatrix} -0,62437 & -0,62443 & -0,62447 \\ -0,63513 & 0,26591 & 0,83893 \\ 0,30502 & -0,5427 & 0,93838 \end{pmatrix}$$

$$\Rightarrow D = \begin{pmatrix} 1,0000 & 0 & 0 \\ 0 & 0,7072 & 0 \\ 0 & 0 & 0,4928 \end{pmatrix}$$

Finalmente:

$$T^n = Q \cdot \begin{pmatrix} 1,0000 & 0 & 0 \\ 0 & 0,7072 & 0 \\ 0 & 0 & 0,4928 \end{pmatrix}^n \cdot Q^{-1}$$

$$= Q \cdot \begin{pmatrix} 1,0000^n & 0 & 0 \\ 0 & 0,7072^n & 0 \\ 0 & 0 & 0,4928^n \end{pmatrix} \cdot Q^{-1}$$

Cuando $n \rightarrow \infty$, resulta:

$$\lim_{n \rightarrow \infty} T^n = Q \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot Q^{-1}$$

$$\lim_{n \rightarrow \infty} T^n = \begin{pmatrix} -0,6201 & -0,8137 & 0,3148 \\ -0,7495 & 0,3479 & -0,8098 \\ -0,2319 & 0,4657 & 0,4950 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -0,62437 & -0,62443 & -0,62447 \\ -0,63513 & 0,26591 & 0,83893 \\ 0,30502 & -0,5427 & 0,93838 \end{pmatrix}$$

$$\Rightarrow \lim_{n \rightarrow \infty} T^n = \begin{pmatrix} 0,3872 & 0,3872 & 0,3872 \\ 0,4680 & 0,4680 & 0,4680 \\ 0,1448 & 0,1448 & 0,1448 \end{pmatrix}$$

$$\Rightarrow \lim_{n \rightarrow \infty} T^n \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} 0,3872 & 0,3872 & 0,3872 \\ 0,4680 & 0,4680 & 0,4680 \\ 0,1448 & 0,1448 & 0,1448 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}$$

Ahora bien, si se considera como vector inicial $x_0 = \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} 0,5 \\ 0,35 \\ 0,15 \end{pmatrix}$

Observe que nuestro vector inicial cumple la condición de probabilidad

$$E_1 + E_2 + E_3 = 0,5 + 0,35 + 0,15 = 1.$$

$$\text{Resulta que } \lim_{n \rightarrow \infty} T^n x_0 = \begin{pmatrix} 0,3872 & 0,3872 & 0,3872 \\ 0,4680 & 0,4680 & 0,4680 \\ 0,1448 & 0,1448 & 0,1448 \end{pmatrix} \begin{pmatrix} 0,5 \\ 0,35 \\ 0,15 \end{pmatrix} = \begin{pmatrix} 0,3872 \\ 0,4680 \\ 0,1448 \end{pmatrix}$$

Por lo tanto, a largo plazo la población pobre será de 38,72%, la clase media del 46,8% y la rica del 14,48%.

5.5.4. Aplicación a la Metereología

En esta aplicación profundizaremos los aspectos netamente **Estadísticos** y la implicancia en las Cadenas de Markov para pronosticar a través de eventos finitos los sucesos a largo plazo. Para ello recordemos que los procesos aleatorios en realidad son un caso particular de procesos más generales que son las Cadenas de Markov. En esencia, una cadena es un proceso en tiempo discreto en el que una variable aleatoria X_n va cambiando con el paso del tiempo. Las cadenas de Markov tienen la propiedad de que la probabilidad de que $X_n = j$ sólo depende del estado inmediatamente anterior del sistema: X_{n-1} . Cuando en una cadena dichas probabilidades no dependen del tiempo en que se considere, n

$$P(X_n = j | X_{n-1} = i)$$

se denomina *cadena homogénea*, esto es, las probabilidades son las mismas en cada paso.

5.5.4.1. Probabilidades de Transición

En una cadena homogénea finita con m posibles estados E_1, E_2, \dots, E_m se puede introducir la notación

$$p_{ij} = P(X_n = j | X_{n-1} = i)$$

donde $i, j = 1, 2, \dots, m$. Si $p_{ij} > 0$ entonces se dice que el estado E_i puede comunicar con E_j . La comunicación puede ser mutua si también $p_{ji} > 0$.

Para cada i fijo, la serie de valores p_{ij} es una distribución de probabilidad, ya que en cualquier paso puede ocurrir alguno de los sucesos E_1, E_2, \dots, E_m que son mutuamente excluyentes. Los valores p_{ij} se denominan *probabilidades de transición* que satisfacen la condición

$$\begin{aligned} p_{ij} &> 0 \\ \sum_{j=1}^m p_{ij} &= 1 \end{aligned}$$

para cada $i = 1, 2, \dots, m$.

Todos estos valores se combinan formando una *matriz de transición* T de tamaño $m \times m$, donde

$$T = [p_{ij}] = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix}$$

Se puede observar que cada fila o columna de la matriz es una distribución de probabilidad, es decir,

$$\sum_{j=1}^m p_{ij} = 1$$

Observación 5.10 Como vimos con anterioridad si las matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices estocásticas, entonces $C = A \cdot B$ es también estocástica.

Por la regla de multiplicación de matrices,

$$C = [c_{ij}] = [a_{ij}] \cdot [b_{ij}] = \left[\sum_{k=1}^m a_{ik} b_{kj} \right]$$

De este modo:

$$\sum_{j=1}^m c_{ij} = \sum_{j=1}^m \sum_{k=1}^m a_{ik} b_{kj} = \sum_{k=1}^m a_{ik} \sum_{j=1}^m b_{kj} = \sum_{k=1}^m a_{ik} \cdot 1 = 1$$

Una consecuencia es que cualquier potencia de la matriz T es también una matriz estocástica: T^n .

5.5.4.2. Probabilidad $p_j^{(n)}$

Teorema 5.12 Si los sucesos $A_1, A_2, A_3, \dots, A_m$ son una partición ($\bigcup_{i=1}^m A_i = E$, $A_i \cap A_j = \emptyset$ $i \neq j$) del espacio muestral E y T un suceso de S , entonces

$$P(T) = \sum_{i=1}^m p(A_i) \cdot p(T/A_i)$$

Como se muestra en la siguiente figura (1.1):

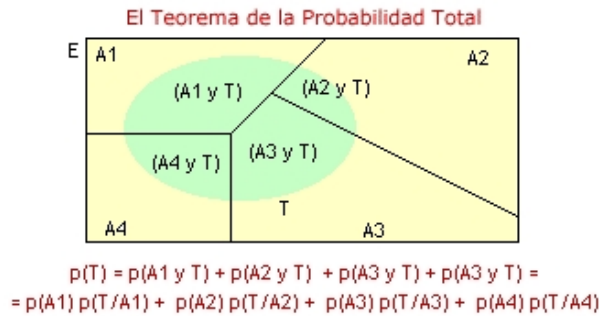


Figura 5.7: Probabilidad Total

Una probabilidad que resulta interesante conocer es la probabilidad de llegar a E_j después de n pasos, dada una distribución de probabilidad $p_i^{(0)}$.

Se observa que $p_i^{(0)}$ es la probabilidad de que el sistema ocupe inicialmente el estado E_i , de modo que

$$\sum_{i=1}^m p_i^{(0)} = 1$$

Si se denomina $p_j^{(1)}$ a la probabilidad de alcanzar E_j en un solo paso, entonces, por el teorema de probabilidad total

$$p_j^{(1)} = \sum_{i=1}^m p_i^{(0)} p_{ij}$$

Esto se puede expresar de forma vectorial: sean $p^{(0)}$ y $p^{(1)}$ los vectores fila de probabilidad dados por

$$\mathbf{p}^{(0)} = (p_1^{(0)}, \dots, p_m^{(0)})$$

y

$$\mathbf{p}^{(1)} = (p_1^{(1)}, \dots, p_m^{(1)})$$

donde $p^{(0)}$ es la distribución de probabilidad inicial y $p^{(1)}$ es la probabilidad de que se alcance cada uno de los estados E_1, \dots, E_m después de un paso.

Con esta notación, se puede expresar

$$\mathbf{p}^{(1)} = [p_j^{(1)}] = \left[\sum_{i=1}^m p_i^{(0)} p_{ij} \right] = \mathbf{p}^{(0)} T$$

donde T es la matriz de transición.

Del mismo modo,

$$\mathbf{p}^{(2)} = \mathbf{p}^{(1)} T = \mathbf{p}^{(0)} T^2$$

y en n pasos,

$$\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} T = \mathbf{p}^{(0)} T^n$$

donde

$$\mathbf{p}^{(n)} = (p_1^{(n)}, \dots, p_m^{(n)})$$

y de forma general,

$$\mathbf{p}^{(n+r)} = \mathbf{p}^{(r)} T^n$$

Observación 5.11 $p_j^{(n)}$ es la probabilidad incondicional de estar en el estado E_j en el n -ésimo paso, dado que la probabilidad inicial es $p^{(0)}$, esto es,

$$P(X_n = j) = p_j^{(n)}$$

que es lo mismo que:

$$\sum_{j=1}^m p_j^{(n)} = 1$$

5.5.4.3. Probabilidad de transición en n pasos $p_{ij}^{(n)}$

Se define $p_{ij}^{(n)}$ como la probabilidad de que la cadena esté en el estado E_j después de n pasos, dado que la cadena empezó en el estado E_i . Se tiene

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i)$$

y además por la propiedad markoviana:

$$p_{ij}^{(n)} = \sum_{k=1}^m P(X_n = j, X_{n-1} = k | X_0 = i)$$

para $n \geq 2$, ya que la cadena debe haber pasado por uno de los m posibles estados en la etapa $n - 1$.

Observación 5.12 Se tiene la siguiente igualdad para tres posibles sucesos A , B y C :

$$P(A \cap B | C) = P(A | B \cap C) \cdot P(B | C)$$

Si se sustituye

$$\begin{aligned} A &\rightarrow (X_n = j) \\ B &\rightarrow (X_{n-1} = k) \\ C &\rightarrow (X_0 = i) \end{aligned}$$

entonces

$$\begin{aligned} p_{ij}^{(n)} &= \sum_{k=1}^m P(X_n = j, X_{n-1} = k | X_0 = i) \\ &= \sum_{k=1}^m P(X_n = j | X_{n-1} = k, X_0 = i) P(X_{n-1} = k | X_0 = i) \\ &= \sum_{k=1}^m P(X_n = j | X_{n-1} = k) P(X_{n-1} = k | X_0 = i) \\ &= \sum_{k=1}^m p_{kj}^{(1)} p_{ik}^{(n-1)} = \sum_{k=1}^m p_{ik}^{(n-1)} p_{kj}^{(1)} \end{aligned}$$

La ecuación anterior se denomina de *Chapman- Kolmogorov*. Haciendo n igual a 2, 3, ... se obtiene que las matrices con esos elementos son

$$\begin{aligned} [p_{ij}^{(2)}] &= [p_{ik}^{(1)} p_{kj}^{(1)}] = T^2 \\ [p_{ij}^{(3)}] &= [p_{ik}^{(2)} p_{kj}^{(1)}] = T^3 \end{aligned}$$

ya que $p_{ik}^{(2)}$ son los elementos de T^2 y así sucesivamente. Así,

$$[p_{ij}^{(n)}] = T^n$$

Ejemplo 5.8 *En una cierta región el tiempo atmosférico sigue la siguiente secuencia: Un día se denomina soleado (S) si el sol luce más de la mitad del día, y se denomina nublado (N), si lo hace menos. Por experiencia, se sabe que si hay un día nublado, es igual de probable que el día siguiente sea también nublado. Si el día es soleado hay una probabilidad de 2/3 de que sea también soleado. Con estos datos se pide:*

1. *Construir la matriz de transición T de este proceso.*
2. *La probabilidad de que dentro de tres días esté también nublado y de que esté soleado; sabiendo en ambos casos que hoy esté nublado.*
3. *Calcular T^5 y T^{10} . ¿Cuál es el comportamiento de T^n cuando $n \rightarrow \infty$?
¿Como se comporta $p^{(n)}$, cuando $n \rightarrow \infty$? ¿Depende el límite de $p^{(0)}$?*

Primero asumimos que el proceso es markoviano y que el tiempo solo depende del día anterior. Se tiene una cadena de Markov con dos estados: E_1 Día nublado N , E_2 Día soleado S . Entonces la matriz de transición es:

	N	S
N	1/2	1/2
S	1/3	2/3

es decir,

$$T = \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}$$

con esto respondemos 1.

Para responder 2, analizamos los datos del día actual. Se tiene que

$$\mathbf{p}^{(0)} = \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} = (1 \quad 0)$$

De este modo, si hoy está nublado dentro de tres días se tiene que

$$\begin{aligned} p^{(3)} &= p^{(0)}T^3 \\ &= (1 \quad 0) \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}^3 \\ &= (1 \quad 0) \begin{pmatrix} 129/72 & 43/72 \\ 43/108 & 65/108 \end{pmatrix} \\ &= (29/72 \quad 43/72) \\ &= (0,403 \quad 0,597) \end{aligned}$$

Así, la probabilidad de que haya un día nublado en el tercer día es $p_1^{(3)} = \frac{29}{72}$ y que sea soleado es $p_2^{(3)} = \frac{43}{72}$. Por lo que se ha dado respuesta a 2.

Ahora, tenemos que calcular T^5 y T^{10} , queda:

$$\begin{aligned} T^5 &= \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}^5 = \begin{pmatrix} 0,40008 & 0,59992 \\ 0,39995 & 0,60005 \end{pmatrix} \\ T^{10} &= \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}^{10} = \begin{pmatrix} 0,4 & 0,60000 \\ 0,40000 & 0,6 \end{pmatrix} \end{aligned}$$

Para responder 3 el lector puede usar un programa computacional y así facilitarse los cálculos o hacerlo manualmente. Es importante mencionar que se debe tener sumo cuidado con las cifras significativas que se utilizan en el desarrollo de la aplicación, debido a que el error aumenta, no obteniendo una solución óptima.

Luego, tenemos

$$T^n \rightarrow \begin{pmatrix} 0,4 & 0,60000 \\ 0,40000 & 0,6 \end{pmatrix} = Q$$

de este modo,

$$\begin{aligned}
 \mathbf{p}^{(n)} = \mathbf{p}^{(0)}T^n &= \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} \begin{pmatrix} 0,4 & 0,60000 \\ 0,40000 & 0,6 \end{pmatrix} \\
 &= \left(\begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} 0,4 \quad \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} 0,6 \right) \\
 &= (0,4 \quad 0,6)
 \end{aligned}$$

ya que $p_1^{(0)} + p_2^{(0)} = 1$.

Se observa que cuando $n \rightarrow \infty$; $\mathbf{p}^{(n)}$ no depende de donde se parte $\mathbf{p}^{(0)}$.

5.5.4.4. Potencias de una matriz de transición

En los ejemplos anteriores se infiere que resulta necesario definir un método para calcular potencias de una matriz; y para esto se necesitarán como conocimientos previos saber diagonalizar una matriz.

Sea T una matriz de transición de orden $m \times m$, los valores propios de T se encuentran a partir de la ecuación característica.

$$|T - \lambda I_m| = 0$$

donde I_m es la matriz identidad de orden $m \times m$. Supongamos que los valores propios son distintos y los denotamos por $\lambda_1, \dots, \lambda_m$. Los correspondientes vectores propios cumplen las siguientes ecuaciones

$$|T - \lambda_i I_m| r_i = 0$$

para $i = 1, \dots, m$.

Ahora si construimos la matriz Q

$$Q = \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix}$$

donde v_1, v_2, \dots, v_m son los respectivos vectores propios y que están escritos en forma de columnas. Multiplicando las matrices:

$$\begin{aligned}
 TQ &= T \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix} \\
 &= \begin{pmatrix} Tv_1 & Tv_2 & \dots & Tv_m \end{pmatrix} \\
 &= \begin{pmatrix} \lambda_1 v_1 & \lambda_1 v_2 & \dots & \lambda_1 v_m \end{pmatrix} \\
 &= \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{pmatrix} \\
 &= QD \\
 \Rightarrow T &= QDQ^{-1}
 \end{aligned}$$

donde D es la matriz diagonal con los valores propios situados en la diagonal principal.

Se puede calcular

$$T^2 = (QDQ^{-1})(QDQ^{-1}) = QD^2Q^{-1}$$

en general

$$T^n = QD^nQ^{-1}$$

donde

$$D^n = \begin{pmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m^n \end{pmatrix}$$

Aunque este método es bastante útil, resulta tedioso a la hora de hacer los cálculos, incluso para cadenas a partir de $m = 4$ ó 5 .

Ejemplo 5.9 *Calcular los valores y vectores propios de la matriz estocástica*

$$T = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

Determinar T^n y el $\lim_{n \rightarrow \infty} T^n$.

Si se calcula

$$|T - \lambda I_m| = 0$$

se obtiene

$$\begin{vmatrix} 1/4 - \lambda & 1/2 & 1/4 \\ 1/2 & 1/4 - \lambda & 1/4 \\ 1/4 & 1/4 & 1/2 - \lambda \end{vmatrix} = 0$$

El polinomio característico es

$$\lambda^3 - \lambda^2 + \frac{1}{16}\lambda - \frac{1}{16} = 0$$

Es fácil corroborar que el polinomio característico escrito como producto de factores es:

$$(1 - \lambda) \left(\frac{1}{4} - \lambda\right) \left(\frac{-1}{4} - \lambda\right) = 0$$

por lo que se obtienen los valores propios

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= \frac{1}{4} \\ \lambda_3 &= \frac{-1}{4} \end{aligned}$$

Del mismo modo, los vectores propios se calculan a partir de

$$[T - \lambda I_m]r_i = 0$$

Para $\lambda_1 = 1$, tenemos

$$\begin{aligned}
 [T - \lambda_1 I|0] &= \left(\begin{array}{ccc|c} -3/4 & 1/2 & 1/4 & 0 \\ 1/2 & -3/4 & 1/4 & 0 \\ 1/4 & 1/4 & -1/2 & 0 \end{array} \right) \begin{array}{l} E_1(4) \\ E_2(4) \\ E_3(4) \end{array} \sim \left(\begin{array}{ccc|c} -3 & 2 & 1 & 0 \\ 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array} \right) \begin{array}{l} E_{13}(3) \\ E_{23}(-2) \\ \sim \end{array} \\
 &\quad \left(\begin{array}{ccc|c} 0 & 5 & -5 & 0 \\ 0 & -5 & 5 & 0 \\ 1 & 1 & -2 & 0 \end{array} \right) \begin{array}{l} E_1(\frac{1}{5}) \\ E_2(\frac{1}{5}) \\ \sim \end{array} \quad \left(\begin{array}{ccc|c} 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array} \right)
 \end{aligned}$$

$$\begin{aligned}
 y - z &= 0 \Rightarrow y = z \\
 x + y - 2z &= 0 \Rightarrow x = z \\
 S_{\lambda_1=1} &= \{(x, x, x)/x \in \mathbb{R}, x \neq 0\}
 \end{aligned}$$

Luego si $x = 1 \Rightarrow v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ es un vector propio.

Para $\lambda_2 = \frac{1}{4}$, tenemos

$$\begin{aligned}
 [T - \lambda_2 I|0] &= \left(\begin{array}{ccc|c} 0 & 1/2 & 1/4 & 0 \\ 1/2 & 0 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 0 \end{array} \right) \begin{array}{l} E_1(4) \\ E_2(4) \\ E_3(4) \end{array} \sim \left(\begin{array}{ccc|c} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right) \begin{array}{l} E_{23}(-2) \\ \sim \end{array} \\
 &\quad \left(\begin{array}{ccc|c} 0 & 2 & 1 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right) \begin{array}{l} E_{12}(1) \\ \sim \end{array} \quad \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right)
 \end{aligned}$$

$$\begin{aligned}
 -2y - z &= 0 \Rightarrow z = -2y \\
 x + y + z &= 0 \Rightarrow x = y \\
 S_{\lambda_2=\frac{1}{4}} &= \{(x, x, -2x)/x \in \mathbb{R}, x \neq 0\}
 \end{aligned}$$

Luego si $x = -1 \Rightarrow v_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$ es un vector propio.

Para $\lambda_3 = \frac{-1}{4}$, tenemos

$$\begin{aligned}
 [T - \lambda_3 I|0] &= \left(\begin{array}{ccc|c} 1/2 & 1/2 & 1/4 & 0 \\ 1/2 & 1/2 & 1/4 & 0 \\ 1/4 & 1/4 & 3/4 & 0 \end{array} \right) \begin{array}{l} E_1(4) \\ E_2(4) \\ E_3(4) \end{array} \sim \left(\begin{array}{ccc|c} 2 & 2 & 1 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 1 & 3 & 0 \end{array} \right) \begin{array}{l} E_{12}(-1) \\ \sim \end{array} \\
 &\quad \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 1 & 3 & 0 \end{array} \right) \begin{array}{l} E_{23}(-2) \\ \sim \end{array} \quad \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 0 \\ 1 & 1 & 3 & 0 \end{array} \right) \begin{array}{l} E_2(\frac{1}{5}) \\ \sim \end{array} \quad \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 3 & 0 \end{array} \right)
 \end{aligned}$$

$$\begin{aligned}
 -z &= 0 \Rightarrow z = 0 \\
 x + y + 3z &= 0 \Rightarrow x = -y \\
 S_{\lambda_3=\frac{-1}{4}} &= \{(x, -x, 0)/x \in \mathbb{R}, x \neq 0\}
 \end{aligned}$$

Luego si $x = 1 \Rightarrow v_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ es un vector propio.

De este modo

$$Q = (v_1 \ v_2 \ \dots \ v_m) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix}$$

$$\begin{aligned} T^n &= QD^nQ^{-1} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} (1)^n & 0 & 0 \\ 0 & (\frac{1}{4})^n & 0 \\ 0 & 0 & (\frac{-1}{4})^n \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/6 & -1/6 & 1/3 \\ 1/2 & -1/2 & 0 \end{pmatrix} \end{aligned}$$

luego, cuando $n \rightarrow \infty$, queda

$$\begin{aligned} \lim_{n \rightarrow \infty} T^n &= \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} (1)^n & 0 & 0 \\ 0 & (\frac{1}{4})^n & 0 \\ 0 & 0 & (\frac{-1}{4})^n \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/6 & -1/6 & 1/3 \\ 1/2 & -1/2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/6 & -1/6 & 1/3 \\ 1/2 & -1/2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} = Q \end{aligned}$$

Supongamos que T es la matriz de transición de una cadena de Markov con 3 estados y que la probabilidad inicial es $\mathbf{p}^{(0)}$, entonces después de n etapas

$$\mathbf{p}^{(n)} = \mathbf{p}^{(0)}T^n$$

y la distribución límite es

$$\mathbf{p} = \lim_{n \rightarrow \infty} \mathbf{p}^{(n)} = \lim_{n \rightarrow \infty} \mathbf{p}^{(0)}T^n = \mathbf{p}^{(0)}Q = \left(\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3} \right)$$

El vector \mathbf{p} da las probabilidades de los estados a largo plazo, y esta distribución es independiente del estado inicial $\mathbf{p}^{(0)}$. Finalmente \mathcal{P} también se ha resuelto.

5.5.5. Sistemas Oscilantes

5.5.5.1. Sistemas oscilantes de muelles y partículas

En esta aplicación podremos observar el análisis del movimiento oscilatorio de una cadena monoatómica lineal para algunos casos particulares que son de gran utilidad por su aplicación a múltiples fenómenos, tanto de la física experimental y tecnología industrial que se dan en gran parte del entorno en que convivimos, como en el estudio teórico de la física y en especial de las partículas.

Los sistemas oscilantes se encuentran en múltiples campos de las ciencias exactas (física, química y biología), no solo en el universo microscópico, sino también, en el ámbito de los movimientos de sistemas estelares, tratados por la Astrofísica. Por ello, describiremos los conocimientos previos necesarios para dar forma a las soluciones, las ecuaciones de movimiento que se utilizarán preliminarmente para resolver sistemas formados por muelles y partículas, luego se describirán los Modos normales de vibración. Posteriormente, se dará una versión generalizada para sistemas oscilantes formados por dos (N=2), tres (N=3) y

cuatro ($N=4$) partículas de una cadena monoatómica lineal. Finalmente se describirá la diagonalización de una matriz simétrica A de gran importancia para la resolución de los problemas, indicando las coordenadas y modos normales de vibración de las partículas unidas a muelles.

Estudiaremos los modos normales de vibración de un sistema formado por muelles y partículas como continuación y generalización del sistema formado por tres osciladores acoplados. Este ejemplo nos ayudará a comprender los modos normales de vibración de una cuerda fija por sus extremos, también denominados ondas estacionarias.

5.5.5.2. Conocimientos previos

En el caso del movimiento oscilatorio, un modelo muy apropiado para comenzar es un objeto de masa m conectado a un resorte o elástico horizontal, si este no está deformado el objeto se dice que está en reposo sobre una superficie sin rozamiento en su posición de equilibrio, en donde $x = 0$. Para ahondar en nuestro tema utilizaremos el modelo de una partícula, para que los efectos de la masa m no influyan en el movimiento oscilatorio de la partícula a través de los muelles. Para mayor comprensión representaremos matemáticamente el movimiento armónico simple, que más adelante nos ayudará a formar las soluciones.

Comenzaremos mencionando que cuando una partícula está unida a un muelle elástico ideal sin masa y es desplazada desde la posición x , el muelle ejerce una fuerza, cuya magnitud viene dada por **la ley de Hooke**

$$F_s = -km \quad (5.12)$$

donde k es la constante de fuerza del elástico o resorte y F_s es la componente de la fuerza ejercida por el resorte, debido a que (5.12) es una ecuación proporcional al desplazamiento con relación a la posición de equilibrio y va dirigida siempre hacia la posición de equilibrio, en sentido opuesto al desplazamiento. Es decir, cuando una partícula es desplazada hacia la derecha, x es positivo y la fuerza ejercida por el resorte tiene una componente negativa, hacia la izquierda. Cuando la partícula se desplaza hacia la izquierda de $x = 0$, x es negativo y la fuerza ejercida por el resorte tiene una componente positiva, hacia la derecha. Cuando una partícula está bajo el efecto de una fuerza, el movimiento de la partícula corresponde a un tipo especial de movimiento oscilatorio denominado **movimiento armónico simple (M.A.S)**. Al sistema que experimenta un movimiento armónico simple se le denomina oscilador armónico simple.

Suponemos que si a una partícula sometida a una fuerza similar indicada en la ecuación de Hooke, se le aplica la segunda ley de Newton en la dirección x se tiene:

$$\sum F = F_s \quad (5.13)$$

$$= ma \quad (5.14)$$

$$-km = ma \quad (5.15)$$

$$a = -\frac{k}{m}x \quad (5.16)$$

Es decir, la aceleración es proporcional al desplazamiento de la partícula con relación a la posición de equilibrio y va dirigido en posición opuesta. Cabe destacar que si desplazamos la partícula hasta una posición $x = A$ y luego, con la partícula en reposo, la soltamos, su aceleración inicial es $-kA/m$. Cuando la partícula pasa por la posición de equilibrio $x = 0$, su aceleración es cero. En ese instante, su rapidez alcanza un valor máximo, ya que su aceleración cambia de signo. Si la partícula sigue su desplazamiento de la posición de equilibrio, alcanza la posición $x = -A$. En ese instante, su aceleración $\pm kA/m$ y su

rapidez es de nuevo, cero. Si esta sigue desplazándose oscila entre los puntos límites $\pm kA$ en ausencia de rozamiento, dado que la fuerza que se ejerce es conservativa.

Para mayor comprensión del M.A.S y de las soluciones representaremos matemáticamente el movimiento que se realiza.

Recordemos que por definición tenemos que $a = dv/dt = d^2x/dt^2$, por lo que podemos expresar la ecuación (5.16) como:

$$\frac{d^2x}{dt^2} = \frac{-k}{m}x \quad (5.17)$$

Si sustituimos el coeficiente k/m por el símbolo ω^2 , se elige ω^2 en lugar de ω para hacer que la solución tenga una forma más simple, entonces

$$\omega^2 = \frac{k}{m} \quad (5.18)$$

y de tal modo podemos escribir la ecuación (5.17) de la forma

$$\frac{d^2x}{dt^2} = -\omega^2x \quad (5.19)$$

Lo que necesitamos ahora encontrar es una solución matemática para la ecuación (5.19), es decir, una función $x(t)$ que satisfaga esta ecuación diferencial de segundo orden. Esto nos dará la representación matemática de la posición de la partícula en función del tiempo. Buscamos, una función $x(t)$, tal que la segunda derivada de la función sea igual a la función original, pero con signo negativo y multiplicada por ω^2 . Las funciones trigonométricas seno y coseno muestran este comportamiento, de modo que podemos construir una solución basándonos en una de ellas (o en ambas). La función coseno siguiente es una solución a la ecuación diferencial (5.19).

$$x(t) = A \cos(\omega t + \phi) \quad (5.20)$$

donde A , ω y ϕ son constantes. Para ver esto de forma más explícita, observe que

$$\begin{aligned} \frac{dx}{dt} &= A \frac{d}{dt} \cos(\omega t + \phi) \\ &= -\omega A \sin(\omega t + \phi) \end{aligned}$$

$$\frac{d^2x}{dt^2} = -\omega A \frac{d}{dt} \sin(\omega t + \phi) \quad (5.21)$$

$$= -\omega^2 A \cos(\omega t + \phi) \quad (5.22)$$

Ahora, si comparamos las ecuaciones (5.20) y (5.22), vemos que

$$\frac{d^2x}{dt^2} = -\omega^2x \quad (5.23)$$

lo que satisface lo que estábamos buscando.

Como observación general cabe destacar que los parámetros A, ω y ϕ son constantes del movimiento. Al hacer una análisis gráfico podemos ver que el parámetro A denominado la **amplitud** del movimiento, es sencillamente **el valor máximo de la posición de la partícula, tanto en la dirección positiva como negativa de x** . A la constante ω se le denomina **frecuencia angular** denotada como:

$$\omega = \sqrt{\frac{k}{m}}$$

Al ángulo ϕ se le denomina **constante de fase**(o ángulo de fase) y al igual que la amplitud A , queda determinado solamente por la posición y rapidez de la partícula en el instante $t = 0$. Si la partícula está en su posición máxima $x = A$ en el instante $t = 0$, la constante de fase es $\phi = 0$ y por último a la cantidad $(\omega t + \phi)$ se le denomina **fase de movimiento**. Observe que la función $x(t)$ es periódica y que su valor es el mismo cada vez que ωt se incrementa en 2π radianes.

Con toda esta información podemos formular nuestros sistemas oscilantes de muelles y partículas.

A partir de lo anterior supongamos el siguiente modelo de la figura (5.8) que indica que tenemos un conjunto de partículas de masas $m_1, m_2, m_3, \dots, m_{N-1}, m_N$ unidas por muelles elásticos de constantes $k_0, k_1, k_2, k_3, \dots, k_{N-1}, k_N$; las cuales se representan a continuación:

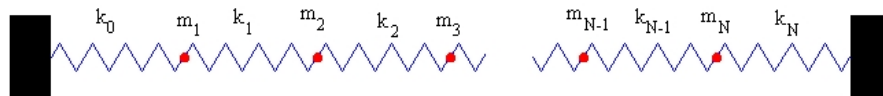


Figura 5.8: Muelle

Para formular un modelo matemático más preciso, se procederá a construir las Ecuaciones del Movimiento, necesarias para formalizar este modelo.

5.5.5.3. Ecuaciones del Movimiento

Para estudiar un sistema formado por dos osciladores acoplados vamos a tomar como modelo el sistema formado por dos partículas iguales de masa m situadas en los extremos de dos muelles de idéntica constante elástica k . El acoplamiento se efectúa uniendo las dos partículas mediante un muelle de constante k_c , tal como se puede ver en la figura (5.9).

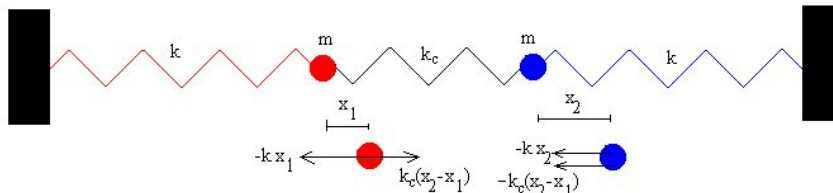


Figura 5.9: Muelles acoplados

Llamemos x_1 y x_2 a los desplazamientos de cada una de las partículas a partir de su posición de equilibrio, medidos como positivos cuando están a la derecha. El muelle de la izquierda se ha estirado x_1 ,

el de la derecha se ha comprimido x_2 y el central se ha deformado $x_2 - x_1$. Las fuerzas sobre cada una de las partículas se indican en la figura.

- Sobre la partícula de la izquierda se ejerce una fuerza hacia la izquierda $-kx_1$ y una fuerza hacia la derecha debido a la deformación del muelle central $k_c(x_2 - x_1)$, suponemos que x_2 es mayor que x_1 .
- Sobre la partícula derecha se ejerce una fuerza hacia la izquierda $-kx_2$ y otra fuerza hacia la izquierda debido a la deformación del muelle central $-k_c(x_2 - x_1)$.

Observación 5.13 *El muelle central ejerce fuerzas iguales y de sentido contrario sobre cada una de las partículas.*

En las siguientes figuras se muestran las fuerzas sobre la primera partícula de masa m_1 , la partícula i y la última partícula N de masa m_N

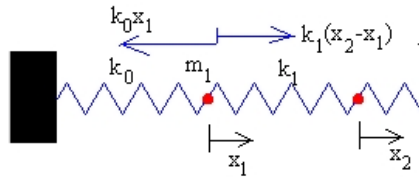


Figura 5.10: Muelle desde x_1 a x_2

$$m_1 \ddot{x}_1 = -k_0 x_1 + k_1(x_2 - x_1)$$

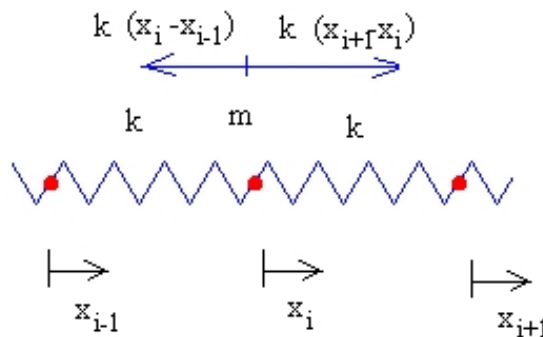
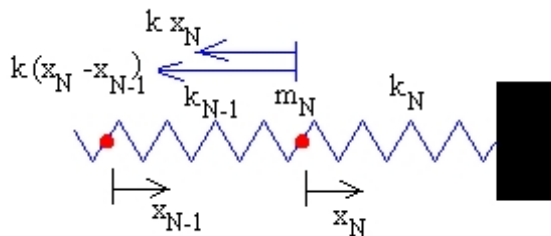


Figura 5.11: Muelle desde x_{i-1} a x_{i+1}

$$m_i \ddot{x}_i = -k_{i-1}(x_i - x_{i-1}) + k_i(x_{i+1} - x_i)$$


 Figura 5.12: Muelle desde x_{N-1} a x_N

$$m_N \ddot{x}_N = -k_{N-1}(x_N - x_{N-1}) + k_N x_N$$

Resumiendo, las ecuaciones del movimiento de las partículas son:

$$\begin{aligned} m_1 \ddot{x}_1 &= -k_0 x_1 + k_1(x_2 - x_1) \\ m_2 \ddot{x}_2 &= -k_1(x_2 - x_1) + k_2(x_3 - x_2) \\ m_3 \ddot{x}_3 &= -k_2(x_3 - x_2) + k_3(x_4 - x_3) \\ &\vdots \\ m_i \ddot{x}_i &= -k_{i-1}(x_i - x_{i-1}) + k_i(x_{i+1} - x_i) \\ &\vdots \\ m_{N-1} \ddot{x}_{N-1} &= -k_{N-2}(x_{N-1} - x_{N-2}) + k_{N-1}(x_N - x_{N-1}) \\ m_N \ddot{x}_N &= -k_{N-1}(x_N - x_{N-1}) + k_N x_N \end{aligned}$$

Buscamos una solución de la forma: $\ddot{x}_i = x_i \cos(\omega t)$

Al ser la fase inicial $\frac{\pi}{2}$, las partículas parten del reposo en el instante $t = 0$, de la posición x_{0i}

De las ecuaciones anteriormente descritas, obtenemos la siguiente matriz

$$A = \begin{pmatrix} \frac{-(k_0+k_1)}{m_1} & \frac{k_1}{m_1} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \frac{k_1}{m_2} & \frac{-(k_1+k_2)}{m_2} & \frac{k_2}{m_2} & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \frac{k_{i-1}}{m_i} & \frac{-(k_{i-1}+k_i)}{m_i} & \frac{k_i}{m_i} & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \frac{k_{N-1}}{m_N} & \frac{-(k_{N-1}+k_N)}{m_N} \end{pmatrix}$$

5.5.5.4. Casos Particulares

Un caso particular importante es una cadena monoatómica lineal cuando las constantes de los muelles elásticos tienen el mismo valor $k_0 = k_1 = k_2 = k_3 \dots = k_{N-1} = k_N = k$, y las partículas tienen la misma masa $m_0 = m_1 = m_2 = m_3 = \dots = m_{N-1} = m_N = m$, estos casos que veremos a continuación nos representarán sistemas formados por partículas $N = 2$, $N = 3$ y $N = 4$.

Calculamos los valores propios ω^2 y los vectores propios de la matriz simétrica A , que en forma general queda:

$$A = \begin{pmatrix} \frac{-2k}{m} & \frac{k}{m} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \frac{k}{m} & \frac{-2k}{m} & \frac{k}{m} & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \frac{k}{m} & \frac{-2k}{m} & \frac{k}{m} & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \frac{k}{m} & \frac{-2k}{m} \end{pmatrix}$$

A continuación veremos la forma generalizada para partículas de $N = 2, 3$ y 4, luego se procederá a resolver el caso para $N = 3$.

Observación 5.14 *El caso para $N = 2$ y 4 se dejará como una actividad para el lector, para ello deberá utilizar los procedimientos descritos en el texto.*

1. Sistema formado por dos partículas $N = 2$

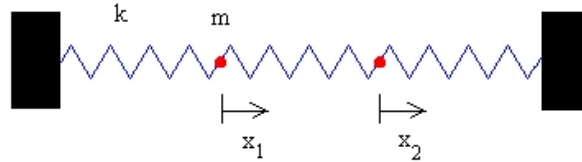


Figura 5.13: Muelle con 2 partículas

Ecuaciones del Movimiento

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 + k(x_2 - x_1) \\ m\ddot{x}_2 &= -k(x_2 - x_1) - kx_2 \end{aligned}$$

Buscamos una solución de la forma:

$$\begin{aligned} \ddot{x}_1 &= x_1 \cos(\omega t) \\ \ddot{x}_2 &= x_2 \cos(\omega t) \end{aligned}$$

En forma matricial el sistema de ecuaciones se expresa:

$$\begin{pmatrix} \frac{-2k}{m} & \frac{k}{m} \\ \frac{k}{m} & \frac{-2k}{m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \omega^2 \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix}$$

Las frecuencias ω^2 son los valores propios de la matriz cuadrada descrita en las ecuaciones del movimiento para $N = 2$ y los vectores propios son las amplitudes de los modos normales de vibración.

Valores propios ω^2	Vectores propios
$\frac{k}{m}$	(x_{11}, x_{11})
$\frac{3k}{m}$	$(x_{12}, -x_{11})$

Los coeficientes x_{11} y x_{12} se determinan a partir de las condiciones iniciales.

2. Sistema formado por tres partículas $N = 3$

Las ecuaciones del movimiento son:

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 + k(x_2 - x_1) \\ m\ddot{x}_2 &= -k(x_2 - x_1) + -k(x_3 - x_2) \\ m\ddot{x}_3 &= -k(x_3 - x_2) - kx_3 \end{aligned}$$

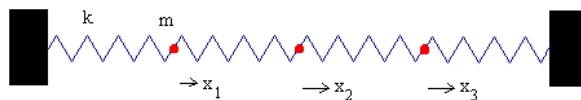


Figura 5.14: Muelle con 3 partículas

Buscamos una solución de la forma:

$$\begin{aligned}\ddot{x}_1 &= x_1 \cos(\omega t) \\ \ddot{x}_2 &= x_2 \cos(\omega t) \\ \ddot{x}_3 &= x_3 \cos(\omega t)\end{aligned}$$

En forma matricial el sistema de ecuaciones se expresa:

$$\begin{pmatrix} \frac{-2k}{m} & \frac{k}{m} & 0 \\ \frac{k}{m} & \frac{-2k}{m} & \frac{k}{m} \\ 0 & \frac{k}{m} & \frac{-2k}{m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \omega^2 \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix}$$

Las frecuencias ω^2 son los valores propios de la matriz cuadrada descrita en las ecuaciones del movimiento para $N = 3$ y los vectores propios son las amplitudes de los modos normales de vibración.

Valores propios ω^2	Vectores propios
$\frac{2k}{m}(1 - \frac{\sqrt{2}}{2})$	$(x_{11}, x_{11}\sqrt{2}, x_{11})$
$\frac{2k}{m}$	$(x_{12}, 0, x_{12})$
$\frac{2k}{m}(1 + \frac{\sqrt{2}}{2})$	$(x_{13}, -x_{13}\sqrt{2}, x_{13})$

Los coeficientes x_{11} , x_{12} y x_{13} se determinan a partir de las condiciones iniciales.

3. Sistema formado por cuatro partículas $N = 4$

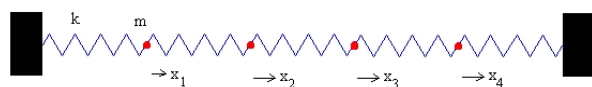


Figura 5.15: Muelle con 4 partículas

$$\begin{aligned}m\ddot{x}_1 &= -kx_1 + k(x_2 - x_1) \\ m\ddot{x}_2 &= -k(x_2 - x_1) + -k(x_3 - x_2) \\ m\ddot{x}_3 &= -k(x_3 - x_2) - k(x_4 - x_3) \\ m\ddot{x}_4 &= -k(x_4 - x_3) - kx_4\end{aligned}$$

Buscamos una solución de la forma:

$$\begin{aligned}\ddot{x}_1 &= x_1 \cos(\omega t) \\ \ddot{x}_2 &= x_2 \cos(\omega t) \\ \ddot{x}_3 &= x_3 \cos(\omega t) \\ \ddot{x}_4 &= x_4 \cos(\omega t)\end{aligned}$$

En forma matricial el sistema de ecuaciones se expresa:

$$\begin{pmatrix} \frac{-2k}{m} & \frac{k}{m} & 0 & 0 \\ \frac{k}{m} & \frac{-2k}{m} & \frac{k}{m} & 0 \\ 0 & \frac{k}{m} & \frac{-2k}{m} & \frac{k}{m} \\ 0 & 0 & \frac{k}{m} & \frac{-2k}{m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \omega^2 \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \ddot{x}_4 \end{pmatrix}$$

Las frecuencias ω^2 son los valores propios de la matriz cuadrada y los vectores propios, las amplitudes de los modos normales de vibración.

Valores propios ω^2	Vectores propios
$\frac{k}{m}(2 - \sqrt{\frac{3+\sqrt{5}}{2}}) = \frac{k}{m}(2 - \frac{\sqrt{5}+1}{2})$	$(x_{11}, \frac{\sqrt{5}+1}{2}x_{11}, \frac{\sqrt{5}+1}{2}x_{11}, x_{11})$
$\frac{k}{m}(2 - \frac{\sqrt{5}-1}{2})$	$(x_{12}, \frac{\sqrt{5}-1}{2}x_{12}, -\frac{\sqrt{5}-1}{2}x_{12}, -x_{12})$
$\frac{k}{m}(2 + \frac{\sqrt{5}-1}{2})$	$(x_{13}, -\frac{\sqrt{5}-1}{2}x_{13}, -\frac{\sqrt{5}-1}{2}x_{13}, x_{13})$
$\frac{k}{m}(2 + \frac{\sqrt{5}+1}{2})$	$(x_{14}, -\frac{\sqrt{5}+1}{2}x_{14}, \frac{\sqrt{5}+1}{2}x_{14}, -x_{14})$

Los coeficientes x_{11} , x_{12} , x_{13} y x_{14} se determinan a partir de las condiciones iniciales.

5.5.5.5. Ejemplo de un Sistema formado por tres partículas

Consideremos unos muelles acoplados de longitud $4k$ y de masa despreciable, cargada con tres partículas, cada una de masa m . Supongamos que las masas se encuentran a distancias fijas k , $2k$ y $3k$ de uno de los muelles. Las partículas se desplazan ligeramente de su posición de equilibrio y se sueltan. Analice el movimiento de estas, al suponer que se lleva a cabo en un plano horizontal a través de los muelles.

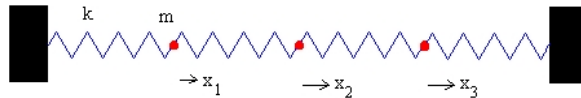


Figura 5.16: Sistema formado por tres partículas

Sean x_1, x_2 y x_3 los desplazamientos horizontales de las partículas en cualquier instante durante el movimiento. Sea T la tensión de la cuerda, además supongamos que los desplazamientos son pequeños y que la tensión permanece inalterada durante el movimiento. A partir de las ecuaciones del movimiento indicadas con anterioridad tenemos que:

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 + k(x_2 - x_1) \\ m\ddot{x}_2 &= -k(x_2 - x_1) + -k(x_3 - x_2) \\ m\ddot{x}_3 &= -k(x_3 - x_2) - kx_3 \end{aligned}$$

Al aplicar **la segunda ley del movimiento de Newton** (fuerza = masa \times aceleración) que describe el movimiento de cada una de las tres partículas a estas tres ecuaciones, se logran combinar en una sola ecuación matricial.

$$\omega^2 \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix} = \frac{T}{ma} \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

es decir,

$$\omega^2 \ddot{x} = \frac{T}{ma} Ax \quad (5.24)$$

Se observa que la matriz A es simétrica y se ajusta a los casos particulares descritos con anterioridad. Por ende, puede ser diagonalizada. Sea C la matriz que tiene como vectores columna a los vectores propios de la matriz A , o sea,

$$D = C^{-1}AC \quad (5.25)$$

Si realizamos algunos arreglos en la ecuación (5.24), esta se puede reescribir de la siguiente forma:

$$\begin{aligned} \omega^2 C^{-1} \ddot{x} &= C^{-1} \frac{T}{ma} ACC^{-1}x \\ \omega^2 C^{-1} \ddot{x} &= \frac{T}{ma} C^{-1} ACC^{-1}x \end{aligned}$$

Sea D la matriz en (5.25). Entonces

$$\omega^2 C^{-1} \ddot{x} = \frac{T}{ma} DC^{-1}x$$

Si definimos un nuevo sistemas de coordenadas tenemos:

$$x' = C^{-1}x \quad (5.26)$$

Luego, nuestro nuevo sistema se transforma en

$$\omega^2 \ddot{x}' = \frac{T}{ma} Dx' \quad (5.27)$$

Sin embargo, x' en la ecuación matricial tiene la siguiente relación

$$x = Cx' \quad (5.28)$$

que nos lleva a la solución en términos del sistema original, posteriormente debemos encontrar los valores y los vectores propios de la matriz simétrica A . Podemos utilizar directamente el método descrito en los casos particulares, para no abusar de la generalidad de este ejemplo se procederá a resolver por los métodos usuales de diagonalización, es decir:

▪ **Valores propios de A**

Sea la matriz simétrica

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

y sea C la matriz columna que contenga los vectores propios de A .

La ecuación característica de la matriz A es: $\det(A - \lambda I) = 0$

$$\begin{aligned}
 \det(A - \lambda I) &= \begin{vmatrix} -2 - \lambda & 1 & 0 \\ 1 & -2 - \lambda & 1 \\ 0 & 1 & -2 - \lambda \end{vmatrix} \\
 &= (-2 - \lambda) \begin{vmatrix} -2 - \lambda & 1 \\ 1 & -2 - \lambda \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 0 & -2 - \lambda \end{vmatrix} \\
 &= (-2 - \lambda)[(-2 - \lambda)^2 - 1] - (-2 - \lambda) \\
 &= (-2 - \lambda)[4 + 4\lambda + \lambda^2 - 1] \\
 &= (-2 - \lambda)[\lambda^2 + 4\lambda + 3] + 2 + \lambda
 \end{aligned}$$

Luego el polinomio característico es:

$$\lambda^3 + 6\lambda^2 + 10\lambda + 4 = 0$$

Ahora, tenemos que resolver el polinomio de grado 3, encontrando sus raíces que son:

$$\begin{aligned}
 \lambda_1 &= -2 \\
 \lambda_2 &= -2 - \sqrt{2} \\
 \lambda_3 &= -2 + \sqrt{2}
 \end{aligned}$$

Observación 5.15 Para resolver el polinomio de grado 3, se encontró por tanteo que $\lambda_1 = -2$ era una raíz, con esto solo falta resolver la ecuación cuadrática que se obtiene al utilizar división sintética.

■ Vectores Propios

1. Para $\lambda = -2$ resolveremos el sistema con la matriz aumentada $[A - \lambda I|0]$

$$[A - (-2)I|0] = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right)$$

$$\begin{aligned}
 y &= 0 \\
 x + z &= 0 \Rightarrow x = -z \\
 S_{\lambda=-2} &= \{(-z, 0, z) / z \in \mathbb{R}, z \neq 0\}
 \end{aligned}$$

Luego si $z = -1 \Rightarrow v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ es un vector propio.

2. Para $\lambda = -2 + \sqrt{2}$ resolveremos el sistema con la matriz aumentada $[A - \lambda I | 0]$

$$[A - (-2 + \sqrt{2})I | 0] = \left(\begin{array}{ccc|c} -\sqrt{2} & 1 & 0 & 0 \\ 1 & -\sqrt{2} & 1 & 0 \\ 0 & 1 & -\sqrt{2} & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$x - z = 0 \Rightarrow x = z$$

$$y - \sqrt{2}z = 0 \Rightarrow y = \sqrt{2}z$$

$$S_{\lambda=-2+\sqrt{2}} = \{(z, \sqrt{2}z, z) / z \in \mathbb{R}, z \neq 0\}$$

Luego si $z = 1 \Rightarrow v_2 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$ es un vector propio.

3. Para $\lambda = -2 - \sqrt{2}$ resolveremos el sistema con la matriz aumentada $[A - \lambda I | 0]$

$$[A - (-2 - \sqrt{2})I | 0] = \left(\begin{array}{ccc|c} \sqrt{2} & 1 & 0 & 0 \\ 1 & \sqrt{2} & 1 & 0 \\ 0 & 1 & \sqrt{2} & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$x - z = 0 \Rightarrow x = z$$

$$y + \sqrt{2}z = 0 \Rightarrow y = -\sqrt{2}z$$

$$S_{\lambda=-2-\sqrt{2}} = \{(z, -\sqrt{2}z, z) / z \in \mathbb{R}, z \neq 0\}$$

Luego si $z = 1 \Rightarrow v_3 = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$ es un vector propio.

Entonces

$$C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\sqrt{2} & \sqrt{2} \\ -1 & 1 & 1 \end{pmatrix} \text{ y } D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 - \sqrt{2} & 0 \\ 0 & 0 & -2 + \sqrt{2} \end{pmatrix}$$

es decir, que $C^{-1}AC = D$

$$D = \frac{1}{4} \underbrace{\begin{pmatrix} 2 & 0 & -2 \\ 1 & -\sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\sqrt{2} & \sqrt{2} \\ -1 & 1 & 1 \end{pmatrix}}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ \frac{-1}{2} - \frac{\sqrt{2}}{4} & \frac{1}{2} + \frac{\sqrt{2}}{2} & \frac{-1}{2} - \frac{\sqrt{2}}{4} \\ \frac{\sqrt{2}}{4} - \frac{1}{2} & \frac{1}{2} - \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{4} - \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\sqrt{2} & \sqrt{2} \\ -1 & 1 & 1 \end{pmatrix}}$$

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 - \sqrt{2} & 0 \\ 0 & 0 & -2 + \sqrt{2} \end{pmatrix}$$

Con esto se comprueba que las soluciones son correctas y procedemos a formar la ecuación matricial del movimiento.

$$\omega^2 \begin{pmatrix} \ddot{x}'_1 \\ \ddot{x}'_2 \\ \ddot{x}'_3 \end{pmatrix} = \frac{T}{ma} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 - \sqrt{2} & 0 \\ 0 & 0 & -2 + \sqrt{2} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

Esta ecuación matricial representa el siguiente sistema de tres ecuaciones diferenciales, cada una de estas describe el movimiento armónico simple.

$$\begin{aligned} \ddot{x}'_1 &= \frac{-2T}{ma} x'_1 \\ \ddot{x}'_2 &= \frac{(-2 - \sqrt{2})T}{ma} x'_2 \\ \ddot{x}'_3 &= \frac{(-2 + \sqrt{2})T}{ma} x'_3 \end{aligned}$$

Estas ecuaciones tienen las siguientes soluciones:

$$\begin{aligned} \ddot{x}'_1 &= b_1 \cos\left(\sqrt{\frac{2T}{ma}}t + \gamma_1\right) \\ \ddot{x}'_2 &= b_2 \cos\left(\sqrt{\frac{(2 + \sqrt{2})T}{ma}}t + \gamma_2\right) \\ \ddot{x}'_3 &= b_3 \cos\left(\sqrt{\frac{(2 - \sqrt{2})T}{ma}}t + \gamma_3\right) \end{aligned}$$

Observación 5.16 Para encontrar con mayor facilidad la solución diferencial de estas ecuaciones véase las ecuaciones (5.19) y (5.20) de la sección de conocimientos previos de esta aplicación.

En las soluciones $b_1, b_2, b_3, \gamma_1, \gamma_2$ y γ_3 son constantes de integración en el tiempo ($t = 0$). Estas constantes en la física se interpretan como el estándar del movimiento armónico simple (M.A.S), donde b_1, b_2, b_3 y γ_3 son amplitudes y γ_1, γ_2 y γ_3 son fases. Estas no se explicitarán de modo físico, solo se nombrarán para dar un mayor sentido al tipo de soluciones a las cuales nos referimos.

A las coordenadas x'_1, x'_2 y x'_3 en las que el movimiento de las partículas se describen con más facilidad se les llaman coordenadas normales del movimiento. estas se pueden obtener de la siguiente forma:

Sea

$$x = Cx'$$

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\sqrt{2} & \sqrt{2} \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \ddot{x}'_1 \\ \ddot{x}'_2 \\ \ddot{x}'_3 \end{pmatrix} \\ &= x'_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + x'_2 \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} + x'_3 \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \end{aligned}$$

Reemplazando x'_1, x'_2 y x'_3 se obtiene:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = b_1 \cos\left(\sqrt{\frac{2T}{ma}}t + \gamma_1\right) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + b_2 \cos\left(\sqrt{\frac{(2 + \sqrt{2})T}{ma}}t + \gamma_2\right) \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} \\ + b_3 \cos\left(\sqrt{\frac{(2 - \sqrt{2})T}{ma}}t + \gamma_3\right) \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$$

Luego, el movimiento es la combinación de los siguientes tres modos:

▪ **Modo 1**

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = b_1 \cos\left(\sqrt{\frac{2T}{ma}}t + \gamma_1\right) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

donde $x_2 = 0$ y $x_3 = -x_1$

▪ **Modo 2**

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = b_2 \cos\left(\sqrt{\frac{(2 + \sqrt{2})T}{ma}}t + \gamma_2\right) \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$$

donde $x_2 = -\sqrt{x_1}$ y $x_3 = x_1$

▪ **Modo 3**

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = b_3 \cos\left(\sqrt{\frac{(2 - \sqrt{2})T}{ma}}t + \gamma_3\right) \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$$

donde $x_2 = \sqrt{x_1}$ y $x_3 = x_1$

La combinación lineal determinada por b_1, b_2 y b_3 nos da la combinación real. Cada una de estos es un M.A.S de una partícula. Los modos descritos se les llaman modos normales de oscilación.

5.6. Ejercicios Propuestos

1. Dada la ecuación de la cónica $x^2 + 2xy + y^2 - 8x + 8y = 0$, identificar la cónica, expresarla en su forma canónica y graficarla.
2. Dada la ecuación de la cónica $5x^2 + 6xy + 5y^2 - 4x + 4y = 4$, identificar la cónica, expresarla en su forma canónica y graficarla.

3. Demuestre que si $a \neq c$, el término xy en la ecuación cuadrática $ax^2 + bxy + cy^2 = d$ se elimina rotando en un ángulo θ , si θ está dado por

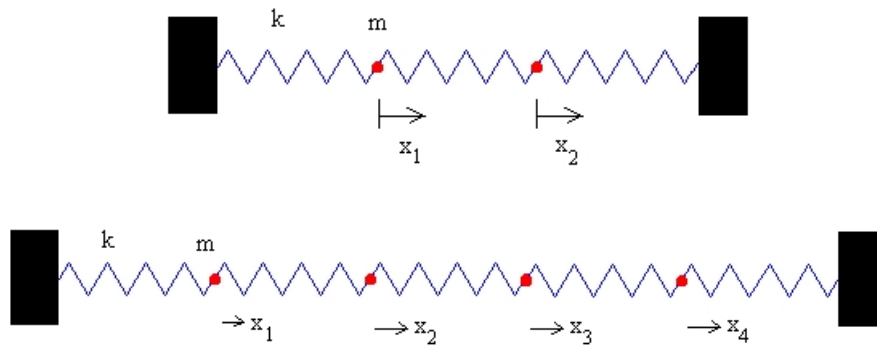
$$\cot 2\theta = \frac{a - c}{b}$$

4. Dada la ecuación de segundo grado en dos variables $x^2 + 10xy + y^2 - 1 = 0$, demuestre que esta representa una cónica degenerada.
5. Dada la ecuación $x^2 + 2yz + 4x + 1 = 0$, representa un hiperboloide de una hoja.
6. Considere la población de pájaros dada por:

$$A = \begin{pmatrix} 0 & 3 \\ 0,4 & 0,6 \end{pmatrix} \quad p_0 = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

- a) Encuentre el número de pájaros hembras adultos y jóvenes después de 2, 5, 10 y 20 años. Después encuentre las razones a la larga de $p_{j,n}/p_{a,n}$ y de T_n/T_{n-1} y demuestre que a la larga la razón $p_{j,n}/p_{a,n}$ se acerca al valor límite k/λ_1 .
- b) Verifique que el valor propio de mayor magnitud λ_1 es positivo con multiplicidad algebraica 1, que existe un vector propio asociado cuyas componentes son todas positivas y que el otro valor propio es estrictamente menor en magnitud. Compare este valor propio mayor con $\lim_{n \rightarrow \infty} T_n/T_{n-1}$. Explique por qué estos números indican que la población está creciendo.
- c) Demuestre que si $\alpha = \beta$, entonces, a la larga, la población de pájaros aumentará siempre si cada hembra adulto produce una hembra entre sus críos.
7. En una ciudad existen dos partidos políticos, uno de derechas y otro de izquierdas. Los alcaldes son elegidos por un período de un año y se ha observado que la probabilidad de que a un alcalde de derechas suceda otro del mismo signo político es $3/5$ y que a un alcalde de izquierdas siga otro de izquierdas es $1/2$. ¿Cuál serán las probabilidades de cada partido en el año 2010?
8. La región de ventas de un vendedor la componen tres ciudades A, B y C. Nunca vende en la misma ciudad en días seguidos. Si vende en la ciudad A, entonces al día siguiente vende en la ciudad B. Sin embargo, si vende en una de las dos B o C, entonces al día siguiente está en doble posibilidad tanto para vender en A como en la otra ciudad. A la larga con qué frecuencia (en
9. En Smalltown, al 90 % de los días soleados siguen días soleados, y al 80 % de los días nublados siguen días nublados. Con esta información modelar el clima de Smalltown como cadena de Markov.
10. En relación con el problema anterior, suponga que el tiempo de mañana en Smalltown depende del tiempo que haya prevalecido los últimos dos días, como sigue; (1) Si los últimos dos días han sido soleados, entonces el 95% será soleado. (2) Si ayer estuvo nublado y hoy soleado, entonces el 70 % de las veces mañana estará soleado. (3) Si ayer estuvo soleado y hoy está nublado, entonces el 60 %, de las veces mañana estará nublado. (4) Si los últimos dos días fueron nublados, entonces el 80 % de las veces mañana será nublado. Con esta información modele el clima de Smalltown como cadena de Markov. Si el tiempo de mañana dependiera del de los últimos tres días, ¿cuántos estados se necesitarían para modelar el clima como cadena de Markov?

Nota: El método que se usa en este problema se puede aplicar para modelar un proceso estocástico de tiempo discreto como cadena de Markov. aun si X_{t+1} depende de los estados anteriores a X_t , tal como X_{t-1} en este ejemplo.



11. Resuelva los ejercicios para 2 y 4 partículas (como muestran las figuras siguientes) indicando:
- Ecuaciones del Movimiento
 - Forma Matricial
 - Valores y Vectores Propios
 - Modos Normales de Vibración y Soluciones respectivamente.

CAPÍTULO 6

Programación Lineal

6.1. Introducción

Como el lector puede observar en el índice, con este capítulo se finaliza el Seminario, el cual se inicia respondiendo a la pregunta ¿Qué es la Programación Lineal?. Después de responder a esta pregunta se recordarán de cursos de Enseñanza Media las inecuaciones de 2 variables, y para cursos más avanzados, sistemas y problemas de optimización de una función sujeta a restricciones en su forma general.

En la sección (6.6) y (6.7) veremos la forma geométrica y algebraica del modelo de dos variables que describiremos, luego se verá una aplicación *Nutrición*.

En otra sección, analizaremos el Método Símplex en su forma geométrica, y se enunciarán sugerencias para formular un modelo de Programación Lineal. Posteriormente se definirán los Tipos de Soluciones: Factible, Básica y Factible Básica.

En la sección (6.13) se describirá el Algoritmo Símplex para construir la tabla inicial, incluir las variables de holgura necesarias y utilizar el procedimiento descrito para obtener la solución óptima del problema que presentaremos.

Finalmente, analizaremos el Análisis Matricial del Método Símplex, para el caso de minimizar la función objetivo y construir la forma generalizada de la tabla inicial. Así se determinará la condición de entrada – salida, y como extraer la forma matricial de cada iteración de la tabla inicial hasta determinar cuando terminan las iteraciones.

6.2. ¿Qué es la programación lineal?

La programación lineal es una técnica matemática de optimización. En todo problema de programación lineal hay que tomar ciertas decisiones. Estas se representan con variables de decisión x_j que se utilizan en el modelo de programación lineal. La estructura básica de un problema de este tipo es maximizar o minimizar la función objetivo, satisfaciendo al mismo tiempo un grupo de condiciones restrictivas. Dicha función es una representación matemática de la meta global formulada en función de las variables de decisión x_j . Las restricciones de un problema de programación lineal pueden representarse con ecuaciones o con desigualdades (de tipo \leq y/o \geq).

Nos centraremos para empezar en aquellos problemas simples de programación lineal, los que tienen solamente 2 variables, problemas bidimensionales. Para los problemas de más variables, el procedimiento no es tan sencillo y se resuelven por el llamado método Simplex (ideado por G.B.Danzig, matemático estadounidense en 1951). Recientemente (1984) el matemático indio establecido en Estados Unidos, Narendra Karmarkar, ha encontrado un algoritmo, llamado algoritmo de Karmarkar, que es más rápido que el método simplex en ciertos casos, el lector que requiera profundizar más en este tema se le recomienda investigar el procedimiento de Karmarkar.

Así los problemas del tipo en que intervienen gran número de variables son implementados en ordenadores para su resolución, debido a que la dificultad para encontrar una solución óptima o factible cualquiera sea el caso aumenta considerablemente.

6.3. Inecuaciones Lineales

6.3.1. Inecuaciones lineales con 2 variables

Una inecuación lineal con 2 variables es una expresión de la forma:

$$ax + by \leq c$$

(donde el símbolo \leq puede ser también \geq , $<$ o bien $>$), donde a, b y c son números reales y x e y las incógnitas.

Para resolver estas inecuaciones, recordará de otros cursos que debe representar gráficamente en el plano la recta dada por la correspondiente ecuación lineal y marcar una de las dos regiones en que dicha recta divide al plano.

Ejemplo 6.1 Si queremos resolver la inecuación: $2x + 3y \geq 3$, representamos en primer lugar la recta $2x + 3y = 3$

La recta divide al plano en dos regiones, una de las cuales es la solución de la inecuación. Para saber qué parte es, hay dos procedimientos:

1. Se despeja y de la inecuación, poniendo cuidado en que si en una inecuación multiplicamos o dividimos por un número negativo, la desigualdad cambia de sentido. En este caso tendríamos que:

$$y \geq \frac{-3 - 2x}{3}$$

Observando el dibujo vemos que la recta divide al eje de ordenadas y en dos partes. La solución de la inecuación sería aquella parte en la que y sea mayor que la recta, es decir, la parte superior.

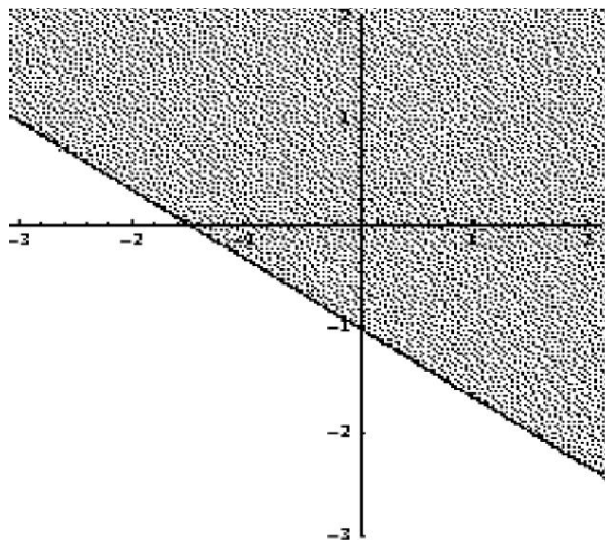


Figura 6.1: Solución de la inecuación lineal

- Se toma un punto cualquiera que no pertenezca a la recta, por ejemplo el (1,2). Para que dicho punto sea solución, se debe cumplir la desigualdad, por lo que sustituimos en la inecuación inicial el punto (1,2): $2 \cdot 1 + 3 \cdot 2 \geq 3$, es decir, $8 \geq 3$.

Como esta última desigualdad es evidentemente cierta, concluimos que el punto (1,2) es solución y por tanto el semiplano que contiene al (1,2) es la solución, es decir el semiplano superior, como habíamos obtenido antes.

Cualquiera de los procedimientos es válido.

6.3.2. Sistema de inecuaciones lineales con 2 variables

Un sistema de inecuaciones lineales es un conjunto de inecuaciones del tipo anterior y resolverlo consistirá en resolver gráficamente cada inecuación (como en el caso anterior), representar la solución en un mismo gráfico y la solución finalmente será la intersección de todas las soluciones.

Ejemplo 6.2 Resolver el sistema de inecuaciones siguiente:

$$\begin{aligned} 2x + 3y &\geq -3 \\ 2x - y - 9 &\leq 0 \\ 2x - 5y - 5 &\geq 0 \end{aligned}$$

Si representamos las rectas:

$$\begin{aligned} 2x + 3y &= -3 && (\text{recta } r) \\ 2x - y - 9 &= 0 && (\text{recta } s) \\ 2x - 5y - 5 &= 0 && (\text{recta } t) \end{aligned}$$

El triángulo achurado es la solución del sistema.

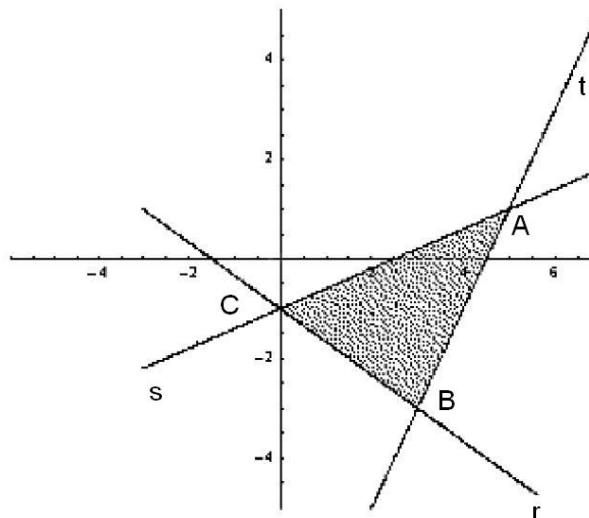


Figura 6.2: Solución del sistema de inecuaciones lineales

Además, para los problemas de programación lineal es necesario calcular las coordenadas de los vértices de la región solución. Su cálculo se reduce a resolver sistemas de ecuaciones lineales con dos incógnitas, que provienen de igualar las ecuaciones de las rectas correspondientes. En el ejemplo anterior si queremos el punto intersección de las rectas r y t tendremos que resolver el sistema formado por:

$$\begin{aligned} 2x + 3y &= -3 \\ 2x - y - 9 &= 0 \end{aligned}$$

luego

$$\begin{aligned} -2x - 3y &= 3 \\ 2x - y - 9 &= 0 \end{aligned}$$

Sumando $-4y = 12 \implies y = -3$.

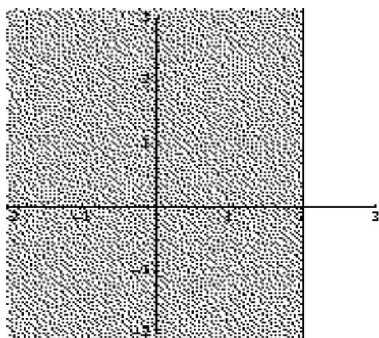
Sustituyendo queda $2x + 3(-3) = -3$, es decir $2x - 9 = -3$, y entonces $x = 3$. Luego r y t se cortan en el punto $(3, -3)$. Los otros dos puntos se dejan como ejercicio al lector.

Rectas horizontales y verticales

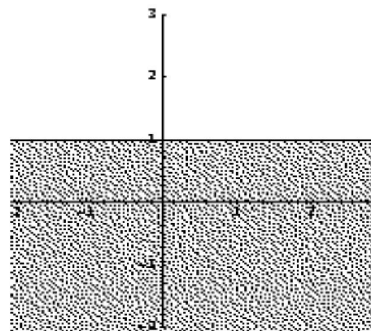
En ocasiones, en estos sistemas, aparecen inecuaciones del tipo $x \geq k$ o bien $y \geq k$, donde falta alguna de las dos incógnitas.

Estas inecuaciones en realidad corresponden a rectas horizontales y verticales (su representación es sencilla). Por ejemplo, la inecuación $x \leq 2$ no es más que el conjunto de puntos a la izquierda de la recta vertical que pasa por el punto $x = 2$, gráficamente:

Lo mismo ocurre con $y \leq 1$, que será en este caso el semiplano inferior a la recta horizontal $y = 1$.



Gráfica $x \leq 2$



Gráfica $y \leq 1$

6.4. Problemas de optimización de una función sujeta a restricciones

En un problema de programación lineal de dos variables x e y se trata de optimizar (hacer máxima o mínima, según los casos) una función (llamada función objetivo) de la forma:

$$f(x, y) = A \cdot x + B \cdot y$$

sujeta a una serie de restricciones dadas mediante un sistema de inecuaciones lineales del tipo:

$$\begin{aligned} a_1x + b_1y & (\leq, \geq, =) c_1 \\ a_2x + b_2y & (\leq, \geq, =) c_2 \\ & \vdots \\ a_mx + b_my & (\leq, \geq, =) c_m \end{aligned}$$

Los puntos del plano que cumplen el sistema de desigualdades forman un recinto convexo acotado (poligonal) o no acotado, llamado región factible del problema. Todos los puntos de dicha región cumplen el sistema de desigualdades. Se trata de buscar, entre todos esos puntos, aquel o aquellos que hagan el valor de $f(x, y)$ máximo o mínimo, según sea el problema.

Los puntos de la región factible se denominan soluciones factibles. De todas esas soluciones factibles, aquellas que hacen óptima (máxima o mínima) la función objetivo se llaman soluciones óptimas. En general, un problema de programación lineal puede tener una, infinitas o ninguna solución.

Lo que si se verifica es la siguiente proposición:

Proposición 6.1 *Si hay una única solución óptima, ésta se encuentra en un vértice de la región factible (Punto esquina), y si hay infinitas soluciones óptimas, se encontrarán en un lado de la región factible.*

Observación 6.1 *Es posible que no haya solución óptima, pues cuando el recinto es no acotado, la función objetivo puede crecer o decrecer indefinidamente.*

Para resolver el problema, podemos abordarlo de dos formas, pero antes de aplicar cualquiera de ellas siempre hay que dibujar la región factible, resolviendo el sistema de inecuaciones lineales correspondiente, como se ha visto en los ejemplos anteriores.

6.4.1. Forma geométrica

En este caso se representa el vector director de la recta que viene dada por la ecuación de la función objetivo,

$$f(x, y) = A \cdot x + B \cdot y$$

que hay que maximizar o minimizar.

El vector director de la recta $A \cdot x + B \cdot y$ viene dado por $v = (-B, A)$. Como lo que nos importa mayormente es su dirección y no su módulo (longitud), podemos dividir a las coordenadas del vector si los números son muy grandes, puesto que vectores con coordenadas proporcionales tienen la misma dirección.

Posteriormente, se trazan rectas paralelas a este vector que pasen por los vértices de la región factible (si es acotada), o por todo el borde de la región factible (cuándo no es acotada) y se observa en qué vértice la función f se hace máxima (o mínima) sin más que tener en cuenta cuál de las rectas tiene mayor (o menor) ordenada en el origen, es decir, qué recta corta en un punto mayor o menor al eje y .

Ejemplo 6.3 Maximizar la función $f(x, y) = 2000 \cdot x + 5000 \cdot y$ sujeta a las restricciones:

$$\begin{aligned} 2x + 3y &\geq -3 \\ 2x - y - 9 &\leq 0 \\ 2x - 5y - 5 &\geq 0 \end{aligned}$$

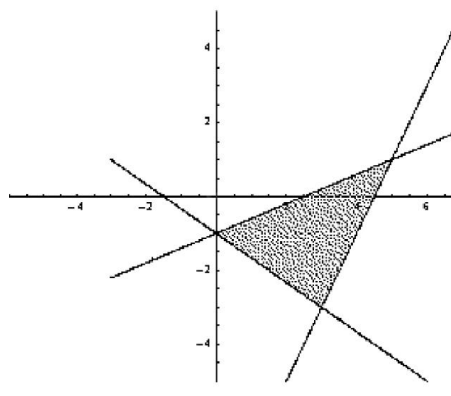


Figura 6.3: Región Factible

Después de graficar las rectas

$$\begin{aligned} 2x + 3y &= -3 \\ 2x - y - 9 &= 0 \\ 2x - 5y - 5 &= 0 \end{aligned}$$

achuramos las soluciones de las inecuaciones y obtenemos la intersección de todas las soluciones, llamada región factible. En este caso la figura siguiente denota la región factible; cuyos vértices son los puntos $(0, -1)$, $(5, 1)$ y $(3, -3)$.

Como la función es $f(x, y) = 2000 \cdot x + 5000 \cdot y$, el vector director es $v = (-5000, 2000)$, que tiene la misma dirección que el $v = (-5, 2)$. Su representación se muestra en la figura:

Se trata ahora de trazar paralelas al vector que pasen por los vértices anteriores, es decir:

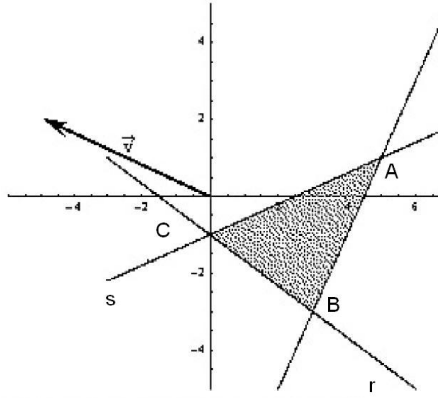


Figura 6.4: Región factible y vector de la función objetivo

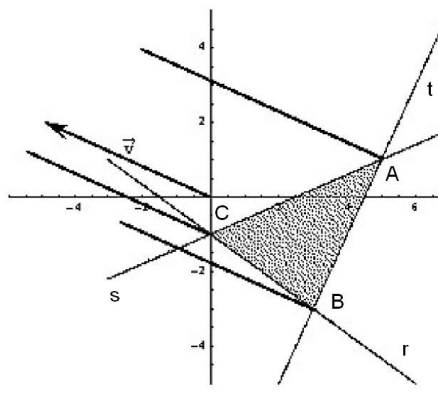


Figura 6.5: Solución gráfica. Paralelas al vector por los vértices.

Se observa gráficamente que de las tres paralelas trazadas, la que corta al eje y en un punto mayor es la que pasa por el punto $(5, 1)$, que por tanto será la solución óptima al problema de máximos planteado.

Para saber cuál es este valor máximo, sustituimos en la función:

$$f(5, 1) = 2000 \cdot 5 + 5000 \cdot 1 = 10000 + 5000 = 15000$$

Luego la función tiene su solución óptima en $(5, 1)$ donde toma el valor 15000.

6.4.2. Forma algebraica

Consiste en sustituir cada uno de los vértices de la región en la función objetivo. La solución óptima vendrá dada por aquel que tome el mayor (o menor) valor.

Considerando el mismo ejemplo:

Ejemplo 6.4 Maximizar la función $f(x, y) = 2000 \cdot x + 5000 \cdot y$ sujeta a las restricciones:

$$\begin{aligned} 2x + 3y &\geq -3 \\ 2x - y - 9 &\leq 0 \\ 2x - 5y - 5 &\geq 0 \end{aligned}$$

obviamente se obtiene la misma región factible, cuyos vértices eran los puntos $(0, -1)$, $(5, 1)$ y $(3, -3)$. Resolver el problema mediante la forma algebraica consiste en sustituir los vértices:

$$\begin{aligned} f(5, 1) &= 2000 \cdot 5 + 5000 \cdot 1 = 10000 + 5000 = 15000 \\ f(0, -1) &= 2000 \cdot 0 + 5000 \cdot (-1) = 0 - 5000 = -5000 \\ f(3, -3) &= 2000 \cdot 3 + 5000 \cdot (-3) = 6000 - 15000 = -9000 \end{aligned}$$

Se observa que el valor máximo se alcanza en el vértice $(5, 1)$ y que dicho valor es 15000. Como puede notar es la misma solución que se obtuvo mediante la forma geométrica.

6.5. Aplicación a problemas concretos de 2 variables

El verdadero valor de las técnicas de la programación lineal consiste en poder aplicarlas a problemas reales. Resolveremos un problema, mostrando los pasos a seguir para encontrar la solución óptima que lo satisfaga.

Ejemplo 6.5 *Una fábrica de muebles fabrica dos tipos de sillones, S_1 y S_2 . La fábrica cuenta con dos secciones; carpintería y tapicería. Hacer un sillón de tipo S_1 requiere 1 hora de carpintería y 2 de tapicería, mientras que uno de tipo S_2 requiere 3 horas de carpintería y 1 de tapicería. El personal de tapicería trabaja un total de 80 horas, y el de carpintería 90. Las ganancias por las ventas de S_1 y S_2 (unidad) son, respectivamente 60 y 30 euros. Calcular ¿cuántos sillones de cada tipo hay que hacer para maximizar las ganancias?*

Este es un problema típico en el que hay que usar las técnicas de programación lineal para encontrar la solución que optimiza la función objetivo y así responder a la pregunta del ejemplo. Intentaremos seguir el siguiente esquema:

1. Leer el enunciado, determinar la función objetivo y definir las variables.
 En este caso, queremos hacer máximo el beneficio, es decir, queremos maximizar una función. Como queremos determinar las cantidades de sillones S_1 y S_2 respectivamente, llamemos
 $x = n^\circ$ de unidades de S_1 e
 $y = n^\circ$ de unidades de S_2 .
 La función beneficio a maximizar será : $B(x, y) = 60 \cdot x + 30 \cdot y$, que es la función objetivo.
2. Reordenar los datos del problema y escribir las inecuaciones correspondientes.
 En este paso es conveniente el uso de tablas:

Tiempo(horas)	Carpintería	Tapicería
S_1	1	2
S_2	3	1
Disponible	90	80

Tiempo(horas)	Cantidad	Carpintería	Tapicería
S_1	x	x	2x
S_2	y	3y	y
Necesario		x+3y	2x+y
Disponible		90	80

De aquí se deduce que:

$$x + 3y \leq 90$$

$$2x + y \leq 80$$

y además

$$x \geq 0$$

$$y \geq 0$$

pues el n° de unidades producidas no puede ser negativo. Ya tenemos por tanto las restricciones.

3. Representar gráficamente la región factible, calcular sus vértices y el vector si usamos el método geométrico.

En este caso, representando la región factible:

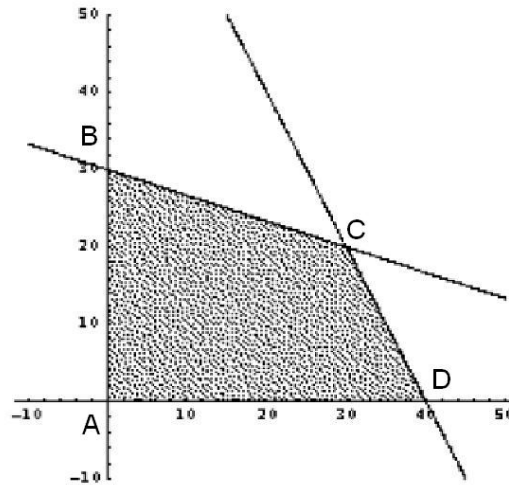


Figura 6.6: Región Factible

Siendo los vértices $A = (0, 0)$, $B = (0, 30)$, $C = (30, 20)$, $D = (40, 0)$. El vector será $(-30, 60)$, equivalente a $(-10, 20)$.

Gráficamente se observa que la solución no es única, sino, se encuentran infinitas soluciones en el lado \overline{CD} , sobre la recta $2x + y = 80$, con x variando desde 30 hasta 40, todas las soluciones son válidas.

4. Sustituir las coordenadas en la función objetivo y dar la solución correcta.

En nuestro ejemplo:

$$f(0, 0) = 0$$

$$f(0, 30) = 900$$

$$f(30, 20) = 2400$$

$$f(40, 0) = 2400$$

como $f(30, 20) = 2400 = f(40, 0)$ se entiende que hay infinitas soluciones y el beneficio que se obtiene es 2400 euros.

- Analizar la solución obtenida en el contexto del problema: ¿tiene sentido?.

Debemos interpretar que en el contexto del problema no todas las soluciones son válidas, sino que sólo sirven soluciones enteras.

Encontramos por tanto sólo 11 soluciones las cuales se enuncian en la tabla:

S_1	30	31	32	33	34	35	36	37	38	39	40
S_2	20	18	16	14	12	10	8	6	4	2	0

En cualquiera de estas soluciones el beneficio es de 2400 euros, que es el máximo, bajo las condiciones del problema.

6.6. ¿Qué es el Método Simplex?

Definición 6.1 (Método Simplex) *El Método Simplex es una técnica algebraica para resolver los sistemas de ecuaciones a partir de problemas de programación lineal en que ha de optimizarse la función objetivo. Es un proceso iterativo que indica la solución factible inicial, posteriormente se busca la existencia de una mejor solución, que se mide según pueda o no mejorarse el valor de la función objetivo. Si se encuentra una solución más satisfactoria, se reanuda la búsqueda. Para generar cada solución sucesiva se requiere resolver un sistema de ecuaciones lineales. La búsqueda continúa hasta que la función objetivo ya no admite mejoramiento.*

Antes de iniciar las explicaciones del método simplex, se dará una fórmula generalizada de un modelo de programación lineal, con las siguientes definiciones:

- x_j =j-ésima variable de decisión
- c_j =coeficiente en la j-ésima variable de decisión dentro de la función objetivo
- a_{ij} =coeficiente en la i-ésima restricción de la j-ésima variable
- b_i =constante del miembro derecho para la i-ésima restricción

el modelo generalizado de la programación lineal puede formularse así:

$$\begin{array}{rcl}
 \text{Optimice(Máx o Mín)} & z = & c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 \\
 \text{sujeta a} & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & (\leq, \geq, =) b_1 \quad (1) \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & (\leq, \geq, =) b_2 \quad (1) \\
 & & \vdots = \vdots \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & (\leq, \geq, =) b_m \quad (m) \\
 & x_1 & \geq 0 \\
 & x_2 & \geq 0 \\
 & \vdots & \vdots \\
 & x_n & \geq 0
 \end{array}$$

Este modelo generalizado tiene n variables de decisión y m restricciones estructurales. Nótese que cada restricción estructural no contiene más de una de las condiciones de tipo $(\leq, \geq, =)$ asignada a ella. Al describir el método simplex, se utilizará esporádicamente la notación de este modelo.

Para describir mejor el método simplex vamos a resolver un problema siguiendo una serie de pasos con el fin de encontrar una solución óptima.

$$\text{Maximizar } Z = f(x, y) = 3x + 2y$$

sujeto a

$$\begin{array}{rcl}
 2x + y & \leq & 18 \\
 2x + 3y & \leq & 42 \\
 3x + y & \leq & 24 \\
 x & \geq & 0 \\
 y & \geq & 0
 \end{array}$$

Se consideran las siguientes fases:

1. Convertir las desigualdades en igualdades.

Se introduce una variable de holgura por cada una de las restricciones, para convertirlas en igualdades, resultando el sistema de ecuaciones lineales:

$$\begin{array}{rcl}
 2x + y + h & = & 18 \\
 2x + 3y + s & = & 42 \\
 3x + y + d & = & 24
 \end{array}$$

2. Igualar la función objetivo a cero

$$-3x - 2y + Z = 0$$

3. Escribir la tabla inicial simplex

En las columnas aparecerán todas las variables del problema y en las filas los coeficientes de las igualdades obtenidas, una fila para cada restricción y la última fila con los coeficientes de la función objetivo:

Tabla de iteración N°1

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
h	2	1	1	0	0	18
s	2	3	0	1	0	42
d	3	1	0	0	1	24
Z	-3	-2	0	0	0	0

4. Encontrar la variable de decisión que entra en la base y la variable de holgura que sale de la base.

- Para escoger la variable de decisión que entra en la base, nos fijamos en la última fila, la de los coeficientes de la función objetivo y escogemos la variable con el coeficiente negativo mayor (en valor absoluto). En nuestro caso, la variable x de coeficiente -3 . Si existiesen dos o más coeficientes iguales que cumplan la condición anterior, entonces se elige uno (cualquiera de ellos).

Si en la última fila no existiese ningún coeficiente negativo, significa que se ha alcanzado la solución óptima. Por tanto, lo que va a determinar el final del proceso de aplicación del método del simplex, es que en la última fila no haya elementos negativos.

La columna de la variable que entra en la base se llama columna pivote.

- Para encontrar la variable de holgura que tiene que salir de la base, se divide cada término de la última columna (valores solución) por el término correspondiente de la columna pivote, siempre que estos últimos sean mayores que cero. En nuestro caso: $\frac{18}{2} = 9$, $\frac{42}{2} = 21$ y $\frac{24}{3} = 8$

Si hubiese algún elemento menor o igual que cero no se hace dicho cociente. En el caso de que todos los elementos fuesen menores o iguales a cero, entonces tendríamos una solución no acotada y no se puede continuar.

El término de la columna pivote que en la división anterior de lugar al menor cociente positivo, el número 3 (en nuestro caso; ya que 8 es el mínimo), esto indica que la fila de la variable de holgura que sale de la base es d . Esta fila se denomina fila pivote.

Si al calcular los cocientes dos o más son iguales, indica que cualquiera de las variables correspondientes pueden salir de la base.

- En la intersección de la fila pivote y columna pivote tenemos el elemento pivote operacional, 3.

5. Encontrar los coeficientes de la nueva tabla.

Los nuevos coeficientes de x se obtienen dividiendo todos los coeficientes de la fila d por el pivote operacional 3, que es el que hay que convertir en 1.

A continuación mediante la eliminación gaussiana hacemos ceros los términos restantes de la columna, con lo que obtenemos los nuevos coeficientes de las otras filas incluyendo los de la función objetivo Z .

También se puede hacer utilizando el siguiente esquema:

Fila del pivote:

$$\text{Nueva fila del pivote} = (\text{Vieja fila del pivote}) : (\text{Pivote})$$

Resto de las filas:

$$\text{Nueva fila} = (\text{Vieja fila}) - (\text{Coeficiente de la vieja fila en la columna de la variable entrante}) \times (\text{Nueva fila del pivote})$$

Veámoslo con un ejemplo una vez calculada la fila del pivote (fila de x en la Tabla de iteración N°2):

Vieja fila de s	2	3	0	1	0	42
	-	-	-	-	-	-
Coeficiente	2	2	2	2	2	2
	×	×	×	×	×	×
Nueva fila pivote	1	1/3	0	0	1/3	8
	=	=	=	=	=	=
Nueva fila de s	0	7/3	0	1	-2/3	26

Tabla de iteración N°2

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
h	0	1/3	1	0	-2/3	2
s	0	7/3	0	1	-2/3	26
d	1	1/3	0	0	1/3	8
Z	0	-1	0	0	1	24

Como en los elementos de la última fila hay un número negativo, -1, significa que no hemos llegado todavía a la solución óptima. Hay que repetir el proceso:

- La variable que entra en la base es y , por ser la variable que corresponde al coeficiente -1
- Para calcular la variable que sale, dividimos los términos de la última columna entre los términos correspondientes de la nueva columna pivote: $\frac{2}{1/3} = 6$, $\frac{26}{7/3} = \frac{78}{7} \approx 11,1428$ y $\frac{8}{1/3} = 24$ y como el menor cociente positivo es 6, tenemos que la variable de holgura que sale es h .
- El elemento pivote, que ahora hay que hacer 1, es 1/3.

Operando de forma análoga a la anterior obtenemos la tabla:

Tabla de iteración N°3

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
h	0	1	3	0	-2	6
s	0	0	-7	0	4	12
d	1	0	-1	0	1	6
Z	0	0	3	0	-1	30

Como en los elementos de la última fila hay uno negativo, -1, significa que no hemos llegado todavía a la solución óptima. Hay que repetir el proceso:

- La variable que entra en la base es d , por ser la variable que corresponde al coeficiente -1
- Para calcular la variable que sale, dividimos los términos de la última columna entre los términos correspondientes de la nueva columna pivote: $\frac{6}{-2} = -3$, $\frac{12}{4} = 3$ y $\frac{6}{1} = 6$ y como el menor cociente positivo es 3 , tenemos que la variable de holgura que sale es s .
- El elemento pivote que ahora hay que hacer 1 , es 4 .

Obtenemos la tabla:

Tabla de iteración N°4

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
h	0	1	-1/2	0	0	12
s	0	0	-7/4	0	1	3
d	1	0	-3/4	0	0	3
Z	0	0	5/4	0	0	33

Como todos los coeficientes de la fila de la función objetivo son positivos, hemos llegado a la solución óptima.

Los solución óptima viene dada por el valor de Z en la columna de los valores solución, en nuestro caso 33 . En la misma columna se puede observar el vértice donde se alcanza, observando las filas correspondientes a las variables de decisión que han entrado en la base: $D(3, 12)$.

Observación 6.2 ▪ *Si en el problema de maximizar apareciesen como restricciones inecuaciones de la forma: $ax + by < c$; multiplicándolas por -1 se transforman en inecuaciones de la forma $-ax - by > -c$ y estamos en el caso anterior.*

- *Si en lugar de maximizar se trata de un problema de minimizar se sigue el mismo proceso, pero cambiando el sentido del criterio, es decir, para entrar en la base se elige la variable (cuyo valor en la fila de la función objetivo) sea el mayor de los positivos. El proceso (iteraciones) se finaliza cuando todos los coeficientes de la fila de la función objetivo son negativos.*

6.6.1. Interpretación geométrica del método símplex

Las tablas sucesivas que hemos construido van proporcionando el valor de la función objetivo en los distintos vértices, ajustándose a la vez los coeficientes de las variables iniciales y de holgura.

En la primera iteración (Tabla N°1) han permanecido todos los coeficientes iguales, se ha calculado el valor de la función objetivo en el vértice $A(0, 0)$, siendo este 0 .

A continuación se desplaza por la arista AB , calculando el valor de f hasta llegar a B . Este paso aporta la Tabla N°2.

En esta segunda iteración se ha calculado el valor que corresponde al vértice $B(8, 0) \rightarrow Z = f(8, 0) = 24$. Sigue por la arista BC hasta llegar a C , donde se para y despliega los datos de la Tabla N°3. En esta tercera iteración se ha calculado el valor que corresponde al vértice $C(6, 6) \rightarrow Z = f(6, 6) = 30$.

Continúa haciendo cálculos a través de la arista CD hasta llegar al vértice D . Los datos que se reflejan son los de la Tabla N°4. Concluye con esta tabla, advirtiendo que ha terminado (antes ha comprobado

que la solución no mejora al desplazarse por la arista DE) El valor máximo de la función objetivo es 33, y corresponde a $x = 3$ e $y = 12$ (vértice D).

Si calcula el valor de la función objetivo en el vértice $E(0, 14)$, su valor no supera el valor 33.

Para mayor claridad de la interpretación geométrica del Método Simplex veremos a continuación un ejemplo orientado a la nutrición.

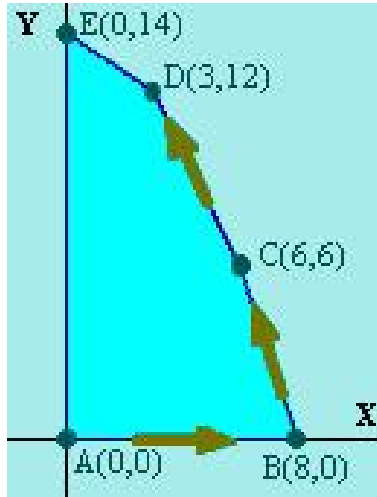


Figura 6.7: Solución gráfica

6.6.1.1. Aplicación a la Nutrición

Ejemplo 6.6 Imaginemos que las necesidades semanales mínimas de una persona en proteínas, hidratos de carbono y grasas son respectivamente 8, 12 y 9 unidades. Supongamos que debemos obtener un preparado con esa composición mínima mezclando dos productos A y B, cuyos contenidos por Kg son los que se indican en la siguiente tabla

	Proteínas	Hidratos	Grasas	Coste/Kg
A	2	6	1	600
B	1	1	3	400

Responder:

- a) ¿Cuántos Kg de cada producto deberán comprarse semanalmente para que el costo de preparar la dieta sea mínimo? (Solución factible)
- b) ¿Cuántos Kg de cada producto deberíamos comprar si el precio de A subiera a 1.000 pts/Kg?

Para dar una solución a este problema típico de dieta, se procederá a resolver no por tabulación o método algebraico, sino, por el método geométrico.

Primero definimos las variables, llamamos x al número de kilogramos usados del producto A e y al número de kilogramos usados del producto B.

El problema nos indica que tenemos que encontrar el mínimo de la función objetivo con sus correspondientes restricciones.

Nuestro problema en la forma estándar queda:

$$\text{Min } f(x, y) = z = 600x + 400y$$

$$2x + y \geq 8$$

$$6x + y \geq 12$$

$$x + 3y \geq 9$$

$$x, y \geq 0$$

Para poder dar solución a este problema, se graficarán las inecuaciones, resultando la siguiente gráfica:

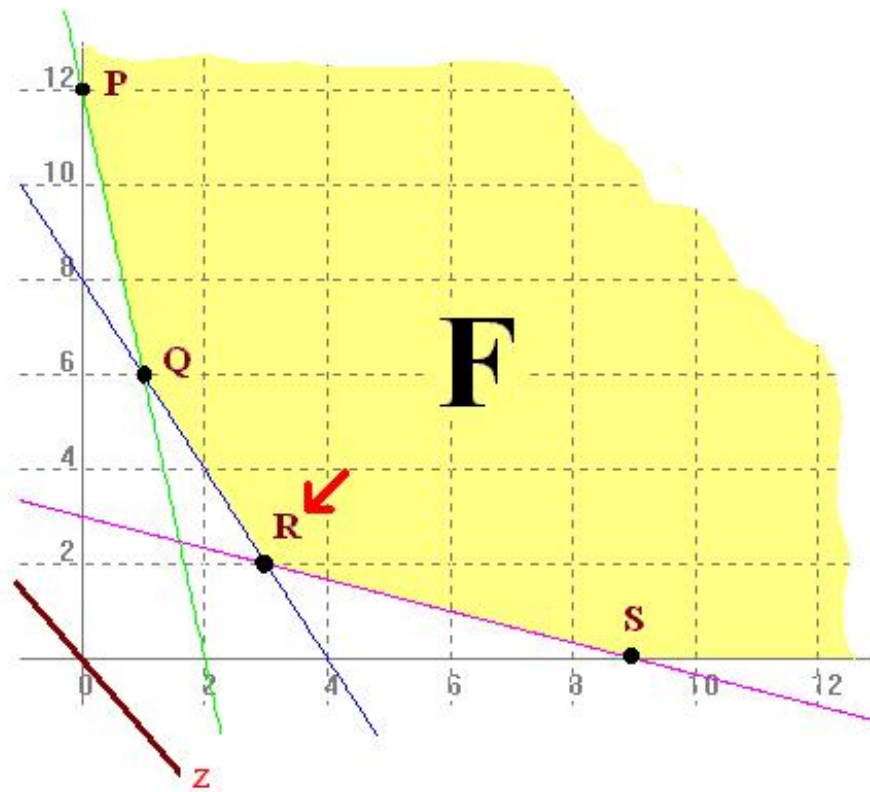


Figura 6.8: Solución gráfica

a) Solución Factible Óptima

Si observamos el gráfico nos podremos fijar que existen 3 puntos de la región F que son soluciones factibles y por paralelismo con la recta de beneficio nulo Z vemos que $R(3,2)$ es el punto mínimo. Por tanto, deben comprarse 3 kg. de A y 2 kg. de B para que el gasto sea mínimo.

b) Valor del Programa Lineal

Cuando la función objetivo es $F(X) = 600x + 400y$ el valor del programa lineal (gasto) es 2.600 pts. Si la función objetivo es $F(X) = 100x + 400y$ la solución óptima está en el punto $Q(1,6)$ y el valor del programa lineal (gasto) es 3.400 pts.

6.7. Sugerencias para formular modelos de programación lineal

Las sugerencias para formular modelos de programación lineal son:

1. Leer comprensivamente el problema.
2. Identificar las variables de decisión y luego enumerarlas para dar una definición escrita que describan las variables del problema.
3. Identificar el objetivo que ha de maximizar o minimizar.
4. Identificar las restricciones estructurales, para eso responder ¿Qué condiciones deben satisfacerse cuando se asignan valores a la variables de decisión? ¿Qué restricciones prohíben llevar a la infinidad (positivas o negativas) el valor de la función objetivo?
5. Formular por escrito el modelo matemático, sin olvidar las restricciones de no negatividad.

6.8. Tipos de soluciones

Definiciones:

- **Definición 6.2 (Solución Factible)** *Es cualquier conjunto de valores de k variables que satisfacen tanto las restricciones estructurales como la no negatividad.*
- **Definición 6.3 (Solución Básica)** *Es cualquier solución que se obtiene al hacer igual a 0 las variables $(k - m)$, o aquellas a las que se han asignado valores 0, se llaman variables no básicas.*
- **Definición 6.4 (Solución Factible Básica)** *Es una solución básica que además cumple con las restricciones de no negatividad.*

Una vez que los problemas de programación lineal han sido convertidos en la forma estándar en que todas las restricciones se reformulan como igualdades y se han agregado variables complementarias. El sistema resultante tiene más variables que ecuaciones. Supóngase que tenemos un problema que posea m restricciones de tipo (\leq) y n variables de decisión. Al ser transformado en su forma estándar, habrá m restricciones estructurales, n variables de decisión y m variables de holgura. Por tanto, el sistema de restricciones puede considerarse un sistema $m \times (n + m)$ donde $m < n + m$. Es decir, el número de ecuaciones es menor que el de variables.

6.9. Algoritmo Simplex

1. Escribir la matriz aumentada del sistema de ecuaciones. Ésta recibe el nombre de tabla inicial simplex.
2. Localizar el elemento negativo en la última fila, que no sea el último elemento con la mayor magnitud (si dos o más entradas comparten esta propiedad, cualquiera de ellas se puede elegir). Si todas las entradas son no negativas, la tabla se encuentra en su forma final.

3. Dividir cada elemento positivo de la columna definida por esta entrada negativa entre el elemento correspondiente de la última columna.
4. Elegir el divisor que arroje el cociente más pequeño. Este elemento recibe el nombre de elemento pivote (si dos o más elementos comparten esta propiedad, cualquiera de ellos se puede elegir como pivote).
5. Utilizar operaciones por filas para crear un 1 en la ubicación del pivote y 0 en cualquier otra parte de la columna pivote.
6. Repetir los pasos 2 y 5 hasta eliminar los elementos negativos de la última fila. La matriz final recibe el nombre de tabla final símplex, esta permite obtener la solución óptima.

Observación 6.3 *Un problema relacionado con la determinación del valor mínimo de una función objetivo f se puede resolver al considerar el valor máximo de $-f$, el valor negativo de f , en la misma región factible, el resultado general es el siguiente: El valor mínimo de una función f en una región S ocurre en él o los puntos de valor máximo de $-f$, lo cual corresponde al negativo de dicho valor máximo.*

6.10. Aplicación

Una aplicación de suma importancia para este texto, es el Análisis matricial del método símplex, y la forma de la base (tabla inicial), para encontrar una solución óptima del problema; a continuación se dará la explicación de este.

6.10.1. Análisis Matricial del método símplex

Recordemos que todo problema de Programación lineal (P.L) podemos expresarlo mediante el siguiente formato estándar para desigualdades.

$$\min Z = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

$$\begin{array}{rcccc} a_{11}x_1 + \dots + a_{1n}x_n & \leq & b_1 & & \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & \leq & b_m & & \end{array}$$

En notación matricial

$$\begin{array}{l} \min \vec{c} \cdot \vec{x} \\ A\vec{x} \leq \vec{b} \end{array}$$

La forma estándar una vez agregadas las variables de holgura y exceso necesarias, es:

$$\begin{array}{rcl} c^T x & & \\ Ax & = & b \\ x & \geq & 0 \end{array}$$

Observación 6.4 *Los componentes del vector b deben ser mayor o igual a cero.*

2. En cada iteración los componentes del vector x se dividen en:

- Las variables básicas (V.B) : X_B y
- Las variables no básicas (V.N.B) : X_D

Con esta nueva nomenclatura el problema se puede reescribir como:

$$\begin{aligned} \min \quad & c_B^T X_B + c_D^T X_D \\ & BX_B + DX_D = b \\ & X_B, X_D \geq 0 \end{aligned}$$

Es decir, la matriz B es una matriz de $m \times m$ que debe ser de rango máximo. Si multiplicamos la restricción en ambos lados por B^{-1} nos queda:

$$\begin{aligned} B^{-1}(BX_B + DX_D) &= B^{-1}b \\ B^{-1}BX_B + B^{-1}DX_D &= B^{-1}b \\ I_n X_B + B^{-1}DX_D &= B^{-1}b \\ \Rightarrow X_B &= B^{-1}b - B^{-1}DX_D \end{aligned}$$

Reemplazando en la función objetivo queda:

$$\hat{v} = c_B^T B^{-1}b + (c_D^T - c_B^T B^{-1}D)X_D$$

Ahora nuestro problema adquiere la siguiente forma

$$\begin{aligned} \min \quad & c_B^T B^{-1}b + (c_D^T - c_B^T B^{-1}D)X_D \\ & X_B + B^{-1}DX_D = B^{-1}b \\ & X_B, X_D \geq 0 \end{aligned}$$

Si todos los elementos de $B^{-1}b$ son no negativos, la solución que se obtuvo en esa iteración corresponde a una solución factible del problema

O sea, $X_B = B^{-1}b, X_D = 0$

El valor de la función objetivo es: $\hat{v} = c_B^T B^{-1}b$

Si analizamos el problema en una tabla que la llamaremos tabla inicial queda:

X_B	X_D	
B	D	b
c_B^T	c_D^T	0

Luego nuestra tabla inicial correspondiente a una solución básica es

X_B	X_D	
I	$B^{-1}D$	$B^{-1}b$
0	$c_D^T - c_B^T B^{-1}D$	$- c_B^T B^{-1}b$

Si definimos los elementos del vector que llamaremos vector de costos reducidos:

$$r_D^T = c_D^T - c_B^T B^{-1} D$$

estos son no negativos, la solución básica factible es también óptima.

Observe que cada elemento de este vector puede expresarse como $c_j - c_B^T B^{-1} D_j$ donde c_j y D_j corresponden al coeficiente asociado a la variable X_j , respectivamente.

Para facilitar el entendimiento de los enunciados anteriores se procederá a resolver un ejemplo ilustrativo, caracterizando la forma matricial de este, hasta encontrar la solución óptima.

Ejemplo 6.7

$$\text{Min } Z = -X_1 - 2X_2 - 3X_3$$

$$\begin{aligned} \text{s.a } -3X_1 + 2X_2 + X_3 &\leq 2 \\ -X_1 + 2X_2 - X_3 &\leq 4 \\ X_1 + X_2 + 4X_3 &\leq 5 \\ X_1, X_2, X_3 &\geq 0 \end{aligned}$$

Para transformar el problema a formato estándar hay que agregar tres variables de holgura, que forman además la base inicial.

$$\text{Min } Z = -X_1 - 2X_2 - 3X_3$$

$$\begin{aligned} \text{s.a } -3X_1 + 2X_2 + X_3 + X_4 &= 2 \\ -X_1 + 2X_2 - X_3 + X_5 &= 4 \\ X_1 + X_2 + 4X_3 + X_6 &= 5 \\ X_1, X_2, X_3, X_4, X_5, X_6 &\geq 0 \end{aligned}$$

Sean las variables básicas y no básicas $X_B = X_4, X_5, X_6$ y $X_D = X_1, X_2, X_3$

Luego las matrices iniciales son:

$$B = B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 2 & 1 \\ -1 & 2 & -1 \\ 1 & 1 & 4 \end{pmatrix}$$

$$b = \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} \quad c_B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad c_D = \begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix}$$

Por consiguiente, la tabla inicial es:

X_1	X_2	X_3	X_4	X_5	X_6		
-3	2	1	1	0	0	2	X_4
-1	2	-1	0	1	1	4	X_5
1	1	4	0	0	1	5	X_6
-1	-1	-3	0	0	0	0	

Observación 6.5 El vector c_B es nulo, por lo tanto $r_D^T = c_D^T$

2. Todas las variables no básicas (V.N.B) tienen costo reducido negativo
3. $\min\{-1, -2, -3\} = -3 \Rightarrow X_3$ entra a la base.

4. Para la variable saliente debemos realizar el cociente entre los elementos de la columna de $B^{-1}b$ y los elementos positivos de la columna de D asociado a X_3 : $\min\{\frac{2}{1}, \frac{5}{4}\} = \frac{5}{4} \Rightarrow X_6$ sale de la base.
5. La próxima tabla la podemos obtener pivoteando (mediante operaciones filas) en la posición de X_3 de la tercera fila, o bien simplemente redefiniendo las matrices del problema acorde a la nueva base.

$$X_B = X_4, X_5, X_3$$

6. Las variables no básicas (V.N.B) son $X_D = X_1, X_2, X_6$. Las nuevas matrices asociadas a esta nueva base son:

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{pmatrix} \Rightarrow B^{-1} = \begin{pmatrix} 1 & 0 & 1/4 \\ 0 & 1 & -1/4 \\ 0 & 0 & 1/4 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 2 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$$b = \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}, \quad c_B = \begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix}, \quad c_D = \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}$$

Observación 6.6 Es importante mantener muy claro el orden en que las variables forman el conjunto X_B y X_D de modo de garantizar consistencia.

Las matrices de la nueva tabla son:

$$B^{-1}D = \begin{pmatrix} -13/4 & 7/4 & -1/4 \\ -3/4 & 9/4 & 1/4 \\ 1/4 & 1/4 & 1/4 \end{pmatrix},$$

$$B^{-1}b = \begin{pmatrix} 3/4 \\ 21/4 \\ 5/4 \end{pmatrix}$$

$$\Rightarrow r_D = (c_D^T - c_B^T B^{-1}D)^T = \begin{pmatrix} -1/4 \\ -5/4 \\ 3/4 \end{pmatrix}$$

Luego

$$c_B^T B^{-1}b = \frac{-15}{4}$$

Esto es equivalente a la siguiente tabla:

X_1	X_2	X_3	X_4	X_5	X_6		
-13/4	7/4	0	1	0	-1/4	3/4	X_4
-3/4	9/4	0	0	1	1/4	21/4	X_5
1/4	1/4	1	0	0	1/4	5/4	X_3
-1/4	-5/4	0	0	0	3/4	15/4	

Observación 6.7 Esta tabla es la misma si se hubiera realizado operaciones fila en el casillero de la tercera fila y tercera columna.

Pero, ¿Es esta una solución óptima? **No**

El $\min\{\frac{-1}{4}, \frac{-5}{4}\} = \frac{-5}{4} \Rightarrow X_2$ entra a la base

El $\min\{\frac{3}{7}, \frac{21}{9}, 5\} = \frac{3}{7} \Rightarrow X_4$ sale de la base

Pivoteando obtenemos

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆		
-13/7	1	0	4/7	0	-1/7	3/7	X ₂
24/7	0	0	-9/7	1	4/7	30/7	X ₅
5/7	0	1	-1/7	0	2/7	8/7	X ₃
-18/7	0	0	5/7	0	4/7	30/7	

observamos que las (V.B) y (V.N.B) son $X_B = X_5, X_3, X_2$ y $X_D = X_1, X_6, X_4$

Ahora ¿Estamos en la solución óptima? **Aún No**

$\frac{-18}{7} \Rightarrow X_1$ entra a la base

El $\min\{\frac{5}{4}, \frac{30}{24}, \frac{8}{5}\} = \frac{5}{4} \Rightarrow X_5$ sale de la base

Al lector se le dejará que formule las nuevas tablas en forma matricial y llegue a la solución óptima del problema.

Para llegar a la solución debemos pivotar nuevamente:

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆		
0	1	0	-1/8	-13/24	1/6	11/4	X ₂
1	0	0	-3/8	7/24	1/6	5/4	X ₁
0	0	1	1/8	-5/24	1/6	1/4	X ₃
0	0	0	-1/4	3/4	1	15/2	

~

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆		
0	1	1	0	1/3	1/3	3	X ₂
1	0	3	0	-1/3	2/3	2	X ₁
0	0	8	1	-5/3	4/3	2	X ₄
0	0	2	0	1/3	4/3	8	

Si observamos, todos los costos reducidos en la última fila de la tabla son mayores o iguales a cero, nos dice que hemos encontrado nuestra solución óptima, cuales son:

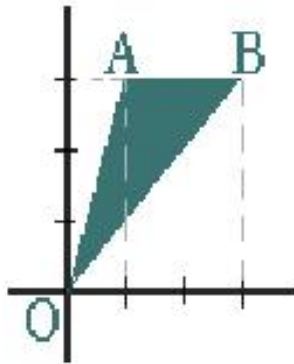
$$\begin{aligned} X_1 &= 2 \\ X_2 &= 3 \\ X_3 &= 0 \\ X_4 &= 2 \\ X_5 &= 0 \\ X_6 &= 0 \end{aligned}$$

Es decir el valor óptimo es $\hat{v} = -8$.

El lector puede deducir que en la penúltima tabla $\hat{v} = \frac{-15}{4}$ y la última tabla $\hat{v} = -8$, son soluciones posibles, pero $\hat{v} = -8$, es el valor óptimo más factible.

6.11. Ejercicios Propuestos

1. Representar el conjunto de puntos que satisfacen simultáneamente las inecuaciones: $x \leq 2$; $x \geq -2$; $y \leq 1$.
2. Dada la región del plano definida por las inecuaciones: $x + y - 1 \geq 0$; $0 \leq x \leq 3$; $0 \leq y \leq 2$. ¿Para qué valores de la región es máxima la función $Z = 5x + 2y$?
3. Se considera el recinto plano de la figura en el que están incluidos los tres lados y los tres vértices de las rectas asociadas a las siguientes desigualdades:



- a) Hallar las inecuaciones que definen el recinto.
 - b) Maximizar la función $Z = 3x - 6y$ sujeta a las restricciones del recinto.
4. Resolver gráficamente y algebraicamente el siguiente problema de programación lineal:

$$\begin{aligned}
 \text{Maximizar } Z &= 0,75x + y \\
 \text{s.a. } x + 3y &\leq 15 \\
 5x + y &\leq 20 \\
 3x + 4y &\leq 24 \\
 x &\geq 0 \\
 y &\geq 0
 \end{aligned}$$

5. Desde dos almacenes A y B, se tiene que distribuir fruta a tres mercados de la ciudad. El almacén A dispone de 10 toneladas de fruta diarias y el B de 15 toneladas, que se reparten en su totalidad. Los dos primeros mercados necesitan, diariamente, 8 toneladas de fruta, mientras que el tercero necesita 9 toneladas diarias. El coste del transporte desde cada almacén a cada mercado viene dado por el siguiente cuadro:

Almacén	Mercado 1	Mercado 2	Mercado 3
A	10	15	20
B	15	10	10

Planificar el transporte para que el coste sea mínimo.

6. Los siguientes ejercicios resuélvalos utilizando la forma matricial del método símplex y compruebelos con las tablas de iteración, hallando la solución óptima.

a)

$$\begin{aligned} \text{Maximizar } Z &= 10X_1 + 14X_2 \\ \text{s.a. } 4X_1 + 6X_2 &\leq 24 \\ 2X_1 + 6X_2 &\leq 20 \\ X_1 &\geq 0 \\ X_2 &\geq 0 \end{aligned}$$

b)

$$\begin{aligned} \text{Maximizar } Z &= 3X_1 - 9X_2 - 5X_3 + 4X_4 \\ \text{s.a. } X_1 + 4X_2 + 5X_3 + 8X_4 &\leq 8 \\ X_1 + 2X_2 + 6X_3 + 4X_4 &\leq 4 \\ X_1 &\geq 0 \\ X_2 &\geq 0 \\ X_3 &\geq 0 \\ X_4 &\geq 0 \end{aligned}$$

Conclusiones

Durante el desarrollo de las distintas etapas de esta investigación llegamos a la conclusión que:

- Buscamos de alguna manera poder entregar un aporte a la Carrera de Pedagogía en Educación Matemática, como una forma de agradecer a la Universidad del Bío-Bío, por los saberes que aquí hemos encontrado.
- Ampliamos nuestros conocimientos más allá del ámbito teórico conocido mientras realizamos el curso Álgebra Lineal en nuestra carrera, más aún, profundizamos en aplicaciones pertenecientes a áreas tan diversas como son: la Arqueología, Sociología, Física, Criptografía, Predicción del tiempo, entre otros. Lo que nos permitió indagar y darnos cuenta que los tópicos específicamente del Álgebra Lineal se pueden aplicar a otros campos y no sólo a la Matemática como Ciencia.
- También podemos sugerir el seminario como texto guía para docentes que están dictando el curso de Álgebra Lineal o a aquellos docentes de Enseñanza Media que incluyen algunos tópicos de esta en sus contenidos, por ejemplo si desean introducir los Sistemas de Ecuaciones, Matrices y sus Propiedades, Espacios Vectoriales, Espacio con Producto Interno, Valores y Vectores Propios y Programación Lineal como una motivación interesante, llamativa y didáctica en el transcurso del curso.
- Además el seminario está abierto a profundizar en cada uno de los tópicos aquí tratados, así también, en las aplicaciones al poder desarrollar los ejercicios propuestos que están de complemento en el texto o elaborar nuevos ejemplos.

BIBLIOGRAFÍA

- [1] Williams Gareth. *Álgebra Lineal con Aplicaciones*, Cuarta Edición, Mc Graw-Hill.1996
- [2] Miguel Hormazabal Díaz-Muñoz, Oscar Bravo Lutz y Luz María Gazzola Torrealba. *Manual de Ciencias Físicas* Ed Pontificia Universidad Católica, año 2004-2006.
- [3] Grossman Stanley. *Álgebra Lineal* Ed Mc Graw-Hill, Tercera Edición, año 1996.
- [4] Francisco Javier Cobos Cavala. *Apunte de Álgebra Lineal* Universidad de Sevilla, departamento de matemática aplicada, año 2000.
- [5] John H. Mathews y Kurtis D.Fink. *Métodos Numéricos utilizando MATLAB* Ed Pearson Education, Tercera Edición, año 1999.
- [6] Gilbert Strang. *Álgebra Lineal y sus Aplicaciones*, Fondo Educativo Interamericano S.A, año 1982.
- [7] George Nakos; David Joyner. *Álgebra Lineal con Aplicaciones* U.S naval academy; International Thomson Editores, Primera Edición, año 1998.
- [8] Ben Noble; James W. daniel. *Álgebra Lineal Aplicada* University of Winsconsin and Texas at Austin P.H.H prentice Hall, Tercera Edición, año 1997.
- [9] Serway, Raymond A.,James Madison University; Beichner, Robert J., James Madison University. *Física Volumen I y II* Ed McGraw-Hill, Quinta Edición, año 2001
- [10] Humberto Eco. *Como se hace una Tesis*, Técnicas y procedimientos de estudio, investigación y escritura. Departamento de matemática aplicada, año 2000.
- [11] Carlos Sabino. *Como hacer una Tesis*, Segunda Edición, año 1993.
- [12] Web-Seite: <http://www.wikipedia.org>
- [13] Web-Seite: <http://www.omerique.net/calcumat/estocasticas0.htm>
- [14] Web-Seite: <http://docencia.udea.edu.co/ingenieria/plineal/simplexmatricial1.htm>
- [15] Web-Seite: <http://www.investigacion-operaciones.com/Curso>