2.8 Корреляционная функция

2.8.1 Ковариация и корреляция

Пусть $X = [x_1,\ x_2, \ldots, x_N]$ и $Y = [y_1, y_2, \ldots, y_N]$ – случайные величины (или последовательности). Требуется определить: являются ли X и Y независимыми либо между ними существует связь?

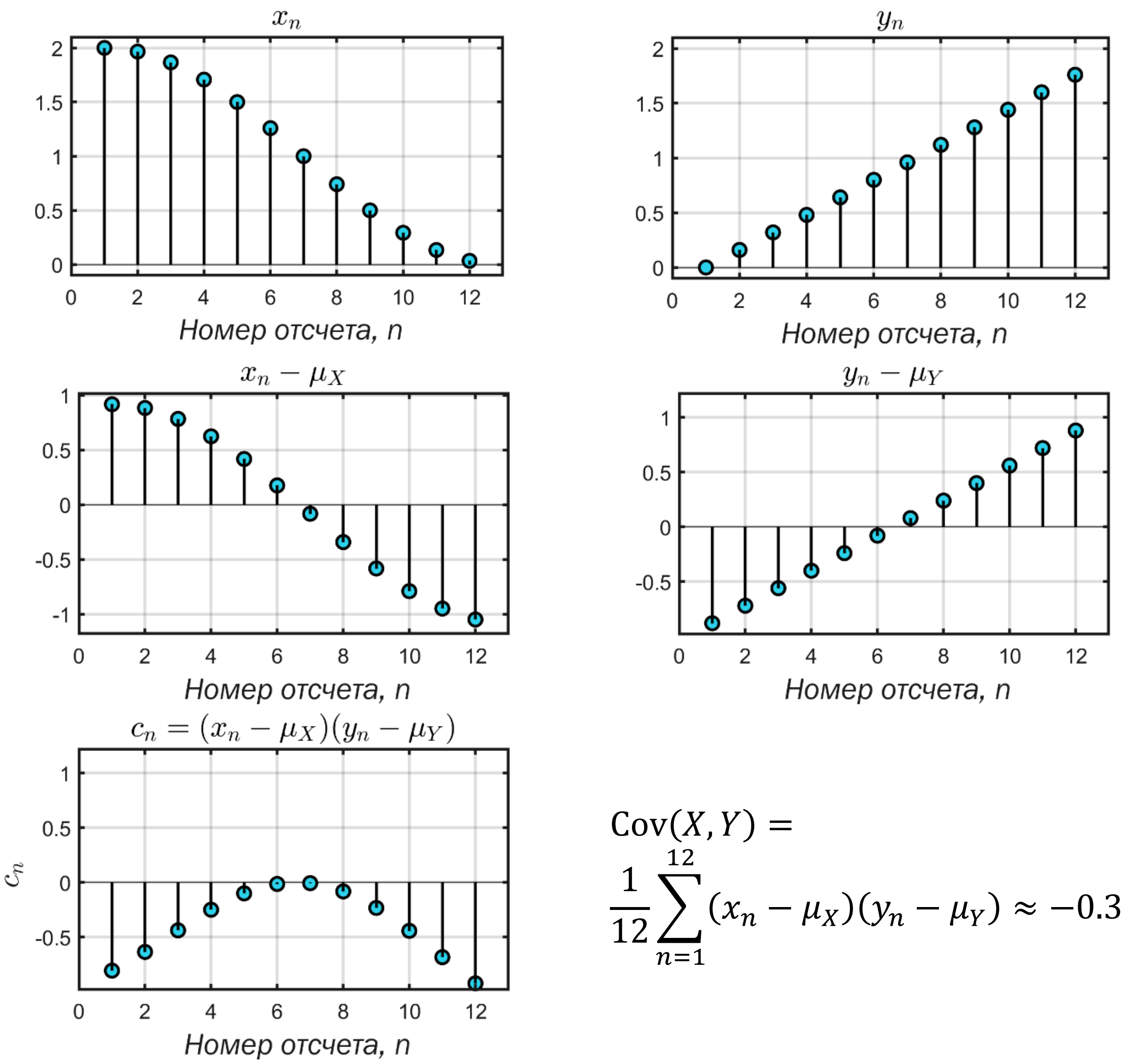
Это можно выяснить, рассчитав ***ковариацию*** между случайными величинами X и Y:

$$\mathrm{Cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu_X)(y_n-\mu_Y), \tag{2.24}$$

где $\mu_X = E\{X\} = \frac{1}{N}\sum_{n=1}^{N} x_n$ – математическое ожидание X, а μ_Y – математическое ожидание Y.

Удобно рассматривать данную величину как показатель зависимости частного вида.

На рисунке 2.12 приведен пример расчета ковариации для двух последовательностей.



$$\mathrm{Cov}(X,Y) = \frac{1}{12}\sum_{n=1}^{12}(x_n-\mu_X)(y_n-\mu_Y) \approx -0.38.$$

Рисунок 2.12 – Расчет ковариации для двух последовательностей

Отрицательное значение ковариации говорит о том, что имеет место обратная зависимость (при увеличении X значение Y уменьшается).

Если случайные величины изначально имеют нулевое мат. ожидание, т.е. $\mu_X = 0$ и $\mu_Y = 0$, то выражение для ковариации упрощается:

$$\mathrm{Cov}(X, Y) = \frac{1}{N}\sum_{n=1}^{N} x_n y_n. \tag{2.25}$$

Очевидно, что если две последовательности независимы, то сумма произведений $x_n y_n$ стремиться к исчезающе малому случайному числу по мере увеличения пар точек. Это объясняется тем, что все числа положительные и отрицательные равновероятны, так что пары произведений компенсируются при сложении

Если в качестве показателя зависимости использовать ковариацию $\mathrm{Cov}(X, Y)$, то серьезным недостатком является то, что её значение может меняться при изменении единиц, в которых измеряются исходные случайные величины. Данный эффект можно исключить, если разделить ковариацию произведение среднеквадратических отклонений (СКО) $\sigma_X \sigma_Y$:

$$r_{X,Y} = \frac{E\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y} = \frac{\sum_{n=1}^{N}(x_n - \mu_X)(y_n - \mu_Y)}{\sqrt{\sum_{n=1}^{N}(x_n - \mu_X)^2 \sum_{n=1}^{N}(y_n - \mu_Y)^2}}. \tag{2.26}$$

Полученное отношение называется ***коэффициентом (линейной) корреляции*** между X и Y. Или просто ***корреляцией***.

2.8.2 Корреляционная функция

В ЦОС дискретные сигналы в большинстве случаев нормированы, т.е. заключены в диапазоне $[-1, \quad 1]$, имеют среднее значение равное 0 и приблизительно одинаковые СКО, поэтому для оценки меры корреляции используют сумму произведений элементов последовательностей

$$r_{x,y} = \frac{1}{N}\sum_{n=1}^{N} x_n y_n$$

Часто сигналы являются неоднородными, т.е. на различные интервалы несут различную информацию. Для того, чтобы определять схожесть сигналов на различных временных интервалах вводится корреляционная функция:

$$r_{x,y}(\ell) = \sum_{n=-\infty}^{\infty} x(n+\ell)y^*(n) = \sum_{n=-\infty}^{\infty} x(n)y^*(n-\ell) \tag{2.27}$$

Функция корреляции (2.27) не зависит от времени n, но зависит от временной задержки ℓ (англ. *time lag*). В случае, когда в (2.27) сигналы $x(n)$ и $y(n)$ различны функцию называют ***кросс-корреляционной***. Если $y(n) = x(n)$ функцию называют ***автокорреляционной***.

Чем более похожи два сигнала при конкретной временной задержке ℓ, тем больше значение корреляционной функции.

Корреляционная функция для последовательностей длины N позволяет оценить зависимость между их отсчетами при различных временных сдвигах по времени[9] ℓ:

$$r_{xy}(\ell) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)y^*(n-\ell), \qquad -(N-1) \le \ell \le (N-1).$$

Пример 2.7 Рассчитайте корреляционную функцию для сигналов:

$$x(n) = u(n) - u(n-4) \text{ и } y(n) = u(n-3) - u(n-9).$$

Построить график корреляционной функции ($N = 10$).

Решение.

$r_{xy}\big(-(N-1)\big) = r_{xy}(-9) = 0$,

…

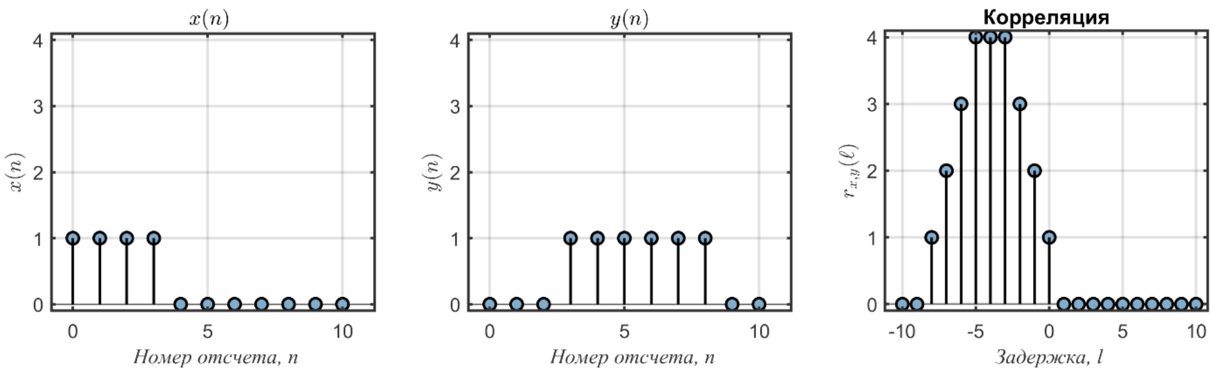


Рисунок 2.13 – Пример вычисления корреляции

Интерпретация: сигнал $y(n)$ опережает сигнал $x(n)$. Наибольшие значения корреляции достигается при значениях $\ell = \{-5, -4, -3\}$. Это значит, что если задержать сигнал $y(n)$ на 5, 4 или 3 отсчета, то произойдет его максимальное совмещение (согласование) с сигналом $x(n)$.

2.8.3 Автокорреляционная функция (АКФ)

АКФ последовательности длины N позволяет оценить зависимость между её отсчетами при различных временных сдвигах по времени[10] ℓ:

$$r_{xx}(\ell) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)x(n+\ell), -(N-1) \le \ell \le (N-1).$$

Для простоты будем считать, что сигнал $x(n)$ принимает только действительные значения.

Говорят, что АКФ является мерой самоподобия (self-similarity, самоповторения) сигнала на различном временном удалении ℓ. Когда значение

---

[9] Солонина А.И. и др. Цифровая обработка сигналов и Matlab, стр. 97.
[10] Солонина А.И. и др. Цифровая обработка сигналов и Matlab, стр. 97.

$r_{xx}(\ell)$ велико для какого-то ℓ, то говорят, что отсчеты расположенные на расстоянии ℓ имеют высокую корреляцию.

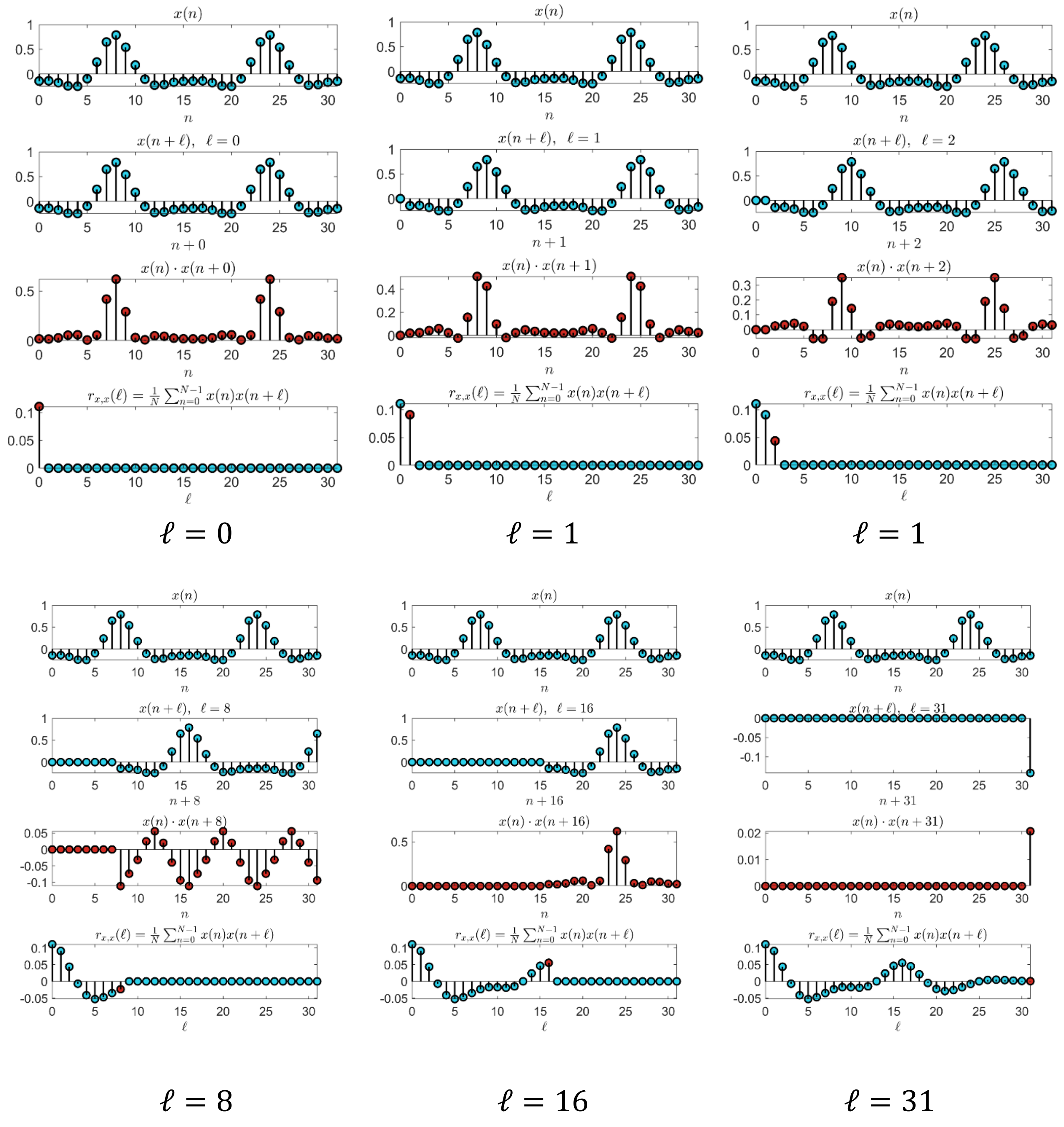


Рисунок 2.14 – Пример вычисления корреляционной функции дискретного сигнала

Рассмотрим процесс вычисления АКФ. На рисунке 2.14 показан сигнал $x(n)$ и процесс расчета для него автокорреляционной функции. При временном лаге $\ell = 0$ сигнал самосовмещается сам с собой. При $\ell = 1$ сигнал $x(n)$ сопоставляется со своей сдвинутой версией $x(n+1)$, в последовательности $x(n)x(n+1)$ появляются отрицательные члены, поэтому значение АКФ $r_{xx}(1) < r_{xx}(0)$. При дальнейшем увеличении ℓ вплоть до величины $\ell = 5$ значения АКФ падает, что говорит о том, что в при увеличении сдвига сигнал вс меньше и меньше напоминает сам себя. Однако после $\ell = 5$ АКФ начинает нарастать и достигает максимума при $\ell = 16$. В этой точке происходит полное совмещение первого и второго периода исходного сигнала. При дальнейшем увеличении ℓ значение АКФ убывает. Это обусловлено двумя факторами: 1) уменьшается число ненулевых отсчетов последовательности $x(n)x(n+\ell)$; 2) уменьшается степень схожести сигнала с самим собой при увеличении ℓ.