Отношение сигнал/шум в астрофизических наблюдениях

Лекция 12

Отношение сигнал/шум в астрофизических временных рядах. Случайные и систематические вариации

КазНУ имени аль-Фараби

Цель и план лекции

Цель

Изучение определения S/N во временных рядах, разделения случайных и систематических вариаций, методов удаления трендов и оценки значимости переменных сигналов.

1. Введение
2. Определение S/N сигнала во временном ряду
3. Систематические вариации
4. Методы удаления трендов
5. Бинирование и шкала времени
6. Статистики переменности
7. Постановка наблюдений временных рядов

1 Введение

- Многие астрофизические задачи требуют измерения изменений во времени: это затмения двойных звёзд, транзиты экзопланет, пульсации, вспышки и изменения блеска активных ядер галактик.
- Космические миссии CoRoT, Kepler и TESS показали, что при низком уровне случайного шума главной проблемой становятся систематические эффекты, связанные с движением космического аппарата и изменением теплового режима.

2 Определение S/N сигнала во временном ряду

- Рассмотрим ряд $y_i = s(t_i) + n_i$ с белым шумом σ. Оптимальная оценка амплитуды A для шаблона g(t) задаётся выражением:

$$\hat{A} = \Sigma\, g_i\, y_i \,/\, \Sigma\, g_i^2, \quad \sigma_A = \sigma \,/\, \sqrt{(\,\Sigma\, g_i^2\,)}.$$

- Следовательно, $S/N = A\sqrt{(\Sigma\, g_i^2)}/\sigma$. Для транзита глубиной δ, содержащего N отсчётов внутри транзита, получаем при учёте равномерного уровня вне транзита:

$$S/N = \delta\, \sqrt{N} \,/\, \sigma.$$

2 Определение S/N сигнала во временном ряду (продолжение)

- Если наблюдается n транзитов, то $N = n \cdot N_1$, и S/N растёт как $\sqrt{n}$. В миссии Kepler порог обнаружения кандидата был установлен на уровне 7,1σ, что учитывало количество проверяемых периодов и фаз: число независимых вариантов достигает порядка 10^{12} для одной звезды, поэтому ожидается не более одного ложного события на весь обзор.

- Для периодического сигнала с амплитудой A и N равномерно распределёнными отсчётами погрешность амплитуды равна $\sigma\sqrt{(2/N)}$, поэтому:

$$S/N = (A/\sigma)\,\sqrt{(N/2)}.$$

- Погрешность частоты определяется формулой Монтгомери и О'Донохью (1999): $\sigma_f = (\sqrt{6}/\pi)\,\sigma / (\sqrt{N}\,A\,T)$, где T – полная длина ряда.

3 Систематические вариации

Систематические изменения могут быть внешними и внутренними.

Атмосфера

Изменение прозрачности и воздушной массы, сцинтилляции.

Прибор

Изменение положения звезды на пикселях, изменение фокуса, тепловые колебания.

Источник

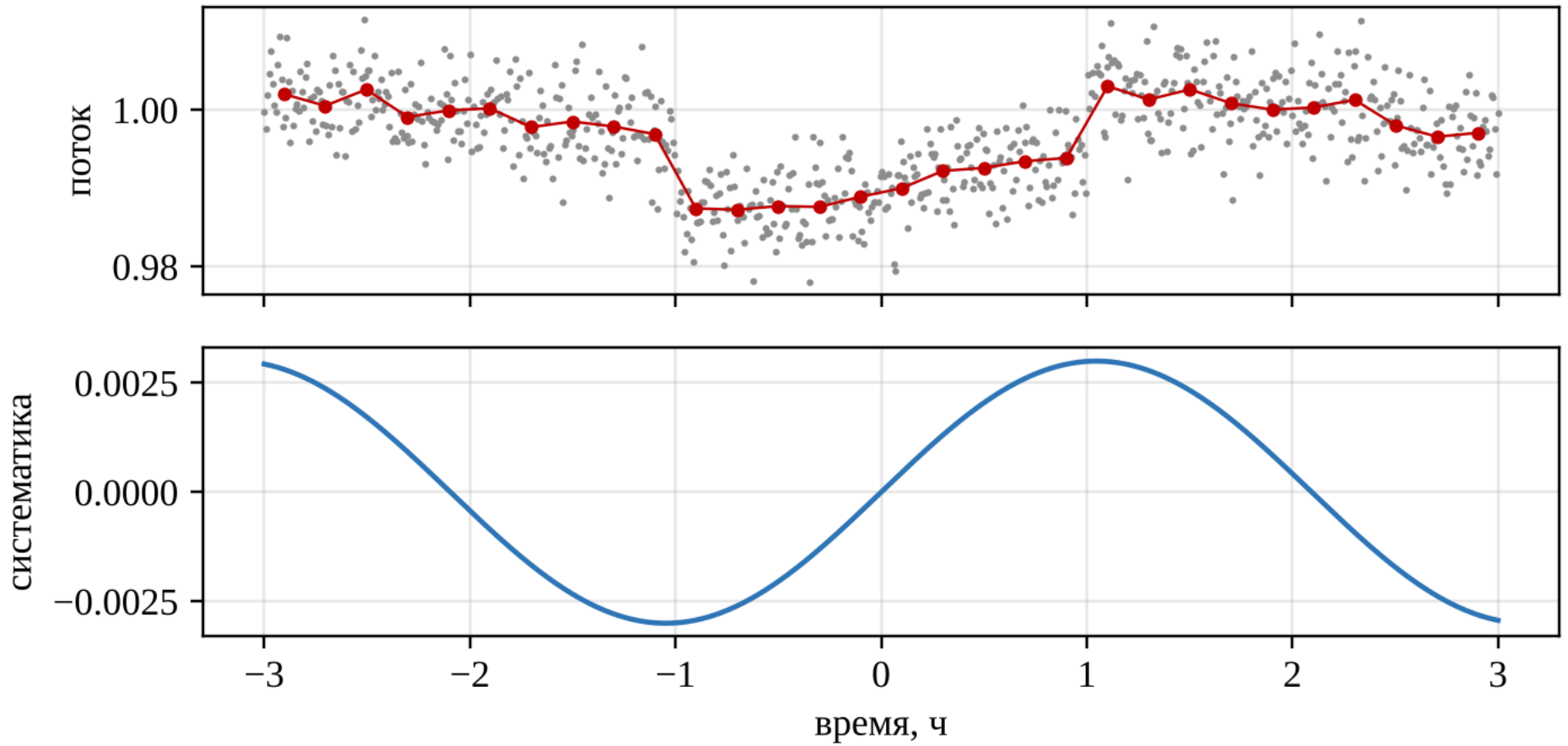
Звёздная активность, пятна, грануляция и другие вариации.

Обработка

Неточные плоские поля, изменение фона, неудачный выбор опорных звёзд.

Систематические эффекты изменяются медленно, и поэтому имеют длительную корреляцию.

3 Систематические вариации (продолжение)



Модель транзитной кривой блеска, бинированные значения и систематический тренд

4 Методы удаления трендов

Дифференциальная фотометрия

Устраняет общие для поля изменения прозрачности.

Полиномиальная подгонка

Простая модель медленных трендов, применяется до и после события. Подгонка без исключения области сигнала частично подавляет сам сигнал.

Метод декорреляции

Линейная регрессия с параметрами состояния (положение, FWHM, воздушная масса).

SysRem

Алгоритм Тамуза и др. (2005) находит систематические эффекты, общие для многих звёзд, не требуя знания их причин.

Основная опасность удаления тренда заключается в подавлении реального сигнала.

4 Методы удаления трендов (продолжение)

TFA

Алгоритм фильтрации трендов (Ковач и др., 2005) аппроксимирует кривую каждой звезды как линейную комбинацию кривых других звёзд.

Гауссовы процессы

Метод обеспечивает вероятностное моделирование коррелированных вариаций и позволяет определить неопределённости.

Основная опасность удаления тренда заключается в подавлении реального сигнала.

5 Бинирование и шкала времени

- Временной ряд часто бинируют, поскольку транзит длится часы, а период регистрации – минуты.

5 S/N транзита и влияние числа отсчётов

Листинг 12.1.

```
import numpy as np

rng = np.random.default_rng(12)
n = 720
t = np.linspace(-3, 3, n)
depth, sig = 0.005, 0.004
intr = np.abs(t) < 1.0
model = 1 - depth * intr

est = []
for _ in range(3000):
    y = model + rng.normal(0, sig, n)
    est.append(y[~intr].mean() - y[intr].mean())
est = np.array(est)

print(f"Точек в транзите: {intr.sum()}")
print(f"Оценка глубины: {est.mean():.5f}")
print(f"S/N (моделирование): {depth / est.std():.1f}")
Nin, Nout = intr.sum(), (~intr).sum()
print(f"S/N (формула): {depth / (sig * np.sqrt(1 / Nin + 1 / Nout)):.1f}")
print(f"Упрощённо d*sqrt(N)/sigma: {depth * np.sqrt(Nin) / sig:.1f}")
```

5 Результат: S/N транзита и влияние числа отсчётов

```
Точек в транзите: 240
Оценка глубины: 0.00500
S/N (моделирование): 15.8
S/N (формула): 15.8
Упрощённо d*sqrt(N)/sigma: 19.4
```

5 Удаление тренда и подавление транзита

Листинг 12.2.

```
trend = 0.006 * t**2 - 0.002 * t
y = model + trend + rng.normal(0, sig, n)

def depth_of(res):
    return res[~intr].mean() - res[intr].mean()

p1 = np.polyfit(t, y, 2)
r1 = y / np.polyval(p1, t)
p2 = np.polyfit(t[~intr], y[~intr], 2)
r2 = y / np.polyval(p2, t)

print(f"Без удаления тренда: {depth_of(y):.5f}")
print(f"Подгонка по всем точкам: {depth_of(r1):.5f}")
print(f"Подгонка вне транзита:   {depth_of(r2):.5f}")
print(f"Заданная глубина:        {depth:.5f}")
```

5 Результат: Удаление тренда и подавление транзита

```
Без удаления тренда: 0.02861
Подгонка по всем точкам: 0.00233
Подгонка вне транзита:   0.00458
Заданная глубина:        0.00500
```

Подгонка полинома по всем точкам частично устраняет транзит, поскольку он коррелирует с параболой.

6 Статистики переменности

- Для определения переменных объектов используют:
- приведённое значение χ^2 относительно постоянного значения;
- индекс Стетсона, использующий совместные изменения в парах наблюдений;
- нормированный избыток дисперсии;
- отношение фон Неймана: квадрат разности последовательных значений к дисперсии, которое чувствительно к гладким изменениям.
- Для прерывистых и неравномерных серий применяют структурную функцию, показывающую среднеквадратичную разность блеска как функцию лага времени.

7 Постановка наблюдений временных рядов

- Качество временного ряда определяется не только шумом одного отсчёта, но и условиями наблюдения.
- Моменты времени необходимо приводить к единой шкале. Барицентрическая поправка, равная до ±8,3 минуты, компенсирует движение Земли вокруг Солнца.
- Погрешность определения времени события зависит от S/N и формы его ингресса и эгресса.
- Для анализа множества объектов применяют конвейеры, включающие фотометрию, удаление систематики, поиск сигналов и ветирование кандидатов.
- Упражнение: на данных космического телескопа TESS (доступны через MAST) измерьте S/N известного транзита при разных методах удаления тренда и сравните глубину транзита с опубликованным значением.

8 Разбор примера

1. Пусть глубина транзита составляет 0,5 %, шум отсчёта 0,4 %, а в транзите имеется 240 отсчётов.

2. Тогда S/N транзита приближённо равен $0{,}005 \cdot \sqrt{240} / 0{,}004 = 0{,}005 \cdot 15{,}5 / 0{,}004 \approx 19{,}4$.

3. Если доступно 9 транзитов, S/N увеличится в 3 раза и составит примерно 58.

4. Для определения частоты при A = 1 %, σ = 0,4 %, N = 1000, T = 30 сут получаем $\sigma_f = (\sqrt{6}/\pi) \cdot 0{,}4 / (\sqrt{1000} \cdot 1 \cdot 30) = 0{,}78 \cdot 0{,}4 / (31{,}6 \cdot 30) \approx 3{,}3 \cdot 10^{-4}$ сут$^{-1}$.

9 Вопросы для обсуждения на занятии

Детрендинг

Как удостовериться, что процедура удаления тренда не исказила научный сигнал?

Пороги

Почему в обзоре Kepler был использован порог 7,1σ?

Планирование

Что важнее для определения периода: количество измерений или длина ряда?

Контрольные вопросы

1. Как определить S/N сигнала заданной формы во временном ряду?
2. От чего зависит S/N транзита?
3. Как определяется погрешность частоты периодического сигнала?
4. Какие источники систематических вариаций существуют?
5. Какие методы применяются для удаления трендов?
6. Почему удаление тренда может подавлять реальный сигнал?
7. Что такое CDPP?
8. Какие статистики применяются при поиске переменных объектов?

Список литературы

1. Tamuz O., Mazeh T., Zucker S. Correcting systematic effects in a large set of photometric light curves // Monthly Notices of the Royal Astronomical Society. – 2005. – Vol. 356. – P. 1466–1470.
2. Kovács G., Bakos G., Noyes R.W. A trend filtering algorithm for wide-field variability surveys // Monthly Notices of the Royal Astronomical Society. – 2005. – Vol. 356. – P. 557–567.
3. Montgomery M.H., O'Donoghue D. A derivation of the errors for least squares fitting to time series data // Delta Scuti Star Newsletter. – 1999. – Vol. 13. – P. 28–32.
4. Jenkins J.M., et al. Overview of the Kepler science processing pipeline // The Astrophysical Journal Letters. – 2010. – Vol. 713. – P. L87–L91.
5. Ivezić Ž., Connolly A.J., VanderPlas J.T., Gray A. Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data. – Updated ed. – Princeton: Princeton University Press, 2020.
6. Basu S., Chaplin W.J. Asteroseismic Data Analysis: Foundations and Techniques. – Princeton: Princeton University Press, 2017.