Problem set 2
The Exoplanet 51 Peg b

Paul van der Werf, Henk Hoekstra, Jarle Brinchmann

2015

1 Introduction

Humans have long looked into space and wondered whether there is life elsewhere but for most of that time we did not even know there were planets outside our solar system. In the last 20 years this has changed dramatically with the discovery of a planet orbiting a main sequence star (why this qualification?) in 1995.

The most direct way to observe these planets, direct imaging, is very challenging since many of these planets are very close to their parent stars and many orders of magnitude fainter than them. Instead the main progress in observing extra-solar planets (exo-planets in the following) was through indirect techniques.

For a number of years, the main technique was the radial velocity detection technique - where the reflex motion of a star in response to the gravitation attraction from an orbiting planet can be observed. This is also the way the first "normal" exo-planet was detected - 51 Pegasus b, normally known as 51 Peg b (51 Peg a is the star). Although we cannot see this planet directly, we can measure the motion of the star around the planet-star barycenter and hence the radial velocity $v_{\rm R}$.

From Newton's law of gravity we can then derive that the variation in radial velocity, $v_{\rm R}$, of the star caused by a planet in circular orbit with period P, has an amplitude k given by

$$k = \left(\frac{2\pi G}{P}\right)^{1/3} \frac{M_{\rm P} \sin i}{(M_{\rm P} + M_\star)^{2/3}} \tag{1}$$

Problem 1

Derive equation 1. *Hint: Make use of Kepler's third law.* Write equation 1 with P in units of days, $M_\star$ in units of solar masses ($\mathrm{M}_\odot$), and $M_{\rm P}$ in units of Jupiter masses ($M_{\rm J}$). (NB: $M_{\rm P} \ll M_\star$). What is the amplitude of the variation in $v_{\rm R}$ of the sun caused by the Earth? And by Jupiter?

2 Measuring the minimal mass of 51 Peg b

The first exoplanet around 51 Peg was discovered in 1995 through the radial velocity method. In this problem we will carry out this analysis and measure the period of the $v_{\rm R}$ variations and the amplitude of the variations in order to understand the properties of the system. The radial velocity data are provided on the course web site.

The file provides the original dataset from 1995 which was used to find the exo-planet. The final consists of three columns of data. The first column gives the Julian date of the observations relative to some reference date. The second column gives v_R measured in meters per second and the third column gives the estimated uncertainty on v_R.

In this problem set we will calculate the period and amplitude using Python. To do this we will adopt the following model for v_R:

$$v_\mathrm{R} = k \sin\left(\frac{2\pi(t-t_0)}{P}\right) + v_0 = k\sin\left(2\pi(f+f_0)\right) + v_0, \tag{2}$$

where we want to estimate k and P because they are required to estimate $M_\mathrm{P}\sin i$ (c.f. equation 1).

We will fit this model by minimising χ^2, where χ^2 is given by

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{v_{\mathrm{R},i} - \mathrm{model}_i}{\sigma_{v_{\mathrm{R},i}}}\right)^2. \tag{3}$$

In this equation N is the number of measurements, $v_{\mathrm{R},i}$ is the v_R measurement and $\sigma_{v_{\mathrm{R},i}} = \sigma_i$ gives the uncertainty estimate for $v_{\mathrm{R},i}$.

A χ^2 fit is quite simply a *weighted* least-squares fit. If the uncertainty on each data point were the same, the χ^2 expressions reduces to a standard least-squares fit. In general the lower that χ^2 is the better the fit. The most common approach, and the one we will take, is to take the model with the lowest χ^2 as providing the best estimate of the period, P. To get this estimate we also need to estimate the other parameters of the model, k, v_0 and f_0. Think carefully about how you can get a good initial estimate of v_0.

Problem 2

Read the file into Python and plot v_R against time. Give an approximate estimate of the period and amplitude by eye.

Use the χ^2 formulation to estimate P and k. Let the period vary over ± 1 day from the period you estimated by eye, use steps of 0.01 day. Make sure you translate these values into phase, f. Vary now for each P also k by $\pm$10m/s from the value estimated by eye in steps of 1 m/s. Let f_0 vary between 0 and 1 in steps 0.01 and vary v_0 using your own estimate of range and steps. For each P calculate the lowest χ^2 and put this in an array.

Plot χ^2 as a function of P and indicate the minimum value of χ^2. Which period gives the best fit? Improve this fit further by using smaller steps around the best-fit values. Use this to derive the best fitting values of P, k and f_0. Show this solution in a diagram of v_R against phase, f, with the best-fit indicated.

What is the minimal mass of the planet? *Hint: You need to find the mass of the main star yourself.*

3 Is the fit statistically acceptable?

In preceding problem we fit the radial velocity variation of the star caused by a planet. From this fit we derived an amplitude and a period and together these give us an estimate of the minimal mass of the planet. But how much should you trust this calculation?

This is a very common question in astronomy and a key aspect to realise is that we can never be completely sure that we have the right model. There are many possible reasons for why the model perhaps is not the correct one in this case:

- The planet might move on an elliptical orbit.
- Perhaps are there multiple planets in the planetary system.
- Perhaps there are systematic errors in the observations.
- ...

A similar situation can be had in many investigations and one of the ways to say more about this is to check whether the fit of the model to the data is a "good" fit. To be more precise, this means that we will check whether the deviations between the model and the data are within that expected from a statistical analysis.

3.1 Statistical tests

Imagine that you a set of measurements, $y_{\mathrm{obs},i}$, with uncertainty estimates σ_i, and that we want to test whether a model $y_{\mathrm{mod},i}$ is a good fit to the measurements. To do this there are a series of approaches in the literature — here will use a fairly simple approach for which some cautionary remarks are needed and we will touch on these later.

The approach we take is known as hypothesis testing. We can get a feeling for how this work by simply looking at possible outcomes. We will imagine that there are only two possibilities for our fitting:

1. model gives a good fit to the data: model is accepted = "positive"
2. model gives a bad fit to the data: model is rejected = "negative"

In other words: Either the model gives a good fit to the data, or it does not and we want to distinguish between these two options.

To check whether this is the case we now need to use a statistical test. There are no tests that always give the correct answers, so again we need to consider various possible outcomes. For either of the two answers the test can either give a true or false outcome so we have in total four possible outcomes:

1. The model is good + the test accepts the model = "true positive"
2. The model is good + the test rejects the model = "false negative"
3. The model is wrong + the test accepts the model = "false positive"
4. The model is wrong + the test rejects the model = "true negative"

So clearly in the cases 2 and 3 are problematic. In the statistic literature it is common to refer to case 2 as a Type II error and case 3 as a Type I error. In any statistical test there is balance between these two types of error and we most choose which one we want to minimise. In some situations this is clear — imagine devising a test for whether someone has a serious disease — and we typically try to minimise the case 2, the Type II errors. A typical value is that we only want a 5% chance to have a false negative.

This kind of significance test can be very helpful but also has a potential for being misused, and indeed it frequently is. The two most important points to keep in mind are:

- If the test finds that the model is a good fit, we say that *the model provides a good description of the data*. It does not say and should not be taken to mean that the model is *true*. That is a very different and much harder question to answer.
- When the test rejects the model, we usually say that the model is rejected with 95% significance, or with a 95% confidence limit. That 95% comes from 100%-5%, where the 5% is the chance that the model was incorrectly rejected.

3.2 The χ^2 test

Above we used the χ^2 as a way to find the best fitting model for the data. However χ^2 has another important function as well as we use it to test for the quality of a fit/model. If we go back to the equation for χ^2, we can write this as

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_{\mathrm{obs},i} - y_{\mathrm{mod},i}}{\sigma_i} \right)^2, \tag{4}$$

where $y_{\mathrm{obs},i}$ are the observables with associated uncertainty estimate σ_i and $y_{\mathrm{mod},i}$ are the model estimates. Now in this equation the observables are random variables due to the noise and if the uncertainty estimates, σ_i are known, then χ^2 is a random number that is distributed according to the χ^2 distribution. This analytic form for this distribution is a Gamma distribution

$$f_{\chi^2}(\nu) = \frac{1}{2^{n/2}\Gamma(n/2)} \nu^{(n/2)-1} e^{-\nu/2}, \tag{5}$$

where n is the number of degrees of freedom of the distribution, a number we'll get back to below. We do of course need to adjust for summing over more data when calculating χ^2. If you want to read more about the χ^2 test, then chapter 15 in *Numerical recipes* or section 7.4 in "Modern Statistical Methods for Astronomy" by Feigelson & Babu, are good places to go. The Wikipedia article on Pearson's chi square test is also ok.

We use the χ^2 test to assess whether the fit of the data to the model is acceptable or not. The idea is that if χ^2 is very large, then the fit of models to the data is not very good, while a very small value might also indicate a possible problem. What then is the approach? It turns out that there is a particular value for χ^2, χ^2_{lim}, above which the fit is sufficiently bad that we should reject our model. Thus the approach becomes:

- Hypothesis: The model provides a good fit to the data.
- If $\chi^2 < \chi^2_{\mathrm{lim}}$: Accept the hypothesis, or better: you cannot reject the hypothesis.
- If $\chi^2 > \chi^2_{\mathrm{lim}}$: Reject the hypothesis.

What then is χ^2_{lim}? In our example it is the value so that the chance of finding a χ^2 value *larger* than χ^2_{lim}, is less than 5%. Mathematically we can express this as:

$$P\left(\chi^2 > \chi^2_{\mathrm{lim}}\right) = 0.05. \tag{6}$$

To calculate this number, we need to know the number of degrees of freedom, n in equation 5. Since we will assume that our measurements are Gaussian random variables and that each measurement is independent, we have a total of N measurements. However we also need the

χ^2 value that we calculated, and that depends on the measurements, this therefore reduces the number of free parameters. You should therefore not use N as the number of degrees of freedom, but rather $N - m$ where m is the number of free parameters in your model (the number of fit parameters).

In Python you can find various routines related to χ^2 in the `scipy.stats` package. The likelihood distribution of χ^2 from equation 4 above, is provided by the `chi2.pdf` function and the cumulative distribution is given by `chi2.cdf`.

Problem 3

Adapt the χ^2-test for the model you have fitted to the observations. Give the limit $\chi^2_{\rm lim}$ for a chance of false negative of 5%, and indicate the value of n, the number of degrees of freedom, and how you decided on this value. What are your conclusions?

4 Uncertainties in the results

In the preceding we tested how good the fit is that we have obtained for the radial velocity data of 51 Peg. What is still missing is an estimate of the uncertainty on the estimated parameters and here we will find that this also can be found using χ^2.

To start we go back to the idea of the likelihood of a random variable x which is distributed like a Gaussian with mean μ and standard deviation σ. In this case the likelihood of the random variable, P, satisfies

$$P(\mu - \sigma < x < \mu + \sigma) = 0.68, \tag{7}$$

which you sometimes see written as

$$P(|x - \mu| > \sigma) = 0.32. \tag{8}$$

We often say that $\pm\sigma$ gives the 68% confidence interval on x, where 68% or 0.68, provides the confidence level. Note that people have a tendency to confuse these terms — try to avoid that!.

In the case of the χ^2 distribution we use a similar approach to get confidence interval for the parameters of the fit. It turns out that the distribution $\chi^2 - \chi^2_{\rm min}$ follows a χ^2 distribution, but we do have account for the various degrees of freedom. What are these degrees of freedom?

When you calculate $\chi^2_{\rm min}$, you have N data points and M parameters in general. Let us now say that you want to calculate χ at a different point in parameter space. But now you want to keep ν parameters fixed — for instance because you want to make a plot of χ^2 as a function of these parameters. In that case N is fixed and you vary a total of $M - \nu$ parameters. In this case it turns out that $\Delta\chi^2 =$ is distributed as a χ^2 distribution with ν degrees of freedom. If you want more details on this particular issue, Chapter 15.6 in Numerical Recipes provides a traditional discussion.

Problem 4

1. Calculate $\Delta\chi^2$ defined as

$$P\left(\chi^2 - \chi^2_{\rm min} > \Delta\chi^2\right) = 0.32 \tag{9}$$

for a number of free parameters of 1, 2 and 4.

2. Calculate the first the minimum value of χ^2 and call this χ^2_{min}. Use thereafter Python to produce a plot with k and P (see problem 1) on the axes with a contour level of constant χ^2 overplotted (using for instance `pyplot.contour`) where

$$\chi^2 = \chi^2_{\text{min}} + \Delta\chi^2 \tag{10}$$

Note that in this expression χ^2_{min} is the minimum χ^2 for the particular values of k and P. Since we fix these two parameters, we have two free parameters in the fit. The 68% confidence limit for the amplitude of the radial velocity and the orbital period is now given by the region enclosed by this contour.

In your answer, please also give the values of $\Delta\chi^2$ and χ^2_{min}.

Problem 5

To close off, please calculate the uncertainty on the estimate of the minimal mass of the planet 51 Peg b.