$$S = \min_{N} \min_{C_n^N} \sum_{i=1}^{N} s_i \ . \quad (11)$$

Не касаясь пока способа определения глобального минимума на множестве сочетаний C_n^N при фиксированном N, отметим, что характер зависимости $S^*(N)$ как огибающей локальных экстремумов открывает возможность построить алгоритм целенаправленного перебора, основанный на последовательном вычислении функции (8) при значениях N = 1, 2, . . ., до нахождения минимума.

Рассмотрим вычислительный аспект определения центра группы. При фиксированном числе групп N зависимость критерия (7) от местоположения их центров является многоэкстремальной решетчатой функцией. Чтобы найти ее глобальный экстремум, нужно решить комбинаторную задачу, связанную с перебором C_n^N сочетаний и определением для каждого из них значения функции. Такая задача очень громоздка в вычислительном отношении и при больших n практически неразрешима. Поэтому необходимо разработать эвристические процедуры, которые хотя и не обеспечивают в общем случае достижения глобального экстремума, но дают решения, достаточно близкие к оптимальным, при гораздо меньших затратах времени. Здесь возможны разные подходы, в частности, основанные на различных модификациях методов последовательно-одиночного размещения, случайного поиска и т.д.

В данной работе предлагается использовать аналог метода покоординатного спуска на множестве сочетаний C_n^N.

Рассматриваемый алгоритм осуществляет поиск местоположения заданного числа центров классов, а следовательно, и разбиение на классы, т. е. определение их состава, минимизирующего выбранный критерий качества. Задача решается путем последовательного перемещения одного центра (при фиксированном положении других) в пространстве возможных точек размещения до тех пор, пока не будет вычислен локальный по данной переменной экстремум. Затем процесс повторяется для следующей переменной (центра) и т.д. до нахождения глобального экстремума. Если не наложены ограничения на число шагов или радиус сферы возможного движения каждого центра, то алгоритм реализует полный перебор сочетаний C_n^N возможного размещения N центров на n допустимых точках. Но это требует больших затрат времени. Поэтому необходимо ввести эвристические процедуры, ограничивающие перебор. Возможны различные подходы к организации такой процедуры. В данной работе принято ограничение на число шагов перемещения каждого центра, не приводящих к улучшению функции цели. Кроме того, поскольку исследуемая поверхность многоэкстремальна, для поиска глобального экстремума используется традиционный подход, основанный на многократной реализации алгоритма спуска из различных начальных приближений (точек пространства).

Описываемый алгоритм спуска может быть использован самостоятельно для решения задачи классификации при заданном числе классов N и как подпрограмма алгоритма решения более общей задачи классификации, когда N – один из оптимизируемых параметров.

Литература: 1. *Рао С.Р.* Линейные статистические методы и их применение. М.: Наука, 1968. 240 с.2. *Овезгельдыев А.О.* Взаимосвязь задач оценивания и классификации в проблеме принятия решений // Тр. 3 Междунар. конф. “Теория и техника передачи, приема и обработки информации”. Туапсе: Харьков, ХТУРЭ. 1997. С. 271. 3. *Айвазян С.А., Бежаева З.И., Староверов О.В.* Классификация многомерных наблюдений. М.: Статистика, 1974. 240 с. 4. *Орлов А.И.* Устойчивость в социально-экономических моделях. М.: Наука, 1979. 296 с. 5. *Болч Б., Хуань К. Дж.* Многомерные статистические методы для экономики: Пер. с англ. М.: Статистика, 1979. 318 с. 6. *Ту Дж., Гонсалес Р.* Принципы распознавания образов. М.: Мир, 1978. 412 с. 7. *Петров Э.Г., Аннамухамедов О.Б., Овезгельдыев А.О. и др.* Синтез информационно–вычислительного обеспечения распределенных АСПИ. Ч. 1. Методологические и инструментальные основы синтеза ИВС. Ашхабад, Изд-во АН ТССР “Ылым”, 1988. 198с.

Поступила в редколлегию 19.12.1998

Рецензент: д-р техн. наук Комяк В.М.

Овезгельдыев Атагельды Оразгельдыевич, канд. техн. наук, докторант кафедры системотехники ХТУРЭ. Научные интересы: многофакторное оценивание и многокритериальная оптимизация. Адрес: Украина, 310726, Харьков, пр. Ленина. 14, тел. 40-93-06.

Писклакова Валентина Петровна, канд.техн. наук, ст. научн. сотр., директор Центра информатизации органов управления ХТУРЭ. Научные интересы: информатизация процессов управления. Адрес: Украина, 310726, Харьков, пр. Ленина. 14, тел. 30-24-29.

УДК 519.6

ВЫБОР ПОРЯДКА МОДЕЛИ РЕГРЕССИИ

ГРИЦЮК В. И.

Рассматривается задача восстановления функции по её измерениям, содержащим случайную ошибку. Приводятся выражения для среднеквадратической ошибки прогноза при различных предположениях. Предлагаются методы выбора порядка модели, уменьшающие время сходимости.

Исследуем применение метода стохастической аппроксимации к построению оценки вектора параметров $\hat{g}(n,m) \in R^m$. Этот метод, относящийся к вероятностным итеративным методам, позволяет решать задачи, связанные с определением условия равенства нулю математического ожидания случайной функции

$$\hat{g}(n,m) = \hat{g}(n-1,m) + \gamma_n \varepsilon_n ,$$
$$\varepsilon_n = Y_n - \varphi^m(x_n)\, \hat{g}(n-1,m) , \quad (1)$$

где $\varepsilon_n, Y_n \in R^p$, $\varphi^m(x_n) \in R^{p \times m}$, $\hat{g}(n,m) \in R^m$. Используя адаптивный вариант метода, приводящий к сокращению числа итераций, определим условия, которым должны удовлетворять коэффициенты γ_n. При условии обеспечения сходимости адаптивный алгоритм становится асимптотически эффективным, если

$$\gamma_n = G / n = n^{-1}\{E[\nabla_\beta f(g)]\}^{-1},$$

здесь $\nabla_\beta(\cdot)$ обозначает частное дифференцирование по g, вычисляемое при $g=\beta$. Так как β неизвестен, то матрица [1] G заменяется на каждой итерации последней располагаемой оценкой

$$G(g_n) = \{E[\nabla_{gn} f(g)]\}^{-1}.$$

Метод будет сходиться, если известно , что β и g_n принадлежат известной прямоугольной области и удовлетворяют определенным условиям.

Итак, точки x_i случайны, независимы (между собой и с помехой), равномерно распределены на $[a,b]$ и каждая из функций $\varphi_k(x)$ ограничена на $[a,b]$. Тогда адаптивный вариант алгоритма (1) принимает вид

$$\hat{g}(n,m) = \hat{g}(n-1,m) + n^{-1}(\varphi^m(x_n))^+ \times \times (Y_n - \varphi^m(x_n)\hat{g}(n-1,m)). \quad (2)$$

Будем считать m фиксированным, обозначим $\Delta_n = \hat{g}(n,m) - g^{*m}$. Тогда из (2) получим

$$\Delta_n = \Delta_{n-1} + n^{-1}(\varphi^m(x_n))^+ \times \times (r_m(x_n) + \zeta_n - \varphi^m(x_n)\Delta_{n-1}), \quad (3)$$

где $r_m(x_n) = \varphi^{N-m}(x_n) g^{*N-m}$.

Поскольку

$E\left[(\varphi^m(x))^T \varphi^m(x)\right] = I$, $E\left[\varphi^m(x) r_m(x)\right] = 0$, $E[\zeta_n] = 0$, ζ_n не зависит от Δ_{n-1} и x_n, то

$$E_x E_\zeta \|\Delta_n\|^2 = E_x \left\|(I - n^{-1}(\varphi^m(x_n))^+ \varphi^m(x_n))\Delta_{n-1}\right\|^2 +$$

$$+ n^{-2} E_x \left\|(\varphi^m(x_n))^+ r_m(x_n)\right\|^2 +$$

$$+ n^{-2} E_x E_\zeta \left\|(\varphi^m(x_n))^+ \zeta_n\right\|^2 \le$$

$$\le \left(1 - \frac{2}{n} + o\left(\frac{1}{n}\right)\right)\|\Delta_{n-1}\|^2 + n^{-2}(\alpha_n + m\sigma^2),$$

где

$$E\left[\zeta_i \zeta_j^T\right] = \delta_{ij}\sigma^2 I. \quad (4)$$

Обозначая $v_n = E\|\Delta_n\|^2$, получаем

$$v_n \le (1 - 2n^{-1} + o(n^{-1}))v_{n-1} + n^{-2}(\alpha_n + m\sigma^2). \quad (5)$$

Из леммы Чжуна следует , что

$$v_n = n^{-1}(\alpha_m + m\sigma^2) + o(n^{-1}).$$

Тогда средняя интегральная ошибка прогноза

$$\bar{L}_m = E\left\|(F(x) - \hat{F}_m(x))\right\|^2 = = E\left\|(\varphi^m(x)(g^{*m} - \hat{g}(n,m)) + r_m(x))\right\|^2 = = v_n + \rho_m = \rho_m + n^{-1}(\alpha_m + m\sigma^2) + o(n^{-1}), \quad (6)$$

где $\rho_m = \sum_{i=m+1}^{N}(g_i^*)^2$, $E\|\varepsilon_n\|^2 = \bar{L}_m + \sigma_p^2$.

Если σ^2, ρ_m известны, а n фиксировано, то можно выбрать m по критерию

$$m^* = \arg\min_{1 \le m \le \bar{m}} A_m,\ A_m = \frac{m\sigma^2}{n} + \rho_m \quad (7)$$

и для этого m выполнить итерационный процесс (1) от 1 до n . Можно уточнить критерий (7). Если неизвестная функция $F(x)$ представляется по системе функций $\varphi_k(x_i)$ на отрезке $[a,b] \in R$, а в качестве функций используются тригонометрические полиномы с

$$\alpha_m \le 2m\rho_m \quad (8)$$

или алгебраические с $\alpha_m \le m^2\rho_m$, то, например, для (8) получаем

$$\bar{L}_m \approx \left(1 + \frac{m}{n}\right)\rho_m + \frac{m\sigma^2}{n},$$

$$E\varepsilon_n^2 \approx \left(1 + \frac{m}{n}\right)(\sigma^2 + \rho_m),\ p=1,$$

$$m = \arg\min_{1 \le m \le \bar{m}} \left(1 + \frac{m}{n}\right)\rho_m + \frac{m\sigma^2}{n}.$$

В случае, если σ^2, ρ_m неизвестны, то для оценки $E\|\varepsilon_n\|^2$ и $\bar{L}_m$ используем $\|\varepsilon_n\|^2$:

$$\|\hat{\varepsilon}(n,m)\|^2 = \frac{1}{k}\sum_{i=1}^{k}\|\varepsilon_{n-i}\|^2,\ k \le n.$$

Более точная оценка $E\varepsilon_n^2$ для системы функций, восстанавливаемой по тригонометрическим полиномам:

$$E\varepsilon_n^2 \approx \alpha_n\ E\varepsilon_{n-1}^2,$$

$$\alpha_n = \left(1 + \frac{m}{n}\right)\left(1 + \frac{m}{n-1}\right)^{-1}.$$

Поэтому, применяя формулы калмановской фильтрации, получаем

$$\hat{\varepsilon}^2(n,m) = \alpha_n \hat{\varepsilon}^2(n-1,m) + \bar{\gamma}_n\left(\varepsilon_n^2 - \alpha_n \hat{\varepsilon}^2(n-1,m)\right),$$

$$\bar{\gamma}_n = \frac{\bar{\gamma}_{n-1}\alpha_n^2}{1 + \bar{\gamma}_{n-1}\alpha_n^2}. \quad (9)$$

Выполняя рекуррентный алгоритм одновременно для всех $\underline{m} \le m \le \bar{m}$, выбираем порядок по критерию

$$m^* = \arg\min_{\underline{m} \le m \le \bar{m}} \hat{\varepsilon}^2(n,m).$$

Для вычисления $(\varphi^m(x))^+$ применяем методы факторизации, где в целях сокращения времени счета используется модифицированный метод Гивенса без квадратных корней [2].

Литература: 1. *Фильтрация* и стохастическое управление в динамических системах / Под ред. К. Т. Леондеса. М.: Мир, 1980. 408 с. 2. *Грицюк В. И.* Улучшенные алгоритмы для оценки методом наименьших квадратов // Радиоэлектроника и информатика. 1998. № 2(3). С 64-66.

Поступила в редколлегию 07.12.1998

Рецензент: д-р техн. наук Шабанов-Кушнаренко С.Ю.

Грицюк Вера Ильинична, канд. техн. наук, докторант ХТУРЭ. Научные интересы: стохастические системы управления. Хобби: музыка, литература. Адрес: Украина, 310726, Харьков, пр. Ленина, 14. Тел. 40-93-06.