«Оптика атмосферы», **1**, № 10 (1988) УДК 551.576

Б.Д. Белан, З.Ф. Идрисов, К.Т. Протасов

АДАПТИВНАЯ ИДЕНТИФИКАЦИЯ ОПТИКО-МЕТЕОРОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК АТМОСФЕРЫ

В работе в байесовской постановке рассматривается задача синтеза алгоритма распознавания оптико-метеорологических ситуаций, которые можно описать параметрическими кривыми в гауссовом шуме. Синтезированный алгоритм, обладает свойством адаптации к реальным ситуациям.

Одной из важных задач обработки экспериментальных данных оптико-метеорологических исследований атмосферы является задача восстановления тех или иных характеристик атмосферы с одновременным уточнением ситуаций, порождающих эти экспериментальные данные.

Зачастую априорной информации для таких задач достаточно, чтобы сформировать параметрические модели ситуаций, которые с приемлемой для практики точностью описывают экспериментальный материал.

В условиях параметрической априорной информации для синтеза байесовых решающих правил распознавания ситуаций естественно воспользоваться идеологией адаптивного подхода, устраняя априорную неопределенность из выборочных данных обучающих последовательностей [1].

Будем полагать, что параметризованные из априорных соображений модели распознаваемых ситуаций имеют следующий вид:

$$\xi(t) - \sum_{i=1}^{n+m} \theta_\lambda^i \varphi_i^\lambda(t) = \eta^\lambda(t), \tag{1}$$

где $\xi(t)$ — наблюдаемые оптико-метеорологические параметры атмосферы; $t \in [0, T]$ — интервал времени наблюдения; $\{\varphi_\lambda^i(t)\}$ — линейно независимые функции базиса модели; $\{\theta_\lambda^i\}$ — параметры модели; $\eta^\lambda(t)$ — гауссов процесс, описывающий уклонения наблюдений и модели, и который для простоты центрирован и имеет дисперсию σ_λ^2, $\lambda \in \Lambda$.

Наряду с описанием ситуаций в виде непрерывных реализаций $\xi(t)$ процесса будем использовать эквивалентное, в смысле допустимой практикой точности (например, рекомендуемое теоремой отсчетов), оцифрованное дискретное представление с тем же обозначением $\xi(t)$:

$$\xi(t) \cong (\xi(t_j)) = \xi, \ j = 1, \dots, s.$$

В связи с этим употребляемые далее записи условных вероятностных распределений дискретных сигналов $\xi(t)$ отличаются от своих непрерывных аналогов (функционалов вероятностных распределений) монотонным преобразованием, что не нарушает смысла используемых далее решающих правил.

Предположим, что вектор неизвестных параметров модели (1) имеет следующую структуру:

$$\Theta = \left(\frac{(Y)^{n\times 1}}{(Z)^{m\times 1}} \right),$$

где $(\boldsymbol{Y})^{n\times 1}$ — подвектор информативных параметров с функцией плотности $g_\lambda(\boldsymbol{y})$, которая предполагается неизвестной, а $(\boldsymbol{Z})^{m\times 1}$ — подвектор влияющих (мешающих) параметров, имеющих распределение $g(\boldsymbol{z})$, которое не зависит от номера ситуации $\lambda \in \Lambda$, где Λ — пространство ситуаций-классов.

Аналогичную структуру имеют и базисные функции, а именно $\{\varphi_i^\lambda(t)\}_n^1$ — совокупность из n базисных функций специфична для соответствующего класса, оставшаяся часть базиса $\{\varphi_i(t)\}_{n+m}^{n+1}$ — едина для всех классов.

Как известно из теории статистических решений, оптимальное (байесово) решающее правило принятия одной из взаимоисключающих гипотез из Λ имеет вид

$$u = \arg\max_{\lambda\in\Lambda} P_\lambda f_\lambda(\xi), \tag{2}$$

где P_λ — априорная вероятность, а $f_\lambda(\xi)$ — условная функция плотности ситуации $\lambda \in \Lambda$, которые неизвестны.

Для оценивания P_λ можно воспользоваться постулатом Лапласа—Байеса и положить их равными (максимум априорной неопределенности). Для восстановления $f_\lambda(\xi)$ воспользуемся параметрической моделью (1).

При наличии априорной неопределенности параметрического типа возрастает роль вновь полученной информации как совокупности наблюдений, которые могут быть представлены в виде обучающих выборок, классифицированных «учителем»:

$$\xi_1^\lambda(t),\ldots,\ \xi_{N_\lambda}^\lambda(t),\ \ t\in[0,\ T],\ \lambda\in\Lambda,$$

где N_λ — объем выборки класса λ причем $\sum_{\lambda\in\Lambda} N_\lambda = N$, где N — общий объем выборки.

Эти данные естественно использовать для восстановления функционального соответствия между наблюдениями $\{\xi_i^\lambda(t)\}_{N_\lambda}$, состояниями природы $\lambda\in\Lambda$ и ожидаемыми потерями от принимаемых решений. В подобных условиях байесовы решающие правила (2) можно синтезировать с привлечением идеи адаптации [1] в условиях выборочных данных [2].

Используя модели ситуаций (1), представим искомые распределения следующим образом:

$$f_\lambda(\xi) = \int\limits_{Y^n}\int\limits_{Z^m} f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})\, g_\lambda(\boldsymbol{y})\, g(\boldsymbol{z})\, d\boldsymbol{z}d\boldsymbol{y}, \tag{3}$$

где $f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})$ — условная функция плотности фиксированных $\lambda\in\Lambda$, $\boldsymbol{y}\in Y^n$, $\boldsymbol{z}\in Z^m$, $d\boldsymbol{y} = dy^1\ldots dy^n$, $d\boldsymbol{z} = dz^1\ldots dz^m$.

Пусть

$$f_\lambda(\xi/\boldsymbol{y}^*,\ \boldsymbol{z}^*) = \max_{\{\Theta\in\Xi\}} f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z}), \tag{4}$$

где $\boldsymbol{\Theta}^* = \begin{pmatrix}\boldsymbol{y}^*\\ \boldsymbol{z}^*\end{pmatrix}$ — оценки максимального правдоподобия векторов $\boldsymbol{y}$ и $\boldsymbol{z}$, и пусть $\ln f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})$ — дважды дифференцируема по $\boldsymbol{y}$ и $\boldsymbol{z}$. Воспользуемся техникой приближенного интегрирования Лапласа [1], учитывая тождество

$$f(\xi/\boldsymbol{y},\ \boldsymbol{z}) = \exp\{\ln f(\xi/\boldsymbol{y},\ \boldsymbol{z})\},$$

разлагая функцию $\ln f(\xi/\boldsymbol{y},\ \boldsymbol{z})$ (считается, что зафиксировано наблюдение $\xi(t)$ и $\ln f(\xi/\boldsymbol{y},\ \boldsymbol{z})$ — есть функция переменных $\boldsymbol{y}$ и $\boldsymbol{z}$) в окрестности точки $\begin{pmatrix}\boldsymbol{y}^*\\ \boldsymbol{z}^*\end{pmatrix}$ максимального правдоподобия в ряд Тейлора и сохраняя первые три члена этого разложения, будем иметь

$$f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z}) = f_\lambda(\xi/\boldsymbol{y}^*,\ \boldsymbol{z}^*)\exp\left\{-\frac{1}{2}\begin{pmatrix}\boldsymbol{z}-\boldsymbol{z}^*\\ \boldsymbol{y}-\boldsymbol{y}^*\end{pmatrix}^T D_\lambda \begin{pmatrix}\boldsymbol{z}-\boldsymbol{z}^*\\ \boldsymbol{y}-\boldsymbol{y}^*\end{pmatrix}\right\}, \tag{5}$$

где

$$D_\lambda = \left\| \begin{array}{c|c} -\dfrac{\partial^2 \ln f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})}{\partial z^i \partial z^j} & -\dfrac{\partial^2 \ln f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})}{\partial z^i \partial y^l} \\ \hline -\dfrac{\partial^2 \ln f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})}{\partial y^t \partial z^j} & -\dfrac{\partial^2 \ln f_\lambda(\xi/\boldsymbol{y},\ \boldsymbol{z})}{\partial y^t \partial y^l} \end{array} \right\|_{\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{pmatrix} = \begin{pmatrix}\boldsymbol{y}^*\\ \boldsymbol{z}^*\end{pmatrix}} =$$

$$= \frac{1}{\sigma_\lambda^2}\left\| \begin{array}{c|c} P & q_\lambda \\ \hline q_\lambda^T & Q_\lambda \end{array} \right\|,\ \lambda\in\Lambda.$$

С учетом (5) выражение для условной функции плотности (3) будет иметь следующий вид:

$$f_\lambda(\xi) \cong f_\lambda(\xi/\boldsymbol{y}^*,\ \boldsymbol{z}^*)\int\limits_{Y^n}\int\limits_{Z^m}\exp\left\{-\frac{1}{2\sigma_\lambda^2}\left[(\boldsymbol{z}-\boldsymbol{z}^*) + P^{-1}q_\lambda^T(\boldsymbol{y}-\boldsymbol{y}^*)\right]\cdot P\times\right.$$

$$\times\left[(z-z^*)+P^{-1}q_\lambda^T(y-y^*)\right]\Big\}\times$$
$$\times\exp\left\{-\frac{1}{2\sigma_\lambda^2}(y-y^*)^T(Q_\lambda-q_\lambda Pq_\lambda^T)(y-y^*)\right\}g_\lambda(y)\,g(z)\,dzdy. \tag{6}$$

Доопределив нормирующим множителем выражение в (6) до гауссовой функции плотности и производя интегрирование по несобственному распределению $\boldsymbol{g}(\boldsymbol{z})$ в (R^m), будем иметь:

$$f_\lambda(\xi)\cong f_\lambda(\xi/y^*,\ z^*)\cdot\frac{(\sqrt{2\pi})^m}{\sigma_\lambda^{-m}(\det P)^{1/2}}\int\limits_{Y^n}\exp\left\{-\frac{1}{2\sigma_\lambda^2}(y-y^*)^T\sum\nolimits_\lambda(y-y^*)\right\}g_\lambda(y)\,dy, \tag{7}$$

где $\sum\lambda=(Q_\lambda-q_\lambda Pq_\lambda^T)$.

Имея в распоряжении линейно независимые базисные функции $\{\varphi_i^\lambda(t)\}_{n+m}$, естественно перейти к эквивалентной модели в (1) с взаимно ортогональными базисами $\{\Phi_i^\lambda(t)\}_{n+m}$, в этом случае матрица D_λ будет иметь вид

$$D_\lambda=\frac{1}{\sigma_\lambda^2}\begin{array}{c} m \\ \\ n \end{array}\left\|\begin{array}{c|c} \int\limits_0^T\Phi_i(t)\,\Phi_j(t)\,dt & 0 \\ \hline 0 & \int\limits_0^T\Phi_i^\lambda(t)\,\Phi_j^\lambda(t)\,dt \\ m & n \end{array}\right\|$$

и

$$f_\lambda(\xi)\cong f_\lambda(\xi/y^*,\ z^*)\cdot\frac{1}{\sigma_\lambda^{-m}}\,\frac{(\sqrt{2\pi})^m}{(\det P)^{1/2}}\int\limits_Y\exp\left\{-\frac{1}{2\sigma_\lambda^2}(y-y^*)^TQ_\lambda(y-y^*)\right\}g_\lambda(y)\,dy. \tag{8}$$

В отличие от параметрического случая, [1], будем рассматривать, квазипараметрический случай, суть которого в том, что распределение $g_\lambda(\boldsymbol{y})$ параметров $\boldsymbol{y}\in Y^n$, $\lambda\in\Lambda$ на практике неизвестно и для оценивания интеграла в (8) следует попытаться привлечь выборочные данные [2].

Учитывая, что значение интеграла в (8) для широкого класса априорных распределений слабо зависит от истинного вида $g_\lambda(\boldsymbol{y})$, целесообразно усреднение в (8) провести по несколько «размазанному» эмпирическому распределению оценок максимального правдоподобия параметра $\boldsymbol{y}$.

Эмпирическое распределение соответствующего класса может быть получено с использованием обучающей последовательности наблюдений и модели (1) путем вычисления оценок максимального правдоподобия параметра $\boldsymbol{y}\in Y^n\equiv R^n$ для каждого выборочного наблюдения $\xi_i^\lambda(t)$, что, в свою очередь, порождает выборку оценок максимального правдоподобия

$$\hat{y}_1^\lambda,\ldots\ \hat{y}_{N_\lambda}^\lambda,\ \lambda\in\Lambda.$$

С учетом полученных статистик будем иметь следующее выражение для условной функции плотности соответствующего класса:

$$f_\lambda(\xi)\cong\frac{1}{\sigma_\lambda^{-m+s}}\,\frac{(\sqrt{2\pi})^{m-s}}{(\det P)^{1/2}}\exp\left\{-\frac{1}{2\sigma_\lambda^2}\left\|\xi(t)-\sum_{i=1}^{n+m}\Theta_i^*\Phi_i^\lambda(t)\right\|^2\right\}\times$$
$$\times\frac{1}{N_\lambda}\sum_{j=1}^{N_\lambda}\exp\left\{-\frac{1}{2\sigma_\lambda^2}(\hat{y}_j^\lambda-y^*)^TQ_\lambda(\hat{y}_j^\lambda-y^*)\right\},\ \lambda\in\Lambda,$$

что формально и решает задачу синтеза байесова решающего правила (2), наделяя последнее свойством адаптивности.

В том случае когда дисперсии невязок σ_λ^2 неизвестны, их можно доопределить из условий минимума эмпирического риска по тем же выборочным данным.

Выбранная модель (1) естественно обобщается на другие типы описания ситуаций, когда наблюдениями являются пространственные поля $\xi(\boldsymbol{r},\,t)$, где $\boldsymbol{r}$ — пространственный вектор, изображения $\xi(x,\,y)$ и другие функции, заданные на регулярной сети наблюдений или в случайно выбранных точках.

Рассмотренный алгоритм используется для обработки экспериментального материала по оптике атмосферы, полученного с помощью самолета-лаборатории.

1. Репин В. Г., Тартаковский Г. П. Статистический синтез при априорной неопределённости и адаптация информационных систем. М.: Сов. радио. 1977. 432 с.
2. Протасов К. Т.//В кн.: Математическая статистика и её приложения. Вып. II. Томск: Изд-во ТГУ. 1987. С. 199—203.

Институт оптики атмосферы
СО АН СССР, Томск

Поступила в редакцию
4 июля 1988 г.

B. D. Belan, Z. F. Idrisov, K. T. Protasov. **Adaptive Identification of Atmospheric Optometeorological Characteristics.**

The problem of synthesizing an algorithm for the identification of optometeorological situations describable by parametric Gaussian noise curves is considered within the framework of Bayesian approach. The proposed procedure is capable of adaptation; to real atmospheric situations.