



Л е к ц и я № 3

Корреляция и свертка

Разработал профессор П. М. Васильев
Кафедра фармакологии и биоинформатики

Для студентов, обучающихся по направлению 06.03.01 «Биология»
профили Биохимия, Генетика
при изучении дисциплины «Цифровые технологии в биологии»

П л а н л е к ц и и

- Корреляция
- Свертка
- Ковариация
- Парная корреляция Пирсона
- Ранговая корреляция Спирмена

Корреляция

Степень похожести двух сигналов
или двух случайных величин.

Отражает эмпирическую зависимость
между величинами, при которой
изменение одной из величин
приводит к систематическому
изменению другой величины.

Корреляция

$$r_{12} = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n)$$

n	1	2	3	4	5	6	7	8	9
x_1	4	2	-1	3	-2	-6	-5	4	5
x_2	-4	1	3	7	4	-2	-8	-2	1

$$\begin{aligned} r_{12} &= \frac{1}{9} (4 \times (-4) + 2 \times 1 + (-1) \times 3 + 3 \times 7 + (-2) \times 4 + (-6) \times (-2) + \\ &\quad + (-5) \times (-8) + 4 \times (-2) + 5 \times 1) = \\ &= 5. \end{aligned}$$

Корреляция с фазовой связью

$$r_{12}(j) = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n+j)$$

n	1	2	3	4	5	6	7	8	9
x_1	4	2	-1	3	-2	-6	-5	4	5
x_2	7	4	-2	-8	-2	-1			

$$r_{12}(3) = \frac{1}{9}(4 \times 7 + 2 \times 4 + (-1) \times (-2) + 3 \times (-8) + (-2) \times (-2) + (-6) \times (-1)) = 2,667.$$

Корреляция непрерывных сигналов

$$r_{12}(\tau) = \int_{-\infty}^{\infty} x_1(t)x_2(t + \tau)$$

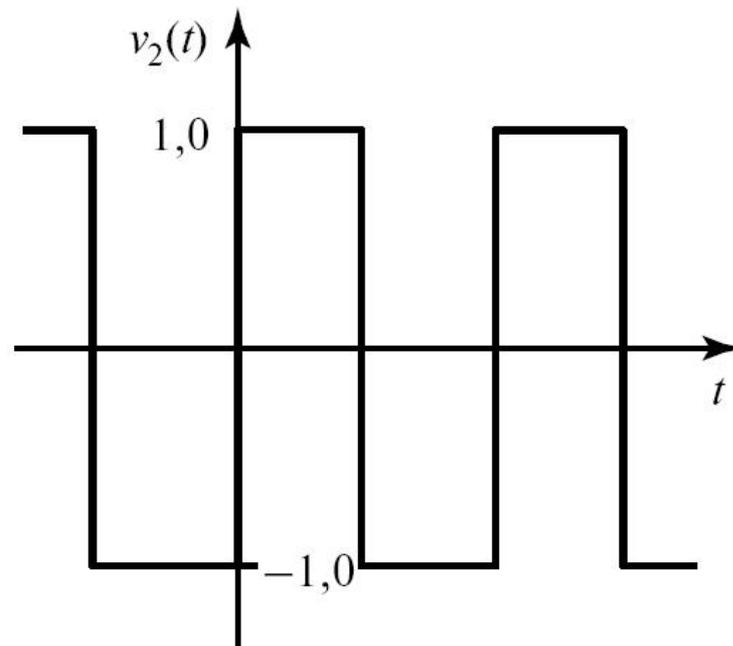
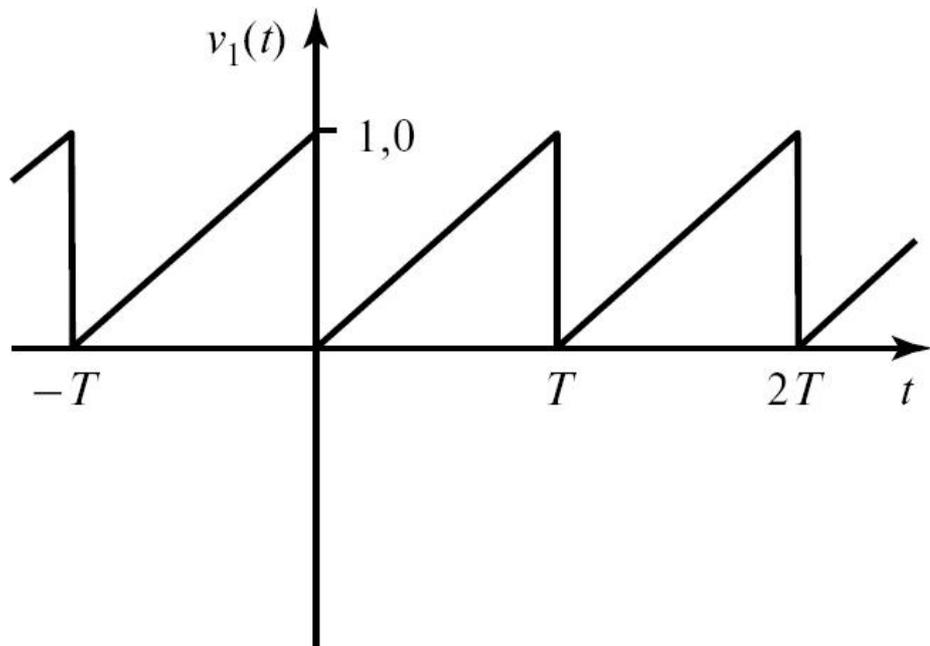
$$r_{12}(\tau) = \frac{1}{T} \int_0^T x_1(t)x_2(t + \tau)dt$$

В случае записей конечной длины

Коэффициент взаимной корреляции

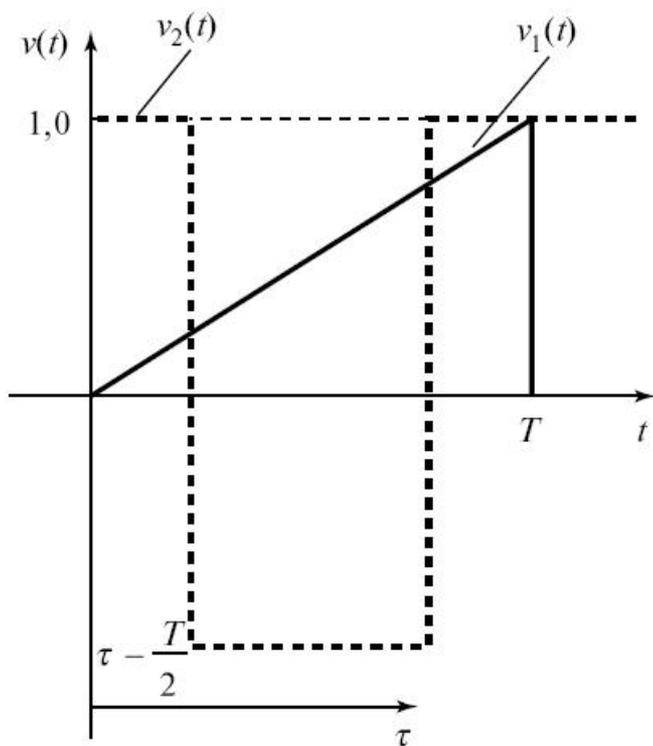
$$\rho_{12}(j) = \frac{r_{12}(j)}{\frac{1}{N} \left[\sum_{n=0}^{N-1} x_1^2(n) \sum_{n=0}^{N-1} x_2^2(n) \right]^{1/2}}$$

Корреляция двух сигналов

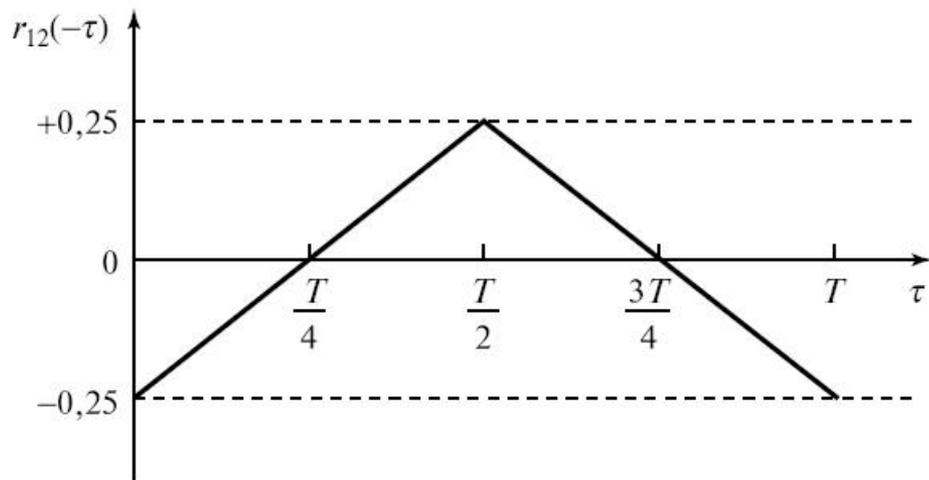


Исходные сигналы

Корреляция двух сигналов



Сегменты сигналов



Результат корреляции

Свертка

Взаимная корреляция одной последовательности с обращенной по времени второй.

Описывает, как выход системы определяется взаимодействием входа с самой системой.

Свертка

$$y(n) = \sum_{m=0}^n h(n-m)x(m)$$

Конечная последовательность

$$y(n) = \sum_{m=-\infty}^{\infty} x(m)h(n-m) = x(n) \circledast h(n)$$

Бесконечная последовательность

Свертка

n	1	2	3	4	5	6	7	8	9
x₁	4	2	-1	3	-2	-6	-5	4	5
x₂	-4	1	3	7	4	-2	-8	-2	1

Исходные данные

n	1	2	3	4	5	6	7	8	9
x₁	4	2	-1	3	-2	-6	-5	4	5
x₂	1	-2	-8	-2	-4	7	3	1	-4

Данные для свертки

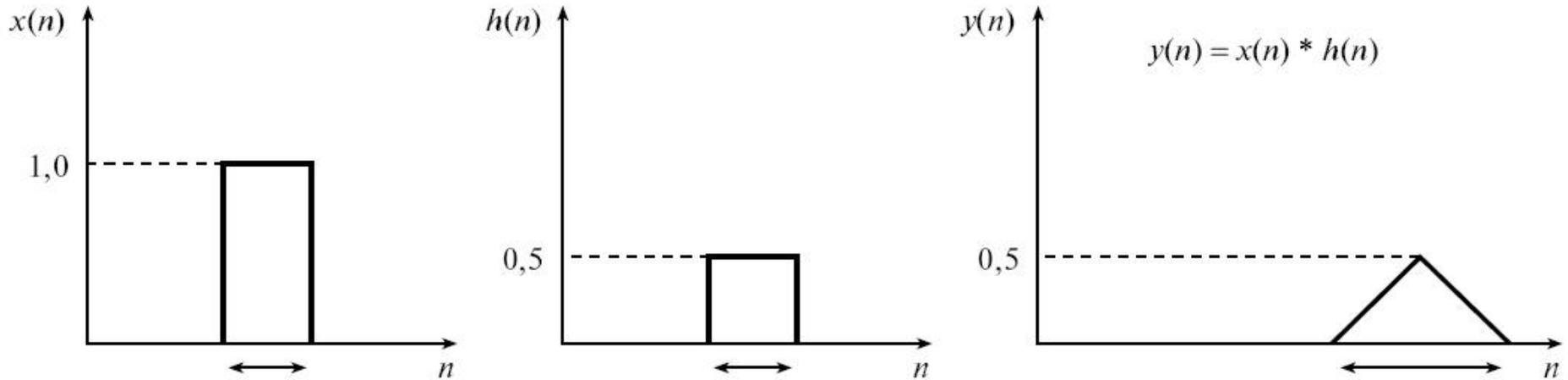
$$\begin{aligned} \mathbf{x_1 \bullet x_2} &= (4 \cdot 1) + (2 \cdot -2) + (-1 \cdot -8) + (3 \cdot -2) + (-2 \cdot -4) + (-6 \cdot 7) + (-5 \cdot 3) \\ &+ (4 \cdot 1) + (5 \cdot -4) = 4 - 4 + 8 - 6 + 8 - 42 - 15 + 4 - 20 = \mathbf{-63} \end{aligned}$$

Свертка непрерывных сигналов

$$y(t) = x(t) \circledast h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau$$

Интеграл свертки

Свертка прямоугольных импульсов



Дисперсия

Характеризует степень разброса случайной величины около её среднего.

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Ковариация

Мера совместной изменчивости двух случайных величин.

$$\bar{s}_{xy} = \text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

Коэффициент корреляции Пирсона

$$R_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

Коэффициент корреляции Спирмена

$$R_{Sp} = 1 - \frac{6 \sum_{i=1}^N (\text{rank}(x_i) - \text{rank}(y_i))^2}{N(N^2 - 1)}$$

To be continued ...

