Statistical Inference: The Eclipse of 1919

Alan Heavens, ICIC, Imperial College London

1 Testing gravity with gravitational lensing

The aim of this exercise is to write an MCMC code to analyse photographic plates from Eddington's 1919 eclipse expedition, to determine whether the data favour Newtonian gravity or Einstein's General Theory of Relativity. Light from stars that passes close to the Sun is deflected, and during an eclipse, these stars can be detected and their displacements measured, when compared with photographs taken when the Sun is far away.

General Relativity predicts that light passing a mass M at distance r will be bent through an angle

$$\theta_{\mathrm{GR}}(r) = \frac{4GM}{rc^2}, \tag{1}$$

whereas an argument based on Newtonian gravity gives half this:

$$\theta_{\mathrm{N}}(r) = \frac{2GM}{rc^2}. \tag{2}$$

We can either treat this as a *parameter inference* problem, modelling the bending as

$$\theta_{\mathrm{N}}(r) = \frac{\alpha GM}{rc^2}, \tag{3}$$

and inferring α, or as a *model comparison* problem (see later in the course). For this exercise we will do the former. For GR, the bending is 1.75 arcsec for light passing close to the limb of the Sun.

We might conclude that equation (3) is the *model*, with a single *parameter* α. We assume that GM, c and r are known perfectly, from measurements of planet orbits etc. However, the experiment has some other parameters as well, which we investigate below.

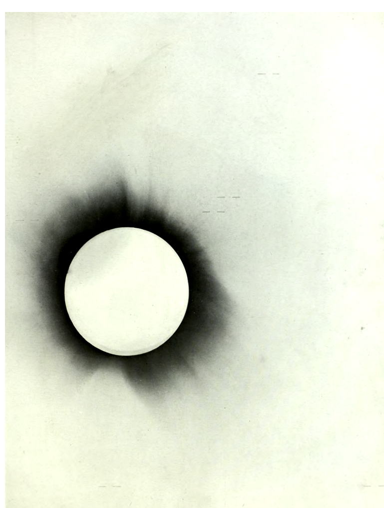

Figure 1: One of the photographs from the expedition.

19. The values of Dx and Dy were equated to expressions of the form

$$ax + by + c + \alpha E_x \, (= Dx)$$

and

$$dx + ey + f + \alpha E_y \, (= Dy),$$

where x, y are the co-ordinates of the stars given in Table I., and E_x, E_y are coefficients of the gravitational displacement.

The quantities c and f are corrections to zero, depending on the setting of the scale plate on the plate measured, a and e are differences of scale value, while b and d depend mainly on the orientation of the two plates. The quantity α denotes the deflection at unit distance (*i.e.*, 50′ from the sun's centre), so that αE_x and αE_y are the deflection in R.A. and Decl. respectively of a star whose co-ordinates are x and y.

The left-hand sides of the equation for the seven stars shown are :—

No.	Right Ascension.	Declination.
11	$c-0{\cdot}160b-1{\cdot}261a-0{\cdot}587\alpha$	$f-1{\cdot}261d-0{\cdot}160e+0{\cdot}036\alpha$
5	$c-1{\cdot}107b-0{\cdot}160a-0{\cdot}557\alpha$	$f-0{\cdot}160d-1{\cdot}107e-0{\cdot}789\alpha$
4	$c+0{\cdot}472b+0{\cdot}334a-0{\cdot}186\alpha$	$f+0{\cdot}334d+0{\cdot}472e+1{\cdot}336\alpha$
3	$c+0{\cdot}360b+0{\cdot}348a-0{\cdot}222\alpha$	$f+0{\cdot}348d+0{\cdot}360e+1{\cdot}574\alpha$
6	$c+1{\cdot}099b+0{\cdot}587a+0{\cdot}080\alpha$	$f+0{\cdot}587d+1{\cdot}099e+0{\cdot}726\alpha$
10	$c+1{\cdot}321b+0{\cdot}860a+0{\cdot}158\alpha$	$f+0{\cdot}860d+1{\cdot}321e+0{\cdot}589\alpha$
2	$c-0{\cdot}328b+1{\cdot}079a+1{\cdot}540\alpha$	$f+1{\cdot}079d-0{\cdot}328e-0{\cdot}156\alpha$

Figure 2: The data model.

2 Underlying theory

3 Data

The Dyson, Eddington and Davidson 1920 paper (Phil. Trans. R. Society, 220, 571) is open access at `http://rsta.royalsocietypublishing.org/content/roypta/220/571-581/291.full.pdf` and contains the data we need. The figures in this document are taken from there.

4 Data model and model parameters

A displacement of the star images will be caused not just by bending of light, but by changes in the scale of the photographic plate (caused e.g. by changes in temperature), rotation of the plate with respect to the comparison plate, and offsets. We are not very interested in these effects, which introduce *nuisance parameters*, but we need to include them. The data model is then given by the equations in Fig 2, i.e.

$$\begin{aligned} Dx &= ax + by + c + \alpha E_x \\ Dy &= dx + ey + f + \alpha E_y. \end{aligned} \tag{4}$$

Thus there are 4 parameters for the displacement in x, and 4 for the y displacement. Only the light bending parameter, α is common. The data use units measured on the plate, and you will do the same. Dyson et al. do not give errors in the displacements. Assume they are 0.05 in the units given.

The positions of the stars are shown in Fig. 3, or you can read them in the table in Fig.2, which also gives the expected bend angle due to General Relativity for each star, as the coefficient of α. This is in a strange unit, so your number α will need to be rescaled to translate it to arcseconds at the limb. **To translate the inferred value of α from the units they used to the bending angle for light grazing the Sun, in arcsec, you need to multiply it by 19.8.**

The displacements are measured and shown in Fig. 4, for 7 plates (I,II,...VIII; VI was not used.)

TABLE I.

No.	Names.	Photog. Mag.	Co-ordinates. Unit = 50′.		Gravitational displacement.			
					Sobral.		Principe.	
			x.	y.	x.	y.	x.	y.
		m.			″	″	″	″
1	B.D., 21°, 641	7·0	+0·026	−0·200	−1·31	+0·20	−1·04	+0·09
2	Piazzi, IV, 82	5·8	+1·079	−0·328	+0·85	−0·09	+1·02	−0·16
3	κ^2 Tauri	5·5	+0·348	+0·360	−0·12	+0·87	−0·28	+0·81
4	κ^1 Tauri	4·5	+0·334	+0·472	−0·10	+0·73	−0·21	+0·70
5	Piazzi, IV, 61	6·0	−0·160	−1·107	−0·31	−0·43	−0·31	−0·38
6	υ Tauri	4·5	+0·587	+1·099	+0·04	+0·40	+0·01	+0·41
7	B.D., 20°, 741	7·0	−0·707	−0·864	−0·38	−0·20	−0·35	−0·17
8	B.D., 20°, 740	7·0	−0·727	−1·040	−0·33	−0·22	−0·29	−0·20
9	Piazzi, IV, 53	7·0	−0·483	−1·303	−0·26	−0·30	−0·26	−0·27
10	72 Tauri	5·5	+0·860	+1·321	+0·09	+0·32	+0·07	+0·34
11	66 Tauri	5·5	−1·261	−0·160	−0·32	+0·02	−0·30	+0·01
12	53 Tauri	5·5	−1·311	−0·918	−0·28	−0·10	−0·26	−0·09
13	B.D., 22°, 688	8·0	+0·089	+1·007	−0·17	+0·40	−0·14	+0·39

Figure 3: The x, y positions of the stars measured. Note that only seven stars are used, numbers $11, 5, 4, 3, 6, 10, 2$.

TABLE II.—Eclipse Plates—Scale.

No. of Star.	I.		II.		III.		IV.		V.		VII.		VIII.	
	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.
	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	−1·411	−0·554	−1·416	−1·324	+0·592	+0·956	+0·563	+1·238	+0·406	+0·970	−1·456	+0·964	−1·285	−1·195
5	−1·048	−0·338	−1·221	−1·312	+0·756	+0·843	+0·683	+1·226	+0·468	+0·861	−1·267	+0·777	−1·152	−1·332
4	−1·216	+0·114	−1·054	−0·944	+0·979	+1·172	+0·849	+1·524	+0·721	+1·167	−1·028	+1·142	−0·927	−0·930
3	−1·237	+0·150	−1·079	−0·862	+0·958	+1·244	+0·861	+1·587	+0·733	+1·234	−1·010	+1·185	−0·897	−0·894
6	−1·342	+0·124	−1·012	−0·932	+1·052	+1·197	+0·894	+1·564	+0·798	+1·130	−0·888	+1·125	−0·838	−0·937
10	−1·289	+0·205	−0·999	−0·948	+1·157	+1·211	+0·934	+1·522	+0·864	+1·119	−0·820	+1·072	−0·768	−0·964
2	−0·789	+0·109	−0·733	−1·019	+1·256	+0·924	+1·177	+1·373	+0·995	+0·935	−0·768	+0·892	−0·585	−1·166
	−1·500*	−0·554	−1·500	−1·324	+0·500	+0·843	+0·500	+1·226	+0·400	+0·861	−1·500	+0·777	−1·300	−1·322

Figure 4: The displacements Dx and Dy, as measured on 7 plates (I,II,... VII). Note that you have to subtract the number at the bottom of each column - e.g. add 1.5 to the numbers in the first column.

5 Exercise

Write an MCMC code to sample from the posterior joint probability of a, b, c and α, using the Dx displacement data for Plate II. The units are weird, so multiply α by 19.8 to translate it to the bending inferred at the limb of the Sun (in arcsec), and compare with the Newtonian result (0.9) and the GR result (1.75). Note that Dyson et al. make a correction of about 10% because of errors arising from measuring indirectly the displacements, but we will ignore that here.

- Assume uniform priors on the parameters (so you compute only the likelihood). What do you assume for the form of the likelihood?
- You might like to start with a very simple 'top-hat' proposal distribution, where the new point is selected from a 4D 'rectangular' region centred on the old point. For this you will need a simple random number generator. Alternatively use a gaussian for each parameter.
- Explore visually the chain when you have (a) a very small proposal distribution, and (b) a very large proposal distribution, for a maximum of 1000 trials. What do you conclude?
- Show how the acceptance probability changes as you change the size of the proposal distribution from very small (say 0.0001) to very large (say 10).
- Once you have settled on a 'reasonable' proposal distribution, compute the average value of the parameters under the posterior distribution, and their variances and covariance.
- Marginalise over the nuisance parameters and plot the posterior distribution of α.
- Now see what happens if you don't consider the nuisance parameters (i.e. set them all to zero, or include a gaussian prior on each one, that keeps them close to zero). What would you conclude about the light bending at the limb of the Sun, and what would be your conclusion about gravity physics?

6 Extensions

- Include some more data, e.g. Dy for Plate II, or more plates.
- Write and apply a Gelman-Rubin convergence test, and deduce roughly how long the chains should be for convergence.
- Extend to perform Hamiltonian Monte Carlo. You might like to try to compare the performance of MCMC and HMC; you will need to decide what the right criterion is.
- Actually, we have two separate theories, with two different discrete values of α. How would you do model comparison in this case? If you have time, you might like to do it, and compute the relative probabilities of GR and Newtonian gravity, given these data.

For HMC, the algorithm is (from Hajian 2006):

Hamiltonian Monte Carlo

initialize $\mathbf{x}_{(0)}$
for i = 1 to $N_{samples}$
    $\mathbf{u} \sim \mathcal{N}(0,1)$ (Normal distribution)
    $(\mathbf{x}^*_{(0)}, \mathbf{u}^*_{(0)}) = (\mathbf{x}_{(i-1)}, \mathbf{u})$
    for j = 1 to N
        make a leapfrog move: $(\mathbf{x}^*_{(j-1)}, \mathbf{u}^*_{(j-1)}) \rightarrow (\mathbf{x}^*_{(j)}, \mathbf{u}^*_{(j)})$
    end for
    $(\mathbf{x}^*, \mathbf{u}^*) = (\mathbf{x}_{(N)}, \mathbf{u}_{(N)})$
    draw $\alpha \sim$ Uniform(0,1)
    if $\alpha < \min\{1, e^{-(H(\mathbf{x}^*,\mathbf{u}^*)-H(\mathbf{x},\mathbf{u}))}\}$
        $\mathbf{x}_{(i)} = \mathbf{x}^*$
    else
        $\mathbf{x}_{(i)} = \mathbf{x}_{(i-1)}$
end for

$H = -\ln L + K$, where $K = \mathbf{u} \cdot \mathbf{u}/2$. If the derivatives are hard, you might try Sympy (https://www.sympy.org/en/index.html) to differentiate U automatically and produce (quite a few lines of!) python code, and there are other possibilities (Jax, autograd). Speak to demonstrators for suggestions.

Aside: an alternative approach is to approximate U by a bivariate gaussian with covariances estimated from the MCMC code:

$$U = \frac{1}{2}(\theta - \theta_0)_\alpha C^{-1}_{\alpha\beta}(\theta - \theta_0)_\beta.$$

Since it's approximate, H will not be conserved, but the Metropolis step sorts everything out. The derivatives are then easy and analytic (but approximate).

You should use the leapfrog algorithm (which is forward-backward symmetric, as required for detailed balance)

$$\begin{aligned} u_i\left(t + \frac{\epsilon}{2}\right) &= u_i(t) - \frac{\epsilon}{2}\left(\frac{\partial U}{\partial x_i}\right)_{\mathbf{x}(t)} \\ x_i(t+\epsilon) &= x_i(t) + \epsilon u_i\left(t + \frac{\epsilon}{2}\right) \\ u_i(t+\epsilon) &= u_i\left(t + \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\left(\frac{\partial U}{\partial x_i}\right)_{\mathbf{x}(t+\epsilon)}. \end{aligned} \tag{5}$$

Issues to consider are how many integration steps per point in the chain, and how big those steps should be. For some discussion, see Hajian (2006), astroph/0608679.

Alan Heavens
March 2021