MAS212 Assignment #3: Detecting extrasolar planets

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

An astronomer has been observing a distant star for the last 60 days, and has found evidence for a periodic Doppler shift in its spectrum of light (see Fig. 1). He 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 in the vicinity of the star. Now he needs your help to test the hypothesis.

In this assignment you will:

- Fit a model to a data set;
- Attempt to interpret the data by solving some physically-motivated ODEs;
- Write a report on your findings.

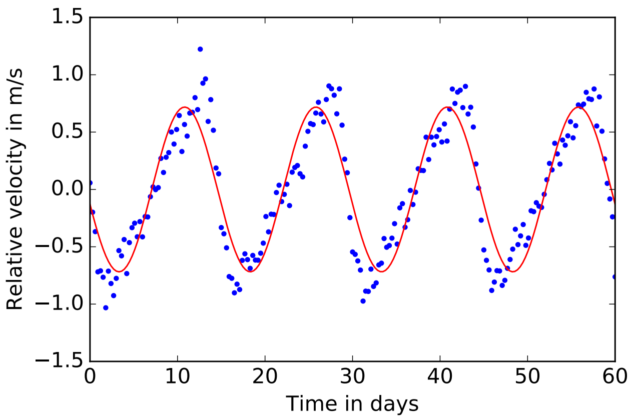

Figure 1: *Blue dots:* Measurements of the relative velocity of the star along our line of sight, as a function of time. *Red line:* the astronomer's crude model (you will improve this). Data file available at: `http://sam-dolan.staff.shef.ac.uk/mas212/data/extrasolar_signal.txt`

The Submission: A completed assignment will comprise a `.pdf` of your report and a single code file (`.py` or `.ipynb`). Your report should comprise three sections corresponding to the four parts below. It should also include a conclusion paragraph, and a list of references. As well as addressing the parts of the brief, your report should also be a coherent document in its own right, which could be read by someone who has not seen this brief. The report should be no more that **eight sides** including figures. If necessary, additional material can be included in an appendix. The report should be accompanied by *one* Python script or Jupyter notebook. Please use LaTeX to write your report. Note that the report 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`.

Part 1: Introduction

You should write this section after attempting Part 2 and Part 3, but it should appear as the first part of your report. It should be approximately one side in length. It is up to you how to introduce your report. For example, you could do some research on the search for extrasolar planets, recent discoveries (e.g. Proxima B), and the Doppler shift method. Alternatively, you could describe the mathematical theory of fitting (linear and non-linear) models to data, as covered in Lec. 7, and thus motivate this investigation as a specific example of a modelling exercise. Or you may have another idea on how to set your work in context. Creativity is welcomed.

Please include least three academic references. Some credit will be given for evidence of outside reading, beyond the course materials.

Part 2: Model Fitting

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

$$f_0(t; \beta_0, \beta_1) = \beta_0 \sin(\omega t) + \beta_1 \cos(\omega t),$$

where β_0 and β_1 are parameters, and $\omega = 2\pi/15$.

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 7 and Lab Class 7).

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

(b) A good fit? Plot the **residuals** r_i defined by

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

If the model fits the data well, then the residuals should be uncorrelated. Is this the case?

(c) A better linear model. Now experiment with alternative linear models, with no more than six parameters. For example, you might start by considering

$$f_1(t; \beta_j) = \beta_0 \sin(\omega t) + \beta_1 \cos(\omega t) + \beta_2 \sin(2\omega t) + \beta_3 \cos(2\omega t)$$

Present some evidence to show that your chosen model is a better fit to the data than the crude model in part 2(a).

(d) A non-linear model. Use `scipy.optimize.curve_fit()` to fit an appropriate non-linear model with no more than six parameters; for example,

$$f_2(t; \beta_j) = \beta_0 \sin(\beta_1 t + \beta_2).$$

Plot your preferred model, and compare and contrast with part (c).

(e) *Extension.* If parts (a)–(d) were straightforward for you then *either*,

1. Instead of using `curve_fit()`, write your own function to fit a non-linear model to data by implementing (e.g.) the Gauss-Newton method; or,
2. Take the **discrete Fourier transform** (DFT) of the data. Apply a frequency filter to remove the majority of the noise. Now take the inverse DFT, and plot the result. Compare with the original data set. `http://sam-dolan.staff.shef.ac.uk/2015/mas212/docs/l10.pdf`

Part 3: A physical model

A theorist argues that the gravitational attraction of an extrasolar planet is causing the star to move on an elliptical orbit. She says that 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. Furthermore, she asserts that we are viewing the 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,

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

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. **Note:** You can convert the $\ddot{r}$ equation into a pair of first-order equations by first introducing a new variable $s = \dot{r}$, with initial condition $s(0) = 0$.

(a) Numerical solution of ODEs with midpoint method. Write code to solve the ODE system numerically using the midpoint method, or another second-order-accurate method (or higher; but *not* the Euler method). Add one paragraph in the report on the key details of your implementation.

(b) Orbits and signals. Use your code to find three numerical solutions, for parameters $(p = 1, e = 0, \phi_0 = 0)$, $(1, 0.4, \pi/6)$ and $(1, 0.8, -\pi/4)$.

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

Plot the relative velocity $v_\parallel$ as a function of time over the domain $t \in [0, 60]$. The relative velocity is given by

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

Comment on similarities and differences with the astronomer's signal.

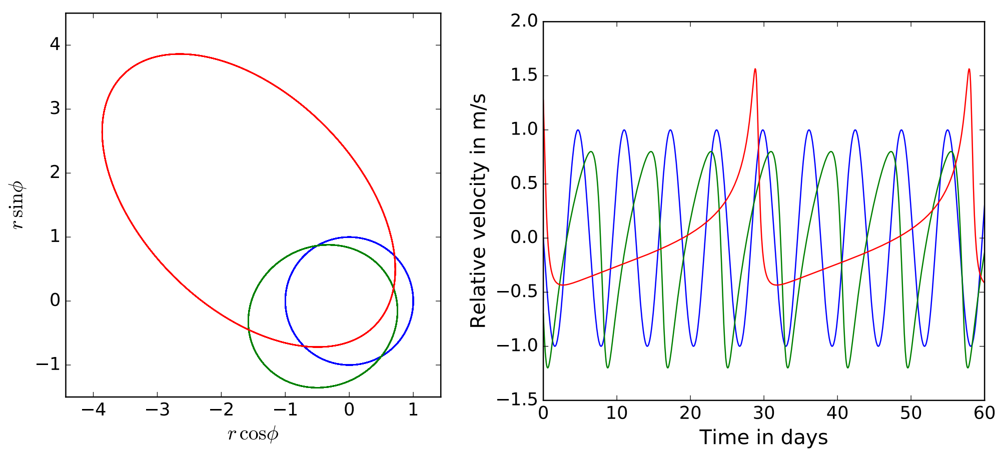

Figure 2: *Left:* Three example orbits. *Right:* The relative velocity as a function of time, for these three orbits.

(c) Fitting the physical model. You showed in 2(b) that, for a given choice of parameters, your code can find the relative velocity 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. Write code to fit the model, and thus find the best-fit parameters. Present your findings in your report.

Extension: 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).

Conclusion and reference list

Add a short conclusion to your report. This should be four sentences or so, summarising the most important findings. Include a list of references.

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, even in the introduction, 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, 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.