Hubble Constant and the Age of the Universe

Abstract:

This is a presentation of the analysis of data for 10 Cepheids as well as the recession velocities and extinction in the Milky Way for 8 different galaxies. Using the Cepheid Period-Luminosity Relationship, obtained by analysing the given data (parallax and its error, pulsating period, magnitude and extinction towards the corresponding star), the distance modulus for 8 nearby galaxies was estimated as well as the errors on these. By propagating the given errors and using model fitting (in Python) the expansion rate of the Universe (Hubble's Law) was estimated to be approximately: 72.0104 ± 3.9625 km/s/Mpc. Finally the age of the Universe was estimated to be 13.580202 ± 0.7472752 billion years.

Overall Analysis:

The whole analysis of the given data in order to reach the final estimation of the age of the Universe was separated into four simple steps.

The goal on the **first step** was to determine the relation between the pulsation period and the luminosity of 10 different Cepheid stars. The luminosity(or brightness) of an Astronomical object is usually measured in magnitudes. In the following analysis apparent and absolute magnitude are both used. The absolute magnitude (M) is equal to the apparent magnitude (m) of an object when viewed from a distance of exactly 10 parsecs (1pc = 3.086×10^{16} m or 1pc = 3.3 light years). More specifically, in astronomy the intrinsic brightness of a star (the brightness it would have if it was not attenuated by distance or intervening gas or dust) is the absolute magnitude and the apparent brightness is the apparent magnitude[1]. Magnitudes are logarithmic units so the difference between the (which is the distance modulus, μ) scales with the logarithm of the distance. For that reason the absolute magnitude was plotted against the logarithm of Period. However, gas and dust between the observers and the object can make the object look dimmer, so it is necessary for the extinction (A) to be taken into account, as well as its error. The relationship between the apparent and absolute magnitudes, the distance (in pc) and the extinction is:

Equation 1:

$$\mu = m - M = log d_{pc} - 5 + A$$

Using equation 1 and the given data for the parallax, the apparent magnitude and the extinction, the absolute magnitude of all 10 Cepheid stars were measured. This way the dependence of brightness was modelled by plotting the absolute magnitude found against the logarithm of Period in days. By finding the best fitting model to that plot, the relationship between the absolute magnitude and the logarithm of period is found to be linear. The best fitting slope (α) in that line and the best intercept (β) found by the model fitting, are used for the following parts.

The **second step** was a transitional necessary step in order to find the rate of expansion of the universe. The same equation was used to determine the distances for a set of 8 nearby galaxies. This time the best slope, α and intercept, β found in the first step were used in combination with the data given for the logarithm of period and the apparent magnitude for each of 25 Cepheids, as well as the extinction of the Milky Way towards each of the 8 galaxies. The desirable result in this part is the estimation of the distance modulus, the distances and its uncertainty for each galaxy. By propagating the errors given, and found in step 1 the uncertainty of the distance moduli and distances were found. It is possible that the resulted distances are affected by the extinction in each galaxy as we only take into account the dimming in our galaxy, Milky Way and not the extinction of the Cepheid's host galaxy.

In the analysis below (step 2) the propagation of the errors and the explanation of the way to find the uncertainties of the distances.

The **third step** was all about finding the best fitting model when plotting the recession velocity against the distances found in step 2. The recession velocity for each galaxy which is the rate at which it recedes from us due to the expansion of Universe is given. This way, using the simple relation between the velocity and distance (Hubble's Law):

Equation 2:

$$v_{rec} = H_0 D_{gal}$$

the expansion rate of the Universe can be determined by finding the best fitting slope in the linear graph of velocity against the distances measured. It is important to mention that the error of the recession velocity is not given so the estimation of velocities might not be very accurate, even though the measurement of the recession velocity is really hard as there are many local effects. In order to get more precise results it could be an advantage to plot the logarithm of both sides of equation 2 in order to find Hubble constant.

The **fourth and final step** is the estimation of the age of the universe using the Hubble constant calculated in the previous step. The theory that the Universe started as a single point and started to expand lets us use the single relation: $d = vt$, where d is distance, v is velocity and t is time. So using the distanced calculated in step 2 and the recession velocities of these nearby galaxies the age of the expanding Universe is:

Equation 3:

$$\tau = \frac{D_{gal}}{v_{rec}} = \frac{1}{H_0}$$

However this is just an approximation as the recessional velocity is not constant with time, it is a really good approximation for the purposes demanded.

Stage 1: The Cepheid Period- Luminosity

From the file named "MW_Cepheids.dat" the absolute magnitude for each Cepheid star was measured using the parallax, the apparent magnitude and the extinction. The absolute magnitude for each Cepheid was first calculated using equation 1. In this file the errors of the parallax and the extinction were given so the errors of each of the magnitudes calculated were propagated the following way.

At first the distance in parsec was found since: $d_{pc} = \frac{1000}{p_{mas}}$*, the absolute magnitude:*

$M = m - A + 5 - \log(d_{pc})$ *and the logarithm of that distance and in each period given was calculated using package* numpy*, as they are used in the analysis.*

After that, using the general formula of the error: *Equation 4:*

$$\sigma_{\hat{f}}^2 = \left(\frac{\partial f}{\partial x}\bigg|_{\hat{x}}\right)^2 \sigma_{\hat{x}}^2 + \left(\frac{\partial f}{\partial y}\bigg|_{\hat{y}}\right)^2 \sigma_{\hat{y}}^2 + \left(\frac{\partial f}{\partial z}\bigg|_{\hat{z}}\right)^2 \sigma_{\hat{z}}^2 + \dots$$

So for the absolute magnitude: $\sigma_M = \sqrt{(\frac{\partial M}{\partial A} \times \sigma_A)^2 + \left(\frac{\partial M}{\partial p_{mas}} \times \sigma_{p_{mas}}\right)^2}$

and after the calculations. , $\sigma_M = \sqrt{\sigma_A^2 + (\frac{5\sigma_{p_{mas}}}{p_{mas}\ln(10)})^2}$.

In the code given the σ_M = sig_M, σ_A = A_e, p_{mas} = p and the σ_p = pe. Another part that is important to mention is that the Periods given do not contain their error so it would fit the line better if the weighted mean was subtracted from the values plotted so that the value of χ2 is closer to the expected value. For that reason the mean of the logarithm of periods was calculated and subtracted from xdata in curve fit. In order to obtain the relationship between the luminosity and the period of the Cepheids the absolute magnitude calculated was plotted against the logarithm of the period given for each star. It is proved that the dependence of brightness on the period is ideally modelled using the logarithm of period. For that reason the logP (logarithm of period with base 10) is calculated and used to plot the absolute magnitude against the logP. It is important to mention that the errors of apparent magnitudes and periods are not provided which means that they can be ignored (mentioned in the assignment quide). Using the function curve_fit the best fitting line in the plot was found as well as the best slope and intercept (α and β), so the relation is:

Equation 5:

$$M = \alpha logP + \beta$$

A starting slope and intercept were set in order to use the curve fit. A simple function (func) was made to calculate the model and was used for the fit. This way the best slope(α) and intercept(β) were calculated as well as the correlation matrix. Apart from that a function for the calculation of χ^2 was defined (called chisq) using the formula: $\chi^2 = \sum_{i=1}^{N} \frac{(M-M_{best})^2}{\sigma_M^2}$ *, where M_{best} are the best fitted values for the absolute magnitude found from the model function. In order to check if the model is good, apart from the degrees of freedom, it is useful to check if:* $\chi^2 \simeq N \pm \sqrt{2N}$*, where N are the degrees of freedom. In this case N = 9, so* $N - \sqrt{2N} = 4.75736$ *and* χ^2 *=4.279452671456861, so the reduced χ_v^2 should be approximately* $1 - \sqrt{\frac{2}{N}} = 0.5286$ *and therefore the model is accepted as: χ_v^2 = 0.475495*

These values are:

The covariance matrix is: [0.00099941 0.00099941]

CURVE_FIT RESULTS:
Best-fitting alpha = -2.4006676323198617
Best-fitting beta = -3.6809135336423595
Corresponding χ^2= 4.279452671456861
Corresponding Reduced χ_v^2 = 0.4754947412729845

Since the values of the covariance matrix are really close to zero and no to 1 we conclude that the two fit result parameters, α and β are not correlated. Also it is important that the value of chi square is the expected for a good fit as it not much smaller than the degrees of freedom (dof). Generally the fitting of the slope is good so the fit can be considered acceptable. The best fit is:

$$M = (2.40066763232 \pm 0.23218943583) \times logP + (3.68091353364 \pm 0.049584242272)$$

And the final graph:

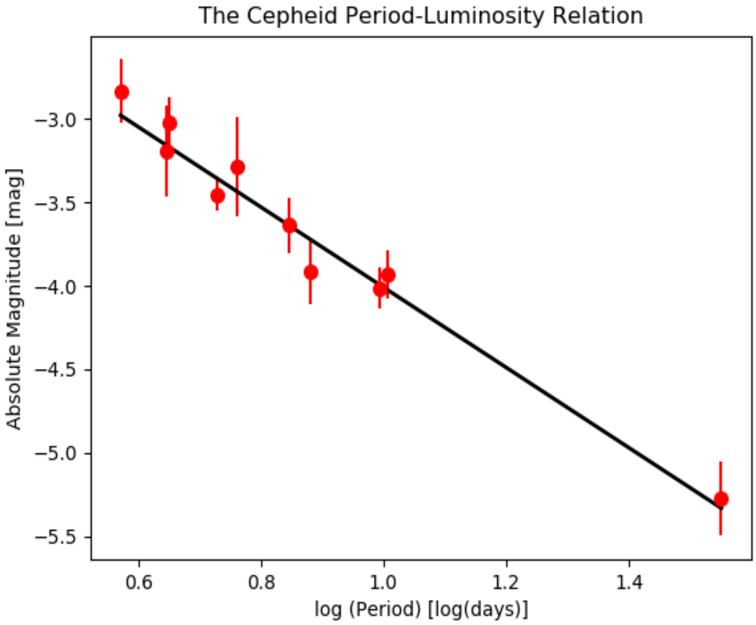

Figure 1

Stage 2: Distances for a Set of Nearby Galaxies

This step is probably the most important part of the analysis as most of our data for the 8 galaxies are used, and the distances that are going to be used for the calculation of the expansion rate of the universe are calculated. The relationship found in step 1 (Equation 5) *was used to measure the absolute magnitude of all the Cepheids for which data were provided. The data file named 'hst_galI_cepheids.dat' , where I is the number of the galaxy included the logarithm of period and the apparent magnitude for 25 Cepheids per galaxy. Also an estimate an estimate of the extinction towards the Milky Way is provided in file 'galaxy_data.dat' for all 8 galaxies, we were able to calculate the distance modulus and the distances. The procedure used to propagate the error of α and β, found in step 1, to distances is:*

Using the same general formula[3] (Equation 4) *the error of the absolute magnitude (from the Period-Luminosity Relation) is found to be:*

$$\sigma_M = \sqrt{\left(\frac{\partial M}{\partial \alpha} \times \sigma_\alpha\right)^2 + \left(\frac{\partial M}{\partial \beta} \times \sigma_\beta\right)^2} = \sqrt{(\sigma_\alpha \times logP)^2 + \sigma_\beta^2}$$

And so the error of the distance (Dgal) is:

$$\sigma_{D_{gal}} = \sqrt{\left(\frac{\partial D_{gal}}{\partial M} \times \sigma_M\right)^2} =$$

$$\sqrt{\left(\frac{ln(10)\times 10^{\frac{\mu-A+5}{5}}}{5}\right)^2 \times ((logP \times \sigma_\alpha)^2 + \sigma_\beta^2)} = D_{gal} \times \frac{ln10}{5} \times \sqrt{((logP \times \sigma_\alpha)^2 + \sigma_\beta^2)}$$

In the code the corresponding values are: logP = logPer (array), Dgal= distances (array), σ_α=alpha_err, σ_β = beta_err, A = A_mw.

The error of the distance moduli is also propagated:

$$\sigma_\mu = \sqrt{\left(\frac{\partial\mu}{\partial M} \times \sigma_M\right)^2} = \sqrt{(-\sigma_M)^2} = \sigma_M$$

Note: all the formulas might have been modified when written in python is order to simplify the code.

So the uncertainty of the distance modulus for each galaxy is equal to that of the distances. The mean of the logarithm of Periods calculated in step 1 was also used in this part as the logarithm of periods given do not have errors. For that reason the mean was subtracted from each value on the array logPer. The following table present the printed values of the code for distance moduli and distances for each galaxy:

The distance for galaxy NGC3627 is: 10.2383921577 ± 0.6308325248 Mpc
and the distance modulus μ for that galaxy is: 30128424.781708352 ± 0.6308325248 pc
The distance for galaxy NGC3982 is: 21.1957004310 ± 1.4232213388 Mpc
and the distance modulus μ for that galaxy is: 31616561.912454434 ± 1.4232213388 pc
The distance for galaxy NGC4496A is: 13.7091370446 ± 0.9413343894 Mpc
and the distance modulus μ for that galaxy is: 30742463.593286466 ± 0.9413343894 pc
The distance for galaxy NGC4527 is: 14.4008507697 ± 0.9927763783 Mpc
and the distance modulus μ for that galaxy is: 30832357.05876294 ± 0.9927763783 pc
The distance for galaxy NGC4536 is: 14.5239175997 ± 1.0708267256 Mpc
and the distance modulus μ for that galaxy is: 30848222.742292807 ± 1.0708267256 pc
The distance for galaxy NGC4639 is: 19.7436256789 ± 1.4077884685 Mpc
and the distance modulus μ for that galaxy is: 31521023.944764968 ± 1.4077884685 pc
The distance for galaxy NGC5253 is: 3.6909356346 ± 0.0868486229 Mpc
and the distance modulus μ for that galaxy is: 27991683.528507292 ± 0.0868486229 pc
The distance for galaxy IC4182 is: 4.1499030085 ± 0.1538146762 Mpc
and the distance modulus μ for that galaxy is: 28115171.68127109 ± 0.1538146762 pc

Although the propagation of the errors in this part is made carefully (and following the formula given in this course) the final errors can not be considered Gaussian. In order to have Gaussian errors with 68% confidence the $\Delta\chi^2$ must be equal to 1, and that is not the case here.

Stage 3: Expansion Rate of the Universe

In this step the distances measured above are used for the calculation of the expansion rate of the Universe i.e. Hubble Constant H_0. Since the Universe started as a single point, as mentioned before, Hubble's Law (Equation 2), can be used to calculate the expansion rate of Universe, as the recession velocity is actual due to the expansion. For that reason the constant of proportionality, H_0, equals. To that rate.

Using the errors found when propagating the uncertainties the recession velocity was plotted against the distances found above. The curve_fit function is used again for the model fitting in this plot but this time the intercept was set equal to zero from the beginning as the equation we use to find the best model (Equation 2) is a simple proportionality equation without intercept. The function used to find the model of the fit is called func2 and is used for the second curve_fit (fit2, covar2), in order to find the best model (best fitted distances to the given velocities). The final plot is:

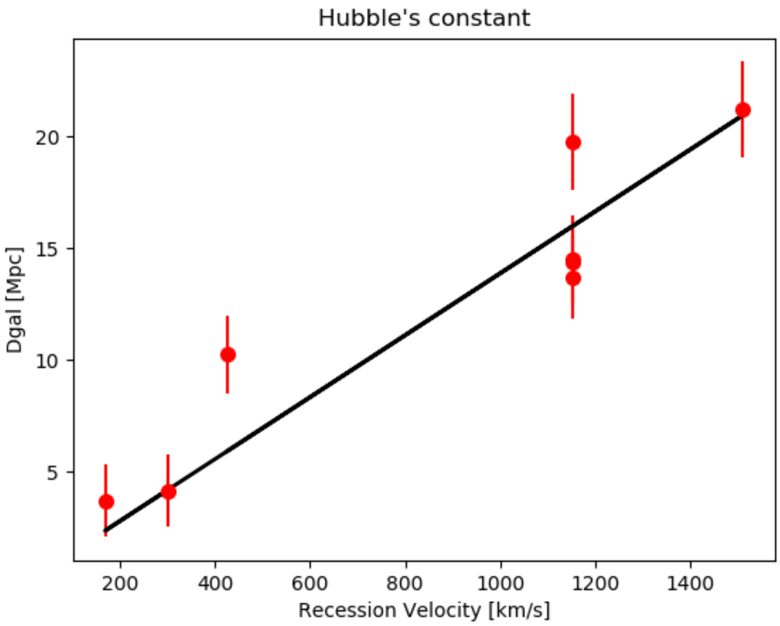

Figure 2

So our final preferred fit is: $v_{rec} = (72.01036392211586 \pm 3.9625005246844793) \times D_{gal}$

At this point it is important to say that the results of the curve fit and the value of chi square was much larger than 1 ($\chi^2 \gg 1$). For that reason the uncertainty has to change in order in order to accept the model. So in order to find a closer value to the expected for the Hubble constant, the intrinsic dispersion is used, so a constant uncertainty was added in the errors found, using the approximation:

$$\sigma_i^2 \rightarrow \sigma_i^2 + \sigma_{int}^2$$

The best value for σ_i is found by trying multiple different values (σ_in), until the reduced χ_v^2 is as closer to 1 as possible. So it is found to be: $\sigma_i = 1.8$. In general the σ_i does not rescale the overall χ^2 and this is the reason the reduced chi square was checked instead[3]. The reason of that non-fitting model is probably due to the error of the recession velocity. All galaxies are fairly nearby so the recession velocity could be really hard to measure, especially with the all effects between the 8 galaxies.

As the distance against the velocity was plotted instead of velocity vs distance the model function had the form: $D_{gal}=c\times v_{rec}$ where c = constant = $1/H_0$, the Hubble's constant was calculated using the 1/α, where α, is the best slope of the fitted line and its error. So the results of the second curve fit are:

CURVE_FIT 2 RESULTS:
Best-fitting alpha = 0.013886889963249853
Corresponding χ^2 = 10.713788947868288
Corresponding Reduced χ^2 = 1.5305412782668983

And finally, by calculating the inverse of slope (α) the Hubble constant and its error is calculated to be:

H_0 = 72.01036392211586 ± 3.9625005246844793 km/s/Mpc

The expected value of Hubble's constant is actually 70 km/s/Mpc[2] while the accepted range of the expansion rate is 65-80km/s/Mpc[4], so the estimated value in this analysis is really close to the theoretical/expected value including the uncertainty. In code: Hubble error = sigma_H, H_0 = 1/best_α (since α =slope=$1/H_0$).

Stage 4: Age of Universe

The final step is a simple relation of the formula (proportionality) found in step 3 with time. The time an object takes to move distance d, with velocity v is: $t = \frac{d}{v}$, so by using the recession velocities of the 8 galaxies which are given, and the distances calculated in the previous stages, the age of the expanding universe can be estimated, using Equation 3:

$$\tau = \frac{D_{gal}}{v_{rec}} = \frac{1}{H_0}$$

So the age of the universe could be estimated easily as the gradient of the graph plotted (Fig. 2) is actually the value $1/H_0$. In order to convert the time in Gigayears (billion years) the best fitting alpha (which is in seconds) is multiplied with 3.086x3.16887646)x(10^2). Also the uncertainty of that estimation is found by the square root of the top left element of the covariance matrix of fit2. In the code the corresponding values are: D_{gal}= distances(array), v_{rec}=vrec, σ_τ = time error = sigma_τ.

The final value of the age of the Universe is:

$$\tau = 13.580201825027327 \pm 0.7472751688242097 \; billion\ years$$

In general the analysis done in this assignment could be improved using other, more advanced methods of data analysis with the given data. However, the final requested values were found to be a good approximation of the measurements already taken, since the Big Bang today is estimated that happened between 12 and 14 billion years ago[4].

References:

1. *''Magnitude (Astronomy).'' Wikipedia, Wikipedia Foundation, 15 May 2020,* *https://en.wikipedia.org/wiki/Magnitude_(astronomy)*
2. *''Hubble's Law.'' Wikipedia, Wikipedia Foundation, 22 May 2020,* *https://en.wikipedia.org/wiki/Hubble%27s_law*
3. *Most of the information used for this analysis were taken from the IPython Notebooks made for the lessons of the course (, as well as the assignment description , noted that the methods for coding was also inspired.*
4. *Wollack, Edward J.''WMAP-Age of the Universe''.NASA, 21 Dec. 2012,* *https://wmap.gsfc.nasa.gov/universe/uni_age.html*