Ay190 – Worksheet 5
Daniel DeFelippis
Date: January 28, 2014

1 Curve Fitting

In this worksheet, we seek to find a relationship between the mass of a supermassive black hole at the center of a galaxy and the the stellar velocity dispersion of that galaxy using data from Greene and Ho.

a

After downloading the data and the AstroPy package, we make a log-log scatterplot of the data - that is, $\log_{10} M_{BH}$ vs $\log_{10} \sigma_*$ - shown in figure 1. There is clearly a positive relationship between the two quantities, which we will quantify in parts b and c.

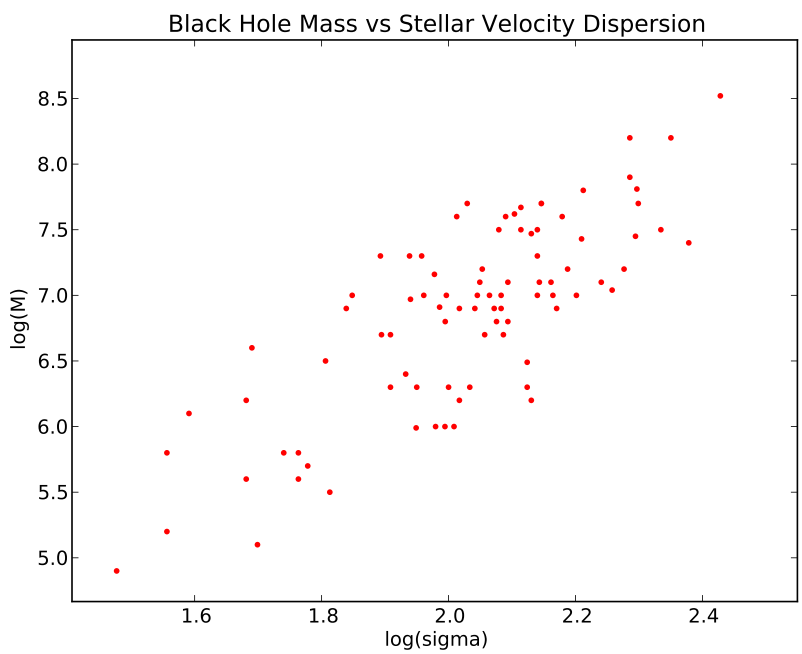

Figure 1: Log of black hole mass vs Log of stellar velocity dispersion

b

We use formulae given in the notes to calculate regression line fit parameters. If we let $x_i = (\log_{10} \sigma_*)_i$ and $y_i = (\log_{10} M_{BH})_i$, then:

$$S = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}, \ \Sigma x = \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}, \ \Sigma y = \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}, \ \Sigma x^2 = \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2}, \ \Sigma xy = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}.$$

Then, performing the fit, we get that $y = a_1 + a_2 x$ where

$$a_1 = \frac{\Sigma y \Sigma x - \Sigma x \Sigma xy}{S\Sigma x^2 - (\Sigma x)^2}, \quad a_2 = \frac{S\Sigma xy - \Sigma x \Sigma y}{S\Sigma x^2 - (\Sigma x)^2}$$

and the uncertainties on the a_i are given by

$$\sigma_{a_1} = \sqrt{\frac{\Sigma x^2}{S\Sigma x^2 - (\Sigma x)^2}}, \quad \sigma_{a_2} = \sqrt{\frac{S}{S\Sigma x^2 - (\Sigma x)^2}}.$$

Since we are ignoring the errors right now, we take $\sigma_i = 1$, so $S = 88$ (there are 88 data points). Calculating these parameters gives

$$a_1 = 0.931069678176 \pm 1.11588027837$$
$$a_2 = 2.92542900883 \pm 0.547497898246$$

The data and the regression line are shown together in figure 2.

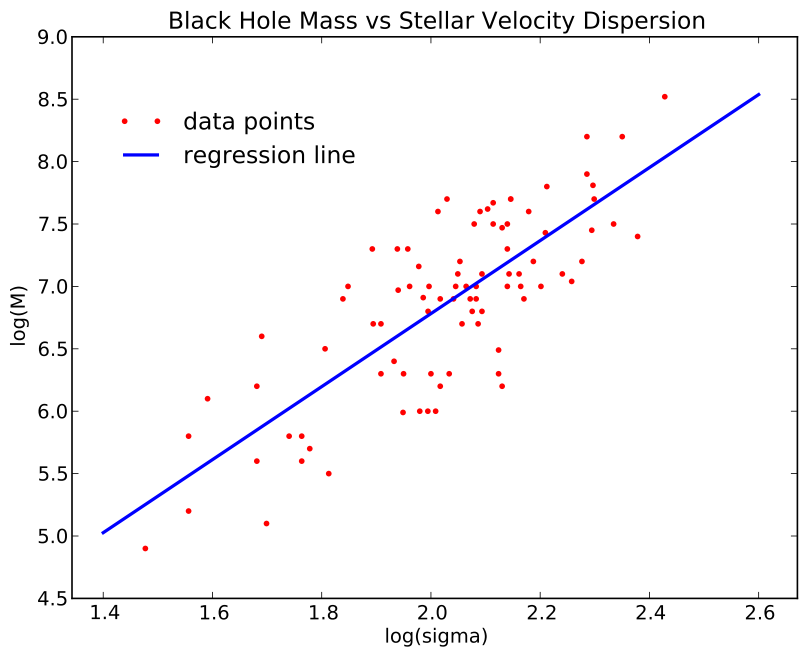

Figure 2: Log-Log data and regression line

We can compare our fit to the one from Greene and Ho. They fit the line

$$\log(M_{BH}/M_\odot) = \alpha + \beta \log(\sigma_* / \sigma_0)$$

where $\sigma_0 = 200$ km/s, and got $\alpha = 7.85 \pm 0.04$ and $\beta = 3.69 \pm 0.13$. Now, our a_1 corresponds to $\alpha - \beta \log \sigma_0$. The data points $log(M_{BH})_i$ are really $\log(M_{BH}/M_\odot)$, so no correction involving $M_\odot$ is needed. Also, our a_2 still corresponds to their β. Comparing the parameters, we see that:

$$a_1 = 0.93 \pm 1.12 \qquad \alpha - \beta \log \sigma_0 = -0.64 \pm 0.14$$
$$a_2 = 2.93 \pm 0.55 \qquad \beta = 3.69 \pm 0.13$$

which agree fairly well with each other.

c

Now, including the errors, we can perform another fit. For the $\log M$ data points, we choose to use the errors in the data file labeled "e_logM" and for the $\log \sigma$ points, we choose the errors labeled "e_sigma*". Now, we need to transform the errors on σ to get errors on $\log \sigma$. To do this, we note that

$$d(\log \sigma_*) = \frac{1}{\ln(10)} \frac{d\sigma_*}{\sigma_*},$$

so the errors on $\log \sigma_*$ are given by

$$\frac{1}{\ln(10)} \frac{e\_sigma^*}{\sigma_*}.$$

Also, we need to combine the x and y errors, and to do that we need a value of $\frac{dy}{dx}$. We choose the slope a_2 from the previous fit as this value for all data points, so that we can write

$$\sigma_{i,total}^2 = (e\_\log M)_i^2 + \left(\frac{dy}{dx}\right)^2 (e\_sigma^*)_i^2.$$

Proceding exactly as before otherwise, we calculate that

$$a_1 = 0.713709687178 \pm 0.226477686432$$
$$a_2 = 3.04393787988 \pm 0.107907788743.$$

The graph of this fit together with the data points and their errors is shown in figure 3.

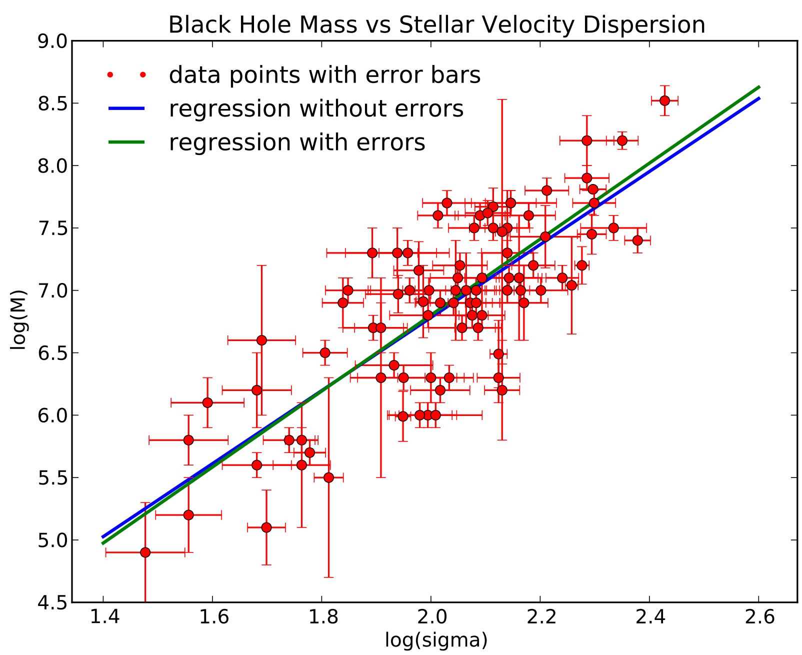

Figure 3: Log-Log data and the two regression lines

It is interesting to note that the uncertainties on both parameters went down, although now they are further from what Greene and Ho calculated, which suggests that there is some subtlety that I am ignoring when I chose the errors I did.