

# Convexidad Global en el Problema del Cajero Viajante Bi-Objetivo

Marcos Villagra

Benjamín Barán

Oswaldo Gómez

Centro Nacional de Computación

Universidad Nacional de Asunción

Email: {mvillagra, bbaran}@cnc.una.py, ogomez@illigal.ge.uiuc.edu

URL: <http://www.cnc.una.py>

9 de enero de 2006

Reporte Técnico RT-001/05

# Nomenclatura

|                        |   |
|------------------------|---|
| $A$                    | Matriz de coeficientes para un LSP                                    |
| $b_i$                  | Cota para la $i$ -ésima restricción de un OP                          |
| $cat(x, P)$            | Categoría de la solución $x$ en la población $P$                      |
| $c^T$                  | Vector de constantes reales para un LPP                               |
| $\delta(P)$            | conjunto de distancias al óptimo global para el TSP                   |
| $\delta^P(P)$          | conjunto de distancias a la población                                 |
| $\delta(x, P)$         | distancia media de una solución a la población                        |
| $\delta(x_p, x_q)$     | distancia, similitud, bondad entre dos soluciones                     |
| $d(i, j)$ o $d_{ij}$   | distancia entre ciudades $v_i$ y $v_j$                                |
| $E$                    | Conjunto de arcos de $G$  |
| $\eta$                 | Entero positivo para la construcción de conjuntos maximales dispersos |
| $f(x)$                 | Función objetivo de un OP   |
| $G$                    | Grafo completo del TSP  |
| $\gamma_i(x)$          | $i$ -ésima restricción de un OP                                       |
| $\gamma(x)$            | Restricciones de un MOP   |
| $h(i, j)$ o $h_{ij}$   | Función vectorial que asigna costos a los arcos de $G$                |
| $k$                    | Número de funciones objetivos   |
| $\lambda$              | Vector de peso para las funciones de escalarización                   |
| $\Lambda$              | Conjunto de vectores de peso  |
| $\lambda_i$            | $i$ -ésimo elemento del vector de peso $\lambda$                      |
| $\langle i, j \rangle$ | Camino del vértice $v_i$ a $v_j$                                      |
| $L(P)$                 | conjunto de soluciones de una población                               |
| $l(x)$                 | Longitud de un <i>tour</i>  |
| $m$                    | Número de restricciones de un OP                                      |

|                              |  |
|------------------------------|--|
| $m$                          | Número de restricciones  |
| $\mathbf{NF}(P)$             | Frontera no dominada de una población $P$                                |
| $\mathbf{R}^n$               | Conjunto de los reales en el espacio $n$ dimensional                     |
| $\mathbf{X}$                 | Espacio de decisión  |
| $\mathbf{X}_f$               | Conjunto de soluciones factibles de un MOP                               |
| $\mathbf{Y}$                 | Espacio objetivo   |
| $\mathbf{Y}_f$               | Conjunto de vectores objetivos factibles de un MOP                       |
| $\mathbf{Z}$                 | Espacio de soluciones alcanzables de MOP                                 |
| $\mathcal{D}(f)$             | Dominio de la función $f$  |
| $n$                          | Número de ciudades de $G$  |
| $P$                          | Población de soluciones  |
| $P(e)$                       | $e$ mejores soluciones de $P$  |
| $\rho$                       | Coefficiente de correlación  |
| $rng(x, P)$                  | Rango de la solución $x$ en la población $P$                             |
| $s_{aTch}$                   | Función de escalarización de sumas ponderadas aumentada de Tchebycheff   |
| $s_{ws}$                     | Función de escalarización de sumas ponderadas                            |
| $\tau(\lambda^i, \lambda^j)$ | Distancia manhattan entre los vectores de peso $\lambda^i$ y $\lambda^j$ |
| $\tau(z^i, z^j)$             | Distancia manhattan entre los vectores de peso $z^i$ y $z^j$             |
| $V$                          | Conjunto de vértices o ciudades de $G$                                   |
| $v_i$                        | $i$ -ésima ciudad de $G$   |
| $w$                          | Número de fila de $A$  |
| $x^*$                        | solución óptima  |
| $x$                          | <i>tour</i> , recorrido, solución del TSP                                |
| $x_q$                        | $q$ -ésima solución  |
| $x \triangleright x'$        | $x$ domina o es igual a $x'$   |
| $x \succ x'$                 | $x$ domina a $x'$  |
| $z^*$                        | Solución ideal de un MOP   |
| $z$                          | Solución alcanzable de un MOP  |

### **Resumen**

El presente trabajo estudia la topología del espacio de soluciones del Problema del Cajero Viajante o Traveling Salesman Problem (TSP) en un contexto de optimización Bi-Objetivo de minimización. Para el análisis, se introducen los conceptos de categoría y rango de una solución, los cuales relacionan a una solución en el espacio con el conjunto de soluciones no dominadas. Los resultados obtenidos son congruentes con la conjetura de un espacio de soluciones globalmente convexo del TSP Bi-Objetivo.

**Palabras Claves:** Problema del Cajero Viajante, Optimización Multiobjetivo, Convexidad Global.

# Capítulo 1

## Introducción

Las metaheurísticas conforman una familia de algoritmos de optimización que constituyen hoy en día la mejor opción para resolver problemas muy complejos. Estos algoritmos tratan de combinar métodos heurísticos básicos en paradigmas de alto nivel orientados a explorar efectiva y eficientemente el espacio de búsqueda [16]. El término metaheurística deriva de la composición de dos palabras griegas, heurística deriva del verbo *heuriskein* que significa "encontrar", mientras el prefijo meta significa "más allá, en un nivel superior".

Esta clase de algoritmos incluye a, Optimización por Colonia de Hormigas (*Ant Colony Optimization, ACO*), Algoritmos Evolutivos (*Evolutionary Algorithms, EC*), Búsqueda Local Iterativa (*Iterated Local Search, ILS*), Templado Simulado (*Simulated Annealing, SA*), Búsqueda Tabu (*Tabu Search, TS*), por citar algunos [10].

El área de investigación sobre metaheurísticas ha evolucionado sobre la base de prueba y error [13], motivado en mejorar las soluciones de un problema dado, y no en identificar las razones de los éxitos y fracasos de estos algoritmos.

El Problema del Cajero Viajante o TSP, del inglés *Traveling Salesman Problem*, ha sido utilizado como problema de prueba o *benchmark* para estudiar el comportamiento de muchas metaheurísticas.

En [1], [2] y [3] se han realizado estudios del espacio de soluciones del TSP. Estos estudios consistieron en hallar óptimos locales mediante distintos algoritmos de búsqueda local, para luego analizar los resultados obtenidos e intentar entender la estructura del TSP para las instancias estudiadas. Estos trabajos fueron realizados de una manera más sistemática en [5], donde se obtuvieron muestras del espacio de búsqueda para compararlas con la solución óptima y con otras soluciones factibles; con eso en mente, se utilizó una heurística de optimización local, y muestras de soluciones uniformemente distribuidas en el espacio de soluciones para cada problema estudiado. Los resultados experimentales de todos estos trabajos indican la existencia de una estructura globalmente convexa del espacio de soluciones del TSP.

La convexidad global no es convexidad en un sentido estricto [13], pero puede ser usada para denotar la observación empírica de que los mejores óptimos locales se encuentran mayormente juntos en una pequeña parte del espacio de soluciones, el cual podría o no incluir al óptimo global. Las metaheurísticas explotan esto concentrando la búsqueda en esa parte del espacio de soluciones.

Se han realizado estudios sobre el espacio de soluciones del TSP Bi-Objetivo

en [4], donde se propone un enfoque para lidiar con problemas de optimización multiobjetivo, presentando a los algoritmos de búsqueda local como punto de referencia para comparar el rendimiento de metaheurísticas en problemas de optimización combinatoria multiobjetivo.

En [13] se presentó un estudio de la convexidad global para el TSP de tres objetivos mediante la utilización de funciones de escalarización, que asignan a cada solución un número, indicando la calidad de la solución. Las funciones utilizadas consistieron en el *weighted sums scalarizing function* y el *augmented Tchebycheff scalarizing function*; mediante estas herramientas se pudo demostrar experimentalmente la presencia de convexidad global para instancias del TSP de tres objetivos.

El presente trabajo, realiza un estudio de la estructura del espacio de soluciones del TSP Bi-Objetivo de una manera más práctica que la realizada en [13], mediante la introducción de dos métricas de calidad, la categoría de una solución y el rango de una solución. El primero relaciona a una solución con el número de soluciones que la dominan; y el segundo establece una jerarquía entre las soluciones. Para dicho estudio fueron creadas instancias con 7, 8, 9 y 10 ciudades, con las que se realizó una búsqueda exhaustiva para hallar el universo de soluciones de los mismos. Luego, se utilizaron problemas de 100 y 150 ciudades, y, se trabajó con subconjuntos de soluciones de cada instancia.

Si la convexidad global existe, esta puede ser utilizada para la creación de nuevas metaheurísticas y explicar el funcionamiento de otras ya usadas con éxito en la actualidad.

El resto del trabajo queda organizado de la siguiente manera, en el capítulo 2 se define el TSP y se introduce la notación que se usará durante todo el trabajo. El capítulo 3 introduce el campo de la optimización convexa, y el concepto de convexidad global. En el capítulo 4 se presentan los problemas de optimización multiobjetivo junto con el TSP Bi-Objetivo y su formulación. En la sección 5 se define la metodología de trabajo, las herramientas utilizadas y los resultados obtenidos. En el capítulo 6 se dan las conclusiones obtenidas y direcciones para trabajos futuros.

## Capítulo 2

# El Problema del Cajero Viajante

Un miembro prominente del rico conjunto de problemas de optimización combinatoria es indudablemente el TSP. Este problema tiene una larga historia, con referencias desde el siglo XIX y es uno de los pocos problemas matemáticos que aparece frecuentemente en la prensa científica e inclusive en diarios [8].

El estudio de este problema ha atraído a muchos investigadores de diferentes campos, e.g., Matemática, Investigación de Operaciones, Física, Biología e Inteligencia Artificial y existe una vasta literatura en el tema. Esto se debe a que, aunque es fácil de formularlo, exhibe todos los aspectos complejos de un problema de optimización combinatoria. El TSP ha servido y continúa sirviendo como problema de prueba para distintos algoritmos, como Búsqueda Local (*Local Search*), Templado Simulado (*Simulated Annealing*), Búsqueda Tabu (*Tabu Search*), Redes Neuronales (*Neural Networks*) y Computación Evolutiva (*Evolutionary Computation*), por citar algunos de ellos [10].

El TSP no es interesante solamente desde un punto teórico, sino que muchas aplicaciones prácticas pueden ser modeladas como una variante del mismo, y por ende, existe una tremenda necesidad de algoritmos que lo resuelvan. Dadas las múltiples áreas de aplicación, también hay una amplia colección de algoritmos para tratar los casos especiales.

En las dos últimas décadas, un enorme progreso se ha alcanzado en la resolución del TSP. Este progreso es debido parcialmente al creciente avance del poder de cómputo de las computadoras, pero sobre todo fue posible al desarrollo de una teoría matemática y de algoritmos eficientes. Aun considerando estos logros, el TSP está lejos de ser resuelto; muchos aspectos del problema todavía necesitan ser considerados y preguntas todavía necesitan ser satisfactoriamente respondidas.

Problemas con algunos cientos de ciudades son todavía difíciles de resolver para los algoritmos de hoy en día y requieren horas de procesamiento en supercomputadores. Esto es fácil de comprender si consideramos que en un problema de tan sólo 100 ciudades existen del orden de  $10^{155}$  recorridos posibles. Además, el crecimiento del número de recorridos en función al número de ciudades es exponencial, poniendo al TSP en la categoría de problemas  $\mathcal{NP}$ -hard [12].

El TSP puede ser representado por un grafo  $G = \langle V, E \rangle$  donde  $V = \{v_i\}$  es

el conjunto de vertices o ciudades con cardinalidad  $|V| = n$ , y

$$E = \{\langle i, j \rangle \mid \forall v_i, v_j \in V \quad v_i \neq v_j\}$$

es el conjunto de arcos que tiene asociado para cada par de ciudades  $\langle i, j \rangle$ , una distancia  $d(i, j)$  o simplemente  $d_{ij}$ . La solución del TSP es conocida como *tour* (recorrido), que representaremos con  $x$ , y está dada por un conjunto ordenado de  $n$  arcos que contiene a todas las ciudades, esto es,  $x = (\langle 1, 2 \rangle, \langle 2, 3 \rangle, \dots, \langle n-1, n \rangle, \langle n, 1 \rangle)$ . Esto es un circuito hamiltoniano [9]. Cada *tour* tiene asociado una longitud formada por la suma de cada uno de sus arcos.

**Definición 2.1** *Longitud de un tour  $x$ .*

$$l(x) = \sum_{i=1}^n d_{ij} \quad \text{donde } \langle i, j \rangle \in x.$$

Sean  $x_1$  y  $x_2$  dos soluciones, entonces decimos que  $x_1 = x_2$  si y solo si, todo arco  $\langle i, j \rangle$  de  $x_1$  es también un arco de  $x_2$ . En consecuencia  $l(x_1) = l(x_2)$ .

El objetivo del TSP es hallar un tour o circuito Hamiltoniano  $x$  de longitud mínima.

Un problema para el cual  $d_{ij} = d_{ji}$  es conocido como TSP *simétrico*. En este contexto, dados dos *tours*  $x_1$  y  $x_2$ , la distancia  $\delta(x_1, x_2)$  entre dos soluciones se define como  $n$  menos la cantidad de arcos comunes a ambos *tours* [11].

**Definición 2.2** *Distancia, similitud o bondad entre dos soluciones  $x_p$  y  $x_q$ .*

$$\delta(x_p, x_q) = n - |\{\langle i, j \rangle \mid \langle i, j \rangle \in x_p, x_q\}|$$

**Definición 2.3** *Una población  $P$  es un conjunto de soluciones de cardinalidad  $|P|$ .*

$P = \{x_1, x_2, \dots, x_{|P|}\}$  donde cada elemento corresponde a una solución válida, también denominado *tour* en adelante.

## Capítulo 3

# Optimización Convexa

La *optimización convexa* es una clase especial de problemas de optimización matemática, la cual incluye a problemas de mínimos cuadrados y programación lineal. Es bien sabido que los mínimos cuadrados y la programación lineal poseen una teoría bien completa, surgen en una variedad de aplicaciones, y pueden ser resueltos numéricamente de forma muy eficiente.

Mientras que las matemáticas de la optimización convexa a sido estudiada durante más de un siglo, varios progresos han estimulado un nuevo interés en el tópico. El primero de ellos, es el reconocimiento de que los métodos de punto-interior, desarrollados en los ochentas para resolver problemas de programación lineal, pueden ser usados también para resolver problemas de optimización convexa.

Como segundo avance aparece el descubrimiento de que los problemas de optimización convexa (más allá de los mínimos cuadrados y programación lineal) prevalecen más en la práctica de lo que era pensado. Desde 1990 [17] muchas aplicaciones han sido descubiertas en áreas como sistemas de control automatizado, procesamiento de señales, comunicaciones y redes, diseño de circuitos electrónicos, análisis de datos y modelado, estadísticas y finanzas. La optimización convexa también posee aplicaciones en optimización combinatoria y optimización global, donde es usado para encontrar cotas para el valor óptimo, así como aproximar soluciones.

Hay muchas ventajas en reconocer o formular un problema como un problema de optimización convexa. La ventaja más simple es que el problema puede ser resuelto de forma muy confiable y eficiente, usando métodos de punto-interior u otros métodos para optimización convexa[17]. Estos métodos son lo suficientemente confiables como para ser insertados en herramientas CAD (*Computer Aided Design*), o inclusive en sistemas automatizados de control de tiempo real.

### 3.1 Optimización Matemática

Un *problema de optimización matemática*, *problema de optimización* o simplemente OP (del inglés, *Optimization Problem*), se define de la siguiente manera:

**Definición 3.1** Dado un vector  $x = \langle x_1, \dots, x_n \rangle$ , definimos un problema de

optimización como:

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeto a} & \gamma_i(x) \leq b_i, \quad i = 1, \dots, m \end{array}$$

donde  $x$  es la variable de optimización o variable de decisión del problema, la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es la función objetivo, las funciones  $\gamma_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  son las funciones de restricción, y las constantes  $b_1, \dots, b_m$  son los límites o cotas de las restricciones.

Por lo general se considera familias o clases de problemas de optimización, caracterizados por formas particulares de la función objetivo y las restricciones. Dos de estas clases son los problemas de mínimos cuadrados y programación lineal.

### 3.1.1 Mínimos Cuadrados

Un problema de mínimos cuadrados o LSP (del inglés *Least-Squares Problem*) es un problema de optimización sin restricciones (i.e.,  $m = 0$ ) con una función objetivo obtenida como la suma de los cuadrados de los términos en la forma  $a_i^T x - b_i$ .

**Definición 3.2** Sea  $A \in \mathbb{R}^{w \times n}$  (con  $w \geq n$ ) la matriz de coeficientes y  $x$  la variable de decisión, definimos un problema de mínimos cuadrados como:

$$\text{Optimizar } f(x) = \|Ax - b\|^2 = \sum_{i=1}^w (a_i^T x - b_i)^2$$

donde  $a_i^T$  son las filas de  $A$ .

El problema de mínimos cuadrados es la base para el análisis de regresión, control óptimo, y muchos métodos de estimación de parámetros y análisis de datos.

Reconocer un problema de optimización como un problema de mínimos cuadrados es muy directo; solamente necesitamos verificar que la función objetivo es una función cuadrática.

Existen buenos algoritmos para resolver LSPs con mucha certeza y muy alta confiabilidad. El LSP puede ser resuelto en un tiempo aproximadamente proporcional a  $n^2 w$ . Una computadora de escritorio puede resolver un LSP con cientos de variables, y miles de términos, en unos cuantos segundos; computadores más poderosos pueden resolver problemas más grandes, o del mismo tamaño, mucho más rápido (aún más, estos tiempos disminuirán exponencialmente en el futuro, según la ley de Moore). Algoritmos y software para resolver problemas de mínimos cuadrados son lo suficientemente confiables para optimización en sistemas empujados [17].

Para problemas extremadamente grandes (digamos, millones de variables), o para problemas con requerimientos de computación en tiempo real precisos, resolver un problema de mínimos cuadrados puede ser un reto. Pero en la mayoría de los casos, podemos decir que los métodos existentes son muy efectivos, y extremadamente confiables. Además, podemos decir que resolver problemas de mínimos cuadrados es una tecnología madura, que puede ser usada por muchas personas que no saben, y que no necesitan saber los detalles.

### 3.1.2 Programación Lineal

Otra clase importante de problemas de optimización es la programación lineal o LPP (del inglés *Linear Programming Problem*), en el cual la función objetivo y todas las restricciones son lineales.

**Definición 3.3** Sea  $x$  el vector de decisión, definimos un problema de programación lineal como:

$$\begin{array}{ll} \text{Optimizar} & c^T x \\ \text{sujeto a} & a_i^T x \leq b_i, \quad i = 1, \dots, m \end{array}$$

Los vectores  $c, a_1, \dots, a_m \in \mathbb{R}^n$  y los escalares  $b_1, \dots, b_m \in \mathbb{R}$  son parámetros del problema que especifican la función objetivo y las restricciones.

No existe una fórmula analítica simple para resolver un problema de programación lineal (así como lo hay para mínimos cuadrados), pero existe una variedad de métodos muy efectivos para resolverlo, incluyendo al método simplex de Dantzig, y los más recientes métodos de punto interior. Como no podemos dar un número exacto de operaciones aritméticas requeridas para resolver un programa lineal (como se puede para mínimos cuadrados), podemos establecer cotas sobre el número de operaciones requeridas para resolver un programa lineal, con una precisión dada, usando un método de punto interior. La complejidad en la práctica es del orden  $n^2m$  (aumiendo  $m \geq n$ ); estos algoritmos son muy confiables, aunque tal vez no tanto como los métodos para mínimos cuadrados. Podemos resolver fácilmente problemas con cientos de variables y miles de restricciones en una computadora de escritorio en unos segundos.

Como ocurre con los mínimos cuadrados, aún es todo un reto resolver programas lineales muy grandes, o con requerimientos de tiempo real precisos. Pero, también podemos decir que resolver problemas de programación lineal es una tecnología madura. Los resolvers de programas lineales pueden ser (y son) insertados en muchas herramientas y aplicaciones.

## 3.2 Optimización Convexa

Antes de definir un problema de optimización convexa o COP (del inglés *Convex Optimization Problem*) presentamos dos definiciones fundamentales para la teoría de la optimización.

**Definición 3.4** Sea  $C$  un subconjunto de  $\mathbb{R}^n$ , i.e.  $C \subset \mathbb{R}^n$ , decimos que  $C$  es convexo si

$$\alpha x + (1 - \alpha)y \in C, \quad \forall x, y \in C, \quad \forall \alpha \in [0, 1]$$

Un conjunto es convexo si cada punto en el conjunto puede ser visto por otro punto del conjunto a lo largo de un camino recto sin obstrucción entre ellos, donde sin obstrucción se refiere a la pertenencia al conjunto. La figura 3.1 muestra algunos conjuntos convexos y no convexos [14].

**Definición 3.5** Sea  $C$  un conjunto convexo de  $\mathbb{R}^n$ . Una función  $f : C \rightarrow \mathbb{R}$  es convexa si

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

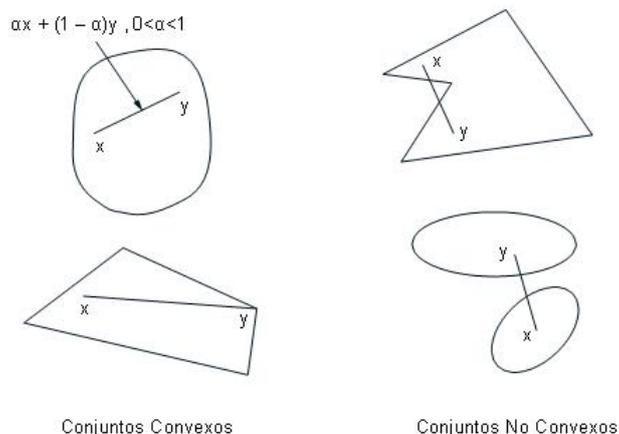


Figura 3.1: Ilustración de la definición de un conjunto convexo. Para la convexidad, la interpolación lineal entre dos puntos del conjunto debe de ser un punto en el conjunto.

La función  $f$  es cóncava si  $-f$  es convexa. Para una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , decimos que  $f$  es convexa para el conjunto  $C$  si se cumple la inecuación de arriba.

Geoméricamente, esta desigualdad significa que el segmento de línea entre  $\langle x, f(x) \rangle$  y  $\langle y, f(y) \rangle$ , se encuentra sobre el gráfico de  $f$  en la figura 3.2 [14].

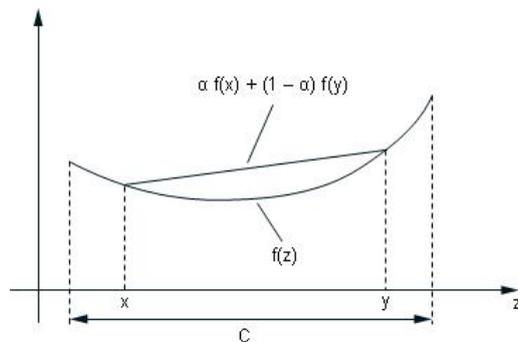


Figura 3.2: Ilustración de la definición de una función convexa. La interpolación lineal  $\alpha f(x) + (1 - \alpha)f(y)$  sobreestima el valor de la función  $f(\alpha x + (1 - \alpha)y)$ . Nótese que el dominio de la función debe ser un conjunto convexo.

Una función  $f$  es *estrictamente convexa* si una desigualdad estricta se da en la definición 3.5. Así mismo decimos que  $f$  es estrictamente cóncavo, si  $-f$  es estrictamente convexo.

También decimos que una función es convexa si y solo si es convexa cuando está restringida a alguna línea que intersecta su dominio. En otras palabras  $f$  es convexa si y solo si  $\forall x \in \mathcal{D}(f)$  y  $\forall v \in \mathbb{R}$ , la función  $g(t) = f(x + tv)$  es convexa (en su dominio,  $\{t \mid x + tv \in \mathcal{D}(f)\}$ ). Esta propiedad es muy útil, porque nos permite verificar si una función es convexa, restringiéndola a una línea.

Las siguientes proposiciones dan otras maneras de reconocer funciones convexas [14].

1. Una función lineal es convexa.
2. Un vector normal es convexo.
3. La suma ponderada de funciones convexas, con pesos positivos, es convexa.

**Definición 3.6** Sea  $x$  el vector de decisión, definimos un problema de optimización convexa como:

$$\begin{array}{ll} \text{Optimizar} & f(x) \\ \text{sujeeto a} & \gamma_i(x) \leq b_i, \quad i = 1, \dots, m \end{array}$$

donde las funciones  $f, \gamma_1, \dots, \gamma_m : \mathbb{R}^n \rightarrow \mathbb{R}$  son convexas.

Mínimos cuadrados (def. 3.2) y programación lineal (def. 3.3) son casos especiales del problema de optimización general (def. 3.6).

Es muy importante notar que tener a las funciones objetivos y restricciones como funciones convexas nos da muchas ventajas que podemos explotar a la hora de resolver COPs. Una de esas ventajas es que no existen óptimos locales que no sean globales, es decir, solo existe una solución óptima global a la que puede converger el problema. También existen dificultades, como reconocer COPs, o reconocer problemas que pueden transformarse a COPs.

En general no existe una fórmula analítica para solucionar un problemas de optimización convexa, pero (como programación lineal) existen métodos muy efectivos para solucionarlos. Los métodos de punto interior funcionan muy bien en la práctica, y en algunos casos puede ser probado que resuelven un problema a una precisión específica, en un número de operaciones que no excede una complejidad polinomial de las dimensiones del problema [17].

Métodos de puntos-interior pueden resolver un COP en un número de pasos que generalmente está en el rango de 10 y 100. Ignorando cualquier estructura en el problema (como esparcidad), cada paso requiere en el orden de  $\max\{n^3, n^2m, F\}$  operaciones, donde  $F$  es el costo de evaluar las primeras y segundas derivadas de la función objetivo y las restricciones.

Como los métodos para resolver problemas de programación lineal, estos métodos de punto-interior son bastante confiables. Fácilmente podemos resolver problemas con cientos de variables y miles de restricciones en una computadora de escritorio, en casi unas cuantas decenas de segundos. Explotando la estructura del problema (como esparcidad), podemos resolver problemas más grandes con miles de variables y restricciones.

Sin embargo, no podemos decir que resolver COPs generales es una tecnología madura, como mínimos cuadrados y programación lineal. Métodos de punto-interior para COPs generales no lineales es aún una área de investigación muy activa [17]. Sin embargo, es razonable decir que la resolución de COPs generales se convertirá en tecnología conocida en algunos años.

### 3.3 Convexidad Global

Un problema de optimización puede ser resuelto con mucha eficiencia si podemos detectar la presencia de convexidad en su función objetivo y restricciones, algo

que puede resultar difícil en muchas situaciones. Siempre es una ventaja conocer, apriori, la estructura del problema a resolver.

En los estudios sobre el TSP realizados por Boese en [1] y [2], se pudo observar la presencia de una estructura muy particular, llamada Convexidad Global o estructura de Gran Valle, donde las mejores soluciones se encuentran concentradas en una región pequeña del espacio de búsqueda, y, mientras mejor sea la solución más central será su posición respecto a las demás soluciones.

Para el estudio de la convexidad global del TSP, una definición de distancia media es presentada. Considérese una población  $P = \{x_1, x_2, \dots, x_{|P|}\}$ , definimos la distancia media de un recorrido  $x$  a una población  $P$  como:

**Definición 3.7** *Distancia media de un recorrido  $x$  a una población  $P$ .*

$\delta(x, P) = \frac{1}{|P|-1} \sum_{i=1}^{|P|} \delta(u, x) \quad \forall u \in P$ . Si  $x = x^*$  se tiene una noción de que tan cerca está una población a la solución óptima  $x^*$ .

Boese estudió en [1] el espacio de soluciones del TSP att532 de 532 ciudades, para lo cual fueron realizadas 2500 corridas de distintas heurísticas de búsqueda local. Para cada heurística,  $|P|$  recorridos diferentes fueron almacenados en un conjunto  $P$ ; y cada recorrido almacenado tenía:

- a. una longitud  $l(x)$  (el conjunto de soluciones se denota  $L(P) = \{l(x_p) \mid x_p \in P\}$ );
- b. una distancia a  $x^*$  denotada como  $\delta(x_p, x^*)$  para algún  $x_p \in P$  (el conjunto de distancias al óptimo global se denota  $\delta(P) = \{q \mid q = \delta(x, x^*) \wedge x \in P\}$ );
- c. una distancia media a las otras soluciones de  $P$  llamada  $\delta(x, P)$  (el conjunto de distancias medias a la población se denota  $\delta^P(P) = \bigcup_{x \in P} \{\delta(x, P)\}$ ).

Boese [1] calculó una correlación positiva entre estas tres variables. Dado que el conjunto  $L(P)$  tiene una correlación positiva con el conjunto  $\delta^P(P)$ , Boese sugirió una estructura globalmente convexa del espacio de soluciones del TSP. En otras palabras, mientras más central es la posición de una solución  $x$  dentro del conjunto de soluciones  $P$ , menor será su distancia media  $\delta(x, P)$  a las otras soluciones, por ende menor es su longitud esperada  $l(x)$ , i.e. mejor es la solución.

La convexidad global no es convexidad en un sentido estricto [7]. Considerando un contexto de minimización, Boese sugirió una analogía con una estructura de gran valle, estructura que vista desde lejos parecería tener un sólo mínimo, pero que al verla desde cerca, se observa que tiene varios mínimos locales [1]. Aunque no existe una definición formal, una interpretación intuitiva de una estructura globalmente convexa puede verse en la figura 3.3.

La idea de convexidad global se basa principalmente en dos hipótesis:

1. Convexidad: Los óptimos locales se encuentran juntos en una región relativamente pequeña del espacio de soluciones.
2. Centralidad: De todos los óptimos locales, los mejores están localizados centralmente respecto a los demás.

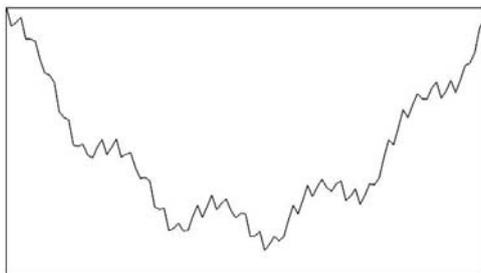


Figura 3.3: Interpretación intuitiva de un espacio de soluciones con estructura globalmente convexa

Si ambas hipótesis son válidas, debemos esperar que los mejores óptimos locales estén más cerca al óptimo global.

Además, solo podemos hablar de una estructura globalmente convexa cuando una topología haya sido inducida en el espacio de soluciones [13]. Una popular técnica de construcción de una topología está basada en un mapeo de proximidad o adyacencia, el cual consiste en definir un conjunto de soluciones vecinas para cada solución. Una vez establecido el mapeo<sup>1</sup>, una métrica basada en este puede ser definida en el espacio de soluciones. La distancia entre dos soluciones puede ser medida; por ejemplo, el mínimo número de veces que necesitamos aplicar el mapeo de manera a obtener una de estas soluciones a partir de la otra.

En particular, cuando se experimenta con heurísticas basadas en búsqueda local, lo que se prueba no es el rendimiento de la estrategia de búsqueda en el problema, sino como funciona en el problema dado una estructura de vecindad específica [13]. Es esa estructura, que define la topología en el espacio de soluciones, la que determinará las características de la superficie de costos, el cual es la piedra angular para la optimización. Un problema puede ser fácilmente resuelto con una estructura de vecindad e intratable con otro [13].

---

<sup>1</sup>También conocido como función de vecindad o estructura de vecindad

## Capítulo 4

# Optimización Multiobjetivo

En el área de optimización multiobjetivo, debido a la naturaleza aún incipiente del ámbito de investigación, no existe una notación estándar, y no ha sido hasta hace muy poco tiempo atrás que los investigadores han empezado a preocuparse por definir con claridad estos aspectos. Sin embargo, en gran parte de los trabajos consultados se percibe aún bastante confusión al respecto. Por lo tanto, es esencial establecer una notación clara antes de iniciar la discusión. A continuación se da la definición formal de un problema de optimización multiobjetivo.

**Definición 4.1** *Un Problema de Optimización Multiobjetivo o Multiobjective Optimization Problem (MOP) incluye un conjunto de  $n$  variables de decisión,  $k$  funciones objetivos y  $m$  restricciones. Las funciones objetivos y restricciones son funciones de las variables de decisión. Esto puede expresarse como:*

$$\begin{array}{ll} \text{Optimizar} & y = f(x) = (f_1(x), f_2(x), \dots, f_k(x)) \\ \text{Sujeto a} & \gamma(x) = (\gamma_1(x), \dots, \gamma_m(x)) \geq 0 \\ \text{donde} & x = (x_1, x_2, \dots, x_n) \in \mathbf{X} \text{ es el vector de decisión} \\ y & y = (y_1, y_2, \dots, y_k) \in \mathbf{Y} \text{ es el vector objetivo} \end{array}$$

donde  $\mathbf{X}$  denota el espacio de decisión mientras que el espacio objetivo es denotado por  $\mathbf{Y}$ . Dependiendo del tipo de problema, optimizar se refiere a minimizar o maximizar. El conjunto de restricciones  $\gamma(x) \geq 0$  determina el conjunto de soluciones factibles  $\mathbf{X}_f \subseteq \mathbf{X}$ , y su conjunto de vectores objetivos correspondiente  $\mathbf{Y}_f \subseteq \mathbf{Y}$ .

Un MOP consiste en encontrar una  $x$  que optimice  $f(x)$ , donde la función objetivo  $f(x)$  mapea  $\mathbf{X}_f$  a  $\mathbb{R}^k$ . Una solución  $z$  es alcanzable si existe una solución  $x \in \mathbf{X}_f$  tal que  $z = f(x)$ . El conjunto de soluciones alcanzables es denotado por  $\mathbf{Z}$ .

La solución ideal  $z^*$  se define como  $z^* = \{\min f_1(x), \min f_2(x), \dots, \min f_k(x)\}$ . Normalmente la solución ideal no es alcanzable, por lo que no existe una única mejor solución, sino que un conjunto de soluciones, dado que ninguna de ellas puede considerarse mejor que las otras cuando todos los objetivos son considerados al mismo tiempo. Esto deriva del hecho que existen relaciones de compromiso o *trade-offs* entre los objetivos. Así, un nuevo concepto de optimalidad debe de

establecerse para MOPs. Dados dos vectores de decisi3n  $x$  y  $x'$  tenemos que:

$$\begin{aligned} f(x) = f(x') &\Leftrightarrow \forall i \in \{1, 2, \dots, k\} \quad f_i(x) = f_i(x') \\ f(x) \leq f(x') &\Leftrightarrow \forall i \in \{1, 2, \dots, k\} \quad f_i(x) \leq f_i(x') \\ f(x) < f(x') &\Leftrightarrow f(x) \leq f(x') \wedge f(x) \neq f(x') \end{aligned}$$

As3, en un contexto de minimizaci3n, los vectores  $x$  y  $x'$  cumplen con una de tres condiciones:

**Definici3n 4.2** *Dominancia Pareto en un contexto de minimizaci3n. Para dos vectores  $x$  y  $x'$ ,*

$$\begin{aligned} x \succ x' \text{ (} x \text{ domina a } x') &\Leftrightarrow f(x) < f(x') \\ x' \succ x \text{ (} x' \text{ domina a } x) &\Leftrightarrow f(x') < f(x) \\ x \sim x' \text{ (} x \text{ y } x' \text{ son no comparables)} &\Leftrightarrow f(x) \not\prec f(x') \wedge f(x') \not\prec f(x) \end{aligned}$$

Alternativamente  $x \triangleright x'$  denotar3 que  $x$  domina o es igual a  $x'$ . Definido el concepto de dominancia Pareto, puede ser introducido el criterio de optimalidad Pareto de la siguiente manera:

**Definici3n 4.3** *Un vector de decisi3n  $x \in \mathbf{X}_f$  es no dominado con respecto a un conjunto  $\mathbf{Q} \subseteq \mathbf{X}_f$  ssi  $x \triangleright x' \quad \forall x' \in \mathbf{Q}$ .*

Cuando  $x$  es no dominado con respecto a todo el conjunto  $\mathbf{X}_f$  es llamado una soluci3n Pareto 3ptima; por lo tanto, el Conjunto Pareto 3ptimo  $\mathbf{X}_{true}$  se define como:

**Definici3n 4.4** *Dado el conjunto de vectores de decisi3n factibles  $\mathbf{X}_f$ , se denomina Conjunto Pareto 3ptimo  $\mathbf{X}_{true}$  al conjunto de vectores de decisi3n no dominados que pertenecen a  $\mathbf{X}_f$ :*

$$\mathbf{X}_{true} = \{x \in \mathbf{X}_f \mid x \text{ es no dominado con respecto a } \mathbf{X}_f\}$$

El correspondiente conjunto de vectores objetivos  $\mathbf{Y}_{true} = f(\mathbf{X}_{true})$  es el Frente Pareto 3ptimo.

## 4.1 El TSP Multiobjetivo

El TSP Multiobjetivo o MOTSP<sup>1</sup> consiste en un grafo completo  $G = \langle V, E, h \rangle$  donde  $V$  es el conjunto de v3rtices o ciudades con  $|V| = n$ ,  $E$  es el conjunto de arcos, aristas o caminos que conecta a todos los v3rtices y  $h$  es una funci3n que asigna a cada arco  $\langle i, j \rangle \in E$  un vector  $h(i, j)$  o simplemente  $h_{ij}$ , donde cada elemento corresponde a una cierta medida (e.g. distancia, costo) entre  $i$  y  $j$ , as3, el problema consiste en hallar un circuito Hamiltoniano de costo m3nimo, donde m3nimo se refiere a la noci3n de optimalidad Pareto. En este trabajo consideramos problemas sim3tricos, i.e.  $h_{ij} = h_{ji}$ .

As3, se define el TSP Bi-Objetivo como un MOP en un contexto de minimizaci3n de la siguiente forma:

<sup>1</sup>del ingl3s Multiobjective TSP

**Definición 4.5** *TSP Bi-Objetivo*

$$\begin{array}{ll} \text{Minimizar} & y = f(x) = (y_1 = f_1(x), y_2 = f_2(x)) \\ \text{sujeto a} & l(x) > 0 \\ \text{donde} & x = (\langle 1, 2 \rangle, \langle 2, 3 \rangle, \dots, \langle n-1, n \rangle, \langle n, 1 \rangle) \in \mathbf{X} \\ y & y = (y_1, y_2) = (f_1(x), f_2(x)) \in \mathbf{Y} \end{array}$$

En el presente trabajo se realiza un estudio del TSP Bi-Objetivo, donde podría considerarse por ejemplo a  $f_1$  como la distancia entre ciudades, y a  $f_2$  a la calidad del camino o el tiempo requerido en recorrerlo.

## 4.2 Convexidad Global en el TSP Multiobjetivo

Hansen y Borges realizaron estudios sobre la convexidad global del TSP de tres objetivos en [13]. El objetivo principal de este trabajo fue mostrar la existencia de convexidad global en un problema de optimización combinatoria usando una estructura de vecindad común. Con esto en mente utilizaron funciones de escalarización<sup>2</sup>, las cuales ordenan el conjunto objetivo según una escala de calidad de las soluciones. Estas funciones consistieron en dos familias de funciones de escalarización, la de sumas ponderadas y la función de escalarización aumentada de Tchebycheff.

Estas funciones de escalarización se definen como:

**Definición 4.6** *Dado una solución alcanzable  $z$  definimos la función de escalarización de sumas ponderadas como:*

$$s_{ws}(z, \lambda) = \sum_{i=1}^k \lambda_i (z_i - z_i^*) \quad (4.1)$$

donde  $\lambda = \langle \lambda_1, \lambda_2, \dots, \lambda_k \rangle$  es un vector de peso con elementos no negativos, y  $k$  es el número de objetivos del problema.

**Definición 4.7** *Dado una solución alcanzable  $z$ , definimos la función de escalarización aumentada de Tchebycheff como:*

$$s_{aTch}(z, \lambda) = \max\{\lambda_i (z_i - z_i^*)\} + \varepsilon s_{ws}(z, \lambda) \quad (4.2)$$

donde  $\lambda = \langle \lambda_1, \lambda_2, \dots, \lambda_k \rangle$  es un vector de peso con elementos no negativos y  $\varepsilon$  es un escalar positivo muy pequeño, y  $k$  es el número de objetivos del problema.

Así un MOP se reduce a la optimización de las funciones de escalarización; para este caso se busca la minimización de estas funciones para el TSP de tres objetivos.

El vector de peso define la dirección de optimización hacia la frontera no dominada. Definimos el conjunto de pesos como:

**Definición 4.8** *Conjunto de pesos*

$$\Lambda = \{\lambda \in \mathbb{R}^k \mid \lambda_i \geq 0 \quad \forall i \wedge \lambda_1 + \lambda_2 + \dots + \lambda_k = 1\}$$

---

<sup>2</sup>Del inglés *achievement scalarizing functions*

Se utiliza exclusivamente vectores de peso pertenecientes a  $\Lambda$ . Además, se mide la distancia entre dos vectores de peso  $\lambda^1$  y  $\lambda^2$  por su distancia Manhattan, esto es:

$$\tau(\lambda^1, \lambda^2) = \|\lambda^1 - \lambda^2\| = \sum_{i=1}^k |\lambda_i^1 - \lambda_i^2| \quad (4.3)$$

La distancia Manhattan también se usa para medir la distancia entre soluciones en el espacio objetivo:

$$\tau(z^1, z^2) = \|z^1 - z^2\| = \sum_{i=1}^k |z_i^1 - z_i^2| \quad (4.4)$$

Una conjunto  $\Lambda' \subset \Lambda$  es *maximal disperso* (con respecto a la distancia Manhattan) si  $\min d(\lambda^i, \lambda^j) \geq \delta$  para todo  $\lambda^i, \lambda^j \in \Lambda', i \neq j$ , y no existe ningún conjunto (mayor)  $\Lambda'' \subset \Lambda$  tal que  $|\Lambda''| > |\Lambda'|$  y  $\min d(\lambda^p, \lambda^q) \geq \delta$  para todo  $\lambda^p, \lambda^q \in \Lambda'', p \neq q$ . Un conjunto maximal disperso puede ser construido a partir de cualquier entero positivo  $\eta$ , incluyendo todos los vectores de peso diferentes cuyos componentes son múltiplos de  $\frac{1}{\eta}$  y cuya suma es igual a uno. Este conjunto tendrá  $\min d(\lambda^i, \lambda^j) = \frac{2}{\eta}$  para  $i \neq j$  [13]. El número de vectores de peso únicos es igual a:

$$\binom{k + \eta - 1}{\eta} = \frac{(k + \eta - 1)!}{\eta!(k - 1)!} \quad (4.5)$$

Los experimentos realizados por Hansen y Borges fueron hechos sobre las instancias kroA100, kroB100 y kroC100 de la TSPLIB, donde cada matriz de costos corresponde a un objetivo [13]. Debido a que las funciones de escalarización pueden ordenar un conjunto de soluciones no dominadas, para  $\eta = 100$  se generaron  $\binom{102}{100} = 5151$  vectores de peso maximales dispersos, y fueron encontrados 1000 óptimos locales para cada uno de estos vectores; estas optimizaciones fueron hechas usando las funciones de escalarización dadas en las definiciones 4.6 y 4.7. Para esta población de soluciones se midió la distancia media de cada individuo a los demás y se comparó contra la distancia al mejor óptimo local. Los experimentos conducidos en [13] indican la presencia de convexidad global de una manera persistente sobre todo el área de la frontera no-dominada para el MOTSP cuando se usa una estructura de vecindad 2-Opt.

Los trabajos realizados por Hansen y Borges conforman el comienzo de un nuevo enfoque para la construcción de heurísticas, para los casos mono-objetivos y multiobjetivos. La presencia de convexidad global nos permite, por un lado caracterizar la topología del espacio de soluciones, y por otro lado describir la aplicabilidad de diferentes metaheurísticas en diferentes tipos de topologías. Además, la utilización de funciones de escalarización es bastante general como para usarlo en cualquier otro problema de optimización combinatoria, aunque convierten al problema multi-objetivo en un conjunto de problemas mono-objetivo.

Al mismo tiempo, los resultados obtenidos en [13] están basados más bien en observaciones que en análisis teórico; y aunque demuestran la presencia de convexidad global en el MOTSP, no tienen aplicabilidad práctica. Esto es debido a que el uso de la solución ideal no es un buen punto de referencia cuando las unidades de medida de las funciones objetivos son distintas. Tomemos como ejemplo el caso de una red de computadores, donde  $f_1$  es el retardo y  $f_2$

es el porcentaje de paquetes perdidos, ambas funciones debe de minimizarse; entonces, el cálculo de la definición 4.6 puede dar resultados distintos para otras unidades de medida de estas mismas funciones objetivos.

En el presente trabajo, se busca mejorar estos aspectos, mediante la introducción de los conceptos de rango y categoría de una solución, presentando de esta manera un marco teórico más sólido y general al problema de la convexidad global en el MOTSP. Esta generalización permitirá el análisis de cualquier otro MOP sin importar el número de objetivos o la unidad de medida.

## Capítulo 5

# Análisis Topológico del Espacio de Soluciones

En los estudios realizados en [1] se utilizaba la longitud de un *tour* como métrica de calidad de una solución, esto es, si  $l(x) < l(x')$  entonces  $x$  es una mejor solución que  $x'$ . Las correlaciones halladas sugieren que los *tours* de menor longitud se encuentran en una posición más central en el espacio de soluciones.

En [13] la calidad de una solución se medía mediante las funciones de escalearización de sumas ponderadas (def. 4.6) y sumas ponderadas aumentada de Tchebycheff (def. 4.7), las cuales ordenan el espacio de solución según un vector de peso.

En este trabajo, se presentan por primera vez, los conceptos de categoría y rango de una solución, las cuales nos permitirán analizar con mayor rigor la presencia de convexidad global en el TSP Bi-Objetivo.

**Definición 5.1** Sea  $P \subset \mathbf{X}_f$  una población y  $x \in \mathbf{P}$  una solución, definimos la categoría de una solución  $x$  en una población  $P$  como:

$$cat(x, P) = |\{u \mid u \in P \wedge u \succ x\}|$$

Tenemos así que la categoría de  $x$  es el número de soluciones en  $P$  que la dominan. Por lo tanto, una solución del frente pareto tendrá siempre una categoría de 0, i.e. si  $u \in \mathbf{X}_{true}$  entonces  $cat(u, \mathbf{X}_f) = 0$ .

**Definición 5.2** Sea  $P \subset \mathbf{X}_f$  una población y  $x \in \mathbf{P}$  una solución, definimos la frontera no dominada de una población como:

$$\mathbf{NF}(P) = \{u \in P \mid cat(u, P) = 0\}$$

Si  $P = \mathbf{X}_f$  entonces  $\mathbf{NF}(P) = \mathbf{X}_{true}$ .

**Definición 5.3** Sea  $P \subset \mathbf{X}_f$  una población y  $x \in P$  una solución, se define rango de una solución en  $P$  (denotado por  $rng(x, P)$ ) conforme al siguiente algoritmo:

$$\begin{array}{ll} \text{si } x \in \mathbf{NF}(P) & \text{entonces } rng(x, P) = 0 \\ \text{si } x \notin \mathbf{NF}(P) & \text{entonces } rng(x, P) = 1 + rng(x, P') \end{array}$$

donde  $P' = P - \mathbf{NF}(P)$ .

De ahora en adelante, se omitirá el uso del parámetro  $P$  en las definiciones de rango y categoría, las cuales serán denotadas como  $rng(x)$  y  $cat(x)$  para alguna solución  $x$ . Solo se usará el parámetro de población cuando ambigüedades puedan surgir en la interpretación.

Siguiendo lo realizado en [1], se almacena en un conjunto  $P$ , recorridos diferentes de un problema de  $n$  ciudades, y cada *tour*  $x$  almacenado tiene:

1. Una categoría  $cat(x)$ .
2. Un rango  $rng(x)$ .
3. Una distancia media a las otras soluciones de  $P$  denotada por  $\delta(x, P)$ .
4. Una distancia a la frontera no dominada denotada por  $\delta(x, \mathbf{NF}(P))$ .

El presente trabajo se realizó en dos partes; la primera, consistió en el análisis de un problema instancia pequeña generado al azar, y la segunda, consistió en el análisis de varios problemas de instancias mayores obtenidos de la TSPLIB<sup>1</sup>.

## 5.1 Estudio del Universo de Soluciones

Se realizó un estudio del universo de soluciones para un problema de ocho ciudades generado al azar, con el objetivo de conocer las correlaciones entre las variables estudiadas en este contexto y así estimar el resultado que debería de esperarse para instancias mayores.

Para mantener el número de soluciones posibles en un valor manejable, se eligió un valor de  $n = 8$  ciudades. Este problema está dado en coordenadas cartesianas en el cuadro 5.1 para sus dos objetivos.

| ciudad | omiA8   | omiB8   |
|--------|---------|---------|
| 1      | (58,12) | (24,83) |
| 2      | (2,73)  | (2,4)   |
| 3      | (14,71) | (98,22) |
| 4      | (29,8)  | (78,45) |
| 5      | (54,50) | (8,50)  |
| 6      | (0,7)   | (48,1)  |
| 7      | (2,91)  | (9,3)   |
| 8      | (44,53) | (37,56) |

Cuadro 5.1: Coordenadas cartesianas para las matrices de costos del problema omiAB8. omiA8 corresponde al objetivo número 1, y, omiB8 corresponde al objetivo número 2.

Debido a la presencia de múltiples soluciones óptimas, definimos la distancia de un *tour* a la frontera de no dominados.

**Definición 5.4** Sea  $P \subset \mathbf{X}_f$  una población y  $x \in P$  una solución, definimos la distancia de una solución a la frontera de no dominados como:

$$\delta(x, \mathbf{NF}(P)) = \min\{d \mid \forall x_i^* \in \mathbf{NF}(P) \quad d = \delta(x, x_i^*)\} \quad (5.1)$$

<sup>1</sup> <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>

La distancia de una solución a la frontera de no dominados de una población consiste en la distancia mínima de la solución respecto a todas las soluciones con categoría 0.

La siguiente definición es necesaria para el estudio de la geometría del problema.

**Definición 5.5**  $P(e)$ . Subespacio conformado por las  $e$  mejores soluciones de  $P$ ; e.g.  $P(100)$  denota el conjunto de las 100 mejores soluciones, esto es, las 100 soluciones con menor valor de categoría.

Para la realización de los cálculos se usó un mecanismo de búsqueda exhaustiva de manera a encontrar todas las soluciones. La población obtenida con el algoritmo consiste en todo el universo de soluciones de un problema de  $n$  ciudades, es decir  $P = \mathbf{X}_f$  donde  $|\mathbf{X}_f| = |P| = \frac{(n-1)!}{2}$  y  $\mathbf{NF}(P) = \mathbf{X}_{true}$ .

Para cada solución  $x \in P$  se calcularon los siguientes valores de correlación:

1. La distancia a la frontera no dominada  $\delta(x, \mathbf{NF}(P))$  y la categoría del tour  $cat(x)$  denotada por  $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$ .
2. La distancia media a la población  $\delta(x, P)$  y la categoría del tour  $cat(x)$  denotada por  $\rho(cat(x), \delta(x, P))$ .
3. La distancia a la frontera no dominada  $\delta(x, \mathbf{NF}(P))$  y la distancia media a la población  $\delta(x, P)$  denotada por  $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ .
4. La distancia a la frontera no dominada  $\delta(x, \mathbf{NF}(P))$  y el rango del tour  $rng(x)$  denotado por  $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$ .
5. La distancia media a la población  $\delta(x, P)$  y el rango del tour  $rng(x)$  denotado por  $\rho(rng(x), \delta(x, P))$ .
6. El rango y la categoría de las soluciones denotado por  $\rho(rng(x), cat(x))$ .

Estos valores son análogos a los estudiados en [5]. En el cuadro 5.2 puede verse un resumen de los resultados obtenidos.

|   | $P$      | $P( P  - 1)$ | $P(\frac{ P }{2})$ | $P(\frac{ P }{4})$ |
|---|----------|--------------|--------------------|--------------------|
| $\rho(cat(x), \delta(x, P))$                    | 0        | 0.645277     | 0.868166           | 0.782142           |
| $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$       | 0.774954 | 0.774225     | 0.627201           | 0.524996           |
| $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ | 0        | 0.549804     | 0.686872           | 0.555586           |
| $\rho(rng(x), \delta(x, P))$                    | 0        | 0.644026     | 0.929499           | 0.853578           |
| $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$       | 0.810660 | 0.810222     | 0.678969           | 0.594514           |
| $\rho(rng(x), cat(x))$                          | 0.976814 | 0.976843     | 0.973864           | 0.959215           |

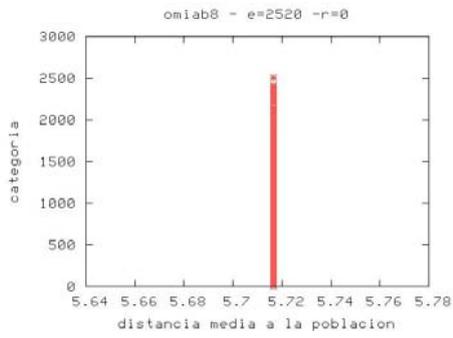
Cuadro 5.2: Resumen de las correlaciones obtenidas para el problema omi8AB.

Las figuras 5.1 al 5.6 muestran las correlaciones para cada par de variables estudiada.

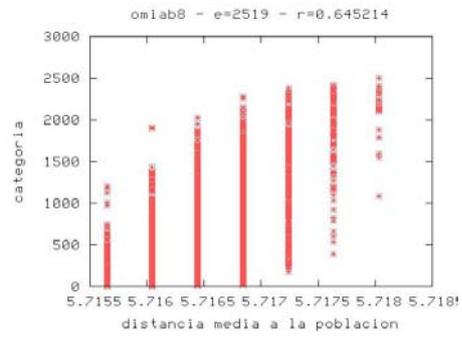
Los conceptos de rango y categoría resultan ser métricas muy similares, según lo observado en la figura 5.6, donde pueden observarse correlaciones mayores a 0.9 entre ambas variables.

Las figuras 5.1.c y 5.4.c, muestran correlaciones de 0.87 y 0.92 respectivamente. Esto sugiere una concentración de las mejores soluciones, donde estas se encuentran cercas unas de otras, cumpliendose así la hipótesis de centralidad de una estructura globalmente convexa. También puede verse de las figuras 5.1.d y 5.4.d que las mejores soluciones se encuentran en una pequeña porción del espacio de soluciones, obteniendo ahora la hipótesis de convexidad.

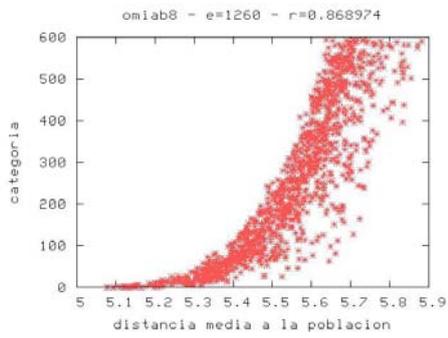
Como ambos supuestos se cumplen, es de esperarse que estas soluciones se encuentren concentrados en una pequeña parte del espacio de soluciones, y así, más cerca del frente pareto óptimo, lo cual es congruente con una estructura globalmente convexa. Esto puede observarse en las figuras 5.2.a y 5.5.a.



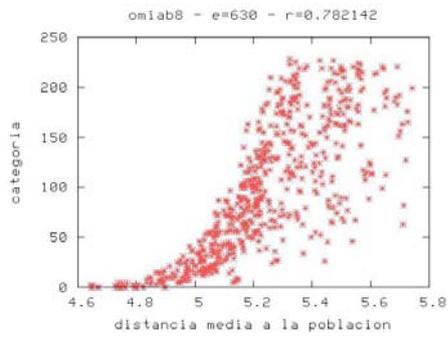
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$

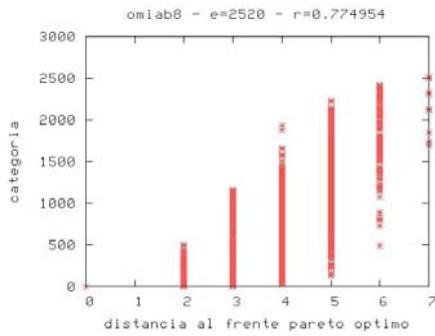


(c)  $P\left(\frac{|P|}{2} = 1260\right)$

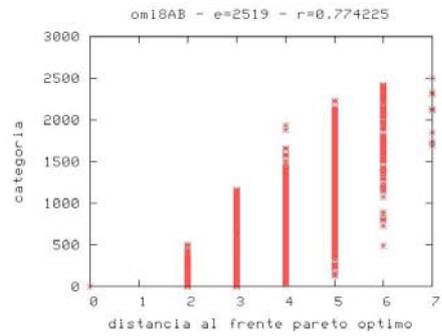


(d)  $P\left(\frac{|P|}{4} = 630\right)$

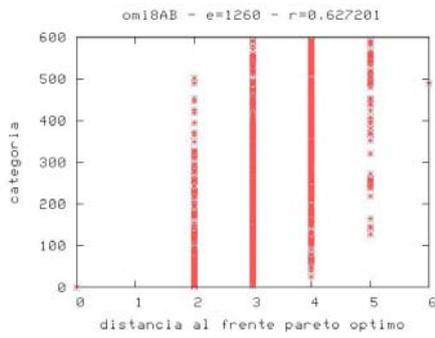
Figura 5.1: Correlación entre la distancia media a la población y la categoría, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .



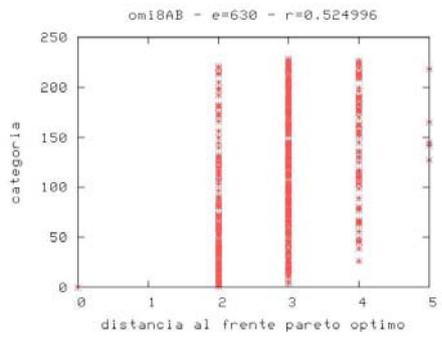
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$

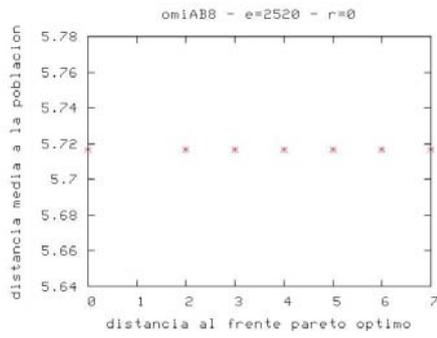


(c)  $P(\frac{|P|}{2} = 1260)$

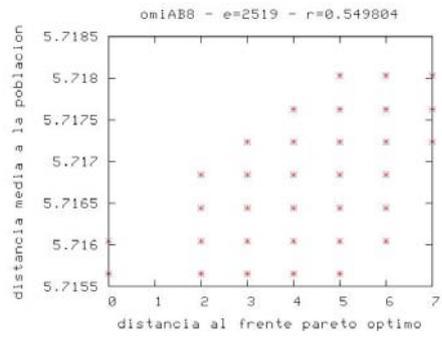


(d)  $P(\frac{|P|}{4} = 630)$

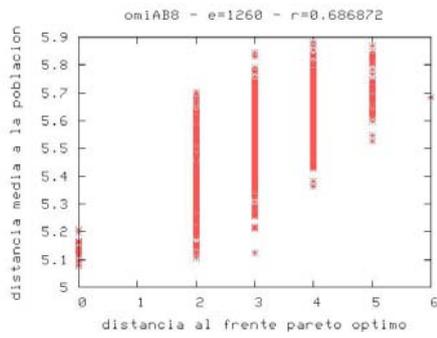
Figura 5.2: Correlación entre la distancia al frente pareto óptimo y la categoría, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .



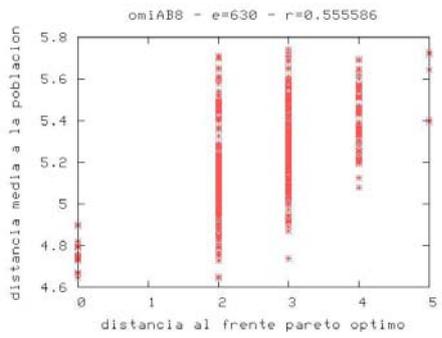
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$

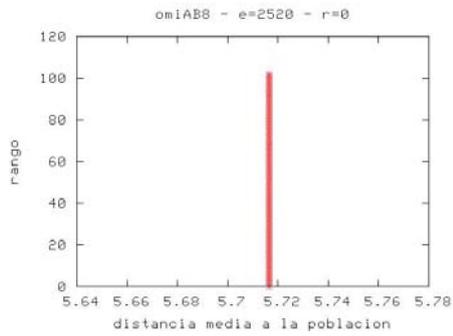


(c)  $P(\frac{|P|}{2} = 1260)$

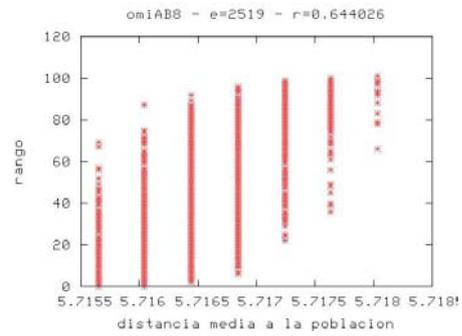


(d)  $P(\frac{|P|}{4} = 630)$

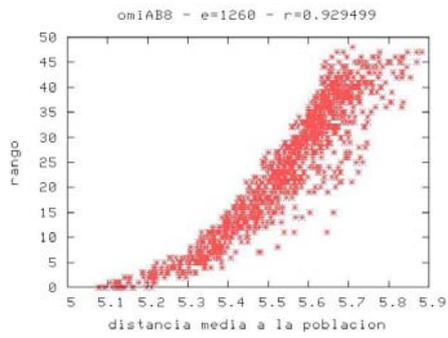
Figura 5.3: Correlación entre la distancia media a la población y la distancia al frente pareto óptimo, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .



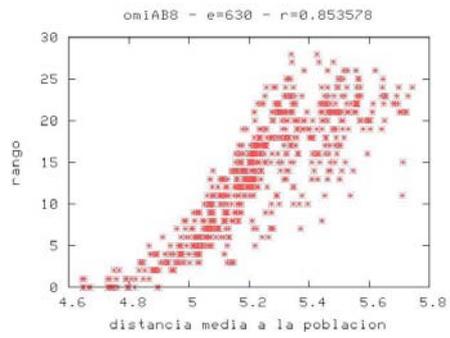
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$

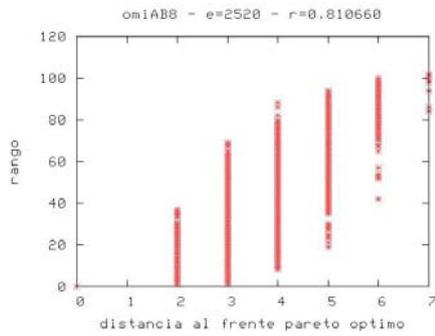


(c)  $P(\frac{|P|}{2} = 1260)$

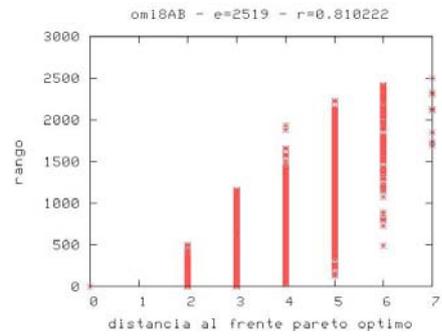


(d)  $P(\frac{|P|}{4} = 630)$

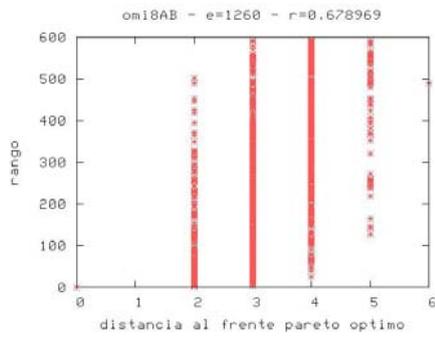
Figura 5.4: Correlación entre la distancia media a la población y rango, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .



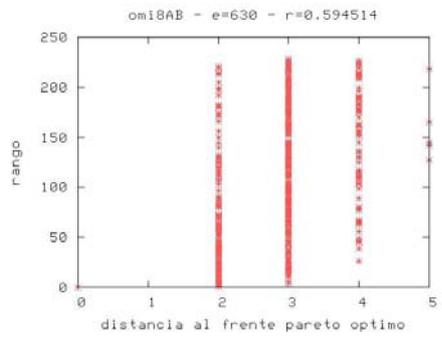
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$

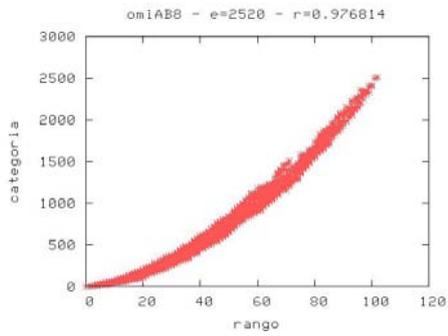


(c)  $P(\frac{|P|}{2} = 1260)$

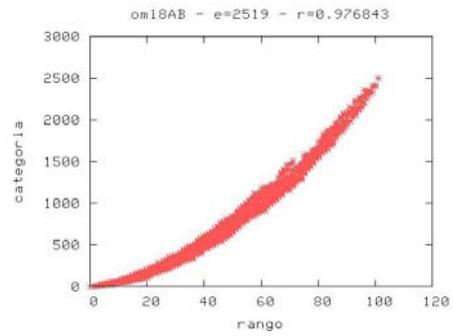


(d)  $P(\frac{|P|}{4} = 630)$

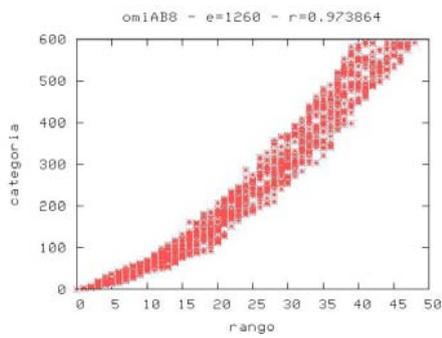
Figura 5.5: Correlación entre la distancia al frente pareto óptimo y rango, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .



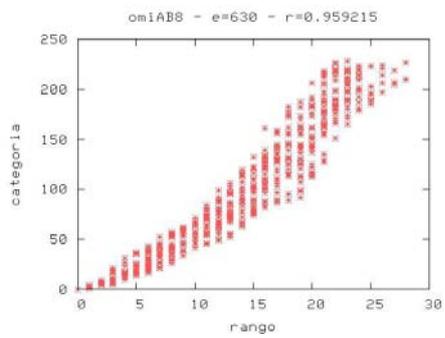
(a)  $P(|P| = 2520)$



(b)  $P(|P| - 1 = 2519)$



(c)  $P(\frac{|P|}{2} = 1260)$



(d)  $P(\frac{|P|}{4} = 630)$

Figura 5.6: Correlación entre rango y categoría, para el problema omiAB8 con un universo  $|\mathbf{X}_f| = 2520$ .

También, la correlaciones encontradas no presentan resultados muy directos como los hechos por Boese en [1], y presentado en la sección 3.3, donde todas las correlaciones eran mayores a 0.9 y presentaban una correlación casi lineal para las variables estudiadas. La razón de la presencia de estas correlaciones más bajas es debido a que no existe una sola solución óptima, sino que todo un frente pareto de soluciones. En la figura 5.7 puede observarse esta situación. Supongamos  $P = \{A, B, C, D, E\}$  y  $\mathbf{NF}(P) = \mathbf{X}_{true} = \{A, B, C\}$ . Tenemos que

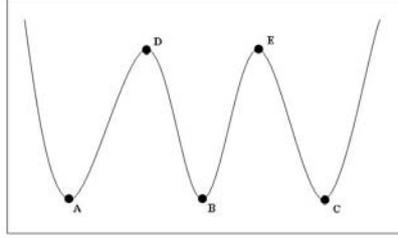


Figura 5.7: Correlación entre soluciones del frente pareto y las demás soluciones

$cat(A) = cat(B) = cat(C) = 0$  y  $rng(A) = rng(B) = rng(C) = 0$ ; también  $cat(D) = cat(E) = 3$  y  $rng(D) = rng(E) = 1$ , para este ejemplo. Se puede observar de este simple gráfico que  $\delta(D, P) < \delta(A, P)$ , esto es, D es más central en la población que A, a pesar de que D sea un peor solución. Esto mismo ocurre con la solución C.

La presencia de soluciones como las presentadas en la figura 5.7 es la razón de las bajas correlaciones encontradas. Pero esto no debe de ser desalentador, puesto que aun así pueden observarse una concentración de las mejores soluciones en una zona central del espacio, lo cual es la idea detras de una estructura globalmente convexa.

El mismo estudio exhaustivo fue realizado usando otros problemas de 7, 9 y 10 ciudades, y los resultados fueron muy similares, haciendo innecesaria la repetición de los mismos en este trabajo. Un resumen de los resultados de estas instancias puede verse en los cuadros 5.3 al 5.5.

|   | $P$      | $P( P  - 1)$ | $P(\frac{ P }{2})$ | $P(\frac{ P }{4})$ |
|---|----------|--------------|--------------------|--------------------|
| $\rho(cat(x), \delta(x, P))$                    | 0        | 0.602793     | 0.761135           | 0.659006           |
| $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$       | 0.767441 | 0.763554     | 0.548112           | 0.585267           |
| $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ | 0        | 0.582333     | 0.633253           | 0.706702           |
| $\rho(rng(x), \delta(x, P))$                    | 0        | 0.628991     | 0.752562           | 0.661246           |
| $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$       | 0.786593 | 0.783295     | 0.567543           | 0.633579           |
| $\rho(rng(x), cat(x))$                          | 0.960048 | 0.960550     | 0.937163           | 0.933828           |

Cuadro 5.3: Resumen de las correlaciones obtenidas para el problema litAB7.

En la siguiente sección pasamos a analizar problemas de mayor número de ciudades, utilizando las mismas métricas, pero con subconjuntos del universo de soluciones.

|   | $P$      | $P( P  - 1)$ | $P(\frac{ P }{2})$ | $P(\frac{ P }{4})$ |
|---|----------|--------------|--------------------|--------------------|
| $\rho(cat(x), \delta(x, P))$                    | 0        | 0.572255     | 0.825523           | 0.697679           |
| $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$       | 0.641862 | 0.641802     | 0.435708           | 0.409532           |
| $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ | 0        | 0.329447     | 0.478431           | 0.371982           |
| $\rho(rng(x), \delta(x, P))$                    | 0        | 0.564819     | 0.894474           | 0.759983           |
| $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$       | 0.672059 | 0.671998     | 0.483163           | 0.492397           |
| $\rho(rng(x), cat(x))$                          | 0.971607 | 0.971621     | 0.974710           | 0.971266           |

Cuadro 5.4: Resumen de las correlaciones obtenidas para el problema encAB9.

|   | $P$      | $P( P  - 1)$ | $P(\frac{ P }{2})$ | $P(\frac{ P }{4})$ |
|---|----------|--------------|--------------------|--------------------|
| $\rho(cat(x), \delta(x, P))$                    | 0        | 0.508309     | 0.830280           | 0.734061           |
| $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$       | 0.712957 | 0.712944     | 0.536178           | 0.488927           |
| $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ | 0        | 0.392366     | 0.577878           | 0.442818           |
| $\rho(rng(x), \delta(x, P))$                    | 0        | 0.493534     | 0.899061           | 0.803484           |
| $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$       | 0.750817 | 0.750809     | 0.583105           | 0.533937           |
| $\rho(rng(x), cat(x))$                          | 0.970935 | 0.970937     | 0.976167           | 0.975501           |

Cuadro 5.5: Resumen de las correlaciones obtenidas para el problema asuAB10.

## 5.2 Estudio de un Subespacio de Soluciones

Los resultados obtenidos en la sección anterior demuestran, por lo menos de forma empírica, la presencia de una estructura globalmente convexa en el TSP Bi-Objetivo. A pesar de que estos resultados son positivos, es necesario el estudio en instancias mayores, pues es en este tipo de problemas donde las metaheurísticas hacen la diferencia, respecto a cualquier otro tipo de algoritmo.

Las pruebas experimentales se basaron en las instancias kroAB100, kroCD100, kroAD100, kroBC100 y kroAB150 de las TSPLIB. De cada uno de estos problemas bi-objetivos fueron obtenidos  $n^2$  muestras aleatorias, teniendo así poblaciones con cardinalidad  $|P| = n^2$ .

Cada población fue optimizada utilizando el algoritmo de búsqueda local 2-Opt, para así obtener poblaciones optimizadas conteniendo exclusivamente óptimos locales. Esta técnica de búsqueda fue escogida debido a que presenta un estructura de vecindad simple, y nos permite estudiar la distribución de los óptimos locales en el espacio de soluciones. La figura 5.8 muestra que la correlación entre óptimos locales es más directa que entre soluciones elegidas al azar, debido a lo explicado en la sección 5.1 y la figura 5.7, lo cual hace necesario el uso de un optimizador local para nuestro estudio.

La metodología utilizada para el análisis sigue siendo el mismo, solo que ahora se trabaja con un subconjunto del universo de soluciones. El cuadro 5.6 presenta un resumen de las correlaciones obtenidas.

Las figuras 5.9 al 5.13 muestran las correlaciones para cada par de variables estudiada.

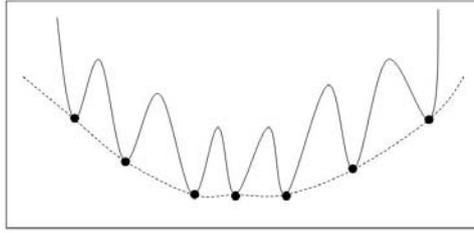


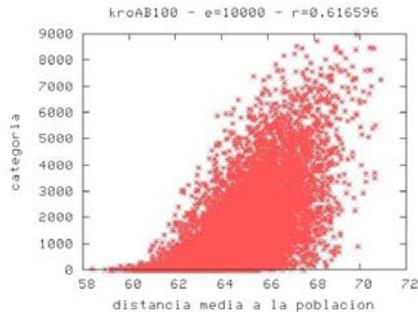
Figura 5.8: Correlación entre óptimos locales

|   | kroAB100 | kroCD100 | kroAD100 | kroBC100 | kroAB150 |
|---|----------|----------|----------|----------|----------|
| $\rho(cat(x), \delta(x, P))$                    | 0.616596 | 0.596558 | 0.603098 | 0.589122 | 0.571318 |
| $\rho(cat(x), \delta(x, \mathbf{NF}(P)))$       | 0.396173 | 0.401700 | 0.391263 | 0.396632 | 0.384648 |
| $\rho(\delta(x, P), \delta(x, \mathbf{NF}(P)))$ | 0.613169 | 0.626019 | 0.611728 | 0.653360 | 0.633580 |
| $\rho(rng(x), \delta(x, P))$                    | 0.639641 | 0.628951 | 0.645090 | 0.617764 | 0.606404 |
| $\rho(rng(x), \delta(x, \mathbf{NF}(P)))$       | 0.432760 | 0.441982 | 0.449073 | 0.437721 | 0.423266 |
| $\rho(rng(x), cat(x))$                          | 0.945914 | 0.941865 | 0.937078 | 0.941890 | 0.938438 |

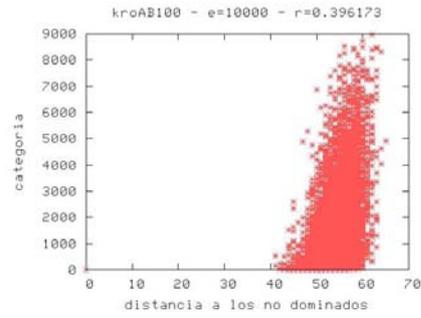
Cuadro 5.6: Correlaciones obtenidas para los problemas kroAB100, kroCD100, kroAD100, kroBC100, kroAB150

La correlación entre categoría y rango sigue siendo superior a 0.9, lo cual confirma que ambos conceptos son muy similares.

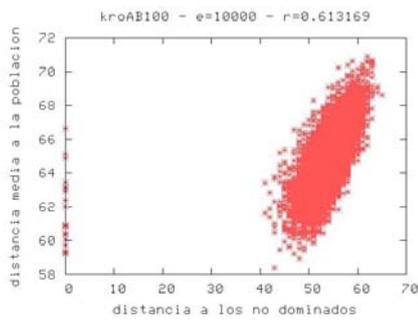
Las figuras 5.9.a y 5.9.d muestran una concentración de las mejores soluciones hacia la zona central de la población, esto es, a medida que mejoran su calidad, su distancia media disminuye. A pesar de la baja correlación presente en las figuras anteriores (alrededor de 0.6), en la figura 5.9.c puede verse que mientras más central este una solución, más cerca se encontrará de la frontera de no dominados. Estos mismos resultados son válidos para las figuras 5.10 al 5.13, y satisfacen las hipótesis de una estructura globalmente convexa.



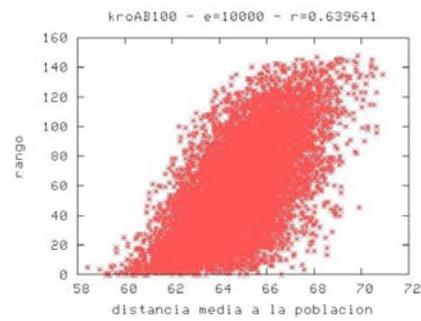
(a)  $|P| = 10000$



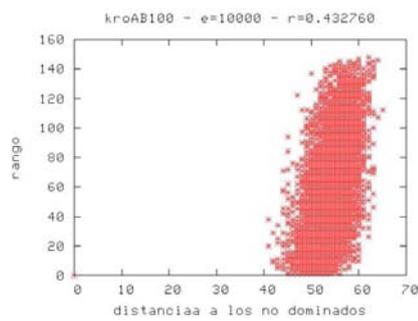
(b)  $|P| = 10000$



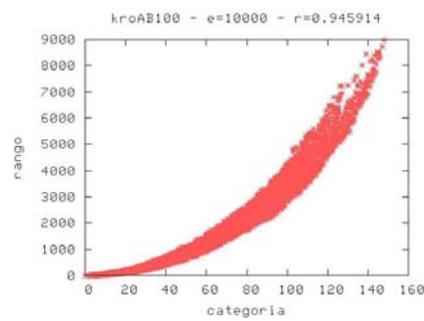
(c)  $|P| = 10000$



(d)  $|P| = 10000$

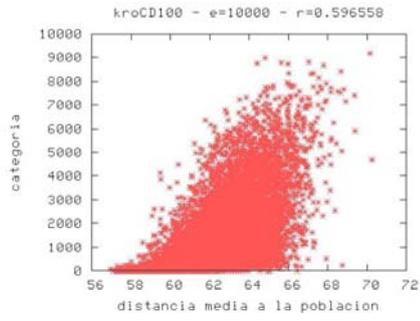


(e)  $|P| = 10000$

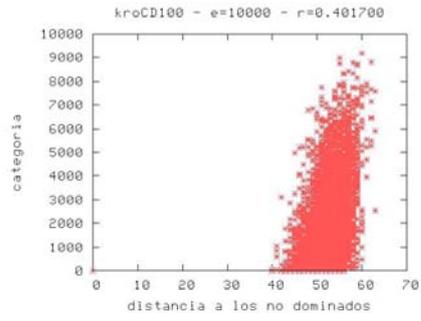


(f)  $|P| = 10000$

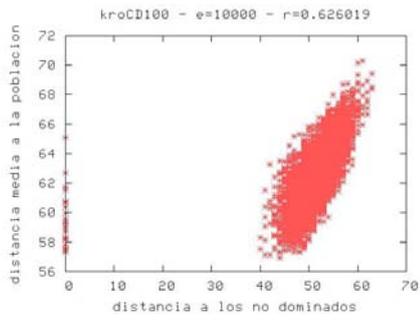
Figura 5.9: Población de 10000 soluciones optimizadas de la instancia kroAB100



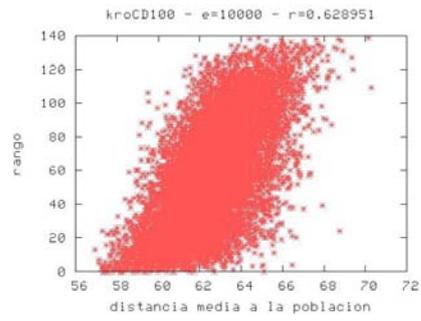
(a)  $|P| = 10000$



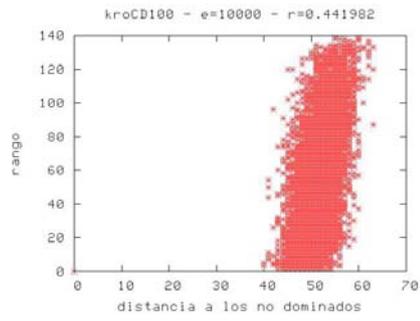
(b)  $|P| = 10000$



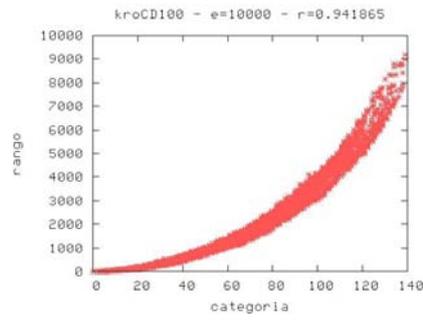
(c)  $|P| = 10000$



(d)  $|P| = 10000$

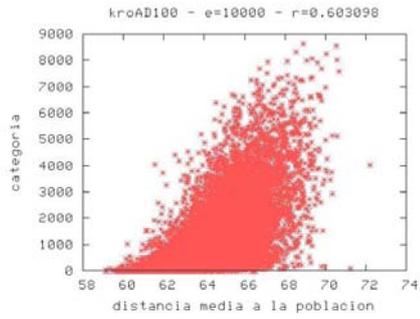


(e)  $|P| = 10000$

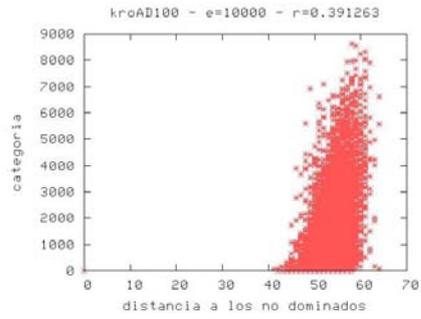


(f)  $|P| = 10000$

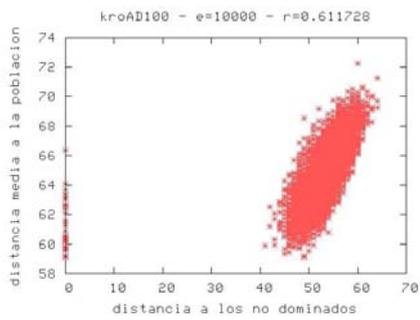
Figura 5.10: Población de 10000 soluciones optimizadas de la instancia kroCD100



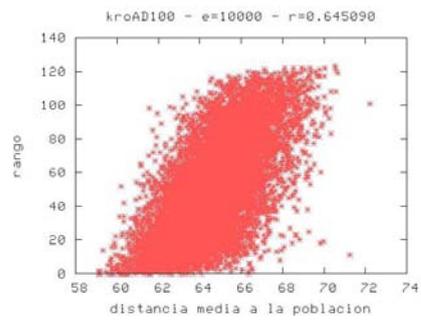
(a)  $|P| = 10000$



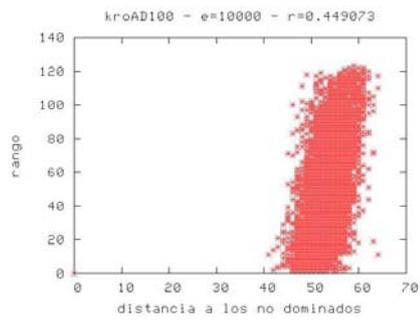
(b)  $|P| = 10000$



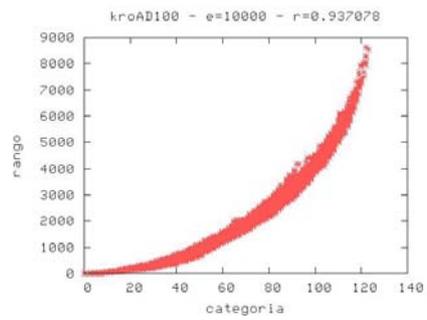
(c)  $|P| = 10000$



(d)  $|P| = 10000$

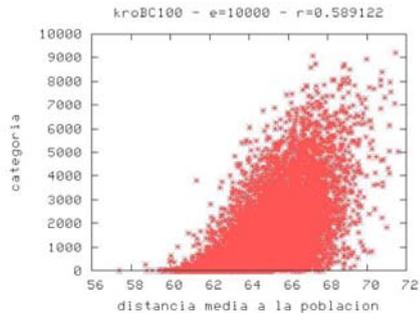


(e)  $|P| = 10000$

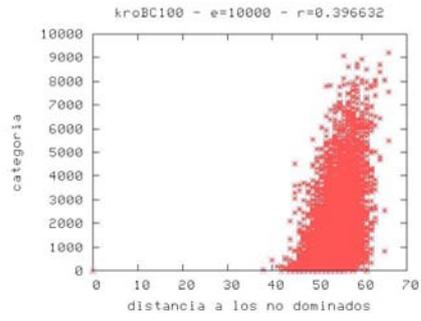


(f)  $|P| = 10000$

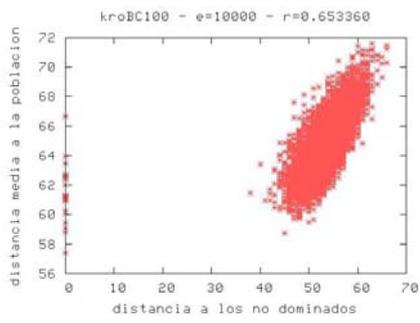
Figura 5.11: Población de 10000 soluciones optimizadas de la instancia kroAD100



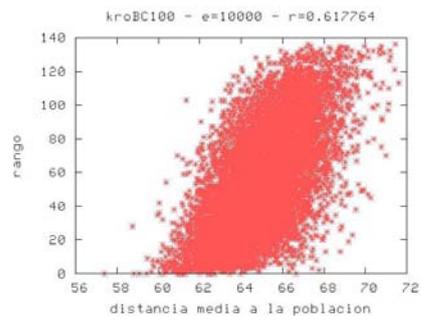
(a)  $|P| = 10000$



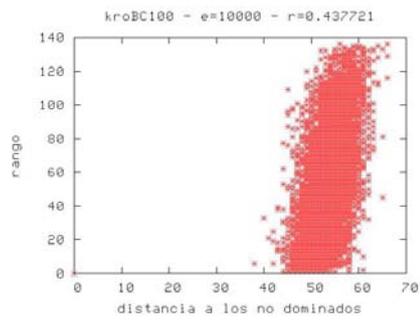
(b)  $|P| = 10000$



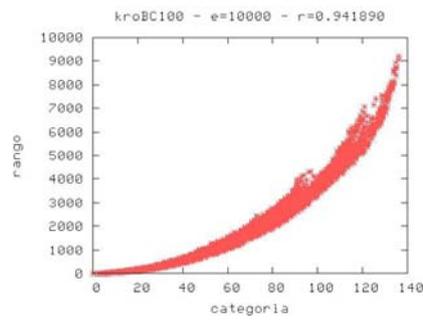
(c)  $|P| = 10000$



(d)  $|P| = 10000$

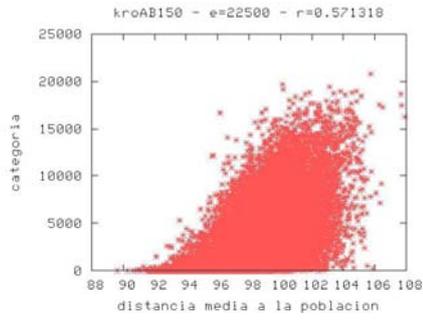


(e)  $|P| = 10000$

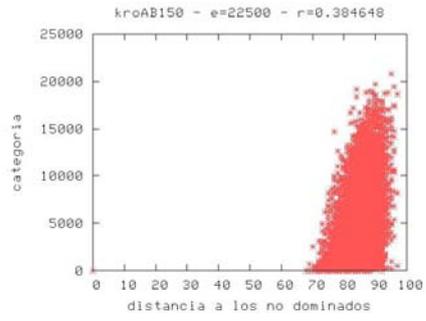


(f)  $|P| = 10000$

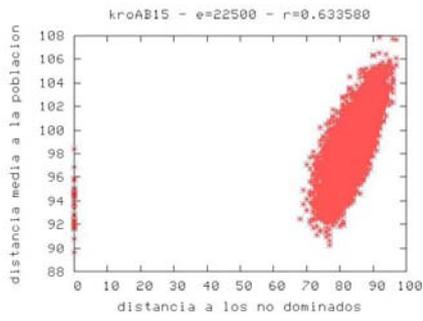
Figura 5.12: Población de 10000 soluciones optimizadas de la instancia kroBC100



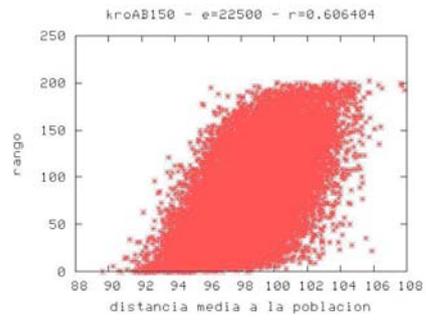
(a)  $|P| = 22500$



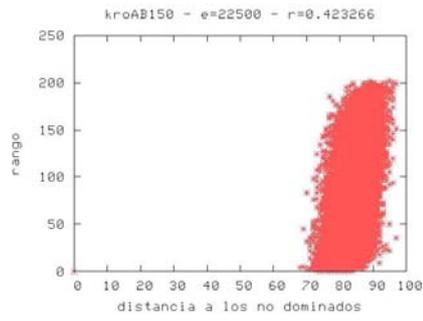
(b)  $|P| = 22500$



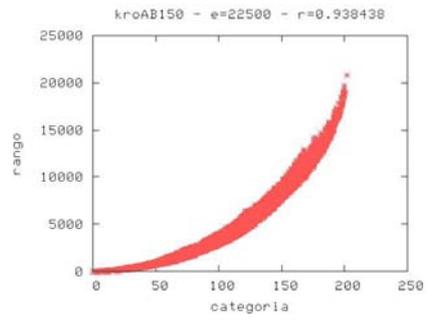
(c)  $|P| = 22500$



(d)  $|P| = 22500$



(e)  $|P| = 22500$



(f)  $|P| = 22500$

Figura 5.13: Población de 22500 soluciones optimizadas de la instancia kroAB150

## Capítulo 6

# Conclusiones y Trabajos Futuros

Los conceptos de rango y categoría mostraron ser métricas de calidad muy eficientes para el MOTSP, cuya generalización a cualquier instancia de un MOTSP puede ser fácilmente realizada. Inclusive pueden usarse en cualquier MOP sin tener que cambiar sus definiciones. Además de que mostraron ser métricas muy parecidas, se observaron mejores resultados para rango, con correlaciones ligeramente más altas.

Las métricas de rango y categoría demostraron de una forma rigurosa, la correlación existente entre estos, la distancia media a la población y la frontera de no dominados, confirmando de esta manera la característica topológica de una convexidad global. Estas métricas podrán usarse para estudiar la presencia de convexidad global en otros MOPs, donde metaheurísticas como las colonias de hormigas y algoritmos evolutivos, resultaron ser muy eficientes, y por supuesto, para la creación de nuevas metaheurísticas que puedan explotar esta estructura del espacio de soluciones.

Aunque aún sea muy difícil encontrar una solución óptima, o, el frente pareto de un TSP bi-objetivo, el conocer la existencia de convexidad global, permitirá acotar la región de búsqueda a un conjunto más pequeño de soluciones, y a partir de ahí se podrá otros algoritmos adecuados, que logren mejores aproximaciones a con las soluciones Pareto.

Hay mucho por hacer en el estudio de funciones globalmente convexas, como la creación de metaheurísticas basadas totalmente en explotar esta característica, y la identificación de problemas donde algoritmos como, las colonias de hormigas y algoritmos evolutivos, han demostrado ser muy eficientes. Además, se puede considerar el estudio de una definición formal de convexidad global en la cual se podrían analizar grados de convexidad.

# Bibliografía

- [1] Boese, K.D.: Cost versus Distance in the Traveling Salesman Problem. Reporte Técnico 950018, University of California, Los Angeles, Computer Science Department, (1995).
- [2] Boese, K.D., A.B. y Muddu, S.: A new Adaptive Multi-Start Technique for Combinatorial Global Optimization. Operations Research Center. Vol 16, pp. 101-113, (1994).
- [3] Stutzle, T. y Hoos, H.H.: Max-Min Ant System. Future Generation Computer Systems. Vol 16, pp. 889-914, (2000).
- [4] Paquete, L., Chiarandini, M. y Stutzle, T.: A Study of Local Optima in the Biobjective Travelling Salesman Problem. Multiobjective Metaheuristics Workshop (MOMH2002). Paris, France. Noviembre 2002.
- [5] Barán, B., Gardel, P. y Gómez, O.: Estudio del Espacio de Soluciones del Problema del Cajero Viajante. En Mauricio Solar, David Fernández y Ernesto Cuadros-Vargas, editores, 30ma Conferencia Latinoamericana de Informática (CLEI2004), páginas 745-756. Sociedad Peruana de Computación, Arequipa, Perú, septiembre 2004.
- [6] B. Barán y D. Pinto. Solving Multiobjective Multicast Routing Problem with a New Ant Colony Optimization Approach. 3rd IFIP/ACM Latin America Networking Conference. pag 125. Santiago de Cali, Colombia. Octubre 2005.
- [7] T. C. Hu, Victor Klee y David Larman. Optimization of globally convex functions. *SIAM Journal on Control and Optimization*, 27(5):1026-1047, septiembre 1989.
- [8] Gerhard Reinelt. The Traveling Salesman: Computational Solutions for TSP Applications. *Lecture Notes in Computer Science* , 840: VIII + 223, 1994.
- [9] Basil Vandegriend. Finding Hamiltonian Cycles: Algorithms, Graphs and Performance. Thesis submitted to the Faculty of Graduate and Research for the degree of Master of Science. Department of Computing Science. University of Alberta. Edmonton, Alberta. Spring 1998.
- [10] O. Gómez. Razones del Éxito de Optimización por Colonia de Hormigas en el Problema del Cajero Viajante. Tesis para Maestría en Ingeniería de Sistemas. Centro Nacional de Computación. Universidad Nacional de Asunción. San Lorenzo, Paraguay. Septiembre 2004.

- [11] S. Kirkpatrick and G. Toulouse. Configuration Space Analysis of Traveling Salesman Problems. *Journal de Physique*, 46:1277-1292, 1985.
- [12] D. Mandrioli y C. Ghezzi. Theoretical Foundations of Computer Science. Krieger Publishing Company. Malabar, Florida. Reprint Edition 1993.
- [13] P.C. Borges and P.H. Hansen. A Study of Global Convexity for a Multiple Objective Traveling Salesman Problem. In C.C. Ribeiro and P. Hansen, editors, *Essays and Surveys in Metaheuristics*, pages 129- 150. Kluwer, 2000.
- [14] D. P. Bertsekas. Nonlinear Programming. Athena Scientific. ISBN: 1-886529-14-0. Belmont, Massachusetts, 1995.
- [15] H. S. Stone. High-Performance Computer Architecture. Addison-Wesley Publishing Company. ISBN: 0-201-16802-2. 1987.
- [16] C. Blum, A. Roli. Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. ACM Computing Surveys (CSUR), Volume 35, Issue 3. ACM Press. September 2003.
- [17] S. Boyd y L. Vandenberghe. Convex Optimization. Cambridge University Press. ISBN: 0-521-83378-7. Cambridge, 2004.