MAS212 Assignment #3 (2017): Extrasolar planets

Dr. Sam Dolan (s.dolan@sheffield.ac.uk)

`http://sam-dolan.staff.shef.ac.uk/2016/mas212/docs/assign3_report.pdf`

An astronomer has been observing a distant star for the last 80 days, and has found evidence for a periodic Doppler shift in its spectrum of light (see Fig. 1). She has found the relative velocity of the star from its Doppler shift, and has inferred that the star is moving, regularly, back and forth. The astronomer is convinced that this implies that an extrasolar planet must be orbiting around the star. Now she needs your help to test the hypothesis.

In this assignment you will:

- Fit a crude model to a data set;
- Build a better model by solving some physically-motivated ODEs;
- Estimate the orbital parameters of the system;
- Prepare a **Beamer presentation** on your findings.

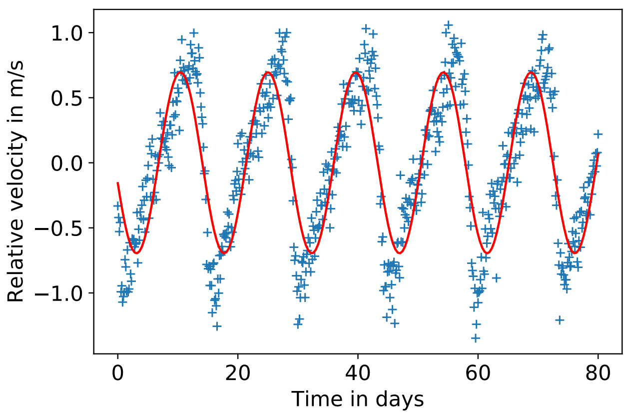

Figure 1: *Blue points:* Measurements of the relative velocity of the star along our line of sight, as a function of time. *Red line:* the astronomer's crude model. Data file available at: `http://sam-dolan.staff.shef.ac.uk/mas212/data/extrasolar_signal.txt`

The Submission: A completed assignment will comprise a `.pdf` of your Beamer presentation and a single code file (`.py` or `.ipynb`). Your presentation will comprise **seven slides** as described below. The report should be accompanied by *one* Python script or Jupyter notebook. Please use LaTeX to write your report. Note that the presentation will carry $\sim 60\%$ credit and the code $\sim 40\%$.

The Deadline: The deadline for submission is found on the course website. Files should be submitted at `http://somas-uploads.shef.ac.uk/mas212`.

The slides

- **Slide 1:** Introduction to extra-solar planets. (Include your registration number somewhere on this slide). **(a)**.
- **Slide 2:** Fitting a linear model. Include the model Eq. (1), a plot similar to Fig. 1, and your best-fit parameters. **(b)**.
- **Slide 3:** Testing the goodness of fit. Include some or all of: (i) a plot of the residuals after fitting; (ii) the RMSD value; (iii) a plot of the autocorrelation for $k = 1$; (iv) the autocorrelation coefficient α_1. **(c)**.
- **Slide 4:** A physical model. On this slide, define the 3D autonomous set of ODEs and initial conditions, and *briefly* describe the midpoint method. **(d)**.
- **Slide 5:** Orbits and signals. Show the three example orbits (left) and the three example signals (right), with the parameter choices. **(e)**.
- **Slide 6:** Fitting the physical model. Present the best-fit model and parameters p, e and ϕ_0; (or describe your attempt to fit, if it didn't work). **(f)**.
- **Slide 7:** Either (i) Conclusions; or (ii) Parameters. (i) Summarise your findings in three bullet points. Add one reference to further reading. (ii) Estimate the mass of the extrasolar planet, and its distance from the star, using Kepler's laws and the assumption that the star has the same mass as the Sun. **(g)**.

The steps

(a) Introduction to extra-solar planets. Read around the subject of extra-solar planets, starting with last year's Assignment 3 report (link above). Read about recent discoveries (Ross 128b, Proxima B, the TRAPPIST planets) of nearby planets in the 'habitable zone'. Read about the methods used for detecting extra solar planets, in particular, the *radial velocity method.* Prepare **one slide** to introduce your talk. (No code is necessary).

(b) Fitting a linear model. The astronomer asks you to fit the data here (`http://sam-dolan.staff.shef.ac.uk/mas212/data/extrasolar_signal.txt`) with a linear model

$$f_0(t; \beta_0, \beta_1) = \beta_0 \sin(\omega t) + \beta_1 \cos(\omega t), \tag{1}$$

where β_0 and β_1 are parameters, and $\omega = 2\pi/14.6$.

Let $\mathbf{X}$ be the $N \times 2$ matrix with rows $[\sin(\omega t_i), \cos(\omega t_i)]$. Solve the *normal equations*

$$(\mathbf{X}^T \mathbf{X})\boldsymbol{\beta} = \mathbf{X}^T \mathbf{y}$$

to find the best-fit parameters (see Lec 8 and Lab Class 8).

Plot the data set and the best-fit model. It should look a bit like Fig. 1.

(c) Testing the goodness of fit. The **residuals** r_i, which describe the difference between the data (t_i, v_i) and the best-fit model, are defined by

$$r_i = v_i - f_0(t_i; \beta_0, \beta_1).$$

One measure of the 'goodness of fit of the model to the data is the *root mean square deviation*

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} r_i^2}. \tag{2}$$

If the model fits the data well, then the residuals should be uncorrelated in time. Is this the case? One way to check this is to plot r_i against r_{i+k} for $k \in \mathbb{N}$. Try this for $k = 1$. If there correlation between successive residuals, then there will be a trend visible in the plot; otherwise it will be randomly distributed. This can be checked using the *autocorrelation coefficient* α_k given by

$$\alpha_k \equiv \frac{\Sigma_{j=1}^{N-k} (r_j - \overline{r})(r_{j+k} - \overline{r})}{\Sigma_{j=1}^{N} (r_j - \overline{r})^2} \tag{3}$$

where $\bar{r}$ is the mean residual.

(d) A physical model. A theorist argues that the gravitational attraction of an extrasolar planet is causing the star to move on an elliptical orbit. The astronomer is measuring (via the Doppler shift) the relative velocity $v_{\parallel} = \mathbf{v} \cdot \hat{\mathbf{n}}$, where $\mathbf{v}$ is the true velocity of the star, and $\hat{\mathbf{n}}$ is the unit vector along our line of sight. We happen to be viewing this system 'edge-on', so that the line-of-sight vector lies in the plane of motion. The theorist proposes a system of ODEs for you to investigate,

$$\ddot{r} = -\frac{1}{r^2} + \frac{p}{r^3}, \qquad \dot{\phi} = \frac{\sqrt{p}}{r^2},$$
$$r(0) = \frac{p}{1+e}, \qquad \dot{r}(0) = 0, \qquad \phi(0) = \phi_0,$$

where $r = r(t)$ and $\phi = \phi(t)$. This model has three dimensionless parameters: p, e, which should determine the size and eccentricity of the ellipse; and ϕ_0, which should determine the orientation of the ellipse.

Write the ODEs as 3 autonomous first-order equations by introducing a new variable $s = \dot{r}$, with initial condition $s(0) = 0$.

Write code to solve the ODE system numerically with the *midpoint method*, or a higher-order method.

(e) Orbits and signals. Use your code to find three numerical solutions, for parameters $(p = 1, e = 0, \phi_0 = 0)$, $(1.7, 0.5, \pi/6)$ and $(1.3, 0.8, -\pi/2)$.

Plot the orbital trajectories in the xy-plane, by plotting $r \cos \phi$ against $r \sin \phi$.

For the same parameter choices, plot the relative velocity $v_\parallel$ as a function of time over the domain $t \in [0, 80]$, where

$$v_\parallel = \dot{r} \cos \phi - r\dot{\phi} \sin \phi.$$

The plots may look similar-but-different to those in Fig. 2.

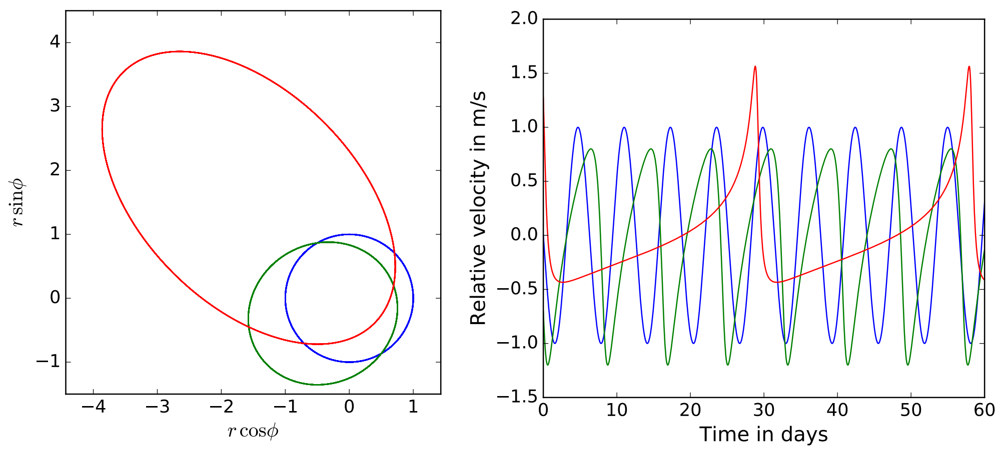

Figure 2: *Left:* Three example orbits. *Right:* The relative velocity as a function of time, for the same three orbits. *NB:* These results are for different parameter choices to in part (d).

(f) Fitting the physical model. You showed in (e) that, for a given choice of parameters, your code can find the relative velocity $v_\parallel$ as a function of time. Thus, you have a three-parameter model, which is non-linear in its parameters. The challenge now is to fit this model to the astronomer's data set to find the best-fit parameters for the physical system: p, e and ϕ_0.

To do this, I recommend that you write a function `calcVres()` that returns an array of residuals (i.e., the difference between the data v_i and the model $v_\parallel$ evaluated at $t = t_i$ for a given choice of parameters p, e and ϕ_0, for each value of t_i); and then use `scipy.optimize.leastsq()` to find the best-fit parameters p, e and ϕ_0. Check that this gives a good fit to the data. Present your key findings in your presentation.

(g) Physical parameters. Assuming that the star has the same mass as the Sun, use Kepler's laws to infer the mass of the extrasolar planet, and its distance from the star (find the semi-major axis). (*Hint:* Read last year's report for the method).

Guidance:

This assignment will count for $\sim 30\%$ of your module mark, and thus it should be a substantial piece of work. For this assignment, **fair means** include: asking the lecturer for advice; reading online materials; using and adapting short code snippets from lectures; etc. Please avoid verbatim copying, and please cite all sources used. **Unfair means** include (but are not limited to): sharing or distributing files; copying-and-pasting from work that is not your own; posting your work online; passing off other's work as your own; asking for help, or code, from last year's cohort; etc.

Note that all submissions will be checked for excessive similarities. Where there is circumstantial evidence of unfair means, I reserve the right to award zero marks for this assignment.