

ОПРЕДЕЛЕНИЕ ДВУМЕРНОЙ ФУНКЦИИ ПО МАЛОЙ ВЫБОРКЕ МЕТОДОМ КОМПЬЮТЕРНОЙ ТОМОГРАФИИ

Рассмотрим задачу определения статистически значимой оценки двумерной функции по выборке малого объема применительно к проблемам радиотехники. Названная задача относится к задачам восстановления функциональной зависимости по эмпирическим данным. Для получения высокой значимости результата необходима достаточно представительная выборка. Но ее объем ограничивается условиями получения эмпирических данных и особенностями используемой специализированной радиотехнической системы. Нужно отметить, что размерность оказывает существенное влияние на требуемый объем выборки. Так, зависимость, связывающая заданную точность ε восстановления функции с вероятностью η уклонения, мерностью пространства n и требуемым объемом выборки $N_{\text{треб}}$, имеет вид [1]

$$N_{\text{треб}} = \frac{8n}{\varepsilon^2} \left(\ln \frac{8n}{\varepsilon^2} - \ln \frac{\eta}{4} \right).$$

Чтобы получить возможность определять значимый результат по выборке малого объема, приходится применять методы снижения размерности. Компьютерную томографию можно рассматривать как метод снижения размерности. Она позволяет восстанавливать двумерную функцию по выборкам малого объема, т.е. в случае, когда для получения статистически значимого результата другие методы использовать нельзя [1].

Нами исследованы метеорные потоки и ассоциации, зафиксированные в результате многолетних радиолокационных наблюдений в Проблемной научно-исследовательской лаборатории радиотехники Харьковского государственного технического университета радиоэлектроники (ПНИЛ РТ). Изучено свыше 260 тыс. орбит индивидуальных метеоров, из которых выделено 5160 потоков и ассоциаций. Выделение потоков и ассоциаций производилось с помощью многоступенчатой процедуры классификации, изложенной в [2].

Для удобства анализа структуры отдельного потока в ПНИЛ РТ было предложено считать сечение потока плоскостью, перпендикулярной к средней орбите потока в точке с заданной истинной аномалией средней орбиты потока. Орбиты потока, пересекая такую плоскость, образуют картину, которая рассматривалась нами как выборка из двумерной функции.

Восстановление производилось следующим образом. Определялись семейства интегралов от искомой функциональной зависимости. В результате размерность пространства снижалась, что упрощало решаемую задачу. Затем восстанавливалась функциональная зависимость по ее линейным интегралам, или, что то же самое, определялось семейство одномерных условных функций и по нему восстанавливалась двумерная функция [3]. Задача получения условных одномерных функций решалась как задача восстановления одномерной функции по статистическим данным методом минимизации среднего риска [3].

Процесс получения двумерной функции по ее условным плотностям — или, что то же самое, восстановление функции по ее интегралам — достаточно часто рассматривается в томографии (см., например, [4]). Ниже описано решение задачи томографическим методом. При решении учитывалось, что восстановление функции по ее интегралам с помощью томографического метода относится, по Адамару, к некорректно поставленным задачам [5]. Задача решена численным методом на ЭВМ.

Таким образом, нами получен алгоритм выявления структуры объекта по заданному набору эмпирических данных. С этой точки зрения данный алгоритм можно рассматривать как алгоритм фильтрации двумерного сигнала из шума при условии неполной априорной определенности [5].

Компьютерной томографией, как известно, называется аппаратная или программная реализация алгоритмов численного восстановления функции по ее линейным интегралам. Проблемы компьютерной томографии стали широко обсуждаться в начале 70-х гг., в частности в связи с рентгенодиагностикой [4]. Истории возникновения и развития томографии и ее использованию для решения различного рода задач посвящен ряд работ. Например, в [6] предложена методика восстановления двумерной плотности вероятности томографическим методом.

Преобразованием Радона (n -мерным) называется функционал, преобразующий функцию во множество ее линейных интегралов по гиперплоскостям [4]. Для пространства мерности $n = 2$ имеем

$$f^*(s, \varphi) = \int_{\Theta} f(x, y) \delta(x \cos \varphi + y \sin \varphi - s) dx dy, \quad (1)$$

где $f(x, y)$ — функция, заданная на R^2 ; $f^*(s, \theta)$ — ее радоновский образ, причем $\theta = (\cos \varphi + \sin \varphi)$; Θ — область задания функции $f(x, y)$; δ — функция Дирака.

Пусть функции $F^*(\omega, \varphi)$ и $F(\omega_1, \omega_2)$ — результаты преобразования Фурье от функций $f^*(z)$ и $f(x, y)$ соответственно, т.е. их спектры. Тогда согласно теореме о центральном сечении справедливо равенство [6; 8]:

$$F^*(\omega, \varphi) = F(\omega \cos \varphi, \omega \sin \varphi). \quad (2)$$

Основываясь на (2), можно утверждать, что если вычислить спектр от радоновского образа, то для каждого значения угла φ будет определено центральное сечение спектра функции $f(x, y)$. Изменяя φ , можно получить все возможные сечения спектра $F(\omega_1, \omega_2)$ функции $f(x, y)$ и затем перейти к функции $f(x, y)$ [4; 6].

Один из классов алгоритмов реконструкции, т.е., обращения (1), реализует обратное преобразование Радона в том виде, в котором его получил Радон [4; 6]:

$$f(x, y) = \frac{1}{\pi} \int_0^\pi d\varphi \int_{-\infty}^{\infty} f^*(s, \varphi) q(x \cos \varphi + y \sin \varphi - s) ds, \quad (3)$$

где q имеет смысл импульсной характеристики фильтра с коэффициентом передачи $K(\omega) = \pi|\omega|$.

Выражение (3) и дает решение уравнения (1) относительно $f(x, y)$. Это решение в случае, если функция $f^*(s, \varphi)$ непрерывная, гладкая и интегрируемая, существует и единственно. Алгоритмы этого класса называются алгоритмами свертки (внутренний интеграл в (3)) и обратной проекции.

Когда функция $f^*(s, \varphi)$ задана приближенно, например:

$$f^* \ni (s, \varphi) = f^*(s, \varphi) + \xi, \quad (4)$$

где ξ — случайная величина, погрешность, получить точное решение для $f(x, y)$, решая (1), очевидно, невозможно, поскольку задача, по Адамару, является некорректно поставленной и решение будет неустойчиво [5].

Нужно получить функцию $\tilde{f}(x, y)$, в каком-то смысле близкую к $f(x, y)$, для чего необходима регуляризация выражения (3).

Для этого надо в рассмотренном алгоритме вместо импульсной функции $q(s)$ использовать функцию $q_\alpha(s)$ некоторого фильтра, позволяющую получить [4; 6]

$$\tilde{f}(x, y) = \frac{1}{\pi} \int_0^\pi d\varphi \int_{-\infty}^{\infty} f^*(s, \varphi) q_\alpha(x \cos \varphi + y \sin \varphi - s) ds; \quad (5)$$

$$\tilde{f}(x, y) \approx f(x, y); \quad (6)$$

$$q_\alpha(s) = \pi \int_{-\infty}^{\infty} |\omega| W_\alpha(|\omega|) \exp\{j2\pi\omega s\} d\omega, \quad (7)$$

где $W_\alpha(|\omega|)$ — некоторая функция, зависящая от параметра α .

Основная проблема — выработка критерия, по которому определяется, является ли $\tilde{f}(x, y)$ близкой к $f(x, y)$. Если такой критерий сформулировать, задача сводится к определению класса функций, среди которых требуется отыскать функцию $W_\alpha(|\omega|)$, а затем определить параметр α . На данном этапе находилось приближение $\tilde{f}(x, y)$ к функции $f(x, y)$ в смысле среднего квадратичного [6]. В качестве функции $W_\alpha(|\omega|)$ в работе [7] рекомендуется использовать фильтр низких частот (ФНЧ), у которого частота среза определяется параметром α и затухание в полосе непрозрачности происходит быстрее, чем возрастает ω^s , где $s \geq 4$. При создании и проверке алгоритма в качестве функции $W_\alpha(|\omega|)$ был выбран идеальный ФНЧ:

$$W_\alpha(|\omega|) = \begin{cases} 1, & \omega \leq \omega_p; \\ 0, & \omega > \omega_p, \end{cases} \quad (8)$$

где ω_p — частота среза ФНЧ, определяемая через α : $\omega_p = \omega(\alpha)$, $\alpha \in [0, 1]$.

При $\alpha = 0$ сигнал на выходе фильтра отсутствует, а при $\alpha = 1$ $W_\alpha(|\omega|) \equiv 1$.

Отладка программы и проверка ее работы осуществлялись на примере выборки размером 500 точек из двумерного нормального закона распределения с коэффициентом корреляции $K = 0,9$. Длина проекции (число дискрет или интервал группирования условного одномерного вариационного ряда) определялась по эмпирической формуле [3]

$$N_2 = 5 \text{Ln}(N_0), \quad (9)$$

после чего N_2 заменяется близким числом вида 2^P . График выборки показан на рис. 1.

Далее определялся набор проекций с помощью выполнения преобразования типа (1). Задавалось число проекций $N_1 = 8$.

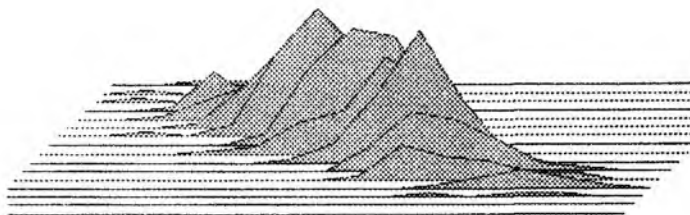


Рис. 1

После получения проекций можно восстанавливать искомую функцию, используя алгоритм фильтра и обратной проекции. Но качество восстанавливаемой функции улучшается с увеличением числа интервалов группирования. Однако, с другой стороны, с точки зрения статистической представительности интервал группирования не может быть выбран произвольно малым: желательно, чтобы число интервалов группирования не превышало оценки, полученной по (9). Поэтому необходимо вначале каким-то образом аппроксимировать проекции, после чего, используя аппроксимирующее выражение, можно увеличить число интервалов в проекции без ущерба для качества.

Если в качестве следующего шага реализовать процедуру обратного проектирования, то, как указывается в работе [9], результат будет представлять собой свертку искомой функции $f(x, y)$ с функцией вида

$$q(x, y) = 1 / (x^2 + y^2)^{1/2}.$$

Поэтому выполнение фильтрации — необходимая операция.

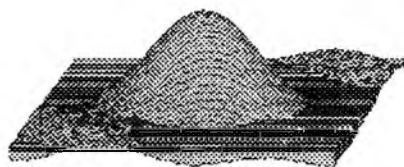


Рис. 2

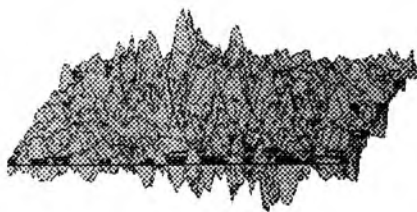


Рис. 3

Параметр фильтра α существенно влияет на результат восстановления, что видно из графиков для отладочного примера с малым (рис. 2) и большим (рис. 3) значением параметра α . Функция на рис. 2 слишком сглажена, не демонстрирует корреляции между переменными; на рис. 3 она практически разрушена шумом.

Для определения оптимального параметра фильтра строится функционал

$$\Psi(\alpha) = \left\| f^*(s, \theta) - f^{**}(s, \theta) \right\| + \alpha \left\| f^*(s, \theta) \right\|, \quad (10)$$

где $f^*(s, \theta)$ — набор сглаженных проекций, радоновский образ;
 $f^{**}(s, \theta)$ — проекции, полученные по искомой функции с помощью (1).
Точка оптимума отыскивалась методом золотого сечения по выражению
(10). Результат восстановления для оптимального параметра α показан на рис. 4.

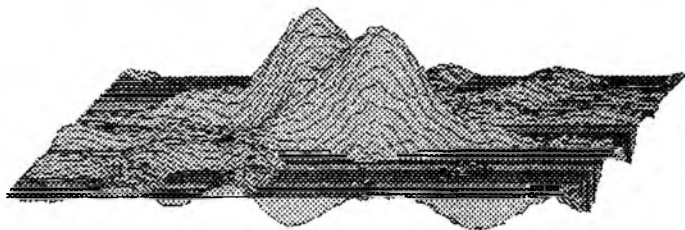


Рис. 4

Таким образом, получен алгоритм восстановления двумерной плотности вероятности по выборке. Снижение размерности позволяет добиться более высокой точности восстановления, чем при других методах (при тех же объемах выборки), что доказывает эффективность выбранного метода. Последний может рассматриваться как метод фильтрации двумерного сигнала из смеси с помехой при неизвестном виде сигнала и ограниченной информации о свойствах шума. Он может быть использован не только для проверки значимости существующей структуры потока, но и для проведения разведочного анализа.

Список литературы: 1. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. М.: Наука, 1984. 415 с. 2. Волощук Ю.И., Кащеев Б.Л., Подолька В.А. Метеорный комплекс вблизи орбиты Земли. 1. Методика выделения потоков и ассоциаций из выборки орбит индивидуальных метеоров большого объема // Астрон. вестн. 1995. Т. 29, № 5. С. 439 — 449. 3. Прикладная статистика. Ч. III. Классификация и снижение размерностей / Под ред. С.А. Айвазяна. М.: Финансы и статистика, 1989. 380 с. 4. Наттерер Ф. Математические аспекты компьютерной томографии: Пер. с англ. М.: Мир, 1990. 288 с. 5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1974. 223 с. 6. Волощук Ю.И., Воргуль А.В. Восстановление функции плотности вероятности методом компьютерной томографии. Х., 1992. 20 с. Деп. в УкрИНТЭИ 15.01.93, № 442Ук93. 7. Тихонов А.Р., Арсенин В.Я., Тимонов А.А. Математические задачи компьютерной томографии. М.: Наука, 1987. 160 с.

Харьковский государственный технический
университет радиоэлектроники

Поступила в редколлегию 18.09.98