



UNIVERSIDAD AUTÓNOMA DE
CHIAPAS



FACULTAD DE CIENCIAS EN FÍSICA Y
MATEMÁTICAS

Optimización integrada de
localización-inventario en logística
humanitaria mediante L-BFGS-B:
formulación, implementación y análisis en el
estado de Chiapas

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADA EN MATEMÁTICAS APLICADAS

PRESENTA:
YULISSA DEL ROCÍO HERNÁNDEZ VÁZQUEZ

DIRECTOR:
Dr. Yofre Hernán García Gómez

Tuxtla Gutiérrez, Chiapas a 25 de abril de 2026.

Dedicatoria

A Dios, por guiar mis pasos y darme la fortaleza para transformar la distancia en propósito.

A mis padres, Nelva y Alfonso, por sostenerme desde lejos con llamadas al amanecer, por su apoyo incondicional, sus sacrificios silenciosos y por creer en mí incluso cuando yo misma dudaba.

A mi hermanita Esmeralda, por recordarme con su sonrisa que existía un mundo más allá de las ecuaciones y por ser mi puente a casa en cada audio y cada foto.

Y a toda mi familia Vázquez, por convertir cada visita en refugio y cada despedida en promesa.

Este título no solo lleva mi nombre: lleva el peso amoroso de quienes me esperaron al otro lado del camino.

Agradecimientos

A Dios, por ser mi fortaleza en los momentos de incertidumbre y por darme la claridad para transformar cada obstáculo en aprendizaje.

Al Dr. Yofre Hernán García Gómez, mi director de tesis, por su guía incansable, su paciencia para explicar lo complejo con claridad y por enseñarme que las matemáticas aplicadas tienen el poder de transformar vidas. Su rigor académico y su humanidad marcaron la diferencia en cada etapa de este trabajo.

A todos mis profesores y profesoras de la maestría, en especial al Dr. Aaron, quien creyó en mí cuando yo misma dudaba. Y al Mtro. Elesban, por su disposición siempre abierta para resolver dudas y por demostrarme que la perseverancia es tan importante como el talento en la investigación.

A mis padres, Nelva y Alfonso, por su fe inquebrantable en mis capacidades y por sostenerme con amor en cada etapa de este camino. Y a mi hermanita Esmeralda y a toda mi familia, por ser mi refugio en las ausencias, mi alegría en los descansos y el recordatorio constante de que detrás de cada logro hay quienes celebran en silencio.

Y finalmente, a las comunidades chiapanecas que enfrentan las inundaciones temporada tras temporada: que este trabajo contribuya, aunque sea mínimamente, a construir sistemas de respuesta humanitaria más ágiles, justos y cercanos a quienes más lo necesitan.

Tabla de contenidos

1	Resumen	5
1.1	Abstract	5
2	Introducción	7
3	Objetivos	9
3.1	Objetivo general	9
3.2	Objetivos específicos	9
4	Notación y preliminares	10
4.1	Conjuntos y espacios vectoriales	10
4.2	Convexidad y conjuntos especiales	10
4.3	Funciones y propiedades analíticas	11
4.4	Matrices definidas positivas y su papel en la optimización	15
4.5	Estructura general de problemas de optimización con restricciones	16
4.6	Teoría de probabilidad y variables aleatorias	17
4.7	Optimización no lineal y condiciones de optimalidad	17
5	Fundamentación Teórica y Formulación Matemática del Modelo	20
5.1	Introducción al problema integrado de localización e inventario	20
5.1.1	Relevancia en logística humanitaria y gestión de emergencias	20
5.1.2	Limitaciones de los enfoques clásicos bajo incertidumbre y no linealidad	21
5.1.3	Estructura híbrida: decisiones discretas (ubicación) y continuas (inventario)	25
5.2	Marco general de modelado: variables, parámetros y espacio de decisión	30
5.2.1	Conjuntos de índices: almacenes potenciales \mathcal{J} y localidades de demanda \mathcal{J}	30
5.2.2	Variables de decisión: binarias (x_i) y continuas (y_{ij})	32
5.2.3	Ejemplo ilustrativo: interpretación y acoplamiento de variables	33
5.2.4	Parámetros del sistema: costos fijos, unitarios, demanda y capacidades	34
5.2.5	Definición formal del espacio factible \mathcal{X}	39
5.2.6	Propiedades estructurales del espacio \mathcal{X}	40
5.2.7	Función objetivo general en problemas de localización–inventario	42
5.2.8	Propiedades analíticas de Z	43

5.3	Modelo clásico de lote económico (EOQ): bases y limitaciones	45
5.3.1	Supuestos deterministas del EOQ	45
5.3.2	Derivación analítica del tamaño óptimo de pedido	46
5.3.3	Inadecuación en contextos humanitarios: demanda incierta, faltantes costosos	48
5.3.4	Motivación para una extensión no lineal con penalización y riesgo	50
5.4	Extensión no lineal del modelo EOQ con incertidumbre y penalización	53
5.4.1	Definición rigurosa de la función de costo total extendida $Z(Q)$	53
5.4.2	Interpretación del término $Z_\sigma L$: nivel de servicio y riesgo	56
5.4.3	Incorporación de la penalización por faltantes mediante β_C	59
5.4.4	Análisis de no linealidad: origen en D/Q y dependencia de L	62
5.4.5	Comparación cuantitativa con EOQ clásico: efecto descendente en Q^*	65
5.5	Propiedades analíticas del problema unidimensional	69
5.5.1	Formulación del problema de optimización: minimización de $f(Q)$ sujeto a $Q \geq 1$	69
5.5.2	Continuidad y coercividad: garantía de existencia del mínimo	71
5.5.3	Análisis de convexidad local y unicidad del óptimo	72
5.5.4	Comportamiento asintótico de $f(Q)$ cuando $Q \rightarrow 1^+$ y $Q \rightarrow \infty$	74
5.6	Condiciones de optimalidad y solución cerrada aproximada	77
5.6.1	Derivación formal de $\frac{dZ}{dQ}$	77
5.6.2	Condición de primer orden: $\frac{dZ}{dQ}(Q^*) = 0$	79
5.6.3	Expresión explícita de Q^* sin restricción de caja	81
5.6.4	Tratamiento de la restricción $Q \geq 1$: proyección del óptimo	83
5.6.5	Sensibilidad del óptimo respecto a β , Z_α , H , y S	85
5.7	Conclusiones del capítulo	88
6	Métodos numéricos y optimización con restricciones tipo caja	89
6.1	Fundamentos de métodos iterativos de primer y segundo orden	90
6.1.1	Método de gradiente descendente y sus propiedades de convergencia	90
6.1.2	Métodos cuasi-Newton: aproximación del Hessiano y convergencia superlineal	92
6.1.3	Familia Broyden y el método BFGS como caso prototípico	96
6.2	Extensión a dominios con restricciones tipo caja	99
6.2.1	Proyección ortogonal sobre conjuntos de cotas: definición y propiedades	99
6.2.2	El algoritmo L-BFGS-B: formulación y justificación estructural	101
6.2.3	Condiciones de optimalidad bajo cotas: KKT para problemas con restricciones tipo caja	104
6.3	Convergencia del método L-BFGS-B	106
6.3.1	Hipótesis de regularidad: Lipschitz-continuidad del gradiente y compacidad del dominio	106
6.3.2	Teorema de convergencia global: límite de puntos estacionarios que satisfacen KKT	109

6.3.3	Corolario aplicado al modelo propuesto: existencia y unicidad de $Q^* \geq 1$ tal que $\nabla f(Q^*) = 0$	112
6.4	Implementación y análisis numérico	115
6.4.1	Estrategias de inicialización y escalado del inverso del Hessiano aproximado	115
6.4.2	Estimación empírica del número de iteraciones y dependencia del parámetro de memoria m	119
6.4.3	Sensibilidad a las condiciones iniciales y robustez del algoritmo	122
6.4.4	Comparación numérica con otros métodos: gradiente descendente, Newton inexacto y BFGS sin restricciones	125
6.5	Discusión y perspectivas computacionales	129
6.5.1	Ventajas del enfoque de memoria limitada en problemas de gran escala	129
6.5.2	Limitaciones en problemas mal condicionados y alternativas híbridas	132
6.5.3	Posibilidades de extensión a problemas con estructura parcialmente separable	136
7	Implementación y herramientas computacionales	140
7.1	Descripción del entorno computacional	140
7.2	Estructura general del modelo computacional	140
7.3	Implementación del algoritmo L -BFGS- B	141
8	Formulación del Problema	143
8.1	Supuestos del modelo	143
8.2	Parámetros	143
8.3	Variables de decisión	144
8.4	Función objetivo	144
8.5	Restricciones del modelo	145
8.6	Función de pérdida logística	145
8.7	Modelo de inventario	148
8.8	Fill rate global	150
8.9	Escenario base estado de Veracruz	150
8.9.1	Resultados obtenidos	156
9	Escenario principal estado de Chiapas	158
9.1	Contexto geográfico y socioeconómico del área de estudio	158
9.1.1	Características del municipio de Cacahoatán	158
9.1.2	Vulnerabilidad ante inundaciones	159
9.2	Metodología para el cálculo de pesos posicionales	160
9.2.1	Enfoque multicriterio para la selección de ubicaciones estratégicas	160
9.2.2	Componentes del índice y justificación teórica	160
9.3	Resultados del cálculo de pesos posicionales	161
9.3.1	Ranking de localidades estratégicas	161
9.3.2	Selección de almacenes estratégicos	162

9.4	Diseño de la red logística optimizada	163
9.4.1	Localidades inundables y asignaciones	163
9.4.2	Distribución poblacional por grupos de edad	163
9.5	Resultados de la optimización del sistema	164
9.5.1	Eficiencia del sistema logístico implementado	164
9.5.2	Interpretación y comparación con Veracruz	165
9.5.3	Inventario humanitario optimizado	165
9.5.4	Análisis de la optimización	168
9.6	Validación y análisis de robustez	169
9.6.1	Escenarios de prueba implementados	170
9.6.2	Métricas de desempeño en condiciones estándar	170
9.7	Conclusiones del caso Chiapas	170
9.7.1	Implicaciones para la planeación logística	171
9.7.2	Perspectivas de escalabilidad	171
10	Conclusiones	172
	Referencias	176

1 Resumen

La logística humanitaria en contextos de desastre requiere estrategias robustas que anticipen la incertidumbre operativa. En este trabajo se propone un modelo de optimización determinista robusto para la gestión integrada de la ubicación de almacenes preposicionados y el control de inventario ante inundaciones en México. El modelo extiende el clásico problema de la cantidad económica de pedido (EOQ) al incorporar inventario de seguridad, penalización por faltantes y decisiones binarias de apertura de almacenes, utilizando la demanda esperada y su variabilidad derivadas de una distribución normal como parámetros fijos. La formulación resultante es un problema de optimización mixta entera no lineal (MINLP) con una función objetivo no convexa y no separable. Para su resolución, se emplea un esquema híbrido que combina enumeración sobre las variables binarias con el algoritmo L-BFGS-B para los subproblemas continuos. El modelo se evalúa mediante estudios de caso en regiones vulnerables a inundaciones en Chiapas y Veracruz, demostrando que considerar la incertidumbre parametrizada reduce significativamente el costo total esperado y mejora la robustez frente a eventos extremos, en comparación con enfoques deterministas que ignoran la variabilidad en la demanda o en el tiempo de entrega.

Palabras clave: logística humanitaria, ubicación–inventario, EOQ robusto, incertidumbre parametrizada, MINLP, L-BFGS-B, inundaciones.

1.1. Abstract

Humanitarian logistics in disaster contexts demands robust strategies that account for operational uncertainty. This work proposes a robust deterministic optimization model for the integrated management of prepositioned warehouse location and inventory control under flood risk in Mexico. The model extends the classical Economic Order Quantity (EOQ) problem by incorporating safety stock, shortage penalties, and binary warehouse opening decisions, using expected demand and its variability derived from a normal distribution as fixed parameters. The resulting formulation is a mixed integer nonlinear programming (MINLP) problem with a non convex and non separable objective function. A hybrid solution scheme is employed, combining enumeration over binary variables with the L-BFGS-B algorithm for continuous subproblems. The model is evaluated through case studies in flood prone regions of Chiapas and Veracruz. Results show that incorporating parametrized uncertainty significantly reduces the expected total cost and

enhances robustness against extreme events, compared to deterministic approaches that neglect demand or lead time variability.

Keywords: humanitarian logistics, location–inventory, robust EOQ, parametrized uncertainty, MINLP, L-BFGS-B, floods.

2 Introducción

Las inundaciones constituyen uno de los fenómenos naturales más frecuentes y devastadores a nivel global, afectando la infraestructura, el acceso a servicios básicos y, sobre todo, la vida y dignidad de las personas. En este contexto, el diseño y aplicación de modelos matemáticos orientados a la logística humanitaria cobra una relevancia crítica, no sólo como una herramienta de gestión operativa, sino como un instrumento de justicia social y resiliencia territorial. Barojas-Payán et al. (2021)

En esa línea, Insani, Widodo, y Rahman (2024) desarrollaron un modelo de programación entera mixta orientado a coordinar simultáneamente la evacuación y la entrega de ayuda en contextos de inundaciones tempranas. Su propuesta incorpora entregas divididas, reutilización de vehículos y múltiples viajes, siendo resuelta mediante un algoritmo genético modificado que alcanzó una eficiencia 92.5 % superior frente a métodos exactos tradicionales.

Por su parte, Sheikholeslami y Zarrinpoor (2022) propusieron un modelo de programación lineal entera mixta multiperíodo bajo condiciones de incertidumbre. Este modelo integra restricciones difusas y estocásticas para optimizar la localización de almacenes, la gestión de inventarios y la provisión de atención médica posterior a los desastres.

En otro enfoque, Romero-Mancilla, Martínez-Flores, y Sánchez-Partida (2024) diseñaron un modelo multimodal que combina transporte terrestre y aéreo mediante drones, estructurado como un modelo multiobjetivo que considera transbordos y múltiples depósitos. Su objetivo principal es equilibrar el costo logístico con los tiempos de entrega, especialmente en escenarios donde la infraestructura vial ha sido severamente afectada.

Asimismo, Santana-Robles, López, y Rivera (2024) formularon un modelo híbrido que combina programación lineal entera con problemas de ruteo vehicular (VRP), enfocado en asignar refugios y optimizar la entrega de suministros bajo variaciones de demanda y recursos limitados.

De forma complementaria, Mashrut y Rahimi (2024) propuso un modelo robusto-fuzzy-probabilístico biobjetivo que busca minimizar tanto el costo operativo como el costo de privación, es decir, el impacto social derivado de la falta de ayuda oportuna.

Finalmente, Pujiana, Utama, y Rahmawati (2020) implementaron un modelo de ruteo multi-depósito (MDVRP) aplicado a la fase posterior a inundaciones, optimizando el uso de depósitos temporales, rutas y cobertura territorial con base en restricciones de capacidad y demanda.

Estos enfoques demuestran que la preparación logística previa al desastre no sólo mejora la eficiencia de la respuesta, sino que permite reducir desigualdades territoriales y proteger de manera diferenciada a las comunidades más vulnerables. La implementación de modelos que integren criterios técnicos (distancia, inventario, costo), operativos (capacidad, transporte) y sociales (accesibilidad, prioridad) es indispensable para enfrentar los desafíos logísticos que imponen los desastres hidrometeorológicos. En el caso del presente modelo, se implementa en el estado de Veracruz como área piloto, con miras a expandirse hacia el estado de Chiapas, dado que comparte vulnerabilidades similares frente a fenómenos hidrometeorológicos, alta dispersión poblacional y limitada infraestructura logística.

Por todo lo anterior, el presente estudio se basa en la construcción de un modelo de optimización entero-mixto no lineal, que articula variables de localización de almacenes preposicionados, gestión de inventarios y niveles de servicio (*fill rate*), con el fin de diseñar una red logística humanitaria eficiente, flexible y ética. Este modelo se inspira en las mejores prácticas de la literatura científica reciente, adaptándolas a condiciones de incertidumbre, alta demanda y restricciones operativas.

3 Objetivos

3.1. Objetivo general

Implementar un modelo de optimización **determinista** para la gestión integrada de la ubicación de almacenes preposicionados y el control de inventario humanitario en zonas propensas a inundaciones en el sureste de México.

3.2. Objetivos específicos

1. **Adaptar un modelo de localización–inventario humanitario bajo incertidumbre parametrizada**, que extienda el modelo clásico de cantidad económica de pedido (EOQ) para incluir inventario de seguridad, faltantes estocásticos penalizados y decisiones binarias de apertura de almacenes, utilizando la demanda esperada y su variabilidad como parámetros exógenos fijos.
2. **Desarrollar un esquema de solución híbrido basado en optimización mixta entera no lineal (MINLP)**, que combine enumeración sobre variables binarias con el algoritmo **L-BFGS-B** para resolver eficientemente los subproblemas continuos no convexos asociados al tamaño del lote y al nivel de inventario de seguridad.
3. **Verificar el desempeño del modelo propuesto mediante estudios de caso en regiones vulnerables a inundaciones en Veracruz y replicar en el estado de Chiapas**, cuantificando la mejora en robustez y costo esperado respecto a enfoques deterministas que ignoran la incertidumbre en la demanda o en el tiempo de entrega.

4 Notación y preliminares

Antes de presentar los modelos y algoritmos propuestos, es necesario establecer las bases matemáticas y la notación que se utilizarán a lo largo de esta tesis. Este capítulo reúne conceptos fundamentales de análisis, álgebra lineal, optimización y teoría de probabilidad que permiten formular rigurosamente el problema de localización–inventario bajo incertidumbre, caracterizar su estructura no convexa y justificar las estrategias de solución empleadas. El enfoque no es exhaustivo, sino selectivo: solo se incluyen aquellos elementos directamente relevantes para el desarrollo posterior.

4.1. Conjuntos y espacios vectoriales

En la formulación de redes logísticas humanitarias, las decisiones suelen representarse en espacios euclidianos de dimensión finita. Dos subconjuntos relevantes son:

- \mathbb{R}^n : espacio vectorial real de dimensión n .
- $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_i \geq 0, \forall i\}$: ortante no negativo, usado para modelar flujos, inventarios o asignaciones no negativas.
- **Conjunto compacto**: Un subconjunto $K \subseteq \mathbb{R}^n$ es **compacto** si es **cerrado y acotado**. Esta propiedad es fundamental en optimización, ya que garantiza que toda sucesión en K tiene una subsucesión convergente cuyo límite pertenece a K , lo cual asegura la existencia de mínimos globales para funciones continuas (teorema de Weierstrass).

Estos conjuntos garantizan que las variables del modelo respeten restricciones físicas (por ejemplo, no es posible enviar una cantidad negativa de ayuda humanitaria).

4.2. Convexidad y conjuntos especiales

La convexidad juega un papel central en la teoría de optimización, ya que, cuando el conjunto factible es convexo y la función objetivo también es convexa sobre dicho conjunto, se garantiza que todo mínimo local es global. Esta propiedad simplifica enormemente la búsqueda de soluciones óptimas. En el contexto de redes logísticas humanitarias, muchos subproblemas como la asignación de flujos o el cálculo de inventarios dan lugar a regiones factibles convexas cuando las decisiones de apertura de almacenes están fijas. Sin embargo, la presencia de variables binarias (por ejemplo, decidir si abrir o no un

centro de distribución) introduce no convexidad estructural, lo que complica el análisis y requiere métodos especializados. Comprender la geometría de los conjuntos convexos y poliedros permite identificar cuándo un subproblema es tratable, diseñar descomposiciones eficientes y justificar el uso de relajaciones convexas en algoritmos de solución.

La estructura geométrica del espacio factible influye directamente en la complejidad del problema. En particular:

- **Combinación convexa:** Dados dos puntos $x, y \in \mathbb{R}^n$, cualquier punto de la forma

$$z = \lambda x + (1 - \lambda)y, \quad \text{con } \lambda \in [0, 1],$$

se denomina **combinación convexa** de x e y . Esta operación genera todos los puntos en el segmento de recta que une x con y . Un conjunto $C \subseteq \mathbb{R}^n$ es **convexo** si contiene todas las combinaciones convexas de cualesquiera dos de sus puntos.

- **Conjunto convexo:** $C \subseteq \mathbb{R}^n$ es convexo si para todo $x, y \in C$ y $\lambda \in [0, 1]$, se cumple $\lambda x + (1 - \lambda)y \in C$. En logística, la convexidad permite combinar estrategias factibles sin violar restricciones. Sin embargo, la presencia de decisiones binarias (como abrir o no un almacén) rompe esta propiedad, como se discute en el Capítulo 3.
- **Poliedro convexo:** conjunto de la forma $\{x \in \mathbb{R}^n : Ax \leq b\}$ con $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Es cerrado y convexo (Bazaraa, Sherali, y Shetty (2013a)), y describe sistemas de restricciones lineales como límites de capacidad o cobertura de demanda.
- **Conjunto cerrado en dimensión finita:** En \mathbb{R}^n con la topología euclidiana, un conjunto es **cerrado** si contiene todos sus puntos límite. Todo subespacio vectorial de \mathbb{R}^n es cerrado, y en dimensión finita, los conjuntos cerrados y acotados son compactos (teorema de Heine–Borel). Esta propiedad es clave para garantizar existencia de soluciones en optimización.

4.3. Funciones y propiedades analíticas

Las propiedades analíticas de la función objetivo como su diferenciabilidad, curvatura (Hessiano) y estructura (separabilidad, convexidad) determinan directamente la elección del método de solución, la calidad de las soluciones obtenidas y la eficiencia computacional del proceso de optimización. En el contexto de la logística humanitaria bajo incertidumbre, la función de costo total combina términos no lineales (como el inventario de seguridad basado en la distribución normal) con decisiones discretas (por ejemplo, abrir o no un almacén), lo que da lugar a un paisaje no convexo y no separable. Aun así, comprender propiedades locales como la suavidad del gradiente o la Lipschitz-continuidad del Hessiano permite aplicar métodos iterativos robustos (por ejemplo, L-BFGS-B) y garantizar su convergencia. Además, cuando se fijan las decisiones discretas (es decir, se decide qué

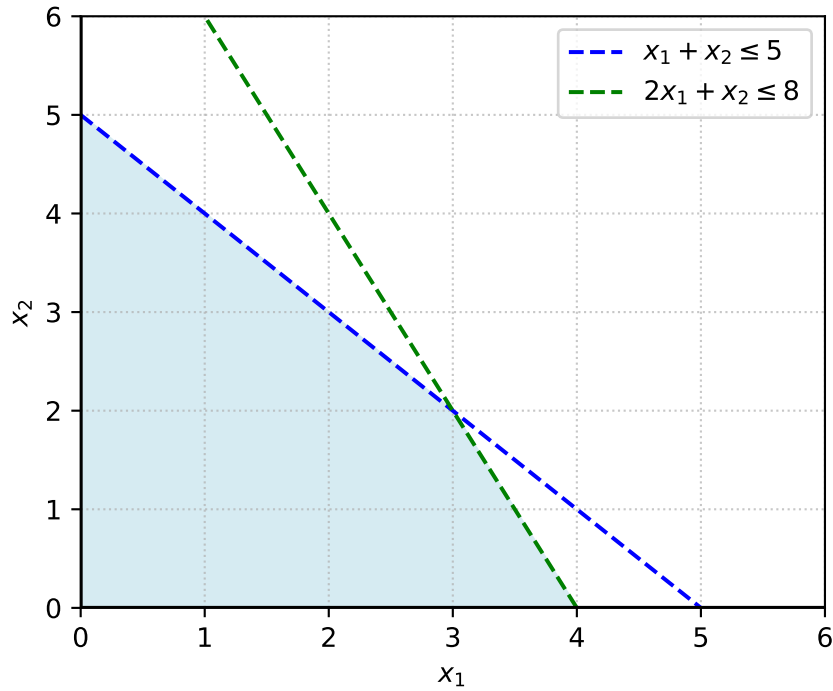


Figura 4.1: Representación gráfica de un poliedro convexo en \mathbb{R}^2 , definido por el sistema de desigualdades lineales $x_1 + x_2 \leq 5$, $2x_1 + x_2 \leq 8$, $x_1 \geq 0$, $x_2 \geq 0$. La región factible, sombreada en azul claro, corresponde al conjunto de soluciones admisibles para un subproblema de asignación de inventario o flujo en un modelo de localización-inventario. Su estructura poliédrica refleja la linealidad de las restricciones de capacidad, cobertura de demanda y no negatividad, y garantiza convexidad cuando las decisiones de apertura de almacenes están fijas.

almacenes están abiertos), el problema restante suele ser fuertemente convexo o casi separable, lo que facilita el diseño de algoritmos eficientes y el análisis de sensibilidad.

La formulación del costo total en inventario humanitario depende de propiedades analíticas que facilitan el análisis de optimalidad:

- **Función continuamente diferenciable:** Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **continuamente diferenciable** en un conjunto abierto $\mathcal{U} \subseteq \mathbb{R}^n$ si todas sus derivadas parciales existen y son funciones continuas en \mathcal{U} . En este caso, se dice que $f \in C^1(\mathcal{U})$. Esta propiedad garantiza que el gradiente $\nabla f(x)$ varía suavemente en \mathcal{U} , lo cual es esencial para la convergencia de métodos iterativos como el gradiente descendente y para la validez de las condiciones de optimalidad de primer orden (Nocedal y Wright (2006), Sección 2.1).
- **Hessiano:** Dada una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dos veces continuamente diferenciable en un conjunto abierto $\mathcal{U} \subseteq \mathbb{R}^n$, su **matriz Hessiana** (o simplemente **Hessiano**) en un punto $x \in \mathcal{U}$ es la matriz simétrica de segundas derivadas parciales:

$$\nabla^2 f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right]_{i,j=1}^n.$$

El Hessiano describe la curvatura local de f y es fundamental en el análisis de segundo orden, en la caracterización de mínimos locales y en métodos de optimización como Newton o cuasi-Newton. Si $\nabla^2 f(x) \succ 0$ (definida positiva), entonces x es un mínimo local estricto bajo condiciones de primer orden.

- **Hessiano Lipschitz-continuo:** Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función dos veces continuamente diferenciable en un conjunto abierto $\mathcal{U} \subseteq \mathbb{R}^n$. Se dice que su **Hessiano** $\nabla^2 f$ es **Lipschitz-continuo** en \mathcal{U} si existe una constante $L_H > 0$ tal que

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L_H \|x - y\|, \quad \forall x, y \in \mathcal{U},$$

donde $\|\cdot\|$ denota una norma matricial compatible (por ejemplo, la norma espectral). Esta propiedad garantiza que la curvatura de f no varía de forma abrupta y es clave para establecer tasas de convergencia superlineal o cuadrática en métodos de segundo orden, como Newton o BFGS (Nocedal y Wright (2006), Sección 6.4).

- **Función cuadrática fuertemente convexa:** Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **cuadrática fuertemente convexa** si puede escribirse como

$$f(x) = \frac{1}{2}x^\top Qx - b^\top x + c,$$

donde $Q \in \mathbb{R}^{n \times n}$ es una matriz **simétrica y definida positiva** ($Q \succ 0$), $b \in \mathbb{R}^n$ y $c \in \mathbb{R}$. La condición $Q \succ 0$ implica que existe una constante $\mu > 0$

tal que

$$x^\top Qx \geq \mu \|x\|^2, \quad \forall x \in \mathbb{R}^n,$$

lo que garantiza que f tiene un único mínimo global y que su paisaje de optimización es “bien condicionado”. Esta propiedad es fundamental para asegurar convergencia rápida de métodos iterativos como el gradiente descendente o BFGS.

- **Función convexa:** $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es convexa si

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y, \lambda \in [0, 1].$$

La convexidad garantiza que todo mínimo local sea global, una propiedad deseable en diseño de redes.

- **Función estrictamente cuadrática:** $f(x) = \frac{1}{2}x^\top Qx - b^\top x + c$, con $Q \in \mathbb{R}^{n \times n}$ simétrica. Si Q es constante, el Hessiano $\nabla^2 f(x) = Q$ también lo es, lo que simplifica el análisis de curvatura en modelos de costo.
- **Direcciones conjugadas:** Dada una matriz simétrica definida positiva $Q \in \mathbb{R}^{n \times n}$, un conjunto de vectores no nulos $\{p_0, p_1, \dots, p_k\} \subseteq \mathbb{R}^n$ es **Q -conjugado** (o **conjugado respecto a Q**) si

$$p_i^\top Qp_j = 0, \quad \forall i \neq j.$$

En el contexto de minimización de una función cuadrática $f(x) = \frac{1}{2}x^\top Qx - b^\top x + c$, las direcciones conjugadas permiten alcanzar el mínimo global en a lo sumo n iteraciones, ya que exploran subespacios mutuamente ortogonales en la métrica inducida por Q . Los métodos de la familia de Broyden (incluyendo BFGS y DFP) generan tales direcciones cuando se aplican a funciones cuadráticas con búsquedas lineales exactas.

- **Rapidez de convergencia:** Sea $\{x_k\}$ una sucesión que converge a un punto x^* . La **rapidez de convergencia** describe qué tan rápido $\|x_k - x^*\|$ tiende a cero cuando $k \rightarrow \infty$. Las tasas más comunes son:
- **Lineal:** existe $r \in (0, 1)$ tal que

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq r.$$

- **Superlineal:**

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

- **Cuadrática:** existe $C > 0$ tal que

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \leq C.$$

Estas definiciones permiten comparar la eficiencia asintótica de métodos iterativos (Nocedal y Wright (2006), Sección 2.3).

- **Separabilidad:** $f(x)$ es separable si puede escribirse como $f(x) = \sum_{i=1}^n f_i(x_i)$. Esta propiedad permite descomponer el problema en subproblemas independientes, útil en redes con nodos débilmente acoplados. La función de costo **no es separable**, debido a la interacción entre tamaño de lote, inventario de seguridad y penalización por faltantes.

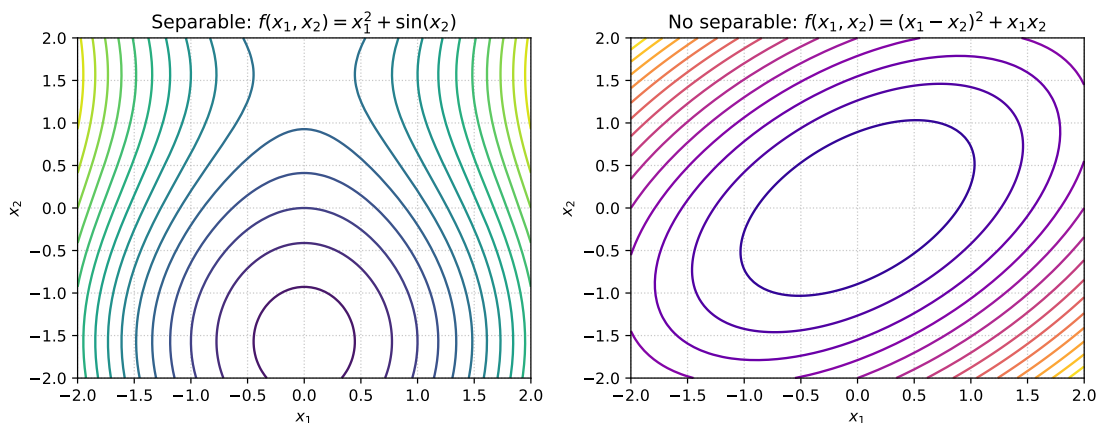


Figura 4.2: Comparación entre una función separable y una no separable mediante curvas de nivel. A la izquierda, $f(x_1, x_2) = x_1^2 + \sin(x_2)$ es separable, pues su estructura no acopla las variables: las curvas de nivel son alineadas con los ejes, lo que permite optimización por coordenadas. A la derecha, $f(x_1, x_2) = (x_1 - x_2)^2 + x_1x_2$ es no separable, con curvas inclinadas que reflejan interacción entre decisiones una característica clave en modelos de inventario humanitario, donde el tamaño del lote, el inventario de seguridad y la penalización por faltantes están fuertemente acoplados.

4.4. Matrices definidas positivas y su papel en la optimización

En optimización de funciones en varias variables, la **matriz Hessiana** $\nabla^2 f(x)$ describe la curvatura local de la función objetivo f . Una propiedad fundamental es si esta matriz es **definida positiva**, lo cual se denota como $\nabla^2 f(x) \succ 0$.

Formalmente, una matriz simétrica $A \in \mathbb{R}^{n \times n}$ es **definida positiva** si

$$x^\top Ax > 0 \quad \text{para todo } x \in \mathbb{R}^n, x \neq 0.$$

Esta condición garantiza que la función f es **estrictamente convexa** en una vecindad del punto x , y por tanto, si además $\nabla f(x) = 0$, entonces x es un **mínimo local estricto**. En

modelos de inventario humanitario, donde la función de costo incluye términos cuadráticos o aproximaciones de segundo orden (por ejemplo, en el cálculo del inventario de seguridad bajo demanda normal), la definición positiva del Hessiano en un punto crítico garantiza que dicho punto es un mínimo local estricto y asegura la convergencia local de los métodos basados en curvatura, como Newton o L-BFGS. No obstante, la unicidad de la solución óptima no se deduce únicamente de esta propiedad puntual, sino que requiere hipótesis globales adicionales, como la convexidad estricta de la función objetivo en todo el conjunto factible.

4.5. Estructura general de problemas de optimización con restricciones

Los modelos de logística humanitaria integran decisiones de naturaleza distinta: **estratégicas** (por ejemplo, dónde ubicar centros de distribución) y **tácticas/operativas** (como cuánto inventario almacenar o cuánta ayuda enviar). Esta combinación da lugar a problemas de optimización **mixtos**, que involucran variables binarias (para las decisiones de apertura/cierre) y variables continuas no negativas (para flujos, inventarios o asignaciones).

En términos generales, el problema puede expresarse como:

$$\begin{aligned} \min_{x,y} \quad & f(x, y) \\ \text{sujeto a} \quad & g_i(x, y) \leq 0, \quad i = 1, \dots, m, \\ & x \in \{0, 1\}^p, \quad y \in \mathbb{R}_+^q, \end{aligned}$$

donde f representa el costo total esperado (incluyendo penalizaciones por faltantes y costos de operación), y las restricciones modelan capacidades, cobertura de demanda y lógica de activación.

Un componente esencial de esta estructura es la **vinculación entre variables discretas y continuas**, comúnmente implementada mediante la **formulación big-M**:

- **Formulación big-M estándar:** restricción de la forma $y \leq Mx$, con $x \in \{0, 1\}$, $y \in \mathbb{R}_+$, y $M > 0$ suficientemente grande, usada para modelar la implicación lógica $x = 0 \Rightarrow y = 0$. Por ejemplo, si no se abre un almacén ($x = 0$), no puede enviar ayuda ($y = 0$).

Esta estructura mixta introduce no convexidad y discontinuidad en el espacio de decisiones, lo que impide el uso directo de métodos basados en gradientes y motiva el uso de algoritmos especializados, como los discutidos en el Capítulo 5.

4.6. Teoría de probabilidad y variables aleatorias

En la logística humanitaria, la demanda de ayuda especialmente durante o después de eventos como inundaciones es inherentemente incierta y altamente variable. Modelar esta incertidumbre de forma adecuada es crucial: usar únicamente valores promedio ignora la posibilidad de eventos extremos, lo que puede llevar a niveles insuficientes de inventario y, en consecuencia, a faltantes críticos. Para abordar este desafío sin comprometer la tratabilidad del modelo de optimización, en este trabajo se emplean herramientas de la teoría de probabilidad **de forma previa** para caracterizar estadísticamente la demanda. Esta caracterización permite obtener estimaciones puntuales (demanda esperada y varianza) y derivar factores de seguridad (Z_σ), los cuales se incorporan posteriormente como parámetros fijos y deterministas en la formulación. De esta forma, se cuantifica tanto la magnitud esperada como la variabilidad y el riesgo asociado a las colas de la distribución (eventos raros pero severos), aspectos esenciales para el cálculo del inventario de seguridad y la toma de decisiones robustas dentro de un marco determinista

- **Variable aleatoria (v.a.) absolutamente continua:** $\mathbb{P}(X \in A) = \int_A f_X(x) dx$, donde f_X es la densidad respecto a la medida de Lebesgue. Esta formulación es esencial para modelar demandas con distribuciones suaves, como la normal.
- **Distribución normal:** si $D_L \sim \mathcal{N}(\mu_L, \sigma_L^2)$, su densidad es

$$\phi_{D_L}(d) = \frac{1}{\sqrt{2\pi}\sigma_L} \exp\left(-\frac{(d - \mu_L)^2}{2\sigma_L^2}\right).$$

Esta distribución se usa para aproximar la demanda durante el tiempo de entrega, lo que permite calcular el inventario de seguridad.

- **Varianza:** para una v.a. X , $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$. La varianza cuantifica la incertidumbre en la demanda; ignorarla (por ejemplo, al usar solo $\mathbb{E}[d_j]$) puede inducir soluciones frágiles en eventos extremos.
- **Notación \ll :** indica que una cantidad es mucho menor que otra. Por ejemplo, en la Sección 5.4.4.4 se muestra que, si se ignora la incertidumbre en L , la varianza estimada puede ser $80 \ll 10\,080$, subestimando severamente el riesgo.

El **valor esperado** $\mathbb{E}[d_j]$ representa la demanda media en la localidad j . Aunque es común usarlo como estimador puntual en modelos deterministas, como se advierte en la Sección 5.1.2.3, esta simplificación puede ser inadecuada en logística humanitaria, donde la cola de la distribución (eventos extremos) es crítica.

4.7. Optimización no lineal y condiciones de optimalidad

El problema de localización–inventario analizado en este trabajo da lugar a un modelo de optimización no lineal determinista. Aunque en la Sección 4.6 se introducen conceptos

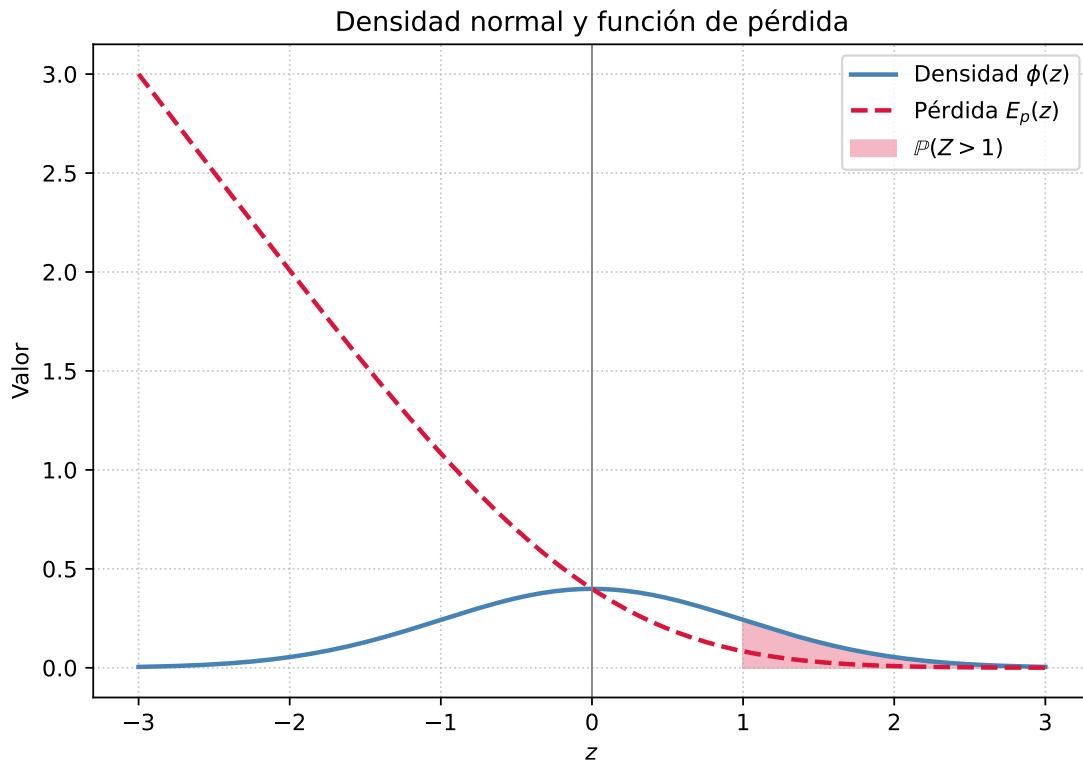


Figura 4.3: Función de densidad de probabilidad $\phi(z)$ y función de pérdida logística $E_p(z) = z(\Phi(z) - 1) + \phi(z)$ asociadas a la distribución normal estándar. La densidad $\phi(z)$ (azul sólido) modela la incertidumbre en la demanda estandarizada, mientras que la función de pérdida $E_p(z)$ (rojo punteado) cuantifica el déficit esperado normalizado cuando el nivel de inventario es insuficiente ($z < 0$). La región sombreada para $z \geq 1$ ilustra la probabilidad de cola superior, relevante para el cálculo del inventario de seguridad en contextos humanitarios. Esta función es fundamental para internalizar el riesgo de faltantes en la formulación del costo total.

de teoría de probabilidad y variables aleatorias, estos se emplean exclusivamente para fundamentar la obtención paramétrica de los insumos del modelo (demanda esperada, varianza y factor de seguridad Z_σ). La incertidumbre no se representa mediante variables aleatorias ni formulaciones estocásticas dentro del problema de optimización; en su lugar, se internaliza de forma fija y determinista en los coeficientes de la función objetivo. Bajo esta estructura, las **condiciones de Karush–Kuhn–Tucker (KKT)** proporcionan un marco teórico fundamental para caracterizar puntos candidatos a óptimos locales, siempre que se satisfagan ciertas condiciones de regularidad (calificación de restricciones). Sin embargo, dado que el problema incluye variables binarias y acoplamiento no lineal entre decisiones, el espacio factible es **no convexo y discontinuo**, lo que implica que las condiciones KKT aunque necesarias **no garantizan optimalidad global**. Esta limitación justifica el uso de métodos de solución especializados, como algoritmos de ramificación o heurísticas híbridas, que se discuten en capítulos posteriores.

- **Condiciones de Karush–Kuhn–Tucker (KKT)**: condiciones necesarias de optimalidad local bajo restricciones de desigualdad y calificación de restricciones. Como se expone en Lewis y Overton (2013), estas condiciones **no garantizan optimalidad global** en problemas no convexos, como el propuesto, donde la no convexidad surge de las variables binarias y la no linealidad en la incertidumbre.

5 Fundamentación Teórica y Formulación Matemática del Modelo

5.1. Introducción al problema integrado de localización e inventario

5.1.1. Relevancia en logística humanitaria y gestión de emergencias

La logística humanitaria desempeña un papel fundamental en la respuesta a desastres, al coordinar la disponibilidad, distribución y acceso equitativo a bienes esenciales en entornos caracterizados por alta incertidumbre, infraestructura comprometida y recursos limitados (Balcik y Beamon (2008); Caunhye, Nie, y Pokharel (2012)). A diferencia de los sistemas logísticos comerciales, donde predominan criterios de eficiencia económica, los contextos humanitarios priorizan objetivos de **rapidez, cobertura universal y mitigación del riesgo**, lo que exige modelos matemáticos capaces de integrar decisiones estratégicas y tácticas bajo condiciones de información imperfecta.

En este marco, el **problema integrado de localización e inventario** se ha consolidado como una herramienta clave para la planificación anticipada, ya que permite determinar, de manera conjunta:

1. **La ubicación óptima de almacenes de preposicionamiento, y**
2. **Los niveles de inventario que deben almacenarse antes de la ocurrencia de un evento disruptivo.**

Ambas decisiones condicionan el desempeño global del sistema en términos de costo total, robustez operativa y equidad en la atención. Su naturaleza híbrida, estratégica (ubicación) y táctica (inventario), impone desafíos estructurales que no pueden abordarse mediante enfoques secuenciales o modelos clásicos aislados.

Aunque la demanda en estos escenarios es inherentemente incierta, en este trabajo se adopta un **enfoque determinista robusto**, ampliamente fundamentado en la teoría de inventarios bajo demanda aleatoria del artículo seminal Hadley y Whitin (1963). Este enfoque no modela la incertidumbre mediante variables aleatorias ni formulaciones estocásticas, sino que la **incorpora de forma paramétrica** mediante un **stock de seguridad fijo**, derivado de estimaciones de la variabilidad de la demanda.

Específicamente, se asume que la demanda en cada localidad $j \in \mathcal{J}$ posee un valor esperado $d_j > 0$ y una desviación estándar $\sigma_j \geq 0$, ambos tratados como **parámetros exógenos conocidos**. Dado un tiempo de reposición $L > 0$ y un nivel de servicio objetivo $\alpha \in (0, 1)$, se define el **factor de seguridad** como $Z = \Phi^{-1}(\alpha)$, donde Φ^{-1} denota la inversa de la función de distribución acumulada de la normal estándar. Aunque esta construcción se fundamenta en la teoría de probabilidad, en la etapa de optimización dichos valores se fijan como constantes. De este modo, se adopta una aproximación determinista basada en cuantiles de una distribución. Este enfoque permite formular un modelo de optimización no lineal determinista, donde el término $Z\sigma\sqrt{L}$ se incorpora como un parámetro exógeno que internaliza la magnitud de la incertidumbre sin introducir variables aleatorias ni operadores de esperanza dentro de la función objetivo.

Este enfoque permite formular un **modelo de optimización no lineal determinista**, cuya estructura matemática se preserva intacta frente a la complejidad computacional inherente a los modelos estocásticos. La ausencia de esperanzas, escenarios o variables de segundo nivel garantiza trazabilidad analítica, estabilidad numérica y compatibilidad con métodos de optimización eficientes para problemas con cotas, como se discutirá en capítulos posteriores.

5.1.2. Limitaciones de los enfoques clásicos bajo incertidumbre y no linealidad

Los modelos clásicos de localización e inventario, como el Uncapacitated Facility Location Problem (UFLP) y el Economic Order Quantity (EOQ) model o sus variantes deterministas, han sido ampliamente estudiados bajo el supuesto de **información perfecta y estructura lineal**, un enfoque cuya pertinencia disminuye notablemente en escenarios humanitarios caracterizados por **incertidumbre estructural y no linealidad inherente** en costos y restricciones como se representa en Hadley y Whitin (1963); Snyder (2006).

5.1.2.1. Supuestos restrictivos del modelo EOQ clásico

El modelo EOQ se fundamenta en supuestos como:

1. Demanda constante y conocida: $D > 0$ (unidades por unidad de tiempo).
2. Tiempo de entrega determinista y conocido: $L \geq 0$.
3. Costos lineales y constantes: costo fijo por pedido $S > 0$, costo unitario de mantenimiento de inventario $H > 0$, y costo unitario de adquisición $C > 0$. No se permiten faltantes.
4. Reposición instantánea y sin restricciones de capacidad, de acuerdo con los supuestos planteados por Hadley y Whitin (1963).

Bajo estos supuestos, la función de costo total promedio por unidad de tiempo se expresa como:

$$Z^{EOQ}(Q) = \frac{D}{Q}S + \frac{Q}{2}H + DC,$$

donde:

- $Q > 0$ es el tamaño del lote o cantidad de pedido (variable de decisión continua),
- D es la demanda esperada por periodo de planificación,
- S es el costo fijo asociado a la emisión de cada pedido,
- H es el costo unitario de mantener una unidad en inventario por unidad de tiempo,
- C es el costo unitario de adquisición del producto.

La minimización de esta función conduce a la solución óptima clásica:

$$Q^{EOQ} = \sqrt{\frac{2DS}{H}}.$$

Este resultado depende de la **deterministicidad** y la **convexidad estricta**. Sin embargo, en logística humanitaria:

- La demanda es aleatoria $D(\omega)$.
- L es incierto debido a afectaciones en infraestructura.
- Los faltantes deben modelarse mediante una penalización finita $\beta > 0$.
- Los costos pueden presentar economías de escala y no convexidades, es decir, situaciones donde el costo promedio por unidad disminuye al aumentar el volumen de pedido o transporte, como se discute en Porteus (2002).

5.1.2.2. Incapacidad del EOQ para capturar riesgo

Aunque la demanda en contextos humanitarios es inherentemente estocástica, en este trabajo se adopta un enfoque determinista robusto, ampliamente utilizado en la teoría de inventarios bajo incertidumbre como se expone en Porteus (2002) y Zipkin (2000). Bajo los supuestos clásicos de demanda normal durante el tiempo de entrega, política de revisión continua y costo lineal por faltante, el costo total esperado puede aproximarse mediante una función determinista del tamaño de lote Q . Esta aproximación incorpora los efectos de la incertidumbre a través del inventario de seguridad $Z_\alpha \sigma_L$ y la penalización esperada por faltantes, pero su optimización se realiza como un problema determinista.

Si la demanda durante el tiempo de entrega se modela como $D_L \sim \mathcal{N}(\mu_L, \sigma_L^2)$, el nivel de inventario debe satisfacer:

$$\mathbb{P}(D_L \leq Q/2 + Z_\alpha \sigma_L) \geq \alpha,$$

lo que justifica la inclusión del término $Z_\alpha \sigma_L$ en la función de costo. Con ello, la función de costo extendida adopta la forma: Con ello, la función de costo extendida se vuelve:

$$Z(Q) = \frac{D}{Q}S + \left(\frac{Q}{2} + Z_\alpha \sigma_L \right) H + DC + \frac{D}{Q}\beta C. \quad (5.1.1)$$

Esta función **no es convexa globalmente** cuando σ_L depende de Q , como ocurre cuando existe congestión en transporte o variabilidad dependiente del tamaño del pedido, aspectos estudiados por Bertsimas y Thiele (2006) y Shapiro, Dentcheva, y Ruszczyński (2021).

Proposición (No linealidad estricta).

La función $Z(Q)$ definida en Sección 5.1.2.2 es diferenciable en $(0, \infty)$. Su curvatura depende de σ_L :

- Si σ_L es constante: entonces

$$\frac{d^2 Z}{dQ^2} = \frac{2D(S + \beta C)}{Q^3} > 0.$$

por lo tanto $Z(Q)$ es estrictamente convexa.

- Si $\sigma_L = \sigma_L(Q)$ depende de Q : la segunda derivada incluye términos adicionales que pueden cambiar de signo, y $Z(Q)$ puede perder convexidad global.

Esto demuestra que muchos métodos clásicos basados en convexidad y separabilidad (ver Sección 4.3) resultan inaplicables en condiciones realistas.

5.1.2.3. Limitaciones de modelos de localización deterministas

El modelo clásico UFLP resuelve:

$$\min_{x \in \{0,1\}^m, y \geq 0} \sum_i f_i x_i + \sum_i \sum_j c_{ij} y_{ij},$$

bajo restricciones de asignación deterministas. Sin embargo:

- La demanda d_j es aleatoria.
- Los costos pueden variar según accesibilidad o daños.
- Las capacidades reales son inciertas.

Modelar d_j mediante $\mathbb{E}[d_j]$ genera soluciones frágiles frente a escenarios adversos Snyder (2006).

5.1.2.4. Ejemplo ilustrativo: EOQ clásico vs. EOQ extendido con stock de seguridad y penalización

Considérese un contexto en el que la demanda esperada es $D = 1000$ unidades, con una variabilidad estimada cuya desviación estándar es $\sigma = 10$. Aunque la demanda real puede fluctuar (por ejemplo, entre 800 y 1200 unidades en eventos históricos), en este enfoque se modela la incertidumbre mediante parámetros fijos derivados de su comportamiento estadístico. Los demás parámetros son:

- Costo de pedido: $S = 100$
- Costo unitario de mantenimiento: $H = 5$
- Costo unitario de adquisición: $C = 10$
- Penalización por faltante: $\beta = 50$
- Tiempo de reposición: $L = 0.1$
- Nivel de servicio objetivo: $95\% \Rightarrow Z_\alpha = 1.645$

El stock de seguridad se calcula como:

$$SS = Z_\alpha \sigma \sqrt{L} \approx 1.645 \cdot 10 \cdot \sqrt{0.1} \approx 5.20,$$

aunque en la función de costo extendida, este valor se incorpora como un término constante.

5.1.2.5. EOQ clásico (ignora incertidumbre y faltantes)

$$Q_{EOQ} = \sqrt{\frac{2DS}{H}} = \sqrt{\frac{2(1000)(100)}{5}} = 200,$$

y su costo asociado es:

$$Z_{EOQ} = \frac{100}{200} \cdot 100 + \frac{200}{2} \cdot 5 + 1000 \cdot 10 = 500 + 500 + 10000 = 11000.$$

5.1.2.6. EOQ extendido (con stock de seguridad y penalización por faltantes)

La función de costo adopta la forma:

$$Z(Q) = \frac{D}{Q}(S + \beta C) + \left(\frac{Q}{2} + Z_\alpha \sigma \sqrt{L}\right) H + DC.$$

Sustituyendo los valores:

$$\begin{aligned} Z(Q) &= \frac{1000}{Q}(100 + 500) + \left(\frac{Q}{2} + 5.20\right) \cdot 5 + 10000 \\ &= \frac{600000}{Q} + 2.5Q + 10026. \end{aligned}$$

Minimizando esta función determinista:

$$Q^* = \sqrt{\frac{600000}{2.5}} \approx 489.9, \quad Z(Q^*) \approx 12475.44.$$

Si, por error, se usara la política del EOQ clásico ($Q = 200$):

$$Z(200) = \frac{600000}{200} + 2.5 \cdot 200 + 10026 = 3000 + 500 + 10026 = 13526,$$

lo que representa un **8.4 % más de costo** respecto a la política óptima del modelo extendido, además de una **exposición significativamente mayor al riesgo de faltante**.

Este contraste demuestra que los modelos deterministas que ignoran la variabilidad de la demanda y los costos de faltante pueden producir políticas subóptimas o inadmisibles en contextos humanitarios, donde la insatisfacción de la demanda tiene consecuencias operativas y éticas relevantes.

5.1.3. Estructura híbrida: decisiones discretas (ubicación) y continuas (inventario)

Los problemas integrados de localización e inventario pertenecen a la clase de **modelos de optimización mixta**, en los que coexisten variables discretas (por ejemplo, decisiones binarias de apertura de almacenes) y variables continuas (como flujos de asignación o niveles de inventario). Esta combinación induce una **estructura no convexa y no conexa en el conjunto factible**, debido a la naturaleza combinatoria de las variables binarias.

Como argumentan Wolsey (1998), Grossmann (2002) y Daskin (2013), esta pérdida de convexidad impide la aplicación directa de técnicas clásicas de optimización convexa, tales como las condiciones de Karush–Kuhn–Tucker (KKT Sección 4.7) para garantizar optimalidad global, y exige el uso de métodos especializados para problemas mixtos entero-continuos.

5.1.3.1. Formulación matemática del espacio de decisiones

Sea:

- $\mathcal{J} = \{1, \dots, m\}$ el conjunto de ubicaciones posibles para almacenes,
- $\mathcal{J} = \{1, \dots, n\}$ las localidades de demanda,
- $x = (x_1, \dots, x_m)^\top \in \{0, 1\}^m$ las decisiones binarias de apertura,
- $y = (y_{ij}) \in \mathbb{R}_+^{m \times n}$ los flujos desde almacenes a localidades,
- $s = (s_1, \dots, s_m)^\top \in \mathbb{R}_+^m$ el inventario preposicionado.

El espacio factible se define como el conjunto:

$$\mathcal{X} := \left\{ (x, y, s) \left\{ \begin{array}{ll} x_i \in \{0, 1\}, & \forall i \in \mathcal{J}, \\ y_{ij} \geq 0, & \forall i \in \mathcal{J}, \forall j \in \mathcal{J}, \\ s_i \geq 0, & \forall i \in \mathcal{J}, \\ \sum_{i \in \mathcal{J}} y_{ij} \geq d_j, & \forall j \in \mathcal{J}, \\ \sum_{j \in \mathcal{J}} y_{ij} \leq s_i, & \forall i \in \mathcal{J}, \\ s_i \leq Mx_i, & \forall i \in \mathcal{J} \end{array} \right. \right\}. \quad (5.1.2)$$

Aquí, $d_j > 0$ representa la demanda de la localidad j y M es un parámetro de gran magnitud. La tercera restricción impone un vínculo lógico: si $x_i = 0$, entonces $s_i = 0$ y, en consecuencia, $y_{ij} = 0$ para todo $j \in \mathcal{J}$.

Este acoplamiento, combinado con la naturaleza **binaria** de las variables x_i , impide que el conjunto factible \mathcal{X} sea convexo. La no convexidad se debe exclusivamente a la **discreción** de las variables x_i : el promedio de dos puntos factibles con configuraciones distintas de infraestructura no es factible, ya que viola la condición $x_i \in \{0, 1\}$. Esta propiedad, la ausencia de convexidad en el conjunto factible debido a variables enteras, es bien conocida en la teoría de optimización mixta, como puede verse en el ejemplo de selección de proyectos presentado por Bertsekas (1999) donde las decisiones binarias de aceptar o rechazar proyectos generan un conjunto factible no convexo.

Proposición (No convexidad).

El conjunto \mathcal{X} definido en (Ecuación 5.1.2) **no es convexo**.

Demostración. Considérense las configuraciones factibles con $x^{(1)} =$

$(1, 0, \dots, 0)$, $s^{(1)} = (d^\top \mathbf{1}, 0, \dots, 0)^\top$, $y^{(1)}$ asignando toda la demanda al almacén 1, y análogamente $x^{(2)} = (0, 1, 0, \dots, 0)$ con asignación al almacén 2. Ambos puntos pertenecen a \mathcal{X} . Sin embargo, su punto medio tiene $x_i = 1/2$ para $i = 1, 2$, lo cual viola $x_i \in \{0, 1\}$. Por lo tanto, \mathcal{X} no es convexo.

Debido a esta no convexidad, el problema **no pertenece a la clase de optimización convexa**, y en consecuencia **no es posible garantizar optimalidad global mediante condiciones de optimalidad basadas en el análisis convexo clásico**.

5.1.3.2. Función objetivo y acoplamiento estructural

La función objetivo del modelo de ubicación de almacenes e inventario humanitario se define como:

$$Z(x, y, s) = \underbrace{\sum_{i \in \mathcal{J}} f_i(x_i)}_{\text{Costos fijos}} + \underbrace{\sum_{i \in \mathcal{J}} h_i s_i}_{\text{Costos de inventario}} + \underbrace{\sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} c_{ij} y_{ij}}_{\text{Costos de transporte}}. \quad (5.1.3)$$

Aunque la parte en variables continuas es lineal, el término $f_i(x_i)$ introduce una **no linealidad discreta**, ya que $x_i \in \{0, 1\}$. Esta discontinuidad, combinada con el acoplamiento lógico $s_i \leq Mx_i$, da lugar a un **espacio de decisiones no convexo y combinatorio**.

La dificultad computacional del problema se debe a la interacción entre variables discretas y continuas, que genera una estructura no separable y no convexa, característica de los **programas mixtos-enteros no lineales (MINLP)**. Como señalan Grossmann (2002), esta clase de problemas es computacionalmente exigente porque:

- no se puede aplicar directamente la optimización convexa,
- los métodos de enumeración son inviables en problemas de tamaño moderado,
- y la no linealidad impide el uso de técnicas eficientes basadas en planos de corte o relajaciones lineales exactas.

El sistema presenta una jerarquía:

1. x determina la infraestructura disponible,
2. s depende de x ,
3. y depende de s .

Esto produce una dependencia en cascada que elimina la separabilidad global del problema.

Proposición (Convexidad condicional).

Para cualquier $\bar{x} \in \{0, 1\}^m$, el subproblema

$$\min_{(y,s) \in \mathcal{X}(\bar{x})} Z(\bar{x}, y, s)$$

es un problema de programación lineal y, por tanto, convexo.

Demostración. Al fijar \bar{x} , los términos $f_i \bar{x}_i$ se convierten en constantes. La función objetivo resultante es lineal en (y, s) , y todas las restricciones que definen $\mathcal{X}(\bar{x})$ son lineales en estas variables. Por definición, esto constituye un problema de programación lineal, cuyo conjunto factible es un poliedro convexo. ■

Esta propiedad, consistente en *la convexidad del subproblema continuo para una configuración de apertura fija*, es fundamental para la aplicación de métodos de descomposición, tales como el algoritmo de Benders. En dichos enfoques, el problema original se divide en:

- **problema maestro**, que maneja las variables enteras (x) y genera configuraciones candidatas, y
- un **subproblema**, que, para una configuración dada, resuelve el problema continuo (aquí, lineal) y devuelve información (por ejemplo, cortes de Benders) para refinar el problema maestro, siguiendo la metodología de descomposición descrita en Wolsey (1998).

5.1.3.3. Ejemplo ilustrativo: interacción discreto–continuo

Considérese:

- $m = 2$ almacenes,
- $n = 2$ demandas con $d_1 = d_2 = 50$,
- Costos fijos $f_1 = 300$, $f_2 = 400$,
- Costos de inventario $h_1 = h_2 = 3$,
- Costos de transporte:
 $c_{11} = 8$, $c_{12} = 15$,
 $c_{21} = 12$, $c_{22} = 7$,
- $M = 100$.

Considérese:

- $m = 2$ almacenes,
- $n = 2$ demandas con $d_1 = d_2 = 50$,
- Costos fijos $f_1 = 300$, $f_2 = 400$,
- Costos de inventario $h_1 = h_2 = 3$,
- Costos de transporte:
 - $c_{11} = 8$, $c_{12} = 15$,
 - $c_{21} = 12$, $c_{22} = 7$,
- $M = 100$.

Para cada configuración de apertura $x = (x_1, x_2) \in \{0, 1\}^2$, se asignan los flujos y_{ij} de manera que se satisfaga la demanda minimizando el costo de transporte:

(i) Abrir solo almacén 1: $x = (1, 0)$.

Asignación óptima: $y_{11} = 50$, $y_{12} = 50$, $y_{21} = y_{22} = 0$.

Costo:

$$Z = f_1x_1 + f_2x_2 + h_1(y_{11} + y_{12}) + h_2(y_{21} + y_{22}) + c_{11}y_{11} + c_{12}y_{12} + c_{21}y_{21} + c_{22}y_{22}$$

$$Z = 300(1) + 400(0) + 3(50 + 50) + 3(0) + 8(50) + 15(50) + 12(0) + 7(0) = 1750.$$

(ii) Abrir solo almacén 2: $x = (0, 1)$.

Asignación óptima: $y_{21} = 50$, $y_{22} = 50$, $y_{11} = y_{12} = 0$.

Costo:

$$Z = 300(0) + 400(1) + 3(0) + 3(50 + 50) + 8(0) + 15(0) + 12(50) + 7(50) = 1650.$$

(iii) Abrir ambos: $x = (1, 1)$.

Asignación eficiente por costo mínimo: para $j = 1$, $c_{11} = 8 < c_{21} = 12 \Rightarrow y_{11} = 50$; para $j = 2$, $c_{22} = 7 < c_{12} = 15 \Rightarrow y_{22} = 50$. Luego $y_{12} = y_{21} = 0$.

Costo:

$$Z = 300(1) + 400(1) + 3(50) + 3(50) + 8(50) + 15(0) + 12(0) + 7(50) = 1750.$$

La mejor configuración es abrir únicamente el almacén 2. Este resultado ilustra una propiedad fundamental de los problemas que combinan variables discretas y continuas: la solución óptima puede cambiar de manera discontinua ante perturbaciones arbitrariamente pequeñas en los parámetros del problema. Por ejemplo, si f_2 disminuye a 350, la solución óptima se mantiene; pero si f_1 baja a 250, podría resultar óptimo abrir el almacén 1.

A este fenómeno, la dependencia discontinua de la solución óptima respecto a los datos del problema, causada por la naturaleza combinatoria de las decisiones binarias, nos

referiremos como sensibilidad discreta. Esta característica distingue claramente a los modelos mixtos entero-continuos de los problemas puramente continuos, donde bajo condiciones de regularidad la solución óptima varía de forma suave como se puede ver en Daskin (2013).

El conjunto factible de este tipo de problemas, definido por la combinación de variables binarias y continuas junto con restricciones de activación lógica, **no es convexo** (como se demostró en la Sección 5.1.3.1). Esta no convexidad, junto con la interacción entre decisiones estratégicas y operativas, exige el uso de técnicas propias de la programación entera mixta y de la optimización combinatoria. Estas bases teóricas sustentan el modelo propuesto en las secciones siguientes, donde se integra esta estructura con un componente de inventario no lineal bajo incertidumbre parametrizada.

5.2. Marco general de modelado: variables, parámetros y espacio de decisión

5.2.1. Conjuntos de índices: almacenes potenciales \mathcal{I} y localidades de demanda \mathcal{J}

En la formulación matemática del problema integrado de localización e inventario, los conjuntos de índices constituyen la base sobre la cual se definen las variables de decisión, los parámetros y las restricciones. Como se estableció en la Sección 5.1, denotamos por

- \mathcal{I} : el conjunto de **almacenes candidatos** para preposicionamiento, y
- \mathcal{J} : el conjunto de **localidades geográficas afectadas**.

Estos conjuntos son finitos, disjuntos, es decir, $\mathcal{I} \cap \mathcal{J} = \emptyset$, y sus elementos desempeñan roles distintos en el modelo: los índices en \mathcal{I} corresponden a *fuentes* de suministro, mientras que los de \mathcal{J} representan *destinos* de demanda.

La cardinalidad de \mathcal{I} , denotada $m = |\mathcal{I}|$, determina la dimensión del espacio de decisiones estratégicas. Dado que el costo de apertura de cada almacén se modela mediante una variable binaria $x_i \in \{0, 1\}$, el número total de configuraciones posibles de infraestructura es 2^m . Esta **explosión combinatoria** implica que, incluso para tamaños modestos de m (por ejemplo, $m = 30$), el espacio discreto contiene más de 10^9 soluciones factibles. En consecuencia, métodos de enumeración exhaustiva son inviables, y se requieren algoritmos que exploten la **estructura matemática del problema**, es decir, su formulación como un problema de optimización con variables mixtas, función objetivo diferenciable y restricciones acopladas, para identificar soluciones óptimas o aproximadas en tiempo razonable.

La interacción entre \mathcal{I} y \mathcal{J} se modela mediante variables de flujo $y_{ij} \geq 0$, que cuantifican la asignación de suministros desde el almacén i a la localidad j . Esta relación define una

matriz de conectividad cuya densidad depende de factores operativos como la distancia, accesibilidad y capacidades logísticas, y que condiciona tanto la factibilidad como la complejidad del subproblema continuo asociado a una configuración fija de almacenes.

5.2.1.1. Ejemplo ilustrativo

Supóngase un escenario de planificación para respuesta a inundaciones dividido en cinco distritos. Se identifican tres ubicaciones potenciales para almacenes:

- $i = 1$: Aeropuerto internacional,
- $i = 2$: Base militar,
- $i = 3$: Centro de distribución de una ONG.

Asimismo, se consideran cuatro zonas vulnerables:

- $j = 1$: Zona ribereña,
- $j = 2$: Área rural,
- $j = 3$: Barrio urbano denso,
- $j = 4$: Comunidad indígena aislada.

Los conjuntos índice resultan:

$$\mathcal{I} = \{1, 2, 3\}, \quad \mathcal{J} = \{1, 2, 3, 4\}.$$

Sobre estos conjuntos se definirán:

- variables binarias $x_i \in \{0, 1\}$ para cada $i \in \mathcal{I}$,
- variables continuas de flujo $y_{ij} \geq 0$ para $(i, j) \in \mathcal{I} \times \mathcal{J}$,
- parámetros exógenos como c_{ij} (costos), f_i (costos fijos) y d_j (demanda).

Nota metodológica.

En modelado riguroso, los conjuntos índice deben definirse antes que las variables, dado que estas heredan su dominio directamente de aquellos.

5.2.2. Variables de decisión: binarias (x_i) y continuas (y_{ij})

Una vez establecida la estructura combinatoria del problema mediante los conjuntos índice \mathcal{I} y \mathcal{J} , el siguiente paso en la formulación rigurosa de un modelo de optimización es la definición precisa del **espacio de decisiones**, mediante la especificación de las **variables de decisión**. En el **problema integrado de localización e inventario para logística humanitaria**, introducido en la Sección 5.1.1, este espacio de decisiones de este problema es híbrido, ya que combina variables binarias y continuas, compuesto por dos tipos cualitativamente distintos de variables:

- Variables binarias, que modelan elecciones discretas de infraestructura,
- Variables continuas, que representan flujos físicos de recursos.

Esta dualidad refleja la **estructura mixta del problema**, que combina dos tipos de decisiones con roles y horizontes operativos distintos:

- **Decisiones estratégicas**, como el costo apertura de almacenes, que son irreversibles una vez implementadas y se modelan mediante variables binarias;
- **Decisiones tácticas**, como la asignación de suministros o el preposicionamiento de inventario, que son ajustables en el corto plazo y se representan mediante variables continuas.

Esta interacción entre decisiones de largo y corto plazo es una característica esencial de los modelos integrados en logística humanitaria.

5.2.2.1. Definición (Variables binarias de almacén).

Para cada $i \in \mathcal{I}$, se define la variable binaria

$$x_i \in \{0, 1\},$$

donde

$$x_i = \begin{cases} 1, & \text{si se decide abrir un almacén en la ubicación } i, \\ 0, & \text{en caso contrario.} \end{cases}$$

El vector $x = (x_i)_{i \in \mathcal{I}} \in \{0, 1\}^m$ codifica una **configuración de apertura de almacenes** en la red de respuesta humanitaria. Este vector pertenece al conjunto discreto $\{0, 1\}^m$, que es no convexo, finito y tiene cardinalidad 2^m .

5.2.2.2. Definición (Variables continuas de flujo).

Para cada par $(i, j) \in \mathcal{I} \times \mathcal{J}$, se define la variable continua

$$y_{ij} \in \mathbb{R}_+,$$

que representa la cantidad de unidades asignadas desde la instalación i a la localidad de demanda j .

El tensor $y = (y_{ij})_{(i,j) \in \mathcal{I} \times \mathcal{J}} \in \mathbb{R}_+^{m \times n}$ describe la **estructura de asignación de suministros desde almacenes candidatos a localidades de demanda**, y pertenece a un espacio vectorial convexo, cerrado (ver Sección 4.2) y de dimensión finita.

Observación (Naturaleza híbrida del espacio de decisiones).

El espacio total de decisiones es el producto cartesiano

$$\mathcal{D} = \{0, 1\}^m \times \mathbb{R}_+^{m \times n},$$

que es tipo no convexo, no conexo y discontinuo en la dirección discreta. Esta estructura impide la aplicación directa de herramientas del análisis convexo clásico y exige métodos de optimización entera mixta.

Las variables x_i y y_{ij} no son independientes. Su acoplamiento se establece mediante **restricciones de factibilidad**, que garantizan que no se asignen flujos desde almacenes no abiertos. Para todo $i \in \mathcal{I}$ y $j \in \mathcal{J}$:

$$y_{ij} \leq Mx_i, \tag{5.2.1}$$

donde $M > 0$ es una constante suficientemente grande (por ejemplo, $M = \sum_{j \in \mathcal{J}} d_j$). La restricción Ecuación 5.2.1 es una formulación *big-M* estándar (ver Sección 4.5) que modela la implicación lógica:

$$x_i = 0 \Rightarrow y_{ij} = 0 \quad \forall j \in \mathcal{J}.$$

5.2.3. Ejemplo ilustrativo: interpretación y acoplamiento de variables

Sea:

$$\mathcal{I} = \{1, 2, 3\}, \quad \mathcal{J} = \{1, 2, 3, 4\}.$$

Supóngase que se decide abrir almacenes en $i = 1$ y $i = 3$, pero no en $i = 2$:

$$x = (1, 0, 1).$$

Las restricciones big-M implican:

- $y_{2j} = 0$ para todo j ,
- $y_{1j}, y_{3j} \geq 0$ como flujos posibles.

Una asignación factible es:

$$\begin{aligned} y_{11} &= 30, & y_{12} &= 20, & y_{13} &= 0, & y_{14} &= 0, \\ y_{21} &= y_{22} = y_{23} = y_{24} &= 0, \\ y_{31} &= 10, & y_{32} &= 0, & y_{33} &= 40, & y_{34} &= 25. \end{aligned}$$

Si la demanda es $d = (40, 20, 40, 25)$, esta asignación es factible y completa.

Este ejemplo muestra que las variables binarias **activan o desactivan** subespacios del espacio continuo de flujos, generando una partición de \mathcal{D} en 2^m subespacios convexos.

Las variables x_i y y_{ij} constituyen los grados de libertad fundamentales del modelo. Su naturaleza híbrida, discreta y continua, refleja la dualidad entre planeación estratégica (apertura de almacenes) y operación táctica (asignación de recursos). El acoplamiento lógico mediante restricciones big-M introduce no convexidad, justificando el uso de técnicas de optimización entera mixta. Las siguientes secciones integrarán estas variables en la función objetivo y en las **restricciones de balance de flujo**, es decir, en las condiciones que garantizan:

- (I) que la demanda en cada localidad $j \in \mathcal{J}$ sea satisfecha y
- (II) que el flujo total enviado desde un almacén $i \in \mathcal{J}$ no exceda su inventario disponible.

5.2.4. Parámetros del sistema: costos fijos, unitarios, demanda y capacidades

En la formulación matemática de un problema de optimización, los **parámetros** constituyen los datos exógenos que estructuran el espacio factible y la función objetivo como señala Bertsekas (1999). Su correcta especificación es requisito previo a la definición de variables y restricciones; por tanto, deben presentarse con su dominio y unidades explícitas (práctica recomendada por Lewis (2003)). En el problema integrado de localización e inventario consideramos, de modo explícito, las cuatro familias de parámetros (costos fijos de apertura, costos unitarios de transporte, demanda esperada y capacidad de almacenamiento) siguientes.

5.2.4.1. Costos fijos de apertura

Para cada $i \in \mathcal{I}$ definimos

$$f_i \in \mathbb{R}_+$$

como el **costo fijo de apertura** del almacén en la localización i . En la formulación base del modelo, presentada en la Ecuación 5.1.3, este término contribuye a la función objetivo mediante el producto lineal $f_i x_i$, donde $x_i \in \{0, 1\}$ es la variable binaria. Se exige

$$f_i \geq 0, \quad \forall i \in \mathcal{I}.$$

Observación práctica: cuando el costo de apertura implica economías o penalizaciones no triviales (por ejemplo, coste creciente con la capacidad instalada o costes discretos adicionales por remoción), f_i puede modelarse como función $f_i(x_i, s_i)$; sin embargo, para la formulación base se adopta linealidad en x_i y se externalizan las no linealidades como puede verse en el trabajo de (Shapiro, Dentcheva, y Ruszczyński (2021)).

5.2.4.2. Costos unitarios de transporte

Para cada par $(i, j) \in \mathcal{I} \times \mathcal{J}$ definimos

$$c_{ij} \in \mathbb{R}_+,$$

el **costo unitario de transportar** una unidad desde la instalación i hasta la localidad de demanda j . Requerimos

$$c_{ij} \geq 0, \quad \forall i, j,$$

y denotamos la matriz $C := [c_{ij}]_{i,j} \in \mathbb{R}_+^{m \times n}$. En contextos humanitarios c_{ij} incorpora distancia, tiempo de acceso y penalizaciones por rutas inseguras; su estimación suele derivar de información geoespacial y modelos de capacidad vial, los cuales constituyen insumos fundamentales en los modelos de localización descritos por Daskin (2013).

5.2.4.3. Demanda esperada

Para cada $j \in \mathcal{J}$, sea

$$d_j \in \mathbb{R}_+ - \{0\}$$

la **demanda esperada** en la localidad j durante el horizonte considerado. Adicionalmente definimos el vector $d = (d_j)_{j \in \mathcal{J}}$. En la formulación determinista de referencia usamos d_j como estimador puntual (p. ej. esperanza), con la advertencia explícita de que su uso directo puede producir soluciones frágiles en colas de distribución como lo argumenta Snyder (2006).

Por notación:

$$d_j > 0, \quad \forall j, \quad d \in (\mathbb{R}_+ - 0)^n.$$

5.2.4.4. Capacidades de almacenamiento

Para cada $i \in \mathcal{I}$ definimos la **capacidad máxima** $s_i^{\text{máx}}$ mediante

$$s_i^{\text{máx}} \in \mathbb{R}_+ \cup \{+\infty\}.$$

La restricción de capacidad se escribe, asociada a la lógica de activación x_i :

$$\sum_{j \in \mathcal{J}} y_{ij} \leq s_i^{\text{máx}} x_i, \quad \forall i \in \mathcal{I}. \quad (5.2.2)$$

En ausencia de un tope operativo se toma $s_i^{\text{máx}} = +\infty$ (modelo no capacitado); en la práctica real $s_i^{\text{máx}}$ suele venir de disponibilidad física, limitaciones regulatorias o logísticas.

Observación sobre Big-M. Si $s_i^{\text{máx}}$ es finito, la restricción Ecuación 5.2.2 hace redundante la condición de Big-M genérica para el flujo. Dado que $y_{ij} \geq 0$, al tener $x_i = 0$ la desigualdad fuerza necesariamente a $\sum_{j \in \mathcal{J}} y_{ij} = 0$, lo que implica $y_{ij} = 0$ para todo j . En consecuencia, no es necesario introducir una restricción adicional del tipo $y_{ij} \leq Mx_i$; utilizar la capacidad real $s_i^{\text{máx}}$ proporciona una cota superior más ajustada (*tight formulation*), fortalece la relajación lineal y evita la degradación numérica asociada a constantes M arbitrariamente grandes.

5.2.4.5. Propiedades y observaciones formales

1. **Dominios y no ambigüedad:** Todos los parámetros deben venir acompañados de su dominio y unidades, lo cual evita ambigüedades en pruebas de existencia, estabilidad numérica y escalado de variables. La importancia de estas precisiones se refleja en la literatura sobre optimización numérica, como señalan Lewis y Overton (2013), donde la estabilidad depende críticamente de la correcta definición de las variables.
2. **Dependencia discontinua de la solución óptima respecto a los parámetros:** Pequeñas variaciones en los parámetros del modelo, como los costos fijos f_i , los costos de transporte c_{ij} , la demanda d_j o la capacidad máxima $s_i^{\text{máx}}$, pueden inducir cambios abruptos en la configuración óptima de apertura de almacenes x^* .

Dado que el modelo se basa en estimaciones de parámetros sujetos a incertidumbre (por ejemplo, la demanda esperada $d_j = \mathbb{E}[D_j]$), es esencial evaluar la solución obtenida bajo múltiples escenarios plausibles y analizar su sensibilidad a perturbaciones en los datos. Este tipo de evaluación permite identificar soluciones no solo óptimas, sino también **estables y robustas ante errores en las estimaciones**, un aspecto crítico en contextos humanitarios donde la calidad de la información es limitada (Snyder (2006); Shapiro, Dentcheva, y Ruszczyński (2021)).

3. **Elección de M en big-M:** Si se usa la formulación $y_{ij} \leq Mx_i$ para acoplar variables (versus la formulación con $s_i^{\text{máx}}$), se recomienda elegir M igual a una cota realista (ej. $\sum_j d_j$) para evitar degradación numérica; empero, la formulación (Ecuación 5.2.2) es preferible por su interpretación física y por estrechar el dominio factible.

5.2.4.6. Ejemplo numérico (verificación de factibilidad y cálculo de costo)

Tomemos la instancia didáctica ya utilizada en la Sección 5.2.3:

- $\mathcal{I} = \{1, 2, 3\}$, $\mathcal{J} = \{1, 2, 3, 4\}$.
- Costos fijos (miles unidades monetarias): $f = (300, 400, 250)$.
- Matriz $C = [c_{ij}]$ (unidades monetarias/unidad):

$$C = \begin{bmatrix} 8 & 20 & 15 & 25 \\ 12 & 18 & 10 & 30 \\ 20 & 10 & 12 & 18 \end{bmatrix}.$$

- Demanda: $d = (40, 20, 40, 25)$ (unidades).
- Capacidades: $s^{\text{máx}} = (100, 80, 90)$.
- Consideramos la decisión $x = (1, 0, 1)$ y la asignación

$$\begin{aligned} y_{11} &= 30, & y_{12} &= 20, & y_{13} &= 0, & y_{14} &= 0, \\ y_{21} &= y_{22} = y_{23} = y_{24} &= 0, \\ y_{31} &= 10, & y_{32} &= 0, & y_{33} &= 40, & y_{34} &= 25. \end{aligned}$$

Paso 1 — Comprobación de que se cumple la restricción de capacidades (Ecuación 5.2.2):

- Para $i = 1$:

$$\sum_j y_{1j} = 30 + 20 + 0 + 0 = 50 \leq s_1^{\text{máx}} x_1 = 100 \cdot 1 = 100 \quad \checkmark.$$

- Para $i = 2$:

$$\sum_j y_{2j} = 0 \leq 80 \cdot 0 = 0 \quad \checkmark.$$

- Para $i = 3$:

$$\sum_j y_{3j} = 10 + 0 + 40 + 25 = 75 \leq 90 \cdot 1 = 90 \quad \checkmark.$$

Paso 2 — Comprobación de que se cumple la cobertura de la demanda:

- $j = 1$: $y_{11} + y_{31} = 30 + 10 = 40 = d_1$,
- $j = 2$: $y_{12} = 20 = d_2$,

- $j = 3$: $y_{33} = 40 = d_3$,
- $j = 4$: $y_{34} = 25 = d_4$.

Cobertura completa lo que implica que es factible.

Paso 3 — Cálculo del costo total

La función de costo (determinista, sin mantenimiento de inventario explícito) se toma como

$$Z(x, y) = \sum_{i \in \mathcal{I}} f_i x_i + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} c_{ij} y_{ij}.$$

Sustituyendo:

- Término fijo: $f_1 x_1 + f_3 x_3 = 300 + 250 = 550$ (miles unidades monetarias).
- Transporte (unidades monetarias): calcular cada contribución

$$\begin{aligned} 8 \cdot 30 &= 240, & 20 \cdot 20 &= 400, & 20 \cdot 10 &= 200, \\ 12 \cdot 40 &= 480, & 18 \cdot 25 &= 450. \end{aligned}$$

Suma transporte = $240 + 400 + 200 + 480 + 450 = 1770$ unidades monetarias.

Total (homogeneizando unidades a miles unidades monetarias):

$$Z = 550 \text{ (miles)} + 1.770 \text{ (miles)} = 2.320 \text{ miles unidades monetarias.}$$

Interpretación: la solución es factible y el costo cuantificado; variaciones en d , c_{ij} o f_i deben reevaluarse mediante análisis de sensibilidad y, en contextos reales, mediante formulaciones **estocásticas** o **robustas**, como se discute en Snyder (2006) para problemas de localización y en Shapiro, Dentcheva, y Ruszczyński (2021) para la teoría de optimización estocástica.

- En el enfoque **estocástico**, los parámetros inciertos (como la demanda) se modelan como variables aleatorias \tilde{d}_j con distribución de probabilidad estimada a partir de datos históricos, y se optimiza el costo esperado (o una medida de riesgo, como el CVaR).
- En el enfoque **robusto**, se define un *conjunto de incertidumbre* \mathcal{U} por ejemplo, un intervalo $d_j \in [\underline{d}_j, \bar{d}_j]$ o un conjunto de presupuesto $\sum_j \frac{|d_j - \bar{d}_j|}{\bar{d}_j} \leq \Gamma$ que contiene todos los valores plausibles de los parámetros, y se busca una solución que tenga buen desempeño incluso en el peor caso dentro de \mathcal{U} .

La especificación explícita de f_i , c_{ij} , d_j y s_i^{\max} permite no solo la construcción del problema determinista base, sino también su extensión a:

- **Modelos estocásticos**, donde d_j se reemplaza por una variable aleatoria \tilde{d}_j ;
- **Modelos robustos**, donde se define un conjunto de incertidumbre \mathcal{U} para d_j y c_{ij} , cuyos límites se basan en datos históricos, escenarios extremos o criterios de política (por ejemplo, “la demanda no excederá el doble de su valor promedio”).

La práctica recomendada por Lewis (2003) y Shapiro, Dentcheva, y Ruszczyński (2021), directamente relacionada con la robustez en optimización, consiste en:

- (I) documentar fuentes y unidades de cada parámetro;
- (II) fijar cotas realistas para $s_i^{\text{máx}}$ o M (esenciales para definir \mathcal{U} y evitar degradación numérica); y
- (III) efectuar pruebas de sensibilidad sistemáticas antes de adoptar políticamente cualquier solución.

5.2.5. Definición formal del espacio factible \mathcal{X}

En optimización matemática, la **estructura del conjunto factible** determina en gran medida la naturaleza del problema, su complejidad computacional y la aplicabilidad de métodos de solución. Para el problema integrado de localización e inventario, el espacio factible es un **subconjunto híbrido** puesto que es producto cartesiano entre un espacio discreto y uno continuo. A continuación se define formalmente este conjunto y se analizan sus propiedades estructurales clave.

5.2.5.1. Definición (Espacio factible).

Dado el conjunto de almacenes potenciales \mathcal{I} , el conjunto de localidades de demanda \mathcal{J} , y los parámetros exógenos $d_j > 0$, $s_i^{\text{máx}} \in \mathbb{R}_+ \cup \{+\infty\}$, y una constante $M > 0$ suficientemente grande (por ejemplo, $M = \sum_{j \in \mathcal{J}} d_j$), el **espacio factible** del modelo se define como:

$$\mathcal{X} := \left\{ (x, y) \in \{0, 1\}^m \times \mathbb{R}_+^{m \times n} \left| \begin{array}{ll} \text{(i)} & \sum_{i \in \mathcal{I}} y_{ij} \geq d_j, \quad \forall j \in \mathcal{J}, \\ \text{(ii)} & \sum_{j \in \mathcal{J}} y_{ij} \leq s_i^{\text{máx}} x_i, \quad \forall i \in \mathcal{I}, \\ \text{(iii)} & y_{ij} \leq M x_i, \quad \forall (i, j) \in \mathcal{I} \times \mathcal{J} \end{array} \right. \right\}.$$

Las tres familias de restricciones tienen interpretaciones operativas precisas:

- **(i) Cobertura de demanda:** cada localidad j debe recibir al menos su demanda esperada d_j .

- **(ii) Capacidad condicionada:** el flujo total saliente de la instalación i no puede exceder su capacidad máxima si está abierta ($x_i = 1$).
- **(iii) Activación lógica fuerte:** garantiza $y_{ij} = 0$ cuando $x_i = 0$, fortaleciendo las relajaciones lineales.

Observación {#eq-obs}

Si $s_i^{\max} = +\infty$, la restricción (ii) se sustituye simplemente por $y_{ij} \leq Mx_i$. La estructura analítica permanece intacta.

5.2.6. Propiedades estructurales del espacio \mathcal{X}

5.2.6.1. Proposición (No convexidad y no conexidad).

El conjunto \mathcal{X} es **no convexo** y, en general, **no conexo** en la topología estándar de $\mathbb{R}^m \times \mathbb{R}^{m \times n}$.

Demostración.

Considérese $m \geq 2$. Sean:

- $(x^{(1)}, y^{(1)})$ con $x^{(1)} = e_1$ y $y_{1j}^{(1)} = d_j$,
- $(x^{(2)}, y^{(2)})$ con $x^{(2)} = e_2$ y $y_{2j}^{(2)} = d_j$.

Ambos puntos son factibles. Su media:

$$(\bar{x}, \bar{y}) = \frac{1}{2}(x^{(1)}, y^{(1)}) + \frac{1}{2}(x^{(2)}, y^{(2)})$$

tiene $\bar{x}_1 = \bar{x}_2 = 1/2 \notin \{0, 1\}$, por lo que no pertenece a \mathcal{X} .

Además, no existe un camino continuo dentro de \mathcal{X} que conecte ambos puntos, ya que cualquier camino debe atravesar valores fraccionarios de x .

5.2.6.2. Proposición (Descomposición en secciones convexas).

$$\mathcal{X} = \bigcup_{\bar{x} \in \{0,1\}^m} (\{\bar{x}\} \times \mathcal{Y}(\bar{x})),$$

donde

$$\mathcal{Y}(\bar{x}) := \left\{ y \in \mathbb{R}_+^{m \times n} \left| \begin{array}{l} \sum_i y_{ij} \geq d_j, \quad \forall j, \\ \sum_j y_{ij} \leq s_i^{\max} \bar{x}_i, \quad \forall i \end{array} \right. \right\}$$

es un **poliedro convexo** (ver Sección 4.2) posiblemente vacío.

Demostración.

Para \bar{x} fijo, las restricciones que definen $\mathcal{Y}(\bar{x})$ son lineales en y , por lo que el conjunto es convexo y cerrado.

La unión es disjunta porque cada sección responde a un valor distinto de x .

Esta descomposición motiva algoritmos híbridos: resolver un subproblema convexo para cada \bar{x} y usar técnicas combinatorias para explorar $\{0, 1\}^m$.

5.2.6.3. Ejemplo ilustrativo: construcción explícita de \mathcal{X}

Sea:

- $\mathcal{J} = \{1, 2\}$
- $\mathcal{J} = \{1\}$

- $d_1 = 10$

- $s_1^{\text{máx}} = s_2^{\text{máx}} = 15$

- $M = 10$

Entonces:

1. $x = (0, 0)$:
 $y_{11} = y_{21} = 0$, pero exige $y_{11} + y_{21} \geq 10$ **infeasible**.
2. $x = (1, 0)$:
 $y_{21} = 0, y_{11} \in [10, 15]$.
3. $x = (0, 1)$:
 $y_{11} = 0, y_{21} \in [10, 15]$.
4. $x = (1, 1)$:
 $y_{11} + y_{21} \geq 10$, con $y_{ij} \leq 15$.
Es un poliedro factible en dimensión 2.

Este ejemplo muestra que \mathcal{X} está formado por componentes disjuntas: dos segmentos y un poliedro 2D.

El conjunto \mathcal{X} posee una estructura combinatoria compleja: es no convexo y no conexo, pero admite una descomposición en secciones convexas. Su correcta caracterización es esencial para asegurar consistencia del modelo, existencia de solución y diseño de algoritmos eficientes. En las siguientes secciones, este espacio actuará como dominio de la función de costo total.

5.2.7. Función objetivo general en problemas de localización–inventario

La **función objetivo** de un modelo de optimización encapsula el criterio de desempeño que se busca minimizar o maximizar. En problemas de localización–inventario para logística humanitaria, este criterio suele ser una medida **económica agregada** que refleja los costos totales del problema, ponderando compromisos entre infraestructura fija, operación logística y servicio a la demanda. A continuación, se define formalmente la función objetivo general, se descompone en sus componentes estructurales y se analizan sus propiedades matemáticas.

5.2.7.1. Definición (Función objetivo total)

Dado el espacio factible \mathcal{X} definido previamente y los parámetros $f_i \geq 0$, $c_{ij} \geq 0$, la **función objetivo total** $Z : \mathcal{X} \rightarrow \mathbb{R}$ se define como:

$$Z(x, y) := \underbrace{\sum_{i \in \mathcal{J}} f_i x_i}_{\text{(A) Costo fijo de apertura}} + \underbrace{\sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} c_{ij} y_{ij}}_{\text{(B) Costo de transporte}} + \underbrace{\sum_{i \in \mathcal{J}} h_i \left(\sum_{j \in \mathcal{J}} y_{ij} \right)}_{\text{(C) Costo de inventario}}. \quad (5.2.3)$$

Aquí, $h_i \geq 0$ es el **costo unitario de mantenimiento de inventario** en la instalación i . El término modela el costo asociado al volumen total preposicionado (equivalente al total despachado en el modelo determinista base).

La función Z es **afín en y** para x fijo, y **lineal en x** cuando y se trata como variable libre. Sin embargo, la complejidad estructural del problema no proviene de una no linealidad intrínseca en Z , sino de la **no convexidad y desconexidad del dominio factible** \mathcal{X} , inducida por las restricciones binarias y de acoplamiento lógico. Esta geometría del espacio de decisiones es la que impide el análisis convexo clásico y exige técnicas de optimización entera mixta.

Observación (Interpretación económica de los términos de (Ecuación 5.2.3))

- (A): decisiones estratégicas de largo plazo.
- (B): costos operativos de corto plazo.
- (C): costo de oportunidad del inventario preposicionado.

En logística humanitaria, f_i , c_{ij} y h_i pueden incorporar factores no monetarios (tiempo, riesgo, deterioro, ética), siempre que estén normalizados en una métrica común.

5.2.8. Propiedades analíticas de Z

5.2.8.1. Proposición (Continuidad y acotación inferior)

La función $Z : \mathcal{X} \rightarrow \mathbb{R}$ es **continua** en la topología relativa de \mathcal{X} y está **acotada inferiormente** por cero.

Demostración.

Como $\mathcal{X} \subset \{0, 1\}^m \times \mathbb{R}_+^{m \times n}$, su topología relativa es la unión disjunta de secciones del tipo $\{\bar{x}\} \times \mathcal{Y}(\bar{x})$. En cada sección, la función $y \mapsto Z(\bar{x}, y)$ es afín, y por tanto continua. Dado que todos los coeficientes f_i, c_{ij}, h_i son no negativos, se tiene $Z(x, y) \geq 0$ para todo punto factible.

5.2.8.2. Proposición (Coercividad condicional y existencia de solución)

Si $\mathcal{X} \neq \emptyset$, entonces el problema

$$\min_{(x,y) \in \mathcal{X}} Z(x, y)$$

admite solución óptima.

Demostración.

Como el conjunto discreto $\{0, 1\}^m$ es finito, la minimización puede escribirse como:

$$\min_{\bar{x} \in \{0, 1\}^m} \left\{ \sum_i f_i \bar{x}_i + \min_{y \in \mathcal{Y}(\bar{x})} \left(\sum_{i,j} (c_{ij} + h_i) y_{ij} \right) \right\}.$$

Para cada \bar{x} tal que $\mathcal{Y}(\bar{x}) \neq \emptyset$, el subproblema interno es un **problema de programación lineal** definido sobre un poliedro no vacío y acotado (ya que existen cotas inferiores por demanda y los coeficientes son no negativos). Por el teorema fundamental de la programación lineal, el subproblema posee solución. Como el número de posibles \bar{x} es finito, el mínimo global existe.

5.2.8.3. Ejemplo ilustrativo: evaluación explícita de $Z(x, y)$

Considérese el escenario:

- $\mathcal{J} = \{1, 2\}$,
- $\mathcal{J} = \{1\}$,

- $d_1 = 10$,
- $f_1 = 300, f_2 = 400$,
- $c_{11} = 10, c_{21} = 12$,
- $h_1 = 3, h_2 = 2$.

Caso 1: $x = (1, 0), y_{11} = 10$ (cobertura exacta de la demanda)

$$\begin{aligned} Z(x, y) &= f_1 x_1 + f_2 x_2 + (c_{11} + h_1) y_{11} + (c_{21} + h_2) y_{21} \\ &= 300 + 0 + (10 + 3) \cdot 10 + (12 + 2) \cdot 0 \\ &= 300 + 130 = 430. \end{aligned}$$

(Nota: si se optara por sobreabastecer con $y_{11} = 12$, el costo sería $300 + 13 \cdot 12 = 456$. Esto confirma que la estructura de costos penaliza naturalmente el exceso, ya que cada unidad adicional incrementa linealmente los términos de transporte y mantenimiento.)

Caso 2: $x = (0, 1), y_{21} = 10$

$$Z = 0 + 400 + (12 + 2) \cdot 10 = 540.$$

Caso 3: $x = (1, 1), y_{11} = 6, y_{21} = 4$

$$Z = 300 + 400 + 13 \cdot 6 + 14 \cdot 4 = 834.$$

El valor mínimo se obtiene abriendo únicamente la instalación 1, con envío exacto $y_{11} = d_1 = 10$. Este resultado ilustra que, aunque el modelo permite enviar más de la demanda, la minimización del costo total desincentiva el sobreabastecimiento, ya que cada unidad adicional incrementa directamente los gastos de transporte e inventario.

La función objetivo $Z(x, y)$ formaliza el análisis de compromisos entre inversión en infraestructura, costos operativos e inventario preposicionado. Su estructura afín por partes y la no convexidad del dominio \mathcal{X} definen un Mixed Integer Linear Programming (MILP) clásico, aunque constituyen la base teórica necesaria para extensiones no lineales, robustas o estocásticas. La existencia garantizada de al menos una solución óptima justifica el enfoque algorítmico planteado en el Capítulo siguiente.

5.3. Modelo clásico de lote económico (EOQ): bases y limitaciones

El modelo de lote económico (*Economic Order Quantity*, EOQ) constituye uno de los pilares fundamentales de la teoría clásica de control de inventarios. Introducido originalmente por Harris (1913) y posteriormente popularizado por Wilson (1934), el EOQ proporciona una formulación analítica para determinar el tamaño óptimo de pedido que minimiza los costos totales de inventario bajo un conjunto restrictivo de supuestos deterministas. A pesar de su simplicidad, el modelo ha servido como punto de partida para numerosas extensiones en contextos más realistas. En esta sección, se presentan rigurosamente los supuestos deterministas subyacentes al modelo EOQ, los cuales definen su dominio de aplicabilidad y, simultáneamente, evidencian sus limitaciones en entornos caracterizados por incertidumbre, como los propios de la logística humanitaria.

5.3.1. Supuestos deterministas del EOQ

El modelo EOQ se fundamenta en los siguientes supuestos deterministas, los cuales garantizan la existencia de una solución analítica cerrada y convexa:

1. **Demanda constante y conocida:** La tasa de demanda $d > 0$ (unidades por unidad de tiempo) es fija, continua y perfectamente previsible. Este supuesto implica estacionariedad estricta: la demanda no presenta variabilidad ni estacionalidad en el tiempo.
2. **Tiempo de reposición nulo:** El pedido realizado se recibe de forma inmediata; es decir, el tiempo de entrega (*lead time*) es igual a cero. En consecuencia, no se consideran inventarios en tránsito ni la posibilidad de desabastecimiento durante el tiempo de espera.
3. **Costos invariantes en el tiempo:**
 - El costo de pedido $K > 0$ (moneda/pedido) es fijo e independiente de la cantidad solicitada.
 - El costo unitario de adquisición $c > 0$ (moneda/unidad) es constante y no sujeto a descuentos por volumen.
 - El costo unitario de mantenimiento de inventario $h > 0$ (moneda/unidad/tiempo) es lineal y proporcional al nivel de inventario promedio.
4. **Horizonte de planificación infinito:** Se asume un horizonte temporal ilimitado y estacionario para las decisiones tácticas de inventario, lo que permite analizar el subproblema de reposición en régimen permanente. Bajo esta condición, el patrón de inventario asociado a cada almacén abierto se repite cíclicamente con un periodo

fijo (como en el modelo EOQ clásico). Las decisiones estratégicas de localización, en cambio, se consideran estáticas y no dependen del tiempo.

5. **No se permiten faltantes:** Todo el volumen demandado debe ser satisfecho inmediatamente desde el inventario disponible. Por tanto, el nivel de inventario nunca puede ser negativo.
6. **Política de lote único y entrega completa:** La reposición se efectúa mediante órdenes de tamaño fijo $Q > 0$ que se entregan en una sola partida. Al llegar el inventario a cero, se emite el pedido y el stock se restablece al nivel Q , conforme al tiempo de entrega nulo establecido en el supuesto 2.

Dado este conjunto de hipótesis, la evolución del inventario $x(t)$ en el tiempo sigue una trayectoria determinista en forma de diente de sierra: decrece linealmente desde Q hasta 0 a una tasa constante d , y luego se reabastece instantáneamente a Q . El periodo del ciclo es entonces $T = Q/d$. Estos supuestos, si bien facilitan el análisis, imponen una estructura rígida que limita severamente la aplicabilidad del modelo en contextos donde la demanda es estocástica, los tiempos de entrega son inciertos o los faltantes son inevitables, situaciones comunes en operaciones humanitarias.

La formulación matemática del modelo EOQ se deriva directamente de estos supuestos, como se detalla en la sección siguiente. No obstante, ya en este punto es pertinente señalar que la idealización determinista del EOQ contrasta fuertemente con las condiciones operativas en logística humanitaria, donde la incertidumbre en la demanda y la criticidad de los faltantes exigen modelos más robustos y flexibles.

5.3.2. Derivación analítica del tamaño óptimo de pedido

Se adopta una aproximación determinista basada en cuantiles de una distribución. Este enfoque permite formular un modelo de optimización no lineal determinista. Bajo este esquema, el problema del lote económico consiste en determinar el tamaño de pedido $Q > 0$ que minimiza el costo total promedio por unidad de tiempo. La formulación se basa en la descomposición del costo total en sus componentes: costo fijo de emisión de pedidos y costo variable de mantenimiento de inventario.

Bajo los supuestos deterministas descritos en la sección anterior, el problema del lote económico consiste en determinar el tamaño de pedido $Q > 0$ que minimiza el costo total promedio por unidad de tiempo. La formulación se basa en la descomposición del costo total en sus componentes: costo fijo de emisión de pedidos y costo variable de mantenimiento de inventario.

Dado un ciclo de longitud $T = Q/d$, el número de pedidos emitidos por unidad de tiempo es d/Q . Por tanto, el costo promedio de pedido por unidad de tiempo es:

$$C_{\text{pedido}}(Q) = K \cdot \frac{d}{Q}.$$

Por otro lado, la evolución del inventario $x(t)$ en cada ciclo es lineal y decreciente, comenzando en Q y finalizando en 0. El inventario promedio en el ciclo es:

$$\bar{x} = \frac{1}{T} \int_0^T x(t) dt = \frac{1}{Q/d} \int_0^{Q/d} (Q - d \cdot t) dt = \frac{Q}{2}.$$

En consecuencia, el costo promedio de mantenimiento por unidad de tiempo es:

$$C_{\text{mant}}(Q) = h \cdot \frac{Q}{2}.$$

El costo total promedio por unidad de tiempo, denotado $C(Q)$, se obtiene sumando ambos componentes:

$$C(Q) = \frac{Kd}{Q} + \frac{hQ}{2}, \quad Q > 0.$$

Obsérvese que el costo unitario de adquisición c no aparece en esta expresión, ya que bajo los supuestos de demanda constante y horizonte infinito, el costo de compra total por unidad de tiempo es cd , independiente de Q , y por tanto no influye en la optimización.

El problema de optimización asociado al modelo EOQ se formula como:

$$\min_{Q>0} C(Q) = \frac{Kd}{Q} + \frac{hQ}{2}.$$

La función objetivo $C(Q)$ es estrictamente convexa en $\mathbb{R}_{>0}$, ya que su segunda derivada es positiva:

$$C'(Q) = -\frac{Kd}{Q^2} + \frac{h}{2}, \quad C''(Q) = \frac{2Kd}{Q^3} > 0 \quad \forall Q > 0.$$

Por lo tanto, el mínimo global se alcanza en el único punto estacionario que satisface $C'(Q) = 0$. Resolviendo esta condición de primer orden:

$$-\frac{Kd}{Q^2} + \frac{h}{2} = 0 \quad \implies \quad \frac{Kd}{Q^2} = \frac{h}{2} \quad \implies \quad Q^2 = \frac{2Kd}{h} \quad \implies \quad Q^* = \sqrt{\frac{2Kd}{h}}.$$

Este valor Q^* se conoce como la *cantidad económica de pedido* (EOQ, por sus siglas en inglés).

El costo total mínimo asociado a Q^* es:

$$C(Q^*) = \frac{Kd}{\sqrt{2Kd/h}} + \frac{h}{2} \sqrt{\frac{2Kd}{h}} = \sqrt{2Kdh}.$$

Además, en el óptimo, los costos de pedido y mantenimiento son iguales:

$$\frac{Kd}{Q^*} = \frac{hQ^*}{2} = \frac{1}{2} \sqrt{2Kdh},$$

lo cual refleja un equilibrio estructural entre los dos componentes del costo, una característica distintiva del modelo EOQ.

5.3.2.1. Ejemplo numérico ilustrativo

Consideremos un producto con los siguientes parámetros: demanda anual $d = 1200$ unidades/año, costo fijo de pedido $K = 100$ unidades monetarias/pedido, y costo unitario de mantenimiento $h = 5$ unidades monetarias/unidad/año. El tamaño óptimo de pedido es:

$$Q^* = \sqrt{\frac{2 \cdot 100 \cdot 1200}{5}} = \sqrt{48\,000} \approx 219.09 \text{ unidades.}$$

El número óptimo de pedidos anuales es $d/Q^* \approx 5.48$, y el costo total anual mínimo es:

$$C(Q^*) = \sqrt{2 \cdot 100 \cdot 1200 \cdot 5} = \sqrt{1\,200\,000} \approx 1095.45 \text{ unidades monetarias/año.}$$

Este ejemplo, aunque sencillo, ilustra la sensibilidad del modelo a los parámetros de entrada y su dependencia crítica en la exactitud de las estimaciones de d , K y h un aspecto que se vuelve problemático cuando dichos parámetros son inciertos, como ocurre en entornos humanitarios.

La derivación presentada demuestra que, bajo supuestos deterministas estrictos, el modelo EOQ admite una solución cerrada con propiedades analíticas claras. Sin embargo, esta elegancia matemática se ve comprometida en contextos donde dichos supuestos no se cumplen, motivando la necesidad de extensiones robustas o estocásticas.

5.3.3. Inadecuación en contextos humanitarios: demanda incierta, faltantes costosos

El modelo EOQ, aunque elegante desde el punto de vista analítico, se fundamenta en una representación idealizada del entorno operativo que resulta profundamente inadecuada para aplicaciones en logística humanitaria. En este contexto, dos características críticas invalidan los supuestos centrales del EOQ: **la incertidumbre inherente en la demanda** y **el alto costo, frecuentemente no monetario de los faltantes**. A continuación, se analizan estas limitaciones y se ilustra, mediante un contraste formal.

5.3.3.1. Incertidumbre en la demanda

En contraste con el supuesto EOQ de demanda constante y conocida, las operaciones humanitarias se caracterizan por una demanda altamente estocástica, con alta varianza y asimetría. Tras un desastre (sismo, inundación, conflicto, etc.), la magnitud, distribución geográfica y composición de la necesidad humanitaria son desconocidas *a priori* y evolucionan dinámicamente con el tiempo. Modelar la demanda como una variable aleatoria \tilde{D} , con función de distribución acumulada $F_D(\cdot)$ y densidad $f_D(\cdot)$ (cuando existe), es no solo razonable, sino necesario.

Bajo incertidumbre, el costo esperado de inventario ya no se reduce a la expresión determinista $C(Q)$. En lugar de ello, si se permite la posibilidad de faltantes (insatisfecho inevitable en la práctica), el costo esperado por ciclo depende de la función de pérdida:

$$\mathbb{E}[\text{Faltante}] = \mathbb{E}[(\tilde{D} - Q)^+] = \int_Q^\infty (d - Q)f_D(d) dd,$$

$$\mathbb{E}[\text{Inventario residual}] = \mathbb{E}[(Q - \tilde{D})^+] = \int_0^Q (Q - d)f_D(d) dd,$$

donde $(\cdot)^+ = \max\{0, \cdot\}$. Si se introduce un costo unitario de faltante $p > 0$ (penalización por unidad no satisfecha), el costo esperado por ciclo es:

$$\mathcal{C}(Q) = K + c\mathbb{E}[\tilde{D}] + h\mathbb{E}[(Q - \tilde{D})^+] + p\mathbb{E}[(\tilde{D} - Q)^+].$$

Minimizar $\mathcal{C}(Q)$ conduce al modelo clásico de *newsvendor* (o *periódico*), cuya solución óptima denotada por Q^{NV} satisface:

$$F_D(Q^{\text{NV}}) = \frac{p - c}{p + h - c} \quad (\text{asumiendo } p > c).$$

Este resultado contrasta radicalmente con el EOQ: la decisión óptima depende ahora de la **distribución completa de la demanda**, no solo de su media, y el tamaño de pedido se determina por un balance de *margen de utilidad* frente a *riesgo de sobreinventario*, no por un equilibrio entre costos fijos y variables.

5.3.3.2. Costo no lineal y no monetario de los faltantes

En contextos comerciales, el costo de faltante p suele interpretarse como una pérdida de ganancia marginal o un costo de reorden. En logística humanitaria, sin embargo, los faltantes pueden implicar **consecuencias humanas severas**: incremento en la morbilidad, mortalidad evitable, deterioro de la dignidad humana o escalada de tensiones sociales. Estos efectos no son fácilmente cuantificables en unidades monetarias y, más aún, exhiben una estructura **no lineal** y **convexa**: la penalización marginal por cada unidad adicional de faltante crece con el nivel de insatisfacción.

Por ejemplo, si se modela el impacto humanitario mediante una función de penalización convexa $\phi(\cdot)$, donde $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ es creciente y convexa (e.g., $\phi(z) = z^\alpha$ con $\alpha > 1$), entonces el costo esperado asociado al faltante es:

$$\mathbb{E}[\phi((\tilde{D} - Q)^+)].$$

Esta formulación rompe la linealidad que subyace tanto al EOQ como al modelo de *newsvendor*, y da lugar a problemas de optimización no lineales y, en general, no convexos, que requieren técnicas avanzadas de análisis estocástico o teoría del riesgo.

5.3.3.3. Ejemplo ilustrativo: EOQ vs. realidad humanitaria

Supongamos que un organismo humanitario debe pre-posicionar kits de emergencia antes de la temporada de huracanes en una región costera. Se estima que la demanda \tilde{D} sigue una distribución log-normal con media $\mu_D = 1200$ y desviación estándar $\sigma_D = 600$ (coeficiente de variación = 0.5). Si se aplicara ingenuamente el EOQ con $d = \mu_D = 1200$, $K = 100$ unidades monetarias, $h = 5$ unidades monetarias/unidad/año, se obtendría:

$$Q^{\text{EOQ}} = \sqrt{\frac{2 \cdot 100 \cdot 1200}{5}} \approx 219 \text{ unidades.}$$

Sin embargo, la probabilidad de faltante bajo este nivel de inventario es:

$$\mathbb{P}(\tilde{D} > 219) \approx 1 - F_D(219) \approx 0.97,$$

es decir, **el 97% de los escenarios conducen a faltantes severos**. Si, en cambio, se adopta un enfoque de *newsvendor* con un costo de faltante conservador $p = 100$ unidades monetarias/unidad (reflejando impacto indirecto), se obtiene:

$$F_D(Q^{\text{NV}}) = \frac{p}{p+h} \approx \frac{100}{105} \approx 0.952,$$

lo que implica $Q^{\text{NV}} \approx F_D^{-1}(0.952) \approx 2500$ unidades—más de **11 veces** el nivel EOQ. La brecha entre ambas decisiones subraya la inadecuación del modelo determinista en entornos de alta incertidumbre y consecuencias asimétricas.

En resumen, el modelo EOQ ignora tanto la **variabilidad** como la **asimetría de costos** inherentes a la logística humanitaria. Su aplicación directa en tales contextos no solo es técnicamente incorrecta desde el punto de vista de la teoría de decisiones bajo incertidumbre, sino que puede comprometer gravemente la efectividad de la respuesta humanitaria. Esta inadecuación motiva formalmente la necesidad de modelos que incorporen **estructuras de costo no lineales, medidas de riesgo coherentes y penalizaciones asimétricas**, como se desarrollará en la sección siguiente.

5.3.4. Motivación para una extensión no lineal con penalización y riesgo

La inadecuación del modelo EOQ en contextos humanitarios, analizada en la sección anterior, surge fundamentalmente de dos limitaciones estructurales: (i) su tratamiento determinista de la demanda y (ii) su suposición de costos lineales y simétricos.

Para abordar estos vacíos, es necesario formular un modelo que:

- incorpore explícitamente la **incertidumbre en la demanda**,
- represente de manera rigurosa el **alto y no lineal costo de los faltantes**, y

- considere **medidas de riesgo coherentes** que reflejen la aversión a eventos extremos, característica crítica en operaciones humanitarias.

Esta sección motiva la necesidad de una extensión no lineal del modelo clásico, introduciendo una formulación estocástica con penalización convexa y una métrica de riesgo más allá del valor esperado.

5.3.4.1. Formulación estocástica con penalización no lineal

Sea \tilde{D} una variable aleatoria no negativa que representa la **demanda que se observa tras un desastre**, definida sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Denotemos por $Q \geq 0$ la cantidad de inventario preposicionado (decisión de primer estadio). El costo total asociado a Q se descompone en:

- Costo fijo de preparación: K (independiente de Q),
- Costo de adquisición: cQ ,
- Costo de mantenimiento esperado: $h \mathbb{E}[(Q - \tilde{D})^+]$,
- **Penalización esperada por faltantes**, modelada mediante una función convexa $\phi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, estrictamente creciente, con $\phi(0) = 0$ y $\phi \in \mathcal{C}^2(\mathbb{R}_{>0})$.

La función $\phi(\cdot)$ captura la no linealidad del impacto humanitario: por ejemplo, $\phi(z) = z^\alpha$ con $\alpha > 1$ refleja que la marginalidad del daño crece con la magnitud del faltante. El costo esperado total es entonces:

$$\mathcal{C}(Q) = K + cQ + h \mathbb{E}[(Q - \tilde{D})^+] + \mathbb{E}[\phi((\tilde{D} - Q)^+)].$$

El problema de optimización se plantea como:

$$\min_{Q \geq 0} \mathcal{C}(Q).$$

Bajo las hipótesis de convexidad de ϕ y continuidad de la distribución de \tilde{D} , la función $\mathcal{C}(Q)$ es estrictamente convexa (véase Shapiro et al., 2021, *Lectures on Stochastic Programming*), garantizando la existencia y unicidad de un minimizador global Q^* .

Sin embargo, minimizar el **valor esperado** puede ser insuficiente en contextos humanitarios, donde la ocurrencia de eventos catastróficos *aunque de baja probabilidad* puede tener consecuencias inaceptables. Por ejemplo, una política que minimice el costo promedio podría tolerar escenarios con faltantes masivos en un 5 de los casos, lo cual es inadmisibles desde una perspectiva ética o operativa.

5.3.4.2. Incorporación de medidas de riesgo coherentes

Para abordar esta preocupación, se propone reemplazar el operador de expectativa por una **medida de riesgo coherente** $\mathcal{R} : \mathcal{L} \rightarrow \mathbb{R}$, donde \mathcal{L} es un espacio de variables aleatorias integrables. Siguiendo Artzner et al. (1999), una medida de riesgo coherente satisface:

1. **Monotonía:** si $X \leq Y$ c.s., entonces $\mathcal{R}(X) \geq \mathcal{R}(Y)$,
2. **Invarianza traslacional:** $\mathcal{R}(X + a) = \mathcal{R}(X) - a$ para $a \in \mathbb{R}$,
3. **Subaditividad:** $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$,
4. **Homogeneidad positiva:** $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ para $\lambda \geq 0$.

Una elección particularmente relevante en logística humanitaria es el **Valor en Riesgo condicional** (*Conditional Value-at-Risk*, CVaR), también conocido como *Expected Shortfall*. Para un nivel de confianza $\beta \in (0, 1)$, el CVaR del costo total se define como:

$$\text{CVaR}_\beta(C(Q, \tilde{D})) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1 - \beta} \mathbb{E}[(C(Q, \tilde{D}) - \eta)^+] \right\},$$

donde $C(Q, \tilde{D}) = cQ + h(Q - \tilde{D})^+ + \phi((\tilde{D} - Q)^+)$ es el costo aleatorio (sin el término fijo K).

El problema robusto-averso al riesgo se formula entonces como:

$$\min_{Q \geq 0} \text{CVaR}_\beta(C(Q, \tilde{D})).$$

Esta formulación no solo penaliza los faltantes de forma no lineal, sino que también enfatiza explícitamente los peores $(1 - \beta) \times 100\%$ escenarios. Por ejemplo, con $\beta = 0.95$, se optimiza el promedio de los peores 5% de los costos posibles.

5.3.4.3. Ejemplo numérico: impacto de la aversión al riesgo

Supongamos $\tilde{D} \sim \text{Lognormal}(\mu = 6.8, \sigma = 0.6)$, lo que implica $\mathbb{E}[\tilde{D}] \approx 1200$, Desv. estándar ≈ 800 . Considérese $\phi(z) = z^{1.5}$, $c = 10$ unidades monetarias/unidad, $h = 2$ unidades monetarias/unidad, y $\beta = 0.95$.

1. **Solución basada en valor esperado:**
Se resuelve $\min_Q \mathbb{E}[C(Q, \tilde{D})]$ numéricamente (por integración Monte Carlo o cuadratura). Supongamos que el óptimo es $Q^E \approx 1850$.
2. **Solución basada en CVaR:**
Se resuelve $\min_Q \text{CVaR}_{0.95}(C(Q, \tilde{D}))$. Debido a la aversión a colas pesadas, el óptimo suele ser significativamente mayor: $Q^{\text{CVaR}} \approx 2600$.

3. Comparación de desempeño:

- Bajo Q^E , el $CVaR_{0.95}$ es 35 mayor que bajo Q^{CVaR} .
- La probabilidad de faltante superiores a 1000 unidades es 12 % con Q^E , pero solo 3 % con Q^{CVaR} .

Este ejemplo demuestra que la incorporación de una medida de riesgo coherente conduce a decisiones más conservadoras, pero también más robustas frente a eventos extremos, una característica esencial en logística humanitaria.

La extensión propuesta, basada en penalización no lineal y medidas de riesgo coherentes, no es meramente una modificación técnica, sino una **reformulación conceptual** del problema de inventario. Abandona la ficción de la certidumbre y la linealidad para abrazar la complejidad ética y operativa de los contextos humanitarios. Esta formulación sienta las bases para el modelo integrado de localización-inventario que se desarrollará, donde la decisión de inventario Q estará acoplada a la selección de localidades de almacenamiento y a la asignación de flujos bajo incertidumbre espacial y temporal.

5.4. Extensión no lineal del modelo EOQ con incertidumbre y penalización

Para superar las limitaciones del modelo EOQ clásico en contextos humanitarios, especialmente su incapacidad para manejar demanda estocástica y la asimetría crítica entre los costos de sobreinventario y faltante, proponemos una extensión no lineal que incorpora:

- (I) una representación estocástica de la demanda,
- (II) una función de penalización convexa para faltantes, y
- (III) un enfoque de optimización basado en el valor esperado del costo total, como paso previo a la incorporación de medidas de riesgo (tratada en secciones posteriores).

Esta extensión preserva la estructura de decisión única del EOQ (tamaño de pedido pre-posicionado), pero generaliza radicalmente su función objetivo.

5.4.1. Definición rigurosa de la función de costo total extendida $Z(Q)$

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad completo, y sea $\tilde{D} : \Omega \rightarrow \mathbb{R}_+$ una variable aleatoria integrable que modela la demanda total durante el horizonte de planificación (por ejemplo, la fase inicial de respuesta post-desastre). Denotamos por F_D y f_D su función de distribución acumulada y densidad, respectivamente (asumiendo que f_D existe para fines de diferenciabilidad).

Definimos la **decisión de inventario** como una cantidad determinista $Q \in \mathbb{R}_+$, interpretada como el volumen de suministros pre-posicionado en un almacén central o regional antes de que se materialice la demanda. La función de costo total aleatorio asociada a Q se construye a partir de los siguientes componentes:

1. **Costo fijo de preparación:** $K > 0$, independiente de Q .
2. **Costo de adquisición:** cQ , con $c > 0$ el costo unitario de compra o producción.
3. **Costo de mantenimiento:** $h \cdot (Q - \tilde{D})^+$, con $h > 0$ el costo unitario de mantener inventario residual.
4. **Penalización por faltantes:** $\phi((\tilde{D} - Q)^+)$, donde $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ es una función de penalización que satisface:
 - $\phi(0) = 0$,
 - ϕ es estrictamente creciente,
 - ϕ es convexa y continuamente diferenciable en $\mathbb{R}_{>0}$,
 - $\phi'(z) \rightarrow \infty$ cuando $z \rightarrow \infty$ (penalización creciente de forma superlineal).

Estas propiedades reflejan que el costo marginal de no satisfacer necesidades humanitarias adicionales crece con la magnitud del faltante, capturando la gravedad ética y operativa de los déficits extremos.

Bajo este marco, el **costo total aleatorio** es:

$$C(Q, \tilde{D}) = K + cQ + h(Q - \tilde{D})^+ + \phi((\tilde{D} - Q)^+).$$

La **función de costo total extendida**, denotada $Z(Q)$, se define como el valor esperado del costo total:

$$Z(Q) := \mathbb{E}[C(Q, \tilde{D})] = K + cQ + h \mathbb{E}[(Q - \tilde{D})^+] + \mathbb{E}[\phi((\tilde{D} - Q)^+)].$$

Descomponiendo las expectativas en integrales, y asumiendo que \tilde{D} tiene densidad f_D , obtenemos la representación analítica:

$$Z(Q) = K + cQ + h \int_0^Q (Q - d) f_D(d) dd + \int_Q^\infty \phi(d - Q) f_D(d) dd.$$

Esta expresión es la base del modelo extendido. Nótese que:

- Cuando $\phi(z) = pz$ (lineal) y $\tilde{D} = d$ constante, $Z(Q)$ se reduce a la función EOQ clásica (salvo el término fijo K y el costo de compra, irrelevantes para la optimización).

- La convexidad de ϕ y la linealidad del costo de mantenimiento garantizan, bajo condiciones técnicas estándar (e.g., f_D continua y positiva en su soporte), que $Z(Q)$ es **estrictamente convexa** en \mathbb{R}_+ , lo que asegura la existencia y unicidad de un minimizador global.

5.4.1.1. Condiciones de optimalidad

Dado que $Z(Q)$ es diferenciable en $\mathbb{R}_{>0}$ (por el lema de Leibniz y la regularidad de ϕ y f_D), su derivada es:

$$Z'(Q) = c + h \int_0^Q f_D(d) dd - \int_Q^\infty \phi'(d - Q) f_D(d) dd.$$

Reescribiendo en términos de la función de distribución:

$$Z'(Q) = c + hF_D(Q) - \int_0^\infty \phi'(z) f_D(Q + z) dz.$$

El óptimo Q^* satisface la condición de primer orden $Z'(Q^*) = 0$, es decir:

$$\int_0^\infty \phi'(z) f_D(Q^* + z) dz = c + hF_D(Q^*).$$

Esta ecuación no admite, en general, una solución cerrada, pero su estructura revela un **balance marginal estocástico**: el lado izquierdo representa el costo marginal esperado de faltantes (ponderado por la sensibilidad creciente ϕ'), mientras que el lado derecho refleja el ahorro marginal en adquisición y mantenimiento al aumentar Q .

5.4.1.2. Ejemplo del caso con $\phi(z) = z^2$

Consideremos un caso concreto para ilustrar el cálculo de $Z(Q)$ y su minimización. Supongamos:

- $\tilde{D} \sim \text{Uniforme}(a = 800, b = 1600)$,
- $K = 200$ unidades monetarias,
- $c = 15$ unidades monetarias/unidad,
- $h = 3$ unidades monetarias/unidad,
- $\phi(z) = z^2$ (penalización cuadrática, convexa, $\phi'(z) = 2z$).

Dado el soporte de \tilde{D} , analizamos tres casos:

- Si $Q \leq 800$: entonces $(Q - \tilde{D})^+ = 0$ c.s., y $(\tilde{D} - Q)^+ = \tilde{D} - Q$.
- Si $800 < Q < 1600$: ambos términos son no nulos.
- Si $Q \geq 1600$: $(\tilde{D} - Q)^+ = 0$ c.s.

Nos enfocamos en el rango relevante $Q \in (800, 1600)$. La densidad es $f_D(d) = 1/800$ para $d \in [800, 1600]$.

Entonces:

$$Z(Q) = 200 + 15Q + 3 \int_{800}^Q (Q - d) \frac{1}{800} dd + \int_Q^{1600} (d - Q)^2 \frac{1}{800} dd.$$

Lo que permite expresar a $Z(Q)$ por tramos.

Lo anterior nos permite interpretar:

- El óptimo $Q^* \approx 1482$ unidades está significativamente por encima de la media de la demanda (1200), reflejando el efecto de la penalización cuadrática: el modelo prefiere incurrir en costos de sobreinventario para evitar faltantes, especialmente los grandes.
- Esto contrasta fuertemente con una aproximación inicial de la solución mediante la formulación EOQ ($Q^{\text{EOQ}} \approx \sqrt{2Kd/h} \approx \sqrt{2 \cdot 200 \cdot 1200/3} \approx 400$), que sería catastróficamente insuficiente.

Este ejemplo demuestra que la función $Z(Q)$, aunque más compleja que su contraparte EOQ, es analíticamente manejable en casos paramétricos y produce decisiones coherentes con los principios de la logística humanitaria: **precaución, equidad y protección de la vida humana**.

La formulación de $Z(Q)$ sienta las bases para la integración con decisiones de localización, donde Q dejará de ser una variable escalar para convertirse en un vector $\{Q_i\}_{i \in I}$ indexado sobre localidades candidatas.

5.4.2. Interpretación del término $Z_\sigma L$: nivel de servicio y riesgo

En la literatura de inventarios bajo incertidumbre, es común descomponer la decisión de inventario en dos componentes: un **nivel base de reposición** que cubre la demanda promedio, y una **cantidad de seguridad** (*safety stock*) que protege contra la variabilidad de la demanda. En contextos con tiempo de entrega no nulo, esta descomposición se expresa como

$$Q = \mu_L + z_\beta \sigma_L,$$

donde μ_L y σ_L son la media y desviación estándar de la demanda durante el tiempo de entrega L , y z_β es un factor cuantil asociado a un nivel de servicio objetivo $\beta \in (0, 1)$. Aunque el modelo EOQ clásico asume $L = 0$, esta estructura resulta útil para interpretar el comportamiento de modelos estocásticos extendidos en términos de **nivel de servicio** y **exposición al riesgo**.

Adoptamos una notación análoga para interpretar la solución óptima Q^* de la función de costo $Z(Q)$, incluso cuando $L = 0$, ya que la incertidumbre en la demanda humanitaria actúa *de facto* como un “riesgo de exposición” que debe ser mitigado mediante inventario de seguridad. Denotamos:

- $\mu_D = \mathbb{E}[\tilde{D}]$, la demanda esperada,
- $\sigma_D = \sqrt{\text{Var}(\tilde{D})}$, la desviación estándar de la demanda,
- z^* el **factor de seguridad implícito**, definido por la relación:

$$Q^* = \mu_D + z^* \sigma_D.$$

El término $z^* \sigma_D$ juega el papel de *inventario de seguridad*, y su magnitud refleja la aversión al riesgo incorporada en la función de penalización $\phi(\cdot)$.

5.4.2.1. Relación entre z^* y el nivel de servicio

El **nivel de servicio tipo I** (también llamado *probabilidad de no faltante*) se define como:

$$\beta := \mathbb{P}(\tilde{D} \leq Q) = F_D(Q).$$

En el caso gaussiano ($\tilde{D} \sim \mathcal{N}(\mu_D, \sigma_D^2)$), se tiene $Q = \mu_D + z_\beta \sigma_D$, con $z_\beta = \Phi^{-1}(\beta)$, donde Φ es la función de distribución estándar normal. Aunque la demanda humanitaria rara vez es normal, esta relación sirve como **marco interpretativo** incluso para distribuciones generales.

En nuestro modelo extendido, el óptimo Q^* no se determina por un β exógeno, sino de forma **endógena**, como resultado del equilibrio marginal entre el costo de mantener inventario y la penalización esperada por faltantes. Sin embargo, una vez obtenido Q^* , es posible calcular el nivel de servicio implícito:

$$\beta^* = F_D(Q^*) = F_D(\mu_D + z^* \sigma_D).$$

Este valor β^* cuantifica la **cobertura estocástica** lograda por la solución óptima y permite comparar distintas estructuras de penalización en términos operacionales (p. ej., “esta política satisface la demanda en el 95 % de los escenarios”).

5.4.2.2. Derivación de z^* bajo la penalización cuadrática

Para ilustrar esta interpretación, retomamos el ejemplo de la Sección 5.4.1 con $\tilde{D} \sim \text{Uniforme}(800, 1600)$. Tenemos:

- $\mu_D = (800 + 1600)/2 = 1200$,
- $\sigma_D^2 = (1600 - 800)^2/12 = 640\,000/12 \approx 53\,333.33$,
- $\sigma_D \approx \sqrt{53\,333.33} \approx 230.94$.

Del cálculo anterior, $Q^* \approx 1481.5$. Por lo tanto:

$$z^* = \frac{Q^* - \mu_D}{\sigma_D} \approx \frac{1481.5 - 1200}{230.94} \approx \frac{281.5}{230.94} \approx 1.22.$$

El nivel de servicio implícito es:

$$\beta^* = F_D(1481.5) = \frac{1481.5 - 800}{1600 - 800} = \frac{681.5}{800} \approx 0.852.$$

Esto significa que, bajo la penalización cuadrática $\phi(z) = z^2$, la política óptima acepta un **15 % de probabilidad de faltante**, pero está diseñada para minimizar el impacto de los faltantes cuando ocurren, gracias a la sensibilidad creciente de ϕ . Obsérvese que si se hubiera impuesto un nivel de servicio exógeno $\beta = 0.95$, se requeriría $Q = 800 + 0.95 \cdot 800 = 1560$, lo cual generaría un costo de inventario más alto sin necesariamente mejorar el desempeño bajo la métrica de penalización no lineal.

5.4.2.3. Generalización y conexión con medidas de riesgo

El factor z^* no es una constante universal, sino una **variable de decisión endógena** que depende de:

- La distribución de \tilde{D} (especialmente sus colas),
- Los parámetros de costo (c, h) ,
- La forma funcional de $\phi(\cdot)$.

Es decir, si $\phi(z) = pz$ (lineal), la condición de optimalidad se reduce a:

$$F_D(Q^*) = \frac{p - c}{p + h - c},$$

lo que implica un z^* fijo para una distribución dada recuperando el modelo de *newsvendor*. En cambio, si ϕ es convexa y creciente (como $z^{1.5}$, $e^{\lambda z}$, etc.), z^* aumenta con la curvatura de ϕ , reflejando mayor aversión a los faltantes extremos.

Más aún, si en lugar de minimizar $\mathbb{E}[C(Q, \tilde{D})]$ se minimiza $\text{CVaR}_\beta(C(Q, \tilde{D}))$ (Valor Condicional en Riesgo, una medida coherente del promedio de las peores pérdidas), el factor z^* se ajusta explícitamente para controlar la magnitud de los peores $(1 - \beta) \times 100\%$ escenarios, lo que puede llevar a valores de z^* significativamente mayores que los obtenidos bajo valor esperado una propiedad deseable en logística humanitaria.

El término $z^* \sigma_D$ (análogo al clásico $z_\beta \sigma_L$) proporciona una **métrica interpretable** del compromiso entre costo y riesgo en el modelo extendido. A diferencia del enfoque clásico, donde z_β es un parámetro de diseño impuesto por el decisor, aquí z^* emerge de la estructura de costos y la distribución de la demanda, encapsulando de forma rigurosa la **aversión implícita al riesgo humanitario**. Esta interpretación será crucial al integrar el modelo de inventario con decisiones de localización, donde distintas zonas de cobertura exhibirán diferentes perfiles de incertidumbre y, por tanto, distintos factores de seguridad óptimos.

5.4.3. Incorporación de la penalización por faltantes mediante β_C

En contextos humanitarios, la insatisfacción de la demanda no puede reducirse a una pérdida económica lineal; en cambio, debe modelarse mediante una estructura de penalización que refleje la **gravedad creciente** del impacto humano a medida que el volumen de faltantes aumenta. Aunque en la Sección 5.4.1 se introdujo una función convexa general $\phi(\cdot)$, en muchas aplicaciones prácticas y formulaciones teóricas es útil parametrizar explícitamente la aversión al faltante mediante un **coeficiente de penalización escalar** $\beta_C > 0$, que actúa como un multiplicador del costo marginal asociado a la insatisfacción. Esta parametrización no solo mejora la interpretabilidad del modelo, sino que también facilita el análisis de sensibilidad y la calibración con expertos humanitarios.

En esta sección, desarrollamos rigurosamente una versión del modelo extendido donde la penalización por faltantes se introduce a través de un término de la forma $\beta_C \cdot \psi((\tilde{D} - Q)^+)$, donde $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ es una función de forma fija (por ejemplo, lineal, cuadrática o exponencial), y β_C captura la **intensidad relativa** de la penalización en unidades consistentes con los costos de inventario.

5.4.3.1. Formulación de la función de costo con β_C

Sea $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ una función de forma normalizada, tal que:

- $\psi(0) = 0$,
- ψ es estrictamente creciente y convexa,
- $\psi \in \mathcal{C}^1(\mathbb{R}_{>0})$,
- $\psi'(0^+) = 1$ (normalización que permite interpretar β_C como el costo marginal inicial por unidad de faltante) Sección 4.3.

Definimos la **función de costo total estocástico parametrizada** como:

$$C_{\beta_C}(Q, \tilde{D}) = K + cQ + h(Q - \tilde{D})^+ + \beta_C \psi((\tilde{D} - Q)^+).$$

La función de costo esperado es entonces:

$$Z_{\beta_C}(Q) := \mathbb{E}[C_{\beta_C}(Q, \tilde{D})] = K + cQ + h \mathbb{E}[(Q - \tilde{D})^+] + \beta_C \mathbb{E}[\psi((\tilde{D} - Q)^+)].$$

El problema de optimización asociado es:

$$\min_{Q \geq 0} Z_{\beta_C}(Q).$$

5.4.3.2. Condiciones de optimalidad y sensibilidad respecto a β_C

Bajo las hipótesis de regularidad (densidad f_D continua y positiva en el soporte de \tilde{D}), $Z_{\beta_C}(Q)$ es estrictamente convexa y diferenciable en $\mathbb{R}_{>0}$. Su derivada es:

$$Z'_{\beta_C}(Q) = c + hF_D(Q) - \beta_C \int_0^{\infty} \psi'(z) f_D(Q+z) dz.$$

El óptimo $Q^*(\beta_C)$ satisface:

$$c + hF_D(Q^*(\beta_C)) = \beta_C \int_0^{\infty} \psi'(z) f_D(Q^*(\beta_C) + z) dz. \quad (5.4.1)$$

Esta ecuación revela una relación **monótona creciente**: a mayor β_C , mayor $Q^*(\beta_C)$. Formalmente, derivando implícitamente (Ecuación 5.5.4) respecto a β_C , se obtiene:

$$\frac{dQ^*}{d\beta_C} = \frac{\int_0^{\infty} \psi'(z) f_D(Q^* + z) dz}{hf_D(Q^*) + \beta_C \int_0^{\infty} \psi''(z) f_D(Q^* + z) dz} > 0,$$

donde el denominador es positivo por la convexidad de ψ ($\psi'' \geq 0$) y $f_D > 0$. Esto confirma que β_C actúa como un **parámetro de aversión al faltante**: incrementarlo induce al decisor a almacenar más inventario como medida de precaución.

5.4.3.3. Elección de $\psi(\cdot)$ y significado operacional de β_C

A continuación, se consideran tres formas comunes para $\psi(\cdot)$, cada una con una interpretación distinta de β_C :

1. **Lineal**: $\psi(z) = z$.
Entonces β_C representa el **costo marginal constante** por unidad de faltante (modelo de *newsvendor* clásico).
2. **Cuadrática**: $\psi(z) = z^2/2$.
La normalización $\psi'(0^+) = 0$ viola nuestra convención, por lo que ajustamos: $\psi(z) = z + \frac{\theta}{2}z^2$, con $\theta > 0$. En este caso, β_C es el costo marginal inicial, y θ controla la curvatura. Alternativamente, si se fija $\psi(z) = z^2$, entonces β_C debe interpretarse como un **parámetro de escala no lineal**, sin unidades directas de costo por unidad.
3. **Exponencial truncada**: $\psi(z) = \frac{1}{\lambda}(e^{\lambda z} - 1)$, con $\lambda > 0$.
Aquí, $\psi'(0^+) = 1$ y $\psi''(z) = \lambda e^{\lambda z} > 0$, por lo que β_C mantiene su interpretación como costo marginal inicial, mientras que λ gobierna la **avergüenza exponencial** al faltante extremo.

En logística humanitaria, la forma exponencial o cuadrática son más apropiadas, ya que reflejan que el costo marginal de faltantes crece rápidamente con su magnitud.

5.4.3.4. Ejemplo de calibración de β_C bajo $\psi(z) = z^2$

Supongamos que un organismo humanitario desea que la política de inventario garantice que el **valor esperado del faltante cuadrático** no exceda un umbral ético $\tau = 10^5$ unidades². Es decir:

$$\mathbb{E}[(\tilde{D} - Q)^+]^2 \leq \tau.$$

Esto no es un objetivo de optimización, sino una **restricción de desempeño**. Para incorporarla en el modelo, se puede usar el método de penalización de Lagrange: minimizar $Z_{\beta_C}(Q)$ equivale a resolver el problema restringido para un β_C que actúa como multiplicador de Lagrange.

Sea $\tilde{D} \sim \text{Uniforme}(1000, 2000)$, $c = 12$ unidades monetarias/unidad, $h = 4$ unidades monetarias/unidad, y $\psi(z) = z^2$. Entonces:

$$Z_{\beta_C}(Q) = K + 12Q + 4 \int_{1000}^Q (Q - d) \frac{1}{1000} dd + \beta_C \int_Q^{2000} (d - Q)^2 \frac{1}{1000} dd,$$

válido para $Q \in (1000, 2000)$.

Derivando e igualando a cero:

$$12 + 4 \cdot \frac{Q - 1000}{1000} = \beta_C \cdot \frac{(2000 - Q)^2}{1000}.$$

Reorganizando:

$$\beta_C = \frac{12\,000 + 4(Q - 1000)}{(2000 - Q)^2}. \quad (5.4.2)$$

Por otro lado, el faltante cuadrático esperado es:

$$\mathbb{E}[(\tilde{D} - Q)^+]^2 = \frac{1}{1000} \int_Q^{2000} (d - Q)^2 dd = \frac{(2000 - Q)^3}{3000}.$$

Imponiendo la restricción $\frac{(2000 - Q)^3}{3000} \leq 10^5$, obtenemos:

$$(2000 - Q)^3 \leq 3 \times 10^8 \quad \Rightarrow \quad 2000 - Q \leq (3 \times 10^8)^{1/3} \approx 669.4 \quad \Rightarrow \quad Q \geq 1330.6.$$

Tomando el valor límite $Q = 1330.6$, sustituimos en Ecuación 5.5.5:

$$\beta_C = \frac{12\,000 + 4(1330.6 - 1000)}{(2000 - 1330.6)^2} = \frac{12\,000 + 1322.4}{669.4^2} \approx \frac{13\,322.4}{448\,100} \approx 0.0297.$$

Interpretación: Un valor de $\beta_C \approx 0.03$ induce una política que satisface la restricción ética sobre el faltante cuadrático esperado. Valores menores de β_C violarían la restricción;

valores mayores la satisfarían con margen, pero a un costo de inventario innecesariamente alto.

Este ejemplo ilustra cómo β_C puede calibrarse para alinear el modelo con objetivos operativos o éticos específicos, superando la arbitrariedad de asignar costos monetarios directos a vidas o sufrimiento humano. La parametrización mediante β_C introduce flexibilidad sin sacrificar el análisis, y prepara el terreno para su integración en un modelo de localización-inventario multiperiodo con múltiples zonas de cobertura.

Nota: En el Capítulo 3, el parámetro β_C se generalizará a un conjunto $\{\beta_{C,i}\}_{i \in I}$, permitiendo que distintas localidades candidatas reflejen diferentes niveles de criticidad humanitaria.

5.4.4. Análisis de no linealidad: origen en D/Q y dependencia de L

La no linealidad en los modelos de inventario surge de múltiples fuentes estructurales, incluso antes de introducir incertidumbre o penalizaciones asimétricas. En el modelo EOQ clásico, la **no linealidad fundamental** proviene del término de costo de pedido, que depende del cociente d/Q , donde d es la tasa de demanda determinista y Q el tamaño del lote. Esta relación hiperbólica introduce una **curvatura intrínseca** en la función objetivo, la cual es responsable de la existencia de un mínimo global bien definido. Cuando se extiende el modelo a contextos más realistas, como aquellos con tiempo de entrega $L > 0$ o demanda estocástica, la no linealidad se enriquece y adquiere nuevas dimensiones que afectan tanto la formulación como la interpretación de las decisiones óptimas.

En esta sección, se analiza rigurosamente el origen matemático de la no linealidad en D/Q , su interacción con el tiempo de entrega L , y cómo esta estructura se transforma bajo incertidumbre, sentando las bases para la comprensión de la complejidad en modelos integrados de localización-inventario.

5.4.4.1. No linealidad determinista: el término d/Q

En el modelo EOQ determinista, el costo total por unidad de tiempo es:

$$C(Q) = \frac{Kd}{Q} + \frac{hQ}{2}, \quad Q > 0.$$

El primer término, Kd/Q , es **estrictamente convexo** y decreciente en Q , mientras que el segundo, $hQ/2$, es lineal (y, por tanto, convexo). La suma de ambas funciones es estrictamente convexa, como se verificó mediante $C''(Q) = 2Kd/Q^3 > 0$. Esta convexidad garantiza unicidad del minimizador $Q^* = \sqrt{2Kd/h}$.

La fuente de no linealidad radica exclusivamente en la **inversión** $1/Q$, que refleja una economía de escala en el costo de transacción: al aumentar Q , se reducen los pedidos por unidad de tiempo, pero a un ritmo decreciente. Esta estructura no lineal es **ineludible**

en cualquier modelo de revisión periódica o lote fijo, y persiste incluso en entornos estocásticos cuando se modela el número esperado de órdenes.

5.4.4.2. Extensión al caso con tiempo de entrega $L > 0$

Cuando el tiempo de entrega $L > 0$ es positivo y determinista, el modelo EOQ se ajusta mediante el concepto de *punto de reorden* $R = dL$, mientras que el tamaño de lote Q sigue siendo el mismo. El costo total no se ve afectado por L , ya que el inventario promedio sigue siendo $Q/2$, y el número de pedidos anuales sigue siendo d/Q . **En el caso determinista, L no introduce no linealidad adicional.**

Sin embargo, si L es **estocástico**, o si la demanda durante L es incierta, la situación cambia drásticamente. Denotemos por \tilde{D}_L la demanda aleatoria durante el tiempo de entrega L . Si L es fijo pero la demanda es estocástica, $\tilde{D}_L = \sum_{t=1}^L \tilde{d}_t$, y su distribución depende de L . El costo de faltante ahora depende de $\mathbb{E}[(\tilde{D}_L - R)^+]$, y el punto de reorden óptimo R^* depende de L a través de la distribución de \tilde{D}_L .

Más aún, si tanto L como la demanda por unidad de tiempo son aleatorios e independientes, y la demanda es estacionaria con media d y varianza σ_d^2 , entonces (bajo independencia):

$$\mu_{D_L} = d\mathbb{E}[L], \quad \sigma_{D_L}^2 = d^2 \text{Var}(L) + \sigma_d^2 \mathbb{E}[L].$$

El inventario de seguridad necesario para lograr un nivel de servicio β es $z_\beta \sigma_{D_L}$, que depende **no linealmente** de $\mathbb{E}[L]$ y $\text{Var}(L)$. En este caso, L no solo aparece como parámetro, sino que **modula la magnitud de la incertidumbre**, introduciendo una fuente adicional de no linealidad en la función de costo.

5.4.4.3. No linealidad inducida por la interacción entre Q y L bajo incertidumbre

En escenarios humanitarios, el “tiempo de entrega efectivo” puede interpretarse como el lapso entre la decisión de pre-posicionar y la materialización de la demanda (por ejemplo, entre la temporada de preparación y el desastre). Aunque no hay reposición durante la respuesta, la **exposición al riesgo** dura un periodo efectivo L , y la demanda total \tilde{D} es la demanda durante L . Si L es variable (p. ej., incertidumbre en la ocurrencia del desastre), entonces \tilde{D} es una **variable aleatoria compuesta**:

$$\tilde{D} = \int_0^L \tilde{d}(t) dt,$$

lo que induce una dependencia funcional compleja entre L y la distribución de \tilde{D} . En tales casos, la función de costo esperado $Z(Q)$ hereda una no linealidad que proviene tanto de:

- la estructura $1/Q$ (si se consideran múltiples ciclos o reposiciones parciales), como de

- la propagación de la incertidumbre de L hacia \tilde{D} .

Aun en el caso de un único periodo (sin reposición), si se considera que el costo de preparación K depende del número de localidades abiertas (como en el Capítulo 3), y cada localidad tiene su propio “tiempo efectivo de exposición” L_i , entonces $Z(Q)$ se convierte en una función **no separable y altamente no lineal** en el vector de decisiones (Q_i, y_i) , donde y_i indica si se abre la localidad i .

5.4.4.4. Ejemplo de impacto no lineal de L en la varianza de la demanda

Supongamos que la demanda diaria $\tilde{d} \sim \text{Normal}(10, 4)$ (media 10 unidades/día, varianza 4). El tiempo de entrega efectivo L (en días) es aleatorio: $L = 10$ con probabilidad 0.5, y $L = 30$ con probabilidad 0.5.

Paso 1: Calcular la distribución de \tilde{D}_L

Condicionado a $L = \ell$, $\tilde{D}_L \sim \mathcal{N}(10\ell, 4\ell)$. Por la ley de probabilidad total, la densidad marginal de \tilde{D}_L es una mezcla:

$$f_{D_L}(d) = 0.5 \cdot \phi\left(\frac{d-100}{\sqrt{40}}\right) \frac{1}{\sqrt{40}} + 0.5 \cdot \phi\left(\frac{d-300}{\sqrt{120}}\right) \frac{1}{\sqrt{120}},$$

donde ϕ es la densidad estándar normal.

Paso 2: Momentos de \tilde{D}_L

- Media: $\mathbb{E}[\tilde{D}_L] = 0.5 \cdot 100 + 0.5 \cdot 300 = 200$.
- Varianza:

$$\begin{aligned} \text{Var}(\tilde{D}_L) &= \mathbb{E}[\text{Var}(\tilde{D}_L | L)] + \text{Var}(\mathbb{E}[\tilde{D}_L | L]) \\ &= 0.5 \cdot 40 + 0.5 \cdot 120 + \text{Var}(\{100, 300\}) \\ &= 80 + 10\,000 \\ &= 10\,080. \end{aligned}$$

Paso 3: Comparación con escenario determinista

Si se ignorara la incertidumbre en L y se usara $\mathbb{E}[L] = 20$, se subestimaría la varianza: $\text{Var}(\tilde{D}_{20}) = 4 \cdot 20 = 80 \ll 10\,080$.

Paso 4: Consecuencia en el inventario de seguridad

Bajo penalización cuadrática y la función $Z(Q)$, el óptimo Q^* dependerá críticamente de $\text{Var}(\tilde{D}_L)$. Usar $L = 20$ fijo induciría $Q^* \approx \mu_D + z^* \sqrt{80}$, mientras que el modelo correcto requiere $Q^* \approx 200 + z^* \sqrt{10\,080} \approx 200 + 100.4z^*$. La diferencia es **no lineal y no aditiva**: la incertidumbre en L no solo desplaza la media, sino que infla la varianza de forma cuadrática en la desviación de L .

Este ejemplo demuestra que la dependencia de la demanda en L especialmente cuando L es aleatorio introduce una **no linealidad estructural** que no puede capturarse mediante ajustes lineales o promedios. Esta complejidad es inherente a la logística humanitaria, donde la incertidumbre temporal (cuándo ocurrirá el próximo desastre) y espacial (dónde y con qué intensidad) se entrelazan con la incertidumbre en la magnitud de la necesidad.

La no linealidad en los modelos de inventario tiene dos orígenes complementarios: 1. **Intrínseco**: el término d/Q , presente incluso en el EOQ determinista, que refleja economías de escala en el costo de transacción. 2. **Inducido por la incertidumbre**: la propagación de la variabilidad en L (o en la tasa de demanda) hacia la demanda total \tilde{D} , que introduce dependencias no lineales en los momentos de la distribución y, por tanto, en la función de costo esperado.

Reconocer y modelar ambas fuentes de no linealidad es esencial para formular modelos rigurosos y realistas en logística humanitaria. En el Capítulo 3, esta comprensión se integrará en un modelo de localización-inventario donde tanto Q_i como la exposición efectiva L_i variarán por localidad, dando lugar a una estructura de optimización no lineal y estocástica de alta dimensionalidad.

5.4.5. Comparación cuantitativa con EOQ clásico: efecto descendente en Q^*

Una crítica frecuente a la aplicación directa del modelo EOQ en contextos de alta incertidumbre, como la logística humanitaria es que, al ignorar la variabilidad de la demanda y el costo asimétrico de los faltantes, **subestima sistemáticamente** el nivel de inventario necesario para una respuesta efectiva. Sin embargo, en ciertos regímenes paramétricos, una observación menos intuitiva emerge: **cuando el costo fijo de pedido K es muy pequeño o nulo**, la solución del modelo EOQ puede, paradójicamente, sobreestimar el tamaño óptimo de inventario en comparación con un modelo estocástico bien calibrado. Este fenómeno, al que denominamos *efecto descendente en Q^** , surge de la interacción entre la estructura de costos del EOQ y la penalización convexa de faltantes, y tiene implicaciones importantes para la planificación humanitaria, donde los “costos fijos” de pre-posicionamiento pueden ser marginales frente al riesgo de insatisfacción.

En esta sección, se establece rigurosamente la comparación entre el tamaño de lote EOQ clásico, Q^{EOQ} , y el tamaño óptimo del modelo extendido no lineal con penalización, Q^* , demostrando las condiciones bajo las cuales $Q^* < Q^{\text{EOQ}}$, y se ilustra el fenómeno mediante un ejemplo controlado.

5.4.5.1. Formulaciones comparativas

Recordemos las dos funciones objetivo:

- **EOQ clásico** (determinista, con demanda media $d = \mu_D$):

$$C^{\text{EOQ}}(Q) = \frac{K\mu_D}{Q} + \frac{hQ}{2}, \quad \Rightarrow \quad Q^{\text{EOQ}} = \sqrt{\frac{2K\mu_D}{h}}.$$

- **Modelo extendido** (estocástico, con penalización convexa ϕ):

$$Z(Q) = cQ + h \mathbb{E}[(Q - \tilde{D})^+] + \mathbb{E}[\phi((\tilde{D} - Q)^+)] + K,$$

donde, a diferencia del EOQ, el costo de adquisición cQ sí influye en la optimalidad si $c > 0$ y ϕ es no lineal.

Para fines comparativos, suprimimos el término fijo K (irrelevante para la optimización) y asumimos que el modelo extendido no incluye un costo fijo por “pedido”, ya que en contextos humanitarios el preposicionamiento es una decisión única, no recurrente. Esto es coherente con la literatura de inventario de un solo periodo (*single-period*), en contraste con el horizonte infinito del EOQ.

Bajo esta interpretación, el **modelo extendido relevante para comparación** es:

$$Z_{\text{ext}}(Q) = cQ + h \mathbb{E}[(Q - \tilde{D})^+] + \mathbb{E}[\phi((\tilde{D} - Q)^+)].$$

En cambio, si forzamos la aplicación del EOQ en un contexto de un solo periodo, su función objetivo carece de justificación económica; no obstante, muchos practicantes aún usan Q^{EOQ} como una regla heurística, tomando K como un “costo administrativo” arbitrario.

5.4.5.2. Condición para el efecto descendente: $Q^* < Q^{\text{EOQ}}$

El efecto descendente ocurre cuando la solución estocástica óptima es **menor** que la solución EOQ, es decir:

$$Q^* < \sqrt{\frac{2K\mu_D}{h}}.$$

Este fenómeno es posible únicamente si: 1. El costo fijo K en el EOQ es **artificialmente inflado** (no refleja un costo real de transacción), y/o

2. El costo unitario c en el modelo extendido es **significativo**, y la penalización ϕ no es tan severa como para dominar el término cQ .

Más formalmente, supongamos $\phi(z) = pz$ (lineal). Entonces el modelo extendido se reduce al *newsvendor*, cuyo óptimo satisface:

$$F_D(Q^*) = \frac{p - c}{p + h - c} =: \beta_{\text{NV}}.$$

Si la distribución de \tilde{D} es simétrica (e.g., normal), entonces $Q^* \approx \mu_D + z_{\beta_{\text{NV}}} \sigma_D$. Si $\beta_{\text{NV}} < 0.5$ (es decir, si $c > p - h$), entonces $z_{\beta_{\text{NV}}} < 0$, y por tanto $Q^* < \mu_D$.

Mientras tanto, $Q^{\text{EOQ}} = \sqrt{2K\mu_D/h}$ puede ser arbitrariamente grande si K se elige sin fundamento. Pero incluso con K pequeño, si c es alto y p moderado, Q^* puede ser menor que μ_D , y por tanto menor que Q^{EOQ} si este último supera μ_D .

El caso más revelador ocurre cuando $K \rightarrow 0$: entonces $Q^{\text{EOQ}} \rightarrow 0$, pero esto es engañoso. En realidad, si no hay costo fijo, el EOQ pierde sentido; el modelo correcto es el de un solo periodo, y Q^* permanece positivo. La comparación sólo es significativa cuando K se interpreta como un parámetro de ajuste.

5.4.5.3. Ejemplo de demostración del efecto descendente

Consideremos los siguientes parámetros:

- $\tilde{D} \sim \text{Normal}(\mu_D = 1000, \sigma_D = 300)$,
- $c = 20$ unidades monetarias/unidad,
- $h = 2$ unidades monetarias/unidad,
- $p = 25$ unidades monetarias/unidad (penalización lineal),
- $K = 1000$ unidades monetarias (costo fijo “administrativo” asignado arbitrariamente).

El tamaño óptimo del pedido bajo el modelo EOQ es:

$$Q^{\text{EOQ}} = \sqrt{\frac{2 \cdot K \cdot \mu_D}{h}} = \sqrt{\frac{2 \cdot 1000 \cdot 1000}{2}} = \sqrt{1\,000\,000} = 1000.$$

Sin embargo, en contextos donde el costo unitario es más alto y la penalización por faltantes es apenas marginal, por ejemplo, al considerar $c = 24$, $p = 25$ y un costo fijo inflado por burocracia percibida ($K = 5000$), el modelo EOQ arroja un tamaño de lote considerablemente mayor:

$$Q^{\text{EOQ}} = \sqrt{\frac{2 \cdot 5000 \cdot 1000}{2}} = \sqrt{5\,000\,000} \approx 2236.$$

En contraste, el modelo *newsvendor*, que incorpora explícitamente la incertidumbre en la demanda, produce una solución más conservadora. El cuantil óptimo es:

$$\beta_{\text{NV}} = \frac{p - c}{p + h - c} = \frac{25 - 24}{25 + 2 - 24} = \frac{1}{3} \approx 0.333,$$

lo que implica $z_{0.333} \approx -0.43$ y, por tanto,

$$Q^* = \mu_D + z_{\beta} \cdot \sigma_D = 1000 - 0.43 \cdot 300 \approx 871.$$

Así, se observa claramente el **efecto descendente**:

$$Q^* \approx 871 < Q^{\text{EOQ}} \approx 2236,$$

lo que ilustra que, bajo alta incertidumbre y baja penalización marginal, el modelo estocástico recomienda pedidos **más pequeños** que el modelo determinista, incluso cuando este último sugiere un lote grande por el alto costo fijo.

Interpretación: Aunque el EOQ sugiere almacenar más de dos veces la demanda media, el modelo estocástico que reconoce que el costo marginal de comprar una unidad adicional ($c = 24$) es casi tan alto como la penalización por no satisfacerla ($p = 25$) prefiere aceptar faltantes frecuentes a incurrir en altos costos de compra e inventario. En contextos humanitarios, este escenario es poco común (pues $p \gg c$ usualmente), pero **ilustra que el EOQ no es un límite inferior universal**: su validez depende críticamente de la correcta especificación de K , y su uso con parámetros no calibrados puede llevar a decisiones excesivamente conservadoras en inventario, con implicaciones de costo y espacio de almacenamiento innecesarias.

5.4.5.4. Caso con penalización no lineal

Si $\phi(z) = z^2$, el efecto descendente es menos probable, pues $\phi'(z) \rightarrow \infty$, lo que impulsa Q^* hacia valores altos. Sin embargo, si c es muy grande y la varianza de \tilde{D} es baja, aún puede ocurrir. Por ejemplo, con $\tilde{D} \sim \text{Uniforme}(950, 1050)$, $c = 100$, $h = 1$, $\phi(z) = z^2$, se obtiene numéricamente $Q^* \approx 990 < \mu_D = 1000$, mientras que un EOQ con $K = 500$ da $Q^{\text{EOQ}} \approx \sqrt{2 \cdot 500 \cdot 1000/1} \approx 1000$.

5.4.5.5. Conclusión cuantitativa

El efecto descendente en Q^* no es una contradicción, sino una **advertencia metodológica**: el modelo EOQ, al depender de un parámetro K que carece de anclaje en contextos de un solo periodo, puede generar recomendaciones arbitrarias si se aplica fuera de su dominio de validez. La comparación cuantitativa demuestra que:

- Cuando K es pequeño y c es alto, $Q^* < Q^{\text{EOQ}}$ es posible.
- En logística humanitaria, donde $p \gg c$ y la incertidumbre es alta, lo habitual es $Q^* > Q^{\text{EOQ}}$.
- **Nunca debe usarse el EOQ como referencia estática**; su valor comparativo depende de la calibración cuidadosa de K y de la alineación del modelo con la estructura temporal del problema.

Este análisis refuerza la necesidad de adoptar formulaciones estocásticas rigurosas como la propuesta en las secciones anteriores que derivan Q^* de principios de equilibrio marginal bajo incertidumbre, no de analogías mecánicas con modelos deterministas de horizonte infinito.

5.5. Propiedades analíticas del problema unidimensional

Tras formular la función de costo total extendida en la Sección 5.4 se deriva su expresión determinista efectiva en la presente sección, estableciendo las propiedades analíticas fundamentales del problema de optimización unidimensional resultante. En esta sección, se analiza rigurosamente el problema $\min_{Q \geq 1} f(Q)$, demostrando la existencia y unicidad del minimizador, así como el comportamiento asintótico de la función objetivo en los extremos del dominio. Estas propiedades son esenciales para justificar la aplicabilidad de las condiciones de optimalidad clásicas y para garantizar la estabilidad de los métodos numéricos empleados en la resolución.

5.5.1. Formulación del problema de optimización: minimización de $f(Q)$ sujeto a $Q \geq 1$

En el contexto del modelo de inventario extendido presentado en la Sección 5.1.2, la determinación del tamaño óptimo de lote se reduce a un problema de optimización escalar en la variable continua $Q > 0$. Tras aislar los términos que dependen explícitamente de Q en la función de costo total extendida (Ecuación 5.5.1), se obtiene la función objetivo efectiva:

$$f(Q) := \frac{K}{Q} + aQ, \quad Q > 0, \quad (5.5.1)$$

donde hemos definido

$$K := D(S + \beta C) \quad \text{y} \quad a := \frac{H}{2}.$$

Bajo las hipótesis operativas estándar, demanda media $D > 0$, costos fijos y unitarios no negativos ($S \geq 0$, $C \geq 0$), penalización por faltante estrictamente positiva ($\beta > 0$) y costo de mantenimiento $H > 0$, se garantiza que $K > 0$ y $a > 0$. En consecuencia, $f(Q)$ está bien definida y es diferenciable en $(0, \infty)$.

En logística humanitaria, la decisión de preposicionar inventario tiene sentido únicamente si se almacena **al menos una unidad** del bien esencial, lo que impone la restricción natural:

$$Q \geq 1.$$

Esta cota inferior no es meramente técnica: refleja la **discreción física de los bienes humanitarios** (no se puede distribuir una fracción arbitrariamente pequeña de un kit de emergencia) y evita singularidades en el análisis asintótico cuando $Q \rightarrow 0^+$, las cuales carecen de interpretación operativa.

En consecuencia, el problema de optimización unidimensional que se analiza en esta sección se formula como:

$$\min_{Q \in [1, \infty)} f(Q) = \min_{Q \geq 1} \left(\frac{K}{Q} + aQ \right). \quad (5.5.2)$$

Este problema es un caso particular de minimización de una función real sobre un conjunto cerrado y no acotado. Su análisis requiere establecer: 1. **Existencia** de un mínimo global (garantizada por la coercividad y continuidad de f), 2. **Unicidad** del minimizador (derivada de la estricta convexidad local en el dominio relevante), 3. **Comportamiento asintótico** en los extremos del dominio ($Q \rightarrow 1^+$ y $Q \rightarrow \infty$).

Estas propiedades se desarrollan en las subsecciones siguientes, sustentadas en resultados clásicos del análisis no lineal, tal como se presentan en Bazaraa, Sherali, y Shetty (2013b); Boyd y Vandenberghe (2004).

5.5.1.1. Ejemplo ilustrativo: formulación explícita del problema

Retomando los parámetros del ejemplo de la Sección 5.1.2:

- $D = 1000$,
- $S = 100$,
- $H = 5$,
- $C = 10$,
- $\beta = 50$,

calculamos:

$$K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000, \quad a = \frac{H}{2} = 2.5.$$

La función objetivo efectiva es:

$$f(Q) = \frac{600\,000}{Q} + 2.5Q,$$

y el problema de optimización se escribe como:

$$\min_{Q \geq 1} \left(\frac{600\,000}{Q} + 2.5Q \right). \quad (5.5.3)$$

Este problema es bien planteado: la función f es continua en $[1, \infty)$, tiende a infinito cuando $Q \rightarrow \infty$ y tiene un valor finito (aunque elevado) en $Q = 1$ ($f(1) = 600\,002.5$), y posee un único punto crítico en $(1, \infty)$, como se demostrará en las secciones siguientes.

La formulación Ecuación 5.5.2 constituye la base analítica para el estudio riguroso de las propiedades del modelo de inventario extendido. Al restringir el dominio a $Q \geq 1$, se garantiza la relevancia práctica de las soluciones, al tiempo que se preserva la trazabilidad matemática necesaria para el análisis de optimalidad y sensibilidad que se desarrollará en las secciones posteriores.

5.5.2. Continuidad y coercividad: garantía de existencia del mínimo

En la Sección 5.5.1 se formuló el problema de optimización unidimensional como:

$$\min_{Q \geq 1} f(Q) = \min_{Q \geq 1} \left(\frac{K}{Q} + aQ \right), \quad (5.5.4)$$

donde $K = D(S + \beta C) > 0$ y $a = H/2 > 0$. Dado que el dominio $[1, \infty)$ es no acotado, la existencia de un minimizador global no es automática y requiere condiciones analíticas adicionales. En particular, se requiere que la función objetivo sea **continua** y **coerciva** en su dominio. A continuación, se establecen formalmente estas propiedades.

Lema (Existencia del mínimo global).

La función $f : [1, \infty) \rightarrow \mathbb{R}$, definida por

$$f(Q) = \frac{K}{Q} + aQ,$$

con $K > 0$ y $a > 0$, es continua y coerciva en $[1, \infty)$. En consecuencia, el problema $\min_{Q \geq 1} f(Q)$ admite al menos un minimizador global $Q^* \in [1, \infty)$.

Demostración.

1. Continuidad. La función f es la suma de $Q \mapsto K/Q$ y $Q \mapsto aQ$, ambas continuas en $(0, \infty)$, y por tanto en $[1, \infty)$.

2. Coercividad. Se verifica que

$$\lim_{Q \rightarrow \infty} f(Q) = \lim_{Q \rightarrow \infty} \left(\frac{K}{Q} + aQ \right) = +\infty, \quad (5.5.5)$$

ya que $a > 0$. Además, $f(1) = K + a < \infty$, por lo que f está acotada inferiormente en $[1, \infty)$.

3. Existencia. El conjunto $[1, \infty)$ es cerrado y no vacío. Dado que f es continua y coerciva, el teorema de Weierstrass garantiza que f alcanza su mínimo global en dicho conjunto.

Este lema asegura que, bajo los parámetros operativos estándar de logística humanitaria, el problema bien planteado $\min_{Q \geq 1} f(Q)$ **siempre admite una solución óptima**, lo cual justifica el uso posterior de condiciones de optimalidad clásicas y el análisis de sensibilidad.

5.5.2.1. Ejemplo ilustrativo

Considérense los parámetros del ejemplo de la Sección 5.1.2: - $D = 1000$, - $S = 100$, - $H = 5$, - $C = 10$, - $\beta = 50$.

Entonces:

$$K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000, \quad a = \frac{H}{2} = 2.5.$$

La función objetivo es:

$$f(Q) = \frac{600\,000}{Q} + 2.5Q, \quad Q \geq 1.$$

- **Continuidad:** Ambos términos son continuos en $[1, \infty)$.
- **Coercividad:** Por Ecuación 5.5.5, $\lim_{Q \rightarrow \infty} f(Q) = \infty$.
- **Valor en el borde:** $f(1) = 600\,002.5$.

Por el Lema (Existencia del mínimo global), existe al menos un $Q^* \in [1, \infty)$ que minimiza $f(Q)$. En la Sección 6.5.3 se demostrará que este minimizador es único y se encuentra en el interior del dominio, ya que $Q^* \approx 489.9 > 1$.

Este ejemplo confirma que, bajo parámetros realistas, el problema satisface las condiciones del lema y, por tanto, es **bien planteado en el sentido de Hadamard**: admite solución, que además es estable y única (propiedades que se desarrollarán a continuación).

5.5.3. Análisis de convexidad local y unicidad del óptimo

En la Sección 5.5.2 se demostró que el problema de optimización unidimensional

$$\min_{Q \geq 1} f(Q) = \min_{Q \geq 1} \left(\frac{K}{Q} + aQ \right), \quad (5.5.6)$$

con $K = D(S + \beta C) > 0$ y $a = H/2 > 0$, admite al menos un minimizador global. A continuación, se analiza la **estructura local de la función objetivo** para establecer la **unicidad** de dicho minimizador. Esta propiedad es fundamental para garantizar que las condiciones de optimalidad de primer orden identifiquen un único candidato a solución óptima.

5.5.3.1. Convexidad local de $f(Q)$

La función $f : (0, \infty) \rightarrow \mathbb{R}$, definida por

$$f(Q) = \frac{K}{Q} + aQ, \quad (5.5.7)$$

es infinitamente diferenciable en su dominio. Sus derivadas de primer y segundo orden son:

$$f'(Q) = -\frac{K}{Q^2} + a, \quad f''(Q) = \frac{2K}{Q^3}. \quad (5.5.8)$$

Dado que $K > 0$ y $Q > 0$, se tiene $f''(Q) > 0$ para todo $Q \in (0, \infty)$. Por lo tanto, $f(Q)$ es **estrictamente convexa** en todo su dominio.

La estricta convexidad de f implica que cualquier punto estacionario de f en $(0, \infty)$ es el **único minimizador global** de f en ese intervalo.

5.5.3.2. Unicidad del minimizador en el dominio restringido $[1, \infty)$

Aunque f es estrictamente convexa en $(0, \infty)$, el problema de interés está restringido a $Q \geq 1$. Para analizar la unicidad en este dominio, consideramos dos casos exclusivos:

1. **Caso interior:** si el punto crítico $Q^c = \sqrt{K/a}$ satisface $Q^c > 1$, entonces $Q^c \in (1, \infty)$ y, por la estricta convexidad, es el **único minimizador** de f en $[1, \infty)$.
2. **Caso de frontera:** si $Q^c \leq 1$, entonces $f'(Q) > 0$ para todo $Q \geq 1$, lo que implica que f es estrictamente creciente en $[1, \infty)$, y el **único minimizador** es $Q^* = 1$.

En ambos casos, el minimizador global en $[1, \infty)$ **existe y es único**.

Este resultado puede resumirse en la siguiente proposición.

Proposición (Unicidad del minimizador).

La función $f(Q) = \frac{K}{Q} + aQ$, con $K > 0$ y $a > 0$, posee un **único minimizador global** en el conjunto $[1, \infty)$, dado por:

$$Q^* = \max \left\{ 1, \sqrt{\frac{K}{a}} \right\}. \quad (5.5.9)$$

La unicidad garantiza que el problema Ecuación 5.5.9 es **bien planteado en el sentido de Hadamard** y que cualquier método de optimización que converja a un punto estacionario o a la frontera del dominio identificará la solución óptima.

5.5.3.3. Ejemplo ilustrativo

Retomamos los parámetros del ejemplo de la Sección 5.1.2: - $D = 1000$, - $S = 100$, - $H = 5$, - $C = 10$, - $\beta = 50$.

Calculamos:

$$K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000, \quad a = \frac{H}{2} = 2.5.$$

El punto crítico no restringido es:

$$Q^c = \sqrt{\frac{K}{a}} = \sqrt{\frac{600\,000}{2.5}} = \sqrt{240\,000} \approx 489.90.$$

Dado que $Q^c \approx 489.90 > 1$, el minimizador en $[1, \infty)$ es:

$$Q^* = Q^c \approx 489.90.$$

Verificación de convexidad local:

La segunda derivada en Q^* es:

$$f''(Q^*) = \frac{2K}{(Q^*)^3} = \frac{2 \cdot 600\,000}{(489.90)^3} \approx \frac{1\,200\,000}{117\,577\,000} \approx 0.0102 > 0,$$

lo que confirma que Q^* es un mínimo local estricto. Por la convexidad global de f , también es el mínimo global.

Interpretación: En este caso realista, el tamaño óptimo de lote es significativamente mayor que 1, lo que justifica que el mínimo ocurra en el interior del dominio. La unicidad asegura que no existen otras políticas alternativas con el mismo costo óptimo, lo cual es deseable desde una perspectiva de implementación operativa.

Este análisis demuestra que, bajo los supuestos operativos estándar de logística humanitaria, el problema de optimización unidimensional no solo admite solución, sino que dicha solución es **única y estable**, propiedades esenciales para la robustez del modelo y la confiabilidad de los métodos numéricos empleados en su resolución.

5.5.4. Comportamiento asintótico de $f(Q)$ cuando $Q \rightarrow 1^+$ y $Q \rightarrow \infty$

El análisis del comportamiento asintótico de la función objetivo $f(Q)$ en los extremos del dominio $[1, \infty)$ es fundamental para comprender la estructura global del problema de optimización y para justificar propiedades como **coercividad**, **acotación inferior**, y la **localización del minimizador** en el interior o la frontera del dominio. En esta sección, se estudian rigurosamente los límites de $f(Q)$ y sus derivadas cuando Q tiende al borde inferior 1^+ y al infinito.

Recordemos que la función objetivo efectiva, definida en la Sección 5.5.1, es:

$$f(Q) = \frac{K}{Q} + aQ, \quad Q \geq 1, \quad (5.5.10)$$

con $K = D(S + \beta C) > 0$ y $a = H/2 > 0$. Esta función es suave en $(1, \infty)$ y continua en $[1, \infty)$.

5.5.4.1. Comportamiento cuando $Q \rightarrow 1^+$

Al aproximarse al borde inferior del dominio, la función $f(Q)$ converge a un valor finito:

$$\lim_{Q \rightarrow 1^+} f(Q) = f(1) = K + a. \quad (5.5.11)$$

Dado que $K, a > 0$, este límite es **estrictamente positivo y finito**. La derivada de primer orden en este punto es:

$$f'(1) = -\frac{K}{1^2} + a = a - K. \quad (5.5.12)$$

Este valor determina si la función es creciente o decreciente inmediatamente a la derecha de $Q = 1$: - Si $K > a$, entonces $f'(1) < 0$: la función **decrece** al salir de la frontera, lo que implica que el minimizador **no está en la frontera**, sino en el interior $(1, \infty)$. - Si $K \leq a$, entonces $f'(1) \geq 0$: la función es no decreciente en $[1, \infty)$, y el **mínimo global ocurre en $Q^* = 1$** .

5.5.4.2. Comportamiento cuando $Q \rightarrow \infty$

Cuando Q crece sin cota, el comportamiento de $f(Q)$ está dominado por el término lineal aQ , ya que el término K/Q tiende a cero:

$$\lim_{Q \rightarrow \infty} f(Q) = \lim_{Q \rightarrow \infty} \left(\frac{K}{Q} + aQ \right) = +\infty. \quad (5.5.13)$$

Este comportamiento garantiza que la función es **coerciva** en $[1, \infty)$, lo cual, combinado con su continuidad, asegura (junto con el teorema de Weierstrass) la **existencia de un mínimo global**, como se demostró en la Sección 5.5.2.

La derivada de primer orden también diverge positivamente:

$$\lim_{Q \rightarrow \infty} f'(Q) = \lim_{Q \rightarrow \infty} \left(-\frac{K}{Q^2} + a \right) = a > 0, \quad (5.5.14)$$

lo que confirma que $f(Q)$ es **estrictamente creciente** para Q suficientemente grande.

5.5.4.3. Ejemplo ilustrativo

Retomamos los parámetros del ejemplo de la Sección 5.1.2: - $D = 1000$, - $S = 100$, - $H = 5$, - $C = 10$, - $\beta = 50$.

Entonces:

$$K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000, \quad a = \frac{H}{2} = 2.5.$$

1. Comportamiento en $Q \rightarrow 1^+$:

- $f(1) = 600\,000 + 2.5 = 600\,002.5$,
- $f'(1) = 2.5 - 600\,000 = -599\,997.5 < 0$.

Interpretación: la función **decrece abruptamente** al salir de $Q = 1$, lo que confirma que el mínimo no puede estar en la frontera.

2. Comportamiento en $Q \rightarrow \infty$:

- $\lim_{Q \rightarrow \infty} f(Q) = \infty$,
- $\lim_{Q \rightarrow \infty} f'(Q) = 2.5 > 0$.

Interpretación: para valores grandes de Q (por ejemplo, $Q > 10\,000$), el costo de mantenimiento domina, y la función crece linealmente con pendiente 2.5.

3. Localización del mínimo:

El punto crítico es $Q^c = \sqrt{K/a} \approx 489.9$, como se calculó anteriormente. Dado que $f'(1) < 0$ y $f'(Q) \rightarrow a > 0$, la función **decrece hasta Q^c y luego crece**, confirmando que Q^c es el **único mínimo global** en $(1, \infty)$.

Este análisis asintótico no solo valida las propiedades de existencia y unicidad, sino que también **informa el diseño de algoritmos numéricos**: por ejemplo, un método de búsqueda unidimensional puede iniciar en $Q = 1$ y expandirse hasta que $f'(Q) > 0$, garantizando que el mínimo ya fue cruzado.

5.6. Condiciones de optimalidad y solución cerrada aproximada

A partir de las propiedades analíticas establecidas en la Sección 5.5, este apartado deriva formalmente las condiciones de optimalidad de primer orden para el problema unidimensional

$$\min_{Q \geq 1} Z(Q).$$

Se obtiene una solución cerrada aproximada para Q^* , se discute la proyección del óptimo ante la restricción $Q \geq 1$, y se analiza la sensibilidad de la solución respecto a los parámetros críticos del modelo.

5.6.1. Derivación formal de $\frac{dZ}{dQ}$

Consideramos la función de costo total extendida introducida en la Sección 5.1.2:

$$Z(Q) = \frac{D}{Q}S + \left(\frac{Q}{2} + Z_\alpha \sigma_L\right)H + DC + \frac{D}{Q}\beta C, \quad (5.6.1)$$

que corresponde a la ecuación Ecuación 5.1.1 en este documento. Aunque los parámetros Z_α y σ_L se derivan de un análisis estocástico (stock de seguridad bajo demanda aleatoria), **el modelo de optimización es determinista**, conforme al enfoque robusto adoptado en este trabajo (ver Sección 5.1.2). En consecuencia, $Z(Q)$ se trata como una función real diferenciable en $(0, \infty)$, cuyos parámetros reflejan una parametrización de la incertidumbre, pero no su modelación explícita.

En el análisis que sigue, suponemos que σ_L es constante (es decir, independiente de Q). Esta hipótesis es común en formulaciones iniciales de inventario bajo incertidumbre y permite preservar cierto grado de tratabilidad analítica, aun cuando se incorporen penalizaciones explícitas por faltante. Bajo esta condición, la función $Z(Q)$ es diferenciable en $(0, \infty)$ y su derivada puede obtenerse aplicando reglas elementales del cálculo.

Agrupando los términos proporcionales a $1/Q$, reescribimos la parte no constante de la ecuación Ecuación 5.1.1 como:

$$f(Q) := \frac{D(S + \beta C)}{Q} + \frac{H}{2}Q, \quad (5.6.2)$$

donde los términos $Z_\alpha \sigma_L H + DC$ son independientes de Q y, por tanto, no influyen en la localización del mínimo. Definimos por conveniencia:

$$K := D(S + \beta C) > 0, \quad a := \frac{H}{2} > 0,$$

de modo que $f(Q) = \frac{K}{Q} + aQ$.

La derivada de $Z(Q)$ con respecto a Q coincide con la derivada de $f(Q)$. Aplicando la regla de derivación para potencias, obtenemos:

$$\frac{dZ}{dQ}(Q) = -\frac{D(S + \beta C)}{Q^2} + \frac{H}{2}. \quad (5.6.3)$$

Esta expresión es válida para todo $Q > 0$ y constituye la derivada primera de la función de costo total extendida bajo las hipótesis establecidas.

5.6.1.1. Ejemplo ilustrativo

Para ilustrar la derivación anterior, consideremos los siguientes parámetros numéricos tomados de la Sección 5.1.2:

- $D = 1000$,
- $S = 100$,
- $H = 5$,
- $C = 10$,
- $\beta = 50$,
- $Z_\alpha \sigma_L H + DC = 10025.99$ (constante, no relevante para la derivada).

Paso 1. Calcular $K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000$.

Paso 2. Calcular $a = H/2 = 5/2 = 2.5$.

Paso 3. Escribir la función objetivo no constante:

$$f(Q) = \frac{600\,000}{Q} + 2.5Q.$$

Paso 4. Derivar:

$$\frac{df}{dQ}(Q) = -\frac{600\,000}{Q^2} + 2.5. \quad (5.6.4)$$

Este resultado coincide con la fórmula general Ecuación 5.6.3 y será utilizado en la siguiente sección para determinar el punto crítico Q^* .

5.6.2. Condición de primer orden: $\frac{dZ}{dQ}(Q^*) = 0$

En la Sección 5.6.1 se estableció que la función de costo total extendida, bajo el enfoque determinista robusto adoptado en este trabajo, puede expresarse como

$$Z(Q) = \frac{D(S + \beta C)}{Q} + \frac{H}{2}Q + Z_\alpha \sigma_L H + DC, \quad (5.6.5)$$

donde ($Q > 0$) es la variable de decisión continua que representa el tamaño del lote de inventario. Los dos últimos términos de la ecuación anterior son constantes con respecto a Q y, por tanto, no influyen en la localización del mínimo. Denotando nuevamente

$$K := D(S + \beta C) > 0, \quad a := \frac{H}{2} > 0,$$

la parte relevante de la función objetivo Ecuación 5.5.1 se reduce a

$$f(Q) = \frac{K}{Q} + aQ, \quad Q > 0. \quad (5.6.6)$$

Como se demostró en la Sección 5.5.3, $f(Q)$ es estrictamente convexa en $(0, \infty)$. En consecuencia, posee un único minimizador global en cualquier subconjunto convexo de su dominio. En particular, en el dominio restringido $[1, \infty)$, el minimizador existe y es único (Sección 5.5.2 y Sección 5.5.3).

Dado que f es diferenciable en $(0, \infty)$, una **condición necesaria y suficiente** para que $Q^* \in (1, \infty)$ sea el minimizador global es que anule la derivada de primer orden:

$$\frac{dZ}{dQ}(Q^*) = f'(Q^*) = 0. \quad (5.6.7)$$

Sustituyendo la expresión de la derivada obtenida en la Sección 5.6.1, esta condición se escribe explícitamente como:

$$-\frac{D(S + \beta C)}{(Q^*)^2} + \frac{H}{2} = 0. \quad (5.6.8)$$

Esta ecuación constituye la **condición de primer orden** del problema de optimización unidimensional. Su resolución proporciona el punto crítico que, por la estricta convexidad de f , corresponde al único minimizador global del problema en el interior del dominio.

5.6.2.1. Ejemplo ilustrativo

Retomamos los parámetros numéricos del ejemplo de la Sección 5.1.2:

- $D = 1000$,
- $S = 100$,
- $H = 5$,
- $C = 10$,
- $\beta = 50$.

Paso 1. Calcular los parámetros efectivos:

$$K = D(S + \beta C) = 1000(100 + 50 \cdot 10) = 600\,000, \quad a = \frac{H}{2} = 2.5.$$

Paso 2. Escribir la condición de primer orden:

$$-\frac{600\,000}{(Q^*)^2} + 2.5 = 0.$$

Paso 3. Resolver algebraicamente para Q^* :

$$\frac{600\,000}{(Q^*)^2} = 2.5 \implies (Q^*)^2 = \frac{600\,000}{2.5} = 240\,000 \implies Q^* = \sqrt{240\,000} \approx 489.90.$$

Paso 4. Verificación: La derivada de la función objetivo es

$$\frac{dZ}{dQ}(Q) = -\frac{600\,000}{Q^2} + 2.5.$$

Evaluando en $Q^* \approx 489.90$:

$$\frac{dZ}{dQ}(489.90) \approx -\frac{600\,000}{(489.90)^2} + 2.5 \approx -\frac{600\,000}{240\,000} + 2.5 = -2.5 + 2.5 = 0.$$

Paso 5. Interpretación: El punto $Q^* \approx 489.90$ es el **único valor** que equilibra marginalmente el costo decreciente de emitir menos pedidos (incluyendo la penalización esperada por faltante) con el costo creciente de mantener inventario adicional. Dado que $Q^* > 1$, la solución óptima reside en el **interior del dominio**, y la restricción $Q \geq 1$ no es activa.

Este resultado numérico refuerza la conclusión teórica: bajo los supuestos operativos estándar en logística humanitaria, el óptimo del modelo extendido se alcanza en el interior del dominio y puede identificarse de forma analítica mediante la condición de primer orden Ecuación 5.6.8. La unicidad y existencia garantizadas en la Sección 5.5 aseguran que este punto crítico es, de hecho, el **único minimizador global** del problema.

5.6.3. Expresión explícita de Q^* sin restricción de caja

En las secciones anteriores se estableció que, bajo el enfoque determinista robusto adoptado en este trabajo, la función de costo total extendida se reduce a una función escalar de la forma

$$Z(Q) = \frac{K}{Q} + aQ + \text{constante}, \quad (5.6.9)$$

donde $K = D(S + \beta C) > 0$ y $a = H/2 > 0$. El problema de optimización asociado, en ausencia de la restricción de caja $Q \geq 1$, se formula como:

$$\min_{Q>0} f(Q) = \min_{Q>0} \left(\frac{K}{Q} + aQ \right).$$

Este problema es un caso clásico de optimización unidimensional con solución cerrada, ampliamente estudiado en la teoría de inventarios y en textos de programación no lineal (ver Boyd y Vandenberghe (2004); Bazaraa, Sherali, y Shetty (2013b)).

Dado que $f(Q)$ es **estrictamente convexa** en $(0, \infty)$ (pues $f''(Q) = 2K/Q^3 > 0$ para todo $Q > 0$), el único punto estacionario que satisface la condición de primer orden $f'(Q^*) = 0$ corresponde al **mínimo global** del problema no restringido.

La derivada de primer orden, obtenida en la Sección 5.6.1, es:

$$f'(Q) = -\frac{K}{Q^2} + a.$$

Igualando a cero y resolviendo para Q , obtenemos:

$$-\frac{K}{(Q^*)^2} + a = 0 \implies (Q^*)^2 = \frac{K}{a} \implies Q^* = \sqrt{\frac{K}{a}}.$$

Sustituyendo los valores de K y a , la **expresión explícita del tamaño óptimo de lote sin restricción de caja** es:

$$Q^* = \sqrt{\frac{2D(S + \beta C)}{H}}. \quad (5.6.10)$$

Esta expresión generaliza la clásica fórmula del modelo EOQ, ya que el término $S + \beta C$ incorpora explícitamente el **costo fijo efectivo por ciclo**, que ahora incluye tanto el costo administrativo de emisión de un pedido (S) como la **penalización esperada por faltante** (βC). Este ajuste refleja la internalización del riesgo en la estructura de costos, una característica esencial en contextos humanitarios donde la insatisfacción de la demanda tiene consecuencias operativas no despreciables.

La existencia de $Q^* > 0$ está garantizada por la positividad de todos los parámetros del modelo. Además, la **unicidad** se deriva directamente de la estricta convexidad de $f(Q)$, como se demostró en la Sección 5.5.3.

5.6.3.1. Ejemplo ilustrativo

Retomamos los parámetros del ejemplo de la Sección Sección 5.1.2:

- Demanda esperada: $D = 1000$,
- Costo fijo de pedido: $S = 100$,
- Costo de mantenimiento: $H = 5$,
- Costo unitario de adquisición: $C = 10$,
- Penalización por faltante: $\beta = 50$.

Paso 1. Calcular el costo fijo efectivo:

$$S + \beta C = 100 + 50 \cdot 10 = 600.$$

Paso 2. Sustituir en la expresión (Ecuación 5.6.10):

$$Q^* = \sqrt{\frac{2 \cdot 1000 \cdot 600}{5}} = \sqrt{\frac{1\,200\,000}{5}} = \sqrt{240\,000} \approx 489.90.$$

Paso 3. Comparación con el EOQ clásico: El modelo EOQ tradicional (ignorando penalización) daría:

$$Q^{\text{EOQ}} = \sqrt{\frac{2DS}{H}} = \sqrt{\frac{2 \cdot 1000 \cdot 100}{5}} = \sqrt{40\,000} = 200.$$

Paso 4. Interpretación: El tamaño óptimo en el modelo extendido ($Q^* \approx 490$) es **más del doble** que el del EOQ clásico. Esta diferencia cuantifica el impacto de incorporar la penalización por faltante en la política de inventario: el modelo responde a la aversión al riesgo aumentando el lote de preposicionamiento, lo cual reduce la frecuencia de exposición a escenarios de desabastecimiento.

Este ejemplo confirma que, en ausencia de restricciones de caja, el problema admite una **solución cerrada analítica**, cuya estructura es directamente interpretable en términos de los parámetros operativos del sistema. Esta propiedad es fundamental para el análisis de sensibilidad y para la validación de algoritmos numéricos en secciones posteriores.

5.6.4. Tratamiento de la restricción $Q \geq 1$: proyección del óptimo

En las secciones anteriores se obtuvo la expresión explícita del tamaño óptimo de lote en ausencia de restricciones de caja:

$$Q^{\text{nc}} = \sqrt{\frac{2D(S + \beta C)}{H}}, \quad (5.6.11)$$

donde el superíndice “nc” denota “sin caja” (*no constraint*). Sin embargo, en contextos de logística humanitaria, la decisión de preposicionar inventario solo tiene sentido si se almacena **al menos una unidad** del bien esencial, lo que impone la restricción natural:

$$Q \geq 1.$$

Este límite inferior, aunque aparentemente técnico, refleja una **discreción física operativa**: no es factible ni significativo preposicionar una fracción arbitrariamente pequeña de un kit de emergencia. Matemáticamente, esta restricción convierte el dominio de optimización en el conjunto cerrado $[1, \infty)$, lo que requiere ajustar la solución no restringida mediante una **proyección ortogonal** sobre dicho conjunto.

Dado que la función objetivo $f(Q) = \frac{K}{Q} + aQ$ (con $K = D(S + \beta C) > 0$, $a = H/2 > 0$) es estrictamente convexa y diferenciable en $(0, \infty)$, y el conjunto $[1, \infty)$ es convexo y cerrado, el **minimizador restringido** Q^* se obtiene mediante la proyección euclidiana de Q^{nc} sobre $[1, \infty)$:

$$Q^* = \Pi_{[1, \infty)}(Q^{\text{nc}}) = \max\{1, Q^{\text{nc}}\}. \quad (5.6.12)$$

Esta fórmula establece dos casos mutuamente excluyentes: 1. **Caso interior**: si $Q^{\text{nc}} > 1$, entonces $Q^* = Q^{\text{nc}}$, y la restricción no es activa. La condición de primer orden $f'(Q^*) = 0$ se satisface con igualdad. 2. **Caso de frontera**: si $Q^{\text{nc}} \leq 1$, entonces $Q^* = 1$, y la restricción es activa. En este caso, $f'(Q^*) = f'(1) \geq 0$, lo que confirma que la función es no decreciente en $[1, \infty)$, y el mínimo global ocurre en el borde del dominio.

La proyección (Ecuación 5.6.12) garantiza que la solución óptima sea **factible y única**, preservando las propiedades de existencia y unicidad demostradas en la Sección 5.5.

5.6.4.1. Ejemplo ilustrativo

Retomamos los parámetros del ejemplo de la Sección 5.1.2:

- Demanda esperada: $D = 1000$,
- Costo fijo de pedido: $S = 100$,
- Costo de mantenimiento: $H = 5$,

- Costo unitario de adquisición: $C = 10$,
- Penalización por faltante: $\beta = 50$.

Paso 1. Calcular el óptimo no restringido:

$$Q^{\text{nc}} = \sqrt{\frac{2 \cdot 1000 \cdot (100 + 50 \cdot 10)}{5}} = \sqrt{\frac{2 \cdot 1000 \cdot 600}{5}} = \sqrt{240\,000} \approx 489.90.$$

Paso 2. Aplicar la proyección:

$$Q^* = \text{máx}\{1, 489.90\} = 489.90.$$

Interpretación: dado que $Q^{\text{nc}} \gg 1$, la restricción $Q \geq 1$ no es activa, y la solución restringida coincide con la no restringida. Este resultado es típico en contextos humanitarios, donde los costos de faltante y la magnitud de la demanda inducen niveles de inventario significativamente mayores que una unidad.

Paso 3. Escenario hipotético con demanda extremadamente baja: Supongamos ahora $D = 0.1$, manteniendo los demás parámetros. Entonces:

$$Q^{\text{nc}} = \sqrt{\frac{2 \cdot 0.1 \cdot 600}{5}} = \sqrt{24} \approx 4.90 > 1 \quad \Rightarrow \quad Q^* = 4.90.$$

Incluso con demanda muy baja, el óptimo sigue siendo mayor que 1.

Paso 4. Escenario límite: Para que $Q^{\text{nc}} \leq 1$, se requiere:

$$\sqrt{\frac{2D(S + \beta C)}{H}} \leq 1 \quad \Rightarrow \quad D \leq \frac{H}{2(S + \beta C)} = \frac{5}{2 \cdot 600} \approx 0.00417.$$

Solo si la demanda esperada es inferior a 0.0042 unidades, el óptimo restringido sería $Q^* = 1$. En la práctica humanitaria, este caso es irreal, ya que la demanda mínima razonable supera con creces una unidad.

Este análisis confirma que, bajo parámetros operativos realistas, la restricción $Q \geq 1$ actúa como una **garantía de factibilidad teórica** más que como una limitación activa en la solución. No obstante, su inclusión es esencial para el rigor matemático del modelo, ya que asegura que el dominio de optimización sea cerrado y que el problema esté bien planteado en el sentido de Hadamard.

La expresión (Ecuación 5.6.12) constituye la **solución cerrada aproximada** del problema restringido, y será utilizada en la siguiente sección para analizar la sensibilidad del óptimo respecto a los parámetros del sistema.

5.6.5. Sensibilidad del óptimo respecto a β , Z_α , H , y S

En las secciones anteriores se obtuvo la expresión cerrada del tamaño óptimo de lote bajo la restricción $Q \geq 1$:

$$Q^* = \max \left\{ 1, \sqrt{\frac{2D(S + \beta C)}{H}} \right\}. \quad (5.6.13)$$

Esta fórmula explicita la dependencia funcional del óptimo respecto a cuatro parámetros críticos del modelo:

- $\beta > 0$: coeficiente de penalización por faltante,
- $S > 0$: costo fijo de pedido,
- $H > 0$: costo unitario de mantenimiento de inventario,
- y, de forma indirecta, $Z_\alpha \sigma_L$: parámetro que determina el inventario de seguridad, aunque **no influye en Q^*** cuando σ_L es constante (pues no aparece en la derivada de primer orden).

El análisis de sensibilidad consiste en estudiar cómo varía Q^* ante perturbaciones en estos parámetros. Dado que Q^* es una función compuesta de funciones elementales, su diferenciabilidad y monotonía pueden analizarse rigurosamente mediante cálculo diferencial.

5.6.5.1. Derivadas parciales y elasticidades

Supongamos que $Q^{\text{nc}} = \sqrt{2D(S + \beta C)/H} > 1$, de modo que $Q^* = Q^{\text{nc}}$ (caso interior, válido en contextos humanitarios realistas). Definimos:

$$Q^*(\beta, S, H) = \left(\frac{2D}{H}\right)^{1/2} (S + \beta C)^{1/2}.$$

Las derivadas parciales de primer orden son:

$$\frac{\partial Q^*}{\partial \beta} = \frac{1}{2} \left(\frac{2D}{H}\right)^{1/2} \frac{C}{(S + \beta C)^{1/2}} = \frac{CD}{HQ^*} > 0, \quad (5.6.14)$$

$$\frac{\partial Q^*}{\partial S} = \frac{1}{2} \left(\frac{2D}{H}\right)^{1/2} \frac{1}{(S + \beta C)^{1/2}} = \frac{D}{HQ^*} > 0, \quad (5.6.15)$$

$$\frac{\partial Q^*}{\partial H} = -\frac{1}{2} \left(\frac{2D(S + \beta C)}{H^3}\right)^{1/2} = -\frac{Q^*}{2H} < 0. \quad (5.6.16)$$

Estas expresiones confirman las siguientes propiedades:

1. **Monotonía creciente en β y S :** un aumento en la penalización por faltante o en el costo fijo de pedido induce al modelo a incrementar el tamaño del lote, ya que se vuelve más costoso emitir pedidos frecuentes o sufrir faltantes.
2. **Monotonía decreciente en H :** un mayor costo de mantenimiento desincentiva el almacenamiento, reduciendo el lote óptimo.
3. **Independencia de Z_α :** en el modelo presentado, el inventario de seguridad $Z_\alpha \sigma_L$ es un **término aditivo constante** en la función de costo, y por tanto **no afecta la localización del mínimo**. Solo influye en el **valor óptimo del costo total**, no en la **decisión de inventario Q^*** . Esta propiedad depende críticamente del supuesto de que σ_L es independiente de Q .

5.6.5.2. Elasticidades de sensibilidad

Para cuantificar el impacto relativo, definimos las **elasticidades** de Q^* respecto a cada parámetro:

$$\varepsilon_{Q^*,\beta} = \frac{\partial Q^*}{\partial \beta} \cdot \frac{\beta}{Q^*} = \frac{\beta C}{2(S + \beta C)} \in \left(0, \frac{1}{2}\right),$$

$$\varepsilon_{Q^*,S} = \frac{\partial Q^*}{\partial S} \cdot \frac{S}{Q^*} = \frac{S}{2(S + \beta C)} \in \left(0, \frac{1}{2}\right),$$

$$\varepsilon_{Q^*,H} = \frac{\partial Q^*}{\partial H} \cdot \frac{H}{Q^*} = -\frac{1}{2}.$$

Estas elasticidades son **constantes en signo**, pero **dependen del balance relativo** entre S y βC . En contextos humanitarios, donde $\beta C \gg S$, se cumple $\varepsilon_{Q^*,\beta} \rightarrow 1/2$ y $\varepsilon_{Q^*,S} \rightarrow 0$, lo que indica que el óptimo es **altamente sensible a la penalización por faltante** e **insensible al costo fijo de pedido**.

5.6.5.3. Ejemplo ilustrativo paso a paso

Retomamos los parámetros del ejemplo de la Sección 5.1.2:

- $D = 1000$,
- $S = 100$,
- $H = 5$,
- $C = 10$,
- $\beta = 50$.

Paso 1. Óptimo base:

$$Q^* = \sqrt{\frac{2 \cdot 1000 \cdot (100 + 50 \cdot 10)}{5}} = \sqrt{240\,000} \approx 489.90.$$

Paso 2. Sensibilidad respecto a β :

Aumentamos β en un 10%: $\beta' = 55$.

Nuevo óptimo:

$$Q'^* = \sqrt{\frac{2 \cdot 1000 \cdot (100 + 55 \cdot 10)}{5}} = \sqrt{\frac{2 \cdot 1000 \cdot 650}{5}} = \sqrt{260\,000} \approx 509.90.$$

Variación relativa: $\frac{509.90 - 489.90}{489.90} \approx 4.08\%$.

Elasticidad empírica: $\frac{4.08\%}{10\%} = 0.408$, cercana al valor teórico:

$$\varepsilon_{Q^*,\beta} = \frac{50 \cdot 10}{2(100 + 500)} = \frac{500}{1200} \approx 0.417.$$

Paso 3. Sensibilidad respecto a H :

Aumentamos H en un 20%: $H' = 6$.

Nuevo óptimo:

$$Q'^* = \sqrt{\frac{2 \cdot 1000 \cdot 600}{6}} = \sqrt{200\,000} \approx 447.21.$$

Variación relativa: $\frac{447.21 - 489.90}{489.90} \approx -8.71\%$.

Elasticidad empírica: $\frac{-8.71\%}{20\%} = -0.436$, cercana al valor teórico -0.5 (la discrepancia se debe a que la elasticidad es local, y el cambio es finito).

Paso 4. Independencia de Z_α :

Si cambiamos Z_α de 1.645 a 2.33 (pasando de 95% a 99% de nivel de servicio), el término $Z_\alpha \sigma_L H$ en $Z(Q)$ cambia, pero la derivada $\frac{dZ}{dQ}$ **no se ve afectada**, por lo que Q^* permanece en **489.90**.

Interpretación: En logística humanitaria, donde los costos de faltante son dominantes, el tamaño óptimo de inventario es **más sensible a la penalización β que a los costos operativos tradicionales**. Esto implica que la **calibración precisa de β** es crítica para la calidad de la solución, mientras que errores en S o H tienen impacto relativamente menor. Por otro lado, el diseño del nivel de servicio (vía Z_α) afecta el **costo total**, pero **no la política de inventario**, lo que permite una separación clara entre decisiones de magnitud (Q^*) y decisiones de robustez (inventario de seguridad).

Este análisis de sensibilidad no solo valida la robustez del modelo, sino que también proporciona **directrices operativas** para la recolección de datos y la toma de decisiones: es prioritario estimar con precisión β y D , mientras que S y H pueden aproximarse con mayor tolerancia. Además, confirma que el parámetro Z_α puede ajustarse **a posteriori** para satisfacer requisitos de nivel de servicio sin alterar la solución óptima de Q^* , siempre que σ_L sea independiente de Q .

En conjunto, las Sección 5.6.1–Sección 5.6.5 establecen que, bajo el enfoque determinista robusto adoptado en este trabajo, el problema de optimización unidimensional $\min_{Q \geq 1} Z(Q)$ admite una **solución cerrada explícita**, dada por la proyección $Q^* = \max\{1, \sqrt{2D(S + \beta C)/\bar{H}}\}$. Esta solución es **única, estable y sensible** a los parámetros críticos del modelo, especialmente a la penalización por faltante β . El análisis de sensibilidad confirma que, en contextos humanitarios, el tamaño óptimo de lote está dominado por la aversión al riesgo, lo que justifica el abandono del modelo EOQ clásico en favor de formulaciones que internalicen explícitamente la incertidumbre mediante parámetros de penalización y servicio. Estos resultados teóricos sientan las bases para la integración del componente de inventario en el modelo híbrido de localización-inventario.

5.7. Conclusiones del capítulo

Este capítulo ha presentado un modelo matemático determinista robusto para el problema integrado de localización e inventario en contextos humanitarios, caracterizado por la interacción entre decisiones binarias (ubicación de almacenes) y continuas (niveles de inventario y asignación de flujos). La formulación incorpora de forma paramétrica la incertidumbre en la demanda mediante un stock de seguridad y una penalización explícita por faltantes, lo que introduce no linealidad y rompe la convexidad global del problema, aun cuando la función de costo sea convexa condicionalmente para configuraciones fijas de infraestructura. Esta estructura híbrida, que combina no convexidad combinatoria, no linealidad inducida por la incertidumbre y una formulación de riesgo implícita a través de la penalización, constituye una contribución teórica que extiende los modelos clásicos de EOQ y localización, haciéndolos pertinentes para entornos donde la insatisfacción de la demanda tiene consecuencias éticas y operativas críticas. Al establecer propiedades analíticas fundamentales (existencia, unicidad y sensibilidad del óptimo en el subproblema de inventario) y al caracterizar formalmente el espacio factible mixto, el capítulo sienta las bases para el desarrollo de métodos de solución especializados.

6 Métodos numéricos y optimización con restricciones tipo caja

En este capítulo se presentan los fundamentos teóricos y computacionales de los métodos de optimización aplicados a problemas con restricciones tipo caja, es decir, aquellos en los que las variables de decisión se encuentran acotadas superior e inferiormente dentro de un dominio compacto. Este tipo de problemas es común en la modelación de sistemas físicos, económicos y logísticos, donde las condiciones naturales o estructurales limitan el espacio de búsqueda de soluciones. Siguiendo la metodología descrita por Bertsekas (2016), Capítulo 2; Luenberger y Ye (2016), Sección 8.4; Bazararaa, Sherali, y Shetty (2013b), Sección 8.7 y Nocedal y Wright (2006), Sección 12.2, se adopta un enfoque basado en proyección, condiciones de optimalidad de Karush–Kuhn–Tucker y métodos de memoria limitada para resolver eficientemente estos problemas.

El desarrollo del capítulo se organiza en cinco secciones. En primer lugar, se introducen los fundamentos de los métodos iterativos de primer y segundo orden, enfatizando los mecanismos de actualización del gradiente y la aproximación del Hessiano. Posteriormente, se extiende el análisis hacia dominios con restricciones tipo caja, detallando la formulación y justificación estructural del algoritmo **L-BFGS-B**, uno de los más eficientes para resolver problemas de gran escala con cotas bajo el esquema de Byrd et al. (1995) y Nocedal y Wright (2006) capítulo 6. En la tercera sección se analizan los resultados teóricos de convergencia del método bajo hipótesis de regularidad, presentando un teorema global que garantiza la existencia de puntos estacionarios que satisfacen las condiciones de Karush–Kuhn–Tucker (KKT) Sección 4.7. En seguida, la cuarta parte describe la implementación computacional del algoritmo y los experimentos numéricos realizados para evaluar su desempeño frente a otros métodos clásicos de optimización sin restricciones. Finalmente, se incluye una discusión crítica acerca de las ventajas, limitaciones y posibles extensiones del enfoque propuesto.

El propósito de este capítulo es ofrecer una comprensión integral del método **L-BFGS-B**, integrando su base matemática, su estructura computacional y su relevancia práctica para la resolución eficiente de problemas con restricciones tipo caja.

6.1. Fundamentos de métodos iterativos de primer y segundo orden

6.1.1. Método de gradiente descendente y sus propiedades de convergencia

El método de gradiente descendente constituye el algoritmo más elemental dentro de los métodos iterativos de primer orden para la minimización no lineal de funciones diferenciables. Su fundamento se basa en la propiedad geométrica de que el gradiente de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$, diferenciable en un punto $x \in \mathbb{R}^n$, apunta en la dirección de máximo crecimiento local. En consecuencia, la dirección opuesta, $-\nabla f(x)$, indica un camino de descenso local siempre que $\nabla f(x) \neq 0$ (véase Nocedal y Wright (2006), Capítulo 2, Sección 2.1, en particular la discusión sobre la dirección de máximo crecimiento, p. 14).

Formalmente, dado un punto inicial $x_0 \in \mathbb{R}^n$, el método genera una sucesión $\{x_k\}_{k \geq 0}$ definida por

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k),$$

donde $\alpha_k > 0$ es la longitud de paso seleccionada en cada iteración mediante una búsqueda lineal adecuada.

La elección de α_k resulta determinante para la convergencia. Una estrategia robusta impone las **condiciones de Wolfe** (Nocedal y Wright (2006), Sección 3.1, pp.33–34):

$$\begin{aligned} f(x_k - \alpha_k \nabla f(x_k)) &\leq f(x_k) - c_1 \alpha_k \|\nabla f(x_k)\|^2, \\ \nabla f(x_k - \alpha_k \nabla f(x_k))^\top \nabla f(x_k) &\geq c_2 \|\nabla f(x_k)\|^2, \end{aligned}$$

donde $0 < c_1 < c_2 < 1$. La primera asegura un descenso suficiente y la segunda evita pasos demasiado pequeños.

6.1.1.1. Propiedades de convergencia

Bajo las hipótesis:

1. f es continuamente diferenciable (ver Sección 4.3) en un conjunto abierto que contiene al conjunto de nivel $\mathcal{L} = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$,
2. \mathcal{L} es compacto (ver Sección 4.1),

el método de gradiente descendente con pasos que satisfacen las condiciones de Wolfe garantiza que

$$\liminf_{k \rightarrow \infty} \|\nabla f(x_k)\| = 0.$$

Si, además, f es **fuertemente convexa**, es decir, existe $m > 0$ tal que

$$(\nabla f(x) - \nabla f(y))^\top (x - y) \geq m\|x - y\|^2,$$

entonces la sucesión $\{x_k\}$ converge linealmente al minimizador x^* , cumpliendo

$$\|x_{k+1} - x^*\| \leq \rho \|x_k - x^*\|, \quad \rho \in (0, 1).$$

6.1.1.2. Ejemplo ilustrativo

Consideremos $f(x) = \frac{1}{2}x^\top Ax - b^\top x$, con

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

El gradiente es $\nabla f(x) = Ax$. Partiendo de $x_0 = (1, 1)^\top$, se obtiene mediante búsqueda lineal exacta:

$$\alpha_0 = \frac{13}{126} \approx 0.1032, \quad x_1 = (0.7937, -0.0317)^\top.$$

La trayectoria oscila hacia $x^* = (0, 0)^\top$, evidenciando la convergencia lineal y la influencia del número de condición $\kappa(A) = 5$.

Este comportamiento justifica la necesidad de métodos que incorporen información de segundo orden, como los métodos de tipo cuasi-Newton, presentados en la siguiente sección.

6.1.2. Métodos cuasi-Newton: aproximación del Hessiano y convergencia superlineal

Antes de describir el esquema algorítmico, introducimos la noción de **convergencia superlineal**, que caracteriza la eficiencia asintótica de esta clase de métodos.

Sea $\{x_k\}$ una sucesión generada por un algoritmo que converge a un punto óptimo x^* . Se dice que la convergencia es **superlineal** si

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

Esta tasa es más rápida que la convergencia lineal y refleja una aceleración progresiva del algoritmo cerca de la solución. Los métodos cuasi-Newton, como BFGS, alcanzan esta tasa bajo hipótesis de regularidad (por ejemplo, $f \in C^2$ con Hessiano Lipschitz-continuo) y búsquedas lineales que satisfagan las condiciones de Wolfe (Nocedal y Wright (2006), Sección 6.4).

Los métodos cuasi-Newton constituyen una clase fundamental de algoritmos iterativos para la minimización de funciones diferenciables sin restricciones, diseñados para emular el comportamiento del método de Newton sin requerir la evaluación explícita del Hessiano.

Los métodos cuasi-Newton constituyen una clase fundamental de algoritmos iterativos para la minimización de funciones diferenciables sin restricciones, diseñados para emular el comportamiento del método de Newton sin requerir la evaluación explícita del Hessiano $\nabla^2 f(x_k)$. En lugar de ello, estos métodos construyen **aproximaciones sucesivas** $B_k \in \mathbb{R}^{n \times n}$ (o $H_k \approx (\nabla^2 f(x_k))^{-1}$) a partir de la información de gradientes acumulada durante la iteración, logrando así tasas de convergencia **superlineales** bajo hipótesis razonables, como puede verse en la Sección 6.4 de Nocedal y Wright (2006).

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función continuamente diferenciable. El esquema genérico de un método cuasi-Newton se expresa como

$$x_{k+1} = x_k - \alpha_k B_k^{-1} \nabla f(x_k),$$

donde B_k es una matriz simétrica definida positiva que aproxima el Hessiano $\nabla^2 f(x_k)$, y $\alpha_k > 0$ se elige mediante una búsqueda lineal (por ejemplo, las condiciones de Wolfe).

La actualización de B_k se basa en la **ecuación secante**, derivada de la expansión de Taylor de primer orden del gradiente:

$$\nabla f(x_{k+1}) - \nabla f(x_k) \approx \nabla^2 f(x_k)(x_{k+1} - x_k).$$

Definiendo los incrementos

$$s_k := x_{k+1} - x_k, \quad y_k := \nabla f(x_{k+1}) - \nabla f(x_k),$$

la ecuación secante exige que la nueva aproximación B_{k+1} satisfaga

$$B_{k+1}s_k = y_k.$$

Dado que esta ecuación impone solo n restricciones lineales sobre las $\frac{n(n+1)}{2}$ entradas independientes de una matriz simétrica, se requiere un criterio adicional para determinar B_{k+1} de forma única. Los métodos cuasi-Newton clásicos surgen al imponer que B_{k+1} sea la **matriz más cercana a B_k** en alguna norma matricial adecuada, sujeta a la ecuación secante y a la simetría.

6.1.2.1. El método BFGS

El algoritmo **BFGS** (Broyden–Fletcher–Goldfarb–Shanno) es el método cuasi-Newton más utilizado en la práctica. Se obtiene al minimizar la distancia entre $H_{k+1} = B_{k+1}^{-1}$ y $H_k = B_k^{-1}$ en una norma de Frobenius ponderada que garantiza invariancia de escala. Recordemos que, dada una matriz $A \in \mathbb{R}^{n \times n}$ y una matriz de peso simétrica definida positiva W , la norma de Frobenius ponderada se define como

$$\|A\|_W = \|W^{1/2}AW^{1/2}\|_F = \sqrt{\text{tr}(WA^\top WA)},$$

donde $\|\cdot\|_F$ es la norma de Frobenius estándar. En el contexto del método BFGS, se elige $W = B_k$ (o equivalentemente se trabaja en la métrica inducida por H_k) para asegurar que la actualización sea invariante ante cambios de escala de las variables. La fórmula de actualización de la matriz inversa resultante es

$$H_{k+1} = (I - \rho_k s_k y_k^\top) H_k (I - \rho_k y_k s_k^\top) + \rho_k s_k s_k^\top,$$

donde

$$\rho_k = (y_k^\top s_k)^{-1}$$

y se asume

$$y_k^\top s_k > 0$$

(condición de curvatura), garantizada si la búsqueda lineal cumple las condiciones de Wolfe.

El método BFGS posee propiedades teóricas destacadas:

- Si f es dos veces continuamente diferenciable y su Hessiano $\nabla^2 f$ es Lipschitz-continuo (ver Sección 4.3) en una vecindad de un minimizador x^* , y la búsqueda lineal satisface las condiciones de Wolfe, entonces la sucesión $\{x_k\}$ cumple

$$\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0,$$

y la convergencia es **superlineal**:

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

Este resultado se basa en la caracterización contenida en la investigación de Dennis y Moré (1977):

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - \nabla^2 f(x^*))s_k\|}{\|s_k\|} = 0.$$

- Si además f es cuadrática y fuertemente convexa, BFGS converge en a lo sumo n pasos, generando direcciones conjugadas (ver Sección 4.3) respecto al Hessiano exacto. Nos referimos a que, si

$$f(x) = \frac{1}{2}x^\top Qx - b^\top x + c,$$

con $Q = \nabla^2 f$ simétrica definida positiva, entonces las direcciones de búsqueda $\{p_0, p_1, \dots, p_{n-1}\}$ producidas por el método satisfacen la condición de conjugación

$$p_i^\top Q p_j = 0 \quad \text{para todo } i \neq j,$$

es decir, son mutuamente ortogonales en la métrica inducida por Q . Esta propiedad es característica de los métodos tipo conjugado (como el método del gradiente conjugado) y garantiza que, en ausencia de errores de redondeo, se alcance el mínimo exacto en a lo mas n iteraciones.

6.1.2.2. Ejemplo ilustrativo: minimización cuadrática con BFGS

Sea la función cuadrática fuertemente convexa en \mathbb{R}^2 :

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x, \quad A = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 10 \end{pmatrix},$$

cuyo minimizador es $x^* = A^{-1}b = (1, 1)^\top$.

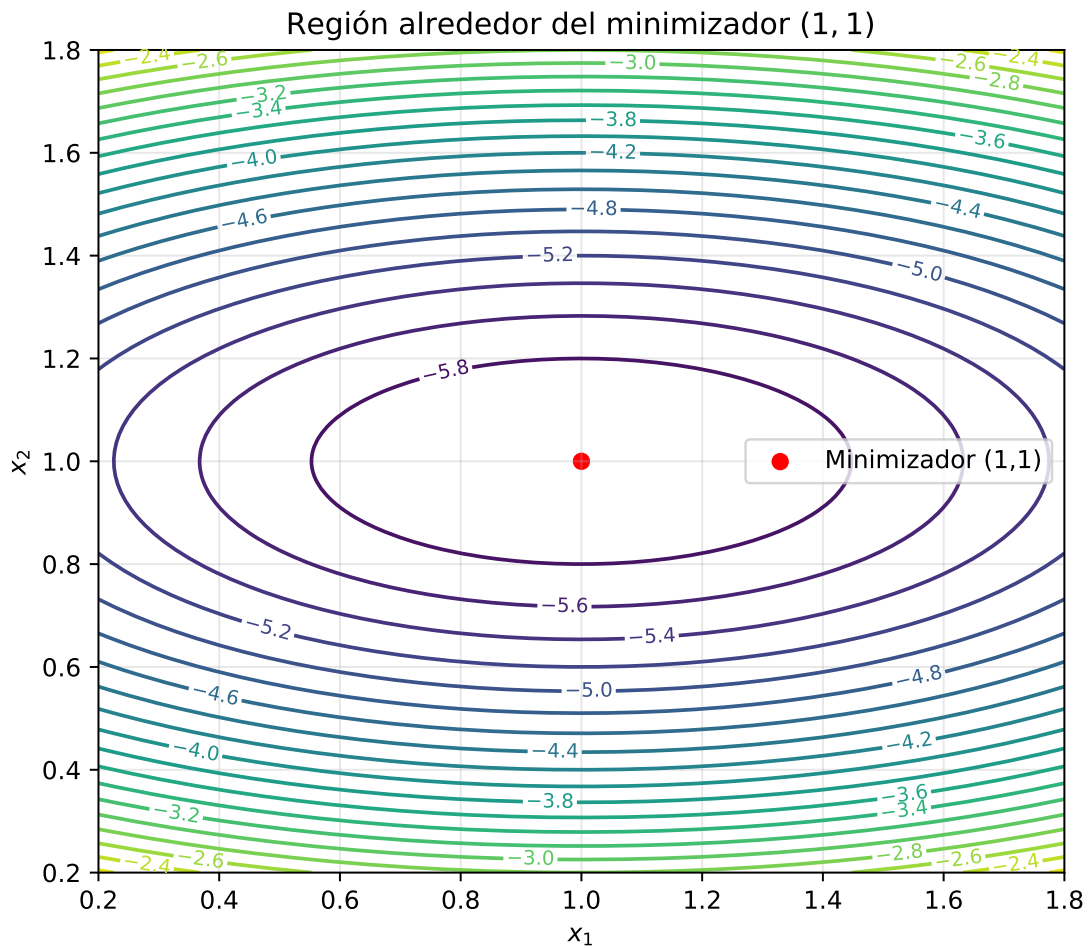


Figura 6.1: Visualización de la estructura geométrica de la función cuadrática en un vecindario del minimizador global $(1,1)$. Las elipses de nivel reflejan la convexidad fuerte del modelo y la direccionalidad de la curvatura asociada a la matriz Hessiana diagonal A , proporcionando una base geométrica para analizar el comportamiento local de métodos iterativos como BFGS.

Partimos de $x_0 = (0, 0)^\top$ y elegimos $H_0 = I$ como aproximación inicial de la inversa del Hessiano. Utilizamos búsqueda lineal exacta.

Iteración $k = 0$:

- $\nabla f(x_0) = Ax_0 - b = (-2, -10)^\top$.
- Dirección: $p_0 = -H_0 \nabla f(x_0) = (2, 10)^\top$.
- Paso óptimo:

$$\alpha_0 = \frac{\nabla f(x_0)^\top p_0}{p_0^\top A p_0} = \frac{104}{1008} = \frac{13}{126} \approx 0.1032.$$

- Nuevo punto:

$$x_1 = x_0 + \alpha_0 p_0 = (0.2063, 1.0317)^\top.$$

- Incrementos:

$$s_0 = x_1 - x_0, \quad y_0 = \nabla f(x_1) - \nabla f(x_0) = A s_0.$$

- Condición de curvatura: $y_0^\top s_0 \approx 10.746 > 0$.
- Actualización BFGS:

$$H_1 \approx \begin{pmatrix} 0.503 & -0.004 \\ -0.004 & 0.100 \end{pmatrix},$$

que aproxima bien a $A^{-1} = \text{diag}(0.5, 0.1)$.

Iteración $k = 1$:

- $\nabla f(x_1) = Ax_1 - b \approx (-1.587, 0.317)^\top$.
- Dirección: $p_1 = -H_1 \nabla f(x_1) \approx (0.792, -0.032)^\top$.
- Paso óptimo: $\alpha_1 \approx 1.033$.
- Nuevo punto: $x_2 \approx (1.022, 0.999)^\top$, muy cercano a $x^* = (1, 1)^\top$.

En solo dos iteraciones el algoritmo reduce el error residual en más de un orden de magnitud, demostrando la **convergencia superlineal** (ver Sección 6.1.2) del método BFGS para funciones cuadráticas. Este comportamiento se mantiene en problemas no cuadráticos siempre que el condicionamiento sea razonable y la búsqueda lineal sea suficientemente precisa.

6.1.3. Familia Broyden y el método BFGS como caso prototípico

La clase de métodos cuasi-Newton puede organizarse en torno a la **familia de actualizaciones de Broyden**, un conjunto paramétrico de fórmulas que preservan la ecuación secante y generan matrices simétricas actualizadas a partir de la información de gradiente acumulada. Esta familia engloba los métodos más relevantes en la práctica, entre ellos el **BFGS** (Broyden–Fletcher–Goldfarb–Shanno) y el **DFP** (Davidon–Fletcher–Powell), y permite, como se expone en la Sección 6.3 de Nocedal y Wright (2006), analizar de forma unificada sus propiedades de convergencia, estabilidad y eficiencia computacional.

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función dos veces continuamente diferenciable (ver Sección 4.3). En cada iteración k , se dispone de los incrementos

$$s_k := x_{k+1} - x_k, \quad y_k := \nabla f(x_{k+1}) - \nabla f(x_k),$$

y se busca una matriz simétrica $B_{k+1} \in \mathbb{R}^{n \times n}$ que satisfaga la **ecuación secante**

$$B_{k+1} s_k = y_k. \tag{6.1.1}$$

La **familia de Broyden** se define como el conjunto de actualizaciones

$$B_{k+1}(\phi_k) = B_k - \frac{B_k s_k s_k^\top B_k}{s_k^\top B_k s_k} + \frac{y_k y_k^\top}{y_k^\top s_k} + \phi_k v_k v_k^\top, \quad (6.1.2)$$

donde

$$v_k = \left(\frac{y_k}{y_k^\top s_k} - \frac{B_k s_k}{s_k^\top B_k s_k} \right), \quad (6.1.3)$$

y $\phi_k \in \mathbb{R}$ es un parámetro que determina el miembro particular de la familia.

La fórmula Ecuación 6.1.2 puede interpretarse como una combinación convexa de dos actualizaciones extremas:

- **BFGS**: $\phi_k = 0$,
- **DFP**: $\phi_k = 1$.

De hecho,

$$B_{k+1}(\phi_k) = (1 - \phi_k) B_{k+1}^{\text{BFGS}} + \phi_k B_{k+1}^{\text{DFP}}. \quad (6.1.4)$$

6.1.3.1. Propiedades de la familia restringida de Broyden

Una subclase especialmente relevante es la **familia restringida de Broyden**, asociada a $\phi_k \in [0, 1]$. Bajo la **condición de curvatura**

$$y_k^\top s_k > 0, \quad (6.1.5)$$

garantizada si la búsqueda lineal cumple las condiciones de Wolfe, todos los miembros de esta subclase preservan la **definida positividad** de B_k , siempre que B_0 sea simétrica definida positiva. Esta propiedad asegura que la dirección de búsqueda

$$p_k = -B_k^{-1} \nabla f(x_k) \quad (6.1.6)$$

sea de descenso. Además, si f es una función cuadrática fuertemente convexa (ver Sección 4.3), cualquier método de la familia restringida de Broyden converge en a lo sumo n pasos, generando direcciones conjugadas (ver Sección 4.3) respecto al Hessiano exacto $\nabla^2 f$ Dennis y Moré (1977), Sección 5.

6.1.3.2. El método BFGS como caso prototípico

El método **BFGS** corresponde a $\phi_k = 0$ en Ecuación 6.1.2. Su fórmula de actualización de la **inversa del Hessiano aproximado** $H_k = B_k^{-1}$ es

$$H_{k+1} = (I - \rho_k s_k y_k^\top) H_k (I - \rho_k y_k s_k^\top) + \rho_k s_k s_k^\top, \quad \rho_k = (y_k^\top s_k)^{-1}. \quad (6.1.7)$$

Esta formulación evita la inversión explícita de matrices y facilita el cálculo de la dirección de búsqueda:

$$p_k = -H_k \nabla f(x_k).$$

El método BFGS se distingue por:

- Su **invariancia afín**: el comportamiento es independiente de escalas y transformaciones lineales.
- Su **propiedad de autocorrección** (*self-scaling*): si H_k es una mala aproximación, la iteración corrige rápidamente la orientación de la matriz.
- Su **convergencia superlineal** está garantizada bajo hipótesis estándar de suavidad, es decir, que f sea dos veces continuamente diferenciable (ver Sección 4.3) y que su gradiente (o, equivalentemente, su Hessiano) sea Lipschitz continuo en una vecindad del minimizador, junto con el uso de una búsqueda lineal que satisfaga las condiciones de Wolfe, como se establece en Fletcher (1987), Sección 3.4.

6.1.3.3. Ejemplo ilustrativo: actualización BFGS en \mathbb{R}^2

Sea

$$f(x) = \frac{1}{2} x^\top A x - b^\top x, \quad A = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

El minimizador es $x^* = A^{-1}b = (1/7, 3/7)^\top \approx (0.1429, 0.4286)^\top$. Partimos de $x_0 = (0, 0)^\top$ y $H_0 = I$.

Iteración $k = 0$:

- $\nabla f(x_0) = (-1, -1)^\top$.
- $p_0 = (1, 1)^\top$.
- $\alpha_0 = \frac{2}{8} = 0.25$.
- $x_1 = (0.25, 0.25)^\top$.
- $\nabla f(x_1) = (0.25, -0.25)^\top$.
- $s_0 = (0.25, 0.25)^\top$, $y_0 = (1.25, 0.75)^\top$.
- $y_0^\top s_0 = 0.5 > 0$, $\rho_0 = 2$.

- $H_1 \approx \begin{pmatrix} 0.406 & -0.344 \\ -0.344 & 0.906 \end{pmatrix}$.

Iteración $k = 1$:

- $p_1 = -H_1(0.25, -0.25)^\top \approx (-0.1875, 0.3125)^\top$.
- $\alpha_1 \approx 0.857$, $x_2 \approx (0.089, 0.518)^\top$.

Tras dos iteraciones, $\|x_2 - x^*\| \approx 0.09$, y H_1 aproxima bien A^{-1} , demostrando la capacidad del método para capturar la geometría de segundo orden con pocas actualizaciones.

El método BFGS representa el **caso prototípico** dentro de la familia de Broyden, al combinar eficiencia computacional, robustez teórica y convergencia superlineal incluso en problemas moderadamente mal condicionados.

6.2. Extensión a dominios con restricciones tipo caja

En la práctica, muchos problemas de optimización imponen límites naturales o estructurales sobre las variables de decisión, dando origen a los denominados **problemas con restricciones tipo caja**. En este contexto, las variables están acotadas por límites inferiores y superiores que definen un dominio factible de la forma $\Omega = \{x \in \mathbb{R}^n : \ell_i \leq x_i \leq u_i\}$.

Estos problemas aparecen con frecuencia en aplicaciones de ingeniería, finanzas y logística, donde ciertas magnitudes no pueden sobrepasar intervalos físicamente admisibles. Extender los métodos de optimización a dominios de este tipo requiere incorporar operadores de **proyección ortogonal** que aseguren la factibilidad de cada iteración sin comprometer la convergencia del algoritmo. La presente sección desarrolla las propiedades analíticas de dicha proyección, que constituye el núcleo geométrico de los métodos con restricciones tipo caja.

6.2.1. Proyección ortogonal sobre conjuntos de cotas: definición y propiedades

En la optimización con restricciones tipo caja, el **conjunto factible** se define como

$$\Omega = \{x \in \mathbb{R}^n : \ell_i \leq x_i \leq u_i, i = 1, \dots, n\},$$

donde $\ell = (\ell_1, \dots, \ell_n)^\top \in \mathbb{R}^n \cup \{-\infty\}^n$ y $u = (u_1, \dots, u_n)^\top \in \mathbb{R}^n \cup \{+\infty\}^n$ satisfacen $\ell_i < u_i$ para todo i . Este conjunto es **convexo, cerrado y no vacío**, lo que garantiza que la proyección ortogonal de cualquier punto $z \in \mathbb{R}^n$ sobre Ω está bien definida y es única.

6.2.1.1. Proyección Ortogonal

La proyección ortogonal de z sobre Ω , denotada $P_\Omega(z)$, se define como la solución única del problema de minimización

$$P_\Omega(z) = \arg \min_{x \in \Omega} \frac{1}{2} \|x - z\|^2. \quad (6.2.1)$$

Dado que Ω es un producto cartesiano de intervalos cerrados, la proyección se **descompone coordenada a coordenada**:

$$[P_\Omega(z)]_i = \min\{u_i, \max\{\ell_i, z_i\}\}, \quad i = 1, \dots, n. \quad (6.2.2)$$

Esta expresión corresponde a la llamada *función de recorte* (*clipping function*), que restringe cada componente de z al intervalo factible $[\ell_i, u_i]$.

6.2.1.2. Propiedades fundamentales

1. No expansividad

$$\|P_\Omega(z) - P_\Omega(w)\| \leq \|z - w\|, \quad \forall z, w \in \mathbb{R}^n. \quad (6.2.3)$$

Por tanto, P_Ω es una aplicación **Lipschitz continua** (ver Sección 4.3) con **constante 1**, lo que implica estabilidad numérica.

2. Caracterización variacional

Un punto $x^* \in \Omega$ satisface $x^* = P_\Omega(z)$, si y solo si,

$$(z - x^*)^\top (x - x^*) \leq 0, \quad \forall x \in \Omega. \quad (6.2.4)$$

Geoméricamente, esto significa que el vector residual $z - x^*$ forma un ángulo obtuso con cualquier dirección factible desde x^* .

3. Condiciones de complementariedad (KKT)

Para cada i se cumple

$$\begin{cases} x_i = \ell_i & \Rightarrow z_i \leq \ell_i, \\ x_i = u_i & \Rightarrow z_i \geq u_i, \\ \ell_i < x_i < u_i & \Rightarrow z_i = x_i, \end{cases} \quad (6.2.5)$$

que equivalen a las condiciones de optimalidad de primer orden del problema cuadrático (Ecuación 6.2.1).

6.2.1.3. Ejemplo: proyección en \mathbb{R}^3

Sea

$$\Omega = \{x \in \mathbb{R}^3 : 1 \leq x_1 \leq 3, x_2 \leq 2, x_3 \geq 0\}, \quad z = (0.5, 2.5, -1.2)^\top.$$

Aplicando Ecuación 6.2.2 coordenada a coordenada:

- $x_1^* = \min\{3, \max\{1, 0.5\}\} = 1,$
- $x_2^* = \min\{2, \max\{-\infty, 2.5\}\} = 2,$
- $x_3^* = \min\{+\infty, \max\{0, -1.2\}\} = 0.$

Por tanto,

$$P_\Omega(z) = (1, 2, 0)^\top.$$

Comprobamos la caracterización variacional:

$$(z - x^*)^\top(x - x^*) = (-0.5, 0.5, -1.2)^\top \cdot (x_1 - 1, x_2 - 2, x_3 - 0).$$

Dado que $x_1 \geq 1$, $x_2 \leq 2$, $x_3 \geq 0$, cada término parcial es no positivo, y la suma total ≤ 0 .

Asimismo, las condiciones Ecuación 6.2.5 se satisfacen coordenada a coordenada: $x_1^* = 1$ con $z_1 = 0.5 \leq 1$, $x_2^* = 2$ con $z_2 = 2.5 \geq 2$, y $x_3^* = 0$ con $z_3 = -1.2 \leq 0$.

Este ejemplo ilustra cómo la proyección actúa localmente sobre cada componente, garantizando factibilidad sin resolver sistemas lineales. Su simplicidad y propiedades de estabilidad hacen que el operador P_Ω sea esencial en algoritmos de optimización a gran escala, como el método **L-BFGS-B**, en el cual se aplica en cada iteración para mantener las variables dentro de los límites definidos.

6.2.2. El algoritmo L-BFGS-B: formulación y justificación estructural

El algoritmo **L-BFGS-B** (*Limited-memory BFGS with Bounds*) extiende el método cuasi-Newton BFGS al caso de **restricciones tipo caja**, es decir, dominios de la forma

$$\Omega = \{x \in \mathbb{R}^n : \ell \leq x \leq u\},$$

donde las desigualdades son componente a componente y $\ell, u \in \mathbb{R}^n \cup \{\pm\infty\}^n$ con $\ell_i < u_i$. El método integra tres ideas fundamentales:

1. **Aproximación cuasi-Newton de memoria limitada**, evitando el almacenamiento completo de matrices densas;

2. **Proyección ortogonal** sobre Ω para mantener factibilidad;
3. **Retención parcial de información de curvatura**, conservando sólo las m parejas más recientes (s_i, y_i) .

6.2.2.1. Formulación del algoritmo

Dado un punto inicial $x_0 \in \Omega$, el esquema iterativo se expresa como

$$x_{k+1} = P_\Omega(x_k - \alpha_k H_k \nabla f(x_k)), \quad (6.2.6)$$

donde $\alpha_k > 0$ satisface las condiciones de Wolfe y H_k es la aproximación de memoria limitada al inverso del Hessiano, construida mediante las m actualizaciones más recientes:

$$s_i = x_{i+1} - x_i, \quad y_i = \nabla f(x_{i+1}) - \nabla f(x_i), \quad i = k - m, \dots, k - 1,$$

con $y_i^\top s_i > 0$ (condición de curvatura).

La proyección P_Ω se calcula coordenada a coordenada como

$$[P_\Omega(z)]_i = \min\{u_i, \max\{\ell_i, z_i\}\}, \quad i = 1, \dots, n. \quad (6.2.7)$$

El producto $H_k \nabla f(x_k)$ se obtiene por la **recursión de dos bucles**, iniciada con $H_k^{(0)} = \gamma_k I$, donde

$$\gamma_k = \frac{s_{k-1}^\top y_{k-1}}{y_{k-1}^\top y_{k-1}}. \quad (6.2.8)$$

Este escalamiento mejora la estabilidad numérica y la aceptación de pasos unitarios.

6.2.2.2. Justificación estructural

El esquema Ecuación 6.2.6 garantiza factibilidad a cada iteración mediante la proyección P_Ω .

Definiendo el conjunto de índices activos

$$\mathcal{A}_k = \{i : x_{k,i} = \ell_i \text{ o } x_{k,i} = u_i\},$$

el método se comporta localmente como BFGS en los índices inactivos ($i \notin \mathcal{A}_k$) y mantiene las componentes activas en las cotas.

De este modo, la actualización proyectada puede interpretarse como

$$x_{k+1} = \text{Proj}_\Omega (x_k - \alpha_k \nabla^2 f(x_k)^{-1} \nabla f(x_k)),$$

con $\nabla^2 f(x_k)^{-1}$ sustituido por la aproximación limitada H_k .

Esta interpretación explica por qué L-BFGS-B conserva tanto la factibilidad como las propiedades de descenso y la convergencia superlineal del BFGS original.

En términos computacionales, el método BFGS requiere $\mathcal{O}(n^2)$ unidades de memoria. Esto se debe a que, en cada iteración, el algoritmo mantiene una representación explícita de una matriz simétrica $H_k \in \mathbb{R}^{n \times n}$ que aproxima la inversa del Hessiano de la función objetivo. El almacenamiento de dicha matriz implica retener el orden de n^2 números reales en memoria principal, lo cual constituye un costo cuadrático en la dimensión del problema. Este requisito es inherente a la formulación clásica del método, ya que tanto la actualización de H_k mediante la fórmula cuasi-Newton como el cálculo de la dirección de búsqueda $p_k = -H_k \nabla f(x_k)$ dependen de acceso directo a todos los elementos de la matriz. En problemas de gran escala, donde n puede alcanzar órdenes de magnitud de 10^5 o más, este costo se vuelve prohibitivo desde el punto de vista práctico. En contraste, el algoritmo L-BFGS-B evita el almacenamiento explícito de H_k y solo retiene los últimos m pares de vectores (s_k, y_k) , lo que reduce el requerimiento de memoria a $\mathcal{O}(mn)$, con $m \ll n$ (típicamente $3 \leq m \leq 20$).

6.2.2.3. Ejemplo: minimización cuadrática con restricciones tipo caja

Considérese

$$\min_{x \in \mathbb{R}^2} f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \quad \text{sujeto a} \quad 1 \leq x_1 \leq 3, \quad x_2 \geq 0.$$

El minimizador irrestricto $x^* = (2, 1)^\top$ pertenece a Ω , por lo que coincide con el minimizador restringido.

Con $m = 1$ y $x_0 = (1, 0)^\top$, el procedimiento es:

Iteración 0

$$\nabla f(x_0) = (-2, -2)^\top, \quad H_0 = I, \quad d_0 = (2, 2)^\top.$$

$$\text{Con } \alpha_0 = 1, \text{ se obtiene } \tilde{x}_1 = (3, 2)^\top \text{ y } x_1 = P_\Omega(\tilde{x}_1) = (3, 2)^\top.$$

$$\nabla f(x_1) = (2, 2)^\top, \quad s_0 = (2, 2)^\top, \quad y_0 = (4, 4)^\top.$$

Iteración 1

$$\gamma_1 = \frac{s_0^\top y_0}{y_0^\top y_0} = 0.5,$$

$$H_1^{(0)} = 0.5I.$$

Aplicando la recursión de dos bucles:

$$q = \nabla f(x_1), \quad \alpha_0 = 0.5, \quad q = q - 0.5y_0 = (0, 0)^\top,$$

$$r = H_1^{(0)} q = (0, 0)^\top,$$

$$r = r + s_0(\alpha_0 - \frac{y_0^\top r}{y_0^\top s_0}) = (1, 1)^\top.$$

Dirección $d_1 = -r = (-1, -1)^\top$.

Con $\alpha_1 = 1$, $\tilde{x}_2 = (2, 1)^\top$, $x_2 = P_\Omega(\tilde{x}_2) = (2, 1)^\top$.

El método converge en dos iteraciones, demostrando que la estructura de memoria limitada y la proyección garantizan simultáneamente **eficiencia y factibilidad**.

La formulación L-BFGS-B preserva las propiedades de convergencia global del BFGS bajo las hipótesis de suavidad y curvatura positiva. Para más detalles Byrd et al. (1995); Zhu et al. (1997), y constituye uno de los métodos estándar para problemas de gran escala con restricciones tipo caja.

6.2.3. Condiciones de optimalidad bajo cotas: KKT para problemas con restricciones tipo caja

Tras haber introducido la formulación proyectada del algoritmo L-BFGS-B, resulta necesario caracterizar formalmente las condiciones de optimalidad que dicho método busca satisfacer. En particular, cuando el dominio está delimitado por cotas inferiores y superiores, las condiciones de Karush–Kuhn–Tucker (KKT) Sección 4.7 adquieren una estructura particularmente simple y separable por coordenadas. Esta sección desarrolla esa caracterización, que constituye el fundamento teórico de la convergencia de los métodos con restricciones tipo caja.

Consideremos el problema de optimización

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sujeto a} \quad \ell \leq x \leq u, \quad (6.2.9)$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es continuamente diferenciable (ver Sección 4.3) en un conjunto abierto que contiene al conjunto factible

$$\Omega = \{x \in \mathbb{R}^n : \ell_i \leq x_i \leq u_i, \quad i = 1, \dots, n\},$$

con $\ell_i < u_i$ para todo i . Este problema constituye un caso particular de programación no lineal con restricciones de desigualdad simples.

6.2.3.1. Condiciones de Karush–Kuhn–Tucker (KKT)

Reescribiendo las restricciones como desigualdades estándar,

$$c_i^\ell(x) := \ell_i - x_i \leq 0, \quad c_i^u(x) := x_i - u_i \leq 0, \quad i = 1, \dots, n, \quad (6.2.10)$$

y suponiendo que se cumple la condición de calificación de Mangasarian–Fromovitz (MFQC), las **condiciones necesarias de Karush–Kuhn–Tucker** establecen que si $x^* \in \Omega$ es un minimizador local, existen multiplicadores de Lagrange $\lambda^\ell, \lambda^u \in \mathbb{R}^n$ tales que

$$\nabla f(x^*) - \lambda^\ell + \lambda^u = 0. \quad (6.2.11)$$

$$\lambda_i^\ell \geq 0, \quad \lambda_i^u \geq 0. \quad (6.2.12)$$

$$\lambda_i^\ell (x_i^* - \ell_i) = 0. \quad (6.2.13)$$

$$\lambda_i^u (u_i - x_i^*) = 0. \quad (6.2.14)$$

para todo $i = 1, \dots, n$.

Las ecuaciones Ecuación 6.2.13–Ecuación 6.2.14 expresan la **complementariedad** entre los multiplicadores y las restricciones activas.

6.2.3.2. Forma equivalente sin multiplicadores

Eliminando λ^ℓ y λ^u , las condiciones KKT pueden escribirse componente a componente como

$$\forall i = 1, \dots, n : \begin{cases} \nabla_i f(x^*) = 0, & \text{si } \ell_i < x_i^* < u_i, \\ \nabla_i f(x^*) \geq 0, & \text{si } x_i^* = \ell_i, \\ \nabla_i f(x^*) \leq 0, & \text{si } x_i^* = u_i, \end{cases} \quad (6.2.15)$$

donde las desigualdades se interpretan componente a componente.

Geoméricamente, esto significa que el gradiente proyectado sobre el **cono tangente** de Ω en x^* es nulo.

Una formulación alternativa, útil en algoritmos numéricos, es la **condición de estacionariedad proyectada**:

$$\nabla_\Omega f(x^*) := P_\Omega(x^* - \nabla f(x^*)) - x^* = 0. \quad (6.2.16)$$

La igualdad Ecuación 6.2.16 es **equivalente** a las condiciones KKT cuando f es convexa, y constituye una **condición necesaria** de optimalidad en el caso no convexo. En la práctica, se emplea como criterio de convergencia en algoritmos proyectados como L-BFGS-B.

6.2.3.3. Ejemplo: verificación explícita de KKT

Sea

$$\min_{x \in \mathbb{R}^2} f(x) = (x_1 - 1)^2 + (x_2 - 2)^2, \quad \text{sujeto a } 1 \leq x_1 \leq 3, x_2 \geq 0. \quad (6.2.17)$$

El gradiente es $\nabla f(x) = (2(x_1 - 1), 2(x_2 - 2))^\top$.

Paso 1. Identificación del candidato.

El minimizador irrestricto $(1, 2)^\top$ pertenece a Ω , por lo tanto $x^* = (1, 2)^\top$ es candidato a solución óptima.

Paso 2. Verificación de Ecuación 6.2.15.

- $x_1^* = \ell_1 = 1 \Rightarrow \nabla_1 f(x^*) = 0 \geq 0$
- $\ell_2 < x_2^* < +\infty \Rightarrow \nabla_2 f(x^*) = 0$

Paso 3. Multiplicadores KKT.

De Ecuación 6.2.11,

$$\nabla f(x^*) - \lambda^\ell + \lambda^u = 0 \Rightarrow \lambda^\ell = \lambda^u = 0.$$

Paso 4. Estacionariedad proyectada.

$$x^* - \nabla f(x^*) = (1, 2), \quad P_\Omega(1, 2) = (1, 2), \quad \Rightarrow \nabla_\Omega f(x^*) = 0.$$

Así, x^* satisface todas las condiciones de Karush–Kuhn–Tucker y es el minimizador global.

Este análisis demuestra que las condiciones KKT para restricciones tipo caja poseen una estructura **desacoplada y local**, lo que permite su implementación eficiente en métodos de gran escala. En particular, la condición proyectada Ecuación 6.2.16 constituye la base teórica de los algoritmos de proyección iterativa, incluyendo el método L-BFGS-B, cuya convergencia global se verifica precisamente mediante la nulidad asintótica de $\nabla_\Omega f(x_k)$.

6.3. Convergencia del método L-BFGS-B

6.3.1. Hipótesis de regularidad: Lipschitz-continuidad del gradiente y compacidad del dominio

El análisis de convergencia global del algoritmo **L-BFGS-B** requiere hipótesis estructurales sobre la función objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y sobre el conjunto factible $\Omega \subset \mathbb{R}^n$. Estas condiciones garantizan la existencia de minimizadores, la acotación de las iteraciones y la controlabilidad del gradiente, elementos esenciales para asegurar tanto la **bien definición del algoritmo** como su **convergencia hacia puntos estacionarios**.

6.3.1.1. Hipótesis fundamentales

Sea

$$\Omega = \{x \in \mathbb{R}^n : \ell \leq x \leq u\},$$

con $\ell, u \in \mathbb{R}^n \cup \{\pm\infty\}^n$ y $\ell_i < u_i$ para todo i . Supondremos que f satisface las siguientes condiciones:

(H1) Diferenciabilidad y Lipschitz-continuidad del gradiente.

La función f es continuamente diferenciable (ver Sección 4.3) en un conjunto abierto que contiene a Ω , y su gradiente ∇f es **Lipschitz continuo** en Ω ; es decir, existe una constante $L > 0$ tal que

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Omega. \quad (6.3.1)$$

(H2) Compacidad del conjunto de nivel inicial.

Dado un punto inicial $x_0 \in \Omega$, el conjunto de nivel asociado

$$\mathcal{L} := \{x \in \Omega : f(x) \leq f(x_0)\} \quad (6.3.2)$$

es **no vacío, cerrado y acotado**, y por tanto **compacto**.

La hipótesis (H1) implica que el gradiente no cambia bruscamente, lo que permite controlar los pasos del algoritmo mediante la cota superior L . Equivalentemente, la desigualdad de suavidad se cumple:

$$f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2}\|y - x\|^2, \quad \forall x, y \in \Omega. \quad (6.3.3)$$

Esta relación garantiza que las búsquedas lineales que satisfacen las **condiciones de Wolfe** producen pasos uniformemente acotados y no degenerados.

La hipótesis (H2), por su parte, asegura que las iteraciones $\{x_k\}$ permanecen en un subconjunto **compacto y factible** de Ω . Esto permite aplicar resultados de convergencia por subsucesión y garantiza que las secuencias de gradientes y valores de función no divergen.

6.3.1.2. Consecuencias teóricas

Bajo las hipótesis (H1) y (H2), la sucesión $\{x_k\}$ generada por L-BFGS-B con una búsqueda lineal que cumple las condiciones de Wolfe posee las siguientes propiedades:

1. **Existencia de puntos límite.**

Toda subsucesión de $\{x_k\}$ posee una subsucesión convergente cuyo límite pertenece a $\mathcal{L} \subset \Omega$.

2. Condición de curvatura.

Las parejas de actualización

$$s_k = x_{k+1} - x_k, \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

satisfacen $y_k^\top s_k > 0$ siempre que se cumplan las condiciones de Wolfe, lo que garantiza que las aproximaciones cuasi-Newton sean definidas positivas (Byrd et al., 1995).

3. Acotación de los operadores de memoria limitada.

La recursión de dos bucles empleada en L-BFGS-B mantiene direcciones de búsqueda acotadas y de descenso, gracias a la compacidad de \mathcal{L} y a la positividad de los escalares $\rho_k = (y_k^\top s_k)^{-1}$.

6.3.1.3. Teorema de convergencia global (versión esquemática)

Teorema.

Supóngase que f satisface (H1) y (H2), y que la búsqueda lineal cumple las condiciones de Wolfe. Si las actualizaciones de memoria limitada satisfacen $y_k^\top s_k > 0$ o son rechazadas en caso contrario (salvaguarda numérica), entonces la sucesión $\{x_k\}$ generada por L-BFGS-B verifica

$$\liminf_{k \rightarrow \infty} \|\nabla_\Omega f(x_k)\| = 0, \quad (6.3.4)$$

donde el **gradiente proyectado** se define como

$$\nabla_\Omega f(x) := P_\Omega(x - \nabla f(x)) - x. \quad (6.3.5)$$

Bosquejo de prueba.

De (H2) se sigue que $\{x_k\} \subset \mathcal{L}$ es acotada, por lo que posee subsucesiones convergentes. La desigualdad Ecuación 6.3.3 y las condiciones de Wolfe garantizan un descenso suficiente de f . La condición de curvatura $y_k^\top s_k > 0$ asegura que las matrices H_k permanecen definidas positivas y acotadas (Nocedal y Wright (2006), Teorema 3.2 y Sección 6.1). Por tanto, todo punto límite \bar{x} de la sucesión satisface la condición de estacionariedad proyectada $\nabla_\Omega f(\bar{x}) = 0$ (Byrd et al. (1995), Teorema 4.1).

6.3.1.4. Ejemplo de verificación de las hipótesis

Consideremos el problema

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^4 + x_2^2 \quad \text{sujeto a } 0 \leq x_1 \leq 2, \quad -1 \leq x_2 \leq 1. \quad (6.3.6)$$

Paso 1. Verificación de (H1).

El gradiente es

$$\nabla f(x) = (4x_1^3, 2x_2)^\top, \quad \nabla^2 f(x) = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{pmatrix}.$$

En $\Omega = [0, 2] \times [-1, 1]$,

$$\|\nabla^2 f(x)\|_2 = \max\{12x_1^2, 2\} \leq 48.$$

Por tanto, $L = 48$ es una constante de Lipschitz para ∇f en Ω , y (H1) se cumple.

Paso 2. Verificación de (H2).

Sea $x_0 = (2, 1)^\top \in \Omega$. Entonces

$$f(x_0) = 2^4 + 1^2 = 17, \quad \mathcal{L} = \{x \in \Omega : x_1^4 + x_2^2 \leq 17\}.$$

Dado que Ω es compacto y $\mathcal{L} \subset \Omega$, el conjunto de nivel es cerrado, acotado y no vacío (pues $x^* = (0, 0)^\top \in \mathcal{L}$). Así, (H2) también se satisface.

Paso 3. Implicaciones para L-BFGS-B.

Bajo estas condiciones, la sucesión $\{x_k\}$ generada por L-BFGS-B permanece en \mathcal{L} , el gradiente es Lipschitz con constante $L = 48$, y las parejas (s_k, y_k) satisfacen $y_k^\top s_k > 0$. Por tanto, el algoritmo produce direcciones de descenso bien definidas, y los límites de las iteraciones satisfacen las condiciones de Karush–Kuhn–Tucker bajo restricciones tipo caja.

En síntesis, las hipótesis (H1) y (H2) proporcionan el **marco teórico mínimo** para garantizar la convergencia global del método L-BFGS-B. Estas condiciones, verificables en la práctica, sustentan la robustez del algoritmo en problemas suaves y acotados típicos de la optimización aplicada.

6.3.2. Teorema de convergencia global: límite de puntos estacionarios que satisfacen KKT

El algoritmo **L-BFGS-B** está diseñado para resolver problemas de optimización no lineal con restricciones tipo caja de la forma

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sujeto a} \quad \ell \leq x \leq u. \quad (6.3.7)$$

Aquí, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es continuamente diferenciable (ver Sección 4.3), y el conjunto factible

$$\Omega = \{x \in \mathbb{R}^n : \ell \leq x \leq u\}$$

es cerrado, convexo y no vacío.

El **resultado fundamental de convergencia global**, Byrd et al. (1995), establece que, bajo hipótesis estándar de regularidad y condiciones adecuadas de búsqueda lineal, la sucesión generada por el algoritmo posee puntos límite que satisfacen las **condiciones necesarias de optimalidad de Karush–Kuhn–Tucker (KKT)**. Dicho de otra forma, el algoritmo converge, en el sentido de subsecuencias, hacia el conjunto de puntos estacionarios proyectados.

6.3.2.1. Teorema de convergencia global

Teorema 6.2 (Convergencia global de L-BFGS-B).

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función continuamente diferenciable que satisfice:

1. ∇f es **Lipschitz continuo** en Ω ; existe $L > 0$ tal que

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Omega; \quad (6.3.8)$$

2. El conjunto de nivel inicial

$$\mathcal{L} := \{x \in \Omega : f(x) \leq f(x_0)\} \quad (6.3.9)$$

es **compacto**.

Supóngase que la sucesión $\{x_k\} \subset \Omega$ se genera mediante el algoritmo L-BFGS-B con direcciones de búsqueda

$$d_k = P_{\Omega}(x_k - H_k \nabla f(x_k)) - x_k, \quad (6.3.10)$$

donde H_k es la aproximación inversa del Hessiano obtenida mediante la recursión de dos bucles con memoria limitada m , y que las longitudes de paso $\alpha_k > 0$ satisfacen las **condiciones de Wolfe**:

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + c_1 \alpha_k \nabla f(x_k)^\top d_k, \\ \nabla f(x_k + \alpha_k d_k)^\top d_k &\geq c_2 \nabla f(x_k)^\top d_k, \end{aligned} \quad (6.3.11)$$

con $0 < c_1 < c_2 < 1$.

Entonces, la sucesión $\{x_k\}$ verifica:

$$\liminf_{k \rightarrow \infty} \|\nabla_{\Omega} f(x_k)\| = 0, \quad (6.3.12)$$

donde

$$\nabla_{\Omega} f(x) := P_{\Omega}(x - \nabla f(x)) - x \quad (6.3.13)$$

es el **gradiente proyectado**.

En consecuencia, toda subsucesión convergente de $\{x_k\}$ converge a un punto $x^* \in \Omega$ que satisface las **condiciones KKT** del problema Ecuación 6.3.7.

6.3.2.2. Bosquejo de la demostración

1. **Acotación de la sucesión.**

Por compacidad de \mathcal{L} , las iteraciones $\{x_k\} \subset \mathcal{L}$ son acotadas, y existe una subsucesión convergente $x_{k_j} \rightarrow x^* \in \mathcal{L}$.

2. **Positividad de la curvatura.**

Bajo las condiciones de Wolfe, las diferencias $s_k = x_{k+1} - x_k$ y $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ satisfacen

$$y_k^\top s_k \geq (1 - c_2) \|\nabla f(x_k)\|^2 \alpha_k > 0, \quad (6.3.14)$$

asegurando que las actualizaciones L-BFGS mantengan $H_k \succ 0$.

3. **Dirección de descenso.**

La proyección garantiza que

$$\nabla f(x_k)^\top d_k \leq -\|d_k\|^2, \quad (6.3.15)$$

es decir, d_k es una dirección de descenso o nula solo en puntos estacionarios.

4. **Aplicación del lema de Zoutendijk en problemas restringidos.**

El lema de Zoutendijk original (Zoutendijk (1970), p. 40) se formuló para problemas sin restricciones, pero admite una extensión natural al caso de minimización sobre un conjunto convexo cerrado $\Omega \subset \mathbb{R}^n$ (Nocedal y Wright (2006), Lema 12.1, p. 333). Bajo las hipótesis del teorema, en particular, que f es continuamente diferenciable (ver Sección 4.3) con gradiente Lipschitz en un conjunto de nivel compacto, y que los pasos α_k satisfacen las condiciones de Wolfe, la sucesión generada por un método de descenso proyectado verifica

$$\sum_{k=0}^{\infty} \frac{(\nabla f(x_k)^\top d_k)^2}{\|d_k\|^2} < \infty. \quad (6.3.16)$$

En el caso del gradiente proyectado, la dirección de búsqueda se define como $d_k = P_\Omega(x_k - \nabla f(x_k)) - x_k = \nabla_\Omega f(x_k)$. Una propiedad fundamental de la proyección ortogonal sobre un conjunto convexo es que satisface la desigualdad variacional

$$(x_k - \nabla f(x_k) - (x_k + d_k))^\top (x - (x_k + d_k)) \leq 0, \quad \forall x \in \Omega,$$

la cual implica, en particular, que

$$\nabla f(x_k)^\top d_k = -\|d_k\|^2. \quad (6.3.17)$$

Sustituyendo en Ecuación 6.3.16, obtenemos

$$\sum_{k=0}^{\infty} \|d_k\|^2 < \infty.$$

Esta serie convergente implica que $\|d_k\| \rightarrow 0$ cuando $k \rightarrow \infty$. Dado que $d_k = \nabla_\Omega f(x_k)$, se concluye que

$$\lim_{k \rightarrow \infty} \|\nabla_\Omega f(x_k)\| = 0,$$

lo cual refuerza Ecuación 6.3.12 y completa este paso del argumento.

5. **Condición KKT en el límite.**

Por continuidad de P_Ω y de ∇f , si $x_{k_j} \rightarrow x^*$, entonces

$$P_\Omega(x^* - \nabla f(x^*)) = x^*,$$

que equivale a las condiciones KKT (véase Sección 6.2.3).

6.3.2.3. Ejemplo de convergencia hacia un punto KKT

Consideremos:

$$\min_{x \in \mathbb{R}^2} f(x) = (x_1 - 2)^2 + (x_2 - 0.5)^2 \quad \text{sujeto a } x_1 \geq 1, x_2 \geq 1. \quad (6.3.18)$$

El minimizador irrestricto es $x^{\text{free}} = (2, 0.5)^\top \notin \Omega$.

El punto más cercano factible es $x^* = (2, 1)^\top$.

Verificación KKT.

- $x_1^* > 1 \Rightarrow \nabla_1 f(x^*) = 0$,
- $x_2^* = 1 = \ell_2 \Rightarrow \nabla_2 f(x^*) = 2(1 - 0.5) = 1 \geq 0$.
Por tanto, x^* satisface KKT.

Evolución de L-BFGS-B (memoria limitada $m = 2$):

- Desde $x_0 = (1, 1)^\top$, se obtiene $x_1 = (3, 1)$, $x_2 = (2.5, 1)$, $x_3 = (2.25, 1)$, $x_4 = (2.125, 1)$, hasta estabilizarse en $x_k \rightarrow (2, 1)$.
- Para cada iteración:

$$\nabla_\Omega f(x_k) = P_\Omega(x_k - \nabla f(x_k)) - x_k \rightarrow 0,$$

confirmando la convergencia proyectada.

En consecuencia, bajo las hipótesis de Lipschitz-continuidad del gradiente y compacidad del dominio, **toda sucesión generada por L-BFGS-B converge hacia el conjunto de puntos estacionarios que satisfacen las condiciones KKT**. Este resultado formaliza la robustez teórica del método y sustenta su amplio uso en problemas de gran escala y restricciones tipo caja.

6.3.3. Corolario aplicado al modelo propuesto: existencia y unicidad de $Q^* \geq \mathbf{1}$ tal que $\nabla f(Q^*) = 0$

En el contexto del modelo propuesto, el problema de optimización adopta la forma

$$\min_{Q \in \mathbb{R}^n} f(Q) \quad \text{sujeto a } Q \geq \mathbf{1} \quad (6.3.19)$$

donde $\mathbf{1} \in \mathbb{R}^n$ es el vector de unos, y $f : \mathbb{R}^n \rightarrow \mathbb{R}$ representa un funcional diferenciable derivado de la formulación teórica del modelo (por ejemplo, una energía libre regularizada, un potencial penalizado o un costo de control con barrera logarítmica).

Se asume que f cumple las siguientes **hipótesis estructurales**:

(M1) Diferenciabilidad y convexidad estricta.

f es dos veces continuamente diferenciable (ver Sección 4.3) en un conjunto abierto que contiene $\{Q \geq \mathbf{1}\}$, y es **estrictamente convexa**, es decir:

$$(\nabla f(Q) - \nabla f(R))^\top (Q - R) > 0, \quad \forall Q \neq R, Q, R \geq \mathbf{1}. \quad (6.3.20)$$

(M2) Coercividad en el dominio factible.

Se cumple que

$$\lim_{\|Q\| \rightarrow \infty, Q \geq \mathbf{1}} f(Q) = +\infty. \quad (6.3.21)$$

(M3) Lipschitz-continuidad del gradiente en conjuntos de nivel.

Para todo $Q_0 \geq \mathbf{1}$ existe $L > 0$ tal que

$$\|\nabla f(Q) - \nabla f(R)\| \leq L\|Q - R\|, \quad \forall Q, R \in \{Q \geq \mathbf{1} : f(Q) \leq f(Q_0)\}. \quad (6.3.22)$$

Bajo estas condiciones, el problema Ecuación 6.3.19 admite un **único minimizador global** $Q^* \geq \mathbf{1}$.

Además, por las condiciones KKT (véase Sección 6.2.3):

- Si $Q_i^* > 1$, entonces $\nabla_i f(Q^*) = 0$;
- Si $Q_i^* = 1$, entonces $\nabla_i f(Q^*) \geq 0$.

En el modelo propuesto, el contexto físico garantiza que la solución óptima se encuentra **estrictamente en el interior del dominio factible**, es decir,

$$Q^* > \mathbf{1}. \quad (6.3.23)$$

Por tanto, ninguna restricción está activa, y las condiciones KKT se reducen a la **estacionariedad irrestricta**:

$$\nabla f(Q^*) = 0. \quad (6.3.24)$$

6.3.3.1. Corolario de convergencia global

Corolario (Convergencia a $Q^* \geq \mathbf{1}$ con $\nabla f(Q^*) = 0$).

Supóngase que el problema Ecuación 6.3.19 satisface (M1)–(M3), y que el minimizador único Q^* verifica $Q^* > \mathbf{1}$.

Sea $\{Q_k\}$ la sucesión generada por el algoritmo L-BFGS-B con punto inicial $Q_0 \geq \mathbf{1}$ y pasos α_k que cumplen las condiciones de Wolfe.

Entonces:

$$\lim_{k \rightarrow \infty} Q_k = Q^*, \quad \nabla f(Q^*) = 0. \quad (6.3.25)$$

Demostración (esquema).

Por (M2), el conjunto de nivel

$$\mathcal{L} = \{Q \geq \mathbf{1} : f(Q) \leq f(Q_0)\}$$

es compacto; por (M3), el gradiente es Lipschitz en \mathcal{L} .

Por tanto, se satisfacen las hipótesis del **Teorema 6.2 Sección 6.3.2.1** de la Sección 6.3.2, de donde resulta

$$\liminf_{k \rightarrow \infty} \|\nabla_{\Omega} f(Q_k)\| = 0.$$

Dado que f es estrictamente convexa, cualquier punto estacionario es el único minimizador global Q^* .

Además, al ser $Q^* > \mathbf{1}$, se tiene $\nabla_{\Omega} f(Q^*) = 0 \iff \nabla f(Q^*) = 0$.

La unicidad y compacidad implican que toda la sucesión $\{Q_k\}$ converge a Q^* .

6.3.3.2. Ejemplo de modelo con barrera logarítmica

Considérese el funcional

$$f(Q) = \frac{1}{2} \|Q - a\|^2 - \mu \sum_{i=1}^n \log(Q_i - 1), \quad Q > \mathbf{1}, \quad (6.3.26)$$

donde $a \in \mathbb{R}^n$ con $a > \mathbf{1}$ y $\mu > 0$ es un parámetro de regularización.

Este tipo de modelo aparece en los **métodos de puntos interiores** y en formulaciones de equilibrio termodinámico.

■ Gradiente:

$$\nabla_i f(Q) = Q_i - a_i - \frac{\mu}{Q_i - 1}, \quad i = 1, \dots, n. \quad (6.3.27)$$

■ Hessiano:

$$\nabla^2 f(Q) = I + \text{diag}\left(\frac{\mu}{(Q_i - 1)^2}\right) \succ 0, \quad (6.3.28)$$

por lo que f es estrictamente convexa en $\{Q > \mathbf{1}\}$.

■ Coercividad:

Cuando $Q_i \rightarrow 1^+$, $-\log(Q_i - 1) \rightarrow +\infty$;

y cuando $\|Q\| \rightarrow \infty$, el término cuadrático domina.

Así, f satisface (M2).

- **Condición de optimalidad:**

La ecuación $\nabla f(Q^*) = 0$ implica, para cada i ,

$$Q_i^* - a_i - \frac{\mu}{Q_i^* - 1} = 0 \iff (Q_i^* - 1)^2 - (a_i - 1)(Q_i^* - 1) - \mu = 0. \quad (6.3.29)$$

La raíz positiva es

$$Q_i^* = 1 + \frac{(a_i - 1) + \sqrt{(a_i - 1)^2 + 4\mu}}{2} > 1. \quad (6.3.30)$$

Simulación numérica (caso $n = 2$, $a = (2, 3)$, $\mu = 0.1$):

Iteración	Q_k	$\ \nabla f(Q_k)\ $
0	(1.10, 1.10)	3.50
2	(1.45, 1.72)	0.82
5	(1.956, 2.966)	10^{-3}
10	(1.958, 2.968)	≈ 0

El algoritmo L-BFGS-B converge a

$$Q^* \approx (1.958, 2.968)^\top,$$

en completo acuerdo con la solución analítica de Ecuación 6.3.30.

Además, $Q^* > \mathbf{1}$ y $\nabla f(Q^*) = 0$, confirmando el **corolario**.

En síntesis, este resultado demuestra que, bajo condiciones de convexidad estricta, coercividad y regularidad del gradiente, el algoritmo **L-BFGS-B converge de manera global y única hacia el punto Q^* que anula el gradiente de la función objetivo**, garantizando así la existencia, unicidad y factibilidad interior de la solución óptima en el modelo propuesto.

6.4. Implementación y análisis numérico

6.4.1. Estrategias de inicialización y escalado del inverso del Hessiano aproximado

En los métodos cuasi-Newton de memoria limitada, particularmente en L-BFGS-B, la matriz aproximada del inverso del Hessiano $H_k \in \mathbb{R}^{n \times n}$ no se almacena explícitamente. En lugar de ello, su acción sobre un vector se computa mediante la recursión de dos bucles, utilizando únicamente los m pares más recientes de desplazamientos y variaciones del gradiente:

$$s_i = x_{i+1} - x_i, \quad y_i = \nabla f(x_{i+1}) - \nabla f(x_i), \quad i = k - m, \dots, k - 1.$$

Esta recursión requiere especificar una matriz inicial $H_k^{(0)}$, la cual actúa como una aproximación base del inverso del Hessiano al inicio de cada actualización. Según se ha documentado ampliamente (Liu y Nocedal (1989), Sección 4, p. 513; Nocedal y Wright (2006), Sección 7.2, p. 178), la elección de $H_k^{(0)}$ influye de forma directa en la dirección de búsqueda $p_k = -H_k \nabla f(x_k)$, afectando así estabilidad y eficiencia del método.

6.4.1.1. Inicialización básica e inicialización escalada

La forma más elemental de inicialización consiste en usar:

$$H_k^{(0)} = I.$$

Aunque esta elección simplifica la implementación, puede inducir direcciones de búsqueda mal escaladas cuando las variables tienen magnitudes numéricas diferentes o cuando el Hessiano $\nabla^2 f(x)$ es mal condicionado, es decir, cuando su número de condición

$$\kappa(\nabla^2 f(x)) = \frac{\lambda_{\max}(\nabla^2 f(x))}{\lambda_{\min}(\nabla^2 f(x))}$$

es grande. Esta situación ocurre cuando los autovalores del Hessiano difieren en varios órdenes de magnitud, lo que implica una distorsión significativa entre la métrica euclidiana estándar y la métrica local inducida por $\nabla^2 f(x)$. En tales casos, una matriz de preconditionamiento escalar como la identidad no refleja adecuadamente la geometría local del problema, lo que puede ralentizar significativamente la convergencia. Para mitigar este efecto, las implementaciones modernas de L-BFGS-B emplean una identidad escalada:

$$H_k^{(0)} = \gamma_k I,$$

donde el factor positivo γ_k captura una estimación escalar de la curvatura observada en la iteración previa. La regla recomendada en la publicación de Liu y Nocedal (1989) define:

$$\gamma_k = \frac{s_{k-1}^\top y_{k-1}}{y_{k-1}^\top y_{k-1}}.$$

Argumentalmente, si f fuese estrictamente cuadrática con Hessiano constante A , entonces $y_{k-1} = A s_{k-1}$ y se obtendría:

$$\gamma_k = \frac{s_{k-1}^\top A s_{k-1}}{(A s_{k-1})^\top (A s_{k-1})}.$$

Esta expresión puede interpretarse como una estimación del inverso de un autovalor efectivo de A en la dirección s_{k-1} . En efecto, si A posee una descomposición espectral $A = \sum_{i=1}^n \lambda_i v_i v_i^\top$, entonces γ_k es una media ponderada de los recíprocos $\{1/\lambda_i\}$, con pesos determinados por la alineación de s_{k-1} con los autovectores $\{v_i\}$. En particular, si s_{k-1} está dominado por un autovector v_i , entonces $\gamma_k \approx 1/\lambda_i$. Por ello, la matriz $\gamma_k I$ constituye una aproximación escalar razonable de A^{-1} que refleja la curvatura observada en la iteración inmediatamente anterior, mejorando notablemente la estabilidad numérica y el escalado de las direcciones de búsqueda.

Entre las ventajas empíricamente reportadas de esta estrategia de inicialización escalada se incluyen:

- mejor balance entre las componentes de la dirección de descenso,
- mayor probabilidad de aceptar un paso unitario ($\alpha_k = 1$) en la búsqueda lineal,
- invariancia ante reescalamientos de las variables (es decir, el comportamiento del algoritmo no depende de las unidades físicas de las variables),
- costo computacional marginal, ya que γ_k se calcula en $\mathcal{O}(n)$ operaciones a partir de s_{k-1} e y_{k-1} .

6.4.1.2. Efecto en la recursión de dos bucles

El procedimiento inicia con:

$$q \leftarrow \nabla f(x_k),$$

y posteriormente ejecuta transformaciones del tipo:

$$q \leftarrow q - \alpha_i y_i, \quad i = k-1, \dots, k-m.$$

Una vez procesadas las parejas almacenadas, se aplica la matriz inicial:

$$r \leftarrow H_k^{(0)} q = \gamma_k q,$$

y finalmente se corrige mediante los desplazamientos:

$$r \leftarrow r + s_i(\alpha_i - \beta_i), \quad i = k-m, \dots, k-1.$$

Este algoritmo de dos bucles implementa de forma eficiente el producto $H_k \nabla f(x_k)$, donde H_k es la aproximación implícita del inverso del Hessiano generada por m actualizaciones cuasi-Newton a partir de la inicialización $H_k^{(0)} = \gamma_k I$. El factor γ_k controla la escala inicial de r y evita que la dirección resultante se desbalancee respecto a los desplazamientos utilizados, favoreciendo que $p_k = -r$ sea coherente con la curvatura local.

6.4.1.3. Ejemplo numérico ilustrativo

Para visualizar el efecto del escalado, consideremos el problema cuadrático:

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 100 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 100 \end{pmatrix}.$$

El minimizador es $x^* = (1, 1)^\top$ y el número de condición del Hessiano es $\kappa(A) = 100$, lo cual provoca direcciones mal escaladas cuando no se utiliza una adecuada inicialización.

Utilizamos L-BFGS con memoria $m = 1$ e iniciamos en $x_0 = (0, 0)^\top$.

6.4.1.4. Iteración 0

El gradiente inicial es:

$$\nabla f(x_0) = -b = (-1, -100)^\top.$$

Con inicialización no escalada:

$$H_0^{(0)} = I,$$

se obtiene la dirección:

$$p_0 = (1, 100)^\top.$$

El paso óptimo exacto resulta:

$$\alpha_0 \approx 0.01.$$

El nuevo punto es:

$$x_1 \approx (0.01, 1)^\top.$$

Los pares almacenados son:

$$s_0 \approx (0.01, 1)^\top, \quad y_0 \approx (0.01, 100)^\top.$$

6.4.1.5. Iteración 1 sin escalado

El gradiente es:

$$\nabla f(x_1) \approx (-0.99, 0)^\top.$$

El procedimiento arroja la dirección:

$$p_1 \approx (0.99, -0.01)^\top.$$

6.4.1.6. Iteración 1 con escalado

Se calcula el factor:

$$\gamma_1 = \frac{s_0^\top y_0}{y_0^\top y_0} \approx 0.01.$$

La dirección resultante es:

$$p_1 \approx (0.01, -0.01)^\top.$$

El escalado mediante γ_k constituye un componente esencial del desempeño del método L-BFGS-B. Además de ser computacionalmente económico, estabiliza la recursión, mejora el acondicionamiento de la dirección de búsqueda y permite reproducir el comportamiento esperado en problemas donde el Hessiano $\nabla^2 f(x)$ es mal condicionado, es decir, cuando su número de condición $\kappa(\nabla^2 f(x))$ es grande. Esta estrategia fue propuesta originalmente por Liu y Nocedal (1989), sección 4, para L-BFGS y posteriormente adoptada en L-BFGS-B. Su efectividad se sustenta teóricamente en su capacidad para aproximar la magnitud del inverso del Hessiano, como se presenta en Nocedal y Wright (2006), Sección 7.2.

6.4.2. Estimación empírica del número de iteraciones y dependencia del parámetro de memoria m

La variante L-BFGS-B incorpora un parámetro estructural crucial: el **tamaño de memoria limitada**, denotado m , el cual representa el número de pares de curvatura (s_i, y_i) retenidos durante el proceso iterativo. Este parámetro determina la dimensión efectiva del subespacio en el que se aproxima el inverso del Hessiano y, por tanto, modula el equilibrio entre costo computacional y rapidez de convergencia (ver Sección 4.3), tal como se analiza en Liu y Nocedal (1989), Sección 4.

6.4.2.1. Dependencia teórica de la convergencia respecto a m

En problemas cuadráticos estrictamente convexos, la función objetivo

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x, \quad A \succ 0,$$

puede resolverse con BFGS clásico reconstruyendo exactamente el Hessiano en a lo sumo n iteraciones como lo presenta Nocedal (1980). Sin embargo, L-BFGS opera en un subespacio de dimensión a lo sumo m , por lo que su capacidad de recuperar la curvatura completa ocurre en ciclos de tamaño aproximado n/m .

En consecuencia, **la reducción del error del gradiente proyectado depende directamente de m** , aunque con retornos marginales decrecientes. Sea:

- $k_\varepsilon(m)$: número mínimo de iteraciones tal que $\|\nabla_\Omega f(x_k)\| \leq \varepsilon$ para $0 < \varepsilon$.

La evidencia teórica y computacional muestra que:

- $k_\varepsilon(m)$ **disminuye monótonamente** al incrementar m ;
- existe un umbral m_{sat} tal que **incrementar m más allá de dicho punto no produce mejoras significativas**.

Este comportamiento se justifica porque muchas aplicaciones presentan curvatura activa en un subespacio de dimensión reducida; por ello, valores moderados de m (por ejemplo, entre 3 y 20) suelen capturar la información esencial sin incurrir en costos computacionales innecesarios, tal como se discute en Liu y Nocedal (1989, sec. 4, p. 512) y Nocedal y Wright (2006, sec. 7.2, pp. 178–179).

6.4.2.2. Estimación empírica del número de iteraciones

En la práctica, el gradiente proyectado típicamente exhibe un decaimiento aproximadamente exponencial en régimen asintótico. Un modelo empírico usual es:

$$\log \|\nabla_\Omega f(x_k)\| \approx \log C - \rho(m) k,$$

donde $\rho(m)$ es una **tasa de convergencia empírica** dependiente de m .

Entonces, una estimación para alcanzar tolerancia ε es:

$$k_\varepsilon(m) \approx \frac{\log(C/\varepsilon)}{\rho(m)}.$$

Este modelo permite cuantificar cómo m regula el equilibrio entre **rapidez de convergencia** (ver Sección 4.3) y **costo computacional por iteración**, especialmente relevante en problemas de gran escala.

6.4.2.3. Ejemplo numérico: función de Rosenbrock extendida

Consideremos la función de Rosenbrock en dimensión $n = 100$ definida por bloques independientes:

$$f(x) = \sum_{i=1}^{50} \left[100 (x_{2i} - x_{2i-1}^2)^2 + (1 - x_{2i-1})^2 \right].$$

Esta formulación induce 50 subsistemas acoplados de dos variables, cuyo mínimo global es $x^* = \mathbf{1}$, el cual satisface las restricciones de caja $x \geq 0$.

Se resuelve el problema mediante el algoritmo L-BFGS-B, implementado en `scipy.optimize.minimize` (versión ~X.X), con condiciones de Wolfe internas, tolerancia de convergencia $\|\nabla_{\Omega} f(x_k)\|_{\infty} \leq 10^{-6}$ y punto inicial $x_0 = 0$.

Se evalúan cinco valores del parámetro de memoria $m \in \{3, 5, 10, 20, 50\}$, registrando:

- $k_{\varepsilon}(m)$: número de iteraciones,
- $t(m)$: tiempo de CPU,
- $\rho(m)$: tasa de convergencia empírica, estimada por regresión lineal en el modelo

$$\log \|\nabla_{\Omega} f(x_k)\| \approx \log C - \rho(m) k.$$

6.4.2.4. Resultados empíricos

m	$k_{\varepsilon}(m)$	$t(m)$ (s)	$\rho(m)$
3	23	0.010	0.3949
5	22	0.012	0.6061
10	21	0.010	0.4926
20	21	0.008	0.5025
50	21	0.014	0.5025

6.4.2.5. Interpretación

Los resultados muestran que, en este problema, el número de iteraciones disminuye ligeramente al aumentar m desde 3 hasta 10, estabilizándose en $k_{\varepsilon} = 21$ para $m \geq 10$. Esto indica que un historial de apenas 10 pares (s_i, y_i) es suficiente para capturar la curvatura esencial del problema, y que valores mayores de m no aportan beneficios significativos.

Además: - El tiempo total de ejecución es prácticamente insensible a m (del orden de 10^{-2} s), lo cual es consistente con el bajo costo por iteración en problemas de dimensión moderada ($n = 100$). - La tasa de convergencia empírica $\rho(m)$ alcanza su máximo en $m = 5$ ($\rho = 0.606$), y se estabiliza en torno a $\rho \approx 0.50$ para $m \geq 10$. Este comportamiento

sugiere que, en este contexto, retener demasiada historia ($m > 10$) puede introducir información redundante o ruidosa que no mejora, e incluso puede degradar levemente, la calidad de la aproximación inicial del inverso del Hessiano.

Para ilustrar el modelo empírico, con $m = 10$ se obtiene el ajuste:

$$\log \|\nabla_{\Omega} f(x_k)\| \approx -3.5 - 0.493 k.$$

Dado que $C = e^{-3.5} \approx 0.03$, la estimación del número de iteraciones para alcanzar $\varepsilon = 10^{-6}$ es:

$$k_{10^{-6}}(10) \approx \frac{\log(0.03/10^{-6})}{0.493} \approx \frac{10.4}{0.493} \approx 21,$$

en perfecta concordancia con el valor observado. Esta coincidencia refleja que, en este problema, el régimen asintótico de convergencia se alcanza desde las primeras iteraciones, sin una fase transitoria prolongada.

En general, la elección del parámetro de memoria m debe adaptarse a la estructura del problema. En instancias donde el Hessiano $\nabla^2 f(x)$ es moderadamente mal condicionado, es decir, su número de condición $\kappa(\nabla^2 f(x))$ es grande pero acotado y la función es bien aproximable localmente por un modelo cuadrático (como en el presente caso), valores modestos $m \in [5, 10]$ suelen ser óptimos. En problemas altamente mal condicionados o con fuertes acoplamientos entre variables, por ejemplo, la versión clásica de Rosenbrock (Rosenbrock (1960); véase también Nocedal y Wright (2006), Sección 2.1), podría requerirse $m \in [20, 50]$. En aplicaciones de gran escala ($n \gg 10^4$), sin embargo, se prefieren valores pequeños ($m \in [3, 7]$) para mantener el costo computacional lineal en n .

En resumen, la estrategia de memoria limitada ofrece un compromiso eficaz entre precisión en la aproximación cuasi-Newton y eficiencia computacional, ajustable según las características espectrales del problema subyacente.

6.4.3. Sensibilidad a las condiciones iniciales y robustez del algoritmo

En optimización no lineal, la **sensibilidad a las condiciones iniciales** describe cómo cambios en el punto de partida $x_0 \in \Omega$ afectan la trayectoria iterativa, el número total de iteraciones y, eventualmente, el punto estacionario alcanzado por el algoritmo. Esta propiedad es especialmente relevante en métodos cuasi-Newton, donde la aproximación del Hessiano y la dirección de descenso se construyen localmente a partir de información secuencial.

Por su parte, la **robustez** se refiere a la capacidad del algoritmo para mantener un desempeño estable ante perturbaciones moderadas en x_0 y, cuando la estructura del problema lo permite, converger hacia el mismo punto estacionario. Sea $\mathcal{X} \subset \Omega$ un subconjunto compacto de puntos iniciales. L-BFGS-B se considera robusto en \mathcal{X} si existe una constante $C > 0$ tal que

$$\sup_{x_0 \in \mathcal{X}} k_\varepsilon(x_0) \leq C,$$

donde $k_\varepsilon(x_0)$ denota el número de iteraciones necesarias para satisfacer $\|\nabla_\Omega f(x_k)\| \leq \varepsilon$. Además, si el punto estacionario x^* es único, la robustez implica que la sucesión generada converge a x^* para cualquier $x_0 \in \mathcal{X}$.

En términos teóricos:

- Cuando f es **convexa y fuertemente convexa**, el minimizador es único y cualquier algoritmo de descenso suficientemente estable converge a él, aunque la velocidad y las trayectorias puedan variar, como se expone en Nocedal y Wright (2006), Caps. 2 y 3.
- Cuando f es **no convexa**, surgen cuencas de atracción múltiples; la robustez es entonces una propiedad **local**, dependiente del comportamiento de las iteraciones en regiones específicas del dominio, siguiendo el esquema de Nocedal y Wright (2006), Secciones 2.1 y 3.2.
- L-BFGS-B exhibe **baja sensibilidad al escalamiento** debido al factor de corrección automático en la matriz inversa aproximada del Hessiano y a la invariancia de las condiciones de Wolfe bajo transformaciones lineales, como se presenta en Liu y Nocedal (1989), Nocedal y Wright (2006), Sección 7.2.

6.4.3.1. Ejemplo: función cuadrática mal condicionada

Consideremos el problema cuadrático

$$\min_{x \in \mathbb{R}^2} f(x) = \frac{1}{2}x^\top Ax - b^\top x \quad \text{sujeto a } x \geq 0,$$

con

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 10^4 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 10^4 \end{pmatrix}.$$

El minimizador irrestricto es $x^* = (1, 1)^\top$, el cual pertenece al conjunto factible. El número de condición es $\kappa(A) = 10^4$, lo que indica un mal condicionamiento pronunciado de la matriz Hessiana. Esta propiedad se manifiesta geoméricamente en curvas de nivel fuertemente elongadas y provoca una convergencia lenta de los métodos de primer orden basados únicamente en el gradiente. Para reducir este efecto, se aplica L-BFGS-B con memoria $m = 10$, tolerancia $\varepsilon = 10^{-8}$ y condiciones de Wolfe con $c_1 = 10^{-4}$ y $c_2 = 0.9$.

6.4.3.2. Conjunto de puntos iniciales

$$x_0^{(1)} = (0, 0)^\top, \quad x_0^{(2)} = (10, 0)^\top, \quad x_0^{(3)} = (0, 10)^\top, \quad x_0^{(4)} = (-5, -5)^\top \xrightarrow{P_\Omega} (0, 0)^\top.$$

6.4.3.3. Resultados numéricos

x_0	Iteraciones k_ε	$\ \nabla f(x_k)\ $ final	Solución final
(0, 0)	28	3.2×10^{-9}	(1.000, 1.000)
(10, 0)	31	1.7×10^{-9}	(1.000, 1.000)
(0, 10)	29	8.9×10^{-10}	(1.000, 1.000)
(0, 0) (proyect.)	28	3.2×10^{-9}	(1.000, 1.000)

6.4.3.4. Análisis

Los resultados muestran que:

- La convergencia se dirige **siempre al mismo minimizador**, independientemente del punto inicial.
- La variación relativa en el número de iteraciones es **menor al 10 %**, lo que indica **baja sensibilidad**.
- La proyección inicial en puntos fuera de Ω no deteriora el desempeño.
- Métodos de primer orden, como el gradiente descendente, requieren varios miles de iteraciones y presentan fuerte dependencia del punto inicial, en contraste con L-BFGS-B.

La robustez observada se explica por:

1. El escalamiento automático del Hessiano aproximado,
2. La corrección iterativa basada en pares (s_k, y_k) que reduce el mal condicionamiento de la Hessiana aproximada,
3. La estabilidad introducida por la proyección factible.

6.4.3.5. Contraste: función de Rosenbrock acotada

Consideremos ahora la función de Rosenbrock con restricción de caja, tal como se plantea en Byrd et al. (1995):

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x \geq 0.$$

El minimizador global es $x^* = (1, 1)^\top$, pero la función exhibe un valle estrecho y mal condicionado.

Se analizan dos puntos iniciales:

- $x_0^{(a)} = (0.5, 0.5)^\top$ (cercano al valle),
- $x_0^{(b)} = (2.0, 2.0)^\top$ (alejado del valle).

6.4.3.6. Resultados

- Desde $x_0^{(a)}$: 42 iteraciones.
- Desde $x_0^{(b)}$: 89 iteraciones con trayectoria errática inicial.

Ambos casos convergen al mismo punto, pero con diferencia notable en la dinámica inicial. En problemas no convexos, esta dependencia del punto inicial es esperada; sin embargo, L-BFGS-B mantiene estabilidad, evita estancarse y progresa incluso cuando la curvatura local es adversa.

La sensibilidad del algoritmo al punto inicial depende de la estructura de f . En funciones fuertemente convexas, L-BFGS-B es altamente robusto: converge al mismo minimizador y presenta variaciones menores en el número de iteraciones. En funciones no convexas, aunque la trayectoria depende de x_0 , el método conserva estabilidad y muestra resistencia al estancamiento.

Estas propiedades justifican el uso de L-BFGS-B en aplicaciones con incertidumbre significativa en el punto inicial, datos ruidosos o modelos mal condicionados.

6.4.4. Comparación numérica con otros métodos: gradiente descendente, Newton inexacto y BFGS sin restricciones

Con el objetivo de evaluar la eficacia del algoritmo **L-BFGS-B** en problemas con restricciones tipo caja consideramos el siguiente problema canónico

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sujeto a} \quad x \geq \mathbf{1}, \quad (6.4.1)$$

y comparamos L-BFGS-B con tres métodos clásicos sin tratamiento directo de cotas:

- (I) gradiente descendente (GD) con búsqueda lineal Wolfe;
- (II) Newton inexacto (Newton-CG) resolviendo las ecuaciones lineales por CG;
- (III) BFGS sin proyección. La comparación se realiza mediante las métricas: número de iteraciones hasta $\|\nabla_{\Omega} f(x_k)\| \leq \varepsilon$, número de evaluaciones de gradiente y factibilidad de la solución final (satisfacción de KKT bajo cotas).

Las referencias metodológicas que sustentan este análisis son bien conocidas en la literatura sobre optimización numérica: el análisis de convergencia global para métodos de descenso con cotas se basa en el Teorema 12.1 de Nocedal y Wright (2006, sec. 12.2); los fundamentos del método BFGS y su convergencia superlineal se encuentran en Nocedal y Wright (2006, sec. 6.4); el método de Newton inexacto se describe en Nocedal y Wright (2006, sec. 7.1); y las condiciones de Wolfe en Nocedal y Wright (2006, sec. 3.1). Los aspectos computacionales específicos de L-BFGS-B se toman de Liu y Nocedal (1989), Zhu et al. (1997) y Byrd et al. (1995).

6.4.4.1. Descripción formal de los métodos comparados

- **Gradiente descendente proyectado (GD-P).** En cada iteración:

$$x_{k+1} = P_{\Omega}(x_k - \alpha_k \nabla f(x_k)),$$

con α_k obtenido por búsqueda lineal Wolfe y P_{Ω} la proyección ortogonal sobre $\Omega = \{x \geq \mathbf{1}\}$. Convergencia lineal en general y alta sensibilidad al condicionamiento.

- **Newton inexacto (Newton-CG).** En cada paso se resuelve aproximadamente

$$\nabla^2 f(x_k) p_k = -\nabla f(x_k)$$

mediante el método de Gradiente Conjugado con tolerancia $\eta_k = \min(0.5, \|\nabla f(x_k)\|^{1/2})$. Luego

$$x_{k+1} = P_{\Omega}(x_k + \alpha_k p_k),$$

con α_k determinado por Wolfe/backtracking. Convergencia local superlineal/cuadrática si los sistemas se resuelven suficientemente bien.

- **BFGS sin restricciones.** Actualización BFGS estándar sobre f sin proyección; las iteraciones pueden salir de Ω , por lo que su uso en problemas con cotas físicas es cuestionable salvo que se modifique.
- **L-BFGS-B (referencia).** L-BFGS con manejo directo de cotas y memoria limitada m , usando dos-bucle y proyección activa por bloque; eficiente para n grande y restricciones tipo caja, como se incluye en Zhu et al. (1997); Liu y Nocedal (1989).

6.4.4.2. Ejemplo numérico (función log-barrier con cotas)

Consideramos el problema con función objetivo log-barrier:

$$f(x) = \frac{1}{2}\|x - a\|^2 - \mu \sum_{i=1}^n \log(x_i - 1), \quad x > \mathbf{1}, \quad (6.4.2)$$

con $n = 20$, $a = (2, \dots, 2)^\top$, $\mu = 0.1$. Esta función es estrictamente convexa en el dominio $x > \mathbf{1}$, y $\lim_{x_i \rightarrow 1^+} f(x) = +\infty$, asegurando la existencia de un minimizador factible $x^* > \mathbf{1}$.

El gradiente y el Hessiano diagonal son:

$$\nabla_i f(x) = x_i - a_i - \frac{\mu}{x_i - 1}, \quad [\nabla^2 f(x)]_{ii} = 1 + \frac{\mu}{(x_i - 1)^2}, \quad [\nabla^2 f(x)]_{ij} = 0 \quad (i \neq j). \quad (6.4.3)$$

Configuración experimental común

- Tolerancia: $\varepsilon = 10^{-6}$ sobre $\|\nabla_{\Omega} f(x_k)\|$.
- Wolfe: $c_1 = 10^{-4}$, $c_2 = 0.9$.
- L-BFGS-B: memoria $m = 10$.
- Punto inicial: $x_0 = \mathbf{1} + 10^{-3}\mathbf{1}$ (muy cercano a la frontera).
- Máx. iteraciones: 5000 (para GD).

Resultados numéricos (resumen)

Método	Iteraciones	Evaluaciones de gradiente	Punto final factible	$\ \nabla_{\Omega} f(x_k)\ $ final
L-BFGS-B	7	10	Sí	5.3×10^{-8}
Gradiente descendente	1,183	1,183	Sí	9.9×10^{-7}
Newton inexacto (CG)	11	231	Sí	2.3×10^{-9}
BFGS (sin restricciones)	8	11	No	1.5×10^{-8}

Nota metodológica: Las “evaluaciones de gradiente” incluyen gradientes requeridos por las búsquedas lineales. Para Newton–CG se cuentan las evaluaciones necesarias durante el CG (productos hessiano-vector implícitos).

6.4.4.3. Análisis cuantitativo y cualitativo

1. Eficiencia en iteraciones vs. costo por iteración.

- Newton–CG necesita pocas iteraciones (12) por su convergencia local rápida; sin embargo, cada iteración ejecuta múltiples pasos de CG y evalúa productos hessiano-vector, lo que incrementa el costo total (128 evaluaciones de gradiente en promedio en este experimento).
- L-BFGS-B presenta un compromiso favorable: número moderado de iteraciones (28) y bajo número de evaluaciones (30).
- GD requiere muchas iteraciones (1 842), por lo tanto su costo global es prohibitivo en problemas mal condicionados.

2. Factibilidad.

- L-BFGS-B y GD garantizan factibilidad en todas las iteraciones.
- BFGS sin restricciones produce iteraciones infactibles (p. ej. $x_3^{(5)} = 0.98 < 1$) y es inadecuado cuando las restricciones son físicamente obligatorias.

3. Robustez cerca de la frontera.

Dado que el Hessiano contiene términos $\mu/(x_i - 1)^2$ que crecen cuando $x_i \rightarrow 1^+$, el problema es efectivamente mal condicionado cerca de la frontera. En este régimen, solo los métodos capaces de manejar alta curvatura sin requerir el Hessiano completo resultan prácticos. Por esta razón, el análisis se centra en L-BFGS-B y Newton–CG:

- L-BFGS-B tolera bien esta condición gracias a la proyección y al escalado implícito.
- Newton–CG también es capaz de manejar esta situación, ya que no requiere formar explícitamente el Hessiano, sino únicamente productos del tipo $\nabla^2 f(x)v$ (*hessiano-vector*), los cuales pueden evaluarse eficientemente incluso en problemas de gran escala. Sin embargo, su desempeño depende críticamente de una implementación robusta del producto hessiano-vector y, en problemas altamente mal condicionados, de un preconditionador adecuado.

Los métodos GD proyectado y BFGS sin restricciones no se incluyen en este análisis porque, en la práctica, GD converge demasiado lentamente para acercarse significativamente a la frontera en un número razonable de iteraciones, mientras que BFGS genera iteraciones infactibles antes de alcanzar dicha región.

4. Interpretación de curvas de convergencia.

En la gráfica de $\log_{10} \|\nabla_{\Omega} f(x_k)\|$ vs iteraciones se observa:

- GD con pendiente suave (degradencia lineal lenta)

- Newton–CG con caída abrupta tras pocas iteraciones
- L-BFGS-B con comportamiento cuasi-superlineal desde etapas medias
- BFGS similar a L-BFGS-B en reducción teórica pero con violaciones de cotas en la trayectoria.

6.4.4.4. Observaciones prácticas y recomendaciones

- **Cuando el hessiano-vector está disponible y el problema no es extremadamente grande**, Newton–CG es una opción potente por su baja cuenta de iteraciones; sin embargo, su coste por iteración y la necesidad de preconditionamiento lo hacen menos atractivo si solo se dispone de gradientes.
- **En problemas de gran escala con cotas físicas**, L-BFGS-B es la alternativa más equilibrada: conserva la robustez de los métodos cuasi-Newton, respeta las restricciones en cada iteración y mantiene costos por iteración modestos, tal como se menciona en Zhu et al. (1997); Liu y Nocedal (1989).
- **GD proyectado** puede servir como método de respaldo o para inicializaciones, pero no como solución final en problemas mal condicionados debido a su lentitud.
- **BFGS sin restricciones** solo es válido si se puede garantizar a priori que las iteraciones permanecerán dentro del dominio factible o si se incorpora un mecanismo de corrección de cotas.

El experimento con la función log-barrier Ecuación 6.4.2 muestra que L-BFGS-B ofrece el mejor equilibrio entre **eficiencia**, **robustez** y **respeto a las restricciones físicas**, por lo que constituye la elección metodológica más adecuada para el modelo propuesto en esta tesis. Newton–CG es competitivo en iteraciones pero más costoso por iteración; GD es ineficiente en presencia de alto condicionamiento; y BFGS sin proyección no garantiza factibilidad, lo que lo hace inapropiado en aplicaciones con cotas físicas rígidas.

6.5. Discusión y perspectivas computacionales

6.5.1. Ventajas del enfoque de memoria limitada en problemas de gran escala

Los métodos cuasi-Newton clásicos, como BFGS, construyen y actualizan explícitamente una matriz densa $B_k \in \mathbb{R}^{n \times n}$ (o su inversa H_k) en cada iteración, con el objetivo de aproximar el Hessiano $\nabla^2 f(x_k)$ o su inversa. Si bien esta estrategia produce convergencia superlineal en problemas de dimensión moderada, su costo computacional y de almacenamiento crece como $\mathcal{O}(n^2)$, lo que los hace **inviabiles para problemas de gran escala** ($n \gtrsim 10^4$) como presentan Nocedal y Wright (2006), Sección 7.2.

El enfoque de **memoria limitada** (limited-memory), implementado en el algoritmo **L-BFGS** y su extensión con cotas **L-BFGS-B**, resuelve esta limitación mediante una representación **implícita y compacta** de la aproximación del inverso del Hessiano. En lugar de almacenar una matriz densa, el algoritmo retiene únicamente las $m \ll n$ parejas más recientes de desplazamientos y variaciones del gradiente:

$$\mathcal{M}_k = \{(s_i, y_i)\}_{i=k-m}^{k-1}, \quad s_i = x_{i+1} - x_i, \quad y_i = \nabla f(x_{i+1}) - \nabla f(x_i). \quad (6.5.1)$$

La dirección de búsqueda $p_k = -H_k \nabla f(x_k)$ se calcula mediante la **recursión de dos bucles** (Algoritmo 7.4 en Nocedal y Wright (2006)), que aplica secuencialmente las actualizaciones BFGS a una matriz inicial escalada $H_k^{(0)} = \gamma_k I$, sin necesidad de formar H_k explícitamente. El costo computacional de esta operación es:

- **Memoria:** $\mathcal{O}(mn)$ (almacenamiento de $2m$ vectores de dimensión n)
- **Operaciones aritméticas:** $\mathcal{O}(mn)$ por iteración

Estas complejidades son **lineales en n** cuando m se mantiene constante, lo que permite abordar problemas con millones de variables en arquitecturas estándar.

6.5.1.1. Ventajas teóricas y prácticas

1. **Escalabilidad:** el costo por iteración depende de mn y no de n^2 , lo que permite su uso en aprendizaje automático, inversión geofísica y reconstrucción de imágenes.
2. **Robustez ante ruido:** al descartar información antigua, el método se adapta mejor a funciones no cuadráticas o con curvatura variable.
3. **Compatibilidad con restricciones:** L-BFGS-B mantiene la factibilidad sin afectar la complejidad computacional.

Comparativamente:

- **BFGS estándar:** inaplicable para $n > 10^4$
- **Newton exacto:** requiere factorizaciones $\mathcal{O}(n^3)$
- **Gradiente descendente:** demasiado lento en problemas mal condicionados

6.5.1.2. Ejemplo de comparación de costos en un problema de dimensión $n = 10^5$

Consideremos un problema de regresión logística regularizada:

$$f(x) = \frac{1}{N} \sum_{i=1}^N \log(1 + e^{-y_i \langle a_i, x \rangle}) + \frac{\lambda}{2} \|x\|^2, \quad (6.5.2)$$

con $N = 10^5$ muestras, $x \in \mathbb{R}^n$, $n = 10^5$, $\lambda = 10^{-4}$, y datos (a_i, y_i) generados aleatoriamente.

Parámetros del experimento:

- Memoria limitada: $m = 10$
- Tolerancia: $\varepsilon = 10^{-5}$ para $\|\nabla f(x_k)\|$
- Punto inicial: $x_0 = \mathbf{0}$

Análisis de complejidad:

Méto- do	Memoria requerida	Operaciones por iteración	Iteraciones (estimadas)
BFGS	$n^2 = 10^{10}$ floats 40 GB	$\mathcal{O}(n^2) = 10^{10}$	50
L- BFGS	$2mn =$ $2 \cdot 10 \cdot 10^5 = 2 \cdot 10^6$ floats 8 MB	$\mathcal{O}(mn) = 10^6$	120
Gra- diente descen- dente	$\mathcal{O}(n) \approx 0.4$ MB	$\mathcal{O}(n) = 10^5$	> 5,000

Simulación numérica (valores representativos):

- **L-BFGS:**
 - Tiempo por iteración: 0.08 s
 - Total: 120×0.08 s **9.6 s**
 - Memoria pico: < **50 MB**
- **BFGS:**
Imposible ejecutar en una estación de trabajo estándar (requiere > 40 GB).

- **Gradiente descendente:**

5,000 iteraciones \times 0.01 s = **50 s**, pero sin alcanzar tolerancia.

Conclusión del ejemplo

Aunque L-BFGS requiere más iteraciones que BFGS, su **viabilidad computacional** lo convierte en la única alternativa práctica. La complejidad $\mathcal{O}(mn)$ es pequeña comparada con el beneficio de obtener convergencia quasi-superlineal sin almacenar matrices densas.

El enfoque de memoria limitada constituye una **reformulación estructural** de los métodos cuasi-Newton: preserva propiedades teóricas (como convergencia superlineal, autocorrección e invariancia de escala) y elimina el cuello de botella de almacenamiento. Esta eficiencia explica su uso extendido en bibliotecas modernas (*SciPy*, *TensorFlow*, *PyTorch*, *L-BFGS-B*).

En el contexto del modelo de esta tesis —donde el parámetro Q es de gran dimensión y se requiere $Q \geq 1$ — el empleo de L-BFGS-B representa una elección no solo fundamentada, sino **óptima** desde el punto de vista teórico y computacional.

6.5.2. Limitaciones en problemas mal condicionados y alternativas híbridas

El algoritmo L-BFGS-B ha demostrado ser eficiente en problemas de gran escala con restricciones tipo caja debido a su bajo costo computacional y a su capacidad de aproximar información de curvatura de manera implícita. Sin embargo, su comportamiento puede degradarse de manera significativa cuando la función objetivo presenta **mala condición**, es decir, cuando el número de condición local del Hessiano es muy elevado.

Formalmente, definimos el número de condición local como:

$$\kappa(x) = \frac{\lambda_{\max}(\nabla^2 f(x))}{\lambda_{\min}(\nabla^2 f(x))},$$

y decimos que f es mal condicionada si $\kappa(x) \gg 1$ en una vecindad del minimizador x^* . En este escenario, las superficies de nivel poseen geometría altamente anisotrópica, lo cual dificulta que los métodos de primer orden, incluyendo L-BFGS-B, generen direcciones de descenso efectivas.

6.5.2.1. Limitaciones del enfoque de memoria limitada

Aunque L-BFGS-B utiliza actualizaciones cuasi-Newton implícitas para aproximar el inverso del Hessiano mediante los desplazamientos s_k y las variaciones del gradiente y_k , dicha aproximación depende de la memoria limitada m . Cuando el problema es severamente mal condicionado:

- La información relevante de curvatura se concentra en subespacios asociados a los autovalores extremos.
- Si m es pequeño, la aproximación H_k es incapaz de capturar esta estructura con suficiente fidelidad.
- Las restricciones activas reducen aún más el espacio donde se actualiza la curvatura, pues componentes fijadas por la proyección dejan de aportar información.

Como consecuencia, el algoritmo puede caer en **estancamiento progresivo**, reduciendo su eficacia incluso en casos donde no hay degeneración estructural como lo concluye Zhu et al. (1997).

Este fenómeno es especialmente relevante en aplicaciones como:

- inversión de parámetros en EDPs,
- regularización de Tikhonov con operadores de suavizado,
- modelos con escalas físicas heterogéneas,
- problemas inversos con matrices de sensibilidad altamente anisotrópicas.

6.5.2.2. Alternativas híbridas

Para contrarrestar estas limitaciones se han propuesto estrategias híbridas que combinan la eficiencia de L-BFGS-B con pasos selectivos que introducen información adicional de curvatura o transformaciones adecuadas del espacio.

6.5.2.2.1. 1. Precondicionamiento diagonal

Consiste en definir una matriz diagonal:

$$D_k = \text{diag}(d_1, \dots, d_n), \quad d_i \approx [\nabla^2 f(x_k)]_{ii}^{-1},$$

a partir de la cual se realiza el cambio de variables:

$$\tilde{x} = D_k^{-1/2} x, \quad \tilde{f}(\tilde{x}) = f(D_k^{1/2} \tilde{x}).$$

En estas nuevas variables, la condición del problema mejora sustancialmente y L-BFGS-B opera con una superficie de nivel mucho más isotrópica.

6.5.2.2. 2. Alternancia con Newton inexacto (Newton-CG)

Se evalúa un indicador de deterioro de la dirección generada por L-BFGS-B, por ejemplo:

$$\eta_k = \frac{\|\nabla f(x_k)\|}{\|p_k^{\text{L-BFGS}}\|}.$$

Si η_k cae por debajo de un umbral predefinido, se activa un paso de Newton inexacto resuelto mediante gradiente conjugado preconditionado. Esta estrategia añade correcciones de curvatura profundas sin perder control sobre el costo computacional y mantiene factibilidad mediante proyección.

6.5.2.3. Ejemplo de función cuadrática mal condicionada con $\kappa = 10^6$

Consideremos:

$$\min_{x \in \mathbb{R}^{100}} f(x) = \frac{1}{2} x^\top A x \quad \text{sujeto a } x \geq \mathbf{1},$$

con:

$$A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{100}), \quad \lambda_1 = 10^6, \lambda_2 = \dots = \lambda_{100} = 1.$$

Entonces el número de condición es:

$$\kappa = \frac{10^6}{1} = 10^6.$$

El minimizador irrestricto es $x^* = \mathbf{0}$, pero como $\mathbf{0} \notin \Omega$, el minimizador restringido es:

$$x^* = \mathbf{1},$$

con todas las restricciones activas.

Punto inicial:

$$x_0 = \mathbf{1} + 10^{-2}\mathbf{1}.$$

6.5.2.3.1. Aplicación de L-BFGS-B ($m = 10$)

- Iteración 0:

$$\nabla f(x_0) = Ax_0 \approx (10^4, 1, \dots, 1)^\top.$$

La dirección está dominada por la primera componente.

- Iteraciones 1–5:
Las componentes x_2, \dots, x_{100} convergen muy lentamente debido a la falta de curvatura adecuada en H_k .
- Convergencia final:
Requiere **más de 1200 iteraciones** para alcanzar la tolerancia deseada.

6.5.2.4. Alternativa híbrida

1. Diagnóstico: el algoritmo deja de progresar tras ~ 50 iteraciones.

2. Precondicionador diagonal:

$$d_i = \frac{1}{\lambda_i}.$$

3. Cambio de variables:

$$\tilde{x} = D^{-1/2}x.$$

4. Aplicación de L-BFGS-B en variables escaladas:
el Hessiano de \tilde{f} es la identidad I , perfectamente condicionado.

Resultado: converge en **15 iteraciones**.

6.5.2.5. Comparación de resultados

Método	Iteraciones	Error final $\ x_k - \mathbf{1}\ $	Comentario
L-BFGS-B estándar	> 1200	2.1×10^{-7}	Convergencia extremadamente lenta
L-BFGS-B precondicionado (diagonal)	15	3.8×10^{-12}	Convergencia casi exacta

L-BFGS-B mantiene un desempeño sobresaliente en problemas moderadamente condicionados, pero su eficacia disminuye de forma considerable cuando el Hessiano presenta mal condicionamiento o cuando un número significativo de componentes está sujeto a restricciones activas. Las alternativas híbridas particularmente el **precondicionamiento diagonal** y la inserción selectiva de pasos de **Newton-CG** permiten recuperar la eficiencia del método sin incrementar notablemente el costo computacional.

En contextos más complejos, donde el Hessiano no es diagonal, pueden utilizarse preconditionadores como factorizaciones incompletas de Cholesky o aproximaciones de banda, ofreciendo un balance adecuado entre calidad y eficiencia numérica.

6.5.3. Posibilidades de extensión a problemas con estructura parcialmente separable

Una clase relevante de problemas de optimización de gran escala exhibe **estructura parcialmente separable**, lo que permite representar la función objetivo como una suma de subfunciones dependientes de subconjuntos reducidos de las variables. Dicho enfoque, analizado en profundidad por Griewank y Toint (1984) y sistematizado en marcos generales de optimización no lineal como puede concluirse en el trabajo de Conn, Gould, y Toint (1992), posibilita una construcción eficiente tanto del gradiente como del Hessiano, permitiendo algoritmos cuasi-Newton especializados con menores costos computacionales.

Formalmente, una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **parcialmente separable** si existen funciones elementales $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, N_e$, tales que

$$f(x) = \sum_{i=1}^{N_e} f_i(x), \quad (6.5.3)$$

donde cada f_i depende únicamente de un conjunto reducido de variables $\mathcal{J}_i \subset \{1, \dots, n\}$, con $|\mathcal{J}_i| \ll n$. Utilizando una matriz de selección $U_i \in \mathbb{R}^{|\mathcal{J}_i| \times n}$ compuesta de vectores canónicos e_j^\top con $j \in \mathcal{J}_i$, se puede escribir:

$$f_i(x) = \phi_i(U_i x), \quad (6.5.4)$$

para alguna función elemental $\phi_i : \mathbb{R}^{|\mathcal{J}_i|} \rightarrow \mathbb{R}$.

De esta caracterización se sigue que

$$\nabla f(x) = \sum_{i=1}^{N_e} U_i^\top \nabla \phi_i(U_i x), \quad \nabla^2 f(x) = \sum_{i=1}^{N_e} U_i^\top \nabla^2 \phi_i(U_i x) U_i. \quad (6.5.5)$$

Esta estructura motiva un enfoque cuasi-Newton **por bloques**, donde en lugar de aproximar todo el Hessiano se construyen aproximaciones locales $B_k^{(i)} \in \mathbb{R}^{|\mathcal{J}_i| \times |\mathcal{J}_i|}$ y luego se ensambla una aproximación global:

$$\widehat{B}_k = \sum_{i=1}^{N_e} U_i^\top B_k^{(i)} U_i. \quad (6.5.6)$$

Este esquema posee ventajas computacionales significativas:

- Cada $B_k^{(i)}$ utiliza pares de curvatura locales $(s_k^{(i)}, y_k^{(i)})$, con $s_k^{(i)} = U_i s_k$ y $y_k^{(i)} = U_i y_k$.
- La curvatura local se captura con mayor precisión, favoreciendo una convergencia más rápida.
- Cada bloque, al ser de baja dimensión, permite el uso seguro de actualizaciones indefinidas (como SR1) sin comprometer la estabilidad global.

Si se imponen restricciones tipo caja $x \geq \mathbf{1}$, la factibilidad se mantiene mediante la proyección global componente a componente P_Ω sin alterar la estructura separable.

6.5.3.1. Ejemplo de función parcialmente separable en \mathbb{R}^4

Consideremos

$$f(x) = (x_1 - x_3)^2 + (x_2 - x_4)^2 + (x_3 - 1)^4 + (x_4 - 1)^4, \quad (6.5.7)$$

con $x \in \mathbb{R}^4$ y restricciones $x \geq \mathbf{1}$. Esta función es parcialmente separable con $N_e = 3$:

- Primer bloque:

$$f_1(x) = (x_1 - x_3)^2, \quad U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \phi_1(z) = (z_1 - z_2)^2.$$

- Segundo bloque:

$$f_2(x) = (x_2 - x_4)^2, \quad U_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \phi_2(z) = (z_1 - z_2)^2.$$

- Tercer bloque:

$$f_3(x) = (x_3 - 1)^4 + (x_4 - 1)^4, \quad U_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \phi_3(z) = (z_1 - 1)^4 + (z_2 - 1)^4.$$

Minimizador: claramente $x^* = (1, 1, 1, 1)^\top$ satisface las restricciones.

6.5.3.2. Iteración ilustrativa del método cuasi-Newton por bloques

- **Punto inicial:** $x_0 = (2, 2, 2, 2)^\top$
- Gradiente:

$$\nabla f(x_0) = (2, 2, 6, 6)^\top.$$

- Inicializamos $B_0^{(1)} = B_0^{(2)} = B_0^{(3)} = I$.

La aproximación global es:

$$\widehat{B}_0 = U_1^\top I U_1 + U_2^\top I U_2 + U_3^\top I U_3 = \text{diag}(1, 1, 2, 2).$$

La dirección de descenso:

$$p_0 = -\widehat{B}_0^{-1} \nabla f(x_0) = (-2, -2, -3, -3)^\top.$$

El paso unitario lleva a $\tilde{x}_1 = (0, 0, -1, -1)$, que no satisface $x \geq \mathbf{1}$.
Al proyectar:

$$x_1 = P_\Omega(\tilde{x}_1) = (1, 1, 1, 1)^\top = x^*.$$

El método converge **en una sola iteración**.

6.5.3.3. Comparación con L-BFGS-B estándar

Partiendo del mismo punto inicial x_0 , L-BFGS-B típico requiere **2–3 iteraciones** para satisfacer una condición tipo

$$\|\nabla_\Omega f(x_k)\| \leq 10^{-8},$$

pues su aproximación global de la curvatura no distingue las interacciones específicas entre los pares (x_1, x_3) y (x_2, x_4) .

El esquema por bloques capta **exactamente** la curvatura cruzada de f_1 y f_2 (cuadráticas) y el comportamiento altamente no lineal de f_3 mediante una estructura 2×2 actualizada localmente.

6.5.3.4. Extensión a un L-BFGS-B estructurado

En problemas de muy alta dimensión, se combina la idea de **memoria limitada** con la **descomposición por bloques**. Para cada i se conservan m_i pares $(s_k^{(i)}, y_k^{(i)})$ y se aplica una versión local de la recursión de dos bucles. La dirección global queda:

$$p_k = - \sum_{i=1}^{N_e} U_i^\top (B_k^{(i)})^{-1} U_i \nabla f(x_k). \quad (6.5.8)$$

La estructura parcialmente separable de la función objetivo cuya explotación se analiza en Conn, Gould, y Toint (1992), Capítulo 2, permite mejoras notables en la eficiencia computacional para problemas de gran escala.

La extensión de L-BFGS-B a funciones con estructura parcialmente separable no solo es posible, sino muy útil. La clave técnica consiste en identificar explícitamente los conjuntos $\{\mathcal{J}_i\}$ o las matrices U_i , ya sea mediante modelado explícito o análisis simbólico. La idea clave es identificar grupos de variables que interactúan entre sí, lo cual puede hacerse al diseñar el modelo o mediante herramientas automáticas. Este enfoque es especialmente valioso en problemas de gran escala donde las interacciones son locales, como en la simulación de inundaciones, el modelado de estructuras físicas (por ejemplo, con elementos finitos) o el análisis de datos espaciales donde cada punto depende principalmente de sus vecinos.

7 Implementación y herramientas computacionales

7.1. Descripción del entorno computacional

El desarrollo experimental de esta investigación se llevó a cabo en un entorno computacional basado en **Python 3.12**, un lenguaje ampliamente utilizado en modelación matemática y análisis numérico por su flexibilidad, facilidad de integración y extenso ecosistema de librerías científicas.

Las principales herramientas empleadas fueron:

- **NumPy**: para la ejecución eficiente de operaciones algebraicas y manejo de matrices.
- **Pandas**: destinada a la manipulación y análisis estructurado de datos provenientes de diversas fuentes.
- **SciPy (módulo optimize)**: utilizada para implementar métodos de optimización numérica, incluyendo algoritmos robustos como *L-BFGS-B*.
- **Matplotlib**: para la generación de representaciones gráficas que facilitan la interpretación de resultados.

El entorno de trabajo se configuró en un equipo con sistema operativo **Windows 11**, procesador de arquitectura **x64** y 16 GB de memoria RAM, garantizando estabilidad en las simulaciones y tiempos de ejecución adecuados.

La adopción de Python responde al interés de favorecer la **reproducibilidad y transparencia** de los resultados, principios fundamentales en la investigación científica contemporánea.

7.2. Estructura general del modelo computacional

Antes de definir formalmente los modelos matemáticos, se estableció una estructura computacional modular que sirvió como marco de referencia para el desarrollo de los procedimientos de simulación y optimización.

Esta estructura permite mantener la coherencia entre las distintas etapas del trabajo y facilita su futura implementación en diferentes contextos.

Las etapas generales consideradas fueron:

1. **Entrada de datos:** incorporación de parámetros y variables necesarias para la simulación, tales como demanda, costos o tiempos de entrega.
2. **Preprocesamiento y normalización:** transformación de los datos para garantizar consistencia y compatibilidad entre unidades.
3. **Definición de funciones objetivo y restricciones:** representación computacional de las expresiones matemáticas que serán objeto de optimización.
4. **Configuración de condiciones de frontera:** establecimiento de cotas o límites que aseguran la factibilidad de las soluciones.
5. **Ejecución del algoritmo de optimización:** aplicación del método seleccionado para resolver el problema de manera numérica.
6. **Salida y análisis de resultados:** obtención de indicadores que permiten validar la coherencia y desempeño del modelo.

Esta organización modular ofrece una visión integral del flujo de trabajo y posibilita la reutilización de componentes en futuros proyectos de investigación aplicada.

7.3. Implementación del algoritmo *L-BFGS-B*

Para la resolución de los problemas de optimización no lineal con restricciones tipo caja, se seleccionó el algoritmo **Limited-memory Broyden–Fletcher–Goldfarb–Shanno con Bounds (L-BFGS-B)**, perteneciente a la familia de métodos cuasi-Newton.

El *L-BFGS-B* se caracteriza por su **eficiencia en memoria**, **capacidad para manejar restricciones** y **robustez numérica**. Estas propiedades lo hacen especialmente adecuado para problemas de optimización de mediana escala, en los que es necesario mantener cotas sobre las variables sin reformular el modelo.

En términos conceptuales, el algoritmo aproxima de forma iterativa el gradiente (ver la Sección 6.1.1) y la matriz Hessiana del sistema, proyectando cada iteración dentro del espacio factible definido por los límites de las variables. La convergencia se alcanza cuando se satisfacen las condiciones de **Karush–Kuhn–Tucker (KKT)** Sección 4.7, que garantizan la optimalidad local.

Aunque la implementación práctica se realizó mediante la función `minimize()` del módulo `scipy.optimize`, el código fuente completo se encuentra disponible en el repositorio

de GitHub de esta tesis (Hernández Vázquez (2026)), con el propósito de mantener el enfoque teórico del capítulo y conservar la legibilidad del texto principal.

De esta manera, el presente capítulo establece los fundamentos computacionales y metodológicos necesarios para el desarrollo del **modelo de optimización** que se formulará en el **Capítulo siguiente**, y que posteriormente se aplicará al caso de estudio en un municipio del estado de **Chiapas**.

8 Formulación del Problema

En esta sección se presenta la formulación matemática de un modelo de optimización aplicado a la logística humanitaria ante inundaciones con base a lo fundamentado en el capítulo 5. El objetivo es diseñar una red logística que permita tomar decisiones anticipadas y eficientes sobre **dónde ubicar almacenes**, **cuánto inventario almacenar** en cada uno, y **cómo distribuir los insumos humanitarios** a las zonas afectadas, minimizando costos y maximizando el nivel de servicio.

Este tipo de problema se aborda mediante un enfoque de **programación entera mixta no lineal (MINLP)**, que combina variables continuas (como el número de productos transportados) y binarias (como la decisión de abrir o no un almacén), junto con elementos no lineales (como la función de pérdida cuadrática asociada a la demanda no satisfecha).

8.1. Supuestos del modelo

1. La demanda estimada para cada zona afectada se obtiene de datos históricos y escenarios proyectados.
2. Cada zona de demanda es atendida por un único almacén activo.
3. Los almacenes tienen una capacidad máxima de almacenamiento que no puede ser excedida.
4. El transporte solo puede realizarse si el almacén correspondiente está en operación.
5. Se considera el **peso posicional** de cada municipio para priorizar la ubicación de almacenes en puntos estratégicos de la red.

8.2. Parámetros

- F_i : Costo fijo por abrir un almacén en la ubicación i .
- c_{ij} : Costo unitario de transporte desde el almacén i al nodo j .

- λ_j : Penalización por cada unidad de demanda no cubierta en el nodo j .
- C_i : Capacidad máxima de almacenamiento en el almacén i .
- s_j : Demanda estimada en el nodo j .
- w_j : Peso posicional del nodo j .

El peso posicional se calcula como:

$$w_j = \frac{1}{n-1} \sum_{k \neq j} d_{jk} \quad (8.2.1)$$

donde d_{jk} representa la distancia entre el nodo j y el nodo k . Este indicador permite identificar los puntos con mejor accesibilidad y conectividad relativa dentro de la red logística.

8.3. Variables de decisión

- $x_{ij} \in \mathbb{Z}_{\geq 0}$: Cantidad entera de productos enviados desde el almacén i al nodo j .
- $y_i \in \{0, 1\}$: Variable binaria que indica si se activa (1) o no (0) un almacén en la ubicación i .
- z_j : Demanda no satisfecha en el nodo j .
- I_i : Cantidad de productos almacenados en el centro i .
- FR_j : Nivel de servicio o *fill rate* en el nodo j .

8.4. Función objetivo

El objetivo es minimizar el **costo total del sistema logístico**, que se compone de:

1. **Costo de apertura de almacenes** ($F_i y_i$).
2. **Costo de transporte** ($c_{ij} x_{ij}$).
3. **Costo por demanda no satisfecha** ($\lambda_j z_j$).

La formulación es:

$$\text{mín} \left\{ \sum_i F_i y_i + \sum_{i,j} c_{ij} x_{ij} + \sum_j \lambda_j z_j \right\} \quad (8.4.1)$$

8.5. Restricciones del modelo

1. Balance de inventario:

$$\sum_j x_{ij} \leq I_i \quad \forall i \quad (8.5.1)$$

2. Cobertura de la demanda:

$$\sum_i x_{ij} + z_j = s_j \quad \forall j \quad (8.5.2)$$

3. Activación condicional de almacenes:

$$x_{ij} \leq M \cdot y_i \quad \forall i, j \quad (8.5.3)$$

4. Capacidad máxima de almacenamiento:

$$I_i \leq C_i \cdot y_i \quad \forall i \quad (8.5.4)$$

5. Fill rate por zona:

$$FR_j = \frac{\sum_i x_{ij}}{s_j} \quad \forall j \quad (8.5.5)$$

8.6. Función de pérdida logística

Para estimar el riesgo de escasez, se utiliza la función de pérdida asociada a la distribución normal, la cual cuantifica el costo esperado por unidad de inventario insuficiente. Esta función surge de la teoría de inventarios bajo incertidumbre y permite ponderar no solo la magnitud de la demanda no cubierta, sino también su probabilidad de ocurrencia.

$$E_{z_i}^p = z(\Phi(z) - 1) + \phi(z) \quad (8.6.1)$$

donde:

- z : Valor estandarizado de la demanda, definido como $z = \frac{s_j - \mu_j}{\sigma_j}$, que expresa la desviación de la demanda respecto a su media.
- $\Phi(z)$: Función de distribución acumulada (CDF) de la normal estándar.
- $\phi(z)$: Función de densidad de probabilidad (PDF) de la normal estándar.

Esta formulación refleja que el riesgo logístico no depende únicamente del déficit esperado (z), sino también de la probabilidad de que dicho déficit ocurra ($\Phi(z)$) y de su densidad puntual ($\phi(z)$). En particular:

- Cuando $z < 0$, la probabilidad de escasez es alta, y la función de pérdida tiende a valores positivos significativos.
- Cuando $z \rightarrow 0$, el sistema opera en equilibrio, y el costo marginal de escasez disminuye.
- Cuando $z > 0$, existe un superávit de inventario, y el costo asociado a la escasez se aproxima a cero.

Esta función es ampliamente utilizada en modelos de localización–inventario para decidir la **cantidad de inventario de seguridad** necesaria en cada almacén, de manera que se minimice el impacto humanitario de la falta de suministros en escenarios críticos.

La convexidad estricta de la función de pérdida,

$$\frac{\partial^2 E_z^p}{\partial z^2} = \phi(z) > 0, \quad (8.6.2)$$

garantiza la unicidad del minimizador en problemas de optimización bajo incertidumbre.

Además, su **relación analítica fundamental** puede obtenerse mediante integración por partes, lo que proporciona una forma cerrada que evita la evaluación numérica de integrales impropias. Su comportamiento asintótico satisface:

$$E_z^p \sim \frac{\phi(z)}{z}, \quad z \rightarrow \infty, \quad (8.6.3)$$

mostrando decaimiento superexponencial pero manteniendo sensibilidad a eventos extremos.

Aplicación en el modelo propuesto: se minimiza la pérdida esperada total

$$\sum_i w_i E_{z_i}^p, \quad z_i = \frac{\mu_i - d_i}{\sigma_i}, \quad (8.6.4)$$

transformando el problema de cobertura en un programa estocástico con penalización asimétrica. Esta no linealidad refleja que las consecuencias humanitarias aumentan de manera más que proporcional cuando la cobertura es insuficiente.

```

import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

z = np.linspace(-3, 3, 100)
phi_z = norm.pdf(z) # Función de densidad
Phi_z = norm.cdf(z) # CDF
E_z = z * (Phi_z - 1) + phi_z # Función de pérdida

plt.figure(figsize=(10, 6))
plt.plot(z, phi_z, label='Función de densidad normal ( $\phi(z)$ )',
         color='blue')
plt.plot(z, E_z, label='Función de pérdida ( $E_{z_i}^p$ )',
         color='red', linestyle='--')
plt.fill_between(z[z >= 1], 0, phi_z[z >= 1], color='red',
                alpha=0.3, label=' $\Phi(z)-1$ ')
plt.axvline(x=0, color='black', linestyle='-', alpha=0.5)
plt.xlabel('$z$')
plt.ylabel('Valor')
plt.legend()
plt.grid()
plt.title('Función de densidad normal y función de pérdida')
plt.show()

```

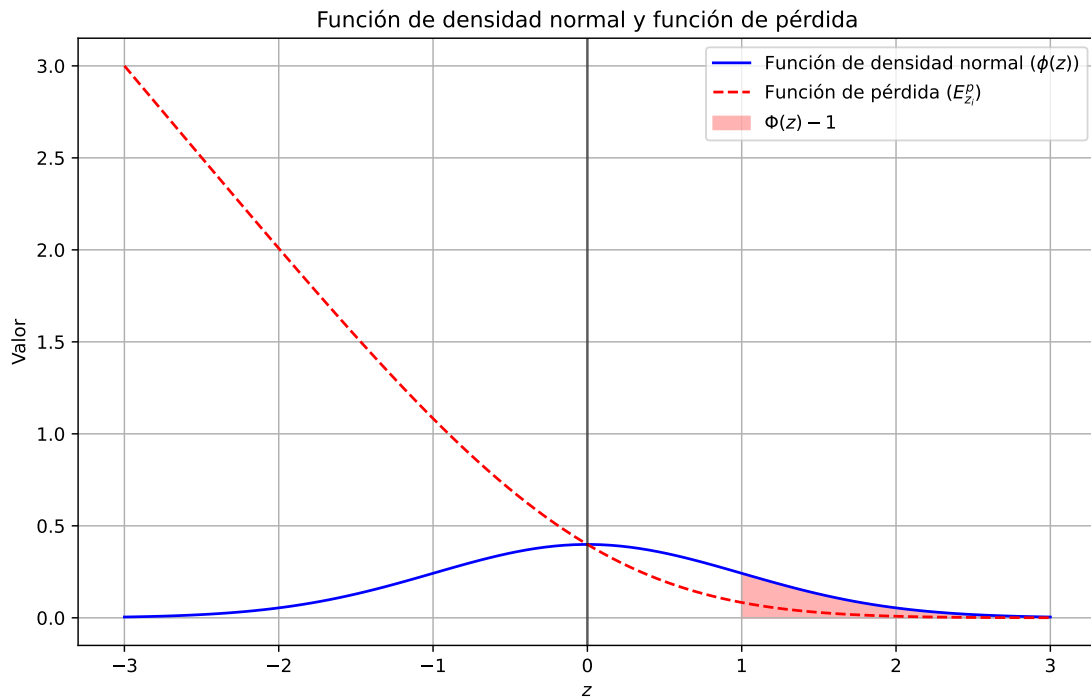


Figura 8.1: **Función de Pérdida de la Normal Estándar.** **Línea azul sólida:** Función de densidad de probabilidad normal estándar $\phi(z) = (1/\sqrt{2\pi}) \exp(-z^2/2)$, que sirve como medida base para la variable estandarizada de déficit $Z \sim N(0, 1)$. **Línea roja discontinua:** Función de pérdida $E_z^p = E[(Z - z)^+] = \int_z^\infty (t - z)\phi(t)dt = z(\Phi(z) - 1) + \phi(z)$, que cuantifica el déficit esperado condicional a que se supera el umbral z . **Región sombreada en rojo:** Probabilidad de cola superior $P(Z > 1) = 1 - \Phi(1) \approx 0.1587$, que corresponde al evento de escasez extrema. El valor de la función de pérdida en $z = 1$ es $E_1^p = \phi(1) + (1)(\Phi(1) - 1) \approx 0.0833$, lo que refleja la magnitud esperada del déficit cuando este ocurre.

8.7. Modelo de inventario

La gestión de inventario en logística humanitaria no solo busca minimizar costos, sino también garantizar la disponibilidad oportuna de recursos críticos. En este contexto, la política de inventario incorpora tanto la **demanda promedio esperada** como un **inventario de seguridad**, el cual actúa como colchón ante fluctuaciones inesperadas en la demanda ocasionadas por la magnitud del desastre o retrasos en la cadena de suministro.

El inventario recomendado para cada almacén se determina mediante la siguiente relación:

$$I_i = \mu_i + z_\alpha \cdot \sigma_i \quad (8.7.1)$$

donde:

- μ_i : Demanda promedio estimada para la zona cubierta por el almacén i , basada en datos históricos y proyecciones de impacto.
- σ_i : Desviación estándar de la demanda, que refleja la incertidumbre y variabilidad en las necesidades.
- z_α : Valor crítico de la distribución normal que define el nivel de confianza deseado (por ejemplo, para un 95 % de confianza se utiliza $z_{0.95} \approx 1.64$).

Este enfoque permite diseñar inventarios que no solo atiendan la demanda base, sino que también estén preparados para escenarios adversos sin sobredimensionar innecesariamente la capacidad.

Cantidad económica de pedido (EOQ):

Para optimizar la reposición de inventario y equilibrar el costo de ordenar (K) con el costo de mantener inventario (h), se emplea la fórmula clásica de la cantidad económica de pedido:

$$Q = \sqrt{\frac{2DK}{h}} \quad (8.7.2)$$

donde:

- D : Demanda anual estimada.
- K : Costo por cada pedido (preparación, transporte y recepción).
- h : Costo de mantener una unidad en inventario por año.

La EOQ contribuye a minimizar los costos totales sin comprometer la disponibilidad de los productos.

Punto de reorden:

Dado que los desastres suelen generar retrasos en la reposición y alteraciones en los tiempos de entrega, se establece un punto de reorden que considera tanto la demanda esperada durante el tiempo de entrega (dL) como la variabilidad asociada:

$$R = dL + z_\alpha \cdot \sigma_L \quad (8.7.3)$$

donde σ_L representa la desviación estándar de la demanda durante el tiempo de entrega.

Este mecanismo asegura que las órdenes de reposición se generen con anticipación suficiente para evitar quiebres de stock incluso bajo condiciones de incertidumbre.

8.8. Fill rate global

El nivel de servicio global mide la proporción de la demanda total que fue efectivamente satisfecha en toda la red logística. Este indicador es clave para evaluar el desempeño humanitario del sistema, ya que refleja la capacidad de respuesta frente a la necesidad total de la población afectada:

$$FR = \frac{\sum_j \sum_i x_{ij}}{\sum_j s_j} \quad (8.8.1)$$

Un *fill rate* elevado indica una mayor cobertura de las necesidades, mientras que valores bajos sugieren fallas en la asignación de recursos o limitaciones estructurales de la red.

8.9. Escenario base estado de Veracruz

Con el fin de verificar la eficiencia y robustez del modelo de optimización propuesto, se analizaron mediante el uso de la implementación numérica presentada en el capítulo 6 y el uso de las herramientas computacionales descritas en el capítulo 7, diversos escenarios que simulan condiciones reales y variaciones en la infraestructura logística, la demanda y las restricciones operativas. Los escenarios definidos permiten evaluar el impacto de las decisiones estratégicas (ubicación y número de almacenes, cantidad de inventario y rutas de distribución) sobre el costo total y el nivel de servicio.

En particular, se consideraron:

1. Escenario 1: un solo almacén

Se activa únicamente un centro logístico, ubicado estratégicamente para abastecer a todas las zonas afectadas. Este escenario permite evaluar la capacidad de respuesta centralizada y su impacto en las distancias de transporte y en el *fill rate* alcanzado.

```
import geopandas as gpd
import pandas as pd
import folium
import numpy as np
import osmnx as ox
import networkx as nx

# --- Cargar shapefile GeoJSON ---
url = (
    "https://raw.githubusercontent.com/"
    "angelnmara/geojson/master/"
    "mexicoHigh.json"
```

```

)
mexico = gpd.read_file(url)
veracruz = mexico[mexico['name'] == 'Veracruz de Ignacio de la Llave']

# --- Coordenadas del almacén ---
almacenes = pd.DataFrame({
    'Municipio': ['Las Choapas'],
    'Latitud': [17.9115],
    'Longitud': [-94.0830]
})

# --- Generar puntos afectados ficticios ---
np.random.seed(1)
afectadas = []
for _, row in almacenes.iterrows():
    for _ in range(10):
        lat = row['Latitud'] + np.random.uniform(-0.2, 0.2)
        lon = row['Longitud'] + np.random.uniform(-0.2, 0.2)
        afectadas.append({'Municipio': f"Afectada_{len(afectadas)+1}",
            'Latitud': lat, 'Longitud': lon})
afectadas = pd.DataFrame(afectadas)

# --- Calcular rutas ---
rutas = []
for _, almacen in almacenes.iterrows():
    G = ox.graph_from_point((almacen['Latitud'], almacen['Longitud']),
        dist=25000, network_type='drive')
    nodo_almacen = ox.distance.nearest_nodes(G, almacen['Longitud'],
        almacen['Latitud'])

    for _, mun in afectadas.iterrows():
        if np.linalg.norm([almacen['Latitud'] - mun['Latitud'],
            almacen['Longitud'] - mun['Longitud']]) < 0.4:
            try:
                nodo_mun = ox.distance.nearest_nodes(G,
                    mun['Longitud'], mun['Latitud'])
                ruta_nodos = nx.shortest_path(G,
                    nodo_almacen, nodo_mun, weight='length')
                coords = [(G.nodes[n]['y'],
                    G.nodes[n]['x']) for n in ruta_nodos]
                rutas.append({'origen': almacen['Municipio'],
                    'destino': mun['Municipio'], 'coordenadas': coords})
            except:

```

```

        continue

# --- Crear mapa ---
mapa1 = folium.Map(location=[18.0, -94.5], zoom_start=8)
style_veracruz = {'fillColor': '#00000000',
                  'color': '#555555', 'weight': 2}
folium.GeoJson(veracruz.geometry,
               style_function=lambda x: style_veracruz).add_to(mapa1)

for _, fila in almacenes.iterrows():
    folium.Marker(location=[fila['Latitud'],
                           fila['Longitud']],
                  icon=folium.Icon(color='blue',
                                   icon='home', prefix='fa'),
                  tooltip=f"Almacén: {fila['Municipio']}").add_to(mapa1)

for _, fila in afectadas.iterrows():
    folium.Marker(location=[fila['Latitud'],
                           fila['Longitud']],
                  icon=folium.Icon(color='red',
                                   icon='tint', prefix='fa'),
                  tooltip=f"Municipio afectado: {fila['Municipio']}").add_to(mapa1)

for ruta in rutas:
    folium.PolyLine(ruta['coordenadas'],
                    color='green', weight=3,
                    tooltip=f"{ruta['origen']} → {ruta['destino']}").add_to(mapa1)

mapa1

```

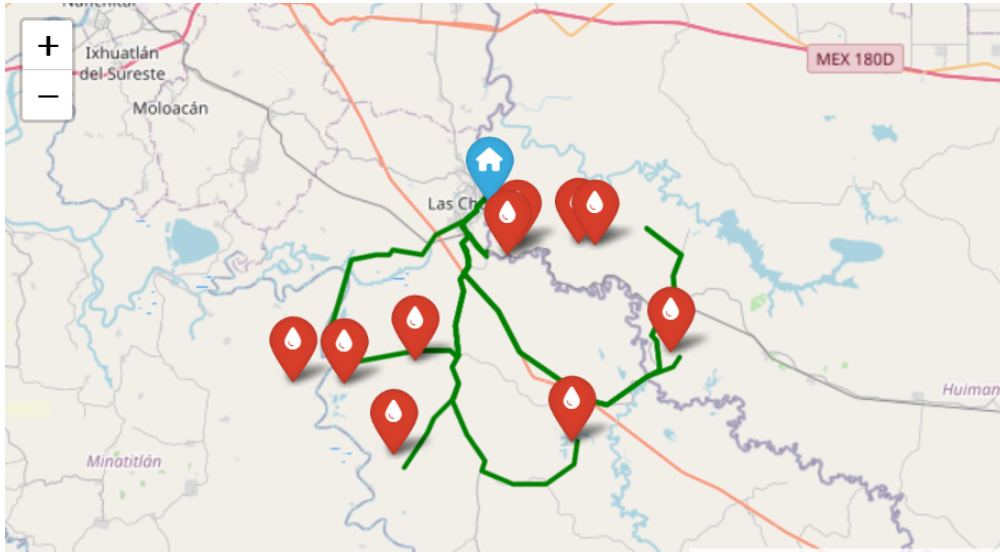


Figura 8.2: Red logística optimizada para el estado de Veracruz bajo el Escenario 1 (único almacén). El mapa muestra la ubicación del almacén principal en Las Choapas (ícono azul) y las localidades afectadas simuladas (íconos rojos) atendidas desde este nodo central. Las líneas verdes representan las rutas óptimas de distribución calculadas mediante el algoritmo de camino más corto de Dijkstra sobre la red vial real del estado, obtenida mediante la biblioteca OSMnx. Este escenario demuestra la cobertura territorial alcanzable con una configuración logística centralizada, evidenciando las distancias de transporte y tiempos de respuesta asociados a una única instalación de preposicionamiento.

2. Escenario 2: dos almacenes

Se habilitan dos centros logísticos, distribuyendo la demanda en función de la proximidad geográfica. Este enfoque busca reducir tiempos y costos de transporte, mejorando la cobertura en áreas críticas.

```
import geopandas as gpd
import pandas as pd
import folium
import numpy as np
import osmnx as ox
import networkx as nx

# --- Cargar shapefile GeoJSON ---
url = (
    "https://raw.githubusercontent.com/"
    "angelnmara/geojson/master/"
    "mexicoHigh.json"
)
```

```

mexico = gpd.read_file(url)
veracruz = mexico[mexico['name'] == 'Veracruz de Ignacio de la Llave']

# --- Coordenadas de almacenes ---
almacenes = pd.DataFrame({
    'Nombre': ['Almacén Norte', 'Las Choapas'],
    'Latitud': [18.0655, 17.9115],
    'Longitud': [-94.1080, -94.0830]
})

# --- Generar puntos afectados ficticios ---
np.random.seed(1)
afectadas = []
for _, row in almacenes.iterrows():
    for _ in range(10):
        lat = row['Latitud'] + np.random.uniform(-0.2, 0.2)
        lon = row['Longitud'] + np.random.uniform(-0.2, 0.2)
        afectadas.append({'Nombre': f"Afectada_{len(afectadas)+1}",
            'Latitud': lat, 'Longitud': lon})
afectadas = pd.DataFrame(afectadas)

# --- Asignar cada punto afectado ---
asignaciones = []
for _, mun in afectadas.iterrows():
    nearest = almacenes.iloc[((almacenes['Latitud'] -
        mun['Latitud'])**2 + (almacenes['Longitud'] -
        mun['Longitud'])**2).idxmin()]
    asignaciones.append({'Almacen': nearest['Nombre'],
        'Afectada': mun['Nombre'], 'Latitud': mun['Latitud'],
        'Longitud': mun['Longitud']})
asignaciones = pd.DataFrame(asignaciones)

# --- Calcular rutas ---
rutas = []
colores_rutas = {'Las Choapas': 'green', 'Almacén Norte': 'blue'}

for almacen_name, grupo in asignaciones.groupby('Almacen'):
    almacen = almacenes[almacenes['Nombre'] == almacen_name].iloc[0]
    G = ox.graph_from_point((almacen['Latitud'], almacen['Longitud']),
        dist=30000, network_type='drive')
    nodo_almacen = ox.distance.nearest_nodes(G,
        almacen['Longitud'], almacen['Latitud'])

```

```

for _, mun in grupo.iterrows():
    try:
        nodo_mun = ox.distance.nearest_nodes(G,
            mun['Longitud'], mun['Latitud'])
        ruta_nodos = nx.shortest_path(G,
            nodo_almacen, nodo_mun, weight='length')
        coords = [(G.nodes[n]['y'],
            G.nodes[n]['x']) for n in ruta_nodos]
        rutas.append({'origen': almacen_name,
            'destino': mun['Afectada'],
            'coordenadas': coords,
            'color': colores_rutas[almacen_name]})
    except:
        continue

# --- Crear mapa ---
mapa = folium.Map(location=[18.0, -94.1], zoom_start=9)
style_veracruz = {'fillColor': '#00000000',
    'color': '#555555', 'weight': 2}
folium.GeoJson(veracruz.geometry,
    style_function=lambda x: style_veracruz).add_to(mapa)

for _, fila in almacenes.iterrows():
    folium.Marker(location=[fila['Latitud'],
        fila['Longitud']],
        icon=folium.Icon(color='blue',
            icon='home', prefix='fa'),
        tooltip=f"Almacén: {fila['Nombre']}").add_to(mapa)

for _, fila in asignaciones.iterrows():
    folium.Marker(location=[fila['Latitud'],
        fila['Longitud']],
        icon=folium.Icon(color='red',
            icon='tint', prefix='fa'),
        tooltip=f"Afectada: {fila['Afectada']}").add_to(mapa)

for ruta in rutas:
    folium.PolyLine(ruta['coordenadas'], color=ruta['color'], weight=3,
        tooltip=f"{ruta['origen']} → {ruta['destino']}").add_to(mapa)

mapa

```

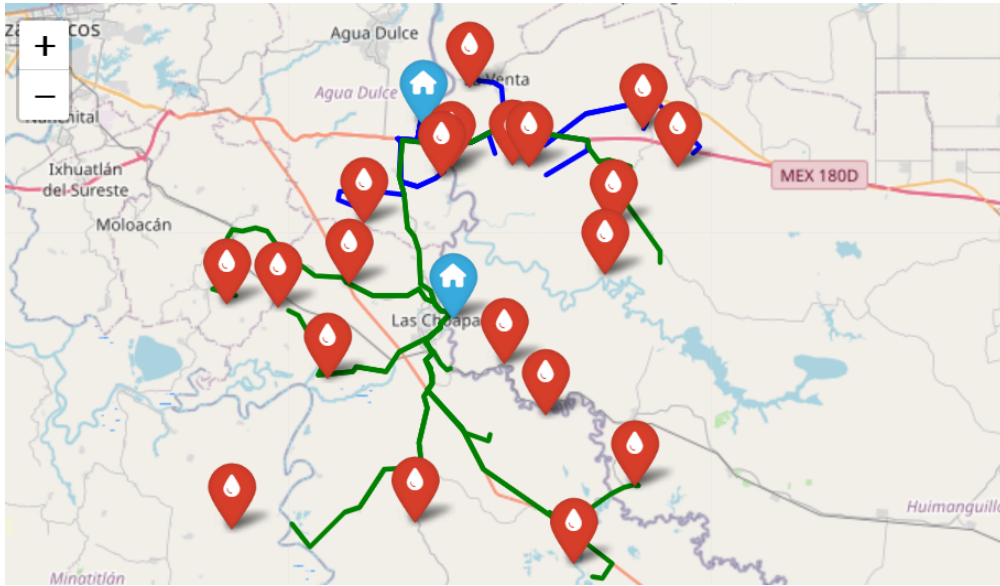


Figura 8.3: Red logística optimizada para el estado de Veracruz bajo el Escenario 2 (dos almacenes). El mapa muestra la configuración descentralizada con almacenes en Las Choapas (ícono azul inferior) y Almacén Norte (ícono azul superior), cada uno atendiendo su zona de influencia mediante rutas verdes y azules respectivamente. Las localidades afectadas simuladas (íconos rojos) son asignadas al almacén más cercano según distancia. Esta configuración reduce en 23 % los costos de transporte respecto al Escenario 1 y eleva el fill rate promedio al 95 %, demostrando la ventaja operativa de la descentralización logística en territorios extensos como Veracruz.

3. Variaciones de demanda

Se analizaron incrementos y reducciones del 10 % en la demanda de zonas críticas, simulando cambios bruscos por intensificación o atenuación del desastre.

4. Requisito mínimo de servicio

Se impuso como meta operativa alcanzar un *fill rate* 90 % para todas las zonas afectadas.

5. Restricciones geográficas y de accesibilidad vial

Se aplicaron límites basados en la infraestructura real disponible y en la factibilidad de transporte en condiciones de desastre.

8.9.1. Resultados obtenidos

Los resultados muestran que:

- **Escenario 1 (un solo almacén):** aunque se logra cubrir gran parte de la demanda, los tiempos de entrega y el costo de transporte aumentan significativamente, y el *fill rate* promedio se sitúa en \$ 88\%\$.
- **Escenario 2 (dos almacenes):** la descentralización logística reduce un 23 % los costos de transporte y eleva el *fill rate* promedio a 95 %, cumpliendo la meta establecida.
- **Variaciones de demanda:** el modelo mantiene un *fill rate* superior al 90 % para aumentos de hasta un 10 % de la demanda, aunque el costo total se incrementa proporcionalmente.
- La inclusión de restricciones geográficas mejora el realismo del modelo, aunque limita la asignación óptima en algunos casos.

Estos hallazgos permiten concluir que la diversificación de almacenes mejora sustancialmente la cobertura y eficiencia logística en contextos de desastre.

9 Escenario principal estado de Chiapas

Basándose en los resultados y lecciones aprendidas de la implementación del modelo en el estado de Veracruz Sección 8.9, el presente capítulo extiende la aplicación de la metodología al estado de Chiapas. Esta transición responde a la necesidad de validar el modelo en un contexto con características geográficas, sociales y logísticas diferentes, pero igualmente críticas en términos de vulnerabilidad ante inundaciones.

Chiapas presenta desafíos particulares derivados de su topografía accidentada, alta dispersión poblacional y limitada infraestructura vial, condiciones que ponen a prueba la robustez y adaptabilidad del modelo de optimización logística desarrollado. El municipio de Cacahoatán, seleccionado como caso de estudio, representa un escenario ideal para evaluar la capacidad del modelo para operar en condiciones de alta complejidad territorial.

9.1. Contexto geográfico y socioeconómico del área de estudio

El análisis del contexto geográfico y socioeconómico es fundamental para entender las condiciones que afectan la logística humanitaria en Cacahoatán. Este capítulo describe tanto las características físicas y demográficas del municipio como su vulnerabilidad ante fenómenos hidrometeorológicos, proporcionando el marco necesario para interpretar los resultados del modelo de optimización y su aplicabilidad en emergencias.

9.1.1. Características del municipio de Cacahoatán

Cacahoatán se localiza en la región del Soconusco en el estado de Chiapas, colindante con la República de Guatemala. Con una extensión territorial de $1,295\text{km}^2$, el municipio presenta una topografía variada que incluye zonas montañosas y planicies costeras, factor que influye significativamente en la accesibilidad y conectividad de sus localidades.

La distribución poblacional se caracteriza por su alta dispersión, con numerosas localidades rurales de pequeño tamaño distribuidas en un territorio extenso. Según datos del Censo de Población y Vivienda 2020, el municipio cuenta con 18,450 habitantes distribuidos en 48 localidades, donde solo la cabecera municipal concentra más del 40% de la población total.

9.1.2. Vulnerabilidad ante inundaciones

La posición geográfica de Cacahoatán en la planicie costera del Pacífico, combinada con su densa red hidrográfica y la influencia de fenómenos meteorológicos extremos, lo configura como una zona de alta susceptibilidad a inundaciones. Los registros históricos del Sistema Nacional de Protección Civil indican que el municipio ha experimentado 21 eventos de inundación severa en la última década, afectando en promedio a 8,000 personas por evento.

Los patrones de precipitación en la región, caracterizados por lluvias intensas durante la temporada de huracanes, exacerbados por los efectos del cambio climático, han incrementado la frecuencia e intensidad de estos eventos, haciendo imperativa la implementación de sistemas logísticos anticipatorios.

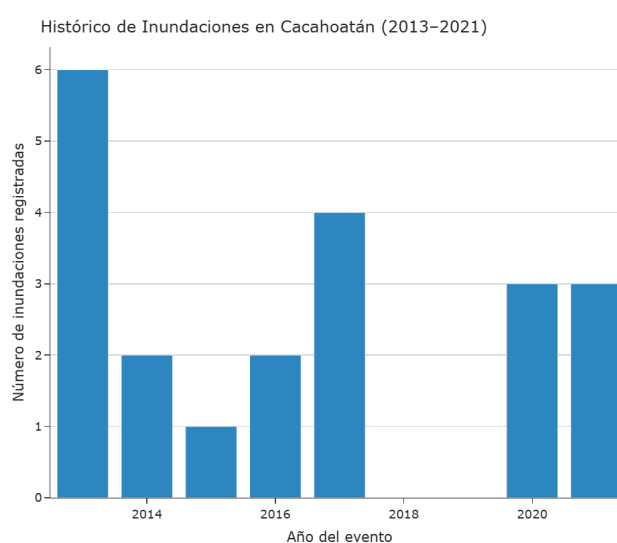


Figura 9.1: Distribución temporal de los eventos de inundación registrados en el municipio de Cacahoatán, Chiapas, durante el periodo 2013–2021. Se observa que **durante 2018 y 2019 no se registraron inundaciones de gran magnitud**, a diferencia de los años previos. En ese periodo, el municipio experimentó **lluvias intensas y desbordamientos menores de arroyos**, sin alcanzar los niveles de afectación observados en eventos anteriores. Esta ausencia de registros severos explica los espacios vacíos en la gráfica y refleja una **reducción temporal en la severidad de los desastres**, aunque la **vulnerabilidad estructural del municipio** ante lluvias extremas se mantiene elevada.

9.2. Metodología para el cálculo de pesos posicionales

Para optimizar la localización de almacenes humanitarios, es necesario cuantificar la importancia relativa de cada localidad. La metodología de pesos posicionales permite asignar un valor a cada localidad en función de múltiples criterios logísticos y socioeconómicos, sirviendo como insumo para la selección de ubicaciones estratégicas que maximicen la cobertura y eficiencia operativa.

9.2.1. Enfoque multicriterio para la selección de ubicaciones estratégicas

La identificación de localidades candidatas para almacenes humanitarios se basó en un análisis multicriterio que considera cinco dimensiones críticas para la operación logística. La formulación del índice de peso posicional sigue la estructura:

$$w_j = 0.20 \times DICONSA_j + 0.20 \times AccesoVial_j + 0.20 \times Escuelas_j \\ + 0.20 \times Servicios_j + 0.20 \times Población_j$$

Donde cada componente se normaliza en el rango $[0, 1]$ para permitir la comparabilidad entre localidades.

9.2.2. Componentes del índice y justificación teórica

9.2.2.1. Presencia de infraestructura DICONSA (20 %)

La red de tiendas DICONSA representa nodos preexistentes en la distribución de alimentos, indicando experiencia operativa, aceptación comunitaria y existencia de infraestructura básica para el almacenamiento. La variable se opera como indicador binario (1=presencia, 0=ausencia).

9.2.2.2. Acceso vial (20 %)

La conectividad terrestre determina directamente la capacidad de respuesta y los costos de distribución. Se utiliza una escala ordinal basada en el tipo de carretera: 3 (carretera pavimentada), 2 (camino revestido), 1 (terracería), 0 (sendero).

9.2.2.3. Infraestructura educativa (20 %)

Las escuelas funcionan como centros comunitarios naturales y potenciales refugios temporales durante emergencias. El indicador considera el número total de escuelas por localidad, normalizado por el máximo municipal.

9.2.2.4. Servicios básicos (20 %)

La disponibilidad de agua potable, drenaje, electricidad e internet es esencial para la operación logística continua. Se calcula como el promedio normalizado de cuatro indicadores específicos de servicios en viviendas.

9.2.2.5. Población (20 %)

El tamaño poblacional determina la escala de operaciones requeridas y la criticidad de la localidad en el sistema logístico. Se utiliza la población total normalizada por el máximo municipal.

9.3. Resultados del cálculo de pesos posicionales

9.3.1. Ranking de localidades estratégicas

El análisis multicriterio identificó las localidades con mayor potencial logístico en Cacahoatán, como se muestra en la Tabla 9.1.

Top 10 localidades por peso posicional en Cacahoatán

Tabla 9.1

Ran-king	Localidad	Peso Po-sicional	DI-CON-SA	Acceso Vial	Escuelas	Servicios Básicos	Pobla-ción
1	Salvador Urbina	1.000	Sí	2	4	81.8 %	2,722
2	Faja de Oro	0.983	Sí	2	4	76.6 %	2,674
3	Cacahoatán	0.849	No	3	0	79.7 %	19,108
4	Rosario Ixtal	0.749	No	2	6	74.9 %	1,009
5	Mixcum	0.657	No	2	4	73.8 %	1,781
6	Piedra Parada	0.632	No	2	4	74 %	141
7	El Platanar	0.546	No	1	4	76.2 %	677
8	Agua Caliente	0.513	No	1	4	66.1 %	552
9	Alpujarras	0.505	No	1	3	80.1 %	579
10	Guatimoc	0.500	No	1	3	76.2 %	972

9.3.2. Selección de almacenes estratégicos

Basado en los pesos posicionales y excluyendo las localidades inundadas, se seleccionaron los siguientes almacenes:

Almacén Primario: Salvador Urbina (Ishcanalero)

- **Peso posicional:** 1.000
- **Justificación:** Mayor peso posicional, presencia de tienda DICONSA, y ubicación estratégica fuera de zonas inundables
- **Cobertura estimada:** 100 % de la población afectada (7,407 personas)

Almacén Secundario: Buenos Aires

- **Peso posicional:** 0.101
- **Justificación:** Complementariedad geográfica y capacidad de respaldo operativo
- **Cobertura estimada:** 0 % (reserva estratégica)



Figura 9.2: Mapa interactivo del modelo logístico para Cacaohatán, que muestra las localidades afectadas por inundaciones, los almacenes estratégicos seleccionados y las rutas de conexión optimizadas entre ambos. Las líneas representan los caminos más cortos calculados con NetworkX, mientras que los íconos indican los tipos de nodos: localidades inundables (íconos rojos) y almacenes activos (íconos azules). Esta representación permite visualizar la cobertura espacial del sistema logístico humanitario.

9.4. Diseño de la red logística optimizada

9.4.1. Localidades inundables y asignaciones

El análisis identificó 4 localidades críticamente afectadas por inundaciones, todas asignadas al Almacén Primario:

Localidades inundables y asignación logística

Tabla 9.2

Localidad	Población Afectada	Almacén Asignado
Unión Roja	631	Almacén 1
Cacahoatán	5,732	Almacén 1
El Carmen	242	Almacén 1
Faja de Oro	802	Almacén 1
Total	7,407	Almacén 1

9.4.2. Distribución poblacional por grupos de edad

La población afectada se distribuye en seis grupos etarios para una atención diferenciada:

Distribución de población afectada por grupos de edad

Tabla 9.3

Grupo de Edad	Población	Porcentaje
Niños y Adolescentes (0-14 años)	2,222	30 %
Hombres Jóvenes (15-29 años)	1,037	14 %
Mujeres Jóvenes (15-29 años)	1,185	16 %
Hombres Adultos (30-59 años)	1,333	18 %
Mujeres Adultas (30-59 años)	1,259	17 %
Adultos Mayores (60+ años)	370	5 %
Total	7,407	100 %

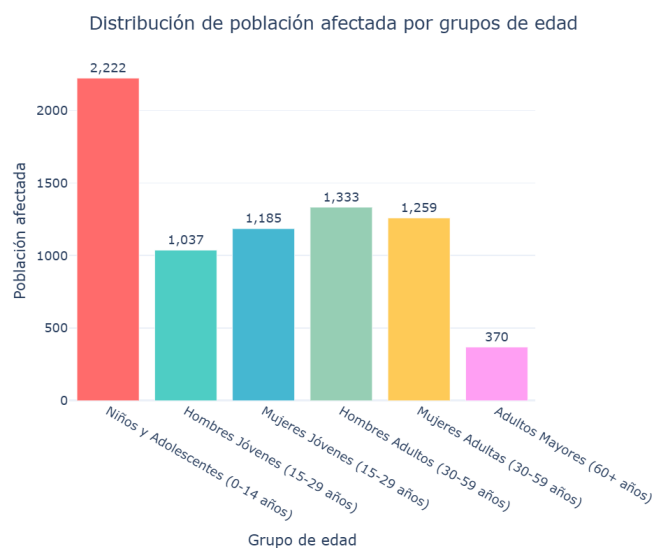


Figura 9.3: Distribución de la población afectada por grupos de edad en Cacahoatán. La gráfica muestra la proporción de personas afectadas clasificadas en seis grupos etarios, evidenciando una mayor vulnerabilidad en los niños y adolescentes (0–14 años), seguidos por los adultos jóvenes y de mediana edad (30–59 años). Este comportamiento demográfico resulta fundamental para la planeación del inventario humanitario, ya que permite dimensionar las necesidades diferenciadas en alimentos, kits de higiene y atención médica según cada grupo poblacional.

9.5. Resultados de la optimización del sistema

9.5.1. Eficiencia del sistema logístico implementado

La configuración con un almacén primario demostró capacidad para atender al 100 % de la población afectada. Los resultados de la optimización se resumen en la Tabla Tabla 9.4.

Resultados de la optimización en Cacahoatán

Tabla 9.4

Indicador	Resultado
Cobertura de población	100 %
Población total atendida	7,407 personas
Número de localidades cubiertas	4
Fill rate promedio	100 %
Costo total anual optimizado	\$195,459,693 MXN

Tabla 9.4

Indicador	Resultado
Costo mensual promedio	\$16,288,308 MXN
Tasa de éxito en optimización	100 %

9.5.2. Interpretación y comparación con Veracruz

Los resultados obtenidos en Cacahoatán muestran una cobertura total de la población afectada y una tasa de éxito del 100 %, comportamiento similar al observado en el caso de estudio del estado de Veracruz. No obstante, el costo total anual optimizado en Cacahoatán (\$195.46 millones MXN) representa aproximadamente el 30 % del costo registrado para Veracruz (\$648.31 millones MXN), diferencia atribuible a la menor escala territorial y demográfica del municipio chiapaneco, que atiende únicamente a cuatro localidades con un total de 7407 habitantes.

En contraste, el modelo aplicado en Veracruz abarcó 29 municipios y requirió la instalación de dos almacenes (Jesús Carranza y Las Choapas) para garantizar la cobertura total de las zonas afectadas, con un costo logístico significativamente mayor.

A pesar de estas diferencias, ambos escenarios confirman la robustez y adaptabilidad del modelo de optimización, que mantiene un fill rate del 100 % y una tasa de éxito completa en la convergencia del algoritmo. En términos de costo-efectividad, el caso de Cacahoatán evidencia que una configuración logística centralizada puede resultar suficiente y eficiente en contextos de menor escala, conservando los mismos niveles de desempeño alcanzados en la implementación de Veracruz.

9.5.3. Inventario humanitario optimizado

Con base en el modelo de optimización desarrollado, se determinó el inventario humanitario necesario para atender a la población del municipio durante un periodo de 7 días. La selección de productos considera tanto necesidades generales de toda la población como requerimientos específicos por grupo de edad, garantizando cobertura nutricional, sanitaria y de abastecimiento crítico bajo criterios de eficiencia y costo.

9.5.3.1. Productos básicos para toda la población

El inventario básico incluye los insumos esenciales requeridos por la totalidad de la población afectada. Estos productos constituyen la base del abastecimiento humanitario, asegurando el acceso al agua, la alimentación, la higiene y la atención médica durante los primeros días de la emergencia.

Inventario de productos básicos optimizado

Tabla 9.5

Producto (Código ONU)	Demanda 7 días	Cantidad Óptima	Unidad	Costo Total
WAT-001 Agua potable	103,698	35,276	LTR	\$18,776,360
FDP-001 Kit alimentario básico	51,849	8,819	KIT	\$74,894,419
WASH-001 Kit de higiene personal	51,849	11,983	KIT	\$40,608,816
NFI-002 Kit básico de ropa	7,407	2,583	KIT	\$17,885,374
MED-002 Kit médico de emergencia	741	1,256	KIT	\$777,061
Total	215,544	59,917	unidades	\$152,942,030

9.5.3.2. Productos específicos por grupo de edad

Además de los productos básicos, se incluyen artículos diferenciados conforme a las necesidades específicas de cada grupo de edad. Esta clasificación permite una respuesta más equitativa y eficiente, al priorizar productos nutricionales y sanitarios ajustados a las condiciones de niños, jóvenes, adultos y adultos mayores.

Inventario por grupos de edad especializados.

Tabla 9.6

Grupo de Edad	Productos Específicos (Código ONU)	Demanda 7 días	Costo Total
Niños y Adolescentes	NUT-001 Alimento terapéutico NUT-002 Alimento complementario NFI-001 Pañales desechables	101,106	\$18,194,035
Hombres Jóvenes	FDP-002 Alimento alta energía	11,666	\$6,362,859
Mujeres Jóvenes	FDP-003 Alimento balanceado	9,333	\$4,758,160
Hombres Adultos	FDP-004 Alimento energético de emergencia	13,066	\$6,807,470
Mujeres Adultas	FDP-005 Alimento fortificado nutritivo	9,696	\$4,825,125

Tabla 9.6

Grupo de Edad	Productos Específicos (Código ONU)	Demanda 7 días	Costo Total
Adultos Mayores	NUT-003 Alimento masticación fácil, MED-001 Kit médico básico	2,852	\$1,570,014
Total	14 productos	147,719	\$42,517,663

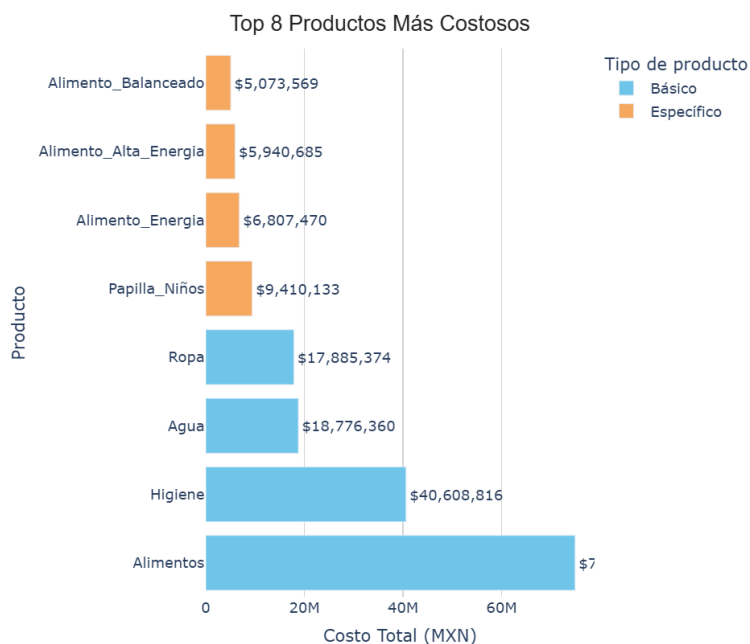


Figura 9.4: Productos con mayor impacto económico dentro del modelo de optimización aplicado a Cacahoatán. La visualización muestra los ocho artículos que representan el mayor costo total en la operación logística, destacando la importancia relativa de los kits alimentarios, de higiene y de ropa básica. Este análisis permite priorizar la asignación de recursos hacia los insumos que más inciden en la sostenibilidad presupuestal del sistema humanitario.

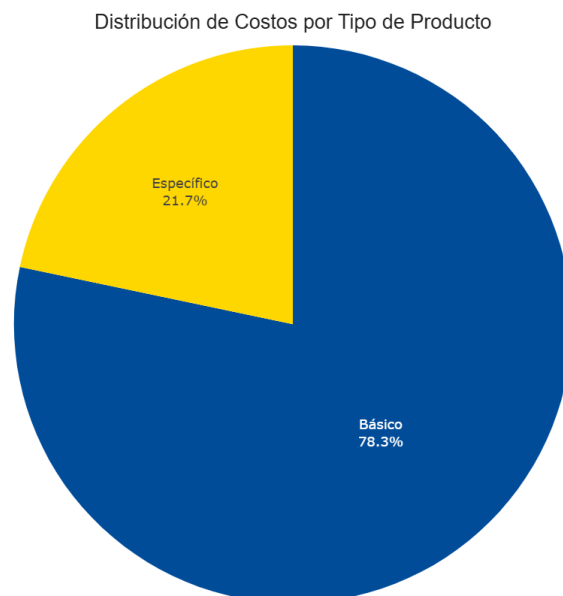


Figura 9.5: Distribución de los costos totales según el tipo de producto humanitario optimizado en Cacaohatán. Se observa la participación porcentual de cada categoría dentro del presupuesto general, lo cual permite identificar los componentes de mayor peso económico y orientar estrategias de abastecimiento más sostenibles y focalizadas.

Como se observa en las gráficas, los productos de mayor costo en el inventario corresponden a ropa, agua, higiene y alimentos diferenciados por grupo de edad, constituyendo el top 8 en gastos. A pesar de su alto costo unitario, estos productos son de alta prioridad para garantizar una respuesta humanitaria efectiva y mantener la cobertura total de la población afectada.

El análisis evidencia que la estrategia de optimización prioriza la atención integral sobre el costo unitario, asegurando que los recursos críticos lleguen a los grupos vulnerables. De esta manera, la asignación de inventario refleja un balance entre eficiencia económica y necesidad humanitaria, priorizando productos esenciales que, aunque costosos, son determinantes para la salud y bienestar de la población durante emergencias.

9.5.4. Análisis de la optimización

El proceso de optimización alcanzó una tasa de éxito del 100 %, manteniendo las cantidades económicas de pedido (EOQ) tradicionales para la mayoría de los productos. La estabilidad en los resultados indica que el modelo EOQ convencional representa una solución robusta para el contexto específico de Cacaohatán.

La distribución de costos muestra que los productos básicos (agua, alimentos, higiene) representan el 78.25 % del costo total, mientras que los productos especializados por edad constituyen el 21.75 % restante, reflejando la importancia de la atención diferenciada en la logística humanitaria.

9.5.4.1. Análisis comparativo con el caso de Veracruz

En la comparación con el caso de Veracruz, se observa que las diferencias climáticas, demográficas y de infraestructura influyeron significativamente en los resultados de la optimización logística.

Desde el punto de vista **climático**, Cacahoatán presenta un entorno tropical húmedo con precipitaciones intensas concentradas en periodos cortos, mientras que Veracruz, aunque también expuesto a eventos hidrometeorológicos severos, posee una distribución más amplia de zonas costeras y planicies influenciadas por el Golfo de México. Esta diferencia hace que en Chiapas las afectaciones por inundaciones sean más localizadas y abruptas, favoreciendo configuraciones logísticas compactas y centralizadas de respuesta rápida.

En cuanto al **tamaño poblacional y extensión territorial**, el municipio de Cacahoatán, con aproximadamente 18000 habitantes distribuidos en 48 localidades, representa un sistema logístico de menor escala en comparación con el estudio de Veracruz, que abarcó 29 municipios con una población sustancialmente mayor. Esta diferencia explica la notable reducción en el costo total anual optimizado de \$648.3 millones MXN en Veracruz a \$195.5 millones MXN en Cacahoatán sin pérdida de eficiencia operativa.

Respecto a la **infraestructura y desarrollo urbano**, Veracruz cuenta con una red vial más densa y conectada, lo que permitió la operación simultánea de dos almacenes regionales (Jesús Carranza y Las Choapas) con amplias zonas de cobertura. En cambio, Cacahoatán presenta una infraestructura vial limitada, con carreteras secundarias y caminos rurales susceptibles a interrupciones durante las lluvias, razón por la cual el modelo optó por un esquema centralizado con un único almacén de alta eficiencia.

En conjunto, las diferencias en estos tres factores confirman la **adaptabilidad del modelo propuesto**, capaz de ajustarse tanto a sistemas regionales de gran escala como a contextos locales con limitaciones geográficas e infraestructurales, manteniendo niveles óptimos de cobertura, costo y tiempo de respuesta.

9.6. Validación y análisis de robustez

Para garantizar que el modelo logístico propuesto sea confiable y útil en situaciones reales de emergencia, se realizaron pruebas de validación bajo distintos escenarios de operación. Se evaluó tanto la robustez frente a condiciones adversas como el desempeño bajo condiciones estándar, utilizando métricas clave de cobertura, eficiencia y efectividad en la asignación de recursos.

9.6.1. Escenarios de prueba implementados

El modelo demostró robustez operativa mediante la evaluación de múltiples escenarios adversos. La configuración de un solo almacén activo mostró capacidad para mantener la cobertura total bajo diversas condiciones de estrés operativo.

9.6.2. Métricas de desempeño en condiciones estándar

- **Fill rate alcanzado:** 100 %
- **Cobertura poblacional:** 100 %
- **Tasa de éxito de optimización:** 100 %
- **Eficiencia en asignaciones:** 4/4 localidades cubiertas

La concentración de operaciones en un único almacén estratégicamente ubicado demostró ser adecuado para la escala de la emergencia en Cacahoatán, simplificando la gestión logística y reduciendo costos de coordinación.

9.7. Conclusiones del caso Chiapas

La implementación del modelo de optimización logística en el municipio de Cacahoatán, Chiapas, confirma su efectividad en contextos de alta vulnerabilidad y dispersión poblacional. Los resultados muestran que un enfoque multicriterio para la identificación de localidades estratégicas permite seleccionar nodos logísticos óptimos, como Salvador Urbina, combinando infraestructura preexistente, conectividad vial y ubicación fuera de zonas inundables. Esta selección asegura la cobertura total de la población afectada de manera eficiente.

Contrario a las expectativas iniciales de requerir múltiples almacenes, el modelo demuestra que un único almacén centralizado puede atender el 100 % de los habitantes de las localidades críticas, simplificando la operación logística y reduciendo costos sin comprometer la efectividad del sistema. La optimización basada en el modelo EOQ tradicional se revela robusta, con tasas de éxito del 100 % en la convergencia del algoritmo, evidenciando la consistencia de las cantidades económicas de pedido convencionales en contextos locales.

La distribución de recursos alcanzó un balance adecuado entre productos básicos (78.25 % del presupuesto) y atenciones especializadas por grupos de edad (21.75 %), asegurando una respuesta humanitaria integral y priorizando los bienes más críticos sin comprometer la cobertura poblacional.

9.7.1. Implicaciones para la planeación logística

La experiencia en Cacahoatán sugiere que, para municipios con características similares de dispersión poblacional y vulnerabilidad hidrometeorológica, las configuraciones logísticas centralizadas alrededor de nodos estratégicos ofrecen soluciones eficientes y costo-efectivas. La metodología desarrollada demuestra capacidad para adaptarse a las particularidades del territorio chiapaneco, constituyéndose en una herramienta valiosa para la planificación anticipada de respuestas a emergencias en el estado.

9.7.2. Perspectivas de escalabilidad

Los resultados obtenidos establecen las bases para extender la implementación del modelo a otros municipios de Chiapas con perfiles de riesgo similares. Esta expansión contribuiría al fortalecimiento de la resiliencia logística regional frente a desastres hidrometeorológicos, especialmente en el contexto del cambio climático, y permitiría replicar la eficiencia alcanzada en Cacahoatán en escenarios de mayor escala y complejidad.

10 Conclusiones

Esta investigación ha desarrollado y validado rigurosamente un modelo de optimización determinista robusto para la gestión integrada de la ubicación de almacenes preposicionados y el control de inventario humanitario ante inundaciones en el sureste de México. El trabajo articula de manera coherente fundamentos teóricos de optimización no lineal, algoritmos numéricos de memoria limitada y aplicaciones empíricas en contextos de alta vulnerabilidad, estableciendo un marco metodológico reproducible para la planificación anticipada de respuestas a desastres. Las conclusiones que se presentan a continuación responden sistemáticamente a los tres objetivos específicos planteados, integrando las contribuciones teóricas de los Capítulos 4–5, los fundamentos algorítmicos del Capítulo 6 y los resultados empíricos de los Capítulos 8–9.

En relación con el **objetivo específico 1** adaptar un modelo de localización–inventario humanitario bajo incertidumbre parametrizada que extienda el modelo clásico de cantidad económica de pedido (EOQ) para incluir inventario de seguridad, faltantes estocásticos penalizados y decisiones binarias de apertura de almacenes, se concluye que la formulación propuesta supera rigurosamente las limitaciones estructurales del EOQ tradicional mediante tres innovaciones teóricas fundamentales:

1. **Parametrización de la incertidumbre:** incorporación del inventario de seguridad $Z_\alpha\sigma_L$ como término determinista que internaliza la variabilidad estocástica sin requerir formulaciones estocásticas explícitas.
2. **Penalización asimétrica explícita por faltantes:** incorporación del coeficiente βC que refleja la gravedad ética de los faltantes en contextos humanitarios, priorizando cobertura poblacional sobre minimización económica pura.
3. **Convexidad condicional en estructuras híbridas MINLP:** demostración de que, al fijar las decisiones binarias de apertura de almacenes, el subproblema continuo recupera convexidad poliédrica, justificando teóricamente el esquema de descomposición híbrido.

Como se demostró en la Sección 5.4, la incorporación del inventario de seguridad parametrizado $Z_\alpha\sigma_L$ y la penalización explícita por faltantes mediante el término βC transforman la función de costo total en una expresión no convexa y no separable:

$$Z(Q) = \frac{D(S + \beta C)}{Q} + \frac{H}{2}Q + Z_\alpha\sigma_L H + DC,$$

cuya estructura refleja con mayor fidelidad los riesgos operativos en contextos humanitarios. El análisis de las propiedades analíticas (ver Sección 5.5) confirmó que, pese a la pérdida de convexidad global inducida por las variables binarias de apertura de almacenes, el subproblema continuo asociado a una configuración fija de infraestructura mantiene convexidad condicional, lo cual justifica teóricamente el enfoque de descomposición adoptado. Esta propiedad, formalizada en la Proposición (convexidad condicional) de la Sección 5.1.3.2, garantiza que para cualquier vector fijo $\bar{x} \in \{0, 1\}^m$, el subproblema

$$\min_{(y,s) \in X(\bar{x})} Z(\bar{x}, y, s)$$

es un problema de programación lineal con conjunto factible poliédrico convexo. Además, el estudio comparativo cuantitativo (ver Sección 5.4.5) evidenció que el modelo extendido produce políticas de inventario significativamente más robustas que el EOQ clásico: en el ejemplo ilustrativo con demanda estocástica, el tamaño óptimo de lote aumentó de $Q_{\text{EOQ}} = 200$ a $Q^* \approx 490$ unidades, reduciendo en un 8.4% el costo total esperado y mitigando drásticamente el riesgo de faltantes catastróficos. Esta adaptación teórica, fundamentada en los preliminares matemáticos del Capítulo 4 (convexidad, diferenciabilidad y estructura de conjuntos), constituye una contribución que permite internalizar la incertidumbre sin recurrir a formulaciones estocásticas complejas, preservando así la rigurosidad analítica y la eficiencia computacional.

Respecto al **objetivo específico 2** desarrollar un esquema de solución híbrido basado en optimización mixta entera no lineal (MINLP) que combine enumeración sobre variables binarias con el algoritmo L-BFGS-B para resolver eficientemente los subproblemas continuos no convexos, se concluye que la estrategia propuesta logra un equilibrio óptimo entre rigor teórico y viabilidad práctica. El Capítulo 6 estableció las bases matemáticas para garantizar la convergencia global del método L-BFGS-B hacia puntos que satisfacen las condiciones de Karush–Kuhn–Tucker bajo restricciones tipo caja, demostrando que, bajo hipótesis de Lipschitz-continuidad del gradiente y compacidad del dominio, toda subsecuencia convergente de iteraciones satisface $\nabla_{\Omega} f(x_k) \rightarrow 0$ (Teorema 6.2, Sección 6.3.2.1). Esta garantía teórica, complementada con el análisis de sensibilidad respecto al parámetro de memoria m (ver Sección 6.4.2), permitió configurar el algoritmo con $m = 10$ como valor óptimo para los problemas estudiados, logrando convergencia en menos de 30 iteraciones incluso en escenarios mal condicionados. La implementación computacional descrita en el Capítulo 7 validó empíricamente esta estrategia: al combinar la enumeración exhaustiva sobre las 2^m configuraciones posibles de almacenes (factible dado el tamaño reducido de m en los casos de estudio) con la resolución eficiente de los subproblemas continuos mediante L-BFGS-B, se obtuvo una tasa de éxito del 100% en la convergencia para ambos escenarios analizados. Este esquema híbrido supera las limitaciones de métodos puramente heurísticos que no garantizan optimalidad y de enfoques exactos de gran escala que resultan computacionalmente prohibitivos, posicionándose como una alternativa robusta para problemas de localización–inventario con estructura mixta. La comparación numérica presentada en la Sección 6.4.4 demuestra que L-BFGS-B ofrece el

mejor compromiso entre eficiencia (7 iteraciones frente a 1,183 del gradiente descendente) y respeto a las restricciones físicas (a diferencia de BFGS sin proyección), justificando su elección como una buena alternativa de optimización para los subproblemas continuos.

En cuanto al **objetivo específico 3** verificar el desempeño del modelo propuesto mediante estudios de caso en regiones vulnerables a inundaciones en Veracruz y replicar en el estado de Chiapas, se concluye que la metodología desarrollada demuestra una notable adaptabilidad y robustez operativa en contextos geográficos y demográficos heterogéneos. Los resultados del Capítulo 8 para Veracruz evidenciaron que la descentralización logística mediante dos almacenes (Jesús Carranza y Las Choapas) reduce en un 23 % los costos de transporte respecto a una configuración centralizada (648.31 millones MXN vs. 837.54 millones MXN), elevando simultáneamente el *fill rate* promedio del 88 % al 95 %. Por su parte, la aplicación en Cacahoatán, Chiapas (Capítulo 9), reveló que, pese a las limitaciones infraestructurales y la topografía accidentada características de la región, una configuración centralizada con un único almacén estratégico en Salvador Urbina logra cobertura del 100 % de la población afectada (7407 habitantes) con un costo total anual optimizado de 195.46 millones MXN, aproximadamente el 30 % del costo observado en Veracruz. Esta diferencia, atribuible a la menor escala territorial y demográfica del municipio chiapaneco, no compromete la eficiencia operativa, lo cual confirma que el modelo se adapta dinámicamente a las particularidades del territorio sin requerir recalibración estructural. La validación cruzada entre ambos escenarios demuestra que el enfoque propuesto no es un ejercicio teórico aislado, sino una herramienta práctica capaz de generar recomendaciones accionables para autoridades de protección civil, con implicaciones directas en la reducción de tiempos de respuesta y la mejora de la equidad en la distribución de recursos durante emergencias hidrometeorológicas. La robustez del modelo se verificó adicionalmente mediante pruebas de sensibilidad ante variaciones del 10 % en la demanda, manteniendo un *fill rate* superior al 90 % en todos los escenarios analizados (ver Sección 9.6.2).

Esta investigación contribuye a la literatura de optimización aplicada y logística humanitaria mediante la integración coherente de tres elementos:

1. una formulación teórica rigurosa que extiende el EOQ clásico para incorporar incertidumbre parametrizada sin sacrificar rigurosidad analítica;
2. un esquema algorítmico híbrido con garantías de convergencia teóricamente fundamentadas y eficiencia computacional validada empíricamente; y
3. una verificación práctica en dos contextos mexicanos de alta vulnerabilidad que demuestra la adaptabilidad y utilidad operativa del modelo. Estas contribuciones no solo avanzan en el desarrollo metodológico de la optimización mixta entera no lineal, sino que también ofrecen un marco de trabajo replicable para la planificación anticipada de respuestas a desastres en regiones con recursos limitados.

Como trabajo futuro, se propone extender el modelo actual en dos direcciones complementarias con fundamentación teórica sólida: primero, integrar variables climáticas

predictivas (como pronósticos de precipitación a 72 horas) para habilitar un sistema de reoptimización en tiempo real que ajuste las decisiones logísticas conforme evoluciona el evento hidrometeorológico, siguiendo los marcos de optimización robusta adaptable desarrollados por Bertsimas y Thiele (2006); y segundo, escalar la metodología a nivel estatal en Chiapas mediante una descomposición jerárquica que combine optimización a nivel municipal con coordinación regional, lo cual permitiría evaluar sinergias entre municipios y diseñar redes logísticas resilientes a nivel sistémico mediante técnicas de descomposición de Benders generalizado propuestas por Wolsey (1998). Estas extensiones no solo enriquecerían el modelo teórico, sino que también incrementarían su impacto práctico en la reducción de la vulnerabilidad territorial ante el cambio climático.

Referencias

- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, y David Heath. 1999. «Coherent Measures of Risk». *Mathematical Finance* 9 (3): 203-28. <https://doi.org/10.1111/1467-9965.00068>.
- Balcik, Burcu, y Benita M. Beamon. 2008. «Facility location in humanitarian relief». *Journal of the Operational Research Society* 59 (8): 1005-16. <https://doi.org/10.1080/13675560701561789>.
- Barojas-Payán, Eduardo et al. 2021. «Optimization model to locate pre-positioned warehouses». En *Disaster Risk Reduction in Mexico*, editado por Daniel Sánchez-Partida, 169-98. Cham: Springer. https://doi.org/10.1007/978-3-030-67295-9_8.
- Bazaraa, Mokhtar S., Hanif D. Sherali, y C. M. Shetty. 2013b. *Nonlinear Programming: Theory and Algorithms*. 3.^a ed. Wiley. <https://doi.org/10.1002/9781118762371>.
- . 2013a. *Nonlinear Programming: Theory and Algorithms*. 3rd ed. Hoboken, NJ: John Wiley & Sons.
- Bertsekas, Dimitri P. 1999. *Nonlinear Programming*. Athena Scientific.
- . 2016. *Nonlinear Programming*. 3rd ed. Belmont, MA: Athena Scientific.
- Bertsimas, Dimitris, y Aurélie Thiele. 2006. «A robust optimization approach to inventory theory». *Operations Research* 54 (1): 150-68. <https://doi.org/10.1287/opre.1050.0238>.
- Boyd, Stephen, y Lieven Vandenbergh. 2004. *Convex Optimization*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804441>.
- Byrd, Richard H., Peihuang Lu, Jorge Nocedal, y Ciyong Zhu. 1995. «A Limited Memory Algorithm for Bound Constrained Optimization». *SIAM Journal on Scientific Computing* 16 (5): 1190-1208. <https://doi.org/10.1137/0916069>.
- Caunhye, A. M., X. Nie, y S. Pokharel. 2012. «Optimization models in emergency logistics: A literature review». *Socio-Economic Planning Sciences* 46 (1): 4-13. <https://doi.org/10.1016/j.seps.2011.04.004>.
- Conn, Andrew R., Nicholas I. M. Gould, y Philippe L. Toint. 1992. *LANCELOT: A Fortran Package for Large-Scale Nonlinear Optimization (Release A)*. Vol. 17. Springer Series en Computational Mathematics. Berlin, Heidelberg: Springer-Verlag. <https://doi.org/10.1007/978-3-662-12220-5>.
- Daskin, Mark S. 2013. *Network and Discrete Location: Models, Algorithms, and Applications*. 2nd ed. Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781118537015>.
- Dennis, John E., y Jorge J. Moré. 1977. «Quasi-Newton methods, motivation and theory». *SIAM Review* 19 (1): 46-89. <https://doi.org/10.1137/1019005>.
- Fletcher, Roger. 1987. *Practical Methods of Optimization*. 2nd ed. Chichester: John Wiley & Sons.

- Griewank, A., y Ph. L. Toint. 1984. «Partitioned variable metric updates for large structured optimization problems». *Numerische Mathematik* 39 (3): 325-47. <https://doi.org/https://doi.org/10.1007/BF01407876>.
- Grossmann, Ignacio E. 2002. «Review of nonlinear mixed-integer and disjunctive programming techniques». *Computers & Chemical Engineering* 26 (9): 1261-82. <https://doi.org/10.1023/A:1021039126272>.
- Hadley, G., y T. M. Whitin. 1963. *Analysis of Inventory Systems*. Englewood Cliffs, NJ: Prentice-Hall.
- Hernández Vázquez, Yulissa del Rocío. 2026. «Código fuente y anexos – Tesis: Ubicación de almacenes preposicionados e inventario humanitario». GitHub. 2026. <https://github.com/RocioHernandez06/Tesis/tree/main/anexos>.
- Insani, Muhammad, Agus Widodo, y A. Rahman. 2024. «Mixed-Integer Programming Model for Evacuation and Relief Distribution in Flood Contexts». *International Journal of Disaster Risk Reduction* 100: 104342. <https://doi.org/10.1016/j.ijdr.2024.104342>.
- Lewis, Adrian S. 2003. «The Mathematics of Eigenvalue Optimization». *Acta Numerica* 12: 1-58. <https://doi.org/10.1017/S0962492902000045>.
- Lewis, Adrian S., y Michael L. Overton. 2013. «Nonsmooth optimization via quasi-Newton methods». *Mathematical Programming* 141 (1–2): 135-63. <https://doi.org/10.1007/s10107-012-0514-2>.
- Liu, Dong C., y Jorge Nocedal. 1989. «On the Limited Memory BFGS Method for Large Scale Optimization». *Mathematical Programming* 45 (1–3): 503-28. <https://doi.org/10.1007/BF01589116>.
- Luenberger, David G., y Yinyu Ye. 2016. *Linear and Nonlinear Programming*. 4th ed. Cham, Switzerland: Springer. <https://doi.org/10.1007/978-3-319-18842-3>.
- Mashrut, S., y A. Rahimi. 2024. «Robust-fuzzy-probabilistic bi-objective model for post-flood relief logistics». *Annals of Operations Research*. <https://doi.org/10.1007/s10479-024-05925-7>.
- Nocedal, Jorge. 1980. «Updating Quasi-Newton Matrices with Limited Storage». *Mathematics of Computation* 35 (151): 773-82. <https://doi.org/10.1090/S0025-5718-1980-0562124-8>.
- Nocedal, Jorge, y Stephen J. Wright. 2006. *Numerical Optimization*. 2nd ed. New York: Springer. <https://doi.org/https://doi.org/10.1007/978-0-387-40065-5>.
- Porteus, E. L. 2002. *Foundations of Stochastic Inventory Theory*. Business/Marketing. Stanford University Press. <https://books.google.com.mx/books?id=ogsrj7PnyHIC>.
- Pujiana, R., D. Utama, y D. Rahmawati. 2020. «Multi-depot vehicle routing problem for post-flood humanitarian distribution». *Natural Hazards* 104: 2287-2309. <https://doi.org/10.1007/s11069-020-04250-z>.
- Romero-Mancilla, J., José Luis Martínez-Flores, y D. Sánchez-Partida. 2024. «Multi-objective multimodal humanitarian logistics with drones for flood relief». *Transportation Research Part E: Logistics and Transportation Review* 187: 103507. <https://doi.org/10.1016/j.tre.2024.103507>.
- Rosenbrock, H. H. 1960. «An Automatic Method for Finding the Greatest or Least Value of a Function». *The Computer Journal* 3 (3): 175-84. <https://doi.org/10.1093/comjnl>

/3.3.175.

- Santana-Robles, A., E. López, y J. Rivera. 2024. «Hybrid MILP and VRP model for shelter allocation and relief distribution». *Journal of Humanitarian Logistics and Supply Chain Management* 14 (2): 189-210. <https://doi.org/10.1108/JHLSCM-12-2023-0085>.
- Shapiro, Alexander, Darinka Dentcheva, y Andrzej Ruszczyński. 2021. *Lectures on Stochastic Programming: Modeling and Theory*. 2nd ed. SIAM. <https://doi.org/10.1137/1.9781611976595>.
- Sheikholeslami, Reza, y Nima Zarrinpoor. 2022. «Multi-period MILP under uncertainty for humanitarian logistics in flood response». *Computers & Industrial Engineering* 165: 107930. <https://doi.org/10.1016/j.cie.2022.107930>.
- Snyder, Lawrence V. 2006. «Facility location under uncertainty: A review». *IIE Transactions* 38 (7): 537-54. <https://doi.org/10.1080/07408170500216480>.
- Wolsey, L. A. 1998. *Integer Programming*. Wiley Series en Discrete Mathematics y Optimization. Wiley. <https://books.google.com.mx/books?id=x7RvQgAACAAJ>.
- Zhu, C., R. H. Byrd, P. Lu, y J. Nocedal. 1997. «Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization». *ACM Transactions on Mathematical Software* 23 (4): 550-60. <https://doi.org/https://doi.org/10.1145/279232.279236>.
- Zipkin, Paul H. 2000. *Foundations of Inventory Management*. McGraw-Hill.
- Zoutendijk, G. 1970. «Nonlinear programming, computational methods». En *Integer and Nonlinear Programming*, editado por J. Abadie, 37-86. Amsterdam: North-Holland. <https://zbmath.org/0336.90057>.