

В. Жданов, д-р физ.-мат. наук,  
А. Сташко, асп.  
Киевский национальный университет имени Тараса Шевченко, Киев

### ПАРАМЕТР ХАББЛА В $f(R)$ -ГРАВИТАЦИИ

*Изучены условия на лагранжианы модифицированной гравитации, имитирующие наблюдаемые величины стандартной космологической модели. Рассмотрен вопрос выбора лагранжиана  $f(R)$ -гравитации на основе диаграммы Хаббла. Получены уравнения для  $f(R)$ , которые позволяют определить эту величину для заданного параметра Хаббла  $H(z)$ . Рассмотрена и обратная задача определения  $H(z)$  для заданной функции  $f(R)$ , которая мало отличается от выражения общей теории относительности; найдено общее приближенное выражение для  $H(z)$ .*

*Ключевые слова: космология, параметр Хаббла,  $f(R)$ -гравитация.*

V. Zhdanov, Dr Hab.,  
O. Stashko, PhD Student  
Taras Shevchenko National University of Kyiv, Kyiv

### HUBBLE PARAMETER IN $f(R)$ -GRAVITY

*In view of the famous problem with the "Hubble constant tension" there is a number of approaches to modify the cosmological equations and correspondingly modify Hubble parameter  $H(z)$  in order to relieve the tension between the "early" and "late" Hubble constants.  $f(R)$ -gravity is one of such possible modifications. We discuss how to choose the Lagrangian in the  $f(R)$ -gravity on account of observational data within the homogeneous isotropic cosmology. The equation is obtained that enable us to derive  $f(R)$  for given Hubble parameter  $H(z)$ .*

$$(z+1)H^2 \frac{d^2 f'}{dz^2} + \left( \frac{1+z}{2} \frac{dH^2}{dz} + 2H^2 \right) \frac{df'}{dz} - f' \frac{dH^2}{dz} = -\frac{\kappa}{3} \frac{dT_{00}}{dz}; \quad z \text{ is the redshift, } f' \equiv df/dR \text{ and } R \text{ expressed in the usual way by}$$

*means of  $H$  and its derivatives. This yields a second order differential equation with corresponding degrees of freedom. If  $H(z)$  corresponds to that obtained from usual Friedmann equations, this equation yields a condition for  $f(R)$  to mimic the observable quantities of the standard  $\Lambda$ CDM with the above-mentioned freedom. To reduce this freedom on needs additional considerations, which involve the other observable quantities, such as those which appear in considerations of cosmological perturbations on the isotropic and homogeneous background. Also, we consider the reverse problem to find  $H(z)$  for given  $f(R)$ . This is fulfilled within an approximation in case of small deviation of  $f(R)$  from the General Relativity value.*

*Key words: cosmology, Hubble parameter,  $f(R)$ -gravity.*

УДК 524.8

С. Парновський, д-р фіз.-мат. наук  
Київський національний університет імені Тараса Шевченка, Київ

### МАТЕМАТИЧНІ ПРОБЛЕМИ, ПОВ'ЯЗАНІ З ПОХИБКАМИ ОЦІНКИ ВІДСТАНЕЙ ДО ГАЛАКТИК

*Розглянуто деякі суто математичні проблеми, що виникають у процесі створення штучних вибірок даних при додаванні похибок до вихідних даних, створених за законом Габбла. Є дві різні можливості за генерації у випадку, коли відхилення у процесі оцінювання відстаней є пропорційними до них і відносна похибка визначення відстані є сталою. Обидві можуть бути застосовані на практиці, але їхні математичні властивості суттєво відрізняються. Крім обговорення проблем, що виникають для обох варіантів, розглянуто формули для опрацювання штучних каталогів за методами найменших квадратів (МНК) та максимальної правдоподібності (ММП). Показано, що формули МНК можна використовувати у ході застосування одного з методів додавання похибок, але кут нахилу оптимальної пропорційної залежності буде недооцінений. Водночас ММП можна використовувати лише для іншого методу додавання похибок. Отримано відповідні формули й оцінки.*

*Ключові слова: статистичні методи опрацювання даних, метод найменших квадратів, метод максимальної правдоподібності.*

**Вступ.** У роботі [1] я розглянув деякі проблеми за визначення величини сталої Габбла, пов'язані з похибками в оцінках відстаней до галактик. Нехай ми маємо вибірку, що містить дані про відстані та швидкості  $N$  галактик з  $z \ll 1$ . Нам треба опрацювати ці дані та визначити з них величину параметра Габбла. Це просто зробити у ідеальному випадку, коли відсутні пекулярні рухи галактик і їх швидкості, визначені за червоним зміщенням, збігаються з габблівськими швидкостями  $v_i$ . А ті є пропорційними відстаням до галактик  $r_i$  згідно із законом Габбла:

$$v_i = Hr_i, \quad (1)$$

де  $H$  – параметр Габбла. Його значення у сучасну епоху є сталою Габбла. Визначити її за набором даних  $v_i$  та  $r_i$  дуже просто. На жаль, за реальних астрономічних спостережень ми визначаємо інший набір даних. До габблівських швидкостей додаються радіальні пекулярні швидкості, які ми вважатимемо випадковими. Як наслідок, для кожної галактики вимірюємо величину  $V_i$ , яка за припущенням така:

$$V_i = v_i + \delta V s_i, \quad (2)$$

де  $s_i$  – випадкова величина з нормальним розподілом, нульовим середнім та одиничною дисперсією, а  $\delta V = \text{const}$ . Натомість, замість справжніх відстаней до галактик  $r_i$  ми маємо їх оцінки  $R_i$ , незалежні від червоного зміщення галактики, отримані за допомогою якоїсь статистичної залежності на кшталт відстані за цефеїдами, червоними

гігантами, за залежностями Таллі-Фішера,  $D_n - \sigma$ , за фундаментальною площиною або за флуктуаціями поверхневої яскравості. Похибка визначення відстані  $R_i - r_i$  збільшується з відстанню до галактики. Вважається, що відносна похибка визначення відстаней приблизно стала та складає 15–30 % залежно від методу оцінювання.

У роботі [1] я для вивчення впливу похибок застосовував математичне моделювання за методом Монте-Карло. Для цього я брав початковий набір даних  $v_i$  та  $r_i$ , що точно задовольняють закон Габбла (1) та додавав до них похибки за законом (2) та

$$R_i = r_i(1 + ap_i), \tag{3}$$

де  $p_i$  – інша випадкова величина з нормальним розподілом, нульовим середнім та одиничною дисперсією, а сталий параметр  $a$  характеризує точність методу оцінювання.

На відміну від згаданої роботи в цій статті розглянуто деякі математичні аспекти опрацювання зашумленого набору даних  $V_i$  та  $R_i$ . Зокрема, результати його опрацювання за методом найменших квадратів (МНК) та можливість застосування методу максимальної правдоподібності (ММП). Також я обговорю деякі проблеми, пов'язані з формулою (3) та її альтернативним варіантом, що має вигляд

$$R_i = r_i(1 - ap_i)^{-1}. \tag{4}$$

За малих значень  $a$  він близький до (3) з практичного погляду. Похибка в (4) дорівнює в середньому певному відсотку від  $R_i$ , а в (3) від  $r_i$ . Але між ними є принципова різниця з погляду математики.

**Проблеми з похибками, визначеними за формулами (3) та (4).** Почнемо з формули (3). Якщо значення  $r_i$  фіксоване, то розподіл величини  $R_i$  є нормальним. Однак є ненульова вірогідність того, що ця величина стане нульовою або від'ємною, що не має сенсу. Проте ця проблема притаманна застосуванню нормального розподілу у будь-якому випадку.

Справжня проблема виникає, якщо ми маємо зі спостережень певну величину  $R_i$  і розглядаємо розподіл справжньої відстані  $r_i$ . З виразу для величини

$$\exp\left(-\frac{p_i^2}{2}\right) = \exp\left(-\frac{(R_i - r_i)^2}{2a^2 r_i^2}\right) \xrightarrow{r_i \rightarrow \infty} \exp\left(-\frac{1}{2a^2}\right) \neq 0 \tag{5}$$

бачимо, що густина ймовірності розподілу  $r_i$  не прямує до нуля при  $r_i \rightarrow \infty$ . Тому її неможливо нормувати.

Розподіл  $r_i$  для (4) не має цієї вади, але має інші. При фіксованому значенні  $r_i$  розподіл величини  $R_i$  не є нормальним та ще й є несиметричним. Середнє значення цієї величини та її степенів виявляється нескінченним. Доведемо це, узявши до відома, що середнє значення непарних степенів величини  $p_i$  дорівнює нулю, а для непарних маємо

$$\langle p_i^{2n} \rangle = (2n - 1)!! . \tag{6}$$

Тут кутові дужки позначають усереднення за нормальним розподілом, а  $(2n-1)!!$  є подвійним факторіалом, тобто добутком усіх непарних чисел від 1 до  $2n-1$ . Тому для середніх значень, отриманих із (4), маємо

$$\begin{aligned} \langle R_i \rangle &= r_i \langle (1 - ap_i)^{-1} \rangle = r_i \langle 1 + ap_i + a^2 p_i^2 + a^3 p_i^3 + a^4 p_i^4 + a^5 p_i^5 + a^6 p_i^6 + \dots \rangle = \\ &= r_i (1 + a^2 + 3a^4 + 15a^6 + \dots) = r_i \sum_{n=0}^{\infty} (2n - 1)!! a^{2n}; \quad \langle R_i^2 \rangle = r_i^2 \sum_{n=0}^{\infty} (2n + 1)!! a^{2n}. \end{aligned} \tag{7}$$

Обидва ці ряди розбігаються. Це видно з того, що відношення двох послідовних членів дорівнює  $(2n-1)a^2$  та  $(2n+1)a^2$  і при ненульовому  $a$  перевищує 1, якщо  $n$  великі.

**Застосування МНК.** Нехай маємо набір даних  $V_i$  та  $R_i$  для змодельованої вибірки, створеної за допомогою (2) та (3) з даних  $v_i$  та  $r_i$ , що задовольняють (1). До цього набору можна застосувати стандартну формулу МНК для проведення оптимальної прямої  $V=AR$ , що проходить через початок координат та апроксимує дані, а саме:

$$A = \frac{\sum_{i=1}^N V_i R_i}{\sum_{i=1}^N R_i^2}. \tag{8}$$

Величина  $A$ , отримана з (8) відрізняється від  $H$  у (1). Покладемо  $A=kH$  та знайдемо середнє значення коефіцієнта  $k$  за формулами (1, 2, 3, 6, 8):

$$k = \left\langle \frac{\sum_{i=1}^N (Hr_i + \delta V s_i) r_i (1 + ap_i)}{H \sum_{i=1}^N r_i^2 (1 + ap_i)^2} \right\rangle \approx \frac{1}{1 + a^2}. \tag{9}$$

Усереднене значення чисельника дорівнює  $H \sum_{i=1}^N r_i^2$ , знаменника –  $H(1 + a^2) \sum_{i=1}^N r_i^2$ , рівняння приблизне через

слабку кореляцію відхилень  $p_i$  у чисельнику та знаменнику.

За спроби застосувати цей підхід до залежності (4) ми отримуємо відношення сум двох рядів, що розбігаються, тобто він є некоректним.

**Застосування ММП.** У цьому випадку метод неможливо застосувати за залежності (3) через брак нормування згідно з (5). Формально його можна застосувати за зв'язку (4) таким чином. Імовірність того, що галактика, яка має справжню відстань  $\xi$  та швидкість  $\eta$ , буде мати оцінку відстані  $R_i$  та швидкість  $V_i$ , за спостереженням

$$w(R_i, V_i, \xi, \eta) = \frac{1}{2\pi\delta VaR_i} \exp\left(-\frac{s_i^2 + p_i^2}{2}\right) \propto \exp\left(-\frac{(V_i - \eta)^2}{2\delta V^2}\right) \exp\left(-\frac{(R_i - \xi)^2}{2a^2 R_i^2}\right). \quad (10)$$

Тому ймовірність спостереження галактики з  $R_i$  та  $V_i$  за умовою, що  $\xi$  та  $\eta$  пов'язані залежністю  $\eta = A\xi$  є

$$P_i(R_i, V_i, A) = \int_0^\infty w(R_i, V_i, \xi, A\xi) d\xi \propto \int_0^\infty \exp\left(-\frac{(V_i - A\xi)^2}{2\delta V^2}\right) \exp\left(-\frac{(R_i - \xi)^2}{2a^2 R_i^2}\right) d\xi. \quad (11)$$

Підінтегральна функція швидко спадає далеко від точок спостереження, тому можна покласти нижню межу інтегрування  $-\infty$ . Повна ймовірність для всіх  $N$  галактик така:

$$P(A) = \prod_{i=1}^N P_i(R_i, V_i, A) \propto \prod_{i=1}^N \int_{-\infty}^\infty \exp\left(-\frac{(V_i - A\xi)^2}{2\delta V^2}\right) \exp\left(-\frac{(R_i - \xi)^2}{2a^2 R_i^2}\right) d\xi. \quad (12)$$

Треба знайти, за якого значення  $A$  ця величина максимальна. Знаходимо похідну від логарифма цього виразу по  $A$  та прирівнюємо її до нуля. Отримуємо умову

$$\sum_{i=1}^N \frac{\int_{-\infty}^\infty (V_i - A\xi) \xi F(\xi, A, V_i, R_i) d\xi}{\int_{-\infty}^\infty F(\xi, A, V_i, R_i) d\xi} = 0, \quad F(\xi, A, V_i, R_i) = \exp\left(-\frac{(V_i - A\xi)^2}{2\delta V^2}\right) \exp\left(-\frac{(R_i - \xi)^2}{2a^2 R_i^2}\right). \quad (13)$$

Її можна розв'язати чисельно.

**Висновки.** Розглянуто два методи додавання похибок до вихідних даних для отримання згенерованих вибірок для застосування методу Монте-Карло. Розглянуто їхні деякі математичні особливості. Вони не заважають використовувати обидва методи, але забороняють застосування певних методів теоретичного опрацювання даних, отриманих таким чином. Зокрема, застосування МНК для похибок вигляду (3) призводить до недооцінювання величини кута нахилу. Відповідні оцінки зроблено у (9). Якщо похибки вигляду (4), то оцінки виконати неможливо через нескінченність деяких математичних очікувань.

Метод максимальної правдоподібності можна використовувати тільки для похибок вигляду (4). Для нього виведено складне нелінійне рівняння (13). У статті не обговорено результати математичного моделювання, а тільки деякі аспекти, пов'язані з їхнім застосуванням та особливостями опрацювання даних.

Ця робота підтримана Національним фондом досліджень України за проектом № 2020.02/0073.

#### Список використаних джерел

1. Парновський С. Л. Вплив статистичних ефектів на величину сталої Габбла, отриману за швидкостями галактик / С. Л. Парновський // Вісник Київ. нац. ун-ту ім. Тараса Шевченка. Астрономія. – 2020. – 61. – С. 20–22.

Надійшла до редколегії 21.10.20

S. Parnovsky, Dr Hab.  
Taras Shevchenko National University of Kyiv, Kyiv

### МАТЕМАТИЧЕСКИЕ ПРОБЛЕМЫ, СВЯЗАННЫЕ С ПОГРЕШНОСТЯМИ ОЦЕНКИ РАССТОЯНИЙ ДО ГАЛАКТИК

*Рассмотрены некоторые чисто математические проблемы, возникающие при создании искусственных выборок данных при добавлении ошибок в исходных данных, подчиняющихся закону Хаббла. Существуют две возможности, каждую из которых можно применять на практике при генерации в случае, когда отклонения при оценке расстояний пропорциональны им и относительная погрешность определения расстояния является постоянной. Их математические свойства существенно отличаются. Кроме обсуждения проблем, возникающих для обоих вариантов, рассмотрены формулы для обработки искусственных каталогов по методам наименьших квадратов (МНК) и максимального правдоподобия (ММП). Показано, что формулы МНК можно использовать при одном из методов добавления ошибок, но угол наклона оптимальной пропорциональной зависимости будет недооценен. В то же время ММП можно использовать только для другого метода прибавления погрешностей. Получены соответствующие формулы и оценки.*

*Ключевые слова: статистические методы обработки данных, метод наименьших квадратов, метод максимального правдоподобия.*

S. Parnovsky, Dr Hab.  
Taras Shevchenko National University of Kyiv, Kyiv

### MATHEMATICAL PROBLEMS ASSOCIATED WITH ERRORS IN ESTIMATING DISTANCES TO GALAXIES

*I generate many mock samples for applying the Monte Carlo method in order to estimate the bias of the Hubble constant because of the use of estimates of distances to galaxies determined from statistical dependences. I add errors to the original sample generated according to the Hubble law. In doing so, I use two possible options for generating errors in distance, having a constant relative error. Both are practical, but there are some math problems with them. I discuss their effect on the properties of the mock sample. The application of the standard least squares method is discussed and shown that it leads to an underestimation of the slope in the Hubble law. A formula is derived for calculating this slope using the maximum likelihood method and it is shown that it is applicable only for one of the variants of the sample noising. All estimates were obtained theoretically, without using the results of mock samples processing, which are described in a separate paper.*

*Key words: statistical methods of data processing, least squares method, maximum likelihood method.*