Техника исследования гравитационных волн

Лекция 9

Информационно-энтропийные характеристики нелинейных и нестационарных сигналов

Усіпов Н.М., PhD · КазНУ имени аль-Фараби

Цель и план лекции

Цель

Изучение энтропийных мер, учитывающих временной порядок значений, методов анализа нестационарных сигналов в скользящем окне и способов различения шума, регулярной и хаотической динамики.

1. Введение
2. Перестановочная энтропия
3. Приближённая и выборочная энтропия
4. Плоскость энтропия–сложность
5. Анализ нестационарных сигналов
6. Энтропия и нелинейная динамика
7. Параметры и статистическая проверка
8. Статистики высших порядков и фрактальные характеристики

1 Введение

- Гравитационно-волновой сигнал и шум детектора являются нестационарными: частота сигнала меняется во времени, а свойства шума подвержены дрейфу и глитчам.
- Важной особенностью таких мер является их чувствительность к нелинейной зависимости, которую не обнаруживают автокорреляционная функция и спектр мощности второго порядка.

2 Перестановочная энтропия

- Предложена Бандтом и Помпе в 2002 году. Сигнал разбивается на векторы длиной m с шагом задержки τ, и каждому вектору ставится в соответствие порядковый паттерн – перестановка, которая сортирует его элементы по величине.

$$H_P = -\sum p(\pi)\,\log p(\pi)\,/\,\log m!.$$

- Метод обладает важными достоинствами: он инвариантен к монотонным преобразованиям амплитуды, робастен и прост в вычислении.

3 Приближённая и выборочная энтропия

Эти меры оценивают условную вероятность того, что близкие фрагменты длиной m остаются близкими при увеличении длины до m+1.

Приближённая энтропия (ApEn, Пинкус, 1991)

Использует сравнение вектора с самим собой, поэтому результат оказывается смещённым.

Выборочная энтропия (SampEn, Ричман и Мурман, 2000)

Исключает самосравнения:

$$SampEn = -\ln(A / B),$$

Многомасштабная энтропия (Коста и др., 2002)

Расчёт SampEn для последовательно огрубляемых рядов. Различает сложный сигнал с многомасштабной структурой от некоррелированного шума.

Спектральная и вейвлет-энтропия

Определяются для распределения энергии по частотным компонентам.

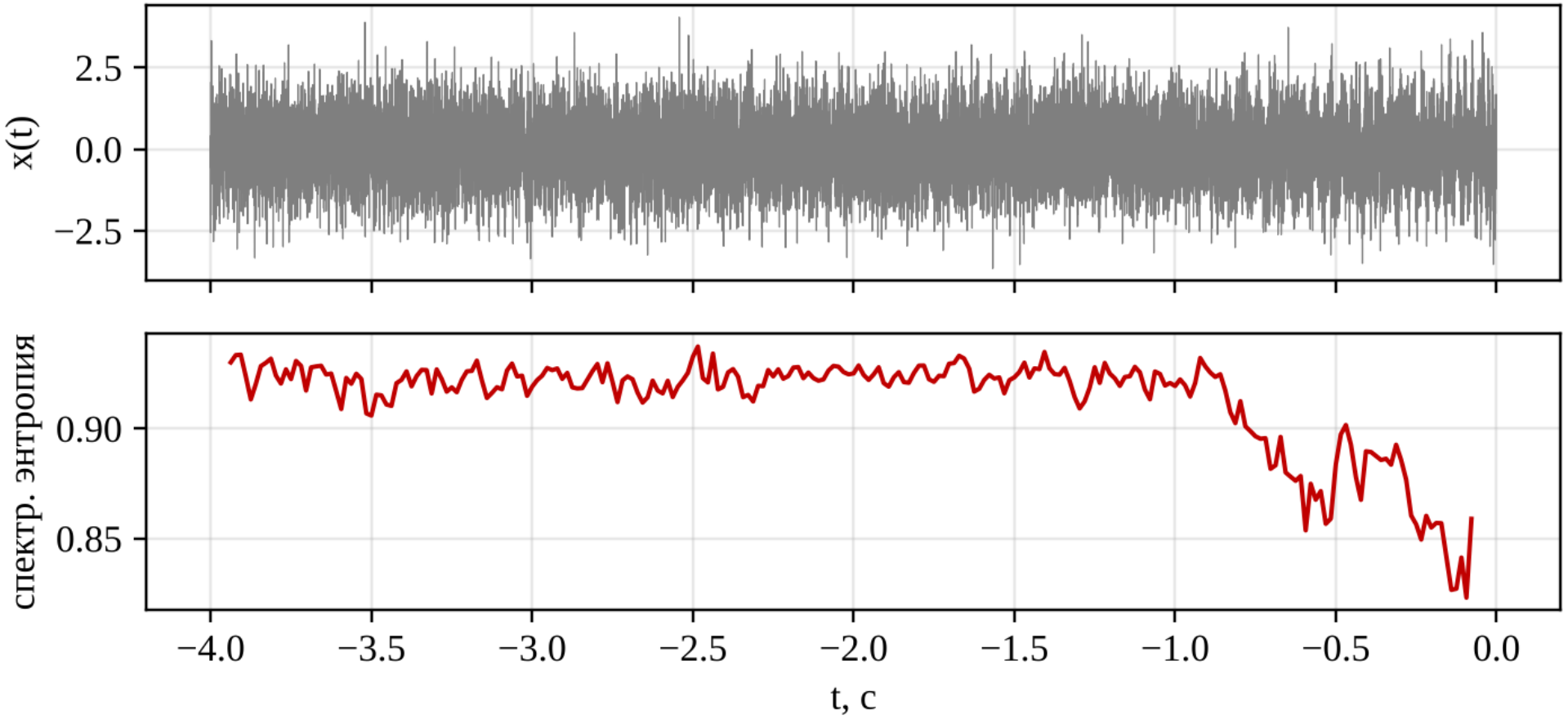
4 Плоскость энтропия–сложность

- Значение энтропии само по себе не различает хаотическую динамику и шум. В плоскости «энтропия–статистическая сложность» (Россо и др., 2007) для каждого ряда вычисляются перестановочная энтропия H и сложность $C = Q_J[P, P_e]\cdot H$, где Q_J – дивергенция Йенсена–Шеннона между распределением паттернов и равномерным распределением.

5 Анализ нестационарных сигналов

- Для анализа временной эволюции применяется расчёт энтропии в скользящем окне.

5 Анализ нестационарных сигналов (продолжение)



Сигнал в белом шуме и спектральная энтропия в скользящем окне

5 Анализ нестационарных сигналов (продолжение)

При этом следует учитывать следующие обстоятельства.

Выбор окна

Длина окна определяет компромисс между временным разрешением и точностью оценки.

Пороговое значение

Порог для события выбирается по шумовым участкам с учётом разброса оценок и требуемой частоты ложных срабатываний.

Подготовка данных

Окрашенный шум снижает спектральную энтропию ещё до появления сигнала, поэтому необходимо предварительное отбеливание.

Тип энтропии

Сочетание нескольких мер обеспечивает более полное описание сигнала.

5 Перестановочная и выборочная энтропия

Листинг 9.1.

```
import numpy as np
from math import factorial
from scipy.spatial.distance import pdist

def perm_entropy(x, m=3, tau=1):
    k = (m - 1) * tau + 1
    w = np.lib.stride_tricks.sliding_window_view(x, k)[:, ::tau]
    _, c = np.unique(np.argsort(w, axis=1), axis=0, return_counts=True)
    p = c / c.sum()
    return -(p * np.log(p)).sum() / np.log(factorial(m))

def samp_entropy(x, m=2, r=0.2):
    r *= x.std()
    def count(k):
        w = np.lib.stride_tricks.sliding_window_view(x, k)[:x.size - m]
        return (pdist(w, "chebyshev") <= r).sum()
    return -np.log(count(m + 1) / count(m))
```

5 Перестановочная и выборочная энтропия (продолжение)

Листинг 9.1.

```
N = 1500
t = np.arange(N) / 200
sine = np.sin(2 * np.pi * 5 * t)
log = np.empty(N)
log[0] = 0.4
for i in range(1, N):
    log[i] = 4 * log[i - 1] * (1 - log[i - 1])
noise = np.random.default_rng(1).normal(size=N)

print("Сигнал              PermEn  SampEn")
for name, x in (("Синусоида", sine), ("Логистич. отобр.", log),
                ("Белый шум", noise)):
    print(f"{name:18} {perm_entropy(x):6.3f} {samp_entropy(x):7.3f}")
```

5 Результат: Перестановочная и выборочная энтропия

```
Сигнал              PermEn  SampEn
Синусоида           0.510   0.281
Логистич. отобр.    0.828   0.642
Белый шум           0.999   2.156
```

Энтропия закономерно возрастает при переходе от регулярной к случайной динамике.

5 Анализ нестационарных сигналов (продолжение)

- Полезно заметить, что величина τ задаёт временной масштаб анализа и соответственно частотную полосу, к которой метод наиболее чувствителен.

5 Перестановочная энтропия в скользящем окне

Листинг 9.2.

```
rng = np.random.default_rng(2)
fs, N = 1024, 4096
t = np.arange(N) / fs
x = rng.normal(size=N)
env = np.exp(-((t - 2.0) / 0.15)**2)
x += 3 * env * np.sin(2 * np.pi * 40 * t)

L, st = 256, 128
pe = np.array([perm_entropy(x[i:i + L], m=4, tau=4)
               for i in range(0, N - L, st)])
tc = (np.arange(pe.size) * st + L / 2) / fs
k = pe.argmin()
print(f"Минимум PermEn = {pe[k]:.3f} при t = {tc[k]:.2f} с")
print(f"Медиана по шуму: {np.median(pe):.3f}")
```

```
Минимум PermEn = 0.886 при t = 2.00 с
Медиана по шуму: 0.986
```

6 Энтропия и нелинейная динамика

- Понятие энтропии в теории динамических систем формализовано Колмогоровым и Синаем.
- Перестановочная энтропия при малом m характеризует только короткие участки временного ряда, а значит, неявно связывает результат с частотой дискретизации.
- Для чирпа при увеличении частоты сигнала отношение частоты дискретизации к частоте сигнала уменьшается, поэтому ординальные паттерны будут медленно изменяться во времени.

7 Параметры и статистическая проверка

- Выборочная энтропия зависит от трёх параметров: длины шаблона m, допуска r и длины ряда N. Как правило, m = 2, r = 0,1–0,25σ, а N должна быть не менее нескольких сотен точек.
- Для проверки значимости нелинейного эффекта применяется метод суррогатных данных (Тейлер и др., 1992).

8 Статистики высших порядков и фрактальные характеристики

- Спектр мощности описывает только статистику второго порядка. Биспектр и триспектр отражают фазовые связи между гармониками, что позволяет выявлять нелинейное взаимодействие.
- Для анализа долговременной памяти применяются показатель Херста и анализ флуктуаций с удалённым трендом (DFA).
- Для нестационарных сигналов применяется анализ распределения энергии в частотно-временной плоскости с последующим расчётом энтропии Реньи.
- Также следует учитывать практический фактор: расчёт энтропии в окне применяется к отбелённым данным, чтобы избежать влияния окрашенного спектра.

9 Разбор примера

1. Вычислим перестановочную энтропию ряда x = (4, 7, 9, 10, 6, 11, 3) при m = 3 (пример из работы Бандта и Помпе).
2. Тройки: (4,7,9) и (7,9,10) возрастающие; (9,10,6) и (6,11,3) имеют порядок $x_3 < x_1 < x_2$; (10,6,11) имеет порядок $x_2 < x_1 < x_3$.
3. Частоты паттернов равны 2/5, 2/5 и 1/5.
4. Энтропия $H = -2 \cdot 0{,}4 \ln 0{,}4 - 0{,}2 \ln 0{,}2 \approx 1{,}055$ нат (1,52 бит).
5. Нормированное значение равно $1{,}055/\ln 6 \approx 0{,}589$.

10 Вопросы для обсуждения на занятии

Параметры

Как выбрать значения m и τ для анализа данных с частотой дискретизации 4096 Гц?

Суррогаты

Почему окрашенный гауссовый шум может быть ошибочно принят за сигнал со структурой?

Сопоставление

Какое преимущество может дать одновременное использование перестановочной и спектральной энтропии?

Контрольные вопросы

1. Почему анализ амплитудной гистограммы не учитывает динамику сигнала?
2. Опишите алгоритм вычисления перестановочной энтропии.
3. В чём заключается различие приближённой и выборочной энтропии?
4. Для чего предназначена многомасштабная энтропия?
5. Что такое плоскость «энтропия–сложность» и как в ней различаются процессы?
6. Как определяется длина окна при анализе нестационарного сигнала?
7. Почему перед вычислением спектральной энтропии данные следует отбеливать?
8. Как параметр задержки τ влияет на анализ сигнала?

Список литературы

1. Bandt C., Pompe B. Permutation entropy: A natural complexity measure for time series // Physical Review Letters. – 2002. – Vol. 88. – Art. 174102.
2. Pincus S.M. Approximate entropy as a measure of system complexity // Proceedings of the National Academy of Sciences. – 1991. – Vol. 88. – P. 2297–2301.
3. Richman J.S., Moorman J.R. Physiological time-series analysis using approximate entropy and sample entropy // American Journal of Physiology – Heart and Circulatory Physiology. – 2000. – Vol. 278. – P. H2039–H2049.
4. Costa M., Goldberger A.L., Peng C.-K. Multiscale entropy analysis of complex physiologic time series // Physical Review Letters. – 2002. – Vol. 89. – Art. 068102.
5. Rosso O.A., Larrondo H.A., Martin M.T., Plastino A., Fuentes M.A. Distinguishing noise from chaos // Physical Review Letters. – 2007. – Vol. 99. – Art. 154102.
6. Fadlallah B., Chen B., Keil A., Príncipe J. Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information // Physical Review E. – 2013. – Vol. 87. – Art. 022911.
7. Shannon C.E. A mathematical theory of communication // Bell System Technical Journal. – 1948. – Vol. 27. – P. 379–423, 623–656.