2.2. Байесовский подход в задачах обнаружения и различения

Суть байесовского подхода заключается в предположении случайного характера неизвестных параметров и введении для них априорных распределений. Одним из преимуществ подхода является существование *безусловно оптимального* решающего правила [8].

Пусть для неизвестных параметров θ введены априорные распределения $\pi_i(\theta), \theta \in \Theta_i$ для каждой из гипотез. В результате статистического усреднения плотностей наблюдений $f_i(y) = f_i(y, \theta)$ получаются безусловные плотности

$$\langle f_i(y) \rangle = \int_{\Theta_i} f_i(y, \theta) \pi_i(\theta) \mathrm{d}\theta .$$

Эти плотности уже не содержат мешающих параметров θ и соответствующие гипотезы становятся простыми.

Важно, что при использовании байесовского подхода исходная задача обнаружения (различения) со сложными гипотезами заменяется новой задачей с другими гипотезами. Эти гипотезы оказываются простыми, поэтому новая задача решается классическими методами. При таком переходе может измениться вид параметра обнаружения, что необходимо учитывать при сравнении качества алгоритмов.

В задаче обнаружения байесовский подход сводится к формированию отношения правдоподобия $\langle \Lambda(y) \rangle = \langle f_1(y) \rangle / \langle f_0(y) \rangle$ с использованием безусловных плотностей, которое сравнивается с порогом Λ_T.

Следует отметить, что введенные априорные распределения $\pi_i(\theta)$ как правило сами зависят от некоторых параметров, которые называются *гиперпараметрами*. Поэтому байесовский подход оказывается конструктивным в тех случаях, когда число гиперпараметров существенно меньше числа исходных параметров и их значения могут быть выбраны с достаточным обоснованием. Гиперпараметры также могут считаться случайными, имеющими свои априорные распределения, зависящие от гиперпараметров более высокого уровня и т. д. Дальнейшее статистическое усреднение имеет целью еще большее снижение мерности неизвестных параметров. Проблема выбора и обоснования априорных распределений при байесовском подходе является ключевой [32].

При выборе априорных распределений часто используются модели, учитывающие физику явлений и результаты экспериментальных исследований. Например, неизвестная начальная фаза принимаемого радиоимпульса часто считается случайной, имею-

щей равномерное распределение, а неизвестная амплитуда полагается случайной величиной с рэлеевским распределением. Это соответствует модели случайного узкополосного сигнала с независимыми гауссовскими квадратурными компонентами, имеющими нулевые математические ожидания.

Ориентация на наименее благоприятные априорные распределения для неизвестных параметров соответствует *минимаксному подходу*. Другие соображения связаны с *выбором наименее информативных* распределений с малым числом гиперпараметров.

Рассмотрим пример обнаружения неизвестного сдвига $\Delta = m_1 - m_0$ гауссовского распределения по единичному наблюдению: $H_0 : y \sim N(m_0, \sigma_0^2)$, $H_1^{(1)} : y \sim N(m_1, \sigma_0^2)$, при известных значениях m_0, σ_0^2. Вместо неизвестного значения m_1 введем случайную величину $\vartheta = m_0 + \Delta$, определенную на интервале $(-\infty, \infty)$ с гауссовской априорной плотностью вероятности

$$\pi(\vartheta) = (2\pi\sigma_\vartheta^2)^{-1/2} \exp(-(\vartheta - m_0)^2 / 2\sigma_\vartheta^2),$$

где σ_ϑ^2 представляет априорную дисперсию. Согласно сказанному ранее σ_ϑ^2 является гиперпараметром и ее значение должно быть задано.

Усредненная плотность наблюдения $\langle f_1(y) \rangle$ при появлении случайного сдвига остается гауссовской. Новая задача обнаружения принимает вид $H_0 : y \sim N(m_0, \sigma_0^2)$, $H_1^{(3)} : y \sim N(m_0, \sigma_A^2)$, где $\sigma_A^2 = \sigma_0^2 + \sigma_\vartheta^2$. Очевидно, что введение предположения о случайности сдвига существенно повлияло на формулировку задачи. Фактически вместо *обнаружения сдвига* плотности требуется обнаружить *увеличение дисперсии*. В качестве параметра обнаружения при этом можно использовать относительное увеличение дисперсии $\gamma = \sigma_\vartheta^2 / \sigma_0^2$. Ясно, что оптимальной статистикой для новой задачи будет $(y - m_0)^2$ или $|y - m_0|$.

Таким образом, байесовский подход в данном случае позволяет избавиться от существенного неизвестного параметра путем перехода к двухпороговой структуре обнаружителя. Априорная дисперсия σ_ϑ^2 задает меру разброса значений неизвестного параметра. Существенная априорная неопределенность значения m_0 заменяется на менее существенную неопределенность, которая связана с разбросом случайных значений параметра положения плотности.

В приведенном примере априорное распределение является *сопряженным* [34, 35] по отношению к плотности наблюдений. В результате апостериорное распределение случайного параметра совпало по виду с априорным. Такие классы распределений называются *самовоспроизводящимися*. Изменения касаются только параметров плотностей, что существенно упрощает получение оценок.

Пусть для параметров $\{\theta_i\}$ существуют достаточные статистики $\{u_i(y)\}$, а $g(\theta,u)$ обозначает *ядро* правдоподобия, т. е. $f(y\,|\,\theta) \sim g(\theta,u)$ означает пропорциональность. В случае воспроизводящегося класса ядро априорной плотности $\pi(\theta) \sim h(\theta,U_0)$ совпадает с ядром апостериорной $w(\theta\,|\,x) \sim h(\theta,U)$, где U и U_0 – параметры плотностей. Поскольку апостериорная плотность равна (с точностью до нормирующего множителя) произведению априорной и функции правдоподобия, то ядро воспроизводящегося класса плотностей должно удовлетворять условию $h(\theta,U_0)g(\theta,u) \sim h(\theta,U)$.

Сопряженные семейства имеют общее ядро, и приведенное выше условие принимает вид $g(\theta,U_0)g(\theta,u) \sim g(\theta,U)$. Используемый знак обозначает пропорциональность. Например, в случае экспоненциальных плотностей $g(\theta,U) = \exp(\theta U)$ параметры апостериорной плотности U выражаются через параметры априорной плотности U_0 и достаточную статистику наблюдения $u(y)$, так что $U = U_0 + u(y)$.

Байесовские оценки в сопряженных классах широко используются для получения рекурсивных (рекуррентных) алгоритмов, которые очень удобны при последовательных и адаптивных процедурах. Общая рекурсивная формула для вычисления статистики u на k-м шаге через ее значение на $(k-1)$-м шаге и выборочноезначение y_k имеет вид $u_k = \vartheta_k(u_{k-1}, y_k)$.