ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

Том XV, № 11 РАДИОФИЗИКА 1972

УДК 512.98

ОЦЕНКА ПОТЕРЬ ИНФОРМАЦИИ И ВОЗРАСТАНИЯ РИСКА ПРИ РЕДУКЦИИ НАБЛЮДАЕМЫХ ДАННЫХ В ЗАДАЧАХ ПОСЛЕДОВАТЕЛЬНОГО ОЦЕНИВАНИЯ

С. Д. Пароджанов

В работе получены асимптотические оценки для приращения риска потерь информации при редукции наблюдаемых данных в задачах последовательного оценивания.

В задачах последовательного анализа при оценке неизвестного параметра $\Theta = (\theta_1, \theta_2, \ldots, \theta_s)$ по результатам наблюдений $x_1, \ldots, x_n, \ldots$ требуется построить решающее правило, минимизирующее риск — математическое ожидание потерь, складывающихся из стоимости проведенных наблюдений и штрафов за неверные решения. Марковская достаточная статистика (если она имеется), на основе которой можно строить оптимальные последовательные решающие правила, является обычно многомерной. Это чрезвычайно затрудняет решение задач. В связи с этим возникает необходимость редукции данных, что, очевидно, приводит к возрастанию байесовского риска.

В работе [1] были получены оценки, связывающие приращение риска и потери информации в задачах последовательного анализа. А именно, при переходе от исходных величин $x_1, \ldots, x_n$ с совместной плотностью $p(x_1, \ldots, x_n, \Theta)$ к величинам $x'_1, \ldots, x'_n$; $x'_l = \varphi_l(x_1, \ldots, x_l)$ с совместной плотностью $p'(x'_1, \ldots, x'_n, \Theta)$, порожденной исходной плотностью, имеет место оценка

$$r' - r \leqslant [2\overline{(r')^2}(I - I')]^{1/2}. \tag{1}$$

Здесь r и r' — байесовские риски, соответствующие исходному и редуцированному пространствам; $\overline{(r')^2}$ — математическое ожидание квадрата суммарных потерь в редуцированной задаче; I — шенноновское количество информации связи между $x_1, \ldots, x_n$ и Θ; аналогично, I' — между $x'_1, \ldots, x'_n$ и Θ, где число наблюдений n — случайная величина, определяемая последовательным решающим правилом.

Если выполнены условия

$$p(x_1, \ldots, x_n | \Theta) = \prod_1^n p(x_j | \Theta),$$

$$p'(x'_1, \ldots, x'_n | \Theta) = \prod_1^n p'(x'_j | \Theta) \tag{2}$$

и, кроме того,

$$p(x_1, \ldots, x_n) = \prod_1^n p(x_j), \tag{3}$$

$$p'(x'_1, \ldots, x'_n) = \prod_1^n p'(x'_i),$$

то оценка (1) принимает вид

$$r' - r \leqslant [2(\overline{r'})^2 \bar{n}(I_1 - I'_1)]^{1/2}.$$

Здесь $\bar{n}$ — математическое ожидание числа наблюдений;

$$I_1 = \int \ldots \int p(x, \Theta) \ln \frac{p(x|\Theta)}{p(x)} dx\, d\Theta,$$

$$I'_1 = \int \ldots \int p'(x', \Theta) \ln \frac{p'(x'|\Theta)}{p'(x')} dx'\, d\Theta.$$

Представляет интерес оценить потери шенноновской информации при редукции данных. Для получения такой оценки воспользуемся асимптотическими свойствами оценки максимального правдоподобия (ОМП).

Пусть $\hat{\Theta}_n$ — ОМП неизвестного параметра $\Theta = (\theta_1, \ldots, \theta_s)$ по результатам наблюдений $x_1, \ldots, x_n$ и $\hat{\Theta}'_n$ — ОМП для параметра Θ по результатам наблюдений в редуцированном пространстве $x'_1, \ldots, x'_n$ n определено последовательным решающим правилом.

Наблюдения предполагаем независимыми, т. е. считаем, что выполнены условия (2). Известно, что ОМП являются асимптотически достаточными. Следовательно, байесовские риски r и r' не изменяются при переходе от величин $x_1, \ldots, x_n$ к статистике $\hat{\Theta}_n$ и от величин $x'_1, \ldots, x'_n$, к статистике $\hat{\Theta}'_n$, т. е.

$$r' - r \leqslant [\overline{2(r')^2}(I(\hat{\Theta}'_n; \Theta) - I(\hat{\Theta}_n; \Theta))]^{1/2},$$

где

$$I(\hat{\Theta}_n; \Theta) = \int \ldots \int p(\hat{\Theta}_n, \Theta) \times$$

$$\times \ln \frac{p(\hat{\Theta}_n|\Theta)}{p(\hat{\Theta}_n)} d\hat{\Theta}_n d\Theta.$$

Здесь $p(\hat{\Theta}_n, \Theta) = p(\hat{\Theta}_n|\Theta) g(\Theta)$, а $g(\Theta)$ — распределение для Θ.

Аналогично записывается $I(\hat{\Theta}'_n; \Theta)$

В [2] было показано, что ОМП $\hat{\Theta}_n$ является асимптотически нормальной с математическим ожиданием Θ и матрицей дисперсий $[\Phi_n(\Theta)]^{-1}$, где $\Phi_n(\Theta)$ — информационная матрица Фишера с элементами $\varphi_{ij}^{(n)}(\Theta)$ $(i, j = 1, 2, \ldots, s)$,

$$\varphi_{ij}^{(n)}(\Theta) = M\left[\frac{\partial \ln p(x_1, \ldots, x_n|\Theta)}{\partial \theta_i} \frac{\partial \ln p(x_1, \ldots, x_n|\Theta)}{\partial \theta_j}\right].$$

Вычислим $I(\hat{\Theta}_n; \Theta)$. Имеем

$$I(\hat{\Theta}_n;\ \Theta) = \int \dots \int p(\hat{\Theta}_n;\ \Theta) \ln p(\hat{\Theta}_n \mid \Theta) d\hat{\Theta}_n d\Theta - \tag{4}$$

$$- \int \dots \int p(\hat{\Theta}_n) \ln p(\hat{\Theta}_n) d\hat{\Theta}_n.$$

Первый интеграл в правой части (4) нетрудно вычислить, используя асимптотическую нормальность ОМП. При вычислении воспользуемся тем, что если выполнено условие (2), то $\varphi_{ij}^{(n)}(\Theta) = \bar{n}\,\varphi_{ij}^{(1)}(\Theta)$, где

$$\varphi_{ij}^{(1)}(\Theta) = M\left[\frac{\partial \ln p(x \mid \Theta)}{\partial \theta_i} \frac{\partial \ln p(x \mid \Theta)}{\partial \theta_j}\right]$$

$$(i, j = 1, 2, \dots, s)$$

и, следовательно, $\det \Phi_n(\Theta) = (\bar{n})^s \det \Phi_1(\Theta)$.
Здесь $\Phi_1(\Theta) = \| \varphi_{ij}^{(1)}(\Theta) \|$. Окончательно

$$\int \dots \int p(\hat{\Theta}_n, \Theta) \ln p(\hat{\Theta}_n \mid \Theta) d\hat{\Theta}_n\, d\Theta =$$

$$= \frac{s}{2} \ln \bar{n} + \frac{1}{2} \int \dots \int g(\Theta) \ln \det \Phi_1(\Theta) d\Theta - \frac{s}{2} (\ln 2\pi + 1).$$

Для вычисления второго интеграла в правой части (4) найдем асимптотику $p(\hat{\Theta}_n)$.

Так как $p(\hat{\Theta}_n \mid \Theta)$ стремится с ростом n к δ-функции, то

$$p(\hat{\Theta}_n) = \int \dots \int p(\hat{\Theta}_n \mid \Theta) g(\Theta) d\Theta \to g(\hat{\Theta}_n).$$

Следовательно,

$$\left| \int \dots \int p(\hat{\Theta}_n) \ln p(\hat{\Theta}_n) d\hat{\Theta}_n - \int \dots \int g(\hat{\Theta}_n) \ln g(\hat{\Theta}_n) d\hat{\Theta}_n \right| \to 0.$$

Изменив обозначения во втором интеграле, получим асимптотическую формулу

$$I(\hat{\Theta}_n;\ \Theta) = \frac{s}{2} \ln \bar{n} + \frac{1}{2} \int \dots \int g(\Theta) \ln \det \Phi_1(\Theta) d\Theta - \tag{5}$$

$$- \int \dots \int g(\Theta) \ln g(\Theta) d\Theta - \frac{s}{2} (\ln 2\pi + 1).$$

Совершенно аналогично

$$I(\hat{\Theta}'_n;\ \Theta) = \frac{s}{2} \ln \bar{n} + \frac{1}{2} \int \dots \int g(\Theta) \ln \det \Phi'_1(\Theta) d\Theta - \tag{5a}$$

$$- \int \dots \int g(\Theta) \ln g(\Theta) d\Theta - \frac{s}{2} (\ln 2\pi + 1).$$

Таким образом, для потерь шенноновской информации получим формулу

$$I(\hat{\Theta}'_n;\ \Theta) - I(\hat{\Theta}_n;\ \Theta) = \frac{1}{2} \int \dots \int g(\Theta) \ln \frac{\det \Phi'_1(\Theta)}{\det \Phi_1(\Theta)} d\Theta. \tag{6}$$

Оценка для приращения риска [1] запишется так

$$r' - r \leqslant \left[\overline{(r')^2} \int \ldots \int g(\Theta) \ln \frac{\det \Phi_1'(\Theta)}{\det \Phi_1(\Theta)} d\Theta \right]^{1/2}. \qquad (7)$$

Заметим, что при выводе (7) условия (3) нам не понадобились. Оценку (7) можно записать в виде

$$\frac{r' - r}{\sqrt{\overline{(r')^2}}} \leqslant \left\{ M \left[\ln \frac{\det \Phi_1'(\Theta)}{\det \Phi_1(\Theta)} \right] \right\}^{1/2}. \qquad (8)$$

Если функция штрафов ограничена величиной W_0, то имеет место оценка (см. [1]):

$$I - I' \geqslant \frac{1}{W_0} \left[r \ln \frac{r}{r'} + (W_0 - r) \ln \frac{W_0 - r}{W_0 - r'} \right].$$

Подставляя найденное выражение (6) для потерь шенноновской информации, получим оценку

$$M \left[\ln \frac{\det \Phi_1'(\Theta)}{\det \Phi_1(\Theta)} \right] \geqslant \frac{2}{W_0} \left[r \ln \frac{r}{r'} + (W_0 - r) \ln \frac{W_0 - r}{W_0 - r'} \right]. \qquad (9)$$

Верхняя оценка (8) для относительного изменения риска, а также оценка (9) позволяют делать заключения о допустимости той или иной редукции данных.

Автор благодарен Э. М. Хазен за ценные советы и постоянное внимание к работе.

ЛИТЕРАТУРА

1. Э. М Хазен, Изв. АН СССР, Техническая кибернетика, № 6, 127 (1969).
2. А. Вальд, Асимптотические минимаксные решения задач последовательных точечных оценок, в прилож. к кн. Вальда «Последовательный анализ», Физматгиз, М., 1960.

Поступила в редакцию
4 июня 1971 г.

ESTIMATION OF INFORMATION LOSS AND RISK INCREASE UNDER REDUCTION OF OBSERVED DATA IN SEQUENTIAL ESTIMATION PROBLEMS

S. D. Parondjanov

Asymptotic estimations are obtained for the increment of information loss risk under reduction of the observed data in sequential estimation problems.