

UNIVERSIDAD NACIONAL DEL CALLAO
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA PROFESIONAL DE MATEMÁTICA



**“MINIMIZACIÓN DE FUNCIONES CUASI-CONVEXAS
USANDO RETRACCIONES EN EL MÉTODO DEL
GRADIENTE SOBRE VARIEDADES RIEMANNIANAS”**
**TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICA**

PRESENTADO POR

Bach. GEAN PIERRE ROSTAING CCAPACCA

ASESOR: Mg. MARISA ELSA QUISPE CÁRDENAS

LÍNEA DE INVESTIGACIÓN: OPTIMIZACIÓN

Callao, 2019

PERÚ



ROSTAING CCAPACCA GEAN PIERRE MINIMIZACIÓN DE FUNCIONES CUASI- CONVEXAS USANDO RETRACCIONES EN EL MÉTODO DEL GRADIENTE SOBRE VARIETADES RIEMANNIANAS



Nombre del documento: ROSTAING CCAPACCA GEAN PIERRE
MINIMIZACIÓN DE FUNCIONES CUASI-CONVEXAS USANDO
RETRACCIONES EN EL MÉTODO DEL GRADIENTE SOBRE VARIETADES
RIEMANNIANAS.pdf
ID del documento: 2a26078570ca95c254206ddf0cbe01ae89471f8
Tamaño del documento original: 1,97 MB

Depositante: FONM PREGRADO UNIDAD DE
INVESTIGACION
Fecha de depósito: 20/6/2025
Tipo de carga: interface
fecha de fin de análisis: 20/6/2025

Número de palabras: 23.389
Número de caracteres: 140.482

Ubicación de las similitudes en el documento:



Fuentes de similitudes

Fuentes principales detectadas

N°	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	repositorio.unac.edu.pe 2 fuentes similares	2%		Palabras idénticas: 2% (337 palabras)
2	repositorio.unac.edu.pe 2 fuentes similares	1%		Palabras idénticas: 1% (286 palabras)
3	repositorio.unac.edu.pe 2 fuentes similares	< 1%		Palabras idénticas: < 1% (180 palabras)
4	Anthony Mendoza Osorio Método de Distancias Proximales Generali... #68076a El documento proviene de mi biblioteca de referencias 1 fuente similar	< 1%		Palabras idénticas: < 1% (53 palabras)
5	repositorio.unac.edu.pe 1 fuente similar	< 1%		Palabras idénticas: < 1% (48 palabras)

Fuentes con similitudes fortuitas

N°	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	TESIS - ANTAURCO_MENDOZA.pdf TESIS - ANTAURCO_MENDOZA #60c31b El documento proviene de mi grupo	< 1%		Palabras idénticas: < 1% (26 palabras)
2	renati.sunedu.gob.pe Registro Nacional de Trabajos de Investigación: Converge... https://renati.sunedu.gob.pe/handle/sunedu/3468600	< 1%		Palabras idénticas: < 1% (20 palabras)
3	repositorio.unac.edu.pe	< 1%		Palabras idénticas: < 1% (19 palabras)
4	ROJAS RODRIGUEZ PATTY ERIKA UNA DEMOSTRACIÓN PURAMENTE AL... #b0d1c El documento proviene de mi biblioteca de referencias	< 1%		Palabras idénticas: < 1% (21 palabras)
5	repositorio.unp.edu.pe Implementación de comandos para el cálculo de la cur... http://repositorio.unp.edu.pe/handle/UNP/2179	< 1%		Palabras idénticas: < 1% (21 palabras)

Fuentes mencionadas (sin similitudes detectadas) Estas fuentes han sido citadas en el documento sin encontrar similitudes.

- <https://doi.org/10.1137/100802529>
- <https://doi.org/10.1007/s10957-013-0307-5>
- <https://doi.org/10.1007/s10957-013>
- <https://doi.org/10.1080/02331939508844139>
- <https://doi.org/10.1023/A:1022673419987>

INFORMACIÓN BÁSICA

- FACULTAD: FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
- UNIDAD DE INVESTIGACIÓN: FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
- TÍTULO: MINIMIZACIÓN DE FUNCIONES CUASICONVEXAS USANDO RETRACCIONES EN EL MÉTODO DEL GRADIENTE SOBRE VARIETADES RIEMANNIANAS.
- ASESORA: MG. MARISA ELSA QUISPE CÁRDENAS / 0009-0009-6158-2132 / DNI: 09837409
- AUTOR: GEAN PIERRE ROSTAING CCAPACCA / 0000-0001-8383-9968 / DNI: 72534155
- LUGAR DE EJECUCIÓN: FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
- UNIDAD DE ANÁLISIS: OPTIMIZACIÓN
- TIPO DE INVESTIGACIÓN: BÁSICA
- TEMA OCDE: 1.01.01 (MATEMÁTICA PURA)

PRÓLOGO DEL JURADO

**“MINIMIZACIÓN DE FUNCIONES CUASI-CONVEXAS USANDO
RETRACCIONES EN EL MÉTODO DEL GRADIENTE SOBRE
VARIETADES RIEMANNIANAS”**

Bach. GEAN PIERRE ROSTAING CCAPACCA

Tesis presentada a consideración del cuerpo docente de la Facultad de Ciencias Naturales y Matemática de la Universidad Nacional del Callao, como parte de los requisitos para obtener el Título Profesional de Licenciado en Matemática.

Aprobado por:

MIEMBROS DEL JURADO DE SUSTENTACIÓN:

DR. WALTER FLORES VEGA	PRESIDENTE
LIC. ELMER ALBERTO LEÓN ZÁRATE	SECRETARIO
LIC. HERMINIA BERTHA TELLO BEDRIÑANA	VOCAL

ASESORA: MG. MARISA ELSA QUISPE CÁRDENAS

Fecha de Aprobación de la tesis: 25 de enero de 2019

Resolución de Jurado de Sustentación: N° 147-2018-D-FCNM



JURADO EVALUADOR PARA LA TITULACIÓN PROFESIONAL EN EL SEGUNDO CICLO DE TESIS
RESOLUCIÓN DECANAL N°147-2018-D-FCNM

ACTA DE SUSTENTACIÓN DE TESIS

En el auditorio de la Facultad de Ciencias Naturales y Matemática, sito Av. Juan Pablo II N° 306, Bellavista – Callao, siendo las 14:50 hrs. del día viernes 25 de enero de 2019, se reunieron los miembros del Jurado Evaluador del II Ciclo de Tesis para la Titulación Profesional por la modalidad de Tesis con Ciclo de Tesis de la Facultad de Ciencias Naturales y Matemática de la Universidad Nacional del Callao:

- Dr. Walter Flores Vega : Presidente
- Lic. Herminia Bertha Tello Bedriñana : Vocal
- Lic. Elmer Alberto León Zárate : Secretario

Designados por Resolución N° 147-2018-D-FCNM de fecha 28 de diciembre de 2018 a fin de proceder al acto de evaluación de la Tesis titulada: **"MINIMIZACIÓN DE FUNCIONES CUASI-CONVEXAS USANDO RETRACCIONES EN EL MÉTODO DEL GRADIENTE SOBRE VARIETADES RIEMANNIANAS"**, presentada por el señor Bachiller **ROSTAING CCAPACCA GEAN PIERRE**.

Contando con la presencia del Supervisor General, Decano de la Facultad de Ciencias Económicas Dr. Pablo Mario Coronado Arrilucea, Supervisor de la FCNM, Mg. Roel Mario Vidal Guzmán, el representante de la Comisión de Grados y Títulos Dr. Richard Saúl Toribio Saavedra y el Coordinador del II Ciclo de Tesis Lic. Absalón Castillo Valdivieso.

A continuación, se dio inicio a la sustentación de la Tesis de acuerdo a lo normado en los numerales del 10.1 al 10.4 del capítulo X de la Directiva para la Titulación Profesional por la modalidad de Tesis con Ciclo de Tesis en la Universidad Nacional del Callao, aprobada por Resolución Rectoral N° 754-2013-R del 21 de agosto de 2013, modificada por la Resolución Rectoral N° 777-2013-R de fecha 29 de agosto de 2013 y la Resolución Rectoral N° 281-2014-R del 14 de abril de 2014 con la que se modifica el Art. 4.5 del capítulo IV de la organización del Ciclo de Tesis, así como lo normado en el Reglamento de Grados y Títulos de Pregrado, aprobado por Resolución de Consejo Universitario N° 135-2017-CU de fecha 22 de junio de 2017 y también lo establecido en el Reglamento de Grados y Títulos de la UNAC aprobado por Resolución N° 309-2017-CU de fecha 24 de octubre de 2017.

Culminado el acto de sustentación, los señores miembros del Jurado Evaluador procedieron a formular las preguntas al indicado bachiller, las mismas que fueron absueltas satisfactoriamente.

Luego de un cuarto intermedio, para la deliberación en privado del Jurado respecto a la evaluación de la Tesis, se acordó calificar la Tesis sustentada por el señor bachiller **ROSTAING CCAPACCA GEAN PIERRE**, para optar el Título Profesional de Licenciado en Matemática por la modalidad de Tesis con Ciclo de Tesis, según la puntuación cuantitativa y cualitativa que a continuación se indica:

CALIFICACIÓN CUANTITATIVA	CALIFICACIÓN CUALITATIVA
16 (DIECISEIS)	MUY BUENO

Finalmente, se procedió a leer en público el acta de la sustentación realizada.

Siendo las 15:35 hrs. del día viernes veinticinco de enero del dos mil diecinueve, el señor Presidente del Jurado Evaluador dio por concluido el acto de sustentación de Tesis.

En señal de conformidad con lo actuado, se levanta la presente acta.

Dr. Walter Flores Vega
Presidente

Lic. Elmer Alberto León Zárate
Secretario

Lic. Herminia Bertha Tello Bedriñana
Vocal

DEDICATORIA

A mis padres, Rosario y Augusto, por el apoyo incondicional.

A mi esposa, por ser mi apoyo incondicional y quien me motivó en cada etapa de este proceso.

AGRADECIMIENTOS

A mi asesora, Mg. Marisa Elsa Quispe Cárdenas, por las observaciones oportunas durante el proceso de elaboración de esta tesis, además por su motivación y paciencia que fueron de gran ayuda para concluir con éxito este trabajo.

A los profesores de la especialidad de Matemática de la Facultad de Ciencias Naturales y Matemática, por la enseñanza de conocimientos tan valiosos que fueron de gran ayuda en la elaboración de este trabajo de investigación.

De forma especial, agradezco a mi esposa, por su paciencia y su amor, por recordarme en todo momento que soy capaz y por ser mi mayor pilar.

A mi perrita Peppa, por ser mi compañía incondicional, haciendo de este proceso algo mucho más ligero con su ternura y alegría.

Y a Noah, nuestro bebé en camino, por acompañarme al cierre de esta etapa con una renovada esperanza y alegría, motivándome a demostrar todo lo que puedo lograr por mí y por él.

GEAN PIERRE ROSTAING CCAPACCA

ÍNDICE

HOJA DE REFERENCIA DEL JURADO Y APROBACIÓN.....	iv
DEDICATORIA.....	v
AGRADECIMIENTO.....	vi
ÍNDICE.....	vii
RESUMEN.....	xi
ABSTRACT.....	xii
INTRODUCCIÓN.....	xiii
I. PLANTEAMIENTO DEL PROBLEMA.....	15
1.1. Descripción de la realidad problemática.....	15
1.2. Formulación del problema.....	16
1.2.1. Problema General.....	16
1.2.2. Problemas específicos.....	16
1.3. Objetivo.....	16
1.3.1. Objetivo general.....	16
1.3.2. Objetivos Específicos.....	16
1.4. Justificación.....	17
1.5. Delimitantes de la investigación.....	17
1.5.1. Teórico.....	17
1.5.2. Tecnológica.....	18
1.5.3. Económico.....	18
II. MARCO TEÓRICO.....	19
2.1.1 Internacional.....	19
2.1.2 Nacional.....	20
2.2 Bases teóricas.....	21
2.2.1 Bases epistémicas.....	21
2.2.2 Bases científicas.....	22
2.3 Marco conceptual.....	24

2.4	Definiciones de términos básicos	74
III..HIPÓTESIS Y VARIABLES		82
3.1	Hipótesis	82
3.2	Operacionalización de las variables.....	82
IV. METODOLOGÍA DEL PROYECTO.....		84
4.1	Diseño metodológico.....	84
4.2	Método de investigación.....	85
4.3	Población y muestra.....	85
4.4	Lugar de estudio y periodo desarrollado	85
4.5	Técnicas e instrumentos para la recolección de información	85
4.6	Análisis y procesamiento de datos	86
4.7	Aspectos éticos en investigación	86
RESULTADOS		87
5.1	Resultados descriptivos	87
5.2	Resultados inferenciales	87
5.3	Otro tipo de resultado, de acuerdo con naturaleza del problema..	87
5.4	Resultados del método de gradiente sobre una variedad	87
5.5	Resultados sobre el análisis de convergencia método gradiente..	91
V. DISCUSIÓN DE RESULTADOS		97
6.1	Contrastación de la hipótesis	97
6.2	Contrastación de los resultados con estudios similares	97
6.3	Responsabilidad ética	98
VI. CONCLUSIONES		99
VII. RECOMENDACIONES		100
VIII. REFERENCIAS BIBLIOGRÁFICAS		101
ANEXO		105

ÍNDICE DE TABLAS

Tabla 1. Operacionalización de la variable.....	83
---	----

ÍNDICE DE FIGURAS

Figura 1. Superficie Regular	25
Figura 2. Cambio de parámetro.....	30
Figura 3. Banda de Moebius.....	35
Figura 4. Aplicaciones diferenciales entre variedades	39
Figura 5. Curva sobre una variedad	40
Figura 6. Difeomorfismo entre variedades.....	41

RESUMEN

Las ideas de Geometría Riemanniana fueron iniciados por G. Riemann un 10 de junio de 1854, por medio de su investigación titulada: Sobre las hipótesis que están en los fundamentos de la geometría. En esta sesión afirma que toda colección continua de fenómenos puede ser considerado como espacio. Esto inició la geometría Riemanniana, siendo más específico, este tópico consiste en el estudio de variedades diferenciables, donde si a estas, se les introduce una forma particular de medir sus longitudes de curvas, estaríamos hablando de variedades Riemannianas, en esta sección se abordan ramas de la matemática, como el análisis, topología, algebra y geometría, esto depende del estudio del problema. Es por ello que, este capítulo presenta las ideas fundamentales que se aplicaran a o largo de esta tesis, tales como: variedades diferenciables y los espacios tangentes a estas variedades, además, se definirá la métrica Riemanniana, geodésica, curvatura, gradiente y Hessiano de funciones o matriz Hessiana de una función en una variedad Riemanniana. Respecto a la métrica, se denotará una clase de métricas Riemannianas diagonales, estas nos ayudaran a obtener interesantes propiedades con la finalidad de obtener nuevos algoritmos en optimización, debido a que, en una variedad Riemanniana se puede derivar un campo vectorial tangente a lo largo de una curva, la cual se llamará derivada covariante a lo largo de curvas que dependen de la métrica. Se formalizará dichos resultados en las siguientes secciones.

ABSTRACT

The ideas of Riemannian Geometry were initiated by G. Riemann on June 10, 1854, through his research entitled: On the hypotheses that are in the foundations of geometry. In this session he states that every continuous collection of phenomena can be considered as space. This initiated Riemannian geometry, more specifically, this topic consists of the study of differentiable varieties, where if a particular way of measuring their curve lengths is introduced to these, we would be talking about Riemannian varieties, in this section branches of mathematics are addressed, such as analysis, topology, algebra and geometry, this depends on the study of the problem. That is why, this chapter presents the fundamental ideas that will be applied throughout this thesis, such as: differentiable varieties and the tangent spaces to these varieties, in addition, the Riemannian metric, geodesic, curvature, gradient and Hessian of functions or Hessian matrix of a function on a Riemannian manifold will be defined. Regarding the metric, we will denote a class of diagonal Riemannian metrics. These will help us obtain interesting properties for the purpose of obtaining new optimization algorithms. This is because, on a Riemannian manifold, a tangent vector field can be derived along a curve. This derivative will be called a covariant derivative along curves that depend on the metric. These results will be formalized in the following sections.

INTRODUCCIÓN

La geometría de Riemann fue aplicada por Einstein en 1915, en su teoría de gravitación universal, la cual originó las diferentes aplicaciones de esta geometría en diferentes campos de la ciencia, las cuales son muy estudiadas en estos tiempos, sobre todo en el campo de la Optimización matemática. El estudio de la geometría Riemanniana con la optimización se origina en el año 1972, con los estudios realizados por Luemberguer, donde aplica el método de menos el gradiente geodésico, obtiene la convergencia del método gradiente para el siguiente modelo:

$$\text{Min } f(x), x \in \mathbb{R}^n$$

$$\text{s. a. } h(x) = 0$$

Donde $f: \mathbb{R} \rightarrow \mathbb{R}^m, h: \mathbb{R} \rightarrow \mathbb{R}^n, n > m$

Estas investigaciones fueron continuadas por Gabay en el año 1982, donde estudia el método del gradiente reducido, generalizando el método de Cuasi-Newton, obteniendo la convergencia superlineal. El método de gradiente o máximo descenso, estudiado por Cauchy en 1847, es uno de los métodos más conocidos para resolver problemas de optimización con funciones objetivos continuamente diferenciables. Sin embargo, para funciones arbitrarias, la convergencia no es muy fuerte ya que es global. Ahora cuando la función objetivo es convexa se tiene convergencia global y local a su vez, porque asumiendo solamente que el conjunto de soluciones óptimas es no vacío, el método del Gradiente con búsqueda de Armijo y con una regularización proximal converge a un punto óptimo. Este método en variedades Riemannianas, considerando un problema de optimización con función objetivo-arbitraria, fue estudiado por Udriste, Smith y Rapcsák, obteniendo los mismos resultados clásicos de convergencia. Para el caso convexo en estas variedades con curvatura

seccional no negativa, la convergencia global usando la regla de búsqueda de Armijo, pasos fijos y una regularización proximal, fue generalizada por Da Cruz Neto y para funciones cuasiconvexas solamente en espacios euclidianos, por Kiwiel y Murty. La dificultad que encontramos es conseguir expresiones geodésicas y precisamente hace casi imposible encontrar el siguiente punto iterativo de acuerdo con el método, razón por la cual, las retracciones sobre variedades Riemanniana son la alternativa para encontrar el óptimo siguiendo el método del gradiente, visto en P. Absil. Hasta el momento no ha sido posible la minimización de funciones cuasi convexas usando retracciones en el método del Gradiente sobre herramientas de variedades Riemanniana. De acuerdo con la literatura revisada solamente se logrado la minimización de funciones cuasiconvexas usando el método del gradiente a través de geodésicas.

I. PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la realidad problemática

En la actualidad, existen muchos problemas en las ciencias e ingeniería orientados a optimizar recursos, en un sentido previamente establecido (maximizar o minimizar), esta situación se puede representar mediante modelos matemáticos clasificados en: teoría de control, la demanda, de economía, de aproximación y otros problemas que son modelados por ecuaciones matemáticas, cuya solución se enmarca en el área de la optimización matemática, ésta a su vez, utiliza distintas metodologías de aproximación en cuanto a su solución que, de acuerdo a las características del modelo, el investigador elige un método apropiado para resolverlo. Estamos interesados en estudiar el siguiente problema:

$$\text{Min } f(x), x \in M$$

Donde M es una variedad Riemanniana y f una función definida sobre M . aquí podremos analizar cuál es su comportamiento mediante la convergencia del algoritmo del Gradiente.

El problema presentado, pertenece al campo de las ciencias básicas aplicadas al área de optimización matemática; en particular la optimización matemática expuestos sobre variedades Riemanniana que surge como una extensión natural de la teoría y métodos de optimización en espacios euclidianos a espacios más generales.

Una de las ventajas de usar herramientas de Geometría Riemanniana es que se puede construir problemas de optimización sin restricciones considerando las propiedades intrínsecas de la propia variedad. Otra ventaja es que, problemas de optimización con funciones objetivo no convexo pasan a ser convexo usando apropiadamente la métrica Riemanniana.

1.2. Formulación del problema

1.2.1. Problema General

¿Cómo minimizar funciones cuasiconvexas usando retracciones en el método del Gradiente usando herramientas de geometría Riemanniana?

1.2.2. Problemas específicos

- ¿Cómo determinar la minimización de funciones cuasiconvexas para obtener convergencia usando retracciones en el método del Gradiente?
- ¿Cómo usar las retracciones en el método del Gradiente para minimizar funciones cuasiconvexas sobre variedades Riemannianas?
- ¿El uso métricas Riemannianas específicas favorecerá la transformación problemas de minimización no convexos en convexos?

1.3. Objetivo

1.3.1. Objetivo general

Demostrar la convergencia del método del Gradiente, para la minimización de funciones cuasiconvexas usando retracciones sobre herramientas de geometría Riemannianas.

1.3.2. Objetivos Específicos

- Obtener la minimización mediante la convergencia de funciones cuasiconvexas usando retracciones en el método del Gradiente, sobre herramientas de variedades Riemannianas.
- Analizar las retracciones en el método del Gradiente para la minimización funciones cuasiconvexas sobre variedades Riemannianas.

- Definir métricas Riemannianas específicas para transformar problemas de minimización no convexos en convexos sobre variedades Riemannianas.

1.4. Justificación

En la presente investigación se analizará en la búsqueda de la convergencia del método del gradiente, con lo cual se utilizarán aproximaciones en la geodésica, a partir de las retracciones R , generando así las sucesiones $x_{k+1} = R_{x_k}(t_k \eta_k)$; donde $\eta_k \in T_{x_k}M$, con la búsqueda de Armijo sea convergente en un punto, nótese que $T_{x_k}M$ es un espacio tangente de la variedad M en el punto x_k , además t_k es una constante, η_k es la dirección de la gradiente. Este concepto de Optimización de funciones convexas aplicando las variedades Riemannianas en general fue estudiado por Udriste (1997). La convergencia que se plantea en la forma general del método de minimización por el método de gradiente sobre funciones cuasiconvexas usando la búsqueda de Armijo, esta investigación fue realizada por Papa Q. et. al. (2008), sin embargo la metodología fue propuesta por Absil P. et al (2012), donde se comienza a buscar la convergencia sobre estas variedades Riemannianas de curva no negativa, de esta forma se logra garantizar un punto candidato para ser mínimo en la función cuasiconvexa.

1.5. Limitantes de la investigación

1.5.1. Teórico

La investigación realizada en este trabajo busca, mediante la aplicación de la teoría y las definiciones básicas de geometría Riemanniana y optimización, encontrar la convergencia de un método, la cual ayudará a obtener el mínimo valor aplicada a una función cuasiconvexa. Esta minimización será de importancia para los estudiantes, docentes e investigadores en general, debido a que se obtendrá resultados que son

más generales de las existentes en la literatura de optimización matemática, debido a que ya se conocen otros métodos para hallar la minimización de funciones, lo que se busca con este trabajo es optimización la acción y el trabajo para hallar dicho valor, a través de las retracciones y la búsqueda de Armijo.

1.5.2. Tecnológica

Los resultados de la generalización de la convergencia del algoritmo del Gradiente contribuirán a generar algoritmos sofisticados, los que, al ser implementados computacionalmente, permitirán aplicaciones tecnológicas importantes y soporte computacional en el área de ciencias e ingeniería.

1.5.3. Económico

El modelamiento del problema de minimización de funciones cuasiconvexas permite reducir costos de implementación computacional, que conlleva al ahorro en tiempo y dinero.

II. MARCO TEÓRICO

2.1. Antecedentes

2.1.1 Internacional

Gabay (1992) donde, del punto de vista de esta teoría, estudia el método de gradiente reducido, generaliza los métodos de Cuasi-Newton obteniendo convergencia superlineal. Gabay también hace un análisis computacional mostrando que la teoría y la práctica interrelacionadas pueden dar buenos resultados. El método de máximo descenso o llamado también método del Gradiente, estudiado por Cauchy en 1847, es uno de los métodos más antiguos y conocidos en la literatura para resolver problemas de optimización con funciones objetivo continuamente diferenciables. Sin embargo, para una función arbitraria los resultados de convergencia no son muy fuertes ya que la convergencia global, como también la existencia de puntos de acumulación no son garantizados. Solamente podemos asegurar que cualquier punto de acumulación, si existe, es un punto crítico del problema. La situación es muy diferente cuando la función objetivo es convexa, porque asumiendo solamente que el conjunto de soluciones óptimas es no vacío, el método del Gradiente con búsqueda de Armijo y con una regularización proximal converge a un punto óptimo. Este método en variedades Riemannianas, considerando un problema de optimización con función objetivo-arbitraria, fue estudiado por Udriste, Smith y Rapcsák, obteniendo los mismos resultados clásicos de convergencia. Para el caso convexo en estas variedades con curvatura seccional no negativa, la convergencia global usando la regla de búsqueda de Armijo, pasos fijos y una regularización proximal, fue generalizada por Da Cruz Neto y para funciones cuasiconvexas solamente en espacios euclidianos, por Kiwiel y Murty. Esta tesis, está orientada a extender los

resultados de Kiwiel y Murty en variedades Riemannianas completas y con curvatura seccional no negativa.

2.1.2 Nacional

Quispe (2008) en su investigación titulada “*El método de máximo descenso para funciones cuasiconvexas en variedades Riemannianas*”, donde el problema que se planteó fue de obtener la minimización en funciones cuasi convexas, donde por medio de propiedades de geométricas y geodésicas, y aplicando el método de máximo descenso se hallará la convergencia usando la regla generalizada de búsqueda de Armijo. El método que se aplicó constituye en el inductivo – deductivo. Los resultados de esta investigación es que se prueba que la sucesión generada por el método, usando la regla de búsqueda generalizada de Armijo, converge a un punto crítico de la función, mostrando la importancia del uso de métricas diagonales que faciliten el análisis de un problema de optimización en variedades Riemannianas. Nuestro resultado extiende la convergencia del método, del caso convexo para el caso cuasiconvexo.

Cruzado (2014) en su investigación titulada “*Método de máximo descenso usando retracciones en variedades Riemannianas*”, donde se plantea el problema de optimizar funciones en una variedad Riemanniana, la cual aplicara el método de máximo descenso con retracciones, para luego resolver un problema de búsqueda lineal, finalmente plantear el método en MATLAB como aplicación para algunos problemas particulares. El estudio de la investigación es de carácter teórico-aplicativo. En el desarrollo del trabajo si se obtiene los resultados deseados, esto será un nuevo aporte científico, muy significativo en optimización en la búsqueda de algoritmos más eficientes tanto teórica como computacionalmente. El método de esta investigación es inductivo – deductivo, finalmente se concluye que aplicando este método probamos que todos los puntos de acumulación de la sucesión de iteraciones generada por el algoritmo dado es un punto crítico de la función objetivo, siendo este hecho de suma importancia para

analizar la convergencia, dicho resultado involucra al menor y al mayor autovalor del Hessiano en un punto crítico. Además, Aplicó el método desarrollado para minimizar el cociente de Rayleigh sobre la esfera unitaria y también para minimizar la función de Brocket sobre la variedad de Stiefel e implementamos sus respectivos algoritmos en MATLAB.

2.2 Bases teóricas

2.2.1 Bases epistémicas

Las bases epistemológicas que orientan esta investigación se fundamentan en los principios de la geometría euclidiana, establecidos por Euclides entre los años 325 a.C. y 265 a.C., considerada un pilar esencial en el desarrollo del pensamiento científico occidental. Este modelo fue cuestionado posteriormente por destacados matemáticos como Bolyai, Gauss y Lobachevski, quien desafió el quinto postulado euclidiano y propuso una alternativa revolucionaria: la geometría hiperbólica. Esta propuesta generó un profundo impacto en la visión científica de la época al evidenciar que era posible construir nuevas formas de entender el espacio. A partir de ese momento, Lobachevski desarrolló formalmente la geometría hiperbólica, sustentándola con rigurosas demostraciones que consolidaron su validez como una nueva forma de geometría. Más adelante, el matemático alemán Félix Klein, en 1870, reinterpreto esta geometría, demostrando que en ella se cumplen todas las propiedades de la geometría euclidiana excepto el quinto postulado. Esto permitió confirmar la existencia de geometrías no euclidianas, como la hiperbólica, la esférica, la elíptica y la geometría diferencial.

En cuanto a la geometría elíptica, fue Bernhard Riemann (1826–1866) quien introdujo sus fundamentos. En este sistema, el plano se concibe como la superficie de una esfera, donde un punto equivale a un par de puntos diametralmente opuestos y las líneas rectas corresponden a círculos máximos o geodésicas. Esta visión, conocida como geometría

Riemanniana, resulta especialmente adecuada para describir fenómenos del mundo físico. De hecho, la teoría de la relatividad adopta esta geometría para explicar la métrica del universo, lo cual impulsó el avance de diversas disciplinas científicas e ingenierías a través del uso de la geometría diferencial.

En tiempos recientes, una de las ramas que ha evolucionado sobre este marco geométrico es la Optimización Matemática. Por ejemplo, Wen H. (2013) en su tesis doctoral desarrolló modelos computacionales basados en funciones generales —diferentes a las utilizadas en este estudio, que emplea funciones cuasiconvexas—. También se destacan investigaciones relacionadas como las de Baygorrea (2017), Papa et al. (2015), Da Cruz Neto (2013) y Bento et al. (2013), realizadas principalmente en el noreste de Brasil, especialmente en la Universidad Federal de Piauí, donde se ha concentrado una importante producción académica en torno a las variedades Riemannianas.

2.2.2 Bases científicas

El sustento teórico de esta investigación se centra, entre otros elementos, en el análisis del problema a partir de modelos matemáticos de carácter abstracto. Estos serán examinados rigurosamente para posteriormente estructurar una metodología basada en procesos tanto deductivos como inductivos, con el fin de obtener resultados clave que permitan resolver el modelo planteado. El problema central gira en torno a la programación no lineal, y su objetivo es hallar soluciones óptimas en el contexto de funciones cuasiconvexas. Este problema se aborda desde tres ejes fundamentales: las variedades Riemannianas, la optimización continua y la implementación del método del gradiente. El estudio se desarrolla sobre el marco de las variedades Riemannianas, con el objetivo de alcanzar criterios de convergencia adecuados que faciliten la identificación de óptimos. La estrategia metodológica se construye a partir del análisis del

comportamiento de algoritmos aplicados previamente a problemas similares, comparando factores como el tiempo de ejecución y la eficiencia en la localización de soluciones óptimas. A partir de este contraste, se valida la hipótesis de convergencia del método del gradiente, hipótesis que serán sometidas a verificación y análisis riguroso para evaluar su efectividad.

Para sustentar este trabajo, se ha elaborado un marco teórico sólido que se apoya en estudios y referencias científicas reconocidas, en las cuales se presentan resultados relevantes que sirvieron como base para la formulación de las hipótesis. Estos fundamentos conceptuales se desarrollan y explican en las siguientes secciones.

Investigación de operaciones

En el campo de la investigación de operaciones, se han desarrollado modelos de programación tanto lineal como no lineal. En la programación lineal, tanto las variables de decisión como la función objetivo presentan una estructura lineal. En cambio, en los modelos no lineales, las variables de decisión y/o la función objetivo o las restricciones se expresan mediante funciones no lineales. Esta flexibilidad permite modelar situaciones en las que los supuestos de proporcionalidad no son aplicables, como en el caso de economías o deseconomías de escala.

Dentro de este enfoque que se analizará en la investigación, el método del gradiente también conocido como método de Cauchy o método del descenso más pronunciado constituye un algoritmo específico para la resolución de problemas de programación no lineal sin restricciones. Se clasifica como un método general de descenso en diversos espacios, especialmente analizado en variedades Riemannianas, donde la búsqueda de soluciones óptimas implica la resolución secuencial de problemas unidimensionales.

Métodos computacionales

Los métodos computacionales se definen como técnicas numéricas que permiten aproximar soluciones de forma eficiente. Estas herramientas son estudiadas desde una perspectiva matemática para evaluar aspectos como su precisión, eficacia y aplicabilidad. Tal como se menciona en Burden (2016), estos métodos permiten abordar una amplia variedad de problemas provenientes de distintas disciplinas científicas. Su implementación se realiza a través de software especializado como Maple, Fortran y Matlab, siendo este último uno de los más utilizados en el ámbito de la ingeniería por su versatilidad y potencia en cálculos numéricos complejos o en diversas ramas.

2.3 Marco conceptual

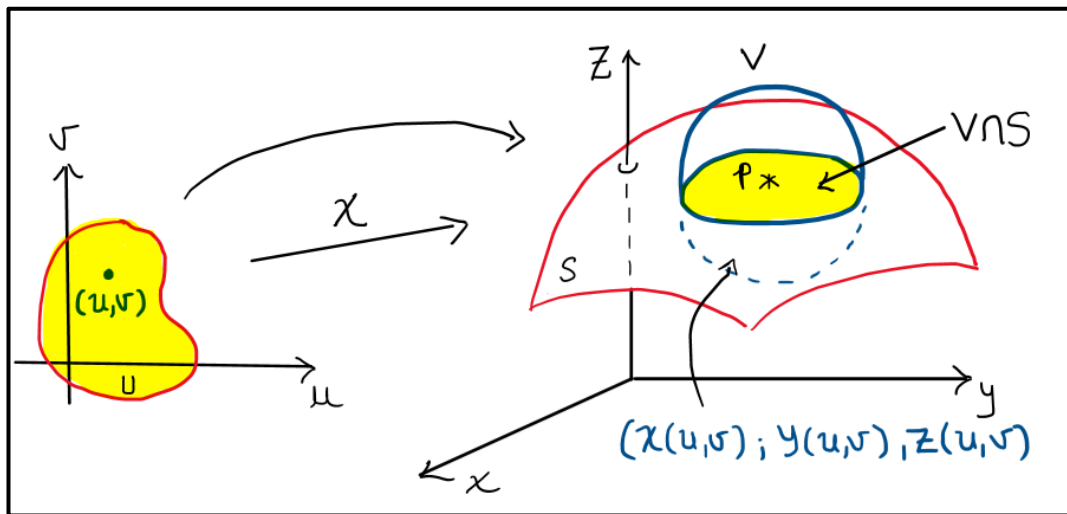
Definición 1. (*Superficie Regular de \mathbb{R}^n*) Un subconjunto $S \subset \mathbb{R}^n$ es una superficie regular de \mathbb{R}^n de dimensión $k \leq n$ si para cada $p \in S$ existe una vecindad V de p en \mathbb{R}^n , un subconjunto abierto $U \subset \mathbb{R}^k$ y una aplicación sobreyectiva $X: U \rightarrow S \cap V$ tal que:

1. X es diferenciable en U
2. X es homeomorfismo, esto significa que si X es diferenciable si y solo si, su inversa es continua.
3. Para todo $q \in U$, $dX_q: \mathbb{R}^k \rightarrow \mathbb{R}^n$ es inyectiva, donde dX_q es la diferencial de X en el punto q .

Para cada $p \in S$, la aplicación $X: U \rightarrow S \cap V$ es llamado parametrización de S en p , o sistema de coordenadas locales en p . $S \cap V$ es llamado vecindad coordenada de p .

En forma particular, podemos realizar un gráfico en \mathbb{R}^3 , tomando $S \subset \mathbb{R}^3$; un punto $p \in S$, existe una vecindad V en \mathbb{R}^3 y una aplicación $X: U \rightarrow S \cap V$ de un subconjunto abierto U de \mathbb{R}^2 sobre $S \cap V$.

Figura 1. Superficie Regular



Fuente: Elaboración propia

Ejemplo 1. Se demostrará que la esfera en unidad es una superficie regular

$$S^2 = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$$

Primero verificaremos que la aplicación $x_1: U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dada por

$$x_1(x, y) = \left(x, y, \sqrt{1 - (x^2 + y^2)}\right); (x, y) \in U$$

Donde $\mathbb{R}^2 = \{(x, y, z) \in \mathbb{R}^3, z = 0\}$ y $U = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 1\}$ es una parametrización de S^2 . Observe que $x_1(U)$ es la parte abierta de S^2 por encima del plano XY .

Debido a que $x^2 + y^2 < 1$, la función $\sqrt{1 - (x^2 + y^2)}$ tiene derivadas parciales continuas de todos los órdenes. Por ende, x_1 es diferenciable y cumple la condición 1.

La condición 3 se verifica inmediatamente puesto que

$$\frac{\partial(x, y)}{\partial(x, y)} = 1$$

Para comprobar la condición 2, se observa que x_1 es inyectiva y que x_1^{-1} es la restricción de la proyección continua $\pi(x, y, z) = (x, y)$ sobre el conjunto $x_1(U)$, por lo tanto, x_1^{-1} es continua en $x_1(U)$.

Ahora recubriremos la esfera con parametrizaciones similares

$$x_2(x, y) = (x, y, -\sqrt{1 - (x^2 + y^2)})$$

$$x_3(x, y) = (x, \sqrt{1 - (x^2 + z^2)}, z)$$

$$x_4(x, y) = (x, -\sqrt{1 - (x^2 + z^2)}, z)$$

$$x_5(x, y) = (\sqrt{1 - (y^2 + z^2)}, y, z)$$

$$x_6(x, y) = (-\sqrt{1 - (y^2 + z^2)}, y, z)$$

Los cuales junto a x_1 recubren totalmente a S^2 , demostrando así que S^2 es una superficie regular.

Ejemplo 2. Sea $\rho = \{(x, y, z) \in \mathbb{R}^3, x = y\}$ un plano y sea $x: U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por

$$x(u, v) = (u + v, u + v, uv)$$

Donde $U = \{(u, v) \in \mathbb{R}^2, u > v\}$ donde $x(U) \subset P$, además x es una parametrización de P , observe que $x(U)$ es la parte abierta de P .

Se puede observar que x es diferenciable, por ello cumple la condición 1. La condición 3 se da debido a que se cumple que la derivada parcial es mayor a cero, como se puede observar

$$\frac{\partial(x, z)}{\partial(u, v)} = u - v > 0$$

Debido a esto podemos probar la condición 2, ya que cumple el teorema de la función inversa. Por lo tanto, x es un homeomorfismo.

Proposición 1. Sea U un subconjunto abierto de \mathbb{R}^n y $F: U \rightarrow \mathbb{R}^m$ una función diferenciable en U con valor regular $a \in \mathbb{R}^m$, entonces $F^{-1}(a)$ es superficie regular de dimensión $n - m$.

Demostración: Sea $p \in f^{-1}(A)$ denotamos así los puntos y las contantes definidas

$$x = (x_1, \dots, x_k); y = (y_1, \dots, y_m); a = (a_1, \dots, a_m)$$

$$(x, y) = (x_1, \dots, x_k, y_1, \dots, y_m)$$

Se denota la función f así $f(x, y) = (f_1(x, y), \dots, f_m(x, y))$. Como a es valor regular de f , se asume, reordenando los ejes si es necesario, donde la derivada parcial se obtiene

$$\frac{\partial(f_1, \dots, f_m)}{\partial(y_1, \dots, y_m)}(p) \neq 0$$

En p se define $F: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ por $F(x, y) = (x_1, \dots, x_k, f_1(x, y), \dots, f_m(x, y))$ entonces:

$$\det(dFp) = \begin{vmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_k} & \frac{\partial f_1}{\partial y_1} & \dots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_k} & \frac{\partial f_m}{\partial y_1} & \dots & \frac{\partial f_m}{\partial y_m} \end{vmatrix} = \frac{\partial(f_1, \dots, f_m)}{\partial(y_1, \dots, y_m)} \neq 0$$

El teorema de la función inversa garantiza la existencia de conjuntos abiertos U de p y V de $f(p)$ tal que $F: U \rightarrow V$ es difeomorfismo y sigue $F^{-1}: V \rightarrow U$ es difeomorfismo y está definido:

$$F^{-1}(x_1, \dots, x_k, t_1, \dots, t_m) = (x_1, \dots, x_k, g(x_1, \dots, x_k, t_1, \dots, t_m))$$

Donde $(x, t) = (x_1, \dots, x_k, t_1, \dots, t_m) \in V$ y además se cumple que el punto evaluado en g es

$$g(x_1, \dots, x_k, t_1, \dots, t_m) = (g_1(x, t), \dots, g_m(x, t))$$

Se denota la función proyección de \mathbb{R}^n sobre \mathbb{R}^k por π esto significa que $\pi(x, y) = x$. Ahora cualquier punto $(x, y) \in f^{-1}(a) \cap U$ tiene la forma

$$(x, y) = F^{-1} \circ F(x, y) = F^{-1}(x_1, \dots, x_k, f(x, y)) = F^{-1}(x, a) = (x, g(x, a))$$

Con x en el abierto $\pi(U)$ de \mathbb{R}^k . Sea $h(x) = g(x, a)$ entonces

$$f^{-1}(a) \cap U = \{(x, h(x)): x \in \pi(U) = \text{graf}(h \cap U)\}$$

Lo que muestra que $f^{-1}(a) \cap U$ es una carta local, por ser gráfica de la función diferenciable, por lo tanto cualquier punto $p \in f^{-1}(a)$ se puede cubrir con una carta local.

Corolario 1. Sea U un subconjunto abierto de \mathbb{R}^n y $f: U \rightarrow \mathbb{R}$ una función diferenciable tal que $(\nabla f)(x) \neq 0$, para todo $x \in f^{-1}(a)$. Entonces es una superficie regular de la forma $S = f^{-1}(a)$.

Demostración: Como se sabe que $(\nabla f)(x) \neq 0$ donde a es un valor regular, luego por la Proposición 1 se obtiene el corolario, esta proposición nos indica que si cumple las condiciones mencionadas será una superficie regular.

Podemos aplicar esta teoría en el campo de la optimización matemática, mostrando así, una familia de superficies regulares en dicho campo.

Ejemplo 3. Consideremos los problemas de optimización lineal

$$\begin{array}{ll}
 (P) \min c^T x & (D) \max b^T \lambda \\
 s \cdot a. Ax = b & s \cdot a. A^T \lambda + s = c \\
 x \geq 0 & s \geq 0
 \end{array}$$

Donde $x, s, c \in \mathbb{R}^n$; $\lambda, b \in \mathbb{R}^m$ y $A \in \mathbb{R}^{m \times n}$ es de rango $m < n$. El problema (P) es el primal y (D) es el dual.

- 1) Si $S = \{x \in \mathbb{R}_{++}^n : Ax = b\}$ es el conjunto de restricciones del problema primal (P), definimos la función $f: \mathbb{R}_{++}^n \rightarrow \mathbb{R}^m$ donde $F(x) = Ax - b$, tenemos que $dFx = A$ para todo $x \in \mathbb{R}_{++}^n$, donde A es una matriz de rango m . Aplicando la Proposición 1 se tiene que $F^{-1}(0) = S$ es superficie regular de dimensión $n - m$.
- 2) Si $S = \{(\lambda, s) \in \mathbb{R}^n \times \mathbb{R}_{++}^n : A^T \lambda + s = c\}$ es el conjunto de restricciones estrictas del problema dual (D), definimos la siguiente función

$$f: \mathbb{R}^n \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}^n$$

Donde

$$F(\lambda, s) = A^T \lambda + s = [A^T I] \begin{bmatrix} \lambda \\ s \end{bmatrix} - C$$

Tenemos que $dF(\lambda, s) = [A^T I]$ con rango n , para todo $(\lambda, s) \in \mathbb{R}^n \times \mathbb{R}_{++}^n$, por la Proposición 1, tenemos que $F^{-1}(0) = S$ es una superficie regular de dimensión m .

Ejemplo 4. Consideremos un problema de programación no lineal PNL de la siguiente forma:

$$\begin{aligned} \min f(x) \\ s \cdot a. \quad h(x) = 0 \\ x \in U \end{aligned}$$

Donde U es un abierto de \mathbb{R}^n , $f: U \rightarrow \mathbb{R}$ y $h: U \rightarrow \mathbb{R}^m$ son funciones, Si h es diferenciable y su $J_n(x)$ matriz jacobiana tiene rango m , el conjunto $\{x \in U; h(x) = 0\}$ es superficie regular, tenemos en particular que $\{x \in \mathbb{R}^n; h(x) = 0\}$ y $\{x \in \mathbb{R}^n; h(x) = 0 \text{ y } x > 0\}$ son superficies regulares en el espacio \mathbb{R}^n .

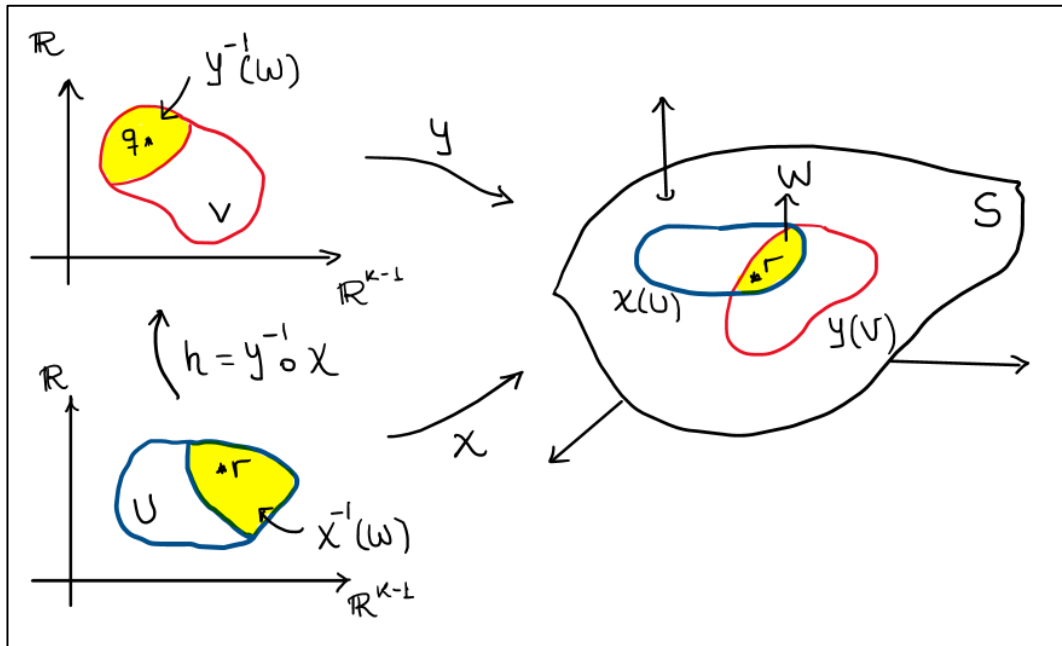
Podemos observar que la regularidad de la superficie no depende de la parametrización, es decir, podemos usar otros parámetros de la superficie en el punto $p \in S$, con sus propiedades diferenciables. Esto nos ayuda a generalizar las definiciones de superficies regulares a variedades diferenciables.

Teorema 1. (*Cambio de parámetro*) Sea $X \rightarrow S$ y $Y \rightarrow S$ dos parametrizaciones de S en el punto p tales que:

$$W = X(U) \cap Y(V) \neq \emptyset$$

La aplicación $Y^{-1} \circ X: X^{-1}(W) \rightarrow Y^{-1}(W)$ es llamado cambio de parámetro, además es un difeomorfismo. Esto es que la función de la composición $Y^{-1} \circ X = h$ es diferenciable y tiene función inversa, h^{-1} es también diferenciable.

Figura 2. Cambio de parámetro



Fuente: Elaboración propia

Demostración: Es una aplicación del teorema de la función inversa. En efecto $h = Y^{-1} \circ X$ es un homeomorfismo, ya que esta compuesta de dos homeomorfismos.

Sea $r \in X^{-1}(W)$ y $q = h(r)$ si

$$(u_1, \dots, u_k) \in V \subseteq \mathbb{R}^k, (v_1, \dots, v_n) \in \mathbb{R}^n$$

Además, sea $y(u_1, \dots, u_k) = (v_1(u_1, \dots, u_k), \dots, v_n(u_1, \dots, u_k))$ definida como una parametrización de S entonces la diferencial de Y en cualquier punto de su dominio tiene rango k , por tanto, se puede asumir renombrando los ejes si es necesario que,

$$\frac{\partial(v_1, v_2, \dots, v_k)}{\partial(u_1, u_2, \dots, u_k)}(q) \neq 0$$

Se extiende y a la función $F: V \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^n$ definida por:

$$F(u_1, \dots, u_k, t_{k+1}, \dots, t_n) = (v_1(u), \dots, v_k(u), v_{k+1}(u) + t_{k+1}, \dots, v_n(u) + t_n)$$

Donde $u = (u_1, \dots, u_k) \in V, t_i \in \mathbb{R}$ es claro que F es diferenciable y que la restricción $F / V \times \{0\}$ se obtiene

$$\det dF_q = \begin{vmatrix} \frac{\partial v_1}{\partial u_1} & \dots & \frac{\partial v_1}{\partial u_k} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial v_k}{\partial u_1} & \dots & \frac{\partial v_k}{\partial u_k} & 0 & \dots & 0 \\ \frac{\partial v_{k+1}}{\partial u_1} & \dots & \frac{\partial v_{k+1}}{\partial u_k} & & & \\ \frac{\partial v_{k+1}}{\partial u_1} & \dots & \frac{\partial v_{k+1}}{\partial u_k} & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial v_n}{\partial u_1} & \dots & \frac{\partial v_n}{\partial u_k} & 0 & \dots & 1 \\ \frac{\partial v_n}{\partial u_1} & \dots & \frac{\partial v_n}{\partial u_k} & & & \end{vmatrix} = \frac{\partial(v_1, \dots, v_k)}{\partial(u_1, \dots, u_k)}(q) \neq 0$$

Por ello es posible aplicar la propiedad de función inversa, la cual garantiza la existencia de un par de conjuntos abiertos V_1 de $Y(q)$ en \mathbb{R}^n y V_2 de $\{q \times 0\}$ en \mathbb{R}^n tal que F es un difeomorfismo.

Por la continuidad de X , existe un conjunto abierto U_1 de $r \in V$ tal que $X(U_1) \subseteq V_1$.

Nótese que sobre $U_1, h|_{U_1} = F^{-1} \circ x|_{U_1}$ es una composición de funciones diferenciables de esta manera se puede aplicar la regla de la cadena para concluir que h es diferenciable en r ; como r es arbitrario, h es diferenciable en $X^{-1}(W)$. De la misma manera se puede demostrar que h^{-1} es diferenciable y así h es un difeomorfismo.

Observación 1. Sea M una k -superficie contenida en \mathbb{R}^n y $F = \{(u_i, \varphi_i)\}$ una estructura diferenciable de M .

- a) Si (u_i, φ_i) y (u_j, φ_j) son elementos de F con $p \in \varphi_i(u_i) \cap \varphi_j(u_j) = W$, entonces el teorema de cambio de parámetro dice

$$h = \varphi_j^{-1} \circ \varphi_i: \varphi_i^{-1}(W) \rightarrow \varphi_j^{-1}(W)$$

Es un difeomorfismo, es decir, si las coordenadas de (u_i, φ_i) y (u_j, φ_j) son x_1, \dots, x_k e y_1, \dots, y_k respectivamente, entonces h se representa por

las funciones $y_1 = y_1(x_1, \dots, x_k)$ e $y_k = y_k(x_1, \dots, x_k)$ y para cada q en el dominio de h

$$\frac{\partial(y_1, \dots, y_k)}{\partial(x_1, \dots, x_k)}(p) \neq 0$$

- b) El teorema de cambio de parámetro garantiza que para cada una de las parametrizaciones (u_i, φ_i) existen subconjuntos abiertos de la forma $U_i \times \mathbb{R}^{n-k}$ de espacio euclidiano \mathbb{R}^n y una función $F_i: U_i \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^n$ tal que F_i es un difeomorfismo de una vecindad abierta en el conjunto $\varphi_i^{-1}(p) \in U_i \times \mathbb{R}^{n-k}$ sobre una vecindad abierta de $p \in M \subseteq \mathbb{R}^n$, con $F_i|_{U_i} = \varphi_i$ lo que indica que cada φ_i^{-1} es diferenciable.
- c) A la familia $\{(v_i, \psi_i)\}$, donde $v_i = \varphi_i(U_i)$ y $\psi_i = \varphi_i^{-1}$, se conoce como atlas para M y al par (v_i, ψ_i) es una carta.

Proposición 2. Sea X una parametrización de la superficie regular S , tal que $X: U \subset \mathbb{R}^2 \rightarrow S$ en el punto $q \in U$, el subespacio de dimensión 2 $dX_q(\mathbb{R}^2) \subset \mathbb{R}^3$ y coincide con el conjunto de vectores tangentes a S en $X(q)$

Demostración: Para la demostración de esta proposición se tomó el libro de Do Carmo (2025) p.98. Donde se puede observar que toman la noción de plano tangente para hablar así, de la diferencial de una aplicación e insertar este concepto.

Definición 2. Sean S_1 y S_2 dos superficies regulares y considerando $\phi: V \subset S_1 \rightarrow S_2$ una aplicación diferenciable de un conjunto abierto V de S_1 en S_2 . Si $p \in V$ se sabe que todo vector tangente $w \in T_p S_1$ es el vector velocidad $\alpha'(0)$ de una curva diferenciable α .

Teniendo $\alpha: (-\epsilon, \epsilon) \rightarrow V$ con $\alpha(0) = p$. La curva $\beta = \phi \circ \alpha$ es tal que $\beta(0) = \phi(p)$, y por tanto $\beta'(0)$ es un vector de $T_{\phi(p)} S_2$.

Proposición 3. En consecuencia con la anterior definición, dado que w ; un vector $\beta'(0)$ no depende del camino α . La aplicación

$$d\varphi_p: T_p S_1 \rightarrow T_{\varphi(p)} S_2$$

Definida por $d\varphi_p(w) = \beta'(0)$ es de la forma lineal

Demostración: Se tomó como referencia la demostración de Do Carmo (2005) p.150.

Proposición 4. Si S_1 y S_2 son superficies regulares y $\varphi: V \subset S_1 \rightarrow S_2$ es una aplicación diferenciable de un conjunto abierto $U \subset S_1$ tal que la diferencial $d\varphi_p$ en $p \in U$ es un isomorfismo, entonces φ es un difeomorfismo local del punto p .

Demostración: Aplicando el teorema de la función inversa se puede obtener este resultado.

Definición 3. (*Variedad diferenciable*) Una variedad diferenciable de dimensión n es un conjunto M y una familia de aplicaciones inyectivas $X_\alpha: U_\alpha \rightarrow M, \alpha \in I$ (conjunto de parámetros), definidos en abiertos U_α de \mathbb{R}^n en M tales que cumplen:

1. $M = \bigcup_{\alpha \in I} X_\alpha(U_\alpha)$
2. $\forall X_\alpha, X_\beta$ con $X_\alpha(U_\alpha) \cap X_\beta(U_\beta) = W \neq \emptyset$ los conjuntos $X_\alpha^{-1}(W)$ y $X_\beta^{-1}(W)$ son abiertos en \mathbb{R}^n y las aplicaciones $X_\beta^{-1} \circ X_\alpha: X_\alpha^{-1}(W) \rightarrow X_\beta^{-1}(W)$ son diferenciables.
3. La familia $\{(U_\alpha, X_\alpha)\}$ es máxima relativamente de las condiciones mencionadas (1) y (2).

La condición (3) aparece por razones solo técnicas. Dada una estructura diferenciable en M , podemos completarla en una máxima, agregando a ella, todas las parametrizaciones que cumplen la condición 2. Por tanto,

podemos decir que una variedad diferenciable es un conjunto pequeño de una estructura diferenciable.

Observación 2. El par (U_α, X_α) con $p \in X_\alpha(U_\alpha)$ es llamada parametrización. Una familia $\{(U_\alpha, X_\alpha)\}$ satisface los ítems 1 y 2 es llamada estructura diferenciable de M . Así la variedad M es un conjunto con estructura diferenciable.

Una estructura diferenciable es una variedad diferenciable M induciendo de forma natural topología en M definida por $A \subset M$ es un abierto en M si $\forall \alpha \in I, X_\alpha^{-1}(A \cap X_\alpha(U_\alpha))$ es abierto en \mathbb{R} . Observemos que la topología es definida de tal manera que los conjuntos $X_\alpha(U_\alpha)$ son abiertos y las aplicaciones X_α son continuas.

Ejemplo 5. La esfera $S^n = \{x \in \mathbb{R}^{n+1} / \|x\| = 1\}$ es una variedad diferenciable de dimensión n donde sus abiertos coordenados son los conjuntos

$$U_{k_j} = \{x \in S^n / (-1)^j x^k > 0\} \quad (j = 0, 1, \dots, k = 1, \dots, n + 1)$$

Los homeomorfismos coordenados son

$$\varphi_{k_i}: U_{k_j} \rightarrow B_0^n(1) = \{x \in \mathbb{R}^n / \|x\| = 1\}$$

$$x \rightarrow (x^1, \dots, x^{k-1}, x^{k+1}, \dots, x^{n+1})$$

Resulta fácil demostrar que este atlas es diferenciable pues la aplicación que se define a continuación

$$\varphi_{k_j}^{-1}: B_0^n(1) \rightarrow S^n$$

Tiene la coordenada k -ésima igual a:

$$(-1)^j \left(1 - \sum_{i \neq k} (x^i)^2 \right)^{1/2}$$

Que es una función diferenciable en el sentido usual $B_0^n(1)$ y φ_{k_j} es la restricción de una aplicación diferenciable de \mathbb{R}^{n+1} en \mathbb{R}^n .

$$W_1 = \{(x, y) \in \mathbb{R}^2 / 2 < x < 8, 0 < y < 2\}$$

$$W_2 = \{(x, y) \in \mathbb{R}^2 / 6 < x < 14, 0 < y < 2\}$$

Como $\tilde{\varphi}_1$ y $\tilde{\varphi}_2$ son compatibles con la relación de equivalencia en R se inducen las aplicaciones $\tilde{\varphi}_1: U_1 \rightarrow W_1 \subset \mathbb{R}^2$ y $\tilde{\varphi}_2: U_2 \rightarrow W_2$ donde se comprueba su homeomorfismo. Para ver que las funciones generan un atlas $\{(U_1, \varphi_1), (U_2, \varphi_2)\}$ de clase C^∞ en M , basta comprobar que $\varphi_2 \circ \varphi_1^{-1}$ y $\varphi_1 \circ \varphi_2^{-1}$ son difeomorfismo. En efecto, por ejemplo

$$\varphi_2 \circ \varphi_1^{-1}: \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2)$$

Este dado por

$$(\varphi_2 \circ \varphi_1^{-1})(x, y) = \begin{cases} (x, y) & \text{si } 6 < x < 8, \quad 0 < y < 2 \\ (10 + x, 2 - y) & \text{si } 2 < x < 4, \quad 0 < y < 2 \end{cases}$$

Proposición 5. Toda superficie regular de \mathbb{R}^n de dimensión m es una variedad diferenciable de la misma dirección.

Demostración: Sea M una superficie regular en el espacio \mathbb{R}^n de dimensión k , entonces por definición, para cada punto $p_\alpha \in M$ existe una vecindad V_α de p_α y una aplicación $X_\alpha: U_\alpha \rightarrow V_\alpha \cap M$ tal que X_α es diferenciable, homeomorfismo y $\forall q_\alpha \in U_\alpha, (dX_\alpha)_{q_\alpha}$ es inyectiva. En consecuencia la familia de parametrizaciones $X_\alpha: U_\alpha \subset \mathbb{R}^n \rightarrow M$ son inyectivas.

La condición 1 es fácil de verificarse, la condición 2 parte de la definición de variedad diferenciable donde se debe considerar un par de parametrizaciones X_α y X_β con $X_\alpha(U_\alpha) \cap X_\beta(U_\beta) = W \neq \emptyset$. Por lo tanto, obtenemos

$$X_\beta^{-1} \circ X_\alpha: X_\alpha^{-1}(W) \rightarrow X_\beta^{-1}(W)$$

Esta expresión es diferenciable, además, $X_\alpha^{-1}(W)$ son abiertos, pues W es abierto en \mathbb{R}^n donde las intersecciones de los abiertos $X_\alpha(U_\alpha) \cap X_\beta(U_\beta)$ es un abierto y continua por X_α .

Observación 3. Toda superficie de dimensión m es de clase C^k , $M^m \subset \mathbb{R}^n$, es una variedad diferenciable de dimensión m de clase C^k con el atlas A formado por los sistemas de coordenadas $X: U \rightarrow \mathbb{R}^m$, inversas de las parametrizaciones $\varphi: U_0 \subset \mathbb{R}^m \rightarrow U \subset M$, de clase C^k . En la realidad A es un atlas máximo de clase C^k . De hecho sea $z: W \rightarrow z(W) \subset \mathbb{R}^m$ un sistema de coordenadas admisibles en relación con A . Entonces la siguiente relación $\psi = z^{-1}: z(W) \subset \mathbb{R}^m \rightarrow W \subset M$ es un homeomorfismo para cada $p \in W$ existe una parametrización $\varphi: U_0 \rightarrow U, p \in U \subset M$ de clase C^k como z es admisible, $\varphi_0^{-1}\psi: z(U \cap W) \rightarrow \varphi^{-1}(U \cap W)$ es un difeomorfismo de clase C^k . Por lo tanto la expresión, $\varphi \circ (\varphi_0^{-1} \circ \psi): z(U \cap W) \rightarrow U \cap W$ es una parametrización de clase C^k de un entorno de p . Como $p \in W$ y es arbitrario, seguido que $\Psi = z^{-1}(W) \rightarrow W$ es una parametrización de clase C^k donde $z \in A$. Entonces A es máximo.

Definición 4. (*Variedad de Hausdorff de base numerable*) Una variedad diferenciable M , es llamado de Hausdorff si, M con la topología dada, es un espacio Hausdorff.

Es de Hausdorff si sea $x, y \in M$, entonces las vecindades V de x y U de y para la intersección es no vacía. Además, las variedades diferenciables M tiene base numerable si ella puede ser cubierta por una cantidad numerable de entornos, esta dado si existe una sucesión $\{(X_n, U_n)\}_{n \in \mathbb{N}}$ de entornos tales que

$$M = \bigcup_{n \in \mathbb{N}} X_n(U_n)$$

Es este capítulo tomaremos a las variedades diferenciables M es de Hausdorff y de base numerable. En resumen, Para decir que (M, A) es una variedad diferenciable de dimensión m de clase C^k diremos que cumple las siguientes condiciones:

1. M es un espacio topológico de Hausdorff con base numerable.

2. A es una colección de homeomorfismos $X:U \rightarrow \mathbb{R}^m$, de conjuntos abiertos $U \subset M$ sobre abiertos $X(U) \subset \mathbb{R}^m$.
3. Los dominios U de los homeomorfismos $x \in A$ recubren todo M .
4. Dados $X:U \rightarrow \mathbb{R}^m$ e $Y:V \rightarrow \mathbb{R}^m$ pertenecen a A con $U \cap V \neq \emptyset$ entonces $\varphi_{XY} = X(U \cap V) \rightarrow Y(U \cap V)$ es un homeomorfismo de clase C^k .
5. Dado un homeomorfismo $z:W \rightarrow \mathbb{R}^m$ de un abierto $W \subset M$ sobre un abierto $z(W) \subset \mathbb{R}^m$, tal que φ_{ZX} y φ_{XZ} son de clase C^k para cada punto $x \in A \rightarrow z \in A$.

Para todo $r \leq k$, una variedad de clase C^k puede ser tomada pues cualquier atlas de clase C^k contenida en un único atlas máximo es de clase C^k .

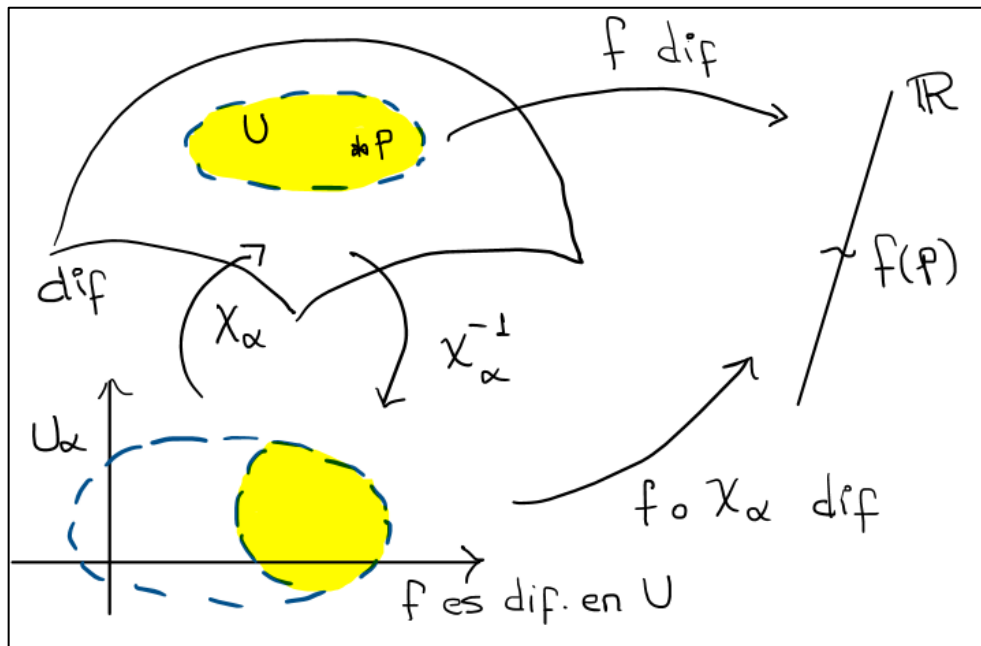
Definición 5. (*Aplicaciones diferenciables entre variedades*) Sea una función $f:U \subset M \rightarrow \mathbb{R}$, donde U es un subconjunto abierto de la variedad diferenciable M . Diremos que f es diferenciable en $p \in U$, si para alguna parametrización $X_\alpha:U_\alpha \subset \mathbb{R}^m \rightarrow M$, con $p \in X_\alpha(U_\alpha) \subset U$, la función compuesta $f \circ X_\alpha:U_\alpha \subset \mathbb{R}^m \rightarrow \mathbb{R}$ es diferenciable en $x_\alpha^{-1}(p)$. Se dice que f es diferenciable en U si, es diferenciable en todo punto del conjunto no vacío U .

Esto es posible debido a la definición de variedad diferenciable, la condición 2, debido a que la diferenciabilidad sobre una variedad diferenciable M no depende de la parametrización, entonces tenemos otra parametrización de la forma $X_\beta:U_\beta \subset \mathbb{R}^n \rightarrow M$ tal que $p \in X_\beta(U_\beta) \subset U$ podemos expresarlo de la siguiente forma:

$$f \circ X_\beta = f \circ X_\alpha \circ X_\alpha^{-1} \circ X_\beta: U_\beta \subset \mathbb{R}^n \rightarrow M$$

Debido a que $f \circ X_\alpha$ es diferenciable por definición y $X_\alpha^{-1} \circ X_\beta$ también lo es por cambio de parámetro, por tanto la función composición $f \circ X_\beta$ es diferenciable.

Figura 4. Aplicaciones diferenciales entre variedades



Fuente: Elaboración propia

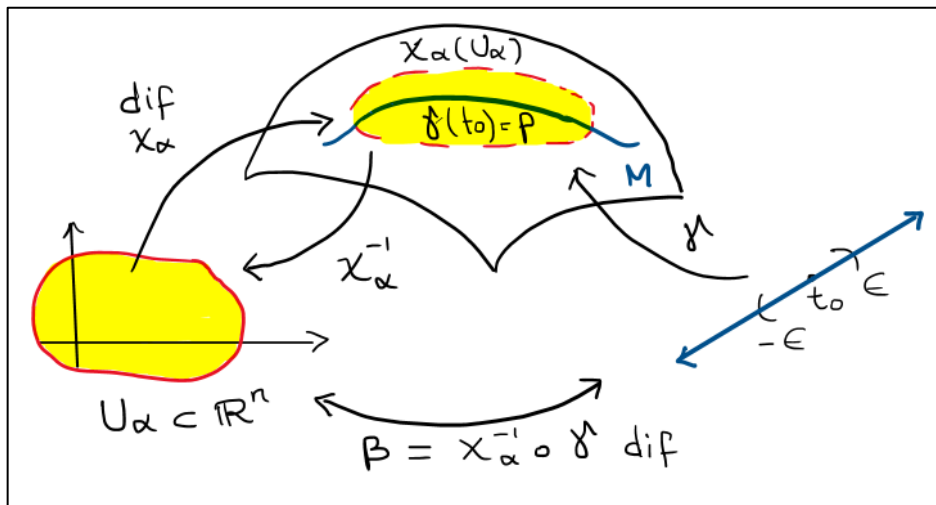
Ejemplo 7. Sea (U, φ) una carta en M y sean (x^1, \dots, x^n) las n funciones coordenadas de la carta; es decir $x^i = r^i \circ \varphi$ donde $i = 1, 2, \dots, n$ considerar que $r^i: \mathbb{R}^n \rightarrow \mathbb{R}$ las proyecciones dadas por $r^i(a_1, \dots, a_n) = a^i$ que son diferenciables.

Entonces cada coordenada $x_i \in F, x \in U$ (F es el conjunto de aplicaciones diferenciables de M). En efecto, basta tener en cuenta la siguiente equivalencia

$$x^i \circ \varphi^{-1} = r_j^i, \quad i = 1, 2, \dots, n$$

Definición 6. Una curva $\gamma: I \rightarrow M$ sobre la variedad, donde $I = (-\epsilon, \epsilon)$ diremos que γ es diferenciable en todo I , si para alguna parametrización $X_\alpha: U_\alpha \subset \mathbb{R}^n \rightarrow M$ con $\gamma(t_0) \in X_\alpha(U_\alpha)$, la función compuesta se genera de la forma $\beta = X_\alpha^{-1} \circ \gamma: I \rightarrow U_\alpha$ es diferenciable en t_0 , donde $\gamma(I) \subset X_\alpha(U_\alpha)$. Si γ es diferenciable en todo $t \in I$, diremos que γ es diferenciable en el conjunto I .

Figura 5. Curva sobre una variedad



Fuente: Elaboración propia

Estas definiciones pueden ser extendidas a teoría de la diferenciabilidad entre variedades.

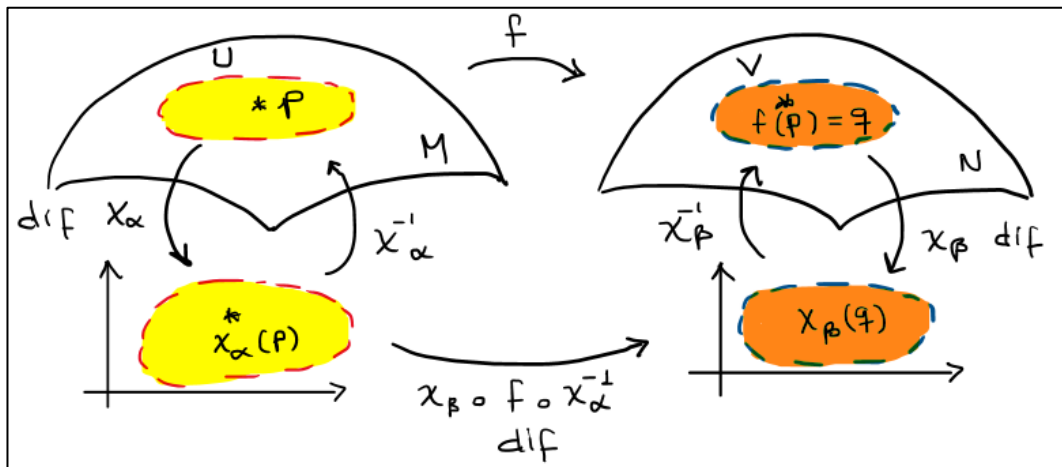
Definición 7. (Difeomorfismo entre variedades diferenciables) Sea M_1 y M_2 variedades diferenciables, se dice que $f: M \rightarrow N$ es diferenciable en el punto $p \in M$ si existe una parametrización $X_\alpha: U \rightarrow \mathbb{R}^m$ en M , $X_\beta: V \rightarrow \mathbb{R}^n$ en N , donde $p \in U$ y $f(U) \subset V$ tal que se tiene que la expresión $X_\beta \circ f \circ X_\alpha^{-1}: X_\alpha(U) \rightarrow X_\beta(V) \subset \mathbb{R}^n$ es diferenciable en el punto $X(p)$.

Debido a que la diferenciabilidad no depende de la parametrización, podemos tomar $x'_\alpha: U' \rightarrow \mathbb{R}^m$ en M , $x'_\beta: V' \rightarrow \mathbb{R}^n$ en N con $p \in U'$, $f(U') \subset V'$, la aplicación $x'_\beta \circ f \circ (x'_\alpha)^{-1}$ es diferenciable en el punto $x'_\alpha(p)$.

Podemos decir que $f: M \rightarrow N$ es diferenciable si f fuera diferenciable en todos los puntos de M .

Además podemos decir que si $f: M \rightarrow N$ es difeomorfismo si f es inyectiva y f^{-1} es diferenciable. f es difeomorfismo local en $p \in M$, si existen vecindades U de p y V de $f(p)$ tal que la función $f: U \rightarrow V$ es un difeomorfismo.

Figura 6. Difeomorfismo entre variedades



Fuente: Elaboración propia

Definición 8. (Vector tangente en un punto de una derivada diferenciable)

Sea M una variedad diferenciable. Consideremos una curva diferenciable

$$\gamma: (-\varepsilon, \varepsilon) \rightarrow M$$

Donde $\gamma(0) = p$ y sea $D_p = \{f: M \rightarrow \mathbb{R}\}$ donde f es diferenciable en p , se define un vector tangente a la curva γ en $t = 0$ como la función

$$\gamma'(0)f \equiv \gamma'(0)f = \frac{d(f \circ \gamma)}{dt}(t)|_t, \quad f \in D_p$$

Un vector tangente en el punto p cuando $t = 0$ de alguna curva de la forma $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$ con $\gamma(0) = p$. Si M es una superficie regular, donde $M \subset \mathbb{R}^n$, se define el vector tangente $\gamma'(0)$.

Definición 9. (Espacio tangente a una variedad diferenciable) El espacio tangente a una variedad M en un punto p es representado por $T_p M$, representa al conjunto de todos los vectores tangente a M sobre el punto p . De esta forma, se define $T_p M = \{v \in \mathbb{R}^m\}$ donde v es el vector tangente en el punto.

Observación 4. Si se tiene una parametrización $X: U \subset \mathbb{R}^n \rightarrow M$ con el punto $p = X(0)$ y el punto $q \in U$, por lo tanto, podemos hacer una restricción sobre $f \in D_p$ y la curva $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$, donde la parametrización es

$$f \circ X(q) = f(g) = f(q_1, \dots, q_n)$$

Además por la inversa lo podemos escribir

$$X^{-1} \circ \gamma(t) = (q_1(t), \dots, q_n(t))$$

Aplicando la definición se tiene

$$\gamma'(0)f = \frac{d(f \circ \gamma)}{dt}(t)|_{t=0} = \frac{d}{dt}, f(q_1(t), q_2(t), \dots, q_n(t))|_t = 0$$

Por lo tanto

$$\gamma'(0)f = \sum_{i=1}^n q'_i(0) \cdot \left(\frac{\partial f}{\partial q_i}(p) \right) = \sum_{i=1}^n q'_i(0) \cdot \left(\left(\frac{\partial}{\partial q_i} \right)_0 \right) f$$

Obteniendo así

$$\gamma'(0) = \sum_{i=1}^n q'_i(0) \cdot \left(\left(\frac{\partial}{\partial q_i} \right)_0 \right) \quad (I)$$

Esta expresión representa la curva γ sobre el punto p con respecto a la parametrización X .

Observación 5. Para la curva en U , $\beta_i(x_i) = (0_1, \dots, x_i, \dots, 0)$, se obtiene que la función que forma la composición de la parametrización $X \circ \beta_i = \gamma_i$ será una curva de M y de la ecuación (I), Específicamente en el caso si vemos $\gamma'(0) = \left(\frac{\partial}{\partial q_i} \right)_0$ donde el vector que sigue la curva γ es $\left(\frac{\partial}{\partial q_i} \right)_0$.

Por otro lado, si escogemos n vectores de la forma $\left(\frac{\partial}{\partial x_i} \right)_p$ donde $i = 1, \dots, n$, estos vectores pueden generar el espacio de los vectores tangentes $T_p M$.

Proposición 6. El espacio tangente de una variedad diferenciable, que es un subconjunto abierto en \mathbb{R}^n es el mismo \mathbb{R}^n

En consecuencia de la proposición se pueden obtener los resultados

a) $T_p \mathbb{R}_{++}^n = \mathbb{R}^n, T_p \mathbb{R}^n = \mathbb{R}^n$

b) Si $M = \{(\lambda, s) \in \mathbb{R}^m \times \mathbb{R}^n : s > 0\}$, entonces $T_p M = \mathbb{R}^{m+n}$

Ejemplo 8. Se define una función $g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, esta función cumple que su jacobiano es de la forma $J_g(x) = dg_x$ con un rango de m . Consideremos la variedad $M = g^{-1}(0) = \{x \in U, g(x) = 0\}$, obteniendo que el espacio de vectores tangentes es el núcleo del jacobiano

$$T_p(g^{-1}(0)) = \ker(J_g(x))$$

Definición 10. (Métrica Riemanniana) Sea S una variedad diferenciable. Una métrica Riemanniana es una aplicación que asocia a cada $p \in M$ un funcional $\langle \cdot, \cdot \rangle_p$

$$\langle \cdot, \cdot \rangle_p: T_p M \times T_p M \rightarrow \mathbb{R}$$

Cumpliendo las siguientes condiciones:

- 1) $\langle \cdot, \cdot \rangle_p$ es un producto interno (Definida positiva, simétrico y bilineal) para cada $p \in M$.
- 2) $\langle \cdot, \cdot \rangle_p$ varía diferenciablemente en el sentido siguiente: Si la aplicación $X: U \subset \mathbb{R}^n \rightarrow M$ es un sistema de coordenadas en torno al punto p , con $X(x_1, \dots, x_n) = q \in X(U)$ y $\frac{\partial}{\partial x_i}(q) = dX_q(0, \dots, 1, \dots, 0)$ entonces la función $g_{ij}: U \rightarrow \mathbb{R}$ definido por

$$g_{ij}(x_1, \dots, x_n) = \left\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \right\rangle_q$$

Es diferenciable

Estas funciones son llamadas métricas Riemannianas dentro del sistema de X y la matriz G usando la métrica Riemanniana, además como se tiene

que X es un difeomorfismo y dX_q es un isomorfismo, por tanto la matriz G es invertible, generando así la matriz relacionada con la métrica Riemanniana tenga una representación invertible.

Definición 11. (*Variedad Riemanniana*) Una variedad diferenciable para la cual se define una métrica Riemanniana se denomina una variedad Riemanniana.

Ejemplo 9. Sea $M = \mathbb{R}^n$, con la parametrización $X: \mathbb{R}^n \rightarrow \mathbb{R}^n$ y se define la métrica $\langle \cdot, \cdot \rangle_p: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, donde $\langle x, y \rangle_p = x^T y$. Sea el punto $q \in \mathbb{R}^n$

$$\frac{\partial}{\partial x_i}(q) = dX_q e_i = e_i$$

De esta forma la matriz $g_{ij}: U \rightarrow \mathbb{R}$ se puede definir como

$$g_{ij}(x) = \left\langle \frac{\partial}{\partial x_i}(x), \frac{\partial}{\partial x_j}(x) \right\rangle_x = \langle e_i, e_j \rangle_x = e_i^T e_j = \delta_{ij}$$

Son diferenciables en \mathbb{R}^n , además si $M = re^n$, con la métrica G sería una variedad Riemanniana.

Ejemplo 10. Sea M la variedad de la siguiente forma

$$M = \{(x_1, x_2) \in \mathbb{R}^2, x_2 > 0\}$$

Donde el producto interno sea

$$\langle (x_1, x_2), (y_1, y_2) \rangle_p = \frac{1}{x_2}(x_1 y_1 + x_2 y_2)$$

Sea el punto $q \in M$ donde sus derivadas parciales sean los vectores unitarios e_i . Además definimos la matriz

$$g_{ij}(x_1, x_2) = \frac{1}{x_2} \langle e_i, e_j \rangle = \frac{1}{x_2} \delta_{ij}$$

Es diferenciable y por lo tanto M es una variedad Riemanniana con la métrica $G(p) = \frac{1}{p_2} (Id)_{2 \times 2}$ la cual se conoce como métrica de Lobatchesky o Poincaré.

Ejemplo 11. Sea la variedad Riemanniana $(\mathbb{R}^2, G(x))$ con la métrica Riemanniana $\langle u, v \rangle_p = u^T G(p)v$ se llamará métrica Riemanniana de Udriste, donde se sabe que la matriz $G(p)$ tiene la forma:

$$G(p) = \begin{bmatrix} 4p_1^2 + 1 & -2p_1 \\ -2p_1 & 1 \end{bmatrix}$$

Definición 12. (*Campo de vectores en una variedad diferenciable*) Un campo de vectores X en una variedad diferenciable M es una correspondencia que a cada punto $p \in M$ asocia un vector $X(p) \in T_p M$.

Se considera la parametrización X , donde es posible escribirlo así

$$X(p) = \sum_{i=1}^n a_i(p) \left(\frac{\partial}{\partial x_i} \right) p$$

Para cada a_i es una función definida en M y $\left(\frac{\partial}{\partial x_i} \right) p$ es una base asociada a la parametrización. Esta será diferenciable si las funciones a_i lo son para alguna parametrización. Considerando los campos vectoriales $\varphi: D \rightarrow F$ donde

$$(\varphi f)(P) = \sum_i a_i(p) \frac{\partial \varphi}{\partial x_i}(p)$$

Donde el conjunto D es el conjunto de funciones diferenciables en M y F , estudiando las trayectorias de M se considera los campos restringidos a una curva.

Definición 13. (*Campo de vectores a lo largo de la curva*) Un campo vectorial V a lo largo de una curva $\alpha: I \rightarrow M$ es una aplicación que a cada $\alpha(t) \in M$ asocia un vector tangente $V(t) \in T_{\alpha(t)} M$. Se dice que V es diferenciable si para cada función diferenciable f en D , la función $V(t)f$ es una función diferenciable en I . Sea X un campo definido en M , el campo X sobre α será denotado $V(t) = X(\alpha(t))$ y diremos que es inducido por X .

Definición 14. (Una conexión afín es una aplicación) Tomando la relación $\nabla: H \times H \rightarrow H$ donde cada par de campos (X, Y) se asocia a otro campo de la forma $\nabla_X Y$ tal que para todo $X, Y, Z \in H$ y $f, g \in D$ se verifica:

$$1) \nabla_{(fX+gY)}Z = f\nabla_X Z + g\nabla_Y Z$$

$$2) \nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$$

$$3) \nabla_X fY = f\nabla_X Y + X(f)Y, \text{ donde } X(f) = \sum_{i=1}^n a_i(\cdot) \frac{\partial f(\cdot)}{\partial x_i}$$

Considerando la curva α en M es diferenciable y se denota el conjunto de campos vectoriales a lo largo de la curva como H_α .

Proposición 7. Sea M una variedad diferenciable con una conexión afín ∇ . Entonces existe una única aplicación donde a cada $V \in H_\alpha$ se asocia otro campo en H_α , denotado por $\frac{DV}{dt}$, tal que para todo $V, W \in H_\alpha$ y la función $f: I \rightarrow \mathbb{R}$ una aplicación diferenciable se cumple:

$$1) \frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$$

$$2) \frac{D}{dt}(fV) = \frac{df}{dt}V + f \frac{DV}{dt}$$

$$3) \text{ Si } V(t) = Y(\alpha(t)), \text{ donde } Y \in H, \text{ entonces } \frac{DV}{dt} = \nabla_{\frac{d\alpha}{dt}} Y.$$

Donde el símbolo de la derivada covariante es $\frac{D}{dt}$

Observación 6. La anterior proposición muestra que cada derivada covariante es producto de la elección de una conexión afín de M , además al realizar la demostración de la proposición se encuentra una característica de la derivada covariante sobre la parametrización

$$\frac{Dv}{dt} = \sum_{j=1}^n \frac{dv^j}{dt} X_j + \sum_{i,i=1}^n v^j \frac{dx_i}{dt} \nabla_{X_i} X_j$$

Además, se debe tener en cuenta que la conexión afín ofrece una forma de derivar sobre el campo de vectores a partir de curvas, particularmente

en el campo $V = \frac{d\alpha}{dt}$ la cual llamaremos la aceleración de una curva en M y se denotará así $\frac{D}{dt}\left(\frac{d\alpha}{dt}\right)$.

Observación 7. Podemos analizar la conexión afín relativa a coordenadas locales donde se puede suponer los campos $X, Y \in H$ son representados en una cierta vecindad, de modo que algún punto p , por simplicidad de notación usaremos $\frac{\partial}{\partial x_i} = X_i$, por lo tanto se tiene que

$$X = \sum_{i=1}^n x_i \left(\frac{\partial}{\partial x_i} \right) = \sum_{i=1}^n x_i X_i \quad , \quad Y = \sum_{i=1}^n y_i \left(\frac{\partial}{\partial x_i} \right) = \sum_{i=1}^n y_i X_i$$

Por las propiedades de la definición de conexión afín

$$\nabla_X Y = \nabla_{\sum x_i X_i} \left[\sum_j y_j X_j \right] = \sum_i x_i \left[\sum_j y_j \nabla_{X_i} X_j \right] + \sum_i x_i \left[\sum_j \left(\frac{\partial y_j}{\partial x_i} X_j \right) \right]$$

Donde $\nabla_{X_i} X_j \in H$ y es representado también como una base local, es decir

$$\nabla_{X_i} X_j = \sum_{k=1}^n \Gamma_{i,j}^k X_k \quad (II)$$

Obteniendo así el resultado siguiente

$$\nabla_X Y = \sum_{k=1}^n \left[\sum_{i,j} x_i y_j \Gamma_{i,j}^k + \sum_{i=1}^n x_i \frac{\partial y_k}{\partial x_i} \right] X_k$$

Definición 15. (*Símbolos de Christoffel*) Los símbolos de Christoffel, o coeficientes de la conexión afín ∇ en U , son las funciones diferenciables definidas anteriormente en la ecuación (II), la cual podemos definir las de la forma $\Gamma_{i,j}^k: U \subset M \rightarrow \mathbb{R}$.

Podemos expresar la derivada covariante en términos de los símbolos de Christoffel, para ello se define un sistema $X: U \rightarrow M$ y un punto $p \in M$.

De la observación 6 podemos reemplazar el resultado de la ecuación (II), obteniendo así

$$\frac{Dv}{dt} = \sum_{j=1}^n \frac{dv^j}{dt} X_j + \sum_{i,i=1}^n v^j \frac{dx_i}{dt} \left(\sum_{k=1}^n \Gamma_{i,j}^k X_k \right)$$

Donde por operaciones con sumatorias podemos presentarla también

$$\frac{Dv}{dt} = \sum_{k=1}^n \left(\frac{dv^k}{dt} + \sum_{i,j=1}^n v^j \frac{dx_i}{dt} \Gamma_{i,j}^k \right) X_k \quad (III)$$

De esta forma, presentamos la derivada covariante en términos de los símbolos de Christoffel y términos locales.

Definición 16. Una curva parametrizada $\alpha: I \rightarrow M$ es una geodésica si el campo tangente $\frac{d\alpha}{dt}$ verifica que

$$\frac{D}{dt} \left(\frac{d\alpha}{dt} \right) = 0$$

Por otro lado, por la definición podemos sobre campos paralelos, lo cual dado una variedad diferenciable M , una conexión afín ∇ y un campo V se dice que es paralelo si $\frac{dV}{dt} = 0$, para todo $t \in I$. Por ende, si α es geodésica entonces $\frac{d\alpha}{dt}$ es paralelo.

Podemos considerar la expresión anterior (III), un campo paralelo V donde se cumple que

$$\sum_{k=1}^n \left(\frac{dv^k}{dt} + \sum_{i,j=1}^n v^j \frac{d\alpha_i}{dt} \Gamma_{i,j}^k \right) X_k = 0$$

Lo cual igualando a cero la parte interior de la sumatoria y como es una geodésica, se tiene que $v^i = \frac{d\alpha_i}{dt}$, transformándose así

$$\frac{d}{dt} \left(\frac{d\alpha_k}{dt} \right) + \sum_{i,j=1}^n \frac{d\alpha_j}{dt} \frac{d\alpha_i}{dt} \Gamma_{i,j}^k = 0 \quad (IV)$$

Donde $k = 1, 2, \dots, n$ y el cual genera un sistema de n ecuaciones diferenciales de segundo orden, las cuales poseen solución única, verificando que $x(0) = \alpha(0) = p$ y $\frac{dx}{dt}(0) = \alpha'(0) = v$.

Definición 17. (Conexión afín en variedades Riemannianas) Sea M una variedad diferenciable con una conexión afín ∇ y una métrica Riemanniana $\langle \cdot, \cdot \rangle$. Se dice que ∇ es compatible con la métrica $\langle \cdot, \cdot \rangle$ sí para todo par de campos de vectores V y W a lo largo de la curva diferenciable $\alpha: I \rightarrow M$ se tiene:

$$\frac{d}{dt} \langle V, W \rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle \quad (V)$$

Proposición 8. Si la conexión afín ∇ es compatible con $\langle \cdot, \cdot \rangle$ y V, W son campos paralelos a lo largo de una curva diferenciable $\alpha: I \rightarrow M$ entonces, $\langle V, W \rangle$ es constante. En el caso de las geodésicas, para un $\alpha(t)$ entonces el producto interno $\left\langle \frac{d\alpha}{dt}, \frac{d\alpha}{dt} \right\rangle$ es constante.

Proposición 9. Sea M una variedad Riemanniana. Una conexión afín ∇ es compatible con $\langle \cdot, \cdot \rangle$ sí y solo sí, $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle$, para todo $X, Y, Z \in H$.

Definición 18. Una conexión afín ∇ es una variedad diferenciable M es llamada simétrico si

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

Donde $[X, Y] = XY - YX$

Observación 7. Podemos notar por la definición anterior que

1) En un sistema de coordenadas (U, X) la simetría de la conexión implica que

$$\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = \nabla_{\frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i}$$

En efecto, para toda función $f \in D$

$$X_i X_j(f) - X_j X_i(f) = 0$$

2) Como resultado de lo anterior se tiene

$$\nabla_{X_i} X_j - \nabla_{X_j} X_i = \sum_{k=1}^n (\Gamma_{i,j}^k - \Gamma_{j,i}^k) X_k = 0$$

Por la independencia lineal de $\{X_k\}$ se obtiene que $\Gamma_{i,j}^k = \Gamma_{j,i}^k$ donde la condición recíproca es inmediata

Por el siguiente teorema nos garantiza la unicidad y existencia de una conexión simétrica, donde es compatible a la métrica de la variedad Riemanniana.

Teorema 2. (Levi-Civita) Dada una variedad Riemanniana M , existe una única conexión afín ∇ en M satisfaciendo las condiciones:

- a) ∇ es simétrica
- b) ∇ es compatible con la métrica Riemanniana

Esta conexión es denominada conexión Riemanniana

Observación 8. Se puede identificar una relación entre la métrica Riemanniana y los símbolos de Christoffel, donde tomando un sistema

(U, X) , los símbolos de Christoffel $\Gamma_{i,j}^k: U \rightarrow \mathbb{R}$ definen los coeficientes de las

conexiones $\nabla_{X_i} X_j = \sum_{k=1}^n \Gamma_{i,j}^k X_k$, obteniendo así que

$$\Gamma_{i,j}^m = \left\{ \frac{1}{2} \sum_k^n \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij} \right\} g^{km}$$

Donde $g_{ij} = \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle$ son elementos de $G(x)$ y g^{ij} son de su inversa, por

lo tanto haciendo un cambio $\frac{\partial}{\partial x_i} = X_i$ y respectivamente para j y k . Se

obtiene que $\langle Z, \nabla_Y X \rangle$ es igual a

$$\frac{1}{2} \{X\langle Y, Z \rangle + Y\langle Z, X \rangle - Z\langle X, Y \rangle - \langle [X, Z], Y \rangle - \langle [Y, Z], X \rangle - \langle [X, Y], Z \rangle\} \quad (VI)$$

Además, se tiene que

$$\langle X_k, \nabla_{X_j} X_i \rangle = \frac{1}{2} \{X_i \langle X_j, X_k \rangle + X_j \langle X_k, X_i \rangle - X_k \langle X_i, X_j \rangle\}$$

Como $\nabla_{X_i} X_j = \sum_{l=1}^n \Gamma_{i,j}^l X_l$ y por la propiedad lineal del producto interno, se tiene que

$$\sum_{l=1}^n \Gamma_{i,j}^l \langle X_k, X_l \rangle = \frac{1}{2} \{X_i \langle X_j, X_k \rangle + X_j \langle X_k, X_i \rangle - X_k \langle X_i, X_j \rangle\}$$

Obteniendo así

$$\sum_{l=1}^n \Gamma_{i,j}^l g_{kl} = \frac{1}{2} \left\{ \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij} \right\}$$

Denominando así a todo como un $b_k, k = 1, 2, \dots, n$ obteniendo un sistema lineal de la forma $Gy = b$ donde $y = (\Gamma_{ij}^1, \Gamma_{ij}^2, \dots, \Gamma_{ij}^n)$ y $b = (b_1, b_2, \dots, b_n)$,

puesto que es G es invertible por la definición de métrica Riemanniana, entonces $y = G^{-1}b$, así la siguiente expresión

$$\Gamma_{i,j}^m = \frac{1}{2} \left\{ \sum_{k=1}^n \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij} \right\} g^{km} \quad (VII)$$

Ejemplo 12. Sea la variedad Riemanniana $M = \mathbb{R}_{++}^n$ con la métrica

$$G(x) = \text{diag} \left(\frac{1}{(h_1(x_1))^2}, \frac{1}{(h_2(x_2))^2}, \dots, \frac{1}{(h_n(x_n))^2} \right)$$

Para funciones $h_i: \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ diferenciables y la matriz inversa $G(x)$ es

$$G^{-1}(x) = \text{diag} \left((h_1(x_1))^2, (h_2(x_2))^2, \dots, (h_n(x_n))^2 \right)$$

Para obtener los símbolos de Christoffel, recordamos la ecuación (VII) y cuando $k \neq m$ la expresión $g^{mk} = 0$, obteniendo la expresión:

$$\Gamma_{ij}^m = \frac{1}{2} \left\{ \frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_j} g_{mi} - \frac{\partial}{\partial x_m} g_{ij} \right\} g^{mm}$$

El primer caso es considerar si $i = j$

$$\Gamma_{ii}^m = \frac{1}{2} \left\{ \frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_i} g_{mi} - \frac{\partial}{\partial x_m} g_{ii} \right\} g^{mm}$$

Si $m = i$ se tiene que $\Gamma_{ii}^i = -\frac{1}{h_i(x_i)} \frac{\partial}{\partial x_i} (h_i(x_i))$ y si $m \neq i$ entonces $\Gamma_{ii}^m = 0$

El segundo caso es considerar si $i \neq j$

$$\Gamma_{ij}^m = \frac{1}{2} \left\{ \frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_j} g_{mi} \right\} g^{mm}$$

Si $m = i$ se tiene que $m \neq j$ y $\Gamma_{ij}^i = 0$, análogamente para $m = j$, de igual forma para $m \neq i \neq j$ se obtiene que $\Gamma_{ij}^m = 0$.

Analizando ambos casos se obtiene

$$\Gamma_{ij}^m = -\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{im} \delta_{ij} \quad (VIII)$$

Es la relación de la métrica $G(x)$ con los símbolos de Christoffel, en forma particular se tiene

- Si $h_i(x_i) = 1$, entonces, $G(x) = 1$. $\Gamma_{ij}^m = 0, \forall i, j, m = 1, \dots, n$
- Si $h_i(x_i) = x_i$, entonces, $G(x) = X^{-2}$. $\Gamma_{ij}^m = -\frac{1}{x_i} \delta_{im} \delta_{ij}, \forall i, j, m = 1, \dots, n$
- Si $h_i(x_i) = x_i^{\frac{r}{2}}$, entonces, $G(x) = X^{-r}$. $\Gamma_{ij}^m = -\frac{r}{2} \frac{1}{x_i} \delta_{im} \delta_{ij}$

Para la obtención de la derivada covariante podemos tomar la ecuación (VIII) y reemplazarla en (III), obteniendo así el siguiente resultado

$$\frac{DV}{dt} = \sum_{i=1}^n \left(\frac{dv^i}{dt} - \frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} v^i \frac{dx_i}{dt} \right) X_i$$

En forma particular si:

- Si $h_i(x_i) = 1$, entonces, $\frac{DV}{dt} = \sum_{i=1}^n \left(\frac{dv^i}{dt} \right) X_i$ es la derivada usual
- Si $h_i(x_i) = x_i$, entonces, $\frac{DV}{dt} = \sum_{i=1}^n \left(\frac{dv^i}{dt} - \frac{1}{x_i} v^i \frac{dx_i}{dt} \right) X_i$
- Si $h_i(x_i) = x_i^{\frac{r}{2}}$, entonces, $\frac{DV}{dt} = \sum_{i=1}^n \left(\frac{dv^i}{dt} - \frac{r}{2} \frac{1}{x_i} v^i \frac{dx_i}{dt} \right) X_i$

Por último, para obtener la ecuación geodésica se toma $p \in \mathbb{R}_{++}^n$ y el vector $v \in T_p \mathbb{R}_{++}^n \rightarrow \mathbb{R}^n$ con la curva definida

$$\alpha: I \rightarrow \mathbb{R}_{++}^n: \alpha(t) = (\alpha_1(t), \alpha_2(t) \dots, \alpha_n(t))$$

Donde $\alpha(0) = p$ y su derivada es v , sustituyendo los símbolos de Christoffel de la ecuación (VIII) en la ecuación (IV), se obtiene

$$\frac{d^2\alpha_i}{dt^2} - \frac{1}{h_i(\alpha_i)} \frac{\partial(h_i(\alpha_i))}{\partial\alpha_i} \left(\frac{d\alpha_i}{dt}\right)^2 = 0, \forall i = 1, \dots, n \quad (IX)$$

Como $\alpha_i(0) = p_i$ y $\alpha'_i(0) = v_i$, la ecuación diferencial se reduce

$$\frac{d\alpha_i}{dt} = h_i(\alpha_i) a_i$$

Donde a_i es una constante, lo cual es análogo a la expresión

$$\int \frac{1}{h_i(\alpha_i)} d\alpha_i = a_i t + b_i \quad (X)$$

Donde a_i y b_i son constantes en \mathbb{R} para todo $i = 1, \dots, n$. De esta forma obtenemos la única geodésica con métrica $G(p)$ sobre el punto p y en la dirección v . En forma particular si:

- Si $h_i(\alpha_i) = 1$, entonces $G(p) = I$ y considerando la expresión (X)

$$\alpha_i(t) = v_i t + p_i$$

Donde las geodésicas son curvas $\alpha: \mathbb{R} \rightarrow \mathbb{R}_{++}^n$ y ara valores de t tal que $v_i t + p_i > 0$.

- Si $h_i(\alpha_i) = \alpha_i$, entonces $G(p) = X^{-2}$ y considerando la expresión (X) las curvas geodésicas son funciones exponenciales

$$\alpha(t) = \left(p_1 \exp\left(\frac{v_1}{p_1} t\right), p_2 \exp\left(\frac{v_2}{p_2} t\right), \dots, p_n \exp\left(\frac{v_n}{p_n} t\right) \right)$$

Donde $p \in \mathbb{R}_{++}^n$ y $v \in \mathbb{R}^n$, la geodésica cumple para todo $t \in \mathbb{R}$.

Definición 19. (Aplicación de curvatura) Sea $A(H, H)$ el conjunto de aplicaciones de H en H y ∇ la conexión afín es una variedad Riemanniana

M , por el teorema de Levi-Civita. La curvatura K de variedad Riemanniana M es una relación de la forma $K: H \times H \rightarrow A(H, H)$ definido por

$$K(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X, Y]} Z$$

Observación 9. Si la variedad $M = \mathbb{R}^n$, entonces se tiene $K(X, Y)Z = 0$, para todo $X, Y, Z \in H$. Esto se cumple tomando $Z = (z_1, \dots, z_n) \in \mathbb{R}^n$ y la conexión afín definida como $\nabla_X Z = (X_{z_1}, X_{z_2}, \dots, X_{z_n})$.

Además si consideramos un sistema en (U, X) centrado en el punto p , donde $\{X_i\}$ es una base del espacio tangente $T_p M$, de esta forma se obtiene

$$K(X_i, X_j)X_k = (\nabla_{X_i} \nabla_{X_j} - \nabla_{X_j} \nabla_{X_i}) X_k$$

Observación 10. La curvatura K es antisimétrica, esto se puede demostrar ya que para todo $Z \in H$ se cumple

$$K(X, Y)Z + K(Y, X)Z = \nabla_{[X, Y]} Z + \nabla_{[Y, X]} Z$$

Además se cumple que $[X, Y] = -[Y, X]$, de esta forma para todo $Z \in H$ se cumple:

$$K(X, Y)Z + K(Y, X)Z = 0$$

Demostrando así la antisimetría ya que $K(X, Y) = -K(Y, X)$

Proposición 10. La curvatura K de una variedad Riemanniana es trilineal ya que cumple dos condiciones

a) K es bilineal en $H \times H$, dado que $f, g \in D(M)$ y $X_1, X_2, Y_1, Y_2 \in H$

$$K(fX_1 + gX_2, Y_1) = fK(X_1, Y_1) + gK(X_2, Y_1)$$

$$K(X_1, fY_1 + gY_2) = fK(X_1, Y_1) + gK(X_1, Y_2)$$

b) Para todo $X, Y \in H$, el operador de curvatura $K(X, Y): H \rightarrow H$ es lineal ya que para todo $f \in D(M)$ y $Z, W \in H$.

$$K(X, Y)(Z + W) = K(X, Y)Z + K(X, Y)W$$

$$K(X, Y)(fZ) = fK(X, Y)Z$$

Proposición 11. Sea (U, X) un sistema de coordenadas en torno a $p \in M$ y $\{X_i\}$ una base de $T_p M$ en este sistema de coordenadas. Entonces:

$$K(X_i, X_j)X_k = \sum_{l=1}^n K_{ijk}^l X_l$$

Donde las componentes K_{ijk}^l está definido como

$$K_{ijk}^l = X_j \Gamma_{ik}^l - X_i \Gamma_{jk}^l + \sum_{s=1}^n \Gamma_{ik}^s \Gamma_{js}^l - \sum_{s=1}^n \Gamma_{jk}^s \Gamma_{is}^l$$

Observación 11. Si en el sistema de coordenadas (U, X) denominamos

$$X = \sum_{i=1}^n u^i X_i, Y = \sum_{j=1}^n v^j X_j, Z = \sum_{k=1}^n w^k X_k$$

Por la propiedad de linealidad de curvatura se puede expresar

$$K(X, Y)Z = \sum_{i,j,k,l=1}^n K_{ijk}^l u^i v^j w^k X_l$$

Ejemplo 13. Sea $M = \mathbb{R}_{++}^n$ y $M = C_0^n$ con estructuras de una variedad Riemanniana $G(x) = \text{diag} \left(\frac{1}{(h_1(x_1))^2}, \frac{1}{(h_2(x_2))^2}, \dots, \frac{1}{(h_n(x_n))^2} \right)$ y con los símbolos de Christoffel definidos, además por la trinealidad de K , en consecuencia a todo lo mencionado tenemos:

$$K(X_i, X_j)X_k = \nabla_{X_j}(\nabla_{X_i}X_k) - \nabla_{X_i}(\nabla_{X_j}X_k) + \nabla_{[X_i, X_j]}X_k$$

Como la conexión cumple el teorema de Levi-Civita, se tiene $[X_i, X_j] = 0$

$$K(X_i, X_j)X_k = \nabla_{X_j}(\nabla_{X_i}X_k) - \nabla_{X_i}(\nabla_{X_j}X_k)$$

Si $i = j$, entonces $K(X_i, X_j)X_k = 0$

Si $i \neq j$, entonces $\nabla_{X_i}X_k = \sum_{j=1}^n \Gamma_{ik}^j X_j$, sustituyendo los símbolos se tiene

$$\nabla_{X_i}X_k = \sum_{j=1}^n \left(-\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{ij} \delta_{ik} \right) X_j = -\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{ik} X_i \quad (XI)$$

Luego tomando la conexión ∇_{X_j} se obtiene

$$\nabla_{X_j}(\nabla_{X_i}X_k) = \nabla_{X_j} \left(-\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{ik} X_i \right)$$

Por la linealidad y definición de la conexión se tiene

$$\nabla_{X_j}(\nabla_{X_i}X_k) = -\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{ik} \nabla_{X_j}X_i + X_j \left(-\frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} \delta_{ik} \right) X_i$$

Usando la ecuación (XI) y como $i \neq j$ toda la expresión se vuelve cero, es decir $K(X, Y)Z = 0$. Luego, por las variedades Riemannianas la curvatura es cero. Ejemplo, la métrica $I, X^{-r} \in \mathbb{R}^n$ y $\text{cosec}^4(\pi x), X^{-r}(I - X)^{-r} \in C_0^n$ son variedades con cero de curvatura.

Proposición 12. (*Curvatura seccional*) Sea $\sigma \subset T_p M$ un subespacio tridimensional, donde $x, y \in \sigma$, se cumple

$$K(x, y) = \frac{\langle K(x, y)x, y \rangle}{Q(x, y)}$$

Donde los vectores son linealmente independientes, y se puede observar que la curvatura no depende de la elección de los vectores.

Por otro lado, podemos definir la gradiente y Hessiana de la variedad, ya que si se tiene una variedad Riemanniana M y una función diferenciable $f: M \rightarrow \mathbb{R}$, donde $p \in M$, por el teorema de Riesz se puede afirmar que existe un único elemento $\nabla_M f(p) \in T_p M$ donde $\forall v \in T_p M$ se cumple

$$df_p(v) = \langle \nabla_M f(p), v \rangle \quad (XII)$$

Además sus módulos son igual, ya que la diferencial se caracteriza por tener el producto interno, definiendo así un campo vectorial $\text{grad } f: M \rightarrow TM$ cumpliéndose que $\text{grad } f(p) = \nabla_M f(p)$, definiendo la expresión (XII) para todo $X \in H$ de la forma

$$df_p(X(p)) = \langle \text{grad } f(p), X(p) \rangle$$

Se puede definir la aplicación $df: H \rightarrow M^* = \mathcal{L}(M, \mathbb{R})$, donde el conjunto de funciones en M esta denotado por $\mathcal{L}(M, \mathbb{R})$, tal que $df(X) = \langle \text{grad } f, X \rangle$, además $f_p(X(p))$ es una composición de la forma $f \circ \gamma$ para alguna curva con $\gamma(0) = p$ y $\gamma'(0) = X(p)$, obteniendo así $df_p(X(p)) = X(f)(p)$, así

$$df(X) = \langle \text{grad } f, X \rangle = X(f)$$

Definición 20. La gradiente de una función diferenciable $f: M \rightarrow \mathbb{R}$ es un campo vectorial $\text{grad } f: M \rightarrow TM$ métricamente equivalente a la diferencial, esto es,

$$df_p(X(p)) = \langle \text{grad } f(p), X(p) \rangle X(p)(f), \quad \forall X \in H$$

Observación 12. Sea $M \subset \mathbb{R}^n$ una variedad Riemanniana con la métrica $\langle v, w \rangle_x = v^T G(x) w$ donde la matriz $G(x)$ es simétrica y definida positiva, donde el campo gradiente tiene la forma

$$\text{grad } f(q) = G^{-1}(q) f'(q)$$

Donde la matriz inversa de G es $(g^{ij}(q))$ y f' es la colección de las derivadas parciales de $f \circ X$, por lo tanto

$$df_q(v) = f'(q)^T v = (G^{-1}(q) f'(q))^T G(q) v = \langle G(q)^{-1} f'(q), v \rangle_q$$

Ejemplo 14. Sea la variedad Riemanniana \mathbb{R}^n con la métrica euclidiana I , entonces la gradiente será la usual, de la forma $\text{grad } f(x) = f'(x)$.

Ejemplo 15. Sea la variedad Riemanniana \mathbb{R}_{++}^n con la métrica $G(x)$ para funciones $h_i: \mathbb{R}_{++}^n \rightarrow \mathbb{R}_{++}^n$

$$G(x) = \text{diag} \left(\frac{1}{(h_1(x_1))^2}, \frac{1}{(h_2(x_2))^2}, \dots, \frac{1}{(h_n(x_n))^2} \right)$$

Donde la gradiente será de la forma

$$\text{grad } f(x) = \text{diag} \left[(h_1(x_1))^2, (h_2(x_2))^2, \dots, (h_n(x_n))^2 \right] f'(x)$$

En forma particular si:

- Si $h_i(x_i) = x_i$, entonces, $\text{grad } f(x) = X^2 f'(x)$
- Si $h_i(x_i) = x_i^{\frac{r}{2}}$, entonces, $\text{grad } f(x) = X^r f'(x)$

Ejemplo 16. Si la variedad Riemanniana C_0^n con la métrica $G(x)$ definida como $\text{csc}^4(\pi x) = \text{diag}(\text{csc}^4(x_1), \text{csc}^4(x_2), \dots, \text{csc}^4(x_n))$ entonces se obtiene la gradiente como $\text{grad } f(x) = \text{Sen}^4(\pi x) f'(x)$.

Definición 21. Sea M una variedad Riemanniana y un punto $p \in M$. Se dice que p es punto crítico si $\text{grad } f(p) = 0$.

Definición 22. (*Hessiana de una función*) Sea $f: M \rightarrow \mathbb{R}$ una función de clase C^k donde $k \geq 2$. La Hessiana de la función, denominada por H^f , esta definida como la covariante del campo gradiente, es decir

$$H^f = \frac{D}{dt}(\text{grad } f)$$

De esta forma, la Hessiana en el punto sobre la dirección de $v \in T_p M$ es:

$$H_p^f(v) = \frac{D}{dt}(\text{grad } f)(p) = \nabla_v \text{grad } f(p)$$

Desde este concepto, se puede definir aplicaciones $H_p^f(v): T_p M \rightarrow T_p M$ y de $H^f: M \rightarrow \mathcal{L}(TM, TM)$ donde el conjunto de aplicaciones lineales de TM en TM es el espacio $\mathcal{L}(TM, TM)$ y $H^f(p) = H_p^f \in \mathcal{L}(T_p M, T_p M)$

Proposición 13. Para cada $p \in M$, el operador $H_p^f: T_p M \rightarrow T_p M$ es lineal y autoadjunto, esto significa que $\langle H_p^f(v), w \rangle_p = \langle v, H_p^f(w) \rangle_p$.

Se puede introducir una forma cuadrática donde para cada punto $p \in M$ se tiene $q_p^f: T_p M \times T_p M \rightarrow \mathbb{R}$, definida por

$$q_p^f(v, w) = \langle H_p^f v, w \rangle_p$$

En forma general, se puede definir $q^f: H \times H \rightarrow \mathcal{L}(M, \mathbb{R})$ dado por

$$q^f(X, Y) = \langle \nabla_X \text{grad } f, Y \rangle \tag{XIII}$$

Esta expresión obtenida en la ecuación (XIII) tiene un punto en contra que depende de la métrica y de la conexión afín, sin embargo, debido al teorema de Levi Civita, convierte a la siguiente proposición en una adecuada caracterización para la obtención de una métrica adecuada y de una conexión afín.

Proposición 14. Para todo $X, Y \in H$ se cumple que

$$q^f(X, Y) = (XY - \nabla_X Y)f = (YX - \nabla_Y X)f$$

Observación 13. En un sistema (X, U) , donde los términos de su base $\{X_k\}$ están definidos de la forma:

$$q^f(X_i, X_j) = \left(X_i X_j - \sum_{m=1}^n \Gamma_{ij}^m X_m \right) f$$

Obteniendo por las proposiciones anteriores el resultado

$$q^f(X_i, X_j) = \langle H_p^f X_i, X_j \rangle = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} - \sum_{m=1}^n \Gamma_{ij}^m \frac{\partial f}{\partial x_m} \right) \quad (XIV)$$

Ejemplo 17. Sea la variedad Riemanniana \mathbb{R}^n con la métrica I , los símbolos de Christoffel son $\Gamma_{ij}^m = 0, \forall i, j, m = 1, \dots, n$, por lo tanto la matriz Hessiana usual es de la forma $H_p^f(p) = f''(p)$.

Ejemplo 18. Sea la variedad Riemanniana \mathbb{R}_{++}^n con $g_{ij} = \frac{\delta_{ij}}{h_i(x_i)h_j(x_j)}$ como la métrica Riemanniana, además por lo estudiado con anterioridad los símbolos de Christoffel son de la forma

$$\Gamma_{ij}^m = \frac{-1}{h_i(x_i)} \frac{\partial h_i(x_i)}{\partial x_i} \delta_{im} \delta_{ij}$$

Entonces por la observación anterior y la expresión (XIV) se tiene que

$$q^f(X_i, X_j) = X_i X_j + \sum_{m=1}^n \frac{1}{h_i(x_i)} \frac{\partial h_i(x_i)}{\partial x_i} \delta_{im} \delta_{ij} X_m$$

En el caso que $m \neq j$, se cumple que $\delta_{im} \delta_{ij} = 0$, obteniendo así con la expresión anterior

$$q^f(X_i, X_j) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} + \frac{1}{h_i(x_i)} \frac{\partial h_i(x_i)}{\partial x_i} \delta_{ij} \frac{\partial}{\partial x_i} \right) f$$

De esta forma la matriz que representa la Hessiana de f es de la forma H_x^f

$$H_x^f = f''(x) + G(x)^{\frac{1}{2}} \left(G(x)^{-\frac{1}{2}} \right)' F'(x)$$

Donde $F'(x) = \text{diag} \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)$, $G(x) = \text{diag} \left(\frac{1}{(h_1(x_1))^2}, \dots, \frac{1}{(h_n(x_n))^2} \right)$ y

$$f''(x) = \text{diag} \left(\frac{\partial^2 f}{\partial x_1^2}, \dots, \frac{\partial^2 f}{\partial x_n^2} \right)$$

Corolario 2. Si el punto $p \in M$ es un punto crítico de la función f , además se tiene que $X, Y \in H$, entonces se cumple que

$$H_x^f(X(p), Y(p)) = X(p)Y(p)f$$

Demostración: Recordemos que:

$$H_x^f(X(p), Y(p)) = X(p)Y(p)f - \left\langle \nabla_{X(p)} Y(p), \text{grad}f(p) \right\rangle$$

Esta fórmula viene de la definición general del Hessiano de una función f sobre una variedad M , cuando X, Y son campos vectoriales y ∇ es una conexión afín. Por definición, como p es un punto crítico de f , entonces $\text{grad}f(p) = 0$, por lo tanto

$$\left\langle \nabla_{X(p)} Y(p), \text{grad}f(p) \right\rangle = 0$$

Cumpliendo así el Corolario 2.

Definición 23. Una variedad Riemanniana M es llamada geodésicamente hablando como completa si para todo $p \in M$, las geodésicas que parten de p están definidas para todos los valores del parámetro $t \in \mathbb{R}$.

Ejemplo 19. En el espacio \mathbb{R}^n con la métrica I , geodésicamente es completa, pues dado un punto $x \in M$ y una dirección arbitraria $v \in T_x M$, donde la i -ésima componente de la geodésica cumple que $\alpha_i(0) = x_i$ y el vector $\alpha_i'(0) = v_i, \forall i = 1, \dots, n$, donde $\alpha_i = x_i + tv_i$ definido $\forall t \in \mathbb{R}$.

Se considerará en el estudio que la variedad Riemanniana tiene la propiedad de conexidad, esto es, para un par de puntos $p, q \in M$. Existe la curva $\gamma: [a, b]: \rightarrow M$, tal que $\gamma(a) = p, \gamma(b) = q$.

Definición 24. Dado los puntos $p, q \in M$, la distancia Riemanniana esta denotada por $d(x, y)$, está definida por

$$d(P_1q) = \inf_{\gamma} \int_a^b \|\gamma'(t)\| dt \quad (XV)$$

Proposición 15. El espacio M es un espacio métrico, esto se debe a la distancia geodésica definida en la expresión (XV)

Definición 25. (*Retracción*) Definimos una retracción R sobre una variedad M como una aplicación regular del plano tangente TM sobre M que satisface las propiedades:

- a) $R_x(0_x) = x$ y 0_x denota el elemento cero del espacio T_xM .
- b) Para la identificación canónica $T_{0_x}T_xM \approx T_xMR_x$ satisface la igualdad

$$DR_x(0_x) = id_{T_xM} \quad (XVI)$$

Donde id_{T_xM} denota la aplicación identidad sobre T_xM .

Como T_xM es un espacio vectorial, existe una correspondencia natural entre el espacio tangente en el vector nulo 0_x dentro de T_xM con el mismo espacio. Es decir, se puede identificar $T_{0_x}(T_xM) \approx T_xM$ lo que permite tratar elementos de T_xM como si fueran puntos de la variedad M . Esta observación destaca la relevancia del uso de retracciones, como se discutirá más adelante.

Además, nótese que R_x representa la restricción de una retracción general R al espacio tangente T_xM en un punto $x \in M$. Esta retracción cumple una condición de rigidez local, ya que conserva la dirección del gradiente en el punto x . En esta formulación, se asume que el dominio de R abarca todo el fibrado tangente TM .

A partir de la propiedad inicial (punto a), se tiene que la retracción R_x lleva el vector nulo O_x al punto base x , es decir, $R_x(O_x) = x$, lo cual será una propiedad clave para lo que sigue.

Definición 26. Considere la aplicación regular F entre dos variedades Riemannianas M, N , esto es $F: M \rightarrow N$. Y sea ξ_x un vector tangente en el punto $x \in M$, se puede mostrar la aplicación $D^F(x)[\xi_x]$ de $\tilde{J}_{F(x)}(N)$ en \mathbb{R} definido por:

$$(D^F(x)[\xi_x]f) := \xi(f \circ F) \quad (XVII)$$

Es un vector tangente a N en $F(x)$

El vector tangente $D^F(x)[\xi_x]$ es obtenido por la composición $f \circ F$ y la curva γ trazada por ξ_x . La relación

$$DF(x): T_x M \rightarrow T_{F(x)} N$$

$$\xi \rightarrow DF(x)[\xi]$$

Es llamada la aplicación diferencial por Absil (2008), también llamadas derivada o aplicación tangente de F . Donde $\tilde{J}_{(x)}(M)$ es el conjunto de funciones suaves sobre una vecindad de x .

Se puede afirmar que F es una inmersión $\Leftrightarrow DF(x): T_x M \rightarrow T_{F(x)} N$ es inyectiva, al ser N un espacio vectorial ξ , se tiene la identidad $T_{F(x)} \xi \cong \xi$, la cual nos da como resultado que

$$D^F(x)[\xi_x] = \sum_i (\xi_x F^i) e_i \quad (XVIII)$$

De igual forma la descomposición de $F(x)$ es de la forma $F(x) = \sum_i F^i(x)$ con la base e_i donde $i = 1, \dots, n$. El resultado se puede expresar cuando $N = \mathbb{R}$, expresando de forma más sencilla de la siguiente manera

$$DF(x)[\xi_x] = \xi_x F \quad (XIX)$$

Aclarando la expresión (XVIII) donde se tiene la identidad canónica $T_{F(x)} \xi \cong \xi$ se puede revisar Absil (2008), pp. 39,40.

Por último, el operador $DF(x)[\xi_x]$, se puede usar $\xi_x F$ para referirse a la derivada. Esto se cumple si M y N son variedades lineales, se puede identificar el espacio tangente a una variedad, es decir $T_x M \cong M$ y $T_y N \cong N$ y usando la definición de la derivada se tiene

$$DF(x)[\xi_x] = \lim_{t \rightarrow 0} \frac{F(x + t\xi_x) - F(x)}{t} \quad (XX)$$

Como $F: M \rightarrow TN$ es una función diferenciable y el campo vectorial ξ sobre la variedad M , denotado así

$$\begin{aligned} DF[\xi]: M &\rightarrow TN \\ x &\rightarrow DF(x)[\xi_x] \end{aligned}$$

En forma específica si $f: \mathbb{R} \rightarrow M$ sobre el campo ξ en la variedad M , de esta forma, se puede escribir $Df(\xi) = f\xi$.

Definición 27. (*Mínimo: global, local y estricto*) Sea M una variedad Riemanniana completa y $f: M \rightarrow \mathbb{R}$ una función

- a) $\bar{x} \in M$ es un mínimo global de f si, $f(\bar{x}) \leq f(x), \forall x \in M$
- b) $\bar{x} \in M$ es un mínimo local de f , si existe $\delta > 0$ tal que

$$f(\bar{x}) \leq f(x), \forall x \in B(\bar{x}, \delta)$$

Donde $B(\bar{x}, \delta) = \{x \in M, d(\bar{x}, x) < \delta\}$

- c) $\bar{x} \in M$ es un mínimo local estricto si, existe $\delta > 0$ tal que $f(\bar{x}) < f(x)$, para todo $x \neq \bar{x}, x \in B(\bar{x}, \delta)$.

El problema de optimización a resolver es el modelo

$$\min_{x \in M} f(x) \quad (XXI)$$

Lo cual significa encontrar los mínimos globales de la función f sobre la variedad Riemanniana M , se llama problema de optimización sobre algunas restricciones en su dominio, este modelo también se puede

expresar como un problema de optimización en forma de maximización de la siguiente forma

$$\max_{x \in M} f(x)$$

Solo es cuestión de definir la función $f(x) = -g(x)$.

Definición 28. (*Existencia de mínimo global*) Una función $f: M \rightarrow \mathbb{R}$ es denominada semicontinua inferior en $\bar{x} \in M$, para toda sucesión $\{x^k\} \in M$ convergente a \bar{x} se obtiene que:

$$\liminf_{k \rightarrow \infty} f(x^k) \geq f(\bar{x})$$

Si f es semicontinua inferior para todo $x \in M$, entonces se dice que f es semicontinua inferior en M . A continuación, el siguiente teorema logra garantizar la existencia de un mínimo global para el problema de optimización.

Teorema 3. (*Weierstrass*) Considerando el problema de optimización, si la función $f: M \rightarrow \mathbb{R}$ es semicontinua inferior y M es compacto, entonces existe un punto mínimo global de f .

Demostración: Se demuestra primero que la función f es limitada inferiormente, esto se debe ya que existe $\alpha \in \mathbb{R}$, tal que $\alpha \leq f(x), \forall x \in M$.

Por contradicción, supongamos que la función f no es limitada inferiormente, por lo tanto existe una sucesión $\{x^k\} \subset M$ tal que

$$\lim_{k \rightarrow +\infty} f(x^k) = -\infty \tag{XXII}$$

Como M es compacto, entonces existe la subsucesión $\{x^{k_j}\} \subset \{x^k\}$ tal que

$$\lim_{j \rightarrow +\infty} x^{k_j} = \hat{x}$$

Por la semicontinuidad de la función tenemos

$$\liminf_{j \rightarrow +\infty} f(x^{k_j}) \geq f(\hat{x})$$

Contradiendo la ecuación (XXII), por lo tanto f es limitada inferiormente en M . De esta forma, existe $f^* \in \mathbb{R}$ tal que $f^* = \inf\{f(x): x \in M\}$. Como es ínfimo, entonces existe una sucesión $\{x^k\} \subset M$ tal que

$$\lim_{k \rightarrow \infty} f(x^k) = f^*$$

Como M es compacto, existe \bar{x} y $\{x^{k_j}\} \subset \{x^k\}$ tal que $\lim_{j \rightarrow +\infty} x^{k_j} = \bar{x} \in M$

De igual forma, por la semicontinuidad inferior de la función

$$\liminf_{j \rightarrow \infty} f(x^{k_j}) \geq f(\bar{x})$$

Como $\{f(x^k)\}$ converge a f^* , la subsucesión $\{f(x^{k_j})\}$ converge a f^* , de esta forma se obtiene

$$f^* \geq f(\bar{x})$$

De esta forma, \bar{x} es un punto mínimo global de f en M

Teorema 4. (Condición necesaria de primer orden) Sea $f: M \rightarrow \mathbb{R}$ de clase C^1 . Si x^* es un punto mínimo local, entonces $\text{grad } f(x^*) = 0$.

Demostración: Se toma un vector $v \in T_{x^*}M$ y una curva $\gamma: \mathbb{R} \rightarrow M$ con condiciones iniciales $\gamma(0) = x^*$ y $\gamma'(0) = v$. Definiendo la aplicación h donde $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$. Como x^* es punto mínimo local para la función, entonces existe $\delta > 0$ tal que

$$h(0) = f(x^p) \leq f(\gamma(0)) = h(t)$$

Donde $\forall t \in (-\delta, \delta)$, lo cual genera en $t = 0$ se tiene un punto mínimo local de h . Debido a la condición necesaria de primer orden se tiene

$$h'(0) = \langle \text{grad } f(x^*), v \rangle = 0$$

En forma particular si se toma $v = \text{grad } f(x^*)$ se tiene que $\text{grad } f(x^*) = 0$

Teorema 5. (Condición necesaria de segundo orden) Sea $f: M \rightarrow \mathbb{R}$ de clase C^2 . Si x^* es punto mínimo local, entonces $\langle v, H_{x^*}^f v \rangle \geq 0, \forall v \in T_{x^*}M$.

Demostración: Sea el vector $v \in T_{x^*}M$ y la curva geodésica $\gamma: \mathbb{R} \rightarrow M$ donde se tiene $\gamma(0) = x^*$ y $\gamma'(0) = v$. Definiendo la aplicación h donde $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$. Del teorema anterior de la condición necesaria de primer orden, se tiene que cuando $t = 0$ se tiene un mínimo local de h , por lo tanto la condición necesaria de segundo orden $h'(0) = 0$ y $h''(0) \geq 0$

$$h'(t) = \langle \text{grad } f(\gamma(t)), \gamma'(t) \rangle$$

$$h''(t) = \frac{d}{dt} \langle \text{grad } f(\gamma(t)), \gamma'(t) \rangle$$

$$h''(t) = \left\langle \frac{D}{dt} (\text{grad } f(\gamma(t))), \gamma'(t) \right\rangle + \left\langle \text{grad } f(\gamma(t)), \frac{D}{dt} (\gamma'(t)) \right\rangle$$

$$h''(t) = \langle H_{\gamma(t)}^f \gamma'(t), \gamma'(t) \rangle$$

$$h''(t) = \langle v, H_{x^*}^f v \rangle = \langle H_{x^*}^f v, v \rangle \geq 0$$

Teorema 6. (Condición suficiente de segundo orden). Sea $f: M \rightarrow \mathbb{R}$ de clase C^2 . Si $x^* \in M$ que satisface:

- a) $\text{grad } f(x^*) = 0$
- b) $H_{x^*}^f$ definida positiva

Entonces, x^* es un punto de mínimo local estricto de f

Demostración: Se demostrará por contradicción, supongamos que x^* no es un punto mínimo local estricto, entonces existe una subsucesión de la forma $\{x^k\} \in B(x^*, \frac{1}{k}) / \{x^*\}$ cumpliendo que

$$f(x^*) \geq f(x^k) \tag{XXIII}$$

Sea la curva geodésica $\gamma_k: [0,1] \rightarrow \mathbb{R}$ tal que los puntos de las condiciones iniciales son $\gamma_k(0) = x^*$, $\gamma_k(1) = x^k$, $\gamma_k'(0) = v_k$ y $d(x^*, x^k) = \|\exp_{x^*} x^k\|$. Se define $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = (f \circ \gamma_k)(t)$ y por el desarrollo de Taylor de segundo orden de h en 0.

$$h(t) = h(0) + th'(0) + \frac{1}{2}t^2h''(0) + \theta(|t|^2)$$

Donde $\lim_{t \rightarrow 0} \frac{\theta(|t|^2)}{|t|^2} = 0$, obteniendo de esta forma

$$f(\gamma_k(t)) = f(x^*) + \frac{t^2}{2} \left\langle \frac{D}{dt} \text{grad } f(x^*) v_k, v_k \right\rangle + \theta(|t|^2)$$

En forma particular, reemplazando en $t = 1$

$$f(x^k) = f(x^*) + \frac{1}{2} \langle v_k, H_{x^*}^f v_k \rangle + \theta(d^2(x^*, x^k)) \quad (XXIV)$$

Donde se cumple que $\lim_{d \rightarrow 0} \frac{\theta(d^2(x^*, x^k))}{(d^2(x^*, x^k))} = 0$

Definiendo $z^k = \frac{v_k}{\|v_k\|}$, la sucesión $\{z^n\}$ es limitada, por lo tanto existe una subsucesión $\{z^{k_j}\} \subset \{z^k\}$ donde $\{z^{k_j}\} \rightarrow \bar{z}$. Reemplazando en la ecuación (XXIV) el valor de k por k_j , se tiene

$$f(x^{k_j}) = f(x^*) + \frac{1}{2} \langle v_{k_j}, H_{x^*}^f v_{k_j} \rangle + \theta(d^2(x^*, x^{k_j})) \quad (XXV)$$

Donde se cumple que $\lim_{d \rightarrow 0} \frac{\theta(d^2(x^*, x^{k_j}))}{(d^2(x^*, x^{k_j}))} = 0$

De la expresión (XXIII) y aplicando el límite cuando $j \rightarrow \infty$ en la expresión (XXV) se obtiene que

$$0 \geq \langle \bar{z}, H_{x^*}^f \bar{z} \rangle$$

Contradiciendo la hipótesis del teorema, de esta forma x^* es un punto mínimo local estricto.

Definición 29. (Convexidad en una variedad) Sea M una variedad Riemanniana completa, se dice que $A \subset M$ es totalmente convexo, si para cualquier par de puntos p y q de A (no distintos necesariamente), las geodésicas que unen estos puntos están contenidos íntegramente en A .

Ejemplo 20. Si la variedad $M = \mathbb{R}^n$ con la métrica de la identidad $G(x) = I$, cualquier conjunto de la variedad que sea convexo en un sentido clásico, es totalmente convexo.

Ejemplo 21. Si $p \in M$, existiendo una relación geodésica en forma no trivial en p , ósea una curva $\gamma: [a, b] \rightarrow M$, donde se cumple $\gamma(a) = p = \gamma(b)$ con $\gamma(t) \neq p$ para algún $t \in [0, 1]$, por lo tanto el conjunto $A = \{p\}$ no es completo en forma total. De esta forma, se puede deducir que no necesariamente los conjuntos unitarios son convexos totalmente.

Definición 30. Decimos que $A \subset M$ es convexo si para todo par de puntos p y q de A existe una geodésica minimal que une los puntos p y q contenido en A .

Ejemplo 22. El mismo espacio M y los conjuntos unitarios son espacios convexos.

Definición 31. Sea $f: M \rightarrow \mathbb{R}$ es llamada función convexa si su restricción a cualquier geodésica de M es una función convexa en \mathbb{R} , es decir, si la curva $\gamma: \mathbb{R} \rightarrow M$ es una geodésica entonces:

$$f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$$

Es convexa

Teorema 7. Sea $f: M \rightarrow \mathbb{R}$ es convexa si, y solamente si, para todo segmento de geodésica $\gamma: [a, b] \rightarrow M$ y para cualquier $\lambda \in [0, 1]$ se verifica

$$f(\gamma(1 - \lambda)a + \lambda b) \leq (1 - \lambda)f(\gamma(a)) + \lambda f(\gamma(b)) \quad (XXVI)$$

Demostración: Sea f convexa y $h: \mathbb{R} \rightarrow \mathbb{R}$, donde $h(t) = f(\gamma(t))$. Para los puntos $a, b \in [a, b]$ y $\lambda \in [0, 1]$ se obtiene que

$$h((1 - \lambda)a + \lambda b) \leq (1 - \lambda)h(a) + \lambda h(b)$$

Obteniendo así la expresión (XXVI), en forma recíproca se define el punto $t = (1 - \lambda)a + \lambda b$ con $\lambda \in [0, 1]$ entonces

$$f \circ \gamma(t) = f(\gamma(1 - \lambda)a + \lambda b) \leq f(\gamma(1 - \lambda)a) + f(\gamma(\lambda b))$$

$$f \circ \gamma(t) \leq (1 - \lambda)f(\gamma a) + \lambda f(\gamma(b))$$

$$f \circ \gamma(t) \leq (1 - \lambda)f \circ \gamma(a) + \lambda f \circ \gamma(b)$$

Observación 14. La definición anterior es la forma general de la definición clásica de convexidad en $M = \mathbb{R}^n$ con su métrica usual. Esto ocurre ya que para un par de puntos p y q de la geodésica $\gamma: [0,1] \rightarrow M$ que los une es

$$\gamma(\lambda) = p + \lambda(q - p) = (1 - \lambda)p + \lambda q$$

Luego por el teorema 7 se puede obtener

$$f(\gamma(\lambda)) = f((1 - \lambda)p + \lambda q) \leq (1 - \lambda)f(p) - \lambda f(q)$$

$$f(\gamma(\lambda)) = f((1 - \lambda)\gamma(0) + \lambda\gamma(1)) \leq (1 - \lambda)f(\gamma(0)) - \lambda f(\gamma(1))$$

$$f(\gamma(\lambda)) \leq (1 - \lambda)f(p) - \lambda f(q)$$

Para $\alpha \in \mathbb{R}$, se define el conjunto de nivel $M^\alpha = \{x \in M; f(x) \leq \alpha\}$

Teorema 8. Si $f: M \rightarrow \mathbb{R}$ es convexa, entonces M^α es totalmente convexo.

Demostración: Sea los puntos p y $q \in M^\alpha$ y la geodésica $\gamma: [a, b] \rightarrow M$ tal que $\gamma(a) = p$ y $\gamma(b) = q$. Se puede probar que $\gamma(t) \in M^\alpha, \forall t \in [a, b]$.

En efecto, sea $t = (1 - \lambda)a + \lambda b$ para algún $\lambda \in [0,1]$, debido a que la función f es convexa y por el teorema 7 se puede obtener

$$f(\gamma(\lambda)) = f(\gamma(1 - \lambda)a + \lambda b) \leq (1 - \lambda)f(\gamma(a)) - \lambda f(\gamma(b))$$

$$f(\gamma(\lambda)) \leq (1 - \lambda)f(p) - \lambda f(q)$$

$$f(\gamma(\lambda)) \leq (1 - \lambda)\alpha + \lambda\alpha = \alpha$$

De esta forma se obtiene que $f(\gamma(\lambda)) \leq \alpha$, por tanto $\gamma(t) \in M^\alpha$

Teorema 9. Sea $f: M \rightarrow \mathbb{R}$, es convexo en p si y solo si, para cualquier geodésica $\gamma: \mathbb{R} \rightarrow M$ con $\gamma(0) = p$, cumple la desigualdad

$$f(\gamma(t)) - f(p) \geq t \langle \text{grad}^f(p), \gamma^1(0) \rangle \quad (XXVII)$$

Demostración: Se define una aplicación $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$, donde h es convexa en 0 desde que f es convexa en p , teniendo así

$$h(t) - h(0) \geq th'(0)$$

Luego se puede obtener

$$f(\gamma(t)) - f(p) \geq \langle \text{grad } f(p), \gamma'(0) \rangle$$

De forma análoga, si $f(\gamma(t)) - f(p) \geq \langle \text{grad } f(p), \gamma'(0) \rangle$, esto significa que

$$h(t) - h(0) \geq th'(0)$$

Por lo tanto h es convexa en 0 y además f es convexa en p

Teorema 10. Si $f: M \rightarrow \mathbb{R}$ es convexa, entonces todo punto crítico de f es un punto mínimo global de f .

Demostración: Sea $x \in M$, por el teorema de Hopf-Rinow se puede considerar una geodésica $\gamma: \mathbb{R} \rightarrow M$ tal que $\gamma(0) = x$ y $\gamma(b) = y$, debido a que f es convexa por el teorema 9

$$f(\gamma(b)) - f(\gamma(0)) \geq b \langle \text{grad } f(\gamma(0)), \gamma'(0) \rangle$$

Esto significa que

$$f(\gamma(b)) - f(x) \geq b \langle \text{grad } f(x), \gamma'(0) \rangle$$

Debido a que $\text{grad } f(x) = 0$ entonces $f(y) \geq f(x)$, para todo $y \in M$. Por lo tanto x es punto de mínimo global de f .

Teorema 11. Sea $f: M \rightarrow \mathbb{R}$ de clase C^2 , f es convexa si y solo si, para todo $p \in M$ la Hessian de f en p

$$H_p^f: T_p M \rightarrow T_p M$$

Es semidefinida positiva

Demostración: Sea el vector $v \in T_p M$, $p \in M$ y la geodésica $\gamma: \mathbb{R} \rightarrow M$ tal que se cumple $\gamma(0) = p$ y $\gamma'(0) = v$. Definiendo la aplicación $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$ se sabe que h es convexa y de clase C^2 . De la condición de convexa se puede obtener que $h''(t) \geq 0$, de esta forma

$$h'(t) = \langle \text{grad } f(\gamma(t)), \gamma'(t) \rangle$$

Además como $h''(0) = \langle v, H_p^f v \rangle \geq 0$

En forma análoga, si $\langle v, H_p^f v \rangle \geq 0$ entonces f es convexo. En efecto para la función $h: \mathbb{R} \rightarrow \mathbb{R}$ convexa, entonces se tiene $h = f \circ \gamma$ es convexa.

Definición 32. Sea la variedad Riemanniana M completa y $f: M \rightarrow \mathbb{R}$ una función real, f es llamada cuasiconvexa en M si para todo $x, y \in M, t \in [0,1]$ se cumple

$$f(\gamma(t)) \leq \max\{f(x), f(y)\}$$

Para toda curva geodésica $\gamma: [0,1] \rightarrow M$, tal que $\gamma(0) = x$ y $\gamma(1) = y$

Teorema 12. Sea $f: M \rightarrow \mathbb{R}$ una función diferenciable y cuasiconvexa en una variedad Riemanniana completa M y sea $x, y \in M$. Si $f(x) \leq f(y)$ por lo tanto

$$\langle \text{grad } f(y), \gamma'(0) \rangle \leq 0$$

Donde $\text{grad } f$ es el gradiente de f y γ es la geodésica donde $\gamma(0) = y$, $\gamma(1) = x$.

Demostración: Sea la geodésica $\gamma: [0,1] \rightarrow M$ tal que $\gamma(0) = y$, $\gamma(1) = x$. Se define $h: \mathbb{R} \rightarrow \mathbb{R}$ tal que $h(t) = f \circ \gamma(t)$, usando la aproximación de Taylor de primer orden de h en $t = 0$ se tiene

$$h(t) = h(0) + th'(0) + \theta t$$

Donde $\lim_{t \rightarrow 0} \frac{\theta(t)}{t}$, entonces se tiene

$$f(\gamma(t)) = f(\gamma(0)) + t\langle \text{grad } f(\gamma(0)), \gamma'(0) \rangle + \theta(|t|)$$

Como f es cuasiconvexa y $f(x) \leq f(y)$ se tiene

$$t\langle \text{grad } f(y), \gamma'(0) \rangle$$

Dividiendo por t y tomando límite cuando $t \rightarrow 0$ se tiene que

$$\langle \text{grad } f(y), \gamma'(0) \rangle \leq 0$$

Definición 33. Una función diferenciable $f: M \rightarrow \mathbb{R}$ es pseudoconvexa si, para todo par de puntos distintos $x, y \in M$ y toda curva geodésica que une x con $(\gamma(0) = x$ e $\gamma(1) = y)$ se tiene

$$\langle \text{grad } f(y), \gamma'(0) \rangle \geq 0 \text{ entonces } f(y) \geq f(x)$$

Teorema 13. Sea $f: M \rightarrow \mathbb{R}$ una función diferenciable y pseudoconvexa. Entonces x^* es un mínimo global de f , si y solo si, $\text{grad } f(x^*) = 0$.

Demostración: Sea la geodésica $\gamma: \mathbb{R}_+ \rightarrow M$ tal que $\gamma(0) = x^*$ y se define la aplicación $h: \mathbb{R} \rightarrow \mathbb{R}$ con $h = f \circ \gamma$, por el teorema de Taylor de primer orden de h en 0

$$h(t) = h(0) + th'(0) + \theta(|t|)$$

Donde $\lim_{t \rightarrow 0} \frac{\theta(|t|)}{|t|}$, entonces se tiene

$$f(\gamma(t)) = f(x^*) + t\langle \text{grad } f(x^*), \gamma'(0) \rangle + \theta(|t|)$$

Como x^* es mínimo global entonces $f(x^*) \leq f(y)$, en forma particular para $y = \gamma(t)$ entonces se tiene

$$f(y) - f(x^*) = t\langle \text{grad } f(x^*), \gamma'(0) \rangle + \theta(|t|)$$

Luego

$$t\langle \text{grad } f(x^*), \gamma'(0) \rangle + \theta(|t|) \geq 0$$

Que cuando el límite $t \rightarrow 0$, $\langle \text{grad } f(x^*), \gamma'(0) \rangle + \theta(|t|) \geq 0$, por último considerando $\gamma'(0) = -\text{grad } f(x^*)$, se tiene $\text{grad } f(x^*) = 0$. El recíproco solo se puede demostrar considerando la definición de la función f ser pseudoconvexa.

2.4 Definiciones de términos básicos

Superficie regular: Una superficie regular es un conjunto de puntos en el espacio tridimensional que, cerca de cada uno de ellos, se puede representar mediante funciones suaves y sin cruces o pliegues. Esto significa que, en una pequeña región alrededor de cualquier punto, la

superficie se comporta como un plano y permite aplicar el cálculo diferencial. Ejemplos comunes incluyen formas como la esfera o el cilindro.

Variedad diferenciable: Una variedad diferenciable es un tipo de espacio que, a pesar de tener una estructura compleja o curva en su totalidad, se asemeja localmente a un espacio euclidiano como \mathbb{R}^n , lo que permite utilizar herramientas del cálculo diferencial. Un ejemplo de esto es la superficie de una esfera, que aunque es curva, en regiones pequeñas se comporta como un plano. Una variedad diferenciable en general es un espacio topológico (no precisamente vectorial) semejante localmente al espacio euclidiano cuya relación tiene el soporte de la diferenciabilidad.

Espacio tangente a la variedad: Un concepto muy importante en la teoría de variedades diferenciables es el espacio tangente. Si imaginamos que la variedad está inmersa en el espacio euclidiano \mathbb{R}^n , se puede hacer corresponder $p \in M$ un cierto espacio vectorial, el espacio de los vectores tangentes a M en p , así la esfera S^2 inmersa en \mathbb{R}^3 , el plano tangente en un punto p , es el conjunto de los vectores ortogonales al vector posición \vec{p} , es decir $\{\vec{v} \in \mathbb{R}^3 / \vec{v} \cdot \vec{p} = 0\}$

Dichos vectores tangentes también son conocidos como vector velocidad en \mathbb{R}^3 . Como en variedades no se dispone del soporte de un espacio ambiente, nos enfocamos de una característica del vector tangente para sustituir la noción de velocidad. Se recuerda algunas nociones en superficies regulares.

Sea $\varepsilon > 0, \gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ una curva regular donde sus elementos tienen la forma $\gamma(t) = (\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t))$, por lo tanto $\gamma(0) = (\gamma_1(0), \dots, \gamma_n(0))$ y $\gamma'(0) = (\gamma_1'(0), \dots, \gamma_n'(0)) = v \in \mathbb{R}^n$

Sea además $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable en \mathbb{R}^n en una variedad de p , entonces se puede restringir f a la curva γ y calcular la derivada direccional de f en dirección hacia v denotado como:

$$\frac{d(f \circ \gamma)}{dt}(t)|_{t=0} = \sum_{i=1}^n \frac{\partial f}{\partial \gamma_i}(\gamma(0)) \frac{d\gamma_i}{dt}(0) = \left(\sum_{i=1}^n \gamma_i'(0) \left(\frac{\partial}{\partial \gamma_i} \right) p \right) f$$

Por lo tanto, la derivada direccional sobre el vector v es un operador sobre las funciones diferenciable que depende de v , sobre esta propiedad se usará la definición del vector tangente en variedades.

Métricas Riemannianas en variedades diferenciables: Estas métricas nos permiten calcular distancias y ángulos y dentro de la variedad. Consiste en asignar, en cada punto, un producto interno al espacio tangente, lo que convierte localmente a la variedad en un espacio similar al euclidiano, pero con una geometría que puede variar de un punto a otro. Esta estructura permite extender conceptos del cálculo vectorial y la geometría clásica a espacios más generales y curvos.

Se definirá la métrica dentro del espacio tangente $T_p M$ donde podemos definir el producto interno:

$$(v, \omega)_p = \sum_i v_i \omega_i \quad y \quad \langle v, \omega \rangle_p = \sum_{ij} g_{ij} v_i \omega_j = (Gv, w)$$

Donde G es una matriz simétrica definida positiva, esta definición aparece al realizar cambios en las coordenadas, ya que si se define las coordenadas x e z pertenecientes a \mathbb{R}^n , donde se cumple que $x(t) = x(z(t))$

Entonces para todo $i = 1, 2, \dots, n$ se cumple que

$$\frac{dx_i}{dt} = \sum_{j=1}^n \left(\frac{\partial x_i}{\partial z_j} \frac{\partial z_j}{\partial t} \right)$$

Si definimos a v^x como las derivadas de cada coordenada de x respecto a t y análogamente definimos a v^z se puede obtener que

$$\|v^x\|^2 = \sum_{i=1}^n \left(\frac{dx_i}{dt}\right)^2 = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial x_i}{\partial z_j} v_j^z\right)^2$$

Haciendo un cambio entre k e i , obtenemos lo siguiente

$$\sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n \frac{\partial x_i}{\partial z_k} v_k^z \frac{\partial x_i}{\partial z_j}\right) v_j^z = \sum_{i,j=1}^n \left(\sum_{k=1}^n \frac{\partial x_k}{\partial z_j} \frac{\partial x_k}{\partial z_i}\right) v_i^z v_j^z$$

Definimos la matriz $G = g_{ij} = \sum_{k=1}^n \frac{\partial x_k}{\partial z_j} \frac{\partial x_k}{\partial z_i}$, obteniendo así

$$(v^x, v^x) = \|v^x\|^2 = \sum_{i,j=1}^n (g_{ij}) v_i^z v_j^z = (Gv^z, v^z)$$

Con este resultado, queda claro el cambio de coordenadas no debe alterar la métrica, ya que si consideramos a $v = G^{1/2}w$ se obtiene que el producto interno $(v, v)_p = (G^{1/2}w, G^{1/2}w)_p = (Gw, w)_p = \langle w, w \rangle_p$.

Campos de vectores: En el estudio de las variedades Riemannianas, se consideran los campos de vectores como asignaciones que a cada punto de la variedad le asocian un vector perteneciente a su espacio tangente. Estos campos resultan especialmente relevantes en el contexto de la Optimización Matemática, ya que permiten modelar las posibles direcciones de desplazamiento que puede seguir un algoritmo iterativo desde un punto inicial. Al analizar la evolución continua de dicho algoritmo, se obtiene una trayectoria cuya forma incluyendo propiedades como la curvatura o la torsión está influenciada directamente por la estructura del campo de vectores que guía su movimiento. Por este motivo, se hace necesario introducir nociones análogas a la derivada de funciones, pero adaptadas a la geometría de las variedades, que permitan medir cómo varía un campo de vectores a lo largo de la variedad. Esto da lugar a

herramientas como la derivada covariante, fundamentales para comprender el comportamiento dinámico de los algoritmos en espacios curvos.

Conexiones afines: es una estructura matemática que permite derivar y comparar vectores en distintos puntos de una variedad diferenciable. Su función principal es definir la derivada covariante, lo que hace posible estudiar cómo varían los campos de vectores a lo largo de otros campos. Gracias a esto, se pueden definir conceptos como geodésicas y curvatura, extendiendo el cálculo diferencial a espacios curvos. Es esencial para analizar trayectorias, variaciones y transportes de vectores en geometría y optimización sobre variedades.

Geodésicas: Una geodésica es la trayectoria más corta (o más recta) entre dos puntos en un espacio curvo. Generaliza la idea de línea recta en geometría euclidiana. En términos técnicos, es una curva que no cambia de dirección según la geometría del espacio, es decir, su aceleración covariante es cero. Es decir que

$$\frac{d}{dt} \left(\frac{d\alpha}{dt} \right) = 0$$

Curvatura de una variedad Riemanniana: Describe cómo varía la forma del espacio en torno a cada punto, comparado con un espacio plano. Informa sobre cómo se curvan las geodésicas, cómo cambia la distancia entre vectores paralelos transportados y cómo se comporta la geometría en diferentes direcciones.

Variedades Completas: En el contexto de la optimización desde la perspectiva de la geometría Riemanniana, es fundamental asumir que la variedad en cuestión sea completa. Esto, en términos prácticos, implica que cualquier geodésica dentro de la variedad puede extenderse indefinidamente a lo largo del parámetro $t \in \mathbb{R}$. Un resultado clave relacionado con esta propiedad es el Teorema de Hopf y Rinow, que

establece que, en una variedad completa, para cualquier par de puntos existe siempre una geodésica que representa la distancia mínima entre ellos, comparada con todas las curvas que los conectan y que son suavemente diferenciables por tramos.

Una inquietud natural que surge es si sería posible formular métodos basados en geodésicas bajo condiciones menos estrictas que la completitud, por ejemplo, definiendo una medida en la variedad mediante el producto interno del espacio tangente, de modo que las geodésicas estén definidas en casi todo punto, salvo en un conjunto de medida nula.

Retracción sobre variedades: Las retracciones son herramientas esenciales en el desarrollo de esta investigación, ya que permiten trasladar conceptos del espacio euclidiano al contexto de variedades diferenciables. De manera general, una retracción se define como una aplicación suave que lleva vectores del espacio tangente en un punto de la variedad de vuelta a la propia variedad. Esta función está estrechamente relacionada con el concepto de exponencial Riemanniana, ya que en muchos casos la retracción se construye a partir de ella, lo cual establece una conexión directa con las geodésicas, como lo señala Boothby (1975).

El problema de optimización y sus condiciones de optimalidad: Los fenómenos observados en la naturaleza pueden ser analizados mediante modelos matemáticos, los cuales, en numerosos casos, se expresan como problemas de optimización. Uno de estos modelos será presentado más adelante. Para abordar adecuadamente tales problemas, es fundamental, en primer lugar, establecer las condiciones que aseguren la existencia y caracterización de soluciones óptimas. Solo entonces se podrá proponer un algoritmo efectivo que permita resolver dichos modelos de optimización.

Con ese propósito, es necesario definir con precisión el espacio sobre el que trabajaremos. En este contexto, introducimos el concepto de variedad

convexa y, dentro de ella, nos enfocaremos en una clase especial de funciones: las funciones convexas y cuasiconvexas.

Análisis convexo y cuasiconvexo: El desarrollo del análisis convexo en el contexto de variedades Riemannianas fue abordado por Rapsack (1997) y Udriste (1997). En su trabajo, Rapsack demuestra que en una variedad diferenciable cuya métrica proviene de \mathbb{R}^n , es posible obtener caracterizaciones de primer y segundo orden aplicables a problemas de optimización definidos sobre dicha variedad. Por su parte, Udriste (1997) estudia esta problemática en el marco más general de una variedad Riemanniana abstracta, extendiendo de forma independiente la teoría de la convexidad.

En cuanto al concepto de convexidad, los estudios orientados al análisis convexo con aplicaciones a la teoría de optimización han experimentado un avance considerable desde entonces. Un ejemplo de ello es el trabajo de Da Cruz Neto et al. (1998), quienes analizan variedades Riemannianas completas con curvatura seccional no negativa. Inspirados en esta línea de investigación, se desarrolla aquí una aproximación al análisis convexo en el contexto de variedades Riemannianas.

Funciones cuasiconvexas: Se dice que una función $f: X \rightarrow \mathbb{R}$ siendo X un subconjunto convexo de un espacio vectorial es cuasiconvexa si se cumple $f(\alpha x + (1 - \alpha)y) \geq \min\{f(x), f(y)\}$ donde $0 \leq \alpha \leq 1$. La función f es cuasiconvexa si y solo si $-f$ es cuasicóncava. Además, por la definición de cuasiconvexidad que f es cuasiconvexa si y solo si se cumple la expresión $f(\alpha x + (1 - \alpha)y) \leq \max\{f(x), f(y)\}$, $0 \leq \alpha \leq 1$ donde X es un subconjunto convexo de \mathbb{R}^n y $f: X \rightarrow \mathbb{R}$.

Se cumplen algunas propiedades de funciones cóncavas las cuales son la desigualdad $f(\sum_{i=1}^n \alpha_i x_i) \geq \sum_{i=1}^n \alpha_i f(x_i)$, es decir que el valor que toma la

función cóncava es una combinación convexa de elementos de X mayor o igual que la combinación convexa de dichos valores.

El subgrafo $s = \{(x, \lambda) \in X \times \mathbb{R} : \lambda \leq f(x)\}$ de una función cóncava, es convexo en $\mathbb{R}^n \times \mathbb{R}$, donde es una combinación lineal no negativa de funciones convexas definidas en un conjunto convexo, lo cual es una función cóncava. Observando que estas definiciones solo han sido desarrolladas en geometría euclidiana, se generalizará en forma particular en variedades Riemannianas.

III. HIPÓTESIS Y VARIABLES

3.1 Hipótesis

Hipótesis general

La geometría Riemanniana tiene herramientas que determinan la demostración de la convergencia del método del Gradiente para la minimización de funciones cuasiconvexas usando retracciones.

Hipótesis Específicas

- El uso de sucesiones convergentes determina la minimización de funciones cuasiconvexas usando retracciones en el método del Gradiente.
- Para minimizar funciones cuasiconvexas sobre variedades Riemannianas se debe reemplazar las geodésicas por retracciones en el método del Gradiente.
- Las métricas Riemannianas específicas favorecen la transformación de problemas de minimización no convexos en cuasiconvexas sobre variedades Riemannianas.

3.2 Operacionalización de las variables

Definición conceptual de variables

Variable Dependiente (D)

Convergencia del Método del Gradiente sobre variedades Riemannianas, es la propiedad que asegura que las sucesivas aproximaciones generadas por este método tienden a un punto óptimo sobre la variedad. Este comportamiento se analiza adaptando el cálculo del gradiente a la estructura geométrica de la variedad, utilizando herramientas como la retracción y el producto interno definido por la métrica Riemanniana.

Variable Independiente (I)

Retracciones en el Método del Gradiente para minimizar funciones cuasi convexas, son aplicaciones que permiten proyectar vectores del espacio tangente de un punto sobre la variedad de regreso a la propia variedad, siguiendo una trayectoria que respeta su geometría. En el caso de funciones cuasiconvexas, las retracciones aseguran que el método del gradiente pueda avanzar hacia una solución óptima sin salirse del dominio de la función, facilitando así la convergencia del algoritmo.

Tabla 1. Operacionalización de la variable

Variables	Dimensiones	Indicadores	Método	Técnica
Independiente Retracciones en el método del Gradiente sobre variedades Riemannianas	Variedades Riemannianas Retracciones en variedades Riemannianas Método del Gradiente	Elementos de variedades Riemannianas Variedades diferenciables Métrica Geodésicas Algoritmo del Gradiente	Método inductivo – Deductivo	Documentos cualitativos Revisión bibliográfica Trabajos con grupos de investigación
Dependiente Minimización de funciones cuasiconvexas	Condiciones de optimalidad	Funciones cuasiconvexas Convergencia de funciones cuasiconvexas	Método inductivo – Deductivo	Documentos cualitativos Revisión bibliográfica

Fuente: Elaboración Propia

IV. METODOLOGÍA DEL PROYECTO

4.1 Diseño metodológico

4.1.1. Tipo de investigación

Se trata de una investigación de tipo básica y descriptiva, cuyo objetivo principal es contribuir al desarrollo del conocimiento en torno al planteamiento del problema. La metodología que se seguirá durante el desarrollo del proyecto responde a un enfoque analítico-deductivo, centrado en el estudio riguroso de conceptos, definiciones, teoremas y corolarios, así como en la elaboración de demostraciones que permitan obtener y discutir los resultados. Esta investigación no está orientada a una aplicación computacional directa.

En este sentido, se clasifica como una investigación demostrativa, fundamentada en marcos teóricos propios de estudios básicos y descriptivos. Su finalidad es ampliar el cuerpo teórico existente, enfocándose en la generación de nuevo conocimiento, sin abordar directamente aplicaciones prácticas derivadas de los análisis teóricos realizados.

4.1.2. Diseño de la investigación

La presente investigación es de tipo no experimental, ya que no se realiza manipulación directa de la variable independiente con el fin de observar efectos sobre la variable dependiente. En este caso específico, la convergencia del método para minimizar funciones está determinada principalmente por la retracción utilizada en el método del gradiente, la cual, a su vez, depende de la métrica Riemanniana seleccionada, tal como se establece en la hipótesis de estudio.

De forma similar, las funciones cuasiconvexas consideradas en esta investigación permiten garantizar la existencia de un óptimo, debido a sus propiedades más generales en comparación con las funciones convexas. Esto respalda el cumplimiento de la hipótesis planteada, la cual será examinada y sustentada a lo largo del desarrollo y análisis de los resultados de esta investigación.

4.2 Método de investigación

En consecuencia, debido con la naturaleza del trabajo de investigación, el método usado es el básico teórico.

4.3 Población y muestra

El problema abordado en esta investigación no posee una dimensión fáctica, ya que no se refiere a hechos observables o medibles en una población específica. Se trata, más bien, de un problema de naturaleza formal, cuya complejidad radica en las dificultades teóricas inherentes a su formulación y análisis. Por esta razón, no implica el estudio de una muestra ni requiere la identificación de una población objetivo.

4.4 Lugar de estudio y periodo desarrollado

Por el tipo de investigación que se ha realizado, el presente trabajo tiene como lugar de estudio la Facultad de Ciencias Naturales y Matemática en el periodo de estudio del año 2018 y 2019.

4.5 Técnicas e instrumentos para la recolección de la información documental

Se empleó una técnica de recopilación de información basada en la revisión documental, utilizando fichas de investigación como herramienta principal para reunir conocimientos vinculados al problema planteado. Este proceso incluyó el uso de diversas definiciones teóricas, teoremas y propiedades, a partir de los cuales se derivaron resultados que sirvieron como base para

la demostración de la hipótesis general, así como de la hipótesis específica que aborda el problema de investigación.

4.6 Análisis y procesamiento de datos

Para fundamentar y desarrollar las hipótesis de esta investigación, se recurrió al uso de definiciones, conceptos, teoremas y propiedades teóricas, los cuales sirvieron como base informativa. Esta información permitió establecer comparaciones entre resultados intermedios y finales, mediante un proceso de análisis y demostración sustentado en razonamientos deductivos y analógicos.

Con ese propósito, se llevó a cabo la recopilación de textos y artículos especializados en el campo de la Optimización Matemática y disciplinas afines. Posteriormente, dicho material fue organizado temáticamente y clasificado de manera cronológica según las áreas a las que pertenecen.

4.7 Aspectos éticos en investigación

La presente investigación cumple con las normas establecidas (artículo 427-438 del código penal, Ley sobre el derecho del autor, decreto legislativo N°822) en el Perú, no incurriendo en el delito contra la fe pública.

RESULTADOS

5.1 Resultados descriptivos

Debido a que el trabajo de investigación es de tipo básica, no cuenta con datos estadísticos o aplicación de herramientas estadísticas en su análisis, por lo tanto, este trabajo no cuenta con resultados descriptivos.

5.2 Resultados inferenciales

Debido a que el trabajo de investigación es de tipo básica, no cuenta con datos estadísticos o aplicación de herramientas estadísticas en su análisis, por lo tanto, este trabajo no cuenta con resultados inferenciales.

5.3 Otro tipo de resultado estadístico, de acuerdo con la naturaleza del problema de hipótesis

Debido a que el trabajo de investigación es de tipo básica, no cuenta con datos estadísticos o aplicación de herramientas estadísticas en su análisis, por lo tanto, este trabajo no cuenta con resultados estadísticos.

5.4 Resultados del método de gradiente sobre una variedad

Según Burden (2016), el método de gradiente fue estudiado inicialmente por Hestenes y Stiefel, método que se usa para la solución de sistema de ecuaciones lineales. Los pasos son los siguientes:

- Paso 1: Considerar un punto inicial $x = x^0$. Haciendo $k = 0$
- Paso 2: Escoger la dirección de descenso $d^k = -\nabla f(x^k)$

- Paso 3: Realizar la búsqueda lineal seleccionando un paso α_k donde se cumple que $g_k(\alpha_k) = f(x^k + \alpha^k d^k) < f(x^k) = f_k(0)$
- Paso 4: Hacer $x^{(k+1)} = x^k + \alpha_k d^k$
- Paso 5: Haciendo la prueba de convergencia donde $\|\nabla f(x^k)\| < \varepsilon$, si converge se detiene el método. Caso contrario se hace $k = k + 1$ y se regresa al paso dos.

En esta investigación se analizó la convergencia, la cual se usó para comprender la convergencia global sobre una variedad Riemanniana, usando la regla de Armijo en forma general, considerando el problema de optimización no lineal

$$\min_{x \in M} f(x) \tag{V.1}$$

Donde la función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ es una función de clase C^1 y M es una variedad Riemanniana completa.

Para generalizar el método de Gradiente de los espacios euclidianos a variedades con retracciones se analiza la siguiente expresión, donde un punto $x_0 \in T_x^k M$ se tiene

$$x_{k+1} = R_{x_k}(t_k \eta_k) \tag{V.2}$$

Donde $\eta_k \in T_{x_k} M$, representando el espacio tangente a M en el punto x_k , t_k es un escalar que sigue la dirección de la búsqueda de Armijo, generando la sucesión x_{k+1} y R es la retracción definida por Absil (2008).

Sabiendo que el método de gradiente sobre una variedad genera una sucesión de puntos $\{x^k\}$ dados por:

$$x^0 \in M \quad (V.3)$$

$$x^{k+1} = \exp_{x^k}(-t_k \text{grad } f(x^k)) \quad (V.4)$$

Donde \exp_{x^k} es una aplicación exponencial en el punto x^k , t_k es parámetro positivo, $-\text{grad } f(x)$ es la gradiente, esto es señalado por Papa (2008). En forma particular si $M = \mathbb{R}^n$ se puede obtener una expresión equivalente a (V.4)

$$x^{k+1} = x^k - t_k \nabla f(x^k)$$

De esta forma se generaliza el método de máximo de máximo descenso en variedades a el máximo descenso en \mathbb{R}^n .

Hay muchas formas de escoger el parámetro t_k , lo cual su elección dependerá de método computacional, mostrando así los métodos más usados:

Método A: Gradiente con búsqueda exacta

1. Dado x^k , calcular el gradiente sobre el plano $T_{x^k}M$
2. Determinar la geodésica $\gamma(t), t \geq 0$, de la variedad M , cumpliendo así que $\gamma(0) = x^k$ y $\gamma'(0) = -\text{grad } f(x^k)$
3. Minimizar $f(\gamma(t)), t \geq 0$ obteniendo el parámetro y definiendo

$$x^{k+1} = \gamma(t_k)$$

Método B: Gradiente con Regla de Armijo

1. Dado x^k , calcular el gradiente sobre el plano $T_{x^k}M$
2. Determinar la geodésica $\gamma(t), t \geq 0$, de la variedad M , cumpliendo así que $\gamma(0) = x^k$ y $\gamma'(0) = -\text{grad } f(x^k)$
3. Haciendo

$$t_k = 2^{-i_k}$$

Donde i_k es el menor entero positivo y $\alpha \in \left(0, \frac{1}{2}\right)$ cumpliendo que

$$f(\gamma(t_k)) \leq f(x^k) - \alpha t_k \|\text{grad } f(x^k)\|^2$$

Definición 34. Una función $f: M \rightarrow \mathbb{R}$ es llamada gradiente Lipschitziana con constante Γ si $\forall p, q \in M$ y la curva geodésica $\gamma: [0, a] \rightarrow M$ cumpliendo que $\gamma(0) = p$ y $\gamma(a) = q$ verificando que

$$\|\text{grad } f(\gamma(t)) - P_{\gamma(t)} \text{grad } f(p)\| \leq \Gamma L(t)$$

Para todo $t \in [0, a]$, donde $P_{\gamma(t)}$ es el transporte paralelo de $\gamma(0) = p$ a $\gamma(t)$

Método C: Gradiente con pasos fijos

1. Dado x^k , calcular el gradiente sobre el plano $T_{x^k}M$
2. Determinar la geodésica $\gamma(t_k), t_k \geq 0$, de la variedad M , cumpliendo así que $\gamma(0) = x^k$ y $\gamma'(0) = -\text{grad } f(x^k)$
3. Dados $\delta_1, \delta_2 > 0$ tales que

$$\delta_1 \Gamma + \delta_2 > 1$$

Donde Γ es la constante Lipschitziana asociada al campo gradiente de f , escogiendo así al parámetro

$$t_k \in \left(\delta_1, \frac{2}{\Gamma} (1 - \delta_2) \right)$$

5.5 Resultados sobre el análisis de convergencia del método gradiente

Resolviendo el problema de optimización:

$$(p) \quad \min_{x \in M} f(x)$$

Donde M es la variedad Riemanniana conexa, completa de dimensión finita y $f: M \rightarrow \mathbb{R}$ es continuamente diferenciable y cuasiconvexa.

La siguiente hipótesis nos ayuda a analizar la convergencia a partir de la generación de sucesiones convergentes en un espacio no vacío.

Hipótesis A1.

El conjunto de puntos óptimos del problema de optimización es llamado X^* , se denota el valor óptimo del problema de optimización por f^* . Definiendo el conjunto

$$U = \left\{ x \in M : f(x) \leq \inf_k f(x^k) \right\}$$

Es necesario introducir una definición importante la cual consiste en la sucesión cuasi Fejér insertado por Ermoliev (1973) y desarrollado por Svaiter (2011) y Burachik (1995) sobre un espacio métrico.

Definición 35. Sea (X, d) un espacio métrico completo. Una sucesión $\{y_k\}, k \geq 0$ de X es cuasi Fejér convergente al conjunto $U \subset X$, donde para cada $u \in U$ existe una sucesión $\{\epsilon_k\} \in \mathbb{R}$ tal que $\epsilon_k \geq 0, \sum_{k=0}^{+\infty} \epsilon_k < \infty$ y además $d^2(y^{k+1}, u) \leq d^2(y^k, u) + \epsilon_k$.

Teorema 14. En un espacio métrico completo (X, d) , si $\{y^k\}$ es Cuasi Fejér convergente para un conjunto $U \subseteq X$, entonces y^k es limitada. Si además un punto de acumulación \bar{y} de $\{y^k\}$ pertenece a U . Entonces $\{y^k\}$ converge

$$\lim_{k \rightarrow \infty} y^k = \bar{y}$$

Demostración: Sea $u \in U$ un valor fijo y usando la desigualdad

$$d^2(y^{k+1}, u) \leq d^2(y^k, u) + \epsilon_k$$

Sumando telescópicamente desde $k = 0$ hasta $n - 1$

$$d^2(y^n, u) \leq d^2(y^0, u) + \sum_{k=0}^{n-1} \varepsilon_k \leq d^2(y^0, u) + \sum_{k=0}^{\infty} \varepsilon_k$$

El lado derecho es infinito, por lo tanto $\{y^k\}$ es acotada

Analizando la convergencia, sea $\bar{y} \in U$ un punto de acumulación de $\{y^k\}$ cómo es acotada y por Bolzano Weierstrass existe una subsucesión convergente $y^{k_j} \rightarrow \bar{y}$. De la desigualdad anterior tenemos

$$d^2(y^{k+1}, \bar{y}) \leq d^2(y^k, \bar{y}) + \varepsilon_k$$

Si consideramos $a_k = d^2(y^k, \bar{y})$ con un punto de acumulación en \bar{y} , el cual cumple que $a_{k+1} \leq a_k + \varepsilon_k$, $\sum \varepsilon_k < \infty$ y además $a_{k_j} \rightarrow 0$, demostrando así que $\lim_{k \rightarrow \infty} y^k = \bar{y}$.

Definición 36. Sea un espacio métrico X y $0 < p < \infty$. La sucesión (x_n) en X es cuasi Fejér convergente a $\Omega \subset X$. Si para cada $x^* \in \Omega$, existe una sucesión no negativa (p_n) tal que $d(x^*, x_n)^p \leq d(x^*, x_{n-1})^p + p_n$, $n = 1, 2, \dots$

El Lema que se muestra a continuación es de suma importancia ya que muestra que la sucesión generada por el método de máximo descenso es cuasi Fejér convergente a U

Lema 1. Sea $f: M \rightarrow \mathbb{R}$ una función continuamente diferenciable y cuasiconvexa en una variedad Riemanniana conexa, completa y de dimensión finita con curvatura seccional no negativa, entonces

$$d^2(x^{k+1}, x) \leq d^2(x^k, x) + t_k^2 \|\text{grad} \cdot f(x^k)\|^2$$

Para todo $x \in U$ y todo $t_k > 0$.

Demostración: Sea $x \in U$ un valor arbitrario. Sea la curva $\gamma_1: [0, l_1] \rightarrow M$ la geodésica minimal que une x^k y x con $\gamma(0) = x^k$. $\|\gamma'(0)\| = 1$ además $\gamma_2: [0, 1] \rightarrow M$ es una geodésica que une x^k y x^{k+1} , esto es $\gamma_2(0) = x^k$,

$\gamma_2(1) = x^{k+1}$ con $\gamma_2'(0) = -t_k \text{grad } f(x^k)$. Por propiedad de homogeneidad de geodésicas, donde γ_2 es parametrizada tal que

$$\gamma_2: [0, t_k \|\text{grad } \cdot f(x^k)\|] \rightarrow M$$

Tal que $\gamma_2(t_k \|\text{grad } \cdot f(x^k)\|) = x^{k+1}$ y teniendo así

$$\|\gamma_2'(0)\| = 1$$

Del teorema 14 se tiene

$$d^2(x^{k+1}, x) \leq d^2(x^k, x) + t_k^2 \|\text{grad } f(x^k)\|^2 + 2t_k d(x^k, x) \langle \text{grad } f(x^k), \gamma_2'(0) \rangle$$

Como f es cuasiconvexa y $f(x) \leq f(x^k)$ del teorema 12 se tiene

$$\langle \text{grad } f(x^k), \gamma_2'(0) \rangle \leq 0$$

Usando el resultado de la desigualdad, se obtiene la demostración.

A continuación se tomará los resultados del algoritmo generalizado de Kiwiel (1996) generalizando así el espacio euclidiano

Método de Armijo generalizado

En 1966, Armijo publicó su trabajo titulado “Minimization of functions having Lipschitz continuous first partial derivatives”, en el cual estableció la convergencia del método del gradiente. Desde entonces, sus aportes han sido ampliamente utilizados como base para desarrollar mejoras y extensiones teóricas y computacionales. En el marco de nuestro estudio, implementamos el método del gradiente sobre variedades Riemannianas, empleando una versión generalizada de la regla de Armijo.

En esta subsección, demostramos la convergencia global del método en el contexto cuasiconvexo. Estos resultados amplían los trabajos de Kiwiel y Murty (1996) al ámbito de las variedades Riemannianas, y a su vez extienden los aportes previos sobre convergencia en el caso convexo realizados por Burachik et al. (1995) y Cruz Neto (1998).

El método del gradiente con regla de Armijo genera una sucesión de puntos $\{x_k\}$, definida según las expresiones (V.3)-(V.4), cumpliendo con ciertas hipótesis que se detallan a continuación.

Hipótesis A2.

Sea $\phi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ una afirmación tal que:

$$\mathbf{A2.1} \exists \alpha \in (0,1), \tau_\alpha > 0, \text{ tal que } \forall t \in (0, \tau_\alpha]: \phi(t) \leq \alpha t$$

$$\mathbf{A2.2} \exists \beta > 0, \tau_\beta \in (0, +\infty], \text{ tal que } \forall t \in (0, \tau_\beta] \cap \mathbb{R}: \phi(t) \leq \beta t^2$$

$$\mathbf{A2.3} \forall k, f(x^{k+1}) \leq f(x^k) - \phi(t_k) \|\eta_k f(x^k)\|^2 \text{ y } 0 < t_k \leq \tau_\beta$$

$$\mathbf{A2.4} \exists \gamma > 1, \tau_\gamma > 0, \text{ tal que } \forall k: t_k \geq \tau_\gamma \text{ o existe } \bar{t}_k \in [t_k, \gamma t_k] \text{ donde } \bar{t}_k: f\left(R_{x^k}\left(-\bar{t}_k \eta_k f(x^k)\right)\right) \geq f(x^k) - \phi(\bar{t}_k) \|\eta_k f(x^k)\|^2$$

Observación 15. Se puede observar que la hipótesis A2 satisface la regla de Armijo para los valores $\phi(t) = \alpha t, \beta = \alpha, \gamma = 2$ y $\tau_\alpha = \tau_\beta = \tau_\gamma = 1$. Además la hipótesis también satisface por el método de gradiente con pasos fijos y generalizo en variedades Riemannianas. Esto ocurre ya que dado δ_1 y δ_2 tal que $\delta_1 \Gamma + \delta_2 < 1$, donde Γ es la constante de Lipschitz asociado a la gradiente, en particular seleccionar

$$t_k \in \left(\delta_1, \frac{2}{\Gamma} (1 - \delta_2) \right)$$

Se define $\phi(t) = \beta t^2$, con $\beta = \frac{\Gamma \delta_2}{2(1-\delta_2)}, \tau_\gamma = \delta_1, \tau_\beta = \frac{2}{\Gamma} (1 - \delta_2), \alpha \in (0,1)$ arbitrariamente y $\tau_\alpha = \alpha / \beta$, de esta forma se garantiza la hipótesis A2.

Proposición 16. Sea $f: M \rightarrow \mathbb{R}$ una función continuamente diferenciable y cuasiconvexa. Suponga que la hipótesis A1 y A2 cumplen. Entonces la sucesión $\{x^k\}$ generado por el método del gradiente con la regla de Armijo es casi Fejér convergente a U .

Demostración: De las hipótesis A2.2 y A2.3 se tiene

$$\beta t_k^2 \|\text{grad} f(x^k)\|^2 \leq f(x^k) - f(x^{k+1}) \quad (V.5)$$

Implicando así

$$\sum_{k=0}^{\infty} t_k^2 \|\text{grad} f(x^k)\|^2 \leq \frac{f(x^0) - f^*}{\beta} < +\infty$$

De Lema 1 y la definición 35 se tiene el resultado.

Teorema 15. Sea $f: M \rightarrow \mathbb{R}$ una función continuamente diferenciable y cuasiconvexa. Suponga que las hipótesis A1 y A2 cumplen. Entonces la sucesión $\{x^k\}$ generado por el método del gradiente con la regla de Armijo es convergente. Además, converge para un punto estacionario ósea un punto \bar{x} donde $\text{grad} f(\bar{x}) = 0$.

Demostración: De la proposición anterior, se tiene que $\{x^k\}$ es cuasi Fejér convergente en U , por lo tanto por el teorema 14 la sucesión $\{x^k\}$ es limitada. Por lo tanto, existe \bar{x} y una subsucesión $\{x^{k_j}\}$ de $\{x^k\}$ que converge a \bar{x} . Como f es continua se obtiene

$$\lim_{j \rightarrow \infty} f(x^{k_j}) = f(\bar{x})$$

Debido a que $\{f(x^k)\}$ es una sucesión no creciente, con una subsucesión que converge para $f(\bar{x})$, se tiene que para toda sucesión

$$f(\bar{x}) \leq f(x^k), \forall k \in \mathbb{N}$$

Implicando así que $\bar{x} \in U$. Por el teorema 14, se puede concluir que $\{x^k\}$ converge para \bar{x} . Por último, se puede probar que $\text{grad} f(\bar{x}) = 0$. Por el método de contradicción, suponemos que $\text{grad} f(\bar{x}) \neq 0$.

Podemos obtener que $\text{grad} f(x^k) \rightarrow \text{grad} f(\bar{x}) \neq 0$ y $f(x^k) \rightarrow f(\bar{x})$

De la expresión (V.5) se puede obtener

$$\lim_{k \rightarrow \infty} t_k = 0 \quad (V.6)$$

Por otro lado, usando A2.4 y A2.1, se tiene que cuando k es suficientemente grande

$$f\left(R_{x^k}\left(-\bar{t}_k \eta_k f(x^k)\right)\right) - f(x^k) \geq -\alpha \bar{t}_k \|\eta_k f(x^k)\|^2 \quad (V.7)$$

Además, por el teorema del valor medio, para cada k , $\exists t_k^* \in [0, \bar{t}_k]$ tal que

$$-\left\langle \eta_k f\left(R_{x^k}\left(-t_k^* \eta_k f(x^k)\right)\right), P_{\gamma_k, 0, t_k^*} \eta_k f(x^k) \right\rangle \geq -\alpha \|\text{grad } f(x^k)\|^2$$

Donde $P_{\gamma_k, 0, t_k^*}$ es el transporte paralelo sobre la geodésica γ_k tal que $\gamma_k(0) = x^k$ y $\gamma_k'(0) = -\eta_k f(x^k)$. Por la expresión (V.6) y la hipótesis A2.4

Se obtiene que:

$$\lim_{k \rightarrow \infty} t_k^* = 0$$

Haciendo que $k \rightarrow +\infty$ por la desigualdad anterior y tomando en cuenta que la gradiente de función f es continua, además la retracción R y el transporte paralelo también lo es, se tiene que $1 \leq \alpha$, lo que contradice A2.1, obteniendo la conclusión que $\eta_k f(\bar{x}) = 0$

V. DISCUSIÓN DE RESULTADOS

6.1 Contrastación de la hipótesis

1. Se logra obtener como resultado a partir del método de búsqueda de línea definido en el algoritmo anterior. Estas afirmaciones están inspiradas en la teoría convencional. La geometría Riemanniana obtienen como resultados la demostración de la convergencia usando retracciones con la finalidad de minimizar funciones cuasiconvexas con la teoría estudiada en la presente investigación.
2. Dentro de la demostración se insertó el concepto de Retracción como una extensión de aplicaciones que usan las métricas específicas tales como lo usa Quispe (2008) p.22. En general, se tomó sucesiones de forma que sean convergentes para determinar la minimización de funciones cuasiconvexas mediante la Gradiente como método de solución.
3. Tal cual como se demostró la convergencia, la presente investigación usó bondades de las funciones cuasiconvexas, lo cual nos permite obtener puntos óptimos que sean útiles para la toma de situaciones ante problemas contextualizados, donde reemplazamos las geodésicas por retracciones del método del Gradiente y en variedades Riemanniana se logra minimizar funciones cuasiconvexas.
4. El uso de métricas específicas Riemannianas favorece considerablemente la transformación de funciones no convexas a cuasiconvexas sobre estas variedades Riemanniana, esta afirmación se puede demostrar solo considerando la métrica clásica euclidiana.

6.2 Contrastación de los resultados con estudios similares

En Absil y Malick (2012), unos de sus resultados son similares al obtenido en la presente investigación, con la diferencia que este autor muestra la convergencia débil, esto se debe, a que este autor trabaja con funciones en general, por lo cual, diversos autores pueden hacer una generalización sobre este planteamiento en situaciones específicas, el resultado obtenido fue el siguiente.

Teorema 16. Sea $\{x_k\}$ una sucesión infinita de iteraciones generada por el algoritmo 1. Entonces todo punto de acumulación $\{x_k\}$ es punto crítico de la función costo f .

Al comparar este resultado con el obtenido en la investigación se puede ver que, en la hipótesis del teorema anterior no se denotan condiciones sobre esta función costo, a diferencia de nuestra investigación, usamos las funciones cuasi convexas. Otro resultado es el siguiente corolario

Corolario 3. Sea $\{x_k\}$ una sucesión infinita de sucesiones generadas por el algoritmo. Supóngase que el conjunto de nivel $L = \{x \in M: f(x) \leq f(x_0)\}$ es compacto. Entonces

$$\lim_{k \rightarrow +\infty} \|\text{grad } f(x_k)\| = 0$$

Lo cual no realiza la búsqueda lineal de Armijo a diferencia del resultado de esta investigación.

6.3 Responsabilidad ética

La presente investigación tiene como responsabilidad ética motivar e incentivar la investigación, donde se brinda un aporte para que pueda ser usado en próximas investigaciones. Se acredita que la investigación es legítima, usando referencias bibliográficas y usando de forma regular los textos.

VI. CONCLUSIONES

1. Se tienen muchas ventajas positivas para la generalización de modelos en la optimización matemática, estos modelos son tratados adecuadamente en funciones cuasiconvexas. En la presente investigación se ha podido obtener la convergencia sobre un punto crítico o candidato para ser punto óptimo, el cual, aplicándose las retracciones se ha podido llegar a su introducción y clasificación de dicho punto.
2. Esta investigación muestra que el método para problemas de optimización se puede extender en una variedad Riemanniana completa de curvatura seccional no negativa, sobre estas se han definido las retracciones, donde a partir de su métrica se llegó a la minimización de la función cuasiconvexa.
3. Los problemas con funciones no convexas se pueden transformar en funciones convexas, a partir de definir una métrica donde se puedan encontrar puntos candidatos a ser óptimos, donde por medio de las sucesiones de Armijo del método gradiente se pueden generar sucesiones convergentes en la variedad Riemanniana.

VII. RECOMENDACIONES

1. Una de las recomendaciones de la presente investigación a partir de los resultados obtenidos es respecto a la búsqueda de expresiones particulares en las retracciones, están son diferentes a la planteadas por Absil (2012)
2. Se recomienda el uso de propiedades geométricas especialmente en el análisis de la geometría Riemanniana con curvatura no positiva especialmente en la optimización.
3. La implementación en forma computacional de los resultados obtenidos en la presente investigación con la finalidad de obtener comparaciones y ver el modelo optimo o modelos similares.

VIII. REFERENCIAS BIBLIOGRÁFICAS

- [1] Absil, P. A., Mahony, R., & Sepulchre, R. (2008). *Optimization algorithms on matrix manifolds*. Princeton University Press.
- [2] Absil, P. A., & Malick, J. (2012). Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 22(1), 135–158. <https://doi.org/10.1137/100802529>
- [3] Aleksandrov, A. D., Kolmogorov, A. N., & Laurentiev, M. A. (1981). *La matemática: Su contenido, métodos y significado*. Alianza Universidad.
- [4] Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, 16(1), 1–3.
- [5] Bazaraa, M. S., & Jarvis, J. J. (2018). *Programación lineal y flujo en redes*. Limusa.
- [6] Baygorrea, C. N. (2017). *Método do ponto proximal inexato para minimização quasi-convexa em variedades de Hadamard* (Tesis doctoral). Instituto Luiz Coimbra de Pós-Graduação e Pesquisa da Engenharia, Brasil.
- [7] Boothby, W. M. (1986). *An introduction to differentiable manifolds and Riemannian geometry*. Academic Press.
- [8] Bento, G. C., & Cruz Neto, J. X. (2013). A subgradient method for multi-objective optimization on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 159(1), 125–137. <https://doi.org/10.1007/s10957-013-0307-5>

- [9] Bento, G. C., da Cruz Neto, J. X., & Santos, P. S. M. (2013). An inexact steepest descent method for multicriteria optimization on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 159(1), 108–124. <https://doi.org/10.1007/s10957-013-0350-2>
- [10] Burachik, R. S., Grana Drumond, L. M., Iusem, A. N., & Svaiter, B. F. (1995). Full convergence of the steepest descent method with inexact line searches. *Optimization: A Journal of Mathematical Programming and Operations Research*, 32(2), 137–145. <https://doi.org/10.1080/02331939508844139>
- [11] Da Cruz Neto, J. X., & Oliveira, P. R. (1998). Geodesic methods in Riemannian manifolds. *Balkan Journal of Geometry and Its Applications*, 3(2), 89–100.
- [12] Da Cruz Neto, J. X., Oliveira, P. R., Soares, P. A., & Soubeyran, A. (2013). Learning how to play Nash, potential games and alternating minimization method for structure nonconvex problems on Riemannian manifolds. *Journal of Convex Analysis*, 20(2), 395–438.
- [13] do Carmo, M. P. (2008). *Geometria diferencial de curvas e superfícies* (2ª ed.). Rio de Janeiro: Sociedade Brasileira de Matemática.
- [14] Do Carmo, M. P. (1988). *Geometria Riemanniana*. Rio de Janeiro, Brasil: IMPA – Projeto Euclides.
- [15] Ferreira, O. P., & Oliveira, P. R. (1998). Subgradient algorithm on Riemannian manifold. *Journal of Optimization Theory and Applications*, 97(1), 93–104. <https://doi.org/10.1023/A:1022673419987>

- [16] Gabay, D. (1982). Minimizing a differentiable function over a differentiable manifold. *Journal of Optimization Theory and Applications*, 37(2), 177–219. <https://doi.org/10.1007/BF00940052>
- [17] Hicks, N. J. (1966). *Notes on differential geometry*. Van Nostrand Reinhold Company.
- [18] Kiwiel, K. C., Murty, K., (1996). Convergence of the steepest descent method for minimization quasiconvex functions. *JOTA, Journal of Optimization. Theory and Applications*, 89(1), 221-223.
- [19] Lages, L. E. (1960). *Introdução às variedades diferenciáveis*. EMMA.
- [20] Lages, L. E. (2011). *Variedades diferenciáveis*. IMPA.
- [21] Luenberger, D. G. (1973). The gradient projection method along geodesic. *Management Science*, 18(1), 620–631.
- [22] Luenberger, D. G. (1973). *Introduction to linear and nonlinear programming*. Addison-Wesley.
- [23] Papa Quiroz, E. A., Quispe, E. M., & Oliveira, P. R. (2008). Steepest method with a generalized Armijo search for quasiconvex functions on Riemannian manifolds. *Journal of Mathematical Analysis and Applications*, 341(1), 467–477.
<https://doi.org/10.1016/j.jmaa.2007.10.009>
- [24] Papa Quiroz, E. A., Mallma, R. L., & Oliveira, P. R. (2015). An inexact proximal method for quasiconvex minimizations. *European Journal of Operational Research*, 246(3), 721–729.
<https://doi.org/10.1016/j.ejor.2015.05.030>

- [25] Rapcsák, T. (1997). Smooth nonlinear optimization in \mathbb{R}^n . Kluwer Academic Publishers.
- [26] Requena, F. A. (2005). Einstein y las matemáticas. Revista SUMA. <https://revistasuma.es/IMG/pdf/50/007-014.pdf>
- [27] Smith, S. T. (1994). Optimization techniques on Riemannian manifolds. Fields Institute Communications, 3, 113–146.
- [28] Strick, D. S. (1950). Geometría diferencial clásica. Toledo, España.
- [29] Tang, G. J., & Huang, N. J. (2013). An inexact proximal point algorithm for maximal monotone vector fields on Hadamard manifolds. Operations Research Letters, 41(6), 586–591. <https://doi.org/10.1016/j.orl.2013.07.011>
- [30] Udriște, C. (1997). Convex functions and optimization methods on Riemannian manifolds. Dordrecht, Países Bajos: Kluwer Academic Publishers.
- [31] Wen, H. (2013). Optimization algorithms on Riemannian manifolds with applications (Tesis doctoral). Florida State University, College of Arts and Sciences.
- [32] Yang, Y. (2007). Globally convergent optimization algorithms on Riemannian manifolds: Uniform framework for unconstrained and constrained optimization. Journal of Optimization Theory and Applications, 132(2), 245–265. <https://doi.org/10.1007/s10957-006-9103-6>

ANEXO

Matriz de Consistencia

PROBLEMA	OBJETIVO	HIPÓTESIS	METODOLOGÍA Y POBLACIÓN
<p>Problema General</p> <p>¿Cómo obtener la convergencia del Método del Gradiente usando retracciones para minimizar funciones cuasi convexas sobre variedades Riemannianas?</p> <p>Problema Específico</p> <p>1. ¿El uso de retracciones con métricas específicas en el Método de Gradiente viabiliza la minimización de funciones cuasi convexas sobre variedades Riemannianas?</p> <p>2. ¿Las funciones cuasi-convexas permiten encontrar óptimos por medio del Método de Gradiente sobre variedades Riemannianas?</p>	<p>Objetivo General</p> <p>Demostrar la convergencia del Método del Gradiente usando retracciones para minimizar funciones cuasiconvexas sobre variedades Riemannianas.</p> <p>Objetivo Específico</p> <p>1. Usar retracciones con métricas específicas en el Método de Gradiente que viabilice la minimización de funciones cuasiconvexas sobre variedades Riemannianas.</p> <p>2. Usar funciones cuasiconvexas permite encontrar óptimos por medio del Método del Gradiente sobre variedades Riemannianas</p>	<p>Hipótesis General</p> <p>La demostración de la convergencia del Método del Gradiente usando retracciones minimiza funciones cuasi convexas sobre variedades Riemannianas.</p> <p>Hipótesis Específica</p> <p>1. Existe retracciones con métricas específicas en el Método de Gradiente que viabiliza minimizar funciones cuasiconvexas sobre variedades Riemannianas.</p> <p>2. El uso de funciones cuasiconvexas permite encontrar óptimos por medio del Método de Gradiente sobre variedades Riemannianas.</p>	<p>Tipo de investigación</p> <p>Es estrictamente una investigación básica.</p> <p>Diseño de investigación</p> <p>Diseño no experimental, hipotético-deductivo</p> <p>Población y muestra</p> <p>El problema presentado no tiene magnitud factual, es un problema que tiene magnitud formal plasmado por el grado de dificultad que tiene la naturaleza propia del problema para su investigación, por tanto carece de población muestral.</p> <p>Técnicas e instrumentos de recolección de datos</p> <p>Para la realización de nuestro trabajo se reunió bibliografía especializada y recopilación de investigaciones relacionadas con el método de retracción y gradiente.</p> <p>Plan de análisis estadísticos de datos</p> <p>La presente investigación no requiere plan de análisis estadístico de datos.</p>