ЛАБОРАТОРНА РОБОТА № 6

КОРЕЛЯЦІЯ

Завдання: Реалізувати програмно розрахунок взаємної кореляції та автокореляції двох послідовностей даних, які записані в файлах **`data1.dat`** та **`data2.dat`**. Результат записати в файл **`correlation.dat`**.

Вступ

Зменшення шумової складової сигналу досить часто виконується за допомогою кореляції і техніки швидкої кореляції з використанням ШПФ.

Інколи необхідно визначити ступінь незалежності одного процесу від іншого, або встановити *схожість одного набору даних з іншим*. Іншими словами, шуканою є ***кореляція*** процесів, або даних, яку можна визначити математично і виміряти. Взагалі, процес кореляції займає значне місце в обробці сигналів. Цей математичний апарат знайшов застосування в обробці зображень у сфері комп'ютерного зору та дистанційного зондування із супутників, в яких порівнюються дані з різних зображень, в радарних й гідроакустичних установках для вимірювання відстані й визначення місцезнаходження (пеленгації), в яких порівнюються передані і відбиті сигнали. Він використовується в детектування та ідентифікації сигналів в шумі, в організації технічного контролю для спостереження за впливом входу на вихід, в ідентифікації двійкових кодових слів у системі з імпульсно-кодовою модуляцією, в звичайних схемах оцінки за методом найменших квадратів та в багатьох інших областях, наприклад, в метеорології. Кореляція також є невід'ємною частиною процесу згортки, яка, по суті є тією ж кореляцією двох послідовностей даних, при обчисленні якої одна з послідовностей обернена в часі. Це означає, що для обчислення кореляції і згортки можуть використовуватися одні й ті ж алгоритми, тільки у випадку згортки одна з послідовностей перевертається.

Опис кореляції

Розглянемо, як можна порівняти дві послідовності даних, що складаються зі значень, одночасно триманих відповідно з двох сигналів. Якщо два сигнали подібно змінюються при переході від точки до точки, то міру їх кореляції можна обчислити за сумою добутків відповідних пар точок. Дана маніпуляція стає аргументованішою, якщо розглянути дві незалежні і випадкові послідовності даних. У цьому випадку сума добутків наближається до малих значень зі збільшенням пар точок. Це пояснюється тим, що всі числа (додатні і від'ємні) є різноймовірними, так що пари добутків компенсуються при додаванні. У той же час, якщо сума є скінченною, то це вказує на наявність кореляції. Від'ємна сума вказує на від'ємну кореляцію, тобто збільшення однієї змінної пов'язано зі зменшенням іншої. Таким чином, взаємну кореляцію $r_{12}(n)$ двох послідовностей даних $x_1(n)$ і $x_2(n)$, що містять по N елементів, можна записати як

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

Таке визначення взаємної кореляції дає результат, який залежить від числа розглянутих точок. Щоб це виправити, результат нормується на число точок (ділиться на *N*). Дану операцію можна також розглядати як усереднення суми добутків. Отже, отримуємо наступний покращений вираз:

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

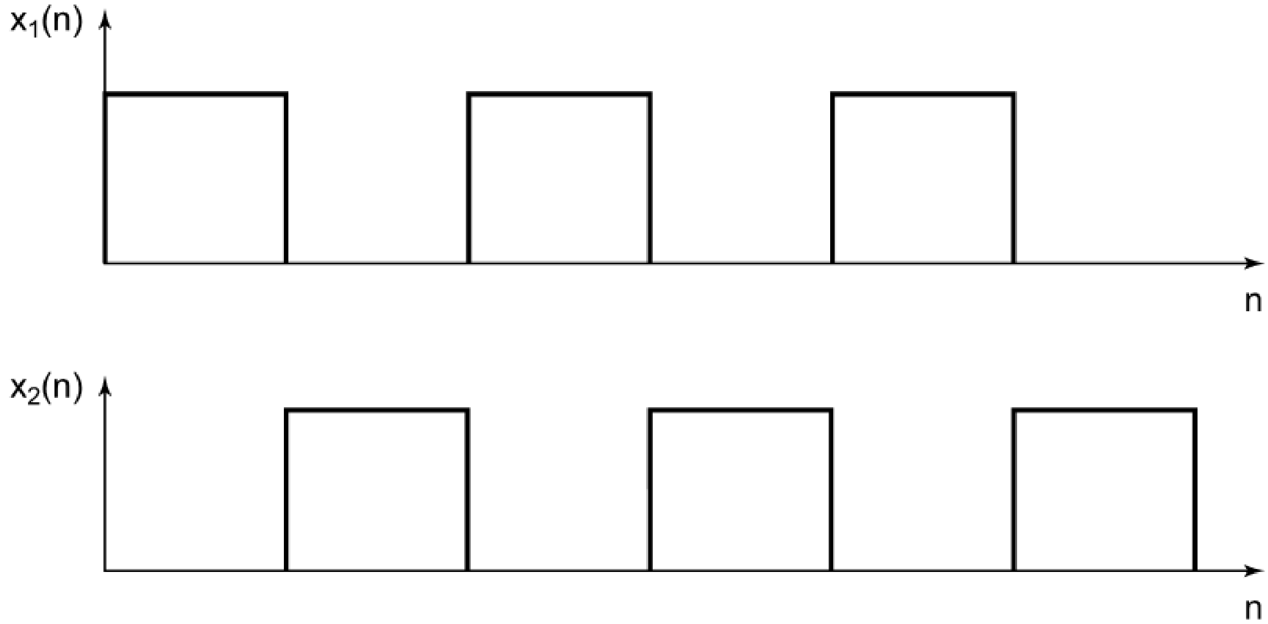

Рис. 1. Сигнали з 100%-ою кореляцією, що йдуть не в фазі, – при нульовій затримці кореляція дорівнює нулю

Приклад 1.

Нижче ілюструється розрахунок $r_{12}(n)$. Кількість точок даних в послідовності дорівнює *n*, використовуються послідовності x_1 і x_2.

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

$$r_{12} = \frac{1}{9}\Big(4\times(-4) + 2\times 1 + (-1)\times 3 + 3\times 7 + (-2)\times 4 + \\ +(-6)\times(-2) + (-5)\times(-8) + 4\times(-2) + 5\times 1\Big) = 5$$

Щоб дане визначення використовувати, його потрібно трохи модифікувати. У деяких випадках кореляція, яка визначена зазначеним вище способом, може бути нульовою, хоча дві послідовності корелюють на 100%. Це може статися, наприклад, коли два сигнали йдуть не в фазі (як часто і буває). Дана ситуація ілюструються сигналами на рис. 1. На рисунку показано, що кожна пара добутків у функції кореляції дорівнює нулю, отже, вся кореляція дорівнює нулю, оскільки нулю завжди дорівнює одне із значень x_1 або x_2. Втім, очевидно, що сигнали сильно корелюють, хоча і йдуть не в фазі. Різниця фаз може, наприклад, пояснюватися тим, що x_1 – якийсь еталонний сигнал, а x_2 – затриманий вихідний сигнал схеми. Щоб подолати подібний зсув фаз, необхідно змістити (або затримати) один із сигналів відносно іншого. Зазвичай, щоб вирівняти сигнали перед визначенням кореляції, x_2 зміщують вліво. Як показано на рис. 2, це еквівалентно заміні $x_2(n)$ на $x_2(n+j)$, де j представляє величину затримки – число точок вибірки, на яке x_2 зміщується вліво. Альтернативною та еквівалентною процедурою є зміщення x_1 вправо. В результаті отримаємо таку формулу для взаємної кореляції:

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

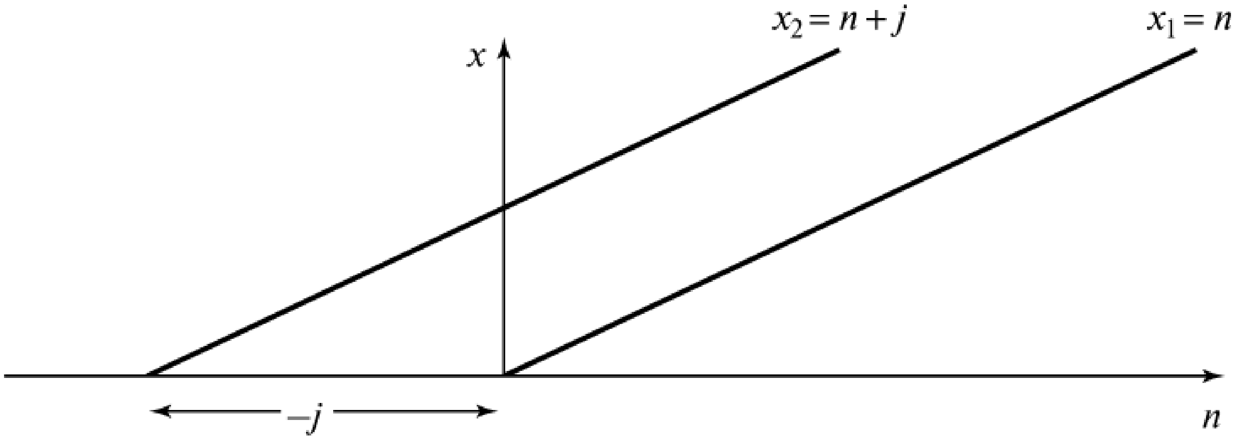

Рис. 2. Сигнал $x_2 = x_2(n+j)$, зміщений на j проміжків часу вліво від сигналу x_1

На практиці, за умови що два сигнали корелюють, їх фазовий зв'язок швидше за все невідомий, тому кореляцію потрібно знаходити для декількох різних затримок, щоб встановити найбільше значення кореляції, яке потім вважається істинним.

Приклад 2.
Розглянемо взаємну кореляцію розглянутих вище двох послідовностей $x_1(n)$ і $x_2(n)$ із затримкою $j = 3$, тобто $r_{12}(3)$. Отже, використовуються наступні послідовності:

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			

так що

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

Кореляцію також можна розглянути в неперервній часовій області, і деякі аналогові схеми кореляції організовані саме так. У неперервній області $n\to t$ та $j\to\tau$

$$r_{12}(\tau)=\lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}x_1(t)\,x_2(t+\tau)\,dt\,. \tag{2}$$

Якщо $x_1(t)$ і $x_2(t)$ – періодичні залежності з періодом *T*, то формула (2) спрощується до

$$r_{12}(\tau)=\frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}}x_1(t)\,x_2(t+\tau)\,dt\,. \tag{3}$$

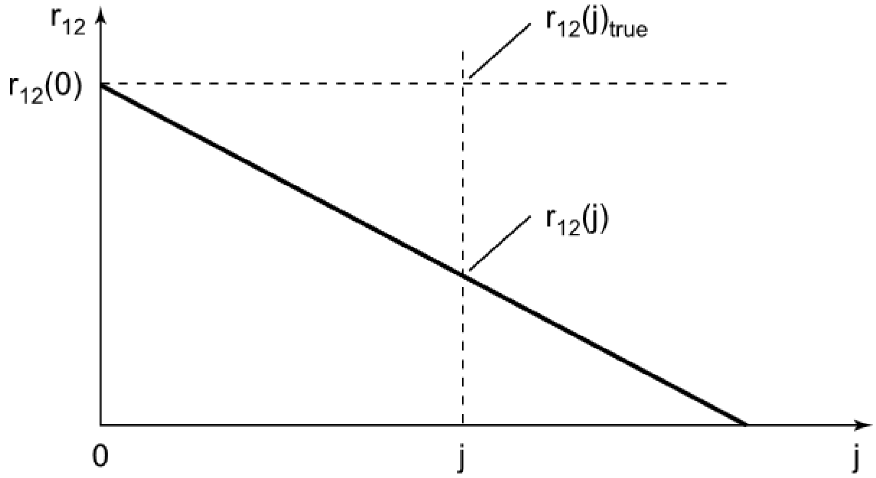

Рис. 3. Вплив крайового ефекту на взаємну кореляцію $r_{12}(j)$

Якщо дані – це сигнали зі скінченною енергією, наприклад, неперіодичні імпульсоподібні сигнали, то усереднення за часом *T* при $T\to\infty$ не виконується, оскільки в цьому випадку $\frac{1}{T}\to 0$ і $r_{12}(\tau)$ було б малою величиною. У такому випадку використовується така формула:

$$r_{12}(\tau)=\int_{-\infty}^{\infty}x_1(t)\,x_2(t+\tau)\,dt\,. \tag{4}$$

На практиці зазвичай обробляють послідовності скінченної довжини, так що застосовують формули (1) або (5):

$$r_{12}(\tau)=\frac{1}{T}\int_{0}^{T}x_1(t)\,x_2(t+\tau)\,dt\,. \tag{5}$$

Є й інша складність, пов'язана з обчисленням взаємної кореляції послідовностей даних скінченної довжини. Це ілюструє *приклад 2*, в якому було знайдено значення $r_{12}(3)=2{,}667$. При зміщенні x_2 вліво сигнали вже не перекриваються, і дані в кінці послідовностей не формують парні добутки – виникає так званий ***крайовий ефект***. У розглянутому прикладі число пар при затримці 3 зменшилася з дев'яти до шести. В результаті спостерігалося лінійне зменшення $r_{12}(j)$ при збільшенні j і отримання сумнівних значень $r_{12}(j)$. Одне з можливих рішень такої проблеми полягає в тому, щоб довжину однієї послідовності зробити вдвічі більшою довжини, необхідної для знаходження кореляції. Для цього можна записати більше даних або, якщо одна з послідовностей періодична, повторити послідовність (особливу увагу слід звернути на узгодження країв). Інше можливе рішення – скорегувати всі розраховані значення. На рис. 3 показано зменшення r_{12} зі зростанням j в результаті крайового ефекту, тобто реальна зміна $r_{12}(j)$ не показана. При $j=0$, $r_{12}(j)=r_{12}(0)$ (це можна порахувати). При $j=N$, $r_{12}(N)=0$, оскільки сигнали вже не перекриваються. У проміжних випадках при деяких значеннях затримки j справжнє значення $r_{12}(j)$ дорівнює $r_{12}(j)_{\text{true}}$, тоді як дійсне значення, спотворене крайовим ефектом, дорівнює $r_{12}(j)$. Далі на основі рисунка отримуємо

$$\frac{r_{12}(j)_{\text{true}}-r_{12}(j)}{j}=\frac{r_{12}(0)}{N},$$

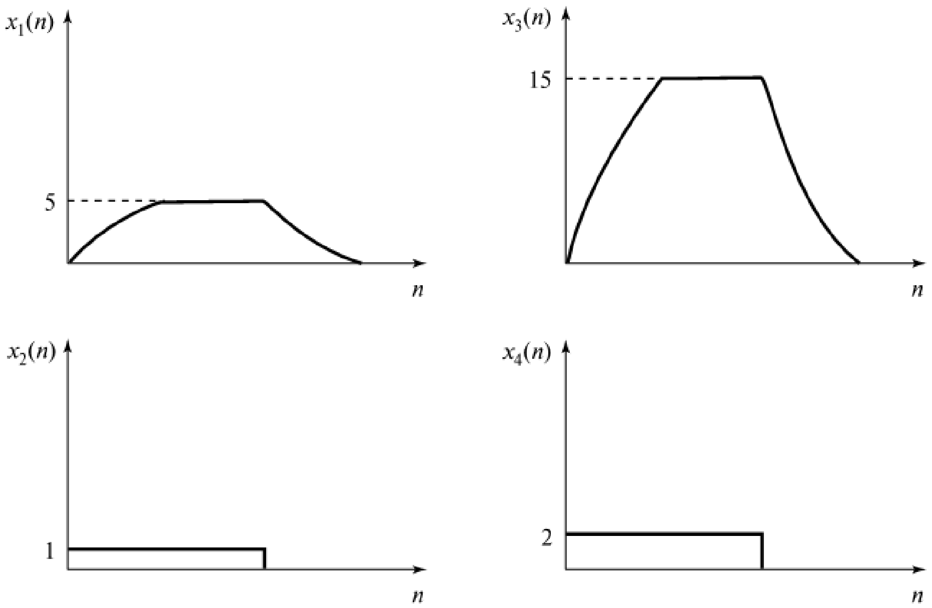

Рис. 4. Пари сигналів що $\{x_1(n), x_2(n)\}$ і $\{x_3(n), x_4(n)\}$ різних амплітуд, але з рівними взаємними кореляціями

звідки

$$r_{12}(j)_{\text{true}} = r_{12}(j) + \frac{j}{N} r_{12}(0). \tag{6}$$

Таким чином, обчислені значення взаємної кореляції легко скорегувати для врахування крайових ефектів, додавши до $r_{12}(j)$ величину $\dfrac{j r_{12}(0)}{N}$.

Значення функції взаємної кореляції обчислюються згідно з наведеними вище формулами в залежності від абсолютних значень даних. Часто буває необхідно виміряти взаємну кореляцію у фіксованому масштабі між –1 і +1. Щоб визначити значення у зазначеному діапазоні, отримані величини нормують на величину, що залежить від енергетичного вмісту даних. Наприклад, розглянемо дві пари сигналів $x_1(n), x_2(n)$ і $x_3(n), x_4(n)$. Значення даних наведені в таблиці нижче.

n	0	1	2	3	4	5	6	7	8
$x_1(n)$	0	3	5	5	5	2	0,5	0,25	0
$x_2(n)$	1	1	1	1	1	0	0	0	0
$x_3(n)$	0	9	15	15	15	6	1,5	0,75	0
$x_4(n)$	2	2	2	2	2	0	0	0	0

Як показано на рис. 4, сигнали $x_1(n)$ і $x_3(n)$ подібні і відрізняються тільки амплітудою. Те ж справедливо для пари $x_2(n)$ і $x_4(n)$. Таким чином, кореляція між $x_1(n)$ і $x_2(n)$ дорівнює кореляції між $x_3(n)$ і $x_4(n)$. У той же час, параметри кореляції r_1 і $r_{34}(1)$ дорівнюють відповідно 1,47 і 8,83. Вони відрізняються, оскільки залежать від абсолютних значень даних. Щоб виправити цю ситуацію, нормуємо взаємну кореляцію $r_{12}(j)$ на коефіцієнт

$$\sqrt{\frac{1}{N}\sum_{n=0}^{N-1} x_1^2(n) \times \frac{1}{N}\sum_{n=0}^{N-1} x_2^2(n)} = \frac{1}{N}\sqrt{\sum_{n=0}^{N-1} x_1^2(n) \times \sum_{n=0}^{N-1} x_2^2(n)}. \tag{7}$$

і подібним чином нормуємо $r_{34}(j)$. В результаті нормований вираз для $r_{12}(j)$ набуває такого вигляду:

$$\boxed{\rho_{12}(j) = \frac{r_{12}(j)}{\dfrac{1}{N}\sqrt{\sum_{n=0}^{N-1} x_1^2(n) \times \sum_{n=0}^{N-1} x_2^2(n)}}}. \tag{8}$$

Величина $\rho_{12}(j)$ відома як ***коефіцієнт взаємної кореляції***. Значення цього коефіцієнта завжди лежить між –1 і +1, причому "+1" означає 100%-ву кореляцію в прямому сенсі, "–1" – 100% -ву кореляцію в протилежному сенсі, наприклад, для сигналів в протифазі. Значення "0" вказує на нульову кореляцію. Це означає, що сигнали абсолютно незалежні, наприклад, якщо один із сигналів абсолютно випадковий. Малі значення $\rho_{12}(j)$ вказують на незначну кореляцію. Нормувальний коефіцієнт для $r_{12}(j)$ з наведеного вище прикладу дорівнює

$$\frac{1}{N}\sqrt{\sum_{n=0}^{N-1} x_1^2(n) \times \sum_{n=0}^{N-1} x_2^2(n)} = \frac{1}{9}\sqrt{88{,}31 \times 6} = 2{,}56;$$

а для $r_{34}(j)$

$$\frac{1}{N}\sqrt{\sum_{n=0}^{N-1} x_1^2(n) \times \sum_{n=0}^{N-1} x_2^2(n)} = \frac{1}{9}\sqrt{794{,}8 \times 24} = 15{,}35.$$

Отже,

$$\rho_{12}(1) = \frac{r_{12}(1)}{2{,}56} = \frac{1{,}47}{2{,}56} = 0{,}57$$

і

$$\rho_{34}(1) = \frac{r_{34}(1)}{15{,}34} = \frac{8{,}83}{15{,}35} = 0{,}58.$$

Тепер $\rho_{12}(1) = \rho_{34}(1)$, звідки видно, що даний процес нормування дійсно дозволяє незалежно порівнювати взаємні кореляції абсолютних значень даних.

Розглянемо окремий випадок $x_1(n) = x_2(n)$, тобто знайдемо *кореляцію сигналу з самим собою*. Даний процес називається ***автокореляцією***. Автокореляційна функція сигналу визначається як

$$\boxed{r_{11}(j) = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)\, x_1(n+j)}.$$

Автокореляційна функція має одну дуже корисну властивість:

$$r_{11}(0) = \frac{1}{N}\sum_{n=0}^{N-1} x_1^2(n) = S,$$

де S – нормована енергія сигналу. В результаті отримуємо метод розрахунку енергії сигналу. Якщо сигнал абсолютно випадковий, наприклад, сигнал, що має характер білого ґаусівського шуму в електричній системі, то його автокореляція буде максимальною при нульовій затримці і зменшуватися до випадкових флуктуацій малої амплітуди біля нуля для затримок, що перевищують одиницю (див. рис. 5). Крім того, справедливе наступне співвідношення:

$$r_{11}(0) \ge r_{11}(j).$$

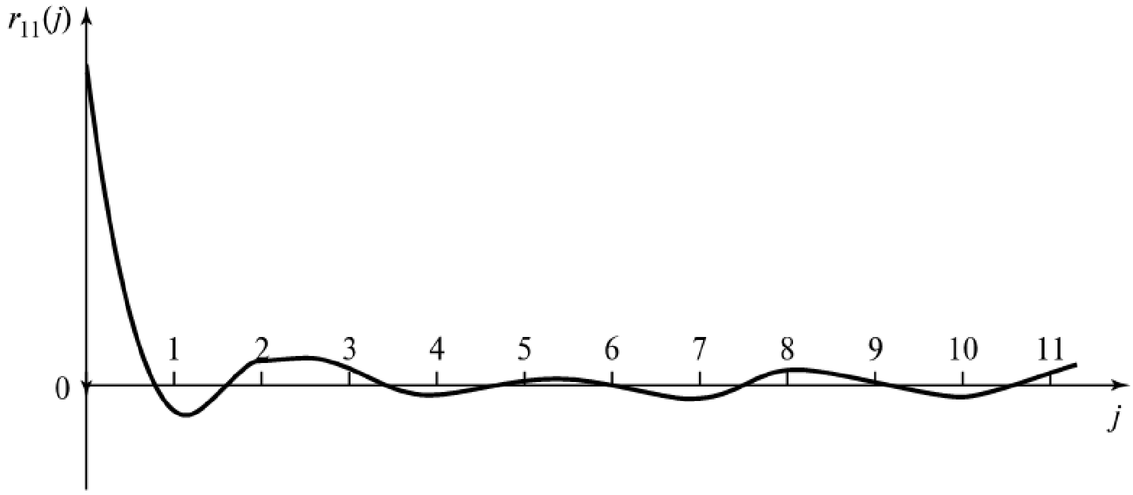

Рис. 5. Автокореляційна функція випадкового сигналу

Взаємна кореляція і автокореляція

Визначення взаємної кореляції двох періодичних послідовностей різної довжини вимагає акуратності. Це пояснюється тим, що результат кореляції буде повторюватися з періодом коротшої послідовності. Цей результат не відображає повної періодичності довшої послідовності, отже, є невірним. Продемонструємо це факт на прикладі: знайдемо взаємну кореляцію $r_{ab}\left(j\right)$ послідовностей $a=\left\{4,3,1,6\right\}$ і $b=\left\{5,2,3\right\}$. Послідовність b записується під a й поетапно зміщується на одну позицію вліво, в останньому стовпці записуються відповідні значення взаємної кореляції.

Послідовність				**Затримка**	$\boldsymbol{r_{ab}(j)}$	
4	3	1	6			
3	5	2	3	0	47	
5	2	3	5	1	59	
2	3	5	2	2	34	
3	5	2	3	3	47	$r_{ab}(j)$ повторюється
5	2	3	5	4	59	

і т.д.

Результат показує, що $r_{ab}(j)$ циклічна з періодом в три затримки, тобто період $r_{ab}(j)$ дорівнює періоду коротшої послідовності b. Описана процедура називається ***циклічною кореляцією***. Щоб отримати правильне значення, в якому кожне значення a множиться на кожне значення b, всі елементи b потрібно послідовно змістити під кожним значенням a:

4	3	1	6		
			5	2	3
		5	2	3	
	5	2	3		
5	2	3			

		5	2	3
	5	2	3	
5	2	3		

Бачимо, що для того, щоб послідовність b стала повторюватися, потрібно 6 затримок. Довжини послідовностей рівні 4 і 3, а число необхідних затримок рівне $4 + 3 – 1 = 6$. Таким чином, отримуємо загальне правило знаходження лінійної взаємної кореляції двох періодичних послідовностей довжини N_1 і N_2: доповнити нулями кожну послідовність, щоб їх довжини були рівні $\boxed{N_1 + N_2 - 1}$ (тобто додати $N_2 - 1$ нулів до послідовності завдовжки N_1 і $N_1 - 1$ нулів до послідовності довжиною N_2). Нижче сказане ілюструється для зазначених вище послідовностей a і b.

Послідовність						**Затримка**	$\boldsymbol{r_{ab}(j)}$	
4	3	1	6	0	0			
5	2	3	0	0	0	0	29	
2	3	0	0	0	5	1	17	
3	0	0	0	5	2	2	12	
0	0	0	5	2	3	3	30	
0	0	5	2	3	0	4	17	
0	5	2	3	0	0	5	35	
5	2	3	0	0	0	6	29	$r_{ab}(j)$ повторюється

Отже, шукана взаємна кореляція a й b дорівнює

$$r_{ab}(j) = \{29, 17, 12, 30, 17, 35\}.$$

Вище передбачалося, що в моменти знаходження взаємної кореляції використовуються тільки цифрові дані, але взаємну кореляцію можна також порахувати аналітично, якщо сигнал записується в явній (аналітичної) формі. На практиці еквівалент описаної аналітичної процедури застосовується в аналогових схемах взаємної кореляції. Нижче наводиться приклад аналітичного розрахунку взаємної кореляції.

Приклад 3.

Знайдіть взаємну кореляцію $r_{12}(-\tau)$ двох сигналів $\upsilon_1(t)$ і $\upsilon_2(t)$, зображених на рис. 6.

Необхідні сигнали легко записати аналітично, розділивши їх на прямолінійні сегменти. Це досить зробити для одного періоду сигналу T, оскільки $r_{12}(-\tau)$ буде періодичним по τ з періодом T. Отже, для

$0 \le t \le T$, $\upsilon_1(t) = \dfrac{t}{T}$,

$0 \le t \le \dfrac{T}{2}$, $\upsilon_2(t) = 1{,}0$,

$\dfrac{T}{2} \le t \le T$, $\upsilon_2(t) = -1{,}0$.

Далі потрібно отримати вираз для $r_{12}(-\tau)$, тобто $\upsilon_2(t)$ (прямокутний сигнал) необхідно змістити вправо відносно $\upsilon_1(t)$. Для $0 \le t \le \dfrac{T}{2}$ відповідна ілюстрація приведена на рис. 7, з якого видно, що потрібно помножити на три послідовних сегмента $\upsilon_1(t)$, в яких значення $\upsilon_2(t)$ дорівнюють –1, 1, –1. Для $\dfrac{T}{2} \le \tau \le T$ необхідна ілюстрація приведена на рис. 8, на якому послідовні значення $\upsilon_2(t)$ змінилися на 1, –1, +1. Це означає, що розв'язок складається з двох частин, які потрібно узгодити в точці $\tau = \dfrac{T}{2}$.

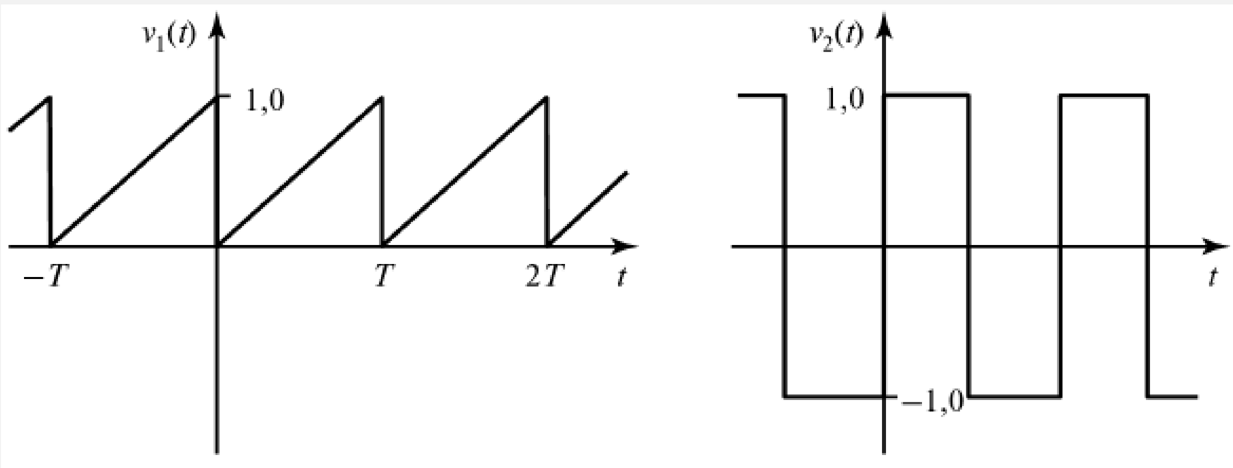

Рис. 6. Сигнали $\upsilon_1(t)$ і $\upsilon_2(t)$ (приклад 3)

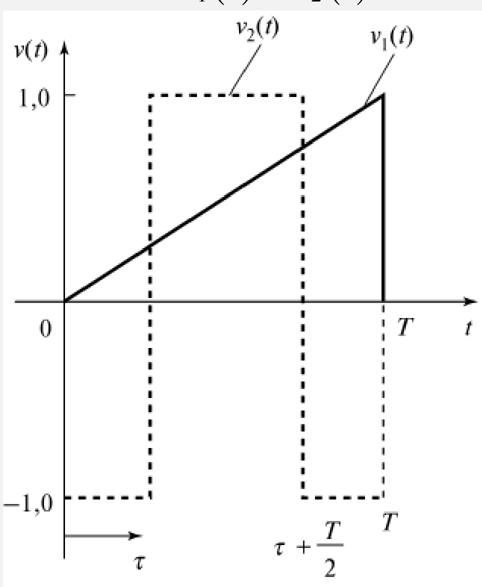

Рис. 7. Сегменти $\upsilon_2(t)$ для $0 \le t \le T$

Розіб'ємо взаємну кореляцію на три сегменти з межами в точках $t=\tau$, $t=\tau+\dfrac{T}{2}$ і $t=T$ (див. рис. 7). Далі отримуємо

$$r_{12}(-\tau)=\frac{1}{T}\int_0^T \upsilon_1(t)\upsilon_2(t-\tau)\,dt=\frac{1}{T}\int_0^T \frac{t}{T}(-1)\,dt+\frac{1}{T}\int_{\tau}^{\tau+\frac{T}{2}}\frac{t}{T}(1)\,dt+\frac{1}{T}\int_{\tau+\frac{T}{2}}^{T}\frac{t}{T}(-1)\,dt=$$
$$=-\frac{1}{T^2}\left[\frac{t^2}{2}\right]_0^{\tau}+\frac{1}{T^2}\left[\frac{t^2}{2}\right]_{\tau}^{\tau+\frac{T}{2}}-\frac{1}{T^2}\left[\frac{t^2}{2}\right]_{\tau+\frac{T}{2}}^{T} \qquad (9)$$

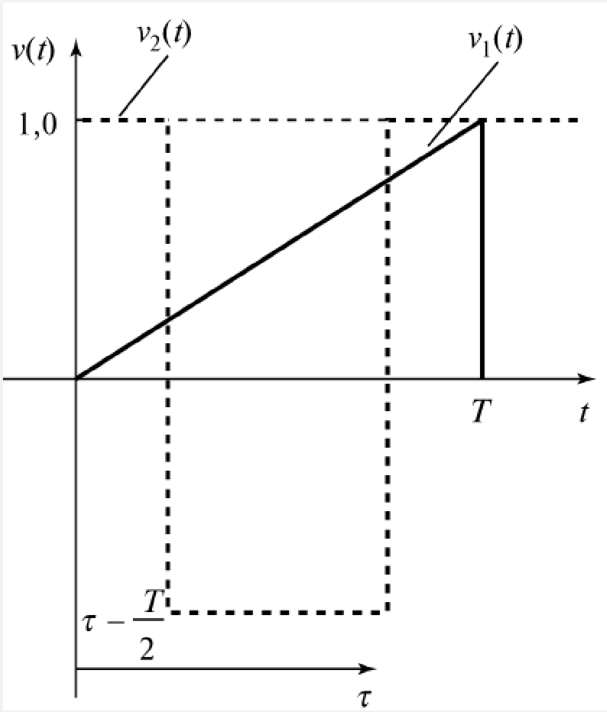

Рис. 8. Сегменти $\upsilon_2(t)$ для $\dfrac{T}{2}\le\tau\le T$

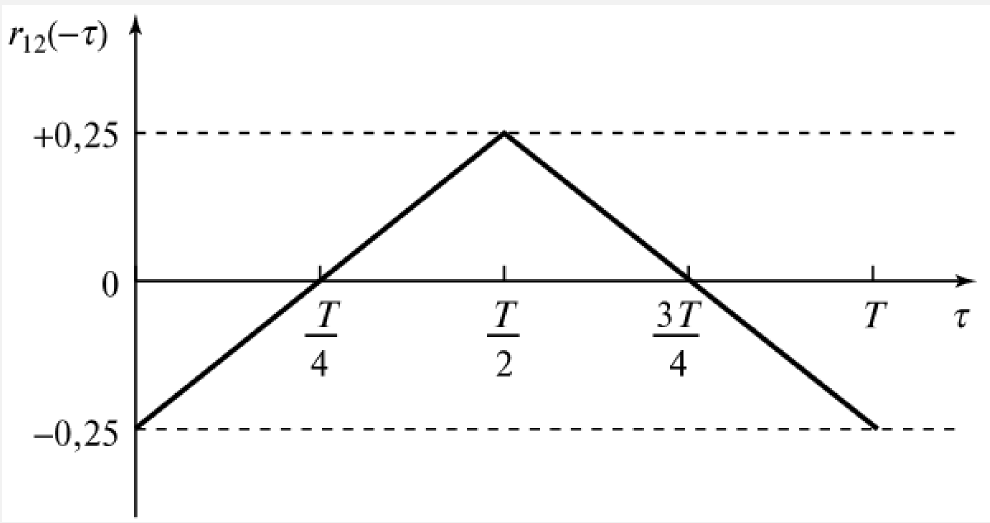

Рис. 9. Величина $r_{12}(-\tau)$ як функція τ

$$r_{12}(-\tau) = -\frac{1}{4} + \frac{\tau}{T} \text{ для } 0 \le \tau \le \tfrac{T}{2}$$

Для $\frac{T}{2} \le \tau \le T$, використовуючи рис. 8, отримуємо

$$r_{12}(-\tau) = \frac{1}{T}\int_0^{\tau-\frac{T}{2}} \frac{t}{T}(1)\,dt + \frac{1}{T}\int_{\tau-\frac{T}{2}}^{\tau} \frac{t}{T}(-1)\,dt + \frac{1}{T}\int_{\tau}^{T} \frac{t}{T}(1)\,dt \tag{10}$$

$$r_{12}(-\tau) = \frac{3}{4} - \frac{\tau}{T} \text{ для } \frac{T}{2} \le \tau \le T$$

Підставляючи $\tau = \frac{T}{2}$ у формули (9) і (10), визначаємо, що в обох випадках $r_{12}(-\tau) = \frac{1}{4}$, звідки випливає, що функції узгоджені правильно. Графік залежності $r_{12}(-\tau)$ від τ для $0 \le \tau \le T$ зображений на рис. 9.

Відзначимо деякі моменти, пов'язані з наслідками використання при обчисленні кореляції даних скінченної довжини. Іншими словами, чим відрізняється використання формули (5), в якій *T* скінченний, замість формули (2)? Для відповіді на це питання можна розглянути лише один синусоїдальний Фур'є-компонент сигналу. Формула (2) дає правильну автокореляції, де $T \ge T_p$ (T_p – період синусоїди). Отже,

$$\begin{aligned} r_{11}(\tau) &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} A\sin(\omega t)\, A\sin(\omega t + \tau)\,dt = \\ &= \lim_{T\to\infty} \frac{A^2}{2}\left[\cos(\omega\tau) - \frac{\cos(\omega T)}{2\omega T}\sin(\omega\tau)\right]. \end{aligned} \tag{11}$$

Вивчаючи дане рівняння, знаходимо, що другий доданок в дужках прямує до нуля при $T \to \infty$, так що за умови $T \ne \infty$ це вказує на помилку. Член $\cos(\omega T)$ представляє періодичний вплив похибки, а член $\frac{1}{2\omega T}$ – тенденцію в похибці. Отже, при розгляді довжини кореляції *T* похибки більші для коротших послідовностей, крім того, вони максимальні для низькочастотних компонентів сигналу. Похибки також періодичні по τ.

Компонента $\cos(\omega T)$ дає найменші похибки при $\omega T = \frac{2n+1}{2}\pi$. Оскільки $\omega = \frac{2\pi}{T_p}$ і шуканими є великі значення *T*, то це відповідає умові

$$T \ge (2n+1)\frac{T_p}{4}. \tag{12}$$

Компонента $sin(\omega T)$ дає найменше значення при $\omega\tau = m\pi$, де τ – ціле. Отже,

$$\tau = \frac{m}{2} T_p . \qquad (13)$$

Тепер варто зробити декілька припущень. Припустимо, що умову великого T задовольняє $N > 10$. Тоді $T \geq \frac{nT_p}{2}$, або

$$T \geq 5T_p . \qquad (14)$$

З рівняння (13) знаходимо, що найбільше дозволене значення τ для найбільш низькочастотного компонента $m = 1$ задовольняє умову

$$\tau < T_p . \qquad (15)$$

Об’єднуючи нерівності (14) і (15), отримуємо

$$\tau \leq \frac{T}{5} .$$

Це означає, що при знаходженні кореляції сигналів похибки, викликані скінченними довжинами послідовностей даних, можна мінімізувати таким чином:
1) гарантувати, що $T \geq 5T_p$, де T_p відповідає найменш значимому частотному компоненту;
2) суміщати дані не більш ніж на 20% їх довжини.

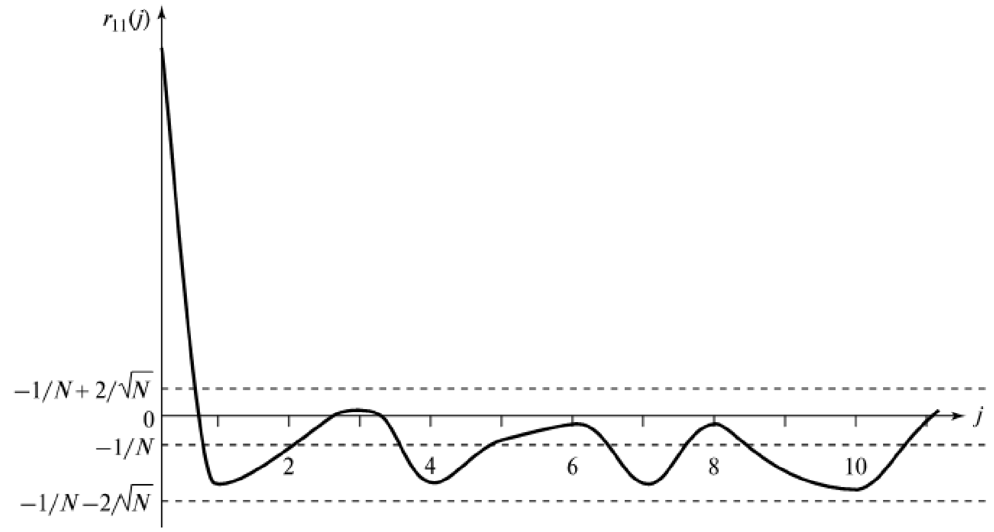

Рис. 10. Автокореляційна функція випадкового сигналу

Розглянемо приклад, коли потрібно знайти кореляцію телефонного звукового сигналу з частотним діапазоном 300–3400 Гц і того ж сигналу, дискретизованого з частотою 40 кГц, $T_p = \frac{1}{300} = 3{,}3 \times 10^{-3}$ с. Найменша прийнятна довжина даних становить $5 \times 3{,}3 \times 10^{-3} = 16{,}7$ мс, а максимальне зміщення при знаходженні кореляції 3,3 мс, або 133 точки даних.

На рис. 10 наведено графік $\rho_{11}(j)$ коефіцієнта автокореляції випадкового сигналу, наприклад, білого шуму. Можна показати, що математичне очікування $r_{11}(j)$ дорівнює $E\left[r_{11} \approx -\frac{1}{N}\right]$, де N – число точок даних, а дисперсія дорівнює $\operatorname{var}\left[r_{11}(j)\right] \approx \frac{1}{N}$. На рисунку показано математичне очікування $-\frac{1}{N}$, а також допустимі межі за рівнем $95\% - \frac{1}{N}$, які дорівнюють $\pm\frac{2}{\sqrt{N}}$. Значення $r_{11}(j)$, які не входять в ці межі, можуть бути значними, тобто вони можуть вказувати, що сигнал не зовсім випадковий. При цьому слід зазначити, що одна точка з 20 може виходити за вказані межі, навіть якщо сигнал абсолютно випадковий. Випадковий сигнал $r_{11}(j)$ повинен входити у визначені межі на рівні 95% за одну-дві затримки. Щоб гарантувати, що сигнал випадковий, потрібен певний досвід і деякі додаткові дії. Інколи рекомендують використовувати попереднє "відбілювання" даних (усунення білого шуму).

Автокореляційна функція періодичного сигналу також є періодичним сигналом. Дане твердження легко довести. Періодичний сигнал $x(t)$ з періодом T задовольняє умові

$$x(t) = x(t + nT),$$

так що

$$r_{11}(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)x(t+\tau)dt = \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)x(t+\tau+nT)dt, \qquad (16)$$

$$r_{11}(\tau) = r_{11}(\tau + nT).$$

Отже, сигнал $r_{11}(j)$ періодичний по τ з періодом T. Дана властивість корисна, оскільки дозволяє детектувати періодичні сигнали в шумі при невеликих відношеннях сигнал-шум. Знаходження автокореляції сигналу зазвичай знижує шум, при цьому виявляючи періодичну автокореляційну функцію сигналу. При необхідності після детектування можна виконувати подальшу обробку сигналу, щоб визначити його форму.

З рівняння (11) бачимо, що автокореляційна функція сигналу $A\,sin(\omega t)$ дорівнює $\frac{A^2}{2}cos(\omega\tau)$. Оскільки амплітуда автокореляційної функції пов'язана просто з амплітудою сигналу, її можна використовувати для оцінки амплітуди сигналу. Наведемо інший поширений приклад – прямокутний сигнал амплітуди A, амплітуда автокореляційної функції якого дорівнює A^2, а сама функція має трикутну форму. Нарешті, слід зазначити, що автокореляційні функції не унікальні. Це означає, що різні сигнали можуть мати однакові кореляційні функції. Таким чином, за знайденою автокореляційною функцією не можна визначити форму сигналу.

Розглянемо задачу, в якій сигнал $\upsilon(t)$ частково випадковий. Отже, вивчається зашумлений сигнал, який можна записати як суму корисного сигналу $s(t)$

і шуму $q(t)$. Таким чином,

$$\upsilon(t) = s(t) + q(t), \tag{17}$$

причому передбачається, що $s(t)$ і $q(t)$ не корелюють. Вибіркова автокореляційна функція $\upsilon(t)$, яка дорівнює $r_{\upsilon\upsilon}(j)$, записується таким чином:

$$r_{\upsilon\upsilon}(j) = \frac{1}{N}\sum_{n=0}^{N-1}\left[s(n)+q(n)\right]\left[s(n+j)+q(n+j)\right] =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1}s(n)s(n+j) + \frac{1}{N}\sum_{n=0}^{N-1}s(n)q(n+j) + \frac{1}{N}\sum_{n=0}^{N-1}q(n)s(n+j) + \frac{1}{N}\sum_{n=0}^{N-1}q(n)q(n+j) = \tag{19}$$

$$= r_{ss}(j) + E\left[s(n)q(n+j)\right] + E\left[q(n)s(n+j)\right] + E\left[q(n)q(n+j)\right] =$$

$$= r_{ss}(j) + E\left[s(n)\right]E\left[q(n+j)\right] + E\left[q(n)\right]E\left[s(n+j)\right] + E\left[q(n)\right]E\left[q(n+j)\right] =$$

$$= r_{ss}(j) + \overline{s(n)q(n)} + \overline{q(n)s(n)} + \overline{q(n)}^2 =$$

$$= r_{ss}(j) + 2\overline{sq} + \overline{q}^2. \tag{20}$$

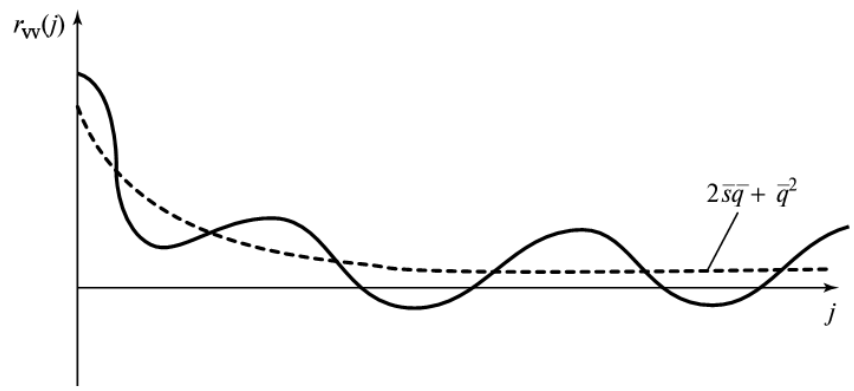

Рис. 11. Автокореляційна функція зашумленого сигналу

Тепер $\overline{q} \to 0$ для великих *N*, для яких

$$r_{\upsilon\upsilon}(j) \to r_{ss}(j). \tag{21}$$

Для менших *N* доданки взаємної кореляції у формулі (19) і автокореляція шумової складової наближаються до нуля зі збільшенням затримки *j*.

Таким чином, видно, що автокореляційна функція частково випадкового або зашумленого сигналу складається з автокореляційної функції сигнального компонента, на яку накладається загасаюча шумова функція, що залежить від випадкового компонента і корисного сигналу, і яка затухає до значення $2\overline{sq} + \overline{q}^2$. Отже, графік залежності $r_{\upsilon\upsilon}(j)$ від *j* відображає періодичність $s(t)$ за умови $\left|r_{\upsilon\upsilon}(j)\right| > \left|2\overline{sq} + \overline{q}^2\right|$ (див. рис. 11). В результаті отримуємо метод визначення періоду сигналу в шумі.

Приклад 4.

Знайдіть функцію взаємної кореляції двох зашумлених сигналів.

Дано два сигнали $\{s_1(t)+q_1(t)\}$ і $\{s_2(t)+q_2(t)\}$. Їх вибіркова взаємна кореляція $r_{12}(j)$ записується таким чином:

$$r_{12}(j)=\frac{1}{N}\sum_{n=0}^{N-1}\left[\{s_1(n)+q_1(n)\}\{s_2(n+j)+q_2(n+j)\}\right]=$$

$$=\frac{1}{N}\sum_{n=0}^{N-1}\left[s_1(n)s_2(n+j)+s_1(n)q_2(n+j)+q_1(n)s_2(n+j)+q_1(n)q_2(n+j)\right]=$$

$$=\frac{1}{N}\sum_{n=0}^{N-1}s_1(n)s_2(n+j)+\frac{1}{N}\sum_{n=0}^{N-1}s_1(n)q_2(n+j)+$$

$$+\frac{1}{N}\sum_{n=0}^{N-1}q_1(n)s_2(n+j)+\frac{1}{N}\sum_{n=0}^{N-1}q_1(n)q_2(n+j)= \tag{22}$$

$$=r_{s_1s_2}(j)+r_{s_1q_2}(j)+r_{q_1s_2}(j)+r_{q_1q_2}(j). \tag{23}$$

Як і в попередній задачі, останні три компоненти в правій частині рівняння (23) наближаються до нуля зі збільшенням затримки j. Для великих N рівняння (23) записується таким чином:

$$r_{12}(j)=r_{s_1s_2}(j)+\overline{s_1q_2}+\overline{q_1s_2}+\overline{q_1q_2} \tag{24}$$

Таким чином, при збільшенні N $r_{12}(j)\to r_{s_1s_2}(j)$, функції взаємної кореляції двох сигналів.

Проведений вище аналіз демонструє, що функції взаємної кореляції і автокореляції дозволяють проявити властивості сигналу, знижуючи шумову складову.

Застосування кореляції

Розрахунок спектральної густини енергії та енергетичного вмісту сигналу

Можна показати, що

$$F\left[r_{11}(\tau)\right]=G_E(f), \tag{25}$$

де $G_E(f)$ – спектральна густина енергії сигналу, тобто спектральна густина енергії і автокореляційна функція є Фур'є-образами одна одної.

Далі можна показати, що

$$r_{11}(0)=E, \tag{26}$$

де E – загальна енергія сигналу.

Приклад 5.

Знайдіть зв'язок між кореляційними функціями з нульовою затримкою двох різних сигналів і їх спільним енергетичним вмістом.

Позначимо сигнали $\upsilon_1(n)$ і $\upsilon_2(n)$, а їх суму $V(n)=\upsilon_1(n)+\upsilon_2(n)$. Автокореляційна функція сигналу з нульовою затримкою $V(n)$ дорівнює

$$r_{\upsilon\upsilon}(0)=E_V=\frac{1}{N}\sum_{n=0}^{N-1}V^2(n)=\frac{1}{N}\sum_{n=0}^{N-1}\left[\upsilon_1(n)+\upsilon_2(n)\right]^2,$$

де E_V – енергія сигнала $V(n)$.

$$E_V=\frac{1}{N}\sum_{n=0}^{N-1}\left[\upsilon_1^2(n)+2\upsilon_1(n)\upsilon_2(n)+\upsilon_2^2(n)\right]=$$

$$=\frac{1}{N}\sum_{n=0}^{N-1}\upsilon_1^2(n)+\frac{1}{N}\sum_{n=0}^{N-1}2\upsilon_1(n)\upsilon_2(n)+\frac{1}{N}\sum_{n=0}^{N-1}\upsilon_2^2(n),$$

так що

$$E_V=r_{\upsilon_1}(0)+r_{\upsilon_2}(0)+2r_{\upsilon_1\upsilon_2}(0). \quad (27)$$

Отримане рівняння є однією з форм шуканого результату. Альтернативним чином результат можна записати так:

$$E_V=E_{\upsilon_1}+E_{\upsilon_2}+2r_{\upsilon_1\upsilon_2}(0). \quad (28)$$

Отже, енергія сигналу $V(n)$ дорівнює сумі енергій його компонент плюс $2r_{\upsilon_1\upsilon_2}(n)$, де $r_{\upsilon_1\upsilon_2}(n)$ – функція взаємної кореляції з нульовою затримкою сигналів $\upsilon_1(n)$ і $\upsilon_2(n)$. Якщо $\upsilon_1(n)$ і $\upsilon_2(n)$ не корелюють, загальна енергія є просто сумою енергій компонентів.

Якщо сигнали $\upsilon_1(n)$ і $\upsilon_2(n)$ зашумлені і $\upsilon_1(n)=\upsilon_1'(n)+q_1(n)$ та $\upsilon_2(n)=\upsilon_2'(n)+q_2(n)$, то легко показати, що

$$E_V=E_{\upsilon_1'}+E_{\upsilon_2'}+E_{q_1}+E_{q_2}+r_{\upsilon_1'\upsilon_2'}(0). \quad (29)$$

Детектування і оцінка періодичних сигналів в шумі

Розглянемо використання взаємної кореляції для детектування та оцінки періодичних сигналів в шумі. Відомо, що сигнал, захований в шумі, можна оцінити, знайшовши його взаємну кореляцію з налаштованим шаблонним (“еталонним”) сигналом. Шаблон налаштовується методом проб і помилок з використанням будь-яких попередніх знань, поки функція взаємної кореляції не досягне максимального значення. Для підкріплення цієї пропозиції можна звернутися до рівняння (22), припускаючи, що для шаблону $q_2(n)=0$. За такої умови рівняння (23) переходить в наступне:

$$r_{12}(j)=r_{s_1s_2}(j)+r_{q_1s_2}(j)= \quad (30)$$

$$=r_{s_1s_2}(j)+\overline{q_1s_2} \quad (31)$$

Оскільки $\overline{q}_1\to 0$ при збільшенні *N*,

$$r_{12}(j)\to r_{s_1s_2}(j). \quad (32)$$

Очевидно, $r_{\upsilon_1\upsilon_2}(j)$ має максимум при $s_2(n)=s_1(n)$, де $r_{s_1s_2}$ – автокореляційна функція сигналу $s_1(n)$. Таким чином, змінюючи форму шаблону $s_2(n)$ з метою максимізації функції взаємної кореляції, можна отримати $s_2(n)$ як оцінку $s_1(n)$.

Оцінка сигналу методом підбору шаблону іноді зручна, наприклад, коли форма сигналу приблизно відома (наприклад, для певних біомедичних потенціалів), але є і більш науковий підхід, який є переважаючим. В цьому методі спочатку оцінюється період сигналу через автокореляційну функцію перешкоди сигналу, а потім знаходиться взаємна кореляція перешкоди сигналу з періодиною серією імпульсів, період якої дорівнює періоду сигналу. Отримана в результаті функція взаємної кореляції вважається оцінкою сигналу.

Позначимо сигнал з періодом N_p точок ($N_p < N$) через $s(n)$, а шум через $q(n)$, так що зашумлений сигнал дорівнює $S(n)=s(n)+q(n)$.

Нехай $\delta(n-kN_p)$ – періодична серія імпульсів, що використовується для знаходження взаємної кореляції, а N_δ – число імпульсів для знаходження взаємної кореляції. Остання величина також дорівнює числу періодів сигналу, за які знаходиться взаємна кореляція сигналу з серією імпульсів. Тоді

$$r_{S_\delta}(-j)=\frac{1}{N}\sum_{n=0}^{N-1}\left[s(n)+q(n)\right]\delta\left(b-kN_p-j\right),\quad k=0,1,2,\ldots \tag{33}$$

Для $j = 0$ і враховуючи, що $\delta(n-kN_p)=0$ для всіх $n\neq kN_p$,

$$r_{S_\delta}(0)=\frac{1}{N_\delta}\left[s(0)+q(0)+s(N_p)+q(N_p)+s(2N_p)+q(2N_p)+\ldots+s(N)+q[N]\right]. \tag{34}$$

Далі через періодичності сигналу $s(n+kN_p)=s(n)$ рівняння (34) набуває вигляду

$$r_{S_\delta}(0)=\frac{1}{N_\delta}\left[Ns(0)+q(0)+q(N_p)+q(2N_p)+\ldots+q[N]\right],$$

або

$$r_{S_\delta}(0)=s(0)+\frac{1}{N_\delta}\sum_{k=0}^{\frac{N}{N_p}}q(kN_p). \tag{35}$$

При $N\to\infty$, $\frac{1}{N_\delta}\sum_{k=0}^{\frac{N}{N_p}}q(kN_p)\to 0$, отже, $r_{S_\delta}(0)\to s(0)$. Подібним чином, для інших значень j

$$r_{S_\delta}(-j)=\frac{1}{N_\delta}\sum_{n=0}^{N-1}\left[s(n)+q(n)\right]\delta\left[(n-j)-kN_p\right],\quad k=0,1,2,\ldots,$$

в результаті також понижується рівень шуму і знаходимо значення $s(n)$ для $n=1,2,\ldots$. Отже, з рівняння (33) одержуємо

$$r_{S_\delta}(-j)=s(0),s(1),\ldots,s(N-1),\qquad j=0,1,2,\ldots,$$

що і є необхідним сигналом. Таким чином, зашумлений сигнал можна оцінити таким чином:

1) знайти автокореляцію сигнала і визначити його період;

2) знайти взаємну кореляцію сигнала з періодичною серією імпульсів, період якої дорівнює періоду сигналу; в ході цієї процедури серія імпульсів зміщується вправо щодо сигналу.

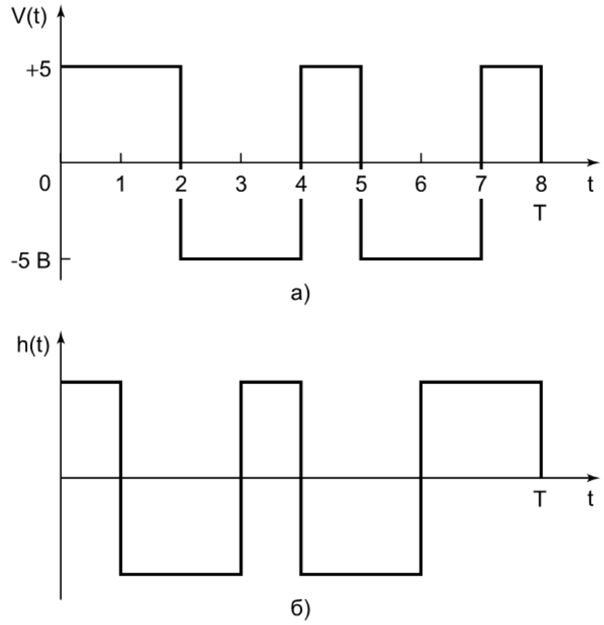

Рис. 12. Сигнал – 8-бітове кодове слово PCM (pulse code modulation) (*а*). Імпульсна характеристика відповідного узгодженого фільтра (*б*)

Швидка кореляція

Розрахунок кореляції можна прискорити, використовуючи теорему про кореляцію, яка зазвичай формулюється таким чином:

$$r_{12}(j) = F_D^{-1}\left[X_1^*(k) X_2(k)\right], \tag{36}$$

хоча коректним є таке формулювання:

$$r_{12}(j) = \frac{1}{N} F_D^{-1}\left[X_1^*(k) X_2(k)\right], \tag{37}$$

де F_D^{-1} означає обернене дискретне перетворення Фур'є. Такий підхід потребує виконання двох дискретних перетворень Фур'є (ДПФ) і одного оберненого ДПФ, що найлегше зробити, використовуючи алгоритм ШПФ. Якщо число компонент в послідовностях досить велике, то даний метод ШПФ дає результат швидше, ніж безпосередній розрахунок взаємної кореляції.

Приклад 6.

Використовуючи теорему про кореляцію, знайдемо взаємну кореляцію двох послідовностей $x_1(n)$ і $x_2(n)$.

$$x_1(n) = \{1;\, 0;\, 0;\, 1\},$$

$$x_2(n) = \{0{,}5;\, 1;\, 1;\, 0{,}5\}.$$

Використовуємо спочатку теорему про кореляцію (рівняння (37)). $X_1(k)$ дорівнює

$$X_1(k) = 2; 1+i; 0; 1-i,$$

так що

$$X_1^*(k) = 2; 1-i; 0; 1+i.$$

Для отримання $X_2(k)$ найпростіше використовувати алгоритм ШПФ. Таким чином, при $x_0 = 0{,}5$, $x_1 = 1$, $x_2 = 1$ і $x_3 = 0{,}5$ отримуємо

$$X_{21}(0) = x_0 + x_2 = 1{,}5$$

$$X_{21}(1) = x_0 - x_2 = -0{,}5$$

$$X_{22}(0) = x_1 + x_3 = 1{,}5$$

$$X_{22}(1) = x_1 - x_3 = 0{,}5$$

$$X_{11}(0) = X_{21}(0) + X_{22}(0) = 3$$

$$X_{11}(1) = X_{21}(1) + (-i)X_{22}(1) = -0{,}5 - 0{,}5i$$

$$X_{11}(2) = X_{21}(0) - X_{22}(0) = 0$$

$$X_{11}(3) = X_{21}(1) - (-i)X_{22}(1) = -0{,}5 + 0{,}5i$$

При складанні значення ШПФ, отримуємо

$$X_1^*(k) = 2; 1-i; 0; 1+i\,,$$

$$X_2(k) = 3; -0{,}5 - 0{,}5i; 0; -0{,}5 + 0{,}5i\,.$$

так що

$$X_1^*(0)X_2(0) = 2 \times 3 = 6,$$

$$X_1^*(1)X_2(1) = (1-i)(-0{,}5 - 0{,}5i) = -1,$$

$$X_1^*(2)X_2(2) = 0 \times 0 = 0,$$

$$X_1^*(3)X_2(3) = 0{,}5(1+i)(-1+i) = -1.$$

Отже,

$$\left[X_1^*(k)X_2(k)\right] = 6; -1; 0; -1.$$

Тепер необхідно до цього результату застосувати обернене ДПФ. Обернене ДПФ отримуємо шляхом заміни знаків експонент (у вагових коефіцієнтах W_N) приведеного вище алгоритму ШПФ і ділення результату на N. Отже, отримуємо такий результат:

$$X_{21}(0) = x_0 + x_2 = 6$$

$$X_{21}(1) = x_0 - x_2 = 6$$

$$X_{22}(0) = x_1 + x_3 = -2$$

$$X_{22}(1) = x_1 - x_3 = 0$$

$$X_{11}(0) = X_{21}(0) + X_{22}(0) = 4$$
$$X_{11}(1) = X_{21}(1) + iX_{22}(1) = 6$$
$$X_{11}(2) = X_{21}(0) - X_{22}(0) = 8$$
$$X_{11}(3) = X_{21}(1) - iX_{22}(1) = 6$$

Компоненти $F_D^{-1}\left[X_1^*(k)X_2(k)\right]$ знаходимо діленням значень $X_{11}(0)$, $X_{11}(1)$, $X_{11}(2)$ і $X_{11}(3)$ на $N = 4$. Таким чином,

$$F_D^{-1}\left[X_1^*(k)X_2(k)\right] = 1; 1{,}5; 2; 1{,}5 \ .$$

Далі з рівняння (37) визначаємо

$$r_{12}(i) = \frac{1}{4}F_D^{-1}\left[X_1^*(k)X_2(k)\right] = \{0{,}25; 0{,}375; 0{,}5; 0{,}375\} \ . \qquad (38)$$

Ця кореляція буде круговою (циклічною), оскільки всі дані періодичні з періодом *N*. Якщо безпосередньо порахувати взаємну кореляцію $r_{12}(i)$, то отримаємо такі значення:

$$r_{12}(0) = \frac{1}{4}(1\times 0{,}5 + 0 + 0 + 1\times 0{,}5) = 0{,}25$$
$$r_{12}(1) = \frac{1}{4}(1\times 1 + 0 + 0 + 1\times 0{,}5) = 0{,}375$$
$$r_{12}(2) = \frac{1}{4}(1\times 1 + 0 + 0 + 1\times 1) = 0{,}5$$
$$r_{12}(3) = \frac{1}{4}(1\times 0{,}5 + 0 + 0 + 1\times 1) = 0{,}375$$

Наступне значення $r_{12}(4)$ знову дорівнює 0,25 і послідовність періодично повторюється. Даний результат узгоджується з даними, отриманими вище за допомогою теореми про кореляцію. Теорему про кореляцію можна використовувати для отримання лінійної кореляції шляхом додавання до двох послідовностей доповнюючих нулів. Отже, якщо послідовність $x_1(n)$ має довжину *N*, а послідовність $x_2(n)$ – довжину N_2, то $x_1(n)$ доповнюється $N_2 - 1$ нулями, а $x_2(n)$ доповнюється $N_1 - 1$ нулями. Далі на основі двох розширених послідовностей розраховується взаємна кореляція. Цей метод обчислення взаємної кореляції за допомогою теореми про кореляцію і ШПФ називається ***швидкою кореляцією***.

Розрахунок взаємної кореляції також можна прискорити, реалізувавши його рекурсивно. Проілюструємо це на прикладі з нульовою затримкою. Взаємна кореляція при нульовій затримці двох дискретних сигналів $x_1(n)$ і $x_2(n)$ дорівнює

$$r_{12}(0) = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)x_2(n). \qquad (39)$$

Ця операція включає обчислення *N* добутків, $N - 1$ суми і одного ділення і може потребувати занадто багато часу при реалізації в реальному часі, коли пари нових даних надходять з частотою дискретизації. Розрахунок необхідно повторювати при надходженні наступної пари даних. Нові обчислення будуть відрізнятися від попередніх тільки тим, що до суми пар добутків буде додано множення нової пари і перший добуток відніметься. Отже, для кожної взаємної кореляції:

(нове значення) = (попереднє значення) +$\frac{1}{N}$(добуток двох нових даних) –
– $\frac{1}{N}$((добуток перших двох даних). (40)

Цей вираз – основа рекурсивного алгоритму. Кожна взаємна кореляція тепер вимагає тільки однієї операції множення, 1 віднімання, 1 додавання і 1 ділення за умови, що в пам’яті зберігаються добутки пар даних. Для *N*-точкової кореляції рекурсивний підхід дає правильні значення після обчислення перших $N-1$ точок.

У багатьох застосунках потрібно, щоб середнє даних дорівнювало нулю, наприклад, для усунення постійного струму з електричних сигналів. При цьому необхідно розрахувати середнє значення сигналу і відняти його від усіх дискретних значень. Це означає, що розрахунок середнього значення також можна провести рекурсивно, оскільки для кожної нової пари даних:

нове середнє = попереднє середнє +$\frac{1}{N}$(новий елемент даних – перший елемент даних).
(41)

Окрім того, віднімання середнього рівня і розрахунок взаємної кореляції можна об’єднати в один рекурсивний алгоритм. Розглянемо величини

$$\overline{x_1}(k)=\frac{1}{N}\sum_{n=0}^{N-1}x_1(n), \quad (42)$$

$$\overline{x_2}(k)=\frac{1}{N}\sum_{n=0}^{N-1}x_2(n). \quad (43)$$

Значення функції взаємної кореляції *k*-го набору з *N* точок дорівнює

$$r_{12}(k)=\frac{1}{N}\sum_{n=0}^{N-1}x_1(n)x_2(n). \quad (44)$$

При видаленні середнього значення функції взаємної кореляції дорівнює $r_{12}^0(k)$, де

$$r_{12}^0(k)=\frac{1}{N}\sum_{n=0}^{N-1}\left[x_1(n)-\bar{x}_1(n)\right]\left[x_2(n)-\bar{x}_2(n)\right], \quad (45)$$

що можна спростити до

$$r_{12}^0(k)=r_{12}(k)-\bar{x}_1(k)\bar{x}_2(k). \quad (46)$$

Об’єднуючи рівняння (40) і (43), отримуємо

$$r_{12}(k) = r_{12}(k-1) + \frac{1}{N}\left[x_1(k)x_2(k) - x_1(k-N)x_2(k-N)\right]. \quad (47)$$

З рівняння (41)

$$\bar{x}_1(k) = \bar{x}_1(k-1) + \frac{1}{N}\left[x_1(k) - x_1(k-N)\right] \quad (48)$$

і

$$\bar{x}_2(k) = \bar{x}_2(k-1) + \frac{1}{N}\left[x_2(k) - x_2(k-N)\right]. \quad (49)$$

Рівняння (46)-(49) складають рекурсивний алгоритм, який об'єднує віднімання середнього з даних з обчисленням взаємної кореляції. Кожне обчислення вимагає тільки 3 операцій множення, 4 віднімань, 3 додавань і 4 ділень. Фахівці-практики відзначають, що при зміні середніх значень даних слід акуратно вибирати N, в іншому випадку можна отримати неточні результати.

Програма 1. Програма на мові С для розрахунку взаємної кореляції та автокореляції даних

```
/*--------------------------------------- */
/*   Програма для розрахунку взаємної     */
/*   кореляції та автокореляції даних     */
/*--------------------------------------- */

#include <stdio.h>
#include <math.h>
#include <dos.h>
#include <graphics.h>

#define size 2048 /* розмір повинен бути >= подвоєній кількості точок даних*/
#define size2 1024
void acf();
void xcf();
void read_data_files();
void save_values();
FILE *in1, *in2, *out, *fopen();
long npoint;
double data1[size2], data2[size2];
int tcorl, ntype;
double ccf[size], acf1[size2];
/*---------------------------------------------------------------------------*/
main()
{
 clrscr();
 printf("select type of correlation \n");
 printf("0     for autocorrelation \n");
 printf("1     for crosscorrelation \n");
 scanf("%d",&tcorl);
 printf("select normalised or unnormalised correlation \n");
 printf("0     for unnormalised correlation \n");
 printf("1     for normalised correlation \n");
 scanf("%d",&ntype);

 read_data_files();
 if(tcorl == 0)
 {
  acf();
 }
 else
```

```
 {
  xcf();
 }
 save_values();
 printf("\n");
 printf("Correlation values have been saved in disk file result.dat \n");
 fcloseall();
 printf("\n");
 printf("press enter to continue \n");
 getch();
}
/*--------------------------------------------------------------------------*/
void read_data_files()
{
 char din1[30], din2[30];

 npoint=0;
 if(tcorl==0)
 {                    /* зчитування даних для автокореляції */

  printf("enter name of data file \n");
  scanf("%s",din1);
  if((in1=fopen(din1,"r"))==NULL)
  {
   printf("cannot open the data file for acf\n");
   exit(1);
  }
  while(fscanf(in1,"%lf",&data1[npoint])!=EOF)
  {
  ++npoint;
  }
  fclose(in1);
 }
 if(tcorl==1)
  {
  printf("enter name of 1st data file \n");
  scanf("%s",din1);
  printf("enter name of 2nd data file \n");
  scanf("%s",din2);
  if((in1=fopen(din1,"r"))==NULL)
  {
   printf("cannot open 1st data file \n");
   exit(1);
  }

  if((in2=fopen(din2,"r"))==NULL)
  {
   printf("cannot open 2nd data file\n");
   exit(1);
  }
  while(fscanf(in1,"%lf",&data1[npoint])!=EOF)
  {
   fscanf(in2,"%lf",&data2[npoint]);
   ++npoint;
  }
  fclose(in1);
  fclose(in2);
 }

 if((out=fopen("result.dat", "w"))==NULL)
 {
  printf("cannot open file result.dat \n");
  exit(1);
```

```
 }
}

/*-------------------------------------------------------*/
/* виконання взаємної кореляції між двома наборами даних */

void xcf()
{
 long j, m, k, i, i1, i2, i3, npt1, npt2;
 double sum1, sum2, rxx, ryy, sf;

 npt1=npoint-1;
 npt2=2*npoint;
 for(i=0; i<npoint; ++i)
 {
  i2=npt1+i;
  i3=npt1-i;
  sum1=0; sum2=0;
  j= npoint-i;
  for(k=0; k<j; ++k)
  {
   sum1=sum1+data1[k]*data2[k+i];
   sum2=sum2+data2[k]*data1[k+i];
  }
  ccf[i2]=sum1/npoint; ccf[i3]=sum2/npoint;
 }
 if(ntype==1)
  {               /* нормалізація значення кореляції */
  sum1=0; sum2=0;
  for(i=0; i<npoint; ++i){
   sum1=sum1+data1[i]*data1[i];
   sum2=sum2+data2[i]*data2[i];
  }
  rxx=sum1/npoint; ryy=sum2/npoint;
  sf=sqrt(rxx*ryy);
  for(i=0; i<npt2; ++i)
  {
   ccf[i]=ccf[i]/sf;
  }
 }
}
/*------------------------------------ */
/* виконання автокореляції набору даних*/

void acf()
{
 long j, m, k, i, i1;
 double sum, sf;

 for(i=0; i<npoint; ++i)
 {
  sum=0;
  j= npoint-i;
  for(k=0; k<j; ++k)
  {
   sum=sum+data1[k]*data1[k+i];
  }
  acf1[i]=sum/npoint;
 }
 if(ntype==1)
 {
  sf=acf1[0];
  for(i=0; i<npoint; ++i)
  {
```

```
    acf1[i]=acf1[i]/sf;
   }
  }
 }

 /*---------------------------------*/
 /* функція друку значень кореляції */

 void save_values()
 {
  long i, npt;
  npt=2*npoint;
  if(tcorl==0)
  {
   for(i=0; i<npoint; ++i)
   {
    fprintf(out,"\n %lf",acf1[i]);
   }
  }
  if(tcorl==1)
  {
   for(i=0; i<npt; ++i)
   {
    fprintf(out,"\n %lf",ccf[i]);
   }
  }
 }
```