

2 ЛАБОРАТОРНАЯ РАБОТА №2. ЭНЕРГЕТИЧЕСКИЕ ХАРАКТЕРИСТИКИ СИГНАЛОВ

ЦЕЛЬ РАБОТЫ – изучить энергетические характеристики сигналов, научиться способы описания дискретных систем при помощи пакета MATLAB.

Изучение алгоритмов вычисления энергии и мощности дискретных сигналов. Автокорреляционная функция (АКФ). Свойства АКФ.

2.1 Теоретические сведения

Мощность и энергия сигнала

Важными характеристиками дискретных сигналов являются их энергия E_x и мощность P_x . Энергия дискретного сигнала определяется как

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2. \quad (2.1)$$

Данное определение справедливо как для действительных так и для комплексных сигналов $x(n)$. Чтобы данное определение имело смысл, энергия сигнала должна иметь конечное значение. Для этого амплитуда сигнала должна стремиться к нулю при $|n| \rightarrow \infty$ (необходимое условие). В противном случае числовой ряд (2.1) не будет сходиться. Сигналы, которые имеют конечную энергию называются *энергетическими сигналами*.

В некоторых случаях амплитуда сигнала $x(n)$ не стремится к нулю при $|n| \rightarrow \infty$ и энергия сигнала является бесконечной. В таких случаях, более разумным является измерение мощности сигнала P_x , которая определяется как усредненное значение энергии:

$$P_x = \lim_{N \rightarrow \infty} \frac{1}{2N + 1} \sum_{n=-N}^N |x(n)|^2. \quad (2.2)$$

В данном уравнении сумма делится на $2N + 1$ поскольку именно такое число отсчетов $x(n)$ попадают в интервал от $-N$ до N . Для N -периодических сигналов усреднение выполняется только на одном периоде:

$$P_x = \frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2. \quad (2.3)$$

Сигнал для которого P_x имеет конечное ненулевое значение называют *мощностным*.

Произвольный сигнал может быть либо энергетическим, либо мощностным, либо ни тем, ни другим, но не может быть одновременно энергетическим и мощностным.

Пример 2.1 На рисунке 2.1 показан дискретный сигнал $x(n)$, а также два периодических сигнала, полученных из $x(n)$ периодическим расширением:

$$x_1(n) = \sum_{k=-\infty}^{\infty} x(n + 6k) \quad \text{и} \quad x_2(n) = \sum_{k=-\infty}^{\infty} x(n + 7k).$$

Определите энергию и мощность каждого из трех сигналов.

Решение. Поскольку $x(n) = n$ при $0 \leq n \leq 5$ и 0 при всех других n , то его энергия может быть найдена используя (2.1)

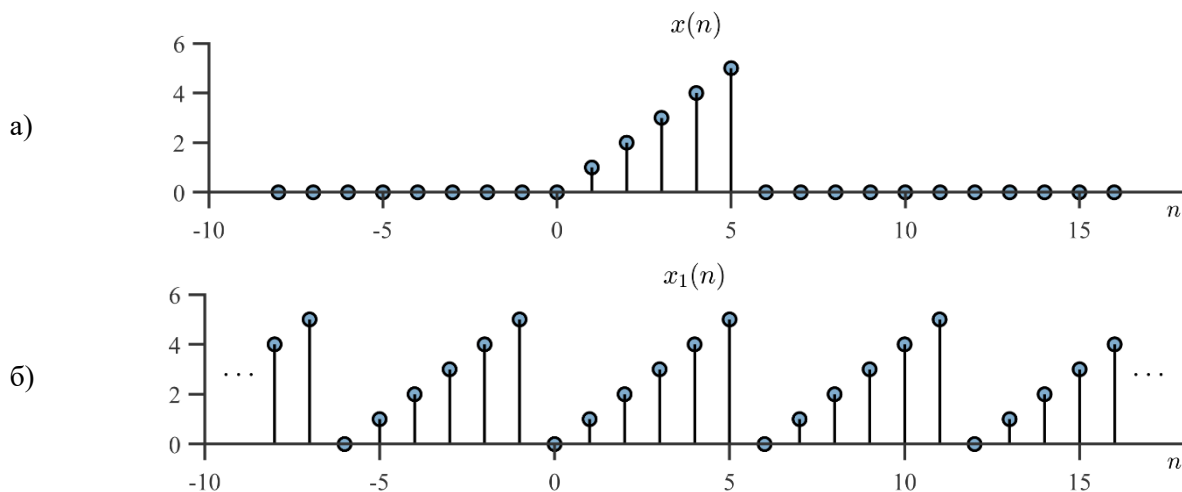
$$E_x = \sum_{n=0}^5 n^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55.$$

Сигнал $x_1(n)$ является периодическим повторением $x(n)$ с периодом 6. Все периодические сигналы имеют бесконечную энергию, т.е. $E_{x_1} = \infty$. Мощность $x_1(n)$ можно найти при помощи выражения (2.3) как

$$P_{x_1} = \frac{1}{6} \sum_{n=0}^5 n^2 = \frac{1}{6} E_x = \frac{55}{6} = 9 \frac{1}{6} \approx 9,1667.$$

Сигнал $x_2(n)$ является периодическим повторением $x(n)$, но только с периодом 7. Его энергия также равна бесконечности. Мощность $x_2(n)$ можно вычислить по формуле (2.3) как

$$P_{x_2} = \frac{1}{7} \sum_{n=0}^5 n^2 = \frac{1}{7} E_x = \frac{55}{7} \approx 7,8571.$$



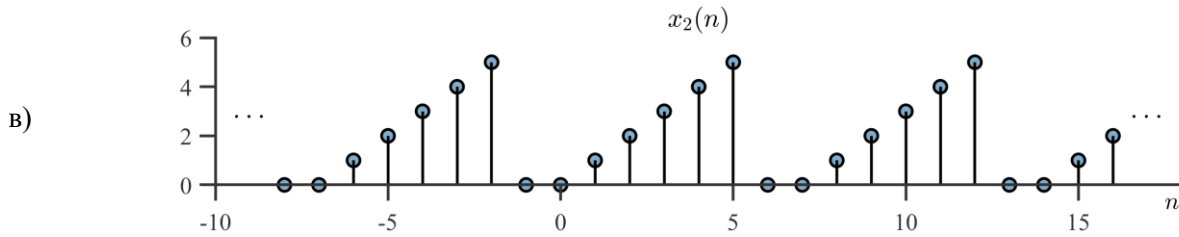


Рисунок 2.1 – Дискретные сигналы из примера 2.1

Два дискретных сигнала **ортогональны**, если их взаимная энергия удовлетворяет условию

$$E_{xy} = \sum_{n=-\infty}^{\infty} x(n)y^*(n) = 0. \quad (2.4)$$

Энергия и мощность ортогональных сигналов аддитивны, т.е.

$$\sum_{n=-\infty}^{\infty} |x(n) + y(n)|^2 = \sum_{n=-\infty}^{\infty} |x(n)|^2 + \sum_{n=-\infty}^{\infty} |y(n)|^2. \quad (2.5)$$

Ковариация и корреляция

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

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

$$\text{Cov}(X, Y) = E\{(X - \mu_X)(Y - \mu_Y)\} = \frac{1}{N} \sum_{n=1}^N (x_n - \mu_X)(y_n - \mu_Y), \quad (2.6)$$

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

На рисунке 2.2 приведен пример расчета ковариации для двух последовательностей.

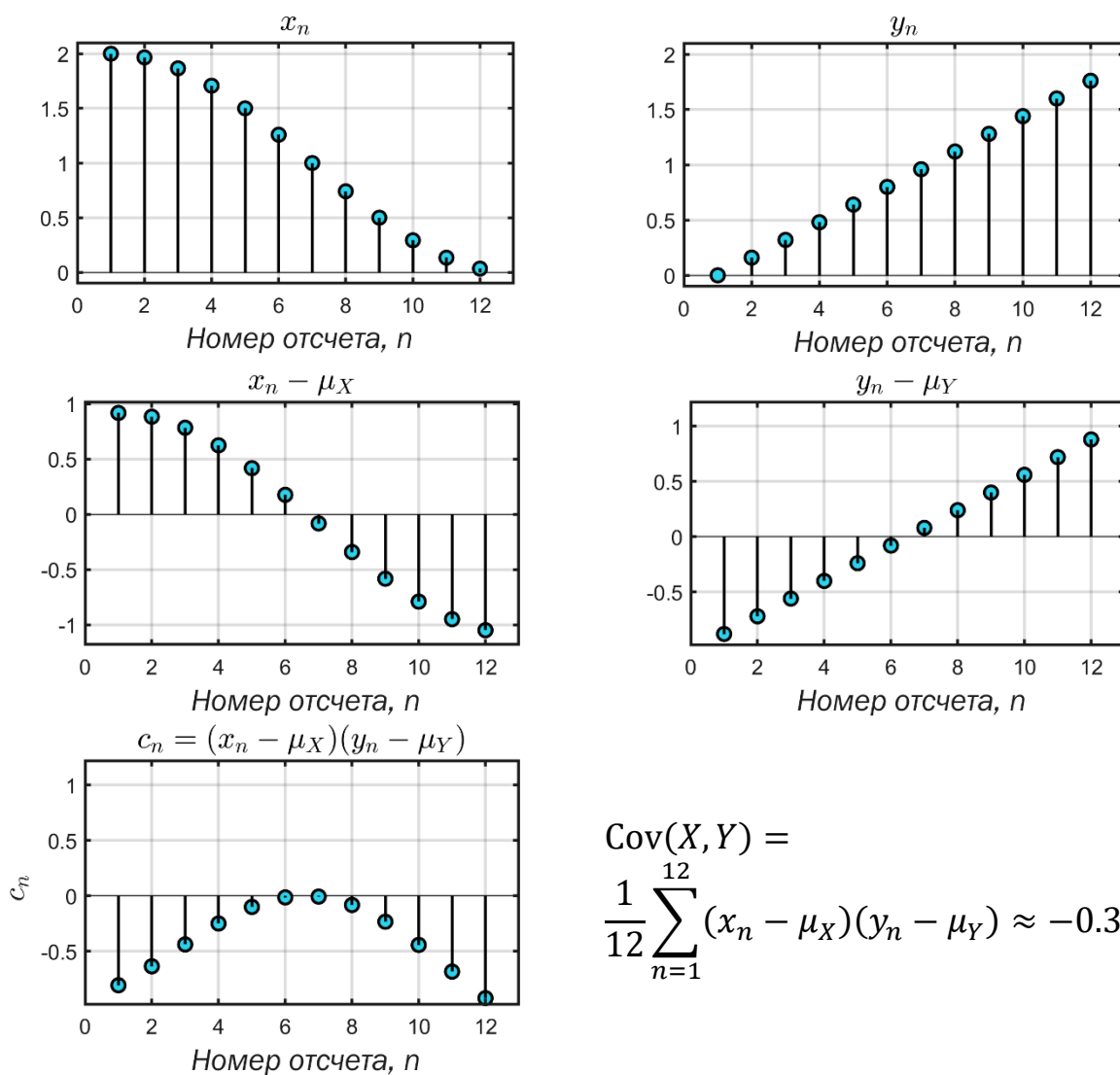


Рисунок 2.2 – Расчет ковариации для двух последовательностей

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

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:

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{n=1}^N x_n y_n. \quad (2.7)$$

Очевидно, что если две последовательности независимы, то сумма произведений $x_n y_n$ стремится 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 $\text{Cov}(X, Y)$, 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) $\sigma_X \sigma_Y$:

$$r_{X,Y} = \frac{E\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y} = \frac{\sum_{n=1}^N (x_n - \mu_X)(y_n - \mu_Y)}{\sqrt{\sum_{n=1}^N (x_n - \mu_X)^2 \sum_{n=1}^N (y_n - \mu_Y)^2}}. \quad (2.8)$$

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

Корреляционная функция

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

$$r_{x,y} = \frac{1}{N} \sum_{n=1}^N x_n y_n \quad (2.9)$$

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:

$$r_{x,y}(\ell) = \sum_{n=-\infty}^{\infty} x(n + \ell) y^*(n) = \sum_{n=-\infty}^{\infty} x(n) y^*(n - \ell) \quad (2.10)$$

Функция корреляции (2.10) не зависит от времени n , но зависит от временной задержки ℓ (англ. *time lag*). В случае, когда в (2.10) 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 принимает корреляционная функция $r_{x,y}(\ell)$.

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

$$r_{xy}(\ell) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)y^*(n-\ell), \quad -(N-1) \leq \ell \leq (N-1). \quad (2.11)$$

Пример 2.2 Рассчитайте корреляционную функцию для сигналов:

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

Построить график корреляционной функции ($N = 10$). Определите энергию и мощность каждого из трех сигналов.

Решение. Воспользуемся для расчета формулой (2.11):

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

...

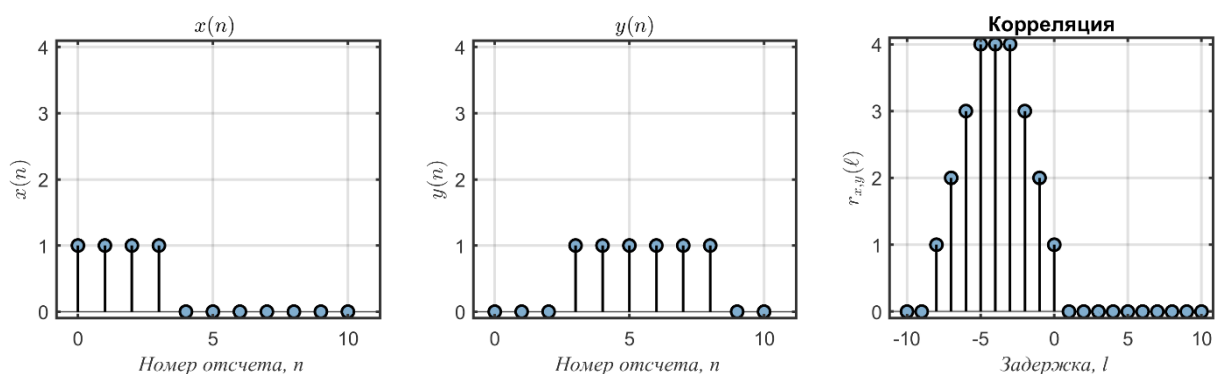


Рисунок 2.3 – Пример вычисления корреляции

Интерпретация примера 2.2: сигнал $y(n)$ запаздывает относительно сигнала $x(n)$. Наибольшие значения корреляции достигается при значениях $\ell = \{-5, -4, -3\}$. Это значит, что если задержать сигнал $x(n)$ на 5, 4 или 3 отсчета, то произойдет его максимальное совмещение (согласование) с сигналом $y(n)$.

Автокорреляционная функция (АКФ)

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

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

Для простоты будем считать, что сигнал $x(n)$ принимает только действительные значения. Заметить, что мощность сигнала сконцентрирована в первом отсчете АКФ:

$$P_x = \frac{1}{N} \sum_{n=0}^{N-1} x^2(n) = r_{xx}(0). \quad (2.12)$$

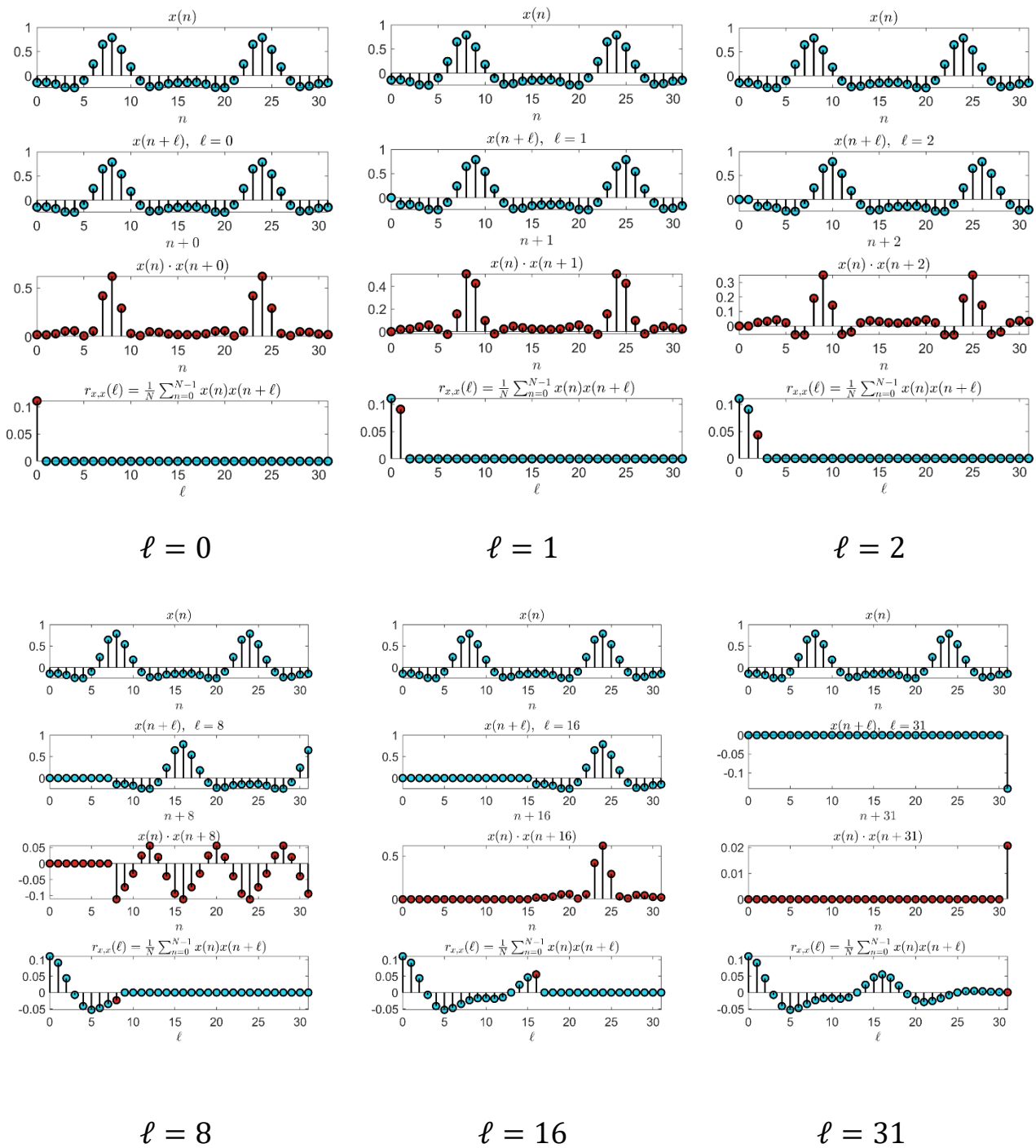


Рисунок 2.4 – Пример вычисления корреляционной функции дискретного сигнала

Говорят, что АКФ является мерой самоподобия (*self-similarity*, самоповторения) сигнала на различном временном удалении ℓ . Когда значение $r_{xx}(\ell)$ велико для какого-то ℓ , то говорят, что отсчеты расположенные на расстоянии ℓ имеют высокую корреляцию.

Рассмотрим процесс вычисления АКФ. На рисунке 2.4 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 все 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:

- уменьшается число ненулевых отсчетов $x(n)x(n+\ell)$;
- уменьшается степень схожести сигнала с самим собой при увеличении ℓ .

Применение АКФ

В примере на рисунке 2.4, при вычислении АКФ мы предполагали, что за пределами интервала $0 \leq n < N$ сигналы $x(n)$ принимают нулевые значения. Однако, для последующих вычислений предположим, что сигнал $x(n)$ имеет за пределами своего определения *периодическое продолжение*. Это предположение весьма обоснованно, особенно для стационарных сигналов.

Повторимся, что АКФ $r_{xx}(\ell)$ показывает нам насколько сильна связь между отсчетами сигнала. Например $r_{xx}(1)$ показывает как сильно зависимы соседние отсчеты в сигнале, $r_{xx}(2)$ обнаруживает связь между отсчетами $x(n)$ и $x(n \pm 2)$. Таким образом, *с помощью автокорреляционной функции можно выявить скрытую в сигнале периодичность*. Допустим, что сигнал содержит гармоническое колебание с периодом в 20 отсчетов. Можно смело предположить, что его АКФ $r_{xx}(\ell)$ будет иметь максимум в точке $\ell = 20$.

Отметим, что $r_{xx}(\ell)$, не содержат информации о фазе сигнала $x(n)$. По этой причине сдвиг $x(n)$ во временной области не влияет на АКФ.

В качестве характерных примеров рассмотрим АКФ гармонического сигнала и белого шума (рисунок 2.5).

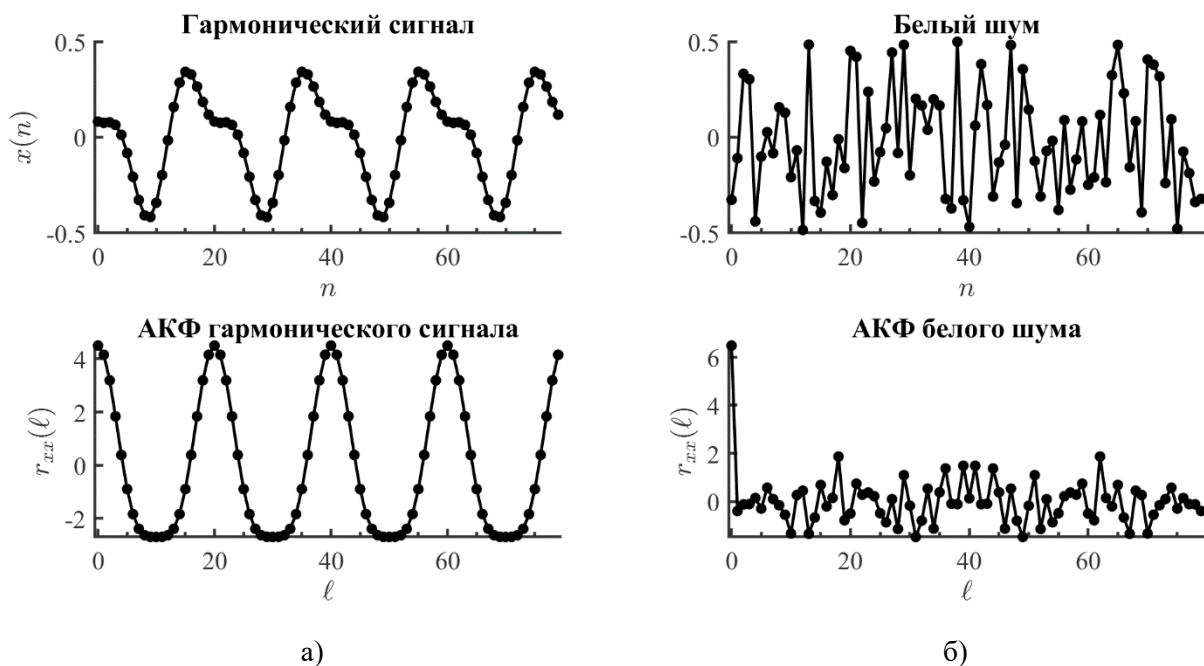


Рисунок 2.5 – АКФ: а) временная форма и автокорреляционная функция гармонического сигнала $x(n) = 0.3 \cos\left(2\pi \cdot 0,05n + \frac{\pi}{4}\right) + 0.15 \sin\left(2\pi \cdot 0,10n - \frac{\pi}{3}\right), n = 0,1, \dots, 79;$ б) временная форма и автокорреляционная функция белого шума

Частота дискретизации сигнала равна 1 Гц. Гармонический сигнал содержит две синусоидальные компоненты: одна с периодом $T_1 = 20$ отсчетов (частота $f_1 = 1/T_1 = 0,05$), другая с периодом $T_2 = 10$ отсчетов (частота $f_2 = \frac{1}{T_2} = 0,10$). Таким образом, в сигнале имеется частота основного тона f_1 и её 2-я гармоника, поскольку $f_2 = 2f_1$. С помощью АКФ можно обнаружить период частоты основного тона, как правило он соответствует первому максимуму $r_{xx}(\ell)$ для $m > 0$. В примере мы наблюдаем максимум $r_{xx}(\ell)$ при $\ell = 20$, что в точности соответствует периоду основного колебания. Белый шум (рисунок 2.5, б) представляет собой случайный сигнал. Известно, что каждый последующий отсчет белого шума не зависит от предыдущего. Именно этот факт отражает его АКФ, в которой имеется пик только при $\ell = 0$. При любом другом сдвиге ℓ АКФ близка к нулю, что говорит о статистической независимости отсчетов белого шума.

АКФ используют для определения частоты основного тона речевого сигнала F_0 . Для определения F_0 от фрейма речевого сигнала вычисляют АКФ (рисунок 2.6). Положение первого пика АКФ для $\tau > 0$ определяет задержку соответствующую периоду основного тона T_0 . Далее частота основного тона определялась как $F_0 = 1/T_0$.

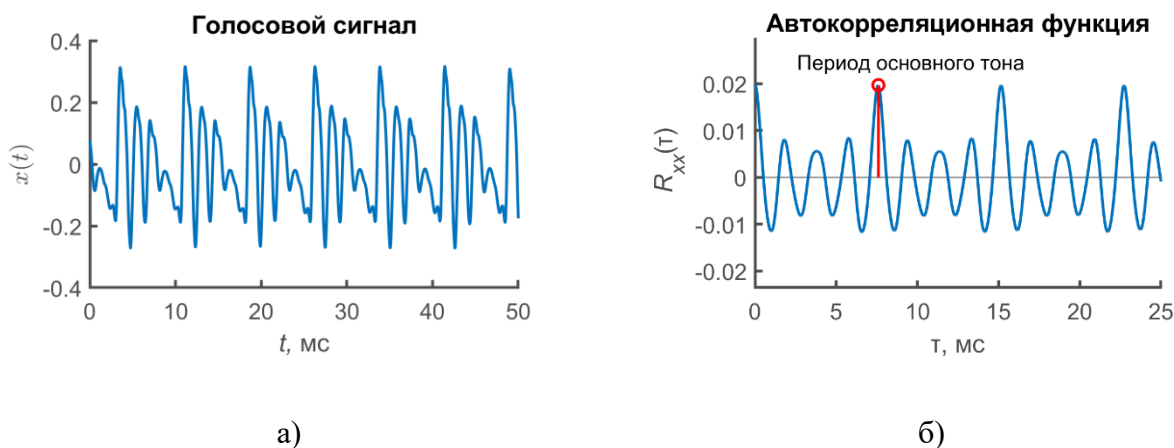


Рисунок 2.6 – а) пример вокализованного фрейма речевого сигнала (звук /а/);
 б) автокорреляционная функция фрейма речевого сигнала

Кратковременный анализ речевых сигналов

На рисунке 2.7 показан типичный речевой сигнал. Из графика видно, что свойства речевого сигнала изменяются во времени. Например, характер возбуждения на вокализованных и невокализованных участках, пиковая амплитуда и др. Вокализованными называются те сегменты речи, где задействованы голосовые связки диктора. Такие звуки как /а/, /о/, /е/ являются вокализованными, в то время как звуки /с/, /ш/, /щ/ являются невокализованными.

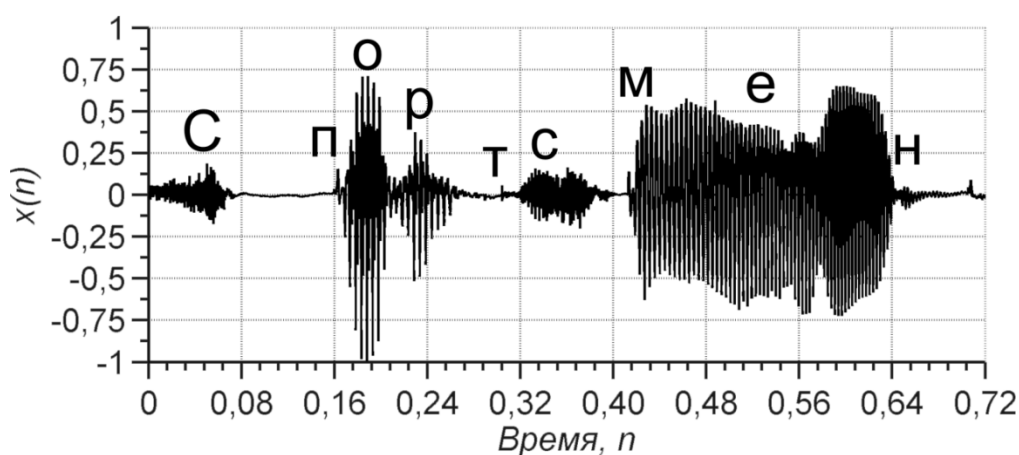


Рисунок 2.7 – Речевой сигнал (слово "Спортсмен")

В основе большинства методов обработки речи лежит предположение о том, что свойства речевого сигнала с течением времени медленно изменяются. Это предположение приводит к методам *кратковременного анализа*, в которых сегменты речевого сигнала выделяются и обрабатываются так, как если бы они были короткими участками отдельных звуков с отличающимися свойствами. Сегменты, которые иногда называют *кадрами анализа* или **фреймами**, обычно

пересекаются. Результатом обработки на каждом фрейме является число или совокупность чисел. Следовательно, подобная обработка приводит к новой временной последовательности, которая может служить характеристикой речевого сигнала.

Простым примером изложенной концепции может служить измерение кратковременной мощности сигнала, которая определяется как

$$P(n) = \frac{1}{N + 1} \sum_{m=n-N/2}^{n+N/2} x^2(m) \quad (2.13)$$

Выражение (2.13) показывает, что кратковременная мощность сигнала $x(n)$ в момент времени n есть просто средняя сумма квадратов $N + 1$ отсчета в диапазоне от $n - N/2$ до $n + N/2$. Обработку (2.13) можно представить себе, как операцию умножения всего сигнала $x^2(n)$ на оконную функцию $w(n - m)$, где

$$w(n) = \begin{cases} 1, & -N/2 \leq n \leq N/2 \\ 0, & \text{в противном случае.} \end{cases} \quad (2.14)$$

Далее отсчеты сигнала $x^2(n)w(n - m)$ суммируются и делятся на $(N + 1)$. Эта идея обработки со *скользящим окном* иллюстрируется на рисунке 2.8.

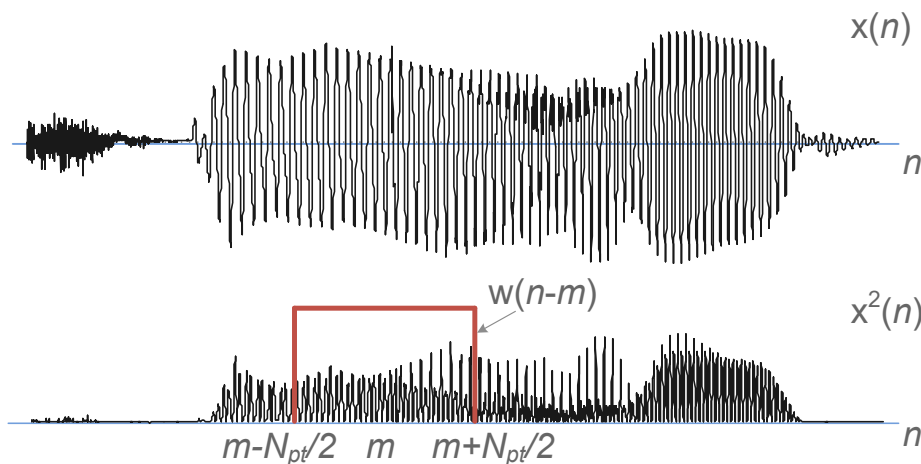


Рисунок 2.8 – Скользящее окно

Отметим, что окно «скользит» вдоль последовательности квадратов значений сигнала, ограничивая длительность интервала, используемого в вычислениях.

Функцию $P(n)$ можно использовать, чтобы отличить вокализованные речевые сегменты от невокализованных. Если мощность сигнала в момент времени n больше порога, то считается, что отсчет $x(n)$ относится к вокализованному

сегменту, иначе к невокализованному. Таким образом, для каждого отсчета сигнала $x(n)$ можно рассчитать признак вокализованности

$$v(n) = \begin{cases} 1, & P(n) > P_{thr}, \\ 0, & P(n) \leq P_{thr}, \end{cases} \quad (2.15)$$

где P_{thr} – порог, определяемый экспериментально.

2.2 Порядок выполнения работы

1. Разработайте в MATLAB функцию для вычисления мощности сигнала P_x по формуле (2.3). Вторым выходным аргументом функция должна выдавать значение мощности в децибелах $P_{dB} = 10 \log_{10} P_a$. Посчитайте мощность следующих сигналов ($n = 0, 1, \dots, N - 1$):

Таблица 1

№ варианта	Сигналы
1	$x_1(n) = 0.5 \cdot \sin(2\pi \frac{2}{N}n), x_2(n) = 0.5 \cdot \sin(2\pi \frac{20}{N}n), N = 64.$
2	$x_1(n) = 0.5 \cdot \sin(2\pi \frac{2}{N}n), x_2(n) = \cos(2\pi 0.1n), N = 50.$
3	$x_1(n) = \cos(2\pi 0.1n), x_2(n) = e^{0.9n}, N = 62.$
4	$x_1(n) = 0.9^n \sin(2\pi \frac{8}{N}n), x_2(n) = (1.05)^n \cos(2\pi \frac{8}{N}n), N = 40.$
5	$x_1(n) = 0.85^n, x_2(n) = (-1.01)^n, N = 48.$
6	$x_1(n) = \sin(\sum_{k=0}^n 0.02\pi k), x_2(n) = \cos(2\pi \frac{3}{N}n + \frac{\pi}{3}), N = 42.$
7	$x_1(n) = \sin(2\pi \frac{2}{N}n), x_2(n) = (0.95)^n \cos(\sum_{k=0}^n 0.02\pi k), N = 56.$
8	$x_1(n) = \sin(2\pi \frac{1}{N}n), x_2(n) = \cos(2\pi \frac{3}{N}n), N = 52.$

Проверьте, являются ли сигналы $x_1(n)$ и $x_2(n)$ ортогональными.

Вычислите коэффициент корреляции сигналов $x_1(n)$ и $x_2(n)$.

2. Постройте АКФ сигнала согласно варианту ($n = 0, 1, \dots, N - 1$).

Таблица 1

№ варианта	Сигналы
1	$x(n) = 0.5 \cdot \sin(2\pi \frac{2}{N}n), N = 40$

2	$x(n) = \cos(2\pi 0.1n), N = 42$
3	$x(n) = e^{0.9n} \cos(2\pi 0.07n), N = 36.$
4	$x(n) = 0.9^n \sin\left(2\pi \frac{8}{N}n + \frac{\pi}{4}\right), N = 44.$
5	$x_1(n) = 0.85^n \cos\left(\frac{\sqrt{2}}{10\pi^2}n\right) N = 50.$
6	$x_1(n) = \sin\left(\sum_{k=0}^n 0.015\pi n\right), N = 64.$
7	$x(n) = (0.95)^n \cos\left(\sum_{k=0}^n 0.02\pi n\right), N = 49.$
8	$x_1(n) = \sin\left(2\pi \frac{10}{N}n\right) \cos\left(2\pi \frac{13}{N}n\right), N = 72.$

В нечетных вариантах считать, что за пределами определения сигнал имеет нулевое расширение, а в четных – периодическое расширение.

Определите период сигнала по АКФ.

3. Разработайте в MATLAB функцию вычисления АКФ сигнала (тип расширения сигнала должен определяться путем передачи в функцию дополнительного параметра, по умолчанию считается, что сигнал имеет нулевое расширение). Примените данную функцию для определения периода основного тона речевого сигнала. При помощи смартфона выполните запись собственного голоса (протяжного произнесения звука /а/, не менее 2 секунд). Сохраните файл в формате wav с частотой дискретизации не ниже $f_s = 16$ кГц. Wav-файл можно загрузить при помощи команды

```
[x, fs]=audioread('<путь_к_файлу>');
```

Входной сигнал разбивается на перекрывающиеся секции по 320 отсчетов. Перекрытие секций составляет 160 отсчетов. Для каждой секции сигнала вычисляется АКФ и находится номер отсчета, содержащий первый максимум функции. Номер этого отсчета и является *оценкой* периода основного тона. Полученные значения отобразите на графике: по оси абсцисс – номер сегмента, по оси ординат – значение периода основного тона.

4. Разработайте в MATLAB функцию для отделения вокализованных участков речи от невокализованных:

```
[x_v, x_u] = vu_separate(x, N, threshold)
```

Сепарация (разделение) происходит на основе анализа кратковременной мощности сигнала. На вход функции поступает речевой сигнал x , длина окна ана-

лиза на котором считается мощность – n и значение порога – `threshold`. Если мощность сигнала в момент времени n больше порога, то считается, что отсчет $x(n)$ является вокализованным, иначе невокализованным. Таким образом сепарация выполняется следующим образом:

а) Для входного сигнала $x(n)$ рассчитывается кратковременная мощность $P(n)$ по выражению.

б) Рассчитывается признак вокализованности:

$$v(n) = \begin{cases} 1, & P(n) > \text{threshold}, \\ 0, & P(n) \leq \text{threshold}. \end{cases}$$

с) Сепарация сигнала:

$$\begin{aligned} x_v(n) &= x(n) \cdot v(n), \\ x_u(n) &= x(n) \cdot (1 - v(n)). \end{aligned}$$

Удобно мощность сигнала посчитать в децибелах

$$P_{dB}(n) = 10 \log_{10} P(n)$$

и порог для сепарации также подбирать в децибелах.

5. (Задание повышенной сложности, не обязательное) Определение задержки случайного сигнала в условиях шума.

а) **Сгенерируйте** 100 точек гауссовского процесса AR(1) с дисперсией 10; $a(1) = \rho = 0,5$. Обозначим полученный сигнал как $x(n)$.

https://fanwangecon.github.io/M4Econ/panel/timeseries/htmlpdfm/fs_autoregressive.html

б) Сформируйте три других сигнала следующим образом:

- задержка $x(n)$ на 10 единиц времени и ослабление его в 2 раза,
- добавление однородного белого шума к процессу с дисперсией 1, 2 и 5.

То есть

$$y_i(n) = 0,5x(n - 10) + \eta_i(n), \quad i = 1,2,3$$

в) Оцените функции взаимной корреляции между $x(n)$ и каждым из $y(n)$.

Чем они отличаются и почему?

2.3 Дополнительные задания/вопросы

1. Докажите выражение (2.5).
2. Что такое фрейм?
3. Что такое вокализованный и невокализованный речевой сигнал?
4. Модифицируйте сигнал $x_2(n)$ из задания 1 таким образом, чтобы его мощность была такой же как у сигнала $x_1(n)$.