



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

PROGRAMACIÓN GENÉTICA APLICADA A LA
ENERGÍA OSCURA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

PRESENTA:

MARIO ATZIN NATAREN ELIZALDE

TUTOR:

DR. JOSÉ ALBERTO VÁZQUEZ GONZÁLEZ



Ciudad Universitaria, CDMX, 2024

Jurado asignado:

- Presidente:
Dr. Raúl Arturo Espejel Morales.
Facultad de Ciencias, UNAM.
- Vocal:
Dra. Katya Rodríguez Vázquez.
Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, UNAM.
- Secretario:
Dr. José Alberto Vázquez González.
Instituto de Ciencias Físicas, UNAM.
- Primer Suplente:
Dr. Sébastien Mickaël Marc Fromenteau.
Instituto de Ciencias Físicas, UNAM.
- Segundo Suplente:
Dra. Roxana Mitzayé del Castillo Vázquez.
Facultad de Ciencias, UNAM.

Sustentante:

Mario Atzin Nataren Elizalde.
Universidad Nacional Autónoma de México.
Facultad de Ciencias.
Física.
315062773.

*A mi madre,
por su apoyo y amor incondicional.*

Agradecimientos

Agradecimientos Familia:
Agradecimientos Amigos 1 , Amigos2
Agradecimientos Academicos
Agradecimientos generales

*Es proponiéndose lo imposible como el hombre ha logrado siempre lo posible,
aquellos que se han ceñido prudentemente a los que les parecía factible
jamás han avanzado un solo paso.*

-La libertad. Mijail Bakunin.

*El progreso debe encontrarse en un equilibrio entre lo que somos
y nunca podremos dejar de ser, y lo que sin sacrificar lo que somos,
tenemos la posibilidad de ser: Jesús, Leonardo o Chimpancé.*

-La región más transparente. Carlos Fuentes.

Resumen

El presente trabajo busca mostrar el uso de la programación genética como una herramienta auxiliar para el análisis de datos cosmológicos. Mediante un regresor simbólico (**SR**), se buscan expresiones matemáticas que describan a un conjunto de datos, ya sean sintéticos u observacionales, sin asumir un modelo específico previo. Ésto permite comparar los modelos cosmológicos existentes y comprender la forma en que se presentan los datos cuando no hay antecedentes claros sobre su comportamiento.

En el capítulo 1 *Cosmología* se presentan los antecedentes teóricos y observacionales de los que se vale el modelo cosmológico estándar actual Λ CDM que incluye la materia oscura además de considerar la energía oscura como la constante cosmológica. Se incluye además, modelos de energía oscura alternativos así como el tipo de observaciones utilizadas para cuantificar los parámetros cosmológicos. En el capítulo 2 *Computo evolutivo* se introduce el concepto de programación simbólica haciendo especial énfasis en las ventajas que este paradigma ofrece para tratar las operaciones y estructuras lógicas de un programa como variables. Además, se presenta el paradigma de la programación genética utilizando los conceptos biológicos asociados a la evolución, selección natural y mutación. El contenido del capítulo 3 *Regresor simbólico programado genéticamente*, está dedicado a comprender cómo las ventajas de la programación simbólica pueden explotarse para evolucionar programas algebraicos mediante la programación genética para obtener una expresión matemática que se ajuste bien a los datos de interés, describiendo cada una de las etapas del proceso. En el capítulo 4 *Funciones de estudio*, se presentan los resultados y se analiza el uso del regresor simbólico para obtener expresiones algebraicas a partir de datos sintéticos generados con diversas funciones de muestra. El capítulo 5 contiene los resultados de haber aplicado el regresor simbólico, primero a un conjunto de observaciones con el fin de determinar la función de Hubble $H(z)$ así como la constante de Hubble H_0 . Posteriormente se muestra el uso del regresor para obtener el tipo de expresión asociada a datos obtenidos de una parametrización libre de reconstrucciones de la densidad de energía, así como de la variable w de la ecuación de estado cosmológica. Finalmente en el capítulo 6 se presentan los resultados así como las conclusiones de este trabajo.

Índice general

Agradecimientos	iii
Resumen	v
Índice general	ix
Índice de figuras	xi
Índice de tablas	xv
1 Cosmología.	1
1.1 Introducción.	1
1.2 Principio Cosmológico	2
1.3 Ecuaciones de Einstein	3
1.4 Universo en expansión.	5
1.5 Métrica FRLW y modelo de Friedmann	7
1.5.1 Ecuación de Friednman	10
1.5.2 Ecuación de continuidad	11
1.6 La composición del universo	12
1.6.1 Materia	12
1.6.2 Radiación	14
1.6.3 Energía Oscura	15
1.7 Parámetros de densidad.	15
1.8 Modelo Λ CDM	16
1.9 Modelo CPL	17
1.10 Modelo PolyCDM	18
1.11 Observaciones	18
2 Computo evolutivo.	21
2.1 Introducción.	21
2.2 Programación simbólica	22

2.2.1	Arboles de aridad-M	24
2.2.2	Programa simbólico de juguete: Derivadas parciales.	27
2.2.3	Cambios en los elementos de un árbol.	30
2.3	Programación genética	31
2.3.1	Introducción	31
2.3.2	Evolución, selección natural y genética: Un recorrido histórico	32
2.3.3	Programación genética como paradigma del computo evolutivo	36
2.3.4	Regresor simbólico	38
3	Regresor simbólico programado genéticamente.	41
3.1	Introducción	41
3.2	Pre-inicialización.	41
3.3	Inicialización.	43
3.4	Selección.	45
3.5	Evolución.	47
3.6	Descendencia.	48
3.7	Finalización.	49
4	Funciones de estudio.	51
4.1	Introducción	51
4.2	La librería <i>gplearn</i>	52
4.3	Función lineal.	53
4.4	Función racional.	57
4.5	Funciones adicionales.	60
4.5.1	Valor absoluto	60
4.5.2	Raiz Cuadrada	61
4.6	Función logaritmo.	64
4.7	Función con asíntota.	66
4.8	Función irracional.	68
4.9	Función trigonométrica.	70
5	Aplicación a variables cosmológicas.	73
5.1	Parámetro de Hubble.	73
5.2	Densidad de energía.	75
5.3	Parametrización de ecuación de estado.	79
6	Conclusiones	83
A	Algoritmos	85

Índice de figuras

1.1	Universo homogéneo bajo un campo magnético constante \vec{B} (izquierda). Universo isotrópico con distribución simétrica esférica visto desde el centro bajo un campo gravitacional \vec{g} (centro). Universo isotrópico y homogéneo (derecha).	2
1.2	Gráfica publicada por Hubble sobre la relación velocidad-distancia para galaxias [16].	7
1.3	Universo esférico con curvatura $k = +1$ (arriba). Universo hiperbólico con curvatura negativa $k = -1$ (medio). Universo plano con curvatura nula $k = 0$ (abajo).	8
1.4	Curvas de rotación Galáctica. La curva A punteada indica la predicción del modelo Kepleriano en el que la velocidad de una distribución de masa disminuye conforme se aleja del centro gravitacional. La curva B sólida indica las observaciones registradas para la velocidad indicando así la presencia de materia oscura [28].	14
1.5	Datos de Hubble en función del corrimiento al rojo z para cronómetros cósmicos [37].	19
2.1	Árbol genealógico	24
2.2	Árbol de 4 ramas (aridad 4)	25
2.3	Árbol de aridad 3 (nodo A) con sub-arboles de aridad 1 (nodo B) ; aridad 2 (nodos C , F y G) y aridad 3 (nodo D) . Los nodos E,I,J,K,L,H1,H2 y H3 son hojas.	25
2.4	Representación de árbol aridad 3 para el polinomio $p(x) = x^2 - x + \frac{\pi}{2}$	26
2.5	Esquema de árbol del programa <i>Derivadas Parciales</i>	28
2.6	Condición inicial del programa	28
2.7	Simplificación del árbol al realizar las instrucciones en los nodos inferiores.	29
2.8	Simplificación del programa <i>Derivada parcial</i>	29
2.9	Generación de un nuevo algoritmo al sustituir uno de los nodos originales.	30

2.10 Evolución de las jirafas según Lamarck (arriba). Evolución según Darwin (abajo) [59].	34
3.1 Creación de un árbol completo de $prof\ max = 3$ mediante el método <i>Full</i> para formar al individuo $\frac{3}{x} + 2x$	44
3.2 Creación de un árbol asimétrico de $prof\ max = 3$ mediante el método <i>Grow</i> para formar al individuo $\frac{x}{1+x}$	45
3.3 Proceso de evaluación de las soluciones. Población de programas (arriba). Evaluación de la población mediante la función fitness (medio). Fitness o aptitud asociada (abajo).	46
3.4 Cruzamiento	48
3.5 Mutación de sub-árbol	48
3.6 Mutación de punto	49
3.7 Mutación de elevación	49
3.8 Diagrama de un árbol solución al final del algoritmo. Los nodos terminales (hojas) corresponden a variables mientras que los nodos verde oscuro corresponden a operadores de la base.	50
4.1 Funciones resultantes con diferente tamaño de población para la ec(4.2). Para apreciar mejor el comportamiento de las soluciones se presentan solo dos de éstas.	54
4.2 Funciones resultantes con diferente número de generaciones para la ec(4.2). Debido a la gran similitud de las curvas , se presentan solo un par de las soluciones.	55
4.3 Evolución del fitness promedio a través del paso de generaciones para distintos tamaños de población.	55
4.4 Evolución del fitness del mejor individuo a través del paso de generaciones para casos con distinto tamaño de población.	56
4.5 Evolución del fitness del mejor individuo a través del paso de generaciones para casos con un número distinto de generaciones.	56
4.6 Comparación de funciones obtenidas respecto a la ec(4.5) sobre el dominio completo de los datos de la Tabla [B.2].	58
4.7 Comparación de funciones obtenidas respecto a la ec(4.5) en la zona de mayor variación de los datos.	59
4.8 Curvas de las soluciones obtenidas en comparación con la función muestra de la ec(4.7)	61
4.9 Curvas de las soluciones obtenidas en comparación con la función muestra de la ec(4.9)	62

4.10	Comparación de funciones obtenidas respecto a la ec(4.11) sobre el dominio completo de los datos de la Tabla[B.5])	65
4.11	Comparación de funciones obtenidas respecto a la ec(4.11) en la zona de mayor variación de los datos.	66
4.12	Comparación de funciones obtenidas respecto a la ec(4.13) sobre el dominio completo de los datos de la Tabla[4.9]). Debido a la similitud de las curvas solo se presenta un par de éstas.	67
4.13	Comparación de funciones obtenidas respecto a la ec(4.13) en la zona de mayor variación de los datos. Debido a la similitud de las curvas solo se presenta un par de éstas.	68
4.14	Comparación de funciones obtenidas respecto a la función irracional muestra de la ec(4.15).	69
4.15	Comparación de funciones obtenidas respecto a la función trigonométrica muestra de la ec (4.17). Para mayor claridad en la presentación se presenta solo un par de las curvas asociadas a las soluciones.	72
5.1	Curvas obtenidas utilizando el regresor simbólico para los datos del parámetro de Hubble. $H(z)$ en $km \cdot s \cdot Mpc^{-1}$	75
5.2	Reconstrucción de datos para la densidad de energía.	76
5.3	Curvas obtenidas utilizando el regresor simbólico a la reconstrucción de datos para la densidad de energía.	78
5.4	Reconstrucción de datos para la ecuación de estado.	79
5.5	Para evitar una <i>sobrecomposición</i> de funciones, el operador de elevación modificado eliminaría automáticamente cualquier nodo función que sea evaluable en un nodo de la misma función , de éste modo el árbol de la izquierda cuya representación simbólica es $\text{sen}(\text{sen}(x + 1))$ se reduciría al árbol $\text{sen}(x + 1)$	81
5.6	Curvas obtenidas utilizando el regresor simbólico a la reconstrucción de datos para la ecuación de estado.	82

Índice de tablas

4.1	Parámetros de ajuste de una configuración estándar.	52
4.2	Soluciones en función del tamaño de la población.	53
4.3	Soluciones en función del número de generaciones.	54
4.4	Tabla de los parámetros asociados a las soluciones racionales de la Tabla [B.2]	58
4.5	Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.5).	58
4.6	Parámetros y soluciones correspondientes al regresor aplicado a los datos de la función valor absoluto	61
4.7	Parámetros y soluciones correspondientes al regresor aplicado a los datos de la función raíz cuadrada	62
4.8	Funciones y fitness asociado correspondientes al regresor aplicado a los datos de la función logaritmo de la ec(4.11)	64
4.9	Parámetros y soluciones correspondientes al regresor aplicado a los datos de la ecuación asintótica (4.13).	67
4.10	Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.15).	70
4.11	Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.17).	71
5.1	Funciones y fitness correspondientes al regresor aplicado a los cronómetros cósmicos.	74
5.2	Funciones y fitness correspondientes al regresor aplicado a los datos de la Fig[5.2]. Para mayor distinción en los términos se alterna el uso de paréntesis y brackets.	76
5.3	Funciones y fitness correspondientes al regresor aplicado a los datos de la Fig[5.4].	79
B.1	Características de los datos sintéticos asociados a la función lineal de la ec(4.2)	87

B.2	Características de los datos sintéticos asociados a la función racional de la ec(4.5)	87
B.3	Características de los datos sintéticos asociados a la función valor absoluto de la ec(4.7)	87
B.4	Características de los datos sintéticos asociados a la función raíz cuadrada de la ec(4.9)	87
B.5	Características de los datos sintéticos asociados a la función logaritmo de la ec(4.11)	87
B.6	Características de los datos sintéticos asociados a la función con asíntota de la ec(4.13)	87
B.7	Características de los datos sintéticos asociados a la función irracional de la ec(4.15)	88
B.8	Características de los datos sintéticos asociados a la función trigonométrica de la ec(4.17)	88
B.9	Tabla de los parámetros asociados a las soluciones logaritmo de la Tabla [4.8]	88
B.10	Tabla de los parámetros asociados a las soluciones irracionales de la Tabla[4.10].	88
B.11	Tabla de los parámetros asociados a las soluciones trigonométricas de la Tabla[4.11].	88
B.12	Tabla de los parámetros asociados a las expresiones obtenidas para la función de Hubble de la Tabla[5.1].	89
B.13	Tabla de los parámetros asociados a las expresiones obtenidas para la densidad de energía de la Tabla[5.2].	89
B.14	Tabla de los parámetros asociados a las expresiones obtenidas para las ecuaciones de estado de la Tabla[5.3].	89

Capítulo 1

Cosmología.

1.1 Introducción.

La cosmología es la ciencia que estudia la estructura, contenido y evolución del Universo como un todo. Comenzó a estudiarse desde el punto de vista teórico a principios del siglo XX con las ecuaciones de campo de la relatividad general de Einstein mientras que desde el punto de vista observacional, tomó impulso con el trabajo de científicos como Edwin Hubble quien postuló, después de estudiar las observaciones de corrimiento al rojo de varias galaxias, que el universo se expande. Al utilizar las ecuaciones de campo y asumir que el universo es homogéneo e isotrópico a partir de las escalas de 100 Mpc (principio cosmológico) se pueden obtener las ecuaciones de Friedmann y de continuidad las cuales describen el comportamiento del contenido y evolución del universo como función del tiempo y otros parámetros. Estas ecuaciones son fundamentales para entender cómo el universo ha evolucionado desde el Big Bang hasta su estado actual.

Además de lo anterior, se presenta en este capítulo una revisión sobre los componentes del universo incluida la materia y energía oscura. La materia oscura, aunque no interactúa con la luz, ejerce una influencia gravitacional significativa, mientras que la energía oscura es responsable de la aceleración en la expansión del universo. El modelo cosmológico Λ CDM, que incorpora estos componentes, es considerado actualmente el modelo estándar en cosmología. Las observaciones de este modelo estiman un valor para la constante de Hubble de $H_0 = 67.4 \pm 0.5 \frac{km}{Mpc \cdot s}$, el cual se encuentra dentro del rango aceptado actual de $H_0 = 70 \pm 7 \frac{km}{Mpc \cdot s}$.

Este capítulo también explorará otros modelos cosmológicos y teorías que buscan explicar la naturaleza y el destino del universo, proporcionando una visión integral de los avances y desafíos en el campo de la cosmología.

1.2 Principio Cosmológico

El principio cosmológico es una hipótesis fundamental en cosmología que es válida al contemplar el Universo a partir de los 100 Mega-parsec¹ (Mpc) y dicta que éste es homogéneo e isotrópico. La homogeneidad significa que todo el universo se ve y comporta igual sin importar la posición en la que se esté mientras que la isotropía implica que el Universo se ve y comporta sin importar la dirección en la que se mire [2].

El asumir alguna de las dos cualidades no implica la existencia de la otra. Como se puede apreciar en la Fig[1.1] un Universo en el que existiera un campo magnético constante \vec{B} no haría la diferencia al situarse en cualquier punto pues siempre se manifestaría el mismo vector asociado, sin embargo dicho Universo no sería isotrópico pues direcciones paralelas a la dirección del campo serían distinguibles de aquellas perpendiculares a él. De modo contrario, un Universo con distribución simétrica y esférica visto desde el centro sería isotrópico al notar lo mismo en cada dirección radial pero no sería necesariamente homogéneo pues podría distinguirse dependiendo de la posición en la que se encuentre (por ejemplo bajo un campo gravitacional mientras mas lejos del centro menor sería el efecto gravitatorio y la materia tendería a agruparse diferente haciendo que no se vea igual para todos los puntos). Sin embargo si se asume isotropía al rededor de cada punto del espacio entonces sí se obtiene automáticamente homogeneidad [3].

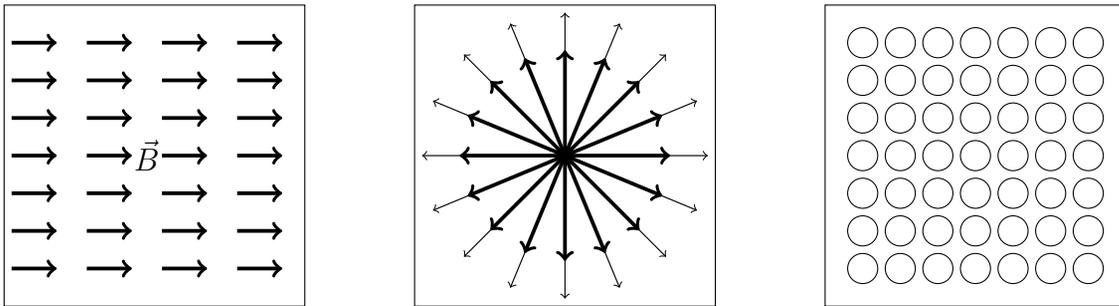


Figura 1.1: Universo homogéneo bajo un campo magnético constante \vec{B} (izquierda). Universo isotrópico con distribución simétrica esférica visto desde el centro bajo un campo gravitacional \vec{g} (centro). Universo isotrópico y homogéneo (derecha).

A pesar de la cantidad de evidencia de que el Universo realmente cumple con el principio cosmológico, gran parte de la investigación actual esta enfocada en mejorar

¹Un parsec es la distancia a la que una unidad astronómica subtende un ángulo de un segundo de arco. Usualmente se utiliza cuando las distancias son de más de miles de años luz por ejemplo las existentes entre galaxias son del orden de Mega parsecs (Mpc), para dimensionar, el diámetro de la Vía Láctea es de alrededor de 0.025 Mpc [1].

$$1pc = 3.2616 \text{ años luz} = 206265 \text{ unidades astronomicas} = 3.0857 \times 10^{16} \text{ metros.} \quad (1.1)$$

la precisión de las pruebas de comprobación o alternativamente encontrar casos en los que no se cumpla y esto guíe a temas nuevos en las cosmología [4, 5]

1.3 Ecuaciones de Einstein

Albert Einstein (1879,1955) Físico, publicó 4 artículos científicos [6,7,8,9] cuya influencia permitió sentar las bases de la Física moderna en el siglo XX. Tal fue la importancia que tuvieron las publicaciones y el hecho de haberlas escrito en un solo año que se le dio nombre a tal suceso como *annus mirabilis* o año milagroso. La tercera de éstas publicaciones *Zur Elektrodynamik bewegter Körper* ("Sobre la electrodinámica de los cuerpos en movimiento") que años mas tarde sería conocida como *Teoría de la Relatividad Especial* concilió las leyes de la Mecánica Clásica de Newton con las ecuaciones de Maxwell del Electromagnetismo. Dicha teoría puede ser resumida en dos postulados:

- El principio de Relatividad Galileano² se cumple para todos los marcos de referencia inercial.
- La rapidez de la luz es igual en todos los marcos de referencia inerciales (en el vacío).

A pesar del éxito que ganó esta teoría al explicar elegantemente las discrepancias que existían con el problema del éter y los experimentos de Michelson-Morley sobre la rapidez de la luz, no podía ser aplicada a sistemas acelerados (bajo acción de la gravedad por ejemplo) o rotatorios.

Durante los casi 10 años posteriores a la publicación de la relatividad especial, Einstein se vio motivado en parte a generalizarla debido a la evidente discrepancia entre su teoría y la interacción gravitacional a distancia instantánea que suponía la teoría de gravitación de Newton. Según ésta última, la fuerza gravitacional que ejerce el Sol sobre la Tierra sería más rápida que la rapidez de la luz al ser instantánea violando uno de los principios relativistas fundamentales.

El primero de los postulados de la relatividad general es el *principio de equivalencia* en el cual, puntualmente la distinción entre un sistema de referencia no inercial acelerado es indistinguible de aquel sistema inmerso en un campo gravitatorio [10]. Dicho de otra forma, localmente un sistema que se encuentra en caída libre y otro que se mueve en una región del espacio-tiempo sin gravedad son sistemas equivalentemente inerciales. Una de las consecuencias más importantes y notables de ésta teoría es que la gravedad

²Galileo Galilei (1564-1642) postuló que todas las leyes físicas (mecánicas en ese entonces) y las ecuaciones que las describen son las mismas en cualquier sistema que se mueva a velocidad constante con respecto a otro marco de referencia en reposo absoluto.

deja de ser una fuerza fundamental y pasa a ser una consecuencia de la distorsión del espacio-tiempo generada por la presencia de materia (o energía), es decir una fuerza aparente.

El segundo de los postulados es considerado una extensión del principio de relatividad especial, el llamado *principio de covariancia* expresa que las leyes así como las ecuaciones fundamentales de la física deben tener la misma forma para cualquier observador sea cual sea su estado de movimiento. Matemáticamente el principio de covariancia implica que las leyes de la física deben ser expresadas por tensores, de manera que las magnitudes medidas por diferentes observadores estén relacionadas a través del cambio de coordenadas (difeomorfismos) de cada observador.

En el marco de la Relatividad General se introduce una variedad 4-dimensional con vectores base e_μ equipada con la métrica $g_{\mu\nu}$, que representa la separación diferencial entre dos eventos (o intervalo) dada por:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (1.2)$$

Se ha utilizado la *convención de Einstein* en la cual, cuando se encuentran índices repetidos arriba y abajo en una expresión se infiere que existe una suma sobre esos índices. Dichos índices (griegos) corresponden a las componentes del espacio-tiempo 4-dimensional cuya primer componente $\mu = 0$ corresponde a la temporal mientras que las otras 3 componentes (expresadas con índices latinos i, j) corresponden a las componentes espaciales. De esta forma la ec(1.2) se expresaría de la forma:

$$\begin{aligned} ds^2 = g_{\mu\nu} dx^\mu dx^\nu &\equiv \sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu} dx^\mu dx^\nu \\ &= g_{00} dx^0 dx^0 + \sum_{i=1}^3 g_{i0} dx^i dx^0 + \sum_{j=1}^3 g_{0j} dx^0 dx^j + \sum_{i=1}^3 \sum_{j=1}^3 g_{ij} dx^i dx^j. \end{aligned}$$

Dicha métrica define los símbolos de Christoffel $\Gamma^\lambda_{\alpha\nu}$ los cuales describen la variación de los vectores base e_μ . Éstos símbolos se obtienen³ mediante [11]:

$$\Gamma^\lambda_{\alpha\nu} = \frac{1}{2} g^{\lambda\mu} (\partial_\nu g_{\mu\alpha} + \partial_\alpha g_{\nu\mu} - \partial_\mu g_{\alpha\nu}). \quad (1.3)$$

Donde ∂_ν representa la derivada parcial respecto a la ν -ésima variable. Por otra parte el llamado *tensor de Riemman* o *tensor de curvatura de Riemman* se define en términos

³En espacios sin torsión.

de los símbolos de Christoffel como :

$$R^{\beta}_{\alpha\mu\nu} = \partial_{\mu}\Gamma^{\beta}_{\alpha\nu} - \partial_{\nu}\Gamma^{\beta}_{\alpha\mu} + \Gamma^{\beta}_{\gamma\mu}\Gamma^{\gamma}_{\alpha\nu} - \Gamma^{\beta}_{\gamma\nu}\Gamma^{\gamma}_{\alpha\mu}. \quad (1.4)$$

A su vez, el tensor y escalar de Ricci como contracciones del tensor de Riemman son respectivamente :

$$R_{\alpha\beta} = R^{\mu}_{\alpha\mu\beta} = g^{\mu\nu}R_{\nu\alpha\mu\beta}, \quad (1.5)$$

$$R = g^{\mu\nu}R_{\mu\nu}. \quad (1.6)$$

Las ecuaciones (1.5,1.6) ayudan a definir el tensor (geométrico) de Einstein :

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R. \quad (1.7)$$

Al considerar el tensor de energía-momento $T^{\mu\nu}$ que contiene la información del contenido de energía y momento en cierta región del espacio, Einstein propuso que fundamentalmente existe una relación entre éstas dos ultimas expresiones dando como resultado las llamadas *ecuaciones de campo de Einstein*:

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \frac{8\pi G}{c^4}T^{\mu\nu}. \quad (1.8)$$

Donde c es la rapidez de la luz y G la constante de gravitación universal. Debido a las simetrías de $R^{\mu\nu}$ y $T^{\mu\nu}$ la ec(1.8) representa un conjunto de 10 ecuaciones diferenciales parciales acopladas.

1.4 Universo en expansión.

A diferencia del carácter experimental que pose la Física desde sus inicios, cuando se trata de 'corroborar' ciertas teorías o registrar información de los sucesos que en nuestro entorno, la principal fuente de ésta información proviene de observaciones realizadas por observatorios, satélites y telescopios. En 1913 Vesto Slipher (1875-1969) observó el desplazamiento de las líneas espectrales de la que entonces se conocía como nebulosa⁴ de Andrómeda mostrando que existía un corrimiento al azul en dicho espectro, por tanto concluyó que Andrómeda se acercaba al sistema Solar [12]. Éste mismo método fue utilizado para estudiar los espectros de una cantidad mayor de nebulosas de las cuales la gran mayoría presentó un corrimiento al rojo [13]. Además, las velocidades que obtuvo de dichas nebulosas resultaron ser unas 30 veces superiores a las de las estrellas por

⁴*Nube* o *niebla* en latín, es como se le conocía a los objetos difusos que se observaban cuando las herramientas de observación no eran muy precisas. Posteriormente con el avance de éstas y de la teoría se diferenció entre galaxias, nebulosas planetarias, nubes moleculares etc.

lo que de este hecho se sugeriría que se trataba de objetos fuera de nuestra galaxia, el Universo podría ser mas grande de lo imaginado [14].

El **corrimiento al rojo** sucede cuando una fuente de luz se aleja haciendo que sea detectada al final del espectro electromagnético. Utilizando un símil con el efecto Doppler cuando la fuente de luz se aleja a una velocidad relativa, la longitud de onda recibida aumenta en comparación con la emitida. De forma contraria el **corrimiento al azul** sucede cuando la fuente de luz se acerca haciendo que la longitud de onda recibida sea menor.

Ya que en general las galaxias se alejan de la Vía Láctea se suele referir puramente al corrimiento al rojo o *redshift* (adimensional) definido como:

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}. \quad (1.9)$$

Donde λ_{em} corresponde a la longitud de onda emitida y λ_{obs} la observada. El redshift de un objeto 'cercano'⁵ que se aleja con una rapidez v se aproxima como :

$$z \approx \frac{v}{c}. \quad (1.10)$$

Durante los años posteriores a las publicaciones de Vesto, Edwin Hubble (1889-1953) al utilizar sistemáticamente el mismo procedimiento para estudiar los espectros y velocidades de más galaxias encontró una relación de proporcionalidad existente entre la distancia a las galaxias y el corrimiento al rojo asociado, este postulado es conocido como la **ley de Hubble-Lemaître**⁶ y se expresa :

$$z = \frac{H_0}{c}r. \quad (1.11)$$

Donde r es la distancia a la galaxia, c la rapidez de la luz, z el corrimiento al rojo y H_0 la constante de proporcionalidad comúnmente conocida como *constante de Hubble*.

Hubble al percatarse de ésta relación lineal interpreto que dicho corrimiento era consecuencia de un efecto Doppler, teniendo las galaxias como fuente de luz alejándose radialmente lo que implicaría una expansión del universo [16]. Al considerar las ec(1.10,1.11) se obtiene la relación velocidad-distancia mostrando que mientras mas lejos se encuentra

⁵Una región de z en la que se desprecian los efectos gravitacionales por objetos cercanos pero no suficientemente larga mayor a $z = 0.1$.

⁶En 2011 el investigador Mario Livio explicó en la revista *Nature* que el astrónomo Belga George Lemaître (1894-1966) descubrió primero la expansión del Universo. Posteriormente en 2018 la Unión Astronómica Internacional decide renombrar el postulado en honor a la labor de los dos científicos [15].

una galaxia, mas rápido se aleja (Fig[1.2]):

$$v = H_0 r. \quad (1.12)$$

Con v la rapidez radial del objeto emisor (galaxia).

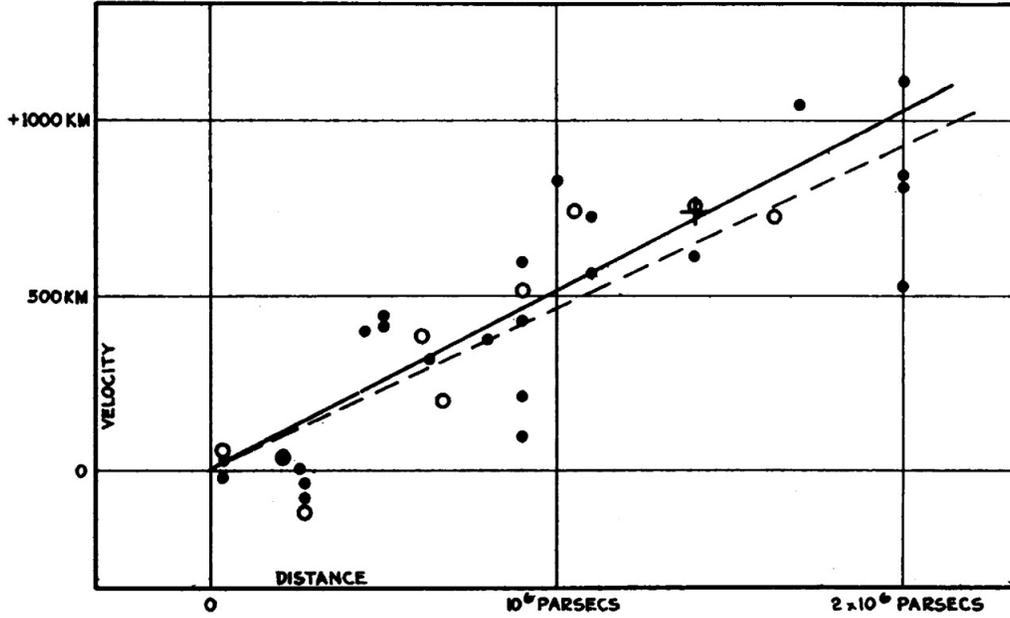


Figura 1.2: Gráfica publicada por Hubble sobre la relación velocidad-distancia para galaxias [16].

Los valores reportados en la literatura para H_0 se pueden encontrar en un rango desde 64.9 [17] hasta [18, 19] $72.5 \text{ kms}^{-1}\text{Mpc}^{-1}$.

No obstante el mejor valor estimado considerando distintas investigaciones [17, 18, 19, 20] es :

$$H_0 = 70 \pm 7 \frac{\text{km s}}{\text{Mpc}}. \quad (1.13)$$

1.5 Métrica FRLW y modelo de Friedmann

La métrica de Friedmann-Robertson-Lemaître-Walker (FRWK) representa la geometría que se adapta mejor al principio cosmológico y a las observaciones de expansión del universo. El intervalo en unidades naturales (donde se asume $c = 1$) está descrito por :

$$ds^2 = dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \text{sen}^2 \theta d\phi^2 \right]. \quad (1.14)$$

Donde $a(t)$ es conocido como el *factor de escala* y es una medida de la tasa de crecimiento de las distancia en o del universo como función del tiempo; las coordenadas espaciales r , θ , y ϕ representan la distancia radial, el ángulo polar y azimutal respectivamente. El parámetro k representa el tipo de curvatura del espacio y puede adquirir los siguientes valores [21]:

$$k = \begin{cases} +1 & \text{Universo cerrado} \\ -1 & \text{Universo abierto} \\ 0 & \text{Universo plano} \end{cases} \quad (1.15)$$

La Fig[1.3] muestra una representación de los posibles espacios con curvaturas diferentes.

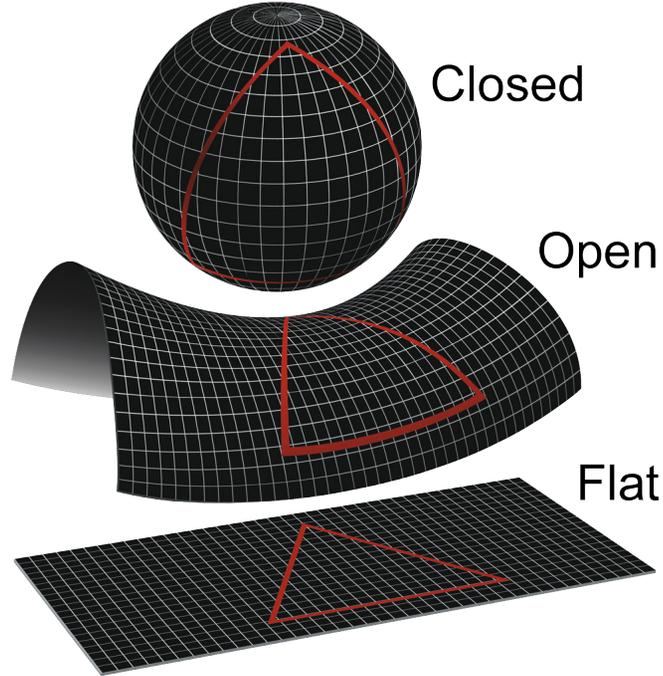


Figura 1.3: Universo esférico con curvatura $k = +1$ (arriba). Universo hiperbólico con curvatura negativa $k = -1$ (medio). Universo plano con curvatura nula $k = 0$ (abajo).

Para modelar al universo se asume un fluido perfecto que cumple las condiciones del principio cosmológico. Éste esta compuesto por partículas que no intercambian momento a través de conducción de calor ni esfuerzos de corte en el marco de referencia en reposo haciendo que las componentes T^{ij}, T^{0i}, T^{i0} con $i \neq j$ del tensor energía-momento asociado sean nulas. Por tanto un fluido perfecto dotado de presión P y densidad de energía ρ es una matriz diagonal :

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix} = \text{diag}(\rho, -P, -P, -P), \quad (1.16)$$

que en términos de la 4-velocidad U^μ y la métrica $g^{\mu\nu}$ se escribe como:

$$T^{\mu\nu} = (\rho + P)U^\mu U^\nu - P g^{\mu\nu}. \quad (1.17)$$

Además se debe considerar que las componentes no nulas del tensor y escalar de Ricci en este caso corresponden a :

$$R^{00} = -3\frac{\ddot{a}}{a}, \quad (1.18)$$

$$R^{ij} = -\left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2}\right]g^{ij}, \quad (1.19)$$

$$R = -6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right]. \quad (1.20)$$

Donde los puntos sobre las variables indican derivadas temporales.

Por otra parte a partir de reescribir la ec (1.9) como :

$$z + 1 = \frac{\lambda_{obs}}{\lambda_{em}}, \quad (1.21)$$

se puede buscar una manera de estudiar los objetos luminosos a grandes escalas. En un universo en expansión, las estrellas, galaxias, cúmulos y otros cuerpos celestes son las fuentes luminosas para observadores actuales. Considerando un rayo de luz proveniente de cierta dirección se cumple de la ec(1.14)

$$ds^2 = dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} \right] = 0 \implies \frac{dt}{a(t)} = -\frac{dr}{\sqrt{1 - kr^2}}, \quad (1.22)$$

por lo que para una posición r al tiempo $t = 0$ donde se emite el rayo, al recibirlo en $r = 0$ al tiempo $t = t_0$ se tiene:

$$\int_t^{t_0} \frac{dt'}{a(t')} = \int_0^r \frac{dr'}{\sqrt{1 - kr'^2}}. \quad (1.23)$$

Ahora bien, si se considera la luz como una onda clásica, se supone la emisión de dos crestas consecutivas a los tiempos t y $t + \Delta t$ que son detectadas a los tiempos t_0 y

$t_0 + \Delta t_0$ es decir:

$$\int_t^{t_0} \frac{dt'}{a(t')} = \int_{t+\Delta t}^{t_0+\Delta t_0} \frac{dt'}{a(t')}, \quad (1.24)$$

Lo que equivale a :

$$\int_t^{t+\Delta t} \frac{dt'}{a(t')} + \int_{t+\Delta t}^{t_0} \frac{dt'}{a(t')} = \int_{t+\Delta t}^{t_0} \frac{dt'}{a(t')} + \int_{t_0}^{t_0+\Delta t_0} \frac{dt'}{a(t')}. \quad (1.25)$$

Por lo tanto, si el factor de escala se mantiene constante durante la emisión de la luz al ser el periodo de la onda asociada suficientemente corto se tiene :

$$\frac{\Delta t}{\Delta t_0} = \frac{\lambda_{obs}}{\lambda_{em}} = \frac{a(t_0)}{a(t)}, \quad (1.26)$$

Entonces al igualar la ec(1.21) se obtiene la expresión para el *corrimiento al rojo cosmológico*:

$$z + 1 = \frac{a(t_0)}{a(t)}. \quad (1.27)$$

De manera que $(z + 1)$ indica cuánto se ha expandido el Universo desde que la luz fue emitida. La luz a su vez se expande en la misma proporción que el mismo Universo. [22]

1.5.1 Ecuación de Friednman

Para obtener la ecuación de Friedmann se emplea la componente $(\mu = 0, \nu = 0)$ de la ec(1.8) de manera que al sustituir $T^{00} = \rho$ y las ecuaciones (1.18,1.20) se obtiene :

$$-3\frac{\ddot{a}}{a} + \frac{1}{2}6 \left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} \right] = 8\pi G\rho, \quad (1.28)$$

que al simplificar se escribe como :

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (1.29)$$

Ahora si se define el *parámetro de Hubble* como función del tiempo t como:

$$H(t) := \frac{\dot{a}}{a} \quad (1.30)$$

se obtiene así finalmente la ecuación de Friedmann:

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (1.31)$$

Ésta ecuación describe como cambia el parámetro de Hubble como función del contenido del universo (la densidad de energía) y la forma que éste tiene.

1.5.2 Ecuación de continuidad

Ya que se esta trabajando con un fluido perfecto, debe respetarse el principio de conservación de la energía el cual utilizando la derivada covariante del tensor de energía momento refleja :

$$T_{\nu;\mu}^{\mu} = 0 , \quad (1.32)$$

que al considerar los símbolos de Christoffel para hacer la corrección al término de la derivada 'ordinaria' se reescribe como :

$$\partial_{\mu}T_{\nu}^{\mu} + \Gamma_{\alpha\mu}^{\mu}T_{\nu}^{\alpha} - \Gamma_{\nu\mu}^{\alpha}T_{\alpha}^{\mu} = 0. \quad (1.33)$$

Ya que la densidad de energía esta presente en la componente $\nu = 0$ se tiene:

$$\partial_{\mu}T_0^{\mu} + \Gamma_{\alpha\mu}^{\mu}T_0^{\alpha} - \Gamma_{0\mu}^{\alpha}T_{\alpha}^{\mu} = 0 , \quad (1.34)$$

y como T_{ν}^{μ} es diagonal, los símbolos de Christoffel relevantes son aquellos con índices repetidos, a saber :

$$\Gamma_{00}^0 = 0 ; \quad \Gamma_{01}^1 = \Gamma_{02}^2 = \Gamma_{03}^3 = \frac{\dot{a}}{a} = H. \quad (1.35)$$

Al sustituir los valores de la ec(1.35) en ec(1.34) se obtiene la conocida *ecuación de continuidad*:

$$\dot{\rho} + 3H(\rho + P) = 0. \quad (1.36)$$

La cual describe el comportamiento y evolución del contenido del Universo como función de la densidad de energía y de la Presión [23]. Sin embargo, aún se necesita una ecuación que relacione al fluido con sus variables termodinámicas. La *ecuación* de estado faltante se asume como la de un fluido perfecto barotrópico , es decir, un fluido en el que las isobaras (lineas de presión constante) coinciden con las isopícnas (lineas de densidad constante) de manera que la presión solo depende de la densidad :

$$P(\rho) = \omega\rho. \quad (1.37)$$

1.6 La composición del universo

La ecuación barotrópica permite diferenciar los distintos componentes que conforman el universo y como cada parámetro asociado modifica la presión existente a través de la ecuación de estado y por ende el comportamiento general del universo. Matemáticamente se expresa de la siguiente forma donde γ_i caracteriza cada uno de los componentes:

$$P_i = (\gamma_i - 1)\rho_i. \quad (1.38)$$

1.6.1 Materia

A la materia se le asigna el valor $\gamma_m = 1$ de manera que la ecuación barotrópica adquiere el valor $P_m = 0$. Ésta presión concuerda con lo esperado en un fluido cuyas partículas se encuentran lo suficientemente separadas para que no exista una interacción entre ellas y como consecuencia no ejerzan ninguna fuerza.

De la ecuación de continuidad (ec1.36) y el valor para la presión se tiene :

$$\dot{\rho}_m + 3H\rho_m = 0 \implies \frac{\dot{\rho}_m}{\rho_m} = -3\frac{\dot{a}}{a}, \quad (1.39)$$

al recordar que $H = \frac{\dot{a}}{a}$. Por lo tanto la densidad de materia está dada por la siguiente proporción:

$$\rho_m \propto \frac{1}{a^3}. \quad (1.40)$$

Resultado esperado, ya que un universo con la misma cantidad de materia reducirá la densidad al aumentar el volumen del espacio debido a la expansión.

Al considerar la densidad de materia al día de hoy $\rho_{m,0}$ se tiene :

$$\rho_m = \frac{\rho_{m,0}}{a^3} \quad (1.41)$$

Existen dos clasificaciones usuales de la materia o simplemente llamada como *polvo*.

Materia Bariónica

La materia bariónica es aquella conformada por bariones (partículas sub-atómicas como los neutrones y protones que están compuestas por 3 quarks) . En cosmología y astrofísica es común incluir también a los leptones como los electrones y otras partículas no relativistas de manera que en general, la materia bariónica es aquella que conforma todos los elementos visibles en el Universo como los planetas, estrellas, galaxias, nubes moleculares etc.

Materia Oscura

Poco tiempo después de la observación de que el universo se expande surgió entre los astrónomos, físicos y cosmólogos un inquietante resultado. Al estudiar la dinámica en el interior del cúmulo de galaxias Coma Berenice [24], Fritz Zwicky (1898-1974) relacionó la masa del conjunto de los cuerpos unidos por interacción gravitacional con sus respectivas velocidades. Los resultados indicaron una masa mucho mayor de lo que en un principio se había obtenido sugiriendo así un *ingrediente extra* que ejercía una fuerza gravitacional lo suficientemente grande para impedir la expansión del cúmulo [26]. Debido a esto Zwicky propuso la existencia de una materia desconocida que no interactúa con la luz pero sí gravitacionalmente con la materia bariónica.

Un par de décadas más tarde Vera Rubin (1928-2016) pionera en la medición de la rotación de las estrellas dentro de una galaxia [25], estimó la velocidad de rotación como función de la distancia al centro de la galaxia descubriendo que ésta no disminuye conforme las estrellas se alejan del centro lo cual es una clara discrepancia con las predicciones Newtonianas/Keplerianas en las que un objeto al estar más lejos de la fuente del campo gravitacional se mueve a mayor lentitud. El modelo de la velocidad considera la galaxia como una distribución de masa en función del radio respecto a su centro como se muestra en la ec(1.42). A grandes distancias dicho modelo predice que la masa se aproxima constante y por tanto, la velocidad de rotación disminuye como la raíz cuadrada de R . Las observaciones por el contrario muestran que la velocidad es mucho mayor a la esperada como se muestra en la Fig[1.4]. Al existir esta discrepancia la hipótesis más aceptada es la existencia de la *materia oscura* la cual al no interactuar con el campo electromagnético no se puede detectar visiblemente pero sí mediante su influencia gravitacional. A pesar de que existen muchos modelos candidatos a materia oscura, la que más tiene relevancia en la actualidad es el de la **materia oscura fría (CDM)** (Cold Dark Matter) por sus siglas en inglés, en el cual, las partículas no se mueven a velocidades relativistas lo que les permite agruparse gravitacionalmente (a éste tipo de materia se le conoce como materia fría); son transparentes en el sentido de no interactuar más que gravitacionalmente con la materia bariónica sin la radiación y tampoco colisionan [27].

$$v(R) = \sqrt{\frac{GM(R)}{R}}. \quad (1.42)$$

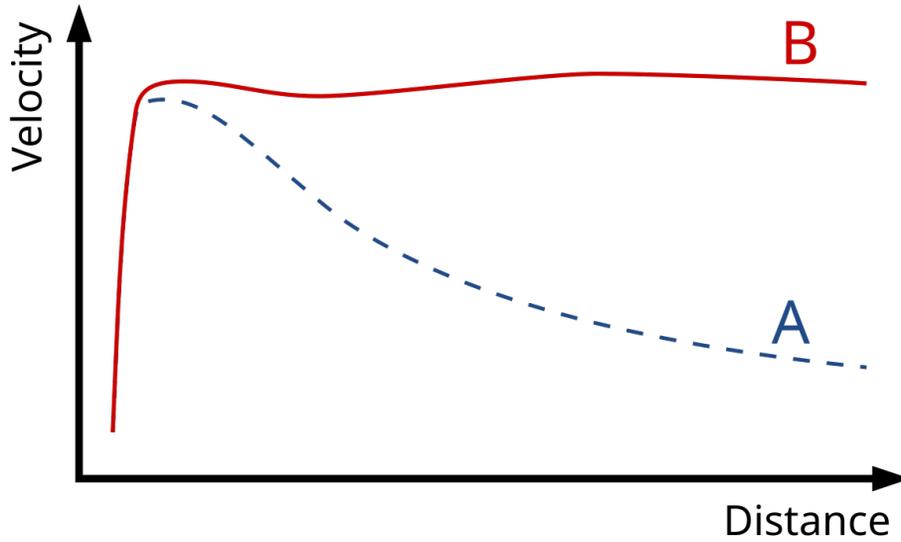


Figura 1.4: Curvas de rotación Galáctica. La curva A punteada indica la predicción del modelo Kepleriano en el que la velocidad de una distribución de masa disminuye conforme se aleja del centro gravitacional. La curva B sólida indica las observaciones registradas para la velocidad indicando así la presencia de materia oscura [28].

1.6.2 Radiación

La radiación incluye a las partículas relativistas como los fotones, un tipo de bosón de masa nula capaz de interactuar con los bariones y electrones, y son las partículas mediadoras de la interacción electromagnética. También los neutrinos son considerados en ocasiones como parte de la radiación, son partículas elementales eléctricamente neutras cuya interacción con la materia es muy débil.

El valor γ asociado para la radiación es $\gamma_r = \frac{4}{3}$ de manera que $P_r = \frac{1}{3}\rho_r$. Por lo que la ecuación de continuidad toma la forma:

$$\dot{\rho}_m + 3H(\rho_m + \frac{1}{3}\rho_r) = 0 \implies \frac{\dot{\rho}_m}{\rho_m} = -4\frac{\dot{a}}{a}. \quad (1.43)$$

Cuya solución considerando la densidad de radiación actual $\rho_{r,0}$ es :

$$\rho_r = \frac{\rho_{r,0}}{a^4}. \quad (1.44)$$

La potencia extra en el factor de escala es debido al *estiramiento* de la longitud de onda de la luz debido a la relación de proporcionalidad entre ésta y la energía de radiación dada por $E = h/\lambda$.

1.6.3 Energía Oscura

Después de formular la teoría de la relatividad general y aplicarla para modelar el universo Einstein introdujo *a mano* un término que evitara la expansión del universo, es decir que éste se mantuviera estático (ver ec(1.48)). Éste término es conocido como constante cosmológica Λ y aunque Hubble mostró que el universo se expandía llevando a pensar a Einstein que dicha constante sería un error, con el paso del tiempo se volvió a interpretar la existencia y consecuencias de la constante Λ relacionándola con un nuevo concepto : La energía oscura.

De la ecuación barotrópica (1.38) a la energía oscura se le asigna el valor $\gamma_\Lambda = 0$ de manera que:

$$P_\Lambda = -\rho_\Lambda, \quad (1.45)$$

lo cual exhibe que la presión que ejerce este tipo de energía es negativa actuando con un comportamiento contrario a la gravedad (repulsivo) y se mantiene constante en el tiempo como se muestra en

$$\dot{\rho}_\Lambda = 0. \quad (1.46)$$

Se cree que la energía oscura esta intrínsecamente ligada al espacio [29]. De la ecuación 1.48 se puede extraer que de describir Λ como un fluido con densidad de energía ρ_Λ ésta se puede escribir como :

$$\rho_\Lambda = \frac{\Lambda}{8\pi G}, \quad (1.47)$$

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}. \quad (1.48)$$

1.7 Parámetros de densidad.

A pesar de que las cantidades descritas en las secciones anteriores son suficientes para estudiar la evolución y estructura del universo, desde el punto de vista observacional no todas se pueden obtener directamente por lo que es conveniente definir ciertos *parámetros observables* con los cuales se pueda comparar la teoría correspondiente.

La densidad crítica ρ_c es el valor necesario para que el parámetro de Hubble H presente una geometría plana $k = 0$, es decir :

$$\rho_c = \frac{3H^2}{8\pi G}. \quad (1.49)$$

Considerando ésto, el parámetro de densidad Ω está definido como el cociente de la densidad de cierta componente del universo entre la densidad crítica. Aunque es adi-

dimensional, Ω varía en el tiempo pues la densidad también lo hace:

$$\Omega = \frac{\rho}{\rho_c}. \quad (1.50)$$

Es posible además definir un parámetro de densidad para la curvatura haciendo :

$$\Omega_k = \frac{-k}{(aH)^2}. \quad (1.51)$$

Por lo que la ecuación de Friedmann (1.31) puede reescribirse en términos de los parámetros de densidad como:

$$\frac{H^2}{H_0^2} = \Omega_{r,0}a^{-4} + (\Omega_{b,0} + \Omega_{DM,0})a^{-3} + \Omega_{k,0}a^{-2} + \Omega_{\Lambda,0}. \quad (1.52)$$

Donde $\Omega_{r,0}$ es el parámetros de densidad de la radiación, $\Omega_{b,0}$ y $\Omega_{DM,0}$ los parámetros de materia bariónica y oscura respectivamente, $\Omega_{k,0}$ el parámetro de densidad de curvatura, $\Omega_{\Lambda,0}$ el asociado con la constante cosmológica y H_0 el parámetro de Hubble. Los subíndices cero indican los valores evaluados al día de hoy [23]. Además, se respeta la relación

$$\Omega_{r,0} + \Omega_{b,0} + \Omega_{DM,0} + \Omega_{k,0} + \Omega_{\Lambda,0} = 1. \quad (1.53)$$

1.8 Modelo Λ CDM

El modelo Λ CDM es considerado el modelo estándar de la cosmología contemporánea por su simplicidad y conformidad con los datos. La energía oscura se considera la constante cosmológica Λ descrita por la ec(1.47) como un fluido perfecto con la ecuación de estado asociada 1.45. El otro componente es la materia oscura fría CDM descrita en la sección (1.6.1). La radiación no es tomada en cuenta ya que la densidad de radiación es muy baja actualmente. Considerando la relación de la ec(1.27) con $a(t=0) = 1$ se tiene la ecuación de Friedmann como:

$$H^2 = H_0^2[(\Omega_{CDM,0} + \Omega_{b,0})(1+z)^3 + \Omega_{\Lambda,0}] \quad (1.54)$$

Utilizando los datos del satélite *Planck* tomados del CMB (fondo cósmico de microondas) se han estimado los parámetros de densidad y de la constante de Hubble con un 68% de confianza. A partir de estas mediciones se estiman los valores de $\Omega_{\Lambda,0} = 0.685 \pm 0.007$, $\Omega_{CDM,0} + \Omega_{b,0} = 0.315 \pm 0.007$ y $H_0 = 67.4 \pm 0.5$ de donde se concluye que la energía oscura abarca aproximadamente el 70% del contenido del universo. La combinación de los valores presentados proporciona indicios de que el parámetro de densidad de cur-

vatura es muy cercano a cero de manera que el universo tendría una geometría plana [30].

A pesar de lo bien que el modelo puede describir el universo, el hecho de que se sepa poco sobre la naturaleza de la materia y energía oscura es un punto débil sobre el que se sigue realizando diversidad de investigaciones. Además, existen diferencias en los valores que debería tener la constante de Hubble (parámetro de Hubble al día de hoy) ya que por ejemplo con los datos obtenidos del CMB se obtiene un valor de $H_0 = 67$ mientras que de los obtenidos utilizando Supernovas tipo I se estima un valor de $H_0 = 71$ [31]; éste problema es conocido como la *tensión de Hubble*.

Lo anterior ha llevado a proponer modelos cosmológicos distintos en los que la energía oscura no figura como la constante cosmológica.

1.9 Modelo CPL

El modelo de Chevallier-Polarski-Linder es una parametrización de la ecuación de estado w considerándola como una función lineal del factor de escala [32] :

$$w(a) = w_0 + w_a(1 - a) = w_0 + w_a \left(\frac{z}{1+z} \right). \quad (1.55)$$

Con :

$$w_a = \left. \frac{dw}{dz} \right|_{z=0} \quad (1.56)$$

Donde w_0 y w_a son números reales con w_0 representando el valor actual de w que proporciona información sobre el comportamiento de la energía oscura en la etapa cosmológica actual, mientras que w_a indica el cambio de la ecuación de estado respecto al redshift z . Para $w_a > 0$ se tiene que w aumenta con el desplazamiento al rojo mientras que para $w_a < 0$ indica decrecimiento. El modelo **CPL** describe una evolución bastante gradual desde el valor $w = w_0 + w_a$ en épocas tempranas al valor presente $w = w_0$ además de reproducir las observables como el parámetro de Hubble con alta precisión [33].

La ecuación de continuidad (1.36) en éste caso toma la expresión:

$$\rho = \rho_0 e^{\int_0^z 3 \frac{1+w(z)}{1+z} dz} = \rho_0 (1+z)^{3(1+w_0+w_a)} e^{-\frac{3w_a z}{1+z}}. \quad (1.57)$$

De manera que la ecuación de Friedmann puede reescribirse como:

$$H(z)^2 = H_0^2 \left[\Omega_{m,0} (1+z)^3 + (1 - \Omega_{m,0}) (1+z)^{3(1+w_0+w_a)} e^{-\frac{3w_a z}{1+z}} \right]. \quad (1.58)$$

1.10 Modelo PolyCDM

Este es un modelo polinomial que parametriza el parametro de Hubble. La densidad de energía de la energía oscura ρ_{DE} es descrita por un polinomio cuadrático en cuyo modelo existe flexibilidad en cuanto a la evolución de la energía oscura a bajo redshift [34]. El parámetro de Hubble es escrito como :

$$H(z)^2 = \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_1(1+z)^2 + \Omega_2(1+z) + (1 - \Omega_m - \Omega_k - \Omega_1 - \Omega_2) \quad (1.59)$$

En donde $\Omega_{m,0}$ corresponde a la materia oscura mientras que $\Omega_{2,0}$ se puede interpretar como *materia faltante* [35].

1.11 Observaciones

En la actualidad se cuenta con cada vez mayor cantidad de datos observacionales para poder estudiar la expansión del universo así como su evolución.

El parámetro de Hubble depende de la edad diferencial del Universo, mide el tiempo que transcurrió entre dos eventos de su historia tales como dos redshifts diferentes. Esta diferencia de edad se relaciona con la expansión cosmológica a través del corrimiento al rojo z y es posible medirla utilizando los cronómetros cósmicos. Al medir las diferencias de edad dt entre dos galaxias formadas al mismo tiempo pero separadas por un corrimiento al rojo pequeño dz se puede aproximar el ratio $\frac{dz}{dt}$. Como se sugiere en [36] las galaxias a analizar mediante espectroscopía pueden funcionar como cronómetros si poseen metalicidad similar y son rojas (galaxias cuya formación de estrellas fue muy rápida en un inicio y en la actualidad la tasa de formación es baja) de manera que las estrellas en promedio tengan una edad significativamente mayor que la diferencia de edad entre las muestras de galaxias. El parámetro de Hubble $H(z)$ depende de la variación diferencial del corrimiento al rojo respecto del tiempo cósmico mediante la relación

$$H(z) = -\frac{1}{1+z} \left(\frac{dz}{dt} \right). \quad (1.60)$$

En este trabajo se utilizaron 26 observaciones de $H(z)$ que pueden verse como función del corrimiento al rojo en la Fig[1.5].

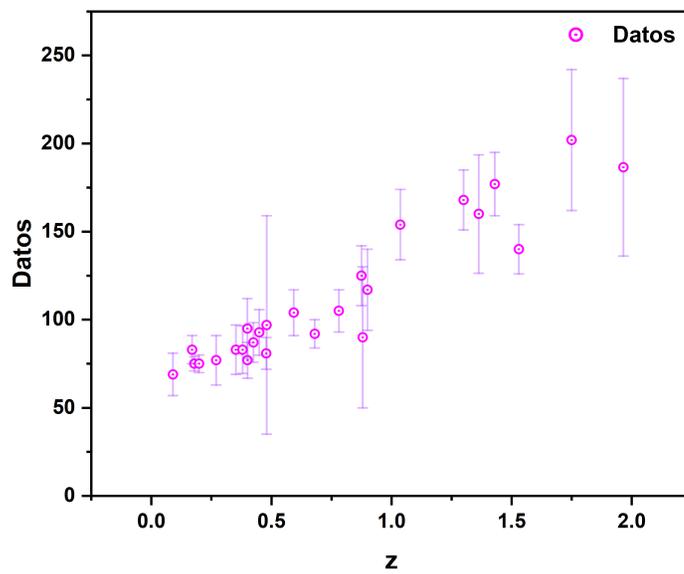


Figura 1.5: Datos de Hubble en función del corrimiento al rojo z para cronómetros cósmicos [37].

Capítulo 2

Computo evolutivo.

2.1 Introducción.

En un mundo donde la información se multiplica en diversas áreas de la ciencia y los problemas se vuelven cada vez más complejos, las herramientas computacionales han encontrado inspiración en los mecanismos naturales para resolver cuestiones de interés de manera sencilla y eficaz. Los algoritmos bioinspirados se pueden clasificar en tres grupos según su fundamento biológico: Algoritmos Evolutivos, Inteligencia de Enjambre y Ecología [38].

Los Algoritmos Evolutivos se basan en las leyes de la evolución, selección natural y mutación. Utilizan el carácter estocástico para la búsqueda de soluciones y resolver problemas de optimización. Ejemplos de éstos métodos incluyen la Evolución Diferencial [39], los Algoritmos Genéticos [40] y la Programación Genética [62].

En este capítulo, se introducen algunos conceptos relacionados con la programación simbólica, así como la representación de árboles de aridad (M). Se estudia cómo un programa puede ser manipulado considerando las instrucciones aritméticas o lógicas como variables. Además, se presentarán los conceptos bioinspirados de la teoría de la evolución de Charles Darwin y cómo es que se aplican al paradigma de la programación genética.

2.2 Programación simbólica

La **programación simbólica** es un paradigma que se centra en la manipulación de símbolos y reglas para definir el comportamiento del programa de manera que puede tratarse a sí mismo, o a material similar, como si fueran datos. Los programas pueden escribir programas (no solo como cadenas de caracteres o textos, sino como las estructuras de datos reales que conforman el programa) [41]. En comparación, la **programación funcional** se centra en la definición de funciones como unidades de código autocontenidas que reciben entradas, realizan cálculos y devuelven salidas [42]. Una vez que estas funciones son definidas no cambian en el tiempo, es decir son inmutables.

La diferencia en ambos paradigmas puede apreciarse desde la forma en que se escriben las variables y funciones a utilizar. Una representación de programación simbólica (utilizada ampliamente en el contexto de programación genética de la sección (2.3.4)) es la notación de prefijo, de *Lukasiewicz* o también conocida por su nacionalidad, representación *Polaca* [43]). Los operadores se escriben a la izquierda de las variables del dominio correspondiente de manera que la ejecución de la expresión se lee de izquierda a derecha.

Para la suma de dos números x y y la representación sería:

```
suma(x,y)--> x suma y
```

De esta manera la ejecución para $x = 5$, $y = 3$ corresponde a :

```
suma(x,y)--> suma(5,y)--> suma(5,3) --> 5 suma 3
```

Al recordar que en general, $\text{suma} \iff \text{" + "}$, se recupera la expresión simbólica conocida de una suma (la última ejecución de asignar 8 a la suma de los números 5 y 3 es en general no necesaria, aunque suele hacerse por motivos de simplificación) :

```
+(x,y)--> +(5,y)--> +(5,3) --> 5 + 3 --> 8
```

En contraste la implementación de una *función matemática* es :

```
def suma(x,y):  
    return x+y
```

Por lo que la ejecución para $x = 5$, $y = 3$ corresponde a :

```
print(suma(5,3))
```

```
8
```

Al considerar un ejemplo mas elaborado se puede comparar con mayor detalle la estructura de ambos tipos de estructuras y su ejecución.

Una representación simbólica del polinomio $p(x) = x^2 - x + \frac{\pi}{2}$ en notación de Lukasiewicz utilizando las operaciones básicas $\{+, -, *, /\}$ \iff $\{\text{sum, rest, prod, div}\}$ puede realizarse de la siguiente manera :

```
sum(sum(prod(x,x),prod(-1,x)),div(PI,2))
```

La cual se ejecuta de la siguiente forma¹ (cada operador en ejecución es representado por la asignación de su simbolo, i.e *sum* \rightarrow +):

```
sum(sum(prod(x,x),prod(-1,x)),div(PI,2))--> sum(prod(x,x),prod(-1,x))
+div(PI,2)--> prod(x,x)+prod(-1,x)+div(PI,2)-->x*x+(-x)+PI/2
```

En contraste, una representación funcional del mismo polinomio es:

```
def p(x) :
    return x*x-x+numeric(PI)/2
```

Las respectivas ejecuciones para $x = 2$ son :

```
sum(sum(prod(2,2),prod(-1,2)),div(3,2))--> sum(prod(2,2),prod(-1,2))
+div(PI,2) --> prod(2,2)+prod(-1,2)+div(PI,2)-->2*2+(-2)+PI/2
--> 4-2+PI/2 -->-2+PI/2
```

```
print(p(2))
3. 57079632679489661923132169163975144209858469968755...
```

Como puede observarse, la notación polaca puede ser complicada de leer e implementar aún para expresiones relativamente sencillas como un polinomio. No obstante a diferencia de la funcional, los términos se mantienen intactos ya que podría considerarse como ejecución final el término $-2 + 2 * 2 + PI/2$. Otro aspecto a considerar es que la expresión funcional recibió la entrada $x = 2$, evaluó numéricamente la expresión definida originalmente y regreso el número correspondiente. Dicha expresión no puede ser modificada por ejemplo al ser utilizada en un programa con mas estructura, no puede cambiarse "el segundo término de la función (a saber $-x$)" por el termino $+x$ sin redefinir la función, situación que si puede realizarse con la programación simbólica. Para hacer mas evidente este detalle es necesario conocer una manera mas intuitiva y gráfica de representar estas estructuras.

¹En realidad las propiedades conocidas de los reales se conservan bajo esta notación , por ejemplo la asociatividad bajo la suma hace que la adición de tres numeros $a+b+c$ pueda representarse como $+(a,b,c)$ en lugar de $+(+(a,b),c)$

2.2.1 Árboles de aridad-M

Un árbol es una estructura *no lineal* que sirve para organizar objetos o representar estructuras algebraicas. Quizás la representación mas conocida de éste tipo es la descendencia a partir de uno de los miembros de una pareja, es decir un árbol genealógico como el de la Fig[2.1]. La **aridad** de un operador, función o árbol es el numero mínimo

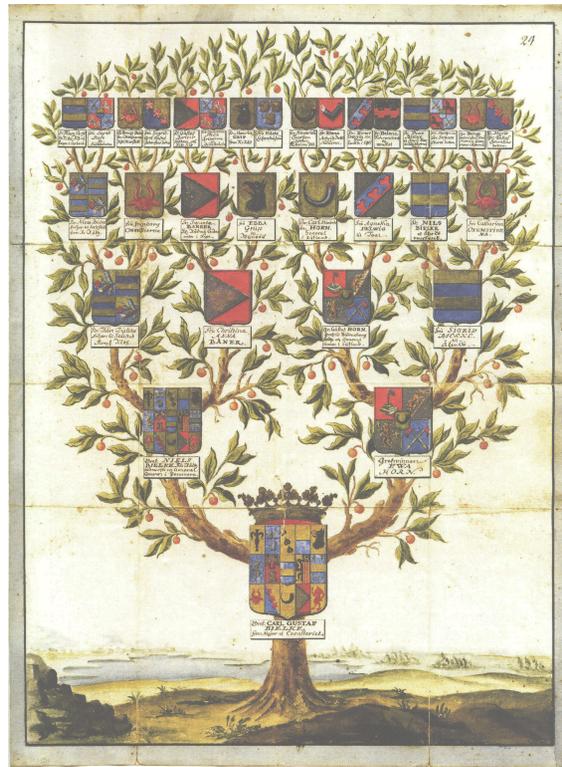


Figura 2.1: Árbol genealógico

de argumentos que requieren para poder ser evaluados o construidos. La operación *producto* de números es de aridad 2 ya que se necesitan dos números para aplicarla. La función trigonométrica seno, $\text{sen}(x)$, es de aridad 1 ya que solo requiere de un elemento para ser evaluada. Por lo tanto, un **árbol de aridad-M** en el contexto computacional es la representación gráfica de una estructura conformada por conjunto finito de nodos conectados por m -ramas. Como se aprecia en el árbol de la Fig[2.2] leído de arriba a abajo, el nodo A que no proviene de ningún otro se le conoce como **raiz** (ya que en esta representación los árboles se esquematizan invertidos). A su vez, los siguientes nodos están conectados por un número finito m de líneas llamadas **ramas**.

La definición de **sub-árbol** tiende a ser recursiva: Es una colección de nodos jerárquicamente estructurados que descienden todos de un nodo perteneciente a un árbol (un árbol dentro de un árbol).

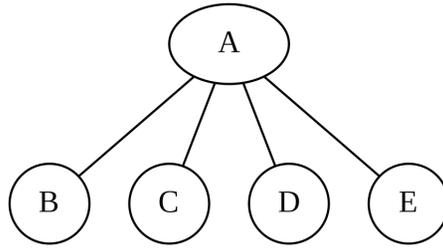


Figura 2.2: Árbol de 4 ramas (aridad 4)

Como se observa en la Figura [2.3], en los árboles con subárboles, existen nodos terminales de la estructura de los que no descienden más nodos. Estos elementos se denominan **hojas** [44].

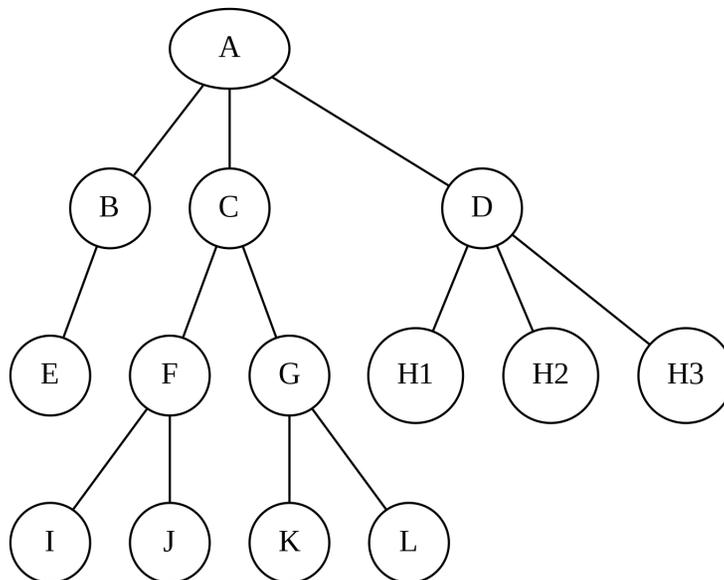


Figura 2.3: Árbol de aridad 3 (nodo A) con sub-árboles de aridad 1 (nodo B) ; aridad 2 (nodos C , F y G) y aridad 3 (nodo D) . Los nodos E,I,J,K,L,H1,H2 y H3 son hojas.

Algunos términos relacionados con los árboles de aridad-M son los siguientes (Refiérase nuevamente a la Fig[2.3]).

- Un **camino** de un nodo n_1 al nodo n_2 es una secuencia de ramas contiguas que van de n_1 a n_2 . El camino ACGL es mas largo que el camino ADH1.
- El **nivel** de un nodo es su distancia a la raíz del árbol. El nivel de A es 0, el nivel de H1 es 2 mientras que el de J es 3.
- La profundidad o **altura** de un árbol es la longitud del camino mas largo de la raíz a una hoja. La longitud del camino ABE es 3, mientras que la del camino ACFJ es 4. La profundidad del arbol de la Fig[2.3] es 4, mientras que la del arbol en Fig [2.2] es 2.

- Los nodos que no son ramas ni raíces se denominan **nodos internos**.

Una terminología ampliamente utilizada y que se usara de manera indistinta en el texto esta basada en el símil con un árbol genealógico.

- Un nodo **padre** es un nodo que tiene nodos sucesores. A es padre de B, C y D.
- Los sucesores de un padre se denominan **hijos**. B, C y D son hijos de A, además cumplen entre ellos ser **hermanos**. Es importante notar que aunque E, F y G se encuentran en el mismo nivel (segundo), E no es hermano de F ni de G.
- Los nodos sucesores a un padre se les conoce como **descendientes**. Por ejemplo, C, F, J y K son descendientes de A, mientras que C y F son **ascendientes** de J e I.

La representación de árbol en la Fig[2.4] corresponde al polinomio de la sec(2.2):

```
sum(sum(prod(x,x),prod(-1,x)),div(PI,2))
```

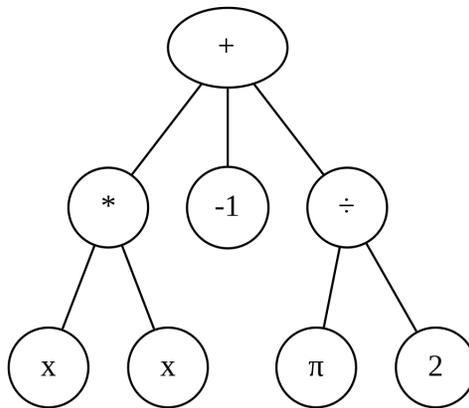


Figura 2.4: Representación de árbol aridad 3 para el polinomio $p(x) = x^2 - x + \frac{\pi}{2}$

2.2.2 Programa simbólico de juguete: Derivadas parciales.

Conceptualmente los programas simbólicos pueden entenderse a través de representaciones de arboles de aridad- m . El programa llamado *Derivadas Parciales* representado en la Fig[2.5] contiene la siguiente notación en sus nodos:

- $\frac{\partial^n}{\partial x_i^n}$:= Calcula la derivada parcial de orden n respecto a la i -ésima variable.
- $\text{RC1}\{0, \dots, n\}$:= Elección aleatoria de un natural de 0 a n
- $\text{if } i \leq |F|$:= Condicional de que el número i sea menor o igual que el número de variables de la función F .
- $i = \text{RC5}$:= asignación de i como un número elegido aleatoriamente
- $F = \{f, \&, f\}$:= Construcción de una función a partir de dos funciones y una operación
- $\text{RC2}, \text{RC4}^2$:= Elección aleatoria de una función.
- RC3 := Elección aleatoria de una operación
- $g(x_i), h(x_i), s(x_i)$ son funciones de varias variables.

La ejecución del programa es la siguiente :

Se ingresa un número n y tres funciones de varias variables (no necesariamente del mismo número y tipo). El programa selecciona aleatoriamente el orden de la derivada en un conjunto de 0 a n . De la misma forma, se eligen aleatoriamente dos de las funciones y una operación a aplicar entre ellas para generar la función F . Además, se elige un entero i el cual de ser menor o igual que el número de variables de la función F , permite la continuación a la ejecución al cálculo de la derivada parcial respecto a la i -ésima variable.

² RC_j de *Random Choice*.

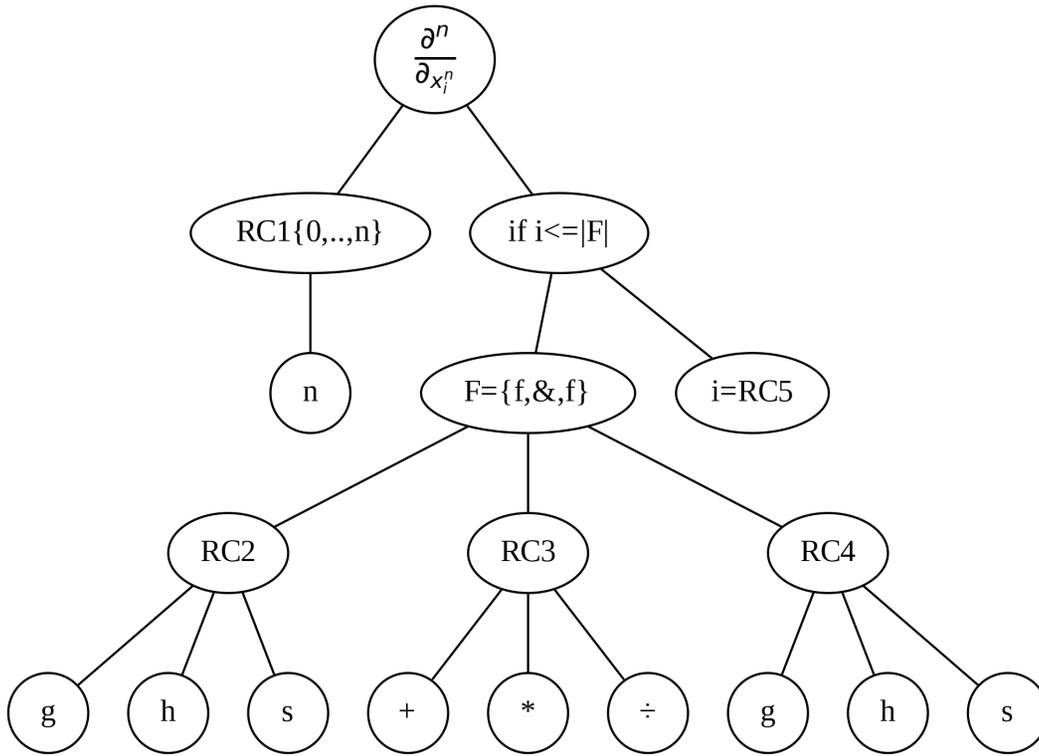


Figura 2.5: Esquema de árbol del programa *Derivadas Parciales*

Para explicitar, en la Fig[2.6] se presenta la ejecución para :

- $n=4$; $RC5=3$; $h(x) = 2x - 3$; $g(x, y, z) = x^2 + yz^5$; $s(z, w) = w(3 + z)$

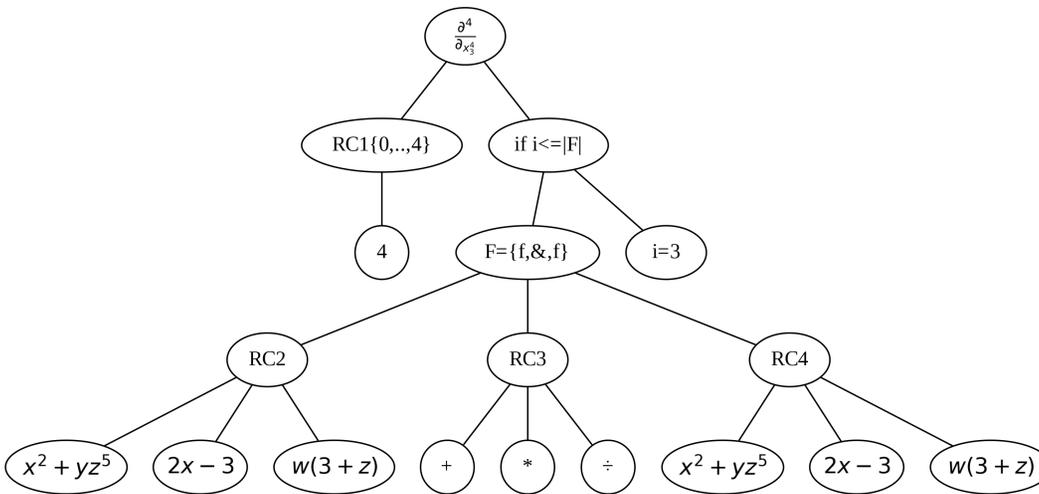


Figura 2.6: Condición inicial del programa .

Suponiendo las siguientes elecciones aleatorias, el árbol se reduce como el mostrado en la Fig[2.7]:

- $RC2=g(x, y, z)$; $RC3= +$; $RC4 =h(x)$

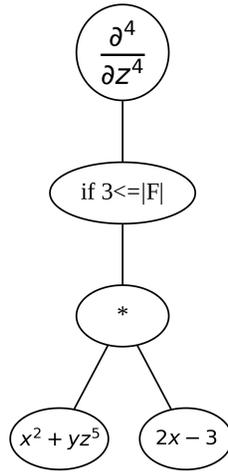


Figura 2.7: Simplificación del árbol al realizar las instrucciones en los nodos inferiores.

De esta manera, ya que $3 \leq 3$, ($F(x_1, x_2, x_3) = F(x, y, z)$) el programa puede ejecutarse completamente dando como resultado:

$$\frac{\partial^4}{\partial z^4} F = \frac{\partial^4}{\partial z^4} (x^2 + yz^5)(2x - 3) = 120(2x - 3)(yz)$$

En resumen, se puede hacer hincapié en dos hechos :

1. Las entradas (o variables) al programa son las tres funciones h, g, s y el entero n .
2. La salida del programa es una función.

Considerando estos hechos, *Derivadas Parciales* se puede expresar con un árbol de dos nodos. El nodo inferior representa las variables de entrada y el nodo superior color azul la salida del programa. Véase Fig[2.8]

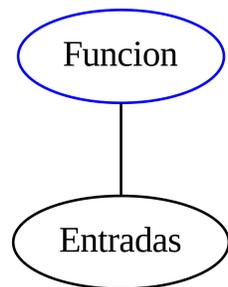


Figura 2.8: Simplificación del programa *Derivada parcial*.

2.2.3 Cambios en los elementos de un árbol.

Para ejemplificar las ventajas del uso de la representación de árbol nótese de la Fig[2.9] el cambio general cuando se sustituye el nodo RC4 en *Calculo de derivada* por una copia de *Calculo de derivada*. Se consideran las mismas variables de entrada (condiciones iniciales) de la sección anterior donde sea el caso.

Las soluciones posibles han cambiado pues ahora una de las funciones del nodo

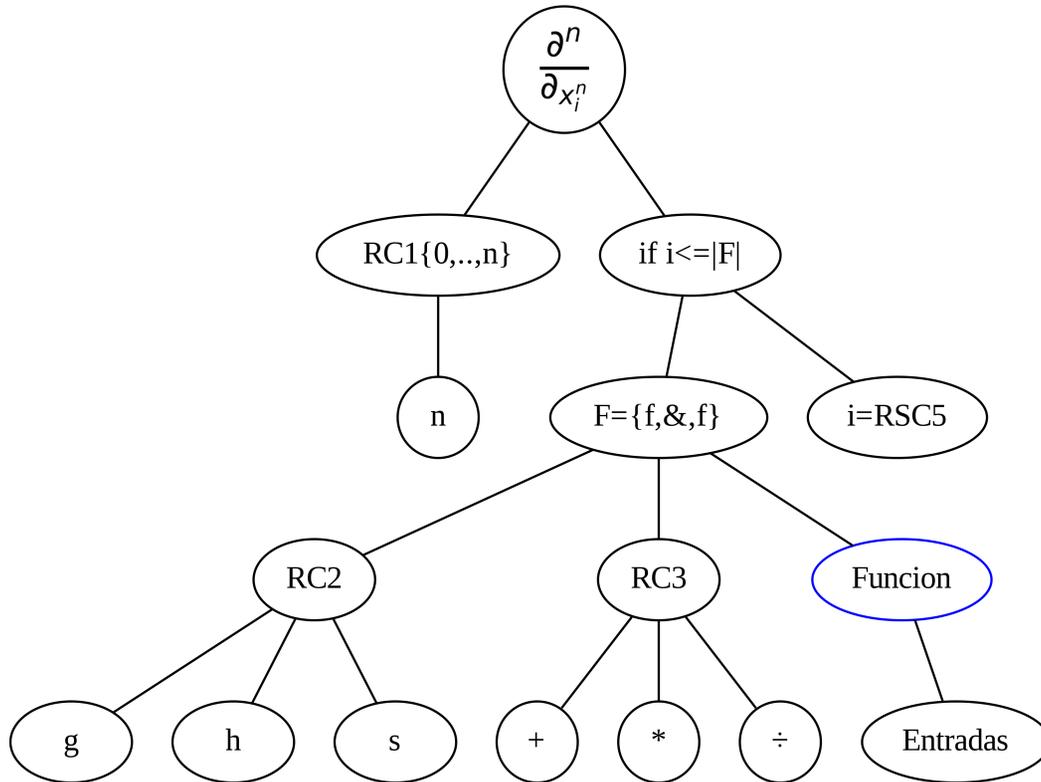


Figura 2.9: Generación de un nuevo algoritmo al sustituir uno de los nodos originales.

$F=\{f,\&,f\}$ corresponde a una n -ésima derivada parcial de una de las tres funciones ingresadas originalmente. En comparación, el resultado de la ejecución con las condiciones iniciales es:

$$\frac{\partial^4}{\partial z^4} [(x^2 + yz^5) \cdot [120(2x - 3)(yz)]] = (120 \cdot 360)(2x - 3)(y^2 z^2). \quad (2.1)$$

Ahora si se modifica la condición del nodo $\text{if } i \leq |F|$ en *Calculo de derivadas* se tienen los sig cambios:

- $\text{if } i \leq |F| \rightarrow \text{if } i = |F|$ El programa solo calcula la derivada respecto a la ultima variable indexada, es decir $\frac{\partial^n}{\partial x_n^n} F$.
- $\text{if } i \leq |F| \rightarrow \text{if } i > |F|$ El programa no calcula nada pues no existe la variable en la

función F^3 .

Por otra parte, existe la posibilidad de sustituir el nodo raíz de *Derivadas Parciales* por alguna otra instrucción final, por ejemplo:

1. Evaluación de F en $x_i = n \forall i$ i.e $F(n, \dots, n) \rightarrow (4^2 + 4 \cdot 4^5)(2 \cdot 4 - 3) = 20560$.
2. Evaluación de ∇F en $x_i = n \forall i$ i.e :

$$\nabla F(n, \dots, n) \rightarrow \left((6x(6x-1)2yz)^5 \hat{i} + (2x-3)z \hat{j} + (10x-15)yz^4 \hat{k} \right) \Big|_{(4,4,4)} = (589824, 5120, 25600).$$

3. Graficación de proyección de $F(x_i)$ en una dimensión para $n = x_i \neq x_2 \rightarrow$ Gráfica de $F(4, y, 4) = -12208 + 8192y$.

Las salidas del algoritmo son un escalar, vector e imagen respectivamente. Los cambios fueron producidos aún conservando las entradas originales, de éste modo **las instrucciones , operadores y subprogramas también pueden tomar el rol de variables y modificarse para cambiar la estructura del programa así como su funcionamiento.**

2.3 Programación genética

2.3.1 Introducción

En la sección (2.2) se discutieron las ventajas de la programación simbólica como una manera de modificar la estructura de un programa considerando las instrucciones y operaciones como variables. En el algoritmo de la sección (2.2.3) se presentó la posibilidad de modificar ciertas partes de su contenido para construir un programa con las mismas entradas pero diferentes salidas, sin embargo, surge la pregunta : ¿Como realizar dichos cambios en el programa sin la necesidad de intervenir directamente? Después de todo, lo que se busca es no tener que "redefinir" el código manualmente como se realiza en la programación funcional. En esta sección se muestra el paradigma de la Programación Genética el cual permite considerar un conjunto de programas y hacerlos evolucionar modificando su estructura hasta que se encuentra el que cumple mejor con el objetivo deseado. Dicho paradigma está basado en las ideas de la selección natural, evolución y genética.

³Una alternativa es que al no existir la variable, la derivada sea cero [45].

2.3.2 Evolución, selección natural y genética: Un recorrido histórico

La interrogante del origen del ser humano y del por qué sus características, fue motivo de gran discusión entre los filósofos de la antigüedad. Por ejemplo, Anaximandro (610-546 a.C) concebía el origen de de la vida terrestre (incluido el ser humano) en el mar. Por su parte, Empédocles (409-430 a.C) consideraba que los primeros seres vivos eran terrestres y que las especies que les prosiguieron surgieron por casualidad sin un fin o poder "supremo" [46]. Aristóteles (384-322 a.C) compartió parte de la teleología de la naturaleza en la que todas las cosas tienen un propósito de existir en un orden cósmico divino. Los seres vivos siendo parte de las cosas naturales debían ser *actualizaciones* de posibilidades naturales fijas [47]. Tomas de Aquino (1224-1274) recupera esta idea mencionando en la *Quinta Via* que los entes naturales siempre o a menudo (muchos sin consciencia de ello), obran intencionadamente con el fin de lo mejor, lo cual solo era posible con la existencia de un ser supremo [48]. De acuerdo con Ron Amundson en [49] existía un poder creador a elementos diferentes de la Naturaleza otorgado por Dios, de manera que fuera posible la creación de vida a lo largo de la historia de la Tierra.

A principios del siglo XVIII muchos embriólogos describieron en el crecimiento del feto un desenvolvimiento secuencial de las primeras formas de la especie, en lugar de la aparición espontanea de la forma adulta.

"Desde el unicelular cigoto, equivalente a un ser unicelular, los organismos van pasando en su desarrollo embrionario por todas las formas anteriores en la serie animal" [51].

De esta forma el concepto de **evolución** fue concebido como un sinónimo de *desarrollo*. Fue Jean Baptiste de Lamarck (1744-1829) el primero en proponer una teoría más formal de la transformación progresiva de las especies (a pesar de no usar explícitamente el término evolución). Dicha transformación se debía a dos *fuerzas* :

1. Los organismos poseen una tendencia natural al cambio.
2. Los organismos deben adaptarse al medio en el que viven.

La primera postulaba que los organismos comenzaron como estructuras muy simples y debido a su esencia natural fueron convirtiéndose en entes mas complejos. Las especies también se transformaron de las mas simples hasta llegar a la especie humana. No obstante, de acuerdo con la segunda fuerza, el entorno natural y volátil en el que viven los organismos los obliga a cambiar para adaptarse a las nuevas condiciones a través

del uso y desuso de sus órganos. En otras palabras, según Lamarck [54] *es el ambiente quien dicta las variaciones que deben existir en las especies para que puedan adaptarse*. Estas variaciones obtenidas serían transmitidas a través de las generaciones mediante la "herencia de los caracteres adquiridos". (Actualmente se conoce que, en general, es imposible que las células de un organismo (células somáticas) transmitan información hereditaria a los genes de manera que se traslade de una primera generación a posteriores. Esto es conocido como *barrera de Weissman*⁴ [52]).

Un par de décadas después el naturalista británico Charles Darwin (1809-1882) comienza una expedición de casi 5 años como naturalista de la tripulación del *Beagle* bajo el mando del capitán Robert Fitzroy [55]. Durante ésta, Darwin recopiló observaciones geográficas, geológicas, de flora y fauna entre otras, las cuales formaron la base para lo que sería el resto de sus investigaciones acerca de la evolución, diversidad y adaptación de las especies comenzando por su escrito *Diario de viaje del Beagle* de 1839 [56]. Fue durante los próximos 20 años donde Darwin comenzó a redactar en una gran y extensa argumentación las ideas, observaciones e incluso objeciones anticipadas a su teoría, las que le condujeron a publicar uno de los escritos más famosos e importantes en la historia de la humanidad. En 1859 se publica *Sobre el origen de las especies mediante la selección natural o la conservación de las razas favorecidas en la lucha por la vida* donde se formula su teoría que puede ser resumida de forma sencilla en la introducción:

Como de cada especie nacen muchos más individuos de los que pueden sobrevivir, y como, en consecuencia, hay una lucha por la vida, que se repite frecuentemente, se sigue que todo ser, si varía, por débilmente que sea, de algún modo provechoso para él bajo las complejas y a veces variables condiciones de la vida, tendrá mayor probabilidad de sobrevivir y, de ser así, será naturalmente seleccionado. Según el poderoso principio de la herencia, toda variedad seleccionada tenderá a propagar su nueva y modificada forma [57].

En contraste con las ideas de Lamarck, Darwin notó una diferencia fundamental en las adaptaciones de los individuos. Los cambios o variaciones que éstos sufren tienen una característica aleatoria, de manera que las variaciones pueden hacerlos más fuertes o débiles para sobrevivir a su ambiente. Una vez que el individuo ha cambiado es el entorno quien lo rechaza o acepta, **no todo cambio es necesariamente adaptativo**. De esta manera Darwin retiró la direccionalidad en el proceso evolutivo. Los organismos (o poblaciones como un conjunto) no cambian sus características *para* poder adaptarse y sobrevivir, sino que aquellos organismos que han sufrido cambios de carácter aleatorio

⁴Friedrich Leopold August Weismann (1834-1914) biólogo alemán. Propuso en 1891 con un experimento de mutilaciones en ratas blancas refutar la idea de los caracteres adquiridos de Lamarck. [53]

(mutaciones) que les permitan adaptarse mejor a su entorno, desplazan a aquellos que no poseen dichas modificaciones haciendo prevalecer estos cambios (y a la especie) a lo largo del tiempo. Esto es **la selección natural, es el mecanismo principal por el cual se produce la evolución.**

Para ejemplificar la disrupción teórica que suponen los postulados de Darwin frente a la antigua teoría de Lamarck considérese la Fig[2.10]. En un inicio, las Jirafas que vivían en el interior de África, donde hay regiones áridas y sin praderas, habrían tenido la dificultad de conseguir su alimento de los árboles debido a su altura. Según Lamarck, la población de Jirafas de una región con dicha dificultad tuvo que modificar sus patas y cuellos, a través del paso de generaciones, para hacerlos mas largos de manera que pudieran alcanzar su alimento. El cambio habría sido gradual y completo, sobre toda la población de las Jirafas, es decir todos los individuos fueron mejorando paulatinamente sus características con un determinado fin (alcanzar el alimento). Lo que en realidad sucedió de acuerdo con Darwin es que en cada generación surgieron aleatoriamente dentro de la población, individuos con características diferentes (tamaño de patas y cuello distinto, color y forma de pelaje diferente etc). Hubo jirafas que gracias a estos cambios pudieron adaptarse mejor a su entorno (sin olvidar que otras igualmente con cambios no se vieron beneficiadas), las cuales mediante la reproducción cambiaron las características de sus descendientes en las siguientes generaciones. De esta forma, aquellas jirafas incapaces de adaptarse a las circunstancias fueron pereciendo y siendo "desplazadas" por aquellas con características favorables permitiéndoles prevalecer.

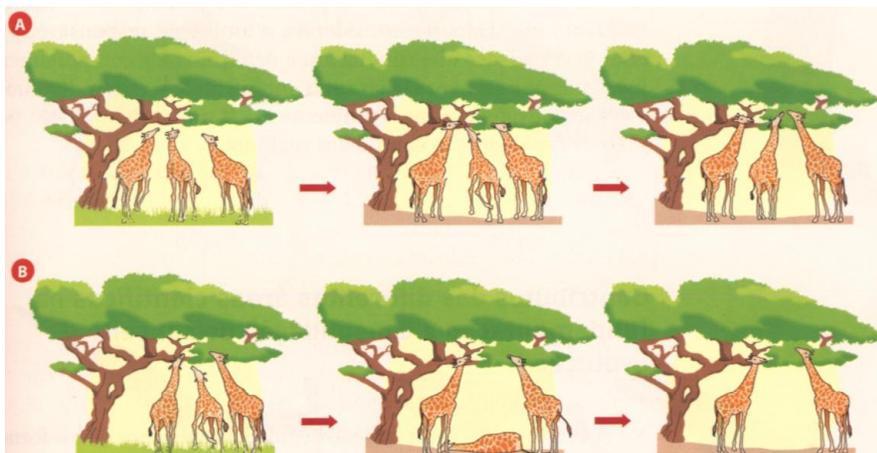


Figura 2.10: Evolución de las jirafas según Lamarck (arriba). Evolución según Darwin (abajo) [59].

¿Cómo se produjeron estos cambios en los individuos? y ¿Cómo es que se transmitieron a través de las generaciones?. La comprensión moderna de la mutación, herencia, genes

e información genética tiene sus raíces en el trabajo pionero de Gregor Mendel a mediados del siglo XIX. Debido a las características fácilmente observables y la capacidad para autofertilizarse, Mendel eligió plantas de guisante para realizar cruzamientos controlados entre plantas con diferentes características (altura, color de la flor etc) y registró cómo es que se transmitía estos rasgos a las generaciones siguientes. Los resultados lo llevaron a descubrir las leyes básicas de la herencia, conocidas ahora como las Leyes de Mendel. Estas leyes sentaron las bases para la genética, demostrando que los rasgos se transmiten de padres a hijos a través de unidades discretas conocidas como "factores hereditarios", hoy llamados genes. Las leyes de Mendel son:

1. Ley de la Segregación: Cada individuo tiene dos alelos para cada gen, uno heredado de cada progenitor. Durante la formación de gametos, los alelos se separan para que cada gameto contenga solo uno de los alelos.
2. Ley de la Distribución Independiente: Los alelos de diferentes genes se segregan independientemente unos de otros durante la formación de gametos, lo que produce una variedad de combinaciones posibles en los descendientes.

Con el avance de la biología, se descubrió que los genes están compuestos por ADN, una molécula que almacena y transmite información genética. Esta información está codificada en la secuencia de nucleótidos del ADN. Los genes contienen las instrucciones para la síntesis de proteínas, cuya importancia es fundamental para el desarrollo y funcionamiento de los organismos. Los genes están localizados en estructuras llamadas cromosomas, que se encuentran en el núcleo de las células.

El genotipo es la composición genética de un individuo, es decir, la combinación y codificación que posee dentro de sus cromosomas. Este genotipo determina el fenotipo, que son las características observables o externas del individuo, como el color de ojos, la altura y otras características físicas que determinan en gran medida su nivel de adaptación al entorno donde se desarrolla. Sin embargo, el genotipo es el único determinante en la transmisión de las características de un individuo a la siguiente generación, de modo que las variaciones adquiridas fenotípicamente no pueden heredarse [58]. Esto explica el porqué la hipótesis de uso y desuso de órganos de Lamarck para describir la adaptación de los individuos es errónea.

Las mutaciones son cambios en la secuencia del ADN de un organismo. Estas pueden ocurrir de manera espontánea o ser inducidas por factores externos como la radiación o sustancias químicas. La teoría evolutiva moderna, establece que los cambios heredables son el resultado de mutaciones en el ADN. Estas mutaciones se producen al azar y

pueden dar ventajas o desventajas a los individuos. La selección natural actúa sobre esta variabilidad genética, favoreciendo a aquellos individuos cuyas mutaciones les otorgan una ventaja adaptativa en su entorno. Esta comprensión permite apreciar cómo las poblaciones pueden adaptarse y evolucionar a lo largo del tiempo, impulsadas por cambios genéticos aleatorios y las presiones selectivas del entorno [60].

En resumen, las mutaciones que producen variaciones hereditarias son fundamentalmente aleatorias pues estas variaciones pueden ser tanto perjudiciales como benéficas para los individuos. Sin embargo, la selección natural elimina aquellos individuos (y por ende cambios genéticos) incapaces de sobrevivir en el entorno permitiendo conservar lo *útil*. Como expresa Ayala en [61], sin mutación no existiría evolución pues no se producirían los cambios que se transmitirían de generación en generación. Pero sin selección natural el proceso de mutación resultaría en desorganización y eventual extinción pues por su naturaleza aleatoria las mutaciones son en su mayoría desventajas.

2.3.3 Programación genética como paradigma del computo evolutivo

La programación genética (**GP** por sus siglas en inglés) es un algoritmo de evolución inspirado en la evolución biológica por selección natural de Darwin y la mutación de estructuras por la genética de Mendel, cuyo objetivo es encontrar estocásticamente programas de computadora⁵ dentro de una población para que resuelvan un propósito en específico definido por el usuario. Uno de los principales impulsores de éste paradigma fue John R. Koza [62] quien a principios de 1990 mostró, gracias al gran incremento en la potencia de los CPU de la época, la capacidad de éste para funcionar como una máquina de invención automatizada, produciendo resultados novedosos y sobresalientes para diversos tipos de problemas. En palabras de Langdon :

La programación genética (GP) es una técnica de computación evolutiva (EC) que resuelve problemas automáticamente sin necesidad de decirle explícitamente al ordenador cómo hacerlo. En el nivel más abstracto, la GP es un método sistemático e independiente del dominio para que los ordenadores solucionen problemas automáticamente a partir de una declaración de alto nivel de lo que se necesita hacer [63].

Para contrastar y utilizar los conceptos vistos en la sección anterior, se presenta la esencia de los términos en este paradigma:

⁵Entendiéndose programas de computadora como entidades que reciben entradas, realizan cálculos y producen salidas con la cualidad de poder ser modificados estructuralmente.

- La **base** del algoritmo esta conformada por todos los operadores, variables, números, funciones e instrucciones lógicas o categóricas posibles para estructurar los programas que se deseen.

$$\text{Base} = \{+, *, -, \sqrt{}, \dots, f(x, y), g(z), \dots, \text{if}, \text{while do}, \text{or}, \dots, x, y, \mathbb{N}\}. \quad (2.2)$$

- Un **individuo** (también llamado candidato o solución) es un programa formado por elementos de la base.
- Una **población** es el conjunto de un número M de individuos pertenecientes a la misma generación.
- Una **generación** es el conjunto de individuos en una etapa particular t del algoritmo.
- Se entiende por **selección** al proceso por el cual los individuos se evalúan por un criterio definido llamado **fitness**.

Una descripción del algoritmo de la programación genética presentado en [1] es la siguiente:

Primero se genera una población de M individuos formados aleatoriamente por combinaciones no lineales de elementos de la base (Inicialización). Posteriormente cada uno de los individuos es puesto a prueba con el criterio elegido (fitness). Este proceso es conocido como selección y emula las condiciones a las que se aspira que los programas se adapten. Una vez que se obtiene el individuo con la mejor adaptación, se utilizan operadores de cruzamiento y mutación para cambiar su estructura aleatoriamente, así como de otros miembros de la población para generar un nuevo conjunto de individuos basados en estos cambios (descendencia). La nueva generación sera puesta a prueba nuevamente por el criterio de selección y entonces se realizará el mismo procedimiento un número N de generaciones mas o hasta que se obtenga un valor aceptado en el proceso de selección dando así final al algoritmo. El individuo seleccionado de la última generación se considera el "mejor adaptado" así como el programa que resuelve mejor la tarea asignada.

A pesar de que se suele hablar de la **GP** como un proceso de evolución de programas, no es común que se use para desarrollar programas en los lenguajes completos de Turing⁶ que los desarrolladores de software utilizan habitualmente. Más bien, es común

⁶Un lenguaje completo de Turing es aquel que puede emular un algoritmo implementado en una máquina de Turing, es decir, puede realizar cualquier cálculo que una computadora sea capaz de ejecutar, siempre que disponga de suficiente tiempo y memoria [72]. Ejemplos de éstos lenguajes son Java, C++ y Excel Macros.

Algoritmo 1 Programación genética

```
1:  $\epsilon \leftarrow$  criterio de suficiencia
2:  $M(t = 0)$  Población inicial de tamaño M ▷ Inicialización
3:  $F \leftarrow fitness(m_i)$  Evaluación de individuos ▷ Evaluación y selección
4:  $S \leftarrow$  mejor individuo de  $M(0)$ 
5:  $m'_i \leftarrow Mutacion(S)$  or  $Mutacion(m_i)$  ▷ Descendencia de S
6: while  $F > \epsilon$  or  $t < N$  do
7:    $t = t + 1$ 
8:    $M(t) \leftarrow \{m'_i\}$  Población de la Generación t
9:    $F \leftarrow fitness(m'_i)$  Evaluación de individuos ▷ Evaluación y selección
10:   $S \leftarrow$  mejor individuo de  $M(t)$ 
11:   $m_i \leftarrow Mutacion(S)$  or  $Mutacion(m'_i)$  ▷ Descendencia de S
12: end while
13: return  $S$ 
```

que se evolucionen expresiones simbólicas o fórmulas dentro de un lenguaje más limitado y frecuentemente especializado en un área concreta. Una de las aplicaciones más utilizadas y de las pioneras en exhibir el potencial de la **GP** es en la implementación de regresores simbólicos (siguiente sección). Otras aplicaciones de la **GP** han resultado en clasificadores y transformadores simbólicos [64, 65], existen también aplicaciones en robótica, medicina, genómica, procesamiento de imágenes, herramientas de detección de plagio [66, 67, 68, 69, 70] entre otras [71].

2.3.4 Regresor simbólico

El principal objetivo de una regresión es en general determinar la relación de una variable dependiente (objetivo) t con un conjunto de variables independientes (variables de entrada) especificadas $\{\theta_i\}$. De esta manera se busca obtener una función $f(\theta_i, w_i)$ que utiliza θ y los coeficientes α_i tales que :

$$t = f(\theta_i, \alpha_i) + \epsilon, \quad (2.3)$$

donde ϵ representa el error.

La forma de f es pre-definida en las técnicas de regresión convencionales como la regresión lineal, polinómica o en redes neuronales artificiales (ANN por sus siglas en inglés).

1. Lineal : $f = \alpha_0 + \alpha_1 x + \epsilon$.
2. Polinómico : $f = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \dots + \alpha_n x^n + \epsilon$.
3. ANN: $f_j = \sigma(\sum_{i=1}^n \alpha_{ji} x_i + \epsilon_j)$ donde g es un función auxiliar de activación.

Cuando se tienen mediciones u observaciones, éste tipo de regresiones busca obtener los parámetros α_i de manera que mejor aproximen al modelo generalmente minimizando el error ϵ :

$$\epsilon = Obs_i - f(\theta, \alpha, t_i), \quad (2.4)$$

donde Obs_i representa la i-ésima medición.

En contraste, un **regresor simbólico** busca encontrar la relación entre las variables y mediciones no sólo como un conjunto de coeficientes en un modelo *a priori* sino como la descripción simbólica de éste [73]. En lugar de asumir la forma del modelo f , se utiliza un conjunto de funciones simples $f_i \in T$ y operadores $\times_i \in N$ para construir la forma general de la función. En este caso T y N son subconjuntos de la base B para desarrollar el regresor simbólico.

Desde un punto de vista más abstracto, las regresiones convencionales *buscan* en el espacio de *parámetros* la combinación $(\alpha_0, \alpha_1 \dots, \alpha_i)$ que mejor adaptan el modelo asumido para describir las mediciones, mientras que el **SR** busca en el espacio de *parámetros* \cup *funciones* \cup *operadores* la mejor combinación $(\alpha_0, \alpha_1 \dots, \alpha_i, f_0, f_1 \dots, f_i, \times_0, \times_1 \dots, \times_i)$. El error a minimizar es precisamente lo que se conoce como función *fitness*⁷ y dependerá del caso que se presente así como de las mediciones u observaciones con las que se cuente. Entre las mas utilizadas en la literatura son:

1. Error absoluto medio :

$$MAE = \frac{1}{n} \sum_{i=1}^n |Obs_{x_i} - f(x_i)|. \quad (2.5)$$

2. Error cuadrático medio:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Obs_{x_i} - f(x_i))^2. \quad (2.6)$$

3. Xi cuadrada :

$$\chi^2 = \sum_{i=1}^n \left(\frac{Obs_{x_i} - f(x_i)}{\epsilon_i} \right)^2. \quad (2.7)$$

En donde Obs_{x_i} representa la observación asociada a la i-esima x , $f(x_i)$ la evaluación del modelo construido en x_i , n el número de observaciones o datos y ϵ_i el error asociado a la i-ésima observación.

El MAE mide el promedio de las diferencias absolutas entre las predicciones de un modelo y los valores observados, es útil cuando los errores deben ser interpretados en

⁷Aunque en general se busca minimizar estas funciones, pueden definirse de otra forma dependiendo el interes de manera que se busque un máximo.

la misma escala que la variable original. El MSE es la media de los cuadrados de las diferencias entre las predicciones y los valores observados, a diferencia del error absoluto medio, el error cuadrático es sensible a los valores atípicos en los datos por lo que es adecuado cuando se intenta dar una medida que refleje la dispersión general de los errores [74]. Por otra parte la χ^2 se utiliza para evaluar qué tan bien las predicciones del modelo se ajustan a las observaciones, teniendo en cuenta las incertidumbres asociadas a cada observación o medición.

Mientras que las ec(2.5,2.6) son utilizadas para regresiones en datos con cierta dispersión pero sin incertidumbres o *ruido* asociado, la ec(2.7) es utilizada en éste último caso.

Capítulo 3

Regresor simbólico programado genéticamente.

3.1 Introducción

En esta sección se describen los conceptos y etapas mas importantes de una técnica que aprovecha el dinamismo estructural bajo el paradigma de la programación simbólica y el uso de la programación genética para optimizar la búsqueda de programas a través de la evolución de éstos: Un **regresor simbólico**. Aunque dichos conceptos son aplicables a varias variables en general, se mostrara la implementación de un regresor simbólico de una variable cuyos cálculos se realizan mediante la notación de Lukasiewicz y cuya representación para interpretarlos son árboles de aridad M.

3.2 Pre-inicialización.

Antes de dar inicio al algoritmo es necesario considerar ciertos aspectos que influyen en la construcción y selección de las soluciones así como en el término del algoritmo. Como se menciona en la sec(2.3.4) la base de donde se obtienen los elementos para construir los individuos esta compuesta por subconjuntos de operadores, funciones y parámetros. Ya que computacionalmente las operaciones pueden considerarse como funciones ¹, se puede escribir la base como la unión de dos subconjuntos:

$$B = \{T \cup V\} = \{sen, cos, |, \sqrt{\quad}, +, -, \div, * \dots, x, y, z, \mathbb{R}\} \quad (3.1)$$

¹Esto debido a que al recibir variables de entrada devuelven un resultado, por ejemplo la función seno devolverá el seno de la variable introducida, de igual manera la "función" suma, devolverá la suma de las variables ingresadas.

Donde T corresponde al conjunto de funciones tanto matemáticas como algebraicas (por ejemplo el seno, coseno, raíz cuadrada, suma, división etc.) y el conjunto V corresponde a las variables (parámetros) del dominio del problema (variables simbólicas como x, y etc, y variables numéricas como \mathbb{N} o \mathbb{R}). Es importante considerar las características de las funciones de manera que puedan ser evaluadas correctamente:

1. **Aridad:** Las funciones al igual que en matemáticas deben contener el dominio apropiado (tanto en tipo de variable como en cantidad) para ser evaluadas aún cuando exista un composición de éstas. Algunos ejemplos de funciones y su aridad son :

- Suma (+) = 2
- División (\div) = 2
- Seno ($\text{sen}(x)$)= 1
- Raíz cuadrada (\sqrt{x}) = 1
- Random² ($\text{Rand}(0,1)$)=0

2. **Cerradura :** Algunas funciones deben ser modificadas de manera que puedan recibir cualquier argumento y aun así devolver un número o expresión válido de manera que el árbol completo pueda ser evaluado en todos los nodos. Ejemplos de estos son:

- *Raíz cuadrada:* Esta función calcula la raíz cuadrada del valor absoluto de la entrada.
- *División:* Si el denominador es cercano a cero (se encuentra por ejemplo entre -0.0001 y 0.0001 ³) la operación regresa un 1.
- *Log natural:* Esta función calcula el logaritmo natural del valor absoluto de la entrada y cuando ésta sea menor a 0.0001 regresa 0.

El contenido de la base influirá tanto en la forma de las soluciones como en el tiempo de cómputo. Por un lado se debe intentar que la base sea **suficiente** para formar las soluciones adecuadas al problema . Como ejemplo, una base $\{+, -, *, x\}$ solo puede formar polinomios más no funciones racionales o con asíntotas (Véase Fig[4.1]). Por el contrario, una base con demasiadas operaciones puede hacer mas compleja la estructura de las expresiones así como requerir un tiempo mayor de cómputo al evaluarlas aunque esto podría ayudar a expandir el espacio de búsqueda.

²Esta función puede devolver aleatoriamente un numero flotante entre 0 y 1 sin la necesidad de ingresar una variable.

³Estos rangos deberían de ser cambiados dependiendo la naturaleza del caso.

Dependiendo el objetivo con el cual se quiere emplear el **SR**, la función fitness debe ser elegida considerando además el tipo de datos que se posee. Ya que el proceso de evaluación del fitness es el que requiere mas tiempo de cómputo pero además dicho calculo es independiente del resto de la población, técnicas como la paralelización y distribución de programas han sido implementados con el fin de incrementar la eficiencia de la **GP** en casos donde se cuente con una arquitectura paralela de computadora o con múltiples computadoras con menos y más potentes elementos de procesamiento [75, 76, 77].

Por último se debe considerar el número máximo N de generaciones que se generaran así como el número M de individuos que conformaran la población. Además, un criterio de finalización puede, de ser alcanzado, terminar con el algoritmo antes de llegar al número N de generaciones. Dicho criterio debe ser mayor o menor al valor del fitness del mejor individuo de la última generación según sea el caso.

Cabe recalcar que aunque la representación utilizada en la mayor parte de éste trabajo es la de arboles de aridad- M debido a su fácil interpretación, en la práctica suele utilizarse la notación polaca o de LISP⁴ al momento de realizar los cálculos correspondientes debido a la reducción de tiempo y espacio de memoria.

3.3 Inicialización.

En este proceso se genera la población de la generación inicial tomando aleatoriamente elementos de la base para formar los individuos. En este paso existen 3 métodos simples propuestos por Koza [62] para realizar dicha actividad tomando en cuenta la profundidad⁵ máxima '*prof max*' establecida en los árboles.

1. **Full method:** En éste método primero se seleccionan aleatoriamente funciones de la base en cada nivel hasta que se alcance una profundidad (*prof max*) $- 1$ de manera que posteriormente se puedan ingresar las variables (hojas). En la Fig[3.1] se muestran los pasos para la generación de un individuo con profundidad máxima *prof max* = 3. En $t=2$ se selecciona la función ' \div ' como nodo hijo de la función '+' del nivel 0. Esto obliga a que los nodos del nivel 2 sean variables pues de ser funciones se necesitaría un nivel más para evaluarlos excediendo la

⁴LISP es una familia de lenguajes de programación de tipo multiparadigma con sintaxis homocónica basada en la notación polaca [78].

⁵En algunos textos se considera que las variables al formar parte de las hojas (nodos inferiores) no se pueden clasificar como nodos, haciendo que la profundidad del árbol decrezca en una unidad, de acuerdo con esto los árboles finales de las figuras [3.1,3.2] tendrían profundidad 2, sin embargo por consistencia se mantendrá la definición de profundidad mostrada en la sección (2.2.1) independientemente del tipo de nodo.

profundidad máxima pre-establecida. Los descendientes del lado derecho desde $t=5$ son agregados análogamente.

2. **Grow method:** En éste método después de seleccionar el nodo raíz, los hijos pueden pertenecer a las funciones o a las variables y ser seleccionados hasta que se alcance el máximo de profundidad o se *trunque* una rama al ser seleccionada una variable. En la Fig[3.2] se muestran los pasos para la generación de un individuo con profundidad máxima $prof\ max = 3$. En $t=2$ se selecciona la variable x como nodo hijo de la función $'\div'$ del nivel 0 haciendo imposible que tenga mas descendencia en el lado izquierdo. En $t=3$ se selecciona la función $'+'$ (aunque podría haber sido una variable) lo que obliga a que en el siguiente nivel sean variables debido a la profundidad máxima alcanzada.
3. **Ramp half-and-half:** Una población generada por este método contiene la mitad de individuos construidos mediante el método *Full* y la otra mitad mediante *Grow*. La profundidad máxima se establece en un rango $[a, b]$; $a, b \in \mathbb{N}$ de manera que en la población exista mas variedad de tamaños y formas.

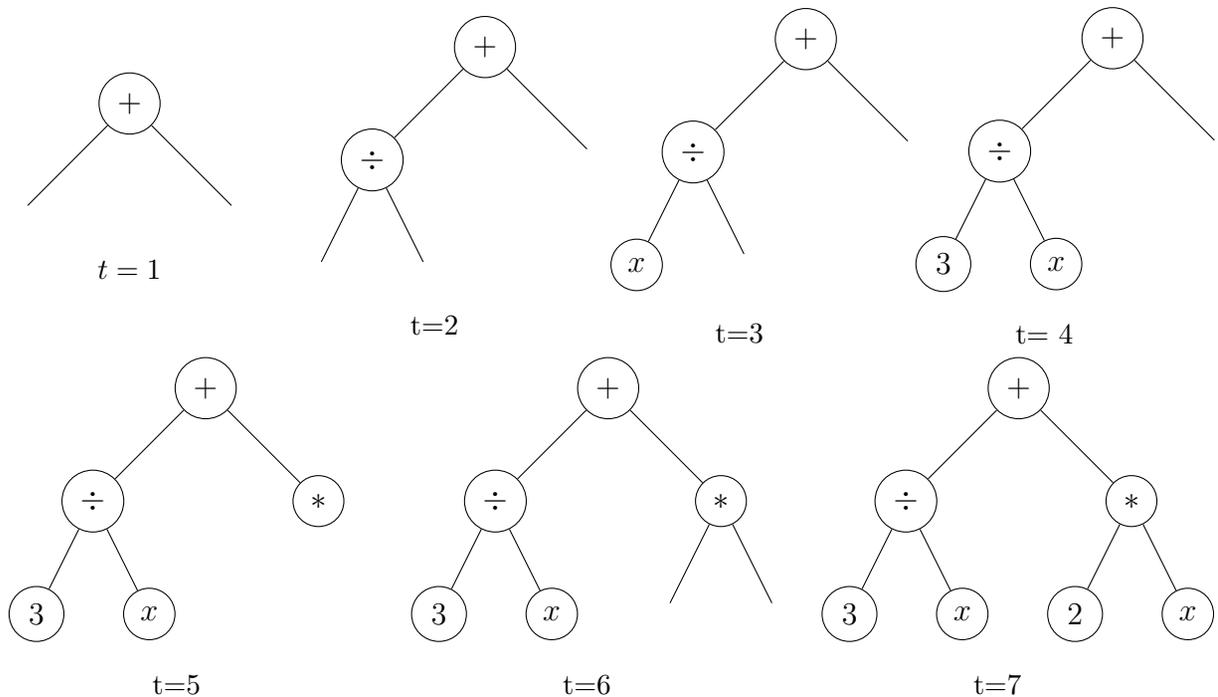


Figura 3.1: Creación de un árbol completo de $prof\ max = 3$ mediante el método *Full* para formar al individuo $\frac{3}{x} + 2x$.

El pseudo-código del algoritmo para la construcción de individuos utilizando los metodos *Full* y *Grow* se encuentra en Algoritmo(2). Una discusión sobre ésta etapa de *GP* así como la exposición de variantes de los tres métodos expuestos de construcción puede encontrarse en [79].

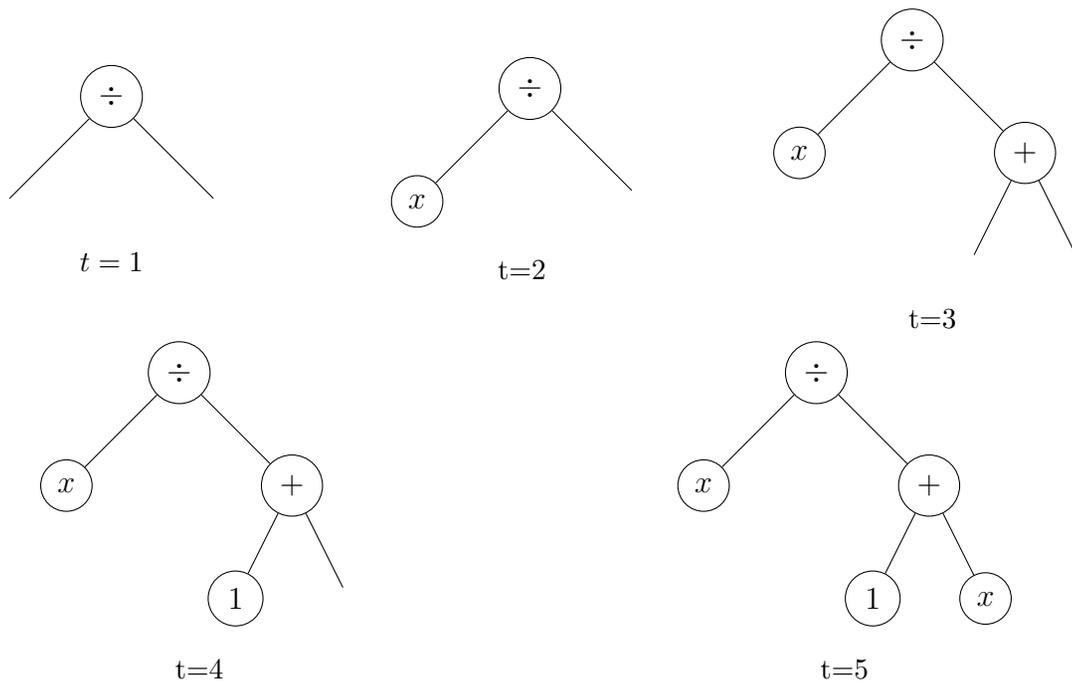


Figura 3.2: Creación de un árbol asimétrico de $prof\ max = 3$ mediante el método *Grow* para formar al individuo $\frac{x}{1+x}$.

3.4 Selección.

En esta etapa se evalúa cada individuo con la función fitness asociándoles así una probabilidad de reproducción proporcional al fitness γ (para las funciones (2.5,2.6,2.7), la probabilidad es inversamente proporcional. Un fitness bajo corresponde a una probabilidad de reproducción alta). Además, se identifica al individuo mejor adaptado en la presente generación (el fitness γ mínimo). Para evitar que durante la evaluación los nodos con operadores se ejecuten antes de conocer el valor de su dominio, es usual que se *recorra* el árbol de manera recursiva comenzando del nodo raíz y posponiendo la evaluación de los nodos hijos hasta que se conozca su valor. En el Algoritmo 3 se muestra el pseudocódigo de éste tipo de técnica para evaluar los árboles .

El siguiente paso es seleccionar un subconjunto de la población de la cual se obtendrán los individuos para realizar un intercambio de genes al aplicar operadores de cruce y así obtener la descendencia que ocupara parcial o totalmente la siguiente generación. Existen dos principales métodos para seleccionar éstos individuos [80]:

1. **K-Torneo de selección:** Ésta es la técnica más utilizada en la literatura y consiste en seleccionar aleatoriamente k individuos de la población y elegir aquel que tenga el mejor fitness. Éste proceso se repite un determinado numero de veces suficiente para obtener la cantidad deseada de padres que formaran la descendencia.

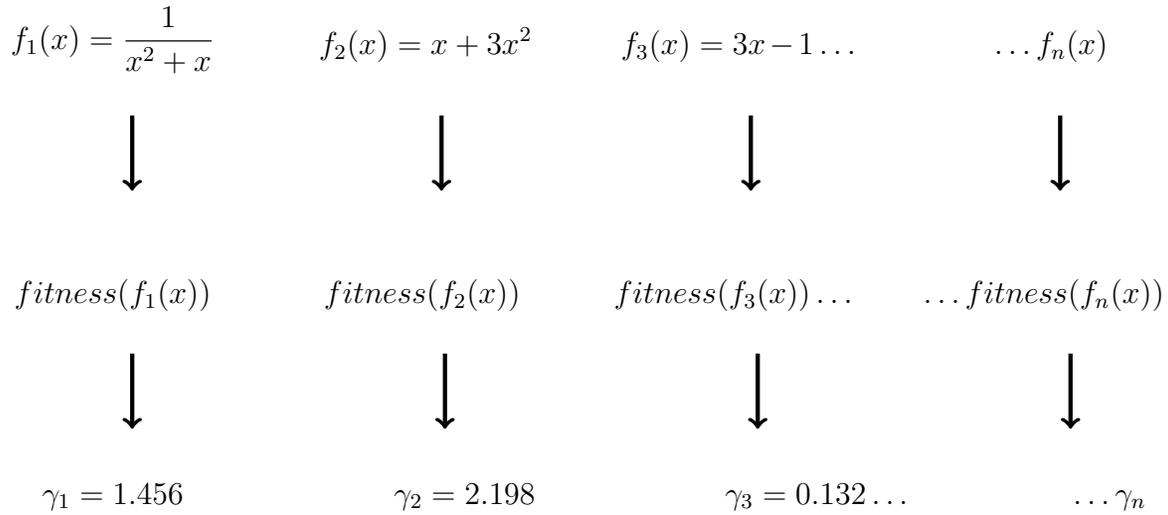


Figura 3.3: Proceso de evaluación de las soluciones. Población de programas (arriba). Evaluación de la población mediante la función fitness (medio). Fitness o aptitud asociada (abajo).

cia. A un mayor valor de k generalmente se le asocia una búsqueda mas rápida de individuos mas adaptados haciendo que el proceso de evolución converja a una solución en menor tiempo (corriendo el riesgo de caer en un mínimo local, proceso conocido como *presión selectiva alta*). Mientras que un valor pequeño de k mantendrá mayor diversidad (pues individuos con menor adaptación podrán heredar sus genes) haciendo que el tiempo de cómputo sea más largo (*presión selectiva alta*).

2. **Selección por ruleta** (Proporcional): En éste método se asocia una probabilidad p a cada i -ésimo individuo dada por⁶ $p_i = \frac{\sum_{j=1}^N \gamma_j}{\gamma_i}$ de manera que el tamaño del segmento correspondiente en la ruleta es proporcional a su probabilidad de selección. Posteriormente se genera un número aleatorio entre 0 y 1 determinando el segmento de ruleta y así el individuo elegido. A pesar de su sencillez presenta dos problemas principales: Cuando las diferencias en los fitness son muy pequeñas puede dar lugar a una menor diversidad genética haciendo caer al algoritmo en un mínimo local. Por otro lado, se pueden crear grupos de individuos con altas aptitudes haciendo que predominen mas llevando también a un estancamiento prematuro.

Una vez que se tiene al grupo de padres que heredaran sus genes para individuos en la siguiente generación, éstos deben pasar por un proceso de mutación y/o cruzamiento. Éste proceso es conocido como evolución.

⁶Recordando que de las ecuaciones de los fitness un mínimo es mejor.

3.5 Evolución.

El grupo de descendientes de la siguiente generación esta conformado por individuos que heredan el material genético de sus padres ya sea por una cruza o por una mutación. Para poder incluir este fenómeno en el algoritmo se deben aplicar aleatoriamente los llamados *Operadores de cruzamiento y mutación*, los cuales pueden variar dependiendo los parámetros del algoritmo [83]. Se presentan a continuación ejemplos de operadores y sus efectos:

- Cruzamiento (Crossover): Este operador necesita de dos generaciones existentes de las cuales se seleccionará al mejor individuo de cada una. El primero de los individuos fungirá como receptor de sub-arboles mientras que el otro será el donante de ellos. De esta manera el material genético se mezcla entre los individuos de la población transmitiendo así la *información* de una posible mejor estructura de adaptación. El cruzamiento es la operador con mayor influencia en la creación de descendientes y posee mayor efectividad en poblaciones grandes, esta es la razón por la que se le asigna en general un mayor porcentaje de probabilidad de ocurrir entre los padres [81].

En la Fig[3.4] de los padres $3x + \frac{x}{2}$ y $10 - x^2$ se retiran las estructuras en blanco para unir las restantes y dar nacimiento al individuo $x^2 + \frac{x}{2}$.

- Mutación de sub-árbol (Subtree mutation): Se selecciona un sub-árbol o rama del individuo que será reemplazada por una generada de la base, éste operador permite mantener la diversidad en la población al introducir posibles características que se hayan perdido entre generaciones.

En la Fig[3.5] al padre $3x + \frac{x}{2}$ se le retira la estructura en blanco para añadir el árbol $5 - x$ proveniente de la base y así dar lugar al individuo $5 - x + \frac{x}{2}$.

- Mutación de nodos o puntual (point mutation): Se seleccionan diferentes nodos al azar del individuo que serán reemplazadas por otras obtenidas de la base. Es posible aumentar el porcentaje de este tipo de mutación cuando no se requiere obtener cambios agresivos (ramas completas) en las soluciones.

En la Fig[3.6] los nodos en blanco del individuo $3x + \frac{x}{2}$ son reemplazados por nodos de la base formando al individuo $3 - x + \frac{x}{7}$.

- Mutación de elevación (Hoist mutation): Se selecciona una rama del individuo que será eliminada, y de ser el caso, los nodos exteriores que no sean eliminados se ‘elevan’ o adhieren al nodo del cual se desprende el sub-árbol. Una de sus funciones principales es evitar el exceso de ramas o ‘follaje’ del árbol [82].

En la Fig[3.7] la estructura en blanco del individuo $3x + \frac{x}{2}$ es removida. El nodo

con valor 2 es *elevado* para ser la variable que se evalúe en el operador superior + dando lugar así al individuo $3x + 2$.

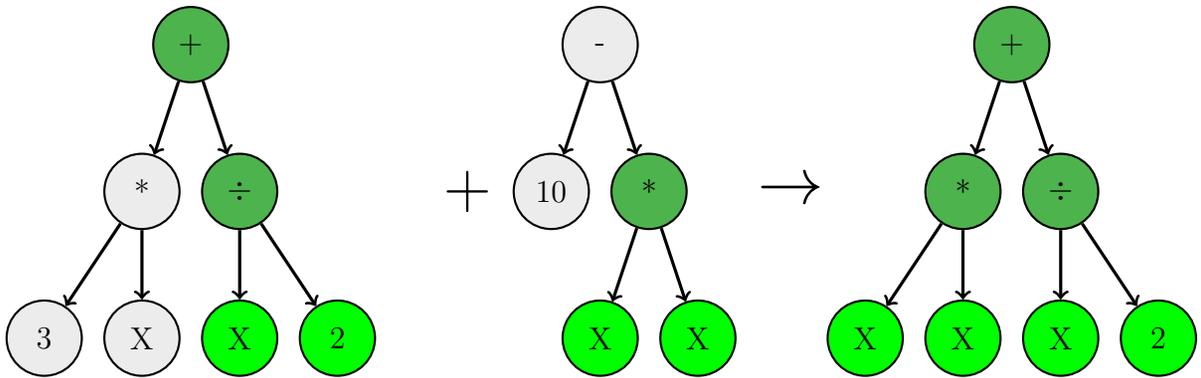


Figura 3.4: Cruzamiento

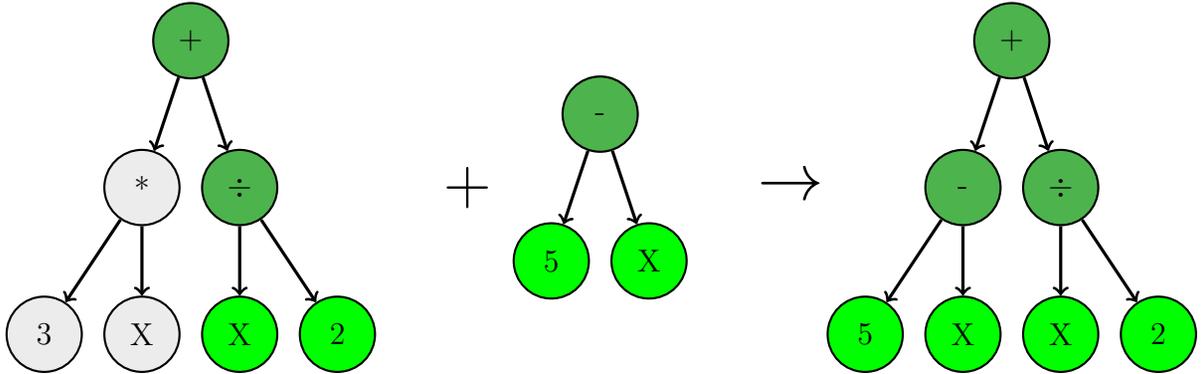


Figura 3.5: Mutación de sub-árbol

En la práctica a los operadores se les debe de asignar una probabilidad de ser aplicados al grupo de padres, generalmente es la cruce la que adquiere un mayor porcentaje mientras que las mutaciones son menos frecuentes.

3.6 Descendencia.

Una vez que se ha generado un grupo de descendientes modificados (el cual debe ser menor o igual al tamaño original M de la población anterior) el siguiente paso es generar la población de la próxima generación reemplazando total o parcialmente a los individuos anteriores. Se presentan algunos de los esquemas de reemplazo encontrados en la literatura [84]:

- **Reemplazo en estado estacionario:** En cada cambio de generación se producen nuevos individuos mientras que se retiran lo que tienen un mayor tiempo

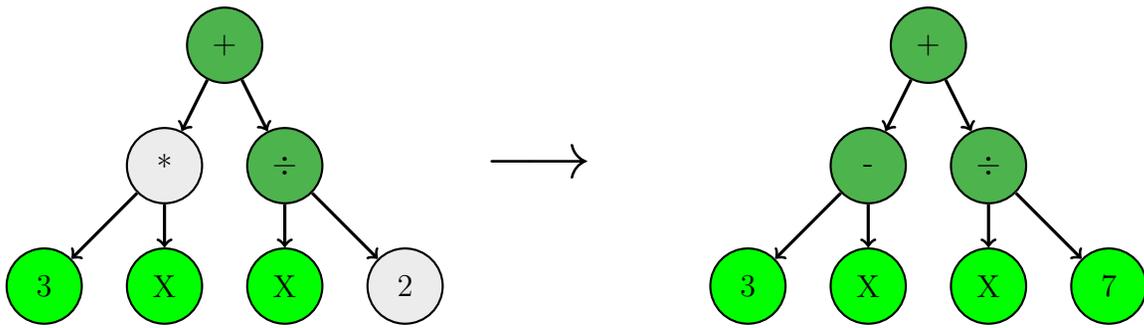


Figura 3.6: Mutacion de punto

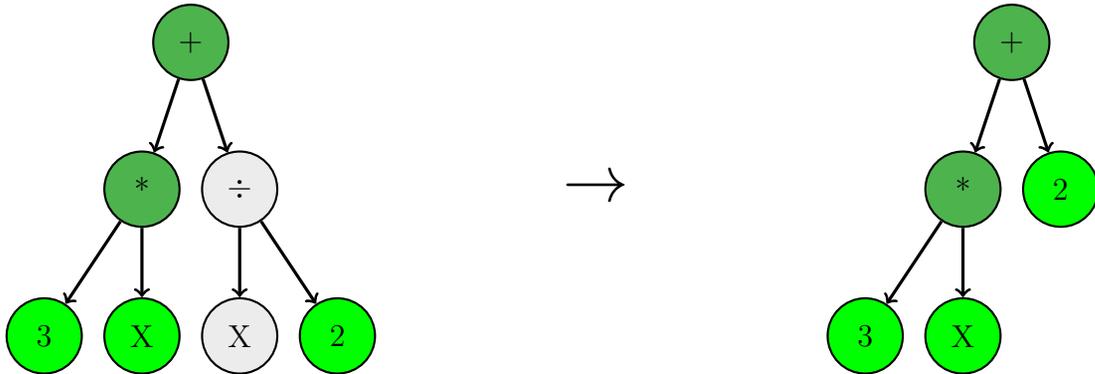


Figura 3.7: Mutacion de elevacion

de existencia (individuos viejos), de ésta forma existe una exploración gradual del espacio de soluciones. Un símil es el nacimiento y muerte continua de personas en una población.

- **Reemplazo general** : En este esquema, la población original entera es reemplazada por los nuevos individuos.
- **Selección de programas reemplazados**: Los individuos a reemplazar pueden ser escogidos aleatoriamente, por edad o peor adaptación.

Una vez que se genera la nueva población, se repite el ciclo hasta terminar el algoritmo entrando en la etapa de finalización.

3.7 Finalización.

En un regresor existen dos tipos de finalización del algoritmo:

- Se alcanza el número máximo N de generaciones.
- Se cumple un criterio de finalización. Por lo general , se espera que el fitness alcance un valor menor (o mayor) a uno establecido en la etapa de pre-inicialización.

El individuo con la mejor aptitud se considera la solución final del problema y puede representarse tanto en notación de prefijo como de árbol de aridad M. Por tanto, usualmente se requiere un proceso de simplificación para obtener una expresión mas sencilla de interpretar.

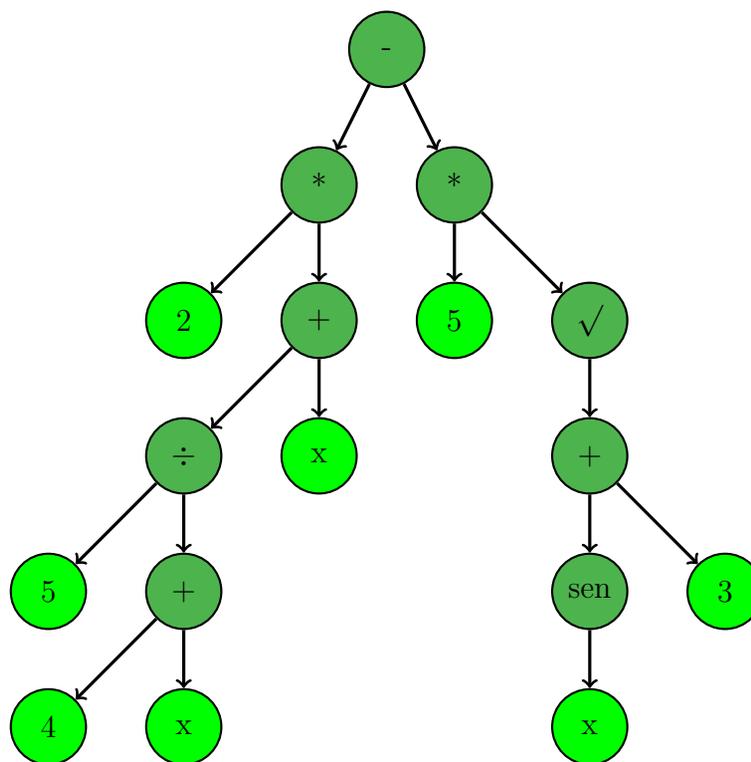


Figura 3.8: Diagrama de un árbol solución al final del algoritmo. Los nodos terminales (hojas) corresponden a variables mientras que los nodos verde oscuro corresponden a operadores de la base.

Las expresiones resultantes del diagrama de árbol de la Fig[3.8] en notación de prefijo así como algebraica después de ser simplificada son respectivamente:

- $\text{subs}(\text{mul}(2, \text{sum}(x, \text{div}(5, \text{sum}(4, x))))), \text{mul}(5, (\text{sqrt}((\text{sum}(3, \text{sen}(x)))))))$
- $f(x) = 2 \left(x + \frac{5}{4+x} \right) - 5\sqrt{3 + \text{sen}(x)}$

Capítulo 4

Funciones de estudio.

4.1 Introducción

Antes de utilizar el **SR** para analizar observaciones de carácter cosmológico, es necesario comprender cómo los parámetros de pre-inicialización afectan la implementación. En este capítulo, se presentan y discuten los resultados del análisis de datos sintéticos basados en funciones con diferentes características de continuidad, dominio, rango entre otras, tomadas de []. Para considerar la influencia de los distintos operadores de mutación y cruzamiento se tomaron en cuenta los datos y anotaciones de Vazquez en [85]. Los datos sintéticos se construyeron a partir de una función $F(z)$ denominada *función muestra* cuyos errores se consideraron normales con media $\mu = 0$ y desviación estándar σ

$$\text{Datos} = F(z) + \mathcal{N}(\mu = 0, \sigma). \quad (4.1)$$

Para calcular y comparar la función fitness denotada por γ se eligió el error absoluto medio de la ec(2.5). Por tanto, una γ menor corresponde a una mejor solución. Las características completas de cada conjunto de datos se encuentran en las tablas del apéndice.

4.2 La librería *gplearn*.

El regresor implementado fue *SymbolicRegressor*, perteneciente a la biblioteca *gplearn* de Python. **gplearn** mantiene la familiar API de *scikit-learn* para ajustar y predecir, y es compatible con los módulos de pipeline y grid search existentes. Además del regresor, ofrece el uso de *SymbolicClassifier* para clasificación binaria y *SymbolicTransformer* para ingeniería automática de características [65].

Durante el proceso de evolución se utilizan los operadores descritos en la sección(3.5) cuya suma de probabilidades debe ser menor a uno para considerar una fracción de los individuos seleccionados en el torneo que pasan a la siguiente generación sin ninguna modificación. La librería además, permite personalizar el rango de donde se seleccionan los números que fungen como variables, por defecto es el rango $(-1, 1)$. La base estándar utilizada contiene las operaciones "add, mul, subs, div" de manera que en los respectivos casos se especifica la adición (\uparrow) o sustracción (\downarrow) de una operación. Además de parámetros como el coeficiente de parsimonia con el que se busca evitar soluciones muy densas (a un nivel alto le corresponden programas mas simples o con menor cantidad de nodos) existen otras características de pre-inicialización como el criterio de finalización, tamaño de los arboles iniciales etc. En la Tabla [4.1] se muestra a manera de referencia una configuración "estándar" de los parámetros más importantes que se utilizaron en todos los casos con excepción de las modificaciones particulares correspondientes. El algoritmo regresa la solución en notación de prefijo por lo que es necesario simplificarla utilizando la función *simplify* de la librería *sympy*.

Parámetro	Valor
Tamaño de la población <i>Pob</i>	1000
Numero de generaciones <i>Gen</i>	35
Función fitness <i>Fit</i>	MAE
P crossover <i>Cros</i>	0.80
P subtree mutation <i>Sub</i>	0.05
P hoist mutaiton <i>Hoi</i>	0.05
P point mutation <i>Poi</i>	0.05
Coficiente de parsimonia <i>ParsC</i>	0.005
Tamaño del torneo de selección <i>Torn</i>	25
Intervalo de variable numérica <i>Ran</i>	$(-1.01, 1.01)$
Profundidad inicial de los árboles <i>Prof</i>	$(2, 6)$

Tabla 4.1: Parámetros de ajuste de una configuración estándar.

4.3 Función lineal.

Se aplico el regresor a un conjunto de datos basados en la función

$$F(z) = 15z - 4 \tag{4.2}$$

cuyas características se encuentran en la Tabla[B.1].

Primero se analizó la variabilidad en las soluciones manteniendo una configuración estándar de los parámetros de entrada pero modificando el tamaño de la población en un rango de 100 a 5000 individuos. Como se muestra en la Tabla[4.2] un número pequeño de la población resulta en una evolución poco satisfactoria de las soluciones haciendo que en este caso una población de 100 individuos posea un fitness γ seis veces mayor que el resto de las funciones obtenidas. En contraste el algoritmo arroja mejores soluciones al aumentar el tamaño de población haciendo que converjan a dos cifras significativas. Con base en la misma tabla, la solución que mejor describe los datos en este caso es :

$$f(z) = 15.086z - 6.124. \tag{4.3}$$

La cual corresponde a una población de 5000 individuos y posee una mejor aptitud que la *función muestra*.

Un procedimiento similar se realizó al variar el número de generaciones posibles manteniendo análogamente una configuración de parámetros estándar. Como se muestra en la Tabla[4.3] a un aumento gradual en el número de generaciones no necesariamente le corresponde una mejora en las aptitudes de las soluciones. Sin embargo, es posible obtener funciones con fitness mas cercanos a los de la *función muestra* con una menor cantidad de tiempo y memoria. La solución que mejor describe los datos en este caso es :

$$f(z) = 15.074z - 5.94. \tag{4.4}$$

f_i	Población	Solución	Fitness γ_i
f_{Muestra}	–	$15z - 4$	7.0085
f_1	100	$z^2 + z$	42.499
f_2	500	$15.042z - 6.112$	6.7649
f_3	1000	$15.105z - 6.305$	6.7640
f_4	2500	$15.02z - 5.657$	6.7691
f_{5000}	5000	$15.086z - 6.124$	6.7593

Tabla 4.2: Soluciones en función del tamaño de la población.

f_i	Generación	Solución	Fitness γ
f_{Muestra}	–	$15z - 4$	7.0085
f_1	15	$15.067z - 3.943$	7.0602
f_2	45	$14.82z - 3.076$	7.1470
f_3	75	$15.074z - 5.94$	6.7540
f_4	100	$15.123z - 6.423$	6.7681
f_5	130	$15.352z - 8.398$	6.8546

Tabla 4.3: Soluciones en función del número de generaciones.

En las Figuras [4.1,4.2], se presentan las gráficas correspondientes a los mejores y peores individuos en comparación con la función de muestra. Estas gráficas están relacionadas con el tamaño de la población y el número de generaciones, respectivamente.

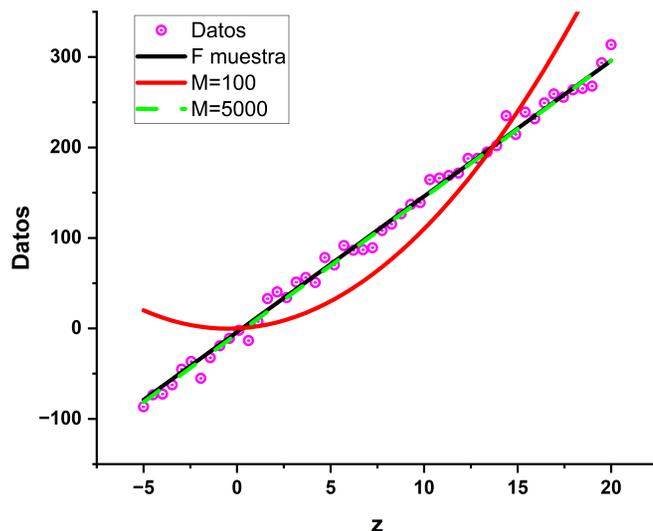


Figura 4.1: Funciones resultantes con diferente tamaño de población para la ec(4.2). Para apreciar mejor el comportamiento de las soluciones se presentan solo dos de éstas.

Al analizar la evolución del fitness a través de cada generación se puede observar el carácter aleatorio en muchas de las soluciones. En la Fig[4.3] la generación cero posee un fitness promedio muy alto en comparación con el resto de las generaciones, esto es porque al inicio la población no ha sido puesta a prueba por ningún criterio; simplemente ha sido construida aleatoriamente. Aunque el promedio de aptitud decrece en ocasiones al evolucionar a través de las generaciones, aún existen promedios con valores muy altos, esto muestra que siempre se está manteniendo diversidad en la población ya sea generando individuos mal adaptados, conservando grupos de miembros no tan bien adaptados producto de la reproducción y/o mutaciones aleatorias.

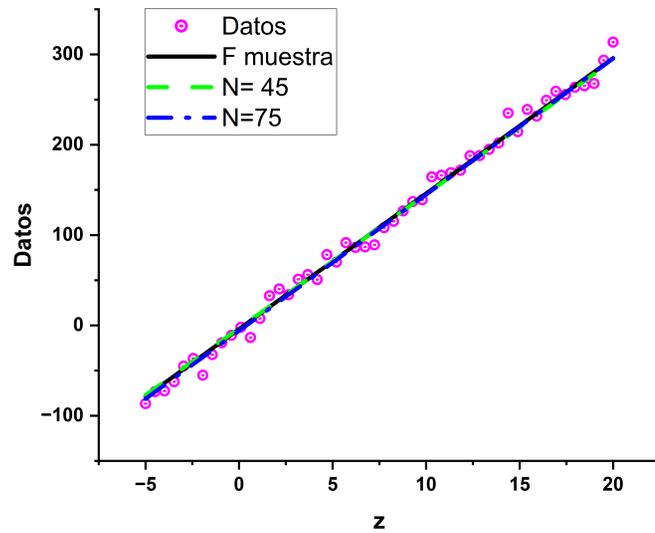


Figura 4.2: Funciones resultantes con diferente número de generaciones para la ec(4.2). Debido a la gran similitud de las curvas, se presentan solo un par de las soluciones.

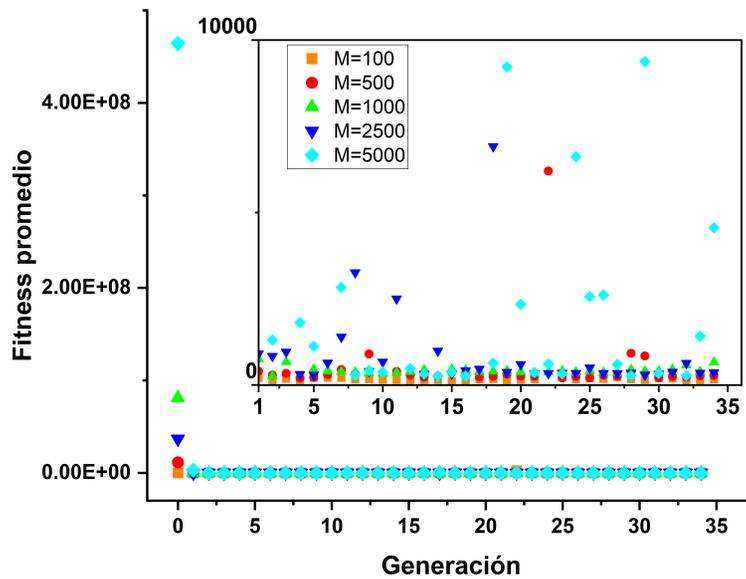


Figura 4.3: Evolución del fitness promedio a través del paso de generaciones para distintos tamaños de población.

En contraste, la adaptación del individuo más sobresaliente de cada generación es en general mejor cada vez. Como se muestra en la Fig[4.4] nuevamente la primera población arroja a los individuos mejor adaptados con un fitness grande en comparación con el resto de las generaciones. Sin embargo es notable la rápida disminución en el error de las soluciones al pasar por la primera prueba de selección. Además, se puede observar

una correlación entre la magnitud del fitness de la primera generación y el tamaño de población pues como es de esperarse en una muestra aleatoria mas grande se espera encontrar por probabilidad, una mejor solución (fitness mas bajo). De acuerdo con este caso, una población de 500 individuos alcanzó de forma mas rápida (número de generación mas pequeño) el valor fitness final sin embargo el comportamiento que tuvo fue convergente en comparación con las poblaciones de 1000, 2500 y 5000 individuos las cuales hacia la última etapa del programa mostraron indicios de una posible reducción en el valor γ .

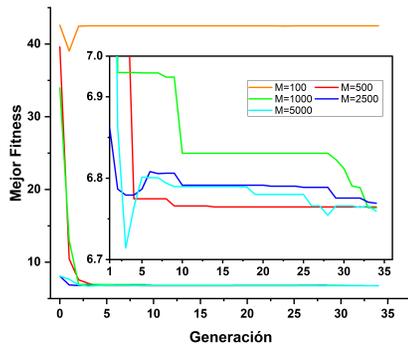


Figura 4.4: Evolución del fitness del mejor individuo a través del paso de generaciones para casos con distinto tamaño de población.

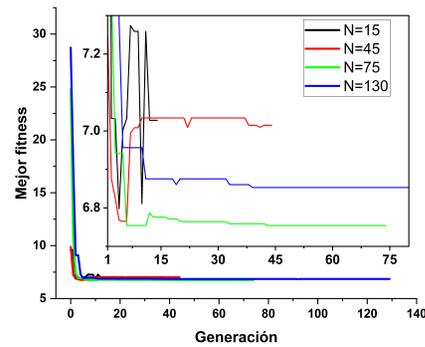


Figura 4.5: Evolución del fitness del mejor individuo a través del paso de generaciones para casos con un número distinto de generaciones.

Con base en la Tabla[4.3] y la Fig[4.5], un aumento en el número de generaciones manteniendo constante la población no garantiza obtener mejores soluciones. Aunque se sigue observando el comportamiento decreciente en la mayoría de los casos, las soluciones a partir de la generación estándar ($N=35$) no mostraron gran cambio en el fitness. Debido a esto, es recomendable variar simultáneamente el tamaño de la población y el número de generaciones considerando posiblemente el tamaño del torneo de selección como se verá mas adelante.

4.4 Función racional.

Se aplico el regresor a un conjunto de datos basados en la función

$$f(z) = -\frac{z^6 + z^5 - 10z^2 - 10z - 12}{z^2 + 6}, \quad (4.5)$$

cuyas características se encuentran en la Tabla[B.2].

Debido al vaor de los coeficientes presentes en la expresión se opto por modificar el rango de las variables numéricas. Además se aumento el tamaño de la población junto con el número de generaciones así como el tamaño del torneo de selección para conocer la influencia que tiene este último en el fitness de los mejores individuos.

Las soluciones obtenidas en general presentaron un gran "follaje", es decir una gran cantidad de nodos en la representación de árbol, sin embargo muchas de las ramas obtenidas corresponden a finalmente a variables numéricas que pueden simplificarse. Por otra parte, como se observa en las Tablas [4.4,4.5] la solución con peor fitness es la función f_3 correspondiente cuyo proceso de evolución le correspondió un menor valor del operador de cruzamiento.

Al considerar las expresiones algebraicas es posible notar una degeneración con polinomios de cuarto grado como lo son las funciones f_2 y f_3 . Sin embargo, también es posible reconstruir funciones cuyos polinomios del denominador tienen una grado cercano al de la función muestra como se observa en las funciones f_4 y f_5 las cuales fueron obtenidas de un procesos cuyo coeficiente de Parsimonia fue disminuido con respecto a la configuración estándar. No obstante, ninguna de las expresiones logro reproducir el grado de los polinomios presentes en la función muestra de manera simultanea. Esto puede ser debido a que aunque en una perspectiva general todas las expresiones se acercan a los datos como se ve en la Fig[4.6], existe una región de mayor variación de éstos en el intervalo $[-1.5, 0.5]$ (Fig[4.7]) propiciando así una mayor variedad de posibles funciones. Por último cabe señalar que existen ciertos puntos de discontinuidad o asíntotas referentes a la división por cero en las expresiones f_1 , f_4 y f_5 que debido al dominio de los datos sintéticos no figuran parte en el proceso de evaluación.

f_i	Pob	Gen	Torn	Ran	PC	Cros	Sub	Hoi	Poin
f_{Muestra}	–	–	–	–	–	–	–	–	–
f_1	2500	60	50	(-4,4)	0.005	0.7	0.1	0.05	0.05
f_2	3500	80	65	(-4,4)	0.005	0.7	0.1	0.05	0.05
f_3	3500	70	45	(-4,4)	0.005	0.6	0.1	0.1	0.1
f_4	5000	70	45	(-4,4)	0.0025	0.7	0.05	0.1	0.1
f_5	5000	90	35	(-4,4)	0.0025	0.7	0.05	0.1	0.1

Tabla 4.4: Tabla de los parámetros asociados a las soluciones racionales de la Tabla [B.2]

f_i	Sol	γ
f_{Muestra}	$-\frac{z^6+z^5-10z^2-10z-12}{z^2+6}$	0.3621
f_1	$\frac{10.129z^4+12.282z^3-26.986z^2-39.972z-36.474}{z^2-0.595z-18.391}$	0.3847
f_2	$-0.7z^4 - 0.602z^3 + 1.991z^2 + 1.793z + 1.84$	0.4345
f_3	$-0.672z^4 - 0.622z^3 + 1.831z^2 + 2.186z + 1.857$	0.5572
f_4	$\frac{2.963z^6-8.482z^5+5.023z^4+33.46z^3-16.606z^2+2.798z-169.373}{z^4-3.480z^3-5.542z^2+42.217z-70.911}$	0.4164
f_5	$\frac{-0.227z^5+1.481z^4+z^3-3.403z^2-0.793z-7.156}{z-3.403}$	0.4613

Tabla 4.5: Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.5).

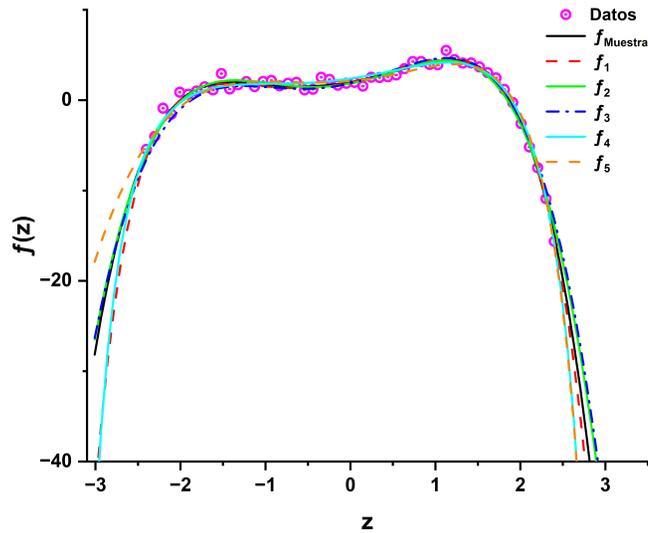


Figura 4.6: Comparación de funciones obtenidas respecto a la ec(4.5) sobre el dominio completo de los datos de la Tabla [B.2].

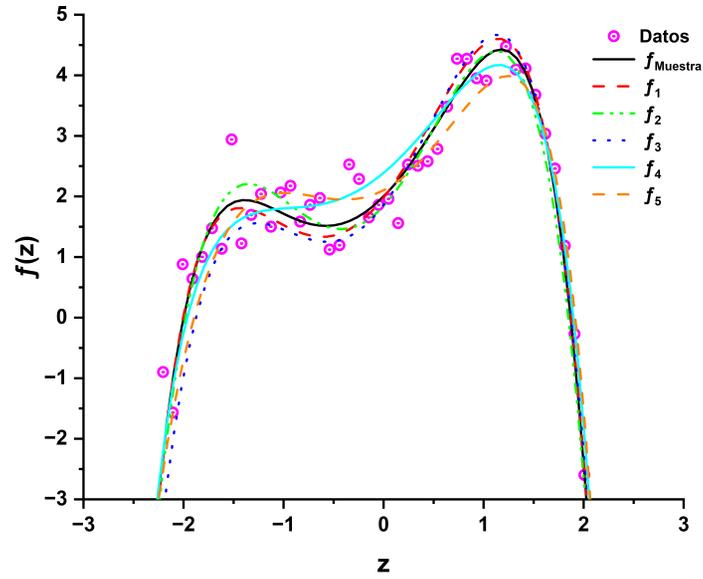


Figura 4.7: Comparación de funciones obtenidas respecto a la ec(4.5) en la zona de mayor variación de los datos.

De acuerdo a los datos de las Tablas mencionadas en esta sección, la solución que mejor describe los datos es :

$$f(z) = \frac{10.129z^4 + 12.282z^3 - 26.986z^2 - 39.972z - 36.474}{z^2 - 0.595z - 18.391}. \quad (4.6)$$

4.5 Funciones adicionales.

4.5.1 Valor absoluto

Después de conocer la influencia del número de individuos así como de generaciones, se dio paso a la adición de otras funciones a la base original y ver que tan bien el algoritmo puede reconstruir las funciones individuales. Para esto se aplicó el regresor a un conjunto de datos basados en la función

$$F(z) = |z|, \quad (4.7)$$

cuyas características se encuentran en la Tabla[B.3].

Primero se optó por intentar obtener una expresión polinomial para describir a los datos al no incluir la función *valor absoluto* `| |` en la base y usar la configuración estándar. Posteriormente se incluyó la función `'abs'` y se aplicó el regresor teniendo como objetivo dos casos :

- Obtener una solución con mejor fitness que la *función muestra*.
- Obtener la expresión de la *función muestra*.

En la Tabla[4.6] puede observarse cómo las soluciones que no contemplan la función valor absoluto no logran describir el conjunto de datos teniendo así una γ alta en comparación con la correspondiente a la función muestra. Una vez que se incluye la nueva función a la base, el algoritmo rápidamente arroja soluciones con ésta haciendo que converja satisfactoriamente no solamente en cuanto al fitness sino también a la expresión original de la función muestra. No obstante, debido posiblemente a los datos dispersos con valores negativos cercanos a $z = 0$ que pueden observarse en la Fig[4.8], una solución con un término extra resultó mejor que la solución original pues la primera contempla mejor dichos datos ignorando el 'espacio' entre éstos debido a su comportamiento asintótico en $z = 0$. Dicha función es :

$$f(z) = |z| - \frac{0.187}{z}. \quad (4.8)$$

f_i	Gen	ParsC	Cross	Sub	Hoi	Poin	Base	Sol	γ
f_{Muestra}	-	-	-	-	-	-	-	$ z $	0.7696
f_1	35	0.001	0.8	0.05	0.05	0.05	-	$0.089z^2 - 0.02z + 1.071$	1.1623
f_2	35	0.005	0.6	0.1	0.1	0.1	-	$0.082z^2 + 2.422$	0.8974
f_3	10	0.005	0.8	0.05	0.05	0.05	$\uparrow $	$ z - \frac{0.187}{z}$	0.7454
f_4	25	0.005	0.8	0.05	0.05	0.05	$\uparrow $	$ z $	0.7696

Tabla 4.6: Parámetros y soluciones correspondientes al regresor aplicado a los datos de la función valor absoluto .

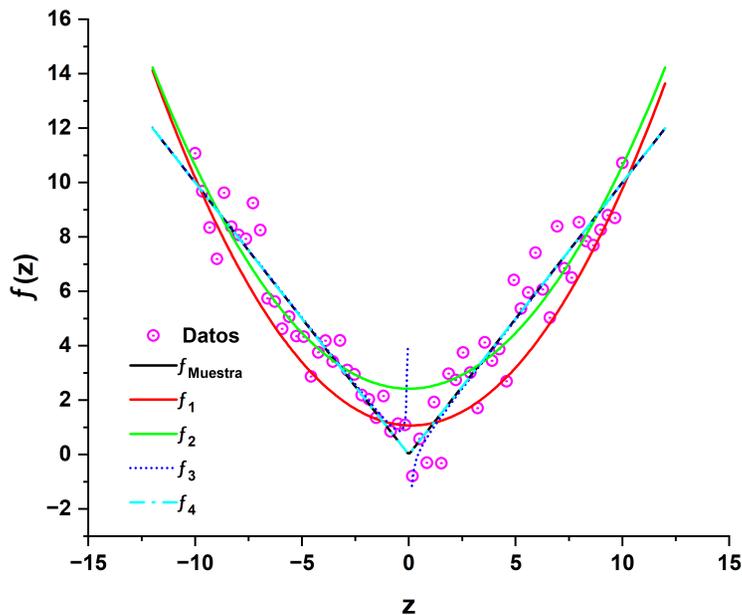


Figura 4.8: Curvas de las soluciones obtenidas en comparación con la función muestra de la ec(4.7) .

4.5.2 Raíz Cuadrada

Análogamente al caso anterior se aplicó el regresor a un conjunto de datos basados en la función

$$F(z) = \sqrt{z}, \quad (4.9)$$

cuyas características se encuentran en la Tabla[B.4].

Como se observa en la Tabla [4.7] , una implementación sin considerar la función raíz cuadrada genera soluciones con un fitness mayor que la función muestra, no obstante es interesante la solución f_3 pues posee el fitness correcto a dos cifras decimales. Dicha función presenta una recursión en el denominador del tercer término lo cual hace que al considerar una extrapolación a $z = 0$ el valor sea $f(z = 0) = 0.657$ en contraste con la función raíz cuadrada cuyo origen sería en $f(z = 0) = 0$. Este comportamiento puede

f_i	Gen	ParsC	Base	Solución	γ
f_{Muestra}	–	–	–	\sqrt{z}	0.1223
f_1	25	0.001	–	$0.25z + 0.904$	0.1564
f_2	25	0.001	–	$\frac{0.26+1.475z+0.208z^2}{0.904+z}$	0.1354
f_3	35	0.005	–	$0.18z + 1.360 - \frac{1.22}{z^2 + \frac{1.22}{z^2 + 0.703}}$	0.1245
f_4	35	0.005	$\uparrow \sqrt{}$	\sqrt{z}	0.1223

Tabla 4.7: Parámetros y soluciones correspondientes al regresor aplicado a los datos de la función raíz cuadrada .

verse reflejado en la Fig[4.9]. Por otra parte en este caso también fue posible obtener la expresión de la función muestra sin mayor problema pues al utilizar la configuración estándar solo fue necesario añadir la función raíz cuadrada. Por lo que la solución que mejor describe los datos es :

$$f(z) = \sqrt{z} \tag{4.10}$$

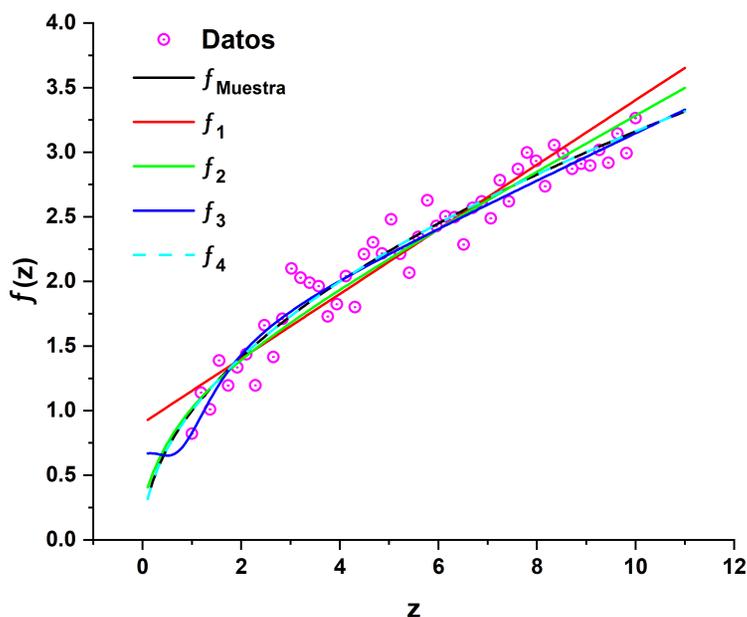


Figura 4.9: Curvas de las soluciones obtenidas en comparación con la función muestra de la ec(4.9) .

En los dos casos anteriores se mostró como la adición de otras funciones puede llevar

a mejores resultados. Es importante notar tanto el comportamiento de los datos como de las soluciones originales. En el caso del conjunto de la Tabla [4.6] el cambio brusco o no suave en los datos sugirió una función con comportamiento como el de la función valor absoluto. Por otra parte, la imposibilidad de un polinomio de describir los datos de la Tabla [4.7] sugirió la adición de una función extra como la raíz cuadrada.

4.6 Función logaritmo.

Se aplico el regresor a un conjunto de datos basados en la función

$$f(z) = z^4(1 - \ln(z)) \quad (4.11)$$

cuyas características se encuentran en la Tabla[B.5].

f_i	Sol	γ
f_{Muestra}	$z^4(1 - \ln(z))$	0.3550
f_1	$z^4(1 - \ln(z)) + 0.012z^2$	0.3659
f_2	$1.0272z^2 + 2.304z^2\ln(z) - 0.460z^4\ln(z)$	0.6579
f_3	$z^4(1.105 - 1.078\ln(z))$	0.7647
f_4	$-0.204z^4(\ln(z^4) + \ln(\ln(z))) - 3.986$	0.4097
f_5	$z^4(1 - \ln(z))$	0.3550

Tabla 4.8: Funciones y fitness asociado correspondientes al regresor aplicado a los datos de la función logaritmo de la ec(4.11) .

Para conocer el alcance del algoritmo con una elección inicial en la base, se incluyo la función logaritmo y se ejecuto con una configuración estándar. Al existir soluciones con composición múltiple como por ejemplo ($\log(\log(1+\log(\log(z))))$) se incremento el coeficiente de Parsimonia asi como se redujo la profundidad de los arboles de la población inicial para evitar una sobrecarga de composiciones y así al evolucionar en cada generación los individuos se ajustaran con menor rapidez pero evitando en mayor medida la *sobrecomposicion*.

Como se observa en la función f_3 de la Tabla[4.8] aunque esta posee la misma forma, las pequeñas diferencias en los parámetros hacen de echo que tenga un fitness mayor que el resto de las soluciones esto es debido al carácter del logaritmo pues el argumento pasa de ser z a $z^{1.078}$ el cual tiene mayor influencia en la curva. Por otra parte la función f_2 posee un termino cuadrático extra que en realidad no contribuye a la curva en la región de los datos como se muestra en la Fig[4.10] haciendo que tenga el fitness mas chico. Posteriormente se retiro la función división de la base para reducir el espacio de soluciones e intentar recuperar la expresión original, esto fue posible utilizando los parámetros que se muestran en la Tabla[B.9]. De esta manera, la función obtenida que mejor describe los datos es :

$$f(z) = z^4(1 - \ln(z)). \quad (4.12)$$

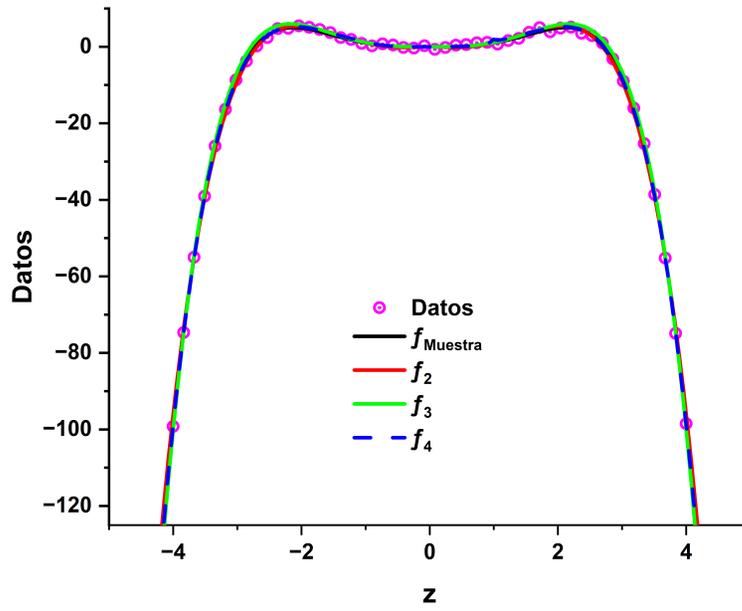


Figura 4.10: Comparación de funciones obtenidas respecto a la ec(4.11) sobre el dominio completo de los datos de la Tabla[B.5])

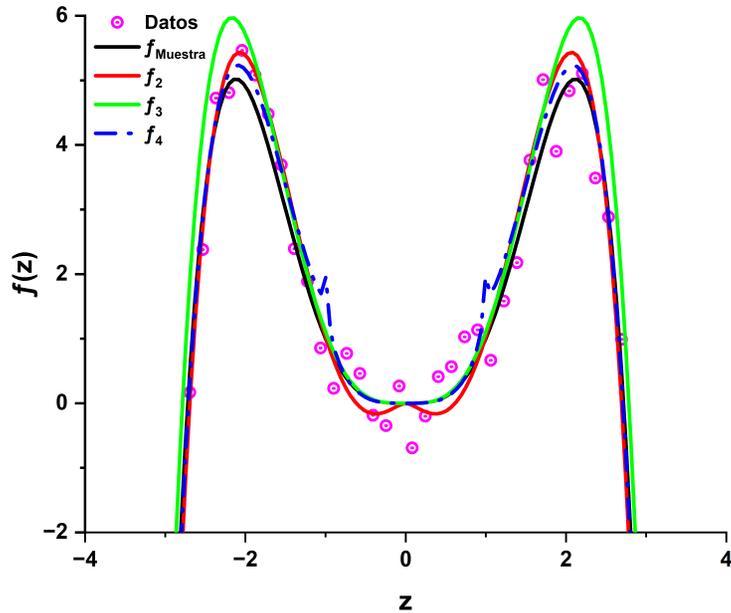


Figura 4.11: Comparación de funciones obtenidas respecto a la ec(4.11) en la zona de mayor variación de los datos.

4.7 Función con asíntota.

Se aplico el regresor a un conjunto de datos basados en la función muestra

$$f(z) = \frac{2z - 1}{z^2} \quad (4.13)$$

cuyas características se encuentran en la Tabla[B.6].

En esta implementación se modifico en baja medida los operadores de mutación con el fin de reconstruir la función muestra. Como se observa en la Tabla[4.9] el algoritmo fácilmente encuentra soluciones con fitness iguales a dos cifras y con expresiones muy similares entre sí. La existencia de una asíntota en $z = 0$ no interfiere en la reconstrucción pues como se observa en la Fig[4.12] los datos mas alejados también son incluidos en las curvas. Se puede conjeturar que una expresión sencilla como esta es mas fácil de reconstruir que una función racional como la de la sec(4.4) debido a que se requieren menos parámetros en comparación. Para poder analizar esto se podría realizar una reconstrucción con funciones cuya asíntota este desplazada a unidades del origen. De acuerdo con los valores de γ la función que mejor describe los datos es :

$$f(z) = \frac{1.979z - 1.003}{z^2}. \quad (4.14)$$

f_i	Pob	ParsC	Cross	Sub	Hoi	Poin	Sol	γ
f_{Muestra}	–	–	–	–	–	–	$\frac{2z-1}{z^2}$	0.2723
f_1	1000	0.001	0.8	0.05	0.05	0.05	$\frac{1.972z-1.003}{z^2}$	0.2722
f_2	2000	0.005	0.6	0.1	0.1	0.1	$\frac{1.979z-1.003}{z^2}$	0.2704
f_3	2000	0.005	0.7	0.1	0.1	0.05	$\frac{2z-1}{z^2}$	0.2723
f_4	2000	0.001	0.8	0.05	0.05	0.05	$\frac{1.996z-1}{z^2}$	0.2714

Tabla 4.9: Parámetros y soluciones correspondientes al regresor aplicado a los datos de la ecuación asintótica (4.13).

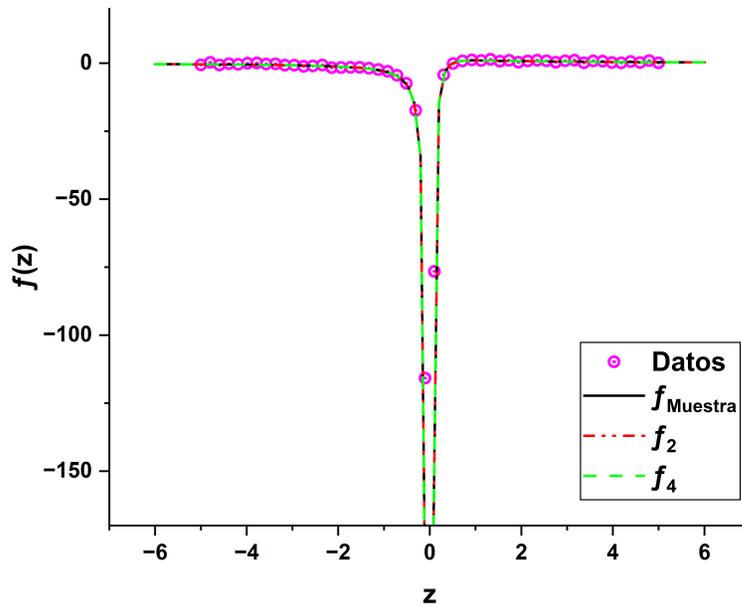


Figura 4.12: Comparación de funciones obtenidas respecto a la ec(4.13) sobre el dominio completo de los datos de la Tabla[4.9]). Debido a la similitud de las curvas solo se presenta un par de éstas.

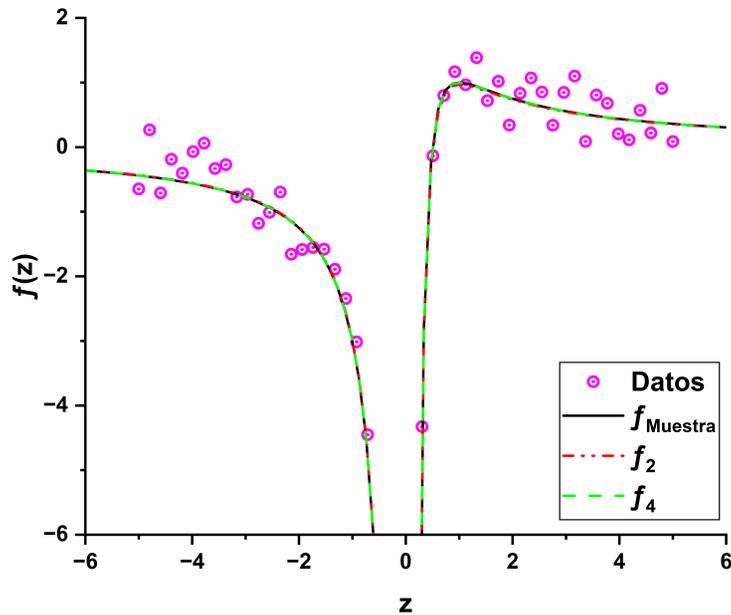


Figura 4.13: Comparación de funciones obtenidas respecto a la ec(4.13) en la zona de mayor variación de los datos. Debido a la similitud de las curvas solo se presenta un par de éstas.

4.8 Función irracional.

Se aplico el regresor a un conjunto de datos basados en la función

$$f(z) = 5 - \sqrt{z^4 + 2z^3 - 3z^2 + 13}, \quad (4.15)$$

cuyas características se encuentran en la Tabla[B.7].

Debido a que en las primeras pruebas al incluir la operación raíz cuadrada las soluciones eran demasiado grandes, se optó por retirar la función división de la base. Se modificó también el rango de las variables numéricas acorde a los presentes en la función muestra. Además con el fin de intentar reducir la composición de raíces en las soluciones, se redujo la profundidad de los arboles de la primera población.

Como se observa en la Tabla [4.10] las expresiones resultantes presentan una alta composición de términos con raíces los cuales no pueden simplificarse y debido a su naturaleza, el dominio de los datos se restringe a considerar el valor absoluto de los radicandos generando picos o cambios bruscos de comportamiento en las curvas asoci-

adas como se muestra en la Fig[4.14].

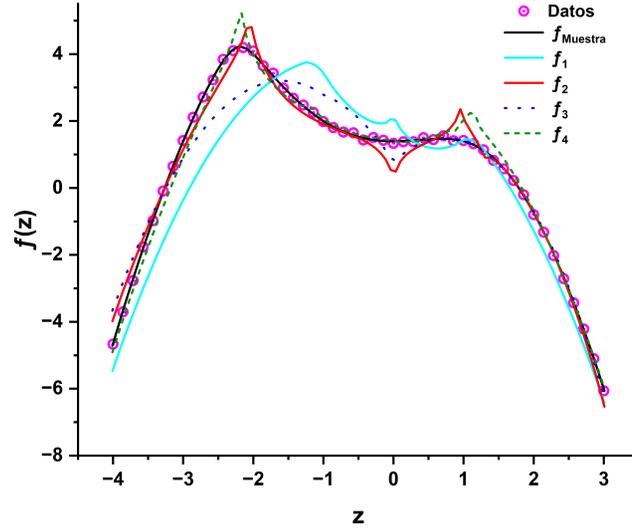


Figura 4.14: Comparación de funciones obtenidas respecto a la función irracional muestra de la ec(4.15).

En la Tabla[B.10] se muestran los parámetros asociados a los procesos de evolución para cada solución obtenida. Las poblaciones fueron incrementadas con el fin de obtener mayor variedad de soluciones razón por la cual, el número de generaciones también se incrementó así como el tamaño del torneo de selección. Las soluciones mostraron alta sensibilidad al coeficiente de Parsimonia de manera que éste se mantuvo fijo en la mayor parte de las ejecuciones del programa. En este caso, se le dio mayor peso al operador de elevación con la intención de eliminar ramas del árbol y así simplificar la solución final. De acuerdo con la Tabla de soluciones y fitness, la solución que mejor representa los datos es :

$$f(z) = 3.466 - z + 1.44 \left(\sqrt{1 - (z - z^2)0.389} \right) \sqrt{\sqrt{1 - (z - z^2)0.363} - 0.6 \sqrt[4]{z^7(-1 + \sqrt[4]{-0.38z^2 - z^3})}}. \quad (4.16)$$

La función irracional de ésta sección es la que mayor complicación presentó en este escrito pues aún considerando los *hints* como la presencia de la operación raíz , eliminación de la operación división y la baja dispersión de los datos, no fue posible obtener expresiones simples ni con un fitness cercano al de la función muestra pues el asociado a la mejor solución es cerca de 8 veces mayor.

f_i	Solución	γ
f_{Muestra}	$5 - \sqrt{z^4 + 2z^3 - 3z^2 + 13}$	0.0382
f_1	$-z + 2.603 - \left[\left(z^2 - \sqrt[4]{z^2 - z^{3/2} + 1.525} \right) \left(-z^2 + \sqrt{z} + \sqrt{z^2 - z^{3/2}} \right) \right]^{1/2}$	0.8458
f_2	$\sqrt{9.05z - 2.71z^2} - \sqrt{\sqrt{3.78z^5 - 4.906z^4} - \sqrt{z^8 + 0.587z^7}}$	0.2526
f_3	$(z + 2.404) \left(-0.417z + \sqrt{0.02z - 0.089} + 0.139\sqrt{z^2 + z} \right) - (0.488z + \sqrt{z})(z - \sqrt{z^2 - 4.427z})$	0.4225
f_4	$3.466 - z + 1.44 \cdot \left(\sqrt{1 - (z - z^2)0.389} \right) \cdot \left[\sqrt{1 - (z - z^2)0.363} - 0.6 \cdot (z^7(-1 + (-0.38z^2 - z^3)^{1/4})) \right]^{1/2}$	0.2205

Tabla 4.10: Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.15).

4.9 Función trigonométrica.

Se aplico el regresor a un conjunto de datos basados en la función

$$f(z) = 3\text{sen}(3z + 1) + 4\text{cos}(z - 5), \quad (4.17)$$

cuyas características se encuentran en la Tabla[B.8].

En este caso se utilizo inicialmente la configuración estándar de los parámetros pero añadiendo las funciones coseno y seno a la base, así como retirando la operación división. Las funciones resultantes al igual que en el caso de la función logaritmo de la sec(4.6) mostraron una alta composición de funciones del tipo $\text{sen}(\text{sen}(\text{sen}(2\text{sen}(z))))$ las cuales se intentaron evitar reduciendo la profundidad de los árboles iniciales así como elevando el coeficiente de Parsimonia (ver Tabla[B.11]). Además, se modificó el rango de elección de las variables numéricas pensando en que el rango presente en la función muestra es mayor al que existe en la configuración estándar. Por último, se conservó la operación división en la base para eliminar el sesgo en la solución al retirarla inicialmente.

Como se observa en la Tabla[4.11] a medida que se eleva el coeficiente de Parsimonia las soluciones comienzan a tener una representación mas 'natural' con menor composición y con sumas de las funciones seno y coseno. La función f_5 resultó en la solución con la estructura mas simple y corta conteniendo solo 3 términos¹ además de poseer argumentos del tipo $(az + b)$. No obstante el fitness en realidad resultó mas grande que la mayoría del resto de las soluciones posiblemente debido a que el espacio

¹En realidad algunas de las soluciones devueltas por el algoritmo contienen mas términos, sin embargo es posible utilizar identidades trigonométricas para obtener expresiones mas cortas.

f_i	Sol	γ
f_{Muestra}	$3\text{sen}(3z + 1) + 4\text{cos}(z - 5)$	0.3866
f_1	$\text{sen}(\text{sen}(3z)) + 1.964(\text{cos}(3z) - \text{sen}(z) + \text{sen}(\text{cos}(z) - 2\text{sen}(z)))$	0.3813
f_2	$-4\text{sen}(z) + \text{sen}(3z) + \text{sen}(\text{sen}(3z + 1.089)) + \text{sen}(z) + 2\text{sen}(3z) - 0.032$	0.3959
f_3	$-2\text{sen}(z) + \text{sen}(3z) - 2\text{sen}(\text{sen}(z)) + \text{sen}(\text{cos}(3z - 0.771)) + \text{cos}(z) + 2\text{cos}(3z)$	0.3503
f_4	$-3.796\text{sen}(z) + 1.511\text{sen}(3z) + \text{cos}(z) + 2.579\text{cos}(3z)$	0.3556
f_5	$2.796\text{sen}(3z + 1.038) - 3.845\text{sen}(z) + \text{cos}(z)$	0.3897

Tabla 4.11: Funciones y fitness correspondientes al regresor aplicado a los datos de la ec(4.17).

de búsqueda resulta mayor al mantener la operación división.

Cabe recalcar que el comportamiento del fitness promedio no tiene grandes desviaciones como lo mostraron las secciones anteriores, esto es debido a que las funciones trigonométricas poseen un rango definido así como el comportamiento de los datos. Además, el valor del fitness de cada función obtenida no es muy diferente el uno del otro. Esto puede llegar a ser una desventaja si no se presta atención a las gráficas correspondientes pues como se ve en la Fig[4.15], la función f_2 que posee una γ mayor en solo una centésima con respecto a la función muestra (4.17) en realidad no se ajusta bien a los datos en comparación con el resto de las soluciones.

De acuerdo con los valores γ de la Tabla[4.11], la función que mejor describe los datos es :

$$f(z) = -2\text{sen}(z) + \text{sen}(3z) - 2\text{sen}(\text{sen}(z)) + \text{sen}(\text{cos}(3z - 0.771)) + \text{cos}(z) + 2\text{cos}(3z). \quad (4.18)$$

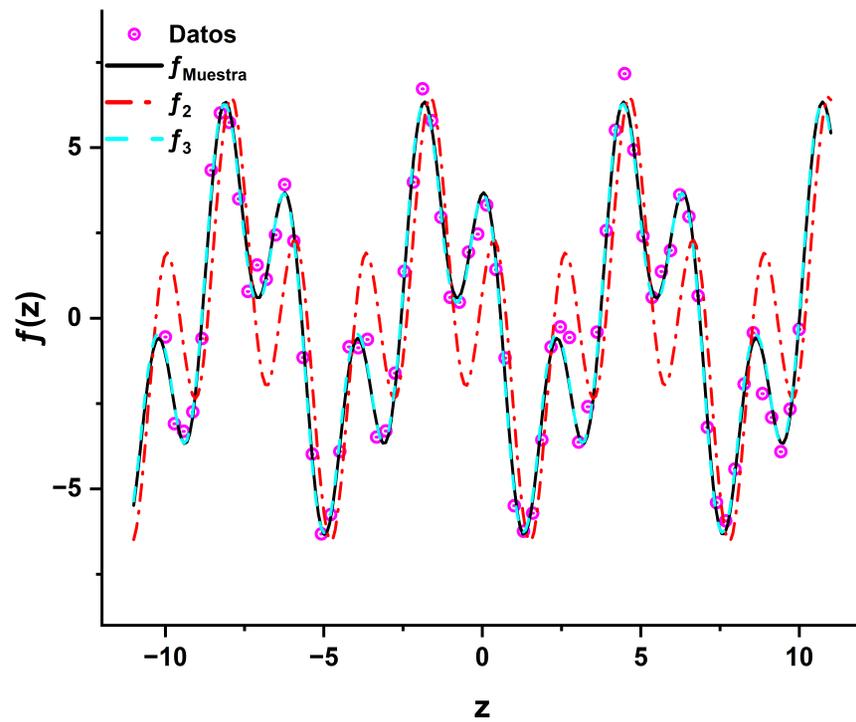


Figura 4.15: Comparación de funciones obtenidas respecto a la función trigonométrica muestra de la ec (4.17). Para mayor claridad en la presentación se presenta solo un par de las curvas asociadas a las soluciones.

Capítulo 5

Aplicación a variables cosmológicas.

5.1 Parámetro de Hubble.

Se buscó obtener la constante de Hubble utilizando las observaciones de la Fig[1.5] así como intentar reconstruir un modelo cosmológico al aplicar el Regresor Simbólico de la librería *gplearn*. Primero se hizo una exploración visual de los datos para conocer su comportamiento así como el rango en el que se encuentran. Debido a que los datos no presentan comportamiento oscilatorio ni logarítmico, dichas funciones no fueron consideradas. Aunque podría asociarse un comportamiento creciente como el de la función raíz cuadrada (que de hecho es el presentado en el modelo Λ CDM), ésta no se incluyó por los problemas presentados en la sección de la función irracional (4.8). Ya que no se observa ningún comportamiento discontinuo ni asintótico se retiró la función división \div de manera que solo se generaron polinomios de grado m . El rango de la variable numérica también se incremento para evitar árboles demasiado grandes. Los parámetros obtenidos así como las soluciones correspondientes se encuentran en las Tablas [B.12,5.1].

En las primeras ejecuciones a pesar de mostrar soluciones con un fitness mas bajo en comparación con por ejemplo el modelo Λ CDM , el comportamiento de las funciones reconstruidas fue de carácter decreciente en la región final de los datos $z \gtrsim 1.62$ como en la función f_4 de la Fig[5.1], debido a esto se ignoraron todas aquellas soluciones con un comportamiento similar . Además, es posible que el no considerar las incertidumbres asociadas a las mediciones exista el problema anterior sobreajustando las observaciones en la región final de los datos ($z \gtrsim 1.5$).

A pesar de que la función f_4 presentó un fitness menor el comportamiento decreciente en la región de los datos tuvo como consecuencia la penalización de dicha solución. Por

f_i	Sol	γ
$f_{\Lambda\text{CDM}}$	$70\sqrt{0.7(1+z)^3 + 0.3}$	9.8545
f_1	$18.209z^2 + 44.345z + 66.03$	9.2858
f_2	$-z^5 - z^4 + 0.131z^3 + 28.219z^2 + 33.154z + 69.16$	8.9475
f_3	$-z^5 - z^4 + 0.98z^3 + 14.23z^2 + 51.635z + 65.408$	9.0615
f_4	$-4.315z^6 + 1.127z^5 + 13.419z^4 + 5.227z^3 + 0.654z^2 + 47.49z + 66.34$	7.7562

Tabla 5.1: Funciones y fitness correspondientes al regresor aplicado a los cronómetros c3smicos.

tanto, la funci3n de Hubble reconstruida $H_R(z)$ es la funci3n :

$$H_R(z) = -z^5 - z^4 + 0.131z^3 + 28.219z^2 + 33.154z + 69.16. \quad (5.1)$$

Al factorizar y completar cuadrados perfectos se obtiene :

$$H_R(z) = 69.16 [-(-3.1 + z)(0.03 + z)(1.142 + z)((z + 1.463)^2 + 6.971)]. \quad (5.2)$$

Y elevando al cuadrado :

$$H_R^2(z) = 69.16^2 [(z - 3.1)^2(z + 0.03)^2(z + 1.142)^2((z + 1.463)^2 + 6.971)^2] \quad (5.3)$$

Es decir se obtiene una expresi3n del tipo:

$$H_R^2(z) = H_0^2 [(z - a)^2(z + b)^2(z + c)^2((z + d)^2 + r)^2]. \quad (5.4)$$

Con a, b, c, d, r numeros reales ,resultando as3 un valor aproximado para la constante de Hubble $H(z = 0) = H_0$ como:

$$H_0 \approx 69.16 \frac{km \cdot s}{Mpc}. \quad (5.5)$$

El cual es un valor dentro del rango aceptado de ec(1.13).

Una primera mejora a 3ste m3todo se realizar3a al considerar la χ^2 como la funci3n fitness de manera que se tomen en consideraci3n los errores asociados a las observaciones. Adem3s , reducir la composici3n de funciones como la ra3z cuadrada podr3a reducir la complejidad de las soluciones y as3 hacer viable la inclusi3n de esta funci3n en la base del algoritmo.

Como puede apreciarse, 3ste m3todo en general no proporciona un rango de error para el valor de los coeficientes presentes en la funci3n reconstruida.

Para obtener errores anal3ticos asociados al aplicar el regresor una cantidad finita de

veces, Nesseris et al. proponen en [86] emplear una integral de camino sobre todas las posibles funciones f resultantes de aplicar el regresor cuya contribución a la verosimilitud :

$$\mathcal{L} = \mathcal{N} \exp(-\chi^2(f)/2), \quad (5.6)$$

sea apreciable. Por otra parte en [87] utilizan el método llamado *Bootstrap Monte Carlo* que implica generar múltiples muestras de los datos originales mediante remuestreo con reemplazo y calcular las estadísticas deseadas para cada muestra.

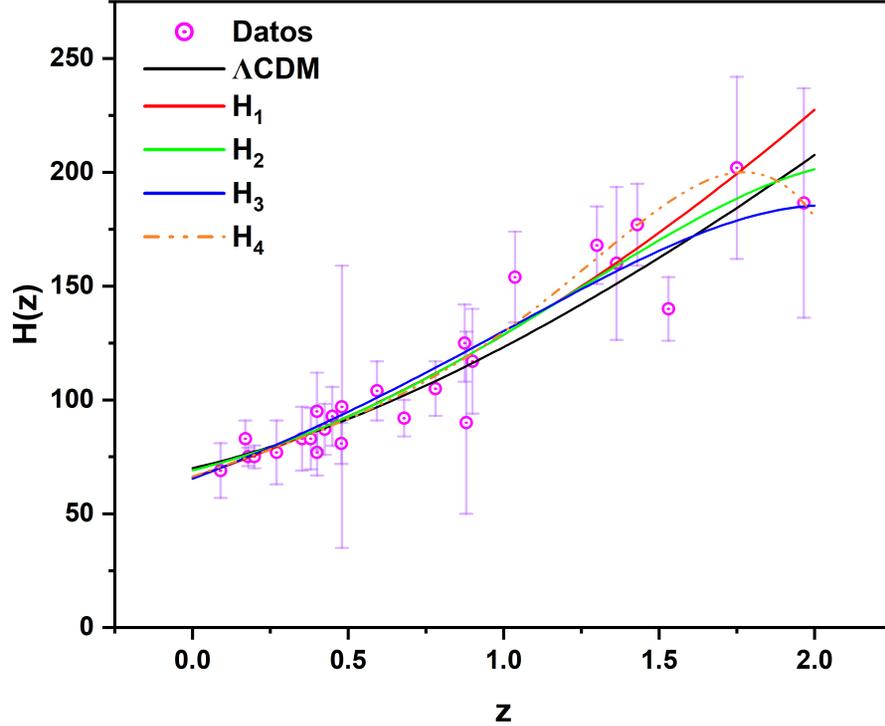


Figura 5.1: Curvas obtenidas utilizando el regresor simbólico para los datos del parámetro de Hubble. $H(z)$ en $km \cdot s \cdot Mpc^{-1}$

5.2 Densidad de energía.

Se buscó obtener una expresión para el conjunto de 29 datos reconstruidos Fig[5.2] utilizando observaciones de Supernovas Tipo I, Oscilaciones Acústicas de Bariones y Hubble (SNIa+BAO+ $H(z)$). Éste conjunto es una propuesta de parametrización libre para las reconstrucciones de la densidad de energía de la que posteriormente se selecciona un modelo a través de Evidencia Bayesiana [88].

Se realizó una metodología similar a la de la sección anterior. Debido a que los datos presentan un comportamiento oscilatorio se optó por incluir las funciones seno y coseno en la base, no obstante, las soluciones iniciales presentaron un sobre ajuste

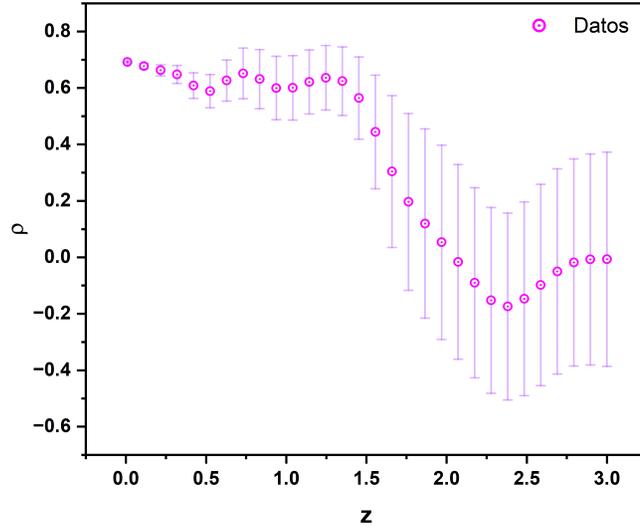


Figura 5.2: Reconstrucción de datos para la densidad de energía.

f_i	Solución	γ
f_1	$\text{sen}(\text{sen}(\text{sen}(z + 0.23)\cos(z - 0.461)))$	0.07299
f_2	$[1.746 - 1.446z + 0.288z^2 + (-1.81 + 2.78z - 0.93z^2)\text{sen}(z)]\text{sen}(0.396 + z)$	0.02906
f_3	$-0.321z - 0.056(z + \text{sen}[z^2] + 2.104)\text{sen}(z[z + \text{sen}(\text{sen}[\text{sen}(z)])] - 0.752)\text{sen}(1.095\text{sen}[z] + 3.76242) + 0.783$	0.02504
f_4	$-0.013z^2 + \text{sen}[z - \text{sen}(z) + 0.972] \cdot \text{sen}(\text{sen}[-0.248z + \text{sen}(\text{sen}(\text{sen}[0.849z + \text{sen}[\text{sen}(z) + \text{sen}(z + 0.695\text{sen}(z))]] + 0.356)]) + 0.474)$	0.02902

Tabla 5.2: Funciones y fitness correspondientes al regresor aplicado a los datos de la Fig[5.2]. Para mayor distinción en los términos se alterna el uso de paréntesis y brackets.

a los datos o en ocasiones simplemente un mal fitness, por esta razón, se decidió retirar la función coseno así como la función división ya que a pesar de ser incluida en las primeras ejecuciones, las soluciones finales no presentaban signos de utilizar dicha operación, de esta manera se redujo a su vez el espacio de búsqueda. Como en la sección de la función trigonométrica, se propuso una población mucho mayor que la de la configuración estándar así como un mayor número de generaciones. El coeficiente de Parsimonia se mantuvo estándar en las primeras ejecuciones, sin embargo las soluciones mostraron un comportamiento *simple* respecto a los datos como la función f_1 en la Fig [5.3]. Por esta razón se redujo el coeficiente de Parsimonia intentando equilibrar tanto la precisión del fitness como la complejidad y extensión de las soluciones finales. Es importante aclarar que se muestran las incertidumbres asociadas a los datos pero éstas no formaron parte del proceso de evolución pues la función de selección corresponde al error absoluto medio ec(2.5) que no toma en cuenta las incertidumbres.

En este caso se encontró que posiblemente debido al tamaño de las soluciones (número de nodos) y la alta composición de funciones, en algunas de las soluciones no fue suficiente con utilizar únicamente la función *simplify* de Python pues llegó a requerir incluso más tiempo para simplificar la notación de prefijo que la ejecución del programa completo. Por esta razón, se utilizó como herramienta auxiliar el modelo de lenguaje ajustado con técnicas de aprendizaje *ChatGPT* para identificar las principales componentes de la notación polaca resultante y así simplificarlas individualmente y por separado con la librería de Python (Algoritmo 4). Los parámetros utilizados así como las soluciones asociadas se encuentran en las Tablas [B.13,5.2]. De acuerdo con estas, la función que representa la densidad de energía es la expresión:

$$\begin{aligned} \rho(z) = & -0.321z - 0.056(z + \text{sen}[z^2] + 2.104) \times \\ & \text{sen}(z^2 + z\text{sen}(\text{sen}[\text{sen}(z)])) - 0.752z \times \\ & \text{sen}(1.095\text{sen}[z] + 3.76242) + 0.783, \end{aligned} \quad (5.7)$$

la cual presenta entonces una forma del tipo :

$$\begin{aligned} \rho(z) = & -az - b(z + \text{sen}[z^2] + c) \times \\ & \text{sen}(z^2 + z\text{sen}(\text{sen}[\text{sen}(z)])) - zd \times \\ & \text{sen}(r\text{sen}[z] + t) + \rho_0, \end{aligned} \quad (5.8)$$

en donde $\rho(z=0) = \rho_0 = 0.783$ y a, b, c, d, r, t son números reales mayores que cero. Esta función presenta 7 parámetros considerando la densidad de energía actual ρ además de una ligera composición del tipo $\text{sen}(\text{sen}(\text{sen}(z)))$. A pesar de tener el valor γ más bajo, en la Fig[5.3] se puede apreciar que el valor actual de la densidad de energía asociado a la ec(5.7) en realidad es más alto que el que sugieren los datos. Esta discrepancia puede ser debido a que la ec MAE (2.5) nuevamente no contempla las incertidumbres y por tanto no proporciona una mayor ponderación a los datos a bajo redshift con menor incertidumbre en contraste con la gran incertidumbre presente a alto redshift. Cabe señalar que en este caso es difícil identificar el comportamiento aparentemente decreciente de la función así como del aumento en su amplitud además de que, debido a las propiedades de las funciones trigonométricas, el considerar derivar dicha función supondría un reto considerable además de una expresión mucho más larga.

Vale la pena analizar la solución $f_2 = \rho_2$ cuya γ asociada no difiere mucho de aquella de mejor solución. En este caso la expresión se puede interpretar como la de una función senoidal modulada por un término compuesto por la suma de dos elementos: Un poli-

nomio de segundo grado y una función seno modulada por un polinomio de segundo grado. La función ρ_2 puede entonces expresarse como:

$$\rho_2(z) = (P[z] + Q[z]\text{sen}(z))\text{sen}(z + a). \quad (5.9)$$

En donde a es un número real, y $P[z]$, $Q[z]$ son polinomios de segundo grado con coeficientes p_0, p_1, p_2 y q_0, q_1, q_2 respectivamente. Esta forma permitiría una manipulación algebraica y analítica más sencilla. Considerando los valores de la función f_2 , la densidad de energía actual corresponde a $\rho_2(z = 0) = \rho_{2,0} = 0.673$, un valor que es mas cercano al sugerido por los datos. Por último, esta expresión al igual que la función f_1 posee 7 parámetros en total.

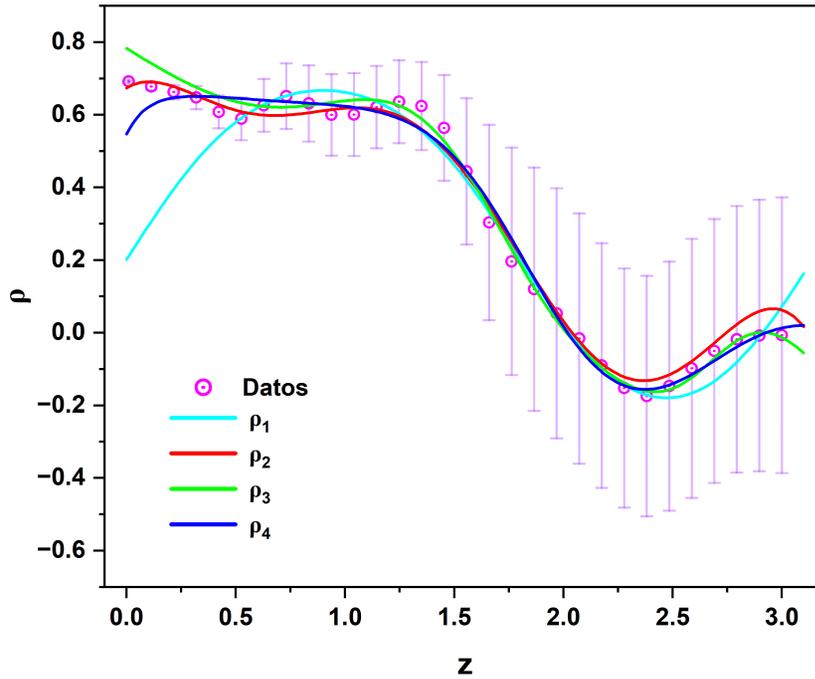


Figura 5.3: Curvas obtenidas utilizando el regresor simbólico a la reconstrucción de datos para la densidad de energía.

5.3 Parametrización de ecuación de estado.

Se buscó obtener una expresión para el conjunto de 29 datos reconstruidos Fig[5.4] utilizando observaciones de Supernovas Tipo I, Oscilaciones Acústicas de Bariones y Hubble (SNIa+BAO+H(z)). Éste conjunto es una propuesta de parametrización libre para las reconstrucciones de la ecuación de estado (EoS) de la que posteriormente se selecciona un modelo a través de Evidencia Bayesiana [88].

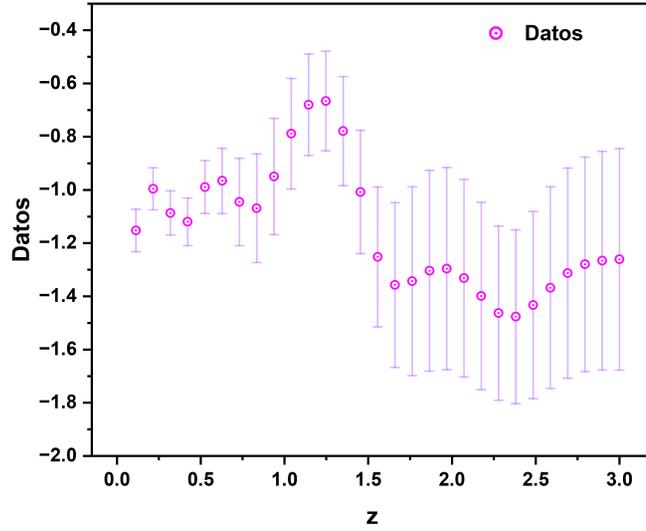


Figura 5.4: Reconstrucción de datos para la ecuación de estado.

f_i	Solución	γ
f_1	$-(\text{sen}(0.602z + 0.550) - \cos(0.947\text{sen}(z)) + 0.626)\cos(\text{sen}(\cos(z + \text{sen}(z)))) + \text{sen}((\cos(z) + 0.951)\text{sen}(\text{sen}(z)) + \cos(\cos(0.951\text{sen}^2(z))\cos(\cos(z(z\cos(0.945z^2) + 1)))) + \cos(\text{sen}(z^2))) - 1.237)$	0.03806
f_2	$-\sin(\sin(0.303z)) - \cos(\sin((z + \sin(0.303z + 0.118)) \cdot \sin(0.924z)))$	0.06987
f_3	$-\cos(\cos(z)\sin^2(z) + 1) - 0.654$	0.08238
f_4	$-0.190 \cdot z - \cos(\cos(z + \text{sen}(z)) + \text{sen}(z \cdot (\text{sen}(z + \text{sen}(z - 1.715) - \text{sen}(\text{sen}(\text{sen}(\text{sen}(\text{sen}(z) \cdot \cos(z + 0.368))))))) + \text{sen}(\text{sen}(\text{sen}(\cos(z)))) - 1.69) + 0.282) + 1.229)$	0.04271

Tabla 5.3: Funciones y fitness correspondientes al regresor aplicado a los datos de la Fig[5.4].

En comparación con la densidad de energía , las soluciones iniciales presentaron un mejor comportamiento así como una mejora en el fitness al incluir las dos funciones

trigonométricas $\text{sen}(z)$ y $\text{cos}(z)$. Similarmente varias de las soluciones presentaron una alta composición de éste tipo de términos razón por la cual se utilizó nuevamente el Algoritmo 4 para obtener las expresiones algebraicas asociadas. Los parámetros utilizados así como las soluciones correspondientes se encuentran en las Tablas[B.14,5.3]. De acuerdo con estas, la función que representa el parámetro de la ecuación de estado es la expresión:

$$\begin{aligned}
w(z) = & -[\text{sen}(0.602z + 0.550) - \text{cos}(0.947\text{sen}(z)) + 0.626] \times \\
& \text{cos}[\text{sen}(\text{cos}(z + \text{sen}(z)))] + \\
& \text{sen}[(\text{cos}(z) + 0.951)\text{sen}(\text{sen}(z))] \\
& + \text{cos}(\text{cos}(0.951\text{sen}^2(z)\text{cos}(\text{cos}(z(z\text{cos}(0.945z^2) \\
& + 1)))) + \text{cos}(\text{sen}(z^2))) - 1.237], \tag{5.10}
\end{aligned}$$

la cual presenta entonces una forma del tipo:

$$\begin{aligned}
w(z) = & -[\text{sen}(az + b) - \text{cos}(c\text{sen}(z)) + d] \times \\
& \text{cos}[\text{sen}(\text{cos}(z + \text{sen}(z)))] + \\
& \text{sen}[(\text{cos}(z) + g)\text{sen}(\text{sen}(z))] \\
& + \text{cos}(\text{cos}(h\text{sen}^2(z)\text{cos}(\text{cos}(z(z\text{cos}(kz^2) \\
& + n)))) + \text{cos}(\text{sen}(z^2))) - m], \tag{5.11}
\end{aligned}$$

en donde $w(z = 0) = w_0 = -1.096$ y $a, b, c, d, g, h, k, n, m$, son números reales.

Esta función presenta 9 parámetros además de una mayor cantidad de composiciones del tipo $\text{sen}(\text{sen}(\text{sen}(z)))$ y $\text{cos}(\text{cos}(\text{cos}(z)))$. A pesar de poseer el valor γ más bajo, en la Fig[5.6] se puede observar que el valor actual de la constante w es ligeramente menor que lo que sugieren los datos aunque más cercana a un modelo de energía oscura con $w = -1$. Por otra parte se aprecia un ligero sobreajuste a partir de redshift $z \gtrsim 1.5$, esto podría ser indicio de una desventaja al no contemplar las incertidumbres a partir de éste rango en la evolución de las soluciones. Al igual que en la sección anterior la expresión obtenida no es prometerdormente fácil de manejar analíticamente por la alta composición de los argumentos de las funciones trigonométricas principales. Una alternativa a explorar sería penalizar o cancelar aquellas composiciones de un mismo tipo de función al modificar el operador de elevación de manera que considere que al existir una evaluación de una función en sí misma, ésta sea *eliminada* y la variable ya sea numérica o simbólica se eleve a la función principal como se muestra en la Fig[5.5].

Otro enfoque para poder aprovechar el potencial del regresor simbólico es en la búsqueda de expresiones con un menor número de parámetros aunque el γ asociado no

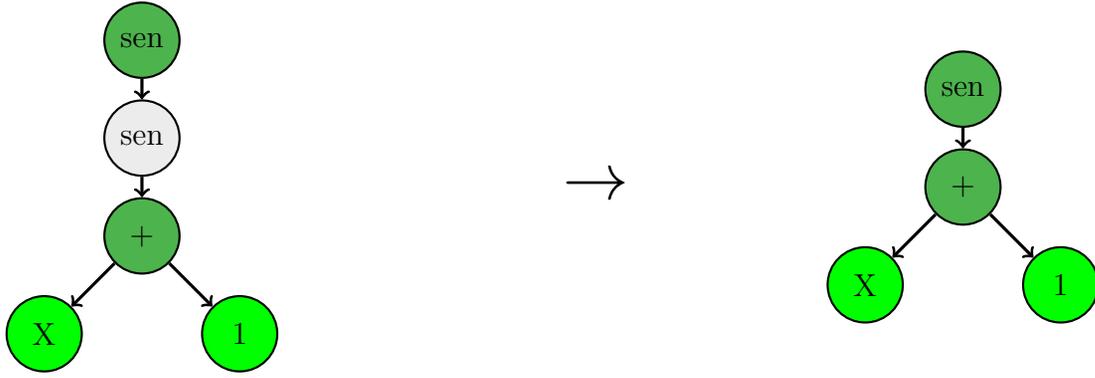


Figura 5.5: Para evitar una *sobrecomposición* de funciones, el operador de elevación modificado eliminaría automáticamente cualquier nodo función que sea evaluable en un nodo de la misma función, de éste modo el árbol de la izquierda cuya representación simbólica es $\text{sen}(\text{sen}(x + 1))$ se reduciría al árbol $\text{sen}(x + 1)$

sea el menor en comparación con otras soluciones. Al tener menos parámetros se puede utilizar un algoritmo de optimización para encontrar los valores que mejor ajusten la expresión. De ésta manera se podría combatir el problema de las expresiones demasiado complejas y así facilitar su uso tanto algebraico como analítico (manteniendo conciencia que al imponer estas limitaciones se está reduciendo el espacio de búsqueda).

La solución $f_2 = w_2$ posee dos características interesantes, la primera es que el valor actual de $w(z = 0) = w_0 = -1$ es el correspondiente de energía oscura como constante cosmológica mientras que la segunda es que posee solo 4 parámetros numéricos en comparación con los 9 de la ec(5.11). En este caso la expresión corresponde a la suma de una función seno y coseno cuyos argumentos son una función seno y una función lineal sumada a otra función seno respectivamente, es decir tiene la forma:

$$w_2(z) = -[\text{sen}(\text{sen}(az)) + \text{scos}(\text{sen}((z + \text{sen}(bz + c)) \cdot \text{sen}(dz)))] \quad (5.12)$$

en donde a, b, c, d son números reales. Como se observa en la Fig[5.6], ésta solución sigue la tendencia decreciente en función del corrimiento al rojo aunque en la región $z \in [1, 1.5]$ presenta una separación considerable respecto a los datos.

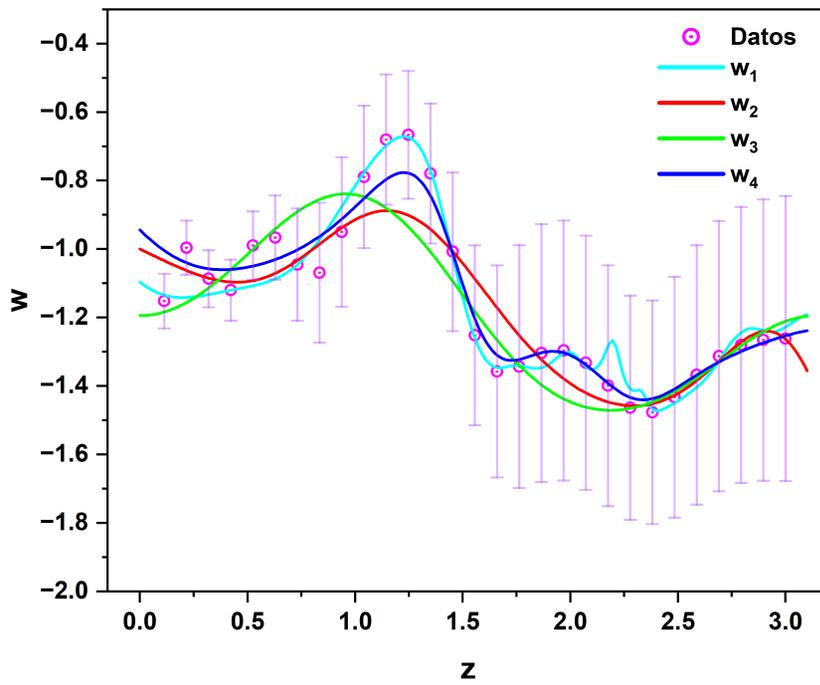


Figura 5.6: Curvas obtenidas utilizando el regresor simbólico a la reconstrucción de datos para la ecuación de estado.

Capítulo 6

Conclusiones

En este trabajo se presentó el paradigma de la programación genética como una herramienta útil para comparar y obtener directamente expresiones matemáticas que describen un conjunto de datos cuando no existe un antecedente conocido o no se quiere asumir un modelo existente, realizándolo a través de la implementación de un regresor simbólico.

La metodología utilizando un **SR** expuesta aquí puede ser aplicada a una variedad de funciones con gran precisión aún cuando el dominio de éstas contenga discontinuidades o esté formada con funciones adicionales a las básicas incluidas en el programa. Para los casos en los que se incluye la operación raíz cuadrada en la base del algoritmo se encontró una sobrecomposición de ésta haciendo que las expresiones resultantes sean muy complicadas de manipular algebraica o analíticamente como las expuestas en la Tabla[4.10]. Situación similar presentaron las soluciones de las Tablas [4.11,5.2,5.3] aunque debido a que en ocasiones es posible simplificar las expresiones mediante identidades trigonométricas se considera una situación más controlable. En el caso de las funciones polinómicas o racionales resultantes de las Tablas [4.5,5.1] se encontró que la forma general obtenida con éste método es del tipo:

$$f(z) = \frac{a_0z + a_1z + a_2z^2 + \dots a_nz^n}{b_0z + b_1z + b_2z^2 + \dots b_kz^k} = \frac{\sum_{k=0}^n a_n z^n}{\sum_{k=0}^m b_k z^k}, \quad (6.1)$$

donde $a_k, b_k \neq 0 \forall k$.

El regresor simbólico fue utilizado para obtener una expresión analítica para el parámetro de Hubble así como para estimar el valor actual de éste (H_0). Los resultados fueron respectivamente :

$$H_R^2(z) = 69.16^2 [(z - 3.1)^2(z + 0.03)^2(z + 1.142)^2((z + 1.463)^2 + 6.971)^2] \quad (6.2)$$

$$H_0 \approx 69.16 \frac{km \cdot s}{Mpc}. \quad (6.3)$$

Además , se buscó una expresión para describir la densidad de energía como función del corrimiento al rojo a un conjunto de datos del cual no se posee un modelo previo, para este caso se obtuvo la función:

$$\begin{aligned} \rho(z) = & -0.321z - 0.056(z + \text{sen}[z^2] + 2.104) \times \\ & \text{sen}(z^2 + z\text{sen}(\text{sen}[\text{sen}(z)])) - 0.752z) \times \\ & \text{sen}(1.095\text{sen}[z] + 3.76242) + 0.783. \end{aligned} \quad (6.4)$$

. Por último se encontró una solución para el comportamiento del parámetro de la ecuación de estado cosmológica en función del redshift como :

$$\begin{aligned} w(z) = & -[\text{sen}(0.602z + 0.550) - \cos(0.947\text{sen}(z)) + 0.626] \times \\ & \cos[\text{sen}(\cos(z + \text{sen}(z)))] + \\ & \text{sen}[(\cos(z) + 0.951)\text{sen}(\text{sen}(z))] \\ & + \cos(\cos(0.951\text{sen}^2(z)\cos(\cos(z(z\cos(0.945z^2) \\ & + 1)))) + \cos(\text{sen}(z^2))) - 1.237]. \end{aligned} \quad (6.5)$$

Una desventaja de éste método es que no se puede obtener directamente una incertidumbre asociada a los modelos resultantes, razón por la cual se sugieren técnicas como integrales de camino o remuestreos con simulaciones de Monte Carlo para ello. Para un trabajo a futuro se sugiere la adición de la χ^2 cómo función fitness en el proceso de evolución para así tomar en cuenta las incertidumbres asociadas a los datos. También, es posible explorar la adición de nuevas funciones con características y comportamientos mas sofisticados a la base del algoritmo. Por último se plantea la idea de modificar o añadir un nuevo operador de elevación de manera que elimine la sobrecomposicion recursiva de funciones idénticas en el proceso de selección con la intención de simplificar las expresiones antes del proceso final de evolución.

Apéndice A

Algoritmos

Algoritmo 2 Pseudo código para la generación recursiva de programas con los métodos Full y Grow

```
1: procedure GEN_RND_EXPR(set_T, set_V, prof_max, metodo)
2:   if prof_max = 0 or (metodo = grow and rand() <  $\frac{(|set\_T|)}{(|set\_T|+|set\_V|)}$ ) then
3:     expr ← CHOOSE_RANDOM_ELEMENT(set_T)
4:   else
5:     func ← CHOOSE_RANDOM_ELEMENT(set_V)
6:     for i = 1 to arity(func) do
7:       argi ← GEN_RND_EXPR(set_T, set_V, prof_max - 1, metodo)
8:     end for
9:     expr ← (func, arg1, arg2, ...)
10:  end if
11:  return expr
12: end procedure
```

13: Notas: **set_T** es el conjunto de funciones, **set_V** es el conjunto de variables y *expr* es la la expresion generada en la notacion polaca.

Algoritmo 3 Interpretación típica para GP

```
1: procedure EVAL(expr)
2:   if expr is a list then
3:     proc ← expr(1)                                ▷ Non-terminal: extract root
4:     if proc is a function then
5:       value ← proc(eval(expr(2)), eval(expr(3)), ...)  ▷ Function: evaluate
        arguments
6:     else
7:       value ← proc(expr(2), expr(3), ...)  ▷ Macro: don't evaluate arguments
8:     end if
9:   else
10:    if expr is a variable or expr is a constant then
11:      value ← expr                                ▷ Terminal variable or constant: just read the value
12:    else
13:      value ← expr()                              ▷ Terminal 0-arity function: execute
14:    end if
15:  end if
16:  return value
17: end procedure
```

18: Notas: **expr** es una expresión en notación de prefijo, **expr(1)** es el nodo primitivo de la raíz de la expresión, **expr(2)** es el primer argumento del nodo primitivo, **expr(3)** representa el segundo argumento y así sucesivamente.

Algoritmo 4 Proceso de Simplificación y Evaluación de Expresiones en Notación de Prefijo

```
1: while tiempo > tiempo_maximo do
2:   Identifica los  $n$  componentes principales de la expresión en notación de prefijo
    $Op(A_1, A_2, \dots, A_n)$ 
3:   ChatGPT devuelve  $A_1, A_1, \dots, A_n$ 
4:   Cambio a Python
5:    $a_1 = \text{simplify}(A_1)$ 
6:    $a_2 = \text{simplify}(A_2)$ 
7:    $\vdots$ 
8:    $a_n = \text{simplify}(A_n)$ 
9: end while
10:  $Op(a, b, \dots, c)$ 
11: Notas:  $Op$  es un operador de aridad  $n$  igual al número de componentes de la expresión.
```

Apéndice B

Tablas

Dominio	$[-5, 20]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	10

Tabla B.1: Características de los datos sintéticos asociados a la función lineal de la ec(4.2)

Dominio	$[-2.4, 2.4]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	0.5

Tabla B.2: Características de los datos sintéticos asociados a la función racional de la ec(4.5)

Dominio	$[-10, 10]$
Numero de datos	60
Media en la desviación μ	0
Desviación estándar σ	1

Tabla B.3: Características de los datos sintéticos asociados a la función valor absoluto de la ec(4.7)

Dominio	$[1, 10]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	0.15

Tabla B.4: Características de los datos sintéticos asociados a la función raíz cuadrada de la ec(4.9)

Dominio	$[-4, 4]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	0.4

Tabla B.5: Características de los datos sintéticos asociados a la función logaritmo de la ec(4.11)

Dominio	$[-5, 5]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	0.3

Tabla B.6: Características de los datos sintéticos asociados a la función con asíntota de la ec(4.13)

Dominio	$[-4, 3]$
Numero de datos	50
Media en la desviación μ	0
Desviación estándar σ	0.05

Tabla B.7: Características de los datos sintéticos asociados a la función irracional de la ec(4.15)

Dominio	$[-10, 10]$
Numero de datos	70
Media en la desviación μ	0
Desviación estándar σ	0.5

Tabla B.8: Características de los datos sintéticos asociados a la función trigonométrica de la ec(4.17)

f_i	Pob	Gen	Torn	Prof	PC	Cros	Sub	Hoi	Poin	Base
f_{Muestra}	-	-	-	-	-	-	-	-	-	-
f_1	3000	50	25	(2,4)	0.05	0.6	0.05	0.1	0.2	$\uparrow \ln$
f_2	3000	50	25	(2,4)	0.05	0.6	0.05	0.1	0.2	$\uparrow \ln$
f_3	3000	25	25	(2,3)	0.05	0.6	0.05	0.1	0.2	$\uparrow \ln$
f_4	2000	25	25	(2,4)	0.05	0.6	0.05	0.1	0.2	$\uparrow \ln$
f_5	5000	40	45	(2,6)	0.055	0.6	0	0.15	0.2	$\uparrow \ln, \downarrow \div$

Tabla B.9: Tabla de los parámetros asociados a las soluciones logaritmo de la Tabla [4.8]

f_i	Pob	Gen	Torn	Ran	Prof	PC	Cros	Sub	Hoi	Poin	Base
f_{M}	-	-	-	-	-	-	-	-	-	-	-
f_1	4000	50	45	(-5,5)	(1,2)	0.004	0.6	0	0.2	0.07	$\uparrow \sqrt{\downarrow \div}$
f_2	4000	65	45	(-5,5)	(1,2)	0.004	0.5	0.05	0.2	0.1	$\uparrow \sqrt{\downarrow \div}$
f_3	7000	90	70	(-5,5)	(1,3)	0.004	0.4	0.05	0.3	0.1	$\uparrow \sqrt{\downarrow \div}$
f_4	7000	90	70	(-7,7)	(1,3)	0.0037	0.5	0.05	0.3	0.1	$\uparrow \sqrt{\downarrow \div}$

Tabla B.10: Tabla de los parámetros asociados a las soluciones irracionales de la Tabla[4.10].

f_i	Pob	Gen	Torn	Ran	Prof	PC	Cros	Sub	Hoi	Poin	Base
f_{M}	-	-	-	-	-	-	-	-	-	-	-
f_1	3000	35	25	(-2,2)	(1,3)	0.005	0.8	0.05	0.05	0.05	$\uparrow \text{'sin'}, \uparrow \text{'cos'}, \downarrow \text{'div'}$
f_2	6000	70	25	(-2,2)	(1,2)	0.0085	0.7	0.05	0.1	0.1	$\uparrow \text{'sin'}, \uparrow \text{'cos'}, \downarrow \text{'div'}$
f_3	6000	70	25	(-2,2)	(2,4)	0.009	0.7	0.05	0.1	0.1	$\uparrow \text{'sin'}, \uparrow \text{'cos'}, \downarrow \text{'div'}$
f_4	8000	70	45	(-2,2)	(2,5)	0.0095	0.7	0.05	0.1	0.1	$\uparrow \text{'sin'}, \uparrow \text{'cos'}, \downarrow \text{'div'}$
f_5	8000	100	35	(-4,4)	(2,5)	0.011	0.7	0.05	0.1	0.05	$\uparrow \text{'sin'}, \uparrow \text{'cos'}$

Tabla B.11: Tabla de los parámetros asociados a las soluciones trigonométricas de la Tabla[4.11].

f_i	Pob	Gen	Torn	Ran	Prof	PC	Cros	Sub	Hoi	Poin	Base
$f_{\Lambda\text{CDM}}$	–	–	–	–	–	–	–	–	–	–	↓div
f_1	2000	50	200	(-35,35)	(2,3)	(0.0025)	0.7	0.05	0.1	0.1	↓div
f_2	2000	80	200	(-35,35)	(2,3)	(0.002)	0.6	0.1	0.1	0.1	↓div
f_3	2000	150	200	(-35,35)	(2,3)	(0.002)	0.6	0.1	0.1	0.1	↓div
f_4	2000	200	200	(-5,5)	(2,3)	0.0025	0.7	0.05	0.1	0.1	0.1↓div

Tabla B.12: Tabla de los parámetros asociados a las expresiones obtenidas para la función de Hubble de la Tabla[5.1].

f_i	Pob	Gen	Torn	Ran	Prof	PC	Cros	Sub	Hoi	Poin	Base
f_1	8000	40	65	(-1,1)	(2,6)	0.005	0.7	0.0	0.1	0.1	↑sin,↑cos,↓div
f_2	8000	50	50	(-2,2)	(2,4)	0.0002	0.7	0.0	0.1	0.1	↑sin,↓div
f_3	8000	35	35	(-2,2)	(2,6)	0.00009	0.7	0.0	0.1	0.1	↑sin,↓div
f_4	8000	100	35	(-1,1)	(2,6)	0.00009	0.7	0.05	0.1	0.1	↑sin,↓div

Tabla B.13: Tabla de los parámetros asociados a las expresiones obtenidas para la densidad de energía de la Tabla[5.2].

f_i	Pob	Gen	Torn	Ran	Prof	PC	Cros	Sub	Hoi	Poin	Base
f_1	4000	35	50	(-1,1)	(2,6)	0.0001	0.7	0.1	0.05	0.1	↑sin,↑cos,↓div
f_2	6000	50	50	(-1,1)	(2,6)	0.0001	0.7	0.1	0.05	0.1	↑sin,↓div
f_3	4000	45	45	(-1,1)	(2,6)	0.00025	0.7	0.0	0.1	0.1	↑sin,↑cos,↓div
f_4	6000	60	50	(-2,2)	(2,6)	0.0001	0.7	0.1	0.05	0.1	↑sin,↑cos,↓div
f_5	8000	40	50	(-2,2)	(2,6)	0.0001	0.7	0.1	0.05	0.1	↑sin,↑cos,↓div

Tabla B.14: Tabla de los parámetros asociados a las expresiones obtenidas para las ecuaciones de estado de la Tabla[5.3].

Bibliografía

- [1] Rizo Franco, S.(2023) .*Funciones de correlación de dos puntos con algoritmos de agrupamientos*. [Tesis de Licenciatura] UNAM.
- [2] Toivo Jaakkola. *The cosmological principle: theoretical and empirical foundations*. Apeiron: Studies in Infinite Nature, 4, 8-31 (1989).
- [3] Liddle, A. (2003). *An Introduction to Modern Cosmology*. Wiley-Blackwell.
- [4] F. K. Hansen, A. J. Banday, K. M. Górski, *Testing the cosmological principle of isotropy: local power-spectrum estimates of the WMAP data*, Monthly Notices of the Royal Astronomical Society, 354(3), 641-665 (2004). <https://doi.org/10.1111/j.1365-2966.2004.08229.x>
- [5] Pavan Kumar Aluri, Paolo Cea, Pravabati Chingangbam, Ming-Chung Chu, Roger G. Clowes, Damien Hutsemékers, Joby P. Kochappan, Alexia M. Lopez, Lang Liu, Niels C. M. Martens, C. J. A. P. Martins, Konstantinos Migkas, Eoin Ó Colgáin, Pratyush Pranav, Lior Shamir, Ashok K. Singal, M. M. Sheikh-Jabbari, Jenny Wagner, Shao-Jiang Wang, David L. Wiltshire, Shek Yeung, Lu Yin, and Wen Zhao. (2023). Is the observable Universe consistent with the cosmological principle? *Classical and Quantum Gravity*, 40(9), 094001. <https://doi.org/10.1088/1361-6382/acbefc>
- [6] Einstein, A. (1905). *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*. *Annalen der Physik*, 322, 132-148.
- [7] Einstein, A. (1905). *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*. *Annalen der Physik*, 322(8), 549-560.
- [8] Einstein, A. (1905). *De la Electrodinámica de los Cuerpos en Movimiento*. *Annalen der Physik*, 17, 891-921.

- [9] Einstein, A. (1923). *Does the Inertia of a Body Depend Upon Its Energy Content?* En: *The Principle of Relativity* (traducido por George Barker Jeffery). Londres: Methuen and Company, Ltd.
- [10] Bonder, Y., & Okon, E.. (2018). Los principios de la relatividad: una introducción pedagógica. *Revista mexicana de física E*, 64(1), 87-91. Recuperado en 15 de julio de 2024, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-35422018000100087&lng=es&tlng=es.
- [11] Ramos-Sánchez, S. (2018). *Geometría en Relatividad* En: *Relatividad para futuros físicos*. Las prensas de Ciencias.
- [12] Slipher, V. M. (1913). “The Radial Velocity of the Andromeda Nebula”. *Observatory Bulletin*, 1, 56-57.
- [13] Slipher, V. M. (1915). “Spectrographic observations of nebulae”. *Popular Astronomy*, 23, 21-24.
- [14] Slipher, V. M. (1917). “A spectrographic investigation of spiral nebulae”. *American Philosophical Society Proceedings*, 56, 403-409.
- [15] International Astronomical Union. (2018). Resolution B4 on a suggested renaming of the Hubble Law. <https://pressbooks.bccampus.ca/astronomy1105/chapter/26-5-the-expanding-universe/>.
- [16] Hubble, E. (1929). *A relation between distance and radial velocity among extragalactic nebulae*. *Proceedings of the National Academy of Sciences*, 15(3), 168-173. <https://doi.org/10.1073/pnas.15.3.168>
- [17] V. Busti, C. Clarkson, and M. Seikel. The value of h_0 from gaussian processes. *Proceedings of the International Astronomical Union*, 10(S306), 25-27. doi:10.1017/S1743921314013751
- [18] V. Busti, C. Clarkson, and M. Seikel. Evidence for a lower value for h_0 from cosmic chronometers data?. *Monthly Notices of the Royal Astronomical Society: Letters*, 441,02 2014.
- [19] Luković, V. V., D’Agostino, R., & Vittorio, N. (2016). Is there a concordance value for (H_0)? *Astronomy & Astrophysics*, 595, A1091
- [20] Bernal J.L., Verde L., Riess. A.G. (2016) The trouble with H_0 . *Journal of Cosmology and Astroparticle Physics*, 2016(10), 019. <https://doi.org/10.1088/1475-7516/2016/10/019>

- [21] Navarro Martínez, R. (2017). *Ecuación de campo. Introducción a la cosmología*. Trabajo de fin de grado. Facultad de Matemáticas, Universidad de Murcia, España.
- [22] Chacon Lavanderos, J.(2018) *.Modelos de energía oscura: Una perspectiva numérica*. [Tesis de Licenciatura] IPN. Carroll S. M. (2004) *Spacetime and Geometry: An Introduction to General Relativity*. Addison Wesley.
- [23] Padilla, L. E.; Tellez, L. O.; Escamilla, L. A.; Vazquez, J. A. (2021). Cosmological Parameter Inference with Bayesian Statistics. *Universe*, 7, 213. <https://doi.org/10.3390/universe7070213>
- [24] Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. *Astrophysical Journal*, 86, 217.
- [25] Rubin, V. C. and Ford, W. Kent, Jr. (1970). Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal*, 159:379.
- [26] Rodríguez, L. (2010). Materia oscura, energía oscura.
- [27] Vázquez-González, A., & Matos, T. (2008). La materia oscura del universo: retos y perspectivas. *Revista mexicana de física E*, 54(2), 193-202. Recuperado en 18 de julio de 2024, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-35422008000200012&lng=es&tlng=es.
- [28] E. Battaner and E. Florido, *The rotation curve of spiral galaxies and its cosmological implications*, *Fundamentals of Cosmic Physics*, vol. 21, pp. 1-154, 2000. <https://doi.org/10.48550/arXiv.astro-ph/0010475>
- [29] Peebles, P. J. E. and Ratra, B. (2003). The Cosmological Constant and Dark Energy. *Reviews of Modern Physics*, 75:559-606.
- [30] Lahav, O. and Liddle, A. R. (2019). The cosmological parameters. *The Review of Particle Physics*, 2019.
- [31] Aloni, D., Berlin, A., Joseph, M., Schmaltz, M., and Weiner, N. (2021). A Step in Understanding the Hubble Tension.
- [32] Chevallier, M. and Polarski, D. (2001). Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D*, 10, 213.
- [33] de Putter, R. and Linder, E. V. (2008). Calibrating Dark Energy. *Journal of Cosmology and Astroparticle Physics*, 0810:042.

- [34] Z. Zhai, M. Blanton, A. Slosar y J. Tinker. *An Evaluation of Cosmological Models from the Expansion and Growth of Structure Measurements*.
- [35] Morales Sanchez, A.(2024) .*Algoritmos genéticos y sus aplicaciones a la cosmología*. [Tesis de Licenciatura] UNAM. *The Astrophysical Journal*, 850(2):183, 2017.
- [36] Jiménez, R., & Loeb, A. (2002). *Constraining Cosmological Parameters Based on Relative Galaxy Ages*. *The Astrophysical Journal*, 573(1), 37–42.
- [37] O. Farooq, F. R. Madiyar, S. Crandall, and B. Ratra. *Hubble parameter measurement constraints on the redshift of the deceleration-acceleration transition, dynamical dark energy, and space curvature*. *The Astrophysical Journal*, 835(1):26, January 2017.
- [38] Riquelme Medina, I., & Luna, J. (2014). Revisión de los Algoritmos Bioinspirados. En *Memorias del XX Congreso Argentino de Ciencias de la Computación* (pp. 1-10). Córdoba, Argentina.
- [39] Liu, J., & Lampinen, J. (2005). A Fuzzy Adaptive Differential Evolution Algorithm. *Soft Computing*, 9(6), 448–462. <https://doi.org/10.1007/s00500-004-0363-x>
- [40] Man, K. F., Tang, K. S., & Kwong, S. (1996). Genetic algorithms: concepts and applications [in engineering design]. *IEEE Transactions on Industrial Electronics*, 43(5), 519-534. <https://doi.org/10.1109/41.538609>
- [41] Michael A. Covington. *CSCI/ARTI 4540/6540: First Lecture on Symbolic Programming and LISP*. University of Georgia, 2010. Archivado desde el original (PDF) el 07-03-2012. Recuperado del 05-06-2024. Disponible en : <https://web.archive.org/web/20120307124013/http://www.ai.uga.edu/mc/LispNotesFirstLectureOnSymbolicProgramming.pdf>
- [42] Maurizio Gabrielli & Simone Martini. *The functional paradigm*. En: *Programming Languages: Principles and Paradigms* . Springer, 2010.
- [43] C. L. Hamblin, *Translation to and from Polish notation*, *Computer Journal*, vol. 5, pp. 210–213, 1962.
- [44] Aguilar, L. J., & Martínez, I. Z. (2008). *Estructuras de datos en Java (Capítulo 13: Árboles. Árboles binarios y árboles ordenados)*. McGraw-Hill
- [45] Spivak, M. (2021). *Cálculo en variedades*. Reverte.

- [46] Museo de Paleontología de la Universidad de California. (s.f.). *Evolution and Paleontology in the Ancient World*. Recuperado el 13-06-2024 de <https://ucmp.berkeley.edu/history/ancient.html>
- [47] Torrey, H. B., & Felin, F. (1937). *Was Aristotle an Evolutionist?*. The Quarterly Review of Biology, 12(1), 1-18. doi:10.1086/394520.
- [48] Parrilla Martínez, D. (2016). *La estructura analógica de la 'quinta vía' de Santo Tomás de Aquino / The analogical structure of the 'Fifth Way' of Saint Thomas Aquinas*. Universidad Católica de Murcia (UCAM).
- [49] Amundson, R. (2005). *The Changing Role of the Embryo in Evolutionary Thought: Roots of Evo-Devo*. Cambridge University Press.
- [50] Coleman, W. (1971). *Biology in the nineteenth century: problems of form, functions and transformation*. The Cambridge History of Science Series. London.
- [51] Ruiz, R. (2009). *Evolución*. Universidad Nacional Autónoma de México, Instituto de Investigaciones Sociales. Recuperado de https://conceptos.sociales.unam.mx/conceptos_final/428trabajo.pdf
- [52] Lamas, S. G. (2023). *Formando y deformando conceptos científicos: el caso de la herencia de caracteres adquiridos*. Philosophy & History of Biology/Filosofia e História da Biologia, 18(2).
- [53] Weismann, A. (1891). *Essays upon heredity and kindred biological problems*. 2nd ed. Vol. 2. Poulton, E., Schonland, S., & Shipley, A. E. (Eds.). Oxford: Clarendon Press.
- [54] Lamarck, J. B. (1809). *Filosofía Zoológica*. Trad. Nuria Vidal Díaz. Barcelona: Editorial Mateu, 1971.
- [55] Darwin Correspondence Project. (1831). Letter 105 — Henslow, J. S. to Darwin, C. R., 24 Aug 1831. Consultado el 17-06-2024.
- [56] Darwin, C. (1839). *Journal of Researches into the Geology and Natural History of the Various Countries visited by H.M.S. Beagle*. Ward, Lock and Bowden, 1894.
- [57] Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (1st ed.). Londres: John Murray. Consultado el 17-06-2024. Disponible en: <https://darwin-online.org.uk/content/frameset?itemID=F373&viewtype=text&pageseq=20>

- [58] Rosario Rodríguez Arnaiz, América Castañeda Sortibrán y María Guadalupe Ordaz Téllez. *Conceptos Básicos de Genética*. Universidad Nacional Autónoma de México, Ciudad de México, México, 1^a edición electrónica, 2016.
- [59] Clelia De La Peña Seaman, *La epigenética y la reivindicación de Lamarck*, Desde el Herbario CICY, vol. 10, pp. 132-136, 2018.
- [60] T. Dobzhansky. *Genetics and the Origin of Species*. Columbia University Press, New York, USA, 3rd edition, 1951.
- [61] Ayala, F. J. (1994). *La naturaleza inacabada. Ensayos en torno a la evolución*. Barcelona: Salvat Editores, S.A.
- [62] Koza, John R. *Genetic Programming: On Programming Computer by Means of Natural Selection and Genetics*. Cambridge, MA: The MIT Press, 1992.
- [63] W.B. Langdon, R. Poli, N.F. McPhee, J.R. Koza, “Genetic Programming: An Introduction and Tutorial, with a Survey of Techniques and Applications,” in *Computational Intelligence: A Compendium*, J. Fulcher, L.C. Jain, Eds., vol. 115, Studies in Computational Intelligence, Springer, Berlin, Heidelberg, 2008.
- [64] M.F. Korn, “Genetic Programming Symbolic Classification: A Study,” in *Genetic Programming Theory and Practice XV*, W. Banzhaf, R. Olson, W. Tozier, R. Riolo, Eds., Genetic and Evolutionary Computation, Springer, Cham, 2018.
- [65] T. Stephens, “gplearn Documentation,” *gplearn: Genetic Programming in Python*, 2024. Available:<https://gplearn.readthedocs.io/en/stable/>
- [66] W. B. Langdon and P. Nordin, “Evolving Hand-Eye Coordination for a Humanoid Robot with Machine Code Genetic Programming,” in *Genetic Programming*, 2001, pp. 313–324.
- [67] W. B. Langdon and B. F. Buxton, “Genetic programming for mining DNA chip data from cancer patients,” *Genetic Programming and Evolvable Machines*, vol. 5, no. 3, pp. 251–257, 2004.
- [68] M. W. Khan and M. Alam, “A survey of application: Genomics and genetic programming, a new frontier,” *Genomics*, vol. 100, no. 2, pp. 65–71, 2012.
- [69] Q. Fan, Y. Bi, B. Xue, and M. Zhang, “Evolving Effective Ensembles for Image Classification Using Multi-objective Multi-tree Genetic Programming,” in H. Aziz, D. Corrêa, and T. French, Eds., *AI 2022: Advances in Artificial Intelligence*, Lecture Notes in Computer Science, vol. 13728, Springer, Cham, 2022.

- [70] M. Cebrian, M. Alfonseca, and A. Ortega, “Automatic generation of benchmarks for plagiarism detection tools using grammatical evolution,” in *Proc. of GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, vol. 2, pp. 2253–2253, 2007.
- [71] M. T. Ahvanooey, Q. Li, M. Wu, and S. Wang, “A survey of genetic programming and its applications,” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 13, no. 4, pp. 1765–1794, 2019.
- [72] Rojas, R. (1998). “How to Make Zuse’s Z3 a Universal Computer.” *IEEE Ann. Hist. Comput.*, 20, 51-54.
- [73] Schmidt, M. D. & Lipson, H. (2006). Co-evolving fitness predictors for accelerating and reducing evaluations. En *Genetic Programming Theory and Practice IV*. Serie *Genetic and Evolutionary Computation*, vol. 5. Springer.
- [74] Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79–82.
- [75] Poli, R. (1996). Parallel distributed genetic programming. Birmingham, UK: University of Birmingham, Cognitive Science Research Centre.
- [76] Fernández, F., Galeano, G., & Gómez, J. A. (2002). Comparing Synchronous and Asynchronous Parallel and Distributed Genetic Programming Models. In *Genetic Programming (EuroGP 2002)* (pp. 326–335). Springer Verlag.
- [77] Tanev, I., Uozumi, T., & Ono, K. (2001). Scalable architecture for parallel distributed implementation of genetic programming on network of workstations. *Journal of Systems Architecture*, 47(7), 557–572.
- [78] Sandewall, E. (1978). Programming in an Interactive Environment: The “Lisp” Experience. *ACM Computing Surveys (CSUR)*, 10(1), 35–71.
- [79] Beadle, L., & Johnson, C. G. (2009). Semantic analysis of program initialisation in genetic programming. *Genetic Programming and Evolvable Machines*, 10(3), 307–337.
- [80] Affenzeller, M., Winkler, S., Wagner, S., & Beham, A. (2009). *Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications*.
- [81] W. M. Spears y V. Anand, *A study of crossover operators in genetic programming*, en *Methodologies for Intelligent Systems*, pp. 409–418, Springer, 1991.

- [82] K. E. Kinnear, Jr., *Alternatives in Automatic Function Definition: A Comparison of Performance*, en *Advances in Genetic Programming*, K. E. Kinnear, Jr. (editor), MIT Press, pp. 119–141, 1994.
- [83] Harries, K., & Smith, P. (1997). Exploring alternative operators and search strategies in genetic programming. *Genetic Programming*, 97, 147-155.
- [84] Langdon, W. B., & Poli, R. (2002). *Foundations of Genetic Programming*. Springer Verlag, Berlin Heidelberg New York.
- [85] Vazquez Uribe, J.(2024) .*Programación genética aplicada a la cosmología observacional..* [Tesis de Licenciatura] UNAM.
- [86] S. Nesseris and J. Garcia-Bellido, “A new perspective on Dark Energy modeling via Genetic Algorithms,” *JCAP*, vol. 11, p. 033, 2012. doi: [10.48550/arXiv.1205.0364](https://doi.org/10.48550/arXiv.1205.0364).
- [87] S. Nesseris and A. Shafieloo, “A model independent null test on the cosmological constant,” *Mon. Not. Roy. Astron. Soc.*, vol. 408, pp. 1879-1885, 2010. doi: [10.48550/arXiv.1004.0960](https://doi.org/10.48550/arXiv.1004.0960).
- [88] Escamilla, Luis A. and Vazquez, J. Alberto, *Model selection applied to non-parametric reconstructions of the Dark Energy*, arXiv preprint arXiv:2111.10457, 2021.