2.4. Адаптивные алгоритмы обнаружения и различения

Адаптивным называется алгоритм обработки, который способен изменять свою *структуру* в соответствии с изменениями свойств сигналов и помех. Такое определение предполагает, что алгоритм способен оценивать условия работы, принимать дополнительные решения об изменениях своей структуры, и реализовывать эти изменения в процессе обработки. В более узком смысле адаптивная система контролирует также качество своей работы [33].

В многоканальных системах изменения структуры могут осуществляться путем выбора определенного канала обработки или вариацией весовых коэффициентов при объединении каналов. Таким образом, структурные изменения достигаются путем изменений некоторой совокупности параметров, влияющих на характеристики алгоритма. Для управления параметрами *основного блока* обработки в структуру обнаружителя вводится *блок адаптации.*

Сигнально-помеховая обстановка (ситуация) обычно описывается своим набором параметров. Адаптивные алгоритмы включают оценку параметров этой неизвестной обстановки, и эти оценки используются в блоке адаптации.

Процесс оценивания условий работы называют *обучением*, в ходе которого по *обучающим* (опорным) выборкам формируются оценки неизвестных параметров «обстановки». Различают *обучение с учителем*, когда используется *классифицированная* обучающая выборка, и *самообучение* по *неклассифицированной* выборке [8, 14, 31].

Оптимальные алгоритмы *адаптивного обнаружения* (различения) для параметрических задач с неизвестными параметрами основаны на решении задачи *совместного обнаружения-оценивания*, которая является задачей существенно более высокого уровня, чем задача обнаружения или задача оценивания. При этом используется *единый* критерий качества, но возникает взаимное влияние не только оценок мешающих параметров на решения, но и решений на оценки. Это приводит к тому, что разделить в структуре *основной блок* и *блок адаптации* бывает затруднительно. В целях упрощения адаптивных алгоритмов эти блоки часто разделяются и синтезируются *раздельно*, каждый по своему критерию. В этом случае выделяют критерий качества оценивания в блоке адаптации как *критерий адаптации*, и он оказывает влияние на качество решения основной задачи обнаружения или различения.

Наряду с *рабочим режимом* для адаптивных устройств часто выделяют *режим обучения*. В режиме обучения принадлежность каждого сигнала сообщается устройству различения (в той или иной форме) вместе с самим наблюдением [13].

Процесс обучения может быть организован несколькими способами [13]. При первом способе предполагается, что обучающая выборка подчиняется тем же закономерностям, что и анализируемая в рабочем режиме. Роль «учителя» в этом случае сводится фактически лишь к тому, чтобы сообщать различителю принадлежность каждого входного наблюдения. При этом «идеальный учитель» точно знает истинную принадлежность каждого сигнала, тогда как «реальный учитель» может допускать ошибки.

При втором способе предусматривается активное участие «учителя» в формировании самой обучающей выборки. Это участие может выражаться в перестановке элементов выборки, изменении пропорций в представительстве разных сигналов и даже исключении «нехарактерных», по мнению «учителя», наблюдений. Действия «учителя» при этом подчиняются стремлению как можно быстрее добиться более эффективного обучения, т. е. такого состояния различителя, которое обеспечило бы наименьшие ошибки различения в рабочем режиме.

Обнаружение и различение сигналов в режиме самообучения также включает оценивание неизвестных параметров, однако при этом отсутствует возможность контроля над ходом самообучения и правильностью принимаемых решений. Кроме того, неизвестные параметры определяются совместно по смешанной выборке наблюдений.

Использование оценок неизвестных параметров для адаптации

В тех случаях, когда при формировании отношения правдоподобия используются *оценки максимального правдоподобия* неизвестных параметров, адаптивный подход сводится к методу *обобщенного отношения правдоподобия*.

Рассмотрим задачу обнаружения неизвестного сдвига гауссовского распределения по независимой выборке $y=(y_1,y_2,...,y_N)$. Максимум функции правдоподобия $f_1(y,\theta)=(2\pi\sigma_0^2)^{-N/2}\exp\{-(1/2\sigma_0^2)\sum_{k=1}^{N}(y_k-\theta)^2\}$ по θ достигается при $\hat{\theta}=\sum y_k/N$. Тогда обобщенное отношение правдоподобия $\Lambda_G(y)=f_1(y,\theta_1^*)/f_0(y,\theta_0^*)$ равно [2] $\Lambda_G(y)=\exp\{N(\hat{\theta}-m_0)^2/2\sigma_0^2\}$. Отсюда решающей статистикой для обнаружения будет

$|\sum y_k / N - m_0|$, что совпадает с РНМ несмещенным двусторонним правилом, рассмотренным в подразд. 1.6.

В рассмотренном случае использование оценки параметра не привело к адаптации в явном виде. Но во многих практических структурах формирование оценок неизвестных параметров осуществляется специальным блоком адаптации, который осуществляет соответствующую подстройку основного блока обработки.

Байесовские процедуры обучения

Использование *байесовских оценок* неизвестных параметров вместо неизвестных значений приводит к уже рассмотренному байесовскому подходу, который связан с заданием априорных распределений. Эта трудность вызывает поиск других подходов, называемых *эмпирическими*.

Предположим, что неизвестные параметры $\{\theta_i\}$ распределений $f_i(y|\theta_i)$, $i = 0, 1, 2, \ldots, M-1$, представляют независимые *случайные* векторы с априорными распределениями $\pi_i(\theta_i)$. Если сведения об априорных распределениях не полны, то байесовский алгоритм различения построить невозможно. Наличие классифицированных обучающих выборок $x_{об} = (x^0, x^1, \ldots x^{M-1})$ для каждого из сигналов позволяет перейти к условным плотностям.

Байесовский различитель по *максимуму апостериорной вероятности* принимает решение в пользу k-го сигнала, если $p_k f_k(y|x^k) > p_i f_i(y|x^i)$ для всех $k \neq i$.

Для вычисления условных плотностей надо найти апостериорные плотности вероятности параметров $\{\theta_i\}$ для каждой гипотезы

$$w_i(\theta_i | x^i) = C_i f_i(x^i | \theta_i)\pi_i(\theta_i),$$

где $C_i = (\int_{\Theta_i} f_i(x^i | \theta_i)\pi_i(\theta_i)\mathrm{d}\theta_i)^{-1}$ – нормирующие множители, $\theta_i \in \Theta_i$. При достаточно большой обучающей выборке эти плотности будут слабо зависеть от неизвестных априорных плотностей $\pi_i(\theta_i)$, что позволяет выбрать их произвольно. Полученные апостериорные плотности $w_i(\theta_i | x^i)$ используются вместо неизвестных априорных для нахождения условных плотностей:

$$f_i(y|x^i) = \int_{\Theta_i} f_i(y|\theta_i) w_i(\theta_i | x^i)\mathrm{d}\theta_i .$$

Задача существенно упрощается, если для параметров $\{\theta_i\}$ существуют достаточные статистики $\{u_i(x^i)\}$, а плотности $w_i(\theta_i \mid x^i)$ принадлежат *сопряженным* семействам распределений.