

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO
MAESTRIA EN OPTIMACION FINANCIERA**

TESIS:

**“COMPARATIVO DE METODOLOGÍAS NO
TRADICIONALES PARA LA SELECCIÓN DE UN
PORTAFOLIO DE INVERSIÓN”**

ALUMNO: ARTURO TREVILLA SERRANO

ASESOR: DR. GUILLERMO SIERRA

Mayo 2011

INDICE DE CONTENIDO

Introducción	7
Capítulo I. Sociedades de Inversión	13
1.1 Definición de las sociedades de inversión.....	13
1.2 Importancia de las sociedades de inversión.....	14
1.3 Clasificación de las sociedades de inversión.....	15
1.4 Artículos de la Ley de Sociedades de Inversión que dan legalidad a las mismas.....	16
Capítulo II. Teoría moderna del portafolio	18
2.1 Introducción.....	18
2.2 Rendimiento Esperado y riesgo.....	19
2.3 Estructura de preferencias bajo incertidumbre.....	21
2.4 Desarrollo de la teoría del portafolio.....	24
2.5 Modelo de Markowitz como Problema de Programación Multiobjetivo.....	26
Capítulo III. Métodos Heurísticos de optimación	29
3.1 Método del Gradiente y del Hill Climbing.....	29
3.2 Búsqueda Iterativa.....	30
3.3 Algoritmo Metrópolis.....	30
3.4 Recocido Simulado.....	31
3.5 Búsqueda Tabú (BT).....	32
3.6 Búsqueda ciega o sin información.....	34
Capítulo IV. Algoritmos Genéticos	37
4.1 Introducción.....	37
4.2 El Algoritmo Genético Simple.....	40
4.3 Extensiones y Modificaciones del Algoritmo Genético Simple.....	45
4.3.1 Población.....	48
4.3.1 Función Objetivo.....	49
4.3.3 Selección.....	51
4.3.4 Cruce.....	53
4.3.5 Mutación.....	56
4.3.6 Reducción.....	57
4.4 Evaluación de Algoritmos Genéticos.....	60

4.5 ¿Por qué funcionan los algoritmos.....	61
4.6 Puntos a considerar en un Algoritmo Genético.....	66
Capítulo V. Análisis Multicriterio.....	71
5.1 Procesos y clasificación de modelos en la toma de decisiones.....	71
5.2 Preferencias y relaciones de orden y función de utilidad.....	79
5.3 Óptimo de Pareto.....	80
5.4 Enfoques no económicos.....	85
5.5 Método de restricciones y de los coeficientes de ponderación.....	105
5.6 Programación Compromiso.....	107
Capítulo VI. Caso Práctico.....	111
6.1 Selección de portafolios de inversión.....	111
6.2 Metodología utilizada para el Algoritmo Genético.....	111
6.3 Metodología utilizada para la Programación Compromiso.....	114
6.4 Análisis de resultados para el Algoritmo Genético.....	117
6.5 Análisis de resultados para la Programación Compromiso.....	120
Capítulo VII. Conclusiones.....	122
7.1 Relativas a los Métodos Heurísticos de Optimación.....	122
7.2 Relativas al Caso Practico.....	122
7.2.1 Algoritmos Genético.....	123
7.2.2 Programación Compromiso.....	123
7.3 Sugerencias a futuras investigaciones.....	124
Bibliografía.....	145
Anexos.....	148

INDICE DE TABLAS

Tabla 1. Diferencia de los algoritmos genéticos y los otros métodos de optimización tradicionales.....	35
Tabla 2. Matriz de decisión en el Análisis Multicriterio.....	92
Tabla 3. Escala fundamental para comparaciones a pares.....	106
Tabla 4. Portafolios seleccionados para el estudio de la tesis.....	134
Tabla 5. Análisis comparativo de la programación compromiso y los algoritmos genéticos.....	140
Tabla 6. Análisis de los resultados al aplicar la programación compromiso.....	141
Tabla 7. Análisis de los resultados al aplicar el algoritmo genético.....	142

INDICE DE FIGURAS

Figura 1. Grafica de una función multimodal.....	26
Figura 2. Bucle básico de un algoritmo genético (AG).....	37
Figura 3. Operador de cruce basado en un punto.....	39
Figura 4. Descomposición generalizada de las tareas de los AE.....	40
Figura 5. Operador de mutación.....	40
Figura 6. Adaptación media y mejor adaptación en un Algoritmo genético.....	41
Figura 7. Gráfica de la repartición más uniforme de la probabilidad de selección.....	50
Figura 8. Método de selección de padres denominado muestreo universal estocástico. El individuo I_1^t se escoge 2 veces, mientras que I_3^t e I_4^t son elegidos una única vez.....	51
Figura 9. Operador de cruce basado en dos puntos, los cromosomas (individuos) pueden contemplarse como un circuito en el cual se efectúa la selección aleatoria de dos puntos	53
Figura 10. Operador de cruce basado en dos puntos.....	54
Figura 11. Operador de cruce uniforme.....	54
Figura 12. "Máscaras de cruce" para los operadores de cruce basados en 1 punto y en 2 puntos.....	56
Figura 13. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en estrella.....	60
Figura 14. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en red....	60
Figura 15. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en anillo. El modelo de islas ha sido utilizado por varios autores (Whitley y Starkweather, 1990; Gorges-Schleuter, 1989; Tanese, 1987).....	61
Figura 16. Grafica que muestra la matriz A con sus arcos y nodos.....	115

INTRODUCCION

Esta investigación brindará la oportunidad de comparar los métodos de programación compromiso y algoritmos genéticos para la selección de portafolios de inversión, además mostrará la aplicación de los mismos en su diseño y conceptualización práctica. Por lo anterior, se está en condiciones de afirmar que aportará información útil a las instancias de crédito, instituciones y organizaciones vinculadas al sector financiero.

El presente estudio utilizará los precios en pesos de las sociedades de inversión de instrumentos de deuda para personas físicas en el período comprendido entre 31 de agosto de 2007 al 30 de septiembre de 2009¹ obtenidos en la Bolsa Mexicana de Valores.

Al pensar en la investigación surgió la pregunta acerca de la existencia en diferencias cuantitativas y cualitativas, al aplicar como en sus resultados en los métodos de programación compromiso y algoritmos genéticos para la selección óptima de un portafolio de inversión.

El presente estudia busca describir las metodologías y diferencias propuestas en los métodos de Programación Compromiso y Algoritmos Genéticos, cuyo fin es la selección eficaz de una cartera de inversiones enfocado a instrumentos de deuda para personas físicas. Asimismo, para alcanzar este objetivo se enfoca en los siguientes apartados:

- Aplicar el modelo de algoritmo genético simple y el de programación compromiso a un conjunto de precios de sociedades de inversión de deuda obtenidos de la Bolsa Mexicana de Valores para optimizar un portafolio.
- Evaluar cual de los métodos resulta más conveniente de acuerdo a las necesidades del usuario y sus expectativas para la selección de un portafolio de sociedades de inversión
- Analizar los factores que se evalúan en ambos métodos que intervienen en la selección de un portafolio de inversión

Los objetivos anteriores se especifican como sigue:

¹ Se consideraron este rango de datos por ser los que existían mas actuales para los fondos de inversión

- Evaluar la aproximación de ambos procedimientos para la solución óptima de los recursos.
- Desarrollar los métodos e implementarlos en un caso práctico orientado a instrumentos de deuda para personas físicas.
- Establecer el tiempo de ejecución de cada método en términos de iteraciones de los mismos.
- Describir los fundamentos teóricos o científicos que sustentan a cada procedimiento
- Determinar las áreas de oportunidad para la aplicación de ambos métodos.

La hipótesis que se tiene en esta investigación pretende resaltar las semejanzas y diferencias en la aplicación entre los métodos de programación compromiso y algoritmos genéticos, en virtud de que ambos se fundamentan en criterios y objetivos distintos

Estado del Arte de la Investigación

Algoritmos Genéticos

Uso de agentes artificiales para la Bolsa de Valores

Yang [66] siguiendo el trabajo de Chan [20] utilizó una Red Neuronal para un problema de doble subasta en el que desarrolló una estrategia que los agentes o jugadores deben seguir para maximizar las ganancias en un mercado artificial. Este mercado artificial lo modeló con una red neuronal la cual reducía la línea existente entre los mercados simulados y los reales.

Shu-Heng Chen [21], propuso un software llamado AIE-ASM, el cual realiza simulaciones de acciones en bolsa basadas en algunos modelos de precios activos. Una característica distinguible de este software es que el comportamiento de los agentes es modelado utilizando la programación genética en vez de los Algoritmos Genéticos.

Jasmina Arifovic [4], fue la primer persona en obtener un doctorado con una tesis relacionada con aplicaciones de los Algoritmos Genéticos en macroeconomía.

Arifovic publicó también varios artículos en el tema de intercambio de divisas [2, 3]. En estos estudios, utilizó los Algoritmos Genéticos para entrenar agentes simulados en varios modelos con dos diferentes monedas.

Francesco Luna [47] se concentra en el problema del rol de las instituciones en un proceso de aprendizaje en agentes económicos. En este trabajo se utiliza una red neuronal para modelar el proceso de aprendizaje, y el algoritmo genético se usa para desarrollar y obtener los pesos en esta red de la mejor forma posible para optimizar el aprendizaje.

David Fogel [28] realizó un estudio sobre la representación de los agentes inteligentes que se suelen utilizar en problemas económicos. Indica que existen clases de modelos que se usan para representar a los agentes inteligentes, y demostró que estos modelos afectan considerablemente el desempeño del mismo.

Ingeniería Financiera

La primera aplicación de los algoritmos evolutivos en econometría fue desarrollada por John Koza [45]. En ella se dice que un problema importante en Economía es encontrar

una relación matemática entre las variables observadas empíricamente para medirlas en un sistema.

George Szpiro [63], propone el uso de Algoritmos Genéticos para buscar dependencias entre conjuntos de datos para problemas en Finanzas y Economía.

En [54], Robert Pereira muestra como es que la computación evolutiva resulta relevante en el estudio de las estrategias comerciales. Comienza con la revisión de dos clases básicas de reglas comerciales, llamadas reglas de moving-average (MA) y reglas de order-statistics (OS).

Raymond Tsang y Paul Lajbcygier [64] ofrecen una alternativa al uso de los Algoritmos Genéticos estándar. Ellos proponen dividir el algoritmo genético en un modelo de islas, en el que la población de soluciones se divide en un número fijo de sub-poblaciones, cada una evolucionando de forma independiente, pero todas resolviendo el mismo problema.

Existe un artículo de Nemara Chidambaran [22] en el que éste lidia con el problema de la elección de precios con Programación Genética (PG). Primero se examina cómo es que la programación genética mejora el modelo de "Black-Scholes" bajo ciertas condiciones, y después utiliza este modelo para resolver un problema con datos del mundo real.

Christian Keber [44] muestra que la programación genética puede usarse para derivar aproximaciones muy precisas y determinar la volatilidad en este tipo de problemas. Keber genera 1.000 soluciones aleatorias y con base en éstas deriva una expresión analítica utilizando programación genética.

Otros desarrollos actuales

Streichert, Ulmer y Zell en 2004 [62] realizó un estudio comparativo de distintos tipos de cruce en un algoritmo evolutivo con representación real para resolver problemas de portafolios de inversión basados en el modelo de media-varianza de Markowitz. Asunción Mochón y co-autores en 2005 [53] hicieron un análisis del uso de la computación evolutiva en problemas de subastas. Se diseñó una simulación para examinar el comportamiento de las subastas, buscando encontrar la estrategia óptima de los apostadores para problemas dinámicos de subastas.

García-Almanza y Tsang en 2006 [38] presenta un método para simplificar los árboles de decisión con ayuda de la programación genética. En él se identifican y se eliminan

las reglas que causan errores en la clasificación y, como consecuencia, se obtiene un árbol más eficiente para tomar decisiones en problemas de mercados financieros.

Jin Li en 2006 [46] presenta un algoritmo que utiliza la programación genética para resolver problemas multi-objetivo para pronosticar sistemas financieros. En este caso, se resuelven problemas de pronósticos financieros en los que se hace uso de la dominancia de Pareto para encontrar varias soluciones eficientes en una misma ejecución, permitiendo al usuario tomar una decisión más completa.

Programación Compromiso

Eduardo Conde (1994) realiza un análisis interactivo de las soluciones del problema lineal múltiple ordenado. En este trabajo se estudian procedimientos para la elección entre las soluciones eficientes del Problema de Programación Lineal con Objetivos Múltiples, cuando el decisor manifiesta preferencias sobre ciertas ordenaciones de las valoraciones de las funciones objetivo, utilizándose como criterio de valoración global funciones basadas en el k-ésimo valor del vector de los objetivos ordenado en cada punto.

Mariano Jiménez López (1998) en un artículo propone un método para resolver un problema de programación lineal participación de los parámetros de distribución difusa, cuya posibilidad se da por números difusos.

Enrique Ballester (1998) propone funciones de utilidad-compromiso y demuestra que una función de utilidad con independencia aditiva (expandida alrededor del punto ideal) es reducible a la suma ponderada de las distancias CP con métricas desde 1 a infinito. Este enlace entre utilidad y compromiso se fundamenta en una hipótesis que cumplen conjuntos de funciones usuales de utilidad, y conduce a algunos resultados destacables: 1) un método para la especificación y optimización de las funciones de utilidad mediante técnicas operacionales; y 2) una reformulación del análisis compromiso que tiene la ventaja de determinar la mejor solución CP desde una perspectiva de utilidad.

Perez Gladish (2001) propone un modelo de Programación Compromiso con parámetros difusos aplicado a la gestión de listas de espera quirúrgicas. Mediante el modelo propuesto se determinan planificaciones óptimas de actividad quirúrgica que posibilitan

el mantenimiento de las listas de espera en los niveles establecidos por el Ministerio de Sanidad para cada momento.

CAPITULO I: SOCIEDADES DE INVERSIÓN

1.1 Definición de las sociedades de inversión

Las sociedades de inversión, mejor conocidas como fondos, son la forma más accesible para que los pequeños y medianos inversionistas puedan beneficiarse del ahorro en instrumentos bursátiles. El inversionista compra acciones de estas sociedades cuyo rendimiento está determinado por la diferencia entre el precio de compra y el de venta de sus acciones. Los recursos aportados por los inversionistas son aplicados por los fondos a la compra de una canasta de instrumentos del mercado de valores, procurando la diversificación de riesgos. Estas instituciones forman carteras de valores o portafolios de inversión con los recursos que captan del público inversionista. La selección de estos valores se basa en el criterio de diversificación de riesgos. Al adquirir las acciones representativas del capital de estas sociedades, el inversionista obtiene ventajas tales como la diversificación de sus inversiones, principio fundamental para disminuir el riesgo y, la posibilidad de participar del Mercado de Valores en condiciones favorables sin importar el monto de los recursos aportados. Para un inversionista pequeño o mediano, adquirir unitariamente instrumentos del mercado de valores, equivaldría a concentrar excesivamente su inversión. Ello, sin considerar que, en muchos casos, son elevados los montos mínimos exigidos para la compra de un instrumento bursátil en particular. En una sociedad de inversión, en cambio, los recursos del inversionista se suman a los de otros, lo que permite ampliar las opciones de valores bursátiles consideradas. Adicionalmente, no todos los inversionistas cuentan con el tiempo o los conocimientos requeridos para participar por cuenta propia en el mercado de valores, por lo que dicha tarea y habilidad queda en manos de los profesionales que trabajan en las operadoras de sociedades de inversión, las cuales funcionan de manera independiente o como subsidiarias de intermediarios financieros. En México, el público interesado en recibir asesoría e invertir en sociedades de inversión puede acudir con cualquiera de los intermediarios siguientes:

- Casas de bolsa: Institución privada que actúa en el mercado de capitales, y que opera por una concesión del Gobierno Federal. Su finalidad principal es la de auxiliar a la bolsa de valores en la compra y venta de diversos tipos de títulos mercantiles tales como bonos, valores, acciones, etc.

- Bancos: Un banco es una institución financiera que se encarga de administrar y prestar dinero. La banca, o el sistema bancario, es el conjunto de entidades o instituciones que, dentro de una economía determinada, prestan el servicio de banco.
- Operadoras independientes de Sociedades de Inversión: Son entidades autorizadas por la CNBV para prestar a las sociedades de inversión, entre otros, los servicios de Administración de Activos, que consisten en la celebración de operaciones de compra, venta o inversión de Activos Objeto de Inversión a nombre y por cuenta de la sociedad de inversión, así como el manejo de carteras de valores en favor de sociedades de inversión y de terceros, ajustándose a lo establecido en la Ley del Mercado de Valores. Tales entidades contarán con todo tipo de facultades y obligaciones para administrar, como si se tratara de un apoderado con poder general para realizar actos de tal naturaleza, debiendo observar en todo caso, el régimen de inversión aplicable a la sociedad de inversión de que se trate, así como su prospecto de información al público inversionista, salvaguardando en todo momento los intereses de los accionistas de la misma, para lo cual deberán proporcionarles la información relevante, suficiente y necesaria para la toma de decisiones.

1.2 Importancia de las sociedades de inversión

Las sociedades de inversión cumplen varias funciones importantes para el conjunto de la actividad económica del país, entre las que se pueden destacar:

- Fomentar el ahorro interno al ofrecer más opciones de inversión atractivas para los ahorradores nacionales.
- Contribuir a captar ahorro externo como complemento del interno al permitir la compra de acciones de sociedades de inversión a inversionistas extranjeros.
- Participar en el financiamiento de la planta productiva al canalizar recursos de los inversionistas a la compra de acciones y títulos de deuda emitidos por las empresas y el gobierno, con los que financian proyectos de modernización y ampliación.

- Fortalecer el mercado de valores al facilitar la presencia de un mayor número de participantes.
- Propician la democratización del capital al diversificar su propiedad accionaria entre varios inversionistas.

1.3 Clasificación de las sociedades de inversión

De acuerdo a la Ley de Sociedades de Inversión existen tres tipos:

SOCIEDADES DE INVERSION EN INSTRUMENTOS DE DEUDA.

Estas sociedades sólo pueden invertir en instrumentos de deuda y cuya utilidad y pérdida neta se asigna diariamente entre los accionistas. Las primeras de estas sociedades iniciaron su operación a finales de 1983, y básicamente se constituyeron como fondos de mercado de dinero, es decir las características básicas que ofrecían estas sociedades eran alta liquidez y rendimiento, y por consiguiente su cartera se encontraba invertida en instrumentos de mercado de dinero. Los posibles adquirentes para este tipo de Sociedades de Inversión para personas físicas son los siguientes:

- a) Personas morales mexicanas o extranjeras
- b) Instituciones de crédito.
- c) Fideicomisos cuyos fideicomisarios sean personas morales
- d) Entidades extranjeras
- e) Agrupaciones de personas morales extranjeras
- f) Dependencias y entidades de la administración pública federal y de los estados,
- g) Municipios
- h) Fondos de ahorro y pensiones
- i) Instituciones de seguros y fianzas
- j) Uniones de crédito
- k) Arrendadoras

Las características de estas son:

- Representan un instrumento de inversión a bajo riesgo.
- Captan recursos adicionales para financiar instrumentos del mercado de dinero y de capitales.

- Por su naturaleza las emisiones adquiridas son tomadas hasta el vencimiento.
- Con las alzas de tasas de interés, al ajustarse los precios de mercado de los instrumentos de inversión, el precio de la sociedad puede disminuir, ajustando al alza sus nuevos rendimientos.
- Reinversión automática.
- Valuación constante de sus activos.

SOCIEDADES DE INVERSION DE RENTA VARIABLE

Fueron las primeras en aparecer en el país y sus activos se invierten en valores de renta variable e instrumentos de deuda. Pueden invertir personas físicas y personas morales. El inversionista obtiene una ganancia de capital que consiste en la diferencia entre el precio de venta y el precio de compra. Esta ganancia es exenta de impuestos para las personas físicas y es acumulable para las personas morales.

SOCIEDADES DE INVERSION DE CAPITALS.

Las Sociedades de Inversión de Capitales (SINCAS) invierten sus recursos de manera temporal en empresas que por sus características particulares presentan viabilidad financiera e importante capacidad de desarrollo productivo que derivan en un retorno sobre el capital invertido de la SINCA.

1.4 Artículos de la Ley de Sociedades de Inversión que dan legalidad a las mismas

El 28 de diciembre de 1992 aparece en el Diario Oficial de la Federación el decreto que reforma, adiciona y deroga diversas disposiciones de la Ley de Sociedades de Inversión, debido a la evidente necesidad de actualizar el marco jurídico que regula este tipo de sociedades con la finalidad de contar con las condiciones adecuadas para flexibilizar su operación y aprovechar aún más las potencialidades que encierra este mecanismo de inversión para la movilización de recurso financieros, por ende se citan y resumen los artículos más significativos para efectos de este trabajo de tesis:

ARTICULO 5. Establece las facultades que tienen las sociedades de inversión a nivel financiero así como sus responsabilidades.²

ARTICULO 6. Determina los tipos de sociedades de inversión

ARTICULO 7. Define las modalidades que las sociedades de inversión podrán adoptar para su operación.

ARTICULO 12. Establece la forma en que pueden constituirse las sociedades de inversión así como de su organización para su operación en el mercado de valores.

ARTICULO 15. Determina las operaciones que las sociedades de inversión podrán realizar así como las disposiciones a las cuales estarán sujetas ante el Banco de México.

ARTICULO 17. Menciona el derecho que las sociedades tienen de recibir una calificación por parte de una calificadora de valores.

ARTICULO 24. Determina de la operación exclusiva que tendrán las sociedades de inversión con los Activos Objetos.

ARTICULO 45. Determina la manera en que los precios actualizados de la valuación de las acciones de las sociedades de inversión, se darán a conocer al público..

ARTICULO 48. Menciona sobre los servicios de calificación de sociedades de inversión que otorgarán las instituciones calificadoras de valores.

² Ley de Mercados de Valores. Cámara de Diputados del H. Congreso de la Unión. Secretaría General. Secretaría de Servicios Parlamentarios. Centro de Documentación, Información y Análisis. Mayo 2009

CAPITULO II. TEORIA MODERNA DEL PORTAFOLIO

2.1 Introducción

Markowitz [1952] desarrolló una teoría en la cual los inversionistas construyen portafolios basados exclusivamente en el riesgo y en el rendimiento esperado. Aquí el riesgo es entendido como la variabilidad del retorno de la inversión, y los inversionistas – en este modelo– prefieren lograr rendimientos con la menor variabilidad posible, es decir, que tienen aversión al riesgo. Cuando se invierte un capital en un portafolio se logra conseguir un rendimiento particular con menor riesgo que el de invertir todo el capital en un solo activo. Este fenómeno es conocido como "diversificación".

Es común pensar que a mayor número de activos, mayor diversificación del portafolio. Por ejemplo, si se invierte un capital en n activos en iguales cantidades, es posible ver que el riesgo de esta inversión disminuye a medida que n se hace más grande. Este tipo de diversificación se conoce como "diversificación ingenua". El modelo de Markowitz no está basado en esta clase de diversificación, sino en las correlaciones de los activos de riesgo. La diversificación basada en la correlación, y no en el número de activos, es llamada "diversificación eficiente".

Se inicia la presentación de los conceptos de riesgo y rendimiento esperado, tanto para un activo individual como para un portafolio. Luego se expone la función de utilidad esperada de un inversionista. La siguiente sección se trata de cómo se compone un conjunto eficiente de portafolios de activos de riesgo que contenga las mejores combinaciones de riesgo y rendimiento esperado (R-R). Más adelante se muestra que al incluir un activo cuyo retorno no es aleatorio es posible construir portafolios con mejores combinaciones de R-R. Bajo la existencia de este activo, además de otras condiciones, es posible mostrar que el portafolio de activos de riesgo es independiente de las preferencias. Esto último se conoce como el "teorema de la separación". Finalmente, se construyen dos portafolios eficientes con cinco acciones de la bolsa de Valores de Colombia (BVC), uno para un inversionista conservador y otro para un inversionista emprendedor, ambos aversos al riesgo.

2.2 Rendimiento Esperado y Riesgo

Dado que el rendimiento futuro de los activos financieros es incierto, éste es considerado como una variable aleatoria. Así, la incertidumbre hace que además de los rendimientos esperados, los analistas deban tener en cuenta el riesgo de los activos financieros. Por este motivo, la teoría moderna de la inversión hace uso de distribuciones de probabilidad para estimar el rendimiento futuro de los activos financieros y el riesgo asociado.

La teoría del portafolio considera que en las decisiones de inversión sólo se tienen en cuenta el retorno esperado y el riesgo. El primer momento de la distribución del retorno es usado como estimación del retorno esperado, y la varianza (o la desviación estándar) del retorno es empleada como medida del riesgo. En el área financiera, la desviación estándar es conocida como la volatilidad.

El retorno esperado de un activo de riesgo es calculado con el primer momento de la distribución de los retornos:

$$\bar{r} = E(r) = \sum_r r f(r) \quad \text{para } r \text{ variable aleatoria discreta} \quad (2.1)$$

$$\bar{r} = E(r) = \int_r^{\infty} r f(r) dr \quad \text{para } r \text{ variable aleatoria continua} \quad (2.2)$$

Donde $f(r)$ representa la función de probabilidad del retorno, la cual puede ser continua o discreta. El segundo momento de esta distribución, con respecto a la media, constituye el riesgo del activo:

$$\sigma^2 = \sum_r (r - \bar{r})^2 f(r) \quad \text{para } r \text{ variable aleatoria discreta} \quad (2.3)$$

$$\sigma^2 = \int_r^{\infty} (r - \bar{r})^2 f(r) \quad \text{para } r \text{ variable aleatoria continua} \quad (2.4)$$

Adicional al riesgo y al rendimiento esperado, la teoría requiere el cálculo de las relaciones entre los rendimientos de estos activos. Para este cálculo se puede emplear la covarianza o el coeficiente de correlación. La covarianza entre retornos de las parejas de activos A y B se determina como:

$$\sigma_{AB} = E[(r_A - \bar{r}_A)(r_B - \bar{r}_B)] = E(r_A r_B) - E(r_A)E(r_B) \quad (2.5)$$

y el coeficiente de correlación de los retornos de esta pareja como:

$$\rho_{AB} = \frac{\sigma_{AB}}{\sigma_A \sigma_B} \quad (2.6)$$

El coeficiente de correlación ρ es una medida que varía en el intervalo $[-1, 1]$. Un valor $\rho = -1$ indica una perfecta relación lineal negativa, mientras que un valor de $\rho = +1$ indica una perfecta relación lineal positiva entre los retornos de A y B.

Un portafolio constituido por n activos individuales se puede representar mediante un vector de n elementos: $[w_1, w_2, \dots, w_n]^T$ tal que

$$\sum_{i=1}^n w_i = 1.$$

El elemento w_i indica la tasa de participación del activo individual i en el portafolio.

La esperanza matemática del rendimiento de un portafolio resulta ser el promedio de los n rendimientos esperados individuales, ponderados por la participación de cada activo individual en el portafolio:

$$E(r_p) = \bar{r}_p = E(\sum_{i=1}^n r_i w_i) = \sum_{i=1}^n E(r_i) w_i = \sum_{i=1}^n \bar{r}_i w_i \quad (2.7)$$

Este también se puede expresar como el producto entre el vector transpuesto de \mathbf{w} y el vector de rendimientos esperados individuales $\bar{\mathbf{r}}$:

$$E(r_p) = \mathbf{w}^T \bar{\mathbf{r}} \quad (2.8)$$

De otro lado, la varianza del portafolio se determina a partir de:

$$\begin{aligned} \sigma_p^2 &= E[(r_p - \bar{r}_p)^2] = E[(\sum_{i=1}^n w_i r_i - \sum_{i=1}^n w_i \bar{r}_i)^2] \\ &= E[(\sum_{i=1}^n w_i r_i - \sum_{i=1}^n w_i \bar{r}_i)(\sum_{i=1}^n w_i r_i - \sum_{i=1}^n w_i \bar{r}_i)] \\ &= E[\sum_{i=1}^n w_i (r_i - \bar{r}_i) \sum_{j=1}^n w_j (r_j - \bar{r}_j)] \\ &= E[\sum_{i=1}^n \sum_{j=1}^n w_i w_j (r_i - \bar{r}_i)(r_j - \bar{r}_j)] \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j E[(r_i - \bar{r}_i)(r_j - \bar{r}_j)] \\ \sigma_p^2 &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \end{aligned} \quad (2.9)$$

Donde i y j representan dos activos individuales del portafolio, w_i y w_j sus participaciones en el portafolio, y σ_{ij} la covarianza entre sus retornos. En la expresión anterior la varianza del portafolio está en función de las participaciones de los activos individuales dentro del portafolio, y de las covarianzas (o correlaciones) entre los retornos medios individuales. Se observa que entre más baja sea la covarianza (o correlación) entre los retornos de los activos individuales, menor será la variabilidad (riesgo) del portafolio. Así, el poder de diversificación de un portafolio está en la correlación. Los portafolios que incluyan activos con la menor correlación posible (especialmente negativa) presentarán menor volatilidad –menor riesgo– que aquellos que incluyan activos altamente correlacionados.

La varianza del portafolio también puede representarse matricialmente como:

$$\sigma_p^2 = \mathbf{w}^T \mathbf{\Omega} \mathbf{w} \quad (2.1)$$

Donde $\mathbf{\Omega}$ representa la matriz de varianzas y covarianzas de los retornos de los activos del portafolio.

2.3 Estructura de preferencias bajo incertidumbre

Bajo incertidumbre, la estructura de preferencias difiere del enfoque tradicional, en el cual, a partir de ciertos axiomas se construye una función de utilidad que depende de la cantidad de bienes no inciertos. La utilidad que deriva un consumidor o inversionista de un activo incierto puede ser tratada por la función de utilidad propuesta por Neumann y Morgenstern [1944]. Estos autores definen el concepto de "utilidad esperada" como la utilidad de un activo incierto o variable aleatoria. Tomemos el caso en que la utilidad de un inversionista se representa por una función que está en términos del rendimiento cierto: $u = u(r)$. Al adquirir un activo incierto cuya función de probabilidad del retorno es $f(r)$, la utilidad esperada de este inversionista será:

$$UE = \sum_r [u(r)f(r)] , \quad \text{con } r \text{ v. a. discreta} \quad (2.11)$$

$$UE = \int_r u(r)f(r)dr , \quad \text{con } r \text{ v. a. continua} \quad (2.12)$$

De este modo, la utilidad esperada queda en términos de la función de utilidad del retorno bajo certeza y de la función de probabilidad del retorno. Markowitz propone que la anterior función de utilidad esperada sea resumida por los dos primeros momentos de la distribución de probabilidad del retorno:

$$UE = u(\sigma_p^2, \bar{r}_p) \quad (2.13)$$

Adicionalmente, Markowitz supone que el inversor es racional y averso al riesgo, es decir, que el inversionista prefiere mayor a menor retorno y, a su vez, menor a mayor riesgo. Asumiendo que $u(\sigma_p^2, \bar{r}_p)$ es diferenciable, entonces:

$$\begin{aligned} \partial u(\sigma_p^2, \bar{r}_p) / \partial \bar{r}_p &> 0 && \text{(inversionista racional)} \\ \partial u(\sigma_p^2, \bar{r}_p) / \partial \sigma_p &< 0 && \text{(inversionista averso al riesgo)} \end{aligned} \quad (2.14)$$

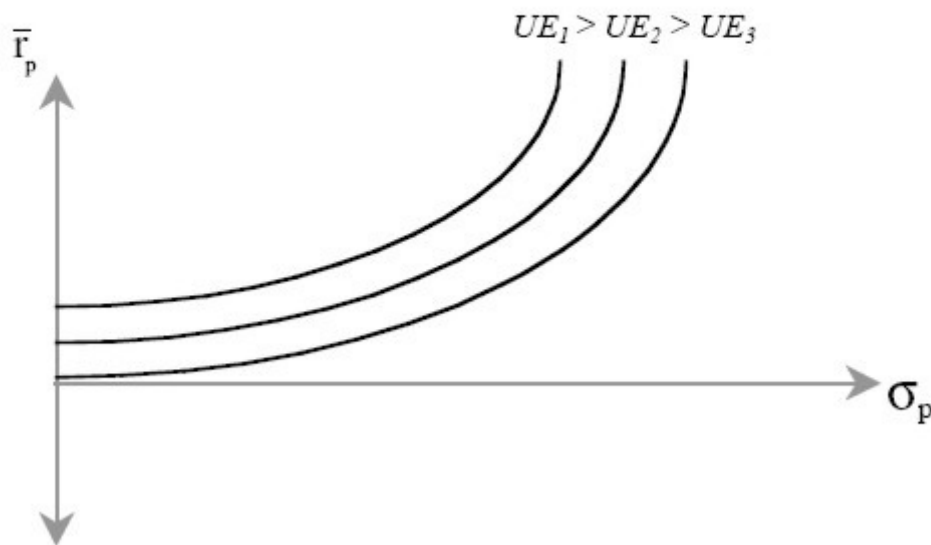
Alternativamente, podemos formalizar este argumento con el siguiente axioma:

Axioma Media-Volatilidad (M-V): sean dos portafolios factibles, A y B, con combinaciones media-volatilidad: (\bar{r}_A, σ_A) y (\bar{r}_B, σ_B) , respectivamente. El portafolio A se prefiere al portafolio B si se cumple que:

$$\begin{aligned} \bar{r}_A > \bar{r}_B \text{ y } \sigma_A \leq \sigma_B & \quad \text{ó} \\ \bar{r}_A \geq \bar{r}_B \text{ y } \sigma_A < \sigma_B \end{aligned} \quad (2.15)$$

Si los principales criterios de inversión son el riesgo y el rendimiento esperado, resulta conveniente representar todos los portafolios en un espacio P generado por la media y la volatilidad: $P: \{(\sigma_p, \bar{r}_p): \sigma_p \in \mathbb{R}^+ \cup \{0\}, \bar{r}_p \in \mathbb{R}\}$. Ver Gráfica 1

GRÁFICA 1 ESPACIO MEDIA-VOLATILIDAD Y CURVAS DE INDIFERENCIA



En el espacio Media-Volatilidad se grafican las curvas de indiferencia de los inversionistas. Cada una de estas curvas representa un nivel de utilidad esperada – razón por la cual son paralelas– y entre más hacia el noroeste se localicen, mayor utilidad esperada representan. Las curvas de indiferencia son de pendiente positiva por el axioma M-V, ya que el riesgo es un mal y el retorno esperado es un bien. Por lo tanto, partiendo de un portafolio inicial, si existe otro con mayor rendimiento esperado, éste debe contener mayor riesgo para que pertenezca al conjunto de indiferencia del portafolio inicial. Adicionalmente, se asume convexidad de las curvas de indiferencia, lo cual implica que el inversor tiene aversión creciente al riesgo con respecto al nivel de riesgo asumido.

En síntesis, el modelo de preferencias propuesto por Markowitz implica que los inversionistas sólo se fijan en el retorno medio y en el riesgo de la inversión. James Tobin [1958] muestra las condiciones bajo las cuales tal argumento es válido. Si la función de utilidad en términos del retorno es cuadrática o el retorno se distribuye normalmente, este modelo de preferencias propuesto por Markowitz es adecuado.

2.4 Desarrollo de la teoría del portafolio

La teoría básica de la selección de portafolios fue desarrollada inicialmente por el Premio Nóbel Harry Markowitz a comienzos de los años 50. Más adelante fue ampliada por otros economistas, que se aportaron aspectos importantes para incrementar su desarrollo teórico.

Esto lo llevo a ampliar la teoría de Williams Sharpe, ya que para Markowitz el valor de capital debería ser el valor presente “esperado” de su flujo de dividendos. Markowitz observó, que un inversionista no sólo debe tener en cuenta la rentabilidad esperada más alta, sino también el riesgo que implica esta inversión.

Esto lo llevó a examinar el problema de encontrar un portafolio con el máximo retorno esperado a un nivel de riesgo dado. Llevándolo a plantear el problema de minimizar las variaciones de un portafolio tomando como restricción el requerimiento de un retorno esperado. (Markowitz, 1952). Es decir, propuso un problema de programación cuadrática, el cual tenía como condiciones de primer orden el aumento marginal en la varianza de invertir un poco más en un activo dado y debería ser proporcional al retorno dado.

Esta variación depende tanto de la varianza del retorno del activo, como de la covarianza del retorno de todos los demás activos del portafolio. Esta se considera como la idea central de la contribución de Markowitz. En 1952 la programación lineal estaba naciendo y la programación cuadrática no estaba definida, sin embargo Markowitz desarrolló métodos prácticos para determinar la *línea crítica* para solucionar los problemas de optimización, describiendo portafolios eficientes con varianza mínima y rentabilidad alta.

Finalmente, Markowitz publicó dos documentos en 1952³ y 1956⁴ y su libro clásico en 1959. El modelo de Markowitz es la base de la mayoría de los modelos de selección de cartera. Sin embargo, su utilización en la práctica es bastante reducida.

El motivo de ello tiene que ver con sus dificultades de cálculo, la inestabilidad de las soluciones que proporciona, los problemas para incluir opiniones de los expertos y la rigidez de la función de riesgo considerada. Este modelo se fundamenta principalmente en recoger de forma explícita los rasgos fundamentales de lo que en un principio se

³ Markowitz, H.M. (March 1952). "Portfolio Selection". *The Journal of Finance* 7 (1): 77–91.

⁴ The optimization of a quadratic function subject to linear constraints", *Naval Research Logistic Quarterly*, 3.)

puede calificar como conducta racional del inversor, consistente en buscar aquella composición de la cartera que haga aportar aspectos importantes para incrementar su desarrollo teórico.

De acuerdo con Varian, (1993), Markowitz observó que en el libro *La teoría del Valor de Invertir* de John Burr Williams, se argumentaba que el valor del capital debería ser el valor presente de sus dividendos (lo que era considerado una teoría nueva). Pero Markowitz vio como esto presentaba un problema, ya que los dividendos futuros no son conocidos con certeza; son variables al azar. Esto lo llevo a ampliar la teoría de Williams Sharpe, ya que para Markowitz el valor de capital debería ser el valor presente “esperado” de su flujo de dividendos.

Markowitz observó, que un inversionista no sólo debe tener en cuenta la rentabilidad esperada más alta, sino también el riesgo que implica esta inversión. Esto lo llevó a examinar el problema de encontrar un portafolio con el máximo retorno esperado a un nivel de riesgo dado.

Llevándolo a plantear el problema de minimizar las variaciones de un portafolio tomando como restricción el requerimiento de un retorno esperado. (Markowitz, 1952). Es decir, propuso un problema de programación cuadrática, el cual tenía como condiciones de primer orden el aumento marginal en la varianza de invertir un poco más en un activo dado y debería ser proporcional al retorno dado. Esta variación depende tanto de la varianza del retorno del activo, como de la covarianza del retorno de todos los demás activos del portafolio. Esta se considera como la idea central de la contribución de Markowitz.

En 1952 la programación lineal estaba naciendo y la programación cuadrática no estaba definida, sin embargo Markowitz desarrolló métodos prácticos para determinar la *línea crítica* para solucionar los problemas de optimización, describiendo portafolios eficientes con varianza mínima y rentabilidad alta.

Markowitz centró su atención en la práctica habitual de la diversificación de carteras y mostró como un inversor puede reducir la desviación típica de las rentabilidades de una cartera eligiendo acciones cuyas oscilaciones no sean paralelas. Markowitz continuó con el desarrollo de los principios básicos de la formación de carteras. Estos principios son el fundamento de todo lo que pueda decirse entre riesgo y rentabilidad.

La rentabilidad de cualquier título o cartera, es una variable aleatoria de carácter subjetivo, cuya distribución de probabilidad para el periodo de referencia es conocido

por el inversor. El valor medio o esperanza matemática de dicha variable aleatoria se acepta como medida de la rentabilidad de la inversión. Se acepta como medida del riesgo la dispersión, medida por la varianza o la desviación estándar, de la variable aleatoria que describe la rentabilidad, ya sea de un valor individual o de una cartera. La conducta del inversor le lleva a preferir aquellas carteras con una mayor rentabilidad y menor riesgo.

En la primera etapa se determina el conjunto de Carteras Eficientes cuando proporciona la máxima ganancia para un riesgo (medido por la varianza) dado, o bien, proporciona el mínimo riesgo para un valor dado de ganancia (Esperanza Matemática). Este supuesto ya fue probado (Jiménez *et al*,2000) y llegaron a la conclusión que en economías poco desarrolladas, esta teoría de cartera no se puede aplicar, es decir, se necesita realizar muchas modificaciones al modelo para encontrar resultados que se puedan comparar con experimentos realizados en Estados Unidos o en los mercados Europeos.

A pesar de estos principios, el modelo en el mercado colombiano no tiene gran aplicabilidad ya que lleva implícitas dos condiciones fundamentales: Para que el modelo se pueda desarrollar gráficamente el número de alternativas de inversión debe ser manejable, es decir, sólo se puede trabajar con dos o tres posibilidades.

A partir de ahí, existen grandes dificultades para extrapolar al caso general de n posibilidades de inversión, y en el caso de aplicación se conformará un portafolio de 10 acciones para el cual este modelo no aplicaría. Se deben asumir determinados postulados de probabilidad estática cuando el estudio trata un análisis dinámico. De esta manera, habría que reconocer que la distribución de probabilidad de los rendimientos de cualquier alternativa de inversión depende del tiempo. Debido a que el análisis se desarrolla en el largo plazo esta condición no es necesaria en el método que se plantea ya que el desarrollo se hará en el corto plazo.

2.5 Modelo de Markowitz como Problema de Programación Multiobjetivo

El modelo de Media- Varianza de Markowitz (1952) se define como la optimización simultánea de dos funciones objetivo. Maximización del promedio del rendimiento del portafolio y la minimización de las correlaciones del portafolio de inversión

$$\text{Max } E(R_p) = \sum_{i=1}^n w_i \times E(R_i)$$

Riesgo del Portafolio

$$\text{Min} \sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n w_i \times w_j \times \rho_{ij} \times \rho_i \times \rho_j$$

s.a

$$\sum_{i=1}^n w_i = 1$$

$$w_i \geq 0$$

(2.16)

Donde R_p = rendimiento del portafolio, w_i =ponderaciones de riesgo, ρ_{ij} =correlacion entre acciones

Después de una revisión de las referencias bibliográficas de la aplicación del enfoque del paradigma evolutivo se plantean las siguientes consideraciones:

La primera pregunta que surge, es por que no diseñar tan solo un algoritmo que solucione el problema cuadrático propuesto inicialmente por Markowitz.

Al respecto Vedejaran et al (1997) señala que para un portafolio de tamaño n , el conjunto de datos de entradas es $(2*n)+(n*(n-1)/2)$, por lo que si n es grande puede ser laborioso seguir la pista de todos los coeficientes de correlación, y se pueden generar problemas por restricciones de almacenamiento de memoria.

Además y mas restrictivo que lo anterior, es que para resolver el problema, es necesario que la matriz de covarianzas sea positiva y si existen discrepancias numéricas (por ejemplo imprecisión por redondeo), esta suposición podría ser violada y no se podría resolver. Esto no es improbable con datos reales particularmente cuando n es muy grande, dado que las covarianzas son estimaciones de series de tiempo de los precios de los instrumentos financieros reales.

Por otro lado, debido al escaso uso del modelo, éste no considera aspectos como el costo de transacción por cambios en el portafolio y una mayor diversificación del portafolio.

Para incorporar la diversificación del portafolio es necesario agregar la restricción

$$\forall i \in \{1, 2, \dots, n\}: w_i \leq W_{\max} \quad \text{donde } W_{\max} \text{ es una constante.}$$

La variante que considera el costo de transacción por cambios en el portafolio (rebalanceo), causa problemas al algoritmo cuadrático, cuestión que puede ser resuelta en el enfoque multi-objetivo introduciendo una tercera función objetivo a ser minimizada.

$$f_{\text{cost}} = \sum_{i=1}^n c_i \times (w_i - w_i^*)^2$$

donde:

$$w_i^* \in \mathbb{R}_+^0 \quad (2.17)$$

es la participación inicial del instrumento financiero i en el portafolio que va a ser potencialmente cambiado, debido a transacciones por rebalanceo y la constante $c_i c_R$ es el costo de transacción del instrumento financiero i .

CAPITULO III. METODOS HEURÍSTICOS DE OPTIMIZACIÓN

El término “optimización” se refiere a un problema complejo de decisión, que incluye la selección de valores para cierto número de variables interrelacionadas, centrando la atención en un solo objetivo diseñado para cuantificar el rendimiento y medir la calidad de la decisión. Este único objetivo se maximiza según las restricciones que pueden limitar la selección de valores de las variables de selección⁵. A continuación se hace una clasificación de los métodos de optimización, mencionando brevemente cómo funcionan algunos de ellos.

3.1 Método del Gradiente y Hill Climbing

Un número de diferentes métodos de optimización se comportan como funciones continuas que han sido desarrolladas para utilizar información del gradiente de la función para guiar la búsqueda del óptimo global. Estos métodos como el del gradiente conjugado, que usa la información de la función objetivo así como de sus derivadas parciales, tiene algunas desventajas, ya que si la derivada de la función no puede ser computada por el hecho de ser discontinua, el método tiene a fallar⁶.

Otro método es el denominado comúnmente como Hill Climbing (escalado de colinas), que no requiere de la información de las derivadas de la función y que funciona adecuadamente sólo si la función contiene un solo máximo (funciones multimodales) tienden a cometer errores ya que el primer máximo local que encuentren en la función lo toman como si fuera el máximo global de la función y por tanto no continúan la búsqueda. Un ejemplo en una sola dimensión se presenta en la figura siguiente:

⁵ Luenberger, 1989

⁶ Beasley, 1993

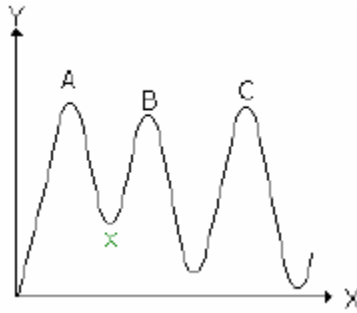


Figura 1. Grafica de una función multimodal

El escalado comienza en un punto aleatorio X y se mueve en forma ascendente donde le máximo B es encontrado, por lo tanto los máximos mejores, A y C no son encontrados⁷.

3.2 Búsqueda Iterativa

La búsqueda aleatoria combinada con el método del gradiente da como resultado la búsqueda iterativa. Cada vez que se encuentra un máximo, la búsqueda comienza de nuevo pero con otro punto escogido al azar. Esta técnica es de enorme simplicidad y su desempeño es bueno en funciones que no contengan muchos máximos locales.

Cada vez que se busca otro punto aleatoriamente, éste es aislado y no se vuelve a tomar como parte del espacio de búsqueda. Esto significa que al estar buscando puntos éstos pueden encontrarse en regiones donde existan máximos locales y no globales⁸.

3.3 Algoritmo Metrópolis

Este método está basado en la analogía con el proceso de recocido de sólidos:

- Se incrementa la temperatura alta.
- Posteriormente se decrementa la temperatura muy lentamente hasta alcanzar un estado base.

La idea principal de este algoritmo es la siguiente:

⁷ Beasley, 1993

⁸ Idem

1. Dado un estado inicial i , con energía E_i ,
2. Genera un nuevo estado j mediante un mecanismo de perturbación (pequeña distorsión del estado i).
3. Calcula la energía del estado E_j .
4. Si $(E_j - E_i) \leq 0$ entonces acepta el estado j como nuevo.
5. Si no, acepta el estado con probabilidad $e^{\left(\frac{E_j - E_i}{K_B T}\right)}$

donde T denota la temperatura y K_B la constante de Boltzmann.

Si se baja la temperatura lentamente se puede alcanzar el equilibrio térmico en cada temperatura. Esto se hace mediante la generación de varias transiciones en cada temperatura.

El equilibrio térmico está caracterizado por la distribución de Boltzmann. La distribución da la probabilidad de que el sólido esté en el estado i con energía E_i a temperatura T y está dada por

$$P_T\{X=i\} = \frac{1}{Z(T)} \exp\left(\frac{-E_i}{K_B T}\right) \quad (3.1)$$

Donde X es la variable estocástica que denota el estado actual del sólido y $Z(T)$ es la función definida como⁹:

$$Z(T) = \sum_j \exp\left(\frac{-E_j}{K_B T}\right) \quad (3.2)$$

3.4 Recocido Simulado

Esta técnica fue inventada por Kirkpatrick en 1982, está basada en el algoritmo Metrópolis propuesto en 1953 y es esencialmente una variación del hill climbing. Kirkpatrick introdujo los conceptos de recocido dentro del mundo de la optimización combinatoria. Comienza con la búsqueda de un punto escogido aleatoriamente y posteriormente con un movimiento aleatorio. Si el punto encontrado es uno de los máximos, entonces es aceptado. Si el punto es un mínimo entonces es aceptado con

⁹ Morales, 1999

probabilidad $p(t)$, donde es aceptado. El valor inicial de la función $p(t)$ es muy cercano a uno y gradualmente se va acercando a cero; esto es análogo al enfriamiento de un sólido.

Inicialmente cualquier movimiento es aceptado, pero conforme la temperatura se va reduciendo, la probabilidad de aceptar un movimiento negativo es mínima. Algunas veces los movimientos negativos son esenciales sobre todo si no se ha encontrado el máximo global, aunque obviamente muchos movimientos negativos pueden desviarnos a la búsqueda del óptimo global. Al igual que la búsqueda aleatoria, el recocido simulado trata con una solución e mismo tiempo, esto ocasiona que no se vaya construyendo la figura a partir del espacio de búsqueda. La información de movimientos previos no es guardada para guiar los nuevos movimientos. Esta técnica está en constante actividad de investigación (recocido rápido, recocido paralelo, etc.)¹⁰.

3.5 Búsqueda Tabú (BT)

La BT es un procedimiento heurístico de “alto nivel” introducido y desarrollado en su forma actual por Fred Glover (1989) y (1990). En términos generales el método de BT puede resumirse de la siguiente manera: Se desea moverse paso a paso desde una solución factible inicial de un problema de optimización combinatoria hacia una solución que proporcione el valor mínimo de la función objetivo C . Para esto se puede representar a cada solución por medio de un punto s (en algún espacio) y se define una vecindad $N(s)$ para cada punto s como un conjunto de soluciones adyacentes a la solución s . El paso básico del procedimiento consiste en empezar desde un punto factible s y generar un conjunto de soluciones en $N(s)$, de estas se elige la mejor s^* y se posiciona en este nuevo punto ya sea que $C(s^*)$ tenga o no mejor valor que $C(s)$.

¹⁰ Beasley, 1993

Hasta este punto se está cercano a las técnicas de mejoramiento local a excepción del hecho de que se puede mover a una solución peor s^* de s .

La característica importante de la BT es precisamente la construcción de una lista tabú T de movimientos: aquellos movimientos que no son permitidos (movimientos tabú) en la presente iteración. La razón de esta lista es la de excluir los movimientos que nos pueden regresar a algún punto de una iteración anterior. Ahora bien, un movimiento permanece como tabu solo durante un cierto número de iteraciones, de forma que se tiene que T es una lista cíclica donde para cada movimiento s a s^* el movimiento opuesto s^* a s se adiciona al final de T donde el movimiento más viejo en T se elimina. Las condiciones tabú tienen la meta de prevenir ciclos e inducir la exploración de nuevas regiones. Pero esto no representa que las restricciones tabú no sean inviolables bajo toda circunstancia. Cuando un movimiento tabú proporciona una solución mejor que cualquier otra previamente encontrada, su clasificación tabú puede eliminarse. La condición que permite dicha eliminación se llama criterio de aspiración. Es así como las restricciones tabú y el criterio de aspiración de la BT, juegan un papel dual en la restricción y guía del proceso de búsqueda. Las restricciones tabú, permiten que un movimiento sea admisible si no está clasificado como tabú, mientras que si el criterio de aspiración se satisface, permite que un movimiento sea admisible aunque este clasificado como tabú. En la BT tres aspectos merecen énfasis:

1. El uso de T proporciona la “búsqueda restringida” de elementos de la aproximación y por lo tanto las soluciones generados dependen críticamente de la composición de T y de la manera como se actualiza.
2. El método no hace referencia a la condición de optimalidad local, excepto implícitamente cuando un optimo local mejora sobre la mejor solución encontrada previamente.
3. En cada paso se elige al “mejor” movimiento.

Para problemas grandes, donde las vecindades pueden tener muchos elementos, o para problemas donde esos elementos son muy costosos de examinar, es de importancia aislar a un subconjunto de la vecindad, y examinar este conjunto en vez de la vecindad completa. Esto puede realizarse en etapas, permitiendo al subconjunto de candidatos expandirse si los niveles de aspiración no se encuentran¹¹.

3.6 Búsqueda ciega o sin información

El orden en que la búsqueda se realiza no depende de la naturaleza de la solución buscada. La localización de la meta(s) no altera el orden de expansión de los nodos¹².

Búsqueda Primero en Anchura

Explora progresivamente en capas del grafo de una misma profundidad. La búsqueda primero en anchura es una búsqueda completa (encuentra una solución si existe) y óptima (encuentra la solución más corta) si el costo del camino es una función que no decrece con la profundidad del nodo. Pero tiene dos desventajas: requiere de mucha memoria, ya que básicamente tiene que guardar la parte completa de la red que está explorando y la segunda es que en búsquedas de complejidad exponencial no se puede utilizar.

La búsqueda en amplitud trabaja de la siguiente forma, si se tiene un factor de arborescencia de b y la meta está a profundidad d , entonces el máximo número de nodos expandidos antes de encontrar una solución es: $1+b+b^2+b^3+b^4+\dots+bd^{13}$.

Búsqueda Primero en Profundidad

Esta búsqueda siempre expande uno de los nodos a su nivel más profundo y solo cuando llega a un camino sin salida se regresa a niveles menos profundos. Aquí cada nodo que es explorado genera todos sus sucesores antes de que otro nodo sea explorado. Después de cada expansión el nuevo hijo es de nuevo seleccionado para expansión.

¹¹ Gutierrez, 1999

¹² Morales, 1999

¹³ Morales, 1999

Si no se puede continuar, se regresa al punto más cercano de decisión con alternativas no exploradas. Esta búsqueda necesita almacenar un solo camino de la raíz a una hoja junto con los “hermanos” no expandidos de cada nodo en el camino. Por lo que con un factor de arborescencia de b y profundidad máxima de m , su necesidad de almacenamiento es a los más b_m . Su complejidad en tiempo (cuánto se tarda) es $O(bm)$ en el peor de los casos.

Para problemas con muchas soluciones la búsqueda en profundidad puede ser más rápida que de amplitud porque existe una buena posibilidad de encontrar una solución después de explorar una pequeña parte del espacio de búsqueda.

La desventaja de la búsqueda en profundidad es que en situaciones donde los arboles de búsqueda tienen una profundidad muy grande puede caerse en ciclos infinitos¹⁴.

Búsqueda Aleatoria

La búsqueda aleatoria basa su trabajo en el uso de números aleatorios para encontrar puntos mínimos. Desde que las librerías de los programas de las computadoras tienen generadores de números aleatorios, esta búsqueda puede ser utilizada convenientemente. Algunas de las ventajas de esta búsqueda aleatoria son:

- Esta búsqueda puede trabajar adecuadamente aun sin la función objetivo es discontinua y no diferenciable en alguno de sus puntos.
- La búsqueda aleatoria puede ser usada para encontrar el mínimo global aun cuando la función posea varios mínimos relativos.
- Esta búsqueda se puede aplicar cuando otros métodos tienden a fallar al presentarse dificultades como: funciones con variaciones muy marcadas y funciones con regiones planas¹⁵.

¹⁴ Idem

¹⁵ Rao, 1996

Algoritmos Genéticos y los Métodos de optimización tradicionales

Algoritmos Genéticos	Métodos de optimización tradicionales
Trabajan con parámetros codificados, es decir que deben codificarse como cadenas de longitud finita sobre algún alfabeto infinito	Trabajan con parámetros mismos
Utilizan poblaciones de puntos, es decir que se usa una base de datos de puntos simultáneamente, de tal forma que la probabilidad de quedar atrapados en óptimos locales se reduce	Operan sobre puntos individuales, ya que sus movimientos en el espacio de búsqueda se hacen de un punto a otro usando reglas de transición determinística. Esto puede ocasionar que se encuentren óptimos locales en lugar de óptimos globales
No necesitan conocimientos auxiliares sobre el problema ya que usan información de la función de evaluación con respecto a los cromosomas.	Requieren de mucha información auxiliar para trabajar adecuadamente.
Utilizan reglas de transición probabilística	Usan reglas determinística

Tabla 1. Diferencia de los algoritmos genéticos y los otros métodos de optimización tradicionales

CAPITULO IV. ALGORITMOS GENETICOS

4.1 Introducción

Los algoritmos genéticos se utilizan frecuentemente en procesos de optimización cuando, bien el espacio de búsqueda, bien la función a optimizar no gozan de las propiedades necesarias para poder aplicar técnicas clásicas del Análisis Matemático. Son algoritmos deterministas que incluyen componentes aleatorias en los procesos de inicialización y de determinación de la mejor dirección de búsqueda.

Estos algoritmos se engloban dentro de la llamada computación evolutiva que se inspira en el principio darwiniano que afirma que los individuos mejor adaptados tienen mayor probabilidad de sobrevivir en sucesivas generaciones.

Hay casi tantos algoritmos genéticos como problemas distintos en los se pueden aplicar, pero todos ellos constan de procesos de selección, evaluación y reproducción. Estos procesos se repiten cíclicamente hasta que se satisface un cierto criterio de parada.

Con respecto a la pregunta: ¿porqué funcionan los algoritmos genéticos? La respuesta la encontramos en el teorema de los esquemas de Holland que afirma que cuando el algoritmo se ejecuta todos los miembros de una población tienden a parecerse al “promedio” de la población, es decir, aumentan los miembros “promedio” de la misma, eliminando así los casos atípicos.

Aunque este resultado justifica de manera teórica el funcionamiento de estos algoritmos, en la practica varios fenómenos pueden dificultar la convergencia; entre ellos cabe citar el orden, el engaño o decepción y la epistaxis.

Los Algoritmos Genéticos (AGs) son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acorde con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin (1859).

Por imitación de este proceso, los Algoritmos Genéticos son capaces de ir creando soluciones para problemas del mundo real. La evolución de dichas soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas. Los principios básicos de los Algoritmos Genéticos fueron establecidos por Holland (1975),

En la naturaleza los individuos de una población compiten entre si en la búsqueda de recursos tales como comida, agua y refugio. Incluso los miembros de una misma especie compiten a menudo en la búsqueda de un compañero. Aquellos individuos que tienen más éxito en sobrevivir y en atraer compañeros tienen mayor probabilidad de generar un gran número de descendientes.

Por el contrario individuos poco dotados producirán un menor número de descendientes. La combinación de buenas características provenientes de diferentes ancestros, puede a veces producir descendientes "superindividuos", cuya adaptación es mucho mayor que la de cualquiera de sus ancestros. De esta manera, las especies evolucionan logrando unas características cada vez mejor adaptadas al entorno en el que viven.

Los Algoritmos Genéticos usan una analogía directa con el comportamiento natural. Trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado.

A cada individuo se le asigna un valor o puntuación, relacionado con la bondad de dicha solución. En la naturaleza esto equivaldría al grado de efectividad de un organismo para competir por unos determinados recursos.

Cuanto mayor sea la adaptación de un individuo al problema, mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma. Este cruce producirá nuevos individuos (descendientes de los anteriores) los cuales comparten algunas de las características de sus padres.

Cuanto menor sea la adaptación de un individuo, menor será la probabilidad de que dicho individuo sea seleccionado para la reproducción, y por tanto de que su material genético se propague en sucesivas generaciones. De esta manera se produce una nueva población de posibles soluciones, la cual reemplaza a la anterior y verifica la interesante propiedad de que contiene una mayor proporción de buenas características en comparación con la población anterior. Así a lo largo de las generaciones las buenas características se propagan a través de la población. Favoreciendo el cruce de los individuos mejor adaptados, van siendo exploradas las aéreas más prometedoras del espacio de búsqueda.

Si el Algoritmo Genético ha sido bien diseñado, la población convergería hacia una solución óptima del problema. El poder de los Algoritmos Genéticos proviene del hecho de que se trata de una técnica robusta, y pueden tratar con éxito una gran variedad de problemas provenientes de diferentes aéreas, incluyendo aquellos en los que otros

métodos encuentran dificultades. Si bien no se garantiza que el Algoritmo Genético encuentre la solución óptima del problema, existe evidencia empírica de que se encuentran soluciones de un nivel aceptable, en un tiempo competitivo con el resto de algoritmos de optimización combinatoria. En el caso de que existan técnicas especializadas para resolver un determinado problema, lo más probable es que superen al Algoritmo Genético, tanto en rapidez como en eficacia.

El gran campo de aplicación de los Algoritmos Genéticos se relaciona con aquellos problemas para los cuales no existen técnicas especializadas. Incluso en el caso en que dichas técnicas existan, y funcionen bien, pueden efectuarse mejoras de las mismas hibridándolas con los Algoritmos Genéticos.

La Computación Evolutiva (CE) es un enfoque alternativo para abordar problemas de búsqueda y aprendizaje a través de modelos computacionales de procesos evolutivos implementado a través de algoritmos evolutivos (AE). El proceso genérico de los AE consiste en guiar la búsqueda estocástica haciendo evolucionar a un conjunto y seleccionar de modo iterativo las más adecuadas.

La CE forma parte a su vez de un conjunto de metodologías de resolución de problemas que imitan comportamiento con mayor o menor exactitud, como las redes neuronales Hertz et al (1991), la solidificación simulada Fonseca y Fleming (1993), todas ellas se engloban bajo el término computación natural.

En general, se llama Algoritmo Evolutivo a los procedimientos estocásticos de búsqueda basado en el principio de evolución. Dicho principio tiene la finalidad última de la supervivencia del más apto y su forma de conseguirlo es por adaptación al entorno. Concretamente, al implementar un algoritmo a una población de individuos que representan un conjunto de candidatos a soluciones de un problema, esta es sometida a una serie de transformaciones con las que actualiza la búsqueda y después a un proceso de selección que favorece a los mejores individuos. Cada ciclo de transformación más la selección, constituye una generación.

Para simular el proceso de evolución, un AE requiere de:

- Una población de posibles soluciones debidamente representadas a través de individuos.
- Un procedimiento de selección basado en la aptitud de los individuos.
- Un procedimiento de transformación, esto es, de construcción de nuevas soluciones a partir de las disponibles actualmente.

Los algoritmos genéticos (AG) que se han desarrollado bajo el esquema evolutivo, hacen evolucionar una población inicial que esta codificada como enteros binarios, sometiéndola a transformaciones unitarias y binarias genéricas, dentro de un proceso de selección. Así dentro de la implantación de un algoritmo genético es crucial definir el proceso de selección, el proceso de reproducción que hacen uso de los operadores genéticos.

4.2 El Algoritmo Genético Simple

El algoritmo comienza con un Begin que indica la generación de una población inicial, esta puede ser generada aleatoriamente. Posteriormente se evalúa con la función a los miembros de la población inicial. Se eligen dos elementos de esa población y se cruza su material genético. Al generarse dos individuos, se computa en forma aleatoria de alguno de los dos nuevos individuos. Después se evalúan los dos nuevos descendientes y se insertan en la población inicial.

```
BEGIN /* Algoritmo Genético Simple */
  Generar una población inicial.
  Computar la función de evaluación de cada individuo.
  WHILE NOT Terminado DO
    BEGIN /* Producir nueva generación */
      FOR Tamaño población/2 DO
        BEGIN /*Ciclo Reproductivo */
          Seleccionar dos individuos de la anterior generación, para el
          cruce (probabilidad de selección proporcional a la función de
          evaluación del individuo).
          Cruzar con cierta probabilidad los dos individuos obteniendo
          dos descendientes.
          Mutar los dos descendientes con cierta probabilidad.
          Computar la función de evaluación de los dos
          Descendientes mutados.
          Insertar los dos descendientes mutados en la nueva
          generación.
        END
      END
    END
  IF la población ha convergido THEN
    Terminado:= TRUE
```

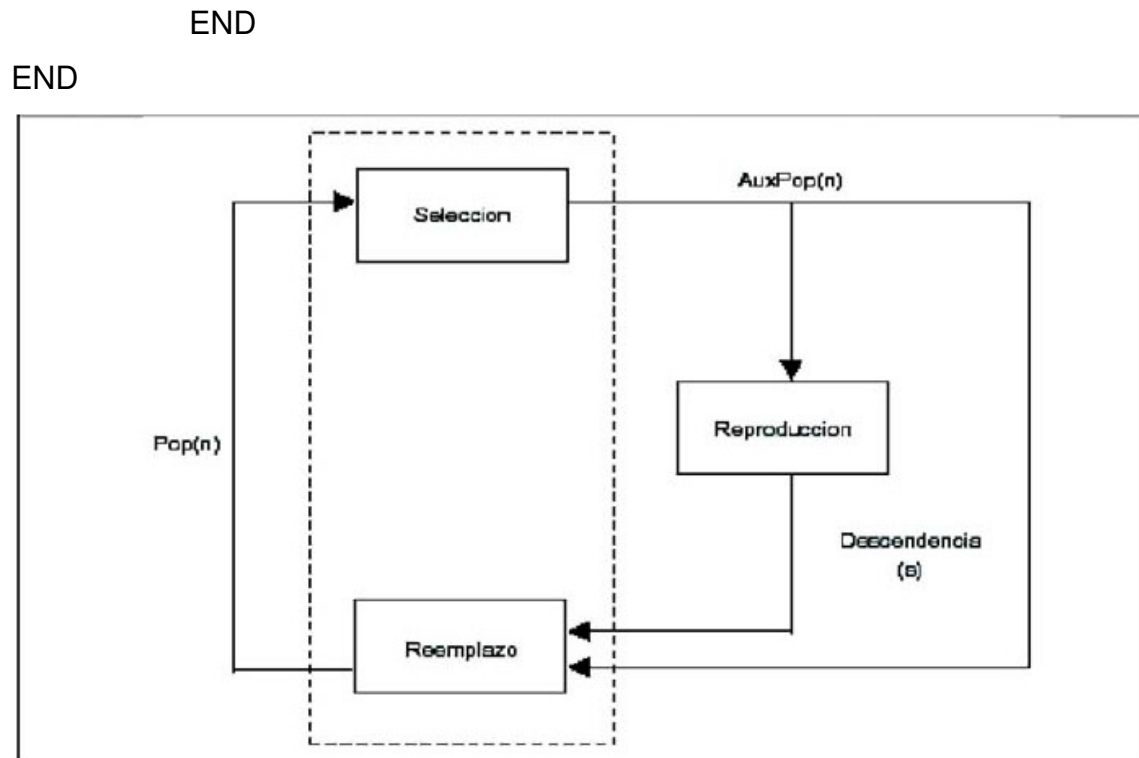



Figura 2. Bucle básico de un algoritmo genético (AG)

Como se verá a continuación, se necesita una codificación o representación del problema, que resulte adecuada al mismo. Además se requiere una función de ajuste o adaptación al problema, la cual asigna un número real a cada posible solución codificada.

Durante la ejecución del algoritmo, los padres deben ser seleccionados para la reproducción bajo un proceso de evaluación usando una función de adaptabilidad que refleje su nivel de calidad; a continuación dichos padres seleccionados se cruzaran generando dos hijos, sobre cada uno de los cuales actuara un operador de mutación, que seleccionará un gen en forma aleatoria y lo mutará.

El resultado de la combinación de las anteriores funciones (de mutación y adaptabilidad) será un conjunto de individuos (posibles soluciones al problema), los cuales en la evolución del Algoritmo Genético formaran parte de la siguiente población.

Codificación

Se supone que los individuos (posibles soluciones del problema), pueden representarse como un conjunto de parámetros (que denominaremos genes), los cuales agrupados forman una ristra de valores (a menudo referida como cromosoma). Si bien el alfabeto utilizado para representar los individuos no debe necesariamente estar constituido por el

{0; 1}, buena parte de la teoría en la que se fundamentan los Algoritmos Genéticos utiliza dicho alfabeto.

En términos biológicos, el conjunto de parámetros representando un cromosoma particular se denomina fenotipo. El fenotipo contiene la información requerida para construir un organismo, el cual se refiere como genotipo. Los mismos términos se utilizan en el campo de los Algoritmos Genéticos. La adaptación al problema de un individuo depende de la evaluación del genotipo.

Esta última puede inferirse a partir del fenotipo, es decir puede ser computada a partir del cromosoma, usando la función de evaluación. La función de adaptación debe ser diseñada para cada problema de manera específica. Dado un cromosoma particular, la función de adaptación le asigna un número real, que se supone refleja el nivel de adaptación al problema del individuo representado por el cromosoma. Durante la fase reproductiva se seleccionan los individuos de la población para cruzarse y producir descendientes, que constituirán, una vez mutados, la siguiente generación de individuos. La selección de padres se efectúa al azar usando un procedimiento que favorezca a los individuos mejor Adaptados, ya que a cada individuo se le asigna una probabilidad de ser seleccionado que es proporcional a su función de adaptación. Este procedimiento se dice que esta basado en la ruleta sesgada. Según dicho esquema, los individuos bien adaptados se escogerán probablemente varias veces por generación, mientras que los pobremente adaptados al problema, no se escogerán más que de vez en cuando. Una vez seleccionados dos padres, sus cromosomas se combinan, utilizando habitualmente los operadores de cruce y mutación.

Las formas básicas de dichos operadores se describen a continuación. El operador de cruce, coge dos padres seleccionados y corta sus ristas de cromosomas en una posición escogida al azar, para producir dos subristras iniciales y dos subristras finales. Después se intercambian las subristras finales, produciéndose dos nuevos cromosomas completos. Ambos descendientes heredan genes de cada uno de los padres. Este operador se conoce como operador de cruce basado en un punto.

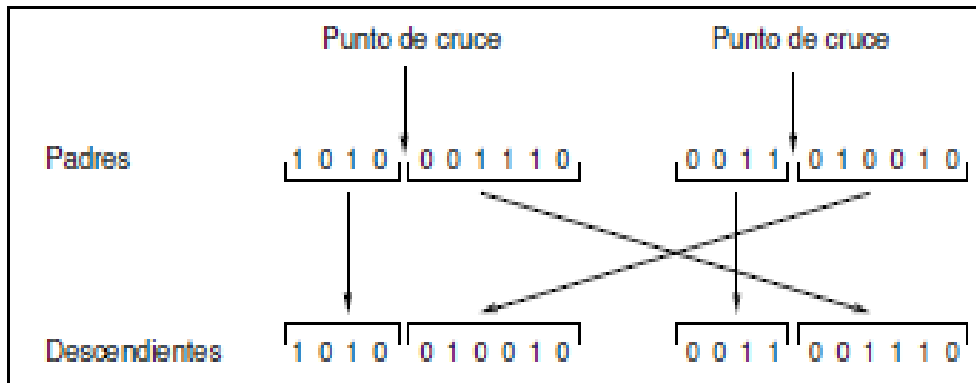


Figura 3. Operador de cruce basado en un punto

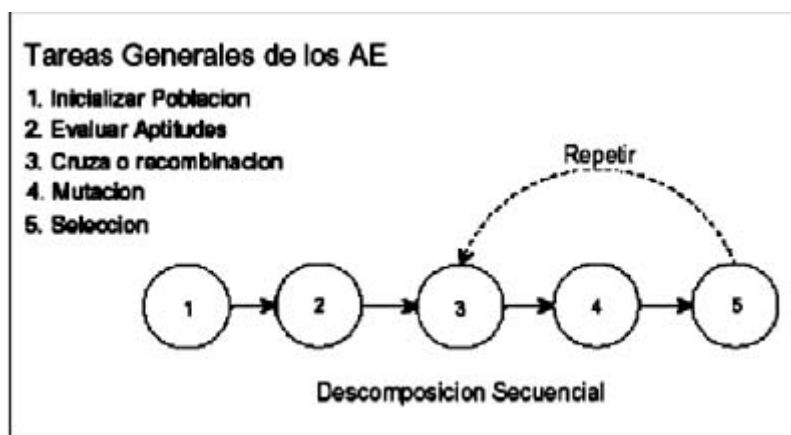


Figura 4. Descomposición generalizada de las tareas de los AE. (AG)

Habitualmente el operador de cruce no se aplica a todos los pares de individuos que han sido seleccionados para emparejarse, sino que se aplica de manera aleatoria, normalmente con una probabilidad comprendida entre 0.5 y 1.0. En el caso en que el operador de cruce no se aplique, la descendencia se obtiene simplemente duplicando los padres. El operador de mutación se aplica a cada hijo de manera individual, y consiste en la alteración aleatoria (normalmente con probabilidad pequeña) de cada gen componente del cromosoma. La figura siguiente muestra la mutación del quinto gen del cromosoma. Si bien puede en principio pensarse



Figura 5. Operador de mutación

que el operador de cruce es más importante que el operador de mutación, ya que proporciona una exploración rápida del espacio de búsqueda, éste último asegura que ningún punto del espacio de búsqueda tenga probabilidad cero de ser examinado, y es de capital importancia para asegurar la convergencia de los Algoritmos Genéticos.

Para criterios prácticos, es muy útil la definición de convergencia introducida en este campo por De Jong (1975) en su tesis doctoral. Si el Algoritmo Genético ha sido correctamente implementado, la población evolucionará a lo largo de las generaciones sucesivas de tal manera que la adaptación media extendida a todos los individuos de la población, así como la adaptación del mejor individuo se irán incrementando hacia el óptimo global.

El concepto de convergencia está relacionado con la progresión hacia la uniformidad: un gen ha convergido cuando al menos el 95 % de los individuos de la población comparten el mismo valor para dicho gen. Se dice que la población converge cuando todos los genes han convergido. Se puede generalizar dicha definición al caso en que al menos un $\beta\%$ de los individuos de la población hayan convergido. La figura siguiente muestra como varía la adaptación media y la mejor adaptación en un Algoritmo Simple Genético Simple típico.

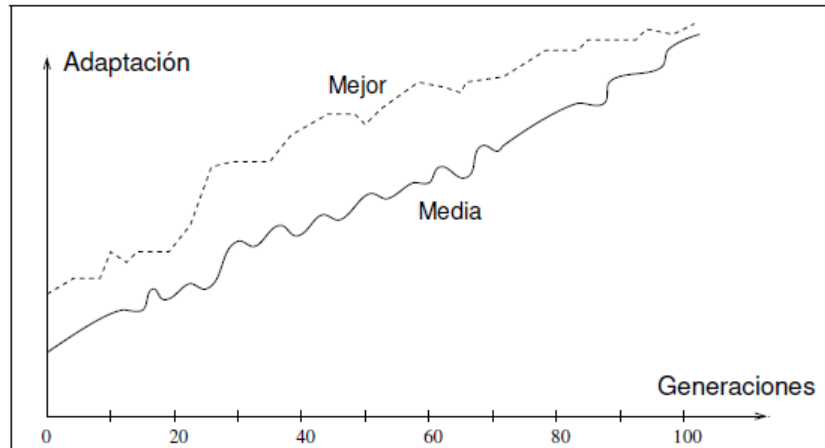


Figura 6. Adaptación media y mejor adaptación en un Algoritmo Genético

A medida que el número de generaciones aumenta, es más probable que la adaptación media se aproxime a la del mejor individuo.

4.3 Extensiones y Modificaciones del Algoritmo Genético Simple

En este apartado se introducirán algunas extensiones y modificaciones del Algoritmo Genético Simple. Se comenzará dando un pseudocódigo para un Algoritmo Genético Abstracto, para a continuación formalizar matemáticamente cada uno de los elementos integrantes del Algoritmo, así como las extensiones y modificaciones que se vayan presentando.

BEGIN AGA

Obtener la población inicial al azar.

WHILE NOT stop DO

BEGIN

Seleccionar padres de la población.

Producir hijos a partir de los padres seleccionados.

Mutar los individuos hijos.

Extender la población añadiendo los hijos.

Reducir la población extendida.

END

END AGA

Figura: Pseudocódigo del Algoritmo Genético Abstracto

Este algoritmo inicia con una población aleatoria finita, de la cual se extraen dos individuos que figurarán como padres progenitores y que a la vez producirán un par de hijos que contendrán su material genético, uno de los progenitores tendrá mejor material genético para que al combinarlo con su cónyuge, se produzca una mejora a la descendencia. Después de engendrar a los hijos, el material genético de uno de los hijos o ambos se mutará en forma aleatoria. Al incorporarse los nuevos hijos a la población se procederá a eliminar los peores elementos de la población inicial, con lo cual se ira mejorando la especie.

Ya que el principal dominio de aplicación de los Algoritmos Genéticos lo constituye la optimización de funciones, se introducen algunos conceptos básicos que se usarán a lo largo de este tema. Dado un dominio Finito D y una función $f: D \rightarrow \mathbb{R}$, el problema de la optimización de la función f se refiere a encontrar el mejor valor de la función f en el dominio D . Se trata por tanto de encontrar $x \in D$ para el cual $f(x) \leq f(y) \forall y \in D$. Ya que $\max\{f(x)\} = -\{\min\{-f(x)\}$ la restricción al problema de minimización no supone ninguna pérdida de generalización. En general, la tarea de optimización se complica debido a la existencia de óptimos locales (mínimos locales en nuestro caso). La función f tiene un mínimo local en $\hat{x} \in D$ si se verifica la siguiente condición:

$\exists E(x)$, entorno de x , tal que si $y \in E(x)$, $\Rightarrow f(x) \leq f(y)$.

Diremos que $e: D \rightarrow S^l$ donde $l \leq \log \frac{|S|}{|D|}$ constituye una codificación, siendo la finalidad de la misma el representar los elementos de D por medio de ristra de elementos de S . A S se le denomina alfabeto, mientras que S^l constituye el espacio de búsqueda. A la función $f(x) = g(e(x))$ se le denomina función objetivo. Es necesario que la función e sea inyectiva, para que los elementos de D sean discernibles. El AGA examinará un subconjunto del espacio de búsqueda, obteniendo una ristra x^* , cuya función objetivo $g(x^*)$ puede considerarse un estimador del $\min_{x \in S^l} g(x)$. Abusando del lenguaje de notación, designaremos por l a los elementos de S^l . En lo que sigue se considerará un AGA como una 10-tupla:

$AGA = (P_0; \lambda; l; f_{sel}; f_{prod}; f_{mut}; f_{ext}; f_{red}; g; ct);$

donde:

- $P_0 = \{I_0^1; \dots; I_0^l\} \in (S^l)^\lambda$ población inicial,
- λ tamaño población,
- l longitud de la representación,
- f_{sel} función de selección,

Fprod	función de producción de hijos,
fmut	función de mutación,
fext	función de extensión,
fred	función de reducción,
g	función objetivo,
ct	criterio de parada.

En principio restringiremos nuestro AGA, imponiéndole la condición de que todas las poblaciones tengan el mismo tamaño. Obviamente una generalización sería el considerar que el tamaño depende de la generación, es decir $\lambda_t = |P_t|$. Denotaremos por P_λ el conjunto de poblaciones de tamaño λ , a las que denominaremos poblaciones bien dimensionadas.

La función de selección global, fsel, selecciona al azar y con reemplazamiento una colección de individuos $y \in P_\lambda$ a partir de una población $x \in P_\lambda$:

$$f_{sel} : (\alpha; x) \rightarrow y;$$

donde α es un vector de dimensión λ constituido por valores escogidos aleatoriamente.

La función de reproducción global, fprod, produce una población de descendientes $z \in P_\lambda$ a partir de individuos seleccionados $y \in P_\lambda$ por medio de un operador de cruce:

$$f_{prod} : (\beta; y) \rightarrow z;$$

donde β es un vector de dimensión $\lambda/2$ de valores escogidos al azar entre los enteros $\{1; \dots; l-1\}$. La función de reproducción se dice que está basada en un punto si los padres $l_i = (s_1; \dots; s_l)$ y $l_j = (b_1; \dots; b_l)$ producen hijos $CH_{i;j;1} = (c_1; \dots; c_l)$ y $CH_{i;j;2} = (d_1; \dots; d_l)$ verificándose:

$$c_j = \begin{cases} s_j & \text{si } j \leq m \\ b_j & \text{si } j > m \end{cases} \quad d_j = \begin{cases} b_j & \text{si } j \leq m \\ s_j & \text{si } j > m \end{cases} \quad (4.1)$$

Donde m es un número entero escogido al azar según una distribución uniforme discreta definida sobre el conjunto $\{1; \dots; l-1\}$. La función de mutación individual find-mut, aplicada a $l = (s_1; \dots; s_l)$, genera otro individuo $MI = (sm_1; \dots; sm_l)$, es decir $find\text{-}mut(l) = MI$, tal que $\forall j \in \{1; \dots; l\}$, $P(sm_j = s_j) = 1 - p_m$, donde p_m es la probabilidad de mutación. La función de extensión, fext, crea a partir de dos poblaciones $x, z \in P_\lambda$, una población $n \in P_{2\lambda}$:

$f_{ext} : (x, z) \rightarrow n$

Denotando por N_i con $i = 1; \dots; 2\lambda$ el i -ésimo individuo en n , por X_k , con $k = 1; \dots; \lambda$ el k -ésimo individuo en x , y por Z_j con $j = 1; \dots; \lambda$ el j -ésimo individuo en z , se tendría:

$$N_i = \begin{cases} X_i & \text{si } i \leq \lambda \\ Z_{i-\lambda} & \text{si } i > \lambda \end{cases} \quad (4.2)$$

La función de reducción global, $fred$, convierte una población $n \in P_{2\lambda}$ en una población $r \in P_\lambda$

$fred : n \rightarrow r$:

Nótese que r denota la población de individuos en el tiempo $t + 1$.

La función de reducción es elitista de grado λ si la población en el tiempo $t + 1$ está constituida por los mejores λ individuos, de entre los λ individuos que constituyen la población en el tiempo t y los descendientes derivados de ellos.

La función de reducción se denomina simple si la población en el tiempo $t+1$ está formada por los descendientes derivados de la población en el tiempo t .

4.3.1 Población

Tamaño de la población

Una cuestión que uno puede plantearse es la relacionada con el tamaño idóneo de la población. Parece intuitivo que las poblaciones pequeñas corren el riesgo de no cubrir adecuadamente el espacio de búsqueda, mientras que el trabajar con poblaciones de gran tamaño puede acarrear problemas relacionados con el excesivo costo computacional.

Goldberg (1989) efectuó un estudio teórico, obteniendo como conclusión que el tamaño óptimo de la población para ristas de longitud l , con codificación binaria, crece exponencialmente con el tamaño de la rista.

Este resultado traería como consecuencia que la aplicabilidad de los Algoritmos Genéticos en problemas reales sería muy limitada, ya que resultarían no competitivos con otros métodos de optimización combinatoria. Alander (1992), basándose en evidencia empírica sugiere que un tamaño de población comprendida entre l y $2l$ es suficiente para atacar con éxito los problemas por él considerados.

Población inicial

Habitualmente la población inicial se escoge generando ristas al azar, pudiendo contener cada gen uno de los posibles valores del alfabeto con probabilidad uniforme. Nos podríamos preguntar que es lo que sucedería si los individuos de la población inicial se obtuviesen como resultado de alguna técnica heurística o de optimización local.

Realmente hay pocos trabajos que existen sobre este aspecto, se constata que esta inicialización no aleatoria de la población inicial, puede acelerar la convergencia del Algoritmo Genético. Sin embargo en algunos casos la desventaja resulta ser la prematura convergencia del algoritmo, queriendo indicar con esto la convergencia hacia óptimos locales.

4.3.2 Función Objetivo

Dos aspectos que resultan cruciales en el comportamiento de los Algoritmos Genéticos son la determinación de una adecuada función de adaptación o función objetivo, así como la codificación utilizada.

Idealmente nos interesaría construir funciones objetivo con "ciertas regularidades", es decir funciones objetivo que verifiquen que para dos individuos que se encuentren cercanos en el espacio de búsqueda, sus respectivos valores en las funciones objetivo sean similares. Por otra parte una dificultad en el comportamiento del Algoritmo Genético puede ser la existencia de gran cantidad de óptimos locales, así como el hecho de que el óptimo global se encuentre muy aislado.

La regla general para construir una buena función objetivo es que ésta debe reflejar el valor del individuo de una manera "real", pero en muchos problemas de optimización combinatoria, donde existen gran cantidad de restricciones, buena parte de los puntos del espacio de búsqueda representan individuos no válidos. Para este planteamiento en el que los individuos están sometidos a restricciones, se han propuesto varias soluciones.

La primera sería la que podríamos denominar absolutista, en la que aquellos individuos que no verifican las restricciones, no son considerados como tales, y se siguen efectuando cruces y mutaciones hasta obtener individuos válidos, o bien a dichos individuos se les asigna una función objetivo igual a cero.

Otra posibilidad consiste en reconstruir aquellos individuos que no verifican las restricciones. Dicha reconstrucción suele llevarse a cabo por medio de un nuevo

operador que se acostumbra a denominar reparador. Otro enfoque está basado en la penalización de la función objetivo. La idea general consiste en dividir la función objetivo del individuo por una cantidad (la penalización) que guarda relación con las restricciones que dicho individuo viola.

Dicha cantidad puede simplemente tener en cuenta el número de restricciones violadas ó bien el denominado costo esperado de reconstrucción, es decir el coste asociado a la conversión de dicho individuo en otro que no viole ninguna restricción. Otra técnica que se ha venido utilizando en el caso en que la computación de la función objetivo sea muy compleja es la denominada evaluación aproximada de la función objetivo.

En algunos casos la obtención de n funciones objetivo aproximadas puede resultar mejor que la evaluación exacta de una única función objetivo (supuesto el caso de que la evaluación aproximada resulta como mínimo n veces más rápida que la evaluación exacta). Un problema habitual en las ejecuciones de los Algoritmos Genéticos surge debido a la velocidad con la que el algoritmo converge.

En algunos casos la convergencia es muy rápida, lo que suele denominarse convergencia prematura, en la cual el algoritmo converge hacia óptimos locales, mientras que en otros casos el problema es justo el contrario, es decir se produce una convergencia lenta del algoritmo.

Una posible solución a estos problemas pasa por efectuar transformaciones en la función objetivo. El problema de la convergencia prematura, surge a menudo cuando la selección de individuos se realiza de manera proporcional a su función objetivo.

En tal caso, pueden existir individuos con una adaptación al problema muy superior al resto, que a medida que avanza el algoritmo "dominan" a la población. Por medio de una transformación de la función objetivo, en este caso una comprensión del rango de variación de la función objetivo, se pretende que dichos "superindividuos" no lleguen a dominar a la población. El problema de la lenta convergencia del algoritmo, se resolvería de manera análoga, pero en este caso efectuando una expansión del rango de la función objetivo.

La idea de especies de organismos, ha sido imitada en el diseño de los Algoritmos Genéticos en un método propuesto por Goldberg y Richardson (1987), utilizando una *modificación de la función objetivo* de cada individuo, de tal manera que individuos que estén muy cercanos entre si devalúen su función objetivo, con objeto de que la población gane en diversidad. Por ejemplo, si denotamos por $d(I_t^j, I_t^i)$ a la distancia de

Hamming entre los individuos I_t^j e I_t^i y por $K \in \mathbb{R}^+$ a un parámetro, podemos definir la siguiente función:

$$h(d(I_t^j, I_t^i)) = \begin{cases} K-d(I_t^j, I_t^i) & \text{si } d(I_t^j, I_t^i) < K, \\ 0 & \text{si } d(I_t^j, I_t^i) > K \end{cases} \quad (4.3)$$

A continuación para cada individuo I_t^j , definimos $\sigma_j^t = \sum_{i \neq j} h(d(I_t^j, I_t^i))$, valor que utilizaremos para devaluar la función objetivo del individuo en cuestión. Es decir, $g^*(I_t^j) = g(I_t^j)/\sigma_j^t$. De esta manera aquellos individuos que están cercanos entre si verían devaluada la probabilidad de ser seleccionados como padres, aumentándose la probabilidad de los individuos que se encuentran más aislados.

4.3.3 Selección

La función de selección de padres más utilizada es la denominada *función de selección proporcional a la función objetivo*, en la cual cada individuo tiene una probabilidad de ser seleccionado como padre que es proporcional al valor de su función objetivo.

Denotando por $p_{j,t}^{\text{prop}}$ la probabilidad de que el individuo I_t^j sea seleccionado como padre, se tiene que:

$$p_{j,t}^{\text{prop}} = \frac{g(I_t^j)}{\sum_{j=1}^{\lambda} g(I_t^j)} \quad (4.4)$$

Esta función de selección es invariante ante un cambio de escala, pero no ante una traslación.

Una de las maneras de superar el problema relacionado con la rápida convergencia proveniente de los superindividuos, que surge al aplicar la anterior función de selección, es el efectuar la selección proporcional al rango del individuo, con lo cual se produce una repartición más uniforme de la probabilidad de selección, tal y como se ilustra en la figura siguiente.

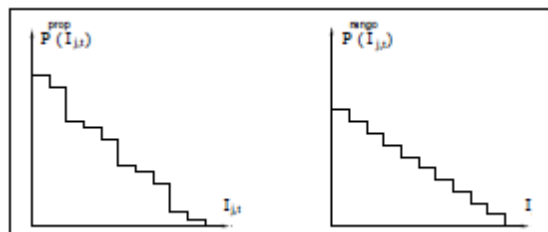


Figura 7. Gráfica de la repartición más uniforme de la probabilidad de selección

Esquemas de selección de padres proporcional a la función objetivo (izquierda) y proporcional al rango de la función objetivo (derecha)

Si denotamos por $\text{rango}(g(I_t^j))$ el rango de la función objetivo del individuo I_t^j cuando los individuos de la población han sido ordenados de menor a mayor (es decir el peor individuo tiene rango 1, mientras que el individuo con mejor función objetivo tiene rango λ), y sea $p_{j,t}^{\text{rango}}$ la probabilidad de que el individuo I_t^j sea seleccionado como padre cuando la selección se efectúa proporcionalmente al rango del individuo, se tiene que

$$p_{j,t}^{\text{rango}} = \frac{\text{rango}(g(I_t^j))}{\lambda(\lambda+1)/2} \quad (4.5)$$

La suma de los rangos, $\lambda(\lambda+1)/2$, constituye la constante de normalización. La función de selección basada en el rango es invariante frente a la translación y al cambio de escala. Otro posible refinamiento del modelo de selección proporcional, es el modelo de selección del valor esperado, el cual actúa de la manera siguiente: para cada individuo I_t^j , se introduce un contador, inicializado en $g(I_t^j)/\bar{g}_t$, donde \bar{g}_t , denota la media de la función objetivo en la generación t . Cada vez que el individuo I_t^j es seleccionado para el cruce, dicho contador decrece en una cantidad c ($c \in (0; 5; 1)$): El individuo en cuestión dejará de poder ser seleccionado en esa generación, cuando su contador sea negativo. Baker (1987) introduce un método denominado muestreo universal estocástico, el cual utiliza un único giro de la ruleta siendo los sectores circulares proporcionales a la función objetivo. Los individuos son seleccionados a partir de marcadores (véase siguiente), igualmente espaciados y con comienzo aleatorio.

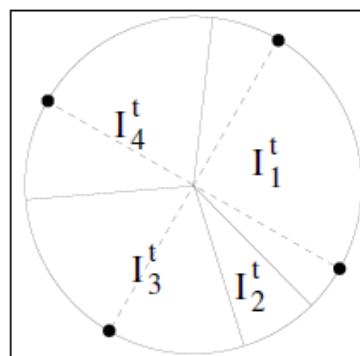


Figura 8. Método de selección de padres denominado muestreo universal estocástico. El individuo I_1^t se escoge 2 veces, mientras que I_3^t e I_4^t son elegidos una única vez

Efectuando un paralelismo con los métodos de muestreo estadísticos, este último tipo de selección de padres se relaciona con el muestreo sistemático, mientras que la selección proporcional a la función objetivo, está basada en el muestreo estratificado

con afijación proporcional al tamaño. También el procedimiento de selección que hemos denominado muestreo estocástico con reemplazamiento del resto, mantiene un paralelismo con el muestreo estratificado con afijación de compromiso. En el modelo de selección elitista se fuerza a que el mejor individuo de la población en el tiempo t , sea seleccionado como padre. La selección por torneo, constituye un procedimiento de selección de padres muy extendido y en el cual la idea consiste en escoger al azar un número de individuos de la población, tamaño del torneo, (con o sin reemplazamiento), seleccionar el mejor individuo de este grupo, y repetir el proceso hasta que el número de individuos seleccionados coincida con el tamaño de la población. Habitualmente el tamaño del torneo es 2, y en tal caso se ha utilizado una versión probabilística en la cual se permite la selección de individuos sin que necesariamente sean los mejores.

Una posible clasificación de procedimientos de selección de padres consistiría en: métodos de selección dinámicos, en los cuales las probabilidades de selección varían de generación a generación, (por ejemplo la selección proporcional a la función objetivo), frente a métodos de selección estáticos, en los cuales dichas probabilidades permanecen constantes (por ejemplo la selección basada en rangos).

Si se asegura que todos los individuos tienen asignada una probabilidad de selección distinta de cero el método de selección se denomina preservativo. En caso contrario se acostumbra a denominarlo extintivo.

4.3.4 Cruce

El Algoritmo Genético Canónico descrito anteriormente, utiliza el cruce basado en un punto, en el cual los dos individuos seleccionados para jugar el papel de padres, son recombinados por medio de la selección de un punto de corte, para posteriormente intercambiar las secciones que se encuentran a la derecha de dicho punto.

Se han investigado otros operadores de cruce, habitualmente teniendo en cuenta más de un punto de cruce. De Jong (1975) investigó el comportamiento del operador de cruce basado en múltiples puntos, concluyendo que el cruce basado en dos puntos, representaba una mejora mientras que añadir más puntos de cruce no beneficiaba el comportamiento del algoritmo.

La ventaja de tener más de un punto de cruce radica en que el espacio de búsqueda puede ser explorado más fácilmente, siendo la principal desventaja el hecho de

aumentar la probabilidad de ruptura de buenos esquemas. En el operador de cruce basado en dos puntos, los cromosomas (individuos) pueden contemplarse como un circuito en el cual se efectúa la selección aleatoria de dos puntos, tal y como se indica en la Figura de abajo.

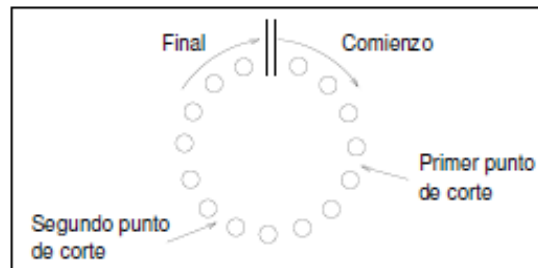


Figura 9. Operador de cruce basado en dos puntos, los cromosomas (individuos) pueden contemplarse como un circuito en el cual se efectúa la selección aleatoria de dos puntos

Desde este punto de vista, el cruce basado en un punto, puede verse como un caso particular del cruce basado en dos puntos, en el cual uno de los puntos de corte se encuentra fijo al comienzo de la ristra que representa al individuo.

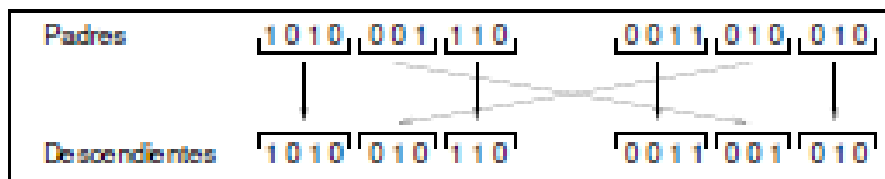


Figura 10. Operador de cruce basado en dos puntos

En el denominado operador de cruce uniforme (Syswerda (1991)) cada gen en la descendencia se crea copiando el correspondiente gen de uno de los dos padres, escogido de acuerdo a una "máscara de cruce" generada aleatoriamente. Cuando existe un 1 en la "máscara de cruce", el gen es copiado del primer padre, mientras que cuando exista un 0 en la máscara, el gen se copia del segundo padre, tal y como se muestra en la Figura siguiente.

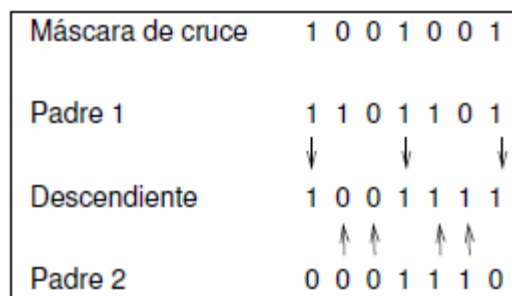


Figura 11. Operador de cruce uniforme

En la literatura, el término operador de cruce uniforme se relaciona con la obtención de la "máscara de cruce" uniforme, en el sentido de que cualquiera de los elementos del alfabeto tenga asociada la misma probabilidad. Hablando en términos de la teoría de la probabilidad la máscara de cruce está compuesta por una muestra aleatoria de tamaño λ extraída de una distribución de probabilidad de Bernouilli de parámetro $1/2$.

Si tuviésemos en cuenta el valor de la función de adaptación de cada padre en el momento de generar la "máscara de cruce", de tal manera que cuanto mayor sea la función de adaptación de un individuo, más probable sea heredar sus características, podríamos definir, véase Larrañaga y Poza (1994), un operador de cruce basado en la función objetivo, en el cual la "máscara de cruce" se interpreta como una muestra aleatoria de tamaño l proveniente de una distribución de Bernouilli de parámetro

$$p = \frac{g(I_t^j)}{g(I_t^j) + g(I_t^i)} \quad (4.6)$$

Donde I_t^j y I_t^i denotan los padres seleccionados para ser cruzados. El concepto de "máscara de cruce" puede también servir para representar los cruces basados en un punto y basados en múltiples puntos, tal y como se muestra en Figura siguiente

Máscara de cruce	1 1 1 0 0 0 0	1 1 0 0 0 1
Padre 1	1 0 1 1 0 0 1	1 0 1 1 0 0
Descendiente	1 0 1 0 1 1 1	1 0 0 0 1 0
Padre 2	1 0 0 0 1 1 1	1 0 0 0 1 1

Figura 12. "Máscaras de cruce" para los operadores de cruce basados en 1 punto y en 2 puntos

Sirag y Weiser (1987), modifican el operador de cruce en el sentido del Simulated Annealing. De esta manera el operador de cruce se modifica definiendo un umbral de energía θ_c , y una temperatura T , las cuales influyen en la manera en la que se escogen los bits individuales. Según el operador propuesto el bit $(i + 1)$ -ésimo se toma del padre opuesto al que se ha tomado el bit i -ésimo, con probabilidad $\exp(-\theta_c/T)$, donde T es el parámetro "temperatura" el cual, al igual que en Simulated Annealing decrecerá lentamente por medio de un programa de enfriamiento.

Con altas temperaturas el comportamiento se asemeja al del operador de cruce uniforme, es decir con probabilidad cercana a la unidad los bits se van escogiendo alternativamente de cada padre. Por otra parte cuando el valor del parámetro temperatura se acerca a cero el hijo resultante coincide prácticamente con uno de los padres. Existen otros operadores de cruce específicos para un determinado problema como son, por ejemplo, los definidos para el TSP, que se tratarán más adelante.

Por otra parte, la idea de que el cruce debería de ser más probable en algunas posiciones ha sido descrita por varios autores (Schaffer y Morishima, 1987; Holland, 1975; Davis, 1985; Levenick, 1991).

4.3.5 Mutación

La mutación se considera un operador básico, que proporciona un pequeño elemento de aleatoriedad en la vecindad (entorno) de los individuos de la población. Si bien se admite que el operador de cruce es el responsable de efectuar la búsqueda a lo largo del espacio de posibles soluciones, también parece desprenderse de los experimentos efectuados por varios investigadores que el operador de mutación va ganando en importancia a medida que la población de individuos va convergiendo¹⁶.

¹⁶ Davis, 1985

Schaffer y col. (1989) encuentran que el efecto del cruce en la búsqueda es inferior al que previamente se esperaba. Utilizan la denominada evolución primitiva, en la cual, el proceso evolutivo consta tan sólo de selección y mutación. Encuentran que dicha evolución primitiva supera con creces a una evolución basada exclusivamente en la selección y el cruce. Otra conclusión de su trabajo es que la determinación del valor óptimo de la probabilidad de mutación es mucho más crucial que el relativo a la probabilidad de cruce.

La búsqueda del valor óptimo para la probabilidad de mutación, es una cuestión que ha sido motivo de varios trabajos. Así, De Jong (1975) recomienda la utilización de una probabilidad de mutación del bit de $1/l$, siendo l la longitud del string. Schaffer y col. (1989) utilizan resultados experimentales para estimar la tasa óptima proporcional a $1/\lambda 0.931810.4535$, donde λ denota el número de individuos en la población. Si bien en la mayoría de las implementaciones de Algoritmos Genéticos se asume que tanto la probabilidad de cruce como la de mutación permanecen constantes, algunos autores han obtenido mejores resultados experimentales modificando la probabilidad de mutación a medida que aumenta el número de iteraciones. Pueden consultarse los trabajos de Ackley (1987), Bramlette (1991), Fogarty (1989) y Michalewicz y Janikow (1991).

4.3.6 Reducción

Una vez obtenidos los individuos descendientes de una determinada población en el tiempo t , el proceso de reducción al tamaño original, consiste en escoger λ individuos de entre los λ individuos que forman parte de la población en el tiempo t , y los λ individuos descendientes de los mismos. Dicho proceso se suele hacer fundamentalmente de dos formas distintas.

O bien los λ individuos descendientes son los que forman parte de la población en el tiempo $t+1$, es lo que se denomina reducción simple, o bien se escogen de entre los 2λ individuos, los λ individuos más adaptados al problema, siguiendo lo que podemos denominar un criterio de reducción elitista de grado λ .

Podemos también considerar otros procedimientos de reducción que se colocan entre los anteriores, por ejemplo, si escogemos los λ_1 mejores de entre padres y descendientes, escogiéndose los $\lambda - \lambda_1$ restantes de entre los descendientes no seleccionados hasta el momento. El concepto de reducción está ligado con el de tasa de

reemplazamiento generacional, trg , es decir en el porcentaje de hijos generados con respecto del tamaño de la población.

Si bien en la idea primitiva de Holland (1975) dicho reemplazamiento se efectuaba de 1 en 1, es decir $trg = \lambda - 1$, habitualmente dicho reemplazamiento se efectúa en bloque, $trg = 1$. De Jong (1975) introdujo el concepto de tasa de reemplazamiento generacional con el objetivo de efectuar un solapamiento controlado entre padres e hijos. En su trabajo, en cada paso una proporción, trg , de la población es seleccionada para ser cruzada. Los hijos resultantes podrán reemplazar a miembros de la población anterior.

Este tipo de Algoritmos Genéticos se conocen bajo el nombre de SSGA (Steady State Genetic Algorithm), un ejemplo de los cuales lo constituye GENITOR (Whitley y Kauth, 1988; Whitley, 1989). Michalewicz (1992) introduce un algoritmo que denomina Algoritmo Genético Modificado, MODGA, en el cual para llevar a cabo el reemplazamiento generacional, selecciona al azar $r1$ individuos para la reproducción, así como $r2$ individuos (distintos de los anteriores) destinados a morir. Estas selecciones aleatorias tienen en consideración el valor de la función objetivo de cada individuo, de tal manera que cuanto mayor es la función objetivo, mayor es la probabilidad de que sea seleccionado para la reproducción, y menor es la probabilidad de que dicho individuo fallezca. El resto de los $\lambda - (r1 + r2)$ individuos son considerados como neutros y pasan directamente a formar parte de la población en la siguiente generación.

Algoritmos Genéticos Paralelos

En este apartado se introducirán tres maneras diferentes de explotar el paralelismo de los Algoritmos Genéticos, por medio de los denominados modelos de islas. Para una profundización sobre el tema puede consultarse Stender (1993).

- Modelos de islas. La idea básica consiste en dividir la población total en varias sub-poblaciones en cada una de las cuales se lleva a cabo un Algoritmo Genético. Cada cierto número de generaciones, se efectúa un intercambio de información entre las sub-poblaciones, proceso que se denomina emigración. La introducción de la emigración hace que los modelos de islas sean capaces de explotar las diferencias entre las diversas sub-poblaciones, obteniéndose de esta manera una fuente de diversidad genética. Cada sub-población es una "isla", definiéndose un procedimiento por medio del cual se mueve el material genético de una "isla" a otra. La determinación de la tasa de migración, es un asunto de

capital importancia, ya que de ella puede depender la convergencia prematura de la búsqueda. Se pueden distinguir diferentes modelos de islas en función de la comunicación entre las sub-poblaciones. Algunas comunicaciones típicas son las siguientes:

- Comunicación en estrella, en la cual existe una sub-población que es seleccionada como maestra (aquella que tiene mejor media en el valor de la función objetivo), siendo las demás consideradas como esclavas. Todas las sub-poblaciones esclavas mandan sus h_1 mejores individuos ($h_1 \geq 1$) a la sub-población maestra la cual a su vez manda sus h_2 mejores individuos ($h_2 \geq 1$) a cada una de las sub-poblaciones esclavas.

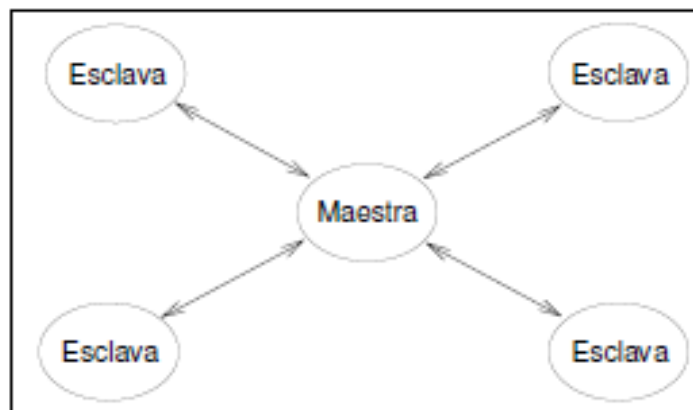


Figura 13. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en estrella

- *Comunicación en red*, en la cual no existe una jerarquía entre las sub-poblaciones, mandando todas y cada una de ellas sus h_3 ($h_3 \geq 1$) mejores individuos al resto de las sub-poblaciones.

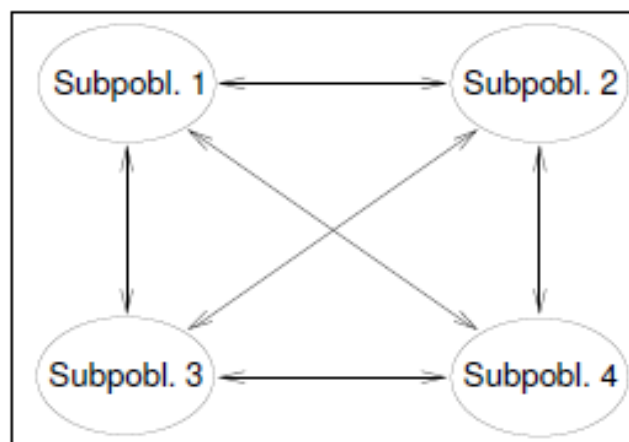


Figura 14. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en red

- *Comunicación en anillo*, en la cual cada sub-población envía sus h_4 mejores individuos ($h_4 \geq 1$), a una población vecina, efectuándose la migración en un único sentido de flujo.

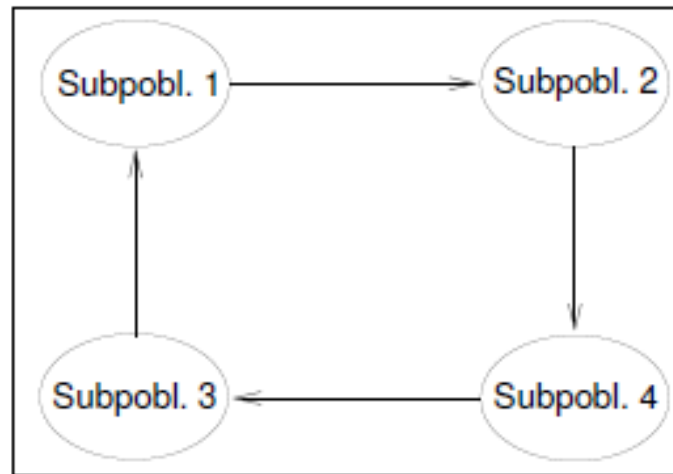


Figura 15. Algoritmo Genético Paralelo. Modelo de Islas. Comunicación en anillo. El modelo de islas ha sido utilizado por varios autores (Whitley y Starkweather, 1990; Gorges-Schleuter, 1989; Tanese, 1987).

4.4 Evaluación de Algoritmos Genéticos

Las tres medidas de evaluación que se tratarán en este apartado, fueron introducidas por De Jong (1975), y se conocen como:

- *evaluación on-line*: mide el comportamiento medio de todas las ristas generadas hasta el tiempo T .
- *evaluación basada en el mejor*: trata de evaluar el Algoritmo Genético por medio del mejor valor de la función evaluación encontrado en la evolución.

Si denotamos por $v_i(t)$ la función objetivo del i -ésimo individuo ($i = 1, \dots, \lambda$) en la t -ésima población, la evaluación on-line, después de T iteraciones, se denota a por $v_{on-line}(T)$, y se define como

$$v_{on-line}(T) = \frac{\sum_{t=1}^T \sum_{i=1}^{\lambda} v_i(t)}{\lambda T} \quad (4.7)$$

Es decir, la evaluación on-line mide el comportamiento medio de todas las ristas generadas hasta el tiempo T . La evaluación on-line se refiere al comportamiento del Algoritmo Genético en su proceso de convergencia hacia el óptimo. Así tendremos que si denotamos por $v^*(t)$ al mejor valor de la función objetivo obtenido hasta el tiempo t (incluyendo dicho tiempo), y si $v_{off-line}(T)$ denota la evaluación off-line después de T generaciones, tenemos que

$$v^{\text{off-line}}(T) = \frac{\sum_{t=1}^T v^*(t)^*}{T} \quad (4.8)$$

La definición de evaluación basada en el mejor trata de evaluar el Algoritmo Genético por medio del mejor valor de la función evaluación encontrado en la evolución. Se trata por tanto de la medida de evaluación usual al tratar de estudiar el comportamiento de cualquier técnica heurística.

4.5 ¿Por qué funcionan los algoritmos genéticos?

En este apartado se tratarán algunas cuestiones relacionadas con el porqué “funcionan” los Algoritmos Genéticos. El significado de la palabra “funcionan” se refiere al hecho de ser capaces de encontrar el óptimo de la función durante el proceso de búsqueda. Se estudiará en primer lugar el denominado Teorema de los esquemas¹⁷, adaptando la notación, para presentarlo en relación con distribuciones binomiales, para a continuación introducir la denominada Hipótesis de bloques, y finalizar referenciando algunos trabajos relativos a la convergencia de los Algoritmos Genéticos.

Teorema de los esquemas

En lo que sigue se desarrollará el denominado Teorema de los esquemas para el caso del Algoritmo Genético Canónico, utilizando además un alfabeto binario. Dicho teorema utiliza el concepto de esquema, que introducimos a continuación. Supongamos un alfabeto binario $S = \{0,1\}$. Para construir un esquema necesitamos ampliar el alfabeto anterior por medio del símbolo *. Un esquema será cualquier ristra de caracteres formada a partir de elementos del alfabeto ampliado S' , siendo $S' = \{0, 1, *\}$. Si la longitud de la ristra es l , el número de esquemas existentes es 3^l , ya que cada posición puede ocuparse por cualquier de los tres elementos del alfabeto extendido S' . Un esquema representa por tanto todas aquellas ristras que se asocian con él, en todas las posiciones distintas de las ocupadas por *. Así por ejemplo considerando ristras de longitud 4, tenemos que el esquema $Q = (*1*0)$ se empareja con los cromosomas (0100), (0110), (1100) y (1110).

¹⁷ Goldberg, 1989

Desde un punto de vista geométrico un esquema equivale a un hiperplano en el espacio de búsqueda. El esquema (0111) representa una única ristra, mientras que el esquema (* * * *) representa todas las ristas de longitud 4. Por otra parte cada esquema se empareja exactamente con 2^r ristas, donde r denota el número de símbolos * contenidos en el esquema. Además cada ristra de longitud l se empareja con 2^l esquemas. Así por ejemplo el ristra (0100) se empareja con los siguientes 24 esquemas: (0100), (*100), (0 * 00), (01 * 0), (010*), (* * 00), (* 1 * 0), (*10*), (0 * * 0), (0 * 0*), (01 * *), (* * * 0), (* * 0 *), (*1 * *), (0 * * *), (* * * *)

Para desarrollar el Teorema de los esquemas necesitamos utilizar los conceptos de orden y longitud de un esquema, que se definen a continuación. El orden del esquema Q , se denota por $o(Q)$, y se refiere al número de posiciones ocupadas por 0-s, ó 1-s dentro del esquema. Se trata por tanto del número de posiciones fijas en el esquema. En otras palabras se trata de la longitud de la ristra menos el número de símbolos *. Por ejemplo, los tres esquemas siguientes,

$$Q1 = (*01*); \quad Q2 = (* * * 1); \quad Q3 = (110*);$$

tienen los siguientes órdenes:

$$o(Q1) = 2; \quad o(Q2) = 1; \quad \text{y} \quad o(Q3) = 3:$$

El concepto de orden del esquema se utiliza, como se verá con posterioridad, para calcular la probabilidad de supervivencia del esquema con relación al operador mutación. La longitud del esquema Q , se denota por medio de $\delta(Q)$, y se define como la distancia entre la primera y las últimas posiciones fijas de la ristra. Dicho concepto define en cierto modo la compactibilidad de la información contenida en el esquema. Así por ejemplo

$$\delta(Q1) = 3 - 2 = 1; \quad \delta(Q2) = 4 - 4 = 0; \quad \text{y} \quad \delta(Q3) = 3 - 1 = 2:$$

El concepto anterior se utiliza para calcular la probabilidad de supervivencia del esquema frente al operador de cruce. Denotaremos por:

λ , el tamaño de la población que se considerará constante a lo largo de la evolución,

p_c , la probabilidad de cruce,

p_m , la probabilidad de mutación.

Por otra parte, $\xi(Q; t)$ denotará el número de ristas que en una determinada población en el tiempo t , se asocian con el esquema Q .

Así por ejemplo si consideramos la población en el tiempo t ,
 $P_t = \{(0100); (1001); (1111); (0001)\}$; y los dos esquemas siguientes
 $Q_1 = (* * 01)$ y $Q_2 = (* * * 0)$;

tenemos que

$$\xi(Q_1; t) = 2 \text{ y } \xi(Q_2; t) = 1:$$

Otra propiedad del esquema es su adaptación en el tiempo t , la cual se denotará por $\text{eval}(Q; t)$ y se define como la media de la función objetivo de todas las ristas que en la población se asocian con el esquema Q . Sea $\overline{F(t)}$, la media de la función objetivo extendida a toda la población en el tiempo t . Es decir $\overline{F(t)} = F(t)/\lambda$, siendo $F(t)$ la suma de las funciones objetivo de todos los individuos en la población en el tiempo t . Con la notación introducida hasta el momento, podemos enunciar el Teorema de los esquemas de la siguiente manera.

Teorema de los esquemas: Si denotamos por $\text{Esel, cru, mut}(\xi(Q; t + 1))$, el número esperado de individuos que se asocian con el esquema Q , en el tiempo $t + 1$, después de efectuar la selección, el cruce y la mutación en un Algoritmo Genético Canónico, se tiene que:

$$\text{Esel, cru, mut}(\xi(Q; t + 1)) \geq \xi(Q; t) \cdot \text{eval}(Q; t) / \overline{F(t)} [1 - pc \frac{\delta(Q)}{l-1} - o(Q)pm]:$$

La demostración se efectuará en tres pasos, en el primero de los cuales se estudiará el efecto de la selección basada en la ruleta sesgada, en el número esperado de individuos que se asocian con la rista Q . Dicho número se denotará por $\text{Esel}(\xi(Q; t+1))$. A continuación se verá el efecto del cruce en dicho número esperado, el cual se denotará por $\text{Esel, cru}(\xi(Q; t + 1))$, para finalmente tratar el efecto de la mutación, obteniéndose una cota inferior para $\text{Esel, cru, mut}(\xi(Q; t + 1))$

Efecto selección

Utilizando la notación anteriormente introducida, se tiene que la probabilidad de que un individuo seleccionado para cruzarse (suponemos que la selección se efectúa de manera proporcional a la función objetivo) se empareje con el esquema Q , se calculará por medio de $\text{eval}(Q; t)/F(t)$. El número esperado de individuos que a partir del anterior individuo seleccionado van a emparejarse con Q , será $(\text{eval}(Q; t)/F(t)) \cdot \xi(Q; t)$: Al efectuarse λ selecciones, se tendrá que:

$$\text{Esel}(\xi(Q; t + 1)) = \lambda \cdot \xi(Q; t) \cdot \text{eval}(Q; t) / F(t):$$

Teniendo en cuenta que $\overline{F(t)} = F(t)/\lambda$, tendremos que:

$$\text{Esel}(\xi(Q; t + 1)) = \xi(Q; t) \cdot \text{eval}(Q; t) / \overline{F(t)}:$$

Si se asume que la evaluación del esquema supera a la evaluación media extendida a toda la población en un $\varepsilon\%$, es decir si

$$\text{eval}(Q; t) = \overline{F(t)} + \varepsilon \overline{F(t)} (\varepsilon > 0),$$

entonces obtendremos que

$$\text{Esel}(\xi(Q; t)) = \xi(Q; 0) \cdot (1 + \varepsilon)^t:$$

Por tanto acabamos de obtener que el número medio esperado de individuos que se asocian con un esquema determinado Q en el tiempo t, crece de forma exponencial en función de la diferencia entre la evaluación media de dicho esquema con respecto a la evaluación media global. La anterior ecuación se conoce bajo el nombre de ecuación de crecimiento reproductivo de los esquemas.

Efecto cruce

Denotando por $pd, \text{cruce}(Q)$, la probabilidad de que el esquema Q sea destruido por el operador de cruce, se tiene que $pd, \text{cruce}(Q) = \frac{\delta(Q)}{l-1}$.

En realidad como el operador de cruce se lleva a cabo con una cierta probabilidad pc, se tiene que $pd, \text{cruce}(Q) = \frac{\delta(Q)}{l-1} \cdot pc$ lo cual equivale a decir que la probabilidad de que el esquema Q sobreviva al cruce es

$$pso, \text{cruce}(Q) = 1 - pd, \text{cruce}(Q) = 1 - \frac{\delta(Q)}{l-1} \cdot pc.$$

En realidad la fórmula anterior nos proporciona una cota para la probabilidad de que sobreviva el esquema Q. Ello es debido a que, puede ocurrir que aunque el punto de corte se encuentre entre los símbolos * comprendidos entre el primer y el último elemento distintos de *, el esquema resultante no muera. Así por ejemplo, si consideramos los dos esquemas: $S1 = 1 * * 0$; $S2 = 1 * * 0$; independientemente del punto de corte, cualquiera de los esquemas resultantes coincide con los anteriores. De ahí que en realidad tenemos que: $pso, \text{cruce}(Q) \geq 1 - \frac{\delta(Q)}{l-1} \cdot pc$

$$\text{obteniéndose que } \text{Esel, cru}[\xi(Q; t + 1)] \geq \xi(Q; t) \cdot \text{eval}(Q; t) / \overline{F(t)} \cdot (1 - \frac{\delta(Q)}{l-1} \cdot pc)$$

Efecto mutación. Denotando por $pso, \text{mut}(Q)$ a la probabilidad de que el esquema Q sobreviva al operador mutación, se tiene que $pso, \text{mut}(Q) = (1 - pm)^o(Q)$. Al ser $pm \ll 1$, la anterior probabilidad de sobrevivir puede ser aproximada por

$$pso, \text{mut} \approx 1 - o(Q)pm:$$

El efecto combinado de la selección, el operador de cruce y el operador de mutación nos proporciona el denominado Teorema de los esquemas. Es decir

$$E_{sel, cru, mut}[\xi(Q; t + 1)] \geq \xi(Q; t) \cdot \frac{eval(Q; t)}{\overline{F}(t)} \cdot \left[1 - \frac{\delta(Q)}{l-1} p_c - o(Q) p_m + \frac{\delta(Q)}{l-1} o(Q) p_c p_m \right]$$

ó bien suprimiendo el último término

$$E_{sel, cru, mut}[\xi(Q; t + 1)] \geq \xi(Q; t) \cdot \frac{eval(Q; t)}{\overline{F}(t)} \left[1 - \frac{\delta(Q)}{l-1} p_c - o(Q) p_m \right]:$$

La fórmula anterior nos proporciona una cota inferior del número esperado de individuos que se asocian con el esquema Q en la siguiente generación, expresado en función del número de individuos que se asocian con el esquema en la actual generación, así como de la adaptación del esquema relativa a la adaptación media de la población, el orden y la longitud del esquema, la longitud de las ristas, así como las probabilidades de cruce y mutación. Tal y como se desprende de la fórmula anterior, esquemas cortos, de bajo orden y con una adaptación al problema superior a la adaptación media, se espera que a medida que evoluciona el Algoritmo Genético, obtengan un incremento exponencial en el número de individuos que se asocian con los mismos.

Paralelismo Implícito

El hecho de que los Algoritmos Genéticos efectúen búsquedas robustas, se debe a que implícitamente muestrean hiperplanos. Es decir cada individuo, representado por medio de una codificación binaria, constituye un vértice del hipercubo que representa el espacio de búsqueda, y es a su vez un miembro de $2^l - 1$ diferentes hiperplanos, donde l es la longitud de la codificación binaria. Por otra parte existen un total de $3^l - 1$ hiperplanos particiones sobre el espacio de búsqueda total. Cada individuo no proporciona gran información si se examina por separado. Es por ello que la búsqueda basada en una población es crítica para los Algoritmos Genéticos. Una población de individuos, proporciona información sobre numerosos hiperplanos, estando los hiperplanos de bajo orden muestrados por numerosos puntos de la población. El hecho de que se muestreen muchos hiperplanos cuando se evalúa una población de ristas, se conoce como el paralelismo implícito o intrínseco de los Algoritmos Genéticos. Dicho paralelismo implícito, significa que muchas competiciones entre hiperplanos se efectúan simultáneamente.

Teoremas de Convergencia

Rudolph (1994) demuestra la no convergencia hacia el óptimo global del algoritmo genético canónico, así como la garantía, de convergencia expresada en términos probabilísticos, del algoritmo genético que mantiene a la mejor solución en la población. Davis y Principe (1993) extrapolan los fundamentos teóricos del algoritmo simulated annealing a un modelo de algoritmo genético basado en cadenas de Markov.

Se efectúa un estudio de las matrices de transición de estados teniendo en cuenta en primer lugar tan sólo la reproducción, a continuación la reproducción y la mutación y finalmente la reproducción, la mutación y el cruce. Suzuki (1993) efectúa un estudio de la convergencia de los algoritmos genéticos por medio de cadenas de Markov. Los algoritmos genéticos estudiados presentan un criterio de reducción elitista modificado, según el cual se genera una población de λ individuos, incluyendo en ella al mejor individuo de la población en la generación anterior, obteniéndose los $\lambda - 1$ individuos restantes por medio de las operaciones genéticas normales.

Liepins (1992) demuestra la convergencia del algoritmo genético hacia poblaciones que contienen al óptimo, en el caso de algoritmos genéticos sin operador de mutación, pero en los cuales el reemplazamiento de individuos es elitista - el mejor individuo no se pierde nunca - y además se efectúa de tal manera que en cada paso cualquier punto del espacio sea potencialmente alcanzable por medio de la operación de cruce. Chakraborty y Dastidar (1993), presentan un modelo de fiabilidad estocástica de un esquema para el algoritmo genético binario con longitud de representación fija, y obtienen una estimación para el número de generaciones necesarias hasta obtener la convergencia. Eiben y col. (1990) modelan la evolución del algoritmo genético por medio de una cadena de Markov, obteniendo condiciones suficientes para la convergencia en probabilidad del proceso evolutivo hacia el óptimo global.

4.6 Puntos a considerar en un Algoritmo Genético

En los algoritmos genéticos es de vital importancia considerar los siguientes puntos:

- Codificación de los parámetros de un problema.
- Dentro de la codificación a veces se usan codificaciones que tengan la propiedad de que números consecutivos varíen a lo más en un bit (i.e., codificación de Gray).

En la codificación se busca idealmente que todos los puntos estén dentro del

espacio de solución (sean válidos). Tradicionalmente se buscan representaciones que favorezcan los esquemas cortos de bajo orden. Pueden existir problemas de interdependencia (problemas para los GA si existe mucha y es preferible usar otro método si es casi nula).

- Función de aptitud.

Es la base para determinar que soluciones tienen mayor o menor probabilidad de sobrevivir. Se tiene que tener un balance entre una función que haga diferencias muy grandes (y por lo tanto una convergencia prematura) y diferencias muy pequeñas (y por lo tanto un estancamiento).

Criterios de tamaño de la población.

Balance entre una población muy pequeña (y por lo tanto convergencia a máximo local) y una población muy grande (y por lo tanto requerimiento de muchos recursos computacionales). Los primeros intentos de tratar de estimar el tamaño de población óptimo basado en el teorema de esquemas resultaban en crecimiento exponenciales de la población con respecto al tamaño de gen. Experimentalmente se vió que esto no es necesario y se buscó el tamaño mínimo para poder alcanzar todos los puntos en el espacio de búsqueda. Se requiere solo que existe al menos cada una de las posibles instancias de los alelos en el conjunto de genes. Para genes binarios, la probabilidad de que existe al menos un alelo en cada punto es:

$$P_2 = (1 - (1/2)^{M-1})^l \quad (4.9)$$

donde M es el tamaño de la población y l es el tamaño del gen. Esto sugiere que con poblaciones de tamaño $O(\log l)$ es suficiente.

Normalmente las poblaciones se seleccionan aleatoriamente, sin embargo, esto no garantiza una selección que nos cubra el espacio de búsqueda uniformemente. Se pueden introducir mecanismos más sofisticados para garantizar diversidad en la población. Aunque normalmente se elige una población de tamaño fijo, también existen esquemas de poblaciones de tamaño variable.

Criterio de selección.

Individuos son copiadas de acuerdo a su evaluación en la función objetivo (aptitud). Los más aptos tienen mayor probabilidad a contribuir con una o más copias en la siguiente generación (se simula selección natural). Se puede implementar de varias formas, sin embargo, la más común es la de simular una ruleta, donde cada cadena tiene un espacio en ella proporcional a su valor de aptitud.

$$\Pr(h) = \frac{\text{amplitud}(h)}{\sum_{j=1}^N \text{aptitud}(h_j)} \quad (4.10)$$

En lugar de seleccionar uno a la vez, se pueden generar N selecciones teniendo en la ruleta N apuntadores que están separados uniformemente, por lo que un giro en la ruleta nos proporciona N individuos. A esto se le conoce como stochastic universal selection. Otra alternativa es usar selección por torneo. Se seleccionan aleatoriamente N individuos que compiten entre sí seleccionando el mejor. En esta alternativa el mejor individuo siempre es seleccionado. Una ventaja de este esquema es que solo necesitamos comparar si un gen es mejor que otro, por lo que posiblemente no tenemos que evaluar la función de aptitud. Sin embargo, haciendo muestreo con reemplazo, existe una probabilidad de aproximadamente 0.368 ($\approx e^{-1}$) de que un gen no sea seleccionado. Finalmente, el otro tipo de selección es por ranqueo. Simplemente se ordenan los genes. La probabilidad de seleccionar un gen ranqueado como el k-ésimo ($P(k)$), en el caso de ranqueo lineal, es:

$$P(k) = \alpha + \beta k \quad (4.11)$$

donde α y β son constantes y

$$\sum_{k=1}^M (\alpha + \beta k) = 1 \quad (4.12)$$

Nueva población.

Se pueden seleccionar individuos de la población actual, generar una nueva población y reemplazar con ella completamente a la población que se tenía (esquema generacional). También a veces se mantienen los N mejores individuos de una población a la siguiente (esto parece ser la mejor opción). Si se conserva solo el mejor individuo se llama estrategia elitista, si se mantiene un subconjunto se habla de poblaciones traslapadas.

Criterio de paro

Normalmente cuando un porcentaje alto de la población converge a un valor o después de un número fijo de evaluaciones en la función de aptitud. Si con ese valor no se llega a la medida esperada (si es que se conoce) o simplemente para tratar de mejorar la solución, entonces se toma una pequeña proporción y se inyecta "diversidad genética" (se generan aleatoriamente nuevos individuos), o inclusive se reemplaza completamente la población.

Operadores genéticos.

Normalmente se hace cruce seguida de mutación, pero se podrá hacer una o la otra, por ejemplo. Algunas personas proponen hacer mucha cruce al principio e incrementar la mutación conforme pasa el tiempo.

- Cruce: tiene una alta probabilidad de ser utilizado y es considerado como el más importante dentro de los AG. Permite la generación de nuevos individuos tomando características de individuos padres. Consiste en seleccionar dos individuos después del proceso de selección, determinar una posición de cruce aleatoria e intercambiar las cadenas entre la posición inicial y el punto de cruce y el punto de cruce y la posición final. Existen diferentes tipos de cruce. (i) Cruce simple: un solo punto de cruce (una máscara de 1's seguida de 0's), (ii) cruce de dos puntos y (iii) cruce uniforme (generando una máscara con 1's y 0's siguiendo una distribución de Bernoulli). Existe evidencia empírica que sugiere que cruce en un punto no es la mejor opción.
- Mutación: tiene baja probabilidad de ser utilizado y permite introducir nueva información no presente en la población. Opera sobre un solo individuo, determina una posición y la invierte con cierta probabilidad. Permite salir de máximos locales. En lugar de generar un número aleatorio para cada alelo y ver si se muta o no, se puede generar un número que nos de el número de mutaciones a realizar (M) y simplemente generar M números aleatorios entre 1 y el tamaño del gen. Lo único que falta por determinar es que valor usar cuando existan más de dos posibles valores.

- Inversión: tiene baja probabilidad. Incrementa la capacidad de exploración. Permite generar cadenas que serían difíciles de obtener con los otros dos operadores. Opera en un individuo, determina dos posiciones dentro de la cadena e invierte la subcadena.

Programación Genética

La programación genética es computación evolutiva que genera programas (Koza, '92). O sea que lo que se tiene es una población de programas.

Los programas típicamente se representan como árboles de "parseo" de un programa. Donde cada función está representada en un nodo y los argumentos de la función en sus ramas. Para aplicar programación genética a un dominio se tiene que especificar:

- Las funciones primitivas a considerar (i.e., \cos , $\sqrt{\quad}$, $+$, $-$ etc.).
- Los nodos terminales (i.e., x ; y ; 2 , etc.).
- Se siguen los siguientes pasos:
- Genera una población inicial mediante una composición (generalmente aleatoria) de funciones y símbolos terminales relacionados al dominio.
- Realiza en forma iterativa los siguientes pasos, hasta llegar a un criterio de terminación:
- Ejecuta cada programa y asígnale un valor de acuerdo a tu función de aptitud.
- Crea nuevos programas haciendo: (i) Reproducción (copia), (ii) Cruza (intercambio aleatorio de dos partes de un programa, (iii) Mutación (cambia una parte aleatoria de un programa), y (iv) Operaciones que cambian la arquitectura (ver abajo),
- Selecciona el mejor programa.
- Cruza selecciona aleatoriamente nodos padres de dos árboles y los intercambia. Las operaciones que alteran la "arquitectura" se usan para determinar: (i) el número de funciones a utilizar, (ii) el número de argumentos por función, y (iii) el tipo de jerarquía (quien llama a quien).

CAPITULO V. ANÁLISIS MULTICRITERIO

5.1 Procesos y Clasificación de modelos en la toma de decisiones

Un proceso de decisión implica, necesariamente, la comparación entre las alternativas, el hecho de comparar elementos se traduce en la necesidad de realizar mediciones que permitan aplicar los criterios de comparación de modo de establecer preferencias entre ellos. Los elementos que participan en un proceso de decisión, por lo general se miden en escalas diferentes (peso, distancia ó tiempo), por lo que, se requiere transformar estas unidades en una unidad abstracta que sea válida para todas las escalas.

En el análisis de decisiones de inversión, normalmente esta diversidad de métricas¹⁸ se resuelve con la transformación de todos los impactos de un proyecto, en un indicador económico, típicamente el VAN, cuya unidad de medida es el valor del dinero en un instante dado del tiempo. Por otro lado, participan también en el proceso muchas variables intangibles imposibles de cuantificar en medidas tradicionales, aspectos políticos, sociales y ambientales por ejemplo, que también deben verse representados por una escala común, y que algunas ocasiones son de difícil medición en términos económicos. ¿Cómo determinamos la importancia de estos factores y sintetizamos luego toda esta información para tomar la mejor decisión? Este es un típico problema de toma de decisiones.

Lo que interesa medir es cuánto más preferible es una alternativa sobre otra y para compararlas necesitamos una escala de evaluación común. Las escalas de evaluación permiten caracterizar los elementos bajo un mismo patrón de comparación pudiendo de esta manera establecer relaciones entre ellas.

¿Por qué una metodología multicriterio? Porque es necesario una metodología que logre combinar las distintas dimensiones objetivos, actores y escalas que se hallan envueltos en el proceso de toma de decisiones, sin sacrificar la calidad, confiabilidad y consenso en los resultados. Una de las características principales de las metodologías multicriterio es la diversidad de factores que se logran integrar en el proceso de evaluación. La particularidad de cada metodología multicriterio está en la forma de transformar las

¹⁸Por métricas se entiende una medida de comparación.

mediciones y percepciones en una escala única, de modo de poder comparar los elementos y establecer ordenes de prioridad.

Una de las metodologías multicriterio más utilizadas, con fundamentos matemáticos, es el Proceso Analítico Jerárquico (Analytic Hierarchy Process: AHP). Los métodos de evaluación multicriterio provienen fundamentalmente del área de Investigación de Operaciones (I.O.). Desde esa disciplina se puede hacer la siguiente clasificación de los modelos multicriterio.

El problema clásico de I. O. es optimizar una función objetivo sujeto a un set de restricciones. Por ejemplo, en evaluación de proyectos: Maximizar VAN dada la disponibilidad de recursos. En los modelos multicriterio se aborda un problema distinto: optimizar un conjunto de funciones objetivo. La “teoría de evaluación multicriterio” comprende en realidad un conjunto de teorías, modelos y herramientas de apoyo a la toma de decisiones, aplicable no sólo al análisis de inversiones, sino a una amplia gama de problemas en la gestión tanto privada como pública tales como: análisis de posicionamiento de marcas en el mercado, medición de percepciones de clientes y selección de tecnologías. Desde el punto de vista de toma de decisiones con base a la rentabilidad de los proyectos, el mundo ideal sería aquel que permite incorporar todos los efectos del proyecto en el VAN. Dado que esto no siempre es posible (hay beneficios y costos que no son medibles), se suele agregar al VAN un listado de beneficios y costos no cuantificables, por ejemplo: efecto en la descentralización, impactos ambientales no cuantificables, efectos redistributivos. Pero lo anterior no resuelve el problema de cómo tomar una decisión a partir de esa información. Los modelos multicriterio permiten agregar afectos de un proyecto en una métrica común. Para ello se debe tener en cuenta los siguientes pasos:

1. Se deben definir los criterios (objetivos intermedios), y sus respectivas restricciones.
2. Definir tipos de variables: discretas o continuas.
3. Modelación de las preferencias. Existen básicamente dos alternativas: optimizar por separado para cada objetivo y luego agregar los subconjuntos de soluciones ó asignar pesos a los distintos objetivos y encontrar una sola solución.
4. Definir si se usan modelos determinísticos (sin incertidumbre) ó aleatorios. En el último caso se aplica la Teoría de preferencias sobre contingencias: programación dinámica, simulación, análisis probabilístico.

5. Si se opta por agregar objetivos se deben definir los métodos de agregación.

Dentro de estos métodos tenemos:

- Método de "juicio de expertos"
- Funciones de utilidad multiatributadas: transforman los múltiples criterios en uno sólo.
- Factor análisis
- Escalamiento Multidimensional
- Analytic Hierarchy Process (AHP)

Otros.

De acuerdo a los pasos anteriores, las características AHP (método utilizado en este trabajo) son las siguientes: método de evaluación multicriterio, de variables discretas, con medición de preferencias por agregación de criterios y determinístico (no considera incertidumbre).

Paradigmas

Podemos entender los paradigmas como meta-preguntas de la teoría, las más sencillas, las que justamente por esa calidad no se explicitan ni se cuestionan, sino que se dan como supuestos. Las preguntas que están en la base de la definición de paradigma son la definición de realidad, de los criterios de verdad, de la relación entre la parte y el todo [1], [7]. El paradigma positivista dice relación con los principios lógicos de la lógica clásica, con el conocimiento objetivo y el método cartesiano. La realidad es un objeto que puede ser estudiado por partes y la lógica inductiva basada en los principios de causalidad, razón necesaria, razón suficiente y determinismo causas para la explicación de los fenómenos formulando regularidades. El análisis y la experimentación requieren de hipótesis que disminuyan lo máximo posible el componente subjetivo de la realidad. El desarrollo científico moderno y la tecnología industrial tienen sus orígenes en el paradigma positivista, esto es, hay partes de la realidad donde su aplicación tiene validez y explicación. El denominador común de todos los problemas que se abordan con evaluación multicriterio, es el reconocimiento explícito de la complejidad en los procesos de toma de decisiones individuales y más aún a nivel grupal. Desde el punto de vista filosófico representa un movimiento desde el paradigma del racionalismo hacia el de "pensamiento lateral o visión periférica".

Conviene resumir las características esenciales de estos dos enfoques. La visión racionalista – positivista – cartesiana, ha dominado el sistema de pensamiento de la civilización occidental desde que la filosofía clásica griega sistematizó el uso del análisis, el juicio y la argumentación. El racionalismo en el mundo moderno asume que para cualquier problema de toma de decisiones existe una solución óptima precisa y que es posible encontrarla razonando respecto al problema y modelándolo adecuadamente. Un supuesto subyacente en esta visión es la tangibilidad de las variables y atributos que inciden en la toma de decisiones.

La teoría tradicional de evaluación de proyectos, se enmarca básicamente dentro de la tradición racionalista, en efecto, hemos supuesto siempre que los individuos maximizan su utilidad y la sociedad maximiza el bienestar social, que podemos conocer toda la información (o al menos la mayor parte) necesaria para la toma de decisiones y que la tangibilidad de esta información nos permite medir (todos los costos y beneficios) para llegar a un criterio único (VAN) que nos permite tomar la decisión en forma racional.

El “pensamiento lateral” ha sido mas propio de la filosofía oriental y comenzó a impregnar la cultura occidental sobre todo en la década de los 80, motivado en parte por la creciente influencia y éxitos de Japón en la economía mundial durante esa época. Podríamos intentar sintetizar este enfoque por contraste con el paradigma racionalista. El enfoque da cuenta de la intangibilidad de muchas de las dimensiones relevantes para el proceso de toma de decisión, incorpora como dato que el proceso de toma de decisión no necesariamente es racional bajo las definiciones antes señaladas, en la medida de que factores subjetivos que el tomador de decisiones no es capaz de reconocer ni explicitar inciden fuertemente en la decisión final, y por último reconoce que la racionalidad varía de una persona a otra y de un grupo a otro.

Para efectos de la evaluación de inversiones, este cambio de paradigma implica al menos incluir en la toma de decisión los aspectos no cuantificables, identificar aspectos subjetivos involucrados en la toma de decisión e incluir las distintas visiones y objetivos de los agentes. A modo de ejemplo, significa considerar en la toma de decisión aspectos tales como correlación de fuerzas entre grupos, intereses no declarados de los agentes, participación de los beneficiarios y restricciones socio-culturales entre otras[2].

Este componente está orientado a desarrollar elementos conceptuales en el campo del Análisis Multicriterio, el cual tiene aplicación en las organizaciones empresariales de cualquier tipología. La Decisión Multicriterio (DM) o Análisis Multicriterio (AM), es una ayuda efectiva en la práctica de la toma de decisiones y de la gestión de las organizaciones, para lo cual es preciso aclarar los aspectos fundamentales ya que constituye una forma de modelizar los procesos de decisión, en los que entran en juego: una decisión a ser tomada, los eventos desconocidos que pueden afectar el o los resultados, los posibles cursos de acción, y el o los resultados mismos. Mediante los modelos multicriterio el decisor podrá estimar las posibles implicaciones que puede tomar cada curso de acción, de modo que se pueda obtener una mejor comprensión de las vinculaciones entre sus acciones y sus objetivos.

Los elementos que son objeto de estudio por el Análisis Multicriterio son de tipo amplio los cuales hacen referencia a la elección entre un número finito de alternativas posibles; como son los proyectos, las políticas, los cursos de acción, las inversiones o la elección de candidatos, entre otras. A pesar de existir una importante literatura científica sobre la Decisión Multicriterio, las herramientas, los métodos y hasta la propia reflexión multicriterio, permanecen casi totalmente desconocidos para los técnicos y los directivos de todos los niveles que hacen parte de las organizaciones o interactúan en ellas.

Se consideran que las causas son de tipo cultural, ya que la cultura científica continúa privilegiando la noción de “la mejor decisión” y el concepto de “óptimo” aspectos que no existen en el Análisis Multicriterio, ni en la mayoría de las situaciones reales de la decisión, los cuales deben tenerse en cuenta casi en forma general en la academia y en la formación de los profesionales y directivos que orientarán en el futuro las organizaciones.

Lo expuesto anteriormente conlleva a que en las empresas y las organizaciones se continúe llevando a cabo el proceso de decisiones fundamentado en la optimización (Paradigma de la Teoría de la Optimización), lo que hace que se limite en muchas ocasiones el análisis de una optimización cualquiera, la cual se sabe perfectamente que no es más que un pálido reflejo de la complejidad de lo real. No es posible ignorar que cada decisión real, consiste de hecho en un compromiso sobre diversas soluciones, cada una con sus ventajas e inconvenientes; dependiendo de la posición que se adopte.

Cada vez será más difícil en las organizaciones en cuanto al proceso de la toma de decisiones, no tener en consideración los diferentes puntos de vista, las motivaciones o los fines de los decisores. Los tiempos de la función monoobjetivo o unicriterio, están finalizando en este mundo tan cambiante; ahora es preciso tener en cuenta los deseos de los diferentes actores y la pluralidad de sus intenciones en la toma de las decisiones, por lo que se hace entonces necesario contar con un mínimo de conocimientos en cuanto hace referencia al Análisis Multicriterio, aspecto que es conocido en el campo de la decisión como, Modelización Multicriterio de las Decisiones o Ayuda Multicriterio a la Decisión. Una aproximación inicial a lo que es el Análisis Multicriterio o la Decisión Multicriterio, se puede observar a través del siguiente modelo de decisión:

Max Z, Min W (Criterios que entran en conflicto)

$$\text{Max } Z = X_1$$

$$\text{Min } W = X_2$$

$$\text{Maximizar } (X_1, X_2) = X_1 - X_2$$

$$\text{Minimizar } (X_1, X_2) = -X_1 + X_2$$

Cuando los deseos entran en conflicto, la decisión resultará un compromiso, en este modelo simple la expresión “decisión multicriterio” no es muy acertada puesto que la decisión, no es unicriterio ni multicriterio, sino elección; es decir acción y casi siempre intención. Lo que es multicriterio es el modelo de ayuda a la decisión. (Roy, 1971). Lo que permite hoy en día, que el Análisis Multicriterio sea considerado como un campo de actividad, en el que la aplicabilidad práctica y las herramientas informáticas son dominantes. Aplicaciones en el campo de los modelos multicriterios pueden ser analizadas a través de los p modelos diferentes:

- P modelos diferentes (Bienes) = B1, B2,, Bn, Criterios en conflicto: Precio y calidad.
- P modelos diferentes (Proyectos) = P1, P2,, Pn, Criterios en conflicto: Rentabilidad, impacto social, impacto en el ambiente.
- Construcción de una vía. Criterios en conflicto: Menor costo, más flujo vehicular, menos impacto en el ambiente.

- Inversión fundamentada en el ahorro, puede estar destinada al préstamo personal, la compra de acciones, adquisición de CDT. Criterios en conflicto: Riesgo, rentabilidad.

Analizando estos modelos propuestos, se puede plantear que toda decisión tiene consecuencia en el futuro, ya que el decisor se ve confrontado con una elección en presencia de múltiples criterios. En una primera aproximación, se puede argumentar que la Decisión Multicriterio (DM), es: aquella acción en la cual el decisor, tiene la disposición de escoger varias posibilidades o alternativas; que conforman su conjunto de elección y del cual, él puede escoger de acuerdo a sus diversos puntos de vista o criterios que percibe en conflicto, el que mejor satisfaga su intención de decisión.

En el desarrollo de la aplicación de un modelo fundamentado en el Análisis Multicriterio, se hace necesario identificar los elementos que le dan la dinámica a dicho patrón para el proceso de la toma de decisiones. (Barba-Romero et. al., 1997). En un orden consecuente estos elementos, son los que a continuación se describen:

- i. Decisor: Persona o grupo de individuos enfrentados a una decisión, que asumen el rol de ser los encargados de analizar la decisión.
 - ii. Toma de decisiones: Proceso a lo largo del tiempo, en el que se identifican las siguientes fases; recoger la información, el diseño, la selección y la revisión. En conjunto conforman la reflexión del decisor.
- a) *Recoger la información*: Obtención de datos respecto a los criterios y alternativas del problema.
 - b) *Diseño*: Determinación precisa de los criterios y sus escalas de medida, para la construcción completa del conjunto de elección, que conforma el agregado de alternativas con sus respectivas evaluaciones para cada criterio.
 - c) *Selección*: Procedimiento de elegir finalmente una de las alternativas.
 - d) *Revisión*: Seguimiento de como fue desarrollado el proceso.
 - i. Analista: Es quien modeliza la situación concerniente al objeto de estudio y hace recomendaciones relativas a la selección final. Para el cumplimiento de su función, se apoya en una herramienta informática que interviene como un modelo predefinido siendo soporte para la formalización, la memorización y la reflexión.

- ii. Conjunto de elección: Conjunto de alternativas que debe elegir el decisor, con las características de ser diferentes, excluyentes y exhaustivas. $A_i = \{A_1, A_2, \dots, A_m\} \quad i = 1 \dots m$

En donde el decisor no puede escoger una solución mixta, es decir, intermedia entre dos alternativas A_i y A_k ; ni tampoco escoger una alternativa que no pertenece al conjunto de elección. Si se introduce una nueva alternativa, se hace necesario empezar el proceso de la toma de decisiones con un nuevo conjunto de elección.

e) Atributos: Son las características o cualidades que poseen las respectivas alternativas, como pueden ser: precio, calidad, estética, solidez entre otras.

f) Criterios: Son las preferencias que el decisor tiene hacia un cierto atributo.

$$C_n; i = 1 \dots n$$

Los criterios deben estar en el mismo plano de igualdad, es decir; serán criterios cuantitativos, cuando estén en una escala numérica o cualitativos, cuando no existe unidad canónica de medida.

g) Matriz de Decisión: El decisor es capaz de asignar un valor numérico o simbólico (A_{ij}), cuantitativo o cualitativo para cada atributo considerado (j) y para cada alternativa del conjunto de elección (A_i).

La siguiente situación refleja la matriz de decisión en el Análisis Multicriterio:

		Criterios					
		C_1	C_2	C_j	C_n
Alternativas	A_1	a_{11}	a_{12}	A_{1j}	a_{1n}
	A_2	a_{21}	a_{22}	A_{2j}	a_{2n}

	A_i	a_{i1}	a_{i2}	a_{ij}	a_{in}

	A_m	a_{m1}	a_{m2}	a_{mj}	a_{mn}

Tabla 2. Matriz de decisión en el Análisis Multicriterio

De esta matriz se puede analizar lo siguiente; primero a_{in} , relaciona las cualidades de la alternativa i respecto a los n atributos o criterios considerados por el decisor y segundo

a_{mj} , corresponde a las evaluaciones hechas por el decisor bajo un preorden de todas las alternativas con respecto al atributo o criterio j .

5.2 Preferencias y relaciones de orden y función de utilidad

En el Análisis Multicriterio, el decisor debe considerar entre alternativas A_i y A_m de su respectivo conjunto de elección, llegándose a presentar las siguientes condiciones:

1. $A_i P A_m (A_i > A_m)$ o $A_m P A_i (A_m > A_i)$: El decisor prefiere estrictamente a A_i que a A_m o A_m que a A_i cuando su elección se efectúa sin ninguna duda sobre la alternativa A_i o A_m .
2. $A_i I A_m (A_i \cong A_m)$ o $A_m I A_i (A_m \cong A_i)$: Para el decisor es indiferente entre la alternativa A_i o A_m igualmente de manera recíproca, cuando el acepta indistintamente una alternativa frente a la otra.
3. $A_i Q A_m (A_i \geq A_m)$ o $A_m Q A_i (A_m \geq A_i)$: El decisor no sabe si prefiere estrictamente a A_i sobre A_m o A_m sobre a A_i , o si su decisión es indiferente entre las alternativas.
4. $A_i NC A_m$: No $(A_i \geq A_m)$ o $A_m NC A_i$: No $(A_m \geq A_i)$: El decisor es incapaz de escoger entre las alternativas, es decir, rechaza escoger entre las alternativas ya que estas no son comparables.

Las condiciones; $>$, \cong , \geq , NC , representan *relaciones de orden* ya que relacionan alternativas del conjunto de elección del decisor. La *función de utilidad*, expresa las preferencias psicológicas reales del decisor respecto a las m alternativas; en donde la función de utilidad ordinal, no indica más que un orden para quien va a tomar la decisión mientras que la función de utilidad cardinal, permite comparar las diferencias en las preferencias del decisor, es decir; respeta las diferencias entre las alternativas además del orden, lo que es conocido en la Decisión Multicriterio como la función de utilidad intervalo.

También es importante considerar la función de utilidad ratio, la cual incluya el valor cero para hacer las respectivas comparaciones. La Decisión Multicriterio, permite al decisor develar progresivamente sus propias preferencias respecto a las m alternativas, ya que es un proceso de decisión con varios actores y aún cuando tales actores no tengan intereses antagonistas, tendrán al menos sensibilidades diferentes que se traducirán en

evaluaciones o criterios de elección diferente, fundamentados en los argumentos de unos y otros, tales como las comparaciones de las utilidades a lo que se puede agregar que el Análisis Multicriterio aporta una metodología y unas herramientas muy interesantes en el campo de las decisiones. (Barba-Romero et. al., 1997).

Los métodos multicriterio requieren de una matriz de decisión de tipo numérico, ya que la transformación del preorden del decisor en una escala numérica supone la construcción de una función de utilidad. Las m alternativas ya se conocen con certeza y el decisor está de acuerdo en evaluarlas según una escala numérica preestablecida, obteniéndose así una función de utilidad del decisor relativa al criterio o atributo considerado.

Los métodos en el Análisis Multicriterio o Decisión Multicriterio, se fundamentan en los aspectos de la dominación y la satisfacción, para así construir los respectivos modelos. A continuación se presentan diferentes aspectos con relación a los métodos que se pueden tener en cuenta en la Decisión Multicriterio para el desarrollo de modelos en este campo de la decisión.

5.3 Óptimo de Pareto

Este método, hace referencia a que para él decisor poder ganar en un criterio o atributo, se hace para él preciso aceptar que perderá en algún otro criterio o atributo. Por lo anterior, se plantea que una alternativa es eficiente u óptima en el sentido de Pareto, cuando ésta es una alternativa que en el proceso de la decisión no está dominada estrictamente. La aplicación de este método, puede ser analizada a través del siguiente proceso que se lleva a cabo en la respectiva Matriz de Decisión:

1. Se observa si existe una alternativa dominante, de serlo así esta será la mejor elección.
2. Las alternativas no dominadas, constituyen los óptimos de Pareto o alternativas eficientes.
3. Se eliminan las alternativas dominadas. Para esto se define un umbral de satisfacción (U_j) por parte del decisor, eliminándose las alternativas que están debajo de este umbral; por defecto el resto de criterios tendrán el nivel más bajo - nivel de satisfacción - Este procedimiento en el campo de la decisión multicriterio es conocido como Preamálisis de Dominación, ya que en él se

identifican las alternativas dominadas por parte del decisor, para ser eliminadas de la matriz de decisión.

Programación Multi-objetivo (PMO)

La programación multi-objetivo consiste en encontrar un vector de variables de decisión que satisfaga ciertas restricciones y optimice una función vectorial cuyos elementos representen las funciones objetivo. Estas funciones forman una descripción matemática de los criterios de desempeño que usualmente están en conflicto entre si y que se suelen medir en unidades diferentes. Por lo tanto, el término optimizar significa "encontrar una solución tal que proporcione valores para todos los objetivos que resulten aceptables para el diseñador" Osyczka (1984).

El problema de optimización multi-objetivo (POM) se define formalmente como:

$$\begin{aligned}
 &\text{Encontrar el vector } \vec{x}^* = [x_1^*, x_2^*, \dots, x_n^*]^T \text{ que} \\
 &\text{satisfaga las } m \text{ restricciones de desigualdad:} \\
 &\quad g_i(\vec{x}) \leq 0 \quad i = 1, 2, \dots, m \\
 &\text{que satisfaga las } p \text{ restricciones de igualdad} \\
 &\quad h_i(\vec{x}) = 0 \quad i = 1, 2, \dots, p \\
 &\text{y optimice la función vectorial} \\
 &\quad \vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \dots, f_k(\vec{x})]^T \\
 &\text{donde: } \vec{x} = [x_1, x_2, \dots, x_n]^T \text{ es el vector de variables} \\
 &\text{de decisión.}
 \end{aligned} \tag{5.1}$$

Las restricciones definen la región factible F y cualquier punto en F forma parte de la solución factible. Optimizar varias funciones objetivos a la vez, no se traduce en encontrar un óptimo para cada función, sino en proponer un conjunto de puntos en la que cada función objetivo contribuya en alcanzar una buena aptitud global, la cual será establecida y evaluada por el diseñador, un inversionista en nuestro caso. Así el concepto de óptimo que se maneja en PMO es diferente a programar una sola función objetivo, ya que se sustenta en un compromiso entre todas las variables que satisfacen las restricciones, es decir, una solución global de la optimización. Esta noción de "óptimo" fue propuesta por Pareto en 1886 y conocida como el "óptimo de Pareto" y su definición formal es la siguiente:

Definición 1: Óptimo de Pareto

Dadas k funciones objetivo del problema, decimos que un punto $\vec{x}^* \in F$ es un óptimo de Pareto si para toda $\vec{x} \in F$, tal que para toda $i \leq 1, 2, \dots, k$

$$f_i(\vec{x}^*) \leq f_i(\vec{x})$$

y, al menos existe una i tal que

$$f_i(\vec{x}^*) < f_i(\vec{x})$$

Esta definición dice que \vec{x}^* es un óptimo de Pareto si no existe dentro del espacio de búsqueda F un vector factible \vec{x} que mejoraría algún criterio, sin hacer que empeore al menos otro de ellos, es decir, que lo domine, por lo que se le conoce a \vec{x}^* como una solución no dominada.

Definición 2: Dominancia de Pareto

Un vector $\vec{u} = (u_1, \dots, u_k)$ se dice que domina al vector $\vec{v} = (v_1, \dots, v_k)$ denotado por $\vec{u} \prec \vec{v}$ si y solamente si es parcialmente menor que v, es decir:

$$\forall i \in \{1, \dots, k\}, u_i \leq v_i \wedge \exists i \in \{1, \dots, k\} : u_i < v_i$$

El óptimo de Pareto casi siempre produce no una, sino un conjunto de soluciones a las que se llaman no dominadas.

Definición 3: Conjunto de Óptimos de Pareto.

Dado un PMO $\vec{f}(\vec{x})$, el conjunto de óptimos de Pareto P_{true} se define como:

$$P_{true} := \{ \vec{x} \in F : \neg \exists \vec{x}' \in F \vec{f}(\vec{x}') \prec \vec{f}(\vec{x}) \}$$

A partir de la definición 3 se define el Frente de Pareto como el contra dominio del conjunto de puntos que forman el conjunto de óptimos de Pareto, es decir, el frente de Pareto son los valores de las funciones objetivo correspondientes a las soluciones que pertenecen al conjunto de óptimos de Pareto.

Resulta imposible determinar de manera analítica la expresión matemática que corresponde al frente de Pareto de un problema arbitrario. Aproximar dicho frente es precisamente el objetivo de la optimización multi-objetivo.

Cualquier decisión involucra la posibilidad de elegir entre varias opciones, lo cual significa necesariamente una comparación de los atributos que las distinguen. Un

modelo multicriterio es la formalización matemática de este proceso de elección. Así, al aplicar un modelo multicriterio el proceso de decisión se descompone en sus partes esenciales:

Identificar las alternativas, las cuales representan las diferentes opciones o cursos de acción que se pueden elegir.

Delinear los criterios de decisión, o sea los estándares o patrones de referencia con las que se evalúan las alternativas con respecto a la meta, los objetivos y los atributos. La meta es un enunciado general de lo que se pretende conseguir; los objetivos son las cuestiones objetivas con las que se mide el grado de cumplimiento de la meta; y los atributos son las características específicas con los que se evalúan las alternativas.

Aplicar un algoritmo o secuencia de operaciones para, por una parte, ponderar los criterios de decisión en cuanto su relevancia y, por la otra, ordenar las alternativas por categorías de preferencia o prioridad.

Es importante señalar que en un modelo multicriterio se pueden considerar criterios de decisión de diferente índole, desde los factores intangibles --como los valores, los principios y las suposiciones-- hasta los datos tangibles --como los costos, especificaciones técnicas, las restricciones financieras o ambientales, las capacidades de la planta productiva, o las características del mercado. De esta forma, un modelo multicriterio permite examinar con detalle todos los elementos de un problema, lo que hace posible discernir la solución que mejor satisface una meta de manera rigurosa.

En la realidad económica, social, política que rodea a las organizaciones, se presentan en forma continua situaciones complejas en las que se ha de elegir entre varios cursos de acción posibles para obtener uno satisfactorio o seguir el óptimo. Para este proceso de toma de decisiones no basta la experticia, el sentido común o la intuición de los agentes expertos, ya que la realidad enseña que en el proceso intervienen múltiples criterios, participan varios decisores, interviene la incertidumbre y en la acción como tal se presentan varias etapas. (Ríos et. al., 1989).

El desarrollo actual de la teoría de los procesos de decisión, teoría esta que puede ser considerada como la columna básica de la Investigación Operativa, denominada también Análisis de Sistemas, es uno de los instrumentos más pujante en el desarrollo actual de la gestión empresarial.

La variabilidad de los problemas reales de decisión humana (toma de decisiones) en las organizaciones ha hecho necesario de la contribución en conjunto de matemáticos, estadísticos, ingenieros, psicólogos, economistas y administradores para la construcción de esquemas coherentes (modelos), cada vez más amplios, pero algunas veces insuficientes y a la espera de nuevos desarrollos, en beneficio del proceso de la Toma de Decisiones. El enfoque de la Decisión Multicriterio abarca la realidad de las decisiones en las organizaciones, ya que considera en forma explícita criterios múltiples para abordar en forma correcta los problemas de decisión a través de la construcción, en forma sucesiva de síntesis de estos criterios en cuya realización se encuentra inmerso el problema central.

La nueva visión de mundo en el campo de la Teoría Decisional, esta representada en la Teoría de la Decisión Multicriterio (Análisis Multicriterio), la cual sustenta que los agentes económicos (organizaciones) no optimizan sus decisiones con base a un solo objetivo, sino que pretenden lograr un equilibrio entre el conjunto de objetivos en conflicto o en la medida de lo posible satisfacer una serie de metas que están asociadas a dichos objetivos. Es así, como hoy en día en otras latitudes la ciencia de la decisión va logrando modelos matemáticos más perfectos, para resolver en la organización situaciones reales ya sean individuales o colectivas (toma de decisiones) soportados en múltiples criterios, la incertidumbre y el riesgo, entre otros. (Romero, 1993).

La Teoría de las Decisiones Multicriterio comprende dos aspectos, la Decisión Multicriterio Discreta (DMD), la cual hace referencia ha que no hay más que un número finito de alternativas posibles (proyectos, inversiones, candidatos) y la Decisión Multicriterio Continua (DMC), para la cual existe un número infinito de alternativas posibles.

Las fases que comprende el proceso de Toma de Decisión Multicriterio, permiten que quien analiza el proceso participe en todas estas, quien actúa como agente o agentes decisores interviene en ellas para recopilar la información o validar las acciones y quien propone la aplicación de la Decisión Multicriterio estará implicado en las dos primeras etapas y en los resultados finales, en las recomendaciones y en el estudio de viabilidad. (Barba-Romero y Pomerol, 1997, p.360). Estas fases son:

1. Exposición formal del problema.
2. Comprensión y aceptación del contexto de la decisión y del estudio. Acuerdo respecto a la problemática.

3. Generación de las alternativas. Modelización de los criterios.
4. Discusión y aceptación de los modelos. Evaluación de las alternativas, respecto a los criterios.
5. Matriz de decisión.
6. Selección de los métodos de agregación. Recopilación de datos de preferencias del decisor: pesos, niveles de satisfacción.
7. Aplicación del método y/o interactividad.
8. Recomendaciones y explicación de los resultados. Análisis de sensibilidad.

Por último, La importancia de esta investigación en el campo de la Decisión Multicriterio radica primero en que, quien realiza la investigación puede evaluar la viabilidad del trabajo e identificar bien los actores y el problema de decisión y como segundo, con la generación de las alternativas y la construcción de los criterios se estructura el modelo de Decisión Multicriterio el cual queda terminado al presentar el prototipo de la matriz de decisión (Jerarquías Analíticas). (Saaty, 1980).

Este primer componente orientado a desarrollar elementos conceptuales en los campos epistemológico, del pensamiento social, de la toma de decisiones y matemático; pretende una profundización teórica que debe generar un esquema de recontextualización respecto al referente matemático como protagonista fundamental en procesos administrativos.

5.4 Enfoques no económicos

A pesar de ser el enfoque económico uno de los más extendidos en la evaluación de proyectos, éste no suele ser apropiado cuando se trabaja con aspectos intangibles difícilmente cuantificables desde un punto de vista económico. Como ya se ha mencionado, la “Teoría de evaluación Multicriterio” comprende un conjunto de teorías, modelos y herramientas de apoyo a la toma de decisiones, aplicable no sólo al análisis de inversiones sino a una amplia gama de problemas en la gestión tanto privada como pública tales como: Análisis de posicionamiento de marcas en el mercado, medición de percepciones de clientes y selección de tecnologías. Por tratarse del tema central de este trabajo, estas teorías se tratan en un capítulo aparte a continuación.

Técnicas de decisión multicriterio

Destacan las tres aproximaciones más extendidas en el campo de la Decisión Multicriterio Discreta: La Teoría de Utilidad Multiatributo (“MAUT”) basada en los trabajos de Keeney y Raiffa (1976); El proceso Analítico Jerárquico (AHP) basado en los trabajos de Thomas L. Saaty [14], [15], y las técnicas de superación “outrankings”, basadas en los trabajos de Bernard Roy (1969, 1985). Se entiende por Técnicas de Decisión Multicriterio el conjunto de herramientas y procedimientos utilizados en la resolución de problemas de decisión, en los que intervienen diferentes criterios, generalmente en conflicto. En esencia, la Decisión Multicriterio es una optimización con varias funciones objetivo simultáneas y un único agente decisor. Puede formularse matemáticamente de la siguiente manera:

$\max F(x)$

x pertenece a X

donde:

x Es el vector $[x_1, x_2, x_3, \dots, x_n]$ de las VARIABLES de decisión. El problema de decisión es el de asignar los “mejores”.

X Es la denominada región factible del problema (el conjunto de posibles valores que pueden tomar las variables)

$F(x)$ Es el vector $[f_1(x), f_2(x), \dots, f_p(x)]$ de las p funciones objetivo que recogen los criterios u objetivos simultáneos del problema.

No obstante existen para ellos métodos específicos como el AHP, entre muchos otros.

Clasificación de Técnicas Multicriterio

Fijándose en el flujo de información existente entre dos de los actores más destacados del proceso de toma de decisiones, el analista y el decisor (Moreno-Jiménez, 1989), las técnicas multicriterio pueden clasificarse en:

1. Técnicas sin información a priori (generadoras): Son aquellas en las que el flujo de información va desde el analista al decisor. Entre estas técnicas destacan: el método de ponderaciones, el de la ε -restricción y el simplex multicriterio.
2. Técnicas con información a priori: El flujo de información es en el sentido contrario, del decisor al analista.

3. Dentro de este grupo de técnicas se suele hacer otra distinción, según el número de alternativas que tenga el problema: finito o infinito. Si el conjunto de alternativas es infinito se suelen aplicar aproximaciones basadas en optimización, en las que se supone que los distintos objetivos pueden ser expresados en un denominador común mediante intercambios. Destacan en este apartado los métodos de Programación por Compromiso o Programación por Metas. Si el conjunto de alternativas es discreto, hacemos la siguiente diferenciación:
 - a. Métodos de Agregación: En este tipo de Métodos se modelan las preferencias a través de una función valor:
 - a) Directos: Teoría de Utilidad Multiatributo (MAUT).
 - b) Jerárquicos: Proceso Analítico Jerárquico (AHP).
 - b. Métodos basados en relaciones de orden: Se modelan las preferencias a través de un sistema de relaciones binarias:
 - o Métodos de Superación (MS)
4. Técnicas en las que el flujo de información es en los dos sentidos, dando lugar a las denominadas técnicas interactivas.

Dentro de este conjunto de métodos, los más utilizados han sido: STEM y Método de Ziots-Wallenius. En la actualidad, casi todos los métodos pueden considerarse dentro de este último grupo, bastando para ello que el decisor revise sus juicios dentro del proceso de toma de decisiones. Respecto de los tres métodos discretos mencionados anteriormente, se puede indicar que, a pesar de los duros enfrentamientos que han tenido los respectivos seguidores, recientemente se está buscando la integración de las mismas, o por lo menos, la integración de las dos técnicas consideradas de la escuela americana (MAUT y AHP).

Aquellos problemas en los que el conjunto de alternativas es finito, además de discreto y, cuya decisión se basará en las diversas características o atributos de las alternativas respecto de los criterios de decisión relevantes, se llaman Decisión Multicriterio Discreta y les son aplicables algunos de los métodos de la Decisión Multiobjetivo. No obstante existen para ellos métodos específicos como el AHP, entre muchos otros.

Metodología del proceso Analítico Jerárquico

El proceso analítico jerárquico (AHP), es una metodología de análisis multicriterio desarrollada a fines de la década del 70 por el doctor en matemáticas Thomas L. Saaty.

Con el tiempo se transformó en una de las metodologías multicriterio de mayor aplicación práctica, ese es el motivo por el cual se seleccionó para la aplicación objeto de este trabajo. El AHP involucra todos los aspectos del proceso de toma de decisiones: Modela el problema a través de una estructura jerárquica, utiliza una escala de prioridades basada en la preferencia de un elemento sobre otro, de este modo combina la multiplicidad de escalas correspondientes a los diferentes criterios, sintetiza los juicios emitidos y entrega un ranking u ordenamiento de las alternativas de acuerdo a los pesos obtenidos (prioridades). Esta metodología propone una manera de ordenar el pensamiento analítico, de la cual destacan tres principios básicos:

- El principio de la construcción de jerarquías
- El principio del establecimiento de prioridades
- El principio de la consistencia lógica

Los sistemas complejos pueden ser mejor comprendidos mediante su descomposición en elementos constituyentes, la estructuración de dichos elementos jerárquicamente, y la composición o sintetización de los juicios, de acuerdo con la importancia relativa de los elementos de cada nivel de jerarquía más simples son lineales, ascendiendo o descendiendo de un nivel a otro.

Las jerarquías que trata el método de AHP son aquellas que conducen un sistema hacia un objetivo deseado como la solución de conflictos, un desempeño eficiente o la felicidad total. Cada conjunto de elementos en una jerarquía como la antes mencionada ocupa un nivel de la jerarquía, El nivel superior llamado Foco, consta solamente de un elemento: el objetivo amplio y global, Los niveles siguientes pueden tener cada uno diversos elementos, aunque su cantidad es generalmente de pequeña –entre cinco y nueve elementos. Debido a que los elementos de un nivel deberán compararse uno con el otro en función de un criterio del nivel superior siguiente, los elementos de cada nivel deben ser del mismo orden de magnitud. El segundo principio que destaca de este método multicriterio es el establecimiento de prioridades entre los elementos de la jerarquía. Se propone una escala de prioridades como forma de independizarse de las diferentes escalas que existen entre sus componentes.

Los seres humanos perciben relaciones entre los elementos que describen una situación, pueden realizar comparaciones a pares entre ellos con respecto un cierto criterio y de esta manera expresar la preferencia de uno sobre otro. La síntesis del conjunto de estos juicios arroja la escala de intensidades de preferencias (prioridad)

entre el total de elementos comparados. De esta forma es posible integrar el pensamiento lógico con los sentimientos, la intuición, (que es reflejo de la experiencia) los juicios que son ingresados en las comparaciones a pares responden a estos factores.

De acuerdo a lo anterior, el primer paso para establecer las prioridades es realizar comparaciones a pares entre elementos de un mismo nivel con respecto del elemento de nivel superior de que dependen. Las matrices de comparación resultan ser la forma más conveniente para esta etapa del proceso, en cada elemento de la matriz se ingresa el valor de la preferencia del elemento, por sobre el elemento. De acuerdo con el procedimiento matemático propuesto por la metodología, una vez completadas las matrices de comparación la obtención de las prioridades se transforma en un problema de vectores y valores propios (la justificación de esta aseveración se señala más adelante) donde el vector propio asociado al mayor valor propio de cada matriz de comparaciones representa el ranking u orden de prioridades mientras que el mayor valor propio es una medida de la consistencia del juicio.

El tercer principio del pensamiento analítico es la consistencia lógica. Los seres humanos tienen la capacidad de establecer relaciones entre los objetos o las ideas, de manera que sean consistentes – es decir, que se relacionen bien entre sí y sus relaciones muestren congruencia. En este sentido consistencia implica dos cosas: transitividad y proporcionalidad; la primera es que deben respetarse las relaciones de orden entre los elementos, es decir, si A es mayor que C y C es mayor que B entonces la lógica dice que A es mayor que B. La segunda es que las proporciones entre los ordenes de magnitud de estas preferencias también deben cumplirse con un rango de error permitido. Por ejemplo si A es 3 veces mayor que C y C es dos mayor que B entonces A debe ser 6 veces mayor que B, este sería un juicio 100% consistente (se cumple la relación de transitividad y de proporcionalidad).

La escala a que se hace referencia existe en el inconsciente, no está explícita y sus valores no son números exactos, lo que existe en el cerebro es un ordenamiento jerárquico para los elementos. Dada la ausencia de valores exactos para esta escala la mente humana no está preparada para emitir juicios 100% consistentes (que cumplan las relaciones de transitividad y proporcionalidad). Se espera que se viole la proporcionalidad de manera tal que no signifique violaciones a la transitividad.

FUENTE: "TOMA DE DECISIONES PARA LIDERES" (THOMAS SAATY).

INTENSIDAD	DEFINICIÓN	EXPLICACIÓN
1	Igual	Dos actividades contribuyen de igual forma al cumplimiento del objetivo
3	Moderada	La experiencia y el juicio favorecen levemente a una actividad sobre la otra
5	Fuerte	La experiencia y el juicio favorecen fuertemente una actividad sobre la otra
7	Muy fuerte o demostrada	Una actividad es mucho más favorecida que la otra; su predominancia se demostró en la práctica
9	Extrema	La evidencia que favorece una actividad sobre la otra, es absoluta y totalmente clara
2,4,6,8	Para transar entre los Valores anteriores	Cuando se necesita un compromiso de las partes entre valores adyacentes
Recíprocos	Si la actividad i se le ha asignado uno de los números distintos de cero mencionados cuando se compara con la actividad j, entonces j tiene el valor recíproco cuando se la compara con i ($a_{ij} = 1/a_{ji}$)	Hipótesis del método

Tabla 3. Escala fundamental para comparaciones a pares

En la tabla 1 se definen y explican los elementos que forman la escala recomendada para las comparaciones a pares entre los elementos de los niveles de la jerarquía, los valores en ella contenidos representan una escala absoluta, con los que se puede operar perfectamente. En todo caso cabe señalar que el método es independiente de la escala utilizada.

De esta manera el Análisis Jerárquico de Procesos integra aspectos cualitativos y cuantitativos en un proceso único de decisión, en el que es posible incorporar simultáneamente valores personales y pensamiento lógico en una estructura única de análisis de modo de convertir el proceso que ocurre naturalmente en nuestra mente en un proceso explícito, facilitando y promoviendo la toma de decisiones bajo escenarios multicriterios, promoviendo resultados más objetivos y confiables.

Un supuesto importante implícito en este modelo es que es necesario partir de la base que no existe la decisión "correcta e inmutable" (como si se tratara de un sistema de ecuaciones donde se deba despejar y encontrar el valor exacto de x), esto queda determinado por las percepciones de quienes participan en el proceso, de esta manera las decisiones son subjetivas y dependen de los valores y objetivos personales, del momento, etc. Hay que recordar que los juicios considerados en la evaluación están sujetos a las condiciones impuestas por el escenario existente al momento de realizar el análisis, las decisiones propuestas por el modelo son válidas para esa realidad, en ese instante de tiempo. Para otras circunstancias (otro escenario) es probable que la importancia relativa de los criterios sea diferente.

Procedimiento del Método Analítico Jerárquico

Para determinar la mejor decisión, el método AHP requiere:

1. Definición del problema: En esta etapa debe quedar claramente definido el objetivo general del proceso de decisión junto con los actores involucrados en él. Además se debe entregar una descripción del ambiente en que se desarrollará el estudio, sus características socio-económicas, ambientales, culturales, etc. dependiendo de los parámetros afectados por los proyectos en cuestión.
2. Definición de actores: Los participantes involucrados en el proceso de decisión, deben ser cuidadosamente seleccionados, ya que de estos depende la representatividad del resultado del modelo.

3. Estructurar el problema de decisión en un modelo de jerarquía (Jerarquizar): En esta etapa se debe construir una estructura jerárquica que involucre todos los aspectos de interés, para la jerarquización de las alternativas
4. Selección de las alternativas factibles: Dentro de todas las posibilidades de proyectos alternativos se seleccionan aquellos que son factibles de realizar bajo un punto de vista de análisis general, donde se consideran criterios tales como la factibilidad técnica o económica.
5. Construcción del modelo jerárquico: Se estructura el problema planteado en una jerarquía de criterios y alternativas. Para esto es necesario definir en una primera instancia los criterios estratégicos que participan en la decisión (Políticos, económicos, sociales, medioambientales, etc.).

Por lo general estos criterios son a nivel macro y representan los objetivos perseguidos por el proyecto. Una vez hecho esto, se procede a desglosar cada uno de los criterios definidos en la etapa anterior hasta llegar a un nivel de especificación que permita un fácil análisis y la comparación de las alternativas.

6. Ingreso de los juicios: En base a la información obtenida o a la percepción de los actores del proceso se ingresan los juicios para cada par de elementos. Se comienza del primer nivel, donde se encuentran los criterios estratégicos, se compara su importancia relativa con respecto del logro del objetivo general, luego se descende en los niveles jerárquicos, siempre realizando comparaciones de a pares referidos al nivel inmediatamente superior, hasta llegar al último nivel donde se encuentran las alternativas, las que son evaluadas en base a criterios técnicos más fáciles de tratar.
7. Síntesis de los resultados: Como se explicó en los párrafos anteriores, por medio de comparaciones entre pares de elementos con respecto a su nivel inmediatamente superior y, gracias a la propiedad de transitividad entre los elementos, es posible establecer un ranking de prioridades para las diferentes alternativas, ranking que, dependiendo de la problemática, enfrentada representa la decisión a adoptar.
8. Validación de la decisión: Para otorgar mayor confiabilidad a la decisión se debe establecer el rango de variación del peso relativo de los criterios estratégicos que soporta la decisión sin cambiar de alternativa propuesta, para esto se realiza un análisis de sensibilidad donde se analizan diversos escenarios posibles, determinando los puntos de corte para el peso de cada uno de los criterios.

Cabe destacar que todo este proceso debe estar muy bien documentado, cada una de las etapas debe contar con la información suficiente para su desarrollo y justificación.

Justificación intuitiva del Método Analítico Jerárquico

De acuerdo a lo establecido anteriormente sobre el proceso analítico jerárquico, una vez que se ha construido el modelo jerárquico, en donde se incorporen los diferentes criterios y alternativas relevantes para el proceso de decisión en cuestión y se han ingresado los juicios correspondientes a las comparaciones a pares entre los diferentes elementos del modelo, el problema se reduce al cálculo de valores y vectores propios los que representarán las prioridades y el índice de consistencia del proceso respectivamente. Por lo general se tiene:

$$A \cdot w = \lambda \cdot w$$

Donde:

A = Matriz recíproca de comparaciones a pares (juicios de importancia/preferencia de un criterio sobre otro)

W = Vector propio

λ = Máximo valor propio

Para la justificación del por qué las prioridades para un proceso de toma de decisiones, estructurado de esta forma, quedan determinadas por la solución de un problema de vectores y valores propios, es necesario recurrir a nociones básicas de matemáticas y de álgebra lineal. En primer lugar supóngase que es posible realizar mediciones exactas para los Juicios. Supongamos un conjunto de n bolitas de distinto peso.

Sea: w_i = Peso de la bolita c_i

a_{ij} = Juicio de las comparaciones emitidas, c_i es a_{ij} veces c_j .

En este caso en particular: $a_{ij} = w_i / w_j$ entonces se tiene:

$$[A] = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \frac{w_1}{w_3} \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \frac{w_2}{w_3} \dots & \frac{w_2}{w_n} \\ \frac{w_3}{w_1} & \frac{w_3}{w_2} & \frac{w_3}{w_3} \dots & \frac{w_3}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \frac{w_n}{w_3} \dots & \frac{w_n}{w_n} \end{bmatrix}$$

En este caso se cumple:

$$A * w = n * w,$$

Donde w es el vector propio de $[A]$.

En esta oportunidad se realizó una simplificación de lo que ocurre en los procesos de decisión, para los cuales es muy poco probable que los a_{ij} sean valores exactos ya que representan juicios humanos emitidos en la mayoría de las veces como variables cualitativas.

El hecho de asumir que a_{ij} es un valor exacto es conceptualmente erróneo si se considera que lo que pretende este valor es representar la percepción de la persona que toma la decisión. Continuando con el ejemplo, supóngase ahora que no es posible obtener un valor exacto para el peso de cada bolita, w_i , en el caso exacto se tiene que:

$a_{ij} = w_i / w_j$ entonces $w_i = a_{ij} * w_j$. En el caso de aproximado se tiene $w_1 \approx a_{12} * w_2 \approx a_{13} * w_3 \approx a_{1n} * w_n$, Al aplicar la ley de los grandes números:

$$\text{En general: } w_i = (1/n) * \sum a_{ij} * w_j$$

(el promedio de la muestra es el mejor valor).

En este caso se cumple con el caso ideal, pero se restringe demasiado el sistema para encontrar el vector de prioridades W que cumpla con estas relaciones, por lo tanto es necesario relajar un poco más el sistema. Para el caso de buenas estimaciones a_{ij} tiende al valor w_i / w_j por lo que se dice que

a_{ij} es una pequeña perturbación de w_i / w_j . Si a_{ij} cambia, w_i y w_j pueden cambiar para ajustarse al nuevo valor, si n también cambia. El problema queda planteado entonces como:

$$W_i = \left(\frac{1}{\lambda_{\max}} \right) * \sum_j a_{ij} * W_j \quad (5.2)$$

Donde λ_{\max} es una relajación de n , de esta manera

$$\sum_j a_{ij} * W_j = \lambda_{\max} * W_i \quad (5.3)$$

Y matricialmente

$$A * w = \lambda * w \quad (5.4)$$

Este es precisamente el problema de valores y vectores propios planteado al comienzo de la sección. Dado que la matriz $[A]$ es recíproca ($a_{ij} = 1/a_{ji} \forall i,j$) el problema tiene solución y es único, además esto asegura su estabilidad (pequeñas perturbaciones en a_{ij} generan modificaciones menores en λ_{\max} y W).

La solución $a_{ij} = w_i / w_j$ al problema $[A] * W = n * W$ supone total consistencia de los juicios, hecho que en la realidad es muy difícil de encontrar, al transformar el problema de modo que $[A'] * W' = \lambda_{\max} * W'$, se está frente a una situación más realista. Aquí λ_{\max} es el mayor valor propio de $[A']$ y $[A']$ es la matriz recíproca de las comparaciones a pares formada por los a_{ij} que constituyen estimaciones de los w_i / w_j .

Axiomas para el Método Analítico Jerárquico

Es importante destacar, que el modelo jerárquico empleado para un determinado proceso de decisión debe cumplir cuatro axiomas principales. Sólo de esta manera son posibles y tienen real sentido los fundamentos teóricos expuestos en el punto anterior.

RECIPROCIDAD.

Axioma 1: Dadas dos alternativas A_i y $A_j \in A \times A$, la intensidad de la preferencia de A_i sobre A_j es inversa a la intensidad de preferencia de A_j sobre A_i .

HOMOGENEIDAD

Axioma 2: Cuando se comparan dos alternativas, el tomador de decisión nunca juzga a una como infinitamente superior a la otra, bajo ningún criterio. De otra forma; para comparar dos elementos de acuerdo a un criterio dado, hay que disponer de una escala acotada.

Bajo el contexto metodológico, la homogeneidad apunta a que los elementos a comparar deben ser de un mismo orden de magnitud. Al construir el modelo jerárquico todos los elementos hijos (que se desprenden) de un determinado elemento deben ser parecidos, si en un modelo, los subcriterios que describen completamente al criterio padre, no cumplen esta condición es necesario agregar niveles intermedios para separar los distintos órdenes de magnitud.

DEPENDENCIA

· Axioma 3: Los problemas de decisión pueden ser formulados como una jerarquía.

Este axioma apunta a la posibilidad de comparar elementos en la jerarquía, es necesario tener controlada la dependencia entre los elementos de dos niveles consecutivos (externo - dependencia) y dentro de un mismo nivel (interno dependencia). Un modelo jerárquico se caracteriza por que sus elementos tienen externo dependencia unidireccional, es decir, los hijos dependen de sus padres y no hay relación entre ellos.

EXPECTATIVAS

Axioma 4: La jerarquía es un modelo que representa todos los criterios y alternativas.

Este axioma está relacionado con la necesidad de agregar o eliminar alternativas a modo de representar fielmente la percepción de los actores involucrados en el proceso de decisión.

Los axiomas anteriores constituyen el marco que delimita como abordar las dos principales tareas del método AHP: formular y resolver el problema como una jerarquía y explicitar los juicios en forma de comparaciones de a pares.

La explicitación de prioridades para cierto set de alternativas bajo un criterio dado, implica completar una matriz de $n \times n$, donde n es el número de alternativas a comparar. Sin embargo, dado que las comparaciones son recíprocas por el Axioma 1, sólo se necesita hacer $n(n-1)/2$ comparaciones de a pares.

Respaldo Teórico del Método Analítico Jerárquico

Además de los axiomas sobre los que se sustenta la metodología, los procedimientos de cálculo de los ponderadores y la verificación de la consistencia de los mismos, están respaldados por teoremas completamente demostrados y aceptados en el ambiente de las ciencias desde hace mucho tiempo.

Los principios del proceso analítico jerárquico derivan de la economía, la investigación de operaciones, teoría de grafos, teoría de organizaciones, teoría de las medidas y el álgebra lineal, siendo esta rama de las matemáticas uno de sus principales pilares; en consideración con esto, a continuación se señalan algunos de los teoremas que respaldan el método.

Def: Una matriz es recíproca si se cumple:

$$[A] = a_{ij}, \text{ donde } a_{ij} > 0 \forall i, j \quad a_{ij} = 1/a_{ji}$$

Def : Una matriz $[A]$ se dice consistente si cumple: $A_{ik} = a_{ij} a_{jk} \forall i, j$

TEOREMA 1: (PERRON – FROBENIUS)

Si $[A]$ es recíproca positiva, entonces:

- a) $[A]$ tiene un valor propio real, positivo y simple λ_{max} , que no es excedido en el módulo por ningún otro valor propio, real o complejo.
- b) Todas las componentes del vector propio w asociado a λ_{max} , son positivas.
- c) λ_{max} , (conocido como la raíz de Perron) cumple con la desigualdad:
- d) $\text{Mini} \{(Aw)_i / w_i\} < \lambda_{max} < \text{Máxi} \{(Aw)_i / w_i\}$

TEOREMA 2:

λ_{max} se encuentra entre los valores de la menor y la mayor de las sumas de las filas de $[A]$ es decir:

$$\text{Mini} (\sum a_{ij}) < \lambda_{max} < \text{Máx} (\sum a_{ij})$$

TEOREMA 3:

Para una matriz $[A]$ en que $a_{ij} = 1 \forall i, j$ se tiene que:

$$\lambda_{max} \geq \text{Rango } [A]$$

TEOREMA 4:

Sea λ_{max} el mayor valor propio de $[A]$, y w el vector propio asociado si $e^t = (1, 1, \dots, 1)$ entonces:

$$\{W\} = \lim_{k \rightarrow \infty} \frac{[A]^{k*} e}{e^{t*} [A]^{k*} e} \quad (5.5)$$

este límite siempre existe para este tipo de matriz.

TEOREMA 5:

a) Si $[A]$ es una matriz recíproca consistente, el vector propio principal está dado por cualquiera de sus columnas.

b) Si $[A]$ es recíproco inconsistente, el vector propio principal w está dado por:

$$\{W\} = \lim_{k \rightarrow \infty} \frac{[A]^{k*} h_j}{\sum_j [A]^{k*} h_j} \quad (5.6)$$

Con $h, j, k = 1, \dots, n$

Este teorema entrega un método aproximado para calcular el vector propio:

a) Multiplicar $[A]$ por sí misma las veces que sea necesario.

b) Sumar sobre una columna.

c) Normalizar.

Significado del Vector Propio Principal

El vector propio principal se asocia con la idea de dominio y el valor propio con la idea de consistencia de los juicios. La teoría de grafos explica también esta situación:

$$\text{Sea } [A] = \begin{bmatrix} 1 & 2 & a \\ 1/2 & 1 & 3 \\ 1/a & 1/3 & 1 \end{bmatrix}$$

El grafo que representa esta matriz, contiene el arco extra entre los nodos i, j los valores a_{ij} esto se denomina intensidad del arco i, j

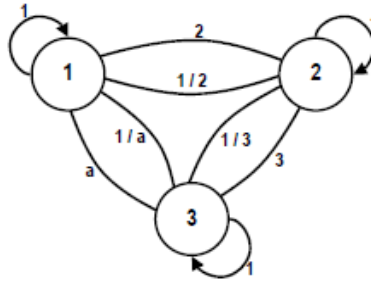
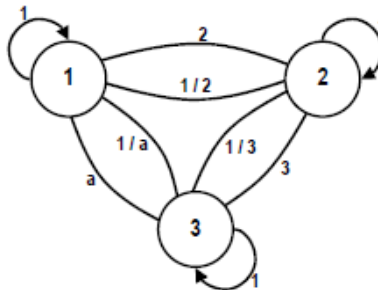


Figura 16. Grafo que muestra la matriz A con sus arcos y nodos

“La prioridad se aprecia como la intensidad del recorrido y es definida como el producto de las intensidades de los arcos del recorrido”. La matriz [A] puede interpretarse como la matriz con la intensidad de los recorridos de largo 1

Si $a = 6$



¿Cómo ir de 1 a 1 en dos pasos?

$$1 - 1 - 1 = 1 * 1 = 1$$

$$1 - 2 - 1 = 2 * 1/2 = 1$$

$$1 - 3 - 1 = 1/6 * 6 = 1$$

Luego la intensidad de preferencia del nodo 1 con respecto del nodo 1, considerando los efectos de 2° orden de los otros nodos es 3.

¿Cómo ir de 1 a 2 en dos pasos?

$$1 - 1 - 2 = 1 * 2 = 2$$

$$1 - 2 - 2 = 2 * 1 = 2$$

$$1 - 3 - 2 = 6 * 1/3 = 2$$

Entonces $a_{12}^2 = 6$, por lo tanto la intensidad de preferencia del nodo 1 con respecto al nodo 2 considerando los efectos de 2° orden es 6.

Denotando $[A]^k = a_{ij}^k$ la matriz de las intensidades de preferencia de recorridos de largo k se tiene:

$$\text{Sea } [A] = \begin{bmatrix} 1 & 2 & 6 \\ 1/2 & 1 & 3 \\ 1/6 & 1/3 & 1 \end{bmatrix}$$

$$\text{Sea } [A]^2 = \begin{bmatrix} 3 & 6 & 18 \\ 3/2 & 3 & 9 \\ 1/2 & 1 & 3 \end{bmatrix}$$

$$\text{Sea } [A]^3 = \begin{bmatrix} 9 & 18 & 54 \\ 9/2 & 9 & 27 \\ 1/2 & 3 & 9 \end{bmatrix}$$

A medida que aumenta k se consideran más y más interacciones entre los nodos.

La columna j de $[A]^k$ contiene la intensidad de preferencia global a lo largo de recorridos de largo k vistos desde el nodo j . La columna j normalizada, representa la importancia relativa de los demás nodos a lo largo de recorridos de largo k con respecto del nodo j . En el caso consistente ($a = 6$), cada columna de $[A]^k$ normalizada representa las preferencias relativas que se obtienen al normalizar las columnas de $[A]^k$.

Si $[A]$ no es consistente, tomando $a = 4$ por ejemplo, entonces los recorridos de largo 1 tienen distintas intensidades relativas, ya que dependen del nodo utilizado como referencia.

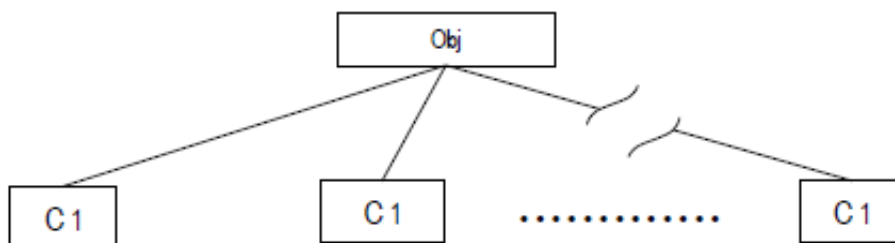
Así a medida que $k \rightarrow \infty$ las columnas normalizadas de $[A]^k$ $\{0,6; 0,3; 0,1\}$ tienden a los pesos w^k $\{0,558; 0,32; 0,122\}$ de los nodos del grafo.

Es decir, el vector principal está dado por el límite de las intensidades normalizadas de recorridos de largo k

$$w_i = \lim_{k \rightarrow \infty} \frac{a_{ij}(k)}{\sum_j a_{ij}(k)} \quad i, h=1, \dots, n \quad (5.7)$$

Relación entre consistencia y prioridades

Sea la jerarquía H :



Y la matriz de comparaciones a pares de criterios $[A]$, correspondiente a:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \quad (5.8)$$

Donde a_{ij} indica la importancia de c_i con respecto a c_j , en cuanto a la característica o propiedad del elemento inmediatamente superior de la jerarquía. Si los juicios del actor fueran exactos se cumpliría:

$A_{ik} = a_{ij} a_{jk} \forall i, j$; es decir, la matriz de las comparaciones a pares $[A]$ sería consistente. La consistencia impone 2 propiedades simultáneamente.

- a) Transitividad de las preferencias: Los juicios emitidos deben respetar las condiciones de transitividad que se producen al comparar más de dos elementos, es decir: si C_1 , es mejor que C_2 , y C_2 es mejor que C_3 entonces se espera que C_1 , sea mejor que C_3 .
- b) Proporcionalidad de las preferencias: Juicios enteramente consistentes implican que además de verificarse las relaciones de transitividad se respeta la proporcionalidad entre ellos, es decir: si C_1 , es 3 veces mejor que C_2 , Y C_2 es 2 veces mejor que C_3 entonces se espera que C_1 sea 6 veces mejor que C_3

$[A]$ es consistente cuando las comparaciones a pares se basan en medidas exactas, es decir, cuando los pesos W_1, \dots, W_n . Son conocidos y se obtiene $a_{ij} = w_i / w_j$. En este caso es necesario comparar $n-1$ objetos y los a_{ij} restantes son factibles de determinar mediante la relación de consistencia. Por ejemplo, de la propiedad b) se puede deducir que si:

$$C_1 = 4C_2 \text{ y } C_1 = 2C_3 \text{ entonces } C_2 = \frac{1}{2} C_3$$

La situación hasta aquí descrita es el caso ideal, pero en la práctica los juicios humanos casi nunca son perfectos, es muy difícil disponer de medidas exactas para los w_i , sobre todo en procesos de toma de decisiones donde, por lo general, existe una gran cantidad de variables cualitativas. Hay que recordar además, que los criterios de primer nivel de una jerarquía son de carácter estratégico.

En este tipo de comparaciones la prioridad no queda determinada por una medida absoluta si no que depende exclusivamente de las apreciaciones de quienes participen en el proceso. Lo que pretende el método es capturar la percepción de los actores, es eso lo que se refleja en los resultados, dos personas enfrentadas a la misma situación no necesariamente tienen la misma percepción de ella, tienen objetivos distintos, valores diferentes, etc. lo que en definitiva se traduce en juicios y prioridades distintas.

El concepto de consistencia va asociado a la existencia en nuestra mente de una escala implícita para cada set de juicios que se utilizan para expresar la intensidad de las preferencias al hacer las comparaciones a pares, los juicios emitidos por los actores son estimaciones de las proporciones que se derivan de dicha escala. Para un set de comparaciones, violaciones a la consistencia, implican generalmente violaciones a la proporcionalidad no significando necesariamente violaciones a la transitividad. Dado que la probabilidad de que una persona emita juicios perfectos es baja (nuestro cerebro no está programado para ser 100% consistente, de otro modo no sería capaz de integrar nueva información y cambiar las relaciones existentes), es necesario permitir un pequeño margen de inconsistencia, a medida que aumenta el número de elementos a comparar (no debe ser superior a nueve, cantidad máxima que es capaz de manejar efectiva y simultáneamente el cerebro) el nivel de inconsistencia permitido aumenta, pero en ningún caso debe exceder el 10%.

La relación de reciprocidad es absolutamente necesaria para que la consistencia se cumpla, si se viola esta relación es imposible obtener consistencia. Si es recíproca y consistente implica que existe un vector (w) donde:

$$W_i / w_j = a_{ij} \text{ y } [A]^*(w) = n^*(w)$$

Dado que $\sum \lambda_i = \sum a_{ij} = n$ (teoría de matrices) y que el número de valores propios no nulos de una matriz $[A]$ recíproca y consistente es el rango $[A] = 1$ entonces se tiene que n es el "único" valor propio de $[A]$. Al considerar la pérdida de consistencia de la matriz $[A]$ como una perturbación se genera $[A']$, para este caso se cumple que:

$[A]^* (w') = \lambda_{\max} (w')$ con (w') vector propio principal y $\lambda_{\max} \geq n$

Pequeñas variaciones en a_{ij} implican pequeñas perturbaciones en λ_{\max} por lo que la desviación de λ_{\max} con respecto de n es una buena medida de la consistencia de los juicios.

Consistencia $\Leftrightarrow a_{ij} \cdot a_{jk} = a_{ik}$

$$\Leftrightarrow \lambda_{\max} = n = \sum_{i=1}^n \lambda_i$$

$$\text{Si } \lambda_i = \lambda_{\max} \Rightarrow n = \sum_{i \neq j} \lambda_i + \lambda_j \Leftrightarrow \sum_{i \neq j} \lambda_i = 0$$

Si no hay consistencia:

$$\sum_{i \neq j} \lambda_i \neq 0 \tag{5.9}$$

Por lo tanto: se define "Índice de Consistencia" como:

$$IC = \frac{\lambda_{\max} - n}{(n-1)} \tag{5.10}$$

Es decir, el promedio de los valores propios distintos de λ_{\max} es un buen indicador del grado de consistencia. IC mide la dispersión de los juicios del actor en la matriz [A].

El 10% de inconsistencia aceptado, es en relación a la comparación entre el del índice de [A] con el valor de un índice de referencia generado en forma estadística de un conjunto extenso de matrices recíprocas, que se han llenado con juicios aleatorios. La inconsistencia del actor (RC) se mide como: IC /IR, donde IR es el dato estadístico del índice IC que se obtiene cuando las comparaciones de a pares se generan al azar. Se considera que la consistencia del actor es aceptable cuando $RC < 10\%$, este valor depende también de la dimensión de la matriz de comparaciones, es decir, del número de comparaciones (NC) necesarias para llenar totalmente la matriz. Este número está dado por:

$$NC = \frac{n \cdot (n-1)}{2} \tag{5.11}$$

Donde:

NC: número de comparaciones

n: dimensión de la matriz.

Las evaluaciones a pares deberán realizarse nuevamente en caso de que la inconsistencia este por sobre el 10%. El ejercicio recurrente de pesar mejora

sustancialmente la consistencia, haciendo que en pocas repeticiones se consiga influir sobre el índice de consistencia.

Métodos de la Programación Multiobjetivo

La programación multiobjetivo u optimización vectorial se enfrenta al problema de optimizar simultáneamente varios objetivos sujetos a un conjunto de restricciones, usualmente lineales. Como desde un punto de vista tanto lógico como operativo es imposible definir un óptimo cuando existen varios objetivos, la programación multiobjetivos pretende establecer el conjunto de soluciones eficientes (no dominadas o Pareto óptimas) en vez de buscar un único óptimo.

El conjunto eficiente está formado por soluciones factibles (esto es, que cumplen las restricciones) tales que no existe otra solución factible que proporcione una mejora en un objetivo sin producir un empeoramiento en al menos otro de los objetivos. Una vez establecido que el propósito de la programación multiobjetivo consiste en generar el conjunto eficiente, el problema puede plantearse dentro del siguiente marco general:

$$\text{Eff}Z(x)=[Z_1(x), Z_2(x), \dots, Z_q(x)] \quad (5.12)$$

sujeto a:

$$x \in F$$

donde Eff significa la búsqueda de soluciones eficientes y F representa el conjunto posible o alcanzable.

Cuando el número de variables de decisión es mayor que dos, no puede aplicarse el método gráfico que acabamos de exponer. En tales casos es necesario recurrir a técnicas multiobjetivo que permitan generar o al menos aproximar el conjunto eficiente.

Básicamente existen tres enfoques para efectuar esta tarea:

- a) el método de las restricciones (constraint method),
- b) el método de los coeficientes de ponderación (weighting method)
- c) el simplex con objetivos múltiples (multiobjective simplex).

De estos tres enfoques el único que permite obtener con toda exactitud el conjunto eficiente es el simplex con objetivos múltiples. No obstante, la aplicabilidad de este procedimiento queda limitada a problemas multiobjetivo de muy reducido tamaño, siendo su interés práctico muy escaso. Ahora bien, antes de proceder a realizar dicha tarea vamos a introducir un concepto esencial en programación multiobjetivo que es la

idea de tasa de intercambio (tradeoff) entre objetivos. La tasa de intercambio entre dos objetivos mide el sacrificio necesario en un cierto objetivo, para poder compensar un incremento unitario en la realización del otro objetivo. Así, si tenemos dos soluciones eficientes x^1 y x^2 la tasa de intercambio T_{jk} entre los objetivos j -ésimo y k -ésimo viene dada por:

$$T_{jk} = \frac{f_j(x^1) - f_j(x^2)}{f_k(x^1) - f_k(x^2)} \quad (5.13)$$

Donde $f_j(x)$ y $f_k(x)$ representan los niveles alcanzados por los correspondientes objetivos. El conocimiento de los valores de las tasa de intercambio puede ser de una gran utilidad a la hora de elegir la solución eficiente preferida por el centro decisor.

5.5 Método de restricciones y de los coeficientes de ponderación

La idea básica de este enfoque consiste en optimizar uno de los objetivos, mientras que a los restantes se les incluye en el conjunto de las restricciones. Por medio de variaciones paramétricas de los valores asignados a los términos independientes de las restricciones en que se han convertido los objetivos, se va generando el conjunto eficiente. Marglin (1967 pp. 24-25) fue el primero en proponer este enfoque. Así, para un problema multiobjetivo considerados la estructura del modelo a formular dentro del método de las restricciones sería:

Maximizar $Z_k(x)$

Sujeto a:

$$x \in F$$

$$Z_j(x) \geq L_j \quad j=1,2,\dots,k-1,k+1,\dots,q$$

Siendo el objetivo k -ésimo el objetivo a optimizar. Por medio de variaciones paramétricas de los términos independientes L_j se va generando el conjunto de soluciones eficientes. Estos valores constituyen los límites del campo de variación del parámetro λ .¹⁹ Por medio de variaciones del valor de este parámetro se van generando puntos eficientes que permiten aproximar el conjunto de soluciones Pareto óptimas.

² En problemas multiobjetivo con más de dos variables de decisión, los límites de variación de cada objetivo pueden obtenerse procediendo a la optimización por separado de cada objetivo. El límite máximo del objetivo j -ésimo vendrá dado, obviamente, por el máximo de dicho objetivo y el mínimo por el valor más pequeño que toma el objetivo j -ésimo cuando se optimizan los otros objetivos.

La idea básica de este método consiste en combinar todos los objetivos en una única función. Con este propósito a cada objetivo se le asocia un peso o coeficiente de ponderación, sumándose a continuación todos los objetivos. Seguidamente se someten los diferentes pesos a un tratamiento paramétrico, generándose de esta forma una aproximación del conjunto eficiente.

Zadeh (1963) fue el primero en proponer este enfoque. Así, para un problema multiobjetivo en el que se desea maximizar los “q” objetivos, la aplicación del método de los coeficientes de ponderación conducirían a la formulación del siguiente modelo:

$$\text{Maximizar } W_1Z_1(x) + W_2Z_2(x) + \dots + W_qZ_q(x)$$

Sujeto a:

$$x \in F$$

$$W \geq 0$$

Por medio de variaciones paramétricas de los coeficientes W se va generando el conjunto eficiente. Igual que sucedía con el método de las restricciones, el método de los coeficientes de ponderación requiere efectuar $pq-1$ “pasadas” del programa lineal correspondiente, siendo p en este caso el número de valores que se asigna a los coeficientes de ponderación.

La fuerte demanda en tiempo de computador que exige este método puede paliarse recurriendo a opciones paramétricas. No obstante, como en este caso la parametrización se realiza con los coeficientes de la función objetivo en vez de con los términos independientes, las dificultades de cálculo son mayores, por lo que puede decirse que en términos operativos el método de las restricciones presenta una cierta superioridad con respecto al método de los coeficientes de ponderación.

Como acabamos de ver los dos métodos que existen para generar o al menos aproximar el conjunto eficiente demandan un tiempo de computadora que medido en número de “pasadas” del programa crece casi exponencialmente con el número de objetivos. Así, un problema con cinco objetivos y en el que el número de valores que se asigne a los coeficientes de ponderación o bien el número de subintervalos en que se divide el campo de variación de los términos independientes fuera de seis, requeriría 64

= 1296 “pasadas” de computadora y generaría 1296 puntos eficientes. Obviamente esta situación no es deseable, tanto por el excesivo coste del proceso de cálculo, como por la sobrecarga de información que supone para el centro decisor, resultando prácticamente imposible para él, en tales casos, elegir la solución óptima de entre el conjunto eficiente.

Para mitigar estos importantes problemas prácticos se han diseñado diferentes métodos. Así, Steuer (1976) ha sugerido utilizar el método de los coeficientes de ponderación en vez de con coeficientes fijos con coeficientes variables pertenecientes a un determinado intervalo. De esta manera, solo se analiza la parte del conjunto eficiente de mayor importancia para el centro decisor. Con este método se consigue ahorrar una gran cantidad de tiempo de computador y reducir considerablemente el tamaño del conjunto eficiente.

Steuer & Harris (1980) han recomendado el uso de técnicas de filtrado con el propósito de descartar soluciones eficientes no suficientemente diferentes con respecto a otras soluciones eficientes previamente calculadas y retenidas por el filtro. De esta forma es posible reducir el tamaño del conjunto eficiente a una dimensión muy manejable (operación de poda). Esta técnica de filtrado ha sido aplicada con éxito por Romero et al. (1987) en el campo de la planificación agraria a un problema de decisión relacionado con un programa de reforma agraria.

Otra forma de mitigar este tipo de problemas consiste en el uso de técnicas interactivas. Este enfoque está basado en la definición progresiva de las preferencias del centro decisor y el modelo. La interacción consiste en una especie de conversación en la que al centro decisor se le preguntan sus preferencias o tasas de intercambio (tradeoffs) entre los diferentes objetivos. Una exposición de las técnicas interactivas excede de las pretensiones de este trabajo. No obstante, una buena revisión de este tipo de enfoques puede verse en Hwang & Masud (1979, páginas 102-226).

5.6 Programación Compromiso

En este modelo se acerca la posición óptima del decisor a un ideal, definido previamente por dicho centro decisor (empresario o analista). El ideal se define como un punto o cesta cuyas coordenadas son las mejores oportunidades para cada una de las variables

consideradas en el modelo (beneficio empresarial, volumen de ventas, penetración en cada mercado, plantilla de trabajadores, etc.). La aproximación del punto o cesta decisional al punto ideal se consigue recurriendo a un modelo de distancias con diversas métricas (entre uno e infinito).

El punto elegido se encuentra obviamente sobre el conjunto de oportunidades (frontera eficiente) y su distancia al ideal es mínima. El problema computacional se resuelve mediante programas de computadora. Las soluciones correspondientes a las distintas métricas se hayan situadas en un subconjunto del conjunto de oportunidades, de acuerdo con la acotación de Yu (1973), cuando se trata de un bicriterio.

Zeleny introdujo en 1973 un interesante método con el nombre de programación compromiso, que ayuda al centro decisor a elegir la solución óptima de entre un conjunto de soluciones eficientes. Este método comienza estableciendo lo que Zeleny llama el punto ideal. Las coordenadas de este punto vienen dadas por los valores óptimos de los diferentes objetivos del centro decisor. El punto ideal es normalmente inalcanzable. Si fuera alcanzable no existiría conflicto entre los objetivos. Cuando el punto ideal es inalcanzable el elemento óptimo o solución compromiso para el centro decisor viene dada por la solución eficiente que se encuentra mas próxima con respecto al punto ideal (axioma de elección de Zeleny). Dependiendo de la medida de distancia utilizada puede establecerse un conjunto de soluciones compromiso.

Como el punto ideal es inalcanzable (esto es, queda fuera el dominio factible F) es necesario obtener soluciones compromiso. Para ello necesitamos calcular la distancia existente entre cada punto de la curva de transformación. Con este propósito se introduce el concepto de grado de proximidad d_j entre el objetivo j -ésimo y su ideal por medio de:

$$d_j = Z_j^* - Z_j(x) \tag{5.14}$$

cuando el objetivo j -ésimo se minimiza. En ambos casos Z_j^* representa el punto ideal. Cuando las unidades en que vienen medidos los objetivos son diferentes, deben utilizarse en el análisis desviaciones relativas en vez de desviaciones absolutas (Zeleny 1973, pagina 299). En tales casos el grado de proximidad viene dado por

$$d_j = \frac{Z_j^* - Z_j(x)}{Z_j^* - Z_{*j}} \tag{5.15}$$

Donde Z^*j es el anti-ideal para el objetivo j -ésimo; esto es, el valor de objetivo j -ésimo cuando el objetivo en conflicto se optimiza. Con objeto de obtener las distancias entre cada solución y el punto ideal, la programación compromiso introduce la siguiente familia de funciones de distancia:

$$L_p(\partial, k) = \left[\left(\sum_{j=1}^k \partial_j d_j \right)^p \right]^{1/p} \dots\dots\dots(5.15)$$

Donde los coeficientes ∂_j ponderan la importancia de la discrepancia entre el objetivo j -ésimo y su valor ideal.

En la programación compromiso la función (7) usualmente se opera para las métricas $p=1$ (la distancia “más larga” en un sentido geométrico), $p=2$ (la distancia “mas corta” en un sentido geométrico; esto es, la línea recta) y $p=\infty$ (la distancia “chebysev”). Para $p>2$ las medidas de distancia son incluso mas cortas que la línea recta. El uso de (7) para $p>2$ tiene sentido, pues las funciones de distancia no se utilizan en este contexto en un sentido geométrico, sino como una medida de proximidad entre cada solución y el punto ideal como indicador de las preferencias humanas²⁰

Es interesante observar que aumentos en el valor de p suponen dar una mayor ponderación a la desviación mayor. Así, cuando $p=\infty$, se procede a minimizar a máxima desviación individual. Dicho con otras palabras, el factor p pondera las diferentes desviaciones de acuerdo con sus magnitudes, mientras que el factor ∂_j pondera las desviaciones de acuerdo con la importancia relativa de cada objetivo. Para diferentes valores de p y ∂_j se pueden generar diferentes soluciones compromiso. La programación compromiso puede ganar en eficiencia y flexibilidad si se opera dentro de un enfoque de tipo continuo en vez de discreto. Así, cuando se utiliza un enfoque continuo, no es necesario calcular previamente el conjunto eficiente, pues las diferentes soluciones compromiso se obtienen directamente resolviendo problemas de programación matemática.

²⁰ Un interesante desarrollo de las funciones de distancia en el campo de la Teoría Económica se debe a Ballesteros & Blanco (1984). Estos autores, recurriendo a la programación por puntos ideales formulan una nueva teoría del consumo no paretiana exenta de supuestos utilitarios

Yu (1973) demostró que las métricas L_1 y L_∞ definen un subconjunto del conjunto eficiente al que Zeleny (1974, página 488) denominó el conjunto compromiso. Las otras “mejores” soluciones compromiso (para $1 \leq p \leq \infty$) están situadas entre las métricas L_1 y L_∞ .

CAPITULO VI. CASO PRÁCTICO

El presente capítulo constituye un núcleo esencial de la tesis. En él se abordarán las técnicas usadas para la selección del portafolio desde la perspectiva de los algoritmos genéticos y la programación compromiso. El capítulo se divide en dos partes, en la primera se aplica el algoritmo genético a un portafolio de inversión seleccionado de un conjunto de operadoras de fondos de inversión, y en la segunda, se aplica la programación compromiso al mismo portafolio de inversión constituido.

6.1 Selección del portafolio de inversión

Recordando que el objetivo principal de la tesis es comparar los algoritmos genéticos con la programación compromiso. Para este trabajo de tesis se seleccionaron 20 operadoras de sociedades de inversión con el mayor rendimiento²¹ en el mercado mexicano de valores según la bolsa Mexicana de Valores (www.bmv.com.mx) en el periodo comprendido entre 31 de agosto de 2007 al 30 de septiembre de 2009.

PORTAFOLIO 1	PORTAFOLIO 2	PORTAFOLIO 3	PORTAFOLIO 4
VALUEF2 A	INGCORP A	ING-1 BM2	INGCORP BOM
VALMX17 A	INGCORP BOI	ING-1 A	APOLO90 A
INGCORP BOE 1	INGCORP BOE2	HZMD A	ACTIREN B5
FINDE1 E	HZMD+ A	INGCORP BOE4	FINDE1 F
MONEXM+ BEC-0	ING-30 A	AMEX-MD A	APOLOM A

Tabla 4. Portafolios seleccionados para el estudio de la tesis

Cabe mencionar que en la construcción de cada portafolio se hizo de manera aleatoria, esta situación no afecta al estudio.

6.2 Metodología utilizada para el Algoritmo Genético

Como se han definido cuatro portafolios se eligen cinco posibles formas de distribuir el capital inicial que es de 5,000,000 millones²² para cada portafolio que se han seleccionado al azar²³, se considera una distribución por cada portafolio y no por cada sociedad. Se establece como función de adaptación la suma de las colocaciones de capital por cada portafolio

²¹ Se considera el mayor rendimiento durante todo el periodo y se construyeron 4 portafolios de 5 opciones

²² Se consideran al inicio del proceso.

²³ La suma de las distribuciones de capital no necesariamente es 5 millones, esto porque se toman en forma aleatoria

A continuación se resume el proceso del algoritmo genético; se elige al mejor individuo y al peor de acuerdo a la función de adaptabilidad, en este caso el mejor portafolio de inversión es la que se acerca más a 5 millones de capital inicial por cada portafolio y el peor el que más se aleja.

Estos datos fungirán como el padre y la madre. Cada asignación por portafolio se convierte a número binario, esto para lograr el intercambio de genes. Para lograr el cruce entre los cromosomas de los padres, es necesario separar los dígitos binarios obtenidos en el paso anterior.

Por la estructura de las asignaciones, es decir los valores que deben tomarse para cada asignación en cada portafolio (1 a 5 millones) los números binarios correspondientes a estas cifras solo tienen 3 dígitos por mucho.

Por lo cual el cromosoma de cada padre deberá tener 12 dígitos (3 dígitos x 4 portafolios). El punto de cruce se establece al final del tercer y noveno cromosoma. Con este punto de cruce definido, tan solo se intercambia la información de los 6 cromosomas centrales de cada padre en cada uno de sus dos hijos. Para realizar una evaluación y mutación se concatena dos o más cadenas en una cadena de texto.

Donde son de 2 a 255 elementos de texto que se unirán en un elemento de texto único. Los elementos de texto pueden ser cadenas de texto, números o referencias a celdas únicas. Posteriormente después de concatenar los dígitos se convierte un número binario en decimal. En los dos hijos se localiza la mayor asignación de capital y aleatoriamente se selecciona el cromosoma a cambiar.

Después de realizar la mutación anterior, solo se convierten los cromosomas a decimales y se evalúan en la función de adaptabilidad. El algoritmo se detiene cuando todas las distribuciones de capital de acuerdo a la función de adaptabilidad son menores que 5 millones para cada portafolio formado por las sociedades de inversión anteriormente citadas.

A continuación se muestra un diagrama del procedimiento anterior:

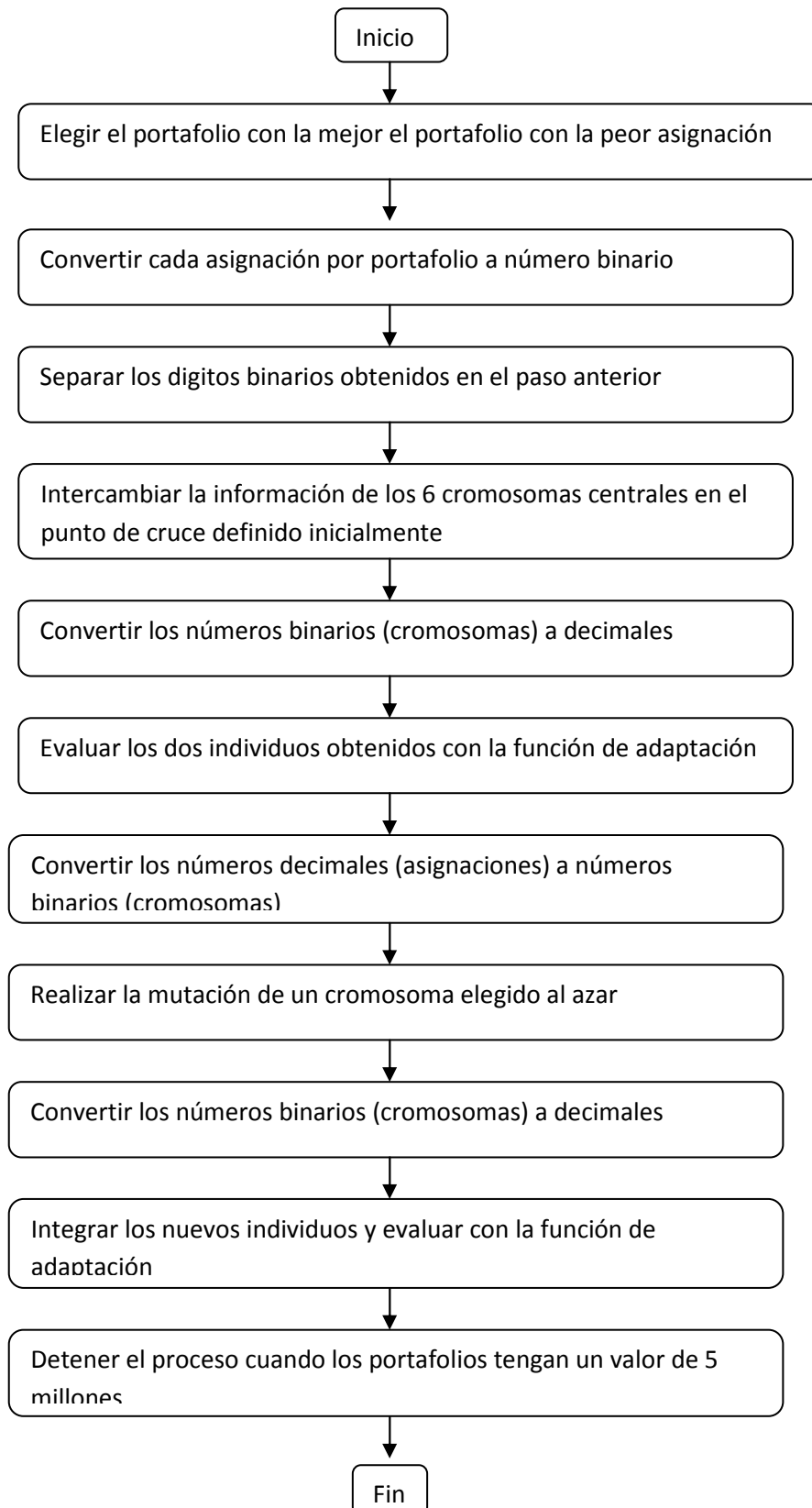


Diagrama de flujo para la Metodología de Algoritmo Genético

6.3 Metodología utilizada para la Programación Compromiso

Antes de aplicar el método de programación compromiso, es indispensable establecer los siguientes índices para cada portafolio de inversión

1. *INDICE BETA*²⁴

Considerando los precios y valores de cada sociedad de inversión para el periodo del 31 agosto de 2007 al 30 de septiembre de 2009, se calculan rendimientos para cada sociedad de inversión y el IPC. Se obtienen promedios de los rendimientos para cada sociedad de inversión y el IPC. Se aplica la siguiente formula para calcular la beta de cada sociedad de inversión $\beta = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(X_i - \bar{X})(Y_i - \bar{Y})}$. Se agrupan las betas por bloques de portafolios para obtener un promedio ponderado²⁵ de betas para cada portafolio

2. *INDICE VPN*

Considerando los precios de las sociedades de inversión y la tasa cete a 28 días para el periodo del 31 de agosto de 2007 al 30 de septiembre de 2009, se obtiene el promedio de la tasa cete en este periodo. Se obtiene el promedio de precios para cada sociedad de inversión. Se obtiene un promedio de los promedios de los rendimientos para cada portafolio de inversión. Se calcula el valor actual de una inversión, que es el valor que tiene actualmente la suma de una serie de pagos que se efectuarán en el futuro.

3. *INDICE TIR*

Antes de calcular la TIR para cada portafolio se hacen los siguientes supuestos. Para cada portafolio se asigna la cantidad de 5 millones y se divide por año la deuda, lo que se obtiene una anualidad de aproximadamente 192 mil. Para calcular el flujo de pagos se utiliza la siguiente fórmula:

$$\text{Flujo de efectivo} = \text{Parte Entera} \left[\frac{\text{Anualidad}}{\text{Suma de precios}} \right] * \text{Suma de precios}$$

Para el cálculo de la TIR utilizamos la formula de Excel, que devuelve la tasa interna de retorno de los flujos de caja representados por los números del argumento valores. Estos flujos de caja no tienen por que ser constantes, como es el caso en una

²⁴ La Beta de una cartera de acciones es el nivel de relación que existe entre el rendimiento de nuestra cartera y el mercado, expresado en un índice.

²⁵ La ponderación se hace en forma aleatoria para cada beta de cada sociedad de inversión

anualidad. Sin embargo, los flujos de caja deben ocurrir en intervalos regulares, como meses o años. La tasa interna de retorno equivale a la tasa de interés producida por un proyecto de inversión con pagos (valores negativos) e ingresos (valores positivos) que se producen en períodos regulares.

4. INDICE RENDIMIENTO

A partir de los datos obtenidos en la BMV, cada sociedad de inversión para cada mes realiza el cálculo de un rendimiento, entonces por cada portafolio de inversión se obtiene un promedio.

5. INDICE FITCHRATING

Se obtuvieron vía internet las calificaciones de las sociedades de inversión²⁶ y de acuerdo a la categoría que tienen asignadas, se etiqueto cada categoría con números y posteriormente se asigno a cada sociedad de inversión. Se promediaron las calificaciones por cada portafolio

6. PROGRAMACION COMPROMISO

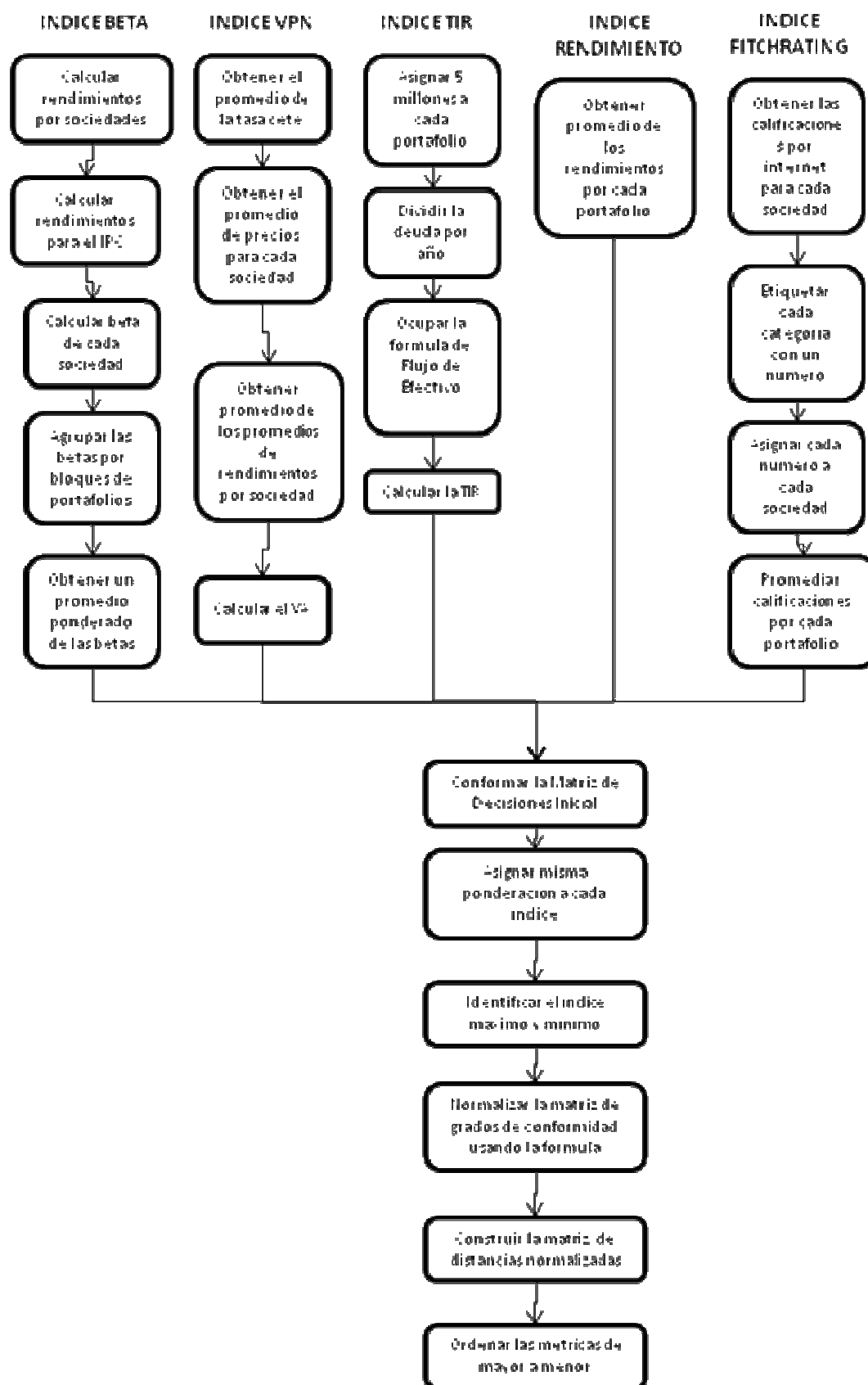
Con los datos de cada índice para cada portafolio se conforma la matriz de decisiones inicial. Para aplicar el método programación compromiso se asigna el mismo peso a cada índice que anteriormente se describió, por lo cual cada índice tendrá un peso de 0.2. Para cada índice se localiza el máximo y mínimo valor por cada portafolio. Para conformar la matriz de grados de proximidad normalizados, se ocupa la siguiente

fórmula: $\frac{X_{Max}-X_i}{X_{Max}-X_{Min}}$ donde

- X_{Max} = El valor mas grande en el índice por cada portafolio
- X_{Min} = El valor mas pequeño en el índice por cada portafolio
- X_i = cada valor en el índice por cada portafolio

Una vez realizado el paso anterior, ahora se construye la matriz de distancias, la cual es calculada en cada celda realizando un promedio ponderado de los datos de la matriz anterior con las ponderaciones asignadas inicialmente. Posteriormente solo se ordenan las métricas de mayor a menor.

²⁶ Las categorías se incluyen en la parte de anexos al final de la tesis y se obtuvieron de www.bmv.com.mx y www.fitchratings.com.



En el cuadro se reflejan las diferencias entre los algoritmos genéticos y la programación compromiso bajo un enfoque cualitativo de ellos.

ALGORITMO GENÉTICO	PROGRAMACIÓN COMPROMISO
Se obtienen resultados cuantitativos	Se obtienen resultados en forma de criterio
La implementación del algoritmo resulta más laboriosa	Es sencilla su implementación
Necesita de una iteración para alcanzar el punto óptimo	En la primera iteración se alcanza el resultado
Puede alcanzar el punto óptimo	No converge a un punto óptimo sino a un conjunto óptimo soluciones cercanos a la mejor solución
Para muchos datos podría llevarse más tiempo alcanzar el punto óptimo	No influyen para efectos de iteración y tiempo el número de datos
Podría no llegar a converger al óptimo, dependiendo si la población inicial es muy grande	Converge aun teniendo mucha información
Es más complicada su conceptualización para lograr la evaluación de los individuos (función de adaptabilidad)	Es muy sencilla su implementación
No necesariamente requiere información del entorno de la problemática	Depende bastante de la información del entorno de la problemática para su implementación
Puede ser complicado reflejar variables de la problemática al aplicar este método	No existe limitación en cuanto al número de criterios a considerar para aplicar el método
Es un método dinámico, ya que itera hasta encontrar la solución óptima	Es un método estático que no requiere iterar para alcanzar el conjunto óptimo de soluciones
Requiere para su implementación de una función de adaptabilidad para controlar la convergencia del algoritmo	No requiere ninguna función para encontrar el conjunto de soluciones cercanas al óptimo
Para varias opciones aumenta el tamaño de los cromosomas ya que haría lenta la convergencia del método	No influye el número de opciones (portafolio) para lograr alcanzar el conjunto óptimo de soluciones

Tabla 5. Análisis comparativo de la programación compromiso y los algoritmos genéticos

6.4 Análisis de resultados para el Algoritmo Genético

Los Algoritmos Genéticos (AGs) son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Cuando el algoritmo se ejecuta todos los miembros de una población tienden a parecerse al “promedio” de la población,

eliminando así los casos atípicos. Los Algoritmos Genéticos trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado. A cada individuo se le asigna un valor o puntuación, relacionado con la bondad de dicha solución. Cuanto mayor sea la adaptación de un individuo al problema, mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma. Este cruce producirá nuevos individuos (descendientes de los anteriores) los cuales comparten algunas de las características de sus padres. De esta manera se produce una nueva población de posibles soluciones, la cual reemplaza a la anterior y verifica la interesante propiedad de que contiene una mayor proporción de buenas características en comparación con la población anterior. Así a lo largo de las generaciones las buenas características se propagan a través de la población. Favoreciendo el cruce de los individuos mejor adaptados y se van explorando las aéreas más prometedoras del espacio de búsqueda.

En la tabla siguiente, se muestran los resultados al aplicar el algoritmo genético a los cuatro portafolios construidos. Se puede observar que la suma de asignaciones coincide con la suma de capital asignada y además que, el portafolio 4 tiene mayor tendencia a serle asignado capital.

	Portaf 1	Portaf 2	Portaf 3	Portaf 4	
Asig1	2	1	0	2	5
Asig2	2	0	1	2	5
Asig3	2	3	0	0	5
Asig4	1	1	0	3	5
Asig5	1	3	0	1	5

Tabla 7. Análisis de los resultados al aplicar el algoritmo genético

En el esquema que a continuación sigue, se muestran, por asignación realizada, las preferencias que el algoritmo genético fue realizando cuando fue iterando en su ejecución.

Asig1: Portafolios 4 y 1 → Portafolio 2 → Portafolio 3

Asig2: Portafolios 4 y 1 → Portafolio 3 → Portafolio 2

Asig3: Portafolios 2 → Portafolio 1 → Portafolios 3 y 4

Asig4: Portafolios 4 → Portafolios 1 y 2 → Portafolios 3

Asig5: Portafolios 2 → Portafolios 4 y 1 → Portafolios 3

De acuerdo a cada asignación de capital para cada portafolio al iterar el algoritmo genético nos puede sugerir lo siguiente:

1. El portafolio 4 tiene cierto predominio sobre los demás portafolios, ya que en cada asignación es el que presenta mayor capital invertido.
2. El portafolio 3 parece ser el menos rentable ya que presenta la menor asignación posible.
3. El portafolio 2 mantiene una presencia en general en las distribuciones de capital ya que en orden de asignación y montos invertidos se presenta en cualquiera de los tres órdenes anteriormente mostrados.
4. El portafolio 1 solo mantiene una presencia en los primeros órdenes de asignación de capital

El algoritmo genético ha convergido en medida de darle mayor peso de inversión al portafolio 4 siendo este el que mas rentable proporcione la mayor inversión posible, mientras que el portafolio 3 no parece ser una opción viable ya que aparece como la menos "asignada". Por otro lado los portafolios 2 y 1 mantienen una presencia en todos los lugares siendo el portafolio 1 el que sólo tenga presencia en los primeros y segundos lugares de asignación.

El algoritmo genético propone una forma más eficiente y veloz para obtener una búsqueda de soluciones para un problema de asignación, en esta tesis de recursos para un portafolio de inversión.

6.5 Análisis de resultados para la programación compromiso

En el modelo de programación compromiso se acerca la posición óptima del decisor a un ideal. El ideal se define como un punto o cesta cuyas coordenadas son las mejores oportunidades para cada una de las variables consideradas en el modelo (beneficio empresarial, volumen de ventas, penetración en cada mercado, plantilla de trabajadores, etc.). La aproximación del punto o cesta de decisiones al punto ideal, se consigue recurriendo a un modelo de distancias con diversas métricas (entre uno e infinito). El punto elegido se encuentra obviamente sobre el conjunto de oportunidades (frontera eficiente) y su distancia al ideal es mínima. El punto ideal es normalmente inalcanzable. Cuando el punto ideal es inalcanzable, el elemento óptimo o solución compromiso para el centro decisor, viene dada por la solución eficiente que se encuentra mas próxima con respecto al punto ideal.

En el cuadro siguiente se puede apreciar el orden que se tienen para cada portafolio de acuerdo a las métricas utilizadas para el estudio. En este mismo, se observa que el portafolio 4 tiene las métricas más pequeñas [$L_1(\text{ATR})=0.818$, $L_2(\text{ATR})=0.168$ y $L_\infty(\text{ATR})=0.218$] con respecto a los otros portafolios y de acuerdo a la teoría de la programación compromiso, es el portafolio más cercano al óptimo.

ORDENACION DE ALTERNATIVAS PARA DIFERENTES METRICAS					
L_1 (ATR)		L_2 (ATR)		L_∞ (ATR)	
Portaf 4	0.818	Portaf 4	0.168	Portaf 4	0.218
Portaf 1	0.902	Portaf 3	0.563	Portaf 3	0.680
Portaf 3	1.285	Portaf 1	0.863	Portaf 2	0.817
Portaf 2	1.497	Portaf 2	0.938	Portaf 1	0.929

Tabla 6. Análisis de los resultados al aplicar la programación compromiso

La métrica L_1 tiene la siguiente precedencia de portafolios, es decir, tomando en cuenta el cuadro anterior de ordenación de alternativas para diferentes métricas, al observar la columna de $L_1(ATR)$ se observa que el portafolio 4 tiene el valor menor con respecto a los otros portafolios y en forma descendente se ordenan, a continuación se describen el resto de las métricas ordenadas:

L_1 : Portaf4=0.818 \rightarrow Portaf1=0.902 \rightarrow Portaf3=1.285 \rightarrow Portaf2=1.497

L_2 : Portaf4=0.168 \rightarrow Portaf3=0.563 \rightarrow Portaf1=0.863 \rightarrow Portaf2=0.938

L_∞ : Portaf4=0.218 \rightarrow Portaf3=0.680 \rightarrow Portaf2=0.817 \rightarrow Portaf1=0.929

Al aplicar programación compromiso y calcular las métricas nos reflejan que

1. El portafolio 2 se muestra mas cercano a la mejor opción de inversión
2. Como segunda opción de inversión cualquiera de los portafolios 1, 2 o 3 resultan estar mas cercanos a la mejor inversión
3. El portafolio 3 tiene un peso fuerte como tercera opción para invertir.
4. El portafolio 4 el mas alejado de la mejor solución

Este análisis puede decir que, en general, optar por invertir en el segundo portafolio, nos dejaría la mejor inversión, ya que presenta las métricas más altas²⁷. Como segunda opción de inversión es viable optar por cualquiera de las opciones 1, 2 o 3, posiblemente explicadas por estar en “igual” cercanía a la mejor inversión. Los portafolios 3 y 4 solo se encontrarían más lejanos aun de la mejor inversión.

La programación compromiso proporciona una metodología para la toma de decisiones bajo una serie de situaciones que afectan o influyen a las mismas. En esta tesis se proporcionó una aplicación al problema de asignación de recursos para un portafolio de inversión.

²⁷ Nótese que se esta considerando la maximización del modelo descrito en el articulo de Carlos Romero “Análisis de las Decisiones Multicriterio” pag 42 formula 2.14

CAPITULO VII. CONCLUSIONES

De acuerdo con el estudio llevado a cabo y los resultados obtenidos con el contraste planteado en la hipótesis- que estableció que existen diferencias de aplicación entre los métodos de programación compromiso y algoritmos genéticos, en virtud de que ambos se fundamentan en criterios y objetivos distintos- se obtuvieron una serie de conclusiones que contribuyen al mejoramiento del manejo de los dos métodos.

7.1 Relativas al caso práctico

La metodología presentada en este trabajo puede ser usada en una gran cantidad de problemas donde se necesite contrastar métodos.

- Al analizar los resultados del caso estudio, los dos métodos se acercan a la misma solución ya que se logra llegar a una distribución más próxima a la asignación inicial (5 millones por cada portafolio)
- El caso estudio requirió que se adaptara de tal manera que fuera aplicable a cualquiera de los dos métodos, ya que para lograr compararlos, se debían tener los portafolios en las mismas condiciones que influían el entorno financiero
- En cuanto al diseño e implementación de los métodos, los algoritmos genéticos resultaron ser más complicados de elaborar en su implementación.
- Aunque el tiempo para iterar de cada modelo fue muy poco y despreciable, se podría afirmar que el de algoritmos genéticos llevaría más tiempo, ya que este modelo si itera en comparación de programación compromiso que no lo hace.

Al aplicar los dos métodos se observaron los siguientes resultados conforme al análisis cualitativo de la tabla 5, 6 y 7 se observan las siguientes conclusiones:

1. Los resultados que se expresan para el algoritmo genético nos proporciona una información cuantitativa de cuanto invertir en que cada portafolio mientras que la programación compromiso nos indica que portafolio de inversión es el mas apropiado para invertir.
2. En general los algoritmos genéticos son mas complicados de implementar
3. La programación compromiso sólo hace una iteración para encontrar su resultado, mientras que los algoritmos genético se necesitan mas de una iteración para alcanzar un resultado favorable.

4. Los algoritmos genéticos encuentran la selección optima para cada portafolio mientras que la programación compromiso solo proporciona un criterio.
5. Haciendo un análisis de la implementación de la programación compromiso, no afectaría cuantos portafolios podríamos tener, siempre seria inmediato el resultado.
6. Para este trabajo de tesis se necesito establecer los portafolios de tal manera que reflejaran las mismas condiciones para aplicarlas a los dos métodos, pero no necesariamente los algoritmos genéticos requieren mucha información para su implementación.
7. Es obvio observar en las tablas ubicadas en el anexo, que los algoritmos genéticos son un método dinámico ya que itera para encontrar una solución satisfactoria mientras que la programación compromiso no lo hace.
8. Si aumentara el número de portafolios, el algoritmo genético tiende a tardar en encontrar una solución satisfactoria, mientras que la programación compromiso es inmediato el resultado.

7.2.1 Algoritmos Genéticos

- Se ocupan para problemas de búsqueda y optimización
- Puede no encontrar la solución optima pero encuentra soluciones a un nivel aceptable
- El gran campo de aplicación de estos algoritmos se relaciona con aquellos problemas para los cuales no existen técnicas especializadas.
- El teorema de esquemas garantiza la convergencia del algoritmo y este teorema tiene fundamentos muy robustos y técnicos
- La aplicación de los algoritmos genéticos en las bolsas de valores es en donde más trabajos se han publicado en los últimos años y donde se muestra el mayor potencial para el uso de los algoritmos evolutivos de los cuales los algoritmos genéticos forman parte

7.2.2 Programación Compromiso

- La programación compromiso forma parte de las metodologías multicriterio ya que logra combinar distintas dimensiones como objetivos, actores y escalas que se encuentran insertados en el proceso de toma de decisiones, sin sacrificar calidad, confiabilidad y consenso en los resultados

- La programación compromiso encuentra un conjunto de valores próximos a la solución óptima conocida con Pareto
- La programación compromiso tiene sus fundamentos en métricas que a su vez están constituidas en diferencias y ponderaciones de las desviaciones individuales de acuerdo a sus magnitudes.
- La programación compromiso resultaría muy eficiente si se opera dentro de un enfoque de tipo continuo
- En la presente investigación se encontraron aplicaciones de la programación compromiso a problemas multicriterio en general

7.3 Sugerencias a futuras investigaciones

Las redes neuronales artificiales y lógica difusa, modelo como el browniano simple o geométrico, son algunas de las herramientas que se están desarrollando para obtener mejores pronósticos del precio de los instrumentos financieros. Se recomienda como trabajo futuro la exploración de estas herramientas para buscar mejorar los pronósticos que generarían a su vez mayor confiabilidad de la metodología propuesta en el presente trabajo; igualmente debería considerarse una exploración más detallada de otras medidas de riesgo como la desviación estándar de los rendimientos.

BIBLIOGRAFIA

1. RIOS, Sixto. RIOS-INSUA, Sixto. RIOS-INSUA, María J. (1989). *Procesos de Decisión Multicriterio*. Ediciones de la Universidad Complutense S.A. EUDEMA. Madrid.
2. ROMERO, Carlos. (1993). *Teoría de la Decisión Multicriterio: Conceptos, técnicas y aplicaciones*. Alianza Editorial S.A. Madrid.
3. BARBA-ROMERO, Sergio y POMEROL, Jean-Charles. (1997). *Decisiones Multicriterio: Fundamentos teóricos y utilización práctica*. Servicio de publicaciones Universidad de Alcalá. Alcalá de Henares.
4. BARBA-ROMERO, Sergio y POMEROL, Jean-Charles. (1997). *Decisiones Multicriterio: Fundamentos teóricos y utilización práctica*. Servicio de publicaciones Universidad de Alcalá. Alcalá de Henares.
5. Beasley, David; Bull, David R. & Martin, Ralph r. [1993], An Overview of Genetic Algorithms: Part 1, fundamentals, University Computing, pp.58-69
6. Gutiérrez Miguel Angel; De los Cobos Sergio & Pérez Blanca. [1999], Búsqueda Tabú: Un Procedimiento Heurístico para Solucionar Problemas de Optimización Combinatoria, División CBI UAM Azcapotzalco e Iztapalapa, <http://www-azc.uam.mx/enlinea2/1-3.htm>
7. Luenberg, David. [1989], Programación lineal y no lineal, Addison-Wesley Publishing, pp. 3-6
8. Morales, Eduardo. [Mayo 1999], Búsqueda, optimización y aprendizaje, Instituto Tecnológico de Estudios Superiores de Monterrey, <http://w3.mor.itesm.mx/emorales/Cursos/Busqueda/principal.html>
9. Rao, Singiresu S. [1996], Engineering Optimization, theory and practice, John Wiley & Sons, Inc., Third Edition, pp.343,348
10. Hernández Sampieri, Roberto. Metodología de la Investigación. McGraw-Hill. Tercera Edición. 2003.638 pp.
11. Ley de Sociedades de Inversión. Cámara de Diputados del H. Congreso de la Unión. Secretaría General. Secretaría de Servicios Parlamentarios. Centro de Documentación, Información y Análisis.

OTRAS REFERENCIAS BIBLIOGRAFICAS

- [4] J. Arifovic, *Learning by Genetic Algorithms in Economic Environments*, University of Chicago, Chicago, 1991.
- [5] W. B. Arthur, *Effective Choice in the Iterated Prisoner's Dilemma*, *Journal Conflict Resolution*, 24 (1994), pp. 3-25.
- [12] R. Bell and S. Beare, *Emuating Trade in Emissions Permits: An Application of Genetic Algorithms*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 159 - 174.
- [13] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, *Journal of Political Economy*, 81 (1973), pp. 637 - 654.
- [20] N. T. Chan, B. LeBaron and A. W. Poggio, *Information Dissemination and Aggregation in Asset Markets with Simple Intelligent Traders*, AIM-1646, 1998.
- [21] S.-H. Chen, C.-H. Yeh and C.-C. Liao, *On AIE-ASM: Software to Simulate Artificial Stock Markets with Genetic Programming*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 107 - 122.
- [22] N. Chidambaran, J. Triqueros and C.-W. J. Lee, *Option Pricing Via Genetic Programming*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 383 - 397.
- [28] D. B. Fogel, K. Chellapilla and P. Angeline, *Evolutionary Computation and Economic Models: Sensitivity and Unintended Consequences*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 245 - 269.
- [38] A. L. García-Almanza and E. Tsang, *Simplifying Decision Trees Learned by Genetic Programming*, *IEEE Congress on Evolutionary Computation (CEC 2006)*, Vancouver, Canada, 2006, pp. 7906-7912.
- [44] C. Keber, *Evolutionary Computation in Option Pricing: Determining Implied Volatilities Based on American Put Options*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 399 - 415.
- [45] J. Koza, *A Genetic Approach to Econometric Modelling*, in P. Bourguine and B. Walliser, eds., *Economics and Cognitive Science*, Pergamon Press, 1992, pp. 57-75.

- [46] J. Li, *Enhancing Financial Decision Making Using Multi-Objective Financial Genetic Programming*, *IEEE Congress on Evolutionary Computation (CEC 2006)*, Vancouver, Canada, 2006, pp. 7935-7942.
- [47] F. Luna, *Computable Learning, Neural Networks and Institutions*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 211 - 232.
- [53] A. Mochón, D. Quintana, Y. Sáez and P. Isasi, *Analysis of Ausubel Auctions by Means of Evolutionary Computation*, *Congress on Evolutionary Computation (CEC 2005)*, Edinburgh, Scotland, 2005, pp. 2645-2652.
- [54] R. Pereira, *Forecasting Ability But No Profitability: An Empirical Evaluation of Genetic Algorithm-Optimised Technical Trading Rules*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, New York, 2002, pp. 287 - 309.
- [55] A. Rapoport, *Optimal Policies for the Prisoner's Dilemma*, Univ. North Carolina, Chapel Hill, 1966.
- [58] T. Riechmann, *Genetic Algorithm Learning and Economic Evolution*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, Physica-Verlag, Heidelberg New York, 2002, pp. 45 - 59.
- [62] F. Streichert, H. Ulmer and A. Zell, *Evaluating a Hybrid Encoding and Three Crossover Operators on the Constrained Portfolio Selection Problem*, *Congress on Evolutionary Computation (CEC 2004)*, Portland, Oregon, 2004, pp. 932-939.
- [63] G. G. Szpiro, *A Search for Hidden Relationships: Data Mining with Genetic Algorithms*, *Computational Economics*, 10 (3) (1997), pp. 267-277.
- [66] J. Yang, *The Efficiency of an Artificial Double Auction Stock Market with Neural Learning Agents*, in S.-H. Chen, ed., *Evolutionary Computation in Economics and Finance*, physica - verlag, New York, 2002, pp. 85 - 105.

ANEXOS

TABLA DE DATOS DE LAS SOCIEDADES DE INVERSION

TIPO	AGRESIVAS	ESPECIALIZADAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS
EMISORA	VALUEF2 A	VALMX17 A	MONEXM+ BEC-0	INGCORP BOE1	INGCORP BOE2	INGCORP BOI	INGCORP A	FINDE1 E	HZMD+ A	ING-30 A
31/08/2007	66.212358	1.091744	21.789102	1.242378	1.241809	1.237316	1.22589	1.618603	77.107267	2.617708
29/09/2007	66.611119	1.104795	21.892663	1.2502	1.249056	1.244625	1.233126	1.628878	77.534332	2.63234
31/10/2007	67.022569	1.117842	22.019261	1.26101	1.259221	1.25485	1.243253	1.639977	78.051636	2.649744
30/11/2007	67.42626	1.119998	22.138731	1.246721	1.263164	1.266272	1.251485	1.648923	78.509615	2.664974
29/12/2007	67.993633	1.127819	22.258597	1.280076	1.277072	1.272823	1.26105	1.658243	78.935109	2.680038
31/01/2008	68.437312	1.138309	22.367679	1.287019	1.28332	1.279163	1.267326	1.669114	79.419754	2.695592
28/02/2008	68.975263	1.155215	22.435469	1.295742	1.29146	1.287353	1.275436	1.678476	79.883626	2.710883
30/03/2008	69.536015	1.168812	22.611853	1.307217	1.302281	1.298233	1.286211	1.68783	80.428681	2.728242
30/04/2008	70.289041	1.173539	22.726027	1.315366	1.309760	1.305787	1.293690	1.697378	80.886540	2.743713
31/05/2008	70.823478	1.173073	22.849129	1.324380	1.318091	1.314193	1.302013	1.706972	81.359480	2.759773
29/06/2008	71.271463	1.170973	22.960424	1.331969	1.325039	1.321215	1.308966	1.715816	81.824182	2.775602
31/07/2008	71.760278	1.180827	23.117723	1.341583	1.333930	1.330186	1.317849	1.726350	82.337161	2.793331
31/08/2008	72.242354	1.190282	23.213661	1.350404	1.342045	1.338381	1.325962	1.736878	82.840540	2.810363
28/09/2008	72.631100	1.199280	23.351205	1.358876	1.349869	1.346277	1.333780	1.745965	83.291384	2.825309
31/10/2008	73.110036	1.213056	23.517969	1.369065	1.359285	1.355777	1.343187	1.757552	83.825687	2.843408
30/11/2008	73.590531	1.220122	23.652946	1.377963	1.367472	1.364043	1.351371	1.768124	84.307384	2.861317
31/12/2008	74.257635	1.228516	23.783772	1.386983	1.375752	1.372407	1.359652	1.779399	84.824594	2.879258
31/01/2009	75.112067	1.235230	23.930871	1.398963	1.386959	1.383309	1.370443	1.790451	85.359101	2.897808
29/02/2009	75.568374	1.243576	24.063482	1.407469	1.394756	1.390812	1.377868	1.801072	85.824100	2.913432
31/03/2009	76.125322	1.252855	24.197433	1.416185	1.402710	1.398450	1.385426	1.812381	86.333720	2.929827
30/04/2009	76.656120	1.260614	24.326272	1.420469	1.406290	1.401732	1.388669	1.822456	86.786490	2.940774
30/05/2009	77.203649	1.265766	24.459420	1.428805	1.413877	1.409004	1.395864	1.832620	87.268097	2.957092
30/06/2009	77.749560	1.269606	24.602813	1.433504	1.417835	1.412647	1.399464	1.843188	87.668283	2.972541
31/07/2009	78.324475	1.284109	24.764330	1.445169	1.428677	1.423144	1.409853	1.854826	88.145588	2.992180
29/08/2009	\$78.892065	\$1.298124	\$24.915424	\$1.456197	\$1.438924	\$1.433064	\$1.419672	\$1.867057	\$88.676652	\$3.008786
30/09/2009	\$79.498004	\$1.307399	\$25.864290	\$1.466363	\$1.448240	\$1.442032	\$1.428546	\$1.880087	\$89.279659	\$3.028496

TABLA DE DATOS DE LAS SOCIEDADES DE INVERSION

TIPO	MERCADO DE DINERO	MERCADO DE DINERO	AGRESIVAS	MERCADO DE DINERO	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS	AGRESIVAS
EMISORA	ING-1 BM2	ING-1 A	HZMD A	AMEX-MD A	INGCORP BOE4	INGCORP BOM	APOLO90 A	FINDE1 F	APOLOM A	ACTIREN B5
31/08/2007	2.417563	2.395929	16.746078	1.689606	1.241238	1.231132	2.041535	1.61424	0.305754	6.242979
29/09/2007	2.430928	2.408975	16.836221	1.698533	1.247909	1.237835	2.051697	1.62392	0.3073	6.275358
31/10/2007	2.446287	2.424117	16.936829	1.70877	1.25743	1.247373	2.0636	1.634355	0.309092	6.310622
30/11/2007	2.459787	2.43745	17.032621	1.717911	1.269872	1.265071	2.075114	1.642782	0.310799	6.343856
29/12/2007	2.47315	2.45059	17.123444	1.727045	1.274072	1.264056	2.088907	1.651594	0.312635	6.376292
31/01/2008	2.488051	2.465273	17.228084	1.737414	1.279651	1.269689	2.096933	1.661872	0.314123	6.413123
28/02/2008	2.501834	2.478806	17.318129	1.746956	1.287188	1.27725	2.108431	1.670723	0.315781	6.446827
30/03/2008	2.51689	2.4936	17.417781	1.757238	1.29736	1.287433	2.120849	1.679515	0.31761	6.482316
30/04/2008	2.531741	2.508244	17.514417	1.767399	1.304172	1.294287	2.132746	1.688472	0.319441	6.521648
31/05/2008	2.547234	2.523487	17.618902	1.777674	1.311829	1.301978	2.144980	1.697467	0.321294	6.558236
29/06/2008	2.562164	2.538218	17.716669	1.787802	1.318141	1.308332	2.157105	1.705743	0.323058	6.593726
31/07/2008	2.577825	2.553722	17.824954	1.798831	1.326317	1.316546	2.170857	1.715636	0.325096	6.631754
31/08/2008	2.592392	2.568142	17.931065	1.808810	1.333734	1.324004	2.183281	1.725524	0.326950	6.668877
28/09/2008	2.605688	2.581305	18.026660	1.818025	1.340918	1.331223	2.195221	1.734024	0.328668	6.702750
31/10/2008	2.621922	2.597376	18.138094	1.829074	1.349569	1.339913	2.209266	1.744925	0.330728	6.741823
30/11/2008	2.638358	2.613649	18.244623	1.840665	1.357057	1.347439	2.221734	1.754874	0.332613	6.779048
31/12/2008	2.654395	2.629525	18.355930	1.851660	1.364606	1.355032	2.234851	1.765495	0.334615	6.818604
31/01/2009	2.670525	2.645487	18.464761	1.862726	1.375053	1.365119	2.251154	1.775408	0.336642	6.858626
29/02/2009	2.684974	2.659785	18.564790	1.872387	1.388442	1.371886	2.263887	1.784917	0.338442	6.897094
31/03/2009	2.700531	2.675178	18.672428	1.882846	1.396020	1.378735	2.276618	1.795025	0.340289	6.937249
30/04/2009	2.714524	2.689023	18.775589	1.891807	1.392248	1.381305	2.286203	1.803935	0.341800	6.977227
30/05/2009	2.728960	2.703306	18.880367	1.901973	1.399100	1.387803	2.297802	1.812920	0.343546	7.014523
30/06/2009	2.743883	2.718071	18.983858	1.912498	1.402334	1.390699	2.307611	1.822257	0.345257	7.053476
31/07/2009	2.761273	2.735279	19.095418	1.924621	1.412369	1.400335	2.320168	1.832636	0.347150	7.093110
29/08/2009	\$2.776845	\$2.750688	\$19.210384	\$1.935249	\$1.421851	\$1.409439	\$2.334170	\$1.843661	\$0.349170	\$7.130828
30/09/2009	\$2.794594	\$2.768252	\$19.338583	\$1.947419	\$1.430337	\$1.417528	\$2.349200	\$1.855353	\$0.351398	\$7.173450

TABLA DE DATOS DE LA TASA DE RENDIMIENTO EN CETES

Periodo	Tasa de rendimiento en CETES (28 días) a/	Periodo	Tasa de rendimiento en CETES (28 días) a/
2007/01	7.04	2008/07	7.93
2007/02	7.04	2008/08	8.18
2007/03	7.04	2008/09	8.17
2007/04	7.01	2008/10	7.74
2007/05	7.24	2008/11	7.43
2007/06	7.2	2008/12	8.02
2007/07	7.19	2009/01	7.59
2007/08	7.2	2009/02	7.12
2007/09	7.21	2009/03	7.03
2007/10	7.2	2009/04	6.05
2007/11	7.44	2009/05	5.29
2007/12	7.44	2009/06	4.98
2008/01	7.42	2009/07	4.59
2008/02	7.43	2009/08	4.49
2008/03	7.43	2009/09	4.48
2008/04	7.44	2009/10	4.51
2008/05	7.44	2009/11	4.51
2008/06	7.56	2009/12	4.5

Tabla de datos del IPC Nominal

Unidad de Medida: Índice base octubre 1978 = 0.78.

Periodo	Índice máximo	Índice mínimo	Último índice del mes	PROM IPC a/
2007/01	27561.49	25783.04	27561.49	26672.265
2007/02	28715.96	26418.82	26638.95	27567.39
2007/03	28747.69	25788.37	28747.69	27268.03
2007/04	29832.48	28996.71	28996.71	29414.595
2007/05	31398.96	29259.92	31398.96	30329.44
2007/06	32271.38	30744.71	31151.05	31508.045
2007/07	32411.84	29996.6	30659.66	31204.22
2007/08	30661.87	27793.16	30347.86	29227.515
2007/09	30932.71	29794.49	30296.19	30363.6
2007/10	32836.12	30855.68	31458.67	31845.9
2007/11	30806.3	27883.01	29770.52	29344.655
2007/12	31268.36	28968.2	29536.83	30118.28
2008/01	29069.56	25284.88	28793.64	27177.22
2008/02	30424.76	27929.29	28918.52	29177.025
2008/03	30912.99	28207.29	30912.99	29560.14
2008/04	32095.04	30270.64	30281.41	31182.84
2008/05	31975.47	30551.47	31975.47	31263.47
2008/06	31726.01	29199.94	29395.49	30462.975
2008/07	29221.11	26732.84	27501.02	27976.975
2008/08	27342.39	26071.46	26290.99	26706.925
2008/09	26522.76	23456.84	24888.9	24989.8
2008/10	25117.32	16868.66	20445.32	20992.99
2008/11	21535.27	18190.7	20534.72	19862.985
2008/12	22572.5	19533.08	22380.32	21052.79
2009/01	23250.96	19213.02	19565.14	21231.99
2009/02	20537.1	17752.18	17752.18	19144.64
2009/03	20542.25	16929.8	19626.75	18736.025
2009/04	22582.17	19880.37	21898.85	21231.27
2009/05	24659.24	23014.05	24331.71	23836.645
2009/06	25460.02	23314.68	24368.38	24387.35
2009/07	27043.5	23359.94	27043.5	25201.72
2009/08	28599.92	27299.58	28129.95	27949.75
2009/09	30017.72	27749.16	29232.24	28883.44
2009/10	31017.73	28588.26	28646.03	29802.995
2009/11	31400.2	28886.53	30957.11	30143.365
2009/12	32626.29	31670.21	32120.47	32148.25

a/ Elaboracion propia (promedio Max y Min)

Tabla de asignación de calificaciones a sociedades de inversión

Num	Letra	Numero	Riesgo Credito	Riesgo Mercado	Num	Letra	Numero	Riesgo Credito	Riesgo Mercado
1	B	1	Mínimo	Extremadamente Baja	22	A	1	Bueno	Extremadamente Baja
2	B	2	Mínimo	Baja	23	A	2	Bueno	Baja
3	B	3	Mínimo	Baja a Moderada	24	A	3	Bueno	Baja a Moderada
4	B	4	Mínimo	Moderada	25	A	4	Bueno	Moderada
5	B	5	Mínimo	Moderada a Alta	26	A	5	Bueno	Moderada a Alta
6	B	6	Mínimo	Alta	27	A	6	Bueno	Alta
7	B	7	Mínimo	Muy Alta	28	A	7	Bueno	Muy Alta
8	BB	1	Bajo	Extremadamente Baja	29	AA	1	Alto	Extremadamente Baja
9	BB	2	Bajo	Baja	30	AA	2	Alto	Baja
10	BB	3	Bajo	Baja a Moderada	31	AA	3	Alto	Baja a Moderada
11	BB	4	Bajo	Moderada	32	AA	4	Alto	Moderada
12	BB	5	Bajo	Moderada a Alta	33	AA	5	Alto	Moderada a Alta
13	BB	6	Bajo	Alta	34	AA	6	Alto	Alta
14	BB	7	Bajo	Muy Alta	35	AA	7	Alto	Muy Alta
15	BBB	1	Aceptable	Extremadamente Baja	36	AAA	1	Sobresaliente	Extremadamente Baja
16	BBB	2	Aceptable	Baja	37	AAA	2	Sobresaliente	Baja
17	BBB	3	Aceptable	Baja a Moderada	38	AAA	3	Sobresaliente	Baja a Moderada
18	BBB	4	Aceptable	Moderada	39	AAA	4	Sobresaliente	Moderada
19	BBB	5	Aceptable	Moderada a Alta	40	AAA	5	Sobresaliente	Moderada a Alta
20	BBB	6	Aceptable	Alta	41	AAA	6	Sobresaliente	Alta
21	BBB	7	Aceptable	Muy Alta	42	AAA	7	Sobresaliente	Muy Alta

Tabla de calificaciones asignándoles un valor

Emisora	Evaludora	Calificacion	Riesgo de Credito	Riesgo Mercado	Num	Promedio a/
VALUEF2 A	FitchRatings	AA/5(mex) F	Alto	Moderada a Alta	33	32.800
VALMX17 A	FitchRatings	AAA/5(mex) F	Sobresaliente	Moderada a Alta	40	
INGCORP BOE 1	investment mangement	BBB/7	Aceptable	Muy Alta	21	
FINDE1 E	FitchRatings	AA/3 F	Alto	Baja a Moderada	31	
MONEXM+ BEC-0	FitchRatings	AAA/4	Sobresaliente	Moderada	39	
INGCORP A	investment mangement	BBB/7	Aceptable	Muy Alta	21	26.800
INGCORP BOI	investment mangement	BBB/7	Aceptable	Muy Alta	21	
INGCORP BOE2	investment mangement	BBB/7	Aceptable	Muy Alta	21	
HZMD+ A	FitchRatings	AAA/4 F	Sobresaliente	Moderada	39	
ING-30 A	investment mangement	AA/4	Alto	Moderada	32	
ING-1 BM2	Moody's	AAA/3	Sobresaliente	Baja a Moderada	38	32.800
ING-1 A	Moody's	AAA/2	Sobresaliente	Baja	37	
HZMD A	FitchRatings	AAA/3 F	Sobresaliente	Baja a Moderada	38	
INGCORP BOE4	investment mangement	BBB/7	Aceptable	Muy Alta	21	
AMEX-MD A	investment mangement	AA/2	Alto	Baja	30	
INGCORP BOM	investment mangement	BBB/7	Aceptable	Muy Alta	21	33.800
APOLO90 A	FitchRatings	AAA/5(mex) F	Sobresaliente	Moderada a Alta	40	
ACTIREN B5	FitchRatings	AAA/3(mex) F	Sobresaliente	Baja a Moderada	38	
FINDE1 F	FitchRatings	AA/3 F	Alto	Baja a Moderada	31	
APOLOM A	FitchRatings	AAA/4(mex) F	Sobresaliente	Moderada	39	