

FUNDAMENTOS MATEMÁTICOS DE DISEÑO EXPERIMENTAL PARA INGENIERÍA PARTE II

**© MOISÉS ARREGUÍN SÁMANO
ANDREA DAMARIS HERNÁNDEZ ALLAUCA
GUICELA MARGOTH ATI CUTIUPALA
ÁNGEL LEYVA-OVALLE**



**FUNDAMENTOS MATEMÁTICOS DE
DISEÑO EXPERIMENTAL PARA
INGENIERÍA**

Parte II

Moisés Arreguín Sámano

Andrea Damaris Hernández Allauca

Guicela Margoth Ati Cutiupala

Ángel Leyva- Ovalle

FUNDAMENTOS MATEMÁTICOS DE DISEÑO EXPERIMENTAL PARA INGENIERÍA

Parte II

© Autores

Moisés Arreguín- Sámano, Docente de la Universidad Estatal de Bolívar. Bolívar, Ecuador.

Andrea Damaris Hernández- Allauca, Docente Investigador Escuela Superior Politécnica de Chimborazo- Grupo de Investigación en Turismo (GITUR). Riobamba, Ecuador.

Guicela Margoth Ati- Cutiupala, Máster en Estadística aplicada por la Universidad de Granada. Doctora (c) en Estadística Multivariante Aplicada. Docente Investigador Escuela Superior Politécnica de Chimborazo- Grupo de Investigación en Ambiente y Desarrollo (GIADE). Riobamba, Ecuador. Docente en la Escuela Superior Politécnica de Chimborazo.

Ángel Leyva- Ovalle, Docente Investigador Universidad Autónoma Chapingo (UACH-DiCiFo), México. Ingeniero forestal con orientación en economía y ordenación por la Universidad Autónoma Chapingo (UACH), con experiencia en la División de Ciencias Forestales (DiCiFo), específicamente en el Departamento de productos forestales.

Casa Editora del Polo - CASEDELPO CIA. LTDA.

Departamento de Edición

Editado y distribuido por:

Editorial: Casa Editora del Polo

Sello Editorial: 978-9942-816

Manta, Manabí, Ecuador. 2019

Teléfono: (05) 6051775 / 0991871420

Web: www.casedelpo.com

ISBN: 978-9942-621-90-0

DOI: <https://doi.org/10.23857/978-9942-621-90-0>

© Primera edición

© Septiembre - 2024

Impreso en Ecuador

Revisión, Ortografía y Redacción:

Lic. Jessica Mero Vélez

Diseño de Portada:

Michael Josué Suárez-Espinar

Diagramación:

Ing. Edwin Alejandro Delgado-Veliz

Director Editorial:

Lic. Henry Darío Suárez Vélez

Todos los libros publicados por la Casa Editora del Polo, son sometidos previamente a un proceso de evaluación realizado por árbitros calificados.

Este es un libro digital y físico, destinado únicamente al uso personal y colectivo en trabajos académicos de investigación, docencia y difusión del Conocimiento, donde se debe brindar crédito de manera adecuada a los autores.

© **Reservados todos los derechos.** Queda estrictamente prohibida, sin la autorización expresa de los autores, bajo las sanciones establecidas en las leyes, la reproducción parcial o total de este contenido, por cualquier medio o procedimiento. parcial o total de este contenido, por cualquier medio o procedimiento.

Comité Científico Académico

Dr. Lucio Noriero-Escalante
Universidad Autónoma de Chapingo, México

Dra. Yorkanda Masó-Dominico
Instituto Tecnológico de la Construcción, México

Dr. Juan Pedro Machado-Castillo
Universidad de Granma, Bayamo. M.N. Cuba

Dra. Fanny Miriam Sanabria-Boudri
Universidad Nacional Enrique Guzmán y Valle, Perú

Dra. Jennifer Quintero-Medina
Universidad Privada Dr. Rafael Bellosó Chacín, Venezuela

Dr. Félix Colina-Ysea
Universidad SISE. Lima, Perú

Dr. Reinaldo Velasco
Universidad Bolivariana de Venezuela, Venezuela

Dra. Lenys Piña-Ferrer
Universidad Rafael Bellosó Chacín, Maracaibo, Venezuela

Dr. José Javier Nuvaéz-Castillo
Universidad Cooperativa de Colombia, Santa Marta,
Colombia

Constancia de Arbitraje

La Casa Editora del Polo, hace constar que este libro proviene de una investigación realizada por los autores, siendo sometido a un arbitraje bajo el sistema de doble ciego (peer review), de contenido y forma por jurados especialistas. Además, se realizó una revisión del enfoque, paradigma y método investigativo; desde la matriz epistémica asumida por los autores, aplicándose las normas APA, Sexta Edición, proceso de anti plagio en línea Plagiarisma, garantizándose así la cientificidad de la obra.

Comité Editorial

Abg. Néstor D. Suárez-Montes
Casa Editora del Polo (CASEDELPO)

Dra. Juana Cecilia-Ojeda
Universidad del Zulia, Maracaibo, Venezuela

Dra. Maritza Berenguer-Gouarnaluses
Universidad Santiago de Cuba, Santiago de Cuba, Cuba

Dr. Víctor Reinaldo Jama-Zambrano
Universidad Laica Eloy Alfaro de Manabí, Ext. Chone

PRÓLOGO

En el dinámico campo de la ingeniería, diseñar, optimizar y mejorar procesos, productos y sistemas es esencial para el avance tecnológico y económico. El diseño experimental es una herramienta crucial para enfrentar estos desafíos con un enfoque meticuloso y sistemático, fundamentado en principios matemáticos y conceptuales sólidos.

El libro "Fundamentos Matemáticos de Diseño Experimental para Ingeniería" busca llenar un vacío en la literatura técnica, integrando principios matemáticos y econométricos aplicados al diseño experimental. Esta obra ofrece una base teórica sólida y aplicaciones prácticas que conectan con los desafíos cotidianos de los ingenieros.

Los lectores encontrarán una rica intersección entre teoría y práctica, con ejemplos concretos que demuestran cómo los principios matemáticos y econométricos pueden resolver problemas reales en ingeniería.

Esperamos que este libro no solo brinde conocimientos técnicos, sino que también inspire una mentalidad analítica y crítica, equipando a los ingenieros con las habilidades necesarias para enfrentar y superar los desafíos futuros con confianza y eficacia.

Agradecemos a los lectores por embarcarse en este viaje de conocimiento y descubrimiento, y esperamos que esta obra se convierta en un recurso valioso y una fuente de inspiración en su desarrollo profesional y académico.

Los autores

ÍNDICE GENERAL

PRÓLOGO	7
ÍNDICE GENERAL	8
INTRODUCCIÓN	10
CAPÍTULO I	12
ECONOMETRÍA	12
1.1. Naturaleza y campo de econometría.....	12
1.2. Concepto de modelo en econometría	14
1.3. Diferencia entre un modelo econométrico y un modelo determinístico	17
1.4. Proceso de construcción y estimación de modelos	22
1.5. Tipos de relaciones económicas	22
1.6. Tipos de variables que intervienen en un modelo econométrico.	24
1.7. Tipos de modelos econométricos.....	25
1.7.1. Modelos uniecuacionales	25
1.7.2. Modelos de ecuaciones simultáneas.....	26
CAPÍTULO II	40
ANÁLISIS ECONOMÉTRICOS	40
2.1. El gradiente	40
2.2. La interpretación geométrica de campo gradiente	40
2.3. La derivada direccional.....	45
2.4. El Vector Proyección-Componente.....	49
2.5. La dirección de máximo y mínimo	49
2.6. Ejercicios Resueltos	56
CAPÍTULO III	59
MÁXIMOS Y MÍNIMOS LOCALES.....	59
3.1. Puntos críticos y extremos locales.....	60
3.2. Extremos con Restricciones.....	69
3.3. Teorema de Multiplicadores de Lagrange	71
3.3.1. Método de Multiplicadores de Lagrange.....	71
3.3.2. Criterios de Clasificación.....	72
3.4. Criterio para Máximos y Mínimos-Taylor	87

3.5.	Integral Doble parte de R	94
CAPÍTULO IV		98
CÁLCULO DE INTEGRALES DOBLES		98
4.1.	Área-Volumen	102
4.2.	El Teorema de Cambio de Variable.....	108
4.3.	El Jacobiano	108
4.4.	Superficies Regulares.....	115
4.4.1.	Área de Superficie	117
4.4.2.	Integral sobre una Superficie.....	122
4.4.3.	Integral de Superficie con Coordenadas Rectangulares	125
4.5.	Campos Escalares y Campos Vectoriales	132
4.6.	Integral de Flujo.....	136
4.7.	Modelos autorregresivos y de rezagos distribuidos.	142
4.8.	Modelos de series de tiempo	145
4.8.1.	Especificación de modelos econométricos consiste en definir	152
CAPÍTULO V		156
MODELOS DE REGRESIÓN		156
5.1.	Mínimos Cuadrados Ordinarios (MCO).....	156
5.1.1.	Teorema de Gauss-Markov.....	163
5.1.2.	Coeficiente de determinación	174
5.1.3.	Coeficiente de correlación r ($\sqrt{r^2}$)	180
5.1.4.	Precisión o errores estándar de estimaciones de mínimos cuadrados ordinarios.....	183
5.1.5.	Propiedades de estimadores de mínimos cuadrados ordinarios según supuestos de normalidad.....	187
5.2.	Aplicaciones con modelos matemáticos-econométricos.....	190
5.2.1.	Modelo matemático-económico lineal	190
5.2.2.	Modelo matemático-económico cuadrático.....	197
5.2.3.	Modelo matemático-económico cúbico	202
BIBLIOGRAFÍA.....		211

INTRODUCCIÓN

En el vasto campo de la ingeniería, el diseño experimental se erige como una herramienta fundamental para el desarrollo y la optimización de procesos, productos y sistemas. Sin embargo, su eficacia y aplicabilidad dependen en gran medida de un sólido fundamento matemático y conceptual que permita comprender, analizar y aplicar los principios subyacentes de manera precisa y efectiva.

Este libro, "Fundamentos Matemáticos de Diseño Experimental para Ingeniería", se adentra en los cimientos teóricos y prácticos que sustentan la disciplina del diseño experimental, enfocándose particularmente en su intersección con la econometría, una rama de las ciencias económicas que emplea métodos matemáticos y estadísticos para analizar y predecir fenómenos económicos.

El contenido de este libro abarca una variedad de temas esenciales para comprender y aplicar el diseño experimental en contextos ingenieriles, desde conceptos básicos de econometría hasta técnicas avanzadas de análisis y modelado. A través de una estructura organizada en capítulos, los lectores serán guiados desde los fundamentos conceptuales hasta la aplicación práctica de modelos matemáticos-econométricos en el contexto ingenieril.

El Capítulo I introduce los conceptos esenciales de econometría, explorando la naturaleza de los modelos, el proceso de construcción y estimación, así como la clasificación de modelos y relaciones económicas.

En el Capítulo II se profundiza en el análisis econométrico, abordando conceptos clave como el gradiente, la interpretación geométrica, la derivada direccional y la dirección de máximo y mínimo, acompañados de ejercicios resueltos que consolidan la comprensión de estos temas.

El Capítulo III se adentra en los máximos y mínimos locales, desde la identificación de puntos críticos hasta la aplicación del teorema de multiplicadores de Lagrange, proporcionando herramientas fundamentales para la optimización de sistemas ingenieriles.

El Capítulo IV se centra en el cálculo de integrales dobles, explorando áreas como el teorema de cambio de variable, el jacobiano y aplicaciones específicas en campos escalares y vectoriales, todo ello con relevancia directa en la modelización y análisis de fenómenos ingenieriles.

El Capítulo V cierra el texto con un enfoque en modelos de regresión, presentando técnicas como mínimos cuadrados ordinarios y su aplicación en modelos matemáticos-económicos lineales, cuadráticos y cúbicos, destacando su utilidad en la predicción y optimización de sistemas ingenieriles.

A lo largo de este libro, se entrelazan los principios teóricos con ejemplos prácticos y aplicaciones ingenieriles, con el objetivo de proporcionar a los lectores una comprensión profunda y una base sólida para abordar desafíos reales en el ámbito del diseño experimental en ingeniería.

Con esta obra, se busca no solo brindar conocimientos técnicos, sino también cultivar una mentalidad analítica y crítica que permita a los ingenieros abordar problemas complejos con confianza y eficacia.

CAPÍTULO I

ECONOMETRÍA

1.1. Naturaleza y campo de econometría

La experiencia acumulada en el ejercicio de la docencia y aplicación de técnicas econométricas a problemas concretos, parecen indicar que este capítulo puede ser muy importante para los estudiantes que se inician en econometría, al menos desde una perspectiva didáctica, puesto que aquí se presentan las bases conceptuales de los métodos econométricos.

Etimológicamente se trata de una palabra compuesta por el vocablo griego *econo*, significa economía y *metria*, significa medir. Es decir, su significado literal es medir la economía, proporcionar las técnicas y métodos para medir los fenómenos económicos. Se trata más bien de un conjunto de métodos de estimación en el sentido amplio.

Este conjunto de técnicas tiene en común que están basadas en modelos que se rigen por las leyes de las probabilidades. En otras palabras, la econometría es un conjunto de técnicas para la medición probabilística de los fenómenos económicos. No obstante, la medición no lo es todo, su interpretación está estrechamente ligada al análisis económico.

¿Cómo se pueden definir los límites teóricos de esta actividad científica?

Algunas opiniones al respecto son de Martínez (1974) que afirma «la econometría es la rama del conocimiento humano que se encuentra entre los límites de la economía y la estadística», Theil (1978) sostiene que «la econometría es la aplicación de métodos estadísticos y datos

estadísticos a los problemas en economía, administración de empresas, etc.»; por su parte, Johnston (1963) dice «el rol esencial de la econometría es la estimación y prueba de modelos económicos», Samuelson, Koopmans y Stone (1954) consideran que es «...el análisis cuantitativo de fenómenos económicos del mundo real a partir del desarrollo concurrente de la teoría y la observación, relacionados mediante métodos apropiados de inferencia...».

Sin embargo, el objetivo fundamental de la econometría consiste entonces en proporcionar los elementos de análisis cuantitativo para explicar el comportamiento de variables económicas observadas y pronosticar aquellas no observadas.

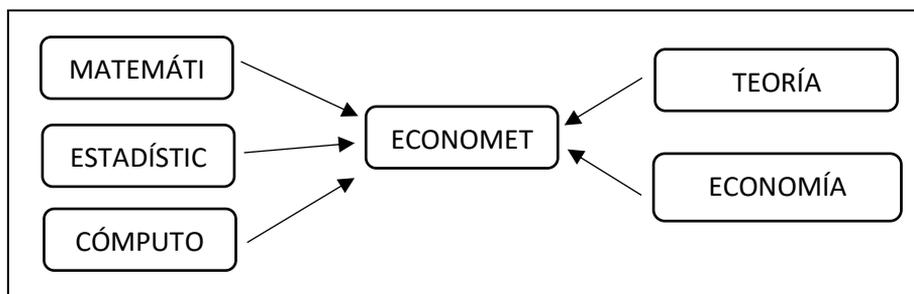
De esta manera, la econometría se convierte en un instrumento eficaz para:

- Estimar relaciones económicas de interés
- Evaluar políticas (análisis de intervención)
- Hacer proyecciones

Dos son las disciplinas científicas que fundamentan el trabajo econométrico: la estadística y la economía. Esto es sin desconocer el importante papel que juega el cómputo para su desarrollo.

Gráficamente se representa como:

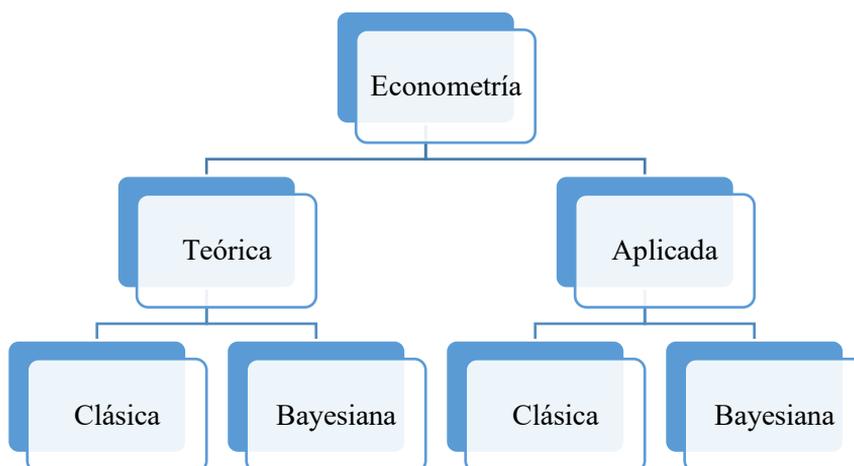
Figura 1.1 Disciplinas científicas que fundamentan el trabajo econométrico



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

La econometría se divide en las categorías siguientes:

Figura 1.2 Categorías de la econometría



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

1.2. Concepto de modelo en econometría

La palabra «modelo» es quizás uno de los vocablos más empleados en econometría. Es un concepto estructurador dentro de esta disciplina y por ello merece una atención especial. Para Malinvaud (1963) un modelo consiste en la representación formal de ideas o de conocimientos relativos a un fenómeno. Christ (1973) sostiene que un modelo es un conjunto de proposiciones a priori. Para Johnston (1963) todos los modelos económicos tienen ciertas características en común:

- a. El comportamiento de las variables económicas está determinado por la acción conjunta y simultánea de ciertas relaciones económicas.
- b. El modelo es una simplificación de complejo real, capta lo fundamental del sector económico en estudio.

El concepto de modelo, como representación ideal de la realidad, se utiliza también en el análisis de muchas disciplinas. Por ejemplo, existen modelos físicos que intentan representar fenómenos como el de la dilatación de los cuerpos al aumentar la temperatura.

Al respecto, la física proporciona una teoría que explica por qué los cuerpos se dilatan al aumentar la temperatura, pero también existe un modelo matemático que lo mide y establece que la longitud final de un cuerpo L_f es igual a la longitud inicial L_i más un pequeño crecimiento por el incremento de la temperatura Δt de acuerdo con el coeficiente de dilatación específico λ :

$$L_f = L_i(1 + \lambda\Delta t)$$

Nótese en este ejemplo la diferencia entre la explicación teórica causal que vincula la dilatación de los cuerpos con el cambio en la temperatura y su explicación molecular, con el modelo matemático que lo mide. Un modelo económico utilizado frecuentemente es la demanda de un bien o satisfactor. La teoría económica establece que la demanda de un bien está en función inversa de su precio, por una implicancia directa que sostiene la teoría económica respecto al comportamiento racional del consumidor. Este fenómeno económico se puede representar por una función lineal de la siguiente forma:

$$Q = \alpha - \beta P$$

Donde,

Q: Cantidad de demanda

α : Coeficiente intercepto

βP : Precio del bien

El signo negativo de α representa el supuesto de racionalidad del consumidor, pues si aumenta el precio de un bien, disminuye la cantidad demandada.

Los modelos puestos como ejemplo se han definido como modelos determinísticos porque las variables guardan una relación exacta. Es decir, conociendo una de ellas se puede calcular exactamente el valor de la otra. Sin embargo, en la representación real del fenómeno, el modelo nunca es capaz de captar todos los factores que inciden en su ocurrencia, solo incluye aquellos aspectos de interés parcial.

Por ello, se dice que el modelo representa a la realidad con error. La fuente de este error radica precisamente en i) la omisión de otras variables o factores relevantes o no, que intervienen en el fenómeno y, naturalmente, en ii) el error de medida.

Con base en esto, es necesario incluir un término que indique el error cometido por el modelo al representar la realidad. En econometría este término se denomina error aleatorio o perturbación y generalmente se representa por alguna letra griega ε . Por consiguiente, el modelo de demanda mencionado se convierte en un modelo econométrico al agregarle el término error:

$$Q = \alpha - \beta P + \varepsilon$$

Todo modelo econométrico debe tener un error aleatorio; caso contrario, no es un modelo econométrico. Los modelos de programación lineal o, en general la economía matemática, no constituyen modelos econométricos.

1.3. Diferencia entre un modelo econométrico y un modelo determinístico

Para ilustrar esta diferencia supóngase la siguiente ecuación de demanda como modelo determinístico:

$$Q_t (\text{Tiempo:Ano,mes,decada,etcetera}) = 10 - 0.5P_t (\text{Tiempo:Ano,mes,decada,etcetera})$$

Con esta ecuación, es posible entonces construir una tabla dando valores a P y calculando el correspondiente de Q_t :

Tabla 1.1. Cálculo de la Cantidad de demanda

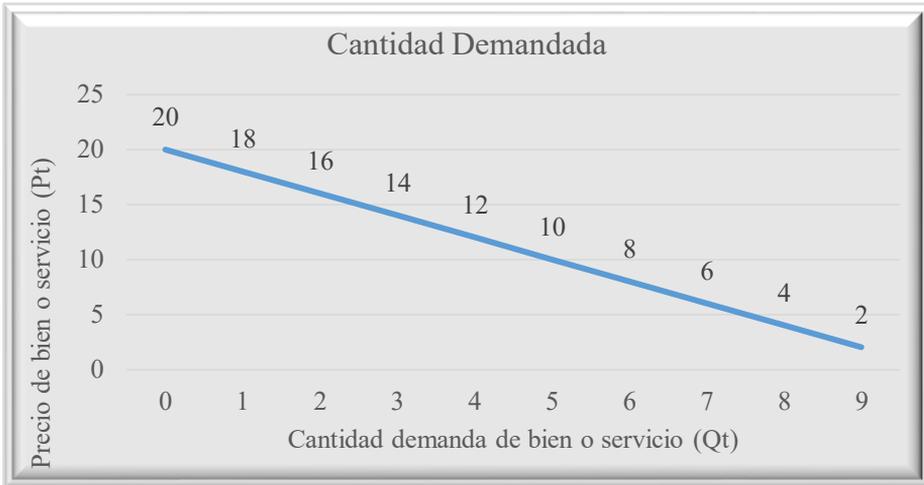
Q_t	P_t
0	20
1	18
2	16
3	14
4	12
5	10
6	8
7	6
8	4
9	2

La tabla 1.1 muestra que cuando el precio del bien es 20 su demanda por parte del público es cero, cuando el precio baja a 18 se demanda 1 unidad, cuando baja a 16 se demandan 2 unidades, etc.

Claramente se ve que a medida que el precio disminuye aumenta la demanda del bien en cuestión, tal como se puede esperar por el

comportamiento racional de los consumidores. Esta relación determinística se puede apreciar en la Figura 1.3.

Figura 1.3 Cantidad de Demanda



Fuente: Adaptado (Gujarati, D. N. y Porter, D. C. 2010).

Figura 1.4 Cantidad de Demanda



Fuente: Adaptado (Gujarati, D. N. y Porter, D. C. 2010).

Por su parte, para construir un modelo estocástico o aleatorio de la demanda se debe incluir un término error aleatorio, para convertirlo en modelo econométrico.

De esta manera la Tabla 1.1 se vería así:

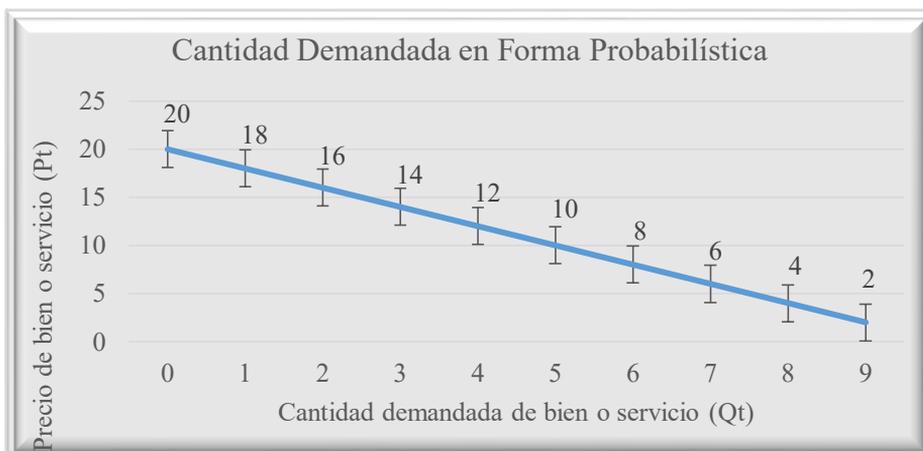
Tabla 1.2. Cantidad de demanda

P_t	Q_t	Prob
18	0	0.25
	1	0.50
	2	0.25
16	1	0.25
	2	0.50
	3	0.25
14	2	0.25
	3	0.50
	4	0.25
12	3	0.25
	4	0.50
	5	0.25
10	4	0.25
	5	0.50
	6	0.25
8	5	0.25
	6	0.50
	7	0.25
6	6	0.25
	7	0.50
	8	0.25
4	7	0.25
	8	0.50
	9	0.25
2	8	0.25
	9	0.50
	10	0.25

La tabla 1.2 indica que, para cada nivel de precio, la demanda se comporta aleatoriamente. Por ejemplo, cuando el precio del bien es de 16 pesos, se pueden comprar 1 bien con 0.25 de probabilidad, 2 bienes con 0.5 de probabilidad o 3 bienes con 0.25 de probabilidad. Nótese que la respuesta no es determinística, sino que fluctúa alrededor de una media.

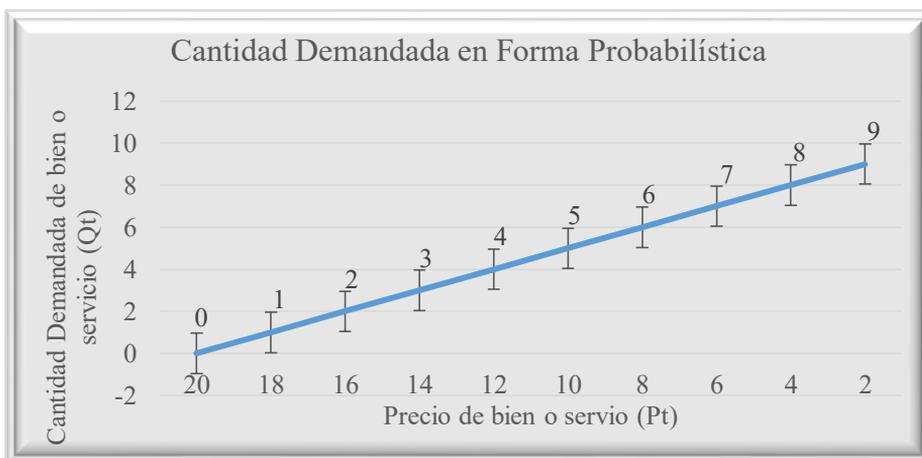
La representación gráfica de este modelo econométrico se puede ver a continuación:

Figura 1.5 Cantidad de Demandada en forma probabilística



Fuente: Adaptado (Gujarati, D. N. y Porter, D. C. 2010).

Figura 1.6 Cantidad de Demandada en forma probabilística



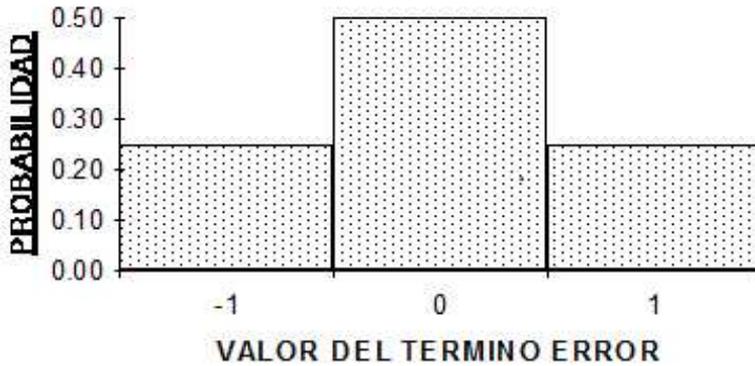
Fuente: Adaptado (Gujarati, D. N. y Porter, D. C. 2010).

En este ejemplo se supuso que para cada nivel de precio el error aleatorio ε puede tomar los valores -1, 0 y 1 con las probabilidades ya anotadas.

Por tanto, su función de distribución sería:

ε_t	$f(\varepsilon_t)$
-1	0.25
0	0.50
1	0.25

Figura 1.7 Distribución del término error



Fuente: Gujarati, D. N. y Porter, D. C. 2010.

Claro está que en este ejemplo se conoce el error aleatorio y su comportamiento, pero generalmente estos son aspectos desconocidos; por ello, se tienen que hacer supuestos en relación con el comportamiento del término error y que se verán más adelante.

Fácilmente se puede demostrar que la media del término error es igual a cero:

$$E(\varepsilon_t) = -1 * (0.25) + 0 * (0.25) + 1 * (0.25) = 0$$

Y, por definición, su varianza es:

$$\text{Var}(\varepsilon_t) = E[\varepsilon - E(\varepsilon)]^2 \text{ siendo } E(\varepsilon) = 0$$

$$\text{Var}(\varepsilon_t) = E(\varepsilon^2) = (-1^2 * 0.25) + (0^2 * 0.25) + (1^2 * 0.25) = 0.50$$

La varianza del término error es la misma para todos los niveles de precios, este es el concepto de «homocedasticidad» que ocupa un lugar destacado en el análisis econométrico.

También, es útil para mostrar que la variable Q tiene una media esperada que depende del valor de P y que su varianza es constante y es la misma que la del término error. Q y ε tienen la misma distribución, aunque ε se centra alrededor de 0 y Q alrededor del valor esperado condicionado a P .

1.4. Proceso de construcción y estimación de modelos

El proceso econométrico se puede definir como una secuencia de pasos en la construcción y estimación de modelos. Evidentemente, existe una interacción entre los hechos, concretizados en datos y la teoría que los vincula e interpreta. Este es el punto de partida en la construcción de modelos. Hay dos vertientes en la construcción que interactúan y que conducen a la especificación del modelo.

Estas vertientes convergen con la teoría estadística para especificar y estimar el modelo para su posterior utilización.

1.5. Tipos de relaciones económicas

En los modelos econométricos pueden intervenir diferentes modalidades de relaciones económicas. Son éstas las bases sobre las que se construyen los modelos y pueden ser:

- a) Relaciones de comportamiento. Típicamente representan el comportamiento de los agentes económicos; por ejemplo: demanda, oferta, ahorro, inversión, etcétera.

Ejemplo:

$$C_{(\text{Consumo})} = \beta_0 + \beta_1 Y_{(\text{Ingreso})}$$

- b) Relaciones técnicas. Se trata de relaciones que representan procesos productivos en que no interviene el comportamiento de los agentes económicos.

El modelo multiplicativo tipo Cobb-Douglas es un ejemplo clásico:

$$P_{(\text{Produccion total})} = cK_{(\text{Capital economico})}^a + L_{(\text{Mano de obra})}^b$$

- c) Definiciones, identidades o relaciones de equilibrio. Son aquellas relaciones en las cuales no hay coeficientes desconocidos a estimar. Generalmente no tienen término error.

Ejemplo:

- 1) Valor monetario =

$$P_{(\text{Precio del bien})} \times Q_{(\text{Cantidad consumida})}$$

- 2) Valor monetario = Capital total – deudas

- 3) Inversion neta = Cambios en stock de capital

- 4) Consumo aparente = producción interna + importaciones – exportaciones – reservas

- 5) Relacion de equilibrio: oferta = demanda

- 6) Definición institucional: IVA = 0.15 Ventas

Estas relaciones anotadas pueden ser estáticas o dinámicas, sea que se tome en cuenta o no el factor tiempo. También, pueden referirse a todo un sistema económico, en forma agregada y, por tanto, serán macro relaciones. Si se refieren a ámbitos más reducidos a nivel regional o local, industria o empresa, se llamarán micro relaciones.

1.6. Tipos de variables que intervienen en un modelo

econométrico

Al desarrollar un modelo econométrico surge una distinción fundamental en relación con las variables que intervienen en el modelo. Sabemos que el objetivo del modelo es representar un fenómeno económico particular, por ello sus variables deben estar estrictamente vinculadas a él, pero algunas de estas variables están simultáneamente determinadas al interior del sistema económico que se está representando y otras «vienen de afuera» por así decirlo.

A las variables determinadas internamente y sobre las cuales se está interesado en estudiar su comportamiento, se les denomina *variables dependientes, explicadas o endógenas*. En cambio, aquellas variables que están determinadas fuera del sistema en estudio se llaman *variables independientes, explicativas o exógenas*. Estas últimas pueden ser exógenas propiamente como tal o variables endógenas rezagadas en el tiempo (predeterminadas).

Ejemplo:

$$\begin{aligned} D_t(\text{Demanda en tiempo}) &= \alpha_0 + \alpha_1 P_t(\text{Precio en tiempo}) \\ &+ \alpha_2 I_t(\text{Ingreso en tiempo}) + \varepsilon_t(\text{Error en tiempo}) \end{aligned}$$

$$\begin{aligned} O_t(\text{Oferta en tiempo}) &= \beta_0 + \beta_1 P_t(\text{Precio en tiempo}) \\ &+ \beta_2 P_{t-1}(\text{Variable predefinida}) \\ &+ v_t(\text{Error en tiempo}) \end{aligned}$$

$$\begin{aligned} \text{Relacion de equilibrio: } O_t(\text{Oferta en tiempo}) \\ = D_t(\text{Demanda en tiempo}) \end{aligned}$$

La cantidad ofertada y demandada, así como el precio, están simultáneamente determinados y, por tanto, son variables endógenas. En cambio, el ingreso de las personas está determinado en un sistema económico más amplio y, en consecuencia, constituye una variable predeterminada, en este caso, exógena.

En los modelos uniecuacionales, las variables endógenas deben aparecer solo en el lado derecho de las ecuaciones. Todas estas son variables observadas, pero ε_t (Error en tiempo) y ν_t (Error en tiempo) son errores aleatorios, pero son variables no observadas en el modelo. Sin embargo, el análisis de residuos es una técnica matemática que permite estudiar a profundidad los errores.

1.7. Tipos de modelos econométricos

1.7.1. Modelos uniecuacionales

La estrategia más común en el análisis económico empírico consiste en la especificación de un modelo uniecuacional para luego estimar sus coeficientes utilizando la información de una muestra, sea serie de tiempo o proveniente de información cruzada. Una de las variables se define como dependiente y se supone que está determinada por una o por un conjunto de variables explicativas independientes. Por ejemplo, se puede establecer el siguiente modelo:

$$\begin{aligned} P_{(\text{Produccion total})} \\ = dK_{(\text{Capital economico})}^a + L_{(\text{Mano de obra})}^b \\ + T_{(\text{Tierra})}^c \end{aligned}$$

1.7.2. Modelos de ecuaciones simultáneas.

Un Modelo de Ecuaciones Simultáneas (M.E.S.) es un conjunto de ecuaciones de regresión donde existe influencia simultánea entre variables y ecuaciones.

De acuerdo con Zill y Wright (2011), el método de Integración de Fracciones Parciales implica que cuando se suman dos funciones racionales, como:

$$g(x) = \frac{2}{(x+5)} \text{ y } h(x) = \frac{1}{(x+1)},$$

Los términos se combinan por medio de un denominador común:

$$g(x) + h(x) = \frac{2}{(x+5)} + \frac{1}{(x+1)} = \frac{2}{x+5} \left(\frac{x+1}{x+1} \right) + \frac{2}{x+1} \left(\frac{x+5}{x+5} \right),$$

Tal que al sumar denominadores en el miembro derecho de esta ecuación se obtiene la función racional simple $f(x) = \frac{3x+7}{(x+5)(x+1)}$.

Suponga que es necesario integrar la función f . Por supuesto, la solución es evidentemente, pues se usa la igualdad de:

$$g(x) + h(x) = \frac{2}{(x+5)} + \frac{1}{(x+1)} = \frac{2}{x+5} \left(\frac{x+1}{x+1} \right) + \frac{2}{x+1} \left(\frac{x+5}{x+5} \right) \text{ y}$$

$$f(x) = \frac{3x+7}{(x+5)(x+1)} \text{ para escribir}$$

$$\int \frac{3x+7}{(x+5)(x+1)} dx = \int \left[\frac{2}{(x+5)} + \frac{1}{(x+1)} \right] dx = 2 \ln|x+5| + \ln|x+1| + C.$$

Entonces, este ejemplo ilustra un procedimiento para integrar ciertas funciones racionales $f(x) = \frac{p(x)}{q(x)}$.

Este método consiste en invertir el proceso de:

$$g(x) + h(x) = \frac{2}{(x+5)} + \frac{1}{(x+1)} = \frac{2}{x+5} \left(\frac{x+1}{x+1} \right) + \frac{2}{x+1} \left(\frac{x+5}{x+5} \right); \text{ es decir,}$$

se empieza con una función racional, como:

$$f(x) = \frac{3x+7}{(x+5)(x+1)} \text{ y se separa en fracciones componentes más simples}$$

$$g(x) = \frac{2}{(x+5)} \text{ y } h(x) = \frac{1}{(x+1)} \text{ denominadas fracciones parciales,}$$

definido como el proceso algebraico para separar una expresión racional como $f(x) = \frac{3x+7}{(x+5)(x+1)}$, en fracciones parciales más simples

llamada descomposición en fracciones parciales; tal que, por conveniencia, se supondrá que la función racional $f(x) = \frac{p(x)}{q(x)}$, $q(x) \neq$

0, es una fracción propia o una expresión racional propia; es decir, el grado de $p(x) < q(x)$ ni que $p(x)$ y $q(x)$ tienen factores comunes.

Nota: Una función racional $\frac{P(x)}{Q(x)}$ es una **fracción propia** si el grado del polinomio $P(x)$ es $< Q(x)$, caso contrario en que el grado del polinomio $P(x)$ es $\geq Q(x)$ la **fracción** se llama **impropia** que se puede expresar, efectuando la división, como la suma de un polinomio más una fracción propia: $\frac{P(x)}{Q(x)} = \{\text{Polinomio}\} + \frac{N_1(x)}{Q(x)}$.

A continuación, evalúa la integral término a término. Existen los siguientes casos de descomposición en fracciones parciales:

Factores lineales distintos $\left(\frac{A}{ax+b} \text{ o } \frac{C}{ax+b} \right)$.

Si el denominador $q(x)$ contiene un producto de n factores lineales distintos $(a_1x + b_1)(a_2x + b_2)(a_3x + b_3)(a_4x + b_4) \dots (a_nx + b_n)$ donde a_i y b_i , $i = 1, 2, 3, 4, \dots, n$ son $\mathbb{R} \Rightarrow$ encontrar constantes reales únicas $C_1, C_2, C_3, C_4, \dots, C_n$ tal que la descomposición en fracciones

parciales de $f(x) = \frac{p(x)}{q(x)}$ contiene la suma $\frac{C_1}{(a_1x+b_1)} + \frac{C_2}{(a_2x+b_2)} + \frac{C_3}{(a_3x+b_3)} + \frac{C_4}{(a_4x+b_4)} + \dots + \frac{C_n}{(a_nx+b_n)}$ o $\frac{A_1}{(a_1x+b_1)} + \frac{A_2}{(a_2x+b_2)} + \frac{A_3}{(a_3x+b_3)} + \frac{A_4}{(a_4x+b_4)} + \dots + \frac{A_n}{(a_nx+b_n)}$; es decir, la descomposición en fracciones parciales que se ha puesto para f contiene una fracción parcial para cada uno de los factores lineales $a_ix + b_i$.

Ejemplos:

a) Evalué $\int \frac{2x+1}{(x-1)(x+3)} dx \approx \frac{A}{(x-1)} + \frac{B}{(x+3)} \Rightarrow 2x + 1 = A(x + 3) +$

$B(x - 1) \Rightarrow 2x + 1 = Ax + 3A + Bx - B \Rightarrow$

$$\begin{array}{c} \xrightarrow{\text{iguales}} \\ 2x + 1x^0 = (A + B)x + (3A - B)x^0 \\ \xleftarrow{\text{iguales}} \end{array} \quad \therefore$$

a) $2 = A + B$

Se obtienen las ecuaciones: b) $1 = 3A - B$. Sustituyendo
 $3 = 4A \Rightarrow \therefore A = \frac{3}{4}$

A en a) o b): $B = 2 - \left(\frac{3}{4}\right) = \frac{5}{4} \Rightarrow \frac{2x+1}{(x-1)(x+3)} = \frac{\left(\frac{3}{4}\right)}{(x-1)} + \frac{\left(\frac{5}{4}\right)}{(x+3)}$.

Por lo tanto, $\int \frac{2x+1}{(x-1)(x+3)} dx = \int \left[\frac{\left(\frac{3}{4}\right)}{(x-1)} + \frac{\left(\frac{5}{4}\right)}{(x+3)} \right] = \frac{3}{4} \text{Ln}|x - 1| + \frac{5}{4} \text{Ln}|x + 3| + C$.

b) Evalué $\int \frac{x^2+2-1}{2x^3+3x^2-2x} dx = \frac{x^2+2-1}{x(x+2)(2x-1)} \approx \frac{A}{x} + \frac{B}{x+2} + \frac{C}{2x-1} =$

$$\frac{A[(x+2)(2x-1)] + B[(x)(2x-1)] + C[(x)(x+2)]}{x(x+2)(2x-1)} =$$

$$\frac{2^a x^2 + 3Ax - 2^a + 2Bx^2 - Bx + Cx^2 + 2Cx}{x(x+2)(2x-1)} =$$

$$\frac{(2^a x^2 + 2Bx^2 + Cx^2) + (3Ax - Bx + 2Cx) - 2^a}{x(x+2)(2x-1)} = \frac{x^2(2^a + 2B + C) + x(3A - B + 2C) - 2^a}{x(x+2)(2x-1)}$$

$\Rightarrow 1x^2 + 2x - 1 \approx (2^a + 2B + C)x^2 + (3A - B + 2C)x - 2^a$

\therefore Se obtienen las ecuaciones:

a) $1 = 2^a + 2B + C$, b) $2 = 3^a - B + 2C$ y c) $-1 = -2A$

$\Rightarrow A = \frac{1}{2}$, sustituyendo A en a): $1 = 2\left(\frac{1}{2}\right) + 2B + C$

$\Rightarrow C = -2B$,

sustituyendo A y C en b): $2 = 3\left(\frac{1}{2}\right) - B + 2(-2B)$

$\Rightarrow B = -\frac{1}{10}$

$\Rightarrow \therefore C = -2B = -2\left(-\frac{1}{10}\right) = \frac{1}{5}$.

Por lo tanto, $\int \frac{x^2+2-1}{2x^3+3x^2-2x} = \frac{1}{2}\text{Ln}|x| - \frac{1}{10}\text{Ln}|x+2| + \frac{1}{5}\text{Ln}|2x-1| + C$.

c) Encuentre el área bajo la gráfica $f(x) = \frac{1}{x(x+1)}$ sobre intervalo

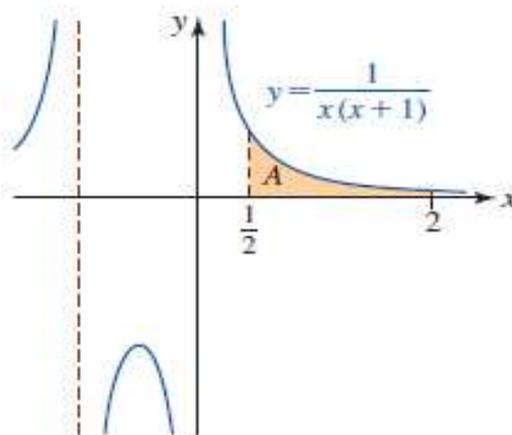
$\left[\frac{1}{2}, 2\right]$. El área en cuestión se muestra enseguida, puesto que

$f(x)$ es positiva para toda x en intervalo, el área es la integral definida $A = \int_{\frac{1}{2}}^2 \frac{1}{x(x+1)} dx$ tal que al usar fracciones parciales

$\frac{1}{x(x+1)} \approx \frac{A}{x} + \frac{B}{x+1} = \frac{A(x+1)+Bx}{x(x+1)} \Rightarrow \therefore 1 = A(x+1) + Bx \Rightarrow x =$

$0, x = -1, A = 1$ y $B = -1$:

Figura 1.8 Área bajo la curva



Fuente: (Mora, W. F. 2016).

$$A = \int_{\frac{1}{2}}^2 \left[\frac{1}{x(x+1)} \right] dx = \int_{\frac{1}{2}}^2 \left[\frac{1}{x} - \frac{1}{(x+1)} \right] dx = [\text{Ln}|x| - \text{Ln}|x+1|]_{\frac{1}{2}}^2 =$$

$$\left[\text{Ln} \left| \frac{x}{(x+1)} \right| \right]_{\frac{1}{2}}^2 = \text{Ln}(2) = 0.6931.$$

Factores lineales repetidos $\left(\frac{A}{(ax+b)^2} \text{ o } \frac{C}{(ax+b)^2} \right)$.

Si el denominador de la función racional $f(x) = \frac{p(x)}{q(x)}$ contiene un factor lineal repetido $(ax + b)^n$, $n > 1$, donde a y b son $\mathbb{R} \Rightarrow$ encontrar constantes reales únicas $C_1, C_2, C_3, C_4, \dots, C_n$ tales que la descomposición en fracciones parciales de f contiene la suma $\frac{C_1}{(ax+b)^1} + \frac{C_2}{(ax+b)^2} + \frac{C_3}{(ax+b)^3} + \frac{C_4}{(ax+b)^4} + \dots + \frac{C_n}{(ax+b)^n}$ o $\frac{A_1}{(ax+b)^1} + \frac{A_2}{(ax+b)^2} + \frac{A_3}{(ax+b)^3} + \frac{A_4}{(ax+b)^4} + \dots + \frac{A_n}{(ax+b)^n}$; es decir, la descomposición en fracciones parciales de f contiene una fracción parcial para cada potencia $ax + b_i$.

Ejemplos:

a) Evalué

$$\int \frac{x^2+2x+4}{(x+1)^3} dx \approx \frac{A}{(x+1)^1} + \frac{B}{(x+1)^2} + \frac{C}{(x+1)^3} =$$

$$\frac{A[(x+1)^2]+B[(x+1)^1]+C[1]}{(x+1)^3} = \frac{A(x^2+2x+1)+Bx+C}{(x+1)^3}$$

$$\Rightarrow x^2 + 2x + 4 = Ax^2 + 2Ax + A + Bx + B + C$$

$$= x^2(A) + x(2A + B) + (A + B + C)$$

$$\Rightarrow \therefore x^2(1) + x(2) + 4 = x^2(A) + x(2A + B) + (A + B + C)$$

$$\Rightarrow \therefore A = 1, \text{ mientras que } 2 = 2A + B \Rightarrow 2 = 2(1) + B$$

$$\Rightarrow B = 0 \text{ y, también, } 4 = A + B + C \Rightarrow 4 = (1) + (0) + C \Rightarrow$$

$$C = 3.$$

$$\text{Por lo tanto, } \int \frac{x^2+2x+4}{(x+1)^3} dx = \int \left[\frac{1}{(x+1)^1} + \frac{0}{(x+1)^2} + \frac{3}{(x+1)^3} \right] dx$$

$$\Rightarrow \int \frac{x^2+2x+4}{(x+1)^3} dx = \int \left[\frac{1}{(x+1)^1} + 3(x+1)^{-3} \right] dx = \ln|x+1| - \frac{2}{3}(x+1)^{-2} + C.$$

b) Evalué

$$\int \frac{4x}{x^3-x^2-x+1} dx \approx \frac{A}{(x-1)} + \frac{B}{(x-1)^2} + \frac{C}{(x+1)} =$$

$$\frac{A[(x-1)(x+1)]+B[(x+1)]+C[(x-1)^2]}{(x-1)^2(x+1)} =$$

$$\frac{A(x^2+x-x-1+1)+B(x+1)+C(x-1)^2}{(x-1)^2(x+1)} = \frac{Ax^2-A+Bx+B+Cx^2-2Cx+C}{(x-1)^2(x+1)} =$$

$$\frac{x^2(A+C)+x(B-2C)-A+B+C}{(x-1)(x-1)(x+1)}$$

$$\Rightarrow 0x^2 + 4x + 0 = (A + C)x^2 + (B - 2C)x + (-A + B + C)$$

⇒ Se obtienen las ecuaciones:

a) $0 = A + C$, b) $4 = B - 2C$ y c) $0 = -A + B + C \Rightarrow A =$

b) $-4 = -0A + B + 2C$

c) $0 = -A + B + C$

$-C$ o $C = -A \Rightarrow$

$-4 = -A + 3C$

a) $0 = A + C$

$\Rightarrow \therefore C = -\frac{4}{4} = -1$

Entonces, si $A = -C = -(-1) = 1$ y $4 = B - 2C \Rightarrow B =$

$4 - 2 = 2$.

Por lo tanto, $\int \frac{4x}{x^3-x^2-x+1} dx = \ln|x-1| - \frac{2}{(x-1)} -$

$1\ln|x+1| + C.$

Factores cuadráticos distintos “parábolas” $\left(\frac{Ax+B}{ax^2+bc+c} \text{ o } \frac{Cx+B}{ax^2+bc+c}\right)$.

Si el denominador de función racional $f(x) = \frac{p(x)}{q(x)}$ contiene un producto de n factores cuadráticos e irreducibles, significa que la expresión cuadrática no se factoriza sobre el conjunto de \mathbb{R} , $(a_1x^2 + b_1x + c_1)(a_2x^2 + b_2x + c_2)(a_3x^2 + b_3x + c_3)(a_4x^2 + b_4x + c_4) \dots (a_nx^2 + b_nx + c_n)$ donde a_i, b_i y $c_i, i = 1, 2, 3, 4, \dots, n$ son $\mathbb{R} \Rightarrow$ encontrar constantes reales únicas $A_1, A_2, A_3, A_4, \dots, A_n, B_1, B_2, B_3, B_4, \dots, B_n$ tal que la descomposición en fracciones parciales para f contiene la suma $\frac{C_1x+B_1}{a_1x^2+b_1x+c_1} + \frac{C_2x+B_2}{a_2x^2+b_2x+c_2} + \frac{C_3x+B_3}{a_3x^2+b_3x+c_3} + \frac{C_4x+B_4}{a_4x^2+b_4x+c_4} + \dots + \frac{C_nx+B_n}{a_nx^2+b_nx+c_n}$ o $\frac{A_1x+B_1}{a_1x^2+b_1x+c_1} + \frac{A_2x+B_2}{a_2x^2+b_2x+c_2} + \frac{A_3x+B_3}{a_3x^2+b_3x+c_3} + \frac{A_4x+B_4}{a_4x^2+b_4x+c_4} + \dots + \frac{A_nx+B_n}{a_nx^2+b_nx+c_n}$; análogamente, en caso que $q(x)$ contiene un producto de factores lineales distintos, la descomposición en fracciones parciales que se ha supuesto para f contiene una fracción parcial para cada uno de los factores cuadráticos $a_ix^2 + b_ix + c_i$.

Ejemplos:

a) Evalúe $\int \frac{x+3}{x^4+9x^2} dx = \int \frac{x+3}{x^2(x^2+9)} dx \approx \frac{A}{x^1} + \frac{B}{x^2} + \frac{Cx+D}{x^2+9} =$

$$\frac{Ax[x^2+9]+B[x^2+9]+x^2[Cx+D]}{x^2(x^2+9)} = \frac{(A+C)x^3+x^2(B+D)+9Ax+9B}{x^2(x^2+9)}$$

$$\Rightarrow x + 3 = Ax(x^2 + 9) + B(x^2 + 9) + (Cx + D)x^2$$

$$= (A + C)x^3 + (B + D)x^2 + 9Ax + 9B$$

$$\Rightarrow 1 = 9A$$

$$\Rightarrow \therefore A = \frac{1}{9}, B = \frac{1}{3} \text{ al hacer } x = 0 \text{ en ecuación } Ax(x^2 + 9) + B(x^2 + 9) + (Cx + D)x^2$$

Tal que con ecuación $(A + C)x^3 + (B + D)x^2 + 9Ax + 9B$

Se obtiene $0 = A + C, 0 = B + D$ y $1 = 9^a \Rightarrow C = -\frac{1}{9}$ y

$$D = -\frac{1}{3}$$

$$\Rightarrow \int \frac{x+3}{x^4+9x^2} dx = \int \left[\frac{\left(\frac{1}{9}\right)}{x^1} + \frac{\left(\frac{1}{3}\right)}{x^2} + \frac{\left(-\frac{1}{9}x - \frac{1}{3}\right)}{x^2+9} \right] dx \approx \int \left[\frac{\left(\frac{1}{9}\right)}{x^1} + \frac{\left(\frac{1}{3}\right)}{x^2} - \right.$$

$$\left. \frac{1}{18} \frac{2x}{x^2+9} - \frac{1}{3} \frac{1}{x^2+9} \right] dx = \frac{1}{9} \text{Ln}|x| - \frac{1}{3} x^{-1} - \frac{1}{18} \text{Ln}(x^2 + 9) -$$

$$\frac{1}{9} \text{Tan}^{-1} \left(\frac{x}{3} \right) + E = \frac{1}{3} \text{Ln} \left(\frac{x^2}{x^2+9} \right) - \frac{1}{3} x^{-1} - \frac{1}{9} \text{Tan}^{-1} \left(\frac{x}{3} \right) + E.$$

b) Evalué $\int \frac{2x^2-x+4}{x^3+4x} dx = \int \frac{2x^2-x+4}{x(x^2+4)} dx \approx \frac{A}{x^1} + \frac{Bx+C}{(x^2+4)} =$

$$\frac{A[x^2+4]+B[x^2+9]+x[Bx+C]}{x^2(x^2+4)} = \frac{(A+B)x^2+(Bx+C)x+4^a}{x^2(x^2+4)} \Rightarrow 2x^2 - 1x +$$

$$4 = (A + B)x^2 + (C)x + 4^a$$

$$\Rightarrow A = \frac{4}{4} = 1, B = 2 - 1 = 1 \text{ y } C = -1$$

$$\Rightarrow \int \frac{1}{x^1} dx + \int \frac{x-1}{x^2+4} dx + \int \frac{x}{x^2+4} dx - \int \frac{1}{x^2+4} dx.$$

Por cambio de variable, si

$$t = x^2 + 4 \Rightarrow dt = 2x dx \Rightarrow \therefore dx = \frac{dt}{2x} \Rightarrow \therefore \text{Ln}|x| +$$

$$\text{Ln}|x + 4| - \frac{1}{x} \text{Tan}^{-1} \left(\frac{2}{x} \right) + \text{Ln}|x + 4| - \frac{4}{x} \text{Tan}^{-1} \left(\frac{2}{x} \right) -$$

$$\frac{1}{x} \text{Tan}^{-1} \left(\frac{2}{x} \right) + C = \text{Ln}|x| + 2\text{Ln}|x + 4| - \frac{6}{x} \text{Tan}^{-1} \left(\frac{2}{x} \right) + C.$$

Factores cuadráticos repetidos $\left(\frac{Ax+B+Cx+D}{(x^2-5)^2} = \frac{Ax+B}{(x^2-5)^1} + \frac{Cx+D}{(x^2-5)^2} \right).$

Si el denominador de la función racional $f(x) = \frac{p(x)}{q(x)}$ contiene un

factor cuadrático irreducible repetido $(ax^2 + bx + c)^n, n > 1$, donde a,

b y c son $\mathbb{R} \Rightarrow$ encontrar constantes reales únicas $A_1, A_2, A_3, A_4, \dots, A_n,$

$B_1, B_2, B_3, B_4, \dots, B_n$ tal que la descomposición en fracciones parciales

para f contiene la suma $\frac{A_1x+B_1}{(ax^2+bx+c)^1} + \frac{A_2x+B_2}{(ax^2+bx+c)^2} + \frac{A_3x+B_3}{(ax^2+bx+c)^3} +$

$\frac{A_4x+B_4}{(ax^2+bx+c)^4} + \dots + \frac{A_nx+B_n}{(ax^2+bx+c)^n}$; es decir, la descomposición en

fracciones parciales que se ha supuesto para f contiene una fracción parcial para cada potencia de $ax^2 + bx + c$.

Ejemplo

$$\text{Evalué } \int \frac{x^2}{(x^2+4)^2} dx \approx \frac{Ax+B}{(x^2+4)^1} + \frac{Cx+D}{(x^2+4)^2} = \frac{Ax+B[(x^2+4)^1]+Cx+D[1]}{(x^2+4)^2} =$$

$$\frac{(A)x^3+(B)x^2+(4^a+C)x+(4B+D)}{(x^2+4)^2} \Rightarrow 0x^3 + 1x^2 + 0x + 0 = (A)x^3 + (B)x^2 + (4^a + C)x +$$

$$(4B + D)$$

$$\Rightarrow \therefore A = 0, B = 1, C = 0 \text{ y } D = -4$$

$$\Rightarrow \int \frac{1}{(x^2+4)^1} dx - 4 \int \frac{1}{(x^2+4)^2} dx =$$

$$\int \frac{1}{x^2+(2)^2} dx - 4 \int \frac{1}{(x^2)^2+(4)^2} dx =$$

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

$$\frac{2}{x} \text{Tan}^{-1} \left(\frac{2}{x} \right) + C.$$

Las Fracciones Impropias indican que en cada uno de los ejemplos precedentes el integrando $f(x) = \frac{p(x)}{q(x)}$ era una fracción propia.

Recuerde que cuando $f(x) = \frac{p(x)}{q(x)}$ es una fracción impropia; es decir, cuando el grado de $p(x)$ es $\geq q(x)$, se procede con una división larga.

Ejemplo 1

$\int \frac{x^3-2x}{x^2+3x+2} dx = x - 3 + \frac{5x+6}{x^2+3x+2}$, pues el numerador se divide entre denominador. Luego, el denominador se factoriza como $x^2 + 3x + 2 = (x + 1)(x + 2)$, el residuo se descompone en fracciones parciales:

$$\frac{5x+6}{x^2+3x+2} = \frac{1}{x+1} + \frac{4}{x+2}.$$

Con esta información, la evaluación de integral es inmediata:

$$\int \frac{x^3 - 2x}{x^2 + 3x + 2} dx = \int \left[x - 3 + \frac{1}{x+1} + \frac{4}{x+2} \right] dx = \frac{1}{2}x^2 - 3x + \ln|x + 1| + 4\ln|x + 2| + C.$$

Además, la derivada de una función definida implícitamente está definida en forma explícita $z = F(x, y) = 0$ o $F(x, y, z) = 0$ en forma explícita, en lugar de la habitual.

Ejemplo 2

$x^2 + y^2 + z^2 = a^2$ donde $z = \pm\sqrt{a^2 - x^2 - y^2}$. No siempre es factible despejar z , pues para definir z_x y z_y se supone que $z = F(x, y)$ donde $F(x, y, z_{(x,y)}) = 0$ en algún conjunto abierto \mathbb{D} . Sea $g(x, y) = F(x, y, z) = 0$, aplicando la regla de la cadena se tiene que $g(x, y) = F(\mu, \nu, \omega)$ con:

$$\mu(x, y) = x$$

$$\nu(x, y) = y$$

$$\omega(x, y) = z(x, y)$$

Se obtiene:

$$\left. \begin{aligned} \frac{\partial g}{\partial x} &= \frac{\partial f}{\partial \mu} * \frac{\partial \mu}{\partial x} + \frac{\partial f}{\partial \nu} * \frac{\partial \nu}{\partial x} + \frac{\partial f}{\partial \omega} * \frac{\partial \omega}{\partial x} \\ \frac{\partial z}{\partial y} &= \frac{\partial f}{\partial \mu} * \frac{\partial \mu}{\partial y} + \frac{\partial f}{\partial \nu} * \frac{\partial \nu}{\partial y} + \frac{\partial f}{\partial \omega} * \frac{\partial \omega}{\partial y} \end{aligned} \right\} \begin{aligned} \frac{\partial g}{\partial x} &= 1 - \frac{\partial f}{\partial \mu} + \frac{\partial f}{\partial \omega} * \frac{\partial z}{\partial x} = 0 \\ \frac{\partial \nu}{\partial x} &= 0 \end{aligned}$$

$\frac{\partial z}{\partial y} = -\frac{\frac{\partial f}{\partial \nu}}{\frac{\partial f}{\partial \omega}}$ en todos los punto \mathbb{D} donde $\frac{\partial f}{\partial \omega}$ con $(x, y, z_{(x,y)}) \neq 0$. Igual es

para $\frac{\partial z}{\partial x} = -\frac{\frac{\partial f}{\partial \mu}}{\frac{\partial f}{\partial \omega}}$

Teorema: Si F es diferenciable en un conjunto abierto \mathbb{D} de \mathbb{R}^2 y si la ecuación $F(x_1, x_2, x_3, x_4, \dots, x_{n-1})$ en algún conjunto abierto de \mathbb{R}^{n-1} , entonces:

$$\frac{\partial f}{\partial x_i} = -\frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial x_n}} \rightarrow \frac{\partial f}{\partial x_n} \neq 0$$

Z es definida de manera implícita por $F(x, y, z) = 0$, pues $F(x, y, z) = 0$ de acuerdo con la hipótesis del teorema $z_x = -\frac{F_x}{F_z}$ y $z_y = -\frac{F_y}{F_z}$ tal que

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}$$

También se cumple si x, z son las variables independientes:

$$y_x = -\frac{F_x}{F_y} \text{ y } y_{zy} = -\frac{F_z}{F_y}$$

Aunque, este teorema se puede generalizar para ecuaciones:

$$F(x, y, z, \mu) = 0$$

Ejemplo a): Sea $F(x, y, z) = xyz + x + y - z = 0$. Tal que $z_x = -\frac{F_x}{F_z} = -\frac{yz+1}{xy-1}$ y $z_y = -\frac{F_y}{F_z} = -\frac{xz+1}{xy-1}$.

Ejemplo 3

Sea $z = x^2 + y^2 + z^2 - 1 = 0$. Encuentre derivadas del cociente z_x, z_y y z_{xx}, z_{yy}, z_{xy} :

$$z_x = -\frac{F_x}{F_z} = -\frac{2x}{2z} = -\frac{x}{z}$$

$$z_y = -\frac{F_y}{F_z} = -\frac{2y}{2z} = -\frac{y}{z}$$

Estas últimas no son implícitas, pues la variable dependiente es explicada por $F(x, y, z) = xyz + x + y - z$.

$$\begin{aligned} z_{xx} &= \frac{\partial(z_x)}{\partial x} = \frac{\partial\left(-\frac{x}{z}\right)}{\partial x} \Rightarrow \left(\frac{\mu}{v}\right)' = \frac{\mu'v - \mu v'}{v^2} \Rightarrow = -\frac{1 * z - (x * z_x)}{z^2} \\ &= -\frac{z - x\left(-\frac{x}{z}\right)}{z^2} = -\frac{z + \frac{x^2}{z}}{z^2} \end{aligned}$$

$$\begin{aligned} z_{yy} &= \frac{\partial(z_y)}{\partial y} = \frac{\partial\left(-\frac{y}{z}\right)}{\partial y} \Rightarrow \left(\frac{\mu}{v}\right)' = \frac{\mu'v - \mu v'}{v^2} \Rightarrow = -\frac{1 * z - (y * z_y)}{z^2} \\ &= -\frac{z - y\left(-\frac{xy}{z}\right)}{z^2} = -\frac{z + \frac{y^2}{z}}{z^2} \end{aligned}$$

$$\begin{aligned} z_{yx} &= \frac{\partial(z_y)}{\partial x} = \frac{\partial\left(-\frac{y}{z}\right)}{\partial x} \Rightarrow \left(\frac{\mu}{v}\right)' = \frac{\mu'v - \mu v'}{v^2} \Rightarrow \\ &= \frac{\frac{\partial(y)}{\partial x} * z(0) - (-y * z_x)}{z^2} = \frac{y z_y}{z^2} = \frac{y\left(-\frac{x}{z}\right)}{z^2} = -\frac{xy}{z^3} \end{aligned}$$

$$\begin{aligned} z_{xy} &= \frac{\partial(z_x)}{\partial y} = \frac{\partial\left(-\frac{x}{z}\right)}{\partial y} \Rightarrow \left(\frac{\mu}{v}\right)' = \frac{\mu'v - \mu v'}{v^2} \Rightarrow \\ &= \frac{\frac{\partial(x)}{\partial y} * z(0) - (-x * z_y)}{z^2} = \frac{x z_y}{z^2} = \frac{x\left(-\frac{y}{z}\right)}{z^2} = -\frac{xy}{z^3} \end{aligned}$$

La derivación implícita (caso de dos ecuaciones), sea $F(x, y, z) \in U \subseteq \mathbb{R}^2 \rightarrow F(x, y) \in \mathbb{R}$ un campo escalar.

La ecuación $F(x, y) = 0$ representa, habitualmente, una curva en el plano y la cuestión es averiguar si se puede despejar la variable y en

función x para representar la curva de manera explícita mediante la función $y = y(x)$. Suponga que $\left. \begin{matrix} \mu = \mu(x, y) \\ \nu = \nu(x, y) \end{matrix} \right\}$ son definidas implícitamente, pues la variable dependiente no es explicada o escrita respecto a las independientes y , también, $F(x, y, \mu, \nu) = 0$ y $G(x, y, \mu, \nu) = 0$ para deducir $\mu_x, \mu_y, \nu_x, \nu_y$ se tiene el sistema:

$$\begin{cases} \delta F = F_x \delta x + F_y \delta y + F_\mu \delta \mu + F_\nu \delta \nu = 0 \\ \delta G = G_x \delta x + G_y \delta y + G_\mu \delta \mu + G_\nu \delta \nu = 0 \end{cases}$$

$$\text{Sea: } J = \begin{vmatrix} F_\mu & F_\nu \\ G_\mu & G_\nu \end{vmatrix} \Rightarrow \delta \mu = -\frac{1}{J} \begin{vmatrix} F_x & F_\nu \\ G_x & G_\nu \end{vmatrix} \delta x - -\frac{1}{J} \begin{vmatrix} F_y & F_\nu \\ G_y & G_\nu \end{vmatrix} \delta y.$$

$$\text{Como: } \mu_x = -\frac{\begin{vmatrix} F_x & F_\nu \\ G_x & G_\nu \end{vmatrix}}{J}, \quad \mu_y = -\frac{\begin{vmatrix} F_y & F_\nu \\ G_y & G_\nu \end{vmatrix}}{J}, \quad \nu_x = -\frac{\begin{vmatrix} F_\mu & F_y \\ G_\mu & G_y \end{vmatrix}}{J} \quad \text{y} \quad \nu_y = -\frac{\begin{vmatrix} F_\mu & F_x \\ G_\mu & G_x \end{vmatrix}}{J}.$$

Ejemplo 4

Si $\left. \begin{matrix} \mu = \mu(x, y) \\ \nu = \nu(x, y) \end{matrix} \right\}$ son funciones implícitas por

$$\begin{cases} F = \mu^2 + \nu^2 - x^2 - y = 0 \\ G = \mu + \nu - x^2 + y = 0 \end{cases}$$

Hallar μ_x y μ_y empleando la Regla del matemático francés Pierre Frédéric Sarrus, que ayuda de manera artificiosa a relacionar palabras, con el objetivo de memorizar conceptos con más facilidad- especial de fórmula de Gottfried Wilhelm Leibniz:

$$J = \begin{vmatrix} F_\mu & F_\nu \\ G_\mu & G_\nu \end{vmatrix} = \begin{vmatrix} 2\mu & 2\nu \\ 1 & 1 \end{vmatrix} = 2\mu - 2\nu$$

$$\begin{aligned}\mu_x &= -\frac{\begin{vmatrix} F_x & F_v \\ G_x & G_v \end{vmatrix}}{2(\mu - v)} = -\frac{\begin{vmatrix} -2x & -2v \\ -2x & 1 \end{vmatrix}}{2(\mu - v)} = -\frac{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{2(\mu - v)} = -\frac{-2x + 4xv}{2(\mu - v)} \\ &= \frac{+2x - (2x * 2v)}{2(\mu - v)} = \frac{2x(1 - 2v)}{2(\mu - v)} = \frac{x(1 - 2v)}{(\mu - v)}\end{aligned}$$

$$\begin{aligned}\mu_y &= -\frac{\begin{vmatrix} F_y & F_v \\ G_y & G_v \end{vmatrix}}{2(\mu - v)} = -\frac{\begin{vmatrix} -1 & -2v \\ 1 & 1 \end{vmatrix}}{2(\mu - v)} = -\frac{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{2(\mu - v)} = -\frac{-1 - 2v}{2(\mu - v)} \\ &= \frac{-1 - (1 * 2v)}{2(\mu - v)} = \frac{1 + 2v}{2(\mu - v)}\end{aligned}$$

Ejemplo 5

Sea $z = f(x, y)$, $z = \mu + v$ $\begin{cases} \mu = \mu(x, y) \\ v = v(x, y) \end{cases}$ son definidas implícitamente

por
$$\begin{cases} F = \mu + e^{\mu+v} - x = 0 \\ G = v + e^{\mu-v} - y = 0 \end{cases}$$

Si $\mu = v = 0 \Rightarrow x = y = 1$. Calcular $z_x(1, 1)$:

$$z_x = \mu_x + v_x \rightarrow \mu_x = -\frac{F_x}{F_\mu} \text{ y } v_x = -\frac{F_x}{F_v} \text{ no son necesarios, pero } \begin{cases} F = 0 \\ G = 0 \end{cases}$$

se derivan parcialmente respecto a x , usando la Regla de la Cadena, tal que:

$$F: \mu_x + e^{\mu+v} * (\mu_x + v_x) - 1 = 0$$

$$G: v_x + e^{\mu-v} * (\mu_x - v_x) - 0 = 0$$

Sin embargo, $\mu = v = 0 \Rightarrow x = y = 1$ implica:

$$F: \mu_x + e^0 * (\mu_x + v_x) - 1 = 0 \Rightarrow 2\mu_x + v_x = 1 = 0$$

$$G: v_x + e^0 * (\mu_x - v_x) = 0 \Rightarrow \mu_x = 0 \Rightarrow 2(0) + v_x = 1 \therefore v_x = 1$$

Finalmente, $z_x = \mu_x + v_x = 0 + 1 \therefore z_x(1, 1) = 1$

CAPÍTULO II

ANÁLISIS ECONOMÉTRICOS

2.1. El gradiente

El gradiente (∇ Derivada Direccional – Vector Gradiente) de un campo escalar (ecuaciones como $2x + 3y = 6$) es un campo vectorial que apunta en dirección de la mayor tasa de aumento del campo escalar, cuya magnitud es la mayor tasa de cambio.

Campo gradiente, sea $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ es una función o campo escalar diferenciable en una región \mathbb{R} , entonces la función o campo gradiente de f es la función vectorial $\nabla_f: R \subseteq (\text{sub conjunto})\mathbb{R}^n \rightarrow \mathbb{R}$ definido por $\nabla_f(x_1, x_2, x_3, x_4, \dots, x_n) = fx_1, fx_2, fx_3, fx_4, \dots, fx_n$ (derivadas parciales).

En caso $f: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ implica $\nabla_{f(x,y)} = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j}$, mientras que en caso $f: D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ definido por $\nabla_f(x, y, z) = fx, fy, fz = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k}$, donde \vec{i}, \vec{j} y \vec{k} son vectores unitarios o vector normalizado (vector cuyo módulo, magnitud o norma es igual a 1).

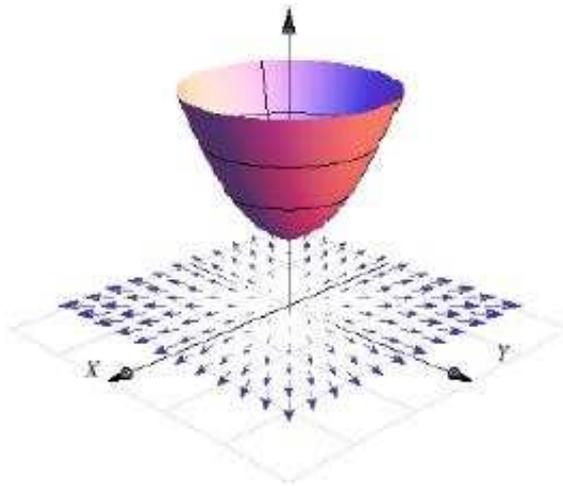
2.2. La interpretación geométrica de campo gradiente

(Curvas y Superficies de Nivel), el gradiente o vector gradiente es denotado por ∇ , $\vec{\nabla}$ o $\text{grad}(f)$ de un campo escalar (representa la distribución espacial de una magnitud escalar, asociando un valor a cada punto del espacio) f es un campo vectorial - representa la distribución espacial de una magnitud vectorial o expresión de cálculo vectorial que asocia un vector a cada punto en el espacio euclidiano-.

El vector gradiente de f evaluado en un punto genérico x del dominio de f , $\nabla f(x)$ indica la dirección en que el campo f varía más rápidamente y su módulo representa el ritmo de variación de f en la dirección de dicho vector gradiente.

Para el efecto se ilustra como ejemplo a la Figura 2.1 (Paraboloide): El gradiente $\nabla_z: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es un campo vectorial conocido como campo gradiente. Sea $z - 1 = x^2 + y^2$ un paraboloide, donde $\nabla_z = (z_x, z_y) = (2x, 2y)$ es un vector en dos dimensiones, tal que
$$\begin{cases} \nabla_f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} \\ \nabla_f = 2x\vec{i} + 2y\vec{j} \end{cases}$$

Figura 2.1 Interpretación geométrica de campo gradiente



Fuente: (Mora, W. F. 2016).

Características:

- Se aprecian algunos de sus vectores «trasladados» en el Plano (II) XY.
- Estos vectores apuntan en dirección del máximo crecimiento del paraboloide.
- Las magnitudes de estos vectores dan una medida de la «intensidad» de esta razón de cambio.

Ejemplo 5

$f(x, y) = \text{Sen}(xy) + x^2y^2$. Calcule: $\nabla_f = (x, y) = (\pi, 1)$.

Entonces, $\nabla_f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} = [y * \text{Cos}(xy) + 2xy^2] \vec{i} + [x * \text{Cos}(xy) + 2yx^2] \vec{j}$

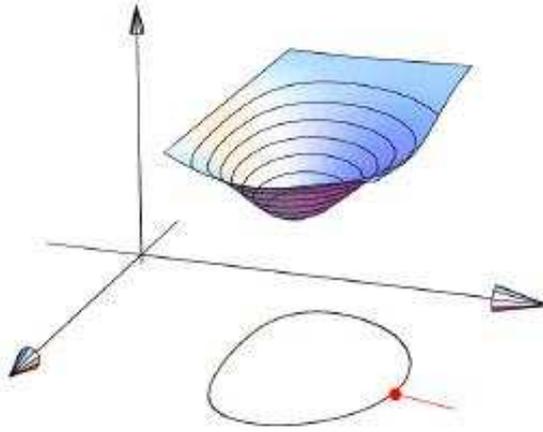
$\Rightarrow \nabla_{f(\pi,1)} = [1 * \text{Cos}(\pi * 1) + 2\pi 1^2] \vec{i} + [\pi * \text{Cos}(\pi * 1) + 21\pi^2] \vec{j} = [-1 + 2\pi] \vec{i} + [-\pi + 2\pi^2] \vec{j}$ debido a que, las curvas y superficies de nivel del gradiente no son funciones, sino conjuntos que pueden verse como «cortes» de gráficos que describen las funciones, tal que a cada corte de función se llama «nivel».

En consecuencia, la curva de nivel se presenta cuando se sabe que $z = f(x, y)$, $z = \zeta$ o $\zeta = f(x, y)$. Entonces, superficie de nivel se presentan si $w = g(x, y, z)$, $w = 0$ o $g(x, y, z) = 0$.

Tal que si S es una superficie de ecuación $g(x, y, z) = 0$ con G derivable con continuidad en plano si $P = (x_0, y_0, z_0) \in S$, entonces:

1. Se cumple en P (Función Implícita) tal que $\nabla_{z(x,y)} = \left(-\frac{G_x}{G_z}, -\frac{G_y}{G_z} \right)$, $\nabla_{z(x_0, y_0)}$ es perpendicular (\perp) a la curva de nivel $z = z_0$. Si se requiere un vector \perp se usa solamente $(-G_x, -G_y)$.
2. Después, si $z = f(x, y)$ es una superficie, se puede calcular el ∇ de manera $G = z - f(x, y) = 0$ y, luego, $G_x = 1$:

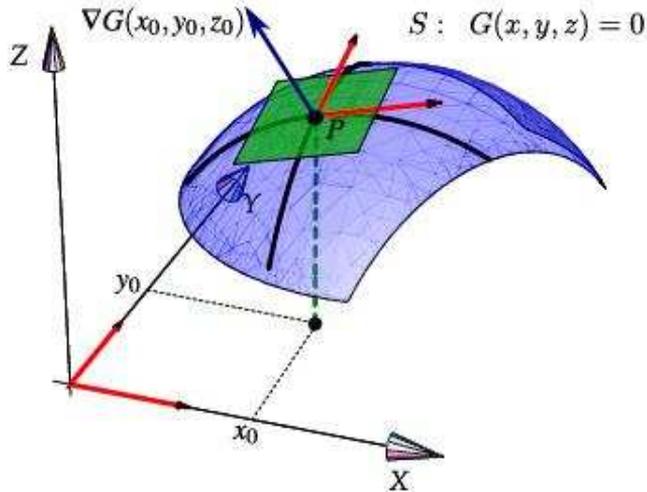
Figura 2.2 Superficie de nivel



Fuente: (Mora, W. F. 2016).

El vector $\nabla_G(x_0, y_0, z_0)$ es \perp a la superficie de nivel. Es decir, $\nabla_G(x_0, y_0, z_0)$ es \perp a cada curva de la superficie S ($w = 0$), que pasa por $P = (x_0, y_0, z_0)$:

Figura 2.3 Curva de superficie



Fuente: (Mora, W. F. 2016).

Ejemplo 6

Sea \mathbb{C} curva $x^2 - y^2(1 + x) = 0$. Sea $P = \left(\frac{1}{6}, \frac{\sqrt{7}}{\sqrt{216}}\right)$ aunque $P \in \mathbb{C}$.

Calcule $\vec{u} \perp$ a curva P . Entonces:

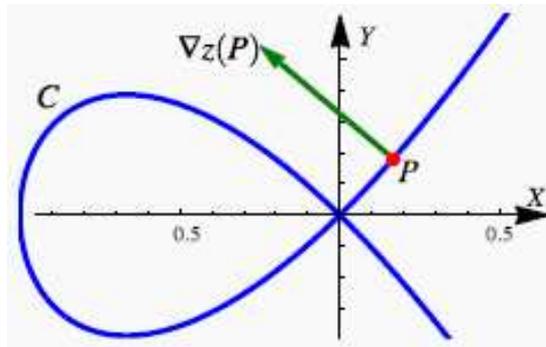
$$\nabla_{z(x,y)} = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right) = (-2x(1+x) - x^2, 2y) = (-2x - 3x^2, 2y)$$

$$\begin{aligned} \nabla_{z(P)} &= \left[-2\left(\frac{1}{6}\right) - 3\left(\frac{1}{6}\right)^2, 2\left(\frac{\sqrt{7}}{\sqrt{216}}\right) \right] = \left[-\frac{5}{12}, \frac{\sqrt{7}}{\sqrt{54}} \right] \\ &= [-0.41667, 0.36004] \end{aligned}$$

Es importante mencionar que $\mathbb{C}: z = x^2 - y^2(1 + x)$ es curva de nivel cuando $z = 0$ tal que $\nabla_{z(P)}$ es \perp a \mathbb{C} .

Entonces, $\nabla_{z(P)} = [-0.41667, 0.36004]$ y $P = (0.16, 0.18)$:

Figura 2.4 Curva de nivel



Fuente: (Mora, W. F. 2016).

La línea azul es la curva de nivel, $0 = x^2(1 + x)$ tal que $x_1 = 0$ y $x_2 = -2$.

Ejemplo 7

Sea $S: \frac{1}{9}(z-1)^2 + (x-2)^2 + (y-2)^2 - 4 = 0$, $P = (3, 2, 1 + 3\sqrt{3}) \in S$.

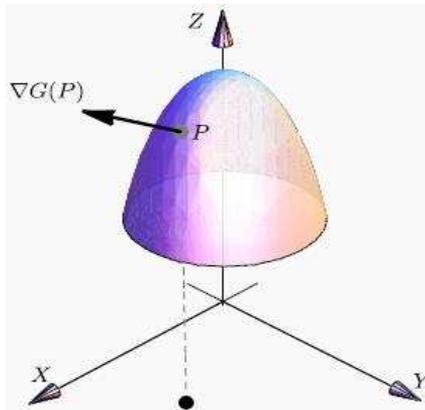
Calcule $\vec{u} \perp$ a S en P . Se sabe que $\nabla G(P)$ es \perp a S (curva de nivel) en

P . Donde: $G_{(x,y,z)} = (x-2)^2 + (y-2)^2 + \frac{1}{9}(z-1)^2 - 4 =$

$(G_x, G_y, G_z) = [2(x-2), 2(y-2), \frac{2}{9}(z-1)] \Rightarrow \nabla G(P) =$

$[2(3-2), 2(2-2), \frac{2}{9}(1+3\sqrt{3}-1)] = (2, 0, \frac{2}{\sqrt{3}}):$

Figura 2.5 Curva de nivel



Fuente: (Mora, W. F. 2016).

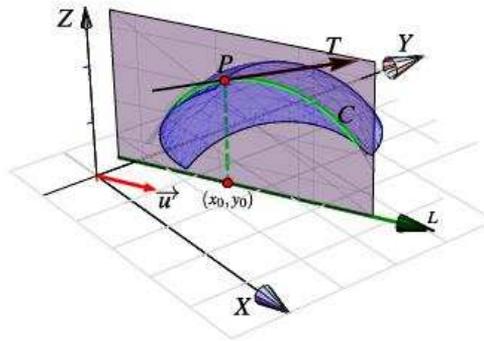
2.3. La derivada direccional

La derivada direccional ($D_{\vec{\mu}}f$) o derivada según una dirección es la derivada direccional en un punto en dirección del vector μ . Es decir, es la dirección de un vector dado, representa la tasa de cambio de la función en dirección de dicho vector. Este concepto generaliza las derivadas parciales, pues son derivadas direccionales según la dirección de respectivos ejes coordenados.

Suponga que se calcula la tasa de cambio $z = f(x, y)$ en el punto (x_0, y_0) en dirección del vector unitario $\vec{\mu} = (a, b)$. Se debe considerar la superficie $S: z = f(x, y)$ y $z_0 = f(x_0, y_0)$. Donde $P = (x_0, y_0, z_0) \in S$.

El plano vertical generado por recta L por $(x_0, y_0, 0)$ en dirección del vector $\vec{\mu}$, interseca a la superficie S en curva \mathcal{C} y la pendiente de la recta tangente T a la curva \mathcal{C} en P es la tasa de cambio de Z en dirección del vector $\vec{\mu}$:

Figura 2.6 Proyecciones ortogonales



Fuente: (Mora, W. F. 2016).

Sea $Q = (x, y, z)$ otro punto sobre \mathcal{C} , sean $P' = (x_0, y_0)$ tal que $Q' = P' + h\vec{\mu}$ proyecciones ortogonales sobre plano XY de puntos P y Q . Entonces, $\overrightarrow{P'Q'} = Q' - P' = h\vec{\mu}$. Para algún escalar h , se tiene:

$$x - x_0 = ha \Rightarrow x = x_0 + ha$$

$$y - y_0 = hb \Rightarrow y = y_0 + hb$$

También, sobre L es $\|\overrightarrow{P'Q'}\|$ (módulo, longitud, magnitud o norma del vector) $= h\|\vec{\mu}\| = h \rightarrow \vec{\mu}$ es unitario. Por lo tanto, la razón de cambio está dada por:

$$\frac{\Delta z}{h\|\vec{\mu}\|} = \frac{\Delta z}{h} = \frac{z - z_0}{h} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

Finalmente, $\lim_{h \rightarrow 0} \left(\frac{\Delta z}{h} \right)$ es la tasa de cambio instantánea de Z en la dirección de $\vec{\mu}$, conocida como derivada direccional de f en dirección de $\vec{\mu}$.

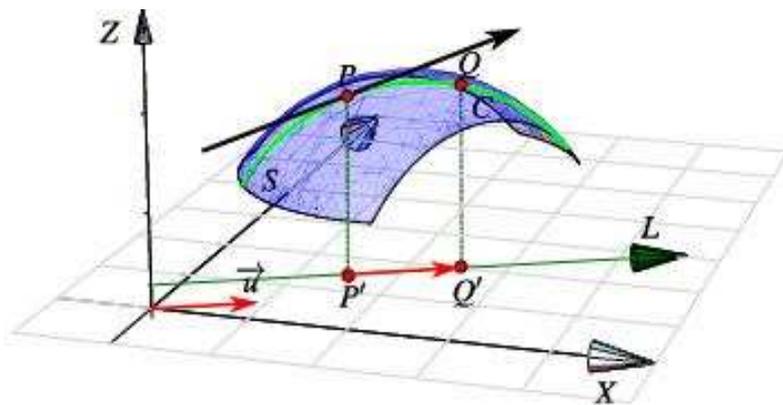
Definición Formal: Sean $f: D \subset (\text{subconjunto})\mathbb{R}^2 \rightarrow \mathbb{R}$ una función escalar y sean $(x_0, y_0) \in D$, $\vec{\mu} = (a, b)$ un vector unitario. Entonces, la derivada direccional de f en (x_0, y_0) en dirección del $\vec{\mu}$ está dada por:

$$D_{\vec{\mu}}f(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

Teorema: Cálculo Derivada Direccional. Sea $f: D \subset (\text{subconjunto})\mathbb{R}^n \rightarrow \mathbb{R}$ una función escalar diferenciable en D . Entonces, f tiene una derivada direccional en dirección de cualquier vector no nulo $\vec{\mu} = (a, b)$ que está dada por:

$$D_{\vec{\mu}}f(x, y) = f_x(x, y) \frac{a}{\|\vec{\mu}\|} + f_y(x, y) \frac{b}{\|\vec{\mu}\|}$$

Figura 2.7 Derivada direccional



Fuente: (Mora, W. F. 2016).

Tal que, la derivada direccional de f en la dirección del vector $\vec{\mu}$ es $\|\overrightarrow{P'Q'}\|$ (módulo, magnitud o norma del vector) = $h\|\vec{\mu}\|$.

Ejemplo 8

Calcular la derivada direccional $D_{\vec{\mu}}f(x, y)$, si $f(x, y) = x^3 - 3xy + 4y^2$ y $\vec{\mu} = (\sqrt{3}, 1)$. Hallar $D_{\vec{\mu}}f(1, 2)$ en $P(x, y)$. Entonces, $\nabla_{f(x,y)} = (F_x, F_y) = (3x^2 - 3y, -3x + 8y) \Rightarrow \nabla_{f(1,2)} = (F_x, F_y) = (3(1)^2 - 3(2), -3(1) + 8(2)) = (-3, 13)$.

Por lo tanto, $D_{\vec{\mu}}f(1, 2) = \nabla_{f(x,y)} * \frac{\vec{\mu}}{|\vec{\mu}|} = (-3, 13) * \left(\frac{\sqrt{3}, 1}{2}\right)$ debido a que

$$\vec{\mu} = (\sqrt{3}, 1), |\vec{\mu}| = \sqrt{(\sqrt{3})^2 + (1)^2} = \sqrt{4} = 2.$$

$$\text{Entonces, } D_{\vec{\mu}}f(1, 2) = (-3, 13) * \left(\frac{\sqrt{3}, 1}{2}\right) = -\frac{3}{2}\sqrt{3} + \frac{13}{2} = -1.3733.$$

Ejemplo 9

Hallar $D_{\vec{\mu}}f(x, y, z)$ en $P(x, y, z) = P(1, 3, 0)$ siendo $f(x, y, z) = x \text{ Sen}(yz)$ y $\vec{\mu} = \vec{i} + 2\vec{j} - \vec{k}$. Entonces, $\nabla_{f(x,y,z)} = (F_x, F_y, F_z) = [\text{Sen}(yz), xz \text{ Cos}(yz), xy \text{ Cos}(yz)]$.

Por lo tanto, $\nabla_{f(1,3,0)} = (F_x, F_y, F_z) = [\text{Sen}(3 * 0), 1 * 0 \text{ Cos}(3 * 0), 1 * 3 \text{ Cos}(3 * 0)] = (0, 0, 3)$.

Paralelamente,
$$\vec{\mu} = \vec{i} + 2\vec{j} - \vec{k} = (1, 2, -1),$$

$$|\vec{\mu}| = \sqrt{(1)^2 + (2)^2 + (-1)^2} = \sqrt{6} = 2.4495.$$

Por lo tanto,
$$D_{\vec{\mu}}f(x, y, z) \Rightarrow D_{\vec{\mu}}f(1, 3, 0) = (0, 0, 3) * \left(\frac{1, 2, -1}{\sqrt{6}}\right) = \left(\frac{0*2}{\sqrt{6}}\right) + \left(\frac{0*1}{\sqrt{6}}\right) + \left(\frac{3*-1}{\sqrt{6}}\right) = \left(\frac{3*-1}{\sqrt{6}}\right) = -1.2247$$

2.4. El Vector Proyección-Componente

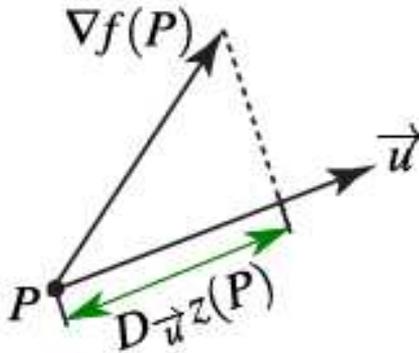
La componente de ∇ en la dirección $\vec{\mu}$ es $\frac{\vec{\mu} \cdot \nabla}{|\vec{\mu}|}$. Esta componente es la longitud de la Proyección Vectorial de ∇ sobre $\vec{\mu}$:

$$P\nabla_{\vec{\mu}} = \frac{\vec{\mu} \cdot \nabla}{|\vec{\mu}|} \vec{\mu}$$

Tal que, la Derivada Direccional en la dirección del vector $\vec{\mu}$ es:

$$D_{\vec{\mu}}f(x, y, z) = \nabla f(x, y, z) \cdot \frac{\vec{\mu}}{|\vec{\mu}|}$$

Figura 2.8 Derivada direccional



Fuente: (Mora, W. F. 2016).

2.5. La dirección de máximo y mínimo

La dirección de máximo y mínimo sea $f: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ una función escalar. El valor máximo de la derivada direccional $D_{\vec{\mu}}f$ en (x, y) es $\|\nabla f(x, y)\|$ y se presenta cuando el vector no nulo \vec{u} tiene la misma dirección que el vector gradiente $\nabla f(x, y)$.

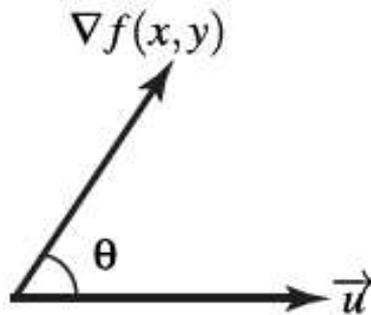
Suponga que tiene una función f de dos o tres variables y considere todas las posibles derivadas direccionales de f en un punto P dado. Esto

proporciona las tasas de cambio de f en todas las posibles direcciones. De modo que se pueden plantear las preguntas «¿en cuál de estas direcciones f cambia con mayor velocidad? y ¿cuál es la máxima razón de cambio?».

Tal que, de manera intuitiva en la anterior figura, la derivada direccional en P aumenta conforme el vector $\vec{\mu}$ se acerca al gradiente. Con base en esto, el teorema dirección de máximo cambio se sustenta tal que el máximo valor se tendría cuando $\theta = 0$; es decir, cuando los vectores $\vec{\mu} \parallel \nabla f(x, y, z)$ porque $\text{Cos } \theta = \text{Cos } 0^\circ = 1$.

Entonces, $\nabla f \cdot \vec{\mu} = |\nabla f| \cdot |\vec{\mu}| \cdot \text{Cos}\theta$:

Figura 2.94 Valor mínimo derivada direccional



Fuente: (Mora, W. F. 2016).

El valor mínimo de la derivada direccional en (x, y, z) es $-\|\nabla f(x, y, z)\|$, ocurre cuando $\vec{\mu}$ tiene la misma dirección de $-\nabla f(x, y, z)$. No obstante, f se mantiene constante sobre curvas de nivel, la dirección $\vec{\mu}$ en que el cambio (instantáneo) de f respecto a P es nulo es la dirección de un vector \perp a $\nabla f(x, y, z)$, tal que la derivada direccional se anule en P en dirección $\vec{\mu}$.

El vector unitario tangente indica que sea $r: I \subseteq (\text{sub conjunto, contenido o igual})\mathbb{R} \rightarrow \mathbb{R}^n$, si la función

vectorial r es continua en \mathbb{I} implica que la gráfica de r se llama curva y está descrita paramétricamente por $r(t)$.

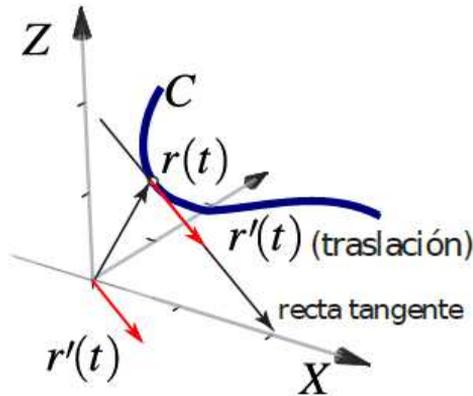
La parametrización de rectas, indica que sea la recta L en \mathbb{R}^3 que pasa por P en dirección de $\vec{\mu}$ implica que $r(t) = P + t\vec{\mu}$ $t \in \mathbb{R}$ y Elipses, representada por $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ y parametrizada mediante $r(t) = \begin{pmatrix} h + a \cos t \\ k + b \sin t \end{pmatrix} \begin{matrix} \vec{i} \\ \vec{j} \end{matrix} t \in [0, 2\pi]$.

Entonces, derivada de $r(t)$ implica que sea $r'(t) = \lim_{h \rightarrow 0} \frac{r(t+h) - r(t)}{h}$ e implica además que:

- a) $r(t) = x(t)\vec{i} + y(t)\vec{j} \Rightarrow r'(t) = x'(t)\vec{i} + y'(t)\vec{j}$ tal que $x(t), y(t)$ y $z(t)$ son derivables en \mathbb{I} .
- b) $r(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k} \Rightarrow r'(t) = x'(t)\vec{i} + y'(t)\vec{j} + z'(t)\vec{k}$.

Es importante mencionar que $r(t)$ parametriza a \mathbb{C} y $r'(t)$ es un vector unitario tangente:

Figura 2.10 Interpretación gráfica de curva



Fuente: (Mora, W. F. 2016).

Una definición de curva (\mathbb{C}) es aquella descrita por la función vectorial continua $r(t)$ $t \in \mathbb{I}$. Si existe la derivada $r'(t)$ y no es nula, la recta que pasa por $r(t)$ y es \parallel a $r'(t)$ se llama tangente a \mathbb{C} en $r(t)$. Entonces, el vector $r'(t)$ se denomina vector tangente a \mathbb{C} en $r(t)$.

El vector $r'(t)$ unitario tangente \mathbb{T} es una función vectorial asociada a curva \mathbb{C} , definida:

$$\vec{\mathbb{T}}(t) = \frac{r'(t)}{\|r'(t)\|} \Rightarrow \|r'(t)\| \neq 0$$

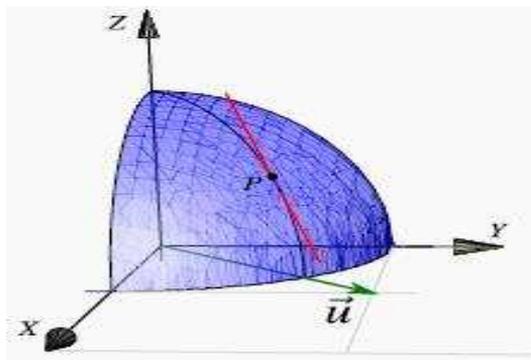
Ejemplos

Sea la pendiente de la recta tangente en P y en dirección $\vec{\mu} = (1, 1)$ es

$\mathbb{D}_{(1,1)}(\mathbf{P})$ (Derivada Direccional Z en punto P en dirección del vector $1, 1$) =

$$\nabla_{z(x,y)} * \frac{\vec{\mu}}{\|\vec{\mu}\|} = \nabla_{z\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)} * \frac{(1,1)}{\sqrt{(1)^2+(1)^2}} = -\sqrt{2}:$$

Figura 2.11 Derivada direccional

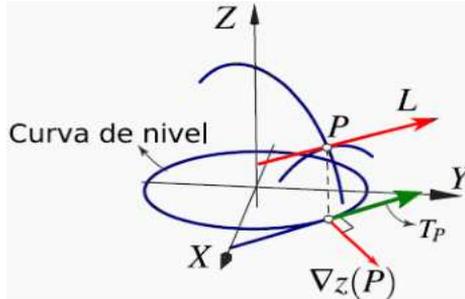


Fuente: (Mora, W. F. 2016).

El vector gradiente “nabla” o “del” $\nabla_{z\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)}$ es \perp a la recta tangente a \mathbb{C} de nivel $z = \frac{1}{\sqrt{3}} \approx 0.57$ en P . La $\mathbb{D}_{\vec{\mu}}(\mathbf{P}) = 0$.

Geoméricamente, L en la siguiente Figura tiene pendiente cero:

Figura 2.12 Curva de nivel



Fuente: (Mora, W. F. 2016).

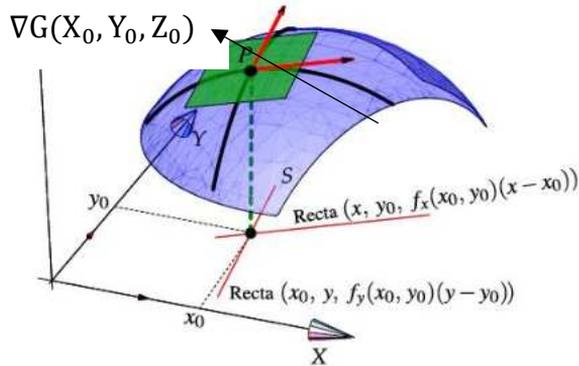
Si $\vec{T}(P)$ es vector unitario tangente a la curva de nivel $\mathbb{Z}: \mathbb{Z} = \frac{1}{\sqrt{3}}$ en $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, por lo que $\mathbb{D}_{\vec{T}(P)}f(P) = \nabla f(P) * \vec{T}(P) = \theta$, pues $\mathbb{D}_{\vec{T}(P)}f(P) = \|\nabla f(P)\| * \text{Cos}(\theta) = \theta$ debido a que $\text{Cos}(\frac{\pi}{2}) = \theta$ y $\text{Cos}(90^\circ) = \theta \Rightarrow \theta = \frac{\pi}{2}$.

Además, el producto escalar es $\vec{A} * \vec{B} = |\vec{A}| * |\vec{B}| * \text{Cos}(\theta)$.

El Plano Tangente y Vector Normal está dado, pues si f es diferenciable implica que el Π Tangente a $Z = f(x,y)$ en $P = (X_0, Y_0, Z_0)$ tiene como ecuación:

$$f_X(X_0, Y_0)(X - X_0) + f_Y(X_0, Y_0)(X - Y_0) = Z - Z_0$$

Figura 2.13 Plano tangente y Vector normal



Fuente: (Mora, W. F. 2016).

Sea Π Tangente tal que $P = (X_0, Y_0, Z_0)$ un punto fijo sobre un Π con vector normal $n = ai + bj + ck$ y si $Q = (x, y, z)$ es otro punto en Π , entonces $\overrightarrow{PQ} = (X - X_0)i + (Y - Y_0)j + (Z - Z_0)k$, como $\overrightarrow{PQ} \perp n$, se tiene que $\overrightarrow{PQ} * n = 0$ implica:

$$a(X - X_0) + b(Y - Y_0) + c(Z - Z_0) = 0$$

Tal que, una manera más común de escribir una ecuación cartesiana de un Π es:

$$aX + bY + cZ = d \Rightarrow d = aX_0 + bY_0 + cZ_0 = \overrightarrow{OP} * n$$

Aunque, en general, se puede tener una ecuación cartesiana del Π Tangente si existiera. Si se tiene $S: G(x, y, z) = 0$, tal que si G es derivable con continuidad en $P = (X_0, Y_0, Z_0) \in S$, si ∇ en P es no nulo. Los vectores tangentes a cada C en S que pasan por P están en el Π Tangente a esta S en P .

Por tanto, $\nabla G(X_0, Y_0, Z_0)$ es vector normal (n) a este Π .

Entonces, una ecuación del Π Tangente en P es:



$$aX + bY + cZ = d \quad \begin{aligned} (a, b, c) &= \nabla G(X_0, Y_0, Z_0) \\ d &= \nabla G(X_0, Y_0, Z_0) \cdot P \end{aligned}$$

El Plano Tangente existe si: 87373_he_mo

- Si S tiene ecuación $Z = f(x, y)$ con f diferenciable, el Π Tangente en $P \in S$ tiene la ecuación cartesiana:

$$f_X(X_0, Y_0)(X - X_0) + f_Y(X_0, Y_0)(X - Y_0) = Z - Z_0$$

- Si la superficie S tiene ecuación $G(x, y, z) = 0$ con G diferenciable, el Π Tangente en $P \in S$ tiene la ecuación cartesiana:

$$G_X(P)X + G_Y(P)Y + G_Z(P)Z = \nabla G(P) \cdot P$$

Entonces, así como la ecuación de una recta en el espacio se obtiene especificando un punto sobre la recta o un vector paralelo a esta recta, pueden derivarse ecuaciones de un Π en el espacio especificando un punto en Π y un vector ortogonal a todos los vectores en el Π . Este vector ortogonal o perpendicular se llama Vector Normal al Π , denotado por n . No hay un vector normal; aunque todos tienen la misma dirección, el tamaño puede variar.

Tal que, si S tiene la ecuación $Z = Z(x, y)$ entonces si se agrega $G(x, y, z) = Z - Z(X, Y)$, un n es $N = (-Z_X, -Z_Y, 1)$. Si S está definida implícitamente por $G(x, y, z) = 0$ entonces un n es $N_1 = (G_X, G_Y, G_Z)$ ó $N_2 = \left(\frac{G_X}{G_Z}, \frac{G_Y}{G_Z}, 1\right) G_Z \neq 0 = \frac{1}{G_Z}(G_X, G_Y, G_Z)$.

2.6. Ejercicios Resueltos

1. Sea S superficie de la ecuación cartesiana explícita $Z = X^2 + 2Y^2$. Obtenga una ecuación cartesiana del Π Tangente a S en

$$P = \begin{pmatrix} 1 & 1 & 3 \\ X_0 & Y_0 & Z_0 \end{pmatrix}:$$

$$\text{Sea } f_X(X_0, Y_0)(X - X_0) + f_Y(X_0, Y_0)(Y - Y_0) = Z - Z_0, \quad f_X(X, Y) = 2x \Rightarrow f_X(1, 1) = 2(1) = 2, \quad f_Y(X, Y) = 4y$$

$$\Rightarrow f_Y(1, 1) = 4(1) = 4$$

$$\text{entonces } f_X(1, 1)(X - 1) + f_Y(1, 1)(Y - 1) = Z - 3$$

$$\Rightarrow 2(x - 1) + 4(Y - 1) = Z - 3$$

$$\Rightarrow 2X - 2 + 4Y - 4 = Z - 3$$

$$\Rightarrow \therefore \frac{2X}{aX} + \frac{4Y}{bY} - \frac{Z}{cZ} = \frac{3}{d}$$

Otra manera de solucionarlo es:

Sea $S: G(x, y, z) = Z - X^2 - 2Y^2 = 0$, entendida como una ecuación implícita.

Tal que un n a S en P es $\nabla G = (-2X, -4Y, 1)$

$$\begin{aligned} \Rightarrow \nabla G_{(1,1,3)} &\Rightarrow -2X - 4Y + 1Z = \nabla G_{(1,1,3)} * P \\ &= (-2, -4, 1) * (1, 1, 3) = -2 - 4 + 3 \end{aligned}$$

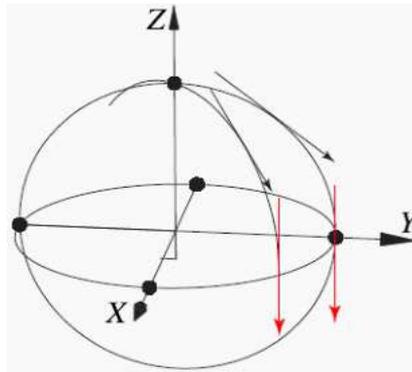
$$\Rightarrow -2X - 4Y + 1Z = -3$$

$$\therefore 2X + 4Y - Z = 3.$$

Entonces, $G_X(P)X + G_Y(P)Y + G_Z(P)Z = \nabla G(P) * P$

2. Sea $S: X^2 + Y^2 + Z^2 = 1$ y $P = (0, 1, 0) \in S$. Halle o estime la ecuación del ΠS en P :

Figura 2.14 Plano tangente y Vector normal



Fuente: (Mora, W. F. 2016).

$$\text{Sea } X^2 + Y^2 + Z^2 - 1 = 0 \text{ tal que } \nabla G = (2X, 2Y, 2Z) \Rightarrow \nabla G_{(0,1,0)} = (2(0), 2(1), 2(0))$$

$$\Rightarrow \nabla G_{(0,1,0)} * P = (0, 2, 0) * (0, 1, 0) \text{ ó, en otros términos, } N = (G_X, G_Y, G_Z) = G(0, 1, 0) = (0, 2, 0).$$

$$\text{Entonces, } 0X + 2Y + 0Z = 2 \Rightarrow Y = \frac{2}{2} = 1 \rightarrow S.$$

Por lo tanto, $\nabla Z_{(X,Y)} = \left(-\frac{X}{Z}, -\frac{Y}{Z}\right) \Rightarrow Z = 0$ tal que sus cocientes son indeterminados. Por lo tanto, no existe derivada direccional en $(0,1,0)$ y, con ello, las líneas son paralelas.

3. Sea $S: X^2 + Y^2 + Z^2 = 1$. Encontrar $Q = (a, b, c) \in S$ tal que Π Tangente en Q sea \parallel al $\Pi: 2X - Y + 3Z = 1$.

Entonces, se tienen 3 incógnitas, $Q \in S: a^2 + b^2 + c^2 = 1$, Π Tangente $\parallel 2X - Y + 3Z = 1$ y sus $\vec{N}_1 \parallel \vec{N}_2$.

Por lo tanto, $\nabla G_{(Q)} = \lambda(2, -1, 3)$. Tal que el sistema de ecuaciones se forma:

$$\begin{cases} a^2 + b^2 + c^2 = 1 \\ \nabla G_{(Q)} = \lambda(2, -1, 3) \\ (2a, 2b, 2c) = \lambda(2, -1, 3) \end{cases} \Rightarrow \begin{cases} a^2 + b^2 + c^2 = 1 \\ 2a = 2\lambda \Rightarrow a = \frac{2}{2}\lambda \\ 2b = -\lambda \Rightarrow b = -\frac{1}{2}\lambda \\ 2c = 3\lambda \Rightarrow c = \frac{3}{2}\lambda \end{cases}$$

Sustituyendo estos valores en:

$$a^2 + b^2 + c^2 = 1 \Rightarrow (\lambda)^2 + \left(-\frac{1}{2}\lambda\right)^2 + \left(\frac{3}{2}\lambda\right)^2 = 1 \Rightarrow \lambda = \pm\sqrt{\frac{2}{7}}.$$

Entonces, $a = \sqrt{\frac{2}{7}} * \frac{\sqrt{2}}{\sqrt{2}} = \frac{1}{\sqrt{2}}$, $b = -\frac{1}{2} * \sqrt{\frac{2}{7}} = -\sqrt{\frac{2}{2*14}} = -\frac{1}{\sqrt{14}}$ y $c = \frac{3}{2} *$

$$\sqrt{\frac{2}{7}} = 3 * \sqrt{\frac{2}{2*2*7}} = \frac{3}{\sqrt{14}}$$

CAPÍTULO III

MÁXIMOS Y MÍNIMOS LOCALES

Los máximos y mínimos locales en varias variables se estiman como en una variable, los extremos locales de una función de varias variables son puntos donde la función alcanza un máximo o mínimo en un entorno del dominio de la función.

Si la función está definida en una región D , los extremos globales son puntos donde la función toma valores máximos o mínimos y puede suceder en cualquier parte de la región en consideración.

En un entorno abierto alrededor $p \in \mathbb{R}^n$ con radio δ es conjunto $D_\delta(p) = \{X \in \mathbb{R}^n: \|X - P\| < \delta\}$, discos sin borde en \mathbb{R}^2 e interior de esferas en \mathbb{R}^3 .

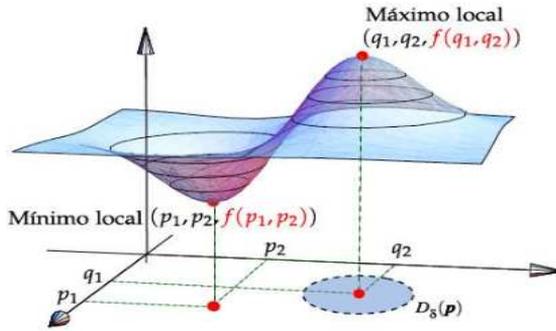
La definición de extremos locales se entiende como f función de n variables, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tal que f tiene un máximo local en $p = (p_1, p_2, p_3, p_4, \dots, p_n) \in \mathbb{R}^n$ si $\forall X \in D_\delta(p) \Rightarrow f(X_1, X_2, X_3, X_4, \dots, X_n) \leq f(p)$ para $\forall (X_1, X_2, X_3, X_4, \dots, X_n) \in D_\delta(p)$.

El punto $(p_1, p_2, p_3, p_4, \dots, p_n, f(p))$ tiene un máximo local de f y el número $f(p)$ es máximo de f en entorno $D_\delta(p)$.

También, f tiene un mínimo local en $p = (p_1, p_2, p_3, p_4, \dots, p_n) \in \mathbb{R}^n$ si $\forall X \in D_\delta(p)$ tal que $f(X_1, X_2, X_3, X_4, \dots, X_n) \geq f(p)$ para $\forall (X_1, X_2, X_3, X_4, \dots, X_n) \in D_\delta(p)$.

El punto $(p_1, p_2, p_3, p_4, \dots, p_n, f(p))$ se dice un mínimo local de f y el número $f(p)$ es mínimo de f en entorno $D_\delta(p)$.

Figura 3.1 Máximos y mínimos locales



Fuente: (Mora, W. F. 2016).

Si las desigualdades de la definición anterior se cumplen para todos los puntos en el dominio de f , implica que f tiene un máximo o un mínimo absoluto en p .

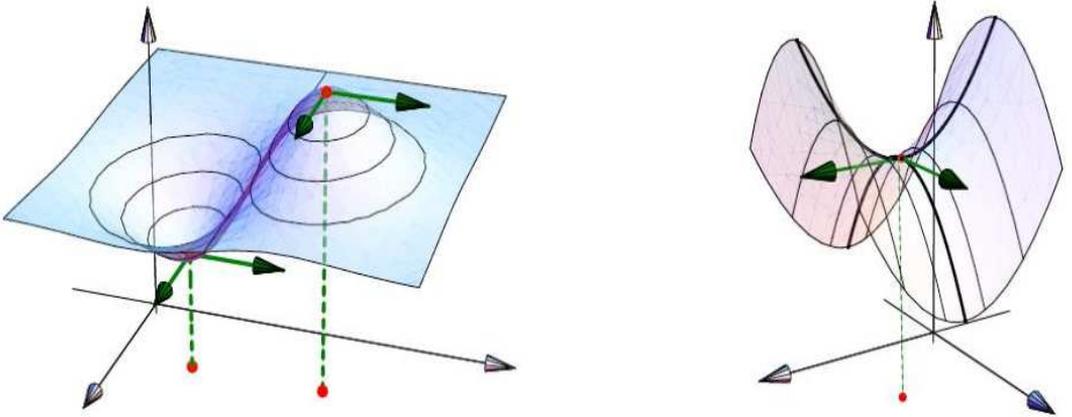
3.1. Puntos críticos y extremos locales

Los puntos críticos y extremos locales se explican pues f es continua. Un punto $p \in \mathbb{R}^n$ es un punto crítico de f si $Df(p) = 0$ o si Df no está definida en este punto; es decir, si $\frac{\partial f}{\partial x_i} = 0, i = 1, 2, 3, 4, \dots, n$. Un punto crítico que no es ni máximo ni mínimo local se llama punto silla ó ensilla (se ubica punto máximo y mínimo simultáneamente).

Como se calculó en una variable, los extremos locales son los puntos críticos; es decir, en caso de que f sea diferenciable, la derivada de f se anula en puntos críticos.

El Teorema que da sustento a esto sostiene que sea $U \subset \mathbb{R}^n$ (perteneciente, contenido o está incluido) un conjunto abierto y $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable, si $p \in \mathbb{R}^n$ es un extremo local de $f \Rightarrow Df(p) = 0$ ó p es punto crítico de f .

Figura 3.2 Extremos locales



Fuente: (Mora, W. F. 2016).

Es importante mencionar que en los extremos locales las derivadas parciales se anulan y en puntos de silla las derivadas parciales se anulan.

La clasificación de los puntos críticos se establece mediante la fórmula de Taylor de segundo orden en n variables, pues dice que si $f: U \subseteq (\text{subconjunto, contenido o igual}) \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en $X \in U \Rightarrow$ si $h = (h_1, h_2, h_3, h_4, \dots, h_n) \in \mathbb{R}^n, \in 0 < \xi < 1$ tal que $f(X + h) = f(X) + \sum_{i=1}^n h_i \frac{\partial f}{\partial X_i}(X) + R_1(X, h)$ con $R_1(X, h) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f}{\partial X_i \partial X_j}(X + \xi h)$. Si se define

$$D^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial X_1 \partial X_1} & \dots & \frac{\partial^2 f}{\partial X_1 \partial X_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial X_n \partial X_1} & \dots & \frac{\partial^2 f}{\partial X_n \partial X_n} \end{bmatrix} \Rightarrow \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f}{\partial X_i \partial X_j}(X + \xi h) = h *$$

$D^2 f(X + \xi h) * h^T$ (matriz transpuesta).

Por lo tanto, la fórmula de Taylor de segundo orden puede escribirse como:

$$f(X + h) = f(X) + Df(X) * h^T + \frac{1}{2}h * D^2f(X + \xi h) * h^T, 0 < \xi < 1.$$

La Hessiana de f en X es la forma cuadrática $\frac{1}{2}h * D^2f(X) * h^T$. Evaluando en un punto crítico p , $Df(p) = 0$ y la fórmula de Taylor de segundo orden queda $f(p + h) - f(p) = h * D^2f(X + \xi h) * h^T, 0 < \xi < 1$.

El signo de resta $f(p + h) - f(p)$ es el signo de $h * D^2f(X + \xi h) * h^T$. Si las derivadas $\frac{\partial^2 f}{\partial X_i \partial X_j}$ son continuas en un vecindario de $p \Rightarrow h * D^2f(X + \xi h) * h^T$ conserva el signo en un entorno de este punto, así $h * D^2f(X + \xi h) * h^T$ tiene igual signo que $h * D^2f(p) * h^T$ si h es suficientemente pequeño y, por tanto, el signo de $h * D^2f(p) * h^T$, siempre y cuando no se anule, decide si en p la función f alcanza un máximo o un mínimo local.

La matriz definida positiva o negativa es una forma cuadrática

$g: \mathbb{R}^n \rightarrow \mathbb{R}, g(h) = h * A_{n \times n} * h^T$ es definida positiva si $g(h) \geq 0$ para $\forall h \in \mathbb{R}^n$ y $g(h) = 0 \Leftrightarrow g(h) = 0$.

Paralelamente, g es definida negativamente si $g(h) \leq 0$ para $\forall h \in \mathbb{R}^n$ y $g(h) = 0 \Leftrightarrow g(h) = 0$.

Del álgebra lineal, $A = (a_{ij})_{n \times n}$, $D_1 = a_{11}$, $D_2 = \text{Det } |A| = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \dots, D_n = \text{Det} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \Rightarrow h * A_{n \times n} * h^T$ es

definida positiva si $D_i > 0$ para $i = 1, 2, \dots, n$ y $h * A_{n \times n} * h^T$ es definida negativa si $\text{sgn}(D_i) = (-1)^i$ para $i = 1, 2, \dots, n$.

El Teorema de Condición Suficiente afirma que si $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^3 y $p \in U$ un punto crítico de f . Si $h * D^2f(p) * h^T$ es definida

positivamente, tal que p es un mínimo relativo de f . Similarmente, si $h * D^2f(p) * h^T$ es definida negativamente, entonces p es máximo relativo de f .

Demostración: Si $h * D^2f(p) * h^T$ es definida positiva entonces en la fórmula de Taylor se tiene $f(p + h) - f(p) > 0$ en un entorno de p o $f(p)$ es un valor mínimo local. Si $h * D^2f(p) * h^T$ es definida negativa entonces en la fórmula de Taylor se tiene $f(p + h) - f(p) < 0$ en un entorno de p o $f(p)$ es un valor máximo local.

Por último, si $h * D^2f(p) * h^T$ no es definida positiva ni negativamente, se tiene un punto de silla o ensilla.

La clasificación de los puntos críticos para el caso de 2 variables es sencilla determinar si $h * D^2f(p) * h^T$ es definida positiva o negativamente:

$$h * D^2f(p) * h^T = (h_1 \ h_1) \begin{pmatrix} f_{xx}(p) & f_{xy}(p) \\ f_{yx}(p) & f_{yy}(p) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

Si f tiene derivadas parciales de segundo orden continuas, las derivadas mixtas son iguales: $D_1(p) = f_{xx}(p)$, $D_2(p) = f_{xx}(p) * f_{yy}(p) - [f_{xy}(p)]^2$.

El Teorema de Condición Suficiente sostiene que $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ de clase C^3 en un conjunto abierto U de \mathbb{R}^2 .

Sea $D_2(X, Y) = f_{xx}(X, Y) * f_{yy}(X, Y) - [f_{xy}(X, Y)]^2$ si $(X_0, Y_0) \in U$ es un punto crítico de f :

- a) Si $D_2(X_0, Y_0) > 0$ (+) y $f_{xx}(X_0, Y_0) > 0$ (+) $\Rightarrow f$ alcanza un mínimo local en (X_0, Y_0) .

- b) Si $D_2(X_0, Y_0) > 0 (+)$ y $f_{xx}(X_0, Y_0) < 0 (-) \Rightarrow f$ alcanza un máximo local en (X_0, Y_0) .
- c) Si $D_2(X_0, Y_0) < 0 (-) \Rightarrow (X_0, Y_0, f(X_0, Y_0))$ es un punto silla o ensilla.
- d) Si $D_2(X_0, Y_0) = 0$ se usan otros métodos para clasificar usando $f_{yy}(p)$ en vez de $f_{xx}(p)$, pues si $D_2(X_0, Y_0) > 0 (+)$ ambas tienen igual signo.

Ejemplos:

1. Calcule y clasifique con puntos críticos de $f(X, Y) = X^4 + Y^4 - 2X^2 + 4XY - 2Y^2$.

$$f_x = 4X^3 - 4X + 4Y \Rightarrow f_{xx} = 12X^2 - 4, \quad f_y = 4Y^3 + 4X - 4Y \Rightarrow f_{yy} = 12Y^2 - 4 \text{ y } f_{xy} = 4.$$

$$\text{Los puntos críticos } \begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \Rightarrow \begin{cases} 4X^3 - 4X + 4Y = 0 \quad 1) \\ 4Y^3 + 4X - 4Y = 0 \quad 2) \\ \hline X^3 + Y^3 = 0 \Rightarrow X = -Y \quad 3) \end{cases}$$

Sustituyendo 3) en 1) tal que:

$$4X^3 - 4X + 4(-X) = X^3 - X - X = 0 \Rightarrow X^3 - 2X = 0 \text{ ó en 2)}$$

$$4Y^3 - 4Y + 4(-Y) = Y^3 - Y - Y = 0 \Rightarrow Y^3 - 2Y = 0$$

obteniendo puntos críticos $(0, 0)$, entonces

$$X(X^2 - 2) = 0 \Rightarrow X_1 = 0 \text{ e implica } (\sqrt{2}, -\sqrt{2}), X_{2-3} = \pm\sqrt{2} \text{ e implica } (-\sqrt{2}, \sqrt{2}).$$

$$\text{Entonces, } D_1(X_0, Y_0) = f_{xx}(X_0, Y_0) = 12(0)^2 - 4 = -4,$$

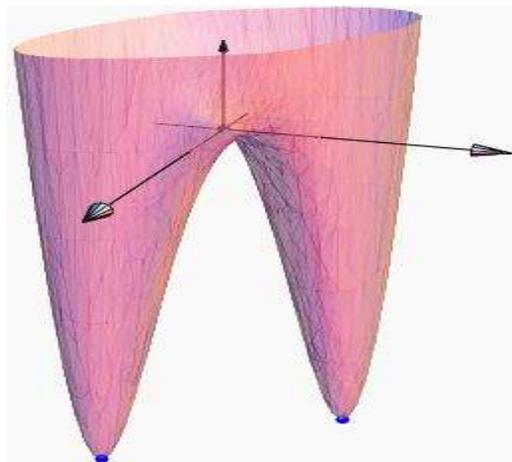
$$D_2(X_0, Y_0) = f_{xx}(X, Y) * f_{yy}(X, Y) - [f_{xy}(X, Y)]^2 = (12(0)^2 - 4)(12(0)^2 - 4) - [4]^2 = 16 - 16 = 0 \text{ y } (X, Y) = X^4 + Y^4 - 2X^2 + 4XY - 2Y^2$$

$$\Rightarrow Z = (\sqrt{2})^4 + (\sqrt{2})^4 - 2(\sqrt{2})^2 + 4(\sqrt{2} * \sqrt{2}) - 2(\sqrt{2})^2 = -8:$$

Tabla 2.1 Cálculo y clasificación con puntos críticos

(X_0, Y_0)	$D_1(X_0, Y_0) = f_{xx}(X_0, Y_0)$	$D_2(X_0, Y_0) = f_{xx}(X, Y) * f_{yy}(X, Y) - [f_{xy}(X, Y)]^2$	Clasificación
(0, 0)	-4 (-)	0	Criterio no decide, pero por figura es punto silla o ensilla
$(\sqrt{2}, -\sqrt{2})$	20 (+)	384 (+)	Mínimo local $(\sqrt{2}, -\sqrt{2}, -8)$
$(-\sqrt{2}, \sqrt{2})$	20 (+)	384 (+)	Mínimo local $(-\sqrt{2}, \sqrt{2}, -8)$

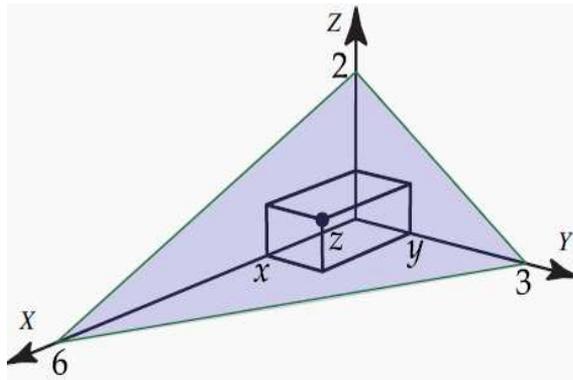
Figura 3.3 Representación gráfica de los puntos críticos



Fuente: (Mora, W. F. 2016).

2. Calcule el volumen de la caja rectangular más grande que esté en I octante, con tres caras en Π s coordenados y un vértice en $\Pi: X + 2Y + 3Z = 6$:

Figura 3.4 Volumen de la caja rectangular



Fuente: (Mora, W. F. 2016).

El objetivo es maximizar el volumen tal que $V = X * Y * Z$ 1) y $X + 2Y + 3Z = 6$

$$\Rightarrow Z = \frac{6 - X - 2Y}{3}$$

reemplazando esta ecuación en 1) $\Rightarrow V = X * Y * \left(2 - \frac{X}{3} - \frac{2}{3}Y\right) =$

$$2XY - \frac{X^2Y}{3} - \frac{2}{3}XY^2 \quad X, Y > 0.$$

Donde: $V_x = 2Y - \frac{2}{3}XY - \frac{2}{3}Y^2$

$$\Rightarrow V_{xx} = -\frac{2}{3}Y, \quad V_y = 2X - \frac{X^2}{3} - \frac{4}{3}XY$$

$$\Rightarrow V_{yy} = -\frac{4}{3}X \quad \text{y} \quad V_{xy} = 2 - \frac{2}{3}X - \frac{4}{3}Y.$$

Los puntos críticos $\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \Rightarrow \begin{cases} 2Y \left(1 - \frac{X}{3} - \frac{Y}{3}\right) = 0 \quad (1) \\ X \left(2 - \frac{X}{3} - \frac{4}{3}Y\right) = 0 \quad (2) \end{cases} \Rightarrow$

$$\begin{cases} 3 - X - Y = 0 = 0 \quad (1) \\ \underline{6 - X - 4Y = 0} \quad (-1) \\ -3 + 3Y \Rightarrow Y = 1 \text{ y } X = 2 \end{cases}.$$

Ahora bien,

$$D_1(X_0, Y_0) = V_{xx}(2, 1) = -\frac{2}{3}(1) = -\frac{2}{3}(1) = -\frac{2}{3} < 0 \quad (-),$$

$$D_2(X_0, Y_0) = V_{xx}(X, Y) * V_{yy}(X, Y) - [V_{xy}(X, Y)]^2$$

$$\begin{aligned} \Rightarrow D_2(2, 1) &= -\frac{2}{3}Y * -\frac{4}{3}X - \left[2 - \frac{2}{3}X - \frac{4}{3}Y\right]^2 = \left(\frac{8}{9}XY\right) - \\ \left[2 - \frac{2}{3}X - \frac{4}{3}Y\right]^2 &= \left(\frac{8}{9}2 * 1\right) - \left[2 - \frac{2}{3}(2) - \frac{4}{3}(1)\right]^2 = \frac{16}{9} - \left(\frac{2}{3}\right)^2 = \frac{4}{3} > \\ 0 \text{ y } Z &= 2 - \frac{2}{3} - \frac{2}{3}(1) = \frac{2}{3}. \end{aligned}$$

Por lo tanto, en P(2, 1) el volumen es máximo y $V_{\text{Máx}} = \frac{4}{3} U^3$.

3. Calcule y clasifique con puntos críticos de $f(X, Y) = 6XY - 2X^2Y - 3XY^2$.

$$f_x = 6Y - 4XY - 3Y^2$$

$$\Rightarrow f_{xx} = -4Y, f_y = 6X - 2X^2 - 6XY$$

$$\Rightarrow f_{yy} = -6X \text{ y } f_{xy} = 6 - 4X - 6Y.$$

Los puntos críticos

$$\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \Rightarrow \begin{cases} Y(6 - 4X - 3Y) = 0 \quad (1) \Rightarrow Y = 2 - \frac{4}{3}X \quad (1') \\ \underline{X(6 - 2X - 6Y) = 0} \quad (2) \\ -9Y = 0 \Rightarrow Y = 0 \quad (3) \end{cases}$$

Por lo tanto, el primer punto crítico es $(0, 0)$.

Sustituyendo 3) en 2)

$6 - 2X - 6(0) = 0 \Rightarrow X = 3$ se obtiene el segundo punto crítico $(3, 0)$, el tercer punto crítico se estima reemplazando valor $X = 0$ en 1') tal que $Y = 2 - \frac{4}{3}(0) \Rightarrow Y = 2$ y,

finalmente, se estima el cuarto punto crítico sustituyendo 1') en

2) $6 - 2X - 6\left(2 - \frac{4}{3}X\right) = 0 \Rightarrow X = 1$ y reemplazando este valor en

2) $6 - 4(1) - 3Y = 0 \Rightarrow Y = \frac{2}{3}$.

Entonces, $D_1(X_0, Y_0) = f_{xx}(X_0, Y_0) = -4(0) = 0$, $D_2(X_0, Y_0) =$

$f_{xx}(X, Y) * f_{yy}(X, Y) - [f_{xy}(X, Y)]^2 = (-4 * 0)(-6 * 0) -$

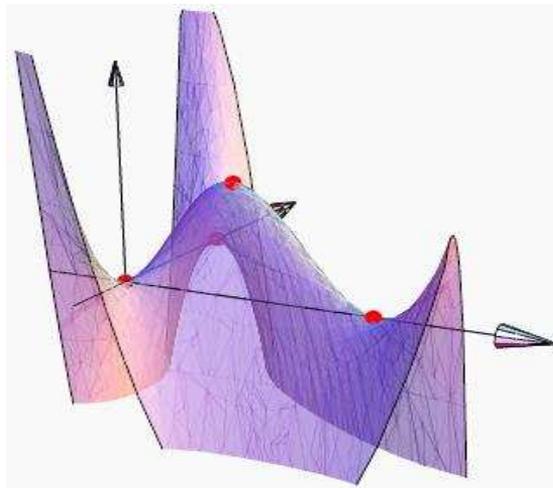
$[6 - 4(0) - 6(0)]^2 = -36$ y $(X, Y) = 6XY - 2X^2Y - 3XY^2 \Rightarrow Z =$

$6(0)(0) - 2(0)^2(0) - 3(0)(0)^2 = 0:$

Tabla 2.2 Cálculo y clasificación con puntos críticos

(X_0, Y_0)	$D_1(X_0, Y_0)$ $= f_{xx}(X_0, Y_0)$	$D_2(X_0, Y_0)$ $= f_{xx}(X, Y) * f_{yy}(X, Y)$ $- [f_{xy}(X, Y)]^2$	Clasificación
(0, 0)	0	-36	Criterio no decide, pero por figura es punto silla o ensilla (0, 0, 0)
(3, 0)	0	-36	Criterio no decide, pero por figura es punto silla o ensilla (3, 0, 0)
(0, 2)	-8	-36	Criterio no decide, pero por figura es punto silla o ensilla (0, 2, 0)
$(1, \frac{2}{3})$	$-\frac{8}{3}$	12	Máximo local $(1, \frac{2}{3}, \frac{4}{3})$

Figura 3.5 Representación gráfica de los puntos críticos



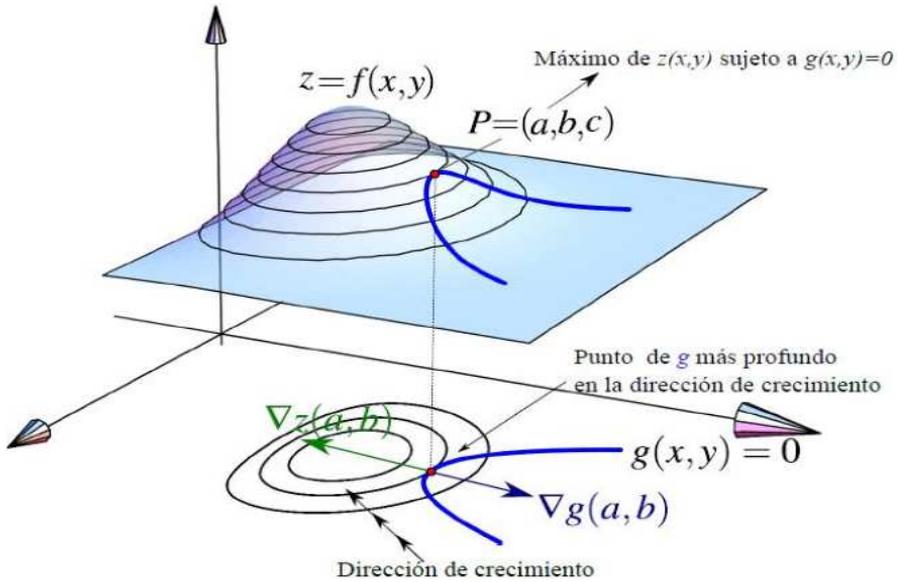
Fuente: (Mora, W. F. 2016).

3.2. Extremos con Restricciones

Multiplicadores de Lagrange (λ_{Lambda}) supone que el investigador quiere hallar los máximos y mínimos relativos de $z = f(X, Y)$ sujeto a restricciones $g = f(X, Y) = 0$.

Significa que la función solo podrá ser evaluada en puntos (X, Y) que estén en la curva de nivel $g = f(X, Y) = 0$; es decir, $f(X, Y)$ está sujeta o restringida a $g = f(X, Y) = 0$. Una manera de resolver este problema es con análisis geométrico:

Figura 3.6 Máximo local



Fuente: (Mora, W. F. 2016).

En las cercanías de un máximo local se desplaza sobre g en dirección de crecimiento f hasta el punto más profundo que puede alcanzarse sobre g en esta dirección. Este punto puede ser el máximo local con restricciones que se busca.

Si $P(a, b, c)$ es un punto que se quiere determinar con una ecuación se «viaja a este punto más profundo» atravesando curvas de nivel, entonces la «última» curva de nivel será $z = c$ tangente a g en P , si P no es un punto terminal de g .

Si estas curvas son tangentes indica que sus gradientes son \parallel o $\nabla z_{(a,b)} = \lambda \nabla g_{(a,b)}$ ($\lambda = 0 \Leftrightarrow$ extremos locales con restricciones coincide con sin restricciones), usada en la determinación de P.

Paralelamente, el análisis es similar en determinaciones de mínimos locales, a diferencia que se desplaza sobre g en dirección de decrecimiento hasta el punto más profundo que se puede alcanzar.

3.3. Teorema de Multiplicadores de Lagrange

El Teorema de Multiplicadores Lagrange de Condición de Primer Orden indica que sea $U \subseteq \mathbb{R}^2$ conjunto abierto, $f, g: U \rightarrow \mathbb{R}$ funciones C^1 y X^* extremo local de f en conjunto $D = \{X \in U | g(X) = 0\} \Rightarrow \nabla g(X^*) \neq (0, 0)$ existe $\lambda \in \mathbb{R}$ (puede ser 0) $\Rightarrow \nabla f(X^*) - \lambda \nabla g(X^*) = (0, 0)$.

Tal que, los extremos locales X^* de f sujetos a la restricción $g(X, Y) = 0$ y $\nabla g(X^*) \neq (0, 0)$ son puntos críticos de la función "lagrangiana" $L(X, Y, \lambda) = f(X, Y) - \lambda g(X, Y)$ aunque no necesariamente viceversa, también puede suceder que algunos puntos críticos de L no sean extremos locales de f sujetos a la restricción $g(X, Y) = 0$.

En tres variables se pueden hallar puntos críticos del problema de optimización como soluciones del sistema:

$$L(X, Y, Z, \lambda) = f(X, Y, Z) - \lambda g(X, Y, Z)$$

3.3.1. Método de Multiplicadores de Lagrange

El Método de Multiplicadores de Lagrange con una restricción indica que para minimizar o maximizar $f(X_1, X_2, X_3, X_4, \dots, X_n)$ sujeta a la condición $g(X_1, X_2, X_3, X_4, \dots, X_n) = 0$, se busca puntos críticos de

$$L(X_1, X_2, X_3, X_4, \dots, X_n, \lambda) = f(X_1, X_2, X_3, X_4, \dots, X_n) - \lambda g(X_1, X_2, X_3, X_4, \dots, X_n) = 0.$$

Tal que, para hallar puntos críticos de $L(X_1, X_2, X_3, X_4, \dots, X_n, \lambda)$ se resuelve el sistema:

$$\left\{ \begin{array}{l} LX_1 = 0 \\ LX_2 = 0 \\ LX_3 = 0 \\ LX_4 = 0 \\ \vdots \\ LX_n = 0 \\ g(X_1, X_2, X_3, X_4, \dots, X_n) = 0 \end{array} \right.$$

3.3.2. Criterios de Clasificación

Los Criterios de Clasificación se usan para determinar si los puntos críticos son máximos, mínimos o ninguno de ambos, se podría recurrir al criterio de Hessiana Orlada. No obstante, en los problemas siguientes, los puntos críticos se pueden clasificar de manera directa, usando la geometría del problema o una comparación.

Ejemplos:

1. Minimizar $Z = X^2 + Y^2$ sujeto a $X - Y = 0$.

$$\text{Sea } L(X, Y, \lambda) = f(X, Y) - \lambda g(X, Y) = X^2 + Y^2 - \lambda(X - Y) = X^2 +$$

$$Y^2 - \lambda X + \lambda Y \Rightarrow \begin{cases} L_X = 0 \\ L_Y = 0 \\ L_\lambda = 0 \end{cases} \Rightarrow \left\{ \begin{array}{l} 2X - \lambda = 0 \rightarrow X = \frac{\lambda}{2} \\ 2Y - \lambda = 0 \rightarrow Y = \frac{\lambda}{2} \\ -X + Y = 0 \rightarrow X = Y \Rightarrow \frac{\lambda}{2} = \frac{\lambda}{2} \end{array} \right.$$

$\Rightarrow \therefore X = 0, Y = 0$ y $\lambda = 0$ e indica que el máximo con restricciones coincide con el mínimo local de Z .

2. Determine 3 números \mathbb{R}^+ X, Y, Z, cuya suma sea 10 y su producto sea máximo.

Maximizar $P = XYZ$ con restricción $X + Y + Z = 10$ tal que sea
 $L(X, Y, \lambda) = f(X, Y) - \lambda g(X, Y) = XYZ - \lambda(X + Y + Z - 10) = XYZ - \lambda X - \lambda Y - \lambda Z + 10\lambda$.

Por lo tanto, sus puntos críticos son:

$$\begin{aligned}
 L_X = 0 \\
 L_Y = 0 \\
 L_Z = 0 \\
 L_\lambda = 0
 \end{aligned}
 \Rightarrow
 \left\{
 \begin{array}{l}
 YZ - \lambda = 0 \\
 XZ - \lambda = 0 \\
 XY - \lambda = 0 \\
 \underbrace{-X + Y + Z - 10}_{g(X, Y, Z) = 0}
 \end{array}
 \Rightarrow
 \left.
 \begin{array}{l}
 \lambda = YZ \\
 \lambda = XZ \\
 \lambda = XY \\
 \therefore 3X = 10 \Rightarrow X = Y = Z = \frac{10}{3}
 \end{array}
 \right\} X$$

$$= Y = Z \in \mathbb{R}^+$$

El producto máximo estará dado por $P = \frac{10}{3}(X) * \frac{10}{3}(Y) * \frac{10}{3}(Z) = \frac{1000}{27}$.

3. Determine el valor máximo y mínimo de $f(X, Y) = X^2 + Y^2$ sujeto a $X^4 + Y^4 = 1$.

$$L(X, Y, \lambda) = f(X, Y) - \lambda g(X, Y) = X^2 + Y^2 - \lambda(X^4 + Y^4 - 1).$$

Siendo sus puntos críticos:

$$\begin{aligned}
 L_X = 0 \\
 L_Y = 0 \\
 L_\lambda = 0
 \end{aligned}
 \Rightarrow
 \left\{
 \begin{array}{l}
 2X - 4\lambda X^3 = 0 \rightarrow 2X(1 - 2\lambda X^2) \Rightarrow X = 0 \\
 2Y - 4\lambda Y^3 = 0 \rightarrow 2Y(1 - 2\lambda Y^2) \Rightarrow Y = 0 \\
 -X^4 - Y^4 = 0 \rightarrow X^4 + Y^4 - 1 = 0 \Rightarrow 1 \neq 0
 \end{array}
 \right.$$

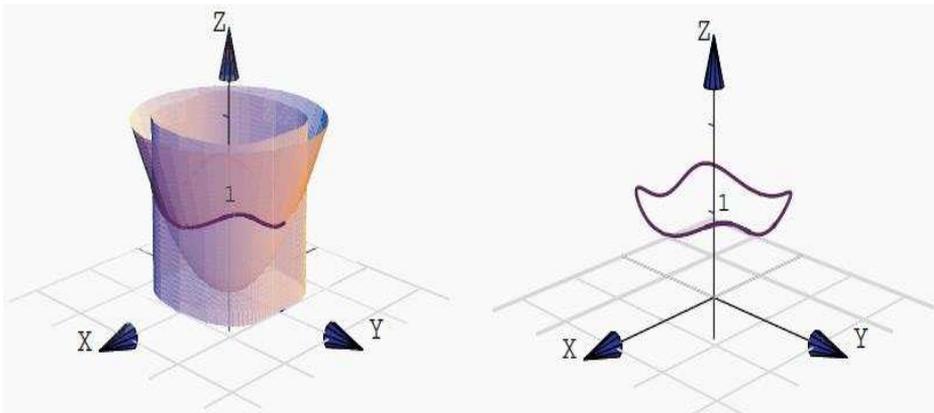
Por lo tanto, no hay puntos críticos. No obstante, se consideran los siguientes casos para anular las tres ecuaciones:

Tabla 2.3 Criterios de Clasificación

Puntos críticos	Ecuación	Resultados	Interpretación
$(X = 0, Y = 0)$	$X^4 + Y^4 - 1 = 0$	$1 \neq 0$	No hay puntos críticos
$(X = 0, Y \neq 0)$	$1 - 2\lambda Y^2$	$1 - 2\lambda Y^2 \Rightarrow Y^2 = \frac{1}{2\lambda}$ y $\lambda = \frac{1}{2Y^2} = \frac{1}{2}$ Sustituyendo en $X^4 + Y^4 - 1 = 0 \Rightarrow Y = \pm 1$	Los puntos críticos son $(0, 1), (0, -1)$ con $\lambda = \frac{1}{2}$, pero sin valor de Z.
$(X \neq 0, Y = 0)$	$1 - 2\lambda X^2$	$1 - 2\lambda X^2 \Rightarrow X^2 = \frac{1}{2\lambda}$ y $\lambda = \frac{1}{2X^2} = \frac{1}{2}$ Sustituyendo en $X^4 + Y^4 - 1 = 0 \Rightarrow X = \pm 1$	Los puntos críticos son $(1, 0), (-1, 0)$ con $\lambda = \frac{1}{2}$, pero sin valor de Z.
$(X \neq 0, Y \neq 0)$	$1 - 2\lambda X^2$	$1 = 2\lambda X^2$ elevando al cuadrado ambos términos $1^2 = 2^2\lambda^2 X^4 \Rightarrow X^4 = \frac{1}{4\lambda^2}$.	Reemplazando los valores de X^4 y Y^4 en ecuación $X^4 + Y^4 - 1 = 0$ se tiene que $\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 1 \Rightarrow \lambda^2 = \frac{1}{2}$. Entonces, $X^4 = \frac{1}{4(\frac{1}{2})} = \pm^4\sqrt{\frac{1}{2}}$, $Y^4 = \frac{1}{4(\frac{1}{2})} = \pm^4\sqrt{\frac{1}{2}}$, $\lambda^2 = \frac{1}{2}$ y sin valor de Z.
	$1 - 2\lambda Y^2$	$1 = 2\lambda Y^2$ elevando al cuadrado ambos términos $1^2 = 2^2\lambda^2 Y^4 \Rightarrow Y^4 = \frac{1}{4\lambda^2}$.	

Para clasificar los puntos «empíricamente» se evaluará la función f y para visualizar la situación se dibuja la curva de intersección entre la superficie $Z = X^2 + Y^2$ y la superficie generada por curva $X^4 + Y^4 = 1$:

Figura 3.7 Máximos Relativos



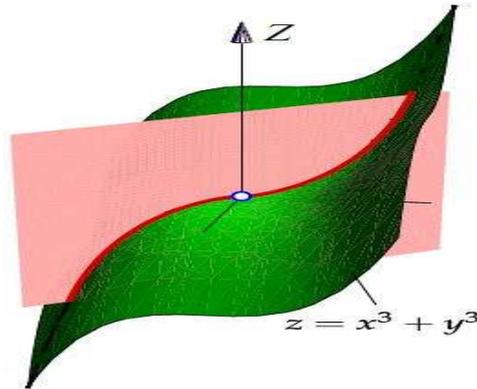
Fuente: (Mora, W. F. 2016).

Se obtienen cuatro puntos máximos relativos, $(\pm \frac{1}{\sqrt[4]{2}}, \pm \frac{1}{\sqrt[4]{2}}, \frac{2}{\sqrt{2}})$ y cuatro puntos mínimos relativos, $(\pm 1, 0, 1)$ y $(0, \pm 1, 1)$.

Cuando las condiciones de primer orden fallan sucede cuando los puntos críticos L no necesariamente son solución del problema de optimización, a pesar de que el método de multiplicadores de Lagrange es muy eficiente, dan origen a L. Dicho teorema se emplea solo cuando se indican las condiciones necesarias.

Ejemplo:

Considere el problema minimizar $f(X, Y) = X^3 + Y^3$ sujeto a la restricción $g = X - Y = 0$:

Figura 3.8 Representación gráfica de maximización

Fuente: (Mora, W. F. 2016).

Tal que $(0, 0)$ no es máximo ni mínimo local de f en D , pues $\forall \varepsilon > 0$, $(\varepsilon, \varepsilon) \in D$ y $(-\varepsilon, -\varepsilon) \in D$, pero $f(0, 0) = 0 > f(-\varepsilon, -\varepsilon) = -2\varepsilon^3$ y $f(0, 0) = 0 < f(\varepsilon, \varepsilon) = 2\varepsilon^3$.

Sin embargo, $(0, 0)$ satisface $\nabla g(0, 0) = (1, -1) \neq (0, 0)$, siendo la única solución con $\lambda = 0$ del sistema $\nabla L(X, Y, \lambda) = 0$:

$$\begin{aligned} L_X &= 3X^2 - \lambda = 0 \\ L_Y &= 3Y^2 - \lambda = 0 \\ L_\lambda &= X - Y = 0 \end{aligned}$$

Cuando ∇g se anula, el método de multiplicadores de Lagrange requiere que ∇g no se anule en puntos críticos de f sobre D para que el conjunto de puntos críticos de L contenga al conjunto de puntos críticos de f sobre D .

Si $\nabla g(X)$ se anula podrían pasar varios resultados de cuidado:

- El sistema no tiene soluciones en \mathbb{R} y puntos críticos de L no detectan el mínimo local $(1, 0, 1)$.
- El sistema no tiene solución, pues el método de multiplicadores de Lagrange no detecta el óptimo.

- Varias soluciones para un mismo λ y no detecta el óptimo.
- Tiene infinitas soluciones con $\lambda \in \mathbb{R}$.

El criterio de clasificación para tres variables o más no siempre funciona, se debe recurrir a otras técnicas y se aplica cuando hay restricciones.

La forma cuadrática general, con n variables es $F(X_1, X_2, X_3, X_4, \dots, X_n) = a_{11}X_1^2 + a_{22}X_2^2 + a_{33}X_3^2 + a_{44}X_4^2 + \dots + a_{nn}X_n^2 + 2a_{12}X_1X_2 + 2a_{13}X_1X_3 + 2a_{14}X_1X_4 + \dots + 2a_{1n}X_1X_n + \dots + 2a_{(n-1)n}X_{(n-1)}X_n$ aunque para dos y tres variables se tiene $F(X, Y) = aX^2 + 2bXY + cY^2$ y $F(X, Y, Z) = aX^2 + bY^2 + cZ^2 + 2a_1XY + 2a_2XZ + 2a_3YZ$.

La forma matricial $2a_{ij}X_iX_j = a_{ij}X_iX_j + a_{ji}X_jX_i$ tal que $F(X_1, X_2, X_3, X_4, \dots, X_n)$ se puede reescribir como $F(X_1, X_2, X_3, X_4, \dots, X_n) = a_{11}X_1^2 + a_{12}X_1X_2 + a_{13}X_1X_3 + a_{14}X_1X_4 + \dots + a_{1n}X_1X_n + a_{21}X_2X_1 + a_{22}X_2^2 + \dots + a_{2n}X_2X_n + \dots + a_{n1}X_nX_1 + a_{n2}X_nX_2 + a_{22}X_2^2 + \dots + a_{nn}X_n^2 \Rightarrow F(X_1, X_2, X_3, X_4, \dots, X_n) = (X_1, X_2, X_3, X_4, \dots, X_n)A(X_1, X_2, X_3, X_4, \dots, X_n)^T = X * A * X^T \Rightarrow A_{(\text{Simétrica})} = (a_{ij})_{n \times n}$

Ejemplos:

1. Sea $F(X, Y, Z) = X^2 + 4XY + 2XZ - 7Y^2 - 6YZ + 5Z^2$ tal que al ser una expresión cuadrática implica una matriz 3 (hileras)*3 (columnas)

$$\begin{aligned}
 F(X, Y, Z) &= \begin{pmatrix} \text{Coeficiente } \frac{X^2}{2} & \text{Coeficiente } \frac{XY}{2} & \text{Coeficiente } \frac{XZ}{2} \\ \text{Coeficiente } \frac{YX}{2} & \text{Coeficiente } \frac{Y^2}{2} & \text{Coeficiente } \frac{YZ}{2} \\ \text{Coeficiente } \frac{ZX}{2} & \text{Coeficiente } \frac{ZY}{2} & \text{Coeficiente } \frac{Z^2}{2} \end{pmatrix} \\
 &= (X, Y, Z) \begin{pmatrix} 1 & 2 & 1 \\ 2 & -7 & -3 \\ 1 & -3 & 5 \end{pmatrix} (X, Y, Z)^T
 \end{aligned}$$

El estudio algebraico de las formas cuadráticas está centrado en determinar si una forma tiene siempre el mismo signo; es decir, si las formas $F(X, Y, Z)$ pueden ser (+) o (-):

- $F(X_1, X_2, X_3, X_4, \dots, X_n)$ se dice definida positiva si $F(X_1, X_2, X_3, X_4, \dots, X_n) > 0, \forall X_1, X_2, X_3, X_4, \dots, X_n$ no todos = 0.
- $F(X_1, X_2, X_3, X_4, \dots, X_n)$ se dice definida negativa si $F(X_1, X_2, X_3, X_4, \dots, X_n) < 0, \forall X_1, X_2, X_3, X_4, \dots, X_n$ no todos = 0.

Teorema: $F(X, Y) = aX^2 + 2bXY + cY^2 > 0, \forall X, Y$ no todos = 0 $\Leftrightarrow a > 0, \text{Det} (|A|) \begin{pmatrix} a & b \\ b & c \end{pmatrix} > 0$ (+) y $F(X, Y) = aX^2 + 2bXY + cY^2 < 0, \forall X, Y$ no todos = 0 $\Leftrightarrow a < 0, \text{Det} (|A|) \begin{pmatrix} a & b \\ b & c \end{pmatrix} > 0$, (-), respectivamente.

Este teorema se puede probar complementando cuadrados: $F(X, Y) = c(X + \frac{b}{a}Y)^2 + (\frac{ac-b^2}{a})Y^2 \Rightarrow F(X, Y) > 0, \forall X, Y \text{ no todos } = 0 \Leftrightarrow a > 0 \text{ y } (\frac{ac-b^2}{a}) > 0$; es decir, $a > 0$ y $ac - b^2 > 0$.

$$\text{Generalizando: } D_n = |A| \rightarrow D_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots & a_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \dots & a_{nn} \end{vmatrix} \text{ y}$$

$$D_i \text{ puede ser definido como } D_i = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1i} \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2i} \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3i} \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots & a_{4i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} & \dots & a_{ii} \end{vmatrix} \text{ tal}$$

que D_i son los menores principales de D_n .

Nota: Dada una matriz cuadrada A , un menor principal es el determinante de una submatriz cuadrada de A en que los elementos de su diagonal principal pertenecen a la diagonal principal de la matriz A . Ejemplo: $A =$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 3 \\ 0 & 3 & -2 \end{pmatrix} \text{ siendo los menores principales de orden 2 los}$$

determinantes, eliminando una columna y una fila, $\begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ 0 & -2 \end{vmatrix}$ y $\begin{vmatrix} 0 & 3 \\ 3 & -2 \end{vmatrix}$.

Teorema: $F(X_1, X_2, X_3, X_4, \dots, X_n)$ es definida positiva (+) si $D_1 > 0, D_2 > 0, D_3 > 0, \dots, D_n > 0$ y definida negativa (-) si $D_i > 0$ para i par y $D_i < 0$ para i impar.

2. Sea $F(X, Y, Z) = 2X^2 - 4XY + 4XZ + 6Y^2 - 4YZ + 8Z^2$ tal que al ser una expresión cuadrática implica una matriz 3 (hileras)*3 (columnas):

$$F(X, Y, Z) = \begin{pmatrix} \text{Coeficiente } X^2 & \text{Coeficiente } \frac{XY}{2} & \text{Coeficiente } \frac{XZ}{2} \\ \text{Coeficiente } \frac{XY}{2} & \text{Coeficiente } Y^2 & \text{Coeficiente } \frac{YZ}{2} \\ \text{Coeficiente } \frac{XZ}{2} & \text{Coeficiente } \frac{YZ}{2} & \text{Coeficiente } Z^2 \end{pmatrix} \Rightarrow$$

$$D_3 = \begin{pmatrix} 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \end{pmatrix} = 48 > 0 \text{ que puede estimarse mediante estos}$$

métodos:

Asociados a matrices 3x3: Si A es una matriz de 3x3 su determinante se estima

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = [(a_{11}a_{22}a_{33}) + (a_{23}a_{12}a_{31}) + (a_{21}a_{32}a_{13})] - [(a_{13}a_{22}a_{31}) + (a_{21}a_{12}a_{33}) + (a_{23}a_{32}a_{11})]$$

por

$$\Rightarrow \text{Det}(A) = \begin{vmatrix} 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \end{vmatrix} =$$

$$[(2 * 6 * 8) + (-2 * -2 * 2) + (-2 * -2 * 2)] - [(2 * 6 * 2) + (-2 * -2 * 8) + (-2 * -2 * 2)] = 48$$

Estrella de David: Método usado comúnmente, que no es muy conocido, pero tiene gran aplicación en cálculos mentales.

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = [(a_{11}a_{22}a_{33}) +$$

$$(a_{23}a_{12}a_{31}) + (a_{21}a_{32}a_{13})] - [(a_{13}a_{22}a_{31}) + (a_{21}a_{12}a_{33}) +$$

$$(a_{23}a_{32}a_{11})] \Rightarrow \text{Det}(A) = \begin{pmatrix} 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \end{pmatrix} = [(2 * 6 * 8) +$$

$$(-2 * -2 * 2) + (-2 * -2 * 2)] - [(2 * 6 * 2) + (-2 * -2 * 8) + (-2 * -2 * 2)] = 48$$

Método por menores, determinantes o regla de Cramer

Se parte de un sistema de ecuaciones

X	Y	Z	Término Independiente	
2X -	2Y +	2Z	= -3	. Enseguida, se estima
-2X +	6Y -	2Z	= 13	
2X -	2Y +	8Z	= 8	

el determinante del sistema de ecuaciones

$$|A| = \Delta_s = \begin{vmatrix} X & Y & Z \\ 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \end{vmatrix}$$

Aplicando la Regla de Sarrus

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} \quad \text{[7]}$$

o

$$\text{Det}(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{vmatrix} \quad \text{[7]}$$

se obtenga el determinante de la matriz.

A continuación, se calculan los determinantes de las incógnitas:

$$\Delta_x = \begin{vmatrix} \text{Término Independiente} & Y & Z \\ -3 & -2 & 2 \\ 13 & 6 & -2 \\ 8 & -2 & 8 \end{vmatrix}$$

$$\Rightarrow |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} \therefore$$

$$\Rightarrow \begin{vmatrix} \text{Término Independiente} & Y & Z \\ -3 & -2 & 2 \\ 13 & 6 & -2 \\ 8 & -2 & 8 \\ -3 & -2 & 2 \\ 13 & 6 & -2 \end{vmatrix} =$$

$$\begin{aligned} &= [(a_{11}a_{22}a_{33}) + (a_{21}a_{32}a_{13}) + (a_{31}a_{12}a_{23})] \\ &- [(a_{13}a_{22}a_{31}) + (a_{23}a_{32}a_{11}) + (a_{33}a_{12}a_{21})] = \\ &= [(-3 * 6 * 8) + (13 * -2 * 2) + (8 * -2 * -2)] \\ &- [(2 * 6 * 8) + (-2 * -2 * -3) + (8 * -2 * 13)] \\ &= -40 \end{aligned}$$

$$\Delta_Y = \begin{vmatrix} X & \text{Término Independiente} & Z \\ 2 & -3 & 2 \\ -2 & 13 & -2 \\ 2 & 8 & 8 \end{vmatrix} \Rightarrow \therefore$$

$$\Rightarrow \text{Det}(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \end{vmatrix}$$

$$= \begin{vmatrix} X & \text{Término Independiente} & Z \\ 2 & -3 & 2 & 2 & -3 \\ -2 & 13 & -2 & -2 & 13 \\ 2 & 8 & 8 & 8 & 2 & 8 \end{vmatrix}$$

$$\begin{aligned} &= [(a_{11}a_{22}a_{33}) + (a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32})] \\ &- [(a_{12}a_{21}a_{33}) + (a_{11}a_{23}a_{32}) + (a_{13}a_{22}a_{31})] \end{aligned}$$

$$\begin{aligned}
 &= [(2 * 13 * 8) + (-3 * -2 * 2) + (2 * -2 * 8)] \\
 &\quad - [(2 * 13 * 2) + (2 * -2 * 8) + (-3 * -2 * 8)] \\
 &= 120
 \end{aligned}$$

$$\Delta_z = \begin{vmatrix} X & Y & \text{Término Independiente} \\ 2 & -2 & -3 \\ -2 & 6 & 13 \\ 2 & -2 & 8 \end{vmatrix}$$

$$\Rightarrow |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \therefore$$

$$\Rightarrow \begin{vmatrix} X & Y & \text{Término Independiente} \\ 2 & -2 & -3 \\ -2 & 6 & 13 \\ 2 & -2 & 8 \\ 2 & -2 & -3 \\ -2 & 6 & 13 \end{vmatrix} =$$

$$\begin{aligned}
 &= [(a_{11}a_{22}a_{33}) + (a_{21}a_{32}a_{13}) + (a_{31}a_{12}a_{23})] \\
 &\quad - [(a_{13}a_{22}a_{31}) + (a_{23}a_{32}a_{11}) + (a_{33}a_{12}a_{21})] = \\
 &= [(2 * 6 * 8) + (-2 * -2 * -3) + (2 * -2 * 13)] \\
 &\quad - [(-3 * 6 * 2) + (13 * -2 * 2) + (8 * -2 * -2)] \\
 &= 88
 \end{aligned}$$

Con base en lo anterior, el valor de las literales o incógnitas son

$$X = \frac{\Delta_x}{\Delta_s} = -\frac{40}{48} = -0.833\bar{3}, \quad Y = \frac{\Delta_y}{\Delta_s} = \frac{120}{48} = 2.5 \quad \text{y} \quad Z = \frac{\Delta_z}{\Delta_s} = \frac{88}{48} =$$

1.833 $\bar{3}$, respectivamente.

Regla de Sarrus

Este método repite las dos primeras filas o columnas de la matriz debajo de la misma o del lado derecho, respectivamente, de manera

que queden cinco filas o columnas. Después sumar los productos de las diagonales descendentes (en línea continua) y sustraer los productos de las diagonales ascendentes (en trazos).

$$[A] = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} \quad \text{?}$$

$$= [(a_{11}a_{22}a_{33}) + (a_{21}a_{32}a_{13}) + (a_{31}a_{12}a_{23})] \\ - [(a_{13}a_{22}a_{31}) + (a_{23}a_{32}a_{11}) + (a_{33}a_{12}a_{21})] \\ \Rightarrow \therefore \text{Det}(A)$$

$$= \begin{vmatrix} 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \\ 2 & -2 & 2 \\ -2 & 6 & -2 \end{vmatrix}$$

$$= [(2 * 6 * 8) + (-2 * -2 * 2) + (2 * -2 * -2)] \\ - [(2 * 6 * 2) + (-2 * -2 * 2) \\ + (8 * -2 * -2)] = 48$$

$$\text{Det}(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{vmatrix} \quad \text{?}$$

$$= [(a_{11}a_{22}a_{33}) + (a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32})] \\ - [(a_{12}a_{21}a_{33}) + (a_{11}a_{23}a_{32}) + (a_{13}a_{22}a_{31})]$$

$$\Rightarrow \therefore \text{Det}(A) = \begin{vmatrix} 2 & -2 & 2 & 2 & -2 \\ -2 & 6 & -2 & -2 & 6 \\ 2 & -2 & 8 & 2 & -2 \end{vmatrix} \\ = [(2 * 6 * 8) + (-2 * -2 * 2) \\ + (2 * -2 * -2)] \\ - [(-2 * -2 * 8) + (2 * -2 * -2) \\ + (2 * 6 * 2)] = 48$$

El cálculo del determinante de esta matriz 3x3, usando *software* diferentes, es:

Tabla 2.4 Cálculo del determinante de una matriz 3x3

<i>Software</i>	Siglas	Programación	Resultado
<i>Matrix Laboratory</i>	MatLab	A=[2 -2 2; -2 6 -2; 2 -2 8] det(A)	48
<i>Statistician R</i>	R	clm1<-c(2, -2, 2) clm2<-c(-2, 6, -2) clm3<-c(2, -2, 8) ma<-cbind(clm1, clm2, clm3) ma det(ma)	
<i>Statistical Analysis System</i>	SAS	Proc IML; A = {2 -2 2, -2 6 -2, 2 -2 8}; Print A; d = det(A); Print d;	
<i>Wolfram System Modeler</i>	Wolfram ®	A={{2,-2,2},{-2,6,-2},{2,-2,8}} MatrixForm[A] Det[A]	

Formas cuadráticas con restricciones lineales suponen que $F(X_1, X_2, X_3, X_4, \dots, X_n)$ está restringida a que sus variables cumplan la restricción lineal $\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \dots + \alpha_n X_n = 0$.

Tal que con el determinante D_i se obtiene un nuevo determinante Orlado $\bar{D}_i, i \geq 2$:

$$\bar{D}_i = |A| \text{ ó Det Orlado } \begin{pmatrix} 0 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \dots & \alpha_i \\ \alpha_1 & a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1i} \\ \alpha_2 & a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2i} \\ \alpha_3 & a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3i} \\ \alpha_4 & a_{41} & a_{42} & a_{43} & a_{44} & \dots & a_{4i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \alpha_i & a_{i1} & a_{i2} & a_{i3} & a_{i4} & \dots & a_{ii} \end{pmatrix} \Rightarrow$$

$$\bar{D}_n = |A| \text{ ó Det Orlado } \begin{pmatrix} 0 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \dots & \alpha_n \\ \alpha_1 & a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\ \alpha_2 & a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\ \alpha_3 & a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\ \alpha_4 & a_{41} & a_{42} & a_{43} & a_{44} & \dots & a_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \alpha_n & a_{n1} & a_{n2} & a_{n3} & a_{n4} & \dots & a_{nn} \end{pmatrix}$$

Teorema: Sea $F(X_1, X_2, X_3, X_4, \dots, X_n)$ restringida a que sus variables cumplan la relación lineal $\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \dots + \alpha_n X_n = 0$ es definida positiva si $\bar{D}_2 < 0, \bar{D}_3 < 0, \bar{D}_4 < 0, \dots, \bar{D}_n < 0 (+)$ y definida negativa si $\bar{D}_i > 0$ para $i \geq 2$ par y $\bar{D}_i < 0$ para impar (-).

Ejemplos:

- a) La forma cuadrática $f(X, Y, Z) = X^2 - Y^2 - 7Z^2 + XY$ sujeta a la relación lineal $X + Y + 2Z = 0$ es definida negativa pues:

$$\bar{D}_3 = \begin{vmatrix} 0 & Cg(X) & Cg(Y) & Cg(Z) \\ Cg(X) & Cf(X^2) & Cf(\frac{XY}{2}) & Cf(\frac{XZ}{2}) \\ Cg(Y) & Cf(\frac{YX}{2}) & Cf(Y^2) & Cf(\frac{YZ}{2}) \\ Cg(Z) & Cf(\frac{ZX}{2}) & Cf(\frac{ZY}{2}) & Cf(Z^2) \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1/2 & 0 \\ 1 & 1/2 & -1 & 0 \\ 2 & 0 & 0 & -7 \end{vmatrix} = -2 \text{ par } < 0 \bar{D}_i > 0$$

$$\bar{D}_2 = \begin{vmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1/2 & 0 \\ 1 & 1/2 & -1 & 0 \\ 2 & 0 & 0 & -7 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 1 & 1/2 \\ 1 & 1/2 & -1 \end{vmatrix} = -1 \text{ impar } < 0 \bar{D}_i < 0$$

Por lo tanto, esta forma cuadrática con restricción lineal es negativa en su determinante.

Clasificación de Puntos Críticos, recordando la definición de extremos locales:

- Un punto P es Mínimo Local de $F(X_1, X_2, X_3, X_4, \dots, X_n)$ si existe un vecindario V_P alrededor de P en que cumple $f(P) \leq F(Q) \forall Q \in V_P$.
- Un punto P es Máximo Local de $F(X_1, X_2, X_3, X_4, \dots, X_n)$ si existe un vecindario V_P alrededor de P en que cumple $f(P) \geq F(Q) \forall Q \in V_P$.
- Un punto P es Ensilladura o Silla de $F(X_1, X_2, X_3, X_4, \dots, X_n)$ si existe un vecindario V_P alrededor de P en que cumple tanto $f(P) \leq F(Q)$ como $f(P) \geq F(R)$ para diferentes Q, R de V_P .

3.4. Criterio para Máximos y Mínimos-Taylor

Criterio para Máximos y Mínimos -Taylor puede generalizarse a varias variables como sea V un conjunto convexo abierto, si f es continua y tiene derivadas parciales continuas de segundo orden, sobre V , entonces existe $t \in [0, 1] \Rightarrow \forall$ dos puntos $P, Q \in V; Q = P + h$:

$$f(P + h) = f(P) + \nabla f(P) \cdot h + \frac{1}{2} h^T H [tQ + (1 - t)P] h \quad t \in [0, 1] \rightarrow$$

$P, Q \in V; Q = P + h; h = h_1, h_2, h_3, h_4, \dots, h_n$ y H es Hessiana $|H_{[R]}| = |A|$ o Det.

Del teorema de Taylor y teoría previa de formas cuadráticas, se obtienen condiciones suficientes para un máximo o un mínimo local según sea $D_1(P), D_2(P), D_3(P), D_4(P), \dots, D_n(P)$, n determinantes definidos:

$$D_3(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) & f_{XZ}(P) \\ f_{YX}(P) & f_{YY}(P) & f_{YZ}(P) \\ f_{ZX}(P) & f_{ZY}(P) & f_{ZZ}(P) \end{pmatrix} \\ = \begin{pmatrix} 2 & -2 & 2 \\ -2 & 6 & -2 \\ 2 & -2 & 8 \end{pmatrix} = 48 > 0$$

$$D_2(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) & \cancel{f_{XZ}(P)} \\ f_{YX}(P) & f_{YY}(P) & \cancel{f_{YZ}(P)} \\ \cancel{f_{ZX}(P)} & \cancel{f_{ZY}(P)} & \cancel{f_{ZZ}(P)} \end{pmatrix} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) \\ f_{YX}(P) & f_{YY}(P) \end{pmatrix} \\ = \begin{pmatrix} 2 & -2 \\ -2 & 6 \end{pmatrix} = 8 > 0$$

$$D_1(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & \cancel{f_{XY}(P)} \\ \cancel{f_{YX}(P)} & \cancel{f_{YY}(P)} \end{pmatrix} = (2) = 2 > 0$$

Por lo tanto, en $P(0, 0, 0)$ f alcanza un mínimo local.

b) $f(X, Y, Z) = X^2 - Y^2 - YZ.$

Primer punto crítico ($\nabla(X, Y, Z) = (0, 0, 0)$):

$$\left. \begin{array}{l} f_X = 2X = 0 \\ f_Y = -2Y - Z = 0 \\ f_Z = -Y \end{array} \right\} P(0, 0, 0)$$

Para estimar $D_i(P)$:

$$D_3(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) & f_{XZ}(P) \\ f_{YX}(P) & f_{YY}(P) & f_{YZ}(P) \\ f_{ZX}(P) & f_{ZY}(P) & f_{ZZ}(P) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & 0 \end{pmatrix} \\ = -2 < 0$$

$$D_2(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) & \cancel{f_{XZ}(P)} \\ f_{YX}(P) & f_{YY}(P) & \cancel{f_{YZ}(P)} \\ \cancel{f_{ZX}(P)} & \cancel{f_{ZY}(P)} & \cancel{f_{ZZ}(P)} \end{pmatrix} \\ = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) \\ f_{YX}(P) & f_{YY}(P) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = -4 < 0$$

$$D_1(P) = |A| \text{ ó Det} = \begin{pmatrix} f_{XX}(P) & f_{XY}(P) \\ f_{YX}(P) & f_{YY}(P) \end{pmatrix} = (2) = 2 > 0$$

Por lo tanto, en $P(0, 0, 0)$ f podría haber o no un extremo local.

Puntos Críticos para problemas con una o más restricciones considere f sujeto a restricción $g = c$ o, es decir, «optimizar la función objetivo $f(X_1, X_2, X_3, X_4, \dots, X_n)$ sujeta a restricción $g(X_1, X_2, X_3, X_4, \dots, X_n) = c$ », cuya función lagrangiana será:

$$\begin{aligned} L(X_1, X_2, X_3, X_4, \dots, X_n, \lambda_1) \\ &= f(X_1, X_2, X_3, X_4, \dots, X_n) - \lambda(c \\ &\quad - g(X_1, X_2, X_3, X_4, \dots, X_n)) \end{aligned}$$

Así, los puntos críticos se obtienen resolviendo el sistema, con condiciones de primer orden:

$$\left\{ \begin{array}{l} LX_1 = f_{X_1} - \lambda g_{X_1} = 0 \\ LX_2 = f_{X_2} - \lambda g_{X_2} = 0 \\ LX_3 = f_{X_3} - \lambda g_{X_3} = 0 \\ LX_4 = f_{X_4} - \lambda g_{X_4} = 0 \\ \vdots \\ LX_n = f_{X_n} - \lambda g_{X_n} = 0 \\ L_\lambda = g(X_1, X_2, X_3, X_4, \dots, X_n) = 0 \end{array} \right.$$

El criterio que se usará para clasificar puntos críticos difiere del criterio que se usa en problema de optimización sin restricciones.

Por lo tanto, se considera el Hessiano Orlado:

$$\bar{D}_n(P) = \begin{pmatrix} 0 & g_{X_1}(P) & g_{X_2}(P) & g_{X_3}(P) & g_{X_4}(P) & \dots & g_{X_n}(P) \\ g_{X_1}(P) & L_{X_1X_1}(P) & L_{X_1X_2}(P) & L_{X_1X_3}(P) & L_{X_1X_4}(P) & \dots & f_{X_1X_n}(P) \\ g_{X_2}(P) & L_{X_2X_1}(P) & L_{X_2X_2}(P) & L_{X_2X_3}(P) & L_{X_2X_4}(P) & \dots & L_{X_2X_n}(P) \\ g_{X_3}(P) & L_{X_3X_1}(P) & L_{X_3X_2}(P) & L_{X_3X_3}(P) & L_{X_3X_4}(P) & \dots & L_{X_3X_n}(P) \\ g_{X_4}(P) & L_{X_4X_1}(P) & L_{X_4X_2}(P) & L_{X_4X_3}(P) & L_{X_4X_4}(P) & \dots & L_{X_4X_n}(P) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ g_{X_n}(P) & L_{X_nX_1}(P) & L_{X_nX_2}(P) & L_{X_nX_3}(P) & L_{X_nX_4}(P) & \dots & L_{X_nX_n}(P) \end{pmatrix}$$

⇒ Sus menores principales

$$\bar{D}_i(P) = \begin{pmatrix} 0 & g_{X_1}(P) & g_{X_2}(P) & g_{X_3}(P) & g_{X_4}(P) & \dots & g_{X_n}(P) \\ g_{X_1}(P) & L_{X_1X_1}(P) & L_{X_1X_2}(P) & L_{X_1X_3}(P) & L_{X_1X_4}(P) & \dots & L_{X_1X_i}(P) \\ g_{X_2}(P) & L_{X_2X_1}(P) & L_{X_2X_2}(P) & L_{X_2X_3}(P) & L_{X_2X_4}(P) & \dots & L_{X_2X_i}(P) \\ g_{X_3}(P) & L_{X_3X_1}(P) & L_{X_3X_2}(P) & L_{X_3X_3}(P) & L_{X_3X_4}(P) & \dots & L_{X_3X_i}(P) \\ g_{X_4}(P) & L_{X_4X_1}(P) & L_{X_4X_2}(P) & L_{X_4X_3}(P) & L_{X_4X_4}(P) & \dots & L_{X_4X_i}(P) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ g_{X_i}(P) & L_{X_iX_1}(P) & L_{X_iX_2}(P) & L_{X_iX_3}(P) & L_{X_iX_4}(P) & \dots & L_{X_iX_i}(P) \end{pmatrix}$$

Si P es un punto crítico de f sujeto a restricción $g(X_1, X_2, X_3, X_4, \dots, X_n) = c$ con $D_1(P) < 0$:

- En P f es Mínimo Local si $D_2(P) < 0, D_3(P) < 0, D_4(P) < 0, \dots, D_n(P) < 0$.
- En P f es Máximo Local si todos los $D_i(P) > 0$ si $i \geq 2$ es par ó pares son + y todos los $D_i(P) < 0$ si i es impar ó impares son -.

Cuando aparece más de una restricción se debe considerar un Hessiano con más de una orla (restricción), pues si hay n variables y m restricciones tal que $m < n$ de forma $g^i(X_1, X_2, X_3, X_4, \dots, X_n) = c_i \Rightarrow$ Lagrangina será:

$$L = f(X_1, X_2, X_3, X_4, \dots, X_n) + \sum_{i=1}^m \lambda_i [c_i - g^i(X_1, X_2, X_3, X_4, \dots, X_n)]$$

Y su hessiano orlado será:

$$\begin{pmatrix} 00 \dots 0 & g_1^1 & g_2^1 & g_3^1 & g_4^1 & \dots & g_n^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 00 \dots 0 & g_1^m & g_2^m & g_3^m & g_4^m & \dots & g_n^m \\ g_1^1 g_1^2 \dots g_1^m & L_{X_1 X_1} & L_{X_1 X_2} & L_{X_1 X_3} & L_{X_1 X_4} & \dots & L_{X_1 X_n} \\ g_2^1 g_2^2 \dots g_2^m & L_{X_2 X_1} & L_{X_2 X_2} & L_{X_2 X_3} & L_{X_2 X_4} & \dots & L_{X_2 X_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ g_n^1 g_n^2 \dots g_n^m & L_{X_n X_1} & L_{X_n X_2} & L_{X_n X_3} & L_{X_n X_4} & \dots & L_{X_n X_n} \end{pmatrix}$$

Ejemplos:

1. Halle extremos de $Z = X^2 + Y^2$ sujeto a restricción $X + 4Y = 2$.

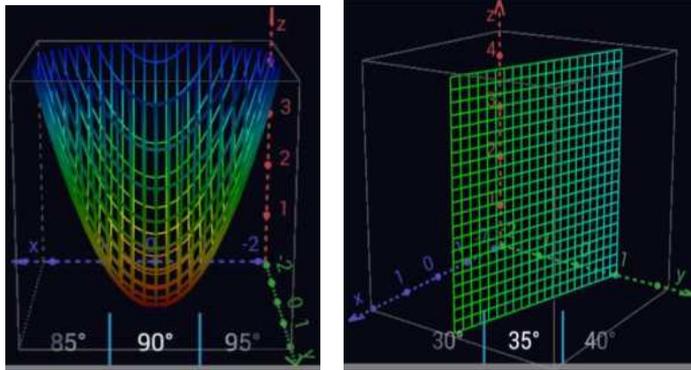
$$\begin{aligned} \text{Puntos críticos} \quad & \begin{cases} L_X = 2X - \lambda = 0 \\ L_Y = 2Y - 4\lambda = 0 \\ L_\lambda = 2 - X - 4\lambda = 0 \end{cases} \Rightarrow \lambda = \frac{4}{17}, X = \frac{2}{17} \text{ y } Y \\ & = \frac{8}{17} \end{aligned}$$

Tal que el único punto crítico es $P = (\frac{2}{17}, \frac{8}{17})$. Se usa el teorema para clasificar los puntos críticos que, en este caso, se calcula el Hessiano Orlado (\bar{D}_2):

$$\bar{D}_2(P) = \begin{vmatrix} 0 & C g(X) & C g(Y) \\ C g(X) & C f(L_{X_1 X_1}) & C f(L_{X_1 X_2}) \\ C g(Y) & C f(L_{X_2 X_1}) & C f(L_Y) \end{vmatrix} = \begin{vmatrix} 0 & 1 & 4 \\ 1 & 2 & 0 \\ 4 & 0 & 2 \end{vmatrix} = -33 < 0$$

Por lo tanto, $P = (\frac{2}{17}, \frac{8}{17}, z(\frac{2}{17}, \frac{8}{17}))$ es mínimo local. Gráficamente:

Figura 3.9 Representación gráfica del mínimo local



Fuente: Adaptación (Mora, W. F. 2016).

2. Maximice $f(X, Y) = 2Y - X$ sujeto a $Y = \text{Sen } X, 0 \leq X \leq 2\pi$.

Sea $F(X, Y, \lambda) = 2Y - X - \lambda(Y - \text{Sen } X)$

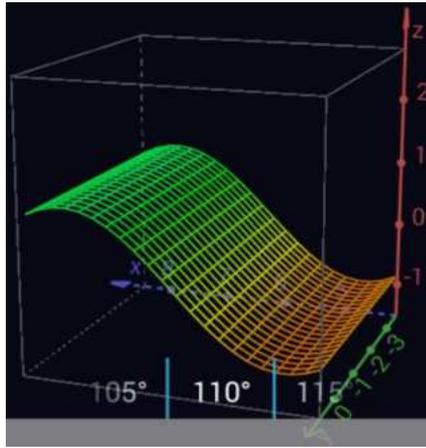
$$\text{Puntos críticos } \begin{cases} F_X = -1 + \lambda \text{Cos } X = 0 \\ F_Y = 2 - \lambda = 0 \\ F_\lambda = -Y + \text{Sen } X = 0 \end{cases} \Rightarrow \lambda = 2, \text{Cos } X = \frac{1}{2}, X = \frac{\pi}{3}, X = \frac{5}{3}\pi \text{ y } Y = \frac{\sqrt{3}}{2}$$

Tal que $P_1 = (\frac{\pi}{3}, \frac{\sqrt{3}}{2})$ y $P_2 = (\frac{5}{3}\pi, -\frac{\sqrt{3}}{2})$. Se usa el teorema para clasificar los puntos críticos que, en este caso, se calcula el Hessiano Orlado (\bar{D}_2):

$$\bar{D}_2(P) = \begin{vmatrix} 0 & C g(X) & C g(Y) \\ C g(X) & C f(L_{X_1 X_1}) & C f(L_{X_1 X_2}) \\ C g(Y) & C f(L_{X_2 X_1}) & C f(L_Y) \end{vmatrix} = \begin{vmatrix} 0 & -\text{Cos } X & 1 \\ -\text{Cos } X & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0$$

Por lo tanto, $\bar{D}_2(P_1) = \bar{D}_2(P_2) = 0$ tal que el criterio no da información. Gráficamente el plano restrictivo es:

Figura 3.10 Representación del plano restrictivo



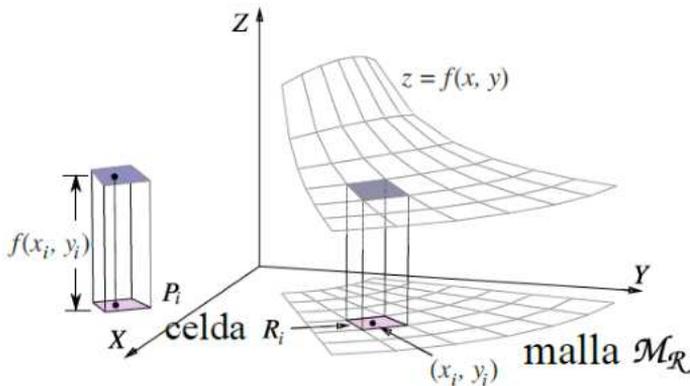
Fuente: Adaptación (Mora, W. F. 2016).

3.5. Integral Doble parte de \mathbb{R}

Integral Doble parte de \mathbb{R} con región acotada y cerrada del plano de área $A(\mathbb{R})$ y $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ función definida acotada sobre \mathbb{R} .

Suponga que $M_{\mathbb{R}} = \{R_1, R_2, R_3, R_4, \dots, R_n\}$ es un conjunto de n celdas que conforman una malla que cubre \mathbb{R} , tal que el área de cada celda R_i se denomina ΔA_i

Figura 3.11 Representación gráfica integral doble



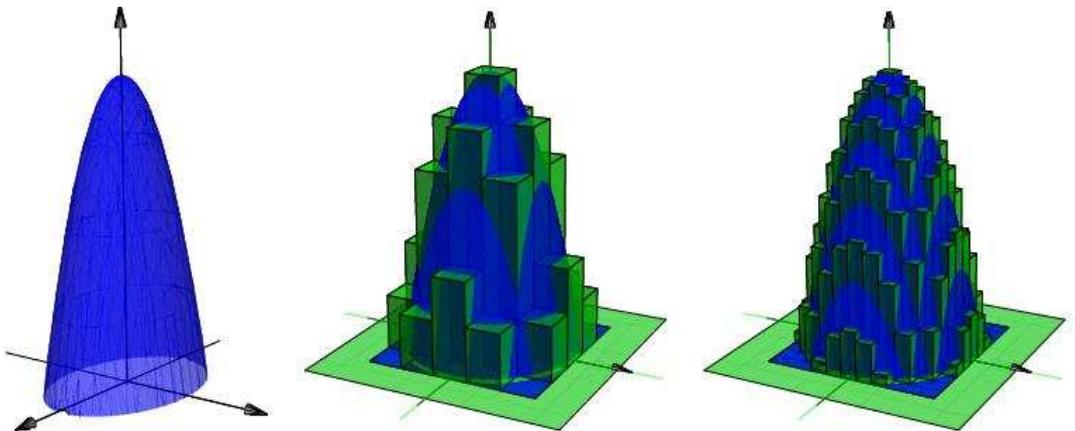
Fuente: (Mora, W. F. 2016).

La suma de Riemann de f sobre R (cuadrado) es una expresión de forma:

$$\sum_{i=1}^n f(X_i, Y_i) \Delta A_i$$

Donde $(X_i, Y_i) \in R_i$. Si f es continua y positiva sobre $R \Rightarrow f(X_i, Y_i) \Delta A_i$ aproxima el volumen de cada prisma P_i de base R_i y altura $f(X_i, Y_i)$ tal que la suma de Riemann aproxima el volumen del sólido entre la región R y gráfico de f :

Figura 3.12 Volumen del sólido



Fuente: (Mora, W. F. 2016).

El diámetro de malla de cada celda R_i es la máxima distancia entre todas las distancias entre cualesquiera dos puntos en R_i , denotado por $\|R_i\|$.

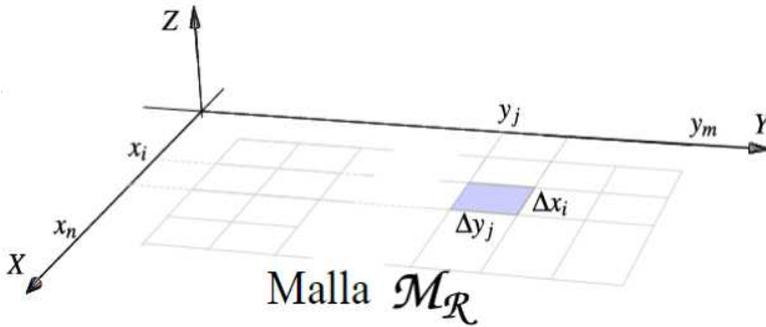
El diámetro de malla M_R es $\|M_R\| = \text{Sup}_i\{\|R_i\|\}$. Entonces, la Función Integrable sostiene que si las sumas de Riemann de f sobre M_R tiene un límite, independiente de los (X_i, Y_i) , conforme $\|M_R\| \rightarrow 0$ entonces se dice que f es integrable sobre R y la integral es este límite:

$$\iint_R f(X, Y) dA = \lim_{\|M\| \rightarrow 0} \sum_{i=1}^n f(X_i, Y_i) \Delta A_i \Rightarrow n = \text{Card}(M)$$

Si R es región rectangular la malla M_R se puede tomar como un conjunto de rectángulos $R_{ij} = [X_i, X_{i+1}] * [Y_j, Y_{j+1}]$ de área $\Delta A_{ij} = \Delta X_i \Delta Y_j$ reemplazando dA por $dX dY$ escribiendo su límite como:

$$\iint_R f(X, Y) dX dY = \lim_{\|n, m\| \rightarrow \infty} \sum_i^n \sum_j^m f(X_i, Y_j) \Delta X_i \Delta Y_j$$

Figura 3.13 Diámetro de malla



Fuente: (Mora, W. F. 2016).

Teorema Propiedades de Funciones Integrables, similares a las de una variable:

- Si f es continua sobre $R \Rightarrow f$ es integrable sobre R .
- Sea $k \in \mathbb{R}$. Si f y g son integrables sobre $R \Rightarrow kf$ y $f \pm g$ son integrables sobre R y $\iint_R kf(X, Y) dA = k \iint_R f(X, Y) dA$ y $\iint_R f(X, Y) \pm g(X, Y) dA = \iint_R f(X, Y) dA \pm \iint_R g(X, Y) dA$.
- Si f y g son integrables sobre regiones R y S que no se traslapan, entonces f es integrable sobre $R \cup S$ y $\iint_{R \cup S} f(X, Y) dA = \iint_R f(X, Y) dA + \iint_S f(X, Y) dA$.
- Si f y g son integrables sobre regiones R y $f(X, Y) \leq g(X, Y)$ para $\forall (X, Y) \in R$ entonces $\iint_R f(X, Y) dA \leq \iint_R g(X, Y) dA$.
- Si f es integrable sobre R y $M \leq f(X, Y) \leq m$ para $\forall (X, Y) \in R \Rightarrow MA(R) \leq \iint_R f(X, Y) dA \leq mA(R)$.

CAPÍTULO IV

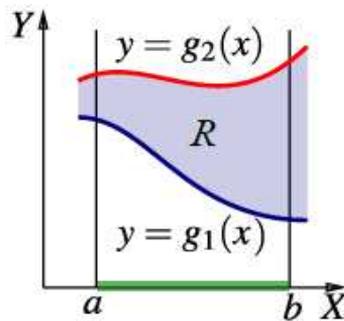
CÁLCULO DE INTEGRALES DOBLES

Cálculo de Integrales Dobles, la integral iterada indica que el teorema de Fubini establece que si f es continua sobre R , la integral doble se puede evaluar por «integración parcial» respecto a cada variable, una a la vez.

Éste es el método de «integrales iteradas» tal que se puede especificar dos maneras equivalentes de describir una misma región:

1. Región entre curvas $Y = g_1(X)$ y $Y = g_2(X)$ tal que $R = \{(X, Y) \in \mathbb{R}^2 \Rightarrow a \leq X \leq b \text{ y } g_1(X) \leq Y \leq g_2(X)\}$ con g_1 y g_2 funciones continuas en $[a, b]$:

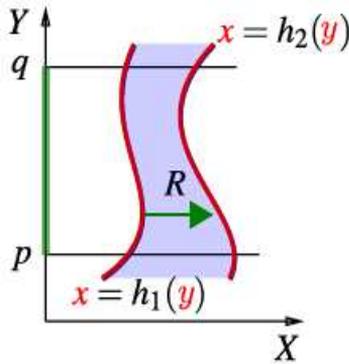
Figura 4.1 Región entre curvas



Fuente: (Mora, W. F. 2016).

2. Región entre curvas $X = h_1(Y)$ y $X = h_2(Y)$ tal que $R = \{(X, Y) \in \mathbb{R}^2 \Rightarrow P \leq Y \leq Q \text{ y } h_1(Y) \leq X \leq h_2(Y)\}$ con h_1 y h_2 funciones continuas en $[P, Q]$:

Figura 4.2 Región entre curvas



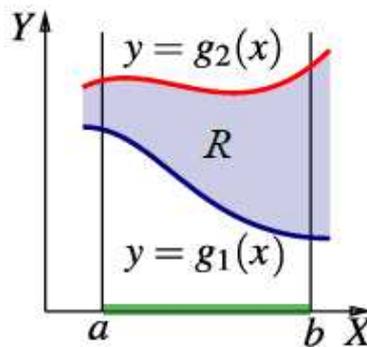
Fuente: (Mora, W. F. 2016).

El Teorema de Fubini indica, por un lado, sea $R = \{(X, Y) \in \mathbb{R}^2 \Rightarrow a \leq X \leq b \text{ y } g_2(X) \leq Y \leq g_1(X)\}$ con g_1 y g_2 funciones continuas en $[a, b]$. Si f es continua en $R \Rightarrow$:

$$\iint_R f(X, Y) \, dA = \int_a^b \int_{g_1(X)}^{g_2(X)} f(X, Y) \, dy \, dx = \int_a^b \left[\int_{g_1(X)}^{g_2(X)} f(X, Y) \, dy \right] dx$$

Gráficamente ($Y = f(X)$, en otras palabras, Y está en función de X):

Figura 4.3 Región entre curvas



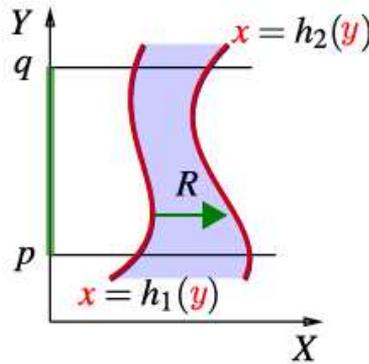
Fuente: (Mora, W. F. 2016).

Por otro lado, sea $R = \{(X, Y) \in \mathbb{R}^2 \Rightarrow p \leq Y \leq q \text{ y } h_1(Y) \leq X \leq h_2(Y)\}$ con h_1 y h_2 funciones continuas en $[p, q]$. Si f es continua en $R \Rightarrow$:

$$\iint_R f(X, Y) \, dA = \int_p^q \int_{h_1(Y)}^{h_2(Y)} f(X, Y) \, dx \, dy = \int_p^q \left[\int_{h_1(Y)}^{h_2(Y)} f(X, Y) \, dx \right] dy$$

Gráficamente ($X = f(X)$), en otras palabras, X está en función de Y):

Figura 4.4 Región entre curvas

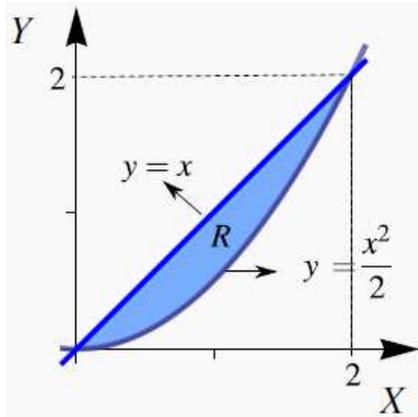


Fuente: (Mora, W. F. 2016).

Ejemplo:

1. Sea R región de figura. Calcule $\iint_R XY \, dA$ usando el orden de integración “ $dydx$ ” ($Y = f(X)$) y teniendo $0 \leq X \leq 2$ y $\frac{X^2}{2}(X) \leq Y \leq X$:

Figura 4.5 Región entre curvas



Fuente: (Mora, W. F. 2016).

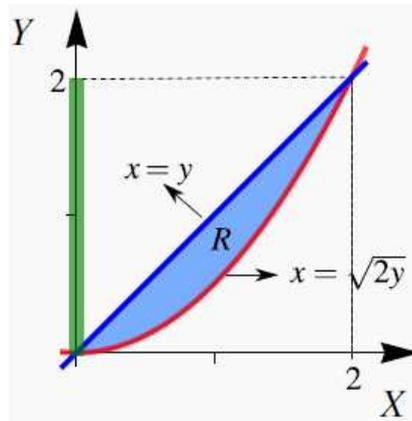
$$\iint_R XY \, dA = \int_0^2 \int_{\frac{x^2}{2}}^x XY \, dy \, dx = \int_0^2 \left[\int_{\frac{x^2}{2}}^x XY \, dy \right] dx$$

$$\begin{aligned} \iint_R XY \, dA &= \int_0^2 X \left(\frac{Y^2}{2} \right) \Big|_{\frac{x^2}{2}}^x dx = \int_0^2 \left(\frac{X^3}{2} - \frac{X^5}{8} \right) dx = \left(\frac{X^4}{8} - \frac{X^6}{48} \right) \Big|_0^2 \\ &= \left[\frac{(2)^4}{8} - \frac{(2)^6}{48} \right] - \left[\frac{(0)^4}{8} - \frac{(0)^6}{48} \right] = \frac{2}{3} \end{aligned}$$

Usando el orden de integración “ $dx \, dy$ ” ($X = f(Y)$) y teniendo $0 \leq Y \leq 2$ y $Y \leq X \leq \sqrt{2Y}$:

Para encontrar este planteamiento se hace un giro de 90° en sentido de manecillas del reloj (derecha) tal que $a = q$, $b = p$ y los límites superior-inferior no sufren cambios; aunque, $Y = g_1(X) \approx X = h_1(Y)$ y $Y = g_2(X) \approx X = h_2(Y)$.

Figura 4.6 Región entre curvas



Fuente: (Mora, W. F. 2016).

$$\iint_R XY \, dA = \int_0^2 \int_Y^{\sqrt{2Y}} YX \, dx dy = \int_0^2 \left[\int_Y^{\sqrt{2Y}} YX dx \right] dy$$

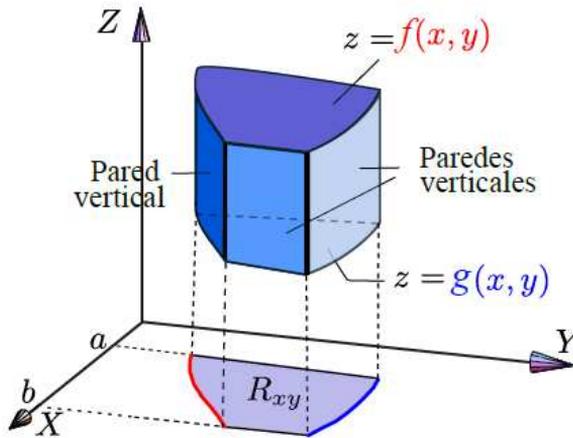
$$\begin{aligned} \iint_R XY \, dA &= \int_0^2 Y \left(\frac{X^2}{2} \right) \Big|_Y^{\sqrt{2Y}} dy = \int_0^2 \left(\frac{2Y^2}{2} - \frac{Y^2}{2} Y \right) dy \\ &= \left(\frac{2Y^3}{2 \cdot 3} - \frac{Y^4}{8} \right) \Big|_0^2 = \left[\frac{(2)^3}{3} - \frac{(2)^4}{8} \right] - \left[\frac{(0)^3}{3} - \frac{(0)^4}{8} \right] \\ &= \frac{2}{3} \end{aligned}$$

4.1. Área-Volumen

Según definición de integral doble, el área A_R de una región R se puede calcular con integral doble “área de base x altura”: $A_R = \iint_R 1 \, dA \Rightarrow V_Q = \iint_R f(X, Y) \, dA$ pues $f(X, Y) \geq 0$ y continua en una región cerrada R , V_Q volumen del sólido Q que tiene a R como base una altura de medida $f(X, Y)$ en cada $(X, Y) \in \mathbb{R} \Rightarrow V_Q =$

$\iint_R f(X, Y) dA$ y sólido Q está limitado, sobre región cerrada R, por dos superficies de ecuaciones $Z = f(X, Y)$ y $Z = g(X, Y)$ con $f-g$ continuas y $f(X, Y) - g(X, Y) \geq 0$ sobre R $\Rightarrow V_Q = \iint_R f(X, Y) - g(X, Y) dA$:

Figura 4.7 Área A_R de una región R



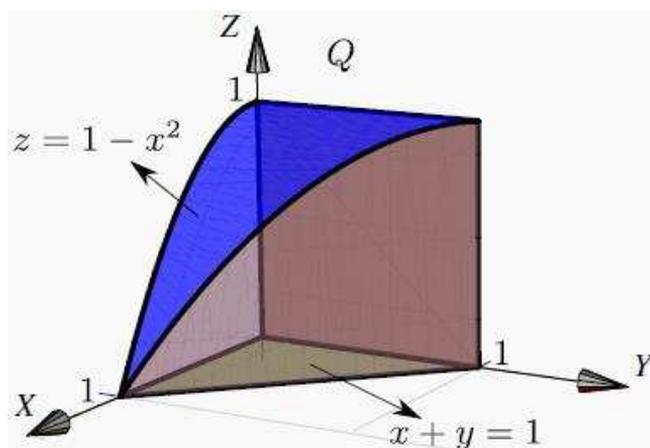
Fuente: (Mora, W. F. 2016).

Y muchas veces es conveniente considerar la región R como proyección del sólido sobre Π : XZ o YZ.

Ejemplos:

- a) Sea Q un sólido limitado por las superficies $Z = 1 - X^2, X + Y = 1$ en el primer octante. Calcule V_Q usando como región R cada una de las proyecciones del sólido sobre Π : XY, YZ y XZ.

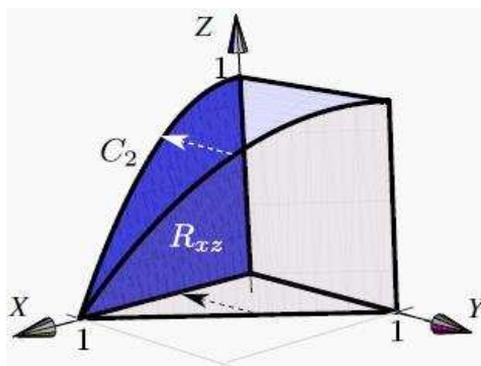
Figura 4.8 Volumen V_Q de una región R



Fuente: (Mora, W. F. 2016).

El cálculo de V_Q proyectado sobre $\Pi: XZ$ es:

Figura 4.9 Cálculo de V_Q proyectado sobre $\Pi: XZ$



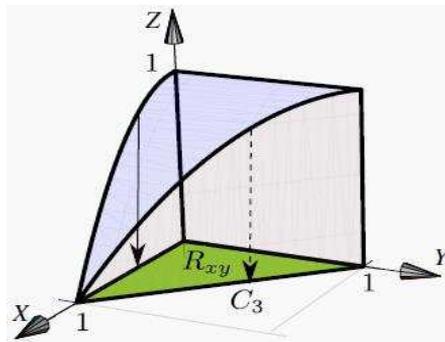
Fuente: (Mora, W. F. 2016).

La región de interés está entre la curva $C_2: Z = 1 - X^2, X = 0$ y $Y = 1 - X$ tal que integrando en el orden “ $dzdx$ ”:

$$\begin{aligned}
 V_Q &= \int_0^1 \int_0^{1-X^2} 1 - X - 0 \, dz dx = \int_0^2 Z - ZX \Big|_0^{1-X^2} dx \\
 &= \int_0^2 (1 - X)(1 - X^2) dx = X - \frac{1}{3}X^3 - \frac{1}{2}X^2 + \frac{1}{4}X^4 \Big|_0^1 \\
 &= \frac{5}{12}
 \end{aligned}$$

El cálculo de V_Q proyectado sobre $\Pi: XY$ es:

Figura 4.10 Cálculo de V_Q proyectado sobre $\Pi: XY$



Fuente: (Mora, W. F. 2016).

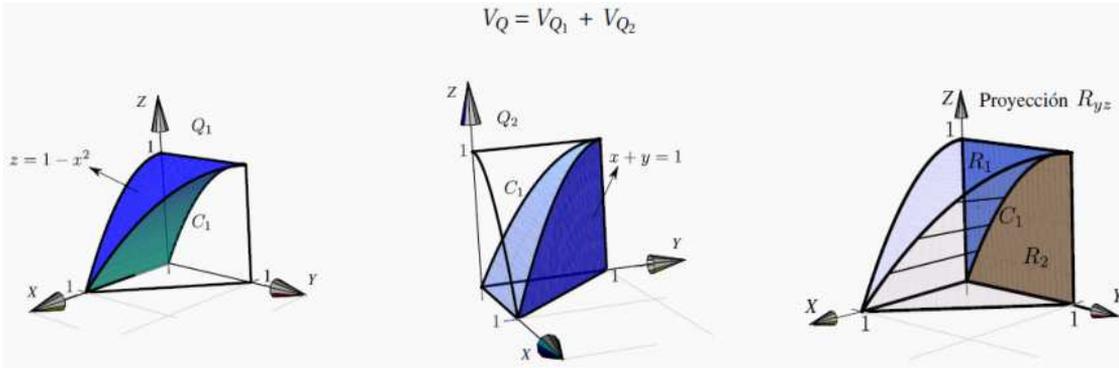
La región de interés está entre la curva $C_3: Y = 1 - X, X \in [0, 1], Z = 0$ y $Z = 1 - X^2$ tal que integrando en el orden “dydx”:

$$\begin{aligned}
 V_Q &= \int_0^1 \int_0^{1-X} 1 - X^2 - 0 \, dy dx = \int_0^2 Z - ZX \Big|_0^{1-X^2} dx \\
 &= \int_0^2 (1 - X)(1 - X^2) dx = X - \frac{1}{2}X^2 - \frac{1}{3}X^3 + \frac{1}{4}X^4 \Big|_0^1 \\
 &= \frac{5}{12}
 \end{aligned}$$

El cálculo de V_Q proyectado sobre $\Pi: YZ$ el sólido no está entre dos superficies, se tiene un sólido Q_1 que está entre $X = 0$ y $Z = 1 - X^2$ en

R_1 y un sólido Q_2 entre $X = 0$ y $\Pi: X + Y = 1$ en R_2 tal que $Q = Q_1 \cup Q_2$:

Figura 4.11 Cálculo de V_Q proyectado sobre $\Pi:YZ$



Fuente: (Mora, W. F. 2016).

La curva C_1 es la proyección de YZ de la curva de intersección entre la superficie $Z = 1 - X^2$ y $\Pi: X + Y = 1$ tal que C_1 tiene ecuación en términos X e Y : $Z = 1 - X^2 \cap X + Y = 1 \Rightarrow Z = 1 - (1 - Y)^2 \Rightarrow Y \in [0, 1]$.

En consecuencia, C_1 divide la región de integración en dos partes (regiones R_1 y R_2), $X = \sqrt{1 - Z}$, $X = 0$ sobre la región R_1 y $X = 1 - Y$, $X = 0$ sobre la región R_2 .

Por lo tanto, el orden de integración “ $dzdy$ ” es:

$$\begin{aligned}
 V_Q &= \int_0^1 \int_{2Y-Y^2}^1 \sqrt{1-Z} - 0 \, dzdy + \int_0^1 \int_0^{2Y-Y^2} 1 - Y - 0 \, dzdy \\
 &= \int_0^1 \frac{2}{3} (1 - 2Y + Y^2)^{\frac{3}{2}} dy + \int_0^1 (2Y - 3Y^2 + Y^3) dy \\
 &= \int_0^1 \frac{2}{3} - (Y - 1)^3 dy + \int_0^1 (2Y - 3Y^2 + Y^3) dy \\
 &= -\frac{1}{6}Y^4 + \frac{2}{3}Y^3 - Y^2 + \frac{2}{3}Y \Big|_0^1 + Y^2 - Y^3 + \frac{1}{4}Y^4 \Big|_0^1 = \frac{5}{12}
 \end{aligned}$$

Para el cambio de variable en una integral doble se plantea una variable, si f tiene una derivada continua en $[a, b]$ y $X = X(U)$ está definida en $[u_1, u_2]$ con $a = X(u_1)$ y $b = X(u_2)$, si $f(X(U))$ es continua en $[u_1, u_2]$ entonces:

$$\int_a^b f(x) dx = \int_{u_1}^{u_2} f(X(U)) \frac{dx}{du} du$$

No obstante, existe una fórmula análoga para integrales dobles:

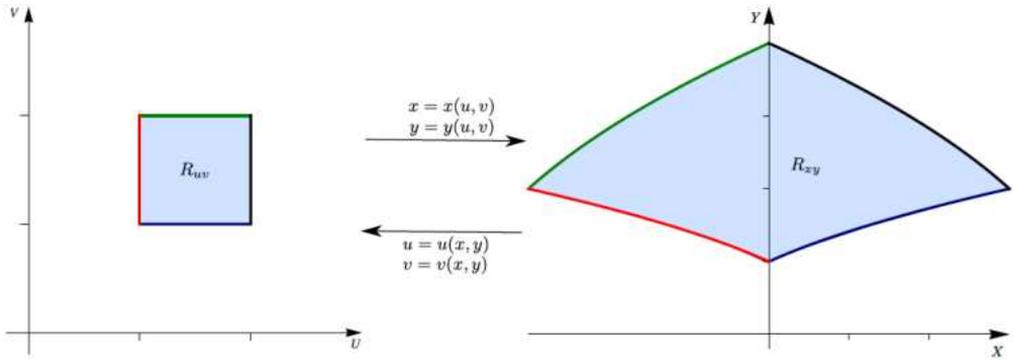
$$\begin{aligned} & \iint_{R_{XY}} f(X, Y) dx dy \\ &= \iint_{R_{UV}} f(X(U, V), Y(U, V)) \left| \frac{\partial(X, Y)}{\partial(U, V)} \right| dU dV \text{ con } \frac{\partial(X, Y)}{\partial(U, V)} \\ &= \text{Det } \acute{o} |A| \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix} \end{aligned}$$

Se asume que las funciones $X = X(U, V), Y = Y(U, V)$ están definidas y tienen derivadas parciales continua en región de integración R_{UV} en $\Pi: UV$. En este caso si se asume que las funciones inversas $U = U(X, Y), V = V(X, Y)$ están definidas y son continuas en R_{XY} , hay un mapeo invertible entre el interior de R_{XY} e interior de R_{UV} .

La función $f(X, Y)$ se asume continua en R_{XY} y así $f(X(U, V), Y(U, V))$ es continua en R_{UV} . También, se asume que el Jacobiano $J(U, V) = \frac{\partial(X, Y)}{\partial(U, V)}$ es no nulo en el interior de R_{UV} .

La restricción que el cambio de variable sea invertible en el interior de R_{XY} , $J(U, V)$ no se anule en el interior de R_{UV} , es necesaria para poder usar el cambio de variable con coordenadas polares en las regiones que contienen el origen.

Figura 4.12 Cambio de variables



Fuente: (Mora, W. F. 2016).

4.2. El Teorema de Cambio de Variable

El Teorema de Cambio de Variable indica que sea R_{UV} una región compacta y conexa en el plano contenida en un conjunto abierto A de \mathbb{R}^2 . Sea $r: A \rightarrow \mathbb{R}^2$ con $r(U, V) = (X(U, V), Y(U, V))$, una función continua con derivadas parciales continuas tal que r es invertible en el interior de R_{UV} y $J(U, V) = \text{Det } \left| \begin{matrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{matrix} \right|$ es no nulo en interior de R_{UV} .

R_{UV} .

4.3. El Jacobiano

El Jacobiano $J(U, V)$ va en valor absoluto dentro de la integral. Solo requiere que $r(U, V)$ sea invertible en el interior de R_{UV} y, por tanto, $|J(U, V)|$ no se anule en el interior de R_{UV} .

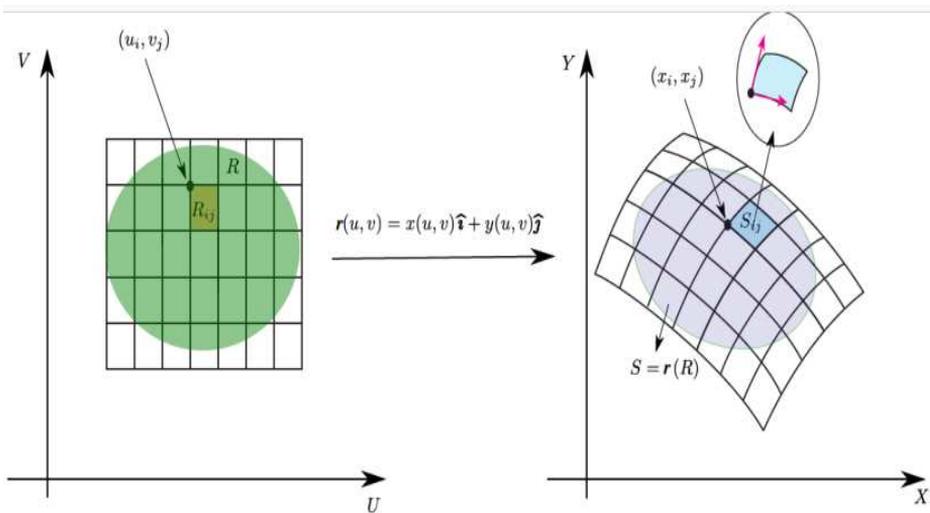
Sea $R_{XY} = r(R_{UV})$ y $f: R_{XY} \rightarrow \mathbb{R}$ una función continua.

Entonces:

$$\iint_{R_{XY}} f(X, Y) dx dy = \iint_{R_{UV}} f(X(U, V), Y(U, V)) |J(U, V)| dU dV$$

La Idea Geométrica considera el cambio de variable $X(U, V), Y(U, V)$ que transforma R en S y cumplen las condiciones del teorema. Este cambio de variable define una función invertible $r(U, V) = X(U, V)\hat{i} + Y(U, V)\hat{j}$ en el interior de S y $S = r(R)$:

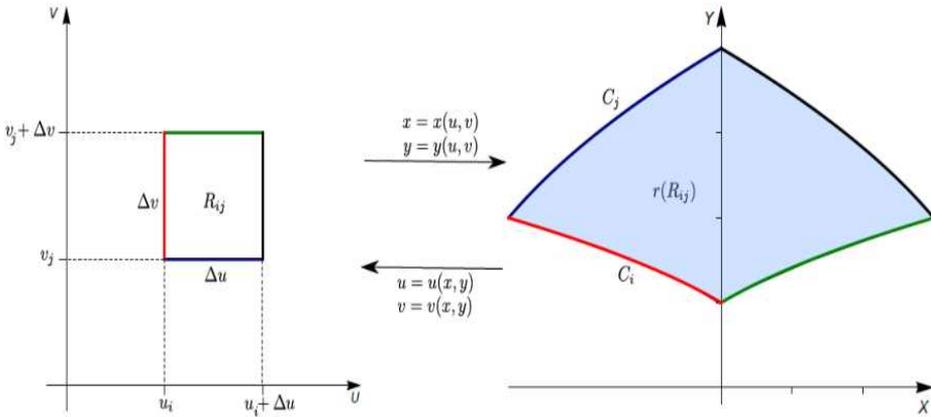
Figura 4.13 El Jacobiano



Fuente: (Mora, W. F. 2016).

Si se toma un rectángulo R_{ij} de una malla M de R . r transforma el lado $u = u_i$ en curva $C_i: r(u_i, V), V \in [V_j, V_j + \Delta V]$ y el lado $V = V_i$ en curva $C_j: r(U, V_i), V \in [U_j, U_j + \Delta U]$. Además, si r^{-1} es continua, r es un homeomorfismo y la frontera del rectángulo R_{ij} es “mapeada” en frontera $S = r(R_{ij})$ en el interior:

Figura 4.14 Mapeo sin frontera



Fuente: (Mora, W. F. 2016).

Si $r(U_i, V_j) = (X_i, Y_i)$ un vector tangente en (X_i, Y_i) en C_i es $\left. \frac{\partial r(U_i, V)}{\partial V} \right|_{V=V_j}$ representando la velocidad a que se desplaza el punto $r(U_i, V)$ cuando V va de U_j a $U_j + \Delta V$ tal que la curva C_i , (X_i, Y_i) se desplaza, en el tiempo ΔV , una distancia aproximada $\left. \frac{\partial r(U_i, V)}{\partial V} \right|_{V=V_j} \cdot \Delta V$.

Usando el teorema del valor medio para derivadas se tiene:

$$r(U_i, V + \Delta V) - r(U_i, V_j) \approx \Delta U r_V$$

Análogamente, un vector tangente en (X_i, Y_i) en C_j es $\left. \frac{\partial r(U, V_j)}{\partial U} \right|_{U=U_i}$, representando la velocidad a que se desplaza el punto (U, V_j) cuando U va de U_i a $U_i + \Delta U \Rightarrow C_j$, (X_i, Y_i) se desplaza, en el tiempo ΔU , una distancia aproximada $\left. \frac{\partial r(U, V_j)}{\partial U} \right|_{U=U_i} \Delta U$.

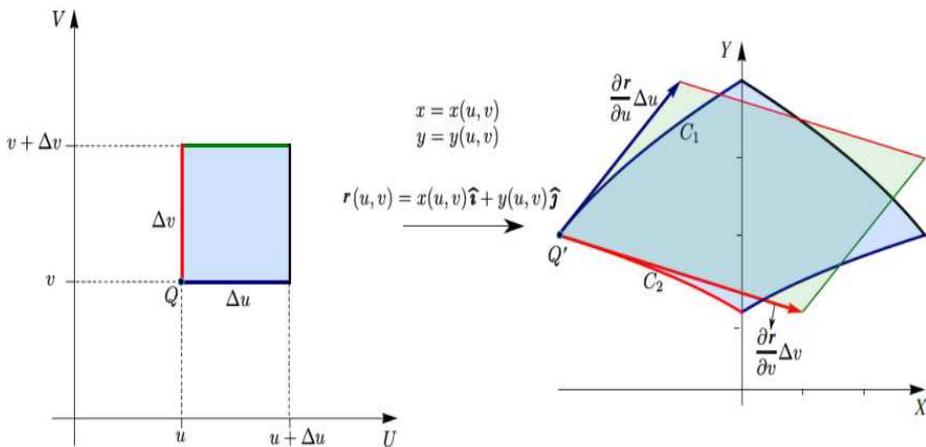
Por tanto, el rectángulo R_{ij} en R se transforma en una porción del $\Pi: XY$ que es casi paralelogramo de lados $\left. \frac{\partial r(U_i, V)}{\partial V} \right|_{V=V_j} \Delta V$ y

$\left. \frac{\partial r(U, V_j)}{\partial U} \right|_{U=U_i} \Delta U$ tal que el área de este paralelogramo es, en producto vectorial:

$$\left| \frac{\partial r(U_i, V)}{\partial V} \right|_{V=V_j} * \left. \frac{\partial r(U, V_j)}{\partial U} \right|_{U=U_i} \Delta U \Delta V$$

Gráficamente con un punto genérico (U, V) es:

Figura 4.15 Área del paralelogramo



Fuente: (Mora, W. F. 2016).

El área del paralelogramo «curvilíneo» es aproximadamente el área del paralelogramo de lados $\frac{\partial r}{\partial V} \Delta V$ y $\frac{\partial r}{\partial U} \Delta U$.

El área de este último paralelogramo es:

$$\left| \frac{\partial r}{\partial V} * \frac{\partial r}{\partial U} \right| \Delta U \Delta V = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial X}{\partial U} & \frac{\partial Y}{\partial U} & 0 \\ \frac{\partial X}{\partial V} & \frac{\partial Y}{\partial V} & 0 \end{vmatrix} = \begin{vmatrix} \frac{\partial X}{\partial U} & \frac{\partial Y}{\partial U} \\ \frac{\partial X}{\partial V} & \frac{\partial Y}{\partial V} \end{vmatrix} = |J(U, V)| \hat{k}$$

Tal que, si $J(U, V) = 1$, el cambio de variable conserva las áreas. Si no, el área de cada paralelogramo en XY es aproximadamente el área de cada rectángulo en UV , multiplicada por $|J(U, V)|$.

Entonces, $|J(U, V)|$ opera como un factor de compensación por la «deformación» sufrida por región ante un cambio de variable. Si la integral existe se debe tener:

$$\begin{aligned} A_S &= \iint_S 1 \, dx \, dy \approx \sum_{i=1}^m \sum_{j=1}^n |J(U, V)|_{U=U_i, V=V_j} \Delta U \Delta V \\ &= \iint_R |J(U, V)| \, dU \, dV \end{aligned}$$

Y en general:

$$\begin{aligned} A_S &= \iint_S f(X, Y) \, dx \, dy \\ &\approx \sum_{i=1}^m \sum_{j=1}^n f(X_i, Y_j) A_{S_{ij}} \\ &\approx \sum_{i=1}^m \sum_{j=1}^n f(r(U_i, V_j)) |J(U, V)|_{U=U_i, V=V_j} \\ &= \iint_R f(U, V) |J(U, V)| \, dU \, dV \end{aligned}$$

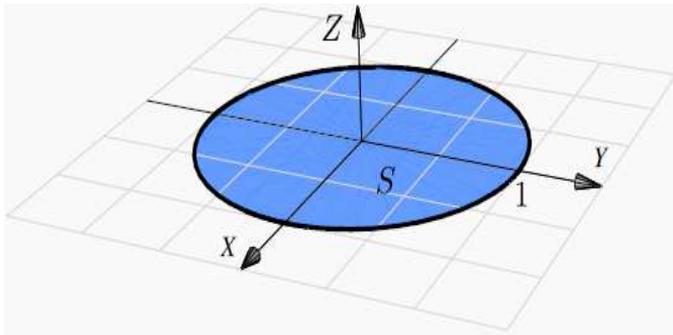
Superficies Parametrizadas indican que una parametrización de una superficie S en \mathbb{R}^3 es una función inyectiva $r: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $r(u, v) = (X(u, v), Y(u, v), Z(u, v))$ con $X(u, v)$, $Y(u, v)$ y $Z(u, v)$ funciones continuas sobre D , con imagen S ; es decir, $r(D) = S$ tal que la superficie $S = r(D)$ se escribe:

$$S: r(u, v) = X(u, v)\hat{i} + Y(u, v)\hat{j} + Z(u, v)\hat{k}, (u, v) \in D$$

Ejercicios propuestos

1. Caso $S: Z = f(X, Y)$ implica que esta superficie en un dominio D , se parametriza como $r(X, Y) = X \hat{i} + Y \hat{j} + F(X, Y) \hat{k}$, $(X, Y) \in D \Rightarrow D$ es proyección de S en $\Pi_{(XY)}$.
2. Considere la superficie $S: X^2 + Y^2 \leq 1, Z = 0$. Claramente, S es círculo de radio 1 en $\Pi_{(XY)}$, centrado en origen:

Figura 4.16 Ejercicio propuesto 2



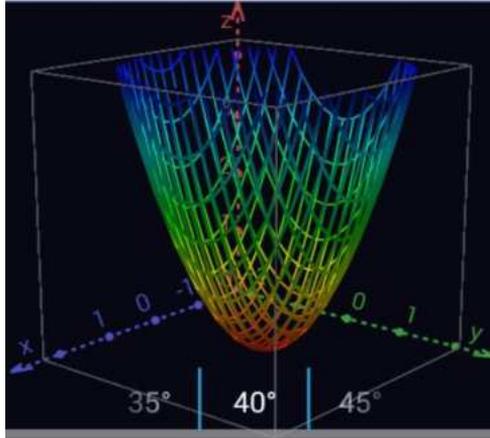
Fuente: (Mora, W. F. 2016).

Se puede describir a $S: Z = 0$ en $D = \{X^2 + Y^2 \leq 1\}$, pero más conveniente es parametrizar esta superficie como $r(X, Y) = X \hat{i} + Y \hat{j} + 0 * \hat{k}$, $(X, Y) \in D$

3. Sea S la porción del paraboloides $Z = X^2 + Y^2$ entre $Z = 0$ y $Z = 1 \Rightarrow S$ se puede parametrizar como $S: r(X, Y) = X \hat{i} + Y \hat{j} + (X^2 + Y^2) \hat{k}$, $(X, Y) \in D = \{(X, Y): X^2 + Y^2 \leq 1\}$. También, $Z = X^2 + Y^2$ se podría ver como circunferencias de radio $\sqrt{Z} \Rightarrow$:

$$S: r(\theta, Z) = \sqrt{Z} \cos(\theta) \hat{i} + \sqrt{Z} \sin(\theta) \hat{j} + Z \hat{k}, \theta \in [0, 2\pi[\text{ y } Z \in [0, 1]$$

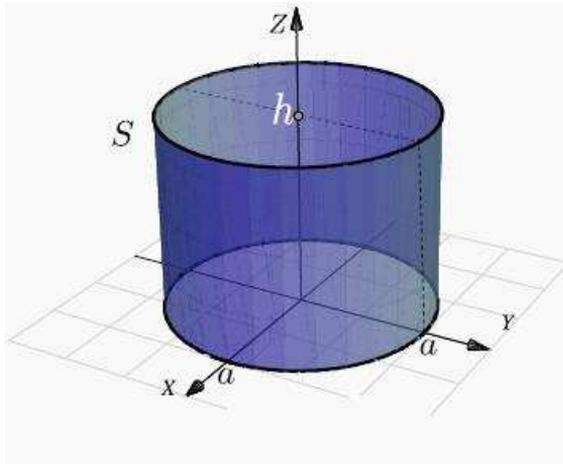
Figura 4.17 Cilindro parametrizado



Fuente: Adaptado (Mora, W. F. 2016).

4. Sea $S_1: X^2 + Y^2 = a^2, 0 \leq Z \leq H$. S es un cilindro que se parametriza como $r(\theta, Z) = a \cos(\theta) \hat{i} + a \sin(\theta) \hat{j} + Z \hat{k}$ con $(\theta, Z) \in D = [0, 2\pi] \times [0, H]$

Figura 4.18 Ejercicio propuesto 4



Fuente: (Mora, W. F. 2016).

4.4. Superficies Regulares

Superficies Regulares señala que sea $S: r(u, v) = X(u, v)\hat{i} + Y(u, v)\hat{j} + Z(u, v)\hat{k}$, $(u, v) \in D$. r es de clase C^1 si $X(u, v)$, $Y(u, v)$ y $Z(u, v)$ son de clase C^1 , funciones continuamente diferenciables. En este caso, considere los vectores $\frac{\partial r}{\partial u} = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right)$ y $\frac{\partial r}{\partial v} = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right)$. Los vectores $\frac{\partial r}{\partial u}\Big|_P$ y $\frac{\partial r}{\partial v}\Big|_P$ son tangentes a S en P .

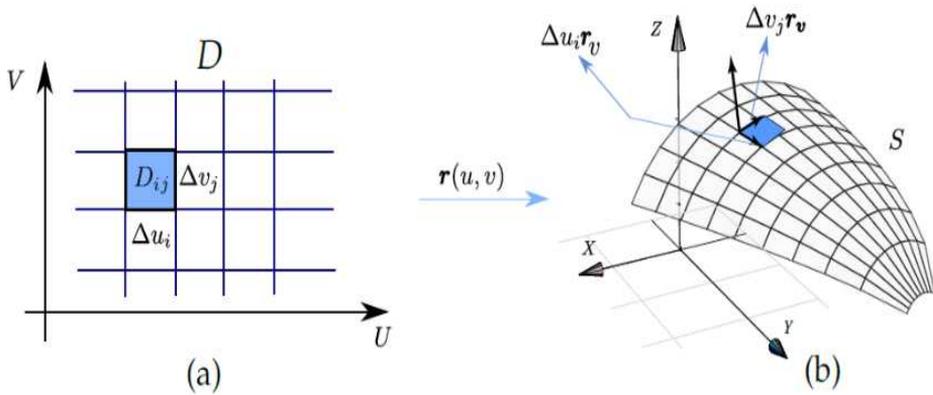
Definición: Sea D abierto y sea S una superficie parametrizada por $r: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ de clase C^1 . Se dice que S es una superficie regular en (u, v) si $\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \neq 0$. Si S se puede partir en un número finito de elementos regulares se dice regular a trozos.

Ejemplos:

1. Caso $S: Z = f(X, Y)$ entonces
 $r(X, Y) = X\hat{i} + Y\hat{j} + F(X, Y)\hat{k}$, $(X, Y) \in D$ con f_x y f_y continuas,
 $\frac{\partial r}{\partial X} \times \frac{\partial r}{\partial Y} = (-f_x, -f_y, 1) \neq \hat{0}$ tal que S es regular en D .

Área de una superficie presenta la definición de área de una superficie paramétrica consistente en aproximar el área de S sumando las áreas de «sus planos tangentes»; es decir, el área de paralelogramos generados por vectores escalados $\Delta u_i r_u$ y $\Delta v_j r_v$. Tal que, tomando el límite cuando $\Delta v_j \rightarrow 0$ y $\Delta u_j \rightarrow 0$:

Figura 4.19 Área de superficie



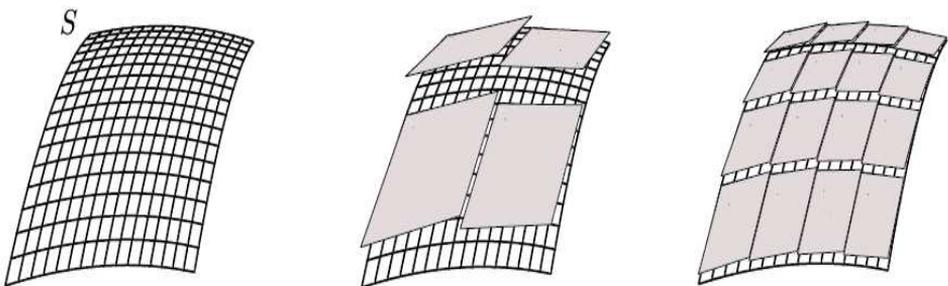
Fuente: (Mora, W. F. 2016).

Se parte que D en n rectángulos como (a), tal que en cada D_{ij} se toma el punto (u_i, v_j) más cercano al origen. En punto $r(u_i, v_j)$ de superficie S_1 , el Π Tan tiene ecuación $T_1: r(u_i, v_j) + tr_u(u_i, v_j) + sr_v(u_i, v_j)$ con $t, s \in \mathbb{R}$.

Tal que, la porción de superficie de S que es imagen de $a D_{ij}$ se puede aproximar con un paralelogramo en Π Tan de lados $\Delta u_i r_u, \Delta v_j r_v$. Ahora bien, el paralelogramo tiene área $\|\Delta u_i r_u \times \Delta v_j r_v\|$.

Sacando escalares y sumando el área de todos los paralelogramos se tiene que área de la superficie $S \approx \sum \|r_u \times r_v\| \Delta u_i \Delta v_j$:

Figura 4.206 Área de la superficie



Fuente: (Mora, W. F. 2016).

4.4.1. Área de Superficie

El Área de Superficie define que S es una superficie regular definida sobre un conjunto abierto medible D .

Entonces:

$$S: r(u, v) = X(u, v)\hat{i} + Y(u, v)\hat{j} + Z(u, v)\hat{k}, (u, v) \in D$$

Si se llama $dS = \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| dA$, el área A_S de superficie S es $A_S = \int \int_S 1 * dS = \int \int_D \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| dA$. Si $S = S_1 \cup \dots \cup S_m$ es unión infinita de superficies parametrizadas que intersecan a lo sumo en curvas que forman parte de sus fronteras, entonces $A_S = A_{S_1} + A_{S_1} + A_{S_2} + \dots + A_{S_k}$.

Ejemplos:

1. Caso $S: Z = f(X, Y)$, una parametrización es $r(X, Y) = X\hat{i} + Y\hat{j} + f(X, Y)\hat{k}$ y $\left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| = \sqrt{1 + f_X^2 + f_Y^2} \Rightarrow A_S = \int \int_S 1 * dS = \int \int_D \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| dA = \int \int_S 1 * dS = \int \int_D \sqrt{1 + f_X^2 + f_Y^2} dA$.

2. Caso $S: F(X, Y, Z) = 0$, donde S se puede proyectar uno a uno sobre una región D del $\Pi_{(XY)}$ y si F define a Z como función de X e Y y si $F_Z \neq 0 \Rightarrow Z_X = -\frac{F_X}{F_Z}$, $Z_Y = -\frac{F_Y}{F_Z}$ y la fórmula anterior es $A_S =$

$$\int \int_S 1 * dS = \int \int_D \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{|F_Z|} dA.$$

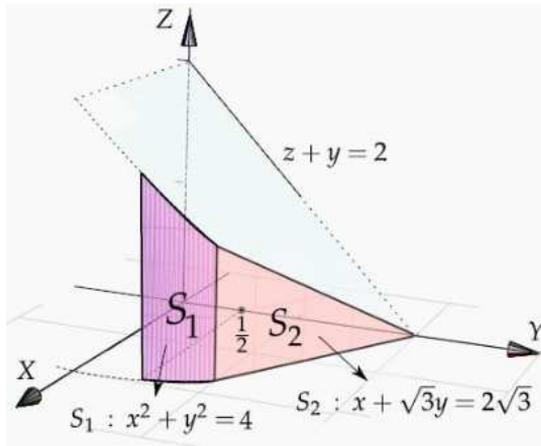
Área de una Superficie-Proyectando sobre varios planos indica que S es una superficie regular y que F es continuamente diferenciable e inyectiva sobre D .

- **Proyectando sobre $\Pi_{(XY)}$:** Si $S: Z = Z(X, Y)$ o $S: F(X, Y, Z = 0)$, con $(X, Y) \in D_{XY}$ tal que $A_S = \int \int_{D_{XY}} \sqrt{1 + Z_X^2 + Z_Y^2} dA$ o, en “versión implícita”, $A_S = \int \int_{D_{XY}} \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_Z} dA$ con $F_Z \neq 0$ en D_{XY} .
- **Proyectando sobre $\Pi_{(XZ)}$:** Si $S: Y = Y(X, Z)$ o $S: F(X, Y, Z = 0)$, con $(X, Z) \in D_{XZ}$ tal que $A_S = \int \int_{D_{XZ}} \sqrt{1 + Y_X^2 + Y_Z^2} dA$ o, en “versión implícita”, $A_S = \int \int_{D_{XZ}} \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_Y} dA$ con $F_Y \neq 0$ en D_{XZ} .
- **Proyectando sobre $\Pi_{(YZ)}$:** Si $S: X = X(Y, Z)$ o $S: F(X, Y, Z = 0)$, con $(Y, Z) \in D_{YZ}$ tal que $A_S = \int \int_{D_{YZ}} \sqrt{1 + X_Y^2 + X_Z^2} dA$ o, en “versión implícita”, $A_S = \int \int_{D_{YZ}} \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_X} dA$ con $F_X \neq 0$ en D_{YZ} .

Ejemplos:

- 1. Plantee integrales que dan el área de superficie $S = S_1 + S_2$ (coordenadas rectangulares):**

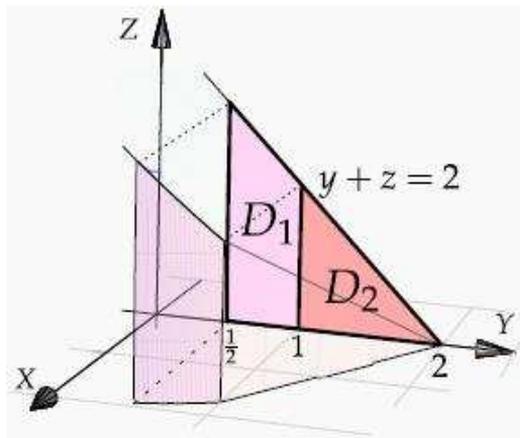
Figura 4.21 Área de superficie (coordenadas rectangulares)



Fuente: (Mora, W. F. 2016).

Se puede proyectar sobre $\Pi_{(YZ)}$ se tiene que $A_S = A_{S_1} + A_{S_2}$ siendo que superficie S_1 tiene ecuación $F(X, Y, Z) = X^2 + Y^2 - 4$, mientras que superficie S_2 tiene ecuación $X = 2\sqrt{3} - \sqrt{3}Y$:

Figura 4.22 Área de superficie con proyecciones

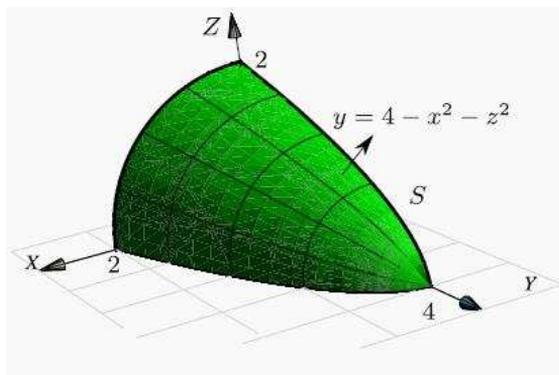


Fuente: (Mora, W. F. 2016).

$$\begin{aligned}
 A_S &= \int \int_{D_1} \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_X^2} dA + \int \int_{D_2} \sqrt{1 + Y_X^2 + Y_Y^2} dA \\
 &= \int_{\frac{1}{2}}^1 \int_0^{2-Y} \frac{\sqrt{4X^2 + 4Y^2 + 0^2}}{4X^2} dZdY \\
 &\quad + \int_1^2 \int_0^{2-Y} \sqrt{3 + 0 + 1} dZdY \\
 &= \int_{\frac{1}{2}}^1 \int_0^{2-Y} \frac{\sqrt{4(4 - Y^2) + 4Y^2 + 0^2}}{4(4 - Y^2)} dZdY \\
 &\quad + \int_1^2 \int_0^{2-Y} 2 dZdY \approx 2.19
 \end{aligned}$$

2. Calcule el área de superficie $S = Y + X^2 + Z^2 = 4$ (coordenadas rectangulares):

Figura 4.23 Área de superficie (coordenadas rectangulares)



Fuente: (Mora, W. F. 2016).

La proyección sobre $\Pi_{(YZ)}$ está limitada por el círculo $X^2 + Z^2 = 4$ y ecuación de superficie es $S: Y = 4 - X^2 - Z^2$.

Entonces,
$$A_S = \int \int_{D_{XZ}} \sqrt{1 + Y_X^2 + Y_Z^2} dA = \int \int_{D_{XZ}} \sqrt{1 + 4X^2 + 4Z^2} dA,$$
 con cambio de variable es:

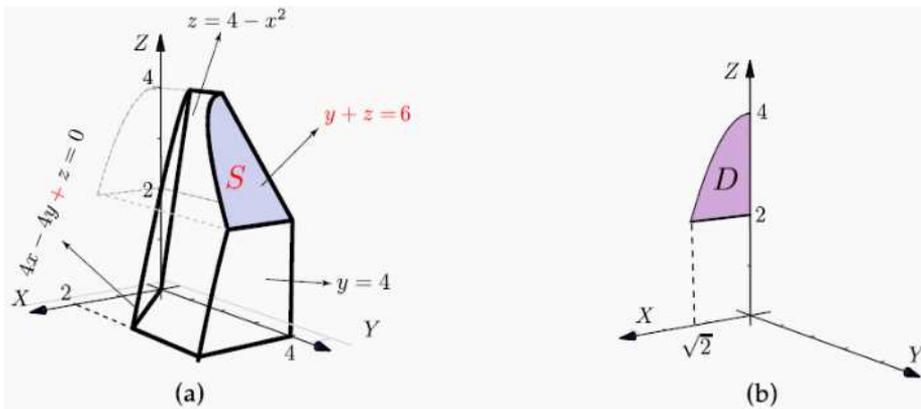
$$\begin{aligned} \begin{cases} X = r \cos(\theta) \\ Z = r \sin(\theta) \end{cases} &= \int_0^{\frac{\pi}{2}} \int_0^2 \sqrt{1 + 4r^2 \cos^2(\theta) + 4r^2 \sin^2(\theta)} r \, dr \, d\theta \\ &= \int_0^{\frac{\pi}{2}} \int_0^2 r \sqrt{1 + 4r^2} \, dr \, d\theta \\ &= \int_0^{\frac{\pi}{2}} \left. \frac{(1+4r^2)^{\frac{3}{2}}}{12} \right|_0^2 d\theta = \frac{\pi}{24} (17\sqrt{17} - 1) \approx 9.04. \end{aligned}$$

De otra forma, se usa la parametrización $r(Y, \theta) = \sqrt{4 - Y} \cos(\theta) \hat{i} + Y \hat{j} + \sqrt{4 - Y} \sin(\theta) \hat{k}$, $Y \in [0, 4]$ y $\theta \in \left[0, \frac{\pi}{2}\right]$ tal que

$$A_S = \int \int_D \left\| \frac{\partial r}{\partial Y} \times \frac{\partial r}{\partial \theta} \right\| dY d\theta = \int_0^{\frac{\pi}{2}} \int_0^4 \sqrt{\frac{17}{4} - Y} dY d\theta = \frac{\pi}{24} (17\sqrt{17} - 1) \approx 9.04.$$

3. Calcule el área de superficie S: $Y + Z = 6$:

Figura 4.24 Área de superficie



Fuente: (Mora, W. F. 2016).

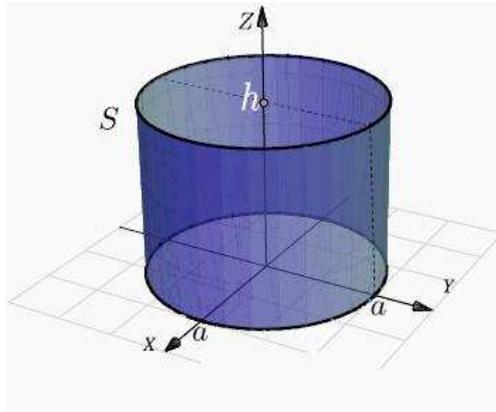
Se sabe que S: $Y(X, Z) = 6 - Z$, se usa la parametrización $r(Y, \theta) = X \hat{i} + (6 - Z) \hat{j} + Z \hat{k}$, sobre región D definida por $X \in [0, \sqrt{2}]$ y $2 \leq Z \leq 4 - X^2$.

Entonces, $Y_X = 0$ y $Y_Z = -1$. La proyección sobre D_{XZ} se ve en b) tal que

$$A_S = \int \int_{D_{XZ}} \sqrt{1 + Y_X^2 + Y_Z^2} \, dA = \int_0^{\sqrt{2}} \int_2^{4-X^2} \sqrt{2} \, dZ dX = \int_0^{\sqrt{2}} \sqrt{2} (2 - X^2) \, dX = \frac{8}{3}.$$

4. Calcule el área del cilindro $X^2 + Y^2 = a^2$ con altura $h \Rightarrow 0 \leq Z \leq h$:

Figura 4.25 Área de superficie



Fuente: (Mora, W. F. 2016).

La parametrización de esta superficie es $r(\theta, Z) = a \cos(\theta) \hat{i} + a \sin(\theta) \hat{j} + Z \hat{k}$ con $(\theta, Z) \in D = [0, 2\pi] \times [0, H]$ donde $r_\theta = (-a \sin(\theta), a \cos(\theta), 0)$, $r_Z = (0, 0, 1)$ y

$$\left\| \frac{\partial r}{\partial \theta} \times \frac{\partial r}{\partial Z} \right\| = \|a \sin(\theta), a \cos(\theta), 0\| = a \Rightarrow \int \int_D \left\| \frac{\partial r}{\partial \theta} \times \frac{\partial r}{\partial Z} \right\| \, dZ d\theta = \int_0^{2\pi} \int_0^h a \, dZ d\theta = 2 a h \pi.$$

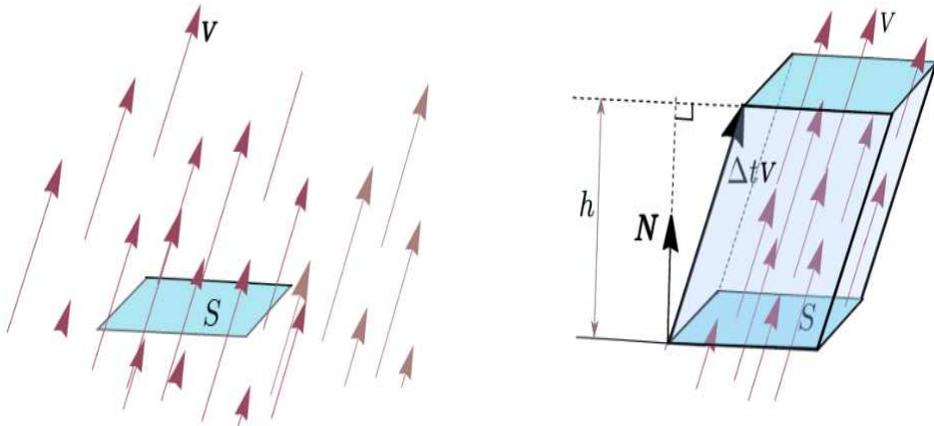
4.4.2. Integral sobre una Superficie

Integral sobre una superficie supone que tiene una región plana S y se quiere determinar la cantidad de «fluido» mediante S , suponiendo que

el fluido puede atravesar la región. La cantidad de fluido es «densidad por volumen».

Si el flujo se mueve con velocidad constante V , entonces durante el intervalo de tiempo Δt llenará un cuboide de base S y «extensión» (arista) ΔtV . El volumen de este cuboide es «área de base» ΔS por «altura» h , calculada con proyección de V sobre vector normal unitario ($\vec{\mu}$) a S , denotado por N . Se sabe que $h = \Delta tV \cdot N \Rightarrow V_S = V \cdot N \cdot \Delta S \cdot \Delta t$ (fluido sobre S_k ocupa un sólido cilíndrico):

Figura 4.26 Integral sobre una superficie



Fuente: (Mora, W. F. 2016).

La masa del fluido es $\Delta M = \rho_{(\text{Densidad})} \cdot V \cdot N \cdot \Delta S \cdot \Delta t$ tal que densidad del fluido es $F = \rho \cdot V$, mientras que flujo total es masa de fluido que pasa a través de S en una unidad de tiempo: $F \cdot N \Delta S - \text{Kg/S} -$.

Asimismo, se tiene una corriente de fluido en espacio con velocidad $V(X, Y, Z)$ y densidad (masa por unidad de volumen) $\rho(X, Y, Z)$ en cada punto (X, Y, Z) . El vector densidad de flujo $F(X, Y, Z) = V(X, Y, Z)\rho(X, Y, Z)$ tiene la misma dirección que la

velocidad y estima cuánta masa de fluido circula por punto (X, Y, Z) en dirección de $V(X, Y, Z)$ por unidad de área y tiempo.

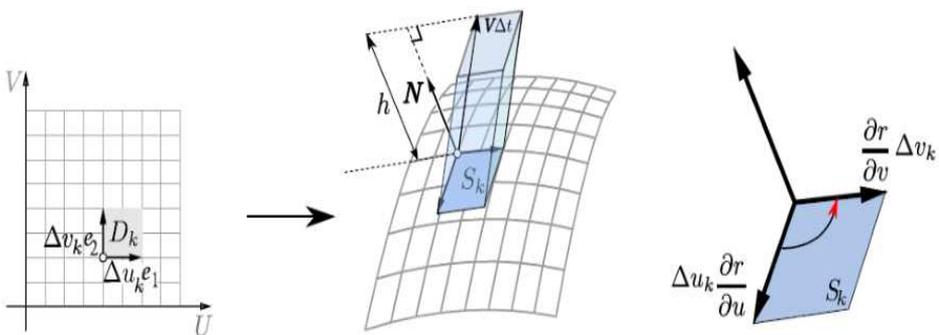
Una definición razonable de cómo medir la masa total de fluido que atraviesa una determinada superficie S en el tiempo, se considera la superficie S parametrizada por $r(u, v)$ en una región rectangular D . Sea N el $\vec{\mu}$ que tiene la misma dirección que el producto vectorial fundamental:

$$N = \frac{\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\|}$$

Para medir la cantidad de fluido que pasa a través S en unidad de tiempo y en dirección de N , se descompone el rectángulo D en m sub rectángulos $D_1, D_2, D_3, D_4, \dots, D_m$.

Sean $S_1, S_2, S_3, S_4, \dots, S_m$ sus correspondientes porciones de superficie en S . Se llamará ΔS_k a k -ésima porción S_k . Si densidad ρ y velocidad V son constantes en S_k , siendo S_k suficientemente plana, el fluido que atraviesa S_k en unidad de tiempo ocupa un sólido cilíndrico oblicuo con base S_k y eje determinado por vector velocidad V :

Figura 4.27 Sólido oblicuo cilíndrico



Fuente: (Mora, W. F. 2016).

Como el área de S_k es $\Delta S_k = \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| \Delta u_k \Delta v_k$, el fluido sobre S_k ocupa un sólido cilíndrico de volumen (base por altura), $\Delta S_k \rho V * N = F * N \Delta S_k \approx F * N \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| \Delta u_k \Delta v_k$.

Sugiere que la suma $\sum_{k=1}^m F * N \Delta S_k$ puede ser una aproximación de masa total de fluido que atraviesa ΔS_k en unidad de tiempo. Si se pone $f(X, Y, Z) = F * N$, se tiene la siguiente definición, tal que \bar{D} significa interior $D \cup \partial D$.

Definición: Sea D un abierto medible y S una superficie regular parametrizada por la función $r(u, v)$ de clase C^1 en \bar{D} , donde $(u, v) \in D$, de modo que $\left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| > 0$ para $\forall (u, v) \in D$ es una biyección entre D y S . Sea $f(X, Y, Z)$ una función definida y acotada sobre \bar{S} .

Se define la integral de superficie de f sobre S por $\int \int_S f(X, Y, Z) dS = \int \int_D f(r(u, v)) \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| dA$. Si $S = S_1 \cup S_2 \cup S_3 \cup S_4 \cup \dots \cup S_m$ es la unión finita de superficies parametrizadas que intersecan a lo sumo en curvas que forman parte de sus fronteras: $\int \int_S g(X, Y, Z) dS = \sum_i^m \int \int_{S_i} g(X, Y, Z) dS$.

4.4.3. Integral de Superficie con Coordenadas Rectangulares

Integral de Superficie con Coordenadas Rectangulares presenta el caso $S = f(X, Y)$ e indica que si $S = f(X, Y)$ con f de clase C^1 sobre \bar{D} , se puede parametrizar

$$r(X, Y) = X \hat{i} + Y \hat{j} + f(X, Y) \hat{k} \Rightarrow \int \int_S g(X, Y, Z) dS = \int \int_D g(X, Y, f(X, Y)) \sqrt{1 + f_X^2 + f_Y^2} dA.$$

Por otro lado, la Integral Superficie-Proyectando sobre varios planos asume que S es una superficie regular y que F es continuamente diferenciable e inyectiva sobre D :

- Proyectando sobre $\Pi_{(XY)}$: Si $S: Z = Z(X, Y)$ o $S: F(X, Y, Z) = 0$,
con $(X, Y) \in D_{XZ}$

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{XZ}} g((X, Y, Z(X, Y))) \sqrt{1 + Z_X^2 + Z_Y^2} dA$$

o, “versión implícita”,

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{XZ}} g((X, Y, Z(X, Y))) \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_Z^2} dA.$$

- Proyectando sobre $\Pi_{(XZ)}$: Si $S: Y = Y(X, Z)$ o $S: F(X, Y, Z) = 0$,
con $(X, Z) \in D_{XZ}$

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{XZ}} g(X, Y(X, Z), Z) \sqrt{1 + Y_X^2 + Y_Z^2} dA \text{ o,}$$

“versión implícita”,

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{XZ}} g(X, Y(X, Z), Z) \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_Y^2} dA.$$

- Proyectando sobre $\Pi_{(YZ)}$: Si $S: X = X(Y, Z)$ o $S: F(X, Y, Z) = 0$,
con $(Y, Z) \in D_{YZ}$

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{YZ}} g(X(Y, Z), Y, Z) \sqrt{1 + X_Y^2 + X_Z^2} dA \text{ o,}$$

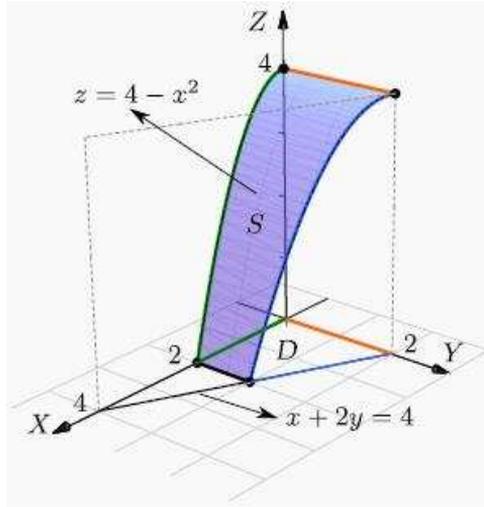
“versión implícita”,

$$\int \int_S g(X, Y, Z) dS = \int \int_{D_{YZ}} g(X(Y, Z), Y, Z) \frac{\sqrt{F_X^2 + F_Y^2 + F_Z^2}}{F_X^2} dA.$$

Ejercicios:

1. Calcule la integral de superficie $\int \int_S \frac{Z+X^2}{\sqrt{1+4X^2}} dS$ con S la porción de superficie $Z = 4 - X^2$ limitada por $\Pi_{(X+2Y=4)}$:

Figura 4.28 Ejercicio 1



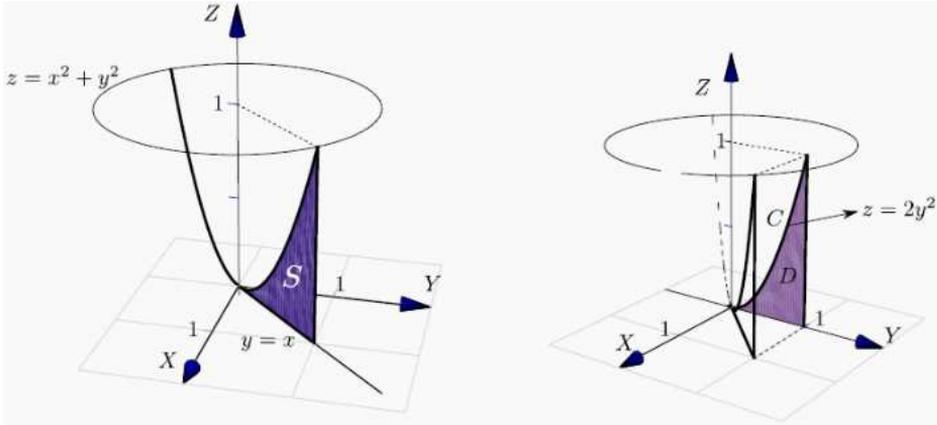
Fuente: (Mora, W. F. 2016).

Las coordenadas rectangulares $\sqrt{1 + Z_X^2 + Z_Y^2} = \sqrt{1 + 4X^2}$,

$$\iint_S \frac{Z+X^2}{\sqrt{1+4X^2}} dS = \iint_D \frac{4-X^2+X^2}{\sqrt{1+4X^2}} \sqrt{1 + 4X^2} dA = \int_0^2 \int_0^{2-\frac{x}{2}} 4 dy dx = 12.$$

2. Calcule la integral de superficie $\int_S 2XYZ \, dS$ con S la parte del $\Pi_{(Y=X)}$ limitado por $\Pi_{(Z=X^2+Y^2)}$:

Figura 4.29 Ejercicio 2

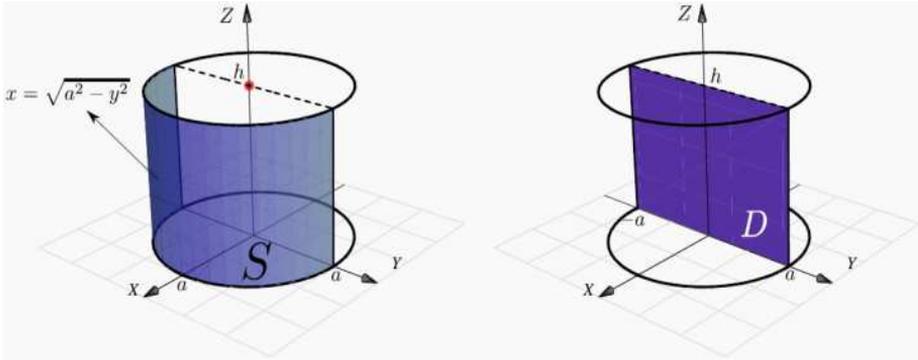


Fuente: (Mora, W. F. 2016).

La superficie S sólo se puede proyectar en $\Pi_{(XZ)}$ o $\Pi_{(YZ)}$. La curva C de proyección en $\Pi_{(YZ)}$ se obtiene como intersección del $\Pi_{(Y=X)}$ y paraboloides: $C: Y = X \cap Z = X^2 + Y^2 \Rightarrow C: Z = 2Y^2$. La proyección en $YZ \Rightarrow S: X = Y$ y $\sqrt{1 + X_Y^2 + X_Z^2} = \sqrt{2}$. Luego, $\int_S 2XYZ \, dS = \int_D 2XYZ \sqrt{1 + X_Y^2 + X_Z^2} \, dA = \int_0^1 \int_0^{2-Y^2} 2X^2Z \sqrt{2} \, dZ \, dY = \frac{4\sqrt{2}}{7}$.

3. Sea $a > 0, I = \int \int_S 2XYZ \, dS$ con S el cilindro $X^2 + Y^2 = a^2$ limitado por $\Pi_{(Z=0)}$ y $\Pi_{(Z=h)} > 0$:

Figura 4.30 Ejercicio 3



Fuente: (Mora, W. F. 2016).

- Proyectando sobre $\Pi_{(YZ)}$, $S: X = \sqrt{a^2 - Y^2}$.

$$\text{Tal que, } \sqrt{1 + X_Y^2 + X_Z^2} = \frac{a}{\sqrt{a^2 - Y^2}}$$

$$\int \int_S \frac{1}{a^2 + Z^2} \, dS = \int \int_D \frac{1}{a^2 + Z^2} \frac{a}{\sqrt{a^2 - Y^2}} \, dY dZ =$$

$$\int_{-a}^a \frac{a}{\sqrt{a^2 - Y^2}} \, dY \int_0^h \frac{1}{a^2 + Z^2} \, dZ \text{ (primera integral impropia),}$$

$$= \lim_{\varepsilon \rightarrow 0^+} a \operatorname{Arcsen}\left(\frac{Y}{a}\right) \Big|_{-a+\varepsilon}^{a-\varepsilon} * \frac{1}{a} \operatorname{Arctan}\left(\frac{Z}{a}\right) \Big|_0^h = \left(a \frac{\pi}{2} +$$

$$a \frac{\pi}{2}\right) \frac{1}{a} \operatorname{Arctan}\left(\frac{h}{a}\right).$$

- La manera más fácil, usando parametrización uno – uno:

$$S_1: r(\theta, Z) = a \operatorname{Cos}(\theta) \hat{i} + a \operatorname{Sen}(\theta) \hat{j} + Z \hat{k}, (\theta, Z) \in D =$$

$$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times [0, h] \text{ tal que}$$

$$r_\theta = (-a \operatorname{Sen}(\theta), a \operatorname{Cos}(\theta), 0), r_Z = (0, 0, 1),$$

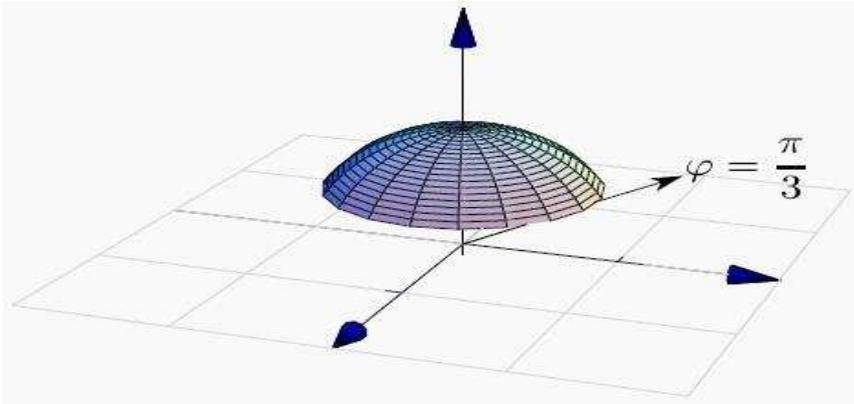
$$\left\| \frac{\partial r}{\partial \theta} \times \frac{\partial r}{\partial Z} \right\| = \|a \operatorname{Cos}(\theta), a \operatorname{Sen}(\theta), 0\| = a$$

$$\iint_S \frac{1}{a^2+z^2} dS = \iint_D \frac{1}{a^2+z^2} \left\| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial z} \right\| dz d\theta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^h \frac{1}{a^2+z^2} dz d\theta = \pi \operatorname{Arctan} \left(\frac{h}{a} \right).$$

Finalmente, usando parametrización no se tienen problemas de singularidades.

4. Considere la integral de superficie, $I = \iint_S \operatorname{Ln} Z dS$ con S el casquete de esfera $X^2 + Y^2 + Z^2 = 1, \frac{1}{2} \leq Z \leq 1$:

Figura 4.31 Ejercicio 4



Fuente: (Mora, W. F. 2016).

- Calcular I usando parametrización (coordenadas esféricas) si:

$$S_1: \mathbf{r}(\theta, \varphi) =$$

$$(\operatorname{Sen}(\varphi)\operatorname{Cos}(\theta), \operatorname{Sen}(\varphi)\operatorname{Sen}(\theta), \operatorname{Cos}(\varphi)), (\theta, \varphi) \in [0, 2\pi[\times \left[0, \frac{\pi}{3}\right]$$

:

Se usa parametrización del casquete de esfera basada en coordenadas esféricas.

Los parámetros son θ y φ . En este caso, $\rho = 1$:

$$\begin{cases} X = \text{Sen}(\varphi)\text{Cos}(\theta) \\ Y = \text{Sen}(\varphi)\text{Sen}(\theta) \Rightarrow r(\theta, \varphi) = (\text{Sen}(\varphi)\text{Cos}(\theta), \text{Sen}(\varphi)\text{Sen}(\theta), \text{Cos}(\varphi)), (\theta, \varphi) \in [0, 2\pi[\times [0, \frac{\pi}{3}] \\ Z = \text{Cos}(\varphi) \end{cases}$$

El valor $\varphi = \frac{\pi}{3}$ se obtiene de resolver $Z = 1 * \text{Cos}(\varphi) = \frac{1}{2}$ tal que

$$r_{\theta} = (-\text{Sen}(\theta)\text{Sen}(\varphi), \text{Cos}(\theta)\text{Sen}(\varphi), 0),$$

$$r_{\varphi} = (\text{Cos}(\theta)\text{Cos}(\varphi), \text{Cos}(\varphi)\text{Sen}(\varphi), -\text{Sen}(\varphi)) \text{ y}$$

$\left\| \frac{\partial r}{\partial \theta} \times \frac{\partial r}{\partial \varphi} \right\| = \text{Sen}(\varphi) > 0$ en $\left[0, \frac{\pi}{3}\right]$, las variables de integración son φ y θ , así que se sustituye Z en integración.

Para resolver la integral se sustituye $u = \text{Cos}(\varphi)$ tal que

$$\begin{aligned} \iint_S \text{Ln } Z \, dS &= \int_0^{2\pi} \int_0^{\frac{\pi}{3}} \text{Ln}(\text{Cos}(\varphi)) \text{Sen}(\varphi) \, d\varphi d\theta \\ &= - \int_0^{2\pi} \int_0^{\text{Cos}(\frac{\pi}{3})} \text{Ln}(u) \, du d\theta = \pi(\text{Ln}2 - 1) \end{aligned}$$

➤ Calcular I usando parametrización

$$S_1: r(Z, \theta) = \sqrt{1 - Z^2} \text{Cos}(\theta) \hat{i} + \sqrt{1 - Z^2} \text{Sen}(\theta) \hat{j} + Z \hat{k}, \frac{1}{2} \leq$$

$$Z \leq 1 \text{ y } \theta \in [0, 2\pi[.$$

Se sabe que $X^2 + Y^2 = 1 - Z^2$ con $\frac{1}{2} \leq Z \leq 1$ se parametriza el

$$\text{casquete con } r(Z, \theta) = \sqrt{1 - Z^2} \text{Cos}(\theta) \hat{i} + \sqrt{1 - Z^2} \text{Sen}(\theta) \hat{j} + Z \hat{k}, \frac{1}{2} \leq$$

$$Z \leq 1 \text{ y } \theta \in [0, 2\pi[\text{ y}$$

$$\left\| \frac{\partial r}{\partial Z} \times \frac{\partial r}{\partial \theta} \right\| = \left\| -\sqrt{1 - Z^2} \text{Cos}(\theta), -\sqrt{1 - Z^2} \text{Sen}(\theta), -Z \right\| = 1 \text{ tal que,}$$

en este caso, las variables de integración son Z y θ , así que no hay nada

$$\text{que sustituir en integral } \iint_S \text{Ln } Z \, dS = \int_0^{2\pi} \int_{\frac{1}{2}}^1 \text{Ln}(Z) \, dZ d\theta =$$

$$\int_0^{2\pi} \int_{\frac{1}{2}}^1 \text{Ln}(Z) * 1 \, dZ d\theta = \pi(\text{Ln}2 - 1).$$

➤ Calcule I usando coordenadas rectangulares

En coordenadas rectangulares $S: Z = \sqrt{1 - X^2 - Y^2}$ con $Z \in \left[\frac{1}{2}, 1\right] \Rightarrow$ proyección sobre $\Pi_{(XY)}$ está entre circunferencias $X^2 + Y^2 = \frac{3}{4}$ y $X^2 + Y^2 = 1$.

Las variables de integración son X e Y, así que se sustituye Z en

$$\text{integral} \quad \int \int_S \text{Ln}(Z) \, dS = \int \int_D \text{Ln}(Z) \sqrt{1 + Z_X^2 + Z_Y^2} \, dA =$$

$$\int \int_D \text{Log}(\sqrt{1 - X^2 - Y^2}) \sqrt{1 + \frac{X^2 + Y^2}{1 - X^2 - Y^2}} \, dA \text{ tal que se pasa a pasa a}$$

$$\text{coordenadas polares} = \int_0^{2\pi} \int_{\sqrt{\frac{3}{4}}}^1 \text{Log}(\sqrt{1 - r^2}) \frac{r}{\sqrt{1 - r^2}} \, dr \, d\theta, \text{ usando la}$$

sustitución $u^2 = 1 - r^2$, $= \pi(\text{Ln}2 - 1)$, siendo impropia la integral, tal que se calcula $u \rightarrow 0$.

4.5. Campos Escalares y Campos Vectoriales

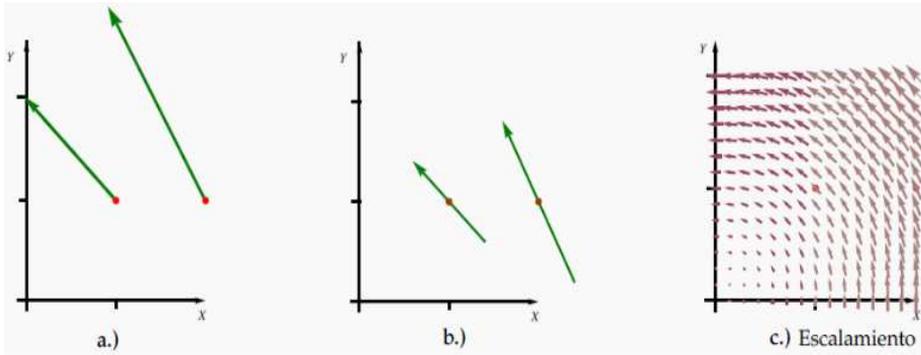
Campos Escalares y Campos Vectoriales se definen como sea $U \subseteq \mathbb{R}^n$ un conjunto abierto. Una aplicación $f: U \rightarrow \mathbb{R}$ se denomina campo escalar o función escalar. Una función $f: U \rightarrow \mathbb{R}^n$ se denomina campo vectorial.

Su representación gráfica consiste en anclar cada punto (X, Y) el respectivo vector $F(X, Y)$, pues se traslada desde el origen. Aunque, se puede anclar el vector tal que el punto quede en medio del vector, como si el vector fuera parte de una recta tangente.

En general, su representación gráfica se hace anclando el vector de esta segunda manera y escalando el tamaño de vectores tal que unos no se superpongan sobre otros, pues para tener una mejor visualización de la

dirección de «flujo» del campo vectorial. Por ejemplo: considere el campo $F(X, Y) = (-Y, X)$ con $F\left(\cos\left(\frac{\pi}{4}\right), \sin\left(\frac{\pi}{4}\right)\right)$:

Figura 4.32 Campos escalares y campos vectoriales



Fuente: (Mora, W. F. 2016).

Tal que en Figura 32 a.) se dibujan dos vectores anclados en punto, Figura b.) se dibujan dos vectores anclados con punto medio y en Figura c.) se grafica el campo escalando los vectores, como usualmente se hace.

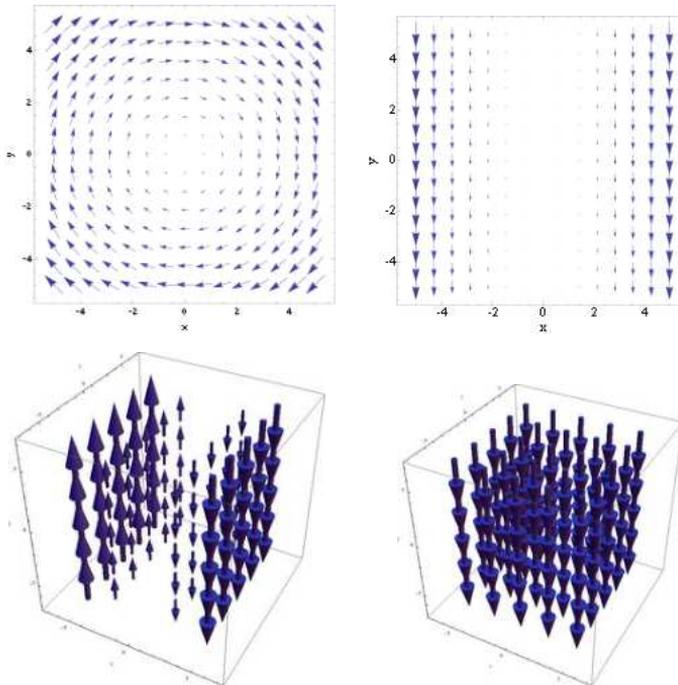
Complementariamente, el vector rotacional ($\overline{\text{Rot}}$) realiza una operación del campo vectorial que da como respuesta un vector ($\overline{\mu}$); es decir, $\text{Rot}(\overline{F}) = \nabla_X * \overline{F}$ que origina un producto multiplicador o vectorial que da un vector.

Es definido como un operador vectorial, operador lineal que actúa sobre campos vectoriales definidos sobre una variedad diferenciable, sobre campos vectoriales definidos en un abierto de \mathbb{R}^3 que muestra la tendencia de un campo vectorial a inducir rotación alrededor de un punto.

Matemáticamente, $U * \text{Rot}(\vec{F}) = U * \nabla_X * \vec{F} = \lim_{\Delta S \rightarrow 0} \frac{1}{\Delta S} \oint_C \vec{F} \, dr$ tal que ΔS es área de superficie apoyada en curva C , que se reduce a un punto.

Su resultado no es el rotacional completo, un vector, sino solo su componente según la dirección normal a ΔS y orientada según la regla de la mano derecha tal que para obtener el rotacional completo se calculan tres límites, considerando tres curvas situadas en $\Pi \perp$:

Figura 4.33 Campo vectorial

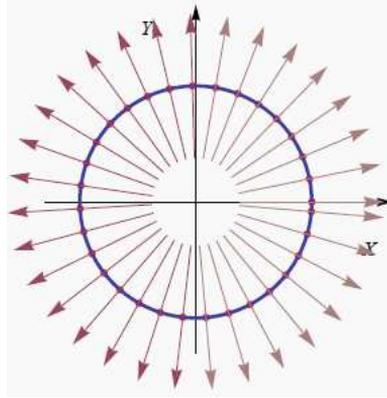


Fuente: (Mora, W. F. 2016).

Ejercicios propuestos

1. Represente gráficamente el campo vectorial $F(X, Y) = (2X, 2Y)$ sobre circunferencia $X^2 + Y^2 = 1$ tal que si $Z = X^2 + Y^2 \Rightarrow F(X, Y) = \Delta Z$ tal que los vectores son perpendiculares a esta circunferencia (curva de nivel $Z = 1$):

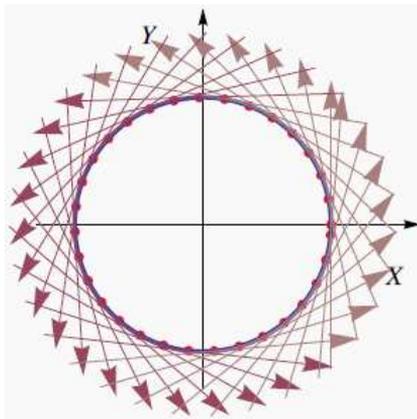
Figura 4.34 Ejercicio propuesto 1



Fuente: (Mora, W. F. 2016).

2. Represente gráficamente el campo vectorial $F(X, Y) = (-Y, X)$ sobre circunferencia $X^2 + Y^2 = 1$:

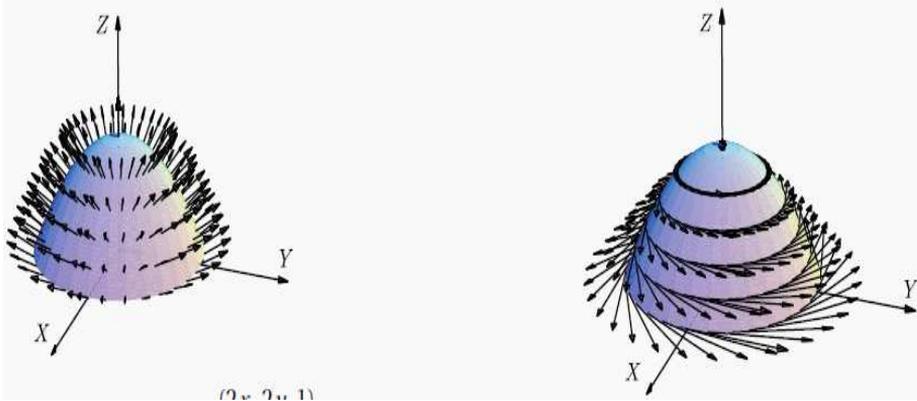
Figura 4.35 Ejercicio propuesto 2



Fuente: (Mora, W. F. 2016).

3. Sea S la superficie de ecuación $Z = -X^2 - Y^2 + 1$, represente gráficamente los campos vectoriales sobre S :

Figura 4.36 Ejercicio propuesto 3



Fuente: (Mora, W. F. 2016).

En la primera figura se representa $F_1(X, Y, Z) = \frac{(2X, 2Y, 1)}{3\sqrt{4X^2 + 4Y^2 + 1}}$ y, en la segunda figura, $F_2(X, Y, Z) = (-Y\sqrt{1-Z}, X\sqrt{1-Z}, 0)$.

4.6. Integral de Flujo

Integral de Flujo sostiene que si F es densidad de flujo de una corriente

de fluido y N es vector unitario normal a S definido por $N = \frac{\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\|} \Rightarrow$

masa total de fluido que pasa por S por unidad de tiempo en dirección

N es $\iint_S F * N \, dS = \iint_D F(r(u, v)) * \frac{\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\|} \left\| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right\| \, dA =$

$\iint_D F(r(u, v)) * \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \, dA.$

Presenta el Caso $Z = f(X, Y)$ que indica que como consecuencia si

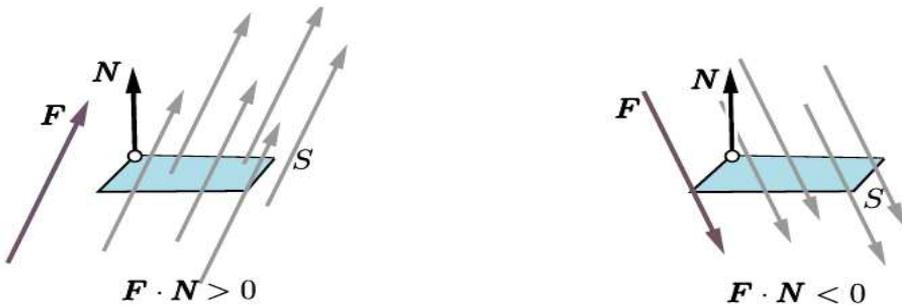
$S: Z = f(X, Y)$ con f de clase C^1 sobre \bar{D} se puede parametrizar S con

$$r(X, Y) = X\hat{i} + Y\hat{j} + f(X, Y)\hat{k} \Rightarrow \iint_S \mathbf{F} \cdot \mathbf{N} \, dS = \iint_{D_{XY}} \mathbf{F}(X, Y, Z) \cdot (-f_X, -f_Y, 1) \, dA.$$

La Orientación es la expresión para el flujo total que lleva implícita la selección de uno de dos $\vec{\mu}$ (elección de \mathbf{N} decide el signo de $\mathbf{F} \cdot \mathbf{N}$).

Escoger un vector unitario para le región S es equivalente a «orientar» la región:

Figura 4.37 Integral de Flujo-Proyectando



Fuente: (Mora, W. F. 2016).

Entonces, en adelante se escoge $\frac{\frac{\partial r}{\partial u} \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \frac{\partial r}{\partial v} \right\|}$. La Integral de Flujo-Proyectando sobre varios planos:

- Proyectando sobre $\Pi_{(XY)}$: Si $S: Z = Z(X, Y)$ o $S: G(X, Y, Z) = 0$ con $(X, Y) \in D_{XY}$ $\iint_S \mathbf{F} \cdot \mathbf{N} \, dS = \iint_{D_{XY}} \mathbf{F}(X, Y, Z(X, Y)) \cdot (-Z_X, -Z_Y, 1) \, dA$ o, en versión “implícita”, $\iint_S \mathbf{F} \cdot \mathbf{N} \, dS = \iint_{D_{XY}} \mathbf{F}(X, Y, Z(X, Y)) \cdot (G_X, G_Y, G_Z) \frac{1}{G_Z} \, dA$.
- Proyectando sobre $\Pi_{(XZ)}$: Si $S: Y = Y(X, Z)$ o $S: G(X, Y, Z) = 0$ con $(X, Z) \in D_{XZ}$ $\iint_S \mathbf{F} \cdot \mathbf{N} \, dS = \iint_{D_{XZ}} \mathbf{F}(X, Y(X, Z), Z) \cdot$

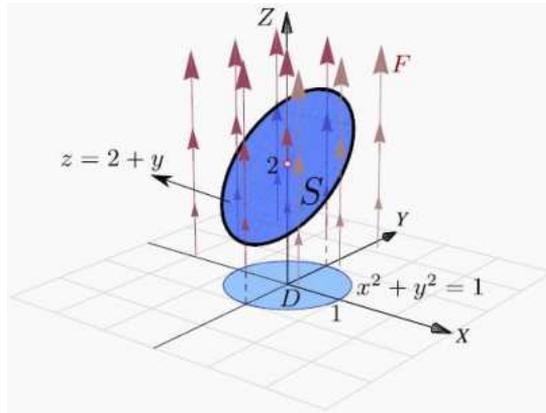
$$(-Y_X, 1, -Y_Z) dA \text{ o, en versión "implícita", } \iint_S F * N dS = \iint_{D_{XZ}} F(X, Y(X, Z), Z) * (G_X, G_Y, G_Z) \frac{1}{G_Y} dA.$$

- **Proyectando sobre $\Pi_{(YZ)}$:** Si $S: X = X(X, Z)$ o $S: G(X, Y, Z) = 0$ con $(Y, Z) \in D_{YZ}$ $\iint_S F * N dS = \iint_{D_{YZ}} F(X(Y, Z), Y, Z) * (1, -X_Y, -X_Z) dA$ o, en versión "implícita", $\iint_S F * N dS = \iint_{D_{YZ}} F(X(Y, Z), Y, Z) * (G_X, G_Y, G_Z) \frac{1}{G_X} dA.$

Ejemplos:

- 1. Calcule $\iint_S F * N dS$ si $F(X, Y, Z) = (Z + 1)\hat{K}$ y S es superficie $Z = 2 + Y$ con $X^2 + Y^2 = 1$:**

Figura 4.38 Ejemplo 1



Fuente: (Mora, W. F. 2016).

Se sabe que la superficie S tiene ecuación $Z = 2 + Y$ tal que D_{YZ} es círculo de radio 1.

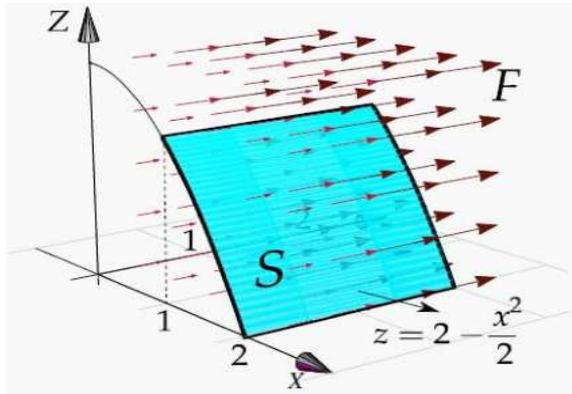
Entonces, $\iint_S F * N dS = \iint_{D_{XY}} F(0, 0, Z + 1) * (-Z_X, -Z_Y, 1) dA$

$$= \int \int_{D_{XY}} F(0, 0, 2 + Y + 1) * (0, -1, 1) dA$$

$$= \int \int_{D_{XY}} Y + 3 dA = \int_0^{2\pi} \int_0^1 (3 + r \text{Sen}(\theta)) dr d\theta = \int_0^{2\pi} \frac{6 + \text{Sen}(\theta)}{4} d\theta = 3\pi.$$

2. Sea $F(X, Y, Z) = (0, Y, 0)$ y S el cilindro $Z = 2 - \frac{X^2}{2}$, desde $Y = 0$ hasta $Y = 2$:

Figura 4.39 Ejemplo 2



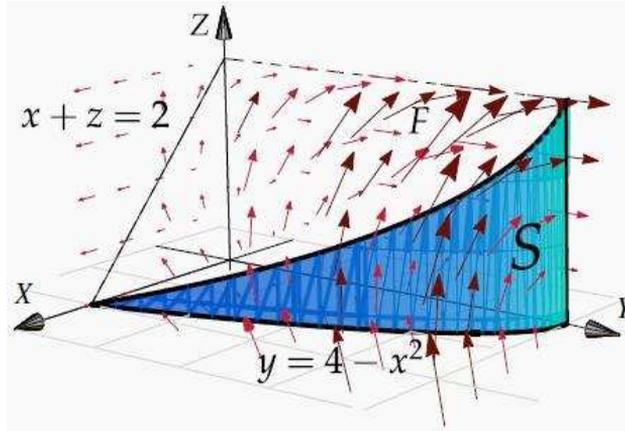
Fuente: (Mora, W. F. 2016).

Calcule $\int \int_S F * N dS$. Intuitivamente, el flujo no pasa a través de la superficie S tal que la integral de flujo sería 0 y, en consecuencia, solo se puede proyectar sobre $\Pi_{(XY)}$ o $\Pi_{(YZ)}$.

La proyección sobre $\Pi_{(YZ)}$ es un rectángulo, tal que sea $S: G(X, Y, Z) = 0$ con $G(X, Y, Z) = Z - 2 + \frac{X^2}{2} \Rightarrow \int \int_S F * N dS = \int \int_{D_{YZ}} F * \frac{\Delta G}{G_X} dA = \int_0^2 \int_0^{\frac{3}{2}} (0, Y, 0) * (X, 0, 1) \frac{1}{X} dZ dY = \int_0^2 \int_0^{\frac{3}{2}} 0 dZ dY = 0.$

3. Calcular $\int \int_S F * N dS$ si $F(X, Y, Z) = (X, YZ, XY)$ y S el cilindro de ecuación $Y = 4 - X^2$ limitado por $\Pi_{(X+Z=2)}$:

Figura 4.40 Ejemplo3



Fuente: (Mora, W. F. 2016).

Se sabe que la superficie S tiene ecuación $G(X, Y, Z) = 0$ con $G(X, Y, Z) = Y + X^2 + 4$ tal que la proyección de S sobre $\Pi_{(XZ)}$ es un triángulo.

Entonces,

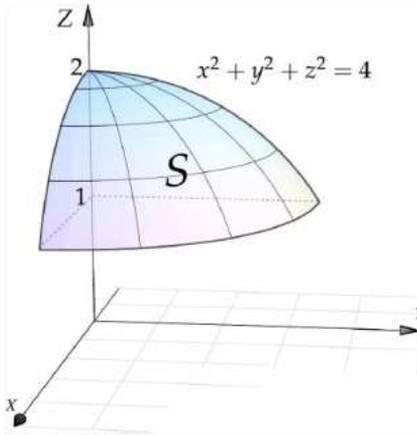
$$\int \int_S \mathbf{F} \cdot \mathbf{N} \, dS = \int \int_{D_{XZ}} \mathbf{F} \cdot \frac{\Delta G}{G_x} \, dA = \int \int_{D_{XZ}} (X, YZ, XY) \cdot \frac{\Delta G}{G_x} \, dA$$

donde $S: G(X, Y, Z) = Y + X^2 + 4 = 0$

$$\begin{aligned} &\Rightarrow \int \int_{D_{XZ}} (X, YZ, XY) \cdot (2X, 1, 0) \, dA \\ &= \int_0^2 \int_0^{2-X} 2X^2 + YZ \, dZ \, dX \\ &= \int_0^2 \int_0^{2-X} 2X^2 + (4 - X^2)Z \, dZ \, dX = \frac{112}{15}. \end{aligned}$$

4. Calcule $\int \int_S \mathbf{F} * \mathbf{N} \, dS$ si $\mathbf{F}(X, Y, Z) = (X, YZ, XY) = 4XZ \hat{i} + YZ^3 \hat{j} + Z^2 \hat{k}$ y S es la superficie $Z^2 + Y^2 + X^2 = 4$ entre $Z = 1$ y $Z = 2$:

Figura 4.41 Ejemplo 4



Fuente: (Mora, W. F. 2016).

La superficie S tiene ecuación $G(X, Y, Z) = 0$ con $G(X, Y, Z) = Z^2 + Y^2 + X^2 - 4$.

La proyección D_{XY} es círculo $Y^2 + X^2 = 3$

$$\begin{aligned} \Rightarrow \int \int_S \mathbf{F} * \mathbf{N} \, dS &= \int \int_{D_{XY}} (4XZ, YZ^3, Z^2) * \frac{\Delta G}{G_Z} \, dA \\ &= \int \int_{D_{XY}} (4XZ, YZ^3, Z^2) * \left(\frac{X}{Z}, \frac{Y}{Z}, 1 \right) \, dA \\ &= \int \int_{D_{XY}} 4X^2 + Y^2 Z^2 + Z^2 \, dA \\ &= \int \int_{D_{XY}} 4X^2 + Y^2(4 - X^2 - Y^2) + 4 - X^2 - Y^2 \, dA \end{aligned}$$

$$= \frac{1}{2} \int_0^{\frac{\pi}{2}} \int_0^{\sqrt{3}} (8 + 6r^2 - r^4 + r^4 \cos(2\theta)) r \, dr \, d\theta = \frac{21}{4} \pi.$$

Con base en esto y de forma complementaria, la teoría económica sugiere relaciones causales entre las variables que son muy complejas como para ser representadas en un modelo uniecuacional.

Por esta razón, si las variables están mutuamente determinadas o, en determinación simultánea, se construyen los llamados modelos de ecuaciones simultáneas.

El modelo de oferta y demanda recién visto es un ejemplo clásico:

$$D_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \varepsilon_t$$
$$O_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + v_t$$

$$\text{Relacion de equilibrio: } O_t = D_t$$

4.7. Modelos autorregresivos y de rezagos distribuidos.

Se trata de modelos en que variables en el tiempo t son una función de sí mismas; es decir, son rezagadas en el tiempo $t - 1$, $t - 2$, etcétera. Según Gujarati y Porter (2010), en análisis de regresión con datos de series de tiempo, si el modelo de regresión incluye no solo valores actuales sino valores rezagados de variables explicativas (X 's) el modelo se llama «Modelo de Rezagos Distribuidos», cuyo modelo es $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$. Si el modelo incluye uno o más valores rezagados de variable dependiente (Y) entre sus variables explicativas, el modelo recibe el nombre de «Modelo Autorregresivo» o «Modelos Dinámicos», cuyo modelo es $Y_t = \alpha + \beta X_t + \gamma X_{t-1} + u_t$ y señalan la trayectoria en tiempo de variable dependiente en relación con su valor (s) pasado (s).

En economía, la dependencia de una variable Y respecto de otra u otras X 's pocas veces es instantánea. Usualmente, Y responde a X en un lapso, denominado rezago. En forma más general, el modelo de rezagos distribuidos con un rezago finito de periodos es k es $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + \dots + \beta_k X_{t-k} + u_t$.

El coeficiente β_0 es conocido como "Multiplicador de corto plazo o de impacto", pues da el cambio en valor medio de Y que sigue a un cambio unitario en X en mismo periodo. Si el cambio en X se mantiene igual desde el inicio, entonces $(\beta_0 + \beta_1)$ da el cambio en valor medio de Y en periodo siguiente $(\beta_0 + \beta_1 + \beta_2)$ en que le sigue y, sucesivamente, así. Estas sumas parciales se llaman «Multiplicadores ínterin o intermedio».

Finalmente, después de k periodos se tiene:

$$\sum_{i=0}^k \beta_i = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \dots + \beta_k = \beta$$

Conocido como «Multiplicador de rezagos distribuidos de largo plazo o total» siempre que exista la suma β tal que si se define $\beta_i^* = \frac{\beta_i}{\sum \beta_i} = \frac{\beta_i}{\beta}$ se obtiene β_i "estandarizado". Las sumas parciales del β_i estandarizado dan proporción del impacto de largo plazo o total, sentido durante cierto periodo.

Existen tres razones principales porque existen los rezagos:

- **Psicológicas.** Es resultado del hábito inerte, pues la gente no cambia sus hábitos de consumo de inmediato tras una reducción de precios o un incremento en ingreso, quizá debido a que el proceso de cambio conlleve alguna desventaja

inmediata. Asimismo, la gente puede no saber si un cambio es «permanente» o «transitorio» tal que su reacción a un incremento en su ingreso depende que el incremento sea permanente o no.

- **Tecnológicas.** En un caso suponga que el precio del capital relativo al trabajo, tal que es económicamente factible sustituir mano de obra por capital. La adición de capital toma tiempo, gestación y, además, si espera la caída de precios sea temporal, las empresas pueden no apurarse a sustituirla por capital, especialmente si esperan que después de la caída temporal el precio del capital tal vez aumente más allá de su nivel anterior.

- **Institucionales.** Un ejemplo son las obligaciones contractuales pueden impedir que empresas cambien de una fuente de trabajo o de materias primas a otra o quienes colocan fondos en cuentas de ahorro de largo plazo con término fijo están «atrapados» aunque las condiciones del mercado de dinero ahora permitan rendimientos más altos en otras partes.

La Estimación de Modelos de Rezagos Distribuidos desempeña un papel muy útil en economía por lo que su estimación en una variable explicativa es $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$ donde no se ha definido la longitud del rezago; es decir, cuán atrás se desea ir. Tal modelo se llama «Modelo de Rezagos Infinito», mientras que un modelo $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + \dots + \beta_k X_{t-k} + u_t$ se denomina “Modelo de Rezagos Distribuidos -Rezagos-Finito”, pues la longitud del rezago k está especificada.

La Estimación de Modelos Autorregresivos presentan las siguientes opciones:

- Koyck: $Y_t = \alpha(1 - \lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t$.
- Expectativas adaptativas: $Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1 - \gamma)Y_{t-1} + [u_t - (1 - \gamma)u_{t-1}]$.
- Ajuste parcial: $Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1 - \delta)Y_{t-1} + \delta u_{t-1}$

Todos estos modelos tienen la forma común $Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 X_{t-1} + v_t$ por lo que todos son autorregresivos por naturaleza. Entonces, el problema de estimación de dichos modelos se presenta, pues los Mínimos Cuadrados Clásicos pueden no ser aplicables directamente a estos debido a presencia de variables explicativas estocásticas y posibilidad de correlación serial.

Esto se debe a que, si una variable explicativa en un modelo de regresión está correlacionada con término de perturbación estocástico, los estimadores MCO están sesgados y no son consistentes; es decir, aunque el tamaño de la muestra aumente indefinidamente, los estimadores no se aproximarán a sus valores poblacionales verdaderos y, por ende, la estimación de modelos Koyck y expectativas adaptativas mediante el procedimiento usual MCO puede producir resultados muy erróneos.

4.8. Modelos de series de tiempo

Existen métodos para pronosticar que han adquirido mucha popularidad, como Autorregresivo Integrado de Promedios Móviles (ARIMA) o conocido como metodología Box-Jenkins, Vectores Autorregresivos (VAR), modelos caracterizados por un fenómeno conocido como «Acumulación de Volatilidad», en que existen lapsos en que muestran amplias variaciones durante prolongados periodos, seguidos por un intervalo de tranquilidad relativa, los llamados Modelos con Heterocedasticidad Condicional Autorregresiva (ARCH)

o Modelos con Heterocedasticidad Condicional Autorregresiva Generalizada (GARCH) representan la Acumulación de Volatilidad.

No obstante, hay cinco enfoques de pronósticos económicos basados en series de tiempo:

- **Métodos de suavizamiento exponencial.** Son métodos para ajustar una curva apropiada a datos históricos de una determinada serie de tiempo. Existen diversos tipos, como suavizamiento exponencial simple, método lineal de Holt y método de Holt-Winters, así como variaciones, aunque aún se emplean en varias áreas para pronósticos en negocios y economía.

- **Modelos de regresión uniecuacionales.** Un ejemplo es la función de demanda de automóviles. La teoría económica postula que está en función de sus precios, gasto de publicidad, ingreso del consumidor, tasa de interés, como medida de costo de endeudamiento, tamaño de familia o distancia de trabajo. A partir de series de tiempo se estima un modelo apropiado de demanda de automóviles, como lineal, log-lineal o no lineal, que sirve para el pronóstico de demanda futura.

- **Modelos de regresión de ecuaciones simultáneas.** En las décadas de 1960 y 1970 estos modelos elaborados para describir la economía de EUA basados dominaron el campo de los pronósticos económicos, pero termino por las crisis del precio del petróleo de 1973 y 1979, así como la crítica de Lucas, pues argumenta que los parámetros estimados de un modelo econométrico dependen de la política prevaleciente en el momento en que se estima el modelo y cambian conforme la

política aunque los parámetros no son invariantes ante cambios políticos tal que con base en información pasada tendría poco valor predictivo con el nuevo régimen.

- **Modelos autorregresivos integrados de promedios móviles (ARIMA).** La publicación de *Box* y *Jenkis* titulada “*Time Series Analysis: Forecasting and Control*” (Análisis de Series de tiempo: Predicción y Control) marcó el inicio de la metodología de *Box-Jenkis*, aunque es conocida como metodología ARIMA.

Su interés en pronósticos no está en construcción de modelos uniecuacionales o simultáneas, sino en análisis de propiedades probabilísticas o estocásticas de series de tiempo económicas por sí mismas según la filosofía “que los datos hablen por sí mismos” que, a diferencia de modelos de regresión en que Y_t se explica por k regresoras $X_1, X_2, X_3, X_4 \dots, X_k$ en modelos de series de tiempo tipo *Box-Jenkis* tal que Y_t se explica por valores pasados o rezagados de sí misma y por términos de error estocásticos.

Por esta razón, los modelos ARIMA reciben el nombre de modelos ateóricos debido a que no se derivan de teoría económica alguna y, por ende, a menudo las teorías económicas son a menudo la base de modelos de ecuaciones simultáneas. Finalmente, los modelos ARIMA pueden pertenecer a una sola serie de tiempo o, incluso, hay ARIMA multivariados.

- **Modelos de vectores autorregresivos (VAR).** La metodología VAR se asemeja, a primera vista, a modelos de ecuaciones simultáneas, pues considera diversas variables endógenas

conjuntas. No obstante, cada variable endógena se explica por sus valores rezagados o pasados y, también, por valores rezagados de todas las demás variables endógenas en modelo, pero usualmente no hay variables exógenas en modelo.

Son modelos probabilísticos que se suponen dan origen a una serie de valores muestrales de una variable en el tiempo. Modelos clásicos son los autorregresivos, en particular en series temporales un modelo autorregresivo integrado de promedio móvil o ARIMA, acrónimo del inglés *Autoregressive Integrated Moving Average*, es un modelo dinámico estadístico de series temporales creado por Box y Jenkins que utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para una predicción hacia el futuro, modelos autorregresivos de media móvil ARMA, formado por una parte autorregresiva (AR) y otra de media móvil (MA) o, en inglés, *Autoregressive Moving Average models*, también llamados Modelos *Box-Jenkins*, se aplican a series temporales de datos como una herramienta para entender y, aún más, para predecir futuros valores de la serie y promedios móviles.

Sin embargo, la Creación de modelos AR, PM y ARIMA para series de tiempo estacionarios se puede modelar de diversas formas:

- **Proceso Autorregresivo (AR).** Si Y_t es logaritmo del PIB en periodo t , se puede modelar como $(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + u_t$ tal que δ es media de Y y u_t término de error aleatorio no correlacionado con media cero y varianza constante σ^2 conocido como ruido blanco, tal que se dice que Y_t sigue un proceso estocástico autorregresivo de primer orden o AR(1) por lo que el valor de Y en el tiempo t depende de su valor en periodo anterior y un determinado aleatorio.

Los valores Y están expresados como desviaciones de su valor medio. Es decir, este modelo dice que el valor de pronóstico de Y en periodo t es simplemente alguna proporción ($= \alpha_1$) de su valor inicial ($t - 1$) más un «choque» o perturbación aleatoria en tiempo t ; aunque, los valores Y están expresados entorno del valor de su media.

No obstante, si se considera el modelo $(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + u_t$ se dice que Y_t sigue un Proceso Autorregresivo de Segundo Orden o AR(2). Es decir, el valor Y en tiempo t depende de sus valores en dos periodos anteriores, valores de Y expresados entorno de valor de su media δ .

Sin embargo, en general, se tiene $(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \alpha_3(Y_{t-3} - \delta) + \alpha_4(Y_{t-4} - \delta) + \dots + \alpha_p(Y_{t-p} - \delta) + u_t$ tal que Y_t es un Proceso Autorregresivo de Orden p o AR(p). En este sentido, se dice que “los datos hablan por sí mismos”, pues solo consideran valores actuales y anteriores de Y .

- **Proceso de Medias Móviles (MA).** El proceso AR no es el único mecanismo que pudo generar a Y . Suponga que se hace un modelo de Y : $Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1}$. Donde μ es una constante y u es término de error estocástico de ruido blanco, Y en periodo t es una constante más promedio móvil de términos de error presente y pasado.

Entonces, Y sigue un proceso de promedio móviles de primer orden o MAE. En cambio, si Y sigue la expresión $Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2}$ es un proceso MA (2). Aunque, en

forma más general $Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \beta_3 u_{t-3} + \beta_4 u_{t-4} + \dots + \beta_q u_{t-q}$ es un proceso MA(q).

Por lo tanto, un proceso de promedios móviles es tan solo una combinación lineal de términos de error de ruido blanco.

➤ **Proceso Autorregresivo y de Promedios Móviles (ARMA).**

Es muy probable que Y tenga características de AR y MA a la vez y, por consiguiente, sea ARMA. Así, Y_t sigue un proceso ARMA(1,1) si se escribe como $Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1}$ tal que θ representa un término constante.

En general, es un proceso ARMA (p, q) donde habrá p términos autorregresivos y q términos de promedios móviles.

➤ **Proceso Autorregresivo Integrado de Promedios Móviles (ARIMA).**

Los modelos de series de tiempo analizados se basan en supuesto que las series de tiempo consideradas son, débilmente, estacionarias en sentido que el trabajo empírico en series de tiempo supone que la serie de tiempo en cuestión es estacionaria respecto a la autocorrelación se debe a varias causas, como series de tiempo subyacentes no estacionarias, al efectuar la regresión de una variable de serie de tiempo sobre otra variable de serie de tiempo con frecuencia se obtiene una R^2 muy elevada (≥ 0.9) aunque no hay una relación significativa entre ambas.

En ocasiones no se espera ninguna relación entre las dos variables; sin embargo, una regresión de una variable sobre la otra a menudo muestra una relación significativa tal que esta situación ejemplifica el problema de regresión espuria o disparada.

Algunas series de tiempo financieras, como precios de acciones, muestran el fenómeno de caminata aleatoria, indicando que la mejor predicción para el precio de una acción,

como IBM, es igual a su precio actual más un choque puramente aleatorio (término error). De ser así, el pronóstico del precio de acciones sería un ejercicio inútil.

Los modelos de regresión que consideran series de tiempo son muy comunes para pronósticos tal que si se desea saber tal pronóstico es válido cuando las series de tiempo sobre las que se basa no son estacionarias y, por último, las pruebas de casualidad de *Granger* y *Sims* suponen que las series de tiempo del análisis son estacionarias.

En consecuencia, las pruebas para la estacionariedad que las de causalidad deben efectuarse. Con base en esto, la media y varianza de una serie de tiempo débilmente estacionaria son constantes y su covarianza es invariante en el tiempo.

Entonces, si se debe diferenciar una serie de tiempo d veces para hacerla estacionaria y, luego, aplicarle el modelo ARMA (p, q) se dice que la serie de tiempo original es ARIMA (p, d, q) ; es decir, es una serie de tiempo autorregresiva integrada de promedios móviles, donde p denota el número de términos autorregresivos, d el número de veces que la serie debe diferenciarse para hacerse estacionaria y q el número de términos de promedios móviles.

Así, una serie de tiempo ARIMA $(2, 1, 2)$ se diferencia una vez ($d = 1$) antes que se haga estacionaria y la serie de tiempo estacionaria, en primeras diferencias, puede modelarse como un proceso ARIMA $(2, 2)$; es decir, tiene dos términos AR y dos términos MA.

Si $d = 0$ (si para empezar la serie es estacionaria, ARIMA $(p, d = 0, q) = \text{ARIMA}(p, q)$). Entonces, un proceso ARIMA $(p, 0, 0)$ significa un proceso estacionario AR(p) puro; un ARIMA $(0, 0, q)$ señala un proceso estacionario MA(q) puro.

Con valores p , d y q se sabe de qué proceso se está haciendo el modelo.

Finalmente, el punto importante es que, para usar la metodología *Box-Jenkins*, se tendrá una serie de tiempo estacionaria o una serie de tiempo que sea estacionaria después de una o más diferenciaciones.

La razón para suponer estacionariedad radica en que el objetivo de BJ (Box – Jenkins) es identificar y estimar un modelo estadístico que se interprete como generador de datos muestrales. Entonces, si se va a pronosticar con este modelo estimado se debe suponer que sus características son constantes a través del tiempo y, particularmente, en periodos futuros.

Así, la sencilla razón para requerir datos estacionarios es que todo modelo que infiera a partir de estos datos pueda interpretarse como estacionario o estable en sí mismo y, también, proporcione por consiguiente una base válida para pronosticar.

4.8.1. Especificación de modelos econométricos consiste en definir

- Relaciones que intervienen en el modelo, indicando el número de ecuaciones.
- Variables que intervienen en cada relación. Especificando qué variables son endógenas y cuáles predeterminadas, así como el tiempo de rezago.
- Forma funcional de cada relación.
- Forma en que interviene el término error (multiplicativo o aditivo) en cada relación.

Es frecuente que, la disponibilidad y calidad de esta información marque la diferencia entre poder estimar o no un modelo o siquiera

poder especificarlo. Se conoce antemano el papel que juegan estos datos en análisis econométrico, pues son la materia prima de la econometría, sin datos no hay práctica de econometría. Además, la calidad de los datos marca el nivel de calidad de los modelos ajustados. No se puede hacer milagros con los datos, la «basura» que se mete es la «basura» que se obtiene.

Conforme al período de tiempo que cubre una muestra de datos se generan dos grandes clases de información estadística: la información proveniente de series de tiempo e información recabada en un mismo momento en el tiempo, llamada información cruzada o en cruz. La notación de una variable de información cruzada normalmente lleva el subíndice *i*.

Por otro lado, las series de tiempo constituyen información que se va registrando sistemáticamente cada cierto período de tiempo bien definido: anual, mensual, semanal y diario. Lógicamente, se trata para cada variable de un vector ordenado de acuerdo con el tiempo. No puede haber dos observaciones en el mismo período de tiempo. La notación de una variable en serie de tiempo usualmente lleva el subíndice *t*.

Los registros temporales de las variables económicas de interés en la gran mayoría de los casos es una tarea realizada por instituciones gubernamentales. En el caso de México, esta función la realiza el Instituto Nacional de Estadísticas, Geografía e Informática (INEGI) dependiente de la Secretaría de Programación y Presupuesto.

Sin embargo, el Banco de México desempeña también un importante papel en el registro y divulgación de numerosas estadísticas económicas y sociales en forma mensual y anual, especialmente los principales índices de la economía nacional. Paralelamente, las

instituciones y entidades gubernamentales llevan sus propias estadísticas temporales de aquellas variables de su incumbencia.

Por ejemplo, CONASUPO es una buena fuente para precios de garantía, volúmenes importados de cereales, precios medios rurales, etc. Por su parte, BANRURAL, FIRA, AGROASEMEX, llevan las estadísticas financieras del sector agrícola. La Secretaría de Agricultura y Recursos Hidráulicos es la encargada de registrar la producción del sector, superficies cosechadas, rendimientos, etc.

Probablemente la disponibilidad y calidad de la información de series temporales en México es, así como en la gran mayoría de países en desarrollo, dispersa, poco confiable, muy agregada, muy desagregada, incompleta, poco comparable en períodos no muy largos de tiempo y, sobre todo, existen notorias discrepancias en los valores reportados sobre un mismo tema por diferentes instituciones. Los investigadores y practicantes de la econometría deben tener esto presente en cada momento.

Respecto a información cruzada, que se puede entender como aquella información proveniente de muestras en un mismo período en el tiempo, el panorama no es muy alentador. Las encuestas que se han aplicado por parte del Estado se reducen a las encuestas sobre ingreso y gastos de las familias aplicadas por la SPP cada cierto tiempo. La SARH aplica, desde 1985, una controvertida encuesta para estimar costos de producción en aquellos productos que estaban sujetos a precios de garantía.

A nivel privado, muchas instituciones de investigación y enseñanza superior, así como investigadores en particular, han aplicado encuestas socioeconómicas para medir algún fenómeno de su interés. Desgraciadamente, este esfuerzo esta cada día más lejos del alcance de

universidades y centros de investigación por altos costos que implican. La información es costosa y su calidad depende de la magnitud del presupuesto disponible.

CAPÍTULO V

MODELOS DE REGRESIÓN

De acuerdo con Grossman y Flores (2012), en múltiples problemas de las ciencias biológicas, físicas, sociales, agronómicas, agroindustriales, forestales, económicas, estadísticas, etcétera resulta útil describir la relación entre las variables de estos por medio de una expresión matemática.

Por ejemplo, se puede describir la relación entre el costo, el ingreso y la ganancia con la fórmula $G_{(\text{Ganancia})} = I_{(\text{Ingreso})} - C_{(\text{Costo})}$. Sin embargo, en un contexto físico, se puede representar la relación entre la aceleración debida a la gravedad, el tiempo que un objeto ha caído y la altura a la que estaba mediante la ley física

$$S_{(\text{Aceleración por gravedad})} = S_0 (\text{Altura inicial de objeto}) - v_0 (\text{Velocidad inicial}) \cdot t_{(\text{Tiempo})} - \frac{1}{2} G_{(\text{Gravedad})} t_{(\text{Tiempo})}^2$$

Asimismo, de acuerdo con Gujarati y Porter (2010), el problema de estimación mediante Modelo de Regresión con dos variables se analiza mediante dos métodos de estimación frecuentes:

5.1. Mínimos Cuadrados Ordinarios (MCO)

El método de MCO es el más común en el análisis de regresión, específicamente por ser mucho más intuitivo y matemáticamente más sencillo que el método de máxima verosimilitud. Además, por lo general los dos métodos proporcionan resultados similares.

El método de mínimos cuadrados ordinarios se atribuye al matemático alemán Carl Friedrich Gauss (1777-1855), considerado el matemático

más grande del siglo XIX, además de uno de los tres matemáticos más importantes de todos los tiempos (Arquímedes y Newton son los otros dos).

A partir de los siguientes supuestos el método de mínimos cuadrados presenta propiedades estadísticas muy atractivas que lo han convertido en uno de los más eficaces y populares del análisis de regresión, partiendo de la idea que el modelo de Gauss, modelo clásico o estándar de regresión lineal (MCRL) es el cimiento de la mayor parte de la teoría econométrica y plantea siete supuestos clásicos, en sentido que Gauss lo empleó por primera vez en 1821 y desde esta fecha sirve como norma o patrón con que compara modelos de regresión que no satisfacen los supuestos Gaussianos:

- I. Modelo de Regresión es Lineal en los Parámetros. Aunque, la variable regresada, dependiente o explicada (Y) y la regresora, independiente o explicativa (X) pueden o no ser lineal, inclusive puede incluir más variables explicativas, como se muestra enseguida:

$$Y_i = \beta_1 + \beta_1 X_i + u_i$$

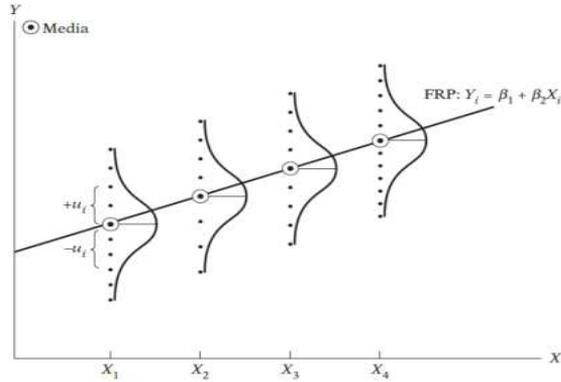
- II. Valores fijos de X, valores de X independientes del término error. Los valores que toma la variable regresora, independiente o explicativa (X) pueden considerarse fijos en muestras repetidas (regresora fija, no aleatoria, Modelos de Mínimos Clásicos de Regresión Lineal –MCRL- o Regresora Fija) o haber sido muestreados junto con la variable regresada, dependiente o explicada (Y) (regresora estocástica, aleatoria o Modelos Noeclásico de Regresión Lineal –MNRL- o Regresora

Estocástica). En el segundo caso, se supone que la(s) variable(s) X y término error son independientes; es decir, $\text{cov}(X_i, u_i) = 0$.

- III.** El valor medio de perturbación $u_i = 0$ (no hay error de especificación en modelo de regresión elegido). Dado el valor de X_i , la media o valor esperado del término de perturbación aleatoria u_i es cero. Simbólicamente es $E(u_i|X_i) = 0 \Rightarrow E(Y_i|X_i) = \beta_1 + \beta_1 X_i$ o, si X no es estocástica, equivale a $E(u_i) = 0$, pues si la media condicional de una variable aleatoria, dada otra variable aleatoria, es cero, la covarianza entre las dos variables es cero y, por tanto, las dos variables no están correlacionadas o X_i y u_i no están correlacionadas.

Cuando la Función de Regresión Poblacional (FRP) se expresa en una ecuación, se supone que X y u , representando la influencia de todas las variables omitidas, ejercen influencias independientes y aditivas, en Y , pero si X y u están correlacionadas, no es posible evaluar los efectos de cada una sobre Y tal que si X y u tienen correlación positiva, X aumenta cuando u aumenta y, viceversa, disminuye cuando u disminuye. Asimismo, si X y u tienen correlación negativa, X se incrementa cuando u se reduce, disminuye cuando u aumenta. Geométricamente, este supuesto se representa mediante la siguiente gráfica:

Figura 5.1 Función de Regresión Poblacional



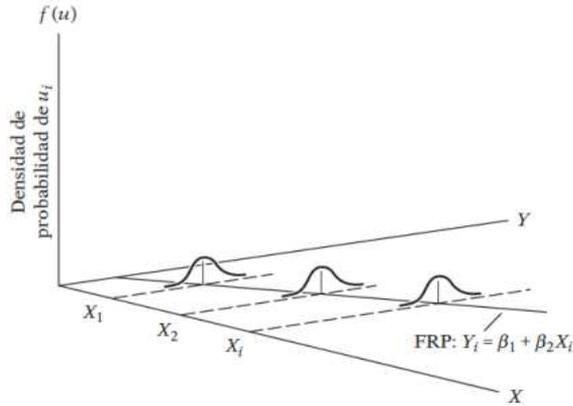
Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

- IV. Homocedasticidad o varianza constante de u_i (del griego skedanime que significa dispersar o esparcir tal que homo es igual y cedasticidad significa dispersión o, en otras palabras, igual varianza). La varianza del término error o de perturbación es la misma sin importan el valor de X . Simbólicamente, se tiene:

$$\begin{aligned} \text{Var}(u_i) &= E[u_i - (Y_i|X_i)]^2 = E(u_i^2|X_i) \text{ por supuesto 3} \\ &= E(u_i^2) \text{ si } X_i \text{ son variables no estocásticas} \\ &= \sigma^2 \end{aligned}$$

Esta ecuación establece que la varianza de u_i para cada X_i , varianza condicional de u_i , es algún número positivo constante igual a σ^2 . Por lo tanto, esta ecuación representa el supuesto de homocedasticidad. En términos llanos, la variación alrededor de la línea de regresión (relación promedio entre X y Y) es la misma para todos los valores de X tal que no aumente ni disminuye conforme varía X :

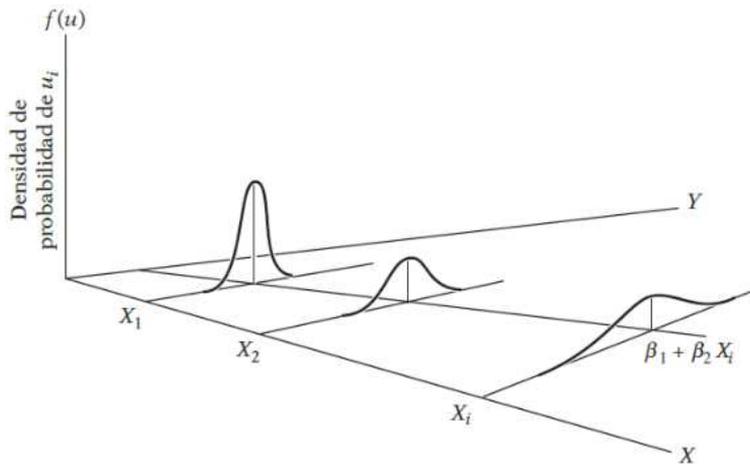
Figura 5.2 Homocedasticidad



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Caso contrario, si se considera la siguiente figura como varianza condicional de la población Y varia con X se conoce apropiadamente como heterocedasticidad o dispersión desigual o varianza desigual, escrita como $E(u_i^2|X_i) = \sigma_i^2$ e indica con subíndice i que la varianza de la población Y no es constante:

Figura 5.3 Varianza condicional



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

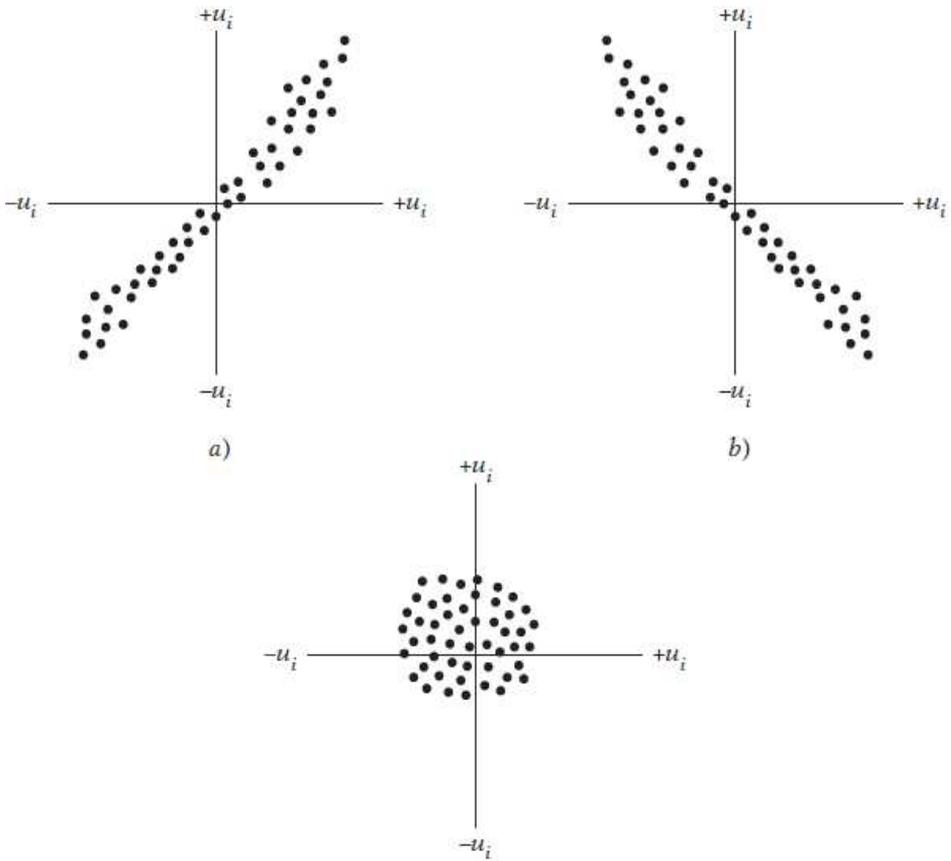
- V. No hay autocorrelación entre perturbaciones (perturbaciones u_i y u_j no están correlacionadas, supuesto de no correlación serial o no autocorrelación). Dados dos valores cualesquiera de X, X_i y

X_j ($i \neq j$), la correlación entre dos u_i y $u_j = 0 \forall (i \neq j)$. En otras palabras, estas observaciones se muestran independiente y simbólicamente así:

$$\text{cov}(u_i, u_j | X_i, X_j) = 0 \Rightarrow \text{cov}(u_i, u_j) = 0 \text{ si } X \text{ no es estocástica}$$

Esto significa que, dado X_i , las desviaciones de dos valores cualesquiera de Y de sus valores promedio no muestran patrones como en figura a en que las u están correlacionadas positivamente, pues a una u positiva sigue una u igual o viceversa, mientras que en figura b las u están correlacionadas negativamente, pues a una u positiva sigue una u negativa y viceversa:

Figura 5.4 Perturbaciones



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Si las perturbaciones (desviaciones) siguen patrones sistemáticos, como figuras a y b, hay correlación serial o autocorrelación mientras que figura c muestra que no hay un patrón sistemático para las u , indica cero correlación.

- VI. Número de observaciones n debe ser mayor que número de parámetros por estimar. Sucesivamente, el número de observaciones será $n \geq$ número de variables explicativas.
- VII. Naturaleza de variables X (variables deben variar). No todos los valores X en una muestra determinada deben ser iguales.

Técnicamente, $\text{Var}(X)$ debe ser un número positivo. Además, no puede haber valores atípicos de variable X ; es decir, valores

VIII. uy grandes en relación con el resto de las observaciones.

Esto tiene base según ecuación $\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} =$

$\frac{\sum x_i y_i}{\sum x_i^2}$, pues si todos los valores X son idénticos, $X_i = \bar{X}$ y el

denominador de esta ecuación es cero (indeterminación) imposibilita la estimación de $\hat{\beta}_2$ y, por consiguiente, de $\hat{\beta}_1$. Por lo tanto, la variación tanto en Y como X es esencial para utilizar el análisis de regresión como herramienta de investigación. Entonces, ¡las variables deben variar! El requisito que no existan valores atípicos de X es para evitar que resultados de regresión estén dominados por estos valores. Si hay algunos valores X que, por ejemplo, sean x veces el promedio de valores X , las líneas de regresión estimadas con o sin dichas observaciones serían muy diferentes. Con frecuencia estos valores atípicos son resultado de errores humanos de aritmética o de mezclar muestras de diferentes poblaciones.

5.1.1. Teorema de Gauss-Markov

Aunque se conoce como Teorema de Gauss-Markov, el método de Gauss de mínimos cuadrados en 1821 antecede al de Markov de varianza mínima en 1900. En contexto de regresión puede probarse que estimadores de MCO son MELI, basado en Teorema de Gauss-Markov: «Dados los supuestos del modelo clásico de regresión lineal, los estimadores de mínimos cuadrados, dentro de la clase de estimadores lineales insesgados, tienen varianza mínima o, en otras palabras, son MELI».

Al resolver ecuaciones normales al mismo tiempo se obtiene $\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$, que, con identidades algebraicas simples, esta fórmula se usa para estimar β_2 como:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum X_i^2 - n\bar{X}^2} = \frac{\sum X_i Y_i}{\sum X_i^2 - n\bar{X}^2}$$

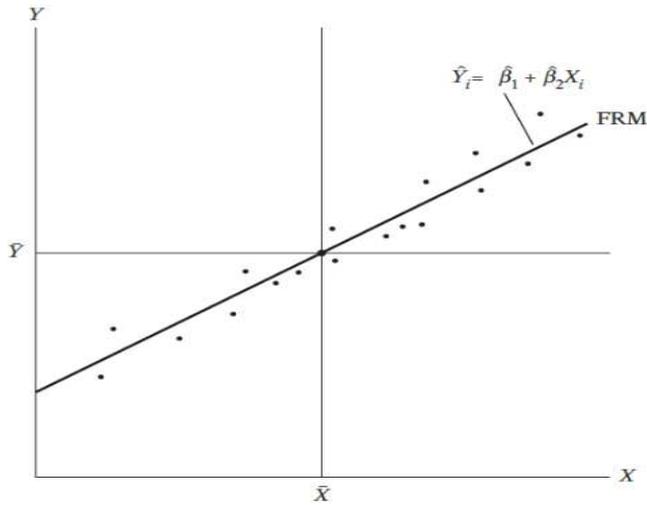
Donde $\sum x_i^2 = \sum (X_i - \bar{X})^2 = \sum X_i^2 - 2 \sum X_i \bar{X} + \sum \bar{X}^2 = \sum X_i^2 - 2\bar{X} \sum X_i + \sum \bar{X}^2$, pues \bar{X} es una constante. Además, $\sum X_i = n\bar{X}$ y $\sum \bar{X}^2 = n\bar{X}^2$, pues \bar{X} es una constante.

Finalmente, $\sum x_i^2 = \sum X_i^2 - n\bar{X}^2$. Los estimadores obtenidos se conocen como estimadores de mínimos cuadrados, pues se derivan del principio de mínimos cuadrados.

Las propiedades numéricas de estimadores obtenidos con método de MCO, que se mantienen como consecuencia del uso de mínimos cuadrados ordinarios sin considerar la forma como se generaron los datos, son:

- Estimadores de MCO se expresan únicamente en términos de cantidades (X y Y) observables (muestras). Por consiguiente, se calculan con facilidad.
- Son estimadores puntuales. Dada la muestra, cada estimador proporciona un solo valor (puntual) del parámetro poblacional pertinente, mientras que estimadores por intervalos proporcionan un intervalo de valores posibles para parámetros poblacionales no conocidos.
- Obtenidos los estimadores de MCO de datos muestrales, se obtiene sin problemas la línea de regresión muestral:

Figura 5.5 Estimadores de MCO



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Presenta las siguientes propiedades:

FRM se limita a la población de valores Y que corresponden a valores fijos X, tal que con toda deliberación evita consideraciones muestrales; es decir, datos poblacionales. Sin embargo, en la práctica lo que se obtiene al alcance no es más que una muestra de valores Y que corresponden a algunos valores fijos de X.

Por lo tanto, la labor es estimar la FRP ((Función de Regresión Poblacional)) con base en información muestral. Se pueden tomar dos o más muestras y representarse mediante lineal de regresión muestral tal que se obtendrían N FRM diferentes para N muestras diferentes y estas FRM no por fuerza son iguales. Al igual que FRP, basada en línea poblacional, se desarrolla el concepto de función de regresión muestral (FRM) para representar la línea de regresión muestral que se escribe como \hat{Y}_i (Estimador de $E(Y | X_i)$) = $\hat{\beta}_1 + \hat{\beta}_2 X_i$.

Un estimador o estadístico (muestral) es una regla, fórmula o método para estimar el parámetro poblacional a partir de información

suministrada por la muestra disponible. La FRP se expresa de dos maneras, $E(Y | X_i) = \beta_1 + \beta_2 X_i$ o $Y_i = E(Y | X_i) + u_i = \beta_1 + \beta_2 X_i + u_i$, la FRM se expresa en su forma estocástica como $Y_i = \beta_1 + \beta_2 X_i + u_i$.

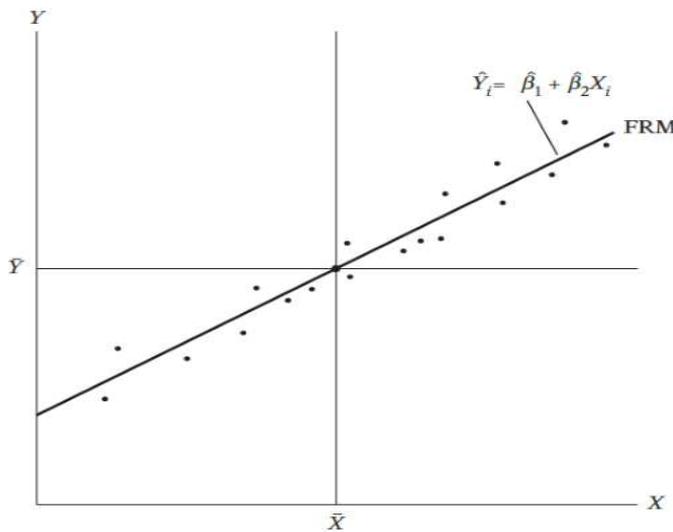
Por lo tanto, el objetivo del análisis de regresión es estimar FRP, $Y_i = \beta_1 + \beta_2 X_i + u_i$, con base en FRM, $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$ aunque, por fluctuaciones, la estimación de FRP basada e FRM es, en mejor de los casos, una aproximación

1. Pasa a través de medias muestrales de Y y X, evidente por ecuación

$$\hat{\beta}_1 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y}{n \sum X_i^2 - (\sum X_i)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}, \text{ pues esta ecuación puede}$$

escribirse $\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}$:

Figura 5.6 Función de Regresión Poblacional



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

2. El valor medio de Y estimada es igual a \hat{Y}_i es igual al valor medio de Y real para:

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) + \hat{\beta}_2 X_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) + \hat{\beta}_2 (X_i - \bar{X})$$

Al sumar ambos lados de esta última igualdad sobre valores muestrales y dividir por el tamaño n de muestra se obtiene que $\hat{Y} = \bar{Y}$, tal que se aprovecha que $\sum(X_i - \bar{X}) = 0$.

3. El valor medio de residuos $\hat{u}_i = 0$, pues al diferenciar parcialmente $\sum \hat{u}_i^2 = \sum(Y_i - \hat{Y}_i)^2 = \sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$ respecto de $\hat{\beta}_1$ y $\hat{\beta}_2$ se obtiene:

$$\frac{\partial(\sum \hat{u}_i^2)}{\partial \hat{\beta}_1} = -2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = -2 \sum \hat{u}_i$$

Complementariamente:

$$\frac{\partial(\sum \hat{u}_i^2)}{\partial \hat{\beta}_2} = -2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) X_i = -2 \sum \hat{u}_i X_i$$

Si se igualan a cero, se simplifican y manipulan algebraicamente se obtiene estimadores dados en ecuaciones:

$$\hat{\beta}_2 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2} \text{ y } \hat{\beta}_1 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y}{n \sum X_i^2 - (\sum X_i)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}.$$

Entonces, la ecuación $2 \sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$, pero como $\hat{u}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$, la ecuación anterior se reduce a $2 \sum \hat{u}_i = 0 \Rightarrow \hat{u}_i = 0$, aunque requiere que el término del intercepto β_1 esté presente en el modelo.

Como resultado de la propiedad anterior, la regresión muestral es:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}$$

Puede definirse de forma en que Y y X se expresan como desviaciones de sus medias.

Con base en esto, las propiedades de linealidad e insesgamiento de estimadores de mínimos cuadrados se muestran:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \sum k_i Y_i$$

Donde $k_i = \frac{x_i}{(\sum x_i^2)}$ muestra que $\hat{\beta}_2$ es un estimador lineal, pues es una función lineal de Y ; de hecho, es un promedio ponderado de Y_i donde k_i representa las ponderaciones. Igualmente, se demuestra que $\hat{\beta}_1$ es un estimador lineal.

Ahora sustituya la Función de Regresión Poblacional (FRP) $Y_i = \beta_1 + \beta_2 X_i + u_i$ en ecuación $\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \sum k_i Y_i$ para obtener:

$$\begin{aligned} \hat{\beta}_2 &= \sum k_i (\beta_1 + \beta_2 X_i + u_i) = \beta_1 \sum k_i + \beta_2 \sum k_i X_i + \sum k_i u_i \\ &= \beta_2 \sum k_i u_i \end{aligned}$$

Donde se emplean las propiedades de k_i :

1. Como se supuso que X_i son no estocásticas, las k_i también son no estocásticas.
2. $\sum k_i = 0$.
3. $\sum k_i^2 = \frac{1}{\sum x_i^2}$.
4. $\sum k_i X_i = \sum k_i Y_i = 1$

Estas propiedades se verifican directamente con definición de k_i . Por ejemplo:

$$\sum k_i = \sum \left(\frac{x_i}{\sum x_i^2} \right) = \frac{1}{\sum x_i^2} \sum x_i = 0$$

Pues para una muestra dada se conoce $\sum x_i^2$ y es = 0 debido a que $\sum x_i$, suma de desviaciones de la media, es siempre cero. Tal que, al obtener los valores esperados de ecuación $\hat{\beta}_2 = \sum k_i(\beta_1 + \beta_2 X_i + u_i) = \beta_1 \sum k_i + \beta_2 \sum k_i X_i + \sum k_i u_i = \beta_2 \sum k_i u_i$ para ambos lados y advertir que las k_i , al ser no estocásticas, pueden tratarse como constantes, por lo que se obtiene:

$$E(\hat{\beta}_2) = \beta_2 + \sum k_i E(u_i) = \beta_2$$

Esto es porque $E(u_i) = 0$ por sustitución. En consecuencia, $\hat{\beta}_2$ es un estimador insesgado de β_2 e, igualmente, se demuestra que $\hat{\beta}_1$ también lo es de β_1 . Para apreciar lo anterior, sume $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$ en ambos lados para obtener:

$$\sum Y_i = n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i + \sum \hat{u}_i = n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i + 0$$

Al dividir la anterior ecuación entre n , se obtiene:

$$\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X} = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y}{n \sum X_i^2 - (\sum X_i)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}$$

Si se resta la ecuación $\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}$ de $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$, se obtiene:

$$Y_i - \bar{Y} = \hat{\beta}_2 (X_i - \bar{X}) + \hat{u}_i \text{ o } y_i = \hat{\beta}_2 x_i + \hat{u}_i$$

Donde y_i y x_i según lo escrito, representan desviaciones de valores respectivos de sus medias muestrales y esta ecuación se llama forma de desviación. $\hat{\beta}_1$ puede estimarse mediante $\hat{\beta}_1 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y}{n \sum X_i^2 - (\sum X_i)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}$, pues la línea de regresión muestral pasa mediante medias

muestrales de Y y X. Una ventaja de la forma de desviación es que está simplificada a menudo los cálculos de fórmulas.

Es importante mencionar que la forma de desviación de Función de Regresión Muestral (FRM) es $\hat{y}_i = \hat{\beta}_2 x_i$ mientras que las unidades de medición originales, dicha expresión era $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$.

4. Residuos \hat{u}_i no están correlacionados con el valor pronosticado de Y_i , que se verifica de la siguiente manera:

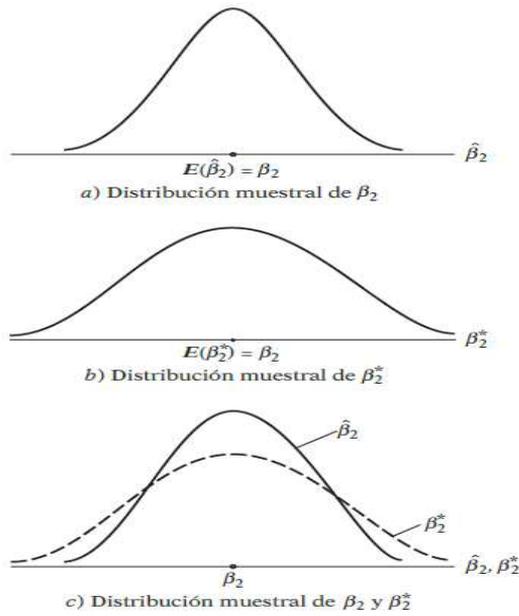
$$\begin{aligned} \sum \hat{y}_i \hat{u}_i &= \hat{\beta}_2 \sum x_i \hat{u}_i = \hat{\beta}_2 \sum x_i (y_i - \hat{\beta}_2 x_i) = \hat{\beta}_2^2 \sum x_i^2 - \hat{\beta}_2^2 \sum x_i^2 \\ &= 0 \end{aligned}$$

Donde se aprovecha que $\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$.

5. Los residuos \hat{u}_i no están correlacionados con X_i ; es decir, $\sum \hat{u}_i X_i = 0$. Esto se desprende de la ecuación $\frac{\partial(\sum \hat{u}_i^2)}{\partial \hat{\beta}_2} = -2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) X_i = -2 \sum \hat{u}_i X_i$.

El Teorema de Gauss-Markov tiene importancia teórica y práctica a la vez. Enseguida se presenta la distribución muestral del estimador de MCO $\hat{\beta}_2$; es decir, la distribución de valores asumidos por $\hat{\beta}_2$ en experimentos repetidos de muestreo (Distribución muestral de estimador de MCO $\hat{\beta}_2$ y estimador alterno β_2^*):

Figura 5.7 Distribución muestral del estimador de MCO



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

La media de los valores $\hat{\beta}_2$, $E(\hat{\beta}_2)$, es igual al verdadero valor β_2 . Por lo que, se dice que $\hat{\beta}_2$ es un estimador insesgado de β_2 . Por conveniencia, suponga que β_2^* , al igual que $\hat{\beta}_2$ es insesgado; es decir, que su valor promedio o esperado es igual a β_2 . Además, suponga que $\hat{\beta}_2$ y β_2^* son estimadores lineales; es decir, funciones lineales de Y .

No obstante, si bien $\hat{\beta}_2$ y β_2^* son insesgados, la distribución β_2^* es más difusa o dispersa alrededor del valor de la media que la distribución de $\hat{\beta}_2$. Es decir, la varianza de β_2^* es mayor que la varianza de $\hat{\beta}_2$. Asimismo, dados dos estimadores a la vez lineales e insesgados, sería preferible el estimador con la menor varianza, pues es probable que esté más cerca de β_2 , en comparación del alterno. Por lo tanto, se escogería el Mejor Estimador Lineal Insesgado (MELI).

El Teorema de Gauss-Markov no hace ninguna suposición respecto de distribución de probabilidad de variable aleatoria \hat{u}_i y, por

consiguiente, tampoco respecto de Y_i . En medida que satisfagan los supuestos de Mínimos Cuadrados de Regresión Lineal (MCRL), el teorema será válido. No se requiere buscar otro estimador insesgado lineal, pues no habrá otro estimador cuya varianza sea más pequeña que la del estimador de MCO.

Por supuesto, si no se cumple una o más de tales suposiciones, el teorema ya no es válido. Las propiedades antes vistas se conocen como propiedades de muestras finitas, pues se mantienen sin importar el tamaño de muestra en que se basen los estimadores.

Sin embargo, para entender este teorema se requiere considerar la propiedad del mejor estimador lineal insesgado: “ $\hat{u} \hat{u} = y' y - 2\hat{\beta}' X' y + \hat{\beta}' X' X \hat{\beta}$ con reglas de diferenciación se obtiene $\frac{\partial(\hat{u} \hat{u})}{\partial \beta} =$

$$-2X'y + 2X'X\hat{\beta} \Rightarrow \frac{2X'y}{2} = X'X\hat{\beta} \Rightarrow X'X\hat{\beta} = X'y \therefore \hat{\beta} = (X'X)^{-1}X'y$$

siempre que exista la inversa. Como $(X'X)^{-1}X'y$ es una matriz de números fijos, $\hat{\beta}$ es una función lineal de Y tal que, por definición, es una estimador lineal.

Es importante recordar que la Función de Regresión Poblacional (FRP) es $y = X\beta + u$, si se sustituye esta última ecuación en $\hat{\beta} = (X'X)^{-1}X'y$ se obtiene $\hat{\beta} = (X'X)^{-1}X'(X\beta + u) = \beta [(X'X)^{-1}X'X] + (X'X)^{-1}X'u = \beta + (X'X)^{-1}X'u$ y, tomando el valor esperado de esta última ecuación, se tiene $E(\hat{\beta}) = E(\beta) + (X'X)^{-1}X'E(u) = \beta$, pues $E(\beta) = \beta$ y $E(u) = 0$, según supuestos indica que $\hat{\beta}$ es un estimador insesgado de β . Sea $\hat{\beta}^*$ cualquier otro estimador lineal de β , escrito como $\hat{\beta}^* = [(X'X)^{-1}X' + C]y$, donde C es una matriz de constantes. Al sustituir y de $y = X\beta + u$ en $\hat{\beta}^* = [(X'X)^{-1}X' + C]y$ se obtiene

$\hat{\beta}^* = [(X'X)^{-1}X' + C](X\beta + u) = \beta + CX\beta + (X'X)^{-1}X'u + Cu$ tal que si $\hat{\beta}^*$ es un estimador insesgado de β se tiene que $CX = 0$.

Entonces, con $\hat{\beta}^* = \beta + CX\beta + (X'X)^{-1}X'u + Cu$ y $CX = 0$ se escribe $\hat{\beta}^* - \beta = (X'X)^{-1}X'u + Cu$ y, por definición, la matriz de Var – Cov($\hat{\beta}^*$) = E($\hat{\beta}^* - \beta$)($\hat{\beta}^* - \beta$)' = E[($X'X$)⁻¹X'u + Cu][(X'X)⁻¹X'u + Cu] empleando propiedades de inversión y transposición de matrices, luego de simplificación algebraica se obtiene Var – Cov($\hat{\beta}^*$) = $\sigma^2(X'X)^{-1} + \sigma^2CC'$ = Var – Cov($\hat{\beta}$) + σ^2CC' ”.

Esto indica que matriz de varianza – covarianza del estimador lineal e insesgado alterno $\hat{\beta}^*$ es igual a matriz varianza – covarianza del estimador MCO, $\hat{\beta}$ más σ^2 veces CC' , que es una matriz semidefinida positiva. Por tanto, las varianzas de un elemento dado de $\hat{\beta}^*$ deben ser necesariamente iguales o mayores al elemento correspondiente de $\hat{\beta}$, que demuestra que $\hat{\beta}$ es Mejor Estimador Lineal Insesgado (MELI).

Por supuesto, si C es una matriz nula (matriz cuyos elementos son todos cero y se denota por 0), entonces $\hat{\beta}^* = \hat{\beta}$, equivalente a que si encuentra un estimador MELI debe ser el estimador de mínimos cuadrados $\hat{\beta}$.

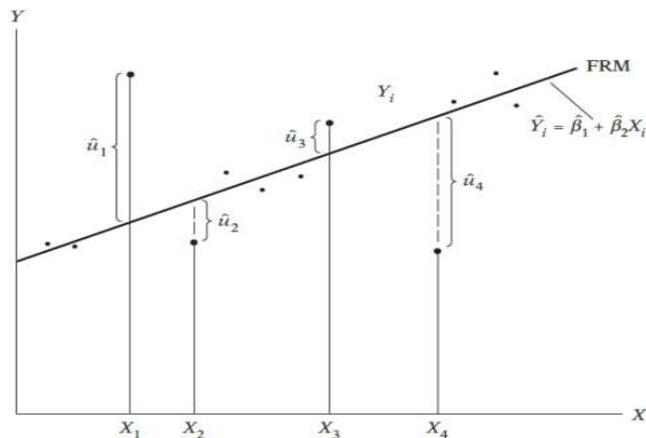
Con base en lo anterior, dados los supuestos del modelo clásico de regresión lineal, las estimaciones de mínimos cuadrados poseen algunas propiedades ideales u óptimas, contenidas en el Teorema de Gauss-Markov y siendo el estimador de MCO $\hat{\beta}_2$ es el mejor estimador lineal insesgado de β_2 o $\hat{\beta}^*$ de $\hat{\beta}$:

1. Es lineal. Función lineal de una variable aleatoria, como variable dependiente Y en modelo de regresión.
2. Es insesgado. Su valor promedio o esperado, $E(\hat{\beta}_2)$, es igual al valor verdadero, β_2 .
3. Estimador eficiente. Tiene varianza mínima dentro de la clase de todos los estimadores lineales insesgados. Un estimador insesgado con varianza mínima se conoce como estimador eficiente.

5.1.2. Coeficiente de determinación

La bondad del ajuste de la línea de regresión de un conjunto de datos refiere a cuán «bien» se ajusta la línea de regresión a los datos. Con base en la siguiente figura, es claro que si todas las observaciones caen en la línea de regresión se obtiene un ajuste «perfecto», pero rara ocasión se presenta:

Figura 5.8 Bondad del ajuste



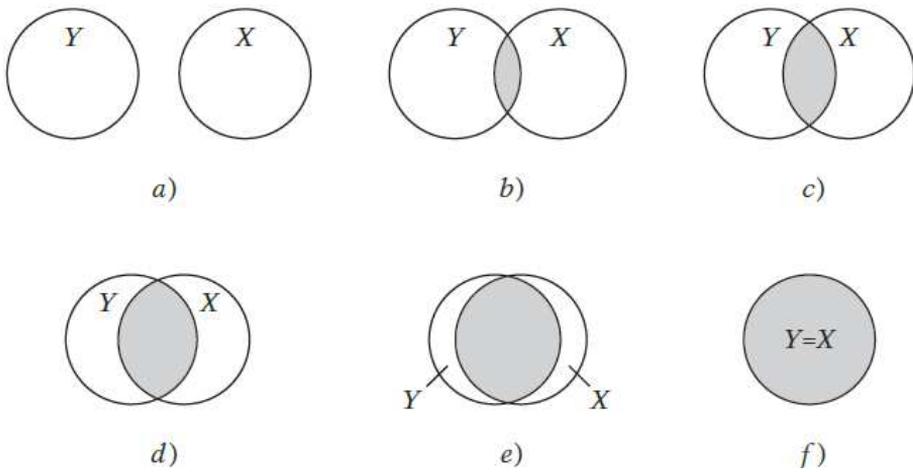
Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

En general, hay algunas \hat{u}_i positivas y algunas \hat{u}_i negativas. Se tiene la esperanza que estos residuos alrededor de la línea de regresión sean lo más pequeños posibles. Entonces, el coeficiente de determinación r^2

(dos variables) ó R^2 (regresión múltiple) es una medida comprendida que indica cuán bien se ajusta la línea de regresión muestral a los datos.

Una explicación del significado de heurística de r^2 ó R^2 en términos gráficos, conocida como Diagrama de Venn, Euler o Ballentine ($r^2: a) = 0; f) r^2 = 1$) es:

Figura 5.9 Diagrama de Venn



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

El círculo Y representa la variación en variable explicada, dependiente o endógena Y, el círculo X es la variación en la variable explicativa, independiente o exógena X

($\text{Varianza}_{(\text{Suma de cuadrados dividida por grados de libertad apropiados})} = \text{Variación}_{(\text{Suma de cuadrados de desviaciones de una variable respecto a su media})} / \text{Gl}$), mientras que la intersección de círculos (área sombreada) señala la media en que la variación en Y se explica por variación en X, como regresión de MCO.

A mayor medida de intersección, mayor será la variación en Y que se explica por X, pues a medida que va de izquierda a derecha, el área de

intersección aumente o, en otras palabras, hay una proporción cada vez mayor de la variación en Y explicada por X. Entonces, r^2 ó R^2 es una medida numérica de esta intersección.

Por lo tanto, cuando no hay intersección r^2 ó $R^2 = 0$, cuando es completa r^2 ó $R^2 = 1$, entendida como 100% de variación de Y se explica por X y, por lógica, r^2 ó R^2 varía ± 1 .

El cálculo de r^2 ó R^2 se hace de la siguiente forma, pues si:

$$Y_i = \hat{Y}_i + \hat{u}_i \text{ o, en términos de desviación, } y_i = \hat{y}_i + \hat{u}_i$$

Donde se emplean ecuaciones $y_i = \hat{\beta}_2 x_i + \hat{u}_i$ y $\hat{y}_i = \hat{\beta}_2 x_i$ tal que al elevar al cuadrado $y_i = \hat{y}_i + \hat{u}_i$ en ambos lados y sumar sobre la muestra se obtiene:

$$\begin{aligned} \sum y_i^2 &= \sum \hat{y}_i^2 + \sum \hat{u}_i^2 + 2 \sum \hat{y}_i \hat{u}_i = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 \\ &= \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{u}_i^2 \end{aligned}$$

Pues $\sum \hat{y}_i \hat{u}_i = 0$ y $\hat{y}_i = \hat{\beta}_2 x_i$. Las diversas sumas de cuadrados en ecuación anterior se describen de la forma siguiente:

$$\sum y_i^2 = \sum (Y_i - \bar{Y})^2$$

Hace referencia a variación total de valores reales de Y respecto de su media muestral, denominada Suma de Cuadrados Total (SCT).

$$\sum \hat{y}_i^2 = \sum (\hat{Y}_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 = \hat{\beta}_2^2 \sum x_i^2$$

Refiere a la variación de valores de Y estimador alrededor de su media ($\hat{Y}_i = \bar{Y}$), que apropiadamente puede llamarse Suma de Cuadrados por Regresión, debida a variables explicativas, independientes o exógenas,

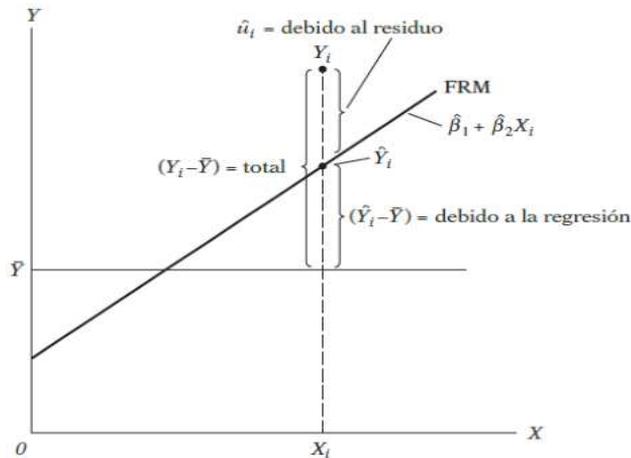
o, simplemente, Suma de Cuadrados explicada (SCE). Adicionalmente, $\sum \hat{u}_i^2$ es variación residual o variación no explicada de valores de Y alrededor de la línea de regresión o, también, llamada Suma de Cuadrados de Residuos (SCR). Así, ecuación:

$$\begin{aligned} \sum y_i^2 &= \sum \hat{y}_i^2 + \sum \hat{u}_i^2 + 2 \sum \hat{y}_i \hat{u}_i = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 \\ &= \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{u}_i^2 \Rightarrow \\ &\therefore \text{SCT} = \text{SCE} + \text{SCR} \end{aligned}$$

Esta ecuación muestra que la variación total en valores Y observados alrededor del valor de su media puede dividirse en dos partes, una atribuible a la línea de regresión y otra a fuerzas aleatorias, pues no todas las observaciones Y caen sobre línea ajustada.

Geoméricamente es:

Figura 5.10 Ajuste de la recta



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Al dividir la ecuación $\text{SCT} = \text{SCE} + \text{SCR}$ entre SCT en ambos lados, se obtiene:

$$\frac{SCT}{SCT} = \frac{SCE}{SCT} + \frac{SCR}{SCT} \Rightarrow 1 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} + \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2}$$

Ahora bien, se define r^2 :

$$r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{SCE_{(Error)}}{SCT_{(Total)}}$$

ó

$$r^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{SCR_{(Regresión)}}{SCT_{(Total)}}$$

La cantidad definida de esta manera se conoce como «Coeficiente de Determinación Muestral», entendida como la medida más común de bondad de ajuste de una lineal de regresión o, verbalmente, r^2 mide la proporción o por ciento de variación total en Y explicada por el modelo de regresión.

Asimismo, r^2 observa dos propiedades:

1. Es una cantidad no negativa, pues al elevar al cuadrado cualquier número, sea positivo o negativo, el resultado es positivo (ley de signos).
2. Sus límites son $0 \leq r^2 \leq 1$. $r^2 = 1$ indica un ajuste perfecto; es decir, $\hat{Y}_i = Y_i$ por cada i . Por otro lado, $r^2 = 0$ significa que no hay una relación alguna entre la variable regresada, dependiente, explicada o endógena y variable regresora, independiente, explicativa o exógena (coeficiente de pendiente o de variable exógena $\hat{\beta}_i = 0$).

En este caso, según $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) + \hat{\beta}_2 X_i = \bar{Y} + \hat{\beta}_2 (X_i - \bar{X})$, $\hat{Y}_i = \hat{\beta}_1 = \bar{Y}$; es decir, la mejor predicción de

cualquier valor de Y es simplemente el valor de su media. En consecuencia, la línea de regresión será horizontal al eje X.

No obstante, r^2 puede calcularse directamente a partir de la definición o, más rápido, con la fórmula siguiente:

$$\begin{aligned} r^2 &= \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{SCR_{(Regresión)}}{SCT_{(Total)}} \\ &= \frac{SCE_{(Error)}}{SCT_{(Total)}} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum y_i^2} = \hat{\beta}_2^2 \left(\frac{\sum x_i^2}{\sum y_i^2} \right) \end{aligned}$$

Si se dividen el numerado y denominador de la ecuación anterior por el tamaño n muestral (n - 1 si la muestra es pequeña), sabiendo que S_x^2 - S_y^2 son varianzas muestrales X - Y respectivamente, se obtiene:

$$r^2 = \hat{\beta}_2^2 \left(\frac{\frac{\sum x_i^2}{n}}{\frac{\sum y_i^2}{n}} \right) = \hat{\beta}_2^2 \left(\frac{S_x^2}{S_y^2} \right)$$

Como $\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum x_i^2}$, la ecuación $\hat{\beta}_2^2 \left(\frac{\sum x_i^2}{\sum y_i^2} \right)$ también se expresa:

$$r^2 = \frac{(\sum X_i Y_i)^2}{\sum x_i^2 \sum y_i^2} = \frac{(\sum y_i \hat{y}_i)^2}{(\sum y_i^2)(\sum \hat{y}_i^2)}$$

Con la definición de r^2 , SCE y SCR explicadas antes, se expresan así:

$$SCE = r^2 * SCT = (r^2) \left(\sum y_i^2 \right)$$

Por lo tanto:

$$SCR = SCT - SCE = SCT \left(1 - \frac{SCR}{SCT} \right) = \left(\sum y_i^2 \right) (1 - r^2)$$

Por lo tanto, se escribe:

$$\begin{aligned} \text{SCT} &= \text{SCE} + \text{SCR}_{()} \\ &\text{ó} \\ \sum y_i^2 &= r^2 \sum y_i^2 + (1 - r^2) \left(\sum y_i^2 \right) \end{aligned}$$

5.1.3. Coeficiente de correlación r ($\sqrt{r^2}$)

Cantidad estrechamente relacionada con el concepto de Coeficiente de Determinación r^2 (dos variables) ó R^2 (regresión múltiple) como medida de «bondad de ajuste». El Coeficiente de Correlación Muestral explica el grado de asociación, mide la fuerza o grado de asociación lineal, entre dos variables. Su cálculo es a partir de:

$$\begin{aligned} r &= \pm \sqrt{r^2} = \frac{\sum X_i Y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} \\ &= \frac{n \sum X_i Y_i - [(\sum X_i)(\sum Y_i)]}{\sqrt{[n \sum x_i^2 - (\sum X_i)^2][n \sum y_i^2 - (\sum Y_i)^2]}} \end{aligned}$$

El Coeficiente de Correlación Poblacional ρ (letra griega minúscula Rho o ro) es:

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{[\text{Var}(X)\text{Var}(Y)]}} = \frac{\text{Cov}(X, Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Entonces, ρ es una medida de Asociación Lineal entre dos variables y su valor se sitúa entre ± 1 , donde -1 indica una perfecta asociación negativa y $+1$ es una perfecta asociación positiva. Con base en lo anterior, se deduce:

$$\text{Cov}(X, Y) = \rho(\sigma_X \sigma_Y)$$

Algunas propiedades de r son:

- Tener signo positivo o negativo, según el signo del término en numerador de

$$r = \pm\sqrt{r^2} = \frac{\sum x_i y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} = \frac{n\sum x_i y_i - [(\sum x_i)(\sum y_i)]}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}$$

que

mide la Covariación Muestral de dos variables.

- Se ubica entre límites -1 y $+1$; es decir, $-1 \leq r \leq 1$.
- Es simétrico por naturaleza; es decir, el Coeficiente de Correlación entre X y Y , denotado por r_{XY} .
- Es independiente del origen y escala; es decir, si se define $X_i^* = aX_i + C$ y $Y_i^* = bY_i + d$ tal que $a > 0$, $b > 0$, c y d son constantes. Entonces, r entre X^* y Y^* es igual a r entre variables originales X y Y .
- Si X y Y son estadísticamente independientes, pues dos variables aleatorias X y Y son estadísticamente independientes si y solo si (\Leftrightarrow) $f(x,y) = f(x) f(y)$; es decir, si Función de Densidad de Probabilidad conjunta se expresa como el producto de las FDP marginales, el Coeficiente de Correlación ente ellas es cero, aunque no significa que dos variables sean independientes. En otras palabras, una correlación igual a cero no necesariamente implica independencia (siguiente figura h).
- Es una medida de Asociación Lineal o Dependencia Lineal solamente, su uso es la descripción de relaciones no lineales no tiene significado. Así, en la siguiente figura h, $Y = X^2$ es una relación exacta y $r = 0$, pues es una expresión cuadrática (parábola positiva) y, en consecuencia, una relación no lineal.
- Es una medida de asociación lineal entre dos variables y no implica obligatoriamente una relación causa-efecto: «Una relación estadística por más fuerte y sugerente que sea nunca

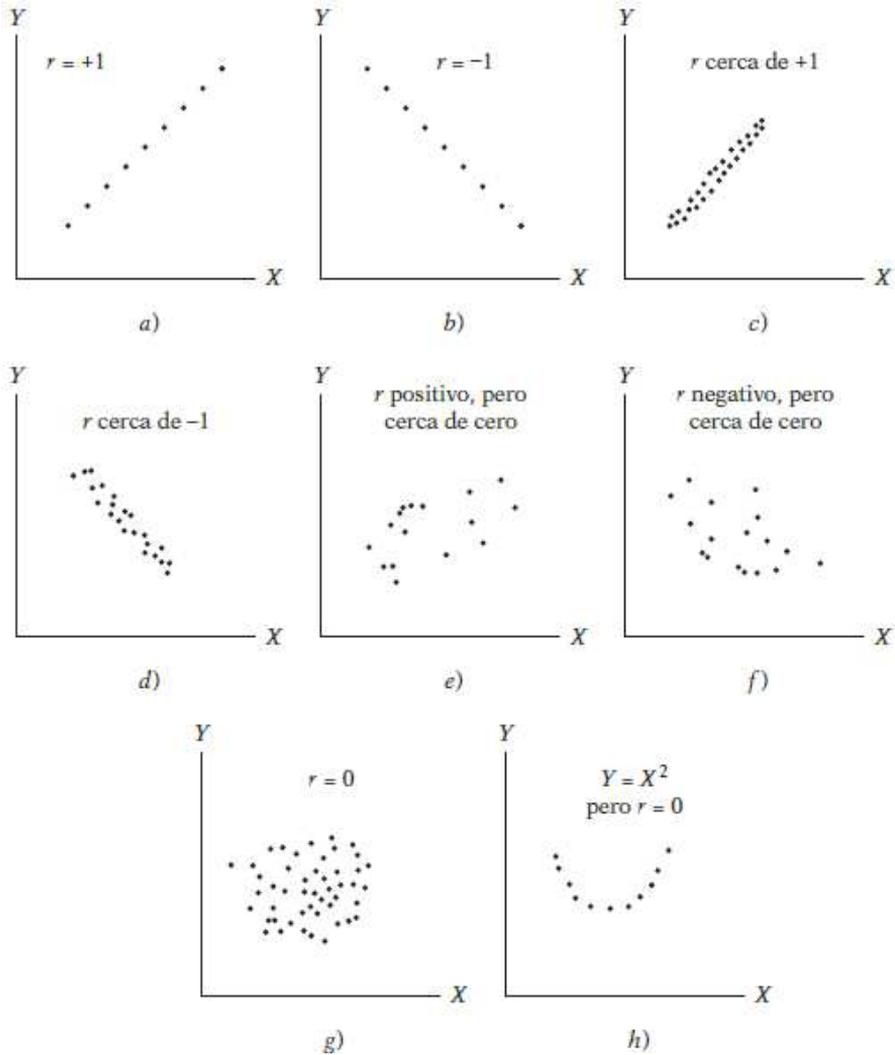
podrá establecer una conexión causal, por lo que ideas de causalidad provendrán de estadísticas externas y, finalmente, de una u otra teoría». Por lo tanto, «una relación estadística por sí misma no puede, lógicamente, implicar causalidad».

En el contexto de regresión, r^2 es una medida con más significado que r , pues la primera indica la proporción de variación en variable dependiente, explicada, predicha, regresada, respuesta, endógena, resultado o controlada explicada por la (s) variable (s) independiente (s), explicativa (s), predictora (s), regresora (s), estímulo (s), exógena (s), covariante (s) o de control (s).

En consecuencia, constituye una medida global del grado en que la variación en una variable determina la variación de otra, r no tiene este valor y, además, la interpretación de r (R) en un modelo de regresión múltiple es de valor dudoso.

Enseguida se muestran los patrones de correlación de Theil citado por Gujarati y Porter (2010):

Figura 5.11 Coeficiente de correlación



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

5.1.4. Precisión o errores estándar de estimaciones de mínimos cuadrados ordinarios

De las ecuaciones $\hat{\beta}_2 = \frac{n \sum x_i Y_i - \sum x_i \sum Y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{\sum (x_i - \bar{X})(Y_i - \bar{Y})}{\sum (x_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$ y

$\hat{\beta}_1 = \frac{\sum x_i^2 \sum Y_i - \sum x_i \sum x_i \sum Y_i}{n \sum x_i^2 - (\sum x_i)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}$ es evidente que estimaciones son

función de datos muestrales. Sin embargo, como es probable que los

datos cambien entre una muestra y otra, los valores estimados cambian ipso facto (expresión latina que significa por el mismo hecho, por el hecho mismo, inmediatamente o en el acto). Entonces, se requiere alguna medida de «confiabilidad» o precisión de estimadores $\hat{\beta}_1$ y $\hat{\beta}_2$.

En estadística, la precisión de un valor estimado se mide por su error estándar, entendido como desviación estándar de la distribución muestral del estimador y la distribución muestral de un estimador es tan solo una probabilidad o distribución de frecuencias del estimador; es decir, una distribución del conjunto de valores del estimador obtenidos de todas las muestras posibles de igual tamaño de una población dada.

Con las distribuciones muestrales se infieren los valores de los parámetros de la población, con base en los valores de los estimadores calculados a partir de una o más muestras. Dados los supuestos gaussianos, se muestra que los errores estándar de estimadores de MCO pueden obtenerse mediante:

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$\text{ee}(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum x_i^2}}$$

$$\text{Var}(\hat{\beta}_1) = \frac{\sum x_i^2}{n \sum x_i^2} \sigma^2$$

$$\text{ee}_{(\text{Error Est\acute{a}ndar})}(\hat{\beta}_1) = \sqrt{\frac{\sum x_i^2}{n \sum x_i^2}} \sigma$$

Todas las cantidades que entran en anteriores ecuaciones, excepto σ^2 , pueden estimarse a partir de datos. Incluso, la misma σ^2 se estima mediante:

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n - 2}$$

Los grados de libertad se entiende como número total de observaciones en la muestra (n) menos el número de restricciones (lineales) independientes o de restricciones que se les impusieron o, como, la cantidad de observaciones independientes de un total de n observaciones y su regla general es $gl = n - \text{número de parámetros estimados}$. Error Estándar de estimación o error estándar de regresión (ee).

No es más que la desviación estándar de valores Y alrededor de la línea de regresión estimada, que suele servir como medida para resumir la «bondad del ajuste» de dicha línea. Entonces, dado X_i , σ^2 representa la varianza condicional de u_i y Y_i .

Por lo tanto, el error estándar de estimación también se denomina desviación estándar condicional de u_i y Y_i . Asimismo, es común, σ_Y^2 y σ_i representan la varianza incondicional y desviación estándar incondicional de Y , respectivamente. Sin embargo, es importante mencionar las siguientes características de varianzas, así como de errores de estimadores $\hat{\beta}_1$ y $\hat{\beta}_2$:

- Varianza de $\hat{\beta}_2$ es directamente proporcional a σ^2 , pero inversamente proporcional a $\sum x_i^2$. Indica que, dada σ^2 , entre más grande sea la variación en valores X menor será la varianza $\hat{\beta}_2$ y, por tanto, mayor será la precisión con que

estimar β_2 y, dado $\sum x_i^2$, entre mayor sea la varianza σ^2 mayor será la de β_2 .

Tal que, a medida que aumenta el tamaño n muestral, también lo hace el número de términos en la suma, $\sum x_i^2$. Es decir, a medida que aumenta n es mayor la precisión para estimar β_2 .

- La varianza de $\hat{\beta}_1$ es directamente proporcional a σ^2 y a $\sum X_i^2$, pero inversamente proporcional a $\sum x_i^2$ y al tamaño n muestral.
- $\hat{\beta}_1$ y $\hat{\beta}_2$ son estimadores, no solo variarán de una muestra a otra, sino que también, en una muestra dada, es probable que dependan entre sí. Tal que, esta dependencia se mide por covarianza entre ellos.

Esto se demuestra con:

$$\begin{aligned} \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) &= E\{[\hat{\beta}_1 - E(\hat{\beta}_1)][\hat{\beta}_2 - E(\hat{\beta}_2)]\} \\ &= E(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2) = -\bar{X}E(\hat{\beta}_2 - \beta_2)^2 = -\bar{X}\text{Var}(\hat{\beta}_2) \\ &= -\bar{X}\left(\frac{\sigma^2}{\sum x_i^2}\right) \end{aligned}$$

Donde $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X}$ y $E(\hat{\beta}_1) = \bar{Y} - \beta_2\bar{X}$ que es igual a $\hat{\beta}_1 - E(\hat{\beta}_1) = -\bar{X}(\hat{\beta}_2 - \beta_2)$. La $\text{Var}(\hat{\beta}_2) = E[\hat{\beta}_2 - E(\hat{\beta}_2)]^2 = E(\hat{\beta}_2 - \beta_2)^2$, pues $E(\hat{\beta}_2) = \beta_2$, implica que $E(\sum k_i u_i)^2$ con ecuación $\hat{\beta}_2 = \sum k_i (\beta_1 + \beta_2 X_i + u_i) = \beta_1 \sum k_i + \beta_2 \sum k_i X_i + \sum k_i u_i = \beta_2 + \sum k_i u_i$ implica que sea $= E(k_1^2 u_1^2 + k_2^2 u_2^2 + k_3^2 u_3^2 + k_4^2 u_4^2 + \dots + k_n^2 u_n^2 + 2k_1 k_2 u_1 u_2 + \dots + 2k_{n-1} k_n u_{n-1} u_n)$.

Por los supuestos $E(u_i^2) = \sigma^2$ para cada i y $E(u_i u_j) = 0, i \neq j \therefore$

$$\text{Var}(\hat{\beta}_2) = \sigma^2_{(\text{Varianza constante u homocedástica de } u_i)} \sum k_i^2 = \frac{\sigma^2}{\sum x_i^2}$$
 con
 definición de k_i^2 .

La varianza de $\hat{\beta}_1$ se obtiene con el mismo razonamiento, pues una vez obtenidas las varianzas de $\hat{\beta}_1$ y $\hat{\beta}_2$ se obtiene también errores estándar correspondientes al tomar raíces cuadradas positivas.

Como $\text{Var}(\hat{\beta}_2)$ es siempre positiva, al igual que varianza de cualquier variable, la naturaleza de covarianza entre $\hat{\beta}_1$ y $\hat{\beta}_2$ depende del signo de \bar{X} , pues si \bar{X} es positiva, la covarianza será negativa. Así, el coeficiente de pendiente de β_2 está sobreestimada (pendiente muy pronunciada), el coeficiente del intercepto β_1 estará subestimado (intercepto será muy pequeño).

La mayor utilidad del estudio de covarianzas entre coeficientes estimados de regresión será en multicolinealidad.

5.1.5. Propiedades de estimadores de mínimos cuadrados ordinarios según supuestos de normalidad

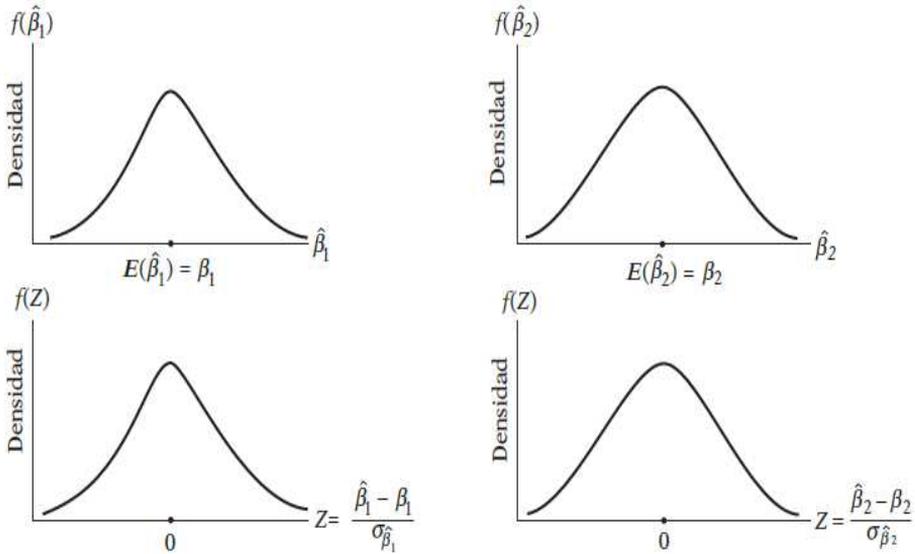
Suponga que u_i sigue una distribución normal, pues supone que cada u_i está Normalmente Distribuida con Media ($E(u_i) = 0$), Varianza ($E[u_i - E(u_i)]^2 = E(u_i^2) = \sigma^2$), $\text{Cov}(u_i, u_j) = (E\{[u_i - E(u_i)][u_j - E(u_j)]\} = E(u_i, u_j) = 0 \text{ } i \neq j)$, expresados en forma resuma como $u_i \sim N(0, \sigma^2)$, por lo que con el supuesto de normalidad la ecuación anterior indica que u_i y u_j no están correlacionadas, sino que están distribuidas independientemente tal que

$u_i \sim$ Normal e Independientemente Distribuido $(0, \sigma^2)$, los estimadores de MCO tienen las propiedades siguientes:

- Son Insesgados.
- Tienen varianza mínima. En combinación con 1 significa que son estimadores insesgados con varianza mínima o eficientes.
- Presentan consistencia. A medida que el tamaño muestral aumenta indefinidamente, los estimadores convergen hacia sus verdaderos valores poblacionales.
- $\hat{\beta}_1$ está Normalmente distribuida al ser función lineal de u_i con Media $(E(\hat{\beta}_1) = \beta_1)$, $\text{Var}(\hat{\beta}_1) \Rightarrow \sigma_{\hat{\beta}_1}^2 = \frac{\sum x_i^2}{n \sum x_i^2} \sigma^2$, en forma más completa $\hat{\beta}_1 \sim N(\beta_1, \sigma_{\hat{\beta}_1}^2)$. Tal que, de acuerdo con las propiedades de distribución normal, variable Z es definida como $Z = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}}$ sigue una distribución normal estándar; es decir, una distribución normal con media cero y varianza unitaria (= 1) ó, en otras palabras, $Z \sim N(0, 1)$.
- $\hat{\beta}_2$ está normalmente distribuida al ser función lineal de u_i con Media $(E(\hat{\beta}_2) = \beta_2)$, $\text{Var}(\hat{\beta}_2) \Rightarrow \sigma_{\hat{\beta}_2}^2 = \frac{\sigma^2}{\sum x_i^2}$, en forma más completa $\hat{\beta}_2 \sim N(\beta_2, \sigma_{\hat{\beta}_2}^2)$. Entonces, $Z = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}}$ sigue una distribución normal estándar.

Geoméricamente, las distribuciones de probabilidad de $\hat{\beta}_1$ y $\hat{\beta}_2$ son:

Figura 5.12 Distribuciones de probabilidad



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

- Como la distribución Ji Cuadrada (χ^2) está distribuida $(n - 2)$ (Grads de libertad) $\left(\frac{\hat{\sigma}^2}{\sigma^2}\right)$. Esto ayuda a hacer inferencias respecto a la verdadera σ^2 a partir de $\hat{\sigma}^2$.
- $(\hat{\beta}_1, \hat{\beta}_2)$ se distribuyen de manera independiente respecto a $\hat{\sigma}^2$.
- $\hat{\beta}_1$ y $\hat{\beta}_2$ tienen varianza mínima entre todas las clases de estimadores insesgados, lineales o no lineales. Es eficaz debido a que, a diferencia del Teorema de Gauss-Markov, no se limita a la clase de estimadores lineales.

Por tanto, los Estimadores de Mínimos Cuadrados son los Mejores Estimadores Lineales Insesgados (MELI), pues tienen varianza mínima en toda clase de estimadores insesgados.

Finalmente, el supuesto de normalidad permite derivar las distribuciones de probabilidad o muestrales de $\hat{\beta}_1$ y $\hat{\beta}_2$ (ambas normales) y de $\hat{\sigma}^2$ relacionadas con Ji Cuadrada (χ^2).

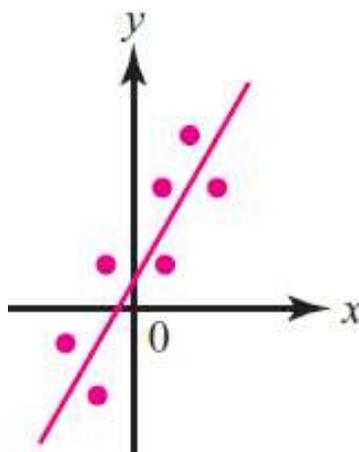
5.2. Aplicaciones con modelos matemáticos-económicos

Por desgracia, no es fácil obtener fórmulas como las anteriores. Muy a menudo los científicos o los economistas tienen que trabajar con grandes cantidades de datos para encontrar relaciones entre las variables de un problema. Una manera común de hacer esto es ajustar una curva entre los distintos puntos de datos. Esta curva puede ser un modelo matemático-económico, suponiendo en cada caso que existen n puntos de datos $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$ y la principal razón de su uso es que permiten obtener máximos o mínimos en las curvas (extremos relativos y/o extremos locales):

5.2.1. Modelo matemático-económico lineal

Suponga que se busca la recta de la forma $y = a + bx$ o $y = b + mx$ que mejor represente a n datos $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$:

Figura 5.13 Modelo matemático-económico lineal

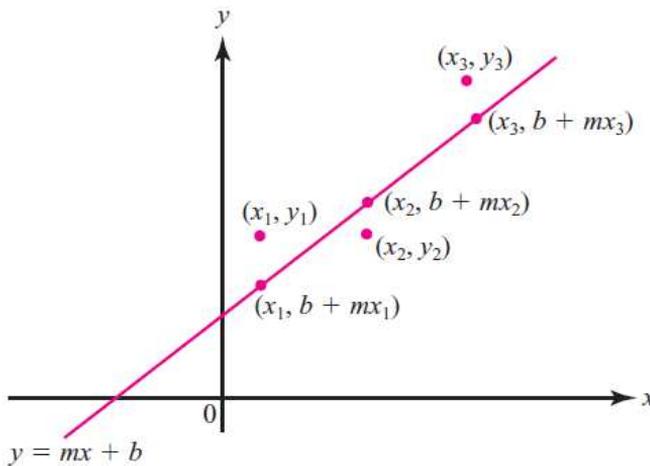


Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

La siguiente figura muestra lo que sucede usando tres datos, suponiendo que las variables x e y están relacionadas por la fórmula $y = b + mx$.

Por ejemplo, para $x = x_1$ el valor correspondiente de y es $b + mx_1$. Sin embargo, esto es diferente del valor “real”, $y = y_1$:

Figura 5.14 Distancia entre puntos



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

En \mathbb{R}^2 la distancia entre puntos (a_1, b_1) y (a_2, b_2) está dada por la ecuación $d = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$.

En consecuencia, al determinar la manera de elegir la recta $y = a + bx$ o $y = b + mx$ que mejor se aproxima a datos dados, es razonable usar el criterio de seleccionar aquella que minimiza la suma de cuadrados de diferencias entre valores y de puntos y el valor y correspondientes a la recta.

La distancia entre (x_1, y_1) y $(x_1, b + mx_1)$ es $y_1 - (b + mx_1)$, el problema para n datos se puede establecer como “problema de mínimos cuadrados en caso de una recta”, en que se encuentra

números m y b tales que la suma $[y_1 - (b + mx_1)]^2 + [y_2 - (b + mx_2)]^2 + [y_3 - (b + mx_3)]^2 + [y_4 - (b + mx_4)]^2 + \dots + [y_n - (b + mx_n)]^2$ sea mínima.

Para estos valores de m y b , la recta $y = a + bx$ o $y = b + mx$ se llama “Aproximación por recta de Mínimos Cuadrados a datos $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$ ”.

Una vez definido el problema se busca un método para encontrar la aproximación de mínimos cuadrados. Lo más sencillo es escribir en forma matricial. Si los puntos $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$ están todos sobre la recta $y = a + bx$, llamados colineales (dos o más elementos que se encuentran en una misma línea), entonces se tiene:

$$\begin{array}{ll} y_1 = a + bx_1 & y_1 = b + mx_1 \\ y_2 = a + bx_2 & y_2 = b + mx_2 \\ y_3 = a + bx_3 & y_3 = b + mx_3 \\ y_4 = a + bx_4 & y_4 = b + mx_4 \\ \dots & \dots \\ y_n = a + bx_n & y_n = b + mx_n \end{array} \quad \text{ó}$$

O, con otra nomenclatura, $y = Au$.

Donde:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} y u = \begin{pmatrix} b \\ m \end{pmatrix}$$

Si los puntos no son colineales, entonces $y - Au \neq 0$ y el problema se convierte en «forma vectorial del problema de mínimos cuadrados», que encuentra un vector \vec{u} tal que la forma euclídeana $|y - Au|$ es mínima.

Es importante mencionar que en \mathbb{R}^2 , la magnitud, longitud, módulo o norma de un $\vec{\mu}$ es $|(x, y)|_{(\text{Nomenclatura Física})}$ o $\|(x, y)\|_{(\text{Nomenclatura Matemática})} = \sqrt{x^2 + y^2}$, mientras que en \mathbb{R}^3 , la magnitud, longitud, módulo o norma de un $\vec{\mu}$ es $|(x, y, z)|$ o $\|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$, etcétera.

Entonces, minimizar $|y - Au|$ equivale a minimizar la suma de cuadrados en $[y_1 - (b + mx_1)]^2 + [y_2 - (b + mx_2)]^2 + [y_3 - (b + mx_3)]^2 + [y_4 - (b + mx_4)]^2 + \dots + [y_n - (b + mx_n)]^2$.

Encontrar el vector \vec{u} que minimiza no es tan difícil como parece. Como A es una matriz de $n_{(\text{Número de filas})} * 2_{(\text{Número de columnas})}$ y u una matriz de $2_{(\text{Número de filas})} * 1_{(\text{Número de columnas})}$, Au es un vector en \mathbb{R}^n perteneciente a la imagen A , entendida como un subespacio de \mathbb{R}^n cuya dimensión es a lo sumo dos, pues cuando mucho dos columnas de A son linealmente independientes.

Entonces, por Teorema de Aproximación de Norma \mathbb{R}^n , “sea H un subespacio de \mathbb{R}^n , u un vector en \mathbb{R}^n . Entonces $\text{PROY}_H(\text{Imagen de } A)^u$ es la mejor aproximación para u en H tal que si h es cualquier otro vector

en H implica que $|\mathbf{v} - \text{proy}_H(\text{Imagen de } A)\mathbf{v}| < |\mathbf{v} - \mathbf{h}|$ ”, Teorema de Proyección, “sea H un subespacio de \mathbb{R}^n y $\mathbf{v} \in \mathbb{R}^n$ ”.

Entonces existe un par único de vectores \mathbf{h} y \mathbf{p} tales que $\mathbf{h} \in H$, $\mathbf{p} \in H^\perp$ y $\mathbf{v} = \mathbf{h} + \mathbf{p}$. En particular, $\mathbf{h} = \text{proy}_H(\text{Imagen de } A)\mathbf{v}$ y $\mathbf{p} = \text{proy}_{H^\perp}(\text{Imagen de } A)\mathbf{v}$ tal que $\mathbf{v} = \mathbf{h} + \mathbf{p} = \text{proy}_H(\text{Imagen de } A)\mathbf{v} + \text{proy}_{H^\perp}(\text{Imagen de } A)\mathbf{v}$.

Su demostración implica que sea $\mathbf{h} = \text{proy}_H(\text{Imagen de } A)\mathbf{v}$ y $\mathbf{p} = \mathbf{v} - \mathbf{h}$ y por Definición Proyección, “sea H un subespacio de \mathbb{R}^n con base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \dots, \mathbf{u}_k\}$. Si $\mathbf{v} \in \mathbb{R}^n$, entonces la proyección ortogonal de \mathbf{v} sobre H , denotada por $\text{proy}_H(\text{Imagen de } A)\mathbf{v}$ está dada por $\text{proy}_H(\text{Imagen de } A)\mathbf{v} = (\mathbf{v} * \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} * \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} * \mathbf{u}_3)\mathbf{u}_3 + (\mathbf{v} * \mathbf{u}_4)\mathbf{u}_4 + \dots + (\mathbf{v} * \mathbf{u}_k)\mathbf{u}_k$, siendo que $\text{proy}_H(\text{Imagen de } A)\mathbf{v} \in H$ ”, se tiene que $\mathbf{h} \in H$.

Demostrando que $\mathbf{p} \in H^\perp$, sea $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \dots, \mathbf{u}_k\}$ una base ortonormal para H : $\mathbf{h} = (\mathbf{v} * \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} * \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} * \mathbf{u}_3)\mathbf{u}_3 + (\mathbf{v} * \mathbf{u}_4)\mathbf{u}_4 + \dots + (\mathbf{v} * \mathbf{u}_k)\mathbf{u}_k$ y sea \mathbf{x} un vector H tal que existen constantes $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \dots, \alpha_k$ tales que $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 + \dots + \alpha_k\mathbf{u}_k$ entonces $\mathbf{p} * \mathbf{x} = (\mathbf{v} - \mathbf{h}) * \mathbf{x} = [\mathbf{v} - (\mathbf{v} * \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v} * \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} * \mathbf{u}_3)\mathbf{u}_3 + (\mathbf{v} * \mathbf{u}_4)\mathbf{u}_4 + \dots + (\mathbf{v} * \mathbf{u}_k)\mathbf{u}_k] * [\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 + \dots + \alpha_k\mathbf{u}_k]$ aunque como $\mathbf{u}_i * \mathbf{u}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ es sencillo verificar que el producto escalar $\mathbf{p} * \mathbf{x} = (\mathbf{v} - \mathbf{h}) * \mathbf{x} = [\mathbf{v} - (\mathbf{v} * \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v} * \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} * \mathbf{u}_3)\mathbf{u}_3 + (\mathbf{v} * \mathbf{u}_4)\mathbf{u}_4 + \dots + (\mathbf{v} * \mathbf{u}_k)\mathbf{u}_k] * [\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 + \dots + \alpha_k\mathbf{u}_k]$ está dado por $\mathbf{p} * \mathbf{x} = \sum_{i=1}^k \alpha_i(\mathbf{v} * \mathbf{u}_i) - \sum_{i=1}^k \alpha_i(\mathbf{v} * \mathbf{u}_i) = 0$, así $\mathbf{p} * \mathbf{x} = 0 \forall \mathbf{x} \in H$ indicando que $\mathbf{p} \in H^\perp$.

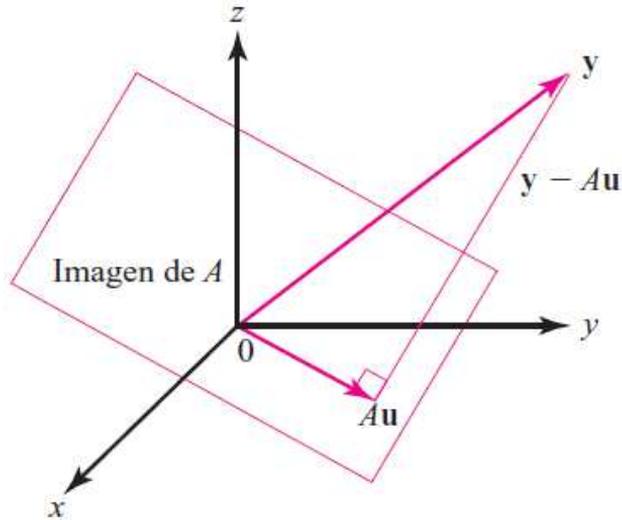
Para demostrar que $p = \text{proy}_H(\text{Imagen de } A)^\perp v$ se amplía $\{u_1, u_2, u_3, u_4, \dots, u_k\}$ a base ortonormal (base ortogonal en que la norma de cada elemento componente es unitaria o tiene vectores unitarios) en \mathbb{R}^n : $\{u_1, u_2, u_3, u_4, \dots, u_k, u_{k+1}, \dots, u_n\}$.

Entonces, $\{u_{k+1}, \dots, u_n\}$ es una base para H^\perp , tal que se puede decir que $v - \text{proy}_H(\text{Imagen de } A)v \in H^\perp$ se escribe $v - h = (v - \text{proy}_H(\text{Imagen de } A)v) + (\text{proy}_H(\text{Imagen de } A)v - h)$ siendo que el primer término de la derecha como el segundo está en H tal que $(v - \text{proy}_H(\text{Imagen de } A)v) * (\text{proy}_H(\text{Imagen de } A)v - h) = 0$ y por Teorema que afirma “sea $B = \{u_1, u_2, u_3, u_4, \dots, u_n\}$ una base ortonormal para \mathbb{R}^n y sea $v \in \mathbb{R}^n$ tal que $v = (v * u_1)u_1 + (v * u_2)u_2 + (v * u_3)u_3 + (v * u_4)u_4 + \dots + (v * u_k)u_k$ o $v = \text{proy}_{\mathbb{R}^n} v$, siendo su demostración que considera B como una base, puede escribir v de manera única como $v = c_1 u_1 + c_2 u_2 + c_3 u_3 + c_4 u_4 + \dots + c_n u_n$ por lo que $v * u_i = c_1 (u_1 * u_i) = c_2 (u_2 * u_i) + c_3 (u_3 * u_i) + c_4 (u_4 * u_i) + \dots + c_i (u_i * u_i) + \dots + c_n (u_n * u_i) = c_i$, pues los vectores \vec{u}_i son ortonormales y como esto se cumple para $i = 1, 2, 3, 4, \dots, n$ la demostración queda completa, se puede decir que $v = (v * u_1)u_1 + (v * u_2)u_2 + (v * u_3)u_3 + (v * u_4)u_4 + \dots + (v * u_k)u_k + (v * u_{k+1})u_{k+1} + \dots + (v * u_n)u_n = \text{proy}_H v + \text{proy}_{H^\perp} v$ aunque para probar la unicidad, suponga que $v = h_1 - p_1 = h_2 - p_2$ donde $h_1, h_2 \in H$ y $p_1, p_2 \in H^\perp$ entonces $h_1 - h_2 = p_1 - p_2$ aunque $h_1 - h_2 \in H$ y $p_1 - p_2 \in H^\perp$ de manera que $h_1 - h_2 \in H \cap H^\perp = \{0\}$, así que $h_1 - h_2 = 0$ y $p_1 - p_2 = 0$ completando la prueba”, por lo que”, tal que $|y - Au|$ es un mínimo cuadrado pues $Au = \text{proy}_H(\text{Imagen de } A)y$ donde H es la imagen de A .

En \mathbb{R}^3 la imagen de A será un plano o recta que pasa por el origen, pues son los únicos subespacios de dimensión uno o dos.

En la siguiente figura, el vector que minimiza se denota por u :

Figura 5.15 Subespacios de dimensión



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Con base en esta figura y el Teorema de Pitágoras, «en todo triángulo rectángulo, el cuadrado de la longitud de la hipotenusa es igual a la suma de los cuadrados de las respectivas longitudes de los catetos», se deduce que $|y - Au|$ es mínima cuando $y - Au$ es ortogonal a imagen A .

Es decir, si \vec{u} es vector que minimiza implica que para todo vector $u \in \mathbb{R}^2$ por lo que $Au \perp (y - Au)$. Usando la definición de producto escalar en \mathbb{R}^n , se encuentra que $Au \perp (y - Au)$ se vuelve $Au * (y - Au) = 0$, $(Au)^T * (y - Au) = 0$ debido a que $a * b = a^T b$, $(A^T u^T) * (y - Au) = 0$ pues por Teorema ii) que supone en i) $(A^T)^T$, ii) $(AB)^T = A^T B^T$, iii) Si A y B son de $n * m$, entonces $(AB)^T = A^T + B^T$ y iv) Si A es invertible, entonces A^T es invertible y $(A^T)^{-1} = (A^{-1})^T$ o $a^T(A^T y - A^T * Au) = 0$ se cumple $\forall u \in \mathbb{R}^2 \Leftrightarrow A^T y - A^T * Au = 0$, pero al despejar \vec{u} de $A^T y - A^T * Au = 0$ se obtiene que $y - Au \perp Au$.

Entonces, la solución al problema de mínimos cuadrados para ajuste por línea recta se da si A y y son:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} \text{ y } u = \begin{pmatrix} b \\ m \end{pmatrix}$$

Entonces, la recta $y = a + bx$ o $y = b + mx$ da el mejor ajuste en sentido de mínimos cuadrados para puntos $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$ cuando:

$$\begin{pmatrix} b \\ m \end{pmatrix} = \vec{u} = (A^T A)^{-1} A^T y \text{ y } y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}$$

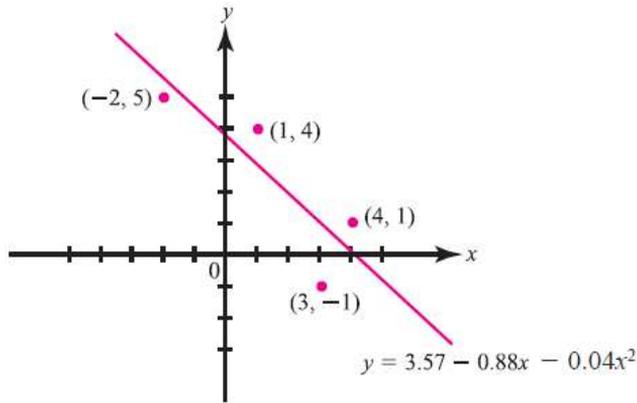
Se ha puesto que $A^T A$ (matriz A transpuesta por matriz A) es invertible. Este siempre es el caso si nos n datos no son colineales.

5.2.2. Modelo matemático-econométrico cuadrático

Ahora se desea ajustar una curva cuadrática a n datos. Una curva cuadrática en x es cualquier expresión de forma $y = a + bx + cx^2$ o $y = b + mx + cx^2$ tal que es la ecuación de una parábola en el plano.

Si los n datos estuvieran sobre la parábola se tendría:

Figura 5.16 Ajuste de una función cuadrática



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

$$\begin{array}{ll}
 y_1 = a + bx_1 + cx_1^2 & y_1 = b + mx + cx_1^2 \\
 y_2 = a + bx_1 + cx_2^2 & y_2 = b + mx + cx_2^2 \\
 y_3 = a + bx_1 + cx_3^2 & y_3 = b + mx + cx_3^2 \\
 & \text{ó} \\
 y_4 = a + bx_1 + cx_4^2 & y_4 = b + mx + cx_4^2 \\
 \dots & \dots \\
 y_n = a + bx_n + cx_n^2 & y_n = b + mx + cx_n^2
 \end{array}$$

Aunque, este sistema se puede escribir como $y - Au$ con:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ \dots & \dots & \dots \\ 1 & x_n & x_n^2 \end{pmatrix} y u = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

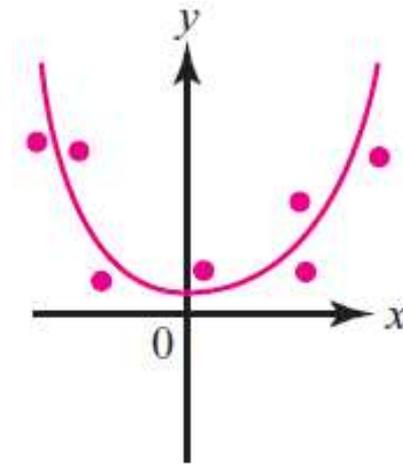
Al igual que antes, si todos los datos no se encuentran sobre la misma parábola implica que $y - Au \neq 0$ para cualquier vector \vec{u} y, de nuevo,

el problema es “encontrar un vector u en \mathbb{R}^2 tal que $|y - Au|$ es mínima”.

Usando un razonamiento similar a éste, se puede demostrar que cuando menos tres de las x_i son diferentes, entonces $A^T A$ es invertible y el vector que minimiza al vector \vec{u} está dado por $\vec{u} = (A^T A)^{-1} A^T y$.

La siguiente imagen muestra la figura de un modelo matemático-económico cuadrático:

Figura 5.17 Modelo matemático-económico cuadrático



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Teorema: Sea $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), \dots, (x_n, y_n)$ puntos en \mathbb{R}^2 , suponga que no todas las x_i son iguales.

Entonces, si A está dada como la siguiente forma es matriz $A^T A$ es invertible de 2×2 :

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} \text{ y } u = \begin{pmatrix} b \\ m \end{pmatrix}$$

Es importante mencionar que si $x_1 = x_2 = x_3 = x_4 = \dots = x_n$, entonces todos los datos sobre la recta vertical $x = x_1$ y la mejor aproximación lineal es dicha recta.

Su demostración es que se tiene:

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}$$

Como no todas las x_i son iguales, las columnas de A son linealmente independientes. Ahora bien:

$$\begin{aligned} A^T_{(\text{Matriz } A \text{ transpuesta})} A &= \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & x_4 & \dots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} \\ &= \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \end{aligned}$$

Si $A^T_{(\text{Matriz } A \text{ transpuesta})} A$ no es invertible, entonces $A^T A = 0$. Esto significa que:

$$n \sum_{i=1}^n x_i^2 = \left(\sum_{i=1}^n x_i \right)^2$$

$$\text{Sea } u = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \text{ y } x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \dots \\ x_n \end{pmatrix}.$$

En consecuencia, $|u|^2 = u * u = n$, $|x|^2 = \sum_{i=1}^n x_i^2$ y $u * x = \sum_{i=1}^n x_i$.

Tal que la ecuación $n \sum_{i=1}^n x_i^2 = (\sum_{i=1}^n x_i)^2$ puede establecerse como $|u||x|^2 = |u * x|^2$, sacando raíz cuadrada se puede obtener $|u * x| = |u||x|$.

La desigualdad de Cauchy-Schwarz en \mathbb{R}^n sostiene que si u y v son dos vectores en \mathbb{R}^n :

- $|u * v| \leq |u||v|$. Su demostración es si $u = 0$, $v = 0$ o ambos, entonces $|u * v| \leq |u||v|$ se cumple, pues ambos lados son iguales a 0. Si se supone que $u \neq 0$ y $v \neq 0$, entonces:

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

De manera semejante con:

$$0 \leq \left| \frac{u}{|u|} - \frac{v}{|v|} \right|^2 \Rightarrow \frac{u * v}{|u||v|} \geq -1 \text{ ó } u * v \geq -|u||v|$$

Tal que con estas dos desigualdades se obtiene:

$$-|u||v| \leq u * v \leq |u||v| \text{ ó } |u * v| \leq |u||v|$$

- $|u * v| = |u||v| \Leftrightarrow u = 0$ o $v = \lambda u$ para algún número $\mathbb{R} \lambda$. Su demostración indica que si $u = \lambda v$, entonces $|u * v| = |\lambda v * v| = |\lambda||v|^2$ y $|u||v| = |\lambda v||v| = |\lambda||v||v| = |\lambda||v|^2 = |u * v|$.

Inversamente, suponga que $|u * v| = |u||v|$ con $u \neq 0$ y $v \neq 0$, entonces:

$$\left| \frac{\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \right| = 1 \Rightarrow \frac{\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = \pm 1$$

Entonces:

➤ Caso 1: $\frac{\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = 1$.

$$\begin{aligned} \left| \frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right|^2 &= \left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right) * \left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right) \\ &= \frac{\mathbf{u} * \mathbf{u}}{|\mathbf{u}|^2} - \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{\mathbf{v} * \mathbf{v}}{|\mathbf{v}|^2} = \frac{|\mathbf{u}|^2}{|\mathbf{u}|^2} - \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2} \\ &= 1 - \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + 1 = 2 - \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = 2 - 2(1) = 0 \\ &\Rightarrow \frac{\mathbf{u}}{|\mathbf{u}|} = \frac{\mathbf{v}}{|\mathbf{v}|} \quad \text{ó} \quad \mathbf{u} = \frac{\mathbf{u}}{|\mathbf{u}|} * \mathbf{v} = \lambda \mathbf{v} \end{aligned}$$

➤ Caso 2: $\frac{\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = -1$.

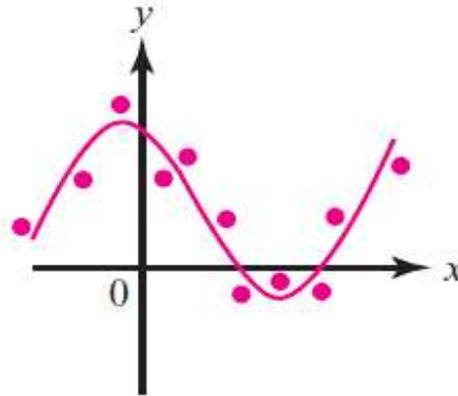
$$\begin{aligned} \left| \frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right|^2 &= \left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right) * \left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right) \\ &= \frac{\mathbf{u} * \mathbf{u}}{|\mathbf{u}|^2} + \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{\mathbf{v} * \mathbf{v}}{|\mathbf{v}|^2} = \frac{|\mathbf{u}|^2}{|\mathbf{u}|^2} + \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2} \\ &= 1 + \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + 1 = 2 + \frac{2\mathbf{u} * \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = 2 - 2(1) = 0 \\ &\Rightarrow \frac{\mathbf{u}}{|\mathbf{u}|} = -\frac{\mathbf{v}}{|\mathbf{v}|} \quad \text{ó} \quad \mathbf{u} = -\frac{\mathbf{u}}{|\mathbf{u}|} * \mathbf{v} = \lambda \mathbf{v} \end{aligned}$$

5.2.3. Modelo matemático-econométrico cúbico

El objetivo es encontrar la curva del tipo específico que se ajuste «mejor» a los datos dados. Complementariamente, un sistema de

ecuaciones se escribe $Ax = b$ tal que A es una matriz de $m \times n$, $X \in \mathbb{R}^n$ y $b \in \mathbb{R}^m$.

Figura 5.18 Modelo matemático-económico cúbico



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Asimismo, la función $Ax = b$ tiene las propiedades que caracteriza las transformaciones lineales, $A(\alpha x) = \alpha Ax$ si α es una escalar y $A(x + y) = Ax + Ay$.

Entonces, la definición de Transformación Lineal es que sean V y W espacios vectoriales reales tal que una transformación lineal T de V en W es una función que asigna a cada vector $v \in V$ un vector único $Tv \in W$ y que satisface, para cada u y v en V y cada escalar α , $T(u + v) = Tu + Tv$ y $T(\alpha v) = \alpha Tv$.

Es importante mencionar que las definiciones y teoremas se cumplen también para espacios vectoriales complejos (espacios vectoriales donde escalares son números complejos); aunque, solo se manejan espacios vectoriales reales y, por ende, se eliminará la palabra «real» en análisis de espacios vectoriales y transformaciones lineales, conocidas como operadores lineales.

Ejemplos:**1. Transformación lineal de \mathbb{R}^2 en \mathbb{R}^3**

Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por $T\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X + Y \\ X - Y \\ 3Y \end{pmatrix}$.

Por ejemplo:

$$\begin{aligned} T\begin{pmatrix} 2 \\ -3 \end{pmatrix} &= \begin{pmatrix} -1 \\ 5 \\ -9 \end{pmatrix} \Rightarrow T\left[\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} + \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}\right] \\ &= T\begin{pmatrix} X_1 + X_2 \\ Y_1 + Y_2 \end{pmatrix} = \begin{pmatrix} X_1 + X_2 + Y_1 + Y_2 \\ X_1 + X_2 - Y_1 - Y_2 \\ 3Y_1 + 3Y_2 \end{pmatrix} = \begin{pmatrix} X_1 + Y_1 \\ X_1 - Y_1 \\ 3Y_1 \end{pmatrix} + \begin{pmatrix} X_2 + Y_2 \\ X_2 - Y_2 \\ 3Y_2 \end{pmatrix} \end{aligned}$$

$$\text{Pero } \begin{pmatrix} X_1 + Y_1 \\ X_1 - Y_1 \\ 3Y_1 \end{pmatrix} = T\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \text{ y } \begin{pmatrix} X_2 + Y_2 \\ X_2 - Y_2 \\ 3Y_2 \end{pmatrix} = T\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}$$

$$\text{Así, } T\left[\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} + \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}\right] = T\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} + T\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}$$

$$\approx T\left[\alpha \begin{pmatrix} X \\ Y \end{pmatrix}\right] = T\begin{pmatrix} \alpha X \\ \alpha Y \end{pmatrix} = \begin{pmatrix} \alpha X + \alpha Y \\ \alpha X - \alpha Y \\ 3\alpha Y \end{pmatrix} = \alpha \begin{pmatrix} X + Y \\ X - Y \\ 3Y \end{pmatrix}$$

$$= \alpha T\begin{pmatrix} X \\ Y \end{pmatrix} \Rightarrow \therefore T \text{ es una transformación lineal.}$$

2. Transformación cero

Sean V y W espacios vectoriales y definida $T: V \rightarrow W$ por $Tv = 0$ para $\forall v$ en $V \Rightarrow T(v_1 + v_2) = 0 + 0 = Tv_1 + Tv_2$ y $T(\alpha v) = 0 = \alpha 0 = \alpha Tv$.

En este caso, T se denomina transformación cero.

3. Transformación identidad.

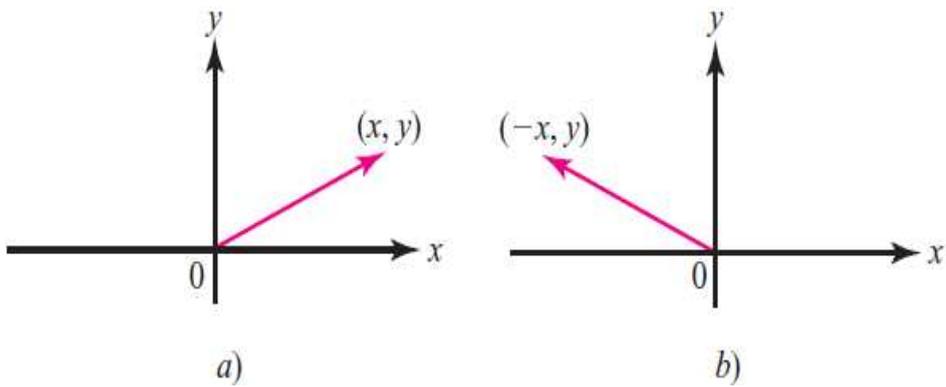
Sea V un espacio vectorial y definida $I: V \rightarrow V$ por $Iv = v$ para $\forall v$ en V . ES obvio que I es una transformación lineal, llamada transformación identidad u operador identidad.

4. Transformación de reflexión.

Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $T \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} -X \\ Y \end{pmatrix}$. Es fácil verificar que T es lineal.

En términos geométricos, T toma un vector en \mathbb{R}^2 y lo refleja respecto al eje Y (vector $(-X, Y)$ es reflexión respecto a eje Y del vector (X, Y)):

Figura 5.19 Transformación de reflexión



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

5. Transformación lineal de $\mathbb{R}^n \rightarrow \mathbb{R}^m$ dada por multiplicación de matriz $m \times n$.

Sea A una matriz de $m \times n$, definida $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ por $Tx = Ax$. Como $A(x + y) = Ax + Ay$ y $A(\alpha x) = \alpha Ax$ si x y y están en \mathbb{R}^n , se observa que T es una transformación lineal.

Entonces, toda matriz A de $m \times n$ se puede usar para definir una transformación de \mathbb{R}^n en \mathbb{R}^m . Por lo tanto, toda transformación lineal entre espacios vectoriales de dimensión finita se puede representar por una matriz.

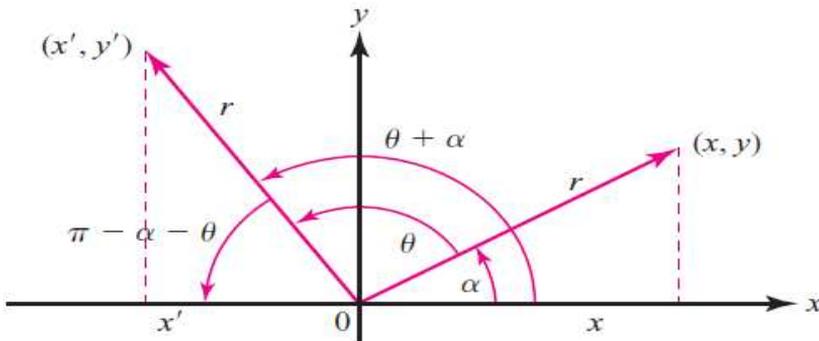
6. Transformación de rotación

Suponga que $\vec{v} = \begin{pmatrix} X \\ Y \end{pmatrix}$ en $\Pi_{(XY)}$ se rota un $\neq \theta$, medido en grados ($^\circ$) y radianes (rad) en sentido contrario a manecillas del reloj.

Llame a este nuevo vector rotado $\vec{v}' = \begin{pmatrix} X' \\ Y' \end{pmatrix}$ (si r denota longitud de v , que no cambia por rotación tal que $x = r \cos(\alpha) \Rightarrow x' = r \cos(\theta + \alpha)$ y $y = r \sin(\alpha) \Rightarrow y' = r \sin(\theta + \alpha)$, pues la definición estándar de $\cos(\theta)$ y $\sin(\theta)$ como las coordenadas X y Y de un punto en círculo unitario.

Si (X, Y) es un punto en círculo de radio r con centro en origen, entonces $x = r \cos(\varphi)$ y $y = r \sin(\varphi)$, donde φ (Φ o Φ) es ángulo que forma $\vec{v}(X, Y)$ con lado positivo del eje X . Además, (X', Y') se obtiene rotando (X, Y) un ángulo θ :

Figura 5.20 Transformación de rotación



Fuente: (Gujarati, D. N. y Porter, D. C. 2010).

Pero $r \cos(\theta + \alpha) = r \cos(\theta)\cos(\alpha) - r \sin(\theta)\sin(\alpha) \Rightarrow x' = x \cos(\theta) - y \sin(\theta)$.

Paralelamente, $r \sin(\theta + \alpha) = r \sin(\theta)\cos(\alpha) + r \cos(\theta)\sin(\alpha) \Rightarrow y' = x \sin(\theta) + y \cos(\theta)$.

Sea $A_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \Rightarrow x' = x \cos(\theta) - y \sin(\theta) \quad y \quad y' = x \sin(\theta) + y \cos(\theta)$ denota que $A_\theta \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X' \\ Y' \end{pmatrix}$.

La transformación lineal $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $Tv = A_\theta v$, donde A_θ está dado por $A_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ y recibe el nombre de transformación de rotación.

7. Transformación de proyección ortogonal (perpendicular:)

Sea H un subespacio de \mathbb{R}^n tal que la transformación de proyección ortogonal está dada por $P: V \rightarrow H$ se define por $Pv = \text{Proy}_H v$. Sea la sucesión $\{u_1, u_2, u_3, u_4, \dots, u_k\}$ una base ortogonal para H .

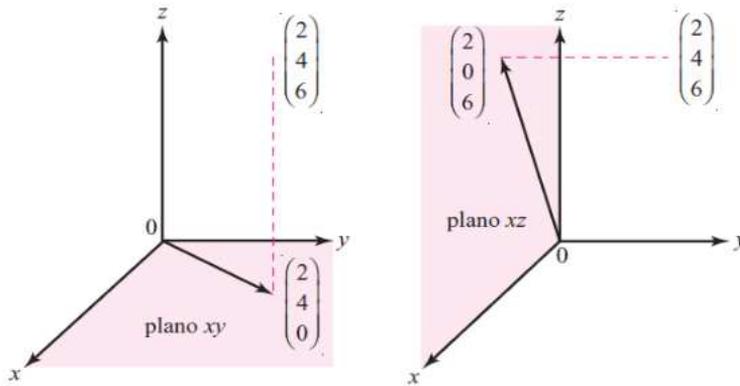
Entonces, con base en Definición de Proyección Ortogonal, que indica sea H un subespacio de \mathbb{R}^n con base ortonormal $\{u_1, u_2, u_3, u_4, \dots, u_k\}$ tal que si $v \in \mathbb{R}^n \Rightarrow$ proyección ortogonal de v sobre H , denotada por $\text{Proy}_H v$ está dada por $\text{Proy}_H v = (v * u_1)u_1 + (v * u_2)u_2 + (v * u_3)u_3 + (v * u_4)u_4 + \dots + (v * u_k)u_k$ tal que $\text{Proy}_H v \in H$, se tiene que $Pv = (v * u_1)u_1 + (v * u_2)u_2 + (v * u_3)u_3 + (v * u_4)u_4 + \dots + (v * u_k)u_k$, como $(v_1 + v_2) * u = v_1 * u + v_2 * u$ y $(\alpha v) * u = \alpha(v * u) \Rightarrow P$ es una transformación lineal.

8. Dos operadores de proyección

Se define $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ por $T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \Rightarrow T$ es operador de proyección que toma un \vec{v} de tres dimensiones y se proyecta en $\Pi_{(XY)}$.

Análogamente, $T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix}$ proyecta un \vec{v} en espacio en $\Pi_{(XZ)}$:

Figura 5.21 Dos operadores de proyección



Fuente: (Mora, W. F. 2016).

9. Operador de transposición

Se define $T: M_{mn} \rightarrow M_{nm}$ por $T(A) = A^T$. Tal que, $(A + B)^T = A^T + B^T$ y $(\alpha A)^T = \alpha A^T \Rightarrow T$, denominado operador de transposición es una transformación lineal.

10. Operador integral

Sea $J: C[0, 1] \rightarrow \mathbb{R}$ definida por $Jf: \int_0^1 f(x) dx$ tal que para $f, g \in C[0, 1]$, como $\int_0^1 [f(x) + g(x)] dx = \int_0^1 f(x) dx + \int_0^1 g(x) dx$ y $\int_0^1 \alpha f(x) dx = \alpha \int_0^1 f(x) dx \Rightarrow J$ es un operador lineal. Por ejemplo: $J(X^3) = \frac{1}{4} \Rightarrow J$ se denomina operador integral.

11. Operador diferencial

Suponga que $D: C^1[0, 1] \rightarrow C[0, 1]$ se define por $Df = f'$. Para $f, g \in C^1[0, 1]$, como $(f + g)' = f' + g'$ y $(\alpha f)' = \alpha f'$ se denota que D es un operador lineal, por lo que D se denomina operador diferencial.

12. Transformación no lineal.

Suponga que $T: C[0, 1] \rightarrow \mathbb{R}$ está definida por $Tf = f(0) + 1 \Rightarrow T$ no es lineal. Es decir: $T(f + g) = (f + g) + 1 = f(0) + g(0) + 1 \Rightarrow Tf + Tg = [f(0) + 1] + [g(0) + 1] = f(0) + g(0) + 2$. Sin embargo, este ejemplo da una idea clara que una transformación puede parecer lineal, pero en realidad no es así.

Complementariamente, las transformaciones de Box – Cox son una familia de transformaciones potenciales usadas en estadística para corregir sesgos en distribución de errores, corregir varianzas desiguales (diferentes valores de la variable predictora) y, especialmente, corregir la no linealidad en la relación (mejorar correlación entre las variables). Esta transformación recibe el nombre de los estadísticos George Box y David Cox.

La transformación potencial está definida como una función continua que varía respecto a la potencia lambda (λ).

Para datos $(Y_1, Y_2, Y_3, Y_4, \dots, Y_n)$ se realiza la transformación $Y_i' = Y_i^\lambda$ tal que $Y_i^{(\lambda)} = \begin{cases} K_1(Y_i^\lambda - 1) & \text{si } \lambda \neq 0, \\ K_2 \text{Ln}(Y_i) & \text{si } \lambda = 0 \end{cases}$ tal que K_2 es media geométrica de valores $Y_1, Y_2, Y_3, Y_4, \dots, Y_n$; en consecuencia,

$$K_2 = \left(\prod_{i=1}^n (\text{Producto desde 1 hasta } n) Y_i \right)^{\frac{1}{n}} = (Y_1 * Y_2 * Y_3 * Y_4 * \dots * Y_n)^{\frac{1}{n}}, \text{ mientras que } K_1 = \frac{1}{\lambda * K_2^{\lambda-1}}.$$

Entonces, para realizar una transformación potencial, dado un valor de lambda λ , se calcula primero la media geométrica de valores $Y_1(K_2)$. Después se sustituye este valor para calcular el parámetro K_1 .

El procedimiento para seleccionar el mejor valor λ consiste en que primero se selecciona el rango de valores lambda λ , se selecciona el que logra que la transformación se acerque al máximo a los datos tal que para cada valor de λ se hace la transformación del paso anterior.

Por último, se sustituyen de la variable (s) explicativas en diferentes funciones y se calculan los cuadrados de residuales estadísticos. Aquella que tenga el menor valor de suma de residuales será la mejor opción tal que K_2 es valor fijo para todos los casos y que solo hay que calcular de nuevo el valor K_1 .

BIBLIOGRAFÍA

- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211-252.
<https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- Christ, C. F. (1973). *The statistical implications of econometric models*. Addison-Wesley.
- Gujarati, D. N. y Porter, D. C. 2010. *Econometría*. McGraw Hill. Tercera Edición. Impreso en Colombia.
- Johnston, J. (1963). *The elements* (T. L. Heath, Trans.). Dover Publications. (Original work published ca. 300 BCE)
- Markov Wooldridge, J. M. (1821). The use of the Hessian matrix in optimization. *Journal of Mathematical Optimization*, 34(2), 45-59.
- Markov Wooldridge, J. M. (2019). *Introducción a la econometría: Métodos y aplicaciones*.
- Malinvaud, E. (1963). *Statistical methods in econometrics*. North-Holland Publishing Company.
- Martínez, C. M., & Sepúlveda, M. A. R. (1974). *Introducción al análisis factorial exploratorio*. Revista colombiana de psiquiatría, 41(1), 197-207.
- Mora, W. F. 2016. *Cálculo en Varias Variables*. Instituto Tecnológico de Costa Rica. Escuela de Matemática.
- Porter, C. 2007. *Diseños Experimentales con Aplicación a Agricultura y Agronomía*. Editorial McGrawHill.
- Porter, C. 2010. *Diseños Experimentales con Aplicación a Agricultura y Agronomía*. Editorial McGrawHill.
- Samuelson, P. A., Koopmans, T. C., & Stone, R. (1954). *Mathematical methods in the social sciences*. McGraw-Hill.

Taylor, S. Nocedal, J., & Wright, S. (2006). *Numerical optimization* (2ª ed.). Springer.

Theil, H. (1978). *Econometrics: Principles and applications*. North-Holland Publishing Company.

Theil, H. (2010). *Econometrics: Principles and applications*. North-Holland Publishing Company.

Zill, D. G., & Wright, W. S. (2011). *Elementary differential equations and boundary value problems* (9ª ed.). Brooks/Cole.



"Fundamentos Matemáticos de Diseño Experimental para Ingeniería" ofrece un enfoque integral que combina la teoría económica con herramientas matemáticas y estadísticas para abordar el diseño experimental en ingeniería. Desde los principios básicos de la econometría hasta técnicas avanzadas de análisis y modelado, este libro proporciona a los lectores una comprensión profunda de cómo aplicar métodos matemáticos-económicos en la optimización de procesos, productos y sistemas ingenieriles.



Moisés Arreguín Sámano

Profesor e investigador de la Universidad Estatal de Bolívar (UEB), Ecuador.



Andrea Damaris Hernández Allauca

Doctora en Matemáticas. Profesora de la Facultad de Recursos Naturales de la Escuela Superior Politécnica de Chimborazo(ESPOCH), además forma parte del Grupo de Investigación en Turismo (GITUR) del Instituto de Investigaciones de la ESPOCH.



Guicela Margoth Ati Cutiupala

Máster en Estadística aplicada por la Universidad de Granada. Doctora (c) en Estadística Multivariante Aplicada de la Facultad de Ciencias de la Universidad de Salamanca. Experiencia en el campo de la bioestadística, biometría, diseño experimental, servicios ecosistémicos, geoestadística. Forma parte del grupo de Investigación en Ambiente y desarrollo (GIADE) de la Escuela Superior Politécnica de Chimborazo en calidad de docente investigador.



Ángel Leyva- Ovalle

Docente Investigador Universidad Autónoma Chapingo (UACH-DiCiFo), México. Ingeniero forestal con orientación en economía y ordenación por la Universidad Autónoma Chapingo (UACH), con experiencia en la División de Ciencias Forestales (DiCiFo), específicamente en el Departamento de productos forestales.

ISBN: 978-9942-621-90-0

