КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

УДК 621.397

ОБ АНАЛИЗЕ ДАННЫХ НА ОСНОВЕ КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ [1]

А.А. ЧЕРНОМОРЕЦ
Е.В. БОЛГОВА

Белгородский государственный национальный исследовательский университет

e-mail:
chernomorets@bsu.edu.ru

В работе рассматривается математическая модель анализа двумерных наборов данных на основе косинус-частотных представлений, которая позволяет осуществлять одновременный анализ изменения состояния анализируемых объектов или явлений как в пространстве, так и во времени.

Ключевые слова: косинусное преобразование Фурье, косинус-частотные представления, квази-энергия, интегральная оценка коэффициентов ДКП.

При решении многих научных и хозяйственных задач возникает потребность анализа и обработки результатов измерений, наборов зарегистрированных данных.

В настоящее время для анализа зарегистрированных данных используются различные методы: экспертные оценки, линейные и нелинейные регрессионные методы, авторегрессионные методы, методы экспоненциального сглаживания, искусственные нейронные сети, методы на базе цепей Маркова, методы на базе классификационно-регрессионных деревьев, методы вейвлет-анализа, преобразование Фурье, дискретное косинусное преобразование и др. Известно, что косинусное преобразование Фурье преобладает свойством концентрации значений коэффициентов преобразований в области низких частот. Поэтому представляет интерес разработка метода обработки данных на основе косинус-частотных представлений [1].

В работе в качестве данных, иллюстрирующих необходимость разработки математической модели анализа зарегистрированных данных, используются сведения о загрязненности рек. В настоящее время существует потребность обработки данных на примере информации о загрязненности рек как в различные моменты времени, так и в различных точках русла реки. В данной работе предлагается математическая модель анализа двумерных наборов данных о геоэкологическом состоянии малых рек, представленных в виде значений функции $f(x_1, x_2)$, которую в дальнейших исследованиям будем рассматривать в дискретном виде в виде матрицы $\Phi = (f_{ik})$, $i = 1,2,\ldots,N_1$, $k = 1,2,\ldots,N_2$. Двумерная мо-

[1] *Работа выполнена при поддержке гранта РФФИ 14-47-08052*

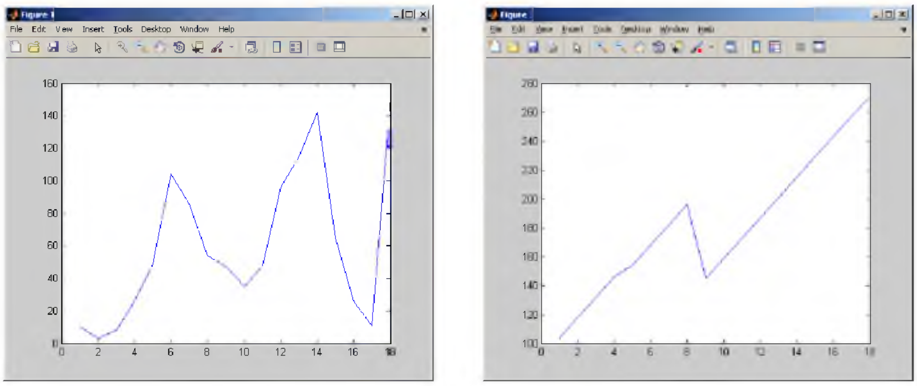
дель позволяет осуществлять одновременный анализ изменения состояния рек как в пространстве (первая координата матрицы), так и во времени (вторая координата матрицы).

В большинстве работ по анализу состояния рек используется одномерное представление значений зарегистрированных данных. Так, на рисунке 1а представлен график, отображающий данные, которые были зарегистрированы в конкретных точках русла «модельной» реки в различные моменты времени. При этом интервал времени между моментами регистрации может составлять 1 час, 1 день, неделю, месяц, год или другой промежуток времени.

На рисунке 1б представлен график, отображающий данные, которые были зарегистрированы в заданный момент времени в зависимости от местоположения точки регистрации данных.

Во многих случаях, указанное на рисунке 1 представление зарегистрированных данных не позволяет представить целостную картинку изменения характеристик состояния реки в зависимости от времени и местоположения проведения измерений.

В работе предлагается хранить и обрабатывать данные результатов измерений в виде матрицы, размерность которой соответствует количеству измерений и, в которой по вертикали (столбцы в матрице) указаны значения зарегистрированных измерений в различных пространственных точках (различное местоположение на русле реки) в одно и то же время, по горизонтали (строки матрицы) указаны значения зарегистрированных измерений в одной точке в разное время.



а б

Рис. 1. Пример отображения данных: а – зарегистрированных в одной точке русла «модельной» реки в зависимости от времени; б – зарегистрированных в одно время в зависимости от местоположения точки регистрации

Для многих задач обработки двумерных наборов зарегистрированных данных, представленных в виде матрицы $\Phi=(f_{ik})$, $i=1,2,\ldots,N_1$, $k=1,2,\ldots,N_2$, значений результатов измерений, адекватной математической основой служат частотные (косинус-частотные) представления [2], позволяющие анализировать повторяемость, периодичность изменений регистрируемой величины f_{ik},

$$f_{ik}=\frac{2}{\pi}\int_0^{\pi}\int_0^{\pi} F_C^{\Phi}(u,v)\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2}))dudv, \tag{1}$$

где частотная характеристика $F_C^{\Phi}(u,v)$ – результат косинус-преобразования Фурье,

$$F_C^{\Phi}(u,v)=\frac{2}{\pi}\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} f_{ik}\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2})), \tag{2}$$

u,v – пространственные частоты (ПЧ).

Значительное количество методов частотной обработки наборов данных основаны на анализе так называемых энергетических характеристик [3]. Данный анализ во многих случаях базируется на следствиях из фундаментального соотношения, называемом равенством Парсеваля: энергия двумерного набора зарегистрированных данных Ф определяется соотношением

$$\|\Phi\|^2 = \int_0^\pi\int_0^\pi (F_c^\Phi(u,v))^2 dudv. \qquad (3)$$

Справедливость соотношения (3) можно показать на основе следующих преобразований. Подставим выражение (2) для вычисления косинусного преобразование в левую часть соотношения (3) и выполним следующие преобразования

$$\int_0^\pi\int_0^\pi (F_C^\Phi(u,v))^2 dudv =$$

$$= \int_0^\pi\int_0^\pi \frac{2}{\pi}\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} f_{ik}\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2}))\frac{2}{\pi}\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} f_{ik}\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2}))dudv =$$

$$= \sum_{i_1=1}^{N_1}\sum_{k_1=1}^{N_2}\sum_{i_2=1}^{N_1}\sum_{k_2=1}^{N_2} f_{i_1k_1}f_{i_2k_2}\int_0^\pi\int_0^\pi \frac{4}{\pi^2}\cos(u(i_1-\frac{1}{2}))\cos(v(k_1-\frac{1}{2}))\cos(u(i_2-\frac{1}{2}))\cos(v(k_2-\frac{1}{2}))dudv =$$

$$= \sum_{i_1=1}^{N_1}\sum_{k_1=1}^{N_2}\sum_{i_2=1}^{N_1}\sum_{k_2=1}^{N_2} f_{i_1k_1}f_{i_2k_2}I_\pi, \qquad (4)$$

где

$$I_\pi = \int_0^\pi\int_0^\pi \frac{4}{\pi^2}\cos(u(i_1-\frac{1}{2}))\cos(v(k_1-\frac{1}{2}))\cos(u(i_2-\frac{1}{2}))\cos(v(k_2-\frac{1}{2}))dudv. \qquad (5)$$

Представим интеграл (5) в виде следующего произведения,

$$I_\pi = g_{i_1i_2}^\pi g_{k_1k_2}^\pi, \qquad (6)$$

где

$$g_{i_1i_2}^\pi = \int_0^\pi \frac{2}{\pi}\cos(u(i_1-\frac{1}{2}))\cos(u(i_2-\frac{1}{2}))du \ ,$$

$$g_{k_1k_2}^\pi = \int_0^\pi \frac{2}{\pi}\cos(v(k_1-\frac{1}{2}))\cos(v(k_2-\frac{1}{2}))dv.$$

Тогда, имеем,

$$g_{i_1i_2}^\pi = \int_0^\pi \frac{2}{\pi}\cos(u(i_1-\frac{1}{2}))\cos(u(i_2-\frac{1}{2}))du = \int_0^\pi \frac{2}{\pi}\frac{1}{2}(\cos(u(i_1-i_2))+\cos(u(i_2+i_2-1)))du =$$

$$= \frac{1}{\pi}\int_0^\pi \cos(u(i_1-i_2))du + \frac{1}{\pi}\int_0^\pi \cos(u(i_2+i_2-1))du = \begin{cases}1, & i_1 = i_2, \\ 0, & \text{в противномслучае.}\end{cases} \qquad (7)$$

Тогда, аналогично соотношению (7) можно показать, что

$$g_{k_1k_2}^\pi = \begin{cases}1, & k_1 = k_2, \\ 0, & \text{в противномслучае.}\end{cases}$$

Таким образом, значение интеграла в (14) имеет следующее значение,

$$I_\pi = \begin{cases}1, & i_1 = i_2, k_1 = k_2, \\ 0, & \text{в противномслучае.}\end{cases}$$

подстановка которого в (13) позволяет получить следующее соотношение

$$\int_0^\pi\int_0^\pi (F_C^\Phi(u,v))^2 dudv = \sum_{i_1=1}^{N_1}\sum_{k_1=1}^{N_2} f_{i_1k_1}f_{i_1k_1} = \|\Phi\|^2.$$

Тем самым показана справедливость соотношения (3).

Соотношение (3) можно рассматривать как равенство Парсеваля для косинусного преобразования Фурье.

Следствием соотношения (3) можно считать утверждение, что квадрат отдельного коэффициента косинусного преобразования Фурье двумерных наборов зарегистрированных данных является так называемой косинус-энергией (квазиэнергией) анализируемого двумерного набора зарегистрированных данных, соответствующего заданным нормированным пространственным частотам (u,v).

Представляет интерес исследование возможности применения указанного следствия равенства Парсеваля к результатам косинусного преобразования Фурье, которое наряду с дискретным преобразованием Фурье используется при решении задач частотной обработки данных [4].

Косинусное преобразование Фурье двумерных наборов данных определяется следующим соотношением

$$F_C^{\Phi}(u,v)=\frac{2}{\pi}\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} f_{ik}\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2})),$$

где $F_C^{\Phi}(u,v)$ – результаты косинусного преобразования набора зарегистрированных данных Ф, представленного в виде матрицы $\Phi=(f_{ik})$, $i=1,2,..N_1$, $k=1,2,..N_2$. При этом в качестве области определения косинусного преобразования обычно рассматривается область нормированных пространственных частот D_π^2,

$$D_\pi^2=\{(u,v)\,|\,0\le u,v<\pi\}. \qquad (8)$$

Можно привести различные формы записи косинусного преобразования Фурье двумерных наборов данных, представленных в виде матрицы Ф [5].

1. Косинусное преобразование Фурье двумерного набора данных Ф может быть представлено в виде произведения матрицы Ф и векторов $\vec{c}_u$ и $\vec{c}_v$,

$$F_C^{\Phi}(u,v)=\frac{2}{\pi}\vec{c}_u^{\,T}\Phi\vec{c}_v, \qquad (9)$$

где

$$\vec{c}_u=(\cos(u(1-\frac{1}{2})),\cos(u(2-\frac{1}{2})),\ldots,\cos(u(N_1-\frac{1}{2})))^T, \qquad (10)$$

$$\vec{c}_v=(\cos(v(1-\frac{1}{2})),\cos(v(2-\frac{1}{2})),\ldots,\cos(v(N_2-\frac{1}{2})))^T,$$

T – операция транспонирования матриц.

2. Коэффициенты косинусного преобразования Фурье можно вычислить на основании следующего соотношения

$$F_C^{\Phi}(u,v)=\frac{2}{\pi}tr(\Phi X_{uv}^T), \qquad (11)$$

где *tr* – операция вычисления следа матрицы,

$$X_{uv}=\vec{c}_u\vec{c}_v^{\,T}, \qquad (12)$$

некоторый элементарный набор данных, образованный векторами $\vec{c}_u$ и $\vec{c}_v$ (матрица размерности $N_1 x N_2$).

Соотношение (11) следует из преобразований:

$$F_C^{\Phi}(u,v)=\frac{2}{\pi}\vec{c}_u^{\,T}\Phi\vec{c}_v=\frac{2}{\pi}tr(\Phi\vec{c}_v\vec{c}_u^{\,T})=\frac{2}{\pi}tr(\Phi(\vec{c}_u\vec{c}_v^{\,T})^T)=\frac{2}{\pi}tr(\Phi X_{uv}^T).$$

Рассмотрим разбиение частотной области D_π^2 на прямоугольные равновеликие подобласти $\Delta_{r_1r_2}$, $r_1=1,2,\ldots,R_1$, $r_2=1,2,\ldots,R_2$,

$$\Delta_{r_1r_2}=\{(u,v)\,|\quad u_1^{r_1}\le u<u_2^{r_1},\quad v_1^{r_2}\le v<v_2^{r_2}\}, \qquad (13)$$

$$u_1^{r_1} = (r_1 - 1)\frac{\pi}{R_1}, \quad u_2^{r_1} = r_1 \frac{\pi}{R_1}, \; r_1 = 1,2,\ldots,R_1,$$

$$v_1^{r_2} = (r_2 - 1)\frac{\pi}{R_2}, \quad v_2^{r_2} = r_2 \frac{\pi}{R_2}, \; r_2 = 1,2,\ldots,R_2.$$

При указанном разбиении в подобласти $\Delta_{r_1r_2}$ переменная u принимает значения из интервала D_{r_1} (субполосы) оси абсцисс плоскости ПЧ

$$D_{r_1} = [u_1^{r_1}, u_2^{r_1}), \tag{14}$$

тогда как одновременно переменная v попадает в интервал G_{r_2} (субполосу) оси ординат

$$G_{r_2} = [v_1^{r_2}, v_2^{r_2}). \tag{15}$$

Подобласть пространственных частот $\Delta_{r_1r_2}$ схематично изображена на рисунке 2.

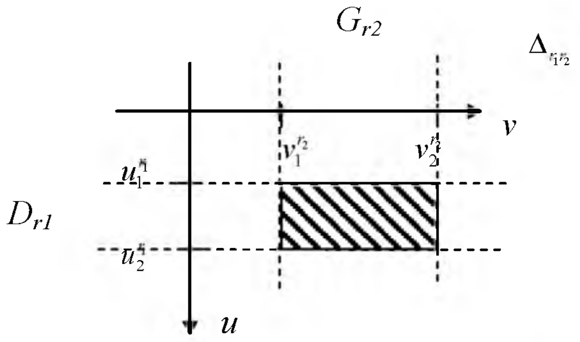


Рис. 2. Подобласть пространственных частот $\Delta_{r_1r_2}$

Учитывая разбиение (13) частотной области, равенство Парсеваля можно записать в следующей форме,

$$\|\Phi\|^2 = \iint\limits_{(u,v)\in D_\pi^2} (F_C^\Phi(u,v))^2 dudv = \sum_{r_1=1}^{R_1} \sum_{r_1=1}^{R_2} \iint\limits_{(u,v)\in\Delta_{r1r2}} (F_C^\Phi(u,v))^2 dudv, \tag{16}$$

Тогда, значение интеграла в правой части (16) можно интерпретировать как часть косинус-энергии (квазиэнергия, интегральная оценка коэффициентов дискретного косинусного преобразования Фурье) $E_{r_1r_2}^c(\Phi)$ анализируемого двумерного набора зарегистрированных данных Φ, соответствующей подобласти $\Delta_{r_1r_2}$ (13),

$$E_{r_1r_2}^c(\Phi) = \iint\limits_{(u,v)\in\Delta_{r1r2}} (F_C^\Phi(u,v))^2 dudv. \tag{17}$$

Многие задачи анализа и синтеза двумерных наборов зарегистрированных данных можно решать, используя разбиение области определения косинус-трансформант Фурье,

$$0 \le u < \pi, \quad 0 \le v < \pi,$$

на ряд подобластей $\Delta_{r_1r_2}$, $r_1 = 1,2,\ldots,R_1$, $r_2 = 1,2,\ldots,R_2$, (13) пространственных частот (ППЧ), так что (1) принимает вид

$$f_{ik} = \frac{2}{\pi} \sum_{r_1=1}^{R_1} \sum_{r_1=1}^{R_2} \iint\limits_{(u,v)\in\Delta_{r1r2}} F_C^\Phi(u,v)\cos(u(i-\frac{1}{2}))\cos(v(k-\frac{1}{2}))dudv, \tag{18}$$

Равенство Парсеваля в виде (16) позволяет построить множество различных ортонормированных базисов [6], обеспечивающих разложение анализируемого двумерного набора зарегистрированных данных на его компоненты, соответствующие различным подобластям пространственных частот.

Одной из основных характеристик, используемых при анализе набора данных, является распределение его энергии в частотной области. Косинус-частотные представления позволяют получить количественные оценки распределения энергии набора данных по подобластям пространственных частот.

Соотношения (16), (17) и (18) представляют теоретическую основу для построения методов и алгоритмов анализа и синтеза двумерных наборов зарегистрированных данных на основе косинус-частотных представлений.

Литература

1. Малоземов В. Н., Машарский С. М. Основы дискретного гармонического анализа. СПб.: НИИММ, 2003. 288 с.
2. Рабинер, Л. Теория и применение цифровой обработки сигналов [Текст] / Л. Рабинер, Г. Голд. – М.: Мир, 1988. – 512 с.
3. Сергиенко, А. Б. Цифровая обработка сигналов [Текст]: учеб. пособие для студ. вузов / А. Б. Сергиенко. – СПб.: Питер, 2002. – 603с. : ил.
4. Голд, Б. Цифровая обработка сигналов: пер. с англ. [Текст] / Б. Голд, Ч. Рейдер. – М.: Сов. радио, 1973.
5. Беллман, Р. Введение в теорию матриц [Текст] / Р. Беллман. –М.: Мир, 1990. – 368 с.
6. Жиляков, Е.Г. О субполосных свойствах изображений [Текст] / Е.Г. Жиляков, А.А. Черноморец, А.С. Белов, Е.В. Болгова // Научные ведомости БелГУ. Сер. История. Политология. Экономика. Информатика. – 2013. – № 8 (151). – Вып. 26/1. – С. 175-182.

ON THE ANALYSIS OF DATA BASED ON THE COSINE TRANSFORMATION

A.A. CHERNOMORETS
E.V. BOLGOVA

Belgorod State National Research University

e-mail: chernomorets@bsu.edu.ru

In this paper the mathematical model of the analysis of two-dimensional data sets based on cosine-frequency representation, which allows the simultaneous analysis of changes in the state of analyzed objects or phenomena, both in space and in time.

Keywords: cosine Fourier transform, cosine-frequency representation, quasi-energy, integrated assessment of DCT coefficients.