



Departamento de Matemáticas, Facultad de Ciencias
Universidad Autónoma de Madrid

Aplicaciones del aprendizaje de máquinas al modelo Λ CDM

TRABAJO DE FIN DE GRADO

Grado en Matemáticas

Autor: Eduardo González García

Tutor: Ernesto Nungesser

Curso 2022-2023

Resumen

El modelo Λ CDM es el modelo cosmológico estándar usado para describir la evolución del universo a gran escala. Este se cimenta en la teoría de la relatividad general y se desarrolla al aplicar las ecuaciones de Einstein bajo ciertas simplificaciones. A partir de soluciones analíticas y numéricas del modelo exploraremos las predicciones que este describe para distintos tipos de universos. A su vez, estimaremos los parámetros que modelizan nuestro universo usando mediciones modernas de la variable de Hubble mediante tres métodos distintos: el cálculo de la mediana, métodos de Cadenas de Markov Monte Carlo, y la introducción de redes neuronales. Por último, compararemos nuestros resultados con los del experimento Planck y observaremos las discrepancias que están surgiendo en el modelo Λ CDM entre el estudio de supernovas y el fondo cósmico de microondas.

Abstract

The Λ CDM model is the standard cosmological model used to describe the evolution of the universe on a large scale. It is based on the theory of general relativity and is developed by applying the Einstein equations under certain simplifications. By analyzing analytical and numerical solutions of the model, we will explore the predictions it makes for different types of universes. Additionally, we will estimate the parameters that model our universe using modern measurements of the Hubble variable through three different methods: median calculation, Markov Chain Monte Carlo methods, and the introduction of neural networks. Finally, we will compare our results with those of the Planck experiment and observe the discrepancies that arise in the Λ CDM model between the study of supernovae and the cosmic microwave background.

Índice general

1	Introducción	1
2	Derivación del modelo ΛCDM	3
2.1	Métrica	3
2.2	Ecuaciones de campo de Einstein	4
2.2.1	Cálculo del tensor y escalar de Ricci	4
2.2.2	Primera ecuación de Friedmann	5
2.2.3	Segunda ecuación de Friedmann	5
2.2.4	Evolución de la densidad de energía	6
2.2.5	Ecuación de estado lineal de un fluido perfecto	7
2.3	Modelo Λ CDM	7
2.3.1	Parámetros	7
2.3.2	Ecuación de Friedmann con parámetros Λ CDM	8
2.3.3	Segunda ecuación de Friedmann con parámetros Λ CDM	9
3	Análisis cualitativo de soluciones FLRW y predicciones del modelo ΛCDM	10
3.1	Existencia y unicidad de soluciones	10
3.2	Resoluciones analíticas	10
3.2.1	Universos espacialmente planos de un solo componente	11
3.2.2	Universos con materia y curvatura	12
3.2.3	Universos espacialmente planos con materia y constante cosmológica Λ	16
3.3	Resolución numérica	18
3.3.1	Autovalores del problema linealizado	18
3.3.2	Método de Bogacki-Shampine	19
3.3.3	Resultados	20
4	Estimación de parámetros del modelo ΛCDM	21
4.1	Corrimiento al rojo	21
4.2	Mediciones de $H(z)$	22
4.3	Estimación de la mediana	23
4.4	Estimación mediante el algoritmo de Metropolis-Hastings	24
4.4.1	Cadena de Markov	25
4.4.2	Descripción del algoritmo de Metropolis-Hastings	25
4.4.3	Convergencia del algoritmo de Metropolis-Hastings	26
4.4.4	Aplicación a las mediciones de $H(z)$	27
4.4.5	Resultados	28

4.5	Estimación mediante Redes Neuronales	32
4.5.1	Propagación hacia delante en una red neuronal densa	32
4.5.2	Propagación hacia atrás en una red neuronal densa	32
4.5.3	Descenso de gradiente	33
4.5.4	Proceso de entrenamiento	35
4.5.5	Hiperparámetros de la red neuronal	36
4.5.6	Resultados	36
5	Conclusión	38
	Bibliografía	39
A	Mediciones de $H(z)$	41

CAPÍTULO 1

Introducción

La base teórica de la cosmología es la teoría de la relatividad general [6, 7] desarrollada en 1915 donde se postulan las ecuaciones de Einstein. Dos años más tarde, en 1917, Einstein escribe un artículo [8] en el que describe el universo en su conjunto añadiendo la constante cosmológica Λ , una constante incluida para producir en sus soluciones un universo estático. Sin embargo, esto demostró ser un gran error, ya que en 1929 Hubble publica sus resultados [13] que demuestran la expansión de nuestro universo a través de la observación de diversas galaxias. Con el fin de explicar un universo en evolución se desarrolló, entre otras, la teoría del Big Bang, la cual ganó mucha aceptación cuando fue detectada la radiación cósmica de fondo en 1965 por Penzias y Wilson [19].

A partir de nuestras mediciones, a grandes escalas el universo parece ser homogéneo (igual en todo punto) e isótropo (igual en toda dirección). Tomando estas suposiciones surge como solución de las ecuaciones de Einstein el universo Friedmann-Lemaître-Robertson-Walker (FLRW) [10, 11, 16, 17, 22, 23, 24, 26]. Dando por hecho además que el universo está modelado por un fluido perfecto (que carece de viscosidad) con una ecuación de estado lineal, se puede describir la evolución de un universo FLRW a través del modelo Λ CDM, que es una parametrización de un modelo cosmológico compuesto de: materia oscura fría [cold dark matter (CDM)], es decir, un nuevo tipo de materia detectable a través de sus interacciones gravitatorias, materia bariónica u ordinaria, radiación y una constante cosmológica Λ asociada a la energía oscura.

Este trabajo contará con una estructura dividida en tres partes. En la primera, usando la métrica de un universo FLRW y las ecuaciones de Einstein, derivaremos las ecuaciones de Friedmann. Con ellas podremos describir la evolución del tamaño del universo y la variable de Hubble. Partiendo de modelizar la materia y radiación mediante un fluido perfecto con ecuación de estado lineal, obtendremos las ecuaciones diferenciales que definen el modelo Λ CDM. Estas ecuaciones dependerán de 5 parámetros: la densidad de radiación y materia (tanto oscura como ordinaria), la constante de Hubble, la curvatura y la constante cosmológica Λ .

En la segunda parte exploraremos las soluciones del modelo Λ CDM, que nos permitirán entender la evolución de distintos tipos de universos según su composición y geometría. Comenzaremos calculando soluciones analíticas de tres familias de universos con restricciones sobre sus parámetros: universos de un solo componente, universos

de materia con una cierta curvatura y universos planos de materia con constante cosmológica Λ . Después, pasaremos a estudiar soluciones numéricas para universos con cualquier combinación de parámetros. Para ello, emplearemos el método de Bogacki-Shampine [2], que hace uso de dos métodos de Runge-Kutta (uno explícito y uno implícito), para hacer predicciones con un buen orden de convergencia, pero controlando su error manteniendo una estabilidad razonable. Aplicando este método a distintas combinaciones de parámetros, podremos observar las diferentes tendencias que predice el modelo Λ CDM. Utilizando en concreto los valores de los parámetros obtenidos por el experimento Planck [20], tendremos una gráfica descriptiva del proceso expansivo de nuestro universo.

En la tercera parte estimaremos los parámetros de nuestro universo usando mediciones actuales de la variable de Hubble. Para ello probaremos tres métodos distintos. El primero será una implementación simple del cálculo de la mediana [18]: puesto que tendremos una fórmula que relacione directamente los parámetros y las mediciones, despejando un sencillo sistema de ecuaciones obtendremos el valor de los parámetros que ajustan la fórmula a tres puntos, y, calculando entonces los valores para todas las combinaciones de puntos, podremos establecer intervalos de confianza para las medianas de los resultados. El segundo método será una implementación clásica de métodos de Cadenas de Markov Monte Carlo, específicamente, el algoritmo de Metropolis-Hastings [21] (véase capítulo 7): extendiendo la idea de Cadena de Markov a un espacio continuo y usando las mediciones con sus errores asociados, podremos definir una cadena cuya distribución límite sea la distribución a posteriori de nuestros parámetros, que reconstruiremos al tomar muestras de nuestro recorrido por la cadena. El tercer método será una implementación innovadora de redes neuronales [27]: a partir de la fórmula de la variable de Hubble crearemos mediciones artificiales para parámetros conocidos, después, entrenaremos redes neuronales sobre estos datos para que aprendan la conexión intrínseca entre mediciones y parámetros, por último, aplicaremos dichas redes sobre las mediciones reales para obtener estimaciones de los parámetros de nuestro universo.

Para terminar, compararemos los resultados obtenidos en este trabajo con los del experimento Planck [20] y presentaremos una discrepancia en el valor de la densidad de materia de nuestro universo. Esto supondrá un indicio más de los problemas que están surgiendo con el modelo Λ CDM, ya que, conforme nuestras predicciones se vuelven más precisas, más discrepancias están apareciendo entre los resultados del estudio de supernovas y el estudio del fondo cósmico de microondas.

El código usado en este trabajo se puede consultar en: <https://github.com/Darmator/Applications-of-Machine-Learning-to-the-LambdaCDM-model>.

CAPÍTULO 2

Derivación del modelo Λ CDM

2.1. Métrica

Si consideramos un universo isótropo la geometría tiene que ser la misma en cualquier punto, por tanto la curvatura espacial debe ser constante. Puede haber curvatura positiva, negativa o nula que corresponderían a una 3-esfera, a un hiperboloide o a la geometría euclídea respectivamente. Lo que no queda fijado es el ‘tamaño’ del universo, que no puede depender de cada punto, pero sí del tiempo. Vamos a utilizar las unidades naturales, donde ciertas constantes universales, como la constante gravitatoria G o la velocidad de la luz c , serán iguales a 1. La variable adimensional relacionada con el tamaño se llama factor de escala y se denota por a . La métrica que describe un universo FLRW se puede expresar como (véase página 14 de [9]):

$$(2.1) \quad ds^2 = -dt^2 + a(t)^2 d\Omega_k,$$

donde $d\Omega_k$ describe una geometría de curvatura constante:

$$d\Omega_k = \left(\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right),$$

conocida como la métrica Robertson-Walker, donde k es un parámetro para describir las tres geometrías mencionadas con $k \in \{1, -1, 0\}$, y define un tensor métrico:

$$(2.2) \quad g_{\mu\nu} = \text{diag} \left(-1, \frac{a^2}{1 - kr^2}, a^2 r^2, a^2 r^2 \sin^2 \theta \right) \quad \text{con } \mu, \nu \in \{0, 1, 2, 3\},$$

donde 0 representa una coordenada temporal y $\{1, 2, 3\}$ coordenadas espaciales. En particular, en coordenadas esféricas:

$$(2.3) \quad (x^0, x^1, x^2, x^3) = (t, r, \theta, \phi).$$

Podemos asumir que a sea positivo, por lo que siempre lo será, dado que a no puede cambiar de signo sin que eso conlleve a una singularidad debido a (2.1).

2.2. Ecuaciones de campo de Einstein

Las ecuaciones de campo de Einstein relacionan la curvatura del espacio-tiempo con la energía y momento de la materia. Matemáticamente quedan escritas de la manera siguiente (véase ecuación (225) en [9]):

$$(2.4) \quad G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}.$$

En esta ecuación tensorial:

- Λ es la constante cosmológica.
- $T_{\mu\nu}$ es el tensor de energía-momento que describe la materia.
- $G_{\mu\nu}$ es el tensor de Einstein, que viene definido por el tensor y escalar de Ricci (véase ecuaciones (194) y (221) en [9]):

$$(2.5) \quad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \quad R = R_{\mu\nu}g^{\mu\nu},$$

donde usamos la convención de Einstein, por la que se suma sobre los índices repetidos:

$$x_\alpha v^\alpha := \sum_{\alpha=0}^3 x_\alpha v^\alpha.$$

Además, dado que la métrica g es diagonal, su notación con los índices superiores será el inverso de sus elementos diagonales:

$$g^{\mu\nu} = \text{diag} \left(-1, \frac{1 - kr^2}{a^2}, \frac{1}{a^2 r^2}, \frac{1}{a^2 r^2 \sin^2 \theta} \right).$$

Por último, considerando el universo isotrópico como un fluido perfecto, el tensor $T_{\mu\nu}$ queda descrito a partir del tensor métrico como:

$$(2.6) \quad T_{\mu\nu} = \text{diag} (\rho, Pg_{11}, Pg_{22}, Pg_{33}),$$

donde $\rho(t)$ es la densidad de energía y $P(t)$ la presión.

2.2.1. Cálculo del tensor y escalar de Ricci

El tensor y escalar de Ricci vienen definidos por los símbolos de Christoffel (véase ecuación (219) en [9]):

$$(2.7) \quad R_{\alpha\beta} = R^\delta_{\alpha\delta\beta} = \Gamma^\delta_{\beta\alpha,\delta} - \Gamma^\delta_{\delta\alpha,\beta} + \Gamma^\epsilon_{\beta\alpha}\Gamma^\delta_{\delta\epsilon} - \Gamma^\epsilon_{\delta\alpha}\Gamma^\delta_{\beta\epsilon},$$

donde el índice tras la coma indica tomar la derivada parcial, es decir:

$$\Gamma^\delta_{\beta\alpha,\delta} = \frac{\partial}{\partial x^\delta} \Gamma^\delta_{\beta\alpha}.$$

A su vez, los símbolos de Christoffel quedan determinados por la métrica (2.2) (véase ecuación (216) en [9]):

$$(2.8) \quad \Gamma_{\beta\gamma}^{\alpha} = g^{\alpha\delta} \Gamma_{\delta\beta\gamma}, \quad \Gamma_{\delta\beta\gamma} = \frac{1}{2} [g_{\delta\gamma,\beta} + g_{\beta\delta,\gamma} - g_{\beta\gamma,\delta}].$$

De modo que, aplicando la fórmula (2.8), los símbolos de Christoffel quedan calculados como:

$$\begin{aligned} \Gamma_{\mu\nu}^0 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{a\dot{a}}{1-kr^2} & 0 & 0 \\ 0 & 0 & a\dot{a}r^2 & 0 \\ 0 & 0 & 0 & a\dot{a}r^2 \sin^2 \theta \end{pmatrix}, & \Gamma_{\mu\nu}^2 &= \begin{pmatrix} 0 & 0 & \frac{\dot{a}}{r} & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ \frac{\dot{a}}{a} & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin \theta \cos \theta \end{pmatrix}, \\ \Gamma_{\mu\nu}^1 &= \begin{pmatrix} 0 & \frac{\dot{a}}{a} & 0 & 0 \\ \frac{\dot{a}}{a} & \frac{kr}{1-kr^2} & 0 & 0 \\ 0 & 0 & -r(1-kr^2) & 0 \\ 0 & 0 & 0 & -r \sin^2 \theta (1-kr^2) \end{pmatrix}, & \Gamma_{\mu\nu}^3 &= \begin{pmatrix} 0 & 0 & 0 & \frac{\dot{a}}{r} \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot \theta \\ \frac{\dot{a}}{a} & \frac{1}{r} & \cot \theta & 0 \end{pmatrix}. \end{aligned}$$

Aplicando ahora la formula (2.7) se obtiene el tensor y escalar de Ricci:

$$\begin{aligned} R_{\mu\nu} &= \text{diag} \left(-3\frac{\ddot{a}}{a}, \frac{a\ddot{a} + 2\dot{a}^2 + 2k}{1-kr^2}, (a\ddot{a} + 2\dot{a}^2 + 2k)r^2, (a\ddot{a} + 2\dot{a}^2 + 2k)r^2 \sin^2 \theta \right), \\ R &= 6\frac{a\ddot{a} + \dot{a}^2 + k}{a^2}. \end{aligned}$$

2.2.2. Primera ecuación de Friedmann

La primera ecuación de Friedmann describe la evolución del factor de escala del universo en términos de su densidad, curvatura y constante cosmológica. Esta se deriva a partir de las ecuaciones de campo de Einstein (2.4) para $\mu = 0, \nu = 0$:

$$\begin{aligned} G_{00} + \Lambda g_{00} &= 8\pi T_{00} \Rightarrow R_{00} - \frac{1}{2} R g_{00} + \Lambda g_{00} = 8\pi T_{00} \\ &\Rightarrow -3\frac{\ddot{a}}{a} + 3\frac{a\ddot{a} + \dot{a}^2 + k}{a^2} = 8\pi\rho + \Lambda \\ &\Rightarrow 3\frac{\dot{a}^2 + k}{a^2} = 8\pi\rho + \Lambda. \end{aligned}$$

Simplificando esta expresión se llega a la primera ecuación de Friedmann:

$$(2.9) \quad H^2 = \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3},$$

siendo $H = \frac{\dot{a}}{a}$ la magnitud conocida como variable de Hubble.

2.2.3. Segunda ecuación de Friedmann

La segunda ecuación de Friedmann describe la evolución de la variable de Hubble en términos de la densidad, presión y constante cosmológica. Esta se deriva a partir

de estudiar la traza de las ecuaciones de Einstein (2.4) generadas al contraer con el tensor $g^{\mu\nu}$:

$$\begin{aligned} G_{\mu\nu}g^{\mu\nu} &= 8\pi T_{\mu\nu}g^{\mu\nu} - \Lambda g_{\mu\nu}g^{\mu\nu} \Rightarrow R_{\mu\nu}g^{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}g^{\mu\nu} = 8\pi T_{\mu\nu}g^{\mu\nu} - \Lambda g_{\mu\nu}g^{\mu\nu} \\ &\Rightarrow R - 2R = 8\pi(-\rho + 3P) - 4\Lambda \\ &\Rightarrow R = -8\pi(-\rho + 3P) + 4\Lambda. \end{aligned}$$

Multiplicando ahora por $-\frac{1}{2}g_{\mu\nu}$:

$$-\frac{1}{2}g_{\mu\nu}R = \frac{1}{2}g_{\mu\nu}8\pi(-\rho + 3P) - 2g_{\mu\nu}\Lambda.$$

Restando la ecuación de Einstein (2.4):

$$R_{\mu\nu} = 8\pi \left(T_{\mu\nu} - \frac{1}{2}(-\rho + 3P)g_{\mu\nu} \right) + g_{\mu\nu}\Lambda.$$

Evalutando de nuevo para los índices $\mu = 0, \nu = 0$:

$$R_{00} = 8\pi \left(T_{00} - \frac{1}{2}(-\rho + 3P)g_{00} \right) + g_{00}\Lambda \Rightarrow -3\frac{\ddot{a}}{a} = 8\pi \left(\frac{1}{2}\rho + \frac{3}{2}P \right) - \Lambda.$$

Y, por último, despejando queda la segunda ecuación de Friedmann:

$$(2.10) \quad \frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3P) + \frac{\Lambda}{3}.$$

2.2.4. Evolución de la densidad de energía

A partir de la primera y segunda ecuación de Friedmann se puede deducir la evolución de la densidad de energía en el universo. Multiplicando la ecuación de Friedmann (2.9) por a^2 y derivando con respecto a t obtenemos:

$$(2.11) \quad 2\dot{a}\ddot{a} = \frac{8\pi\dot{\rho}}{3}a^2 + \frac{8\pi\rho + \Lambda}{3}2\dot{a}a.$$

Sustituyendo \ddot{a} mediante la segunda ecuación de Friedmann (2.10) se llega a:

$$\frac{2}{3}\dot{a}[-4\pi(\rho + 3P) + \Lambda]a = \frac{8\pi\dot{\rho}}{3}a^2 + \frac{8\pi\rho + \Lambda}{3}2\dot{a}a,$$

y despejando obtenemos una fórmula para la evolución de la densidad de energía:

$$(2.12) \quad \dot{\rho} = -\frac{3\dot{a}}{a}(\rho + P).$$

Esta ecuación no depende ni de k ni de Λ y se obtiene, como era de esperar, que para un universo en expansión la densidad de energía disminuye. Además, nótese que, derivando (2.9) y aplicando la igualdad (2.12), se puede recuperar (2.10). Demostrándose así que soluciones de la primera ecuación de Friedmann también serán soluciones de la segunda, si se cumple (2.12).

2.2.5. Ecuación de estado lineal de un fluido perfecto

Consideremos un fluido perfecto con ecuación de estado lineal que por razones históricas se denota como:

$$(2.13) \quad P = (\gamma - 1)\rho,$$

donde γ con $1 \leq \gamma \leq 2$ es una constante. Hay dos valores de γ que luego van a jugar un papel importante. Para materia tipo polvo, donde no hay presión, el valor de γ es 1 y para radiación el valor de γ es $\frac{4}{3}$.

Sustituyendo en la ecuación de evolución de densidad (2.12) obtenemos

$$(2.14) \quad \dot{\rho} = -\frac{3\dot{a}}{a} [\rho + (\gamma - 1)\rho],$$

$$(2.15) \quad \frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a}\gamma,$$

e integrando

$$(2.16) \quad \rho = \rho_0 a^{-3\gamma},$$

donde ρ_0 es una constante que denota la densidad en el tiempo actual ($\rho_0 = \rho(t_0)$).

Introduciendo la ecuación de estado lineal (2.13) en la segunda ecuación de Friedmann (2.10) se deduce

$$(2.17) \quad \frac{\ddot{a}}{a} = -\frac{4\pi}{3}\rho(3\gamma - 2) + \frac{\Lambda}{3}.$$

2.3. Modelo Λ CDM

2.3.1. Parámetros

El modelo Λ CDM que vamos a estudiar cuenta con 5 parámetros: $H_0, \Omega_m, \Omega_r, \Omega_k$ y Ω_Λ .

- H_0 : La constante de Hubble tiene unidades de $\frac{1}{\text{tiempo}}$ y suele expresarse en ((km/s)/Mpc). Intuitivamente, mide la velocidad por unidad de espacio a la que se separan dos puntos en la actualidad, cuanto más lejos estén, más rápido se separan:

$$H_0 = H(t_0) = \frac{\dot{a}(t_0)}{a(t_0)}, \quad t_0 \text{ el tiempo actual}$$

Dada la variable de Hubble $H(t)$ se puede obtener la densidad crítica ρ_{crit} , que es la densidad que sería necesaria para obtener curvatura $k = 0$ en la ecuación de Friedmann (2.9) asumiendo una constante cosmológica $\Lambda = 0$

$$(2.18) \quad \rho_{crit} = \frac{3H^2}{8\pi},$$

y se usará para normalizar diferentes cantidades.

- Ω_m, Ω_r : Parámetros de densidad de la materia (tanto ordinaria como oscura), y radiación respectivamente. Se definen como:

$$(2.19) \quad \Omega_i = \frac{\rho_i}{\rho_{crit}} = \frac{8\pi\rho_i}{3H^2}, \quad i \in \{b, c, r\}.$$

Nótese que para el Modelo Λ CDM estos parámetros serán Ω_m^0 y Ω_r^0 dados en términos del tiempo actual (usando las densidades y constante de Hubble actuales).

- Ω_Λ : Parámetro de densidad de la Constante Cosmológica definido como:

$$\Omega_\Lambda = \frac{\Lambda}{3H^2},$$

y, de nuevo, este parámetro será Ω_Λ^0 dado en términos actuales.

- Ω_k : Parámetro de curvatura descrito como:

$$(2.20) \quad \Omega_k = -\frac{k}{a^2 H^2}.$$

Cabe destacar que los parámetros Ω_i sobredeterminan el modelo, ya que cualquiera de ellos se puede despejar a partir del resto usando la ecuación de Friedmann (2.9). Recalcamos que se supondrá que no hay interacción entre la materia y radiación, razón por la cual podemos dividir la densidad de energía en dos partes, la de materia ρ_m y la de radiación ρ_r , dejándose (2.9) escrita como:

$$(2.21) \quad H^2 = \frac{8\pi}{3} (\rho_m + \rho_r) - \frac{k}{a^2} + \frac{\Lambda}{3}.$$

Dividiendo por $H(t)^2$

$$1 = \Omega_m + \Omega_r + \Omega_k + \Omega_\Lambda,$$

y sea

$$\Omega_0 = \Omega_m + \Omega_r + \Omega_\Lambda,$$

se llega a

$$\Omega_k = 1 - \Omega_0.$$

Lo cual demuestra que las distintas densidades de materia y energía dictaminan la curvatura del espacio. De esta forma podremos trabajar esencialmente con solo 4 parámetros.

2.3.2. Ecuación de Friedmann con parámetros Λ CDM

Fijándonos en (2.21), recordemos que las densidades ρ_m y ρ_r quedaban determinadas por (2.16):

$$\rho_m = \rho_m^0 a^{-3}, \quad \rho_r = \rho_r^0 a^{-4}.$$

Dividiendo la ecuación (2.21) por H_0^2 y aplicando la igualdad anterior tenemos:

$$\frac{H^2}{H_0^2} = \Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} + \Omega_k^0 a^{-2} + \Omega_\Lambda^0.$$

Con lo que acabamos llegando a la ecuación de Friedman con parámetros Λ CDM

$$(2.22) \quad H(t)^2 = \left(\frac{\dot{a}}{a}\right)^2 = H_0^2 (\Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} + \Omega_\Lambda^0 + (1 - \Omega_0^0) a^{-2}) .$$

Fijándonos en la ecuación se puede ver que, para factores de escala a pequeños (correspondientes a un universo temprano), la densidad de radiación Ω_r será el factor más importante, pero, pasado el tiempo una vez aumente el factor de escala a , la radiación se diluirá rápidamente, cediendo la dominancia primero a la densidad de materia Ω_m , y, más tarde, al factor Ω_Λ en tiempos largos. Esta tendencia se puede observar en el siguiente diagrama:

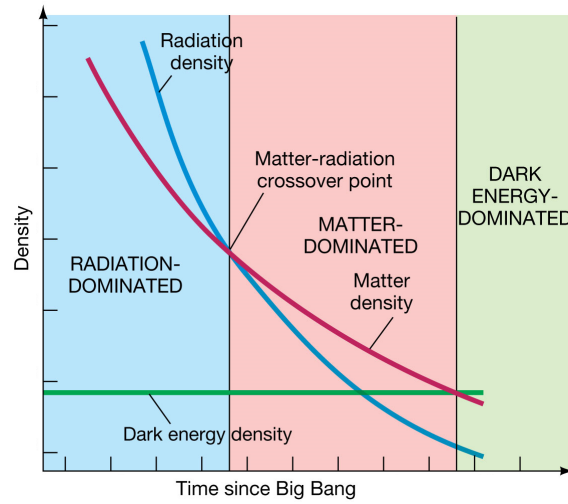


Figura 2.1

2.3.3. Segunda ecuación de Friedmann con parámetros Λ CDM

Sea

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}\rho(3\gamma - 2) + \frac{\Lambda}{3} .$$

Ajustando γ para las densidades de materia y radiación se tiene:

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}\rho_m - \frac{8\pi}{3}\rho_r + \frac{\Lambda}{3} .$$

Repitiendo el procedimiento anterior: dividiendo por H_0^2 y sustituyendo las densidades en términos de la densidad actual se obtiene:

$$\frac{\ddot{a}}{aH_0^2} = -\frac{1}{2}\Omega_m^0 a^{-3} - \Omega_r^0 a^{-4} + \Omega_\Lambda^0 ,$$

y se acaba con:

$$(2.23) \quad \frac{\ddot{a}}{a} = -H_0^2 \left(\frac{1}{2}\Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} - \Omega_\Lambda^0 \right) .$$

CAPÍTULO 3

Análisis cualitativo de soluciones FLRW y predicciones del modelo Λ CDM

Vamos a estudiar la ecuación diferencial de Friedmann:

$$(3.1) \quad \begin{cases} \dot{a} = H_0 \sqrt{\Omega_m^0 a^{-1} + \Omega_r^0 a^{-2} + \Omega_\Lambda^0 a^2 + (1 - \Omega_0^0)} \\ a(t_0) = 1 \quad t_0 \text{ el tiempo actual.} \end{cases}$$

Esta ecuación diferencial al completo es difícil de resolver analíticamente. El modelo Λ CDM requiere $\Omega_m^0, \Omega_\Lambda^0 > 0$, pero para entender las tendencias que pueden surgir en la evolución de un universo FLRW, será más sencillo estudiar diversas simplificaciones donde este requisito pueda no cumplirse.

3.1. Existencia y unicidad de soluciones

Aplicando el teorema de Picard en (3.1) se demuestra la existencia y unicidad local de soluciones, al ser la raíz cuadrada una función continua y localmente Lipschitz respecto a a (para $a \neq 0$). Más adelante veremos que, para la mayoría de combinaciones de parámetros, existirá una singularidad (distinta en cada solución) donde la solución no será diferenciable, ya que a tenderá a 0. En esos puntos singulares será donde situaremos la aparición del Big Bang, y lo usaremos para calcular la edad del universo.

3.2. Resoluciones analíticas

Por simplicidad, tomaremos la constante de Hubble como $H_0 = 67 \text{ (km/s)/Mpc}$ procedente de la medición de la radiación de fondo de microondas por el experimento Planck 2018 [20].

Vamos a explorar tres simplificaciones distintas de universos FLRW expuestas en [25] (véase capítulos 5 y 6): primero, universos espacialmente planos compuestos solo

de materia, radiación o energía oscura (nótese que el universo plano de materia es una de las primeras parametrizaciones de nuestro universo), segundo, universos compuestos solo de materia pero con curvatura positiva o negativa, y, por último, un universo espacialmente plano de materia con energía oscura, que será la simplificación que más se acerque a nuestras mediciones actuales y cumplirá el requisito de parámetros del modelo Λ CDM.

3.2.1. Universos espacialmente planos de un solo componente

3.2.1.1. Universo con materia

Tomamos un universo plano compuesto únicamente de materia. Aplicando la ecuación (3.1) con $(\Omega_k^0 = \Omega_r^0 = \Omega_\Lambda^0 = 0, \quad \Omega_m^0 = 1)$:

$$\dot{a} = H_0 a^{-\frac{1}{2}} \Rightarrow \int_{t_0}^t \dot{a} \sqrt{a} dt = \int_{t_0}^t H_0 dt \Rightarrow \frac{2}{3} a^{\frac{3}{2}} - \frac{2}{3} = H_0 (t - t_0) \Rightarrow a(t) = \left(\frac{3}{2} H_0 (t - t_0) + 1 \right)^{\frac{2}{3}}.$$

Esta solución tiene una única singularidad en $t^* = t_0 - \frac{2}{3H_0}$, ya que $a(t) \xrightarrow{t \rightarrow t^*} 0$. Situaremos en t^* la aparición del Big Bang y lo catalogaremos como el principio de este universo, con una edad de $t_0 - t^* = 9,74$ Ga (Giga años). Nótese que esta solución describirá un universo en una constante expansión que decelera.

3.2.1.2. Universo con radiación

Tomamos un universo plano compuesto únicamente por radiación. Aplicando la ecuación (3.1) con $(\Omega_k^0 = \Omega_m^0 = \Omega_\Lambda^0 = 0, \quad \Omega_r^0 = 1)$:

$$\dot{a} = H_0 a^{-1} \Rightarrow \int_{t_0}^t \dot{a} a dt = \int_{t_0}^t H_0 dt \Rightarrow \frac{1}{2} a^2 - \frac{1}{2} = H_0 (t - t_0) \Rightarrow a(t) = \sqrt{2H_0 (t - t_0) + 1}.$$

Esta solución también tiene una única singularidad en $t^* = t_0 - \frac{1}{2H_0}$, por tanto su edad será de 7,3 Ga. Este universo tendrá un proceso expansivo similar al de materia, pero con una expansión más rápida al principio y más lenta pasado un punto de inflexión, cuando se haya diluido la radiación lo suficiente.

3.2.1.3. Universo vacío dominado por Λ

Tomamos un universo plano compuesto únicamente por energía oscura. Aplicando la ecuación (3.1) con $(\Omega_k^0 = \Omega_m^0 = \Omega_r^0 = 0, \quad \Omega_\Lambda^0 = 1)$:

$$\dot{a} = H_0 a \Rightarrow \int_{t_0}^t \frac{\dot{a}}{a} dt = \int_{t_0}^t H_0 dt \Rightarrow \ln a(t) = H_0 (t - t_0) \Rightarrow a(t) = e^{H_0(t-t_0)}.$$

Este universo define una expansión exponencial sin ninguna singularidad, por tanto, sin principio ni final.

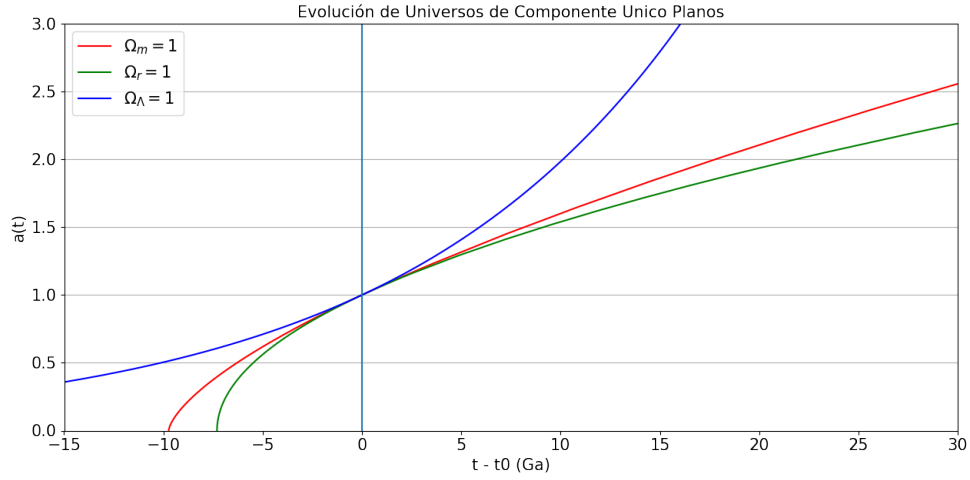


Figura 3.1

3.2.2. Universos con materia y curvatura

3.2.2.1. Universo cerrado ($k = +1$)

Puesto que el universo va a estar únicamente compuesto de materia, usando (2.20) se tiene:

$$(3.2) \quad \Omega_k^0 = -\frac{1}{H_0^2} = 1 - \Omega_0^0 \Rightarrow \Omega_0^0 = \Omega_m^0 > 1, \quad \Omega_r^0 = \Omega_\Lambda^0 = 0.$$

Aplicándolo a la ecuación (3.1):

$$\dot{a} = H_0 \sqrt{\Omega_0^0 a^{-1} + 1 - \Omega_0^0} \Rightarrow \frac{\dot{a}}{H_0 \sqrt{\Omega_0^0 a^{-1} - \frac{1}{H_0^2}}} = 1.$$

Integramos ahora en la variable t y dividimos la igualdad por a

$$\int \frac{\dot{a}}{\sqrt{H_0^2 \Omega_0^0 a - a^2}} dt = \int \frac{dt}{a}.$$

Usando ahora la definición de distancia conforme $\eta = \int \frac{dt}{a}$ se tiene:

$$\eta = \int \frac{\dot{a}}{\sqrt{H_0^2 \Omega_0^0 a - a^2}} = \arcsin \left(\frac{2a - \Omega_0^0 H_0^2}{\Omega_0^0 H_0^2} \right) + C.$$

Vamos a asumir que, para estos universos, existe al menos un punto singular donde $a = 0$. Ajustando la constante de integración para que la distancia conforme η sea 0 cuando $a = 0$, se tiene que $C = \frac{\pi}{2}$. Despejando a :

$$a = \frac{\Omega_0^0 H_0^2}{2} \left(\sin \left(\eta - \frac{\pi}{2} \right) + 1 \right).$$

Usando ahora la identidad (3.2) y sabiendo que $\sin(\theta - \frac{\pi}{2}) = -\cos \theta$ se llega a:

$$(3.3) \quad a = \frac{1}{2} \frac{\Omega_0^0}{\Omega_0^0 - 1} (1 - \cos(\eta)).$$

Definiéndose así una formula paramétrica para a .

Volviendo ahora a la definición de distancia conforme, se tiene recíprocamente que:

$$t = \int a d\eta = \frac{1}{2} \frac{\Omega_0^0}{\Omega_0^0 - 1} \int (1 - \cos(\eta)) d\eta = \frac{1}{2} \frac{\Omega_0^0}{\Omega_0^0 - 1} (\eta - \sin(\eta)).$$

Usando de nuevo la identidad (3.2) se llega a:

$$(3.4) \quad t = \frac{1}{2H_0} \frac{\Omega_0^0}{(\Omega_0^0 - 1)^{\frac{3}{2}}} (\eta - \sin(\eta)).$$

Obteniéndose la formula paramétrica para t .

Estas ecuaciones definen cicloides que describen un comportamiento cíclico de Big Crunches y Big Bangs en los que el universo colapsa y se vuelve a formar respectivamente. Se puede ver en la definición paramétrica de a que el universo alcanzará su expansión máxima en $\eta = \pi$:

$$(3.5) \quad a_{max} = \frac{\Omega_0^0}{\Omega_0^0 - 1}.$$

El momento estimado del Big Crunch ocurriría en $\eta = 2\pi$:

$$t_{crunch} = \frac{\pi \Omega_0^0}{H_0 (\Omega_0^0 - 1)^{\frac{3}{2}}}.$$

La edad del universo t_0 se obtiene despejando η_0 con $a_0 = 1$:

$$\eta_0 = \arccos\left(\frac{2 - \Omega_0^0}{\Omega_0^0}\right),$$

$$t_0 = \frac{1}{2H_0} \frac{\Omega_0^0}{(\Omega_0^0 - 1)^{\frac{3}{2}}} \left(\arccos\left(\frac{2 - \Omega_0^0}{\Omega_0^0}\right) - \sin\left(\arccos\left(\frac{2 - \Omega_0^0}{\Omega_0^0}\right)\right) \right).$$

Usando que $\sin \arccos x = \sqrt{1 - x^2}$ se desarrolla a:

$$t_0 = \frac{1}{H_0} \left(\frac{\Omega_0^0}{2(\Omega_0^0 - 1)^{\frac{3}{2}}} \arccos\left(\frac{2 - \Omega_0^0}{\Omega_0^0}\right) - \frac{1}{\Omega_0^0 - 1} \right).$$

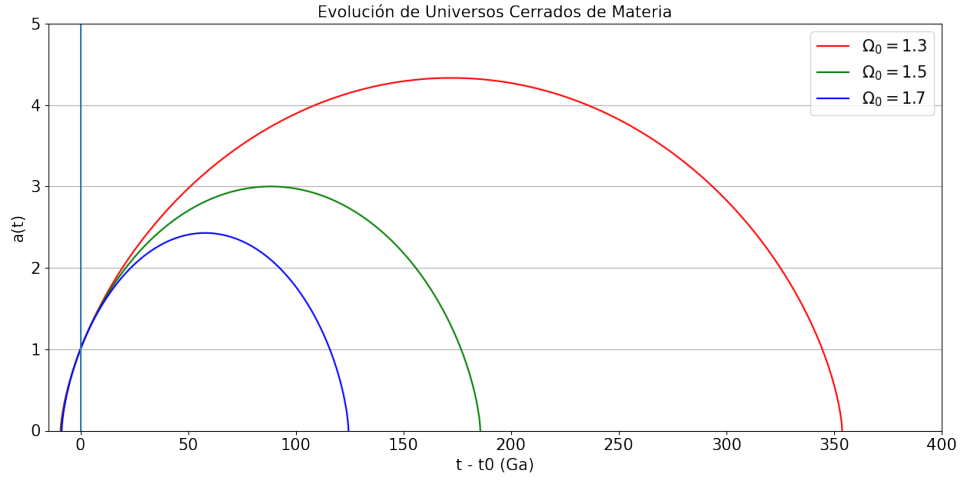


Figura 3.2

3.2.2.2. Universo abierto ($k = -1$)

Los cálculos en esta situación serán parecidos al caso cerrado:

$$(3.6) \quad \Omega_k^0 = \frac{1}{H_0^2} = 1 - \Omega_0^0 \Rightarrow \Omega_0^0 = \Omega_m^0 < 1, \quad \Omega_r^0 = \Omega_\Lambda^0 = 0.$$

Aplicándolo a la ecuación (3.1):

$$\dot{a} = H_0 \sqrt{\Omega_0^0 a^{-1} + 1 - \Omega_0^0} \Rightarrow \frac{\dot{a}}{H_0 \sqrt{\Omega_0^0 a^{-1} + \frac{1}{H_0^2}}} = 1.$$

Integrando ahora en la variable t y dividiendo la igualdad por a :

$$\int \frac{\dot{a}}{\sqrt{H_0^2 \Omega_0^0 a + a^2}} dt = \int \frac{dt}{a}.$$

Usando de nuevo la definición de distancia conforme $\eta = \int \frac{dt}{a}$ en la ecuación:

$$\begin{aligned} \eta &= \int \frac{\dot{a}}{\sqrt{H_0^2 \Omega_0^0 a + a^2}} = \ln \left(\frac{2a + \Omega_0^0 H_0^2 + 2\sqrt{a(\Omega_0^0 H_0^2 + a)}}{\Omega_0^0 H_0^2} \right) + C \\ &= \ln \left(\frac{2a}{\Omega_0^0 H_0^2} + 1 + \sqrt{4 \frac{a}{\Omega_0^0 H_0^2} + \left(2 \frac{a}{\Omega_0^0 H_0^2} \right)^2} \right) + C \\ &= \operatorname{arccosh} \left(\frac{2a}{\Omega_0^0 H_0^2} + 1 \right) + C. \end{aligned}$$

Volviendo a tomar $\eta = 0$ cuando $a = 0$, la constante de integración será $C = 0$. Despejamos a ahora usando la identidad (3.6):

$$(3.7) \quad a = \frac{1}{2} \frac{\Omega_0^0}{1 - \Omega_0^0} (\cosh(\eta) - 1).$$

Definiéndose así una formula paramétrica para a .

De nuevo usando la definición de distancia conforme de manera recíproca se tiene:

$$t = \int a d\eta = \frac{1}{2} \frac{\Omega_0^0}{1 - \Omega_0^0} \int (\cosh(\eta) - 1) d\eta = \frac{1}{2} \frac{\Omega_0^0}{1 - \Omega_0^0} (\sinh(\eta) - \eta).$$

Usando otra vez la identidad (3.6) se consigue:

$$(3.8) \quad t = \frac{1}{2H_0} \frac{\Omega_0^0}{(1 - \Omega_0^0)^{\frac{3}{2}}} (\sinh(\eta) - \eta).$$

Tratándose de la fórmula paramétrica de t .

Este universo tendrá una expansión permanente más rápida que los universos planos y no sufrirá un Big Crunch como los cerrados. Su edad se calcula de nuevo despejando η_0 con $a_0 = 1$:

$$\eta_0 = \operatorname{arccosh} \left(\frac{2 - \Omega_0^0}{\Omega_0^0} \right)$$

$$t_0 = \frac{1}{2H_0} \frac{\Omega_0^0}{(1 - \Omega_0^0)^{\frac{3}{2}}} \left(\sinh \left(\operatorname{arccosh} \left(\frac{2 - \Omega_0^0}{\Omega_0^0} \right) \right) - \operatorname{arccosh} \left(\frac{2 - \Omega_0^0}{\Omega_0^0} \right) \right).$$

Usando ahora

$$\sinh(x) = \frac{e^x - e^{-x}}{2}, \quad \operatorname{arccosh} x = x + \sqrt{x^2 - 1},$$

se llega a:

$$t_0 = \frac{1}{H_0} \left(\frac{1}{1 - \Omega_0^0} - \frac{\Omega_0^0}{2(1 - \Omega_0^0)^{\frac{3}{2}}} \operatorname{arccosh} \left(\frac{2 - \Omega_0^0}{\Omega_0^0} \right) \right).$$

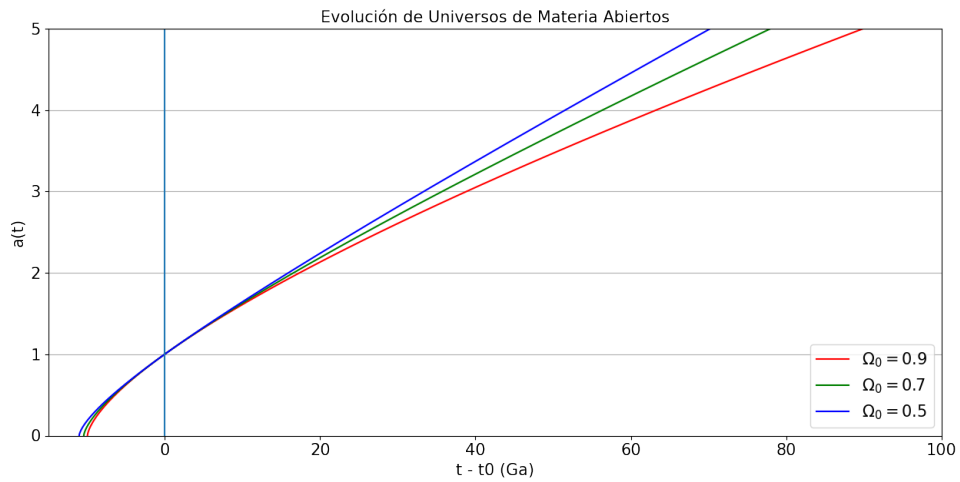


Figura 3.3

3.2.3. Universos espacialmente planos con materia y constante cosmológica Λ

Procedemos a comprobar que ocurre si al modelo del universo de materia plano se le añade la constante cosmológica $\Lambda > 0$. Se debe cumplir

$$0 = \Omega_k^0 = 1 - \Omega_0^0 = 1 - \Omega_\Lambda^0 - \Omega_m^0 \Rightarrow \Omega_\Lambda^0 = 1 - \Omega_m^0.$$

Por tanto, la ecuación (3.1) queda como:

$$\dot{a} = H_0 \sqrt{\Omega_m^0 a^{-1} + (1 - \Omega_m^0) a^2} \Rightarrow \frac{\dot{a}}{\sqrt{\Omega_m^0 a^{-1} + (1 - \Omega_m^0) a^2}} = H_0.$$

Integramos ahora la igualdad obteniendo:

$$\int_{t_0}^t \frac{\dot{a}}{\sqrt{\Omega_m^0 a^{-1} + (1 - \Omega_m^0) a^2}} dt = \int_{t_0}^t H_0 dt \Rightarrow \frac{2 \operatorname{arctanh} \left(\sqrt{\frac{(1 - \Omega_m^0) a^3}{(1 - \Omega_m^0) a^3 + \Omega_m^0}} \right)}{3 \sqrt{1 - \Omega_m^0}} \Bigg|_{t_0}^t = H_0(t - t_0).$$

Usando la definición

$$\operatorname{arctanh} x = \frac{1}{2} (\ln(1 + x) - \ln(1 - x)) = \frac{1}{2} \ln \left(\frac{1 + x}{1 - x} \right),$$

se llega a:

$$\begin{aligned} \frac{3 \sqrt{1 - \Omega_m^0} H_0}{2} (t - t_0) &= \operatorname{arctanh} \left(\sqrt{\frac{(1 - \Omega_m^0) a^3}{(1 - \Omega_m^0) a^3 + \Omega_m^0}} \right) \Bigg|_{t_0}^t \\ &= \frac{1}{2} \ln \left(\frac{\sqrt{(1 - \Omega_m^0) a^3 + \Omega_m^0} + \sqrt{(1 - \Omega_m^0) a^3}}{\sqrt{(1 - \Omega_m^0) a^3 + \Omega_m^0} - \sqrt{(1 - \Omega_m^0) a^3}} \right) \Bigg|_{t_0}^t \\ &= \frac{1}{2} \ln \left(\frac{\left(\sqrt{(1 - \Omega_m^0) a^3 + \Omega_m^0} + \sqrt{(1 - \Omega_m^0) a^3} \right)^2}{\Omega_m^0} \right) \Bigg|_{t_0}^t \\ &= \ln \left(\sqrt{\frac{1 - \Omega_m^0}{\Omega_m^0} a^3 + 1} + \sqrt{\frac{1 - \Omega_m^0}{\Omega_m^0} a^3} \right) \Bigg|_{t_0}^t. \end{aligned}$$

Aprovechándonos ahora de que $\operatorname{arcsinh} x = \ln(\sqrt{x^2 + 1} + x)$ y $a(t_0) = 1$

$$\frac{3 \sqrt{1 - \Omega_m^0} H_0}{2} (t - t_0) = \operatorname{arcsinh} \left(\sqrt{\frac{1 - \Omega_m^0}{\Omega_m^0} a^3} \right) - \operatorname{arcsinh} \left(\sqrt{\frac{1 - \Omega_m^0}{\Omega_m^0}} \right),$$

y despejando a sabiendo que $\Omega_\Lambda = 1 - \Omega_m$ se acaba con:

$$(3.9) \quad a(t) = \sinh \left(\frac{3 \sqrt{\Omega_\Lambda^0} H_0}{2} (t - t_0) + \operatorname{arcsinh} \left(\sqrt{\frac{\Omega_\Lambda^0}{\Omega_m^0}} \right) \right)^{\frac{2}{3}} \left(\frac{\Omega_m^0}{\Omega_\Lambda^0} \right)^{\frac{1}{3}}.$$

Este modelo, para ciertos valores de Ω_m^0 (en torno a 0.3) se acerca bastante a describir la expansión de nuestro propio universo, esto se debe a que: primero, nuestro universo parece ser plano y, segundo, suponer una densidad de radiación nula en el modelo no se aleja de la realidad ya que Ω_r es un factor despreciable fuera de entornos temporales cercanos al Big Bang.

Esta solución tiene una única singularidad en $t^* = t_0 - \frac{2 \operatorname{arcsinh} \left(\sqrt{\frac{1-\Omega_m^0}{\Omega_m^0}} \right)}{3\sqrt{1-\Omega_m^0}H_0}$, por la cual $a(t) \xrightarrow{t \rightarrow t^*} 0$. Para $\Omega_m^0 = 0,317$ (calculado por el experimento Planck 2018 [20]) se tendría una edad de $t_0 - t^* = 13,86$ Ga.

Una característica importante de este modelo es que comienza siendo dominado por la densidad de materia, por lo que empieza pareciéndose al modelo únicamente de materia, hasta que, pasado un punto de inflexión, la constante cosmológica se sobrepone parando la deceleración de la expansión y generando una expansión exponencial. Este momento ocurre cuando $\Omega_m = \Omega_\Lambda = 0,5$. Volviendo de nuevo a la ecuación de fluidos (2.16)

$$\Omega_m = \Omega_m^0 a^{-3} \Rightarrow a = (2\Omega_m^0)^{\frac{1}{3}}.$$

Y calculado t_{inf} usando este valor en (3.9), se llega a que el punto de inflexión es:

$$t_{inf} = t_0 + \frac{2 \left(\operatorname{arcsinh} \left(\sqrt{2\Omega_\Lambda^0} \right) - \operatorname{arcsinh} \left(\sqrt{\frac{\Omega_\Lambda^0}{\Omega_m^0}} \right) \right)}{3\sqrt{\Omega_\Lambda^0}H_0}$$

Para $\Omega_m^0 = 0,317$ se tiene que $t_{inf} < t_0$, lo cual implicaría que ya hemos entrado el proceso de expansión exponencial por la dominancia de la constante Λ .

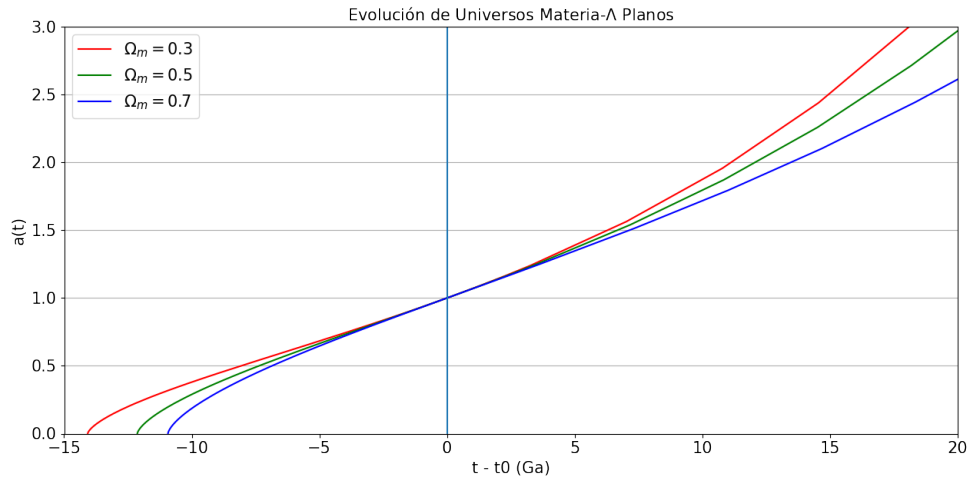


Figura 3.4

3.3. Resolución numérica

Ahora nos interesará estudiar las predicciones del modelo Λ CDM con combinaciones genéricas de parámetros. Para el cálculo numérico de soluciones vamos a emplear la segunda ecuación de Friedmann (2.23). Aunque la ecuación de Friedmann (2.22) es una expresión más sencilla con la que trabajar, su uso en esta situación no es ideal, debido a los problemas que trae la aparición de raíces negativas al usar determinados métodos numéricos para ciertos parámetros.

Comenzamos transformando la ecuación de segundo orden (2.23) en un sistema de ecuaciones de primer orden aplicando el cambio $u = a, v = \dot{a}$:

$$\frac{\dot{v}}{u} = -H_0^2 \left(\frac{1}{2} \Omega_m^0 u^{-3} + \Omega_r^0 u^{-4} - \Omega_\Lambda^0 \right).$$

Los datos iniciales quedan determinados por $u_0 = a(t_0) = 1$ y $v_0 = \dot{a} = H_0$ (nótese que la relación de v_0 es cierta porque $a(t_0) = 1$).

$$(3.10) \quad \begin{cases} \dot{u} = v, & u_0 = 1 \\ \dot{v} = -H_0^2 \left(\frac{1}{2} \Omega_m^0 u^{-2} + \Omega_r^0 u^{-3} - \Omega_\Lambda^0 u \right), & v_0 = H_0. \end{cases}$$

Por último, vamos a escribir la ecuación diferencial con la notación estándar para métodos numéricos $Y(t) = (u(t), v(t))$, $Y_0 = (u_0, v_0)$:

$$(3.11) \quad \begin{cases} Y'(t) = f(t, Y(t)) \\ Y(t_0) = Y_0. \end{cases}$$

3.3.1. Autovalores del problema linealizado

Cabe preguntarse si la ecuación diferencial podría generar un problema Stiff, ya que, de ser cierto, habría que tener cuidado en su resolución numérico debido a la inestabilidad de sus posibles soluciones. Linealizando el problema entorno al punto Y_0 , este queda como:

$$\begin{aligned} f(t, Y(t)) &\approx f(t_0, Y_0) + \frac{\partial f}{\partial t}(t_0, Y_0)(t - t_0) + \frac{\partial f}{\partial u}(t_0, Y_0)(u - u_0) + \frac{\partial f}{\partial v}(t_0, Y_0)(v - v_0) \\ &= \begin{pmatrix} 0 & 1 \\ H_0^2(\Omega_m^0 + 3\Omega_r^0 + \Omega_\Lambda^0) & 0 \end{pmatrix} \begin{pmatrix} u - 1 \\ v - H_0 \end{pmatrix}. \end{aligned}$$

Los autovalores de la matriz asociada son $\pm H_0 \sqrt{\Omega_m^0 + 3\Omega_r^0 + \Omega_\Lambda^0}$. El módulo de estos autovalores no es especialmente grande porque $H_0 < 0,1 (Mpc/Ga)/Mpc$ (equivalente a $100 (km/s)/Mpc$) y $\Omega_m^0 + \Omega_r^0 + \Omega_\Lambda^0 \approx 1$, además, el ratio del módulo de los autovalores es 1, por lo que no nos encontramos ante un problema Stiff. De modo que, métodos numéricos simples serán útiles para alcanzar predicciones estables.

3.3.2. Método de Bogacki-Shampine

Lo importante de la resolución numérica de este sistema son las distintas tendencias de expansión que describe, no los valores exactos de las soluciones, por tanto, a la hora de escoger un método, se prioriza su estabilidad antes que su orden de convergencia. Los métodos de Runge-Kutta explícitos pueden probarse útiles para la resolución de este problema debido a su eficiencia computacional y sencillez de estudio e implementación. Sin embargo, se debe escoger un método de Runge-Kutta de orden bajo, ya que sus regiones de estabilidad tienden a decrecer rápidamente con el orden.

Nos decantamos por el uso del método de Bogacki-Shampine [2]. Este se trata de un algoritmo de pasos adaptativos que hace uso de dos métodos de Runge-Kutta, uno explícito y uno implícito, de orden 3 y 2 respectivamente. El método de orden 3 se usará para hacer predicciones suficientemente precisas, mientras que el de orden 2 se utilizará para controlar el error del método total, debido a su mayor región de estabilidad. Sus tableros de Butcher serán respectivamente:

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ \frac{3}{4} & 0 & \frac{3}{4} & \\ 1 & \frac{2}{9} & \frac{1}{3} & \frac{4}{9} \\ \hline & \frac{2}{9} & \frac{1}{3} & \frac{4}{9} & 0 \end{array} \quad \begin{array}{c|ccc} 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ \frac{3}{4} & 0 & \frac{3}{4} & \\ 1 & \frac{2}{9} & \frac{1}{3} & \frac{4}{9} \\ \hline & \frac{7}{24} & \frac{1}{4} & \frac{1}{3} & \frac{1}{8} \end{array}$$

Siguiendo el primer tablero de Butcher, comenzamos usando el método de orden 3 para realizar predicciones iterativas de los puntos (t_n, Y_n) , dando pasos h_n , empezando por (t_0, Y_0) :

$$\begin{aligned} k_1 &= f(t_n, Y_n), \quad k_2 = f(t_n + \frac{1}{2}h_n, Y_n + \frac{1}{2}h_n k_1), \quad k_3 = f(t_n + \frac{3}{4}h_n, Y_n + \frac{3}{4}h_n k_2) \\ Y_{n+1} &= Y_n + \frac{2}{9}h_n k_1 + \frac{1}{3}h_n k_2 + \frac{4}{9}h_n k_3 \\ t_{n+1} &= t_n + h_n. \end{aligned}$$

Después, procedemos a hacer una segunda predicción con el método de orden 2, con el fin de estimar una cantidad error entre los dos

$$\begin{aligned} k_4 &= f(t_{n+1}, Y_{n+1}) \\ \hat{Y}_{n+1} &= Y_n + \frac{7}{24}h_n k_1 + \frac{1}{4}h_n k_2 + \frac{1}{3}h_n k_3 + \frac{1}{8}h_n k_4 \\ err_n &= \hat{Y}_{n+1} - Y_{n+1}. \end{aligned}$$

El valor err_n no está normalizado. Para normalizarlo lo dividimos por una tolerancia predefinida tol_n que consiste en una tolerancia absoluta $Atol = 10^{-5}$ y una tolerancia relativa $Rtol = 10^{-5}$:

$$tol_n = Atol + Rtol \cdot \max(\|Y_n\|, \|Y_{n-1}\|), \quad E_n = \left\| \frac{err_n}{tol_n} \right\|.$$

Por último, usamos E_n para adaptar el tamaño del paso h_{n+1} con el fin de controlar el error del método con:

$$h_{n+1} = h_n \left(\frac{1}{E_n} \right)^{\frac{1}{3}}.$$

3.3.3. Resultados

Aplicando el método de Bogacki-Shampine [2] al sistema (3.10), se obtienen las siguientes gráficas:

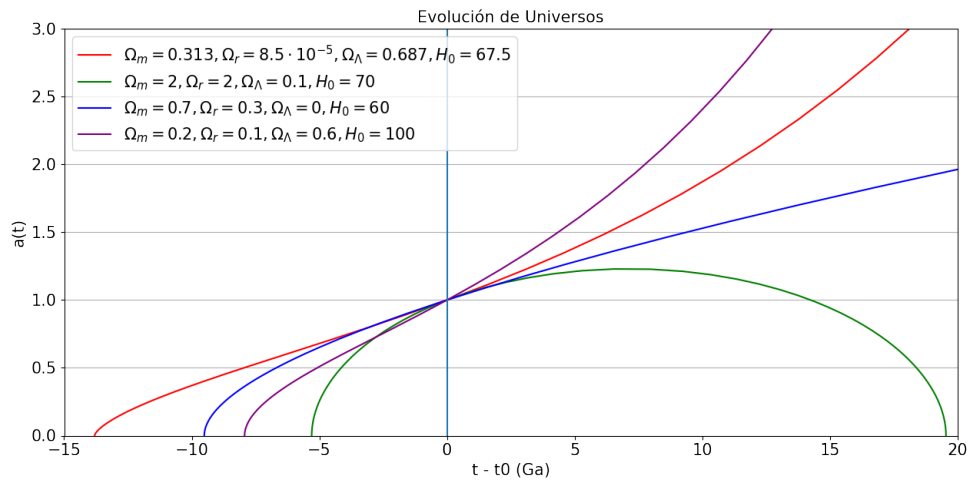


Figura 3.5

- La solución roja usa parámetros aproximados de nuestro propio universo (medidos por el experimento Planck 2018 [20]). En ella se puede apreciar la diferencia entre los primeros momentos, dominados por la materia con una expansión en deceleración, y, pasado un punto de inflexión, el dominio de la constante Λ , caracterizado por una expansión en aceleración como la medida en la actualidad. El modelo calcula una edad de 13.8 Ga.
- La solución azul describe un universo plano sin constante Λ , en permanente crecimiento decelerando, con una edad de 9.5 Ga.
- La solución verde define un universo cerrado, su edad es de 5.3 Ga, y aun teniendo una constante Λ no nula, las densidades de materia y radiación son lo suficientemente grandes como para generar un Big Crunch dentro de 19.5 Ga.
- La solución morada se trata de un universo abierto con una constante de Hubble, que modeliza la expansión actual, mucho mayor que la de los demás universos. Tiene una edad estimada de 7.93 Ga.

CAPÍTULO 4

Estimación de parámetros del modelo Λ CDM

Hasta ahora, hemos estudiado las ecuaciones que rigen la expansión del espacio y hemos calculado algunas de sus posibles soluciones. Por lo tanto, el siguiente paso lógico es tratar de estimar los parámetros de densidad y curvatura que describen nuestro universo. No obstante, antes de hacerlo, necesitamos expresar la ecuación de Friedmann (2.22) en términos de otra variable. Es importante señalar que, debido a la dinámica expansiva que describe el modelo Λ CDM, pasados los tiempos más tempranos la densidad de radiación Ω_r se diluye rápidamente, por tanto, fuera de entornos temporales muy cercanos al Big Bang Ω_r debe ser tan cercana a 0 que tomaremos directamente $\Omega_r^0 = 0$.

4.1. Corrimiento al rojo

El corrimiento al rojo z de una fuente de luz es un proceso mediante el cual su longitud de onda varía desde su emisión hasta su observación (λ_o y λ_e denotan las longitudes de onda al observar y emitir respectivamente):

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e}.$$

En un universo en expansión, este corrimiento puede ser producido por el incremento del espacio por el que se mueve la luz. Usando la métrica Robertson-Walker (2.1) se puede demostrar que existe una relación directa entre z y el factor de escala a que vamos a derivar.

Supongamos que se tiene un rayo de luz que se mueve a través de una geodésica nula ($ds^2 = 0$). Orientando el sistema de coordenadas vamos a tomar que, para el rayo de luz $d\theta = d\phi = 0$, y que el punto desde donde vamos a observarlo es el origen. Usando la métrica (2.1), se debe cumplir:

$$0 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - kr^2} \right) \Rightarrow \frac{dt}{a(t)} = \frac{dr}{\sqrt{1 - kr^2}}.$$

Supongamos que el rayo de luz es emitido en $r = r_e$, consideremos dos picos de su onda separados en el tiempo por Δt , con Δt pequeño. Integrando sobre la trayectoria de la luz desde el tiempo de emisión t_e hasta ahora t_o :

Para el primer pico

$$\int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_0^{r_e} \frac{dr}{\sqrt{1 - kr^2}}.$$

Para el segundo pico

$$\int_{t_e + \Delta t_e}^{t_o + \Delta t_o} \frac{dt}{a(t)} = \int_0^{r_e} \frac{dr}{\sqrt{1 - kr^2}}.$$

Restando las igualdades y descomponiendo $\int_{t_e + \Delta t_e}^{t_o + \Delta t_o}$ en $\int_{t_e}^{t_o} + \int_{t_o}^{t_o + \Delta t_o} - \int_{t_e}^{t_e + \Delta t_e}$:

$$\int_{t_o}^{t_o + \Delta t_o} \frac{dt}{a(t)} - \int_{t_e}^{t_e + \Delta t_e} \frac{dt}{a(t)} = 0.$$

$a(t)$ es prácticamente una constante para cambios de tiempo tan pequeños como Δt , por lo que, sacándola de las integrales se tiene la igualdad

$$\frac{\Delta t_e}{a(t_e)} = \frac{\Delta t_o}{a(t_o)} \Rightarrow \frac{a(t_o)}{a(t_e)} = \frac{\lambda_o}{\lambda_e}.$$

Nótese que, tomando $c = 1$, entonces se tiene que $\lambda = \Delta t$ para el rayo de luz. Usando ahora que $a(t_0) = 1$ se llega a:

$$\frac{1}{a(t_e)} = \frac{\lambda_o}{\lambda_e} = \frac{\lambda_o - \lambda_e}{\lambda_e} + 1 = z + 1 \Rightarrow a(t_e) = \frac{1}{1 + z}.$$

Aplicando esta igualdad, podemos reescribir la ecuación de Friedmann (2.22) como:

$$(4.1) \quad H^2(z) = H_0^2 (\Omega_m^0 (1 + z)^3 + \Omega_\Lambda^0 + \Omega_k^0 (1 + z)^2),$$

$$(4.2) \quad \Omega_k^0 = (1 - \Omega_m^0 - \Omega_\Lambda^0).$$

4.2. Mediciones de $H(z)$

Para estimar los parámetros $\Omega_m, \Omega_k, \Omega_\Lambda$ y H_0 del modelo Λ CDM, utilizaremos mediciones independientes del valor de $H(z)$ para distintos corrimientos al rojo. Estas mediciones provienen de dos métodos diferentes: el primero, llamado “differential age method” (DA), utiliza cronómetros cosmológicos, que son galaxias cuya evolución se estudia para inferir el valor de $H(z)$, el segundo método se basa en observaciones de las Oscilaciones Acústicas de Bariones (BAO), que son ondas provenientes del plasma primigenio liberadas durante la época de recombinación de la materia y que contienen información sobre la expansión del universo.

En total tendremos 53 valores distintos de $H(z)$ para los que realizar nuestras estimaciones (se pueden consultar en el apéndice) y, a continuación, mostraremos tres métodos distintos para realizar dichas estimaciones.

4.3. Estimación de la mediana

Vamos a comenzar con un método básico de estadística frecuentista: estimar mediante el cálculo de la mediana muestral. Para ello, supondremos que los datos que tenemos son independientes y no sufren de un error sistemático entre ellos. Seguiremos el procedimiento desarrollado en [18].

A partir de tres valores distintos $\{H(z) : z = a, b, c\}$ usando la fórmula (4.1) podemos establecer un sistema de 3 ecuaciones, y, añadiendo (4.2), se pueden calcular los 4 parámetros. Despejando el sistema se llega a:

$$A(a, b, c) = (1+a)^2(1+b)^2(b-a) + (1+b)^2(1+c)^2(c-b) + (1+c)^2(1+a)^2(a-c)$$

$$(4.3) \quad \begin{aligned} \Omega_m H_0^2 = & \frac{H^2(a) [(1+b)^2 - (1+c)^2]}{-A(a, b, c)} + \frac{H^2(b) [(1+c)^2 - (1+a)^2]}{-A(a, b, c)} + \\ & + \frac{H^2(c) [(1+a)^2 - (1+b)^2]}{-A(a, b, c)}, \end{aligned}$$

$$(4.4) \quad \begin{aligned} \Omega_k H_0^2 = & \frac{H^2(a) [(1+b)^3 - (1+c)^3]}{A(a, b, c)} + \frac{H^2(b) [(1+c)^3 - (1+a)^3]}{A(a, b, c)} + \\ & + \frac{H^2(c) [(1+a)^3 - (1+b)^3]}{A(a, b, c)}, \end{aligned}$$

$$(4.5) \quad \begin{aligned} \Omega_\Lambda H_0^2 = & \frac{H^2(a)(1+b)^2(1+c)^2(c-b)}{A(a, b, c)} + \frac{H^2(b)(1+c)^2(1+a)^2(a-c)}{A(a, b, c)} + \\ & + \frac{H^2(c)(1+a)^3(1+b)^3(b-a)}{A(a, b, c)}, \end{aligned}$$

$$(4.6) \quad H_0 = \sqrt{\Omega_m H_0^2 + \Omega_k H_0^2 + \Omega_\Lambda H_0^2}.$$

Tomar 3 valores aleatorios de nuestros datos y calcular los parámetros de esta manera no es útil. Debido a los errores inherentes a las medidas de $H(z)$, no podemos calcular los parámetros exactos, sin embargo, si calculamos $\Omega_m H_0^2, \Omega_k H_0^2$ y $\Omega_\Lambda H_0^2$ para las $n = \binom{53}{3} = 23426$ posibles combinaciones de $H(z)$, podemos estimar los parámetros usando la mediana muestral. Usaremos la mediana en vez de la media porque esta es más robusta ante datos atípicos.

Estamos especialmente interesados en el intervalo de 95% de confianza de la mediana m . Sin pérdida de generalidad, calcularemos el intervalo de $X = \Omega_m H_0^2$. Consideremos el estadístico de orden de X : $X_{(0)} \leq \dots \leq X_{(n)}$. Por definición, $P(X_{(i)} < m) = 0,5$, entonces, definiendo W como el número de $X_{(i)} < m$, se tiene que $W \sim B(n, 0,5)$ con $\mu = \frac{n}{2}$ y $\sigma^2 = \frac{n}{4}$. Puesto que $n = 23426$ es suficientemente grande, normalizando, se llega a la aproximación $\frac{W-\mu}{\sigma} = Z \sim N(0, 1)$, que se puede usar para calcular el intervalo de confianza:

$$0,95 \approx P(-1,96 \leq Z \leq 1,96) \approx P\left(\frac{n}{2} - 1,96 \frac{\sqrt{n}}{2} \leq W \leq \frac{n}{2} + 1,96 \frac{\sqrt{n}}{2}\right)$$

$$\xrightarrow{n=23426}$$

$$0,95 \approx P(X_{(11563)} \leq m \leq X_{(11862)}).$$

Siguiendo el método previamente descrito y aplicando este intervalo de confianza se tiene:

$$\Omega_m H_0^2 = 1484,89_{-49,90}^{+41,37}, \quad \Omega_k H_0^2 = -137,26_{-123,03}^{+129,62}, \quad \Omega_\Lambda H_0^2 = 3126,15_{-85,97}^{+133,70}.$$

Por último, usando (4.6) y propagando los intervalos adecuadamente se obtiene:

$$\Omega_m = 0,332_{-0,016}^{+0,017}, \quad \Omega_k = -0,031_{-0,027}^{+0,029}, \quad \Omega_\Lambda = 0,699_{-0,031}^{+0,042}, \quad H_0 = 66,89_{-1,18}^{+1,43}.$$

Comparando estos valores con las estimaciones del experimento Planck, se puede ver que el resultado es bastante satisfactorio teniendo en cuenta la sencillez del método y la escasez de datos utilizados:

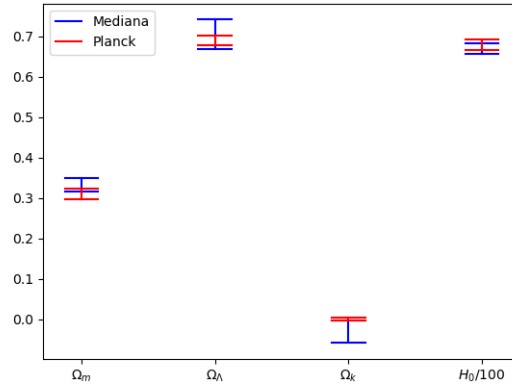


Figura 4.1

4.4. Estimación mediante el algoritmo de Metropolis-Hastings

El algoritmo de Metropolis-Hastings (véase capítulo 7 de [21]) es un método de Cadenas de Markov Monte Carlo que se utiliza para estimar la distribución de ciertos parámetros a través del análisis Bayesiano.

Trataremos al parámetro θ (en nuestro caso, $H_0, \Omega_m, \Omega_\Lambda$) como una variable aleatoria con una densidad posterior $\pi(\theta|Y)$ dependiente de unas muestras Y , y usaremos cadenas de Markov para estimar dicha densidad.

Nótese que aplicando la fórmula de Bayes se tiene que:

$$(4.7) \quad \pi(\theta|Y) = \frac{P(Y|\theta)\pi(\theta)}{\int_{\Theta} P(Y|\theta)\pi(\theta)d\theta} \propto P(Y|\theta)\pi(\theta),$$

donde $\pi(\theta)$ es una probabilidad a priori que definiremos más adelante según el problema.

Ahora vamos a introducir una serie de definiciones básicas para el desarrollo teórico del algoritmo.

4.4.1. Cadena de Markov

Definición 4.1. Una *cadena de Markov* es un proceso estocástico $\{X^{(t)} : t \in \mathbb{N}\}$ en un espacio de estados S discreto que $\forall t \geq 0$ y $\forall x_0, \dots, x_{t+1} \in S$ satisface:

$$P(X^{(t+1)} = x_{t+1} | X^{(0)} = x_0, \dots, X^{(t)} = x_t) = P(X^{(t+1)} = x_{t+1} | X^{(t)} = x_t).$$

Definición 4.2. Una cadena de Markov es *homogénea* si la probabilidad de pasar de un estado i a otro j no depende del tiempo:

$$P(X^{(t+1)} = j | X^{(t)} = i) = P(X^{(1)} = j | X^{(0)} = i) = p_{ij}.$$

p_{ij} es la probabilidad de transición y la matriz $P = (p_{ij})_{|S| \times |S|}$ se conoce como la matriz de transición que cumple que:

$$\forall i, j \in S \quad p_{ij} > 0, \quad \forall i \quad \sum_{j \in S} p_{ij} = 1.$$

Definición 4.3. Un estado $i \in S$ se *comunica con un estado* $j \in S$ si existe una cierta cantidad de pasos t tras la cual la probabilidad de transicionar $p_{ij}(t)$ una vez hechos es $p_{ij}(t) > 0$, esto establece una relación de equivalencia. Si todos los estados de una Cadena de Markov están comunicados entre ellos entonces la cadena es *irreducible*.

Definición 4.4. El *periodo de un estado* $i \in S$ queda descrito como:

$$d(i) = \text{mcd}\{t \geq 1 : p_{ii}(t) > 0\}.$$

Si $d(i) = 1$ entonces i será un estado *aperiódico*. Una cadena de Markov es *aperiódica* si $\forall i \in S \quad d(i) = 1$

Definición 4.5. Una distribución de probabilidad $\pi = (\pi_0, \pi_1, \dots)$ es *estacionaria* para una cadena de Markov con matriz de transición P si:

$$\pi P = \pi.$$

Teorema 4.6. Si una cadena de Markov homogénea es irreducible, aperiódica y cuenta con una distribución estacionaria π , entonces $\forall i, j \in S$:

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j.$$

Su distribución estacionaria es una distribución límite para la cadena.

4.4.2. Descripción del algoritmo de Metropolis-Hastings

El algoritmo de Metropolis-Hastings descrito en [21] (véase capítulo 7) nos permitirá construir una cadena de Markov cuya distribución límite coincidirá con la distribución posterior $\pi(\theta|Y)$. De modo que tomando suficientes muestras de la cadena podremos estimar dicha densidad.

Para realizarlo, se extenderá el concepto de cadena de Markov a un espacio de estados continuo haciendo uso de un kernel de Markov $Q = \{Q(\theta_a, \theta_b) : \theta_a, \theta_b \in \Theta\}$, que será una densidad con respecto a su segunda coordenada:

$$\forall \theta_a \in \Theta \int_{\Theta} Q(\theta_a, \theta) d\theta = 1.$$

Este kernel tomará la función de las probabilidades de transición p_{ij} . Una vez definido Q el algoritmo queda como:

1. Inicializa $\theta^{(0)}$ con un valor predefinido.
2. Desde $t = 0$ hasta $t = T - 1$ repite:
 - Genera una muestra $\theta^* \sim q(\theta|\theta^{(t)}) = Q(\theta^{(t)}, \theta)$.
 - Calcula $\alpha(\theta^{(t)}, \theta^*) = \min \left(1, \frac{\pi(\theta^*|Y)Q(\theta^{(t)}, \theta^*)}{\pi(\theta^{(t)}|Y)Q(\theta^{(t)}, \theta^*)} \right) = \min \left(1, \frac{P(Y|\theta^*)\pi(\theta^*)Q(\theta^{(t)}, \theta^*)}{P(Y|\theta^{(t)})\pi(\theta^{(t)})Q(\theta^{(t)}, \theta^*)} \right)$.
 - Genera una muestra $u \sim \text{Uniforme}(0, 1)$.
 - Si $u < \alpha$: toma $\theta^{(t+1)} = \theta^*$. Si no: toma $\theta^{(t+1)} = \theta^{(t)}$.
3. Elimina las m primeras muestras.

Intuitivamente, con el algoritmo iremos tomando valores siguiendo la distribución Q , pero solo actualizaremos la cadena con esos valores según una cierta probabilidad (dependiendo del valor de la densidad posterior en θ^* y $\theta^{(t)}$). De esta forma, la cadena devolverá más muestras de estados con mayor densidad siguiendo la distribución posterior como distribución límite. Por último, se eliminan las primeras muestras del algoritmo previas a la convergencia a la distribución posterior.

4.4.3. Convergencia del algoritmo de Metropolis-Hastings

Para obtener resultados útiles del algoritmo necesitamos demostrar dos propiedades sobre su Cadena de Markov asociada:

1. La densidad $\pi(\theta|Y)$ es estacionaria con respecto a la cadena.
2. La distribución de la cadena converge a $\pi(\theta|Y)$.

Nótese que al extender el conjunto de estados de un espacio discreto a uno continuo, la condición 1) sobre la distribución estacionaria queda modificada a:

$$(4.8) \quad \pi P = \pi \rightarrow \forall \theta_b \in \Theta \int_{\Theta} \pi(\theta_a|Y) K(\theta_a, \theta_b) d\theta_a = \pi(\theta_b|Y).$$

Donde K se trata del kernel de transición que será la contraparte continua de P , que deberá cumplir el ser una densidad respecto a su segunda variable.

El kernel de transición K asociado a la Cadena de Markov del algoritmo de Metropolis-Hastings es:

$$(4.9) \quad K(\theta_a, \theta_b) = Q(\theta_a, \theta_b) \alpha(\theta_a, \theta_b) + (1 - r(\theta_a)) \delta_{\theta_a}(\theta_b).$$

Siendo $r(\theta_a) = \int_{\Theta} \alpha(\theta_a, \theta_b) Q(\theta_a, \theta_b) d\theta_b$ y $\delta_{\theta_a}(\theta_b)$ la delta de Dirac centrada en θ_a . El primer sumando representa la distribución cuando se acepta un determinado θ^* propuesto, y el segundo sumando, el caso contrario.

Para demostrar que (4.9) cumple la propiedad (4.8) probaremos antes el siguiente lema:

Lema 4.7.

$$\pi(\theta_a|Y)K(\theta_a, \theta_b) = \pi(\theta_b|Y)K(\theta_b, \theta_a).$$

Demostración.

- Si $\theta_a = \theta_b$: el lema es trivial.
- Si $\theta_a \neq \theta_b$:

$$\begin{aligned} \pi(\theta_a|Y)K(\theta_a, \theta_b) &= \pi(\theta_a|Y)Q(\theta_a, \theta_b)\alpha(\theta_a, \theta_b) = \pi(\theta_a|Y)Q(\theta_a, \theta_b) \min\left(1, \frac{\pi(\theta_b|Y)Q(\theta_b, \theta_a)}{\pi(\theta_a|Y)Q(\theta_a, \theta_b)}\right) \\ &= \min(\pi(\theta_a|Y)Q(\theta_a, \theta_b), \pi(\theta_b|Y)Q(\theta_b, \theta_a)) = \pi(\theta_b|Y)Q(\theta_b, \theta_a) \min\left(\frac{\pi(\theta_a|Y)Q(\theta_a, \theta_b)}{\pi(\theta_b|Y)Q(\theta_b, \theta_a)}, 1\right) \\ &= \pi(\theta_b|Y)Q(\theta_b, \theta_a)\alpha(\theta_b, \theta_a) = \pi(\theta_b|Y)K(\theta_b, \theta_a). \quad \square \end{aligned}$$

Usando el lema (4.7) se puede ver que:

$$\begin{aligned} \forall \theta_b \in \Theta \quad \int_{\Theta} \pi(\theta_a|Y)K(\theta_a, \theta_b) d\theta_a &= \int_{\Theta} \pi(\theta_b|Y)K(\theta_b, \theta_a) d\theta_a \\ &= \pi(\theta_b|Y) \int_{\Theta} K(\theta_b, \theta_a) d\theta_a = \pi(\theta_b|Y) \left(\int_{\Theta} K(\theta_b, \theta_a) d\theta_a = 1 \right). \end{aligned}$$

Por tanto, $\pi(\theta|Y)$ será la distribución estacionaria de la cadena del algoritmo.

Durante el algoritmo, para cada punto $\theta^{(t)}$, existe una probabilidad de que se de $\theta^{(t+1)} = \theta^{(t)}$, por lo que, la Cadena de Markov asociada es aperiódica ya que la probabilidad de moverse de un estado a él mismo es no nula. La Cadena de Markov es irreducible ya que, al ser Q una densidad positiva sobre todos los valores, todos los estados están comunicados pues cualquier θ^* puede ser propuesto. Como la cadena es aperiódica e irreducible, por el teorema (4.6) esta terminará convergiendo a la distribución estacionaria $\pi(\theta|Y)$. Por tanto, podremos estimar adecuadamente $\pi(\theta|Y)$ recogiendo el estado de la cadena en cada paso del algoritmo.

4.4.4. Aplicación a las mediciones de $H(z)$

Vamos a estimar los parámetros $(H_0, \Omega_m, \Omega_\Lambda)$. Nótese que no es necesario estimar Ω_k al quedar determinado por los demás. Vamos a suponer que cada medición de $H(z_i)$ es independiente y sigue una distribución normal cuyo centro queda determinado por la ecuación (4.1) y su desviación típica es la reportada por cada experimento. Usaremos el cuadrado de estas mediciones, de modo que:

$$H^2(z_i) \sim N(H^2(z_i; H_0, \Omega_m, \Omega_\Lambda), 2H(z_i)\sigma_i).$$

Nótese que la desviación típica queda reescalada por $2H(z_i)$ para ajustar al uso del cuadrado $H^2(z_i)$.

Tomando esta suposición se llega a:

$$\begin{aligned} P(Y|(H_0, \Omega_m, \Omega_\Lambda)) &\propto \prod_{i=1}^{i=n} \exp \left(-\frac{1}{2} \left(\frac{H^2(z_i) - H^2(z_i; H_0, \Omega_m, \Omega_\Lambda)}{2H(z_i)\sigma_i} \right)^2 \right) \\ &= \exp \left(-\frac{1}{2} \sum_{i=1}^{i=n} \left(\frac{H^2(z_i) - H^2(z_i; H_0, \Omega_m, \Omega_\Lambda)}{2H(z_i)\sigma_i} \right)^2 \right). \end{aligned}$$

Como probabilidad a priori $\pi(H_0, \Omega_m, \Omega_\Lambda)$ tomaremos que los parámetros son independientes y que siguen las siguientes distribuciones:

$$\begin{cases} H_0 \sim \text{Uniforme}(50, 100) \\ \Omega_m, \Omega_\Lambda \sim \text{Uniforme}(0, 1) \\ \Omega_k = 1 - \Omega_m - \Omega_\Lambda \sim \text{Uniforme}(-0.1, 0.1) \end{cases}$$

Estas condiciones provienen de querer buscar los parámetros entre intervalos razonables, sin ninguna preferencia de valores. Nótese que el intervalo de Ω_k es muy estrecho debido al conocimiento de que la curvatura del universo debe ser cercana a 0. Inicializaremos $(H_0^{(0)}, \Omega_m^{(0)}, \Omega_\Lambda^{(0)}) = (65, 0.3, 0.7)$ y tomaremos Q como:

$$Q((H_0^{(t)}, \Omega_m^{(t)}, \Omega_\Lambda^{(t)}), (H_0, \Omega_m, \Omega_\Lambda)) \sim N(H_0^{(t)}, 0.8) \times N(\Omega_m^{(t)}, 0.025) \times N(\Omega_\Lambda^{(t)}, 0.025).$$

Las desviaciones típicas del kernel están seleccionadas balanceando entre una buena exploración de los parámetros y una buena proporción de muestras propuestas aceptadas.

Por último, para evitar errores numéricos en la computación, realizaremos el algoritmo de Metropolis-Hastings con el logaritmo de $P(Y|(H_0, \Omega_m, \Omega_\Lambda))$, nótese que, como Q es simétrico, este se va a cancelar en el cálculo de α . De esta forma podremos estimar la distribución posterior de los parámetros. Cabe destacar que eliminaremos las 5000 primeras muestras de la cadena, ya que éstas no serán representativas de la distribución límite.

4.4.5. Resultados

Con las muestras obtenidas mediante el algoritmo de Metropolis-Hastings, podemos utilizar un estimador de núcleo para reconstruir la distribución de cada parámetro y compararla con los resultados del experimento Planck. Reescalando las densidades por altura:

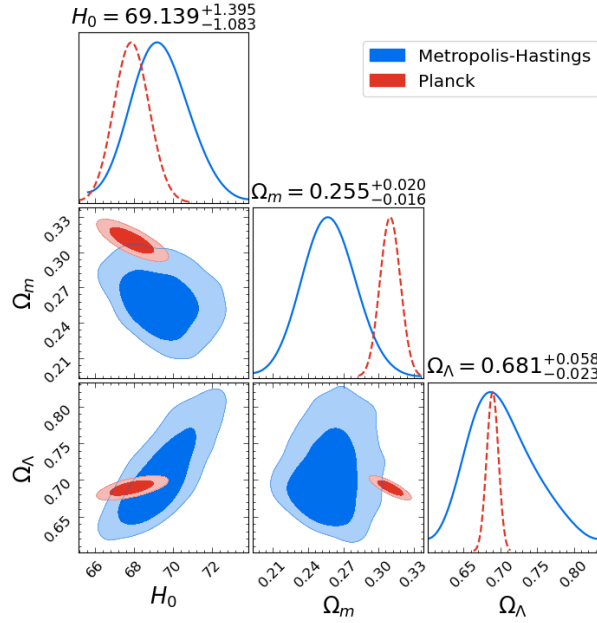


Figura 4.2

El resultado de la distribución de Ω_Λ es lo esperable: una densidad más ancha que la de Planck, debido a la diferencia en la cantidad de datos usados, y picos aproximadamente alineados. Sin embargo, esta no es la tendencia que sigue Ω_m . Calculando los intervalos del 95 % de credibilidad:

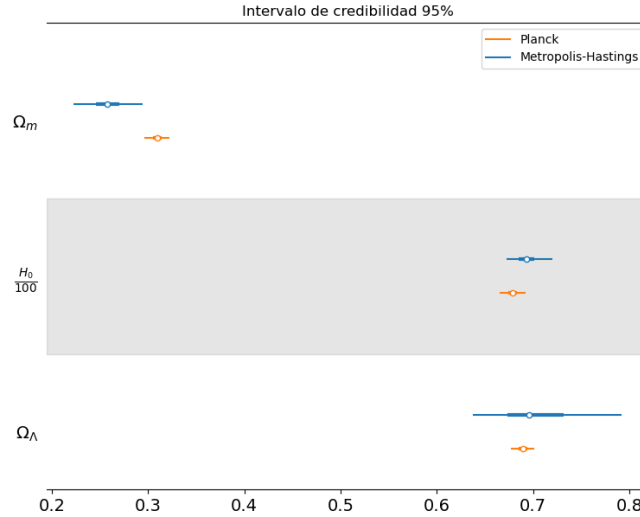


Figura 4.3

Es apreciable que los intervalos de Ω_m son disjuntos. Esto muestra una clara discordancia entre ambas estimaciones de Ω_m . Trazando la ecuación de Friedmann

(4.1) con las medias de los parámetros para ambas estimaciones y añadiendo las barras de error de las mediciones de $H(z)$:

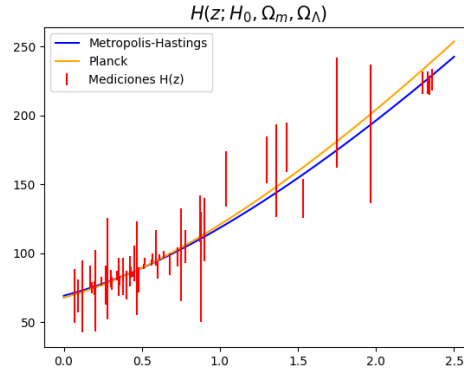


Figura 4.4

Ambas predicciones comienzan similares pero se separan al final. El modelo con los parámetros de Planck no interseca las cuatro últimas mediciones. Estas son importantes debido a su gran precisión y explican la disparidad entre las estimaciones. Dados estos datos, los parámetros de Planck fallan en predecir $H(z)$ para altos valores de corrimiento al rojo con 2σ de error. Esto presenta uno de los grandes problemas del modelo Λ CDM, y es que no parece ajustar correctamente todas nuestras mediciones actuales. Comienza a existir una cierta disonancia entre el estudio de los parámetros usando el Fondo Cósmico de Microondas y el de supernovas. Un ejemplo de este problema es el caso de la tensión de Hubble [5].

Repitiendo el procedimiento excluyendo los cuatro últimos valores se tiene:

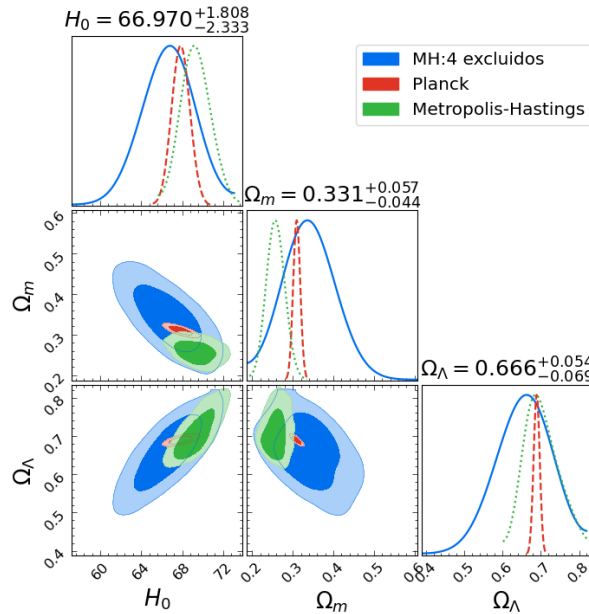


Figura 4.5

Con los datos excluidos obtenemos las distribuciones que nos hubiesemos esperado antes: anchas y aproximadamente centradas en los picos de Planck. Calculando de nuevo los intervalos del 95 % de credibilidad se puede ver como los intervalos del algoritmo con los datos excluidos incluyen a los de Planck:

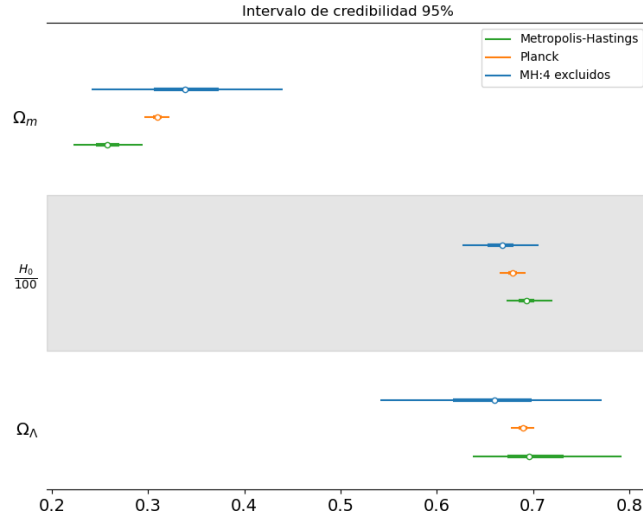


Figura 4.6

Y trazando la función de $H(z)$ nuevo, se vuelve a apreciar el problema en las predicciones en $z \approx 2,4$:

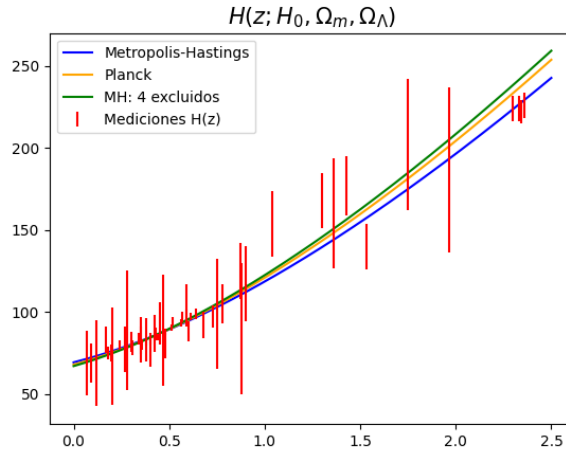


Figura 4.7

Nótese que, observando la fórmula (4.1), se encuentra la explicación sobre la diferencia entre las estimaciones de Ω_m . Esta se debe a que Ω_m multiplica directamente

a un factor de z al cubo, que tiene un mayor efecto en z grandes donde fallan las predicciones de Planck.

4.5. Estimación mediante Redes Neuronales

Una red neuronal es un modelo matemático capaz de aprender relaciones complejas entre datos dados, a partir de métodos numéricos de optimización (véase capítulo 6 de [12]). En este caso, aprovecharemos sus excelentes resultados en el campo de la regresión numérica y los aplicaremos a la estimación de parámetros cosmológicos.

4.5.1. Propagación hacia delante en una red neuronal densa

En una estructura clásica de red neuronal densa, un vector de valores se introduce por la capa de entrada y se va propagando hacia delante hasta la capa de salida.

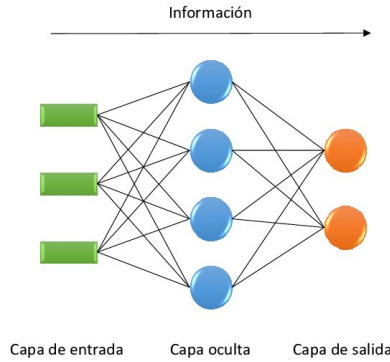


Figura 4.8

Esta propagación se hace a través de transformaciones lineales. Sea una capa l con $n^{(l)}$ neuronas, esta tiene como parámetros: la matriz de pesos $W^{(l)} \in \mathbb{R}^{n^{(l)} \times n^{(l-1)}}$, el vector de coeficientes $b^{(l)} \in \mathbb{R}^{n^{(l)}}$ y una función de activación $\sigma^{(l)}$ (normalmente no lineal). El resultado de salida $a^{(l)} \in \mathbb{R}^{n^{(l)}}$ de dicha capa se calcula como:

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)},$$

$$a^{(l)} = \sigma^{(l)}(z^{(l)}).$$

Donde $\sigma^{(l)}$ se aplica a cada elemento de $z^{(l)}$, excepto en la última capa l' donde $a^{(l')} = z^{(l')}$.

4.5.2. Propagación hacia atrás en una red neuronal densa

El entrenamiento de una red neuronal se realiza con datos de entrada para los cuales conocemos el resultado que deseamos. Se comparan los valores de la red $z^{(l')}$ con los resultados correctos \hat{y} mediante una función de pérdida convexa $\mathcal{L}(z^{(l')}, \hat{y})$ cuyo mínimo es 0 y se alcanza con $z^{(l')} = \hat{y}$. Entrenar la red se trata de un problema de

optimización donde se quiere minimizar \mathcal{L} alterando los parámetros W y b . Sin embargo, no es posible resolver este problema analíticamente, razón por la cual es necesario usar algún método numérico de optimización. Estos métodos requieren información sobre el gradiente de los parámetros $\frac{\partial \mathcal{L}}{\partial w_{ij}^{(l)}}$, $\frac{\partial \mathcal{L}}{\partial b_j^{(l)}}$, pero, estos valores tampoco se pueden calcular directamente, para ello es necesario derivar el algoritmo de propagación hacia atrás:

Aplicando la regla de la cadena se llega a las siguientes igualdades:

$$(4.10) \quad \frac{\partial \mathcal{L}}{\partial w_{ij}^{(l)}} = \frac{\partial \mathcal{L}}{\partial z_j^{(l)}} \frac{\partial z_j^{(l)}}{\partial w_{ij}^{(l)}} = \frac{\partial \mathcal{L}}{\partial z_j^{(l)}} a_i^{(l-1)},$$

$$(4.11) \quad \frac{\partial \mathcal{L}}{\partial b_j^{(l)}} = \frac{\partial \mathcal{L}}{\partial z_j^{(l)}}.$$

Basta con encontrar una fórmula para $\frac{\partial \mathcal{L}}{\partial z_j^{(l)}}$. Usando la regla de la cadena de nuevo, podemos establecer una recursión de $\frac{\partial \mathcal{L}}{\partial z_j^{(l)}}$ en términos de $\left\{ \frac{\partial \mathcal{L}}{\partial z_k^{(l+1)}} \right\}$, esto se debe a que el efecto de $z_j^{(l)}$ en la función se propaga a todos los $z_k^{(l)}$ ($k = 1, \dots, n^{(l)}$):

$$(4.12) \quad \frac{\partial \mathcal{L}}{\partial z_j^{(l)}} = \sum_{k=1}^{n^{(l)}} \frac{\partial \mathcal{L}}{\partial z_k^{(l+1)}} \frac{\partial z_k^{(l+1)}}{\partial z_j^{(l)}}.$$

Por último, sabiendo que $z_k^{(l+1)} = \sum_i w_{ik}^{(l+1)} a_i^{(l)} = \sum_i w_{ik}^{(l+1)} \sigma^{(l)}(z_i^{(l)})$ se tiene:

$$(4.13) \quad \frac{\partial z_k^{(l+1)}}{\partial z_j^{(l)}} = w_{jk}^{(l+1)} \sigma'^{(l)}(z_j^{(l)}).$$

Combinando (4.10), (4.11), (4.12) y (4.13) y gracias a la recursión comenzada por la última capa, con el algoritmo de propagación hacia atrás calcularemos los gradientes necesarios para aplicar un método de optimización que entrene a la red.

4.5.3. Descenso de gradiente

El método de optimización con gradientes más básico es el descenso de gradiente [3]. Es un método recursivo que usaremos para minimizar, modificando los parámetros x_i de la red, una función de pérdida f sobre los datos de entrenamiento. Dado los parámetros $x_i \in \mathbb{R}^n$ y un tamaño de paso $\alpha > 0$, este calcula un punto cercano x_{i+1} aplicando:

$$(4.14) \quad x_{i+1} = x_i - \alpha \nabla f(x_i).$$

Vamos a demostrar que usando este proceso se obtendrá $f(x_{i+1}) \leq f(x_i)$.

Lema 4.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente $\nabla f(x)$ L -Lipschitz

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|x - y\|^2.$$

Demostración.

Definamos la función $g : \mathbb{R} \rightarrow \mathbb{R}$ como $g(\epsilon) = f(x + \epsilon(y - x))$. Usando el teorema fundamental del cálculo se tiene:

$$g(1) = g(0) + \int_0^1 \frac{d}{d\epsilon} g(\epsilon) d\epsilon.$$

Reescribiendo la igualdad en términos de f y desarrollando las derivadas se llega a:

$$f(y) = f(x) + \int_0^1 \langle \nabla f(x + \epsilon(y - x)), y - x \rangle d\epsilon.$$

Sumando y restando $\langle \nabla f(x), y - x \rangle$ de la igualdad

$$= f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \langle \nabla f(x + \epsilon(y - x)) - \nabla f(x), y - x \rangle d\epsilon.$$

Aplicando la desigualdad de Cauchy-Schwarz

$$\leq f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \|\nabla f(x + \epsilon(y - x)) - \nabla f(x)\| \cdot \|y - x\| d\epsilon.$$

Usando la propiedad de Lipschitz del gradiente

$$\begin{aligned} &\leq f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 L \|\epsilon(y - x)\| \cdot \|y - x\| d\epsilon \\ &= f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|x - y\|^2. \end{aligned}$$

□

Aplicando ahora el lema (4.8), suponiendo α lo suficientemente pequeño y con $\nabla f(x)$ L -Lipschitz, se llega a $f(x_{i+1}) < f(x_i)$:

$$\begin{aligned} f(x_{i+1}) &\leq f(x_i) + \langle \nabla f(x_i), x_{i+1} - x_i \rangle + \frac{L}{2} \|x_{i+1} - x_i\|^2 \\ &= f(x_i) + \langle \nabla f(x_i), -\alpha \nabla f(x_i) \rangle + \frac{L}{2} \|\alpha \nabla f(x_i)\|^2 \quad \text{Sustituyendo con (4.14)} \\ &= f(x_i) - \alpha \|\nabla f(x_i)\|^2 + \frac{L\alpha^2}{2} \|\nabla f(x_i)\|^2 = f(x_i) - \alpha \left(1 - \frac{L\alpha}{2}\right) \|\nabla f(x_i)\|^2 \\ &\leq f(x_i) - \frac{\alpha}{2} \|\nabla f(x_i)\|^2 \quad (\text{Si } L\alpha \leq 1). \end{aligned}$$

Al calcular un paso de descenso de gradiente tenemos garantizado que los parámetros modificados reducirán el valor de la función de pérdida. Sin embargo, surgen dos problemas: el primero es que el descenso puede quedarse atascado en un mínimo local, el segundo es que realizar este cálculo sobre todos los datos de entrenamiento puede ser muy costoso. Para reducir el impacto de estos problemas se desarrolla el descenso de gradiente estocástico, donde los cálculos del descenso se producen en cada paso con un subconjunto de los datos. De esta forma la optimización se vuelve más aleatoria, pero se reduce el coste computacional y puede evitarse con una cierta probabilidad el acabar en mínimos locales. De esta idea surge la rama de optimizadores de gradientes estocásticos.

4.5.4. Proceso de entrenamiento

La metodología que vamos a aplicar proviene del paper “ECoPANN” [27], donde se desarrolla el como estimar parámetros cosmológicos usando redes neuronales.

En un problema clásico de regresión con redes neuronales se cuenta con un conjunto de datos de entrenamiento (etiquetados con la respuesta que buscamos) y otro sobre el que queremos predecir soluciones. En este caso, solo contamos con mediciones y sus respectivas estimaciones de error. No tenemos un conjunto de entrenamiento. Sin embargo, a diferencia de un problema de regresión usual, conocemos que la relación entre los datos medidos y los parámetros es (4.1). Aprovechando esto, los pasos a seguir para el entrenamiento son los siguientes:

1. Definimos el conjunto donde vamos a buscar los parámetros. En este caso:

$$\begin{cases} \frac{H_0}{100} \in (0.3, 1) \\ \Omega_m, \Omega_\Lambda \in (0, 1) \\ \Omega_k = 1 - \Omega_m - \Omega_\Lambda \in (-0.1, 0.1) \end{cases}$$

2. Tomamos uniformemente 5×10^3 muestras aleatorias de $(\frac{H_0}{100}, \Omega_m, \Omega_\Lambda)$ siguiendo los intervalos del paso 1. Para cada muestra, generamos grupos de 53 valores artificiales de $H^2(z_i) \sim N\left(H^2(z_i; \frac{H_0}{100}, \Omega_m, \Omega_\Lambda), \frac{2H(z_i)\sigma_i}{100}\right)$ usando la ecuación (4.1) evaluada para los mismos corrimientos al rojo z que las mediciones de $H(z)$ añadiendo sus errores proporcionales.
3. Entrenamos la red neuronal usando como datos de entrada los valores artificiales de $H^2(z)$ del paso 2, y como etiquetas, sus respectivos parámetros $(\frac{H_0}{100}, \Omega_m, \Omega_\Lambda)$.
4. A partir de las 53 mediciones naturales de $H(z)$ generamos 5×10^3 grupos de 53 valores añadiéndoles a cada medición un ruido $\delta \sim N(0, \frac{\sigma_i}{100})$ y elevando el resultado al cuadrado.
5. Para cada grupo de mediciones del paso 4, predecimos con la red neuronal un valor para $(\frac{H_0}{100}, \Omega_m, \Omega_\Lambda)$.
6. Con los 5×10^3 resultados del paso 5, establecemos un intervalo de confianza de 5σ para cada parámetro, actualizamos el conjunto del paso 1 y reentrenamos la red repitiendo los pasos 2 hasta 6, n veces.
7. Tras la última repetición del paso 6, guardaremos los últimos 5×10^3 resultados y los usaremos para estimar la distribución y los intervalos de confianza de los parámetros.

Mediante este procedimiento, la red aprenderá con los datos artificiales las relaciones entre los parámetros y mediciones dadas por el modelo Λ CDM, y, posteriormente, será capaz de predecir acertadamente los parámetros a partir de los datos reales.

Cabe destacar que, con el fin de reducir la variabilidad del método por la escasez de mediciones, repetiremos este proceso de entrenamiento para 10 redes neuronales. De modo que las estimaciones finales se tomarán a partir de 5×10^4 muestras.

4.5.5. Hiperparámetros de la red neuronal

La red neuronal tendrá una estructura con dos capas ocultas densamente conectadas de 128 y 10 neuronas respectivamente. Antes de cada capa oculta se aplicará una normalización de los valores intermedios [14] con el fin de estabilizar el proceso de entrenamiento. Usaremos como función de pérdida la media del error, y como función de activación aplicaremos la conocida como RReLU “Randomized leaky rectified linear unit” [28]:

$$rrelu(x) = \begin{cases} x & \text{if } x > 0 \\ ax & \text{otherwise} \end{cases}, \quad a = \begin{cases} \sim U(\frac{1}{8}, \frac{1}{3}) & \text{(Durante el entrenamiento)} \\ \frac{1/8+1/3}{2} & \text{otherwise} \end{cases}$$

La red será generada y entrenada usando la librería tensorflow [1], en python. Cada llamada de entrenamiento hará 50 iteraciones con lotes (o “batch”) de 512 grupos de datos. Por último, usaremos un optimizador de gradiente estocástico llamado Adam [15], una versión más sofisticada del descenso de gradiente básico que además aprovecha vectores de momento y el segundo orden del gradiente para la optimización.

4.5.6. Resultados

Una vez terminado el entrenamiento de las redes neuronales y el cálculo de predicciones, procedemos a comparar los valores obtenidos con los datos de Planck y las estimaciones de Metropolis-Hastings anteriores:

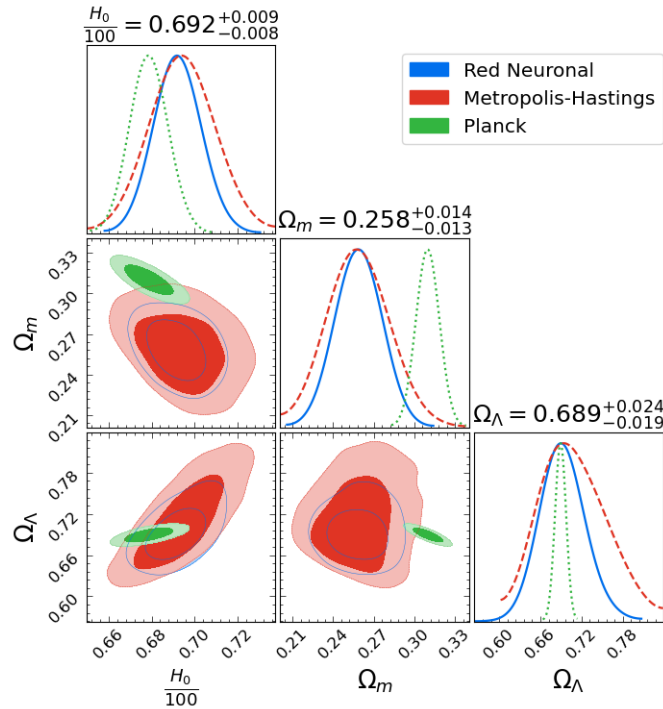


Figura 4.9

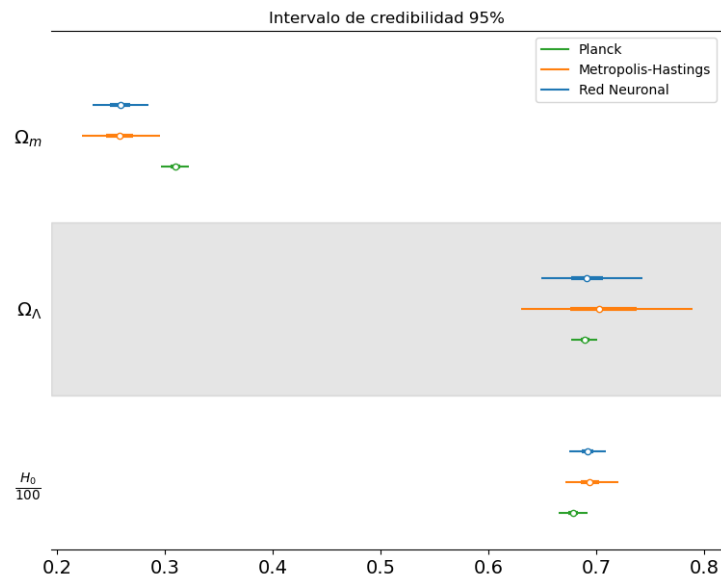


Figura 4.10

Las distribuciones estimadas por las redes neuronales concuerdan con las del algoritmo de Metropolis-Hastings manteniendo una menor varianza, y, vuelven a confirmar la tensión previamente expuesta entre las mediciones de $H(z)$ y los datos de Planck.

CAPÍTULO 5

Conclusión

En este trabajo se han propuesto tres métodos distintos para estimar los parámetros del modelo Λ CDM usando mediciones independientes del valor $H(z)$. El primer método descrito, centrado en el cálculo de la mediana, pese a su simpleza muestra resultados interesantes, pero no es demasiado relevante ya que no hace uso de la información que tenemos sobre los errores de las mediciones y no es ampliable a mayores conjuntos de datos. Comparando los dos últimos métodos, estos producen estimaciones parecidas. Con la pequeña cantidad de datos que tenemos, el uso de redes neuronales es generalmente peor que la aplicación de métodos de Cadenas de Markov Montecarlo. Esto se debe a que es más costoso computacionalmente y requiere de un extensivo ajuste de hiperparámetros. No obstante, la eficiencia y el rendimiento de las redes neuronales escala mejor con el aumento del volumen de datos, razón por la cual, con una cantidad de mediciones más extensa, las redes neuronales pueden acabar siendo una mejor alternativa para la estimación de parámetros cosmológicos.

Debido a la falta de precisión de los datos obtenidos por redes neuronales y Metropolis-Hastings, no se puede realizar ninguna declaración contundente respecto al parámetro Ω_k de la curvatura del universo. Sin embargo, si se puede afirmar que, tanto nuestros resultados como los del experimento Planck [20], predicen (usando el modelo Λ CDM) que nuestro universo continuará con una expansión acelerada indefinidamente. También, es importante destacar la discordancia surgida entre nuestras estimaciones y las de Planck [20], esto supone un indicio de las fracturas que están apareciendo en el modelo Λ CDM entre el estudio de supernovas y el fondo cósmico de microondas.

Por último, cabe recalcar que este trabajo queda claramente limitado por la pequeña cantidad de datos usados. Una mejora evidente sería añadir otros conjuntos de mediciones más extensas a los datos, esto añadiría la complejidad de incluir fórmulas de conversión entre las variables medidas pero mejoraría la precisión de los métodos. Sin embargo, también es importante notar que es necesaria la aparición de más mediciones precisas de la variable de Hubble, ya que esta presenta una conexión muy directa con la historia y el futuro de nuestro universo.

Bibliografía

- [1] M. Abadi et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. <https://www.tensorflow.org/>, 2015.
- [2] P. Bogacki, L.F. Shampine. A 3(2) pair of Runge - Kutta formulas. *Applied Mathematics Letters* Volume 2, Issue 4, 321-325, 1989.
- [3] A. Cauchy. Méthode générale pour la résolution des systèmes d'équations simultanées. *C. R. Acad. Sci. Paris*, 25:536–538, 1847.
- [4] W. de Sitter. On the relativity of inertia: Remarks concerning Einstein's latest hypothesis. *Proc. Kon. Ned. Acad. Wet.* 19, 1217-1225, 1917.
- [5] E. Di Valentino et al. In the Realm of the Hubble tension, a Review of Solutions. *Class. Quantum Grav.* 38, 153001, 2021.
- [6] A. Einstein. Zur allgemeinen Relativitätstheorie. *Sitzungsber. Preuss. Akad. Wiss. Berlin* 778-786, 1915.
- [7] A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin* 844-847, 1915.
- [8] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)*, 142-152, 1917.
- [9] G.F.R. Ellis Course on General Relativity http://www.math.uct.ac.za/sites/default/files/image_tool/images/32/Staff/Emeritus_Professors/Prof_George_Ellis/Overview/GR%20Lectures.pdf 2009.
- [10] A. Friedman. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10 (1), 377-38, 1922.
- [11] A. Friedman. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik* 21 (1), 326-332, 1922.
- [12] I.J. Goodfellow, Y. Bengio, A. Courville. Deep Learning. *MIT Press*. <http://www.deeplearningbook.org>, 2016.
- [13] E.P. Hubble. A relation between distance and radial velocity among extragalactic nebulae. *Proc. Nat. Acad. Sci. U.S.* 15, 69-173, 1929.

- [14] S. Ioffe, C. Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift arXiv:1502.03167, 2015.
- [15] D.P. Kingma, J. Ba. Adam: A Method for Stochastic Optimization arXiv:1412.6980, 2014.
- [16] G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles* A47, 49-59, 1927.
- [17] G. Lemaître. L'Univers en expansion. *Annales de la Société Scientifique de Bruxelles* A53, 51-85, 1933.
- [18] S. Pal, R. Saha. Estimation of Density Parameters Including Curvature and Hubble Constant for Λ CDM Universe Using Hubble Measurements. arXiv:2204.07099, 2022.
- [19] A.A. Penzias, R.W. Wilson. A Measurement Of Excess Antenna Temperature At 4080 Mc/s. *Astrophysical Journal Letters* 142: 419-421, 1965.
- [20] Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. *A&A* 641, A6, 2018.
- [21] C.P. Robert, G. Casella. Monte Carlo Statistical Methods. *Springer New York, NY* <https://doi.org/10.1007/978-1-4757-4145-2>, 1999.
- [22] H.P. Robertson. Kinematics and World-Structure. *Astrophys. J.* 82, 284-301, 1935.
- [23] H.P. Robertson. Kinematics and World-Structure II. *Astrophys. J.* 83, 187-201, 1936.
- [24] H.P. Robertson. Kinematics and World-Structure III. *Astrophys. J.* 83, 257-271, 1936.
- [25] B. Ryden Introduction to Cosmology http://carina.fcaglp.unlp.edu.ar/extragalactica/Bibliografia/Ryden_IntroCosmo.pdf
- [26] A.G. Walker. On Milne's theory of world-structure *Proceedings of the London Mathematical Society* 2 42 (1), 90-127, 1937.
- [27] G. Wang, S. Li, J. Xia. ECoPANN: A Framework for Estimating Cosmological Parameters using Artificial Neural Networks. *Astrophys. J. Suppl.* 249 no. 2, 25, 2020.
- [28] B. Xu et al. Empirical Evaluation of Rectified Activations in Convolution Network arXiv:1505.00853, 2015.

APÉNDICE A

Mediciones de $H(z)$

z	$H(z)$	$\sigma_{H(z)}$	Método
0.07	69	19.6	DA
0.09	69	12	DA
0.12	68.6	26.2	DA
0.17	83	8	DA
0.1791	75	4	DA
0.1993	75	5	DA
0.2	72.9	29.6	DA
0.24	79.69	2.99	BAO
0.27	77	14	DA
0.28	88.8	36.64	DA
0.30	81.7	6.22	BAO
0.31	78.18	4.74	BAO
0.34	83.8	3.66	BAO
0.35	82.7	8.4	BAO
0.3519	83	14	DA
0.36	79.94	3.38	BAO
0.38	81.5	1.9	BAO
0.3802	83	13.5	DA
0.4	82.04	2.03	BAO
0.4004	77	10.2	DA
0.4247	87.1	11.2	DA
0.43	86.45	3.97	BAO
0.44	84.81	1.83	BAO
0.4497	92.8	12.9	DA
0.47	89	34	DA
0.4783	80.9	9	DA

z	$H(z)$	$\sigma_{H(z)}$	Método
0.48	87.79	2.03	BAO
0.51	90.4	1.9	BAO
0.52	94.35	2.64	BAO
0.56	93.34	2.3	BAO
0.57	96.8	3.4	BAO
0.59	98.48	3.18	BAO
0.5929	104	13	DA
0.6	87.9	6.1	BAO
0.61	97.3	2.1	BAO
0.64	98.82	2.98	BAO
0.6797	92	8	DA
0.73	97.3	7	BAO
0.7812	105	12	DA
0.8754	125	17	DA
0.88	90	40	DA
0.9	117	23	DA
1.037	154	20	DA
1.3	168	17	DA
1.363	160	33.6	DA
1.43	177	18	DA
1.53	140	14	DA
1.75	202	40	DA
1.965	186.5	50.4	DA
2.3	224	8	BAO
2.33	224	8	BAO
2.34	222	7	BAO
2.36	226	8	BAO

Tabla A.1: Mediciones de $H(z)$ en $(km/s)/Mpc$ y su correspondiente incertidumbre $\sigma_{H(z)}$, medidas con los métodos BAO y DA a distintos corrimientos al rojo z (véase página 6 de [18]).