



Pontificia Universidad Católica de Chile  
Instituto de Astrofísica — Facultad de Física  
Estadística para Astrónomos - ASP5408  
Primer Semestre 2016

## Interrogación 3

Profesor: Andrés Jordán (ajordan@astro.puc.cl)  
Ayudante: Néstor Espinoza (nespino@astro.puc.cl)

**Tiempo:** 2 días (Entrega: Viernes 17 de Junio, 17:00 hrs.)

---

Los sets de datos adjuntos a esta interrogación contienen observaciones del flujo de 10 estrellas en función del tiempo (usualmente llamada una “curva de luz”), tal y como se muestra en la Figura 1, en donde se grafica el flujo relativo de cada estrella; el ruido de medición en cada punto es de 100 ppm. El flujo de todas estas estrellas fue tomado al mismo tiempo y con el mismo detector y se espera que todas excepto la estrella objetivo, para la cual esperamos observar un tránsito (denominada “Flux Target” en el set de datos adjunto) tuvieran un flujo constante. Aún así, como se puede ver, éstas estrellas (denominadas “estrellas de comparación”) mapean variaciones de flujo importantes. Estas no son variaciones físicas (i.e., de la estrella misma) sino que instrumentales similares debido a efectos sistemáticos (e.g., atmósfera, nubes, variaciones en el detector, etc.) que, creemos, afectan a todas las estrellas de manera similar. Su tarea será extraer y modelar un tránsito de la curva de luz objetivo; en particular, estamos **muy** interesados en extraer la mejor estimación posible del radio del planeta.

- a) El flujo, en general, es afectado por factores multiplicativos en vez de aditivos. Aún así, este problema es fácil de resolver trabajando con el logaritmo de los flujos. De esa manera, un modelo simple para el flujo de la estrella  $i$ -ésima (sin contar la estrella que tiene el tránsito) viene dado por:

$$\log F_i(t) = \sum_{j=1}^p a_{i,j} S_j(t),$$

donde  $S_j(t)$  es el  $j$ -ésimo factor que perturba el flujo de nuestra estrella  $i$ -ésima, y  $a_{i,j}$  es la amplitud de esa señal en dicha estrella. Para resolver el problema, asumiremos que las señales  $S_j(t)$  son no-correlacionadas entre sí. Aplique PCA para obtener las señales  $S_j(t)$  usando las estrellas de comparación, gráfíquelas y responda: basados en los autovalores asociados a cada señal, ¿cuáles son las señales más importantes?

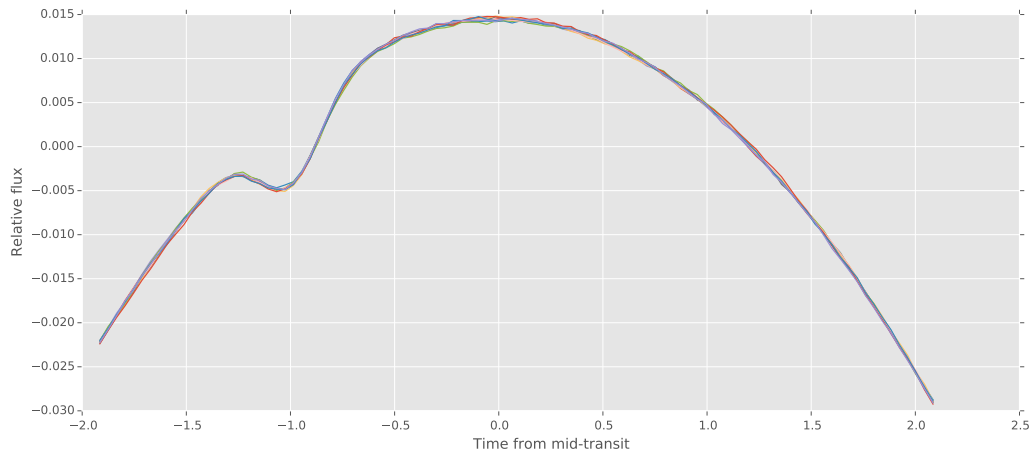


Figure 1 — Curvas de luz de las 10 estrellas que usted modelará en esta interrogación.

- b) Con las señales modeladas en la parte anterior, ahora podemos modelar el flujo de nuestra estrella objetivo como:

$$\log F(t) = c + \log T(t) + \sum_{j=1}^p \alpha_j S_j(t) + \epsilon(t),$$

donde  $c$  es una constante (que modela el “flujo cero” de la estrella),  $T(t)$  es el modelo de tránsito, las señales  $S_j(t)$  fueron las obtenidas en el problema anterior y, para este ítem, asumiremos  $\epsilon(t) \sim N(0, \sigma^2)$ , con  $\sigma^2$  desconocido<sup>1</sup>. El tránsito del planeta se espera dure dos horas exactas y, como se ve, el tiempo entregado en el set de datos ya considera que el centro del tránsito ocurre en  $t = 0$ . Para modelar el tránsito, use la herramienta **batman**<sup>2</sup>, donde los parámetros del mismo son

- **t0**: Tiempo del centro de tránsito. Fijelo en 0.
- **per**: Periodo orbital del planeta. Fijelo a 0.78884 días.
- **rp**: Razón entre el radio del planeta y de la estrella,  $R_p/R_*$  (nuestro parámetro de interés!). Debe ser un parámetro libre del modelo.
- **a**: Razón entre el semi-eje mayor y el radio de la estrella,  $a/R_*$ ; debe ser un parámetro libre del modelo.
- **inc**: Inclinação de la órbita del planeta,  $i$ ; debe ser un parámetro libre del modelo.
- **ecc**: Eccentricidad de la órbita. Fíjela en 0.

<sup>1</sup>Note que, si bien sabemos el ruido de medición para  $F(t)$  (que, por enunciado, era 100 ppm), hacer la conversión al logaritmo es un poco complicado. Aún así, usted notará en sus resultados que el ruido observado es bastante más pequeño que el ruido esperado solamente por la medición, por lo que consideraremos el ruido en esta parte del modelo como completamente desconocido.

<sup>2</sup><http://astro.uchicago.edu/~kreidberg/batman/quickstart.html>

- 
- `w`: Longitud de periastro de la órbita. Fíjela en 90.
  - `u`: Coeficientes de oscurecimiento al limbo de la estrella. Fíjelos en  $[0.1, 0.3]$ .
  - `limb_dark`: Modelo de oscurecimiento al limbo. Fíjelo en “quadratic”.

Note que (i) el modelo es altamente no-lineal (tanto por los parámetros del tránsito como por la varianza del ruido,  $\sigma^2$ , que en este caso ud. debe estimar) y (ii) ud. no sabe cuántas señales  $S_j(t)$  debe tener el modelo y, por tanto, no sabe cuántos coeficientes  $\alpha_j$  debe ajustar. Para resolver (i), use un MCMC para ajustar el modelo completo al flujo de la estrella objetivo. Resuelva (ii) como ud. estime conveniente. Grafique las distribuciones a-posteriori de los parámetros ajustados; compare su mejor ajuste con los datos en un gráfico.

- ¿Cree ud. que el modelo de ruido usado en la parte b) es suficiente para modelar los datos? Para responder esta pregunta, modele  $\epsilon(t)$  como un proceso gaussiano ad-hoc a lo observado en los residuos de la parte a). Compare las distribuciones a-posteriori para la razón entre el radio del planeta y el radio de la estrella,  $R_p/R_*$  ajustado con (i.e., lo hecho en este ítem) y sin (lo hecho en el ítem anterior) un proceso gaussiano.
- Discuta, usando argumentos estadísticos aprendidos en el curso, si se quedaría con el modelo ajustado en b) o en c). Entregue su mejor estimador para  $R_p/R_*$ , calculando la mediana de la distribución a-posteriori, y la banda de credibilidad de 68% alrededor de este parámetro.