UNIVERSIDAD NACIONAL DEL CALLAO

ESCUELA DE POSGRADO

UNIDAD DE POSGRADO DE LA FACULTAD DE CIENCIAS ECONÓMICAS

"CONVERGENCIA DEL MÉTODO DEL GRADIENTE USANDO RETRACCIONES PARA MINIMIZAR FUNCIONES CUASI-CONVEXAS SOBRE VARIEDADES RIEMANNIANAS"

TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN INVESTIGACIÓN Y DOCENCIA UNIVERSITARIA

ELSA MARISA QUISPE CÁRDENAS

CALLAO - 2021

PERÚ

UNIVERSIDAD NACIONAL DEL CALLAO

FACULTAD DE CIENCIAS ECONÓMICAS

UNIDAD DE POSGRADO

MAESTRÍA EN INVESTIGACIÓN Y DOCENCIA UNIVERSITARIA

RESOLUCIÓN N° 0139 - 2021-CD-UPG-FCE-UNAC

JURADO DE SUSTENTACIÓN

- PhD. ALMINTOR GIOVANNI TORRES QUIROZ **Presidente**
- Mg. DAVID DAVILA CAJAHUANCA **Secretario**
- Mg. JOSE ASENCION CORBERA CUBAS **Miembro**
- Mg. SEGUNDO AGUSTÍN GARCÍA FLORES **Miembro**
- Mg. LUIS LEONCIO BARBOZA CARAPE **Suplente**

RESOLUCIÓN N° 012 - 2018-UPG-FCE-UNAC

- Mg. RUBEN ORLANDO ARBAÑIL RIVADENEIRA **ASESOR**
- Dr. ERIK ALEX PAPA QUIROZ **CO-ASESOR**

N° DE LIBRO DE ACTA DE SUSTENTACIÓN: LIBRO 2 FOLIO N○. 50.

N° DE ACTA DE SUSTENTACIÓN: 06-2021

FECHA DE APROBACIÓN DE TESIS: 19-11-2021

DEDICATORIA

A Mauro, Cirila, mis padres.

A Juan, mi hermano, siempre listo y con su apoyo.

A Reder mi otro hermano, ingenioso él, a Luzmila mi cuñada.

y con mucho cariño a mi baby Adrianito!!!.

AGRADECIMIENTO

A mis compañeros de aula MIDU-16, Shilda Chacasanay, Ruth Romero, Ana Quispe, Daniela, Alexander Pacaya, Ruth Vargas, Pedro López, Erick Taipe, Alex Espinoza, Ester, Hugo; compartimos gratos momentos de estudio con ellos. Se sumaron Moises, Sonia, y José con quienes también compartimos gratos momentos, gracias compañeros.

No podia dejar de mencionar a un querido amigo, todos del MIDU-16, lo conocemos, muy solidario, empeñoso, me refiero a Dennis Medrano, amigo, gracias por tu amistad especial, sigue adelante amiguito. **......Compañeros de mi promoción, muchas gracias por su amistad.**

Mis gracias y con todo respeto a los señores Administrativos de la FCE POSGRADO: El señor Mendoza, la Sta. Lilia, la Sra Norma, Sta Tania, Sta Mary, por sus atenciones, muchas gracias.

Mi agradecimiento de manera muy especial a los profesores de la Universidad Nacional del Callao:

Dr. Colonibol Torres Bardales, de vasto conocimiento y carácter estricto.

Dra. Katia Vigo Ingar. Exigente con su ciencia, muy académica.

PHd. Almintor Torres Quiroz, de amplio carácter, cumplido, de exquisito saber.

Dr. Juan Nunura Chully, un experto de su clase, un gestor.

A ellos mis respetos, reconocimiento, mis sinceros agradecimientos, por su dedicación, profesionalismo y entrega de conocimientos. **Mis gracias infinitas**.

Finalmente, también muchas gracias a mis asesores, el Dr. Erik Papa Quiroz y Mg. Rubén Arbañil Rivadeneira, y especialmente. a los desarrolladores del proyecto CONVENIO N° 460-INÓVATE PERÚ-BRI-2015 y al CONCYTEC por la financiación de esta tesis.

ÍNDICE

RESUMEN iv

RESUMO v

INTRODUCCIÓN 1

I. PLANTEAMIENTO DE LA INVESTIGACIÓN 3

1.1. Descripción de la realidad problemática 3

1.2. Formulación del problema 8

1.2.1 Problema general 8

1.2.2 Problemas específicos 8

1.3. Objetivos de la investigación 9

1.3.1 Objetivos general 9

1.3.2 Objetivos específicos 9

1.4. Limitantes 9

1.4.1 Teórico 9

1.4.2 Temporal 10

1.4.3 Espacial 11

II. MARCO TEÓRICO 12

2.1. Antecedentes 12

2.1.1 Internacional 12

2.1.2 Nacional 20

2.2. Bases teóricas 21

2.2.1 Teoría sobre elementos de geometría Riemanniana 21

2.2.2 Teoría de la Investigación de Operaciones 22

2.2.3 Teoría de los métodos computacionales 24

2.3. Marco conceptual 24

2.3.1 Variedades Riemannianas 24

2.3.1.1 Variedades diferenciables 24

2.3.2 Método del Gradiente sobre variedades diferenciables 81

2.3.2.1 Método del Gradiente 81

2.3.2.2 Retracciones sobre variedades 82

2.3.2.3 Algoritmo con Retracciones 88

2.3.3 Condiciones de Optimalidad 88

2.3.3.1 Existencia de mínimo global 90

2.3.3.2 Caracterización de mínimo local 91

2.3.4 Resultados básicos de convexidad y cuasi-convexidad 94

2.3.4.1 Convexidad en una variedad riemanniana 95

2.3.4.2 Funciones cuasi-convexas y pseudoconvexas 99

2.4. Definiciones de términos básicos 102

2.4.1 Símbolos y Notaciones 102

III. HIPÓTESIS Y VARIABLES **105**

3.1. Hipótesis 105

3.1.1 Hipótesis general 105

3.1.2 Hipótesis específica 105

3.2. Definición conceptual de variables 105

3.3. Operacionalización de variables 106

IV. DISEÑO METODOLÓGICO **107**

4.1. Tipos y diseño de investigación 107
4.1.1 Tipo 107
4.1.2 Diseño de investigación 108
4.2. Método de investigación 109
4.3. Población y muestra 110
4.4. Lugar de estudio y periodo de desarrollo 110
4.5. Técnicas e instrumentos para la recolección de la información 111
4.6. Análisis y procesamiento de datos 111

V. RESULTADOS **113**
5.1. Resultados descriptivos 113
5.2. Resultados inferenciales 115
5.2.1 Estudio de la convergencia del método del gradiente 115

VI. DISCUSIÓN DE RESULTADOS **123**
6.1. Contrastación y demostración de la hipótesis con los resultados 123
6.2. Contrastación de resultados con otros estudios similares 124
6.3. Responsabilidad ética 126

VIICONCLUSIONES **127**

VIIIRECOMENDACIONES **129**

REFERENCIAS BIBLIOGRÁFICA **130**

ANEXOS **134**

RESUMEN

Esta investigación, **Convergencia del método del gradiente usando retracciones para minimizar funciones cuasi-convexa sobre variedades Riemanniana** tuvo como finalidad el estudio del comportamiento; en el sentido de convergencia computacional, respecto de las sucesiones numéricas generadas el algoritmo del Gradiente con el objetivo de garantizar puntos candidatos a óptimos, para funciones cuasi-convexas sobre la extensión natural geométrica del espacio euclidiano; la variedad Riemanniana, para ello se usó las llamadas retracciones, siguiendo la propuesta de Absil P. y col. (2012). La naturaleza de esta investigación es básica, teórica, cuyo propósito fue analizar la convergencia del método, a fin de garantizar solución, al menos teóricamente, del problema indicado. En esta investigación se aplicó principalmente el método hipotético-deductivo, a fin de inferir conceptos abstractos e interpretar constructos, propios del problema de investigación, para los cuales se utilizaron fichas de investigación para recopilación de información relevante; los que fueron interpretados y analizados, aplicando la lógica matemática, de las distintas organizaciones matemáticas, logrando demostrar las hipótesis del problema de investigación.

Palabras clave: Variedades Riemannianans, Retracciones, funciones cuasi-convexa, método de Armijo, método del Gradiente.

RESUMO

Esta pesquisa, Convergência do método gradiente usando retrações para minimizar funções quase-convexas em variedades Riemannianas, teve como objetivo o estudo do comportamento; no sentido de convergência computacional, no que diz respeito às sequências numéricas geradas pelo algoritmo do Gradiente de forma a garantir pontos candidatos ótimos para funções quase-convexas na extensão geométrica natural do espaço euclidiano; a variedade Riemanniana, para isso foram utilizadas as chamadas retrações, seguindo a proposta de Absil P. et al. (2012). A natureza desta pesquisa é básica, teórica, cujo objetivo foi analisar a convergência do método, de forma a garantir uma solução, pelo menos teoricamente, do problema apontado. Nesta pesquisa, aplicou-se principalmente o método hipotético-dedutivo, com o objetivo de inferir conceitos abstratos e interpretar construtos, tópicos do problema de pesquisa, para os quais foram utilizados arquivos de pesquisa para coletar informações relevantes; aquelas que foram interpretadas e analisadas, aplicando a lógica matemática das diferentes organizações matemáticas, conseguindo demonstrar as hipóteses do problema de investigação.

Palabras chaves: Variedades Riemanniana, retrações, funções cuasi-convexas, método Armijo, método Gradiente.

INTRODUCCIÓN

El uso de elementos y propiedades de geometría Riemanniana es una de las herramientas importantes en el desarrollo de nuevos modelos en diversos campos de la ciencia y en la actualidad de mucho exito en Optimización Matemática. Esta última se ve en el trabajo desarrollado por Luenberger (1972), en la citada referencia se usa el método de descenso geodésico y obtiene la tasa de convergencia del método del Gradiente proyectado para un problema de Optimización mín $f(x)$ con restricciones $h(x) = 0$, donde $f : \mathbb{R}^n \to \mathbb{R}, \;\; h : \mathbb{R}^n \to \mathbb{R}^m, \;\; n > m.$

Algunos autores como Gabay (1982, 2), Udriste (1994), Smith (1994), Rapcsák (1997) continuaron con la imvestigación de modelos de Optimización, sobre variedades riemannianas en relación a la convergencia de los métodos computacionales para dichos problemas. Cuando el problema de minimización, tiene condiciones sobre la función objetivo es el caso de función objetivo convexa en variedades riemannianas, fué revisada y generalizada por Da Cruz Neto, Lima y Oliveira (1998) y usaron búsqueda de Armijo para el método Proximal. Podemos decir, que en esa línea, la novedad de uso de funciones cuasi-convexa y método del Gradiente lo estudia por primera vez Kiwiel y Murty (1996, 1) analizó la convergencia del método. A partir de ese momento dicho problema es extendido a variedades riemannianas, como en Papa Quiroz, Quispe y Oliveira (2008, 1), resultado que nos motivó a desarrollar nuesta investigación. cuya distribución lo describimos a continuación.

En el Capítulo I, se presenta el planteamiento del problema y una descripción de la realidad problemática, dentro del cual establecemos las condiciones y formulación del problema, así como los objetivos de la investigación.

En el capítulo II, descripción del marco teórico de la investigación, antecedentes, nuestras bases teóricas y marco conceptual del problema de investigación, analizamos elementos básicos de geometría riemanniana en relación con la Optimización, se deducen las propiedades geométricas más importantes como curvatura cero, condiciones suficientes para garantizar que la variedad riemanniana sea completa, el método del Gradiente, condiciones de óptimos y resultados de cuasi-convexidad. En el Capítulo III, las hipótesis y variables de investigación.

En el capítulo IV, desarrollo el diseño de la investigación, en cuanto a la metodología que se aplicó para alcanzar los resultado, las técnicas de investigación apropiadas para la pesquisa.

Para finalizar, en el Capítulo V descripción de los resultados hallados, específicamente sobre la convergencia del método del Gradiente, en el Capítulo VI hacemos la discusión de los resultados, la contrastación con las hipótesis y con otros resultados, en el Capítulo VII, las conclusiones de la investigación, y recomendaciones en el Capítulos VIII.

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1. Descripción de la realidad problemática

La Investigación de Operaciones que una ciencia aplicada en muchos ámbitos del desarrollo científico. Si pudiera retroceder a la época, durante la segunda guerra mundial, fue aplicado en forma de estratégias para resistir el ataque enemigo, en la práctica hoy podemos decir que es una especie de teoría de juego; y pasa a utilizarse como estrategia logística en la administración de recursos para su distribución y otros. Debido a la imperiosa necesidad de optimizar recursos y estratégias ante ataques enemigos (en la guerra), Gran Bretaña se vió en la necesidad de reclutar científicos para apoyar el proceso bélico, y dar solución a sus problemas bélicos, los resultado exitosos obtenidos por la Investigación de Operaciones, a la posteridad de la guerra y con el creciente y acelerado desarrollo industrial, condujeron a buscar alguna metodología, estratégia que sea importante para maximizar o minimizar recursos.
G. Dantzing en (1947), profesor de computación, además físico y matemático, desarrolló un método computacional, una de las mayores contribuciones que hizo Dantzing al desarrollo de las ciencias; lo que en el fondo se quería, era resolver un problemas de optimización lineal y propuso el algoritmo que más tarde se llamaría, Simplex, dicho resultado fué usado intensamente por la fuerza aérea de los EE. UU. en el programa SCOOP (Scientific Computation of Optimum Programs). Y más adelante, la actual Investigación de Operaciones se expande a los intereses de las distintas ciencias.

Con el paso del tiempo se puede afirmar que, existieron muchos problemas en las ciencias e ingeniería orientados por ejemplo a optimizar recursos, como el de maximizar o minimizar inversiones, ganancias, utilidades, u orientado para otras ciencias; en ciencias de la computación, el cual requiere de modelos matemáticos para estudiar por ejemplo la complejidad del sistema, que en general son no lineales, es decir; en términos generales, a través de modelos matemáticos generalizados podemos hacer la toma de desiciones, y resolver problemas que demanda la sociedad.

Se puede, citar a las diferentes áreas de aplicación y que ahora será específicamente es en referencia a la Optimización Matemática, pero también tenemos a por ejemplo; Teoría de control, Teoría de la demanda, en Teoría económica, de aproximación, en ciencias Ambientales, Ciencias de la Computación y todas las áreas con problemas que son modelados por ecuaciones matemáticas, cuya solución se enmarca, en el área de la optimización matemática, esta a su vez, utiliza distintas metodologías computacionales de aproximación en cuanto a su solución y, de acuerdo a las características del modelo, el investigador elige un método apropiado para resolverlo.

Ahora, específicamente al exponer el modelo matemático del cual se ha abordado anteriormente en forma descrptiva; un problema de Optimización clásico, y tiene la siguiente expresión matemática:

$$Minf(x), \qquad x \in \mathbb{R}^n \tag{I.1}$$

el modelo (I.1) es llamado Problema de Programación no Lineal sin restricciones definido sobre el espacio euclidiano n-dimensional $\mathbb{R}^n$.

Una expresión generalizada del modelo anterior, es el que se plantea en Quispe C. (2008), quien define al problema anterior, sobre una variedad Riemanniana y al que también nosotros denotamos con $\mathcal{M}$, entendida como la extensión natural geométrica del espacio euclidiano $n-$dimensional, que se desarrollado y analizado, conceptualizado y esquematizado en las posteriores secciones. El problema al cual nos referimos es el siguiente:

$$Minf(x), \qquad x \in \mathcal{M} \tag{I.2}$$

M es una variedad riemanniana y f una función definida sobre $\mathcal{M}$.

Note la diferencia, en I.1, x pertenece al espacio n-dimensional, $\mathbb{R}^n$ mientras que en I.2 la variable x pertenece a la variedad Riemanniana $\mathcal{M}$. Aquí se analizará cuál es su comportamiento, respecto de sus máximos o mínimos, si los tuviere, mediante el análisis de la convergencia del método o algoritmo del Gradiente.

El problema (I.2), es un modelo que pertenece a la ciencias matemáticas, en general es una ciencia básica; y específicamente pertenece a la línea de investigación de Análisis numérico y Computación Matemática; muy en particular a la Optimización Matemática, y ésta investigación la desarrollamos sobre variedades Riemannianas que como ya dijimos se verá en detalle más adelante.

Los estudios de los modelos matemáticos generalizados son importantes porque, permiten clasificar los problemas y plantear metodologías que permitan establecer una mejor solución y estos, generalizarlos; por ejemplo, desde el punto de vista teórico o respecto al coste computacional de un modelo de máximos o mínimos. Razón por cual con herramientas de geometría Riemanniana, o variedades Riemannianas pueden construirse problemas de optimización sin restricciones considerando las propiedades intrínsecas de la propia geometría o mejor dicho de la variedad.

Por otro lado, también, se tiene problemas de optimización osea de máximo o mínimo, cuya característica de la funciones objetivo son no convexa, a su vez modelados sobre una variedad Riemanniana; esos problemas se convierten en problemas de optimización convexos cuando se usa apropiadamente ciertas métricas Riemanniana, lo que permite la generalización de los métodos clásicos de optimización matemática (Quispe C., 2008).

En un sentido más generalizado respecto lo anterior, Absil, Mahony y Sepulchre (2008), involucró a las construcciones llamadas retracciones, definidos sobre variedades riemannianas particulares. En suma, podemos decir que, encontramos algunas dificultades respecto ciertas estratégias metodológicas, cuando se usa la aproximación de sucesiones a la solución del problema (I.2); esto es, las sucesiones que genera el algoritmo del Gradiente, tienen necesariamente involucrado expresiones geodésicas no explicitas lo que complican la generación de sucesiones explicitas dificultando una buena implementación computacional. (Quispe C., 2008). Al respecto, se muestran resultados computacionales y se generaliza el método, respecto la convergencia al solucionar un problema de minimización con funciones cuasi-convexas. (Papa Quiroz y col., 2008, 1).

En esta nvestigación se estudió y analizó, los conceptos teóricos que fundamentan el planteamiento del problema de investigación, entre ellos: teoremas, corolarios y los resultados principales para la demostración de la convergencia del método del Gradiente con retracciones, usanos funciones cuasi-convexas, en el contexto de la variedad Riemanniana; en ese sentido, el algoritmo del método del Gradiente genera una solución al problema, como una sucesión de puntos $x_k \in \mathcal{M}$ dado por:

$$x_{k+1} = \mathcal{R}_{x_k}(t_k \eta_k),$$

donde $\eta_k \in T_{x_k}\mathcal{M}$ y $T_{x_k}\mathcal{M}$ es el espacio tangente de la variedad $\mathcal{M}$ en el punto x_k, t_k es un escalar mediante el cual se aplica la búsqueda de Armijo para determinar el siguiente punto x_{k+1} y $\mathcal{R}$ es una aplicación llamada retracción, ésta se define como una aplicación suave de la unión de espacios tangentes sobre la variedad, esto es, $\mathcal{R} : T\mathcal{M} \to \mathcal{M}$, en adelante se denotará por $\mathcal{R}$, la restricción de $\mathcal{R}$ al espacio tangente $T_x\mathcal{M}$ en el punto x, debe satisfacer condiciones de identificación por ejemplo, la retracción en x del vector cero del espacio tangente, la identificación con el propio espacio. A su vez, se habla también de la identificación canónica del tangente en el vector cero del espacio tangente, con el espacio tangente, llamada aplicación identidad, esta situación será especificado con detalles rigurosos a partir del capítulo 2.3.

La aplicación a la cual nos referimos; retracción, es una expresión más general respecto las aplicaciones usadas, vistos en (Absil & Malick, 2012). También se tiene como ejemplo, la aplicación exponencial que permite hablar sobre convergencia para el problema de optimización y depende de la búsqueda lineal en el algoritmo, ésta búsqueda fué desarrollado por Luenberger (1972) y propuso el uso de búsqueda direccional a lo largo de geodésicas, obtenidas por proyección del vector gradiente en $\mathbb{R}^n$, sobre el espacio tangente del conjunto de restricciones, en este caso estamos hablando de la derivada covariante, que es a su vez una generalización natural de la derivada normal que conocemos.

También podemos referirnos a otras contribuciones en optimización sobre variedades entre ellos, Gabay (1982, 2) en aquella investigación se generaliza el método Cuasi-Newton obteniendo convergencia superlineal. El investigador Udriste (1994), resalta las métricas sobre las propiedades de una variedad riemanniana y sobre todo el estudio de las funciones convexas orientados a sistemas dinámicos y propone también, algoritmos de descenso computacionales. Por otro lado en Papa Quiroz y col. (2008, 1) propusieron la metodología de búsqueda lineal de Armijo generalizado que, a lo largo de geodésicas usa funciones cuasi convexas para aproximar las funciones, y se determina la convergencia.

Finalmente, mencionar a Cruzado Quispe (2014) en cuya tesis, estudia las retracciones y revela algunas implementaciones computacionales de la metodología propuesta por (Absil y col., 2008). cabe resaltar que en dicha tesis, el autor no usa funciones cuasi-convexas.

Con la información revisada, estamos en capacidad de estudiar, analizar y responder las interrogantes del problema general y específicos, y así tengamos formulado nuestro problema de investigación.

1.2. Formulación del problema

1.2.1. Problema general

De la referencia bibliográfica analizada, aún no se ha planteado la minimización de funciones cuasi convexas usando retracciones en el método del Gradiente sobre herramientas de variedades Riemannianas. De acuerdo a la información obtenida, podemos afirmar que, se ha demostrado la convergencia del problema de minimización de funciones cuasi convexas usando el método del Gradiente cuya búsqueda del parámetro longitud de paso es encontrada a través de curvas geodésicas, que implican formulación de aplicaciones exponenciales, véase en Quispe C. (2008), esto nos permite abordar la investigación alrededor de otra aplicación, como son las llamadas retracciones $\mathcal{R}$ debido a su generalidad teórica. Viendo esas fortalezas y oportunidades, se ha logrado plantear y formular el problema de investigación a seguir, y así, realizar el siguiente cuestionamiento:

¿Cómo obtener la convergencia del método del Gradiente usando retracciones para minimizar funciones cuasi convexas sobre variedades Riemannianas?

1.2.2. Problemas específicos

1. ¿El uso de retracciones con métricas específicas en el método de Gradiente viabiliza la minimización de funciones cuasi convexas sobre variedades Riemannianas?.

2. ¿Las funciones cuasi-convexas permiten encontrar puntos candidatos a óptimos por medio del método de Gradiente sobre variedades Riemannianas?.

1.3. Objetivos de la investigación

1.3.1. Objetivos general

Obtener la convergencia del método del Gradiente usando retracciones para minimizar funciones cuasi convexas sobre variedades Riemannianas

1.3.2. Objetivos específicos

1. Usar retracciones con métricas específicas en el método del Gradiente viabilice la minimización de funciones cuasi convexas sobre variedades Riemannianas.

2. Usar funciones cuasi convexas permite encontrar puntos cadidatos a óptimos por medio del método del Gradiente sobre variedades Riemannianas.

1.4. Limitantes

1.4.1. Teórico

El estudio de la geometría no euclidiana permite la generalización del modelo , la problematica y su solución. Es una herramienta matemática que nos permite generalizar métodos computacionales alrededor de la Optimización matemática. Los resultados que se tienen al respecto, son entonces de mucha importancia teórica para la comunidad científica, estudiantes, docentes e investigadores en el contexto de su aplicación en general, debido a la gran cantidad de información sobre variedades Riemannianas, resultados que son más generales de las existentes en la literatura de Optimización matemática que a su vez podemos decir que es una ciencia de múltiples aplicaciones, como en: Economía, Física, ciencias Ambientales, computación, y todas las ingenierías. Así se tiene la clasificación de los problemas de Optimización matemática, esto es, existen una clase de problemas como

son, lineales convexos, no lineales convexos y los cuasi convexos para los cuales se aplica distintos métodos computacionales de minimización sobre geometría no euclidiana. Los resultados de la generalización de la convergencia del método del Gradiente, contribuirán a generar nuevos algoritmos, que, al ser implementados computacionalmente, permiten aplicaciones tecnológicas importantes y de soporte computacional en las área de ciencias e ingeniería, así se puede ver en (Yang, 2007, 2). Otra investigación que contribuye en el aspecto teórico se puede ver también la tesis de (Wen, 2013) que nos pueden permitir modelamiento para problemas específicos, por ejemplo para obtener imágenes de alta calidad visual debido a los componentes teóricos involucrados, como son las herramientas de geometría no euclidiana que lo sustentan.
Finalmente, se puede decir que esta investigación a pesar de las limitantes teóricas por la distante referencia bibliográfica actualizada, contribuye teóricamemte con el saber científico en el área de la optimización matemática pese al limitado material bibliográfico.

1.4.2. Temporal

Por la narutaleza de ésta investigación, la recopilación de información fue limitada a la fecha en que se realizaron las diferentes investigaciones en relación con esta tesis, por ejemplo el investigador Absil y col. (2008) desarrolla la teoría de retracciones sobre variedades orientados a la optimización, del cual hemos tomado también varios resultados de importancia, para la contrastación de las hipótesis. Otro autor que hemos considerado leer es Boothby (1975) quien trata sobre Variedades, que a pesar del tiempo las teorías desarrolladas por el autor, nos sirven para acompañar nuestro marco teórico. Así podemos citar a diferentes autores, uno de ellos, Do Carmo (1988), Do Carmo (2005), cuyas investigaciones han sido de mucho aporte teórico a pesar del tiempo y que hemos tomado con cuidado de siempre citarlos. En resumen en esta línea de investigación pudimos ver que hay pocas investigaciones de reciente publicación lo que dificulta o es una limitante en todo el proceso que ha tenido ésta investigación

1.4.3. Espacial

Al respecto, se puede decir que se ha tenido también limitante espacial, por razones de que la información que hemos logrado obtener y leer para ser analizado, discutido y procesado, ha sido trabajado en muchos casos, en los propios ambientes de la Escuela de Posgrado de la Facultad de Ciencias Económicas de la Universidad Nacional del Callao, usando también la red de conectividad, y sus laboratorios de cómputo, que finalmente me permitío la digitalización en plataforma en línea free, de los resultados para el presente informe.

CAPÍTULO II. MARCO TEÓRICO

2.1. Antecedentes

2.1.1. Internacional

Wen (2013), en su tesis doctoral sobre " Optimization Algorithms on Riemannian Manifoldswith Applications", realizado en el College of Art and Sciences, de la Universidad del Estado de Florida. En su investigación, se puede ver fundamentalmente que tienen resultado sobre convergencia sobre el cual propuso. Los análisis de convergencia del método de región de confianza de rango uno simétrico de Riemann y los métodos de la familia RBroyden asumen una función objetiva suave. Respecto las condiciones para su función objetivo dice, para funciones objetivas continuas de Lipschitz parcialmente uniformes, se muestra que una variación de uno de los métodos,la familia RBroyden, RBFGS, funciona bien empíricamente. Además afirman que se ha demostrado que el método del Gradiente sobre el espacio Riemanniano funciona bien empíricamente tanto para una función objetivo continua y Lipschitz como para una función de objetivo continua que no es de Lipschitz asociada con la importante aplicación de reducción de dimensión no lineal. Se presentan implementaciones eficientes y efectivas para una variedad en Rn, una variedad cociente de una variedad total en Rn y otros. Sobre los hallazgos de la investigación dicen que, los resultados incluyen representaciones y operaciones eficientes de elementos en una variedad, vectores tangentes, operadores lineales, retracciones y transportes vectoriales. Se derivan técnicas novedosas para construir y calcular múltiples tipos de transportes vectoriales.

Baygorrea (2017), en su tesis dortoral sobre "Método do ponto proximal inexacto para minimización quasi-convexa em variedades de Hadamard", tesis presentada en la COP-

PE/UFRJ, del Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia Facultade de Ciencias Exactas y Naturales, Universidade Federal do Rio de Janeiro, Brasil. En dicha investigación sobre una variedad particular, la llamada variedad de Hadamard, ha sido tomado en cuenta para el desarrollo de este estudio y para plantear el problema de investigación sobre dichas variedades. Según los resultado de la indicada investigación, generalizan el algoritmo del Punto Proximal Escalarizado para problemas de optimización multiobjetivo irrestricto con funciones objetivos cuasi-convexos, además hay una necesidad de considerar la condición de Lipschitz, localmente sobre espacio euclidianos, dichos resultados fueron estudiandos por Apolinario et al. [3], para variedades de Hadamard. Indica Baygorrea además que, considerando hipótesis muy razonables, prueban que cualquier punto de acumulación es un punto de crítico Pareto-Clarke para cualquiera de las versiones inexactas. Y, para calcular las iteraciones del algoritmo que han estudiado, el HSIPP, se necesita hacer una estimativa de los vectores del cono normal, lo que puede ser difícil de conseguir y el reto será, dice Baygorrea, superar ese obstáculo que requieran alguna otra técnica de estudio en futuras investigaciones. Para la investigación que se ha realizado, y desde el punto de vista teórico, fué una oportunidad que se tomó en cuenta y considerar sus resultados como recomendación, pues es un trabajo afín respecto de esta investigación.

Luenberger (1972), en su artículo sobre; El Método del Gradiente Proyectado a lo largo de geodésicas, de la revista Management Science, Theory Series. En vínculo a nuestra formulación o planteamiento del problema de investigación, la relación entre los métodos de Optimización Matemática y la geometría Riemanniana tiene una antigüedad aproximadamente del año 1972, específicamente, ese planteamiento se inicia con la investigación hecha por Luenberger (1972), quien usa el método de descenso geodésico y calcula la tasa de convergencia del método del Gradiente proyectado para el problema de minimización con restricciones de igualdad, definidos en $\mathbb{R}^n$. Exactamente los resultados que obtiene D. Luemberger es de proporcionar resultado en relación al método de proyección del Gradiente para resolver problemas de minimización con restricciones. Muestra que la tasa de convergencia asintótica de dicho método, también está dada por la relación de Kantorovich, con la diferencia que los valores α y β se determinan con el Lagrangiano asociado con el problema. Específicamente, si L es el Hessiano del Lagrangian evaluado en la solución, α y β los valores propios más pequeño y más grandes de L cuando están restringidos a la tangente subespacial a la superficie de restricción. Los resultados hallados por Luemberger, son una extensión natural de problemas sin restricciones unidimensional a multidimensional.

Gabay (1982, 2) en su artículo “Minimizing a differentiable function over a differential manifold” de la revista Journal of Optimization Theory and Applications, aplica el método de Gradiente reducido, generaliza el método Cuasi-Newton y muestra que la convergencia óptima del problema de optimización es superlineal.
El resultado exacto que Gabay indica es que para generalizar los métodos de Descenso de optimización sin restricciones al caso restringido, se define intrínsecamente en el campo de gradiente de la función objetivo en la variedad de restricciones y que analiza los métodos de descenso a lo largo de la geodésica, incluida la proyección de gradiente y los métodos de gradiente reducido para opciones especiales de sistemas de coordenadas. En particular, sostiene Gabay, generalizamos los métodos cuasi-Newton y establecemos su convergencia superlineal; demostramos que solo requieren la actualización de una matriz de tamaño reducido. En la práctica, la búsqueda geodésica se aproxima mediante un paso tangente seguido de una restauración de restricciones o mediante una simple búsqueda de arco nuevamente seguida de un paso de restauración. En lo que respecta a nuestra investigación, los resultados de Gabay, si bien son importantes, lo que se verá más adelante y de otros reportes es que la ecuaciones geodésicas en general son implícitas, lo que dificulta o es una limitante en el aspecto teórico de nuestra investigación razon por la cual proponemos el uso de Retracciones, el que se verá más adelante.

Bento, Cruz Neto y Santos (2013) en el artículo “An inexact steepest descentmethod for multicriteria optimization on Riemannian manifolds”, publicada en la revista J. Optim. Theory.Appl, Los mencionados autores realizan investigaciones de alguna manera relacionadas con nuestra investigación; las investigaciones que se realizan en los institutos del Brasil en particular de la Universidad Federal de Piauí, donde más investigaciones sobre variedades Rimannianas y distintas aplicaciones podemos encontrar, lo mismo que ocurre con los autores Bento y Cruz Neto (2013). En el mencionado artículo, los autores muestran una versión inexacta del método de Descenso más pronunciado con la regla de Armijo para la optimización de múltiples criterios en el contexto riemanniano. Bajo suposiciones simples en la función multicriterio, demostran que cada punto de acumulación (si lo hay) satisface condiciones necesarias de primer orden para la optimización de Pareto. Además, afirman que suponiendo que la función multicriterio es cuasi convexa y la variedad Riemanniana tiene efectos de curvatura no negativa, se muestra la convergencia completa de cualquier suceción generada por el método para un punto crítico Pareto. Como ya se vió en (Baygorrea Cusihuallpa, 2017), la anterior investigación se desarrolla usando una particular variedad, denominada variedad de Hadamard.

Smith (1994), publica su artículo “Optimization Techniques on Riemannian Manifolds” de la revista Fields Institute Communications. Sobre elcual se puede decir que tanto el método de Newton como el método del Gradiente Conjugado fueron estudiados por primera vez sobre variedades riemanniana, sobre ellos se analizó la convergencia y se encontró que era de orden cuadrático y superlineal respectivamente; un ejemplo particular que Smith muestra sobre variedades es la aplicación cociente de Rayleigh definido sobre la esfera y que la iteración del cociente de Rayleigh es una aproximación eficiente del método de Newton. Textualmente el autor dice:

> Las técnicas y el análisis presentados en este documento proporcionan nuevos métodos para resolver los problemas de optimización planteados sobre variedades de Riemann. Se ofrece un nuevo punto de vista para la solución de problemas de optimización restricciones. Algunas técnicas clásicas de optimización del espacio euclidiano se generalizan a variedades riemannianos. Se presentan varios algoritmos y se analizan sus propiedades de convergencia empleando la estructura riemanniana de la variedad. Específicamente, se presentan dos algoritmos aparentemente nuevos que se pueden considerar como el método de Newton y el método de gradiente conjugado en las variedades de Riemann, y poseen, respectivamente, convergencia cuadrática y superlineal. Se dan ejemplos de cada método en ciertas variedades de Riemann con los resultados de experimentos numéricos. El cociente de Rayleigh definido en la esfera es un ejemplo, Se muestra que el método de Newton aplicado a esta función converge cúbicamente, y que la iteración del cociente de Rayleigh es una aproximación eficiente del método de Newton. La versión riemanniana del método de gradiente conjugado aplicado a esta función proporciona un nuevo algoritmo para encontrar los vectores propios correspondientes a los valores propios extremos de una matriz simétrica. En un ejemplo similar, se muestra que el método de Newton aplicado a la suma de los cuadrados de las entradas fuera de la diagonal de una matriz simétrica converge cúbicamente.(Smith, 1994, p. 113).

Como puede verse, en dicho artículo, el investigador encuentra importantes técnicas sobre los métodos de optimización en variedades riemannianas que han dado lugar a interesantes generalizaciones sobre la convergencia y aplicación en variedades.

Burachik, Graña Drummond, Iusem y Svaiter (1996) en el artículo “Full Convergence of the Steepest Descent Method with Inexact Line Searches Optimization”, revista Optimization. Estudiaron y analizaronn el método de Máximo Descenso, con búsqueda lineal Inexacta y son aportes bibliográficos que nos dan luces sobre el análisis del problema de investigación que estamos proponiendo. A diferencia de nuestra invetigación, usamos retracciones además de la variedad de Riemann, los métodos Inexactos generan soluciones del promblema con cierto error de tolerancia, los que al menos teoricamente pueden ser muy bien aprobechados para generalización de diversos métodos.

Armijo (1966), en su articulo "Minimization of functions having Lipschitz continuous first partial derivatives", publicado en el Pacific J. Math. Muestra la convergencia del método del Gradiente con algunas condiciones, y así, se pudo ver el punto de partida que realizaron Kiwiel y Murty (1996, 1), tras publicar el artículo titulado Convergence of the steepest descent method for minimization quasiconvex functions, ellos usan por primera vez la función cuasi-convexa junto con el método de Armijo generalizado y obtiene convergencia del método para el problema de minimización con función objetivo cuasi-convexa, sobre el espacio euclidiano n-dimensional, la misma función que hemos usado para nuestra investigación; es el mencionado autor quien generaliza el método de Armijo usando funciones cuasi-convexas, demuestra el teorema de convergencia general para el método del gradiente y se demuestra bajo ciertas condiciones en las hipótesis del teorema. Demostrando que el método de descenso es el descenso más elevado. Estos algoritmos de descenso convergen bajo algunas hipótesis y permiten la posibilidad de tener tamaño de pasos variable. Consecuentemente, en esta investigación, se hizo similar metodología respecto la convergencia del método, con la diferencia que será cambiada a la geometría, variedades de Riemann.

Papa Quiroz y col. (2008, 1), cuyo artículo "Convergence of the steepest descent method for mi-nimizing quasiconvex functions", publicado en la revista, Journal of Optimization Theory and Applications, resumen todo lo trabajado en la tesis de licenciatura de Quispe

C. (2008) debido al conjunto de resultados de mucha importancia desde el punto de vista teórico, y que son una generalización a variedades riemannianas de la convergencia del método del Gradientes usando funciones cuasi-convexas, comparada con la investigación realizada por (Kiwiel & Murty, 1996, 1).

Yang (2007, 2) en el artículo “Globally convergent optimization algorithms on Riemannian manifolds, realizado en Uniform Framework for Unconstrained and Constrained Optimization1 con su revista Journal of Optimization Theory and Applications, proponen varios métodos globalmente convergente para problemas con restricciones y también sin restricciones en el contexto de variedades riemannianas; el tratamiento que hace Yang es sobre convergencia global de algoritmos sobre variedades riemanniana es importante porque además realiza modelamiento computaciones de distintas aplicaciones.

Absil y Malick (2012) y el artículo Projection-like retractions on matrix manifolds, publicado en el periódico, Journal on Optimization, Society for Industrial and Applied Mathematics, sobre ello el modelamiento computacional de algunas aplicaciones con retracciones sobre el espacio tangente de la variedad riemannian; Absil, demuestra la convergencia para minimizar funciones costo cualesquiera sobre una variedad riemanniana, lo que nos permite plantear una investigación, orientada a obtener la generalización de la coonvergencia del método del Gradiente, pero en este caso, la aproximación debe ser realizada no por geodésica sino a través de aplicaciones llamadas, retracciones, para funciones costo cuasi-convexa, esto, debido a la dificultad que se tiene al tener que usar geodésicas explícitas y que no resultan fáciles de obtener; por tanto afirman que es conveniente definir las retracciones sobre la variedad riemanniana para obtener dicha convergencia del método del gradiente, ésta convergencia es una forma de generalizar lo investigado por (Quispe Cárdenas, Papa Quiroz & Oliveira, 2007, 1).

Cruz Neto y Oliveira (1995) en su trabajo “Geodesic algorithms in Riemannian geometry, publicado en la revista de Balkan Journal of Geometry and Its Applications, estudian las condiciones sobre las funciones, usando las llamadas funciones convexas y respecto la variedad, la condiciona a aquellas de curvatura no negativa, usando el método de Armijo además incluyen una regularización Proximal planteando métodos geodésicos como ruta de aproximación al óptimo. Cabe indicar que éste resultado se une nuestra investigación respecto a la curvatura de la geometría, y el que consideramos para fines aplicaciones prácticas de la teoría que aquí se ha propuesto.

(Papa, Mallma & Oliveira, 2015, 3), en el artícul titulado “An inexact proximal method forquasiconvex minimizations” en la revista European J. of Operational Research, del mismo modo los investigadores brasileños Da Cruz Neto, Lima y Oliveira (2013) sostienen que en los problemas de minimización con restricciones se muestran que si el conjunto de restricciones es una variedad riemanniana de curvatura no positiva y la función objetivo es semicontinua inferior y satisface la propiedad Kurdyka-Lojasiewicz, entonces el algoritmo Proximal alterno en el espacio euclídeo se extiende naturalmente para resolver esa clase de problemas, si bien no aborda el método del Gradiente, la propuesta de dichos investigadores nos puede permitir mayores proposiciones a favor de esta investigación.

Absil y col. (2008) quienes publican su texto titulado “Of Optimization Algorithms on Matrix Manifolds”, en dicho texto, tratan entre otros aspecto la teoría de retracciones y distintos algoritmos para optimización sobre variedades y que sirvió de material bibliográfico, de gran aporte sobre la teoría de retracciones que usamos también para esta investigación.

Rapcsák (1997), en su libro “Smooth nonlinear optimization in Rn”, desarrolla estudios sobre problemas de optimización no lineales, las condiciones de convergencias generales usando funciones cualesquiera, es decir no impone ninguna condición para la maximización o minimización de funciones, con y sin restricciones.

Udriste (1994) en su texto “Convex Functions and Optimization Methods on Riemannian Manifolds”, involucra a las funciones convexas sobre una variedad riemanniana, además estudia también las métricas sobre las variedades, su investigación se generaliza por otros autores, como verá en el desarrollo de este informe.

2.1.2. Nacional

Quispe C. (2008) en su tesis de licenciatura titulada “El método de máximo descenso para funciones cuasi-convexas en variedades riemannianas, presentada en la Facultad de Ciencias Naturales y Matemática de la Universidad Nacional del Callao, en dicha tesis se logra demostrar la convergencia global del método del máximo descenso, para el caso con búsqueda generalizada de Armijo en problemas de minimización con función objetivo cuasi-convexas definidas sobre una variedad riemanniana completa y de curvatura seccional no negativa, los resultados que obtiene generaliza la convergencia del método de espacio euclidiano a espacios como las variedades riemannianas, que es una referencia teoricamente hablando, lo más cercano a la propuesta actual, y es el hecho de involucrar las aplicaciones llamada retracciones para ver el comportamiento del método mediante la convergencia.

Cruzado Quispe (2014) en su tesis de licenciatura “Método de Máximo Descenso Usando Retracciones en variedades riemannianas” también de la Facultad de Ciencias Naturales y Matemática de la Universidad Nacional del Callao. Cabe indicar que en dicha investigación el autor no usa funciones cuasi-convexas, en general realiza fundamentalmente implementación computacional del método usando retracciones sus objetivos son estudiar la aplicación retracción y determinarlo para algunas variedades particulares, también estudiar la extensión del método de máximo descenso usando retracciones en variedades riemannianas, el problema minimizar el cociente de Rayleigh sobre la esfera unitaria. y minimizar la función de Brocket sobre la variedad de Stiefel. En términos generales usa funciones especficas y es la diferencia puntual respecto la investigación que aquí se propone.

2.2. Bases teóricas

Las bases teóricas de la investigacón realizada consiste de, entre otros aspectos del planteamiento de la realidad problemática, sobre modelos matemáticos abstractos en optimización matemática, los cuales fueron estudiados y analizados para plantear una metodología, consistente en la deducción e inducción de los constructos matemáticos fundamentales para solucionar la ecuación o modelo abstracto. Por ejemplo el problema de programación no lineal $PPnL$, del que se quiere hallar un óptimo para funciones cuasi convexas está propuesto en tres grandes campos, estos son: Las construcciones geométricas en variedades Riemannianas, optimización continua y el método del Gradiente; los que se desarrollaron en el marco conceptual de este informe.
Sin embargo, debemos complementar tales constructos para ampliar el marco teŕoico, estos son:

2.2.1. Teoría sobre elementos de geometría Riemanniana

Las nociones de geometría riemanniana fueron introducidas por G. Riemann aproximandamente en 1854, propuesto en su investigación titulada: Sobre las hipótesis que estan en los fundamentos de la geometría. Y afirma que toda colección continua de fenómenos homogéneos puede considerarse como un espacio. Estas ideas dieron origen a lo que hoy se conoce como geometría riemanniana, propuesto por el Nobel en física, Riemann, (Requena Fraile, 2005). Para una mayor precisión, el tópico de la geometría diferencial que consiste en el estudio de las variedades diferenciales y si en estas variedades se introduce una particular forma de medir longitudes de curva o distancia o distancia entre dos puntos, entonces se está hablando de variedades riemannianas, en esta geometría confluyen las ramas de la matemática, como son el álgebra, análisis, topología y la geometría, su abordaje dependera del problema a estudiarse. En las siguientes subsecciones se expone las teorías mas fundamentales en relación a la geometría diferencial y sus elementos como base fundamental de la geometría riemanniana, los elementos a estudiarse son propiamente, la definición y conceptualización de variedades diferenciables, las aplicaciones entre ellas, los espacios tangentes, de gran importancia cono son; geodésicas, curvatura, gradiente y Hessiana.

Respecto las métricas, se ha estudiado las más conocidas en torno a la optimización matemática en particular las métricas diagonales, que permitiran obtener resultados importantes para generalizar algoritmos para minimizar por ejemplo funciones cuasi-convexas. A continuación estudiaremos las variedades diferenciales, construidas a partir de la concepción de una superficie regular.

2.2.2. Teoría de la Investigación de Operaciones

La investigación de operaciones permite diseñar diferentes ecuaciones, llamados por ejemplo, modelo de Programación lineal y no Lineal. En un caso los variables de decisión son lineales y función objetivo lineal también, en el otro, correspondientemente las variables de decisión se expresan como funciones no lineales ya sea en la función objetivo como también las restricciones del modelo dado. Por ejemplo, esta característica particular de los modelos no lineales permite abordar problemas donde existen economías o deseconomías de escala o en general donde los supuestos asociados a la proporcionalidad no se cumplen. En este sentido, el Método del Gradiente (conocido también como método de Cauchy o del descenso más pronunciado) consiste, de un algoritmo específico para la resolución de problemas de Programación no Lineal sin restricciones, perteneciente a la categoría de algoritmos generales de descenso, donde la búsqueda de un mínimo está asociada a la resolución secuencial de una serie de problemas unidimensionales.

Por ejemplo, el siguiente es un Problema de Programación Lineal (PPL)

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

Aquí, f y h son funciones lineales, donde f es llamada función objetivo y h restriciones del problema de programación lineal.
Por lo contrario si se considera el problema anterior donde f y h son funciones no lineales, respectivamente, más aún, se puede hablar de regularidad de superficies:

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

donde U es un abierto de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ y $h : U \longrightarrow \mathbb{R}^m$ son funciones dadas.
Si h es diferenciable y su matriz Jacobiana en el punto x, $J_h(x)$ tiene rango m entonces, el conjunto $\{x \in U : h(x) = 0\}$ es una superficie regular. Como casos particulares se tiene, que los conjuntos $\{x \in \mathbb{R}^n : h(x) = 0\}$ y $\{x \in \mathbb{R}^n : h(x) = 0 \ y \ x > 0\}$ son superficies regulares, vea la proposición (6) cuya verificación lo podemos ver en (Do Carmo, 2005, pp. 69-70). La idea de este ejemplo es que se busque condiciones sobre el programación lineal que cumplan regularidad para establecer vínculo con la geometría diferencial sobre el cual se explica más adelante.

2.2.3. Teoría de los métodos computacionales

Son técnicas de aproximación numérica, analizadas matemáticamente para responder a por ejemplo las bondades de los métodos numéricos, o cuándo es que tenemos un cómputo adecuado de los resultados de aplicación de los algoritmos, Burden y Faires (2011) y que resuelven distintos problemas computacionales de las diversas ciencias, a través del software elegido, entre ellos el Maple, el Fortran, Mathlab, el más usado en la ingeniería, y ultimamente Phyton. Nosotros analizaremos desde el punto de vista teórico abstracto el método del Gradiente del cual el interés particular que en general se estudia, es el aspecto teórico en relación a la convergencia del algortimo.

2.3. Marco conceptual

2.3.1. Variedades Riemannianas

2.3.1.1. Variedades diferenciables

Se muestra a continuación resultados importantes acerca de la geometría diferencial como punto de partida para definir la geometría riemanniana, en la definición formal, para una variedad diferenciable es necesario establecer una superficie regular, sus propiedades y otros resultados, para construir una variedad riemanniana muy semejante localmente hablando, al espacio euclidiano $\mathbb{R}^3$ cuja relación tiene el soporte de la diferenciabilidad. Para formalizar los conceptos estudiados, usamos básicamente la información de Do Carmo (2005) y de quien tomamos algunos resultados importantes, del mismo modo los resultado de (Quispe C., 2008) y (Papa Quiroz y col., 2008, 1).

Una aclaración importante respecto al término diferenciable es que se refiere a que la función es infinitamente diferenciable.

Véase cómo se define formalmente una superficie regular, a la que también llamamos, superficie suave.

Definición 1. *(Superficie regular de $\mathbb{R}^n$). Consideremos un subconjunto S de tal modo que, $S \subset \mathbb{R}^n$, y es, una superficie regular 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 $\mathcal{X} : U \to S \cap V$ tal que:*

1. *$\mathcal{X}$ es diferenciable en U. Esto es:*

$$X(u_1, u_2 \cdots u_k) = (x_1(u_1), x_2(u_2), x_3(u_3), \cdots, x_n(u_k))$$

   Las funciones $x_1(u_1), x_2(u_2), x_3(u_3), \cdots, x_n(u_k)$ tienen derivadas parciales continuas en todas las órdenes de U

2. *$\mathcal{X}$ es homeomorfismo. Pues como $\mathcal{X}$ es continua, entonces podemos decir que la aplicación $\mathcal{X}^{-1} : S \cap V \to U$ es continua.*

3. *Para todo punto $q \in U$, $d\mathcal{X}_q : \mathbb{R}^k \to \mathbb{R}^n$ es inyectiva, donde $d\mathcal{X}_q$ es la diferencial de $\mathcal{X}$ en el punto q.*

Tome en cuenta que $\mathcal{X}$ es una aplicación, parametrización o, sistema de coordenadas locales aplicados en una vecindad del punto p.

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

Demostración. *Ver una demostración análoga como en (Do Carmo, 2005, pp. 69-70).*

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

Demostración. Como $\nabla f(x) \neq 0$ a tener a valor regular, luego aplicando la Proposición (1) tenemos el corolario. ■.

En relación a las superficies regulares, las proposiciones anteriores y los modelos de optimización de funciones podemos exponer un claro ejemplo a continuación:

Ejemplo 1. *Sea el problema de programación lineal (PPL), dado por:*

$$\begin{array}{lll} (PPL) & min & c^T x \\ s.a & & Ax = b \\ & & x \geq 0. \end{array}$$

donde: $b \in \mathbb{R}^m$ *y* $A \in \mathbb{R}^{m\times n}$ *es de rango* $m < n$. *Llamado también el problema Primal.*

Ejemplo 2. *El siguiente el problema dual D del (PPL), dado por:*

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

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* (D) *es el dual de priblema Primal.*

Del problema (P), vemos que, la región de factibilidad es una superficie regular. En efecto: Si $S = \{x \in \mathbb{R}^n_{++} : Ax = b\}$ es el conjunto de las restricciones estrictas del problema (PPL), definiendo la función $F : \mathbb{R}^n_{++} \to \mathbb{R}^m$, por $F(x) = Ax - b$ se tiene que $dF_x = A$, para todo $x \in \mathbb{R}^n_{++}$. Por tener la matriz A rango m y aplicando la Proposición 1 tenemos que $F^{-1}(0) = S$ es una superfície regular de dimensión $n - m$. Análogamente para el problema (D).

La siguiente definición es referente al cambio de parámetros y que gracias a la condición 2 de superficie regular, permite probar que, si queremos definir algunos objetos en términos de una parametrización, ésta no depende de la parametrización tomada sino del conjunto que se tome, de ahi la importancia de la propiedad 2. Este resultados es muy útil para generalizar la definición de superficie regular a variedad diferenciable. En las siguientes definiciones, proposiciones o teoremas se sigue lo propuesto como resultado según (Quispe C., 2008, pp. 9-10).

Definición 2. *(Cambio de parámetros). Sean las aplicaciones* $\mathcal{X} \longrightarrow S$ *y* $\mathcal{Y} \longrightarrow S$ *dos parametrizaciones de* S *en el punto* p *tales que*

$$W = \mathcal{X}(U) \cap \mathcal{Y}(V) \neq \phi.$$

La aplicación $\mathcal{Y}^{-1} o \mathcal{X} : \mathcal{X}^{-1}(W) \longrightarrow \mathcal{Y}^{-1}(W)$ *es llamado cambio de parámetros.*

Proposición 2. *Sea* S *una superficie regular de* $\mathbb{R}^n$ *de dimensión* k*. El cambio de parámetros* $\mathcal{Y}^{-1} o \mathcal{X} : \mathcal{X}^{-1}(W) \longrightarrow \mathcal{Y}^{-1}(W)$ *es un difeomorfismo.*

Demostración.

- $\mathcal{Y}^{-1} \circ \mathcal{X}$ es biyectiva. es sobreyectiva

  En efecto, Sea $q_2 \in \mathcal{Y}^{-1}(W)$ entonces existe $\tilde{p} \in W$ y tal que ahora, como $\tilde{p} = \mathcal{X}(q_1)$ para algun $q_1 \in \mathcal{X}^{-1}(W)$ entonces

  $$\mathcal{Y}^{-1}(\mathcal{X}(q_1)) = \mathcal{Y}^{-1}(\tilde{p}) = q_2$$

  y se tiene la sobreyectividad.

- la inyectividad de $\mathcal{Y}^{-1} \circ \mathcal{X}$ es obvia

- $\mathcal{Y}^{-1} o \mathcal{X}$ es diferenciable en $\mathcal{X}^{-1}(W)$

  Veamos, Sea $q_1 \in \mathcal{X}^{-1}$, entonces existe $p \in W = \mathcal{X}(U) \cap \mathcal{Y}(U)$ tal que

  $$q_1 = \mathcal{X}^{-1}(p).$$

  También, como $p \in (W)$, existe $q_2 \in V$ tal que $q_2 = \mathcal{Y}^{-1}(p)$. Considerando la aplicación proyección $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ definido por:

$$\pi(x) = \pi(x_1, ..., x_k, x_{k+1}, ..., x_n) = (x_1, ..., x_k),$$

y la aplicación $\pi \circ \mathcal{Y} : \mathcal{Y}^{-1}(W) \subset \mathbb{R}^k \to \pi(W) \subset \mathbb{R}^k$ dado por:

$$\pi o \mathcal{Y}(u_1, u_2, ...u_k) = (x_1(u_1, u_2, ...u_k), ..., x_k(u_1, u_2, ...u_k)).$$

De la diferenciabilidad de $\mathcal{Y}$ y π tenemos que $\pi \circ \mathcal{Y}$ es diferenciable en $\mathcal{Y}^{-1}(W)$. Como $q_2 \in \mathcal{Y}^{-1}(W)$, entonces $d\mathcal{Y}_{q_2}$ es inyectiva, luego podemos suponer que

$$det \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_k} \\ \vdots & & \vdots \\ \frac{\partial x_k}{\partial u_1} & \cdots & \frac{\partial x_k}{\partial u_k} \end{bmatrix} = \frac{\partial(x_1, x_2, ..., x_k)}{\partial(u_1, u_2, ..., u_k)} \neq 0.$$

Así,

$$det \left[d(\pi o \mathcal{Y})_{q_2} \right] = \frac{\partial(x_1, x_2, ..., x_k)}{\partial(u_1, u_2, ..., u_k)} \neq 0.$$

dado que $\pi \circ \mathcal{Y}$ es diferenciable en q_2 y $d(\pi \circ \mathcal{Y})_{q_2}$ es un difeomorfismo, entonces por el teorema de la función inversa existen vecindades $\Omega_2 \subset \mathcal{Y}^{-1}(W)$ de q_2 y $\Omega_1 \subset \pi(W)$ de $(\pi \circ \mathcal{Y})(q_2)$ tales que,

$$\pi \circ \mathcal{Y} : \Omega_2 \to \Omega_1$$

es un difeomorfismo.

Luego $(\pi \circ \mathcal{Y})^{-1} : \Omega_1 \longrightarrow \Omega_2$ es diferenciable.

Finalmente, observando que $\mathcal{Y}^{-1} \circ \mathcal{X} = (\pi \circ \mathcal{Y})^{-1}(\pi \circ \mathcal{X})$ tenemos que $\mathcal{Y}^{-1} \circ \mathcal{X}$ es diferenciable para todo $q \in \mathcal{X}^{-1}(W)$ tal que $\pi \circ \mathcal{X}(q) \in \Omega_1$.

Como $\pi \circ \mathcal{X}(q_2) = \pi(\mathcal{X}(q_1)) = \pi(p) = \pi(\mathcal{Y}(q_2)) = (\pi \circ \mathcal{Y})(q) \in \Omega_1$, entonces en particular:

$$\mathcal{Y}^{-1} \circ \mathcal{X} = (\pi \circ \mathcal{Y})^{-1}(\pi \circ \mathcal{X})$$

es diferenciable en q_1, y como q_1 es arbitrario, entonces podemos concluir que $\mathcal{Y}^{-1} \circ \mathcal{X}$ es diferenciable en $\mathcal{X}^{-1}(W)$. $\blacksquare$

Sea $\tilde{p} \in W$ y sean $q_1 = \mathcal{X}^{-1}(p)$ y $q_2 = \mathcal{Y}^{-1}(p)$. Definamos la aplicación proyección $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ y dado $x \in \mathbb{R}^n$ tal que

$$x = (x_1, x_2, ..., x_k, x_{k+1}, ..., x_n)$$

y que por aplicación de π:

$$\pi(x) = \pi(x_1, ..., x_k, x_{k+1}, ..., x_n) = (x_1, ..., x_k)$$

y la aplicación $\pi o \mathcal{Y} : \mathcal{Y}^{-1}(W) \subset \mathbb{R}^k \to \pi(W) \subset \mathbb{R}^k$ definido por:

$$\pi o \mathcal{Y}(u_1, u_2, ...u_k) = (x_1(u_1, u_2, ...u_k), ..., x_k(u_1, u_2, ...u_k)).$$

$\mathcal{Y} o\ \pi$ es diferenciable en $\mathcal{Y}^{-1}(W)$ desde que $\mathcal{Y}$ y π son diferenciables.

Para $q_2 \in \mathcal{Y}^{-1}(W)$ la aplicación $d(\pi o \mathcal{Y})(q_2) : \mathbb{R}^k \to \mathbb{R}^k$ es un isomorfismo Como $d\mathcal{Y}_{q_2}$ es inyectiva vamos suponer que:

$$det \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_k} \\ \frac{\partial x_k}{\partial u_1} & \cdots & \frac{\partial x_k}{\partial u_k} \end{bmatrix} = \frac{\partial(x_1, x_2, ..., x_k)}{\partial(u_1, u_2, ..., u_k)} \neq 0.$$

Luego:

$$det \left[(\pi o \mathcal{Y})_{q_2}\right] = \frac{\partial(x_1, x_2, ..., x_k)}{\partial(u_1, u_2, ..., u_k)} \neq 0.$$

Por el teorema de la función inversa, existen vecindades $\Omega_2 \subset \mathcal{Y}^{-1}(W)$ de q_2 y $\Omega_1 \subset \pi(W)$ de $(\pi o \mathcal{Y})(q_2)$ tales que $\pi o \mathcal{Y} : \Omega_2 \to \Omega_1$ es difeomorfismo.

Observe que $\mathcal{Y}^{-1} \circ \mathcal{X} = (\pi \circ \mathcal{Y})^{-1} \circ \pi \circ \mathcal{X}$

$$: h^{-1}(\Omega_2) \to \Omega_2$$

y por tanto $h = \mathcal{Y}^{-1} o \mathcal{X}$ es diferenciable como la composición de tres aplicaciones diferenciables. En particular $h = \mathcal{Y}^{-1} o \mathcal{X}$ es diferenciable en q_1 y como este punto es arbitrario se tiene que $\mathcal{Y}^{-1} o \mathcal{X}$ es diferenciable en $\mathcal{X}^{-1}(W)$ ■.

Antes de definir una variedad diferenciable, es importante establecer, al menos en $\mathbb{R}^3$ al espacio tangente a una superficie regular, que a decir de la condición 3 de la definición de superficie regular permite establecer un conjunto de vectores tangentes, localmente hablando, a las curvas parametrizadas en un punto dado y constituyen un plano sobre la superficie. Véase la siguiente proposición.

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

Demostración. Para la demostración de esta proposición véase (Do Carmo, 2005, p. 98.) ■

Proposición 4. *Si S_1 y S_2 son superficies regulares y $\varphi : U \subset S_1 \to S_2$ 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 en p.*

Demostración. Por la aplicación del teorema de la función inversa. ■

Devido a la noción de plano tangente se puede hablar de diferencial de una aplicación en la siguiente definición en relación a dos superficies regulares.

Definición 3. *Sean S_1 y S_2 dos superficies regulares y considerando $\phi : V \subset S_1 \to S_2$ una aplicación diferenciable de un conjunto abierto V de S_1 en S_2. Si $p \in V$ se sabe que todo vector tangente $w \in T_pS_1$ es el vector velocidad $\alpha'(0)$ de una curva diferenciable α.*

Teniendo $\alpha : (-\epsilon, \epsilon) \to V$ con $\alpha(0) = p$. La curva $\beta = \varphi \circ \alpha$ es tal que $\beta(0) = \varphi(p)$, y por tanto $\beta'(0)$ es un vector de $T_{\varphi(p)}S_2$ De esta definición tenemos la siguiente proposición.

Proposición 5. *Dado w, un vector $\beta'(0)$ no depende de la elección de α. La aplicación*

$$d\varphi_p : T_pS_1 \to T_{\varphi(p)}S_2$$

definida por $d\varphi_p(w) = \beta'(0)$ es lineal.

Demostración. Análogo a la demostración que puede verse en (Do Carmo, 2005, pp.150-151). ■

En la definición abstracta de variedades es importante establecer los conceptos de cartas y atlas, por la importancia de analizar los objetos relacionados con conjuntos, por ejemplo el conjunto $\mathcal{U}$ que pueden ser estudiados como subconjuntos de $\mathbb{R}^d$. Las definiciones planteadas aquí, se siguen de (Absil, 2012), el que puede leerse a continuación.

Definición 4. *(Carta). Sea $\mathcal{M}$ un conjunto, una carta d-dimensional del conjunto $\mathcal{M}$ es una biyección φ de un subconjunto $\mathcal{U}$ sobre un $S \subset \mathbb{R}^d$, esto es: $\varphi : \mathcal{U} \subset \mathcal{M} \to S \subset \mathbb{R}^d$, para el cual se tiene $\varphi(x)$.*

Se denota a una carta usando el par $(\mathcal{U}, \varphi)$, y siempre que no exista lugar a confusión, podemos escribir solamente φ que indicará la respectiva carta. Los elementos $\varphi(x)$ son las coordenadas en x de la carta φ.

Definición 5. *(Atlas). Un Atlas C^∞ es una colección o familia de cartas $\{(\mathcal{U}_\alpha, \varphi_\alpha) : \alpha \in I\}$ del conjunto $\mathcal{M}$ tal que:*

1. *$\bigcup_{\alpha \in I} \mathcal{U}_\alpha = \mathcal{M}$, reunión de aplicaciones abiertas $\mathcal{U}_\alpha$*
2. *Tomando el par $\alpha, \beta \in I$, las colecciones de cartas $\mathcal{U}_\alpha$, $\mathcal{U}_\beta$ tal que $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$, el conjunto $\varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ y $\varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ son conjuntos abiertos en $\mathbb{R}^d$ y el cambio*

de coordenadas

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \subset \mathbb{R}^d \to \varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \subset \mathbb{R}^d$$

es suave y con inversa también suave, es decir $\varphi_\beta \circ \varphi_\alpha^{-1}$ es un difeomorfismo. Así podemos decir que los elementos de un atlas se superponen suavemente.

Las siguientes definiciones son tomados de (Absil y col., 2008), p. 19.

Definición 6. *Dos atlas $\mathcal{A}_1$ y $\mathcal{A}_2$ son compatibles si $\mathcal{A}_1 \cup \mathcal{A}_2$ es también un atlas, es decir para cada carta $(\mathcal{U}, \varphi)$ en $\mathcal{A}_2$, el conjunto de cartas $\mathcal{A}_1 \cup \{(\mathcal{U}, \varphi)\}$ sigue siendo un atlas.*

Definición 7. *Dado un atlas $\mathcal{A}$, sea $\mathcal{A}^+$ es el conjunto de todas las cartas $(\mathcal{U}, \varphi)$ tal que $\mathcal{A} \cup \{(\mathcal{U}, \varphi)\}$ es también un atlas; el atlas $\mathcal{A}^+$ se llama atlas maximal o atlas completo inducido por el atlas $\mathcal{A}$*

Definición 8. *Un atlas maximal de un conjunto $\mathcal{M}$ es también llamado una estructura diferenciable sobre $\mathcal{M}$.*

Dos atlas son equivalentes si y sólo si son generados por un atlas maximal.

Seguidamente se explica de modo formal una variedad diferenciable a partir de un conjunto $\mathcal{M}$ sobre el cual se cubrirá con aplicaciones cuyas condiciones se explicarán en la siguiente definición a partir de lo planteado por (Do Carmo, 1988).

Definición 9. *(Variedad diferenciable). Sea el conjunto $\mathcal{M}$ y una familia de aplicaciones inyectivas $X_\alpha : U_\alpha \longrightarrow \mathcal{M}$, $\alpha \in I$ (conjunto de parámetros), definidos en abiertos U_α de $\mathbb{R}^n$ en $\mathcal{M}$, $\mathcal{M}$ se llama una variedad diferenciable si se cumplen las siguientes condiciones:*

1. $\mathcal{M} = \bigcup_{\alpha \in I} X_\alpha(U_\alpha)$.

2. *Para todo par $\mathcal{X}_\alpha$, $\mathcal{X}_\beta$ con $\mathcal{X}_\alpha(U_\alpha) \cap \mathcal{X}_\beta(U_\beta) = W \neq \emptyset$, los conjuntos ${\mathcal{X}_\alpha}^{-1}(W)$ y ${\mathcal{X}_\beta}^{-1}(W)$ son abiertos en $\mathbb{R}^n$ y las aplicaciones ${\mathcal{X}_\beta}^{-1} \circ \mathcal{X}_\alpha : {\mathcal{X}_\alpha}^{-1}(W) \longrightarrow {\mathcal{X}_\beta}^{-1}(W)$ son diferenciables.*

Una notación muy usada para la parametrización es, $(U_\alpha, \mathcal{X}_\alpha)$ con $p \in \mathcal{X}_\alpha(U_\alpha)$.
Se puede también definir al conjunto $\{(U_\alpha, \mathcal{X}_\alpha)\}$ como una familia de parametrizaciones, estas son las llamadas sistema de coordenadas, las cuales al satisfacer los items *a).* y *b)* conforma una estructura diferenciable de $\mathcal{M}$. De tal manera que, una variedad M es un conjunto dotado de una estructura diferenciable.
Por otro lado, una variedad diferenciable con estructura diferenciable (llamado también atlas maximal) induce de forma natural una topología en $\mathcal{M}$ definido por:

$A \subset \mathcal{M}$ es abierto en $\mathcal{M}$ si para todo $\alpha \in I$, ${\mathcal{X}_\alpha}^{-1}(A \cap \mathcal{X}_\alpha(U_\alpha))$ es abierto en $\mathbb{R}^n$.

Respecto el item *2* de la Definición 9 podemos decir que la diferenciabilidad de una función real, definida sobre una variedad diferenciable $\mathcal{M}$ no depende de la parametrización elegida.
Seguidamente se prueba que toda superficie regular es una variedad diferenciable. Por la Proposición 2 es que se tiene el siguiente resultado.

Proposición 6. *Toda superficie regular de $\mathbb{R}^n$ de dimensión k es una variedad diferenciable de la misma dimensión.*

Demostración. Sea $\mathcal{M}$ una superficie regular en $\mathbb{R}^n$ de dimensión k, entonces por definición de superficie regular, para cada punto $p_\alpha \in \mathcal{M}$ existe una vecindad V_α de p_α y una aplicación $\mathcal{X}_\alpha : U_\alpha \to V_\alpha \cap \mathcal{M}$ tal que $\mathcal{X}_\alpha$ es diferenciable, homeomorfismo y $\forall q_\alpha \in U_\alpha$, $(d\mathcal{X}_\alpha)_{q_\alpha}$ es inyectiva. En consecuencia la familia de paranetrizaciones $\mathcal{X}_\alpha : U_\alpha \subset \mathbb{R}^n \to M$ son inyectivas.
La condición $\mathcal{M} = \bigcup\limits_{\alpha \in I} \mathcal{X}_\alpha(U_\alpha)$ es facil de verificarse. Para probar la condición 2, de la

definición de variedad diferenciable, consideremos $\mathcal{X}_\alpha$ y $\mathcal{X}_\beta$ un par de parametrizaciones con $\mathcal{X}_\alpha(U_\alpha) \cap \mathcal{X}_\beta(U_\beta) = W \neq \phi$. Por la Proposición 2 las aplicaciones:

$$\mathcal{X}_\beta{}^{-1} o \mathcal{X}_\alpha : \mathcal{X}_\alpha{}^{-1}(W) \rightarrow \mathcal{X}_\beta{}^{-1}(W)$$

son diferenciables, además, $\mathcal{X}_\alpha^{-1}(W)$ son abiertos, pues W es abierto en $\mathbb{R}^n$ desde que la intersección de abiertos de $\mathcal{X}_\alpha(U_\alpha)$ y $\mathcal{X}_\beta(U_\beta)$ son abiertos y por la continuidad de $\mathcal{X}_\alpha$. ■

La proposición 6, es un claro ejemplo de variedad diferenciable, en particular si tenemos $\mathbb{R}^3$. Otro ejemplo de variedad diferenciable y de mucho interés en ser estudiado es, la esferea unitaria S^1 cuya expresión matemática es la siguiente

$$S^1 = \{(x, y) \in \mathbb{R}^2 \text{ tal que: } x^2 + y^2 = 1\}$$

Los siguientes resultados son importantes para la operaciones diferenciables entre variedaded diferenciables y son tomados de (Quispe C., 2008).

Definición 10. *Sea $f : U \subset \mathcal{M} \longrightarrow \mathbb{R}$, donde U es un subconjunto abierto de la variedad diferenciable $\mathcal{M}$. Diremos que f es diferenciable en $p \in U$, si para alguna parametrización $\mathcal{X}_\alpha : U_\alpha \subset \mathbb{R}^n \longrightarrow \mathcal{M}$, con $p \in \mathcal{X}_\alpha(U_\alpha) \subset U$, la función compuesta $f \circ \mathcal{X}_\alpha : U_\alpha \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es diferenciable en $\mathcal{X}_\alpha^{-1}(p)$. Se dice que f es diferenciable en U si es diferenciable en todo punto de U.*

Definición 11. *Una curva sobre una variedad difereciable $\mathcal{M}$ es una función $\gamma : I \rightarrow \mathcal{M}$ donde $I = (-\varepsilon, \varepsilon)$. Diremos que γ es diferenciable en $t_0 \in I$ si para alguna parametrización $\mathcal{X}_\alpha : U_\alpha \subset \mathbb{R}^n \rightarrow \mathcal{M}$ con $\gamma(t_0) \in \mathcal{X}_\alpha(U_\alpha)$, la función compuesta $\beta = \mathcal{X}_\alpha^{-1} \circ \gamma : I \rightarrow U_\alpha$ es diferenciable en t_0, donde $\gamma(I) \subset \mathcal{X}_\alpha(U_\alpha)$. Si γ es diferenciable en todo $t \in I$, diremos que γ es diferenciable en I.*

Si se tiene dos variedades, $\mathcal{M}_1$ y $\mathcal{M}_2$ se puede definir a las aplicaciones diferenciables entre ellas. veáse la siguiente definición.

Definición 12. *Sean $\mathcal{M}_1$ y $\mathcal{M}_2$ variedades diferenciables de dimensión m y n respectivamente. Una aplicación $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$ es diferenciable en $p \in V$, si dados*

$$\mathcal{X}_1 : U_1 \subset \mathbb{R}^n \to \mathcal{M}_1$$

parametrización de $\mathcal{M}_1$ en p y

$$\mathcal{X}_2 : U_2 \subset \mathbb{R}^m \to \mathcal{M}_2$$

parametrización de $\mathcal{M}_2$ en $\varphi(p)$ con $\varphi(\mathcal{X}_1(U_1)) \subset \mathcal{X}_2(U_2)$, la aplicación $\mathcal{X}_2^{-1} \circ \varphi \circ \mathcal{X}_1 : U_1 \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $\mathcal{X}_1^{-1}(p)$.

Esta última aplicación $\mathcal{X}_2^{-1} \circ \varphi \circ \mathcal{X}_1$ es llamada expresión de φ en las parametrizaciones $\mathcal{X}_1$ y $\mathcal{X}_2$. φ es diferenciable en un abierto de $\mathcal{M}_1$ si es diferencible en todos los puntos del abierto.

Otra vez, esta definición de diferenciabilidad entre variedades, no depende de las parametrizaciones elegidas.

En efecto, sean $\mathcal{Y}_1 : Q_1 \to S_1$ y $\mathcal{Y}_2 : Q_2 \to S_2$ otras parametrizaciones de S_1 y S_2 en p y $\varphi(p)$ respectivamente tales que $\varphi(\mathcal{Y}_1(Q_1)) \subset \mathcal{Y}_2(Q_2)$. Entonces

$$\mathcal{Y}_2^{-1} \circ \varphi \circ \mathcal{Y}_1 = (\mathcal{Y}_2^{-1} \circ \mathcal{X}_2) \circ (\mathcal{X}_2^{-1} \circ \varphi \circ \mathcal{X}_1) \circ (\mathcal{X}_1^{-1} \circ \mathcal{Y}_1)$$

Como $(\mathcal{Y}_2^{-1} \circ \mathcal{X}_2)$ e $(\mathcal{X}_1^{-1} \circ \mathcal{Y}_1)$ son diferenciables por ser ellas cambio de parámetros y $\mathcal{X}_2^{-1} \circ \varphi \circ \mathcal{X}_1$ es diferencible por ser φ $\mathcal{Y}_2^{-1} \circ \varphi \circ \mathcal{Y}_1 : Q_1 \subset \mathbb{R}^n \to \mathbb{R}^m$ es también diferenciable.

Observe que, si se tiene $\mathcal{X} : U \to \mathcal{M}$ una parametrización de $\mathcal{M}$ en p entonces $\mathcal{X}^{-1} : \mathcal{X}(U) \subset \mathcal{M} \to \mathbb{R}^n$ es diferenciable.

En efecto, sea $q \in \mathcal{X}(U)$ y $\mathcal{Y} : V \subset \mathbb{R}^n \to S$ una parametrización de S en q, $\mathcal{X}^{-1}$ seran diferenciables en q si la aplicación $Id\, o\mathcal{X}^{-1}\, o\mathcal{Y} : V \subset \mathbb{R}^n \to \mathbb{R}^n$ es diferenciable en $\mathcal{X}^{-1}(q)$ (en el sentido Euclideano). Esto es verdad pues $Id\, o\mathcal{X}^{-1}\, o\mathcal{Y} = \mathcal{X}^{-1}\, o\mathcal{Y}$ es diferenciable en $\mathcal{Y}^{-1}(q)$ por ser cambio de parámetros. Luego $\mathcal{X}^{-1}$ es diferenciable en q y como este punto es arbitrario, $\mathcal{X}^{-1}$ es diferenciable en $\mathcal{X}(U)$.

Definición 13. *(Difeomorfismo entre variedades diferenciables). Sea $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$ una aplicación diferenciable entre dos variedades diferenciables. Decimos que φ es difeomorfismo si φ es bijetiva y φ^{-1} es diferenciable. φ es difeomorfismo local en $p \in \mathcal{M} mathcal M_1$, si existen vecindades U de p y V de $\varphi(p)$ tal que $\varphi : U \to V$ es difeomorfismo.*

Observación 1. *Podemos identificar con $\mathcal{X}(U) \equiv U$, puesto que, cualquier parametrización $\mathcal{X} : U \subset \mathbb{R}^n \to \mathcal{X}(U) \subset \mathcal{M}$ es un difeomorfismo.*

a.1 Tangente a una variedad diferenciable

Se aborda este tópico describiendo a un vector local tangente a una superficie, en particular a superficies en $\mathbb{R}^3$, y cuyo vector a ella, es entendido como el "vector velocidad" en $\mathbb{R}^3$ en un punto p de la superficie. La idea es caracterizar al vector tangente de la variedad tal que se extienda la noción de velocidad y la propia defnición de derivada "normal". La siguiente definición de curva será expresada en términos de $\mathbb{R}^n$.

Definición 14. *Sea $\varepsilon > 0$ suficientemente pequeño y una curva $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ tal que:*

$$\gamma(t) = (\gamma_1(t), ..., \gamma_n(t)),$$

con $\gamma(0) = p$ y $\gamma'(0) = (\gamma_1'(0), ..., \gamma_n'(0)) = v \in \mathbb{R}^n$.

Ahora definimos la derivada direccional.

Definición 15. *Sea f una función tal que $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable definida en una vecindad o entorno de p. La misma, se restringe a la curva dada, γ y se escribe la* ***derivada direccional*** *de f en la dirección de $v \in \mathbb{R}^n$ como:*

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

De la definición (14) y (15) se puede ver que la derivada direccional de v, es un operador sobre funciones diferenciables y podemos notar que básicamente depende sólo de v, así, ya se tiene una caracterización para definir a un vector tangente sobre variedades, como sigue a continuación.

Definición 16. *(Vector tangente sobre variedad diferenciable $\mathcal{M}$). Sea γ una curva diferenciable. tal que $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$, donde $\gamma(0) = p$ y sea $D_p = \{f : \mathcal{M} \longrightarrow \mathbb{R} :$ con f diferenciable en $p\}$. El vector tangente a la curva γ en $t = 0$ es como la función $\gamma'(0) : D_p \to \mathbb{R}$ dada por:*

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

El vector tangente en p es el vector tangente en $t = 0$ de alguna curva $\gamma : (-\varepsilon, \varepsilon) \to M$ con $\gamma(0) = p$.

En particular, si $\mathcal{M}$ es una superficie regular de dimensión $k \leq n$, esto es $\mathcal{M} \subset \mathbb{R}^n$, entonces se puede definir al vector tangente en el punto p como el vector velocidad en $\mathbb{R}^n$, esto es,

$$\gamma'(0) = (\gamma_1'(0), \gamma_2'(0), ..., \gamma_n'(0)).$$

Definición 17. *(El Espacio tangente). Se denota con $T_p\mathcal{M}$, al espacio tangente a una variedad $\mathcal{M}$ en un punto p y es el conjunto de todos los vectores tangentes a $\mathcal{M}$*

en p. Esto es:

$$T_p\mathcal{M} = \{v \in \mathbb{R}^m : v es\ un\ vector\ tangente\ en\ p\}.$$

Observación 2. *Si consideramos una parametrización $\mathcal{X}$ tal que $\mathcal{X} : U \subset \mathbb{R}^n \to \mathcal{M}$ con $p = \mathcal{X}(0)$ y $q \in U$, podemos representar la función $f \in D_p$ y la curva $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ en la parametrización $\mathcal{X}$, del siguiente modo:*

$$f o\mathcal{X}(q) = f(\mathcal{X}(q)) = f(q) = f(q_1, ..., q_n),\ q = (q_1, q_2, q_3..., q_n) \in U$$

Evidentemente: $f o\mathcal{X} \equiv f$.
Análogamente al caso anterior, y usando la parametrización inversa $\mathcal{X}^{-1}$ se tiene:

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

Entonces se puede restringir la función f a la curva γ tenemos:

$$\gamma'(0)f = \frac{d(f \circ \gamma)}{dt}(t)\Big|_{t=0} = \frac{d(f \circ \mathcal{X} \circ \mathcal{X}^{-1} \circ \gamma)}{dt}(t)\Big|_{t=0}$$

$$= \frac{d}{dt}(f(q_1(t), q_2(t), ..., q_n(t))\Big|_{t=0}$$

entonces

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

Se puede ver que el vector γ' se puede escribir como

$$\gamma'(0) = \sum_{i=1}^{n} q_i^{'}(0).\left(\frac{\partial}{\partial q_i}\right)_0 \tag{II.1}$$

y es la expresión del vector tangente a γ en el punto p según la parametrización $\mathcal{X}$.

Observación 3. *De la observación anterior, al elegir una parametrización tenemos n* **vectores** *de la formar ($(\frac{\partial}{\partial x_i})_p$, $i = 1, ..., n$) en el punto p del espacio tangente $T_p\mathcal{M}$, dichos vectores se generan, por causa de la ecuación (II.1), y los mismos están en $T_p\mathcal{M}$.*

Definición 18. *Sea $\mathcal{M}$ una variedad diferenciable, definimos el fibrado tangente de $\mathcal{M}$ como*

$$T\mathcal{M} = \{(p, v);\ p \in \mathcal{M}/\ v \in T_p\mathcal{M}\}.$$

$T\mathcal{M}$ puede ser munido de uma estructura diferenciable, transformandose así en una variedade diferenciable.

Para mayores detalles, vea a (Do Carmo, 1988).

A continuación se tienen dos resultados importantes, la proposición (7) y (8) que muestran ejemplos concretos de espacios tangentes

Proposición 7. *El espacio tangente de una variedad diferenciable que es un subconjunto abierto de $\mathbb{R}^n$ es el propio $\mathbb{R}^n$.*

Demostración. Sea $\mathcal{M}$ una variedad diferenciable, subconjunto abierto de $\mathbb{R}^n$. Claramente, por definición se tiene el espacio tangente en p como $T_p\mathcal{M} \subset \mathbb{R}^n$. Debe probarse que $\mathbb{R}^n \subset T_p\mathcal{M}$. Sea $v = (v_1, v_2, ..., v_n) \in \mathbb{R}^n$. Como $p = (p_1, p_2, ..., p_n) \in \mathcal{M}$ y $\mathcal{M}$ es abierto, entonces existe un $r > 0$ tal que $B(p, r) \subset \mathcal{M}$, donde $B(p, r) = \{y \in \mathbb{R}^n / \|y - p\| < r\}$.
Definiendo $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^n$ tal que $\gamma(t) = p + tv$, es decir,

$$\gamma(t) = (p_1 + tv_1, p_2 + tv_2, ..., p_n + tv_n).$$

Sabiendo que $\gamma(0) = p$ y que $\gamma'(0) = v$, solo falta mostrar si $\gamma(t) \in \mathcal{M}$, desde que $B \subset \mathcal{M}$ entonces $\gamma(t) \in B(p, r)$.

En efecto, sin perdida de generalidad se puede suponer que para un $\epsilon > 0$, existe un $t > 0$ suficientemente pequeño tal que $\|\gamma(t) - p\| < r$, esto es $\|p + tv - p\| < r$ luego:

$$|\mathcal{X}|\|(q_1, q_2, ..., q_n)\| < r$$

$$|t|\|v\| < r.$$

Entonces, qué valores debe tomar t para que $\gamma(t) \in \mathcal{M}$. Luego $-\frac{r}{\|v\|} < t < \frac{r}{\|v\|}$, tomando $\epsilon = \frac{r}{\|v\|}$ se tiene $\gamma(t) \in B(p, r)$ y como $B \subset \mathcal{M}$ se tiene que $\gamma(t) \in \mathcal{M}$). Y así $v \in T_p\mathcal{M}$, luego $\mathbb{R}^n \subset T_p\mathcal{M}$ por tanto, $T_p\mathcal{M} = \mathbb{R}^n$. ■

Los siguientes resultados fueron tomados de (Quispe C., 2008).

Proposición 8. *Sea $\mathcal{M} = F^{-1}(a)$ una variedad de dimensión $n - m$, donde la aplicación $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es una función diferenciable, U es abierto y a es un valor regular de F, entonces: $T_p\mathcal{M} = T_p(F^{-1}(a)) = Ker(dF_p)$.*

Demostración. Sea $v \in T_p\mathcal{M}$, entonces existe $\gamma : (\epsilon, \epsilon) \longrightarrow \mathcal{M}$ con $\alpha(0) = p$. Se sigue que

$$(F \circ \gamma)(t) = a$$

$$\frac{d}{dt}(F \circ \gamma)(t) = 0$$

esto es, $dF_p(\gamma'(0)) = 0$ y así$v = \gamma'(0) \in KerdF_p$.

Reciprocamente, sea $v \in Ker(dF_p)$ entonces $dF_p(v) = 0$. Definamos $z = p + tv + dF_p^T u$, donde $u \in \mathbb{R}^m$, $t \in \mathbb{R}$; p y v fijos. Consideremos el siguiente sistema de m ecuaciones y $m + 1$ incognitas (ya que u y t son variables)

$$F(p + tv + dF_p^t u) = a.$$

Tomando $u = \theta = (0, 0, 0, ..., 0) \in \mathbb{R}^m$ y $t = 0 \in \mathbb{R}$, tenemos una solución particular en $\widehat{x} = (\theta, 0) \in \mathbb{R}^{m+1}$ pues $F(p + 0v + dF_p^t \theta) = F(p) = a$.

Tomando el Jacobiano de F con respecto a las variables de $u = (u_1, ..., u_m)$ y calculando en el punto $\widehat{x} = (\theta, 0)$ tenemos

$$JF = dF_p dF_p^t$$

Como dF_p es de rango m, por ser a su valor regular, entonces JF es una matriz no singular en $\widehat{x} = (\theta, 0) \in \mathbb{R}^{m+1}$. Luego por el Teorema de la función inversa existe una vecindad $(-\epsilon, \epsilon)$ de $t = 0$ y $\alpha : (-\epsilon, \epsilon) \longrightarrow \mathbb{R}^m$ tal que, para todo $i = 1, ..., m$

$$\alpha(0) = (\alpha_1(0), ..., \alpha_m(0)) = \theta \ \ y \ \ F(p + tv + dF_p^T \alpha(t)) = a.$$

Aplicando la diferencial a F en el punto p tenemos:

$$dF_p v + dF_p dF_p^T \alpha'(0) = 0.$$

Como $v \in Ker(dF_p)$ tenemos que $dF_p dF_p^t \alpha'(0) = 0$ y de la inyectividad de $(dF_p)(dF_p)^t$ se tiene que $\alpha'(0) = 0$.

Definiendo $\gamma : (-\epsilon, \epsilon) \longrightarrow F^{-1}(a)$ tal que $\gamma(t) = p + tv + dF_p^T \alpha(t)$ obtenemos $\gamma(0) = p$ y $\gamma'(0) = v + dF_p^T \alpha'(0) = v$. Así, $v \in T_p\mathcal{M}$.
De ambas inclusiones se obtiene que $T_p F^{-1}(a) = Ker(dF_p)$. ∎

Ejemplo 3. *Considérese una variedad diferenciable $\mathcal{M}$ conformada por el conjunto de restricciones factibles de un problema de programación lineal, definidas como $\mathcal{M} = \{x \in \mathbb{R}^n_{++} : Ax = b\}$, donde $A \in \mathbb{R}^{m \times n}$ tiene rango $m < n$, entonces el espacio tangente es*

$$T_p\mathcal{M} = KerA = \{\Delta x \in \mathbb{R}^n : A\Delta x = 0\}.$$

En efecto, la función que define $\mathcal{M}$ es $F : \mathbb{R}^n_{++} \longrightarrow \mathbb{R}^m$ tal que $F(x) = Ax - b$, la diferencial de F en el punto $p \in \mathcal{M}$ es $dF_p = A$, luego aplicando la Proposición 8 *obtenemos el resultado.*

En el siguiente ejemplo, se considerea una variedad particular para encontrar el espacio tangente usando el Jacobiano de una función.

Ejemplo 4. *Sea $h : U \subset \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable con Jacobiano $J_h(x) = dh_x$ con rango m. Consideremos la variedad $\mathcal{M} = h^{-1}(x) = \{x \in U,\ h(x) = 0\}$, entonces:*

$$T_p\left(h^{-1}(0)\right) = Ker(J_h(x)).$$

Proposición 9. *Sean $\mathcal{M}_1$ y $\mathcal{M}_2$ dos variedades diferenciables de dimensión n y m respectivamente y sea $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$ una aplicación diferenciable. Para cada $p \in \mathcal{M}_1$ y cada $v \in T_p\mathcal{M}_1$, elijamos una curva diferenciable $\alpha : (-\epsilon, \epsilon) \to \mathcal{M}$ con $\alpha(0) = p$, $\alpha'(0) = v$. Definiendo $\beta = \varphi \circ \alpha$, la aplicación*

$$d\varphi_p : T_p\mathcal{M}_1 \to T_{\varphi(p)}\mathcal{M}_2$$

dada por $d\varphi_p(v) = \beta'(0)$ es una aplicación lineal que independe de la elección de α.

Demostración. Para la demostración detallada vea, (Do Carmo, 1988) $\blacksquare$.

Lo que el difeomorfismo permite hacer, es establecer una noción natural de equivalencia entre variedades diferenciables, debido a que es un homeomorfismo diferenciable, y cuya inversa también debe ser diferenciable.

La siguiente proposición permite estudiar el comportamiento local de una variedad sin necesidad de conocer la otra.

Proposición 10. *Sea $\mathcal{M}_1$ y $\mathcal{M}_2$ dos variedades diferenciables. Si $\varphi : \mathcal{M}_1 \to \mathcal{M}_2$ es un difeomorfismo, entonces $d\varphi_p : T_p\mathcal{M}_1 \to T_{\varphi(p)}\mathcal{M}_2$ es un isomorfismo. Entonces φ es un difeomorfismo local.*

Demonstración. La demostración es por aplicación del teorema de la función inversa, ver (Do Carmo, 1988).

a.2 Métricas sobre variedades riemannianas

En general, cuando se habla de métricas, inmediatamente nos llega a la mente la idea de "medir". En matemáticas, por supuesto existen distintas expresiones que nos permiten medir por ejemplo: distancias, calcular errores, longitudes de curvas y otros, tal que impacten en la toma de desiciones. Recuerdo aún en pregrado, en el curso de cálculo II me enseñaron la más elemental fórmula para calcular la longitud de arco s está dado por:

$$s = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

"Fácil" de demostrar y más aún, de aplicar a situaciones de contexto real.

Y como ya se visto, lo que se quiere en ésta ocasión es generalizar la forma de la expresión matemática para calcular la longitud de arco, entonces se toma una curva parametrizada en $\mathbb{R}^n$, si bien, en los diferentes textos de geometría encontramos expresiones en $\mathbb{R}^3$, se representa como ya anunciamos por extensión a $\mathbb{R}^n$.

Así, al tomarse una curva parametrizada $\gamma(t) = (\gamma_1(t), \gamma_2(t), ..., \gamma_n(t))$ con t en el intervalo I de $\mathbb{R}$, la longitud de arco de la curva generada por dicha curva $\gamma(t)$ es dado por:

$$\ell(\gamma) = \int_I \|v(t)\| dt$$

donde, $v(t) = (\gamma_1'(t), \gamma_2'(t), ..., \gamma_n'(t))$ y la expresión $\| . \|$ representa la norma euclideana. Como se puede ver, la longitud de la curva depende de la norma del vector velocidad definido por la métrica usual en $\mathbb{R}^n$.

Definición 19. *Dados los puntos* $X = (x_1, ..., x_n)$ *y* $Y = (y_1, ..., y_n)$ *pertenecientes a*

$\mathbb{R}^n$. *Se define la distancia euclidiana entre dichos puntos como*

$$d(X,Y) = ||X - Y|| = \sqrt{(x_1 - y_1)^2 + ... + (x_n - y_n)^2}$$

La función distancia d definida como $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ se denomina distancia o métrica euclidiana, o la métrica usual en $\mathbb{R}^n$.

Para establecer una forma de medir sobre una variedad diferenciable M sobre el cual se traza una curva, entonces la longitud de arco de la curva se calcula a través de la **medida** realizada al vector perteneciente al espacio tangente en un punto fijo. Para ello, se necesita definir una mérica sobre el espacio tangente T_pM para cada $p \in M$. Tómese en cuenta sobre el producto interno de dos vectores en el espacio V el cual es un valor escalar, con las propiedades siguientes

a) $< u,u > \geq 0$, y $< u,u >= 0$, si y sólo si $u = 0$.

b) $< u,v >=< v,u >$, para todo par de vectores $u,v \in V$.

c) $< u,v + w >=< u,v > + < u,w >$, para todo $u,v,w \in V$.

d) $< ku,v >= k < u,v >=< u,kv >$, para todo $u,v \in V$ y $k \in \mathbb{R}$.

Estas propiedades permiten estudiar la geometría del espacio considerado, razón por la cual, el producto interno es muy usado en la generalización en por ejemplo distintas métricas y que más adelante se hicieron los detalles.

Ahora, la siguiente definición de producto interno en $\mathbb{R}^n$

Definición 20. *Definimos producto interno clásico:* $<,>: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *en un punto* p *como:*

$$< v,w >_p = \sum_i v_i w_i,$$

donde $v,w \in \mathbb{R}^n, i = 1,..,n$

Por otro lado, también se pueden definir distintos productos internos, por ejemplo:

$$< v, w >_p = \sum_{i,j} g_{ij} v_i w_i$$

$$< v, w >_p = (Gv, w),$$

donde, identificamos $G = (g_{ij})$ y es una matriz simétrica definida positiva.

Ahora, cómo ocurre esta forma de definir otro producto interno. Basta realizar un cambio de coordenadas. Esto es, sean las coordenadas $x = (x_1, x_2, ..., x_n)$ y $z = (z_1, z_2, ..., z_n)$ tal que expresamos $x(t) = x(z(t))$, y visto en cada componente:

$$x(t) = (x_1(z_1(t), z_2(t), ..., z_n(t)), x_2(z_1(t), z_2(t), ..., z_n(t)), ..., x_n(z_1(t), z_2(t), ..., z_n(t)),$$

derivando la componente i-ésima respecto t se tiene:

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} \left(\frac{\partial x_i}{\partial z_j} \frac{\partial z_j}{\partial t} \right), \text{ para todo } i = 1, 2, ..., n.$$

Sean los vectores v^x, v^z tal que $v^x = (v_1^x, v_2^x, ..., v_n^x)$ y $v^z = (v_1^z, v_2^z, ..., v_n^z)$, donde cada coordenada componente es expresado en términos de su derivada respecto t:

$$v_i^x = \frac{dx_i}{dt} \tag{II.2}$$

y

$$v_i^z = \frac{dz_i}{dt} \tag{II.3}$$

Expresando en norma cuadrado el v^x, se tiene:

$$\| v^x \|^2 = (v^x, v^x) = \sum_{i=1}^{n} \left(\frac{dx_i}{dt} \right)^2 = \sum_{i=1}^{n} \left(\frac{\partial x_i}{\partial t} \right)^2 \tag{II.4}$$

Aquí, elevamos al cuadrado la ecuación (II.2)

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

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

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

entonces de la ecuación (II.4) tenemos que:

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

por conmutatividad de las sumatorias, se tiene:

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

Por comodidad, hacemos un cambio, k por i en los inicios de las sumatorias

$$\| v^x \|^2 = \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n}\frac{\partial x_k}{\partial z_j}\frac{\partial x_k}{\partial z_i}\right)v_i^z v_j^z.$$

La expresión $\sum_{k=1}^{n}\frac{\partial x_k}{\partial z_j}\frac{\partial x_k}{\partial z_i}$ la denotamos con g_{ij} y queda así: $g_{ij} = \left(\sum_{k=1}^{n}\frac{\partial x_k}{\partial z_j}\frac{\partial x_k}{\partial z_i}\right)$ otra vez en norma, tenemos ahora:

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

es decir, hacemos $g_{i,j} = G$ podemos ver que un cambio en el sistema de coordenadas no altera las expresiones métricas.

Por ejemplo, si $v = G^{1\backslash 2}w$ tenemos:

$$(v,v)_p = (G^{1\backslash 2}w, G^{1\backslash 2}w)_p = (Gw,w)_p = \langle\, w,w \rangle_p\,.$$

que es otra expresión de la métrica estudiada.

Ahora veremos la definición formal de métrica riemanniana siguiendo lo abordado en (Quispe C., 2008).

Definición 21. *Sea $\mathcal{M}$ una variedad diferenciable. Una métrica riemanniana es una aplicación que asocia a cada $p \in \mathcal{M}$ un funcional $\langle\,,\rangle_p$*

$$\langle\,,\rangle_p : T_p\mathcal{M} \times T_p\mathcal{M} \longrightarrow \mathbb{R}$$

de modo que se cumplen las siguientes condiciones:

i). $\langle\,,\rangle_p$ es un producto interno (bilineal, simétrica y definida positiva) para cada $p \in \mathcal{M}$.

ii). $\langle\,,\rangle_p$ varía diferenciablemente en el siguiente sentido: Si $\mathcal{X} : U \subset \mathbb{R}^n \longrightarrow \mathcal{M}$ es um sistema de coordenadas en torno de p, con $\mathcal{X}(x_1,x_2,x_3,\ldots.,x_n) = q \in \mathcal{X}(U)$ y $\frac{\partial}{\partial x_i}(q) = d\mathcal{X}_q(0,0,\ldots,0,1,0,\ldots,0,0)$, entonces la función: $g_{ij} : U \longrightarrow \mathbb{R}$ definida por

$$g_{ij}(x_1,x_2,\ldots,x_n) = \left\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \right\rangle_q,$$

es diferenciable.

Esta definición tampoco depende del sistema de coordenadas elegidas, y las funciones g_{ij} son llamadas expresiones de la métrica riemanniana en el sistema coordenado $\mathcal{X}$ y la matriz $G = (g_{ij})$ es la representación de la métrica riemanniana.

La siguiente definición, estable la variedad riemanniana vinculada a la métrica riemanniana definida en 21.

Definición 22. *(Variedad riemanniana). Una variedad diferenciable con una métrica riemanniana se llama* **variedad riemanniana.**

Definición 23. *Sean $\mathcal{M}$ y $\mathcal{N}$ dos variedades riemannianas, un difeomomrfismo $f : \mathcal{M} \to \mathcal{N}$ se llama isometría si:*

$$\langle \, , \rangle_p = \langle \, df_p(u), df_p(v) \rangle_{f(p)} \tag{II.5}$$

para todo $p \in \mathcal{M}$ y $u, v \in T_p\mathcal{M}$

Además, la isometría planteada es local en la variedad

Definición 24. *Sean $\mathcal{M}$ y $\mathcal{N}$ dos variedades riemannianas, para $U \subset \mathcal{M}$ abierto, un difeomorfismo $f : U \to f(U) \subset \mathcal{N}$ que satisface (II.5) se llama isometría local.*

Veamos algunos ejemplos de métrica y variedad riemanniana más conocidas:

Ejemplo 5. *Sea la variedad particular cuando tomamos $\mathcal{M} = \mathbb{R}^n$, y al tomar la parametrización $\mathcal{X} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tal que $\mathcal{X}(x) = (x_1, x_2, ..., x_n)$. donde $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$*
Respecto la métrica, tomemos aquel definida por el producto interno euclidiano en $\mathbb{R}^n$, esta es:

$$\langle \, , \rangle_p : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \textit{ definido por } \langle \, x, y \rangle_p = x^T y \tag{II.6}$$

Sea un punto $q \in \mathbb{R}^n$ entonces al derivar parcialmente sobre la i-ésima componente, la parametrización en dicho punto queda así:

$$\frac{\partial}{\partial x_i}(q) = d\mathcal{X}_q e_i = e_i,$$

y tomando las expresiones de la métrica riemanniana, $g_{ij} : U \to \mathbb{R}$ definida como *el producto interno según (II.6)*

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

son diferenciables en $\mathbb{R}^n$.

Luego $\mathcal{M} = \mathbb{R}^n$, *y métrica riemanniana identidad* $G = Id$, *es una variedad riemanniana, esto significa que el propio espacio euclidiano es un ejemplo de variedad riemanniana.*

Los siguientes dos ejemplos son también variedades riemannianas para las métricas G según se definen, los cuales se demuestran bajo el mismo procedimiento del ejempl anterior.

Ejemplo 6. *Si* $\mathcal{M} = \mathbb{R}^n_{++}$ *y el funcional dado por el producto interno en* $\mathbb{R}$ $\langle\, , \rangle_p : T_p\mathbb{R}^n_{++} \times T_p\mathbb{R}^n_{++} \longrightarrow \mathbb{R}$ *de modo que dicho producto interno se define por medio de G, esto es:*

$$\langle u, v \rangle_p = u^T G(p) v,$$

donde

$$G(p) = diag(1/(h_i(p_i))^2),$$

entonces, la expresión de la métrica riemanniana, toma lasiguiente forma:

$$g_{ij}(x) = \frac{\delta_{ij}}{(h_2(x_i))^2}.$$

Ejemplo 7. *Sea la variedad riemanniana* $(\mathbb{R}^2, G(x))$, *con* $\langle u, v \rangle_p = u^T G(p) v$, *donde la matriz de representación de la métrica riemanniana se define como sigue:*

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

esta métrica fué dado por (Udriste, 1994).

a.3 Campos vectoriales, conexiones y derivada covariante

En este punto, se explica el movimiento o comportamiento de un punto a lo largo de una trayectoria, si es que existe; para ello introducimos la definición de una función especial, llamada campo vectorial, la idea es asociar un punto con un vector cuyas coordenadas varian diferenciablemente en relación al punto. Los distintos métodos computacionales en Optimización matemática, generan una sucesión de puntos digamos x_0, x_1, x_2 $\cdots$, y requiere entre otras cosas, de una dirección para generar la sucesión de puntos y minimizar una función; en el proceso, los vectores dirección que se generen, puede ser visto como un modo de transportar vectores sobre curvas diferenciables; lo que en el fondo se quiere es, derivar campos vectoriales en una variedad riemannina M, (Figueroa, 2004). Por lo que debemos estudiar a los campos vectoriales, conexiones afines, derivadas, curvatura y propiedades alrededor de esos conceptos lo que se verá a continuación.

- **Campo de vectores**

  Definición 25. *Sea X Un campo de vectores o campo vectorial, en una variedad diferenciable $\mathcal{M}$, es una correspondencia tal que a cada punto $p \in M$ asocia un vector $X(p)$ del espacio tangente $\in T_p\mathcal{M}$.*

  El vector $X(p)$ puede ser escrito como

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

  desde que X es una parametrización $\mathcal{X} : U \subset \mathbb{R}^n \to \mathcal{M}$, donde cada $a_i : \mathcal{M} \to \mathbb{R}$ es una función en $\mathcal{M}$ y $\{(\frac{\partial}{\partial x_i})_p\}$ es un operador asociado a $\mathcal{X}$, $1 \leq i \leq n$. Además, X es diferenciable sí, y solamente sí, las funciones componentes a_i son diferenciables también para alguna parametrización.

Y si consideramos la función f, entonces tenemos la composición Xf en el punto p, entonces los campos vectoriales toman la siguiente expresión

$$(Xf)(p) = \sum_i a_i(p).\frac{\partial f}{\partial x_i}(p),$$

Esto es, si $X : D \longrightarrow F$ donde D es el conjunto de las funciones diferenciables sobre v y F es el conjunto de las funciones sobre $\mathcal{M}$.

Como se quieren las trayectorias en $\mathcal{M}$, se consideran los campos restringidos a una curva.

Ahora vamos a definir el campo de vectores a lo largo de una curva en la siguiente definición.

Definición 26. *Sea V la notación de un campo vectorial, a lo largo de una curva $\alpha : I \longrightarrow \mathcal{M}$ es una aplicación donde a cada $\alpha(t) \in \mathcal{M}$ asocia un vector tangente $V(t)$ en la tangente de α, esto es $V(t) \in T_{\alpha(t)}\mathcal{M}$.*

veamos la siguiente definición sobre la diferenciabilidad del campo V.

Definición 27. *Se dise que V es diferenciable si para cada función diferenciable f en D, la función $V(t)f$ es una función diferenciable en I.*

El campo X a lo largo de α denotamos como $V(t) = X(\alpha(t))$ esto es, V es inducido por X.

El campo vectorial $d\mathcal{X}_{(\mathcal{X}^{-1}\circ\alpha)(t)}[(\mathcal{X}^{-1}\circ\alpha)'(t)] = \frac{d(\mathcal{X}o\mathcal{X}^{-1}\circ\alpha)(t)}{dt}$ denotado por $\frac{d\alpha}{dt}$, es llamado campo velocidad o tangente de α.

- **Conexiones Afines**

Las conexiones afines son estructuras que permiten conectar puntos en el conjunto de espacios tangente de la geometría en referencia a la derivada covariante, que más adelante se definirá. Para definir una Conexión Afín, usaremos como notación a $T\mathcal{M}$ como el conjunto de espacios tangentes en $\mathcal{M}$, sobre el cual se definirá al conjunto de campos vectoriales, $\mathcal{H}$ seguidamente.

Definición 28. *Sea* $\mathcal{H} = \mathcal{H}(\mathcal{M}) = \{X : \mathcal{M} \longrightarrow T\mathcal{M} : \text{ para cada } p \in M, X(p) \in T_p\mathcal{M},\ y\ X \in C^\infty\}$ *el conjunto de campo de vectores y* $D = D(\mathcal{M}) = \{f : M \longrightarrow \mathbb{R} : f \in C^\infty\}$ *el conjunto de funciones reales de clase* C^∞.

En resumen, una conexión en una variedad diferenciable, es como una estructura que permite relacionar puntos de la geometría, (localmente hablando), con el fin establecer la derivada covariante.
Véase la siguiente definición tomado de (Quispe C., 2008).

Definición 29. *Definimos una conexión afin a la aplicación* $\nabla : \mathcal{H}\times\mathcal{H} \longrightarrow \mathcal{H}$ *donde a cada par de campos* (X, Y) *se asocia otro campo* $\nabla_X Y$ *tal que para todo* $X, Y, Z \in \mathcal{H}$, *y* $f, g \in D$ *verifique:*

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. $\nabla_X fY = f\nabla_X Y + X(f)Y$, *donde* $X(f) = \sum\limits_{i=1}^{n} a_i(.)\frac{\partial f(.)}{\partial x_i}$.

Al considerar una curva diferenciable en $\mathcal{M}$, $\alpha : I \longrightarrow \mathcal{M}$, y se denota el conjunto de campo de vectores a lo largo de esta curva como $\mathcal{H}_\alpha$.

Proposición 11. *Sea* $\mathcal{M}$ *una variedad diferenciable con una conexión afin* ∇. *Entonces existe una única aplicación* $\frac{D}{dt}$, *donde a cada* $V \in \mathcal{H}_\alpha$ *se asocia otro campo en* $\mathcal{H}_\alpha$, *denotado por* $\frac{DV}{dt}$, *tal que para todo* $V, W \in \mathcal{H}_\alpha$ *y* $f : I \longrightarrow \mathbb{R}$ *una función diferenciable en* I *se cunplen:*

a. $\frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$
b. $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$
c. *Si* $V(t) = Y(\alpha(t))$, *donde* $Y \in \mathcal{H}$, *entonces* $\frac{DV}{dt} = \nabla_{\frac{d\alpha}{dt}}Y$.

$\frac{D}{dt}$ *es llamada* **Derivada Covariante**.

Demostración. Supongamos inicialmente que existe una correspondencia satisfaciendo 1, 2, y 3. Sea $\mathcal{X} : U \subset \mathbb{R}^n \to \mathcal{M}$ un sistema de coordenadas con $\alpha(I) \bigcap \mathcal{X}(U) \neq \phi$ y sea $(x_1(t), x_2(t), ..., x_n(t))$ la expresión local de $\alpha(t), t \in I$. Denotaremos $X_i = \frac{\partial}{\partial x_i}$.

Podemos expresar el campo V localmente como:

$$V = \sum_{j=1}^{n} v^j X_j$$

donde $V(t) = V(\alpha(t))$, $v^j(t) = v^j(\alpha(t))$ y $X_j(t) = X_j(\alpha(t))$. Entonces

$$\frac{DV}{dt} = \sum_{j=1}^{n} \frac{D}{dt}(v^j X_j) = \sum_{j=1}^{n} (\frac{dv^j}{dt} X_j + v^j \frac{DX_j}{dt}) \tag{II.7}$$

por eso tenemos,

$$\frac{DX_j}{dt} = \nabla_{\frac{d\alpha}{dt}} X_j = \sum_{i=1}^{n} (\nabla_{(\frac{dx_i}{dt} X_i)} X_j) = \sum_{i=1}^{n} \frac{dx_i}{dt} \nabla_{X_i} X_j$$

Sustituyendo esta expresión en (II.7) tenemos:

$$\frac{DV}{dt} = \sum_{j=1}^{n} \frac{dv^j}{dt} X_j + \sum_{i,j=1}^{n} v^j \frac{dx_i}{dt} \nabla_{X_i} X_j \tag{II.8}$$

La expresión (II.8) muestra que, si existe una correspondencia satisfaciendo a las condiciones de la proposición entonces tal correspondencia es única (ya que cualquier otra aplicación seguirá dependiendo de v^j, $\frac{dv^j}{dt}$, $\frac{dx_i}{dt}$ y X_j.

Para probar la existencia definamos en la parametrización $\mathcal{X} : U \to \mathcal{M}$

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

y verifiquemos que se cumplen las propiedades 1, 2, y 3.

Dado otro campo $W = \sum\limits_{j=1}^{n} w^j X_j$ se tiene que

$$V + W = \sum_{j=1}^{n} (v^j + w^j) X_j$$

Propriedad 1):

$$\begin{aligned}
\tfrac{D}{dt}(V+W) &= \sum_{j=1}^{n} \tfrac{d(v^j+w^j)}{dt} X_j + \sum_{i,j=1}^{n} (v^j + w^j) \tfrac{dx_i}{dt} \nabla_{X_i} X_j \\
&= \left(\sum_{j=1}^{n} \tfrac{dv^j}{dt} X_j + \sum_{i,j=1}^{n} v^j \tfrac{dx_i}{dt} \nabla_{X_i} X_j \right) \\
&+ \left(\sum_{j=1}^{n} \frac{dw^j}{dt} X_j + \sum_{i,j=1}^{n} w^j \frac{dx_i}{dt} \nabla_{X_i} X_j \right) \\
&= \tfrac{DV}{dt} + \tfrac{DW}{dt}.
\end{aligned}$$

Propriedad 2):

$$\begin{aligned}
\tfrac{D}{dt}(fD) &= \tfrac{D}{dt}\Big(\sum_{j=1}^{n} f v^j X_j\Big) \\
&= \sum_{j=1}^{n} \tfrac{d}{dt}(f v^j) X_j + \sum_{i,j=1}^{n} f v^j \tfrac{dx_i}{dt} \nabla_{X_i} X_j \\
&= \sum_{j=1}^{n} \tfrac{df}{dt} v^j X_j + \sum_{j=1}^{n} f \tfrac{dv^j}{dt} X_j + \sum_{i,j=1}^{n} f v^j \tfrac{dx_i}{dt} \nabla_{X_i} X_j \\
&= \tfrac{df}{dt} \sum_{j=1}^{n} v^j X_j + f \left(\sum_{j=1}^{n} \tfrac{dv^j}{dt} X_j + \sum_{i,j=1}^{n} v^j \tfrac{dx_i}{dt} \nabla_{X_i} X_j \right) \\
&= \tfrac{df}{dt} V + f \tfrac{DV}{dt}.
\end{aligned}$$

Propriedad 3):

$$\begin{aligned}
\nabla_{\frac{d\alpha}{dt}} Y &= \nabla_{(\sum_{i=1}^{n} \frac{dx_i}{dt} X_i)} \left(\sum_{j=1}^{n} v^j X_j \right) \\
&= \sum_{i=1}^{n} \tfrac{dx_i}{dt} \left(\sum_{j=1}^{n} v^j \nabla_{X_i} X_j + \tfrac{\partial v^j}{\partial x_j} X_j \right) \\
&= \sum_{i,j=1}^{n} \big(\tfrac{dx_i}{dt} v^j \nabla_{X_i} X_j\big) + \sum_{i=1}^{n} \tfrac{dv^j}{dt} X_j \\
&= \tfrac{DV}{dt}
\end{aligned}$$

hasta ahora $\frac{DV}{dt}$ sólo fué definida en $\mathcal{X}(U)$, falta definir $\frac{DV}{dt}$ en todo $\mathcal{M}$.

Si $\mathcal{Y}(W)$ es una otra vecindad coordenada, con $\mathcal{Y}(W)\cap\mathcal{X}(U) \neq \phi$, y definimos $\frac{DV}{dt}$ con respecto a la parametrización $\mathcal{Y}$ em $\mathcal{Y}(W)$ por (II.8), las definiciones concuerdan en $\mathcal{Y}(W)\cap\mathcal{X}(U)$ por la unicidad de $\frac{DV}{dt}$. Se sigue que la definición puede ser extendida para todo $\alpha(I) \subset \mathcal{M}$, y esto concluye la prueba. ■

Con esta Proposición se muestra que la elección de una conexión afin de $\mathcal{M}$ origina una única derivada covariante para cada campo vectorial.

Observación 4. *La demostración de la Proposición 11 muestra una caracterización de la derivada covariante para una alguna parametrización, por ejemplo $\mathcal{X}$, cuya expresión es la siguiente:*

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

Observación 5. *Como ya dijimos, la conexión afin, permite establecer una forma de derivar un campo vectoriale. Esto es, si tenemos el campo vectorial $V = \frac{d\alpha}{dt}$ escribimos:*

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

y decimos que es la **aceleración de la curva** α **en** $\mathcal{M}$.

- **Conexión afin relativa a coordenadas locales**

Sean los campos de vectores $X, Y \in \mathcal{H}$ definidos en $\mathcal{M}$ a partir de una vecindad local $\mathcal{X} : U \subset R^n \longrightarrow \mathcal{M}$, para un punto p, por:

$$X = \sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i}, \quad Y = \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i},$$

donde el operador $(\partial/\partial x_i)$ que representan a los vectores de la base para el sistema de coordenadas locales. Denotamos como $\frac{\partial}{\partial x_i} = X_i$.

Así se tiene las expresiones de la conexión afín relativa a las coordenadas locales:

$$X = \sum_{i=1}^{n} x_i X_i, \quad Y = \sum_{i=1}^{n} y_i X_i.$$

En relación a las propiedades conexión afin se puede expresar cada campo vectorial como expresión de coordenadas locales:

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

Ahora, si se tiene $\nabla_{X_i} X_j \in \mathcal{H}$, puede ser representado en térmnos de los símbolos de Christoffel, esto es:

$$\nabla_{X_i} X_j = \sum_{k=1}^{n} \Gamma_{ij}^{k} X_k \tag{II.9}$$

que, substituyendo en la ecuación (II.9), tenemos una expresión más simplificada del campo vectorial:

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

Definimos a continuación las expresiones en coordenadas espaciales que, puede decirse, simplifican o facilitan hacer cálculos o la descripción en expresiones, de los desplazamiento paralelos de vectores a través de los campos tensoriales; estos, son los llamados Símbolos de Christoffel Γ_{ij}^{k}, en honor a Elwin Bruno Christoffel.

Definición 30. *Los símbolos de Christoffel, son funciones Γ_{ij}^{k}, coeficientes de* *la conexión afin ∇ en U tal que*

$$\Gamma_{ij}^k : U \subset \mathcal{M} \to \mathbb{R}$$

tal que se cumple la ecuaón (II.9).

- **La derivada covariante y los símbolos de Christoffel.**

  Como sabemos, la derivada covariante la entendemos como la generalización de la muy conocida derivada parcial, y que nos permite extender el cálculo diferencial sobre $\mathbb{R}^n$. Vamos ahora escribir la derivada covariante usando las coordenadas locales y los llamados símbolos de Christoffel, para tal efecto hacemos una aplicación $\mathcal{X} : U \to \mathcal{M}$ como el sistema de coordenadas locales al rededor del punto $p \in M$. De la ecuación (II.8)

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

  y del campo vectorial

$$\nabla_{X_i} X_j = \sum_{k=1}^{n} \Gamma_{ij}^k X_k,$$

  se tiene:

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

$$= \sum_{j=1}^{n} \frac{dv^j}{dt} X_j + \sum_{k=1}^{n} \sum_{i,j=1}^{n} v^j \frac{dx_i}{dt} \Gamma_{ij}^k X_k.$$

  Así:

$$\frac{DV}{dt} = \sum_{k=1}^{n} \left(\frac{dv^k}{dt} + \sum_{i,j=1}^{n} v^j \frac{dx_i}{dt} \Gamma_{ij}^k \right) X_k \tag{II.10}$$

a.4 **Geodésicas**

La derivada covariante permite definir el transporte paralelo a lo largo de curvas que dependen de la métrica, osea, que cambiando la métrica, cambia en general la manera de derivar campos vectoriales; en particular, nos permite conocer geodésicas, curvas cuyo vector tangente es paralelo a la curva. Si $\alpha : [a,b] \to \mathcal{M}$ es una curva, tal que $\alpha(a) =$ y $\alpha(b) = q$, el transporte paralelo es una aplicación $P_{\alpha(t)} : T_p\mathcal{M} \to T_q\mathcal{M}$ que satizface ser un isomorfismo lineal tal que $P_{\alpha(t)}(v) = V(b)$ que es el único transporte paralelo a lo largo de α.

Si usamos la métrica euclidiana en un espacio euclidiano, digamos $\mathbb{R}^2$, la geodésica es el segmento de recta que une dos puntos p y q en distintas posiciones, lo que caracteriza la trayectoria de menor longitud del segmento de curva que los une.

Otra manera de entender respecto de geodésicas, es definir una curva $\alpha : [a,b] \to \mathbb{R}^n$ diferenciable que pasa por los puntos $p = \alpha(a)$ y $q = \alpha(b)$, siendo el campo $\frac{d\alpha}{dt}$ asociado a la velocidad, por lo cual se tiene la aceleración $\frac{d}{dt}(\frac{d\alpha}{dt})$ en cada punto de $\alpha(t)$ con la propiedad de ser la geodésica tal que se cumple:

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

Luego, para expresar la generalización de esta ecuación a variedades, basta a considerar que la componente tangencial de dicha derivada sea nula.

Ahora, veamos la definición de geodésica, que en términos simples, es la distancia que existe entre dos puntos distintos, ubicados sobre una esfera. La definición matemática la damos a continuación.

Definición 31. *Sea α una curva parametrizada como $\alpha : I \longrightarrow \mathcal{M}$, decimos que α es una geodésica si el campo tangente de la curva dada $\frac{d\alpha}{dt}$ es:*

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

Estudiamos ahora los campos paralelos, estos son campos vectoriales paralelo a una curva α; respecto el paralelismo, podemos decir que es usada para transportar vectores tangentes de un punto de la superficie a otro punto de la misma. Veamos la definición formal de campos paralelos, estudiada de (Quispe C., 2008).

a.6 Campos paralelos

Dado $\mathcal{M}$ una variedad diferenciable, una conexión afin ∇ y un campo V a lo largo de una curva diferenciable $\alpha : I \to \mathcal{M}$, V es denominado campo paralelo si $\frac{DV}{dt} = 0$, para todo $t \in I$.
Así, si α es una geodésica, entonces $\frac{d\alpha}{dt}$ es paralelo.
Otra manera de expresar es: el único campo vectorial paralelo a la curva α se llama transporte paralelo.

- **Ecuaciones geodésicas.**

  Ahora veamos cuáles son las ecuaciones que gobiernan una geodésica, para esto tomamos la expresión (II.10), entonces un campo paralelo V es determinado por las ecuaciones

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

  y como el sistema de las n ecuaciones ecuaciones diferenciables en v_k y donde cada $\frac{\partial}{\partial x_k}$ son linealmente independientes, podemos escribir:

$$\frac{dv^k}{dt} + \sum_{i,j=1}^{n} v_j \frac{d\alpha_i}{dt}\Gamma_{ij}^k = 0, \qquad k = 1, \ldots, n.$$

  Ahora, de la geodésica $\alpha(t) = (\alpha_1(t), \ldots, \alpha_n(t))$, un $v^i = \frac{d\alpha_i}{dt}$, y la ecuación anterior, pasa a ser escrito como sigue:

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

Si ordenamos la ecuación anterior, tenemos a la vista una ecuación diferencial ordinaria:

$$\frac{d^2\alpha_k}{dt^2} + \sum_{i,j=1}^{n} \Gamma_{ij}^{k} \frac{d\alpha_i}{dt}\frac{d\alpha_j}{dt} = 0, \qquad k = 1, \ldots, n \tag{II.11}$$

Como podemos observar, tenemos un sistema de n ecuaciones diferenciales de segundo orden, la cual tiene una única solución en el intervalo $I = [a, b]$, satisfaciendo condiciones iniciales $x(0) = \alpha(0) = p$ y $\frac{dx}{dt}(0) = \alpha'(0) = v$

- **Conexión afín y métrica en variedades riemannianas.**

El siguiente concepto define a una conexión compatible a partir de una conexión afin y así establecer una métrica en una variedade riemanniana. Sea $\mathcal{M}$ una variedad diferenciable con una conexión afin ∇ y una métrica riemanniana $\langle\, , \rangle$. Los resultados aqui expuestos se siguen de (Quispe C., 2008).

Definición 32. *Se dice que* ∇ *es* compatible *con la métrica* $\langle\, , \rangle$ *si para todo par de campos de vectores V y W a lo largo de la curva diferenciable* $\alpha : I \longrightarrow \mathcal{M}$ *se tiene:*

$$\frac{d}{dt}\langle V, W\rangle = \langle \frac{DV}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle \tag{II.12}$$

Proposición 12. *Sea* $\mathcal{M}$ *una variedad riemanniana. Una conexión afin* ∇ *es compatible con* $\langle, \rangle$ *sí y solamente sí,* $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, para\ todo\ X, Y, Z \in \mathcal{H}$.

Demostración. Para la demostración, véase Do Carmo, (1988) pp 54.

De los resultados anteriores, se define que la conexión afín es también una métrica.

Definición 33. *Una conexión afín* ∇ *en una variedad diferenciable* $\mathcal{M}$ *es llamada simétrica si*

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

donde $[X, Y] = XY - YX$.

De las propiedades de la conexión afín se puede deducir algunos hechos, veamos:

Para un sistema de coordenadas $(U, \mathcal{X})$ y por la propiedad de simetría de la conexión afín, podemos escribir dicha conexión en términos de los operadores diferenciales correspondientes, esto es

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

Si tomams una $f \in D$, y aplicando la definición anterior, tenemos el siguiente resultado.

$$X_i X_j(f) - X_j X_i(f) = \frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 f}{\partial x_j \partial x_i} = 0.$$

Es decir que:

$$\nabla_{X_i} X_j - \nabla_{X_j} X_i = \sum_{k=1}^{n} (\Gamma_{ij}^k - \Gamma_{ji}^k) X_k = 0.$$

y como $\{X_k\}$ es linealmente independencia, se tiene que:

$$\Gamma_{ij}^k = \Gamma_{ji}^k.$$

A continuación presentamos un resultado importante sobre la existencia y unicidad de una conexión simétrica y compatible con la métrica en una variedad riemanniana, fundamentalmente que ∇ está univocamente determinado por la métrica $\langle , \rangle$

Teorema 1. *(Levi-Civita). Dada una variedad riemanniana M, existe una única conexión afin ∇ en M satisfaciendo las condiciones:*

a) ∇ *es simétrica.*

b) ∇ *es compatible con la métrica riemanniana.*

Demostración. Para la demostración, véase (Do Carmo, 1988) a.7 **Métrica riemanniana y símbolos de Christoffel.**

Lo que veremos aquí es, la relación que hay entre la métrica riemanniana con los símbolos de Christoffel, para esto, tomamos el sistema de coordenadas $(U, \mathcal{X})$, y las funciones conocidas como símbolos de Christoffel $\Gamma^k_{i,j} : U \to \mathbb{R}$, estas definen los coeficientes de conexión $\nabla_{X_i} X_j = \sum_k^n \Gamma^k_{ij} X_k$. Se muestra que

$$\Gamma^m_{ij} = \{\frac{1}{2}\sum_k^n \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij}\} g^{km}$$

donde $g_{ij} = \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle$ son elementos de la matriz $G(x)$ y g^{ij} los elementos de su inversa $G^{-1}(x)$ respectivamente.

En efecto, tomemos $\frac{\partial}{\partial x_i} = X_i$, $\frac{\partial}{\partial x_j} = X_j$ y $\frac{\partial}{\partial x_k} = X_k$. Usando el resultado siguiente:

$\langle Z, \nabla_Y X \rangle =$

$$= \frac{1}{2}\{X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle [X, Z], Y\rangle - \langle [Y, Z], X\rangle - \langle [X, Y], Z\rangle\} \quad \text{(II.13)}$$

tenemos

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

Como $\nabla_{X_j} X_i = \nabla_{X_i} X_j = \nabla_{X_i} X_j = \sum_{l=1}^n \Gamma^l_{ij} X_l$ y usando a linealidad del producto interno, se tiene:

$$\sum_{l=1}^n \Gamma^l_{ij} \langle X_k, X_l\rangle = \frac{1}{2}\{X_i\langle X_j, X_k\rangle + X_j\langle X_k, X_i\rangle - X_k\langle X_i, X_j\rangle\},$$

y así:

$$\sum_{l=1}^n \Gamma^l_{ij} g_{kl} = \frac{1}{2}\{\frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij}\}.$$

Denotando $b_k = \frac{1}{2}\{\frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij}\}, k = 1, 2, ..., n$ obtenemos un sistema lineal $Gy = b$ con $y = (\Gamma^1_{ij}, \Gamma^2_{ij}, ..., \Gamma^n_{ij})$ y $b = (b_1, b_2, ..., b_n)$. Como $G(x)$ es invertible (ver definición de métrica riemanniana) entonces $y = G^{-1}b$. Así tenemos

$$\Gamma_{ij}^{m} = \frac{1}{2}\sum_{k=1}^{n} g^{mk} b_k.$$

Finalmente sustituyendo el valor de b_k en la expresión anterior se tiene:

$$\Gamma_{ij}^{m} = \frac{1}{2}\sum_{k=1}^{n}\{\frac{\partial}{\partial x_i}g_{jk} + \frac{\partial}{\partial x_j}g_{ki} - \frac{\partial}{\partial x_k}g_{ij}\}g^{km}. \tag{II.14}$$

Ejemplo 8. *Consideremos la variedad riemanniana $\mathcal{M} = \mathbb{R}_{++}^{n}$, con la métrica dada por*

$$G(x) = diag\left(\frac{1}{(h_1(x_1))^2}, \frac{1}{(h_2(x_2))^2}, ..., \frac{1}{(h_n(x_n)^2}\right) \tag{II.15}$$

para funciones $h_i : \mathbb{R}_{++} \longrightarrow \mathbb{R}_{++}$ diferenciables. La inversa de la matriz $G(x)$ es:

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

a) *Hallamos los símbolos de Christoffel.*

Recordemos que la relación de la métrica con los símbolos de Christoffel está dado por la ecuación (II.14), recordemos que estos resultados lo estamos tomando de Quispe C., 2008.

Cuando $k \neq m$ tenemos que $g^{mk} = 0$, así la expresión es reducida a:

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

Consideramos dos casos:

a) Si $i = j$

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

Para $m = i$

$$\Gamma_{ii}^{i} = -\frac{1}{h_i(x_i)}\frac{\partial}{\partial x_i}\left(h_i(x_i)\right).$$

Para $m \neq i$

$$\Gamma_{ii}^{m} = 0.$$

b) *Si $i \neq j$*

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

Para $m = i$ entonces, $m \neq j$ y:

$$\Gamma_{ij}^{i} = 0.$$

Para $m = j$ entonces, $m \neq i$ y:

$$\Gamma_{ij}^{j} = 0.$$

Para $m \neq i$ y $m \neq j$ entonces,

$$\Gamma_{ij}^{m} = 0.$$

De ambos casos tenemos:

$$\Gamma_{ij}^{m} = -\frac{1}{h_i(x_i)}\frac{\partial(h_i(x_i))}{\partial x_i}\delta_{im}\delta_{ij} \tag{II.16}$$

que es la expresión de los Símbolos de Christoffel en relación a la métrica $G(x)$.

Como aplicaciones tenemos:

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

- *Si* $h_i(x_i) = x_i^{\frac{r}{2}}$ *entonces,* $G(x) = X^{-r}$. *Luego:* $\Gamma_{ij}^m = -\frac{r}{2}\frac{1}{x_i}\delta_{im}\delta_{ij}$.
- *Se* $h_i(x_i) = s_i^{\frac{-r}{2}} x_i^{\frac{r}{2}}$, $s_i \in \mathbb{R}_{++}$ *entonces,* $G(x) = S^r X^{-r}$. *Luego* $\Gamma_{ij}^m = -\frac{r}{2}\frac{1}{x_i}\delta_{im}\delta_{ij}$.

b) *Hallando la derivada covariante.*

Vimos que la relación de la derivada covariante con respecto a los símbolos de Christoffel es dada por la ecuación (II.10). Sustituyendo la expresión (II.16) en (II.10) obtenemos:

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

En particular:

- *Si* $h_i(x_i) = 1$, $\Gamma_{ij}^k = 0$, *y así:*

$$\frac{DV}{dt} = \sum_{i=1}^{n}\frac{dv^i}{dt}X_i,$$

que es la propia derivada usual.

- *Si* $h_i(x_i) = x_i$, $\Gamma_{ij}^m = -\frac{1}{x_i}\delta_{im}\delta_{ij}$ *y*

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

- *Si* $h_i(x_i) = x_i^{\frac{r}{2}}$, $\Gamma_{ij}^m = -\frac{r}{2}\frac{1}{x_i}\delta_{im}\delta_{ij}$

$$\frac{DV}{dt} = \sum_{i=1}^{n}\left(\frac{dv^i}{dt} - \frac{r}{2}\frac{1}{x_i}v^i\frac{dx_i}{dt}\right)X_i.$$

c) *Determinación de la ecuación geodésica: Sea* $p = (p_1, p_2, \ldots, p_n) \in \mathbb{R}_{++}^n$ *y*

$v = (v_1, v_2, ..., v_n) \in T_p\mathbb{R}^n_{++} = \mathbb{R}^n$ *con*

$$\alpha : I \longrightarrow \mathbb{R}^n_{++} : \alpha(t) = (\alpha_1(t), \alpha_2(t), ..., \alpha_n(t))$$

donde $\alpha(0) = p$ *y* $\frac{d\alpha(0)}{dt} = v$*, I algún intervalo abierto de* $\mathbb{R}$*. Substituyendo los símbolos de Christoffel (II.16) en la ecuación (II.11) obtenemos:*

$$\frac{d^2\alpha_i}{dt^2} - \frac{1}{h_i(\alpha_i)}\frac{\partial(h_i(\alpha_i))}{\partial\alpha_i}(\frac{d\alpha_i}{dt})^2 = 0, \forall i = 1, ..., n \tag{II.17}$$

$$\alpha_i(0) = p_i, \quad i = 1, ..., n$$

$$\alpha_i'(0) = v_i, \quad i = 1, ..., n.$$

La ecuación diferencial es equivalente a resolver:

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

para alguna constante a_i*, que también es equivalente a resolver la integral:*

$$\int \frac{1}{h_i(\alpha_i)}d\alpha_i = a_i t + b_i, \qquad i = 1, 2, ..., n$$

para algunas constantes a_i *y* b_i *en* $\mathbb{R}$*.*

Entonces, la única geodésica $\alpha(t)$ *de* $\mathbb{R}^n_{++}$*, con métrica* $G(p)$*, pasando por el punto* $\alpha(0) = p$*, en la dirección* $\alpha'(0) = v$*, es obtenida resolviendo el siguiente problema:*

$$\int (\frac{1}{h_i(\alpha_i)})d\alpha_i = a_i t + b_i \qquad i = 1, ..., n \tag{II.18}$$

donde a_i *y* b_i *son constantes reales tales que:*

$$\alpha_i(0) = p_i, \qquad i = 1, ..., n.$$

$$\alpha_i'(0) = v_i, \qquad i = 1, ..., n.$$

En particular:

- *Si $h_i(\alpha_i) = 1$ tenemos que $G(p) = I$ y considerando las condiciones iniciales de (II.18) encontramos la expresión de la curva geodésica*

$$\alpha_i(t) = v_i t + p_i \qquad i = 1, ..., n.$$

Esto es, las geodésicas son curvas $\alpha : \mathbb{R} \longrightarrow \mathbb{R}^n_{++}$ definidas por:

$$\alpha(t) = (v_1 t + p_1, ..., v_n t + p_n).$$

Observemos que la geodésica $\alpha(t)$ está definida para valores de t tal que $v_i t + p_i > 0$.

- *Si $h(\alpha_i) = \alpha_i$ entonces, $G(x) = X^{-2}$ considerando las condiciones iniciales de (II.18), las curvas geodésicas son funciones exponenciales:*

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

Vemos que dados cualquier $p \in \mathbb{R}^n_{++}$ y $v \in \mathbb{R}^n$, la geodésica $\alpha(t)$ está definida para todo $t \in \mathbb{R}$.

a.8 Curvatura de una variedad riemanniana

Un resultado importante sobre una variedad riemanniana es la curvatura, sobre el que podemos referirnos en forma intuitiva; es una medida de la variación de un vector por transportes paralelos, otra manera de expresar lo anteriro es, cuánto se aleja la curva de ser euclidiana o, que la curvatura riemanniana es la generalización natural de la curvatura Gaussiana de las superficies, por ejemplo cuando considere que la variedad $\mathcal{M} = \mathbb{R}^n$ la curvatura $K = 0$. Como veremos en más adelante, en relación a una variedad $\mathbb{R}^n_{++}$ y usando una métrica dada por (II.15) para una función diferenciable $h_i : \mathbb{R}_{++} \longrightarrow \mathbb{R}_{++}$ y $h_i : (0, 1)^n \longrightarrow \mathbb{R}_{++}$ respectivamnte, tiene curvatura cero.

Veamos seguidamente la siguiente definición, si consideramos un $A(\mathcal{H},\mathcal{H})$ como el conjunto de aplicaciones de $\mathcal{H}$ en $\mathcal{H}$ y ∇ la conexión afin en una variedad riemanniana M, dada por el teorema de Levi-Civita.

Definición 34. *Una curvatura K de una variedad riemanniana M es una correspondencia*

$$K : \mathcal{H} \times \mathcal{H} \longrightarrow \mathcal{A}(\mathcal{H},\mathcal{H})$$

definida por:

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

Observación 6. *Si la variedad $M = \mathbb{R}^n$, entonces $K(X,Y)Z = 0$, para todo $X,Y,Z \in \mathcal{H}$. En efecto, basta indicar $Z = (z_1, z_2, ..., z_n)$ las componentes del campo Z en las coordenadas naturales de $\mathbb{R}^n$ y la conexión definida por $\nabla_X Z = (Xz_1, Xz_2, ..., Xz_n)$.*

Observación 7. *Si consideramos un sistema de coordenadas $(U, \mathcal{X})$ en torno del punto p y $\{X_i\}, i = 1, 2, ..., n$ es una base de T_pM obtenemos:*

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

Observación 8. *La curvatura K es antisimétrica. En efecto,*

$$K(X,Y)Z + K(Y,X)Z = \nabla_{[X,Y]} Z + \nabla_{[Y,X]} Z, \; para\,todo \; Z \in \mathcal{H}.$$

Como $[X,Y] = -[Y,X]$, entonces:

$$K(X,Y)Z + K(Y,X)Z = 0, \; para\,todo\, Z \in \mathcal{H},$$

y así,

$$K(X,Y) = -K(Y,X).$$

Proposición 13. *La curvatura K de una variedad riemanniana es trilineal, en el siguiente sentido:*

a. K es bilineal en $\mathcal{H} \times \mathcal{H}$, esto es,

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

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

donde $f, g \in D(M)$ y $X_1, X_2, Y_1, Y_2 \in \mathcal{H}$.

b. Para todo par $X, Y \in \mathcal{H}$, el operador curvatura $K(X,Y) : \mathcal{H} \longrightarrow \mathcal{H}$ es lineal, esto es,

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

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

donde $f \in D(M)$, $\quad Z, W \in \mathcal{H}$.

Proposición 14. *Sea $(U, \mathcal{X})$ un sistema de coordenadas en torno de $p \in \mathcal{M}$ y $\{X_i\}$ una base de $T_p\mathcal{M}$ 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} son dadas por:

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

Observación 9. *Si en las coordenadas $(U, \mathcal{X})$, escribimos: $X = \sum_{i=1}^{n} u^i X_i$, $Y =$*

$\sum\limits_{j=1}^{n} v^j X_j$, $Z = \sum\limits_{k=1}^{n} w^k X_k$, *por la linealidad de K tenemos:*

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

Ejemplo 9. *Tómese los espacios $\mathcal{M}$ como $\mathcal{M} = \mathbb{R}^n_{++}$ y $\mathcal{M} = C^n_0$, con estructura de variedad riemanniana representada por la matriz $G(x)$ cuyos símbolos de Christoffel son:*

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

Escribiendo en coordenadas espaciales $(U,\mathcal{X})$: $X = \sum\limits_{i=1}^{n} u^i X_i$, $Y = \sum\limits_{j=1}^{n} v^j X_j$, $Z = \sum\limits_{k=1}^{n} w^k X_k$, de la tri-linealidad de K tenemos:

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

Aplicando la definición de curvatura:

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

como la conexión es de Levi Civita se tiene $[X_i, X_j] = 0$. Así,

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

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

Supongamos que $i \neq j$, entonces

$$\nabla_{X_i} X_k = \sum_{j=1}^{n} \Gamma_{ik}^{j} X_j.$$

Sustituyendo los símbolos de Christoffel tenemos:

$$\nabla_{X_i} X_k = \sum_{j=1}^{n} \left(-\frac{1}{h_i(x_i)} \frac{\partial (h_i(x_i))}{\partial x_i} \delta_{ij} \delta_{ik} \right) X_j = -\frac{1}{h_i(x_i)} \frac{\partial (h_i(x_i))}{\partial x_i} \delta_{ik} X_i \qquad \text{(II.19)}$$

Aplicando ∇_{X_j} *se tiene:*

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

por definición de conexión afin $\nabla_X(fY) = f\nabla_X Y + X(f)Y$ *donde* $X(f) = \sum_{i=1}^{n} a_i(.)\frac{\partial f}{\partial x_i}$, *entonces tenemos:*

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

Usando (II.19) y dado que i $\neq$ *j, el primero y segundo término de la suma anterior, es igual a cero. Por tanto*

$$\nabla_{X_j}(\nabla_{X_i} X_k) = 0.$$

Análogamente,

$$\nabla_{X_i}(\nabla_{X_j} X_k) = 0.$$

De ambos resultados se tiene:

$$K(X_i, X_j)X_k = 0, \forall\, i, j, k = 1, 2, ...n.$$

Así $K(X,Y)Z = 0$. *Luego las variedades riemannianas* $\mathbb{R}^n_{++}$ *y* C^n_0 *con métrica* $G(x)$ *tienen curvatura cero. En particular, con las métricas* I, X^{-r}, *respectivamente, para* $\mathbb{R}^n_{++}$ *y cosec*$^4(\pi x)$, $X^{-r}(I - X)^{-r}$ C^n_0, *son variedades de curvatura cero.*

- **Curvatura Seccional**

  Como ya vimos, una propiedad intrinseca de las superficies es la curvatura y esta a su vez está ligada a la curvatura riemanniana o curvatura seccional.

Entonces consideremos una variedad riemanniana $\mathcal{M}$ y un subespacio bidimensional σ del espacio tangente $T_p\mathcal{M}$, para definir la forma cuadrática como: $Q : \sigma \to \mathbb{R}$ tal que

$$Q(x, y) = \langle x, x\rangle\langle y, y\rangle - \langle x, y\rangle^2.$$

Proposición 15. *Sea el subespacio σ un subespacio de $T_p\mathcal{M}$ y tomando $x, y \in \sigma$, dos vectores linealmente independientes. Entonces,*

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

La curvatura K no depende de cuales sean los vectores x y y.

Demostración. Para la demostración, véase ver (Do Carmo, 1988).

Ahora veamos la definición de curvatura seccional, que será usada como condición para la aplicación del planteamiento del problema de optimización sobre variedades en lo que respecta a la convergencia de la método.

Definición 35. . *Dado un punto $p \in \mathcal{M}$ y $\sigma \subset T_p\mathcal{M}$. El número $K(x, y) = K(\sigma)$, donde $\{x, y\}$ es una base de σ, es llamado "Curvatura Seccional de $\mathcal{M}$".*

Si $K(x, y) \leq 0$ para todo $x, y \in \sigma$ entonces, la curvatura seccional de la variedad riemanniana es no positiva.

Si $K(x, y) \geq 0$ para todo $x, y \in \sigma$ entonces, la curvatura seccional de la variedad riemanniana es no negativa.

a.9 Gradiente y Hessiana en una variedad riemanniana

La intensión en esta sección es ver algunons operadaores matemáticos, que pueden ser usado para distintas aplicaciones sobre una variedad riemanniana.

Entonces, sea $\mathcal{M}$ una variedad riemanniana y f tal que $f : M \longrightarrow \mathbb{R}$ una función diferenciable. Para un punto $p \in \mathcal{M}$, la diferencial de f en dicho punto p es un funcional lineal definido sobre el espacio $T_p\mathcal{M}$, entonces por el teorema de representación de Riesz existe un único elemento denotado por $\nabla_{\mathcal{M}} f(p) \in T_p\mathcal{M}$ tal que para todo $v \in T_p\mathcal{M}$ se tiene

$$df_p(v) = \langle \nabla_{\mathcal{M}} f(p), v \rangle \tag{II.20}$$

y

$$\|\nabla_M f(p)\| = \|df_p\|,$$

esto es, la aplicación diferencial se puede caracterizar por la aplicación de producto interno. Así, es posible definir un campo vectorial $grad\ f : \mathcal{M} \longrightarrow T\mathcal{M}$, como

$$grad\ f(p) = \nabla_M f(p).$$

La expresión (II.20) puede ser escrita como

$$df_p(X(p)) = \langle grad\ f(p), X(p) \rangle, para\, todo\ X \in \mathcal{H},$$

y así también se puede definir una aplicación $df : \mathcal{H} \longrightarrow \mathcal{M}^* = \mathcal{L}(\mathcal{M}, \mathbb{R})$, donde $\mathcal{L}(M, \mathbb{R})$ es el conjunto de funciones en $\mathcal{M}$ en $\mathbb{R}$, tal que

$$df(X) = \langle grad\ f, X \rangle.$$

Además, $df_p(X(p)) = \frac{d}{dt}(f \circ \gamma)|_{t=0}$ para alguna curva $\gamma : I \longrightarrow M$ con $\gamma(0) = p$ y

$\gamma'(0) = X(p)$, luego tenemos que $df_p(X(p)) = \sum_{i=1}^{n} \gamma_i'(0)\frac{\partial f}{\partial x_i}(p) = X(f)(p)$, por tanto

$$df(X) = \langle grad\ f, X\rangle = X(f).$$

Así llegamos a la siguiente definición.

Definición 36. *El gradiente de una función diferenciable $f : \mathcal{M} \longrightarrow \mathbb{R}$ es un campo vectorial $grad\ f : \mathcal{M} \longrightarrow T\mathcal{M}$ metricamente equivalente a la diferencial, esto es,*

$$df_p(X(p)) = \langle grad\ f(p), X(p)\rangle = X(p)f,\ para\, todo\ X \in \mathcal{H}.$$

Observación 10. *Sea $\mathcal{M} \subset \mathbb{R}^n$ una variedad riemanniana con la métrica definida por $\langle v, w\rangle_x = v^T G(x)w$ donde $G(x)$ es una matriz simétrica definida positiva. Se puede caracterizar el campo gradiente como*

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

donde $G^{-1}(q) = (g^{ij}(q))$ es la matriz inversa de $G(q)$ y $f' = (\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n})$ es el vector de derivadas parciales de la función $f \circ \mathcal{X}$. En efecto,

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

$$= \left\langle G(q)^{-1} f'(q), v\right\rangle_q.$$

Ejemplo 10. *Sea la variedad riemanniana $\mathbb{R}^n_{++}$ con la expresión de la métrica*

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

para funciones $h_i : \mathbb{R}_{++} \longrightarrow \mathbb{R}_{++}$

$$grad\, f(x) = diag(h_1(x_1))^2, ..., (h_n(x_n))^2)f'(x).$$

En particular:

a) Si $h_i(x_i) = x_i$ entonces

$$grad\, f(x)f(x) = X^2 f'(x),$$

donde denotamos $X = diag(x_1, ..., x_n)$.

b) Si $h_i(x_i) = x_i^{\frac{r}{2}}, r \neq 2$ entonces

$$grad\, f(x)f(x) = X^r f'(x).$$

A continuació se calcula el punto crítico a través del Gradiente y la Hessiana, véase.

Definición 37. *Sea $\mathcal{M}$ una variedad riemanniana y $p \in \mathcal{M}$. Decimos que p es punto crítico si $grad\, f(p) = 0$.*

Definición 38. *Sea $f : \mathcal{M} \longrightarrow \mathbb{R}$ una función de clase C^k, $k \geq 2$. La Hessiana de f, denotada por H^f, es definida como la derivada covariante del campo gradiente, esto es,*

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

Así, la Hessiana en el punto p, en la dirección de $v \in T_pM$ es:

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

A partir del concepto de Hessiana podemos definir las aplicaciones $H_p^f : T_p\mathcal{M} \longrightarrow$

$T_p\mathcal{M}$ y $H^f : \mathcal{M} \longrightarrow \mathcal{L}(T\mathcal{M}, T\mathcal{M})$ donde $\mathcal{L}(T\mathcal{M}, T\mathcal{M})$ es el conjunto de aplicaciones lineales de $T\mathcal{M}$ en $T\mathcal{M}$ y $H^f(p) = H_p^f \in \mathcal{L}(T_p\mathcal{M}, T_p\mathcal{M})$.

Proposición 16. *Para cada $p \in \mathcal{M}$, el operador $H_p^f : T_p\mathcal{M} \longrightarrow T_p\mathcal{M}$ es lineal y autoadjunto, esto es, $\langle H_p^f(v), w\rangle_p = \langle v, H_p^f(w)\rangle_p$.*

De esta, para cada $p \in \mathcal{M}$ podemos introducir una forma cuadrática $q_p^f : T_p\mathcal{M} \times T_p\mathcal{M} \longrightarrow \mathbb{R}$ definida por

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

Más generalmente, podemos definir la aplicación $q^f : \mathcal{H} \times \mathcal{H} \longrightarrow \mathcal{L}(\mathcal{M}, \mathbb{R})$ dada por:

$$q^f(X, Y) = \langle \nabla_X grad\ f, Y\rangle. \tag{II.21}$$

La función definida en (II.21) tiene la desventaja de depender del conocimiento de la métrica y de la conexión, cuando sabemos que la métrica determina una conexión afin (Teorema de Levi Civita), por tanto la proposición siguiente es importante para poder obtener una caracterización adecuada.

Proposición 17. *Para todo $X, Y \in \mathcal{H}$*

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

Observación 11. *En un sistema de coordenadas $(\mathcal{X}, U)$ en terminos de la base $\{X_k\}$ tenemos:*

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

esto es:

$$q^f(X_i, X_j) = \langle H_p^f X_i, X_j\rangle = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} - \sum_{m=1}^{n} \Gamma_{ij}^m \frac{\partial f}{\partial x_m}\right). \tag{II.22}$$

Ejemplo 11. *Sea la variedad riemanniana $\mathbb{R}^n$ con métrica $G(x) = I$; como vimos anteriormente, los símbolos de Christoffel son $\Gamma_{ij}^m = 0$, para todo $i, j, m = 1..., n$,* *entonces la matriz Hessiana es la Hessiana usual* $H_p^f(p) = f''(p)$.

Ejemplo 12. *Sea la variedad riemanniana* $\mathbb{R}^n_{++}$ *con la métrica* $g_{ij} = \frac{\delta_{ij}}{h_i(x_i)h_j(x_j)}$. *Sabemos que los símbolos de Christoffel son*

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

entonces

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

Si $m \neq j$ *entonces* $\delta_{im}\delta_{ij} = 0$, *luego se tiene*

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

Así $H_x^f = \left(q^f(X_i)(X_j)\right)$ *es la matriz que representa la Hessiana de la función* f. *Aún podemos dar una representación matricial*

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

donde:

$\mathcal{F}'(x) = diag\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, ..., \frac{\partial f(x)}{\partial x_n}\right)$

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

$f''(x) = diag(\frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_2^2}, ..., \frac{\partial^2 f}{\partial x_n^2})$

En particular:

a) Si $h_i(x_i) = 1$ entonces $H_x^f = f''(x)$ la matriz Hessiana usual.

b) Si $h_i(x_i) = x_i$ entonces $H_x^f = f''(x) + X^{-1}\mathcal{F}'(x)$

c) Si $h_i(x_i) = x_i^{\frac{r}{2}}$, $r \neq 2$, entonces $H_x^f = f''(x) + \frac{r}{2}X^{-1}\mathcal{F}'(x)$

Corolario 2. *Si $p \in M$ es un punto crítico de f y $X, Y \in \mathcal{H}$, entonces*

$$H_p^f(X(p), Y(p)) = X(p)Y(p)f.$$

Demostración. $H_p^f(X(p), Y(p)) = X(p)(Y(p)f) - \langle \nabla_{X(p)}Y(p), grad\ f(p)\rangle$, y como $grad\ f(p) = 0$, se sigue el Corolario. ∎

De este corolario, se deduce que si $p \in M$ es un punto crítico de f entonces la matriz Hessiana de f, calculada en este punto, coincide con la matriz Hessiana usual.

- **Variedades completas**

Para hacer distintas aplicaciones en el ámbito de la optimización sobre geometría riemanniana, es necesario considerar variedades completas es decir si sobre la variedad existen dos puntos conectados por una geodésica, ésta es minimizante entre dichos puntos.

Según Quispe C. (2008) se puede desarrollar métodos geodésicos para optimización matemática, donde la hipótesis de variedad completa sea más suave, la idea definir alguna una forma de medir sobre la variedad a partir del producto interno del espacio tangente donde la geodésica está definida en casi todos los puntos, y los puntos donde la geodésica no sea definida, pertenezca a un conjunto de medida nula.

Se usarán los resultado siguientes sobre variedades completas

Definición 39. *Una variedad riemanniana $\mathcal{M}$ es llamada (geodésicamente) completa si para todo $p \in \mathcal{M}$, las geodésicas que parten de p estan definidas para todos los valores del parámetro $t \in \mathbb{R}$.*

Ejemplo 13. *El espacio euclidiano n-dimensiona, $\mathbb{R}^n$ con la métrica euclidiana $G(x) = I$, es (geodesicamente) completa, pues dado un punto cualquiera*

$x \in \mathcal{M}$ y una dirección arbitraria $v \in T_x\mathcal{M}$, vimos que la i-ésima componente de la geodésica que cumple las condiciones iniciales $\alpha_i(0) = x_i$ y $\alpha_i'(0) = v_i$, para todo $i = 1, ..., n$, es dado por: $\alpha_i(t) = x_i + tv_i$, para todo $i = 1, 2, ..., n$, lo que está definida para todo $t \in \mathbb{R}$.

A continuación definiremos la distancia riemanniana, para ello es necesario considerar que la variedad riemanniana tenga la propriedad de conexidad, esto es, para puntos $p, q \in \mathcal{M}$, existe una curva diferenciable contenida en $\mathcal{M}$, $\gamma : [a, b] :\to \mathcal{M}$, tal que $\gamma(a) = p$ y $\gamma(b) = q$.

Definición 40. *Dados dos puntos p y q en $\mathcal{M}$, la distancia riemanniana de p a q en la variedad, denotada por $d(x, y)$, es definida por*

$$d(p, q) = \underset{\gamma}{Inf} \int_a^b \|\gamma'(t)\| dt \tag{II.23}$$

donde $\gamma : [a, b] \to \mathcal{M}$ es una curva diferenciable tal que $\gamma(a) = p$ y $\gamma(b) = q$.

Proposición 18. *Con la distancia geodésica (II.23) $\mathcal{M}$ es un espacio métrico.*

Teorema 2. *(El teorema de Hopf-Rinow) Tenemos a $\mathcal{M}$ como una variedad riemanniana y sea $p \in \mathcal{M}$. Afirmamos que las siguientes afirmaciones son equivalentes:*

a) La aplicación esp_p está definida en todo el espacio tangente a $\mathcal{M}$ en p

b) Limitados y cerrados son compactos.

c) $\mathcal{M}$ es completo como espacios métrico.

d) $\mathcal{M}$ es geodesicamente completa.

e) Para todo $q \in \mathcal{M}$ existe una geodésica uniendo p y q con

$$d(p, q) = \underset{\gamma}{Inf} \int_a^b \|\gamma'(t)\| dt$$

esto es, el mínimo de (II.23) es obtenida por una geodésica.

Demostración. Puede ver la demostración en (Do Carmo, 1988)

El siguiente ejemplo es estudiado en (Quispe C., 2008).

Ejemplo 14. *Sea la variedad riemanniana $\mathbb{R}^n_{++}$ con expresión de la métrica $G(x) = X^{-2}$. Dados p y q en $\mathbb{R}^n_{++}$, existe una única geodésica uniendo p a q. En efecto, sea $\gamma_i(t) = p_i exp(\frac{v_i t}{p_i})$ y $\alpha_i(t) = p_i exp(\frac{w_i t}{p_i})$ las i-esimas componentes que satisfazen*

$$\gamma_i(0) = \alpha_i(0) = p_i,$$

$$\gamma_i(t_0) = \alpha_i(t_0) = q_i.$$

Se puede verificar que $v_i = w_i$ para todo $i = 1, \ldots, n$. En efecto, tomando el valor $t = t_0$ tenemos que para todo $i = 1, \ldots, n$: $q_i = p_i exp(\frac{v_i t_0}{p_i}) = p_i exp(\frac{w_i t_0}{p_i})$ dividiendo por p_i, tomando logaritmo y multiplicando por p_i/t_0 tenemos $v_i = w_i$ y así $\gamma_i(t) = \alpha_i(t)$ para todo $t \in \mathbb{R}$. Además:

a). Debido a que $q_i = p_i exp(\frac{v_i}{p_i})$ entonces $v_i = p_i ln(\frac{q_i}{p_i})$.

b). $\frac{\gamma_i'(t)}{\gamma_i(t)} = \frac{v_i}{p_i}$ entonces $\left(\frac{\gamma_i'(t)}{\gamma_i(t)}\right) = \frac{1}{t_0} ln^2(\frac{q_i}{p_i})$ así:

$$d(p,q) = \int_0^{t_0} \|\gamma'(t)\| dt = \left\{ \sum_{i=1}^{n} \left[ln\left(\frac{q_i}{p_i}\right)\right]^2 \right\}^{\frac{1}{2}}.$$

Teorema 3. *Ley de cosenos. Sea $\mathcal{M}$ una variedad riemanniana completa con curvatura seccional no negativa, en un triángulo geodésico normalizado tal que $\gamma_1, \gamma_2, \gamma_3$ segmento de geodésicas minimizantes. Vale la desigualdad*

$$c^2 \leq a^2 + b^2 - 2a\, b\, cos\alpha \tag{II.24}$$

donde $\alpha = arg(\gamma_1'(0). - \gamma_3'(l_3))$, $a = L(\gamma_1)$, $b = L(\gamma_3)$, $c = L(\gamma_2)$, L, longitud de geodésica

Demostración. Puede ver la demostración de este resultado en (Da Cruz Neto y col., 1998).

2.3.2. Método del Gradiente sobre variedades diferenciables

2.3.2.1. Método del Gradiente

El método del gradiente, es uno de los métodos computacionales clásicos en el campo de la optimización matemática y otros, para modelos no lineales, estudiado inicialmente por Hestenes y Stiefel según Burden y Faires (2011), fué un método para solucionar sistemas de ecuaciones lineales que se presentan en diversos campos de la ciencias en el que se deba emplear algoritmos para su simulación.
Los pasos asociados a la utilización del método del Gradiente o descenso más pronunciado consiste en:

- Paso 1: Considere un punto inicial $x = x^0$. Hacer $k = 0$
  Paso 2: Escoger una dirección de descenso $d^k = -\nabla f(x^k)$
  Paso 3: Realizar una búsqueda lineal que seleccione un paso α_k tal 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: Hacer un test de convergencia (Por ejemplo $\|\nabla f(x^k)\| < \epsilon$). Si converge se detiene el método. En caso contrario hacer $k = k + 1$ y volver al paso dos.

Lo que vimos en esta investigación es el análisis de su convergencia, usado para extender a su vez la convergencia global del método, sobre una variedad Riemanniana, utilizando la regla de Armijo generalizado, considerando el siguiente problema de Optimización no lineal

$$\min_{x \in \mathcal{M}} f(x) \tag{II.25}$$

donde, $f : \mathbb{R}^n \to \mathbb{R}$ es una función de clase C^1 y $\mathcal{M}$ una variedad riemanniana completa.

Para la generalización del método del Gradiente de espacios euclidianos a variedades Riemannianas con retracciones, vimos en la formulación del problema la siguiente expresión: Dado un punto $x_0 \in T_x^k \mathcal{M}$ se tiene

$$x_{k+1} = \mathcal{R}_{x_k}(t_k \eta_k) \tag{II.26}$$

donde $\eta_k \in T_{x_k}\mathcal{M}$, espacio tangente a $\mathcal{M}$ en el punto x_k, t_k es un escalar tal que siguiendo la dirección de búsqueda de Armijo, genera la sucesión x_{k+1} y $\mathcal{R}$ es la aplicación llamada retracción, (Absil y col., 2008).

Como se sabe, el método del gradiente sobre la variedad, genera una sucesión de puntos $\{x^k\}$ dados por:

$$x^0 \in \mathcal{M}, \tag{II.27}$$

$$x^{k+1} = \exp_{x^k}(-t_k grad\, f(x^k)) \tag{II.28}$$

donde exp_{x^k} es una aplicación exponencial en el punto x^k, t_k es un parámetro positivo, $-grad\, f(x)$ es el gradiente de f, según (Quispe Cárdenas y col., 2007, 1). En el caso de tener $M = \mathbb{R}^n$ (el espacio euclidiano) tenemos que (II.28) es equivalente a:

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

Así, el método de máximo descenso en variedades riemannianas generaliza el método clásico de máximo descenso en $\mathbb{R}^n$.

2.3.2.2. Retracciones sobre variedades

Las retracciones constituyen uno de los elementos fundamentales en el desarrollo de ésta investigación, estos son mapeos, llamados retracciones, entre $\mathcal{M}$ y $T_x\mathcal{M}$, que se pueden usar para transformar localmente un problema de optimización en variedades en un problema de optimización en el espacio vectorial más amigable (Absil y col. (2008), p.

34). Esta se define como una aplicación regular sobre el espacio tangente a una variedad $\mathcal{M}$, y es en particular una aplicación expoencial, vea (Absil y col. (2008), p. 103). Como veremos más adelante, está a su vez relacionado con la geodésica, como se muestra en (Boothby (1975), p. 334). Ahora estudiamos las retracciones y sus propiedades.

Definición 41. *(Retracción). Definimos una retracción $\mathcal{R}$ sobre una variedad $\mathcal{M}$ como una aplicación, regular del fibrado tangente $T\mathcal{M}$ sobre la variedad $\mathcal{M}$, el cual satisface las propiedades siguientes:*

1. *$\mathcal{R}_x(0_x) = x$, y 0_x denota el elemento cero del espacio $T_x\mathcal{M}$*
2. *Para la identificación canonica $T_{0_x}T_x\mathcal{M} \simeq T_x\mathcal{M}$, $\mathcal{R}_x$ satisface la igualdad*

$$\mathrm{D}\mathcal{R}_x(0_x) = \mathrm{id}_{T_x\mathcal{M}} \tag{II.29}$$

$\mathrm{id}_{T_x\mathcal{M}}$ *denota la aplicación identidad sobre $T_x\mathcal{M}$*

Como $T_x\mathcal{M}$ es un espacio vectorial, existe una identificación natural del espacio tangente en cero 0_x del propio espacio tangente $T_x\mathcal{M}$ con el mismo espacio tangente $T_x\mathcal{M}$, esto es $T_{0_x}(T_x\mathcal{M}) \simeq T_x\mathcal{M}$ convirtiendo elementos de $T_x\mathcal{M}$ en puntos de $\mathcal{M}$, de ahi importancia de las retracciones, como lo veremos mas adelante.
Observe también que $\mathcal{R}_x$ es la restricción de la retracción $\mathcal{R}$ sobre el espacio tangente $T_x\mathcal{M}$ en el punnto x, con condición de rigidez local que preseva el gradiente en x. Asumimos que el dominio es todo el fibrado tangente $T\mathcal{M}$.
Del punto a), la Retracción $\mathcal{R}_x$ envía el 0_x a x y se sigue que

$$\mathrm{D}\mathcal{R}_x(0_x)$$

es una aplicación del $T_{0_x}(T_x\mathcal{M})$ a $T_x\mathcal{M}$, veáse la siguiente definición

Definición 42. *Considere la aplicación regular F entre dos variedades riemannianas $\mathcal{M}, \mathcal{N}$, esto es $F : \mathcal{M} \to \mathcal{N}$. Y sea ξ_x un vector tangente en el punto $x \in \mathcal{M}$, se puede* *mostrar que la aplicación* $\mathrm{D}F(x)[\xi_x]$ *de* $\mathfrak{F}_{F(x)}(\mathcal{N})$ *en* $\mathbb{R}$ *definido por:*

$$(\mathrm{D}F(x)[\xi]f) := \xi(f \circ F) \tag{II.30}$$

es un vector tangente a $\mathcal{N}$ *en* $F(x)$

El vector tangente $\mathrm{D}F(x)[\xi_x]$ es calculado por $F \circ \gamma$ y γ trazado por ξ_x. La aplicación

$$\mathrm{D}F(x) : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$$

$$\xi \mapsto \mathrm{D}F(x)[\xi]$$

es una aplicación; valga la redundancia, denominada el diferencial, Absil (2008) también la denomina, aplicación diferencial, derivada o aplicación tangente de F en x.

Aquí, hacemos un paréntesis para explicar que $\mathfrak{F}_x(\mathcal{M})$ es el conjunto de funciones de valores reales, suave definidas sobre una vecindad de x.

Volviendo al punto. Afirmamos que F es una inmersión si y sólo si $\mathrm{D}F(x) : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ es inyectiva para cada $x \in \mathcal{M}$, y al ser $\mathcal{N}$ un espacio vectorial ξ, se tiene la identificación canonica $T_{F(x)}\xi \simeq \xi$ nos da

$$\mathrm{D}F(x)[\xi_x] = \sum_i (\xi_x F^i) e_i \tag{II.31}$$

a su vez $F(x) = \sum_i F^i(x) e_i$ es la descomposición de $F(x)$ en la base $(e_i)_{i=1,\cdots,n}$. Este resultado permite ser expuesto particularmente cuando $\mathcal{N} = \mathbb{R}$, entonces $F \in \mathfrak{R}_x(\mathcal{M})$ y se tiene una expresión más simple como

$$\mathrm{D}F(x)[\xi_x] = \xi_x F \tag{II.32}$$

usando la identificación $T_x\mathbb{R} \simeq \mathbb{R}$.

Para una mayor aclaración de la ecuación (II.31), que viene de la identificación canónica

$T_{F(x)}\xi \simeq \xi$ podemos revisar (Absil y col., 2008).

Finalmente, respecto al operador $\mathrm{D}F(x)[\xi_x]$, podemos usar la notación $\xi_x F$ que expresa con mayor énfasis el referirno a la derivada. Esto es, si $\mathcal{M}$ y $\mathcal{N}$ son variedades lineales, se tiene la identificación del espacio tangente a una variedad con la propia variedad, osea $T_x\mathcal{M} \simeq \mathcal{M}$, del mismo modo para $T_y\mathcal{N} \simeq \mathcal{N}$, y podemos usar la expresión $\mathrm{D}F(x)$ que simplifica la definición natural de derivada de una función, así

$$\mathrm{D}F(x)[\xi_x] = \lim_{\iota \to 0} \frac{F(x + \iota\xi_x) - F(x)}{\iota} \tag{II.33}$$

Ahora, dado una función diferenciable $F : \mathcal{M} \mapsto \mathcal{N}$ y el campo vectorial ξ sobre $\mathcal{M}$ denotamos con:

$$\mathrm{D}F[\xi] : \mathcal{M} \to T\mathcal{N}$$

$$x \mapsto \mathrm{D}F(x)[\xi_x]$$

en particular si tenemos $f : \mathbb{R} \to \mathcal{M}$ y el campo vectorial ξ sobre $\mathcal{M}$, escribimos simplemte $\mathrm{D}f(\xi) = f\xi$.

Las **retracciones** tiene otros aspectos importantes para haberlo propuesto y ser usado para obtener los objetivos planteados en la investigación que estamos presentando, esto es, la retracción $\mathcal{R}_x$ transforma la función costo definida en una vecindad de x de $\mathcal{M}$ y también en funciones costo definido sobre el espacio vectorial $T_x\mathcal{M}$, esta aseveración es formalizada en la siguiente definición.

Definición 43. *Dado una función f de valor real sobre una variedad $\mathcal{M}$ $f : \mathbb{R} \to \mathcal{M}$, desde de que $\mathcal{M}$ está equipada con una retracción $\mathcal{R}$, y sea $\hat{f} = f \circ \mathcal{R}$ que denota la vuelta de f a través de $\mathcal{R}$.*

Restringiendo la definición anterior a $x \in \mathcal{M}$ se tiene

$$\hat{f}_x = f \circ \mathcal{R}_x \tag{II.34}$$

$\hat{f}_x$ es una función de valor real sobre el espacio vectorial, y denota la restricción de $\hat{f}$ al espacio tangente $T_x\mathcal{M}$

Observación 12. *Debido a la condición de rigidez local (II.29) tenemos la identificación respecto la derivada,* $\mathrm{D}\hat{f}_x(0_x) = \mathrm{D}f(x)$ *y si la variedad* $\mathcal{M}$ *está munida con una métrica Riemanniana, dado por el producto interno en* $T_x\mathcal{M}$ *tenemos:*

$$\mathrm{grad}\hat{f}_x(0_x) = \mathrm{grad}f(x) \tag{II.35}$$

Observación 13. *En 2008, Absil sostiene que toda variedad que admite una métrica Riemanniana también admite una retracción definida por la aplicación exponencial Riemanniana.*

Seguidamente, véase cómo se interrelacionan tanto, geodésicas, la propia aplicación exponencial y el transporte paralelo.
Recuerde que una geodésica γ sobre una variedad $\mathcal{M}$ dotado con una conexión afin ∇ es una curva con aceleración nula, esto es

$$\frac{\mathrm{D}^2}{dt^2}\gamma(t) = 0 \tag{II.36}$$

para todo t en el dominio de la curva γ.

Definición 44. *Para cada* $\xi \in T_x\mathcal{M}$ *existe un intervalo* I *alrededor de 0 y una unica geodésica* $\gamma(t, x, \xi) : I \to \mathcal{M}$*, tal que* $\gamma(0) = x$ *y* $\dot{\gamma}(0) = \xi$*.*

La propiedad de homogeneidad $\gamma(t, x, a\xi) = \gamma(at, x, \xi)$ y la aplicación

$$\exp_x : T_x\mathcal{M} \to \mathcal{M}$$

$$\xi \mapsto \exp_x\xi = \gamma(1, x, \xi)$$

llamada la aplicación exponencial en x juegan un rol importante en este desarrollo, veamos esos conceptos, según (Do Carmo, 1988).
El lema siguiente, establece la homogeneidad de geodésicas, en el sentido que se puede aumentar la velocidad de la geodésica, al disminuir su intervalo de definición.

Lema 1. *Si la geodésica $\gamma(t, q, v)$ está definida en el intervalo $(-\delta, \delta)$ entonces la geodésica $\gamma(t, q, av)$ está definida en el intervalo $(-\frac{\delta}{a}, \frac{\delta}{a})$, con $a \in \mathbb{R}$ y $a > 0$, y se tiene la igualdad siguiente:*

$$\gamma(t, q, av) = \gamma(at, q, v)$$

Demostración. Ver (Do Carmo, 1988). ■

Proposición 19. *Dado $q \in \mathcal{M}$, existe un $\epsilon > 0$ tal que $exp_q : B_\epsilon(0) \subset T_q\mathcal{M} \to \mathcal{M}$ es un difeormorfismo de $B_x(0)$ sobre un abierto de $\mathcal{M}$*

Demostración. Ver (Do Carmo, 1988). ■
La siguiente proposición es importante en optimización de algoritmos sobre variedades, en relación a la conexión afin, (que, como hemos visto se usan para definir geodésicas en una variedad, de tal manera que se puede generalizar líneas rectas del espacio euclidiano) con la aplicación retracción que estamos proponiendo y juega un rol importante en la demostración de la convergencia del método del Gradiente. Esta proposición es tomada de (Absil y col., 2008).

Proposición 20. *Sea $\mathcal{M}$ una variedad dotada con una conexión afín ∇. La aplicación exponencial sobre $\mathcal{M}$ inducido por ∇ es una retracción, llamada retracción exponencial.*

Demostración. Puede ver el análisis que realiza Absil, (Absil y col. (2008), 103) ■.
Esta proposición generaliza o extiende el concepto del movimiento lineal siguiendo la dirección del vector tangente y permite de forma natural la actualización de la iteración dada en dirección de búsqueda en el espacio tangente a la variedad dada.

Finalmente, podemos decir que las retracciones permite proporcionar una alternativa generalizada respecto al operador exponencial en el diseño de algoritmos numéricos conservando las propiedades, claves que aseguran los resultados de convergencia; esto debido a las dificultades computacionea en calcular el exponencial que involucra a su vez el cálculo de la solución de una ecuación diferencial ordinaria de segundo orden y ésta a su vez, para obtener los símbolos de Cristoffel. Seguidamente planteamos el algoritmo con retracciones.

2.3.2.3. Algoritmo con Retracciones

Entonces, de las condiciones antes vistas, tenemos el algoritmo de minimización de búsqueda lineal sobre variedades. Este método involucra una aproximación de la búsqueda lineal a través de retracciones, tomando un punto del espacio tangente $T_{x_k}\mathcal{M}$ vía la retracción y lo lleva o traslada a la variedad $\mathcal{M}$.

Algoritmo G: Algoritmo de búsqueda lineal con retracciones

Ingrese: $f : \mathcal{M} \to \mathbb{R}, x_0 \in \mathcal{M}, k = 0$

repetir

Elegir la dirección de descenso $p_k \in T_{x_k}$

Elija la retracción $R_{x_k} : T_{x_k}\mathcal{M} \to \mathcal{M}$

Escoja una longitud de paso $\alpha_k \in \mathbb{R}$

Calcular $x_{k+1} = R_{x_k}(\alpha_k p_k)$

$k \leftarrow k + 1$

Hasta que x_{k+1} alcance el mínimo de f.

2.3.3. Condiciones de Optimalidad

Una candidad de fenómemos en general son estudiados a través de modelos matemáticos, En particular el modelo que representa un problema de optimización a estudiarse en lo sucesivo. Para estudiar dichos modelos se hace imprescindible se garantice qué condiciones se tiene para la existencia y caracterización de mínimos, con fines de plantear un algoritmo conveniente tal que me genere una sucesión de puntos convergentes a la solución del problema modelo de optimización, se define la variedad convexa y estudiamos una clase particular de funciones llamadas convexas y cuasi-convexas. Para ello, veamos las sguiente definiciones.

Para el desarrollo de este Capítulo, proponemos algunas definiciones elementales, muy necesarias para plantear la solución un problema de optimización.

Definición 45. *(Mínimo: global, local, estricto). Sea $\mathcal{M}$ una variedad riemanniana completa y $f : \mathcal{M} \longrightarrow \mathbb{R}$ una función.*

1. *$\bar{x} \in \mathcal{M}$ es un mínimo global de f si, $f(\bar{x}) \leq f(x), \ \ para \ todo \ \ x \in \mathcal{M}$.*

2. *$\bar{x} \in \mathcal{M}$ es un mínimo local de f si, existe $\delta > 0$ tal que*

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

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

3. *$\bar{x} \in \mathcal{M}$ es 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 interés será resolver el siguiente modelo:

$$\begin{aligned} &min \ f(x) \\ &x \in \mathcal{M} \end{aligned} \tag{II.37}$$

que significa encontrar los mínimos globale de una función f sobre $\mathcal{M}$, y es denominado “Problema de Minimización", sujeta generalmente a algunas restricciones sobre su dominio.

2.3.3.1. Existencia de mínimo global

Los siguientes resultados garantizan la existencia de mínimo global de una función.

Definición 46. *Una función $f : \mathcal{M} \longrightarrow \mathbb{R}$ es denominada semicontinua inferior en $\bar{x} \in \mathcal{M}$, si para toda sucesión $\{x^k\}$ de $\mathcal{M}$ convergente a $\bar{x}$ se tiene que:*

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

Si f es semicontinua inferior para todo $x \in \mathcal{M}$, entonces decimos que f es semicontinua inferior en $\mathcal{M}$.

El siguiente Teorema garantiza la existencia de un punto de mínimo global para el problema (II.37).

Teorema 4. *(Weierstrass) Considere el problema (II.37), si $f : \mathcal{M} \longrightarrow \mathbb{R}$ es semicontinua inferior y $\mathcal{M}$ es compacto, entonces existe un punto de mínimo global de f.*

Demostración. Mostraremos inicialmente que f es limitada inferiormente, esto es, existe $\alpha \in \mathbb{R}$ tal que

$$\alpha \leq f(x), \quad para\ todo\ x \in \mathcal{M}.$$

Por contradicción, supongamos que f no es limitada inferiormente, entonces existe una sucesión $\{x^k\} \subset \mathcal{M}$ tal que

$$\lim_{k\to+\infty} f(x^k) = -\infty. \tag{II.38}$$

Dado que $\mathcal{M}$ es compacto, entonces existe una subsucesión $\{x^{k_j}\} \subset \{x^k\}$ tal que

$$\lim_{j\to+\infty} x^{k_j} = \widehat{x},$$

por la semicontinuidad inferior de f tenemos

$$\liminf_{j\to+\infty} f(x^{k_j}) \geq f(\widehat{x}),$$

lo que contradice a (II.38), por lo tanto f es limitada inferiormente en $\mathcal{M}$. De aquí existe $f^* \in \mathbb{R}$ tal que $f^* = inf\{f(x) : x \in \mathcal{M}\}$. Por propiedad de ínfimo, existe una sucesión $\{x^k\} \subset \mathcal{M}$ tal que

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

Por la compacidad de $\mathcal{M}$, existe $\bar{x}$ y $\{x^{k_j}\} \subset \{x^k\}$ tal que $\lim_{j\to\infty} x^{k_j} = \bar{x} \in \mathcal{M}$. Nuevamente, por la semicontinuidad inferior de f

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

Como $\{f(x^k)\}$ converge a f^*, la subsucesión $\{f(x^{k_j})\}$ converge a f^* obteniendo que

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

así, $\bar{x}$ es un punto de mínimo global de f en $\mathcal{M}$. ■

2.3.3.2. Caracterización de mínimo local

Presentamos ahora las condiciones necesarias y suficientes de optimalidad local para el problema (II.37).

Teorema 5. *(Condición necesaria de primer orden). Sea $f : \mathcal{M} \to \mathbb{R}$ de clase C^1. Si x^* es un punto de mínimo local, entonces $grad\, f(x^*) = 0$.*

Demostración. Tomemos $v \in T_{x^*}\mathcal{M}$ y una curva geodósica $\gamma : \mathbb{R} \to M$ con condiciones $\gamma(0) = x^*$ y $\gamma'(0) = v$. Definamos la aplicación $h : \mathbb{R} \to \mathbb{R}$ tal que $h(t) = f(\gamma(t))$. Como

x^* es punto de mínimo local para f, entonces existe $\delta > 0$ tal que

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

para todo $t \in (-\delta, \delta)$ lo que implica que en $t = 0$ tenemos un punto de mínimo local de h. Por la condición necesaria de primer orden en $\mathbb{R}$ se tiene

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

Tomando en particular $v = grad\ f(x^*)$ tenemos que $grad\ f(x^*) = 0$. ■

Teorema 6. *(Condición necesaria de segundo orden). Sea $f : \mathcal{M} \to \mathbb{R}$ de clase C^2. Si x^* es punto de mínimo local, entonces $\langle v, H^f_{x^*} v \rangle \geq 0,\ \forall\ v \in T_{x^*}\mathcal{M}$.*

Demostración. Sea $v \in T_{x^*}\mathcal{M}$, y $\gamma : \mathbb{R} \to \mathcal{M}$ una geodésica con $\gamma(0) = x^*$, $\gamma'(0) = v$. Definimos $h : \mathbb{R} \to \mathbb{R}$ tal que $h(t) = f(\gamma(t))$. Del Teorema 5, en $t = 0$ tenemos un punto de mínimo local de h, entonces por la condición necesaria de segundo orden: $h'(0) = 0$, luego $h''(0) \geq 0$.

Veamos

$$\begin{aligned}
h'(t) &= \langle grad\ f(\gamma(t)), \gamma'(t) \rangle \\
h''(t) &= \tfrac{d}{dt} \langle grad\ f(\gamma(t)), \gamma'(t) \rangle \\
&= \left\langle \tfrac{D}{dt}(grad\ f(\gamma(t))), \gamma'(t) \right\rangle + \left\langle grad\ f(\gamma(t)), \tfrac{D}{dt}(\gamma'(t)) \right\rangle \\
&= \left\langle H^f_{\gamma(t)} \gamma'(t), \gamma'(t) \right\rangle \\
&= \left\langle v, H^f_{x^*} v) \right\rangle = \left\langle H^f_{x^*} v, v) \right\rangle \geq 0. \quad ■
\end{aligned}$$

Teorema 7. *(Condición suficiente de segundo orden). Sea $f : \mathcal{M} \longrightarrow \mathbb{R}$ de clase C^2. Si $x^* \in \mathcal{M}$ que satisface:*

a) $grad\ f(x^*) = 0$.

b) $H^f_{x^*}$ *definida positiva.*

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

Demostración. Por contradicción. Supongamos que x^* no es punto de mínimo local estricto, entonces existe una subsucesión $\{x^k\} \in B(x^*, \frac{1}{k})/\{x^*\}$ tal que

$$f(x^*) \geq f(x^k). \tag{II.39}$$

Sea la geodésica minimal $\gamma_k : [0,1] \to \mathbb{R}$ tal que $\gamma_k(0) = x^*$, $\gamma_k(1) = x^k$, $\gamma_k'(0) = v_k$ y $d(x^*.x^k) = \left\| exp_{x^*} x^k \right\|$. Definimos $h : \mathbb{R} \to \mathbb{R}$ tal que $h(t) = (f \circ \gamma_k)(t)$ y por el desarrollo de Taylor de segundo orden de h en 0 :

$$h(t) = h(0) + th'(0) + \frac{1}{2}t^2 h''(0) + \theta(|t|^2), \quad donde, \lim_{t\to 0} \frac{\theta(|t|^2)}{|t|^2} = 0,$$

esto es,

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

Evaluando en $t = 1$

$$f(x^k) = f(x^*) + \frac{1}{2}\left\langle v_k, H^f_{x^*} v_k \right\rangle + \theta(d^2(x^*, x^k)) \tag{II.40}$$

donde $\lim_{d\to 0} \dfrac{\theta(d^2(x^*, x^k)}{d^2(x^*, x^k)} = 0$.

Definamos $z^k = \frac{v_k}{\|v_k\|}$, la sucesión $\{z^n\}$ es limitada, entonces existe una subsucesión $\{z^{k_j}\} \subset \{z^k\}$ tal que $\{z^{k_j}\} \to \bar{z}$. Substituyendo en (II.40) k por k_j, tenemos

$$f(x^{k_j}) = f(x^*) + \frac{1}{2}\left\langle v_{k_j}, H^f_{x^*} v_{k_j} \right\rangle + \theta(d^2(x^*, x^{k_j})) \tag{II.41}$$

donde $\lim_{d\to 0} \dfrac{\theta(d^2(x^*, x^{k_j})}{d^2(x^*, x^{k_j})} = 0$.

De la relación (II.39) y tomando limite en (II.41) cuando $j \to \infty$, obtenemos

$$0 \geq \left\langle \bar{z}, H^{f}_{x^*}\bar{z} \right\rangle,$$

lo que contradice la hipótesis *b)* del Teorema, por tanto x^* es un punto de mínimo local estricto. ∎

2.3.4. Resultados básicos de convexidad y cuasi-convexidad

Como puede verse en Quispe C. (2008), afirma que sobre análisis convexo sobre variedades riemannianas existen varios investigadores, uno de ellos, Rapcsák, 1997 y Udriste, 1994, demuestran que en una variedad diferenciable con métrica induzida de $\mathbb{R}^n$, se obtiene caracterizaciones de primer y segundo orden para problemas de optimización sobre dicha variedad. Udriste (1994) en su texto, consideró el estudio sobre una variedad riemanniana abstracta generalizando (independientemente) la teoría de convexidad.

Respecto a la convexidad, o mejor expresado, el análisis convexo orientado a resolver problemas de Optimización matemática, se ha ido ampliando, desde entonces, como se puede verse, por ejemplo, el trabajo de investigación de Da Cruz Neto y col. (1998) donde, consideran que la variedad riemanniana, tiene que ser completa y con curvatura seccional no negativa, con estas ideas, se desarrolla y analiza básicamente la convexidad sobre una variedad Riemanniana.

Particularmente, en esta sub subsección se ofrece algunos resultado elementales sobre análisis convexo, importante para el análisis de optimalidad de un problema de programación matemática, en el cual se ha usado retracciones, $\mathcal{R}$, porque si se elije una buena retracción, y posibilita encontrar una aproximación del mapeo exponencial que se puede calcular a un bajo costo computacional, esto; sin afectar negativamente el comportamiento del algoritmo de optimización, como se afirma en Wen, Gallivan y Absil (2015, 3), Qi, Gallivan y Absil (2010).

2.3.4.1. Convexidad en una variedad riemanniana

Existen diversos puntos de vista en la geometria riemannianna para generalizar el concepto de convexidad de $\mathbb{R}^n$, los más importantes son los que presentamos en las siguientes definiciones, siguiendo lo analizado por (Quispe C., 2008).

Definición 47. *Sea $\mathcal{M}$ una variedad riemanniana completa, se dice que $A \subset \mathcal{M}$ es totalmente convexo, si para cualquier par de puntos p y q de A (no necesariamente distintos), las geodésicas que unen dichos puntos, estan integramente contenidos en A.*

Ejemplo 15. *Si $\mathcal{M} = \mathbb{R}^n$ con la métrica identidad $G(x) = I$, cualquier conjunto convexo en el sentido clásico es totalmente convexo.*

Ejemplo 16. *Si $p \in \mathcal{M}$ y existe una relación geodésica no trivial en p, es decir una geodésica: $\gamma : [a, b] \longrightarrow \mathcal{M}$ tal que $\gamma(a) = p = \gamma(b)$ con $\gamma(t) \neq p$ para algun $t \in [0, 1]$, entonces el conjunto $A = \{p\}$ no es totalmente convexo. Se deduce de esto que en general conjuntos unitarios no son totalmente convexos.*

Definición 48. *Decimos que $A \subset \mathcal{M}$ es convexo si para todo par de puntos p y q de A existe una geodésica minimal que une p y q contenido en A.*

Ejemplo 17. *El propio $\mathcal{M}$ y los conjuntos unitarios son conjuntos convexos.*

A continuación se define una función convexa f, sobre una variedad Riemanniana $\mathcal{M}$, la cual está se compone con a geodésica γ.

Definición 49. *$f : \mathcal{M} \longrightarrow \mathbb{R}$ es llamada función convexa si su restricción a cualquier geodésica de $\mathcal{M}$ es una función convexa en $\mathbb{R}$, es decir, si $\gamma : \mathbb{R} \longrightarrow \mathcal{M}$ es una geodésica entonces:*

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

es convexa.

El siguiente teorema nos permite caracterizar las funciones convexas sobre la extensión natural del espacio euclidiano.

Teorema 8. $f : \mathcal{M} \longrightarrow \mathbb{R}$ *es convexa si, y solamente si, para todo segmento de geodésica* $\gamma : [a,b] \longrightarrow \mathcal{M}$ *y para cualquier* $\lambda \in [0,1]$ *se verifica*

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

Demostración. Siendo f convexa, demostraremos que

$$f(\gamma((1-\lambda)a + \lambda b)) \leq (1-\lambda)f(\gamma(a)) + \lambda f(\gamma(b)) \tag{II.42}$$

Sea $h : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$. Para $a, b \in [a,b]$ y $\lambda \in [0,1]$ se tiene

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

De aquí se tiene (II.42).

Recíprocamente, sea $t = (1-\lambda)a + \lambda b$ con $\lambda \in [0,1]$ entonces

$$\begin{aligned} f \circ \gamma(t) &= f(\gamma((1-\lambda)a + \lambda b)) \\ &\leq f(\gamma((1-\lambda)a) + f(\gamma(\lambda b) \\ &= (1-\lambda)f(\gamma(a)) + \lambda f(\gamma(b)) \\ &\leq (1-\lambda)f \circ \gamma(a) + \lambda f \circ \gamma(b). \quad \blacksquare \end{aligned}$$

Observación 14. *La Definición 49 es la generalización natural de la definición clásica de función convexa en* $\mathcal{M} = \mathbb{R}^n$ *con la métrica usual. En efecto, dados p y q la geodésica* $\gamma : [0,1] \longrightarrow M,$ *que los une es*

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

Luego, por aplicación del Teorema 8 tenemos

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

$$\begin{aligned} f((1-\lambda)p + \lambda q) = f(\gamma(\lambda)) = f((1-\lambda)(0) + \lambda(1)) &\leq (1-\lambda)f(\gamma(0)) + \lambda f(\gamma(1)) \\ &= (1-\lambda)f(p) + \lambda f(p). \end{aligned}$$

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

Teorema 9. *Sea la función $f : \mathcal{M} \longrightarrow \mathbb{R}$, convexa, entonces $\mathcal{M}^{\alpha}$ es totalmente convexo.*

Demostración. Sea $p, q \in \mathcal{M}^{\alpha}$ y la geodésica $\gamma : [a, b] \longrightarrow \mathcal{M}$ tal que $\gamma(a) = p$ y $\gamma(b) = q$.

Probaremos que $\gamma(t) \in \mathcal{M}^{\alpha}$, para todo $t \in [a, b]$.

En efecto, sea $t = (1-\lambda)a + \lambda b$ para algun $\lambda \in [0, 1]$, como f es convexa y por el Teorema 8 se tiene

$$\begin{aligned} f(\gamma(t)) = f(\gamma(1-\lambda)a + \lambda b) &\leq (1-\lambda)f(\gamma(a)) + \lambda f(\gamma(b)) \\ &= (1-\lambda)f(p) + \lambda f(q) \\ &\leq (1-\lambda)\alpha + \lambda\alpha = \alpha. \end{aligned}$$

Así $f(\gamma(t)) \leq \alpha$, por tanto $\gamma(t) \in M^{\alpha}$. ∎

Teorema 10. *Sea una función f tal que $f : \mathcal{M} \longrightarrow \mathbb{R}$, decimos que es convexa en p si y sólo si, para cualquier geodésica $\gamma : \mathbb{R} \longrightarrow \mathcal{M}$ con $\gamma(0) = p$ vale la desigualdad*

$$f(\gamma(t)) - f(p) \geq t \langle grad\ f(p), \gamma'(0)\rangle \qquad \text{(II.43)}$$

Demostración. Definimos una aplicación $h : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$, donde h es función convexa en 0, desde que f es convexa en p y así se tiene

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

luego

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

Recíprocamente, si $f(\gamma(t)) - f(p) \geq \langle grad\ f(p), \gamma'(0) \rangle$, esto es,

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

entonces h es convexa en 0 y por tanto f es convexa en p. ■

Teorema 11. *Si $f : \mathcal{M} \longrightarrow \mathbb{R}$ es convexa, entonces todo punto crítico de f es un punto de mínimo global de f.*

Demostración. Sea $x \in \mathcal{M}$, debido al Teorema de Hopf-Rinow consideramos una geodésica $\gamma : \mathbb{R} \longrightarrow \mathcal{M}$ tal que $\gamma(0) = x$ y $\gamma(b) = y$, como f es convexa y del Teorema 10

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

esto es,

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

Como $grad\, f(x) = 0$ entonces $f(y) \geq f(x)$, para todo $y \in \mathcal{M}$. Por tanto x es punto de mínimo global de f. ■

Teorema 12. *Sea $f : \mathcal{M} \longrightarrow \mathbb{R}$ de clase C^2, f es convexa si y solamente si, para todo $p \in \mathcal{M}$, la Hesiana de f en p es*

$$H_p^f : T_p\mathcal{M} \longrightarrow T_p\mathcal{M}$$

es semidefnida positiva.

Demostración. Sea $v \in T_pM$ y $p \in \mathcal{M}$ y la geodésica $\gamma : \mathbb{R} \longrightarrow \mathcal{M}$ tal que $\gamma(0) = p$ y $\gamma'(0) = v$.

Definiendo $h : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $h(t) = f(\gamma(t))$ sabemos que h es convexa y de clase C^2. Del análisis convexo clásico, tenemos que esto es equivalente a $h''(t) \geq 0$, se tiene

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

y

$$h''(0) = \left\langle v, H_p^f v \right\rangle \geq 0.$$

Reciprocamente, si

$$\left\langle v, H_p^f v \right\rangle \geq 0,$$

entonces f es convexo. En efecto, definiendo $h : \mathbb{R} \longrightarrow \mathbb{R}$ convexa, entonces se tiene $h = f \circ \gamma$ es convexa. ■

2.3.4.2. Funciones cuasi-convexas y pseudoconvexas

Se dice que $f : X \to \mathbb{R}$ siendo X un subconjunto convexo de un espacio vectorial es cuasicóncava si $f(\alpha x + (1-\alpha)y) \geq min\{f(x), f(y)\}\ 0 \leq \alpha \leq 1$.

La función f es cuasi convexa si y solamente si $-f$ es cuasicóoncava. Se sigue de la definición de cuasiconvexidad que f es cuasi convexa si y solamente si $f(\alpha + (1-\alpha)y) \leq max\{f(x), f(y)\}\ 0 \leq \alpha \leq 1$.

Sea X un subconjunto convexo de $\mathbb{R}^n$ y $f : X \to \mathbb{R}$.

Algunas propiedades de las funciones cóncavas que, se deducen por su definición en forma inmediata son: Se verifica la siguiente desigualdad: $f(\sum_{i=1}^n \alpha_i x_i) \geq \sum_{i=1}^n \alpha_i f(x_i)$ es decir el valor que toma una función cóncava en una combinación convexa de elementos de X mayor o igual que la combinación convexa de los respectivos valores.

El subgrafo $S = \{(x, \lambda) \in X \times \mathbb{R} : \lambda \leq f(x)\}$ de una función cóncava, es convexo en $\mathbb{R}^n \times \mathbb{R}$.

El epígrafo $S = \{(x, \lambda) \in X \times \mathbb{R} : \lambda \geq f(x)\}$ de una función convexa, es convexo en $\mathbb{R}^n \times \mathbb{R}$. Una combinación lineal no negativa, de funciones cóncavas, definidas en un conjunto convexo, es una función cóncava. Se debe observar que estas definiciones has sido desarrollada solamente sobre geometría euclidiana en general. Para describir estos conceptos, primeramente se desarrollará aspectos de geometría no euclidiana, en particular la variedad Riemannana con retracción $\mathcal{R}$

Definición 50. *Sea $\mathcal{M}$ una variedad riemanniana completa y $f : \mathcal{M} \longrightarrow \mathbb{R}$ una función real. f es llamada cuasi-convexa en $\mathcal{M}$ si para todo $x, y \in \mathcal{M}$, $t \in [0, 1]$, se cumple*

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

para toda curva geodésica $\gamma : [0, 1] \longrightarrow \mathcal{M}$, tal que $\gamma(0) = x$ y $\gamma(1) = y$.

Teorema 13. *Sea $f : \mathcal{M} \to \mathbb{R}$ una función diferenciable y cuasi-convexa en una variedad riemanniana completa $\mathcal{M}$ y sea $x, y \in \mathcal{M}$. Si $f(x) \leq f(y)$ entonces*

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

donde $grad\, f$ es el gradiente de f y γ es la curva geodésica tal que $\gamma(0) = y$ y $\gamma(1) = x$.

Demostración. Consideremos una curva geodésica $\gamma : [0, 1] \longrightarrow \mathcal{M}$ tal que $\gamma(0) = y$ y $\gamma(1) = x$. Definimos $h : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $h(t) = f \circ \gamma(t)$, usando la aproximación de Taylor de primer orden de h en $t = 0$ tenemos

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

donde $\lim_{t\to 0} \frac{\theta(t)}{t}$. Entonces tenemos

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

como f es cuasi-convexa y $f(x) \leq f(y)$ tenemos

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

dividiendo por t y tomando limite cuando $t \to 0$ se tiene $\langle grad\, f(y), \gamma'(0)\rangle \leq 0$. ■

Definición 51. *Una función diferenciable $f : \mathcal{M} \to \mathbb{R}$ es pseudoconvexa si, para todo par de puntos distintos $x, y \in \mathcal{M}$ y toda curva geodésica que une x a y ($\gamma(0) = x$ y $\gamma(1) = y$) tenemos*

$$\langle grad\, f(x), \gamma'(0)\rangle \geq 0, \quad entonces \quad f(y) \geq f(x).$$

Teorema 14. *Sea $f : \mathcal{M} \longrightarrow \mathbb{R}$ una función diferenciable y pseudoconvexa. Entonces, x^* es un mínimo global de f si, solamente si, $grad\, f(x^*) = 0$.*

Demostración. Sea la geodésica $\gamma : \mathbb{R}_+ \to \mathcal{M}$ tal que $\gamma(0) = x^*$ y definimos $h : \mathbb{R} \to \mathbb{R}$ con $h = f \circ \gamma$. Por el desarrollo de Taylor de primer orden de h en 0

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

donde $\lim_{t\to\infty} \frac{\theta(|t|)}{|t|} = 0$, entonces tenemos:

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

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

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

luego

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

que en el limite cuando $t \to 0$, $\langle grad\, f(x^*), \gamma'(0)\rangle \geq 0$, finalmente tomando

$$\gamma'(0) = -grad\, f(x^*),$$

se tiene $grad\, f(x^*) = 0$. El reciproco es inmediato basta usar la definición de f ser pseudoconvexa. ∎

Observación: Un concepto en relación a la retracción y la convexidad, según Wen y col. (2015, 3), y que nos sirvió para el análisis de la convergencia de nuestro problema en estudio es la siguiente:

Definición 52. *Sea $f : \mathcal{M} \to \mathbb{R}$, donde $\mathcal{M}$ es una Variedad con retracción $\mathcal{R}$, donde $f(\mathcal{R}_x(t\eta)) \in \mathbb{R}$, con $x \in \mathcal{M}$ y $\eta \in T_x\mathcal{M}$. Más aún, f se llama función convexa-retractiva, con respecto a la retracción $\mathcal{R}$ en el conjunto S, si para todo $x \in S$ y todo $\eta \in T_x\mathcal{M}$ y $\|\eta\| = 1$, f es convexa para todo t tal que $\mathcal{R}_x(\tau\eta) \in S$, $\tau \in [0, t]$*

Esta definición de convexidad retractiva sobre una variedad $\mathcal{M}$, incluso sobre el propio espacio Euclidiano, generaliza el concepto clásico de convexidad

2.4. Definiciones de términos básicos

2.4.1. Símbolos y Notaciones

Los símbolos y notaciones que aquí, presentamos han sido y se usó de manera formal, estas notaciones simbólicas son las más usadas en el contexto de las mateáticas, en particular la Optimización matemática y sus aplicaciones.

- Al conjunto de vectores, positivos incluído el cero vector, del espacio euclidiano n-dimensional denotamos por $\mathbb{R}^n_+ = \{x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n : x_i \geq 0,\ i = 1, 2, .., n\}$.

- El conjunto de vectores, estrictamente positivos, no incluye el cero vector, del espacio euclidiano n-dimensional denotamos por $\mathbb{R}^n_{++} = \{x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n : x_i > 0\}$, $i = 1, 2, .., n$.

- El producto interno de dos vestores de $\mathbb{R}^2$ se expresa en forma general para $\mathbb{R}^n$ de la siguiente forma: $x, y \in \mathbb{R}^n$, $(x, y) = \sum_{i=1}^n x_i y_i$, llamado, producto interno euclideano en $\mathbb{R}^n$.

- Conjunto de funciones diferenciables un número p veces, es escrito simbólicamente por $C^p(\Omega) = \{f : \Omega \to \mathbb{R} : f$ en un domínio abierto Ω. Si $p = \infty$, entonces $C^\infty(\Omega)$ es el conjunto de funciones infinitamente diferenciables.

- Denotamos con $\mathcal{M}$: variedad diferenciable. Por ejemplo, la esfera de radio uno es una variedad diferenciable.

- La notación usual $T_p\mathcal{M}$, es el espacio tangente a una variedad diferenciable $\mathcal{M}$ en el punto p.

- El símbolo $\mathcal{H}$, es el conjunto de campos de vectores $X \in T_p\mathcal{M}$. que estan en el campo tangente. Por ejemplo el campo magnético de un imán es también un campo vectorial.

- $X(p)$, es un campo vectorial aplicado a un punto p fijo del espacio tangente. Otro ejemplo muy conocido es el campo gravitatorio que se representa por el campo vectorial en la Física newtoniana.

- El símbolo ∇, llamado nabla, representa al objeto matemático conocido como la conexión afin que sirve para representar o especificar la derevada covariante en relación a un campo de vectores $\mathcal{H}$. Se entiende también como la generalización del concepto de derivada parcial de una función.

- El gradiente $grad\, f(x)$, es un símbolo que permite expresar la variación de las variable involucradas de una función respecto un valor o magnitud f. En éste caso, es en el sentido covariante.

- El Hˀ es denominada matriz Hessiana de f. Es una expresión que sirve para denotar la segunda derivada parcial de f.

- PPL, Es una expresión simplificada para poder hacer referencia a Problemas de programación lineal.

- $PPnL$ Es una expresión simplificada para poder hacer referencia a Problem de programación no lineal.

CAPÍTULO III. HIPÓTESIS Y VARIABLES

3.1. Hipótesis

3.1.1. Hipótesis general

La demostración de la convergencia del Método del Gradiente usando retracciones minimiza funciones cuasi convexas sobre variedades Riemannianas

3.1.2. Hipótesis específica

1. Existe retracciones con métricas específicas en el Método de Gradiente que viabiliza minimizar funciones cuasi convexas sobre variedades Riemannianas.

2. El uso de funciones cuasi convexas permite encontrar puntos candidatos a óptimos por medio del Método de Gradiente sobre variedades Riemannianas

3.2. Definición conceptual de variables

A continuación describimos las variables de investigación, esto son, símbolos a los cuales les asignamos valores, los que se podran analizar según la información que se tenga al respecto, para esta investigación las variables las denotamos con X e Y respectivamente

1. Variable X:

   Retracciones en el Método del Gradiente para minimizar funciones cuasi convexas

2. Variable Y:

   Convergencia del Método del Gradiente sobre variedades Riemannianas.

3.3. Operacionalización de variables

Para la demostración y comprobación se hicieron los respectivos análisis del marco teórico y del marco conceptual de la investigación. En la siguiente tabla se presenta las variables de investigación, la interrelación entre ellas, las dimensiones y los respectivos indicadores.

Tabla 1

Cuadro de operacionalización de variables

Variables	Dimensiones	Indicadores
X. Retracciones en el Método del Gradiente para minimizar funciones cuasi convexas	X_1**:** Variedades Riemannianas	- Variedades diferenciables - Tangente a una variedad diferenciable - Métricas sobre variedades riemannianas - Campos vectoriales, conexiones y derivadas covariantes - Geodésicas - Campos paralelos - Métrica riemanniana y símbolos de Crhristoffel - Curvatura de una variedad riemanniana - Gradiente y Hessiana en una variedad riemanniana
	X_2**:** Método del Gradiente sobre variedades Riemannianas	- Algoritmo del Gradiente
Y. Convergencia del Método del Gradiente sobre variedades Riemannianas.	Y_1**:** Condiciones de optimalidad	- Existencia de mínimo global - Caracterización de mínimo - Elemento de análisi convexo
	Y_2**:** Resultados de análisis convexo y cuasi-convexidad	- Convergencia de funciones cuasi convexas - Búsqueda de Armijo generalizado

Elaboración propia

CAPÍTULO IV. DISEÑO METODOLÓGICO

4.1. Tipos y diseño de investigación

4.1.1. Tipo

La naturaleza de esta investigación, es de tipo pura, teórica, o básica, en el ámbito de las ciencias matemáticas, dado que el proceso de investigación, gira en torno, al estudio, análisis de conocimientos respecto la teoría de optimización matemática, dado por un conjunto de conceptos, definiciones, proposiciones, teoremas y resultados teóricos que nos llevaron a demostrar el objetivo principal de nuestra investigación, que, además, se puede decir que, en este caso no se tuvo la necesidad de realizar una aplicación directa a la realidad, sino, más bien hacer, el análisis de los entes matemáticos a ser estudiados, cuyos resultados pueden ser usados para ser aplicados con posterioridad.

En relación a la afirmación anterior, algunos autores sostienen, precisamente que, este tipo de investigación es una investigación básica, de perspectiva teórica así lo detalla, Hernández, Fernández y Baptista (2010), puesto que básicamento se está ante un proceso de sistemas de información dentro de otros conocimientos ya existentes y estos a su vez, pueden estar relacionados al planteamiento del problema, debido a que los resultados esperados se obtienen necesariamente de teorías pre-existentes a fin de expandirlos.

En esa misma línea decimos que esta investigación tiene forma pura, afirmado por Tamayo (2003), dado que en este tipo de investigación se proponen leyes generales, a partir de teorías matemáticas para todo el proceso de análisis de la información recopilada.

Por otro lado, la presente investigación, está a su vez, enmarcado dentro de los problemas conceptuales, tal como se indica en (Torres, 2005).

En resumen, la investigación que hemos propuesto es una investigación principalmente demostrativa, basado en los resultados especializados sobre teorías en torno a la optimización matemática, dados en conceptos, definiciones, proposiciones, teoremas y corolarios para la construcción de algunas demostraciones, estudiados por abstracción con el fin obtener y discutir otros teoremas, cuya interpretación no es de aplicación directa ni en la simulación computacional. Más en particular es una investigación propiamente demostrativa, basado, como se ha indicado, en teorías pre-existentes, porque busca incrementar teorías que se relaciona entre si, para generar nuevos conocimientos, de este modo no se ocupa de las aplicaciones prácticas o directas que puedan hacer referencias los análisis teóricos; buscamos la deduccíon de otros conocimientos;

4.1.2. Diseño de investigación

El conjunto de conocimientos matemáticos analizados, trata de entes formales, y su relación entre ellos. Son objetos ideales o abstractos a los cuales le ha dado interpretación, por ejemplo, al construir el objeto: convergencia del método que minimiza funciones, y lo relacionamos con otro objeto matemático: el objeto métrica. Esto es, tenemos datos cuantificables luego de una interpretación y análisis. Entonces, a lo anterior, se ha buscado interpretar, resumiendolo en un conjunto de conceptos, definiciones, proposiciones, lemas, teoremas y corolarios, de cuyas expresiones formales hemos razonado deductivamente y generalizado a su vez, en otro conjunto de conocimientos o sistemas de conocimientos. Así, se ha de lograr una visibilización del contexto real con finalidad de interpretarlos y hacer toma de desiciones. Siguiendo el sentido anterior, hemos aplicado en esta investigación un **diseño cuantitativo** que ha permitido interrelacionar las variables de investigación, al estar inmersos en una investigación no experimental. Entonces, esto permite describir y diferenciar los objetos matemáticos, y de los resultados de ésta investigación, se hizo la contrastación de las hipótesis, dado que la validez de los entes supuestos, depende también de los resultados hallados.

4.2. Método de investigación

Respecto el método que usé en la investigación, puedo señalar lo siguiente: En principio el método que se usó en esta investigación es el hipotético-deductivo, debido a la naturaleza propia de la investigación, fue necesario inferir de ciertos conceptos teóricos a través de abstracciones para analizar y sintetizar la información tras la recopilación de la fuente bibliográfica, ello implicó inducir resultados a través de amplias deducciones de ciertas hipótesis que se propusieron en su momento.

A su vez, se aplicó diversos métodos más, como el método lógico; muy importante debido a que así se pudo, plantear todo el aspecto teórico que circunda al problema de investigación, ello conlleva el estudio de leyes generales respecto por ejemplo la geometría diferencial, o las condiciones alrededor de la optimización matemática, con este método toda la investigación recae en las teorías recolectadas.

También se aplicó fuertemente el método lógico-inductivo, el cual sirvió para argumentar propiamente a los distintos conceptos formales u objetos abstractos, que se transformaron a su vez en hipótesis teóricas, las cuales se verificaron usando los conocimientos y hechos establecidos, vía la abstracción y el razonamiento lógico que permitió la formación de otras hipótesis, así se llegó a conclusiones importantes como parte de un análisis fundamentados en leyes y principios que fueron validados en su oportunidad.

Otro método usado en todo el proceso de investigación es el método inductivo, con el que se ha logrado amalizar teorías muy particulares, por ejemplo cuando tenemos una superficie en $\mathbb{R}^3$, digamos la esfera unitaria es un caso particular de variedad riemanniana, obtenida a partir de teorías sistematizadas y deducidas con rigurosidad científica y permitió obtener conlusiones generales.

Para complementar el proceso de investigación, se hace énfasis respecto el método científico de la sguiente manera:

> el método científico se suele utilizar para mejorar o precisar teorías previas en función de nuevos conocimientos, donde la complejidad del modelo no permite formulaciones lógicas. Por lo tanto, tiene un carácter predominantemente intuitivo y necesita, no sólo para ser rechazado sino también para imponer su validez, la contrastación de sus conclusiones. (Behar, 2008, p. 41).

4.3. Población y muestra

Debido a la naturaleza de la investigación, el problema de investigación no tiene magnitud factual, podriamos afirmar que es un problema que tiene magnitud formal plasmado a través de proposiciones, teoremas, lemas; entes, figuras estrictamente formales o abstractas de los cuales no es posible hablar de población ni muestra. Por tanto, estrictamente da la defínción de población, el problema de investigación carece de población. En consecuencia para esta investigación el problema de investigación también carece de ella.

4.4. Lugar de estudio y periodo de desarrollo

El lugar de estudio se puede decir que se realizó en una primera etapa, en los ambiente de la escuela de posgrado de Economía de la Universidad Nacional del Callao.

Para la segunda parte, del desarrollo sobre el marco teórico, conceptual, obtención y análisis de los resultado, se realizó en los ambientes de la biblioteca especializada de la Facultad de Ciencias Naturales y Matemática, también, en la Universidad Nacional del Callao.

La compilación de la información procesada, tanto en digitación como impresión, se realizó también en los ambientes de la Escuela de posgrado de Economía, por contar con ambientes confortables y de buena logística, que permitió la culminación de la redacción de esta investigación.

El periodo de ejecución que duró esta investigación oscila entre el 2018 al 2019 aproximadamente.

4.5. Técnicas e instrumentos para la recolección de la información

Técnica para obtener información documental; para ello se usó fichas de investigación para recopilar información de las referencias bibliográficas, relacionados con el problema objeto de investigación. En dichas fichas, se anotó los resultados más importantes del marco conceptual, en relación a los indicadores del problema de investigación.

Eso, implicó la interpretación, el análisis lógico matemático de los entes abstractos expresados en definiciones, propiedades, teoremas, lemas y corolarios.

Se aplicó y en algunos casos se demostró los resultados más importantes, luego del cual se hizo las respectivas interpretaciones para concretar las conclusiones de la investigación.

4.6. Análisis y procesamiento de datos

Al analizar los resultados, que concordante con la sección 4.3, ausencia de magnitud factual, datos estadísticos y tipo de investigación, no se usó ninguna estadística. Al tratarse de una investigación netamente teórica o pura, se planteó sólo el proceso del análisis de la información, mediante el recojo de conocimientos por medio de otro conocimiento, hecho a través de distintas fichas, como son; las fichas de investigación, de tipo textual, de resumen, también las de registro y las fichas bibliográficas.

Entonces sobre los resultados obtenidos, se realizó su respectivo análisis, posteriormente, se sometió a discusión a fin de usar las hipótesis planteadas en ésta investigación.

Entonces se puede ordenar el proceso de análisis de la información del siguiente modo:

1. En las fichas de investigación, se ubica la información precisa y coherente, de manera breve, del tema o tópico seleccionado.

2. En términos generales, dentro de los tópicos o temas de investigación seleccionamos las definiciones, los conceptos, teoremas y propiedades teóricas más importantes que constituyen las organizaciones matemáticas, estos son los indicadores involucrados en la investigación.

3. Estas organizaciones matemáticas o entes abstractos, llevaron a mostrar las hipótesis de la investigación, y fueron usados como fuente de información de tal manera que permitió comparar teoremas o resultados parciales con los finales, y sirvió para deducir las demostraciones, estrictamente analizadas y verificadas.

CAPÍTULO V. RESULTADOS

5.1. Resultados descriptivos

En esta sección se buscó responder a las hipótesis planteadas en esta investigación; en cuyo caso, requerimos del análisis de los algoritmos de búsqueda lineal, el método del Gradiente y las retracciones sobre variedades riemannianas; como principales indicadores para la obtención de los resultados, principalmente en el teorema que refleja la convergencia del algoritmo a través de sucesiones $\{x_k\}$ que, teóricamete genera.

Para aplicar los indicadores, en particular el algoritmo del Gradiente expresado en la dimensión de la variedad Riemanniana, existen diferentes maneras de escoger el parámetro t_k generando consecuentemente diversos submétodos los cuales, para su uso, dependerá exclusivamente de su performance computacional.

A seguir, se describió en primer lugar las reglas de búsqueda lineal más conocidas, o prodemos decir; los más clásicos, que han usado anteriormente como parte del algoritmo del Gradiente.

Algoritmo A: Gradiente con búsqueda exacta

1. Dado x^k, calcule el $grad\ f(x^k)$ sobre el plano tangente $T_{x^k}\mathcal{M}$.

2. Determine la geodésica $\gamma(t),\ t \geq 0$, de $\mathcal{M}$ que verifique $\gamma(0) = x^k$ y $\gamma'(0) = -grad\ f(x^k)$.

3. Minimize $f(\gamma(t)),\ t \geq 0$, obteniendo t_k y defina

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

Algoritmo B: Gradiente con Regla de Armijo

1. Dado x^k, calcule el $grad\ f(x^k)$ en $T_{x^k}\mathcal{M}$.

2. Determine la geodésica $\gamma(t)$, $t \geq 0$, de $\mathcal{M}$ que verifique $\gamma(0) = x^k$ y $\gamma'(0) = -grad\ f(x^k)$.

3. Hacer:

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

donde i_k es el menor entero positivo tal que:

$$f(\gamma(t_k)) \leq f(x^k) - \alpha t_k \left\| grad\ f(x^k) \right\|^2$$

y $\alpha \in (0, \frac{1}{2})$.

Para proponer el algoritmo C usamos la siguiente definición.

Definición 53. *Una funcón $f : \mathcal{M} \longrightarrow \mathbb{R}$ es llamada gradiente Lipschitziana con constante Γ si para todo $p, q \in \mathcal{M}$ y $\gamma : [0, a] \longrightarrow \mathcal{M}$ la geodésica con $\gamma(0) = p$ y $\gamma(a) = q$ se verifica*

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

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

Algoritmo C: Gradiente con Pasos fijos

1. Dado x^k calcule el $grad\ f(x^k)$ en $T_{x^k}\mathcal{M}$.

2. Determine la geodésica $\gamma(t_k)$, $t_k \geq 0$, de $\mathcal{M}$ que verifique $\gamma(0) = x^k$ y $\gamma'(0) = -grad\ f(x^k)$

3. Dados $\delta_1 > 0$ y $\delta_2 > 0$ tales que,

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

donde Γ es la constante de Lipschitz asociada al campo gradiente de f, escoger

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

En seguida se muestra las deducciones alcanzadas para mostrar las hipótesis, en la siguiente sección detallamos los resultados.
Y más adelante se decribirá la búsqueda generalizada de Armijo con retracciones.

5.2. Resultados inferenciales

Luego de realizar el análisis, de un conjunto de teorías, enunciados, teoremas, lemas, y corolarios, existentes en torno al problema de investigación, y se logró inferir en relación a la naturaleza de la investigación, las hipótesis planteadas. Para ello, fue necesario hacer algunos analogías que sirvió para plantear los teoremas que garantizan la convergencia del algoritmo, usando retracciones; se usó y demostro las hipótesis (por ejemplo las hipótesis A1 y A2) en torno del problema de investigación.
En la siguiente subsubsección puntualizamos al respecto.

5.2.1. Estudio de la convergencia del método del gradiente

Para el estudio sobre la convergencia del algortimo que es otro indicador, esto es el modelo que se planteó en la descripción de la realidad problemática. Es decir, el, interés fué resolver el siguiente problema de optimización:

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

donde $\mathcal{M}$ es una variedad riemanniana conexa, completa de dimensión finita y $f : \mathcal{M} \to \mathbb{R}$ es una función continuamente diferenciable y cuasi-convexa.

La siguiente hipótesis permitió establecer la convergencia a través de la generación de sucesiones convergente en un conjunto no vacío U.

Hipótesis A1.

El conjunto de puntos óptimos globales del problema (p), denotado por X^*, es no vacío. Denotamos el valor óptimo de (p) por f^*. Ahora, definamos el siguiente conjunto

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

Antes de proseguir con el análisis de la convergencia del algortimo del gradiente, fue necesario introducir una definición importante en la investigación, la sucesión cuasi Fejér, y algunos resultados desarrollado por Burachik y col. (1996), definido sobre un espacio métrico.

Definición 54. *Sea (X,d) un espacio métrico completo. Una sucesión $\{y_k\}, k \geq 0$, de X es cuasi-Fejér convergente al conjunto $U \subset X$, si 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 $d^2(y^{k+1},u) \leq d^2(y^k,u) + \epsilon_k$, para todo k.*

El siguiente resultado que estudiamos, se puede ver como se involucra al indicador métrica indirectamente.

Teorema 15. *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 $\{y^k\}$ pertence a U. Entonces $\{y^k\}$ converge y*

$$\lim_{k\to+\infty} y^k = \bar{y}$$

Demostración. Ver (Burachik y col., 1996, p. 6).

Entonces estudiemos los siguiente conceptos en torno a la sucesión Quasi-Fejér convergente.

Definición 55. *Let X un espacio métrico y $0 < p < \infty$. La sucesión (x_n) en X is p-Quasi-Fejér convergente a $\Omega \subset X$ Si para cada $x^* \in \Omega$, existe una sucesión no negativa (p_n) tal que $d(x^*, x_n)^p \leq d(x^*, x_{n-1})^p + \rho_n, \ \ n = 1, 2, \ldots.$*

El siguiente Lema es un resultado de gran interés el cual será usado para probar que la sucesión, generada por el método de máximo descenso es cuasi-Fejér convergente a U.

Lema 2. *Sea $f : \mathcal{M} \to \mathbb{R}$ una función continuamente diferenciable y cuasi-convexa en una variedad riemanniana conexa, completa y de dimensión finita con curvatura seccional no negativa, entonces*

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

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

Demostración. Tomemos el conjunto U tal que, $x \in U$ arbitrário. Sea también $\gamma_1 : [0, l_1] \to \mathcal{M}$ la geodésica minimal que une x^k y x con $\gamma(0) = x^k$, $\|\gamma'(0)\| = 1$ y $\gamma_2 : [0, 1] \to \mathcal{M}$ una geodésica que une x^k y x^{k+1}, esto es $\gamma_2(0) = x^k$, $\gamma_2(1) = x^{k+1}$ con $\gamma_2'(0) = -t_k grad\, f(x^k)$. Por propiedad de homogeneidad de las geodésicas, γ_2 es reparametrizada tal que

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

tal que $\gamma_2(t_k||grad\, f(x^k)||) = x^{k+1}$ y ahora tenemos

$$\left\|\gamma_2'(0)\right\| = 1.$$

Del Teorema 3 tenemos:

$$d(x^{k+1}, x)^2 \leq d(x^k, x)^2 + t_k^2||grad\, f(x^k)||^2 + 2t_k d(x^k, x) \left\langle grad\, f(x^k), \gamma_1'(0) \right\rangle.$$

Como f es cuasi-convexa y $f(x) \leq f(x^k)$, del Teorema 13 obtenemos que

$$\left\langle grad\, f(x^k), \gamma_1'(0) \right\rangle \leq 0.$$

Usando este resultado en la desigualdad anterior obtenemos el resultado deseado. ■

Para la demostración de la convergencia del método del Grediente es necesario tomar el algoritmo generalizado de Kiwiel y Murty (1996, 1) y el (**algoritmo G** que estamos proponiendo para generalizar el algoritmo de Armijo involucrado con la llamada retracción $\mathcal{R}$ que se propuso en (II.26).

Veamos alguna reseña sobre la propuesta de Armijo que publica en un artículo.

El algoritmo de Armijo generalizado. En 1966, Armijo, publicó el artículo titulado: “Minimization of functions having lipschitz continuous firts partial derivatives", en dicho artículo, demuestra la convergencia del método del Gradiente, a partir de entonces sus resultados son utilizados convenientemente en la búsqueda de mejoras o extensiones teóricas computacionales. Para el objetivo, se ha estudiado el desarrollado el método del Gradiente, sobre variedades riemannianas utilizando regla de Armijo generalizado.

Enfatizar también que los resultados de Armijo, fueron generalizados a convergencia global para el caso cuasi-convexo sobre variedades Riemannianas, partir del trabajo realizado por Kiwiel y Murty en (1996) que, hasta ese año, se había modelado sobre espacio Euclidiano, y los que a su vez fueron generalizados para condición convexo por (Burachik y col., 1996) y (Cruz Neto & Oliveira, 1995), también sobre espacio Euclidiano

Ahora, la regla de Armijo con retracción (Regla de Armijo-$\mathcal{R}$):

$$t_k = \arg \text{máx}\{t : f\left(\mathcal{R}_{x^k}(-t\eta_k\, f(x^k))\right) \leq f(x^k) - \alpha t ||\eta_k\, f(x^k)||^2\} \tag{V.1}$$

donde $t = 2^{-i}, i = 0, 1, \ldots$ y $\alpha \in (0, 1)$.

Entonces, el método del gradiente vía la regla de Armijo-$\mathcal{R}$, genera una sucesión de puntos $\{x^k\}$ dados por (II.27)-(II.28) donde se satisfacen las siguientes hipótesis:

Hipótesis A2.

Sea $\phi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ una función tal que:

A2.1 Existe $\alpha \in (0, 1), \tau_\alpha > 0$, tal que, para todo $t \in (0, \tau_\alpha] : \phi(t) \leq \alpha t$,

A2.2 Existe $\beta > 0$, $\tau_\beta \in (0, +\infty]$, tal que, para todo $t \in (0, \tau_\beta) \cap \mathbb{R}$: $\phi(t) \geq \beta t^2$,

A2.3 Para todo k, $f(x^{k+1}) \leq f(x^k) - \phi(t_k)||\eta_k f(x^k)||^2$ y $0 < t_k \leq \tau_\beta$ en (II.28),

A2.4 Existe $\gamma > 1$, $\tau_\gamma > 0$, tal que, para todo $k : t_k \geq \tau_\gamma$ o

$\left[\text{ existe } \bar{t}_k \in [t_k, \gamma t_k] : f(\mathcal{R}_{x^k}(-\bar{t}_k \eta_k f(x^k))) \geq f(x^k) - \phi(\bar{t}_k)||\eta_k f(x^k)||^2\right]$.

Observación 15. *Obsérvese que la hipótesis A2 es satisfecha por la regla de Armijo para los mismos valores: $\phi(t) = \alpha t, \beta = \alpha, \gamma = 2$ y $\tau_\alpha = \tau_\beta = \tau_\gamma = 1$.*

Observación 16. *Vemos que la hipótesis A2 también es satisfecha por el método del gradiente con pasos fijos, incluida en Burachik y col. (1996), dicha hipótesis se generalizó para variedades riemannianas en (Da Cruz Neto y col., 1998). Esto es, en las referencias estudiadas, la regla para obtener t_k se calcula de la siguiente manera:*

Dados δ_1 y δ_2 tal que $\delta_1\Gamma + \delta_2 < 1$, donde Γ es la constante de Lipschitz asociada al $grad\, f$, elegimos

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

Aquí, se toma los siguientes parámetros $\phi(t) = \beta t^2$, con $\beta = \frac{\Gamma\delta_2}{2(1-\delta_2)}$, $\tau_\gamma = \delta_1$, $\tau_\beta = (2/\Gamma)(1 - \delta_2)$, $\alpha \in (0, 1)$ arbitrario y $\tau_\alpha = \alpha/\beta$, que nos permite garantizar la hipótesis A2.

Así, se tiene la siguiente proposición que garantiza la convergencia en el que se usa una sucesión del tipo cuasi-Fejér.

Proposición 21. *Sea $f : \mathcal{M} \to \mathbb{R}$ una función continuamente diferenciable y cuasi-convexa. Suponga que las hipótesis $A1$ y $A2$ son satisfechas. Entonces la sucesión $\{x^k\}$ generada por el método del gradiente con regla de Armijo generalizada es cuasi-Fejér convergente a U.*

Demostración. Ver (Quispe C., 2008, p. 63) $\blacksquare$.

Finalmente mostramos el siguiente teorema, el cual conserva hasta cierto punto, las hipótesis de los resultados principales de Quispe C. (2008), salvo las condiciones impuestas sobre el método del gradiente, vía la regla de Armijo-$\mathcal{R}$ y es la diferencia resaltante como un ligero aporte teórico a las investigaciones más próximas relacionadas.

Observe también que, en este teorema se involucra al indicador, funciones cuasi-convexas, junto a los otros indicadores como son los métodos de búsqueda lineal. Entonces se ha planteado el siguiente teorema, que permite encontrar puntos críticos para la minimización de funciones cuasi-convexas usando retracciones.

Teorema 16. *Sea $f : \mathcal{M} \to \mathbb{R}$ una función continuamente diferenciable y cuasi-convexa. Suponga que las hipótesis $A1$ y $A2$ son satisfechas. Entonces la sucesión $\{x^k\}$ generada por el método del gradiente con regla de Armijo-$\mathcal{R}$, converge para un punto estacionário (un punto $\bar{x}$ tal que $grad\, f(\bar{x}) = 0$).*

Demostración. Como ya se tiene la garantía de la convergencia de la sucesión cuasi.Fejér, visto en la Proposición 21; esto es, $\{x^k\}$ es cuasi-Fejér convergente en U, por tanto se puede decir que la sucesión $\{x^k\}$ es limitado gracias al Teorema 15, esto implica que existen puntos $\bar{x}$ y una subsucesión $\{x^{k_j}\}$ de la sucesión $\{x^k\}$ que converge a $\bar{x}$. Entonces, de la continuidad de f se tiene que, $\lim\limits_{j\to+\infty} f(x^{k_j}) = f(\bar{x})$.

También, la sucesión $\{f(x^k)\}$ es una sucesión no creciente, esto, tomando el resultado de la Proposición (21), ecuación (4.4) de (Quispe C., 2008), con una subsucesión que converge para $f(\bar{x})$, y así, se tiene esta relación,

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

Así, se puede afirmar que, efectivamente $\bar{x} \in U$. Ahora, del Teorema 15, puede verse que $\{x^k\}$ converge para $\bar{x}$. Ahora se debe ver que $grad\, f(\bar{x}) = 0$. Esto es por contradicción, véase; supógase que $grad\, f(\bar{x}) \neq 0$.

Y se tiene que $grad\, f(x^k) \to grad\, f(\bar{x}) \neq 0$ y si $f(x^k) \to f(\bar{x})$. se cumple que

$$\lim_{k \to +\infty} t_k = 0. \tag{V.2}$$

Ahora, se considera las hipótesis $A2{,}4$ y $A2{,}1$, y para un k suficientemente grande,

$$f(\mathcal{R}_{x^k}(-\bar{t}_k \eta_k\, f(x^k))) - f(x^k) \geq -\alpha \bar{t}_k ||\eta_k\, f(x^k)||^2. \tag{V.3}$$

Luego en aplicación, del teorema del valor medio, para cada k, existe $t_k^* \in [0, \bar{t}_k]$ tal que

$$-\langle\, \eta_k\, f(\mathcal{R}_{x^k}(-t_k^* \eta_k\, f(x^k))), P_{\gamma_k, 0, t_k^*} \eta_k\, f(x^k)\rangle \geq -\alpha ||\eta_k\, f(x^k)||^2,$$

y recuerde que, $P_{\gamma_k, 0, t_k^*}$ es el transporte paralelo a lo largo de la geodésica γ_k tal que $\gamma_k(0) = x^k$ y $\gamma_k'(0) = -\eta_k\, f(x^k)$. Ahora, (V.2) y $A2{,}4$ implican que

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

y haciendo tender $k \to +\infty$ en la igualdad anterior y tomando en cuenta la continuidad de $grad\, f$, $\mathcal{R}$ y el transporte paralelo, tenemos que $1 \leq \alpha$, esto contradice $A2{,}1$, puesto que $\alpha \in (0, 1)$. Eso implica que $\eta_k\, f(\bar{x}) = 0$. $\blacksquare$

Ahora, de la aplicación de la metodología propuesta:

De manera particular, se considera la variedad riemanniana $\mathcal{M}$ como $(\mathbb{R}^n_{++}, X^{-2})$, y el siguiente problema de investigación:

$$min\{f(x) : x \geq 0\}$$

donde $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función diferenciable.

La variedad considerada es completa finita dimensional y con curvatura seccional nula.

Hacemos los siguientes cálculos:

El gradiente de $\eta_k f$ es $X^2 \nabla f(x)$.

La iteración del método del gradiente es: $x_i^{k+1} = x_i^k \mathcal{R}\left(-x_i^k \frac{\partial f(x^k)}{x_i} t_k\right)$

La distancia riemanniana entre los puntos x^k y x^{k+1} es

$$d(x^k, x^{k+1}) = \left\{\sum_{i=1}^{n}\left[ln\left(\frac{x_i^{k+1}}{x_i^k}\right)\right]^2\right\}^{\frac{1}{2}}$$

Otras variedades que se pueden usar son:

1. $(\mathbb{R}^2, G(x))$, donde
$$G(x) = \begin{pmatrix} 1 + e^{2x_1} & -e^{x_1} \\ -e^{x_1} & 1 \end{pmatrix}$$

2. $(\mathbb{R}^4, G(x))$, donde
$$G(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 + 4x_3^2 & -2x_3 \\ 0 & 0 & -2x_3 & 1 \end{pmatrix}$$

CAPÍTULO VI. DISCUSIÓN DE RESULTADOS

6.1. Contrastación y demostración de la hipótesis con los resultados

Debido a la naturaleza de la investigación, del campo de las matemáticas, las pruebas desarrolladas para la contrastación de las hipótesis se remiten a analizar los aspectos teóricos en torno a las definiciones, teoremas, corolarios, incluyendo lemas, de los cuales se fueron deduciendo el resultado principal de nuestro trabajo, los que fueron detallados en su momento.

1. Seguidamente se realiza la discusión de los resultado obtenidos en relación a los fundamentos teóricos abordados alrededor del marco teórico, el planteamiento del problema y en particular de las hipótesis planteadas. Al respecto, y como se puede ver en la sección 3.1, la hipótesis dice: La demostración de la convergencia del método del Gradiente usando retracciones minimiza funciones cuasi-convexas sobre variedades Riemannianas.

2. Del teorema (16) que fué demostrado, vea pp. 94-95, se concluye que, con el método del Gradiente, se puede alcanzar un punto óptimo, candidato a solución del problema; esto, sólo es posible a través de la convergencia del método para cuando la función objetivo es cuasi-convexas. Note que, en el proceso de demostración del teorema y de sus hipótesis, se consideró el método de gradiente con búsqueda lineal de Armijo generalizado, Armijo-$\mathcal{R}$, es decir con el algoritmo **G**.

   Respecto a la hipótesis: Existe retracciones con métricas específicas en el Método de Gradiente que viabiliza minimizar funciones cuasi convexas sobre variedades Riemannianas. En este caso, mencionar por ejemplo, las métricas propuestas en

$G(x)$, según se puede ver en la página, p. 95, por tanto, se cumple con la prueba de ésta hipótesis.

Para la segunda hipótesis de investigación: El uso de funciones cuasi convexas permite encontrar óptimos por medio del Método de Gradiente sobre variedades Riemanniana. En cuanto a ésta hipótesis, se sobrentiende o infiere que, las funciones cuasi-convexas han sido usadas para presentar la convergencia de la metodología que resuelve al menos teóricamente hablando el problema de optimización sobre variedades, esto se vió en el programa de demostración de la hipótesis general.

3. En el proceso de demostración se introdujo la llamada Retracción ($\mathcal{R}$), ésta resulta ser, una extensión de aplicaciones por medio de geodésicas para la cual se usa métricas específicas, como es el caso de la métrica dado en (Quispe C., 2008, p. 22.).

6.2. Contrastación de resultados con otros estudios similares

Otros estudios similares se encuentra en el artículo de Absil y col. (2008), ellos demuestran un resultado que se ha usado como referente con respecto a la propuesta de esta investigación, que ya se tiene analizado, es decir, los mencionados autores, demuestraron la convergencia del método, usando función costo, cualquiera. Véase el siguiente teorema que propone Absil:

Teorema 17. *Sea $\{x_k\}$ una sucesión infinita de iteraciones generada por el algoritmo 1 (véase p. 63). Entonces todo punto de acumulación $\{x_k\}$ es punto crítico de la función costo f. (Absil y col., 2008, p. 65)*

Como se puede ver, en la hipótesis del teorema 17, no se imponen condiciones sobre la función costo, a diferencia del teorema 16 de nuestra investigación, la función costo, debe ser cuasi-convexa, además usamos búsqueda de Armijo-$\mathcal{R}$, para hallar el parámetro t_k.

Otro resultado es el siguiente corolario:

> **Corolario 3.** *Sea* $\{x_k\}$ *una sucesión infinita de sucesiones generada por el algoritmo 1. Supóngase que el conjunto de nivel* $L = \{x \in \mathcal{M} : f(x) \leq f(x_0)\}$ *es compacto. Entonces* $\lim\limits_{k \to +\infty} \|\mathrm{grad} f(x_k)\| = 0$. *(Absil y col., 2008, p. 66.)*

El resultado de convergencia del Corolario 3 de Absil y col. (2008), es para una búsqueda lineal, según se determine el algoritmo que requiera implementar, está desarrollado, siguiendo el análisis y teoría clásica del espacio euclideano $\mathbb{R}^n$, logrando un resultado bastante general.

Cabe indicar que tanto en Quispe Cárdenas y col. (2007, 1) que a su vez, es un artículo como resultado de la tesis de Quispe C. (2008), se demuestra convergencia local del método del gradiente, para el cual se usa sucesiones, que son aproximaciones generadas a través de aplicaciones exponenciales.

Se puede complementar respecto de Absil y col. (2008) que, proponen algoritmos globalmente convergentes, para problemas de optimización con y sin restricciones, en variedades riemannianas, no usa funciones cuasi-convexas.

Otros estudios similares se puede ver también en el propio Absil y Malick (2012), con la diferencia que el autor usa alguntos tipos particulares de retracción y presentan a su vez modelamientos computacionales.

Wen (2013), estudia y presenta una serie de algoritmos con aplicaciones computacionales sobre la teoría de optimización sobre variedades riemanniana.

Finalmente en la tesis de Cruzado Quispe (2014), realizan un análisis de convergencia del método del Gradiente, implementa computacionalmente algunos algoritmos, el autor aplica retracciones, pero no usa funciones cuasi-convexas.

6.3. Responsabilidad ética

En cuanto a la responsabilidad ética del desarrollo de esta investigación, se ha tenido muy en cuenta el cumplimiento del código de ética y deontología y, según disposición Ley nro. 25239 de creación del Colegio de Matemáticos del Peú. Con prudencia, se ha respetado las teorías, conceptos, lemas, definiciones y otros resultados, y puntos de vistas propuestos de los autores de las referencias bibliográficas que han sido usados durante todo el proceso de investigación.

CAPÍTULO VII. CONCLUSIONES

1. En el proceso de investigación, análisis, realizado, se mostró que efectivamente, se pudo obtener los puntos candidatos a óptimos de funciones cuasi-convexas. Esto es, con el algoritmo vía retracciones se alcanza dicha convergencia a punto crítico.

2. Las condiciones que se utilizó para mostrar la convergencia del algoritmo fueron:

   - Considerar al conjunto de soluciones U como vacío, $U \neq \emptyset$
   - La función costo, es cuasi-convexa.
   - La variedad riemanniana debe ser $(\mathcal{M}, g) \geq 0$, con métrica g.

3. El método del Gradiente es un método clásico, uno de los más conocido en el contexto de la Optimización Matemática, por ofrecer ventajas teóricas que permite la generalización de modelos en problemas de hallar máximos y/o mínimos.

4. Usando las propiedades de geometría riemanniana, sobre el cual se ha planteado el problema de minimización, se puede generalizar el método del Gradiente con retracciones como $x^{k+1} = \mathcal{R}(\lambda_k \eta_x)$ y aquí se usa funciones cuasi-convexas, así, nos permitimos tener una expresión extendida en relación al modelo en estudio, donde el método del Gradiente genera sucesiones sobre los cuales se definió la aplicación exponencial.

5. Al usar retracciones con alguna métrica específica, vemos que encontramos, que el método del Gradiente hace viable la generación de sucesiones convergentes para minimizar funciones de característica cuasi-convexa, también para puntos que sean candidato a óptimo.

6. Mediante la convergencia del método se ha establecido un punto crítico candidato al óptimo para una función cuasi-convexa, esto fue posible, gracias a las fuertes

propiedades sobre las funciones que se transformaron en funciones cuasi-convexas a través del cambio de métrica, dado que nos encontramos sobre una variedad Riemanniana.

CAPÍTULO VIII. RECOMENDACIONES

1. La recomendación principal sobre los resultados de nuestra investigación es, respecto a otras expresiones particulares que puedan ser finidas para las retracciones, y tienen que ser distintas de las que se encuentran en la literatura, así como lo plantea (Absil & Malick, 2012).

2. Realizar el análisis para geometria Riemanniana con curvatura no positiva en el que se puedan aprovechar propiedades intrinsecas de la geometría.

3. Implementar computacionalmente los resultados obtenidos a fin de comparar con otros métodos.

4. Se recomienda extender el procedimiento realizado con otros métodos computacionales de optimización como son: Newton, Cuasi-Newton, Gradiente conjugado, para resolver el problema de optimización con función cuasi-convexa, usando retracciones.

5. Aplicar el método presentado en la solución de problemas matriciales, esto es, problemas de optimización donde las variables son matrices, estos, lo encontramos generalmente en modelos de reconstrucción de imágenes.

6. Aplicar el procedimiento por ejemplo, a la variedad de Hadamard, siendo ésta, una variedad riemaniana completa, es simplemente conexa y de curvatura seccional estrictamente negativa.

REFERENCIAS BIBLIOGRÁFICA

Absil, P. A., Mahony, R. & Sepulchre, R. (2008). *Optimization Algorithms on Matrix Manifolds*. New Jersey: Princeton University Press.

Absil, P. A. & Malick, J. (2012). Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization: Society for Industrial and Applied Mathematics*, *22*(1), 135-158. doi:10.1137/100802529

Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific J. Math. 16*(1), 1-3. doi:10.2140/pjm.1966.16.1

Baygorrea Cusihuallpa, N. (2017). *Método do ponto proximal inexacto para minimización quasi-convexa em variedades de Hadamard* (Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtencão do grau de Doutor em Ciências (D.Sc.) Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia Facultad Ciencias Exactas y Naturales, Universidade Federal do Rio de Janeiro, Rio de Janeiro).

Behar, R. D. S. (2008). *Metodología de la Investigación* (1.ª ed.). Barinas: Ediciones Sahalom. Recuperado desde http://trabajodegradobarinas.blogspot.com/2014/04/methodologia-de-la-investigacion-2008.html

Bento, G. C. & Cruz Neto, J. X. (2013). A subgradient method for multi objective optimization on Riemannian manifolds. *J. Optim Theory Appl*, *159*(1), 125-137. doi:10.1007/s10957-013-0307-7

Bento, G. C., Cruz Neto, J. X. & Santos, P. S. M. (2013). An inexact steepest descent method for multicriteria optimization on Riemannian manifolds. *J. Optim. Theory. Appl. 159*(1), 108-124. doi:10.1007/s10957-013-0305-9

Boothby, W. M. (1975). *An introduction to differentiable manifolds and Riemannian geometry*. New York: Academic Press, Inc.

Burachik, R., Graña Drummond, L. M., Iusem, A. N. & Svaiter, B. F. (1996). Full convergence of the steepest descent method with inexact line searches. *Optimization*, *32*(2), 137-146. doi:10.1080/02331939508844042

Burden, R. L. & Faires, D. J. (2011). *Análisis numérico* (7.ª ed.). México D. F.: Thomson learnig.

Cruz Neto, J. X. & Oliveira, P. R. (1995). *Geodesic methods in riemannian manifolds*. Systems Engineering y Computer Sciences, PESC/COPPE, Federal University of Rio de Janeiro.

Cruzado Quispe, E. F. (2014). *Método de Máximo descenso usando retracciones en variedades riemannianas* (Tesis de Licenciatura en Matemática, Facultad Ciencias Naturales y Matemática, Universidad Nacional del Callao, Perú, Callao).

Da Cruz Neto, J. X., Lima, L. L. & Oliveira, P. R. (1998). Geodesic Algorithms in Riemannian Geometry. *Balkan Journal of Geometry and Its Applications*, *3*(2), 89-100.

Da Cruz Neto, J. X., Lima, L. L. & Oliveira, P. R. (2013). Learning how to play Nash, potential games and alternating minimization method for structure nonconvex problems on Riemannian manifolds. *J. Convex Anal. 20*(2), 395-438.

Do Carmo, M. P. (1988). *Geometria Riemanniana*. Rio de Janeiro: Instituto de Matemática Pura e Aplicada.

Do Carmo, M. P. (2005). *Geomtria Diferencial de Curvas e Superficies*. Rio de Janeiro: Sociedade Brasileira de Matemática.

Figueroa, C. (2004). Curvatura y Topología. En *XXII Coloquio, 19-23 de Julio del 2004*. Sociedad Matemática Peruana,

Gabay, D. (1982). Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, *3*, 177-219. doi:10.1007/BF00934767

Hernández, S. R., Fernández, C. C. & Baptista, L. M. (2010). *Metodología de la investigación*. México D. F.: Mc Graw Hill.

Kiwiel, K. C. & Murty, K. G. (1996). Convergence of the steepest descent method for minimizing quasiconvex functions. *Journal of Optimization Theory and Applications*, *89*, 221-226. doi:10.1007/BF02192649

Luenberger, D. G. (1972). The Gradient Projection Method along Geodesics. *Management Science*. Theory Series, *18*(11), 620-631. doi:10.1287/mnsc.18.11.620

Papa Quiroz, E., Quispe, E. & Oliveira, P. R. (2008). Steepest descent method with a generalized Armijo search for quasiconvex functions on Riemannian manifolds. *Journal of Mathematical Analysis and Applications*, *341*, 467-477. doi:10.1016/j.jmaa.2007.10.010

Papa, Q. E. A., Mallma, R. L. & Oliveira, P. (2015). An inexact proximal method for quasiconvex minimizations. *European J. of Operational Research*, *246*, 721-729. doi:10.1016/j.ejor.2015.05.041

Qi, C., Gallivan, K. A. & Absil, P.-A. (2010). Riemannian BFGS algorithm with applications. *Recent Advances in Optimization and its Applications in Engineering*, 182-192.

Quispe C., E. M. (2008). *El método de máximo descenso para Funciones cuasi-convexas en variedades riemannianas* (Tesis presentada como requisito para optar al título de Licenciado en Matemática, Facultad Ciencias Naturales y Matemática, Universidad Nacional del Callao, Callao).

Quispe Cárdenas, E. M., Papa Quiroz, E. A. & Oliveira, P. R. (2007). El Método de Máximo Descenso para funciones cuasi-convexas en Variedades Riemannianas. *Revista de la Facultad de Ciencias de la Uni. REVCIUNI*, *11*, 53-59.

Rapcsák, T. (1997). *Smooth Nonlinear Optimization in Rn*. Nonconvex Optimization and Its Applications. Boston, MA: Springer USA.

Requena Fraile, Á. (2005). Einstein y las matemáticas. *SUMA*, *50*, 7-14. Recuperado desde https://revistasuma.es/IMG/pdf/50/007-014.pdf

Smith, S. T. (1994). Optimization Techniques on Riemannian Manifolds. *Fields Institute Communications*, *3*, 113-136.

Tamayo, y. T. M. (2003). *El proceso de la investigación científica*. México D. F.: Editorial Limusa S. A.

Torres, B. C. (2005). El proyecto de investigación científica. Ediciones del autor. Lima.

Udriste, C. (1994). *Convex Functions and Optimization Methods on Riemannian Manifolds*. Dordrecht: Kluwer Academic Publishers.

Wen, H. (2013). *Optimization Algorithms on Riemannian Manifolds with Applications* (Tesis doctoral, College of Art y Sciences, Florida State University, Florida).

Wen, H., Gallivan, K. & Absil, P.-A. (2015). A Broyden Class of Quasi-Newton Methods for Riemannian Optimization. *Journal of Optimization*, *25*, 1660-1685. doi:10.1137/140955483

Yang, Y. (2007). Globally Convergent Optimization Algorithms on Riemannian Manifolds: Uniform Framework for Unconstrained and Constrained Optimization1. *Journal of Optimization Theory and Applications*, *132*, 245-265. doi:10.1007/s10957-006-9081-0

ANEXOS

Tabla 2

Matriz de consistencia

Problema	Objetivos	Hipótesis	Metodología	Población
Problema general *¿Cómo obtener la convergencia del Método del Gradiente usando retracciones para minimizar funciones cuasi convexas sobre variedades Riemannianas?*	**Objetivo General** *Demostrar la convergencia del Método del Gradiente usando retracciones para minimizar funciones cuasi convexas sobre variedades Riemannianas.*	**Hipóesis General** *La demostración de la convergencia del Método del Gradiente usando retracciones minimiza funciones cuasi convexas sobre variedades Riemannianas.*	**Nivel** *Es una investigación demostrativa o de comprobación de hipótesis.* **Tipo** *Es estrictamente una investigación básica.*	*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*
Problema específico *1. ¿El uso de retracciones con métricas específicas en el Método de Gradiente viabiliza la minimización de funciones cuasi convexas sobre variedades Riemannianas?.*	**Objetivos Específicos** *1. Usar retracciones con métricas específicas en el Método de Gradiente viabiliza la minimización de funciones cuasi convexas sobre variedades Riemannianas.*	**Hipóesis Específica** *1. Existe retracciones con métricas específicas en el Método de Gradiente que viabiliza minimizar funciones cuasi convexas sobre variedades Riemannianas.*	**Diseño de la investigación** *Diseño no experimental, cuantitativa*	
2. ¿Las funciones cuasi-convexas permiten encontrar puntos cadidatos a óptimos por medio del Método de Gradiente sobre variedades Riemannianas?.	*2. Usar funciones cuasi convexas permite encontrar puntos candidatos a óptimos por medio del Método del Gradiente sobre variedades Riemannianas*	*2. El uso de funciones cuasi convexas permite encontrar puntos candidatos a óptimos por medio del Método de Gradiente sobre variedades Riemannianas.*		

Variables	Dimensiones	Indicadores
X. *Retracciones en el Método del Gradiente para minimizar funciones cuasi convexas*	X_1**:** *Variedades Riemannianas*	*- Variedades diferenciables* *- Tangente a una variedad diferenciable* *- Métricas sobre variedades riemannianas* *- Campos vectoriales, conexiones y derivadas covariantes* *- Geodésicas* *- Campos paralelos* *- Métrica riemanniana y símbolos de Crhristoffel* *- Curvatura de una variedad riemanniana* *- Gradiente y Hessiana en una variedad riemanniana*
	X_2**:** *Método del Gradiente sobre variedades Riemannianas*	*- Algoritmo del Gradiente*
Y. *Convergencia del Método del Gradiente sobre variedades Riemannianas.*	Y_1**:** *Condiciones de optimalidad*	*- Existencia de mínimo global* *- Caracterización de mínimo* *- Elemento de análisi convexo*
	Y_2**:** *Resultados de análisis convexo y cuasi-convexidad*	*- Convergencia de funciones cuasi convexas* *- Búsqueda de Armijo generalizado*

Elaboración propia